Advanced Mathematical Economics

Paulo B. Brito
PhD in Economics: 2020-2021
ISEG
Universidade de Lisboa
pbrito@iseg.ulisboa.pt

Lecture 8
9.12.2020

Contents

Chapter 10

Stochastic differential equations

10.1 Introduction

If we consider again the ordinary differential equation

$$\dot{y} = f(y(t)) \tag{10.1}$$

we can extend it by introducing a random perturbation,

$$\dot{Y} = f(Y(t)) + \epsilon(t) \tag{10.2}$$

and call $f(Y(t))$ the deterministic component (or skeleton) and $\epsilon(t)$ is a random perturbation. However, "noise" can be introduced in a more general form

$$\dot{Y} = f(Y(t), \epsilon(t)). \tag{10.3}$$

While the solution of (10.1) is a mapping $y : \mathbb{R}_+ \to \mathbb{R}^n$, in the cases of equations (10.2) and (10.3) the solution is a mapping $Y : \mathbb{R}_+ \times \Omega \to \mathbb{R}^n$ where $(\Omega, \mathbb{P})$ is a probability space. We denote $Y(t) = y(t) = y_t$ the *realization* of process $Y(t)$ at time $t \geq 0$.

In the previous parts, we studied the behaviour of the solution for the deterministic ODE. We saw that if function $f(.)$ is continuous and differentiable a solution $y(t)$ exists, it is unique, and it is a continuous and differentiable function of time. In addition we characterized the solution as regards the existence of steady states, their stability properties and their bifurcation behavior.

The solution of a stochastic differential equation can be seen as a (very large) family of solutions associated to their deterministic component. This is why we use $Y(t)$ instead of $y(t)$. Indeed if we fix "noise" as $\epsilon(t) = \epsilon_0$ it becomes a deterministic ODE. In this sense, some of the properties associated to the deterministic part $f(.)$, like continuity, differentiable, stability and bifurcation behavior should be checked and analysed. However, the introduction of noise implies that solutions of a stochastic differential equation may need some reinterpretation and some new features of the solutions emerge: they may not be differentiable, they do not converge to a deterministic steady state and even if the deterministic component has a fixed point, the solution may not be stable.

Simplifying, we can view stability for perturbed systems as stability in a distributional sense. We are unaware of a general bifurcation theory for stochastic differential equations. However, we can look at the solutions by trying classify the effects of the perturbation as regards their comparison with a related deterministic model:

- high noise may generate large deviations (from the deterministic solution)
- high noise may generate small deviations
- low noise can generate small deviations
- low noise can generate high deviations

There are several ways to introduce randomness in dynamic models. However, the most common model is called **diffusion equation**

$$dY(t) = f(Y(t), t)dt + \sigma(Y(t), t)dW(t) \tag{10.4}$$

where $(W(t))_{t\geq 0}$ is a **Wiener process** and $f(.)$ and $\sigma(.)$ are continuous and differentiable known functions. The main reason for this formalism is related to the fact that although $Y(t)$ is not differentiable in the classic sense, there some simple stochastic calculus rules provided by the Itô's Lemma, which resemble expanding a Taylor series up until the quadratic deviation term.

Therefore, in general, the term **stochastic differential equation** (SDE) is reserved to equations as (10.4) in the differential form or in the integral form

$$Y(t) = Y(0) + \int_0^t f(Y(s), s)ds + \int_0^t \sigma(Y(s), s)dW(s)$$

where the first integral in the right-hand-side is a Riemmann integral, but the second is an **Itô integral**. In order to solve and/or characterise SDE we have to introduce the properties of the Wiener process and of the Itô's integral.

Next we present a very brief introduction to stochastic differential equations following a heuristic approach and with a view to characterizing analytically and geometrically (when possible) the properties of the solutions.

10.2 Introduction to stochastic calculus

The most common approach to SDE's view "noise" as generated by a Wiener process and builds upon the Itô process. From this we present the basic linear SDE, the diffusion equation, and study its statistical and stability properties.

10.2.1 Stochastic processes

Stochastic processes and Markov processes

Probabilistic, analytical and geometrical characterization of stochastic processes

10.2.2 Wiener process

There are several ways of characterising the Wiener process also called standard Brownian motion.

Definition: Wiener process For our purposes we define the **Wiener process**, $(W(t))_{t\geq 0}$ as a stochastic process, where $W : \Omega \times T \to \mathbb{R}$ with the following properties

1. the initial value is equal to 0 with probability one: $\mathbb{P}[W(0) = 0] = 1$
2. it has a continuous version: i.e., a randomly generated path is a continuous function of time with probability one;
3. the path increments are independent and are Gaussian with zero mean and variance equal to the temporal increment

$$dW(t) = W(t + dt) - W(t) \sim N(0, dt), \;\; \geq 0$$

The conditional probability (or propagator) is

$$\mathbb{P}_{dt}(w^{'}|w) \equiv \mathbb{P}[W(t + dt) = w^{'}|W(t) = w]$$

if $w^{'} = w + dw$ then

$$\mathbb{P}_{dt}(w^{'}|w) = \frac{1}{\sqrt{2\pi dt}}\, e^{-\frac{(dw)^2}{2dt}}$$

Sample path properties

Proposition 1. *The Wiener process is not first-order-differentiable.*

Proof. (Heuristic) Let

$$\left|\frac{dW(t)}{dt}\right| = \left|\frac{W(t + dt) - W(t)}{dt}\right|$$

for a given $0 < t < \infty$ and $dt > 0$.

Then

$$\mathbb{E}\left[\left|\frac{dW(t)}{dt}\right|\right] = \frac{1}{|dt|}\mathbb{E}\left[|W(t + dt) - W(t)|\right]$$

But, if $W(t + dt) - W(t) = x$

$$\begin{aligned}
\mathbb{E}[|x|] &= \int_{-\infty}^{\infty} \frac{|x|}{\sqrt{2\pi dt}}\, e^{-\dfrac{x^2}{2dt}}\, dx \\
&= \frac{\sqrt{2dt}}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{|x|}{\sqrt{2dt}} e^{-\dfrac{x^2}{2dt}} \frac{dx}{\sqrt{2dt}} \\
&\text{(setting } \; y = x/\sqrt{2dt}, \text{ and as}\, dt > 0) \\
&= \frac{\sqrt{2dt}}{\sqrt{\pi}} |y|\, e^{-y^2}\, dy \\
&= \sqrt{\frac{2dt}{\pi}}
\end{aligned}$$

(see the Appendix for the properties of the Gaussian integral) then

$$\mathbb{E}\left[\left|\frac{dW(t)}{dt}\right|\right] = \sqrt{\frac{2}{\pi dt}}$$

which is of order $dt^{-1/2}$. When $dt \to 0$ tends to zero $\mathbb{E}\left[\left|\frac{dW(t)}{dt}\right|\right] \to \infty$ which means that it is not first-order differentiable. $\square$

Therefore, we can write $dW(t)$ or

$$W(t) = \int_0^t dW(t)$$

in the integral form, but

$$\frac{dW(t)}{dt}$$

is not well defined.

This is the reason why we need a particular calculus to deal with functions of Wiener processes.

Statistic properties Figure 10.1 presents one sample path and 100 sample path replications of a Wiener process.

Some properties can be derived from the definition of the Wiener process

Proposition 2. *Assume that the time variation is positive $dt > 0$.*

- *The Wiener process is stationary in expected value*

$$\mathbb{E}[dW(t)] = 0$$

- *The mathematical expectation of the square variation of the Wiener process is equal to the time increment*

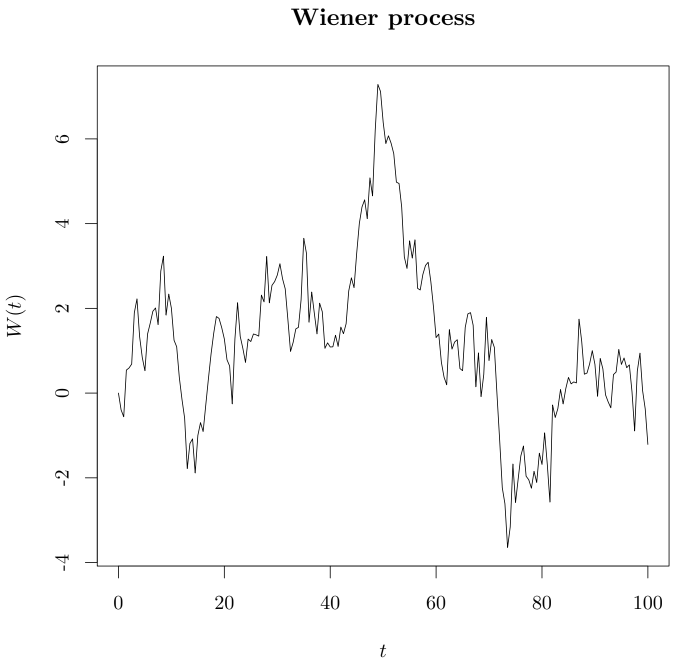
$$\mathbb{E}[(dW(t))^2] = dt$$

- *the variance of the variation is equal to the time increment*

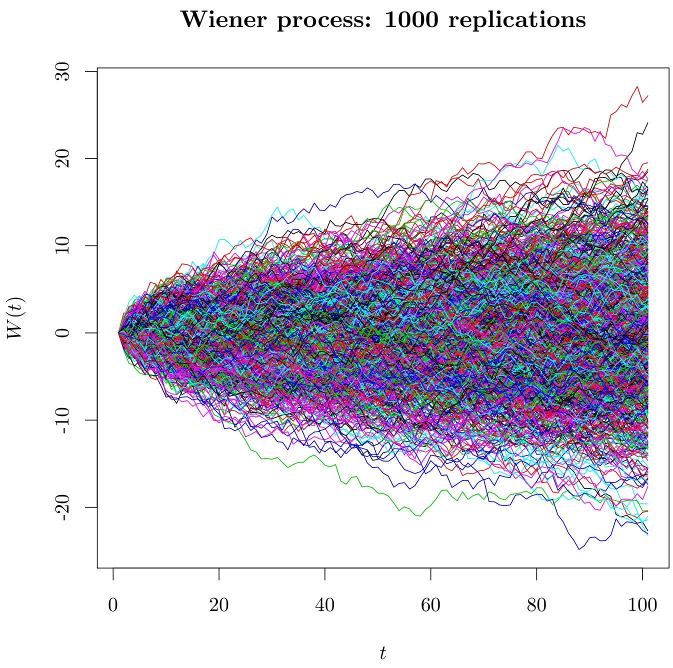
$$\mathbb{V}[dW(t)] = \mathbb{E}[\ dW(t)^2] - \mathbb{E}[\ dW(t)]^2 = dt$$

