Advanced Mathematical Economics

Paulo B. Brito
PhD in Economics: 2021-2022
ISEG
Universidade de Lisboa
pbrito@iseg.ulisboa.pt

Lecture 6
30.11.2022

Contents

Chapter 13

Introduction to stochastic calculus and stochastic differential equations

??

13.1 Introduction

The ordinary differential equation, having time as the independent variable,

$$\dot{y} = f(y(t)), \text{ for } \ t \in \mathbb{R}_+ \tag{13.1}$$

can be extend by introducing a linear random perturbation, $\epsilon(t)$,

$$\dot{Y} = f(Y(t)) + \epsilon(t), \text{ for } \ t \in \mathbb{R}_+ \tag{13.2}$$

where $f(Y(t))$ is the deterministic component (or skeleton), or a mor general non-linear perturbation

$$\dot{Y} = f(Y(t), \epsilon(t)). \tag{13.3}$$

While the solution of (13.1) is a mapping $y : \mathbb{R}_+ \to \mathbb{R}^n$, in the case of equations (13.2) or (13.3), 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, for a particular realization of the random variable, such as as $\epsilon(t) = \epsilon_t$, where ϵ_t is a number, 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 can 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 (linear or non-linear) skeleton 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. In continuous time models applied to economics and finance there are two main ways to introduce a stochastic component[1] which can be seen as the formal counterparts to, the previuosly referred, high and low noise, respectively:

- to model rare events with a local high impact, uncertainty is introduced via a **Poisson process**, $(Q(t)\ )_{t\in \mathrm{T}}$,
$$dY(t) = f(Y(t),t)dt + v(Y(t^-),t^-)\, dQ(t) \tag{13.4}$$
where $Y(t^-) = \lim_{s\uparrow t}\ Y(s)$ and $dQ(t) = 1$ with probability $\lambda\, dt$ and $dQ(t) = 0$ with probability $(1-\lambda)\, dt$, $f(.)$ and $v(.)$ are continuous and differentiable known functions.

- to model frequent events having a local small impact, uncertainty is introduced via a **Wiener process** $(W(t))_{t\in \mathrm{T}}$. The most common model is called **diffusion equation**
$$dY(t) = f(Y(t),t)dt + \sigma(Y(t),t)dW(t) \tag{13.5}$$
where $f(.)$ and $\sigma(.)$ are continuous and differentiable known functions.

The main reason for using the previous formalism is related to the fact that $Y(t)$ is not differentiable in the classic sense in both cases. Before solving any of the previous two equations a specific stochastic calculus has to be developed.

Equation (13.4) is called **stochastic differential equation with jumps** (SDEJ) in the differential form. The equivalent SDEJ in the integral form is

$$Y(t) = Y(0) + \int_0^t f(Y(s),s)ds + \int_0^t v(Y(s^-),s^-)\ dQ(s) \tag{13.6}$$

[1] For a clear discussion see Merton (1982).

where the first integral in the right-hand-side is a Riemmann integral, but the second is a **Poisson integral**. In order to solve and/or characterise SDEJ we have to introduce the properties of the Poisson process and of the Poisson integral. Likewise, equation (13.5) is called **stochastic differential equation** (SDE) in the differential form. The equivalent SDE in the integral form is

$$Y(t) = Y(0) + \int_0^t f(Y(s), s)ds + \int_0^t \sigma(Y(s), s)\, dW(s) \tag{13.7}$$

where the first integral in the right-hand-side is a Riemmann integral, but the second is a **Itô integral**. Stochastic differential equations having both jump and diffusion,

$$dY(t) = f(Y(t), t)dt + \sigma(Y(t), t)dW(t) + v(Y(t^-), t^-)\, dQ(t),$$

also have been dealt in the litarature The most common approach to SDE's view "noise" as generated by a Wiener process and builds upon the Itô process, as in equations (13.5) or (13.7). In the rest of this lecture we will restrain to Itô's SDE's. From this we present the basic linear SDE, the diffusion equation, and study its statistical and stability properties. In order to solve and/or characterise SDE we have to introduce the properties of the Wiener process and of the Itô's integral.

We present a very brief introduction to the Itô's stochastic calculus applied to stochastic differential equation of type (13.5), following an applied and heuristic approach. In particular, we emphasise the connections with ordinary and partial differential equations.

In section 13.2 we define and describe the properties of the Wiener process. In section 13.3 we define and present the properties of the Itô's process and integral. In section 13.4 we present methods for characterizing the solutions to diffusion processes.