- *Let $s = dt + t$. Then the covariance of the Wiener process is*

$$Cov[W(s), W(t)] = s$$



(a) One replication



(b) 100 replications

Figure 10.1: Sample paths for the Wiener process

- *The correlation coefficient is*

$$Corr[W(s), W(t)] = \sqrt{\frac{s}{t}},\ s > t$$

Proof. Let $dW(t) = w$ and $dt > 0$. Then,

$$\mathbb{E}[w] \ = \int_{-\infty}^{\infty} \frac{w}{\sqrt{2\pi dt}} e^{-\frac{(w)^2}{2dt}}\, dw$$

if we introduce a change in variables $w = \sqrt{2dt}x$, implying $dw = \sqrt{2dt}\, dx$, then

$$\mathbb{E}[w] \ = \sqrt{\frac{2dt}{\pi}} \int_{-\infty}^{\infty} x e^{-x^2} dx = 0$$

from the properties of the Gaussian integral (see the Appendix). The variance of a change $\mathbb{V}[w] \ = \mathbb{E}[w^2] - \mathbb{E}[w]^2 = \mathbb{E}[w^2]$. Using the same transformation

$$\begin{aligned}
\mathbb{E}[w^2] \ &= \int_{-\infty}^{\infty} \frac{w^2}{\sqrt{2\pi dt}} e^{-\frac{(w)^2}{2dt}}\, d(w) = \\
&= \frac{2dt}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2} dx = \\
&= \frac{2dt}{\sqrt{\pi}} \frac{\sqrt{\pi}}{2} = \\
&= dt
\end{aligned}$$

For the covariance

$$\begin{aligned}
\mathrm{Cov}[W(s), W(t)] &= \mathrm{Cov}(W(s), W(s) - (W(s) - W(t))) = \\
&= \mathrm{Cov}(W(s), W(s)) - \mathrm{Cov}(W(s), W(s) - W(t))) = \\
&= \mathbb{V}(W(s)) - \mathrm{Cov}(W(s), dW(t))) = s
\end{aligned}$$

□

10.2.3 The Itô's integral

In the definition of the stochastic differential equation, in its integral form, we had the expression (Itô (1951))

$$\int_0^t \sigma(Y(s)) dW(s)$$

which, from the non-differentiability properties of the Wiener process needs to be addressed.

Definition Let $f(t)$ be a bounded function of time. We call **Itô's integral** to

$$I(t) = \int_0^t f(s) dW(s).$$

This definition can be extended to functions of type $f(t, w)$. If the function is bounded in the sense $\mathbb{E}[\int_0^t f(t)^2 dt] < \infty$, a more general definition of an Itô integral is

$$I(t, w) = \int_0^t f(s, w) dW(s)$$

where w is the outcome of a non-anticipating Wiener process, i.e, $w = W(s)$ for $s \leq t$.

The Itô's integral generates an **Itô's process** $(I(s, .))_{s=0}^t$.

Properties of the Itô's integral

- The Itô's integral is stationary in expected value terms, because

$$\mathbb{E}[I(t)] = \mathbb{E}\Big[\int_0^t f(s) dW(s)\Big] = \int_0^t \Big[f(s)\mathbb{E}[dW(s)]ds = 0$$

- The variance variance of the Itô's integral is

$$\mathbb{V}[I(t)] = \mathbb{E}[I(t)^2] = \int_0^t \mathbb{E}[f(s)^2]ds$$

- The integral of a sum is equal to the sum of the integrals

$$\int_0^t (f_1(s) + f_2(s))dW(s) = \int_0^t f_1(s)dW(s) + \int_0^t f_2(s)dW(s)$$

- The Itô integral is additive as regards the time integrand

$$\int_0^T f(s)dW(s) = \int_0^t f(s)dW(s) + \int_t^T f(s)dW(s)$$

  for $0 < t < T$.

10.2.4 The Itô's integral and stochastic calculus

We caw write the Itô's integral in the *differential form* as

$$dI(t) = f(t)dW(t)$$

where $dW(t)$ is a variation of the Wiener process. Even though $f(.)$ is differentialble we readily see that $I(t)$ is not first-order differentiable. However, there is differentiability in a second-order sense.

Itô's formula for a one-dimensional process Assume that $X(t)$ is an Itô's integral and assume a C^2 function $f(X)$. Then the integral $Y(t)$

$$Y(t) = g(t, X(t))$$

satisfies, in its differential form, the **Itô's formula**

$$dY(t) = g_t(t, X(t))dt + g_x(t, X(t))dX(t) + \frac{1}{2}g_{xx}(t, X(t))(dX(t))^2.$$

The following **Itô's rules** are used

$$(dt)^2 = dt\, dW(t) = 0, \; (dW(t))^2 = dt.$$

If $dX(t) = dW(t)$ then $Y(t)$ satisfies

$$dY(t) = \left(g_t(t, X(t)) + \frac{1}{2}g_{xx}(t, X(t))\right) dt + g_x(t, X(t))dW(t)$$

If $dX(t) = f(t)\, dW(t)$, then $Y(t)$ satisfies

$$dY(t) = \left(g_t(t, X(t)) + \frac{1}{2}g_{xx}(t, X(t))\, f^2(t)\right) dt + g_x(t, X(t))\, f(t)\, dW(t).$$

Examples Let $dX(t) = dW(t)$ then

- If $g(x) = a\,x + b$ then $dY(t) = \frac{a}{2}\, dW(t)$
- If $g(x) = x^a$, for $a \neq 0$ then

$$\begin{aligned} dY(t) &= \frac{a(a-1)}{2}\, X(t)^{a-2}dt + aX(t)^{a-1}\, dW(t) \\ &= \frac{a(a-1)}{2}\, Y(t)^{\frac{a-2}{a}}\, dt + aY(t)^{\frac{a-1}{a}}\, dW(t) \\ &= aY(t)^{\frac{a-2}{a}}\, \left(\frac{a-1}{2}dt + Y(t)\, dW(t)\right) \end{aligned}$$

- If $g(x) = e^{\lambda x}$, for $\lambda \neq 0$ then

$$dY(t) = \frac{\lambda^2}{2}Y(t)\, dt + \lambda Y(t)\, dW(t)$$

- If $g(x) = \ln(x)$ then

$$\begin{aligned} dY(t) &= -\frac{1}{2X(t)^2}\, dt + \frac{1}{X(t)}\, dW(t) \\ &= \frac{1}{2}e^{-2Y(t)}dt + e^{-Y(t)}\, dW(t) \end{aligned}$$

Itô's formula for a multi-dimensional process The formula can be extended to a multi-dimensional function,

$$Y(t) = f(\mathbf{X}(t), t)$$

where

$$\mathbf{X}(t) = \begin{pmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{pmatrix}$$

satisfies the variation, in its differential form,

$$dY(t) = f_t(\mathbf{X}(t), t)dt + \nabla_x f(\mathbf{X}(t), t)^\top d\mathbf{X}(t) + \frac{1}{2}(\mathbf{X}(t))^\top \ \nabla_x^2 f(\mathbf{X}(t), t) d\mathbf{X}(t),$$

where

$$\nabla_x f(\mathbf{X}(t), t) = \begin{pmatrix} f_{x_1}(\mathbf{X}(t), t) \\ \vdots \\ f_{x_n}(\mathbf{X}(t), t) \end{pmatrix}, \ \nabla_x^2 f(\mathbf{X}(t), t) = \begin{pmatrix} f_{x_1 x_1}(\mathbf{X}(t), t) & \dots & f_{x_1 x_n}(\mathbf{X}(t), t) \\ \vdots & \ddots & \vdots \\ f_{x_n x_1}(\mathbf{X}(t), t) & \dots & f_{x_n x_n}(\mathbf{X}(t), t) \end{pmatrix}$$

If there are n independent Wiener processes $\mathbf{W}(t) = (W_1(t), \dots, W_n(t))$ we use the rule

$$dW_i(t)dt = dW_i(t)dW_j(t) = 0, \text{for any}, i \neq j.$$

Example: product rule Let $Y(t) = f(X_1(t), X_2(t)) = X_1(t)X_2(t)$. Then

$$dY(t) = X_1(t)dX_2(t) + X_2(t)dX_1(t) + dX_1(t)dX_2(t)$$

To prove this, apply the Itô rule observing that we have the following derivatives of $f(x_1, x_2)$:

$$\nabla f(x_1, x_2) = \begin{pmatrix} f_{x_1}(x_1, x_2) \\ f_{x_2}(x_1, x_2) \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}, \nabla^2 f(x_1, x_2) = \begin{pmatrix} f_{x_1 x_1}(x_1, x_2) & f_{x_1 x_2}(x_1, x_2) \\ f_{x_2 x_1}(x_1, x_2) & f_{x_2 x_2}(x_1, x_2) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Then

$$dY(t) = \begin{pmatrix} X_2 & X_1 \end{pmatrix} \begin{pmatrix} dX_1 \\ dX_2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} dX_1 & dX_2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} dX_1 \\ dX_2 \end{pmatrix}.$$

10.3 The diffusion equation

The general diffusion equation is a stochastic differential equation in the Itô interpretation

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dW(t) \tag{10.5}$$

where the solution $(X(t))_{t \in T}$ is called a **diffusion process**. Next we deal with one-dimensional diffusions, $X : \Omega \times T \to \mathbb{R}$.

There are several results that allow to solve and characterise the properties of the diffusion process

10.3.1 Functions of the diffusion

Proposition 3. *Consider the process* $(Y(t))_{t\in T}$ *such that*

$$Y(t) = f(X(t))$$

where $X(t)$ *is of dimension one and* $f(.)$ *is at least* $C^2(\mathbb{R})$ *and assume it is invertible such that* $X = f^{-1}(Y) = g(Y)$. *Then* $Y(t)$ *is also a diffusion process such that*

$$dY(t) = \mu_Y(Y(t))dt + \sigma_Y(Y(t))dW(t).$$

where

$$\begin{aligned} \mu_Y(Y) &= f_x(g(Y))\mu(g(Y)) + \frac{1}{2}f_{xx}(\sigma(g(Y)))^2 \\ \sigma_Y(Y) &= f_x(g(Y))\sigma(g(Y)). \end{aligned}$$

Proof. To prove this we use the Itô's formula to find $dY(t) = d(f(X(t))$,

$$\begin{aligned} dY(t) &= f_x(X(t))dX(t) + \frac{1}{2}f_{xx}(X(t))s(dX(t))^2 \\ &= f_x(X(t))\left(\mu(X(t))dt + \sigma(X(t))dW(t)\right) + \frac{1}{2}f_{xx}(X(t))(\sigma(X(t)))^2\, dt = \\ &= \left(f_x(X(t))\mu(X(t)) + \frac{1}{2}f_{xx}(\sigma(X(t)))^2\right)dt + f_x(X(t))\sigma(X(t))dW(t). \end{aligned}$$

If the function $f(.)$ is invertible then we substitute $X = f^{-1}(Y) = g(Y)$ into the last equation. □

We can use the Itô's rule to get several properties related to the diffusion equation. In particular, we can characterise statistics for the sample path (or moment) and distribution properties.

10.3.2 Dynamics of the density: the Kolmogorov forward equation

Consider again the diffusion process

$$dX(t) = \mu(X(t))dt + \sigma(X(t))\,dW(t),\ t > 0$$

and assume the initial value is observed $X(0) = x_0$.

Let the unconditional probability, that is ad off time $t = 0$, of the realization of the process at time $t > 0$ be equal to x, $X(t) = x$ at $t > 0$, be denoted by $p(t, x)$, that is

$$p(t, x) = \mathbb{P}[X(t) = x|X(0) = x_0].$$

We can see the initial state as a Dirac-delta distribution $p(0, x) = \delta(x - x_0)$.

Assume that the support of x is $\mathbb{R}$, that $\lim_{x\to\pm\infty} p(t, x) = 0$, that the normalization condition holds

$$\int_{-\infty}^{\infty} p(t, x)dx = 1, \text{for every}\, t \geq 0.$$

Proposition 4 (Kolmogorov forward equation, also called the **Fokker-Planck equation**).
Assume we have If the initial state is x_0 at $t = 0$, that is $X(0) = x_0$, thens the density distribution of $X(t)$ at time $t > 0$, when $X(t)$ follows a diffusion process, with unbounded domain, the solution to

$$p_t(t, x) = G^*[p](t, x) \tag{10.6}$$

*where $G^*[(.)$ is the **adjoint operator***

$$G^*[p](t, x) = -\frac{\partial(\mu(x)\, p(t, x))}{\partial x} + \frac{1}{2}\frac{\partial^2(\sigma(x)^2\, p(t, x))}{\partial x^2}$$

together with $p(0, x) = \delta(x - x_0)$.

Proof. (Heuristic) Let $t \in [0, T]$ and $X = (-\infty, \infty)$ and consider an arbitrary stationary and bounded function $f(t, X(t))$ such that $f(0, X(0)) = f(T, X(T)) = 0$ for $X(t) = x \in (-\infty, \infty)$ and $\lim_{x\to\pm\infty} f(t, x) = 0$. By the Itô's Lemma

$$df(t, x) = \Big[\ \partial_t f(t, x) + \mu(x)\partial_x f(t, x) + \frac{1}{2}\sigma^2(x)\partial_{xx} f(t, x)\Big]\ dt + \Big(\sigma(x)\partial_x f(t, x)\Big)\, dW(t).$$

The variation of f from $t = 0$ to $t = T$ is

$$\int_0^T df(t, x) = \int_0^T \Big[\ \partial_t f(t, x) + \mu(x)\partial_x f(t, x) + \frac{1}{2}\sigma^2(x)\partial_{xx} f(t, x)\Big]\ dt + \int_0^T \Big(\sigma(x)\partial_x f(t, x)\Big)\, dW(t).$$

Taking the unconditional expected value

$$\begin{aligned}
\mathbb{E}\Big[\int_0^T df(t)\Big] &= \mathbb{E}\Bigg[\int_0^T \Big[\ \partial_t f(t, x) + \mu(x)\partial_x f(t, x) + \frac{1}{2}\sigma^2(x)\partial_{xx} f(t, x)\Big]\ dt\Bigg] + \\
&\quad + \mathbb{E}\Bigg[\int_0^T \Big(\sigma(x)\partial_x f(t, x)\Big)\, dW(t)\Bigg] \\
&= \mathbb{E}\Bigg[\int_0^T \Big[\ \partial_t f(t, x) + \mu(x)\partial_x f(t, x) + \frac{1}{2}\sigma^2(x)\partial_{xx} f(t, x)\Big]\ dt\Bigg] = \\
&\quad \text{(because the second integral is an Itô integral)} \\
&= \int_{-\infty}^{\infty}\int_0^T \Big[\ \partial_t f(t, x) + \mu(x)\partial_x f(t, x) + \frac{1}{2}\sigma^2(x)\partial_{xx} f(t, x)\Big]\ p(t, x)dtdx \\
&= I_1 + I_2 + I_3
\end{aligned}$$

Because function $f(\cdot)$ is arbitrary, but with the properties we introduced, we see that the $\mathbb{E}[df(t)]$ is equal to the sum of three integrals. Performing repeatedly integration by parts we find

$$I_1 = \int_{-\infty}^{\infty} p(t, x)\, f(t, x)\, dx\Big|_{t=0}^{T} - \int_{-\infty}^{\infty}\int_0^T \partial_t p(t, x)\, f(t, x)\, dt\, dx,$$

$$I_2 = \int_0^T \mu(x)\, p(t, x)\, f(t, x)\, dt\Big|_{x=-\infty}^{\infty} - \int_{-\infty}^{\infty}\int_0^T \partial_x\big(\mu(x)p(t, x)\big)\, f(t, x)\, dt\, dx$$

and

$$I_3 = \frac{1}{2}\int_0^T \Big[\, \sigma^2(x)\, p(t,x)\, \partial_x f(t,x) - \partial_x(\sigma^2(x)\, p(t,x))\, f(t,x)\Big]\ dt\Big|_{x=-\infty}^{\infty}$$
$$+\frac{1}{2}\int_{-\infty}^{\infty}\int_0^T \partial_{xx}(\sigma^2(x)p(t,x))\, f(t,x)\, dt\, dx$$

With the boundary conditions introduced then

$$\mathbb{E}\left[\int_0^T df(t)\right] = \int_{-\infty}^{\infty}\int_0^T \Big[-\partial_t p(t,x) - \partial_x(\mu(x)\, p(t,x)) + \frac{1}{2}\partial_{xx}(\sigma^2(x)p(t,x))\Big]\ f(t,x)\, dt\, dx$$

Therefore, for an arbitrary stationary process $\mathbb{E}\Big[\int_0^T df(t)\Big] = 0$ if equation (10.6) holds. □

If we determine the probability distribution $p(t,x)$ then we have an alternative method fo find the moments of the diffusion process. For the case in which the support is $\mathbb{R}$ The mathematical expectation is

$$\mathbb{E}[X(t)] \ = \int_{-\infty}^{\infty} x\, p(t,x)\, dx$$

and the variance is

$$\mathbb{V}[X(t)] \ = \int_{-\infty}^{\infty} \big(x - \mathbb{E}[X(t)]\big)^2\, p(t,x)\, dx.$$

A process is called **ergodic** if the asymptotic probability distribution is time independent

$$p^*(x) = \lim_{t\to\infty} p(t,x).$$

This implies that the moments are asymptotically constants

$$\lim_{t\to\infty}\ \mathbb{E}[X(t)] \ = \int_{-\infty}^{\infty} x\, p(t,x)\, dx = \mu_X^*$$

and the variance is

$$\lim_{t\to\infty} \mathbb{V}[X(t)] \ = \int_{-\infty}^{\infty} \big(x - \mathbb{E}[X(t)]\big)^2\, p(t,x)\, dx = \sigma_X^{*2} > 0$$

Intuition: small or large perturbations do not have large long run effects on the value of X.

Example 1 Let $dX(t) = \sigma dW(t)$ and let $X(0) = x_0$. In order to find the $p(t,x) = \mathbb{P}[X(t) = x|X(0) = x_0]$, we set $p(x,0) = \mathbb{P}[X(0)] = \delta(z - z_0)$ is a Dirac delta function with the distribution mass concentrated at x_0. The initial distribution is a probability distribution because

$$\int_{-\infty}^{\infty} \delta(x - x_0)\, dx = 1.$$

As we have $\mu(x) = 0$ and $\sigma(x) = \sigma$ the adjoint operator is

$$G^*[p](t,x) = \frac{1}{2}\frac{\partial^2(\sigma^2 p(t,x))}{\partial x^2} = \frac{\sigma^2}{2} p_{xx}(t,x).$$

To find the $p(t,x)$ we apply the Fokker-Planck equation and solve the problem with a forward parabolic PDE and an initial condition:

$$\begin{cases} p_t(t,x) = \frac{\sigma^2}{2}\, p_{xx}(t,x), & (t,x) \in \mathbb{R}_+ \times \mathbb{R} \\ p(0,x) = \delta(x - x_0), & t = 0. \end{cases}$$

We saw in chapter 9 that the solution to this problem is

$$p(t,x) = \frac{1}{\sigma\sqrt{2\pi t}}\, e^{-\frac{x^2}{2\sigma^2 t}}, \text{ for } t > 0$$

Example 2 Let $dX(t) = \mu\, dt + \sigma dW(t)$ and let $X(0) = x_0$. As we have $\mu(x) = \mu$ and $\sigma(x) = \sigma$ the adjoint operator is

$$G^*[p](t,x) = -\mu p_x(t,x) + \frac{\sigma^2}{2} p_{xx}(t,x).$$

To find the $p(t,x)$ we apply the Fokker-Planck equation and solve the problem with a forward parabolic PDE and an initial condition:

$$\begin{cases} p_t(t,x) = -\mu p_x(t,x) + \dfrac{\sigma^2}{2}\, p_{xx}(t,x), & (t,x) \in \mathbb{R}_+ \times \mathbb{R} \\ p(0,x) = \delta(x - x_0), & t = 0. \end{cases}$$

We saw in chapter 9 that the solution to this problem is

$$p(t,x) = \int_{-\infty}^{\infty} \delta(s - x_0)\, g(t, x - s)\, ds$$

where

$$g(t,y) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\dfrac{(y-\mu t)^2}{2\sigma^2 t}}.$$

Therefore

$$p(t,x) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\dfrac{(x-\mu t - x_0)^2}{2\sigma^2 t}}. \tag{10.7}$$

10.3.3 Moment equations

An alternative method to determine the dynamics of moments, without resorting to the forward Kolmogorov equation is the following.

Consider the one-dimensional diffusion equation in integral form

$$X(t) = X(0) + \int_0^t \mu(X(s))ds + \int_0^t \sigma(X(s))dW(s). \tag{10.8}$$

Proposition 5. *Consider the diffusion integral form in equation* (10.8) *and assume that* $X(0) = x_0$ *is deterministic. Then*

- *the first moment of the diffusion process is*

$$\mathbb{E}[X(t)] = x_0 + \int_0^t \mathbb{E}\left[\mu(X(s))\right] ds$$

- *the second moment of the diffusion process is*

$$\mathbb{E}[X(t)^2] = x_0^2 + \int_0^t \left(2\mathbb{E}[X(s)\mu(X(s))] + \mathbb{E}[\sigma(X(s))^2]\right) ds$$

- *and the variance is*

$$\mathbb{V}[X(t)] = \int_0^t \left(2\mathbb{E}[X(s)\mu(X(s))] + \mathbb{E}[\sigma(X(s))^2]\right) ds - \int_0^t \mathbb{E}\left[\mu(X(s))\right] ds \left(2x_0 - \int_0^t \mathbb{E}\left[\mu(X(s))\right] ds\right)$$

Proof. As $\sigma(X(t))$ is a non-anticipating random variable, if we use the properties of the Wiener process we have

$$\begin{aligned}\mathbb{E}[X(t)] &= \mathbb{E}[x_0] + \mathbb{E}\left[\int_0^t \mu(X(s))ds\right] + \mathbb{E}\left[\int_0^t \sigma(X(s))dW(s)\right] = \\ &= x_0 + \mathbb{E}\left[\int_0^t \mu(X(s))ds\right] = \\ &= x_0 + \int_0^t \mathbb{E}\left[\mu(X(s))\right] ds\end{aligned}$$

because of the properties of the expected value operator. In order to determine the second moment, $\mathbb{E}[X(t)^2]$, we introduce the variable $Y(t) = X(t)^2$. Using the Itô's formula, as

$$\begin{aligned}dY(t) &= 2X(t)dX(t) + (dX(t))^2 \\ &= 2X(t)(\mu(X(t))dt + \sigma(X(t))dW(t)) + (\mu(X(t))dt + \sigma(X(t))dW(t))^2 = \\ &= \left(2X(t)\mu(X(t)) + \sigma(X(t))^2\right) dt + 2X(t)\sigma(X(t))dW(t)),\end{aligned}$$

then in the integral form $Y(t)$ is

$$Y(t) = x_0^2 + \int_0^t (2X(s)\mu(X(s)) + \sigma(X(s))^2)ds + \int_0^t 2X(s)\sigma(X(s))dW(s)).$$