13.2 The Wiener process

13.2.1 Stochastic processes: a brief description

A uni-dimensional **stochastic process** can be seen as a flow of random variables $\Big(X(t, \omega)\Big)_{t \in \mathrm{T}}$, for $X(t, \omega) : \mathrm{T} \times \Omega \to \mathbb{R}$, where $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ is a filtered probability space. A filtered probability space is defined by the sample space Ω, by the set of events $\mathcal{F}$ (i.e., the set of all subsets of Ω), a probability measure over events and the filtration $\mathbb{F} = \Big(\mathcal{F}(t)\Big)_{t \in \mathrm{T}}$.

In non-rigorous terms, we can interpret a filtration as the way in which the flow information allows for determining a probability distribution for events taking place over time. A **non-anticipating** process is a process in which the probability associated to a particular event is determined from past events, meaning that we can only ascertain the probability of a future event on the base of past information. An **anticipating process** is a process in which we condition the probability of present events on the occurrence of future events. A stochastic process is **adapted to a filtration** if it has a probability distribution drawn from a particular filtration, that is, depending on the flow of information implicit in the filtration.

In the rest of the lecture we represent a stochastic process by $\Big(X(t)\Big)_{t\in \mathrm{T}}$, where $X(t) : \mathrm{T} \to \mathbb{R}$ represents the **possible realizations** of the process at time t, that is **before** nature (or a pseudo random number generator of a computer) makes a draw at time t. The **realization** of a stochastic process at time t, denoted by $X(t) = x(t)$, is a particular number, $x(t)$, that is observed **after** nature (or the computer) makes a draw.

From now on, we consider non-anticipating processes, or **processes adapted** to a non-anticipating filtration.

13.2.2 Wiener process: definition

There are several ways of characterising the **Wiener process**, also called **standard Brownian motion**. As we mentioned before, the Wiener process, as a stochastic process, can is defined by a particular type of information on the flow of events and the associated probability distribution.

Definition 1. *A* ***Wiener process****, denoted by $(W(t))_{t\geq 0}$, is a stochastic process, where $W : \Omega \times T \to \mathbb{R}$ has the following properties:*

1. *the initial value is equal to* 0 *with probability one:* $\mathbb{P}[W(0) = 0] = 1$ *(also written as $W(0) = 0$ a.s.)* [2]
2. *it has a continuous version: i.e., a randomly generated path is a continuous function of time with probability one (i.e., there can be discontinuous jumps, but they have probability zero of occurring);*
3. *the path increments are independent and have a Gaussian probability distribution with zero mean and variance equal to the temporal increment*

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

The last property implies that the Wiener process is a **Markovian process**. This means that it is a memory-less process.

A **propagator** can be defined as the conditional probability of the process $(W(t))$ having the realization w' at time $t + dt$, given that it had the realization w at time t,

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

If we write $w' = w + dw$ then the **propagator of a Wiener process** is

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

There are several ways of charaterizing stochastic processes. Next we characterize the Wiener process by its **sample path** and **statistical** (or stochastic) properties.

[2] Almost surely.

Figure 13.1 depicts a sample path, in subfigure (a), and 100 draws of the process, in subfigure (b). It illustrates the stochastic properties of the Wiener process: first, every sample path is strongly jagged, second, the distribution displays increasing dispersion over time, although it tends to be located most of the time close to $W = 0$. Furthermore, although sample paths looks like being continuous they are not smooth enough to allow differentiability (i.e,. at every point in time limits do not look like being unique).

We can present a heuristic confirmation of those perceptions.

Sample path properties

Proposition 1. *The Wiener process is not first-order-differentiable in time, in the classic sense.*

Proof. (Heuristic) Let

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

for a given $0 < t < \infty$. If $dt > 0$, then

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

Writing $W(t+dt) - W(t) = X$, and taking into account the definition of the Wiener process,

$$\begin{aligned}
\mathbb{E}[|X|] &= \int_{-\infty}^{\infty} \frac{|x|}{\sqrt{2\pi dt}}\, e^{-\dfrac{x^2}{2dt}} dx = \frac{\sqrt{2dt}}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{|x|}{\sqrt{2dt}} e^{-\dfrac{x^2}{2dt}} \frac{dx}{\sqrt{2dt}} \\
&\text{(setting } y = x/\sqrt{2dt}, \text{ and as } dt > 0) \\
&= \frac{\sqrt{2dt}}{\sqrt{\pi}} \int_{-\infty}^{\infty} |y|\, e^{-y^2}\, dy \\
&\Big(\text{because } \int_{-\infty}^{\infty} |y|\, e^{-y^2}\, dy = 1 \text{ see the Appendix on the Gaussian integral)} \\
&= \sqrt{\frac{2dt}{\pi}}.
\end{aligned}$$