Then

$$\begin{aligned}\mathbb{E}[X(t)^2] &= x_0^2 + \mathbb{E}\left[\int_0^t (2X(s)\mu(X(s)) + \sigma(X(s))^2)ds\right] + \mathbb{E}\left[\int_0^t 2X(s)\sigma(X(s))dW(s))\right] = \\ &= x_0^2 + \int_0^t \left(2\mathbb{E}[X(s)\mu(X(s))] + \mathbb{E}[\sigma(X(s))^2]\right) ds\end{aligned}$$

The variance is

$$\mathbb{V}[X(t)] = \mathbb{E}[X(t)^2] - (\mathbb{E}[X(t)])^2 =$$

$$= x_0^2 + \int_0^t \left(2\mathbb{E}[X(s)\mu(X(s))] + \mathbb{E}[\sigma(X(s))^2]\right) ds - \left(x_0 + \int_0^t \mathbb{E}\left[\mu(X(s))\right] ds\right)^2 =$$

$$= \int_0^t \left(2\mathbb{E}[X(s)\mu(X(s))] + \mathbb{E}[\sigma(X(s))^2]\right) ds - 2x_0 \int_0^t \mathbb{E}\left[\mu(X(s))\right] ds - \left(\int_0^t \mathbb{E}\left[\mu(X(s))\right] ds\right)^2$$

□

The following properties result

$$\frac{d\mathbb{E}[X(t)]}{dt} = \mathbb{E}[\mu(X(t))].$$

$$\frac{d\mathbb{E}[X(t)^2]}{dt} = 2\mathbb{E}[X(t)\mu(X(t))] + \mathbb{E}[\sigma(X(t))^2]$$

$$\frac{d\mathbb{V}[X(t)]}{dt} = 2\mathbb{E}[X(t)\mu(X(t))] + \mathbb{E}[\sigma(X(t))^2] - \mathbb{E}[\mu(X(t))]^2$$

Example Consider the linear diffusion equation

$$dX(t) = -\gamma X(t)dt + \sigma dW(t)$$

where $X(0) = x_0$, and $\gamma > 0$ and $\sigma > 0$.

The first moment satisfies the ODE

$$\frac{d\mathbb{E}[X(t)]}{dt} = -\gamma\mathbb{E}[X(t)]$$

then the expected value of the process follows the deterministic path

$$\mathbb{E}[X(t)] = x_0 e^{-\gamma t}.$$

The second moment satisfies

$$\frac{d\mathbb{E}[X(t)^2]}{dt} = -2\gamma\mathbb{E}[X(t)^2] + \sigma^2$$

also satisfies the deterministic path

$$\mathbb{E}[X(t)^2] = \frac{\sigma^2}{2\gamma} + \left(x_0^2 - \frac{\sigma^2}{2\gamma}\right) e^{-2\gamma t}.$$

The variance is

$$\begin{aligned}\mathbb{V}[X(t)] &= \mathbb{E}[X(t)^2] \ - \mathbb{E}[X(t)]^2 \ = \\ &= \frac{\sigma^2}{2\gamma} + \left(x_0^2 - \frac{\sigma^2}{2\gamma}\right) e^{-2\gamma t} - \left(x_0 e^{-\gamma t}\right)^2 \\ &= \frac{\sigma^2}{2\gamma}\left(1 - e^{-2\gamma t}\right)\end{aligned}$$

In this case we can determine the asymptotic moments:

$$\lim_{t\to\infty} \mathbb{E}[X(t)] = 0$$

$$\lim_{t\to\infty} \mathbb{V}[X(t)] = \lim_{t\to\infty} \mathbb{E}[X(t)^2] = \frac{\sigma^2}{2\gamma}.$$

This means that the process is asymptotically bounded tends to a limit distribution $N\left(0, \frac{\sigma^2}{2\gamma}\right)$. It is an ergodic process.

10.4 Backward distributions

In some problems, particularly in finance applications, we may be interested in determining the distribution dynamics such that a terminal condition is observed. We continue to assume that a diffusion process

$$dX(t) = \mu(X(t))\, dt + \sigma(X(t))\, dW(t).$$

First, we introduce the concept of a generator of a diffusion

10.4.1 Generator of a diffusion

Definition: Let $f(X(t))$ be a smooth function and let $X(t) = x$. The **infinitesimal generator of** $f(X)$ is a function $G(t,x)[f]$,

$$\begin{aligned}G(t,x)[f] &= \frac{d\mathbb{E}[f(X(t))|X(t)=x]}{dt} \ = \\ &= \lim_{\Delta t\to 0} \frac{\mathbb{E}[f(X(t+\Delta t))|X(t)=x] \ - f(x)}{\Delta t} = \\ &= \frac{\mathbb{E}[df(X(t))|X(t)=x]}{dt}\end{aligned}$$

The generator is defined for every time, t, and is conditional on the realization value at time t, x, that is $X(t) = x$.

The **generator of a function** $f(X)$ **of the diffusion**,

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dW(t)$$

conditional on $X(t) = x$ is the function

$$G(t,x)[f] \;=\; f_x(x)\mu(x) + \frac{1}{2}\sigma(x)^2 f_{xx}(x),\; t \geq 0,$$

We can prove this by just using the Itô's formula.

The generator of a diffusion (over an Itô process), for a differentiable function of a diffusion, allows us to find a directional derivative of f averaged over the paths generated by the diffusion.

10.4.2 Kolmogorov backward equation

The Kolmogorov backward equation allows for the determination of the probability, at time t, conditional on the observable state of the process $X(t) = x$, that the value of the process will belong to a target set ϕ_T at time $T > t$.

We denote the hitting probability by $q(t,x)$

$$q(t,x) = \mathbb{P}[X(T) \in \Phi_T | X(t) = x],$$

where $X(t)$ follows a diffusion process.Then it satisfies

$$q_t(t,x) + G(t,x)[q] = 0.$$

The equation is called **Kolmogorov backward equation**

$$q_t(t,x) = -G(t,x)[q] = -q_x(t,x)\mu(x) - \frac{1}{2}\sigma(x)^2 q_{xx}(t,x)$$

which we want to solve together with with the terminal condition

$$q(T,x) = \begin{cases} \zeta(x) & \text{if } X(T) = x \in \phi_T \\ 0 & \text{if } X(T) \notin \phi_T. \end{cases}$$

Using the Feynman-Kac the probability satisfies

$$\begin{aligned} q(t,x) = \mathbb{P}[X(T) \in \Phi_T | X(t) = x] \; &= \\ = \mathbb{E}[q(T, x(T)) | X(t) = x] &= \end{aligned}$$

Example Let $dX(t) = \sigma dW(t)$ and let $q(T,x) = x^2$. The distribution for $t < T$ follows the PDE

$$q_t(t,x) = -\frac{\sigma^2}{2}\, q_{xx}(t,x),\; 0 < t < T$$

From the Feynman-Kac formula

$$q(t,x) = \mathbb{E}[X(T)^2]$$

We can find $q(t,x)$ by solving the parabolic PDE or by using the Feynman-Kac formula.

Following the second course, we know that the solution of the SDE $dX(t) = \sigma dW(t)$ is

$$X(T) = x + \sigma \int_t^T dW(s) = x\sigma(W(T) - W(t)), \text{ for } T > t,$$

because $W(T) = W(t) + \int_t^T dW(s)$. Computing the moments, we have

$$\mathbb{E}[X(T)] = x, \mathbb{E}[X(T)^2] = \sigma^2(T-t) + x^2$$

Then

$$q(t,x) = \mathbb{E}[X(T)^2] = \sigma^2(T-t) + x^2.$$

If we solve the problem, i.e., a well-posed backward parabolic PDE,

$$\begin{cases} q_t(t,x) = -\frac{\sigma^2}{2}\, q_{xx}(t,x), & 0 < t < T \\ q(t,x) = x^2, & t = T \end{cases}$$

we would reach the same solution.

10.4.3 The Feynman-Kac formula

The Feynman-Kac formula allows us to determine the probability distribution, at time $0 < t < T$, conditional on a known terminal distribution, at time T, for the realization of a diffusion process $(X(t))_{t\in[0,T]}$, when there is a discount factor with discount rate $f(X(t))$.

Let $v(t,x)$ be the probability at time t for a realization $X(t) = x$. Assume that the function $v(t,x)$ is the solution for the partial differential equation boundary value problem

$$\begin{cases} v_t(t,x) = -G(t,x)[v] + v(t,x)f(x), & 0 < t \le T \\ v(T, X(T)), & T \end{cases} \tag{10.9}$$

where $v(T, X(T))$ is known, $f(.)$ is a known function and

$$G(t,x)[v] = v_x(x)\mu(x) + \frac{1}{2}\sigma(x)^2 v_{xx}(x)$$

is the infinitesimal generator of $v(.)$.

Proposition 6. *The solution to the PDE problem* (10.9) *is the* ***Feynman-Kac*** *formula:*

$$v(t,x) = \mathbb{E}\left[\ v(T, X(T))e^{-\int_t^T f(X(s))ds} | X(t) = x\right].$$

Then $v(t,x)$ is the present value of a terminal value $v(T, X(T))$ where the discount rate if $f(X(t))$.

Proof. Write

$$V(t, X(t)) = v(t, X(t))H(t)$$

where $H(t) \equiv e^{-Z(t)} = e^{-\int_s^t f(X(\tau))d\tau}$. As

$$\begin{aligned} dH(t) &= -Z(t)e^{-Z(t)}dZ(t) + \frac{1}{2}Z(t)^2e^{-Z(t)}(dZ(t))^2 = \\ &= -H(t)dZ(t) + \frac{1}{2}Z(t)H(t)(dZ(t))^2 \end{aligned}$$

But because $dZ(t) = f(X(t))dt$ we find, using Itô's rule ,

$$dH(t) = -H(t)f(X(t))dt.$$

Using Itô's formula we obtain

$$\begin{aligned} dv(t, X(t)) &= v_t(t, X(t))dt + v_x(t, X(t))dX(t) + \frac{1}{2}v_{xx}(t, X(t))(dX(t))^2 = \\ &= \left(v_t(t, X(t)) + v_x(t, X(t))\mu(X(t)) + \frac{1}{2}v_{xx}(t, X(t))\sigma(X(t))^2\right)dt + (v_x(t, X(t))\sigma(X(t)))\,dW(t) = \\ &= v(t, X(t))f(X(t))dt + v_x(t, X(t))\sigma(X(t))dW(t) \end{aligned}$$

if we use the PDE in problem (10.9). Then, using the product rule, the previous derivations and Itô's multiplication rules, writing $v(t) = v(t, X(t))$ and $f(t) = f(X(t))$

$$\begin{aligned} dV(t) &= H(t)dv(t) + v(t)dH(t) + dv(t)dH(t) = \\ &= H(t)\left(v(t)f(t)dt + v_x(t)\sigma(t)dW(t)\right) - v(t)H(t)f(t)dt + 0 = \\ &= H(t)v_x(t)\sigma(t)dW(t). \end{aligned}$$

Integrating forward from t, yields

$$V(T) = V(t) + \int_t^T dV(s) = V(X(t)) + \int_t^T e^{-\int_t^s f(X(\tau))d\tau}v_x(s, X(s))\sigma(X(s))dW(s)$$

the initial value plus an Itô's integral. Therefore, the expected value conditional on $X(t) = x$ is

$$\mathbb{E}\left[V(T)|X(t) = x\right] \;= \mathbb{E}\left[V(t)|X(t) = x\right]$$

Seeing $v(t, x)$ as an unconditional expected value $v(t, x) = \mathbb{E}[V(X(t))|X(t) = x]$ and using the expression for $V(T) = v(T, X(T))H(T)$ we have the Feinman-Kac formula. □

10.5 The linear diffusion equation

We apply some of the previous results to obtain explicit solutions of linear scalar Itô stochastic differential equation, which has the general form

$$dX(t) = \left(\mu_0 + \mu_1\, X(t)\right)dt + \left(\sigma_0 + \sigma_1\, X(t)\right)dW(t). \tag{10.10}$$

We will present closed-form solutions for several versions this equation, and characterize their sample path statistical properties and some discussion of its geometrical content.

10.5.1 Brownian motion

The Brownian motion is usual name of a process $\left(X(t), t \in \mathbb{R}_+\right)$ generated by the Itô SDE

$$dX = \mu dt + \sigma dW(t),\ t \in \mathbb{R}_+ \tag{10.11}$$

with $X(0) = x_0 \in \mathbb{R}$ and $\sigma > 0$. This is a special case of equation (10.10) with $\mu_1 = \sigma_1 = 0$ and $\mu_0 = \mu$ and $\sigma_0 = \sigma$.

The solution of equation (10.11), given $X(0) = x_0$ is

$$X(t) = x_0 + \mu\, t + \sigma W(t),\, t \in \mathbb{R}_+.$$

To prove this, writing $X(t)$ in the integral form

$$\begin{aligned} X(t) &= X(0) + \int_0^t dX(s) \\ &= x_0 + \int_0^t \mu\, ds + \int_0^t \sigma dW(s) \\ &= \phi + \mu\, t + \sigma\left(W(t) - W(0)\right) \\ &= \phi + \mu\, t + \sigma\, W(t) \end{aligned}$$

because, form the properties of the Wiener process, $W(0) = 0$.

Figure 10.2 presents one sample path in panel (a) and 100 sample paths for the case in which $\mu = -0.5$ and $\sigma = 1$.

The probability distribution is given by equation (10.7)

$$p(t, x) = \frac{1}{\sigma\sqrt{2\pi t}}\, e^{-\frac{(x - x_0 - \mu t)^2}{2\sigma^2 t}},\ (t, x) \in \mathbb{R}_+ \times \mathbb{R}.$$

which implies that the first and second moments are

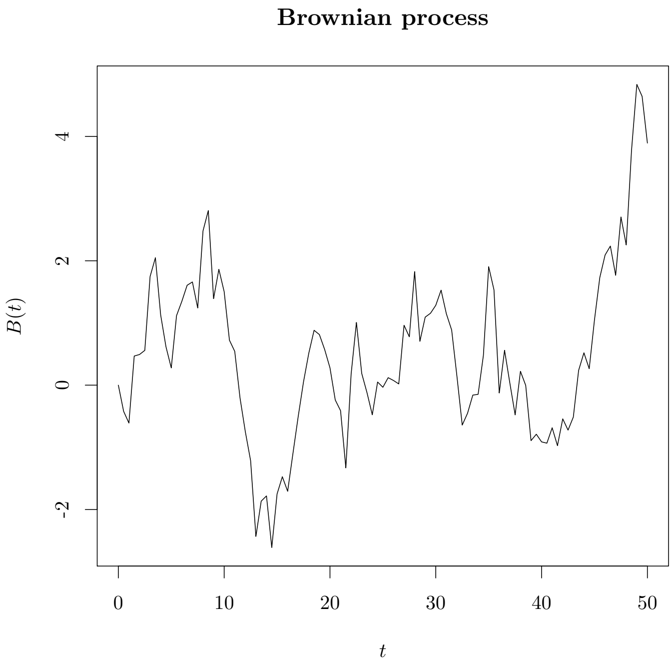
$$\begin{aligned} \mathbb{E}[X(t)] &= \int_{-\infty}^{\infty} x\, p(t, x) dx = x_0 + \mu\, t,\, t \in \mathbb{R}_+, \\ \mathbb{V}[X(t)] &= \int_{-\infty}^{\infty} \left(x - \mathbb{E}[X(t)]\ \right)^2 p(t, x) dx = \sigma^2\, t,\, t \in \mathbb{R}_+. \end{aligned}$$

We observe that the process is not ergodic, because

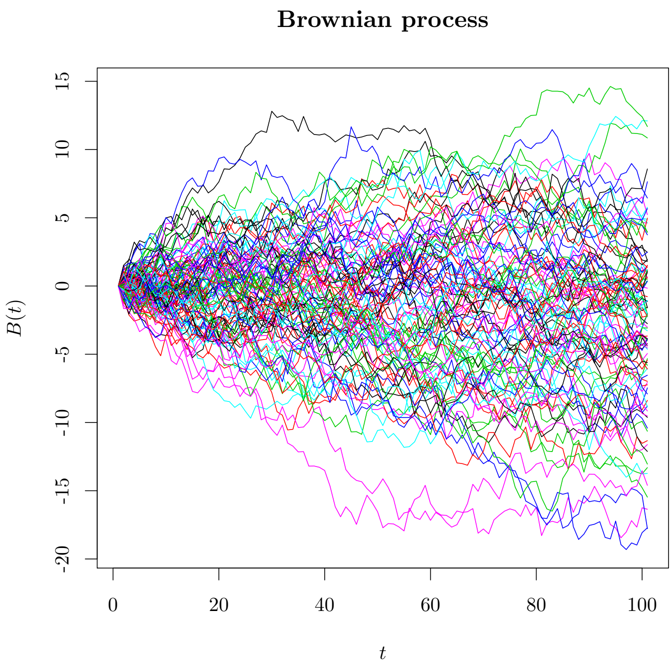
$$\lim_{t \to \infty} \mathbb{E}[X(t)] = \lim_{t \to \infty} \mathbb{V}[X(t)] = \pm\infty$$

if $\mu \neq 0$ and $\sigma \neq 0$.

Observe that the solution of the squeleton $\dfrac{dx(t)}{dt} = \mu$, given x_0 is $x(t) = x_0 + \mu\, t$.



(a) One replication



(b) 100 replications

Figure 10.2: Sample path for the Brownian process for $\mu = -0.5$ and $\sigma = 1$.

10.5.2 Geometric Brownian motion

The geometric Brownian motion is usual name of a process $\big(X(t), t \in \mathbb{R}_+\big)$ generated by the Itô SDE

$$dX(t) = \mu\, X(t)dt + \sigma\, X(t)dW(t),\ t \in \mathbb{R}_+,$$

where $X(0) = x_0$ with $\mathbb{P}[X(0) = x_0] = 1$. This is a special case of equation (10.10) with $\mu_0 = \sigma_0 = 0$ and $\mu_1 = \mu$ and $\sigma_1 = \sigma$.

The explicit solution is

$$X(t) = x_0 e^{\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma\, W(t)},\ t \in \mathbb{R}_+. \tag{10.12}$$

To prove this we define $Y(t) = \ln X(t)$. Using Itô's formula

$$\begin{aligned} dY(t) &= \frac{1}{X(t)} dX(t) + \frac{1}{2}\left(-\frac{1}{X(t)^2}\right)(dX(t))^2 = \\ &= \frac{dX(t)}{X(t)} - \frac{\sigma^2}{2} dt = \\ &= (\mu - \frac{\sigma^2}{2})dt + \sigma dW(t) \end{aligned}$$

Then,

$$\begin{aligned} Y(t) &= y(0) + \int_0^t dY(s) \\ &= y(0) + \int_0^t (\mu - \frac{\sigma^2}{2})\, ds + \int_0^t \sigma\, dW(s) \\ &= y(0) + (\mu - \frac{\sigma^2}{2})t + \sigma\, W(t) \end{aligned}$$

Therefore,

$$\ln X(t) = \ln x_0 + \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W(t)$$

and taking exponential we arrive at equation (10.12).

By using the Kolmogorov forward equation (or Fokker-Planck) we find the probability distribution $p(t, x) = \mathbb{P}[X(t) < x]$ given $X(0) = x_0$ solves the problem

$$\begin{cases} \dfrac{\partial}{\partial t} p(t, x) = -G(t, x)[p] = -\dfrac{\partial}{\partial x}\left(\mu x p(t, x)\right) + \dfrac{1}{2}\dfrac{\partial^2}{\partial x^2}\left(\sigma x p(t, x)\right) \\ p(x_0, 0) = \delta(x - x_0) \end{cases}$$

The solution to this problem is (see the solution of the general linear parabolic equation in chapter 9)

$$p(t, x) = \frac{1}{x\sigma\sqrt{2\pi t}} e^{-\dfrac{\left(\ln\left(x/x0\right) - \left(\mu - \frac{1}{2}\sigma^2\right) t\ \right)^2}{2\sigma^2 t}}$$

The linear diffusion has the moments

$$\begin{aligned}\mathbb{E}[X(t)] &= x_0 e^{\mu t},\ t \in \mathbb{R}_+,\\ \mathbb{V}[X(t)] &= x_0^2 e^{2\mu t}\ \left(e^{\sigma^2 t} - 1\right),\ t \in [0, \infty)\end{aligned}$$

Properties In Figure 10.3 we plot one sample path and several sample paths for the linear diffusion equation where $\mu < 0$ and $\sigma > 0$ and in Figure 10.4 for the case in which $\mu > 0$. We see that in the first case the paths converge to $\lim_{t\to\infty} X(t) = 0$ and in the second case they diverge.

From the moment expressions, we see that:

- if $\mu < 0$, for any $\sigma \neq 0$, then $\lim_{t\to\infty} \mathbb{E}[X(t)] = \lim_{t\to\infty} \mathbb{V}[X(t)] = 0$
- if $\mu > 0$, for any $\sigma \neq 0$, $\lim_{t\to\infty} \mathbb{E}[X(t)] = \text{sign}(x_0)\infty$ and $\lim_{t\to\infty} \mathbb{V}[X(t)] = \infty$.

In the first case, i.e., when $\mu < 0$ the steady state of the skelleton $\dfrac{dx(t)}{dt} = \mu x(t)$, that is $X = x = 0$ is an **absorbing state**, meaning that, although the model is stochastic, all the trajectories converge to a (measure zero) point.

10.5.3 Ornstein-Uhlenback process

An Ornstein-Uhlenback process $(X(t), t \in \mathbb{R}_+)$ is generated by solution to the Itô SDE

$$dX = \theta\,(\mu - X)\,dt + \sigma dW(t) \tag{10.13}$$

where $X(0) = x_0$. This is a special case of equation (10.10) with $\mu_0 = \theta\,\mu$, $\mu_1 = -\theta$, $\sigma_0 = \sigma$ and $\sigma_1 = 0$.