Then

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

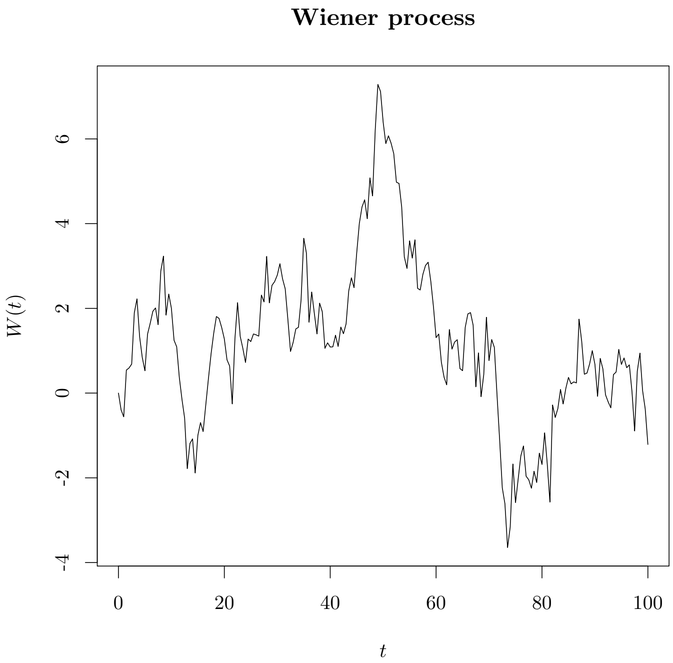
The time derivative is of order $\sqrt{dt}$ meaning that as $\lim_{dt\to\infty} \mathbb{E}\left[\left|\frac{dW(t)}{dt}\right|\right] = \infty$ which means that the sample path of $(W(t))$ is not first-order differentiable in time. □

This means that the derivative

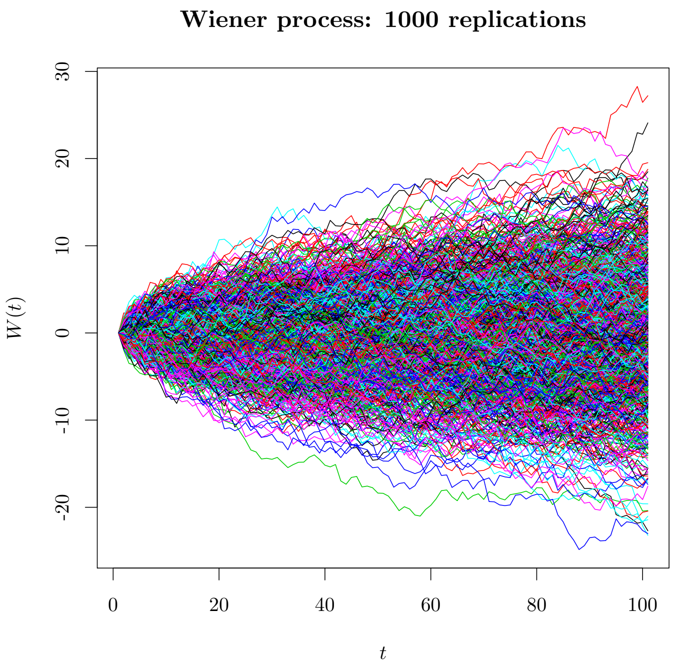
$$\frac{dW(t)}{dt}$$

is not well defined, although the variation $dW(t)$ is well defined. Therefore, from the fundamental theorem of calculus, the process specified in the integral form

$$W(t) = \int_0^t dW(t),\ a.s.$$



(a) One replication



(b) 100 replications

Figure 13.1: Sample paths for the Wiener process

is well defined.

This is the reason why we need a particular calculus to deal with stochastic differential equations which are functions of Wiener processes, as we will see next.

Statistic properties

Looking again to Figure 13.1 we can characterize the statistic properties of the Wiener process. Those 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 has the following statistic properties:*

*1 It is **stationary** in expected value terms, that is*

$$\mathbb{E}[dW(t)] \; = 0, \; \textit{for each} \; t \in \mathrm{T};$$

2 the mathematical expectation of the square variation of the Wiener process is equal to the time increment

$$\mathbb{E}[(dW(t))^2] \; = dt, \; \textit{for each} \; t \in \mathrm{T};$$

3 the variance of the variation is equal to the time increment

$$\mathbb{V}[dW(t)] \; = \mathbb{E}[\; dW(t)^2] \; - \mathbb{E}[\; dW(t)]^2 = dt, \; \textit{for each} \; t \in \mathrm{T}.$$