The solution is

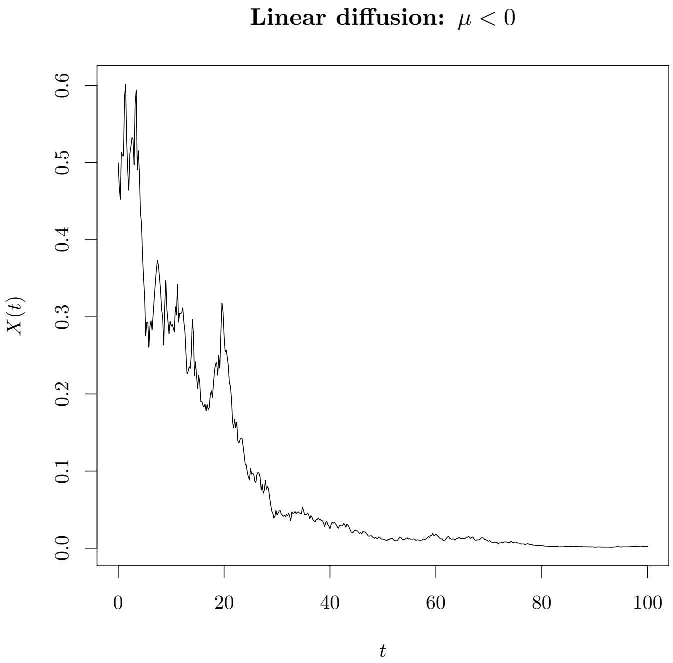
$$X(t) = \mu + (x_0 - \mu)e^{-\theta t} + \sigma \int_0^t e^{-\theta(t-s)}\ dW(s)$$

To prove this, we introduce the change in variables $Y(t) = X(t)\,e^{\theta t}$. Itô's formula yields

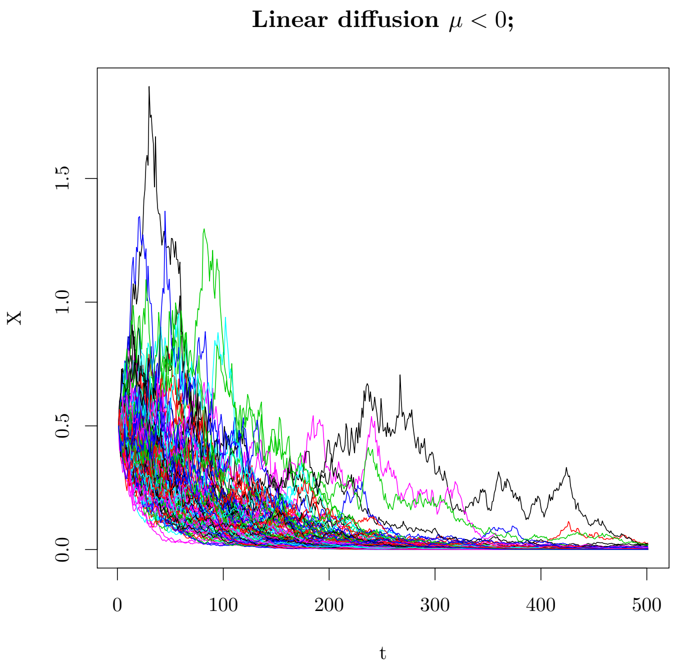
$$\begin{aligned}dY(t) &= \theta\,X(t)\,e^{\theta t}\,dt + e^{\theta t}\,dX(t)\\ &= \theta\,X(t)\,e^{\theta t}\,dt + e^{\theta t}\ (\theta\,(\mu - X(t))\ dt + \sigma\,dW(t))\\ &= e^{\theta t}\,(\theta\mu dt + \sigma dW(t))\,.\end{aligned}$$

Integrating on time we have

$$Y(t) = y_0 + y_0 + \theta\mu \int_0^t e^{\theta s} ds + \sigma \int_0^t e^{\theta s} dW(s).$$

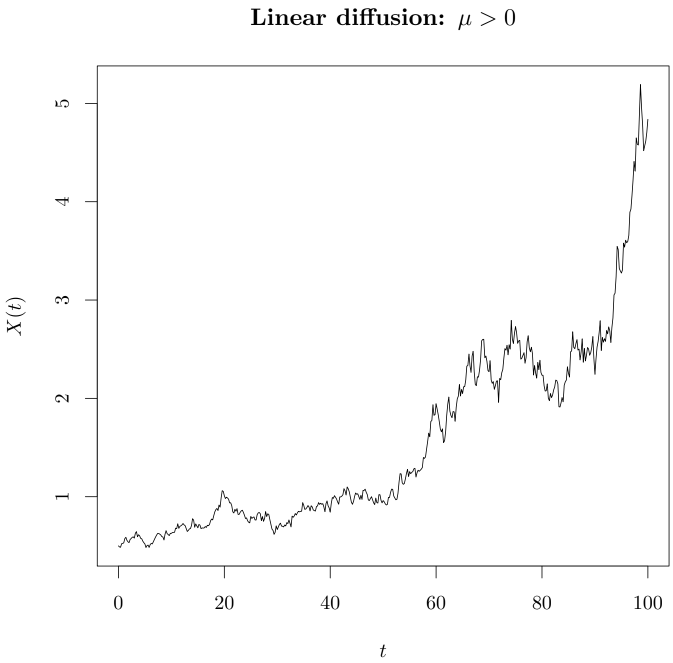


(a) One replication

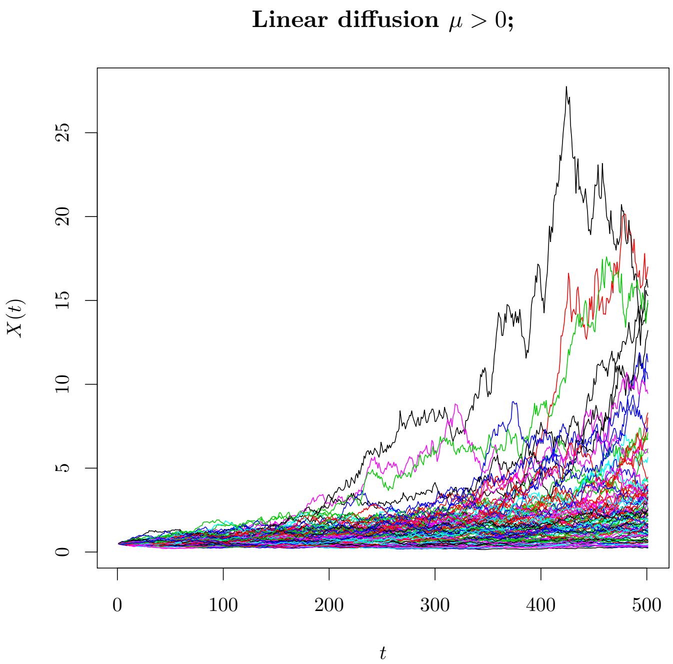


(b) 100 replications

Figure 10.3: Sample paths for the linear diffusion process with $\mu < 0$



(a) One replication



(b) 100 replications

Figure 10.4: Sample paths for the linear diffusion process with $\mu > 0$

Transforming back to the original variable, by making $X(t) = e^{-\theta t}Y(t)$ and $x_0 = y_0$, we obtain the solution to the Itô SDE (10.13)

$$\begin{aligned} X(t) &= e^{-\theta t}\left(y_0 + \theta\mu\int_0^t e^{\theta s}ds + \sigma\int_0^t e^{\theta s}dW(s)\right) \\ &= x_0 e^{-\theta t} + \mu e^{-\theta t}\left(e^{\theta t} - 1\right) + \sigma\int_0^t e^{-\theta(t-s)}dW(s). \end{aligned}$$

By using the Kolmogorov forward equation (or Fokker-Planck) we find the probability distribution $p(t,x) = \mathbb{P}[X(t) < x]$ given $X(0) = x_0$ solves the problem

$$\begin{cases} \frac{\partial}{\partial t}p(t,x) = -\frac{\partial}{\partial x}\Big(\theta\left(\mu - x\right)p(t,x)\Big) + \frac{1}{2}\,\frac{\partial^2}{\partial x^2}\Big(\sigma p(t,x)\Big) \\ p(x_0, 0) = \delta(x - x_0) \end{cases}$$

The solution to this problem is (see the solution of the general linear parabolic equation in chapter 9)

$$p(t,x) = \left(2\pi\,\frac{\sigma^2}{\theta}\left(1 - e^{-2\theta t}\right)\right)^{-\frac{1}{2}} e^{-\frac{\left(x-\mu-(x_0-\mu)\,e^{-\theta t}\right)^2}{2\frac{\sigma^2}{\theta}\left(1-e^{-2\theta t}\right)}}, \; (t,x) \in \mathbb{R}_+ \times \mathbb{R}$$

Therefore, the conditional expected value and variance, for $X(0) = x_0$ are

$$\mathbb{E}^{x_0}\left[X(t)\right] = \mathbb{E}\left[\,X(t)|X(0) = x_0\right] = \mu + (x_0 - \mu)e^{-\theta t}$$

and

$$\mathbb{V}^{x_0}\left[\,X(t)\right] = \mathbb{V}\left[\,X(t)|X(0) = x_0\right] = \frac{\sigma^2}{2\theta}\left(1 - e^{-2\theta t}\right).$$

The properties of the sample paths and of the statistics depend on the sign of θ. Again, assuming that $\sigma \neq 0$ we have the following cases:

- if $\theta > 0$ then the process is ergodic

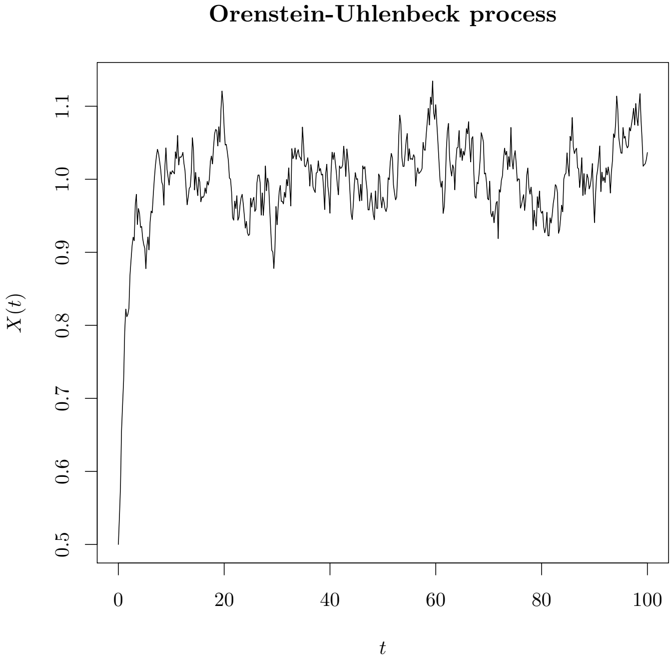
$$\begin{aligned} \lim_{t\to\infty}\ \mathbb{E}^{x_0}\left[\,X(t)\right] &= \mu \\ \lim_{t\to\infty}\ \mathbb{V}^{x_0}\left[\,X(t)\right] &= \frac{\sigma^2}{2\theta} \end{aligned}$$

  and it is asymptotically Gaussian, because

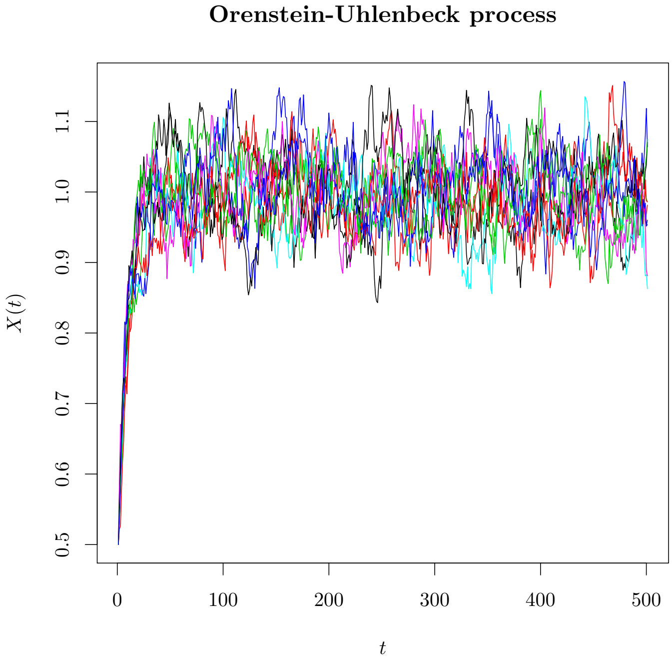
$$\lim_{t\to\infty} X(t) \sim N\left(\mu, \frac{\sigma^2}{2\theta}\right);$$

- if $\theta < 0$ then $\lim_{t\to\infty}\ \mathbb{E}^{x_0}\left[\,X(t)\right] = (x_0 - \mu)\,\infty$ and $\lim_{t\to\infty}\ \mathbb{V}^{x_0}\left[\,X(t)\right] = \infty$

Sample paths for the case $\theta > 0$ are illustrated in figure 10.5



(a) One replication



(b) 100 replications

Figure 10.5: Sample paths for Ornstein-Uhlenbeck process for $\theta > 0$ and $\mu = 1$

10.5.4 The linear diffusion SDE

Now consider equation

The general linear Itô-SDE (10.10) with $X(0) = x_0$.

It can be proved that the explicit solution is

$$X(t) = \Phi(t)\left(x_0 + (\mu_0 - \sigma_0\,\sigma_1)\int_0^t \Phi(s)^{-1}\,ds + \sigma_0\int_0^t \Phi(s)^{-1}\,dW(s)\right)$$

where $\Phi(t)$ is the solution of the geometric Brownian motion

$$d\Phi(t) = \mu_1\,\Phi(t)dt + \sigma_1\,\Phi(t)dW(t)$$

and $\Phi(0) = 1$.

Exercise: prove this. Hint conjecture that $X(t) = \Phi(t)\,Y(t)$, where $\Phi(t)$ follows the geometric Brownian motion. Use the Itô formula to derive $dX(t)$. Match with equation (10.10) to find the process $dY(t)$.

The conditional probability $p(t,x) = \mathbb{P}[X(t) = x|X(0) = x_0]$ is the solution of the FPK equation

$$\begin{cases}\partial_t p(t,x) = -\partial_x\Big((\mu_0 + \mu_1\,x)\,p(t,x)\Big) + \frac{1}{2}\,\partial_{xx}\Big((\sigma_0 + \sigma_1\,x)\,p(t,x)\Big),\ (t,x) \in \mathbb{R}_+ \times \mathbb{R}\\ p(0,x) = \delta(x - x_0),\ (t,x) \in \{t = 0\}\ \times \mathbb{R}\end{cases}$$

It can be proved that the conditional moments are

$$\mathbb{E}[X(t)] = -\frac{\mu_0}{\mu_1} + e^{\mu_1 t}\left(x_0 + \frac{\mu_0}{\mu_1}\right),$$

and

$$\begin{aligned}\mathbb{V}[X(t)] = &-\frac{(\mu_1\sigma_0 - \mu_0\sigma_1)^2}{\mu_1^2\,(2\mu_1 + \sigma_1^2)} + \frac{(\mu_0 + \mu_1 x_0)\,e^{\mu_1 t}}{\mu_1^2}\left(e^{\mu_1 t}(\mu_0 + \mu_1 x_0) + 2\frac{\sigma_1(\mu_0\sigma_1 - \mu_1\sigma_0)}{\mu_1 + \sigma_1^2}\right) +\\ &+ \frac{e^{(2\mu_1 + \sigma_1^2)\,t}}{(\mu_1 + \sigma_1^2)(2\mu_1 + \sigma_1^2)}\Big(2\mu_0(\mu_0 + \sigma_0\sigma_1) + \sigma_0^2(\mu_1 + \sigma_1^2) + 2(x_0 + \mu_0)\sigma_0\sigma_1(2\mu_1 + \sigma_1^2) +\\ &+ x_0^2(\mu_1 + \sigma_1^2)(2\mu_1 + \sigma_1^2)\Big)\end{aligned}$$

If $\mu_1 < 0$

$$\lim_{t\to\infty}\ \mathbb{E}[X(t)] = -\frac{\mu_0}{\mu_1}$$

however the process is ergodic if in addition $\mu_1 + \sigma_1^2 < 0$, which implies $2\,\mu_1 + \sigma_1^2 < 0$ and

$$\lim_{t\to\infty}\ \mathbb{V}[X(t)] = -\frac{(\mu_1\,\sigma_0 - \mu_1\sigma_1)^2}{\mu_1^2\,(2\,\mu_1 + \sigma_1^2)} > 0.$$

10.5.5 Summing up

From the perspective of the asymptotic dynamics, the following behaviors can be expected from a linear Itô-SDE

1. if $\mu_1 + \sigma_1^2 < 0$ the process is ergodic and tends asymptotically to a Gaussian distribution $N\Big(-\frac{\mu_0}{\mu_1}, -\frac{(\mu_1\,\sigma_0 - \mu_1\sigma_1)^2}{\mu_1^2\,(2\,\mu_1+\sigma_1^2)}\Big)$, which means that the steady state is a distribution

2. if $\mu_1 + \sigma_1^2 < 0$ and $\mu_1\,\sigma_0 - \mu_1\sigma_1 = 0$ the dynamic tends to absorbing state $x = -\frac{\mu_0}{\mu_1}$ which is a deterministic steady state

3. if $\mu_1 + \sigma_1^2 \geq 0$ the equation tends to an unbounded distribution in which both moments are asymptotically unbounded.

10.6 The general linear SDE: the non-autonomous case

The general linear SDE has the form

$$dX = (a(t)X + u(t))dt + (b(t)X + v(t))dW(t)$$

where $X(0) = x_0$ with $\mathbb{P}[X(0) = x_0] = 1$, has the explicit solution

$$X(t) = \Phi(t)\left(x_0 + \int_0^t \Phi(s)^{-1}(u(s) - b(s)v(s))ds + \int_0^t \Phi(s)^{-1}v(s)dW(s)\right)$$

where $\Phi(t)$ is the solution of

$$d\Phi(t) = a(t)\Phi(t)dt + b(t)\Phi(t)dW(t)$$

and $\Phi(0) = 1$

10.7 Economic applications

10.7.1 The Solow stochastic growth model

Several papers, starting with Merton (1975) and Bourguignon (1974) (see (Malliaris and Brock, 1982, ch. 3)) study the stochastic Solow model.

Assume that population follows the SDE

$$dL(t) = \mu L dt + \sigma L dW(t)$$

where μ is the rate mean rate of growth of population and σ its variance.

The equilibrium equation for the product market is

$$\frac{dK(t)}{dt} = sF(K, L)$$

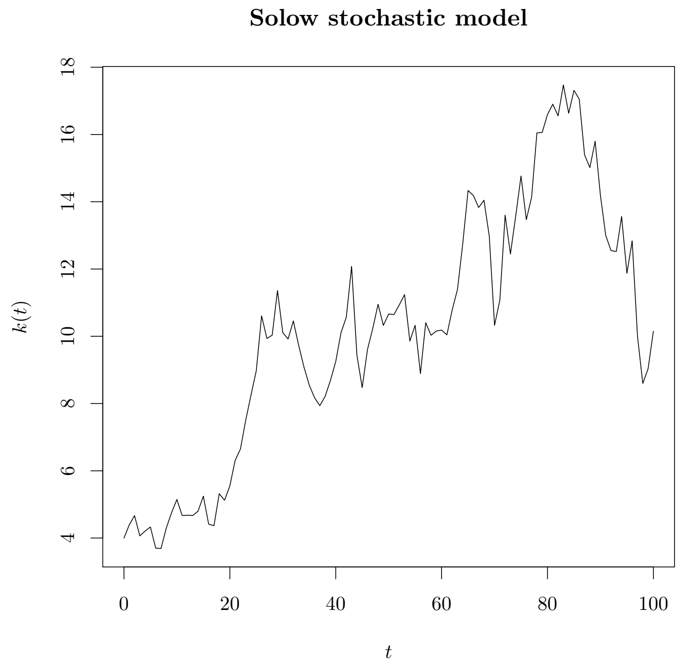


Figure 10.6: Sample path for the capital intensity: $s = 0.1$, $\alpha = 0.3$, $\mu = 0.01$, $\sigma = 0.1$

where $F(.)$ has the neoclassical properties (increasing, concave, homogeneous of degree one and Inada). We define the capitai intensity as usual $k \equiv K/L$. Then $F(K, L) = Lf(k)$. and

$$dK = sLf(k)dt$$

We can write $k = \kappa(K/L)$. Then $\kappa_K = 1/L$, $\kappa_L = -K/(L^2)$, $\kappa_{KK} = 0$, $\kappa_{KL} = \kappa_{LK} = -1/(L^2)$ and $\kappa_{LL} = 2K/(L^3)$. Then, applying the Itô's Lemma

$$\begin{aligned} dk &= \kappa_K dK + \kappa_L dL + \frac{1}{2}\left(\kappa_{KK}(dK)^2 + 2\kappa_{KL} dK dL + \kappa_{LL}(dL)^2\right) \\ &= sf(k)dt - k(\mu dt + \sigma dW) + \frac{1}{2}\left(-sf(k)dt(\mu dt + \sigma dW) + 2k(\mu dt + \sigma dW)^2\right) \end{aligned}$$

Using $(dt)^2 = dtdW(t) = 0$ and $(dW(t))^2 = dt$ then we get the SDE

$$dk = \left(sf(k) - (\mu - \sigma^2)k\right) dt - k\sigma dW(t) \tag{10.14}$$

For a Cobb-Douglas function we have

$$dk = \left(sk^\alpha - (\mu - \sigma^2)k\right) dt - k\sigma dW(t)$$

where $0 < \alpha < 1$. Figures 10.6 and 10.7 present one replication and 100 replications for this equation for a deterministic initial value $k(0) = k_0$

The stationary distribution for the capital intensity is (see Merton (1975) and (Malliaris and Brock, 1982, p. 146)

$$p(k) = \frac{m}{\sigma^2 k^2} \exp\left(2\int^k \frac{sf(\xi) - (n - \sigma^2)\xi}{\sigma^2 \xi^2}\, d\xi\right)$$

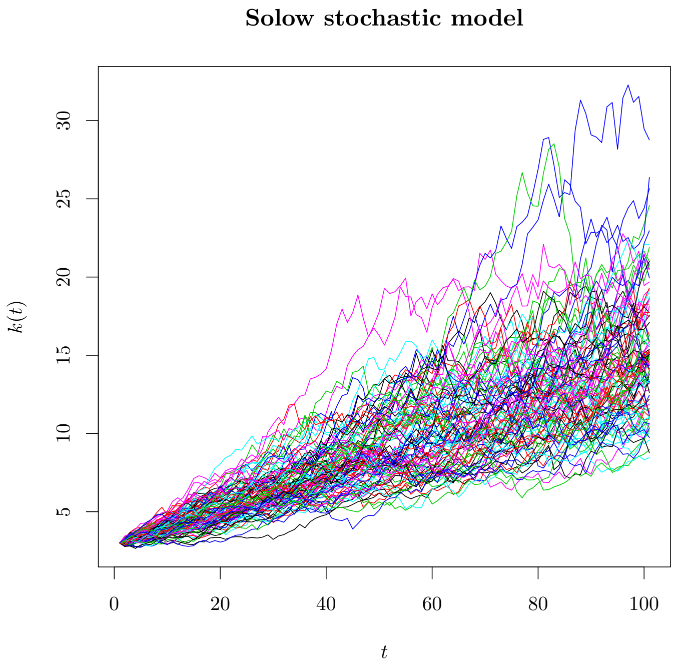


Figure 10.7: Sample paths for the capital intensity: $s = 0.1$, $\alpha = 0.3$, $\mu = 0.01$, $\sigma = 0.1$, 100 replications

where m is chosen such that $\int_0^\infty p(k)dk = 1$. For the Cobb-Douglas case it is

$$p(k) = mk^{-2\mu/\sigma^2} \exp\left(\frac{-2s}{(1-\alpha)\sigma^2} k^{-(1-\alpha)}\right)$$

10.7.2 Derivation of the Black and Scholes (1973) equation

Assume that there are two assets, a risk free asset, with value $B(t)$, following the process

$$dB(t) = rB(t)dt,$$

and a risky asset, with value $S(t)$, and following the diffusion process

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t).$$

The current prices of both assets, $B(0)$ and $S(0)$ are observed.

An European call option is a contract offering the option (but not the obligation) to buy, at the expiration time $T > 0$, the risky asset at a price K. A purchaser would have an interest to exercise the option only if the price of the risky asset at time T, $S(T)$, is higher than the exercise price. If $K < S(T)$ the purchaser would not exercise the option.

Let $V(S,t)$ be the value of the option on the risky asset at time t, for $0 \le t \le T$. The value of the option at time of the exercise T is dependent of $S(T)$ and is

$$V(S,T) = \max\{\ S(T) - K, 0\}.$$

However, the contract would only be possible if there is a payment at time $t = 0$, otherwise the writer would have no incentive in offering the contract. What would be the price of the option at the moment of the contract, i.e., at time $t = 0$, $V(S,0)$?

Using the Itô's formula we obtain the process for the value of the option

$$\begin{aligned} dV(S,t) &= V_t(S,t)dt + V_s(S,t)dS + \frac{1}{2}V_{ss}(S,t)(dS)^2 = \\ &= V_t(S,t)dt + V_s(S,t)\left(\mu S(t)dt + \sigma S(t)dW(t)\right) + \frac{1}{2}V_{ss}(S,t)\sigma^2 S(t)^2 dt = \\ &= \left(V_t(S,t) + \mu S(t)V_s(S,t) + \frac{1}{2}\sigma^2 S(t)^2 V_{ss}(S,t)\right) + \sigma S(t)V_s(S,t)dW(t). \end{aligned}$$

The market data also allows us to obtain a valuation, if we assume that there are **no arbitrage opportunities**. If the markets are complete, the yields generated by the option can also be generated by the yields of a portfolio composed by the available assets with the same value. We call this portfolio the replicating portfolio.

The replicating portfolio is composed of θ units of the risky asset and $(1-\theta)$ units of the risk free asset such that

$$V^r(B(t), S(t)) = (1-\theta(t))B(t) + \theta(t)S(t), \text{ for every } t \in [0,T]$$

Using the Itô's formula, we have

$$\begin{aligned} dV^r(B(t), S(t)) &= (1-\theta)dB + \theta dS = \\ &= (1-\theta)rB(t)dt + \theta S(t)\left(\mu dt + \sigma dW(t)\right) = \\ &= \left(rV^r(B,S) + (\mu - r)S(t)\right)dt + \theta\sigma S(t)dW(t). \end{aligned}$$

In the absence of arbitrage opportunities we should have $dV(S(t),t) = dV(B(t), S(t)$.

Matching the diffusion and the dispersion components of the two differentials for the option and the replicating portfolio values, yields

$$\begin{cases} \theta\sigma S(t) = \sigma S(t)V_s(S,t) \\ rV^r(B,S) + (\mu - r)S(t) = V_t(S,t) + \mu S(t)V_s(S,t) + \frac{1}{2}\sigma^2 S(t)^2 V_{ss}(S,t) \end{cases}$$

From the first equation we obtain the weight of the risky asset in the replicating portfolio composition

$$\theta(t) = V_s(S,t).$$

After setting $V(S,t) = V^r(B,S)$, we obtain from the second equation the Black and Scholes (1973) PDE,

$$V_t(S,t) = -\frac{\sigma^2}{2}S^2 V_{ss}(S,t) - rSV_s(S,t) + rV(S,t),$$

which is backward semi-linear parabolic PDE.

The value of the option, and in particular its price $V(S,0)$ is the solution of the following option valuation problem:

$$\begin{cases} V_t(S,t) = -\frac{\sigma^2}{2}S^2V_{ss}(S,t) - rSV_s(S,t) + rV(S,t), & (S,t) \in (0,\infty)\times[0,T] \\ V(S,T) = \max\{S-K,0\}, & (S,t)\in(0,\infty)\times\{\ t=T\} \end{cases} \tag{10.15}$$

We show in the PDE chapter that the solution of the option valuation problem is

$$V(S,t) = S\Phi(d_+(t)) - Ke^{-r(T-t)}\Phi(d_-(t)),\ t\in[0,T]$$

where $\Phi(\cdot)$ is the Gaussian distribution function (see the Appendix) and

$$d_{\mp}(t) = \frac{\ln\left(\frac{S(0)}{K}\right) + (T-t)(r \mp \frac{\sigma^2}{2})}{\sigma\sqrt{T-t}}$$

The price of the option is

$$V(S,0) = S(0)\Phi(d_+(0)) - Ke^{-rT}\Phi(d_-(0)),$$

with

$$d_{\mp}(0) = \frac{\ln\left(\frac{S(0)}{K}\right) + T\left(r \mp \frac{\sigma^2}{2}\right)}{\sigma\sqrt{T}}.$$

where $S(0)$ is observable at time $t=0$, K and T are specified in the option contract and r and σ are estimated or conjectured.

10.8 References

- Mathematics of SDE's: Øksendal (2003), Pavliotis (2014)
- Dynamic systems theory and SDE's: Cai and Zhu (2017)
- Numerical analysis of SDE Iacus (2010)
- Application to economics and finance: Malliaris and Brock (1982), Stokey (2009)

Appendix: The Gaussian integral

The gaussian kernel is a function

$$g(x) = e^{-x^2}$$

which has the well known bell shape.

A Gaussian integral is an integral of type

$$\int_{-\infty}^{\infty} h(x)g(x)dx$$

if it is finite (I.e. L^2).

Some properties of the Gaussian integral are:

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi},$$

$$\int_{-\infty}^{\infty} xe^{-x^2} dx = 0,$$

$$\int_{-\infty}^{\infty} |x|e^{-x^2} dx = 1,$$

where $|x| = \sqrt{x^2}$

$$\int_{-\infty}^{\infty} x^2 e^{-x^2} dx = \sqrt{\frac{\pi}{4}}$$

If we introduce a parameter $a > 0$

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$$

$$\int_{-\infty}^{\infty} xe^{-ax^2} dx = 0$$

$$\int_{-\infty}^{\infty} x^2 e^{-ax^2} dx = \frac{1}{a} \sqrt{\frac{\pi}{4a}}$$

Gaussian distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{-\frac{s^2}{2}} ds.$$

Bibliography

Bensoussan, A. (1988). *Perturbation Methods in Optimal Control.* Wiley/Gauthier-Villars.

Bielecki, T. R. and Rutkowski, M. (2004). *Credit Risk: Modeling, Valuation and Hedging.* Springer.

Björk, T. (2004). *Arbitrage Theory in Continuous Time.* Finance. Oxford University Press, 2nd edition.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–659.

Bourguignon, F. (1974). A particular class of continuous-time stochastic growth models. *Journal of Economic Theory*, 9:141–58.

Brock, W. A. and Mirman, L. (1972). Optimal economic growth and uncertainty: the discounted case. *Journal of Economic Theory*, 4:479–513.

Cai, G.-Q. and Zhu, W.-Q. (2017). *Elements of Stochastic Dynamics.* World Scientific.

Cvitanić, J. and Zapatero, F. (2004). *Introduction to the Economics and Mathematics of Financial Markets.* MIT Press.

Duffie, D. (1996). *Dynamic Asset Pricing Theory.* Princeton University Press, Princeton, second edition.

Fleming, W. H. and Rishel, R. W. (1975). *Deterministic and Stochastic Optimal Control.* Springer-Verlag.

Iacus, S. M. (2010). *Simulation and Inference for Stochastic Differential Equations.* Springer.

Itô, K. (1951). On stochastic differential equations. *Memoirs of the American Mathematical Society*, 4:289–302.

Kamien, M. I. and Schwartz, N. L. (1991). *Dynamic optimization, 2nd ed.* North-Holland.

Kushner, H. J. (2014). A partial history of the early development of continuous-time nonlinear stochastic systems theory. *Automatica*, 50:303–334.

Malliaris, A. and Brock, W. (1982). *Stochastic Methods in Economics and Finance.* North-Holland.

Merton, R. (1971). Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory*, 3:373–413.

Merton, R. (1975). An Asymptotic Theory of Growth under Uncertainty. *Review of Economic Studies*, 42:375–93.

Merton, R. (1990). *Continuous Time Finance*. Blackwell.

Øksendal, B. (2003). *Stochastic Differential Equations*. Springer, 6th edition.

Pavliotis, G. A. (2014). *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*. Texts in Applied Mathematics 60. Springer-Verlag New York, 1 edition.

Pham, H. (2009). *Continuous-time Stochastic Control and Optimization with Financial Applications*. Stochastic Modelling and Applied Probability. Springer, 1 edition.

Seierstad, A. (2009). *Stochastic control in discrete and continuous time*. Springer.

Stokey, N. L. (2009). *The Economics of Inaction*. Princeton.

Yong, J. and Zhou, X. Y. (1999). *Stochastic Controls. Hamiltonian Systems and HJB Equations*. Number 43 in Applcations of Mathematics. Stochastic Modelling and Applied Probability. Springer.