Proof. Let $dW(t) = dW$, where dW can be seen as a random variable with a $N(0, dt)$ density distribution, and $dt > 0$. Then: (1) the expected value is

$$\begin{aligned}
\mathbb{E}\,[dW] \; &= \int_{-\infty}^{\infty} \frac{dw}{\sqrt{2\pi dt}} e^{-\frac{(dw)^2}{2dt}} \, d(dw) \\
&\text{(changing variables } \; dw = \sqrt{2dt}x \Rightarrow d(dw) = \sqrt{2dt}\, dx) \\
&= \sqrt{\frac{2dt}{\pi}} \int_{-\infty}^{\infty} x e^{-x^2} dx = 0
\end{aligned}$$

from the properties of the Gaussian integral (see the Appendix); (2) the quadratic variation

$(dW(t))^2 = (dW)^2$ has the expected value

$$\begin{aligned}
\mathbb{E}\left[(dW)^2\right] &= \int_{-\infty}^{\infty} \frac{dw^2}{\sqrt{2\pi dt}} e^{-\frac{(dw)^2}{2dt}}\, d(dw) = \\
&\text{(using the same change in variables)} \\
&= \frac{2dt}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2} dx = \\
&\text{(using again the properties of the Gaussian integral)} \\
&= \frac{2dt}{\sqrt{\pi}} \frac{\sqrt{\pi}}{2} = \\
&= dt
\end{aligned};$$

(3) as the he variance of a change $\mathbb{V}\left[dW(t)\right] = \mathbb{E}\left[dW(t)^2\right] - \mathbb{E}\left[dW(t)\right]^2$ and $\mathbb{E}\left[dW(t)\right] = 0$ then $\mathbb{V}\left[dW(t)\right] = \mathbb{E}\left[dW(t)^2\right] = dt$. □

Corollary 1. *Assume that the time variation is positive $dt > 0$.*

1 The expected value for a Wiener process is

$$\mathbb{E}[W(t)] = 0, \text{ for every } t \in \mathrm{T}$$

2 the mathematical expectation of the square of the Wiener process is

$$\mathbb{E}[W(t)^2] = t, \text{ for each } t \in \mathrm{T}$$

3 the variance of the Wiener process is

$$\mathbb{V}[W(t)] = t, \text{ for each } t \in \mathrm{T};$$

4 let $s = dt + t$. Then the covariance of the Wiener process is

$$Cov[W(s), W(t)] = s, \text{ for any } s > t \in \mathrm{T}$$

5 the correlation coefficient of the Wiener process is

$$Corr[W(s), W(t)] = \sqrt{\frac{s}{t}}, \text{ for any } s > t \in \mathrm{T}.$$

Proof. (1) As $W(t) = W(0) + \int_0^t dW(t)$ and $\mathbb{P}[W(0) = 0] = 1$, then $\mathbb{E}[W(t)] = \mathbb{E}\Big[\int_0^t dW(t)\Big] = \int_0^t \mathbb{E}[dW(s)] = 0$, for every $t \in \mathrm{T}$; (2) $W(t)^2 = \int_0^t dW(t)^2$ then $\mathbb{E}[W(t)^2] = \mathbb{E}\Big[\int_0^t dW(t)^2\Big] = \int_0^t \mathbb{E}[dW(s)^2] = \int_0^t ds = t$, for every $t \in \mathrm{T}$; (4) for the covariance, because the increments of the Wiener process are i.i.d

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

□

13.3 The Itô's processes

In equation SDE (**??**) we had the expression

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

which, from the non-differentiability properties of the Wiener process needs to be addressed. Next we present Itô (1951) interpretation of the integral.

Definition 2. *Let $f(t)$ be a bounded function of time and $\big(W(t)\big)_{t\geq 0}$ a Wiener process. We call* ***Itô's integral*** *to*

$$I(t) = \int_0^t f(s)\, dW(s). \tag{13.8}$$

This definition can be extended to functions of type $f(t, \omega)$,where ω is adapted to the Wiener process. 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

$$I(t, w) = \int_0^t f(s, w)dW(s) \tag{13.9}$$

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

The Itô's integral generates an **Itô's process** $(I(s))_{s=0}^t$. It shares some properties with the classic (Riemannian) integral, but same other properties which are related to the fact that it integrates a Wiener stochastic process.

Properties of the Itô's integral

- 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 $0 < t < T$.

Statistic properties of the Itô's integral

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

$$\mathbb{E}[I(t)] \ = \mathbb{E}\Big[\int_0^t f(s)dW(s)\Big] \ = \int_0^t f(s)\mathbb{E}[dW(s)] = 0, \text{ for each } \ t \in \mathrm{T}.$$

- The variance of the Itô's integral is

$$\mathbb{V}[I(t)] \ = \mathbb{E}[I(t)^2] = \int_0^t \mathbb{E}[f(s)^2]ds, \text{ for each } \ t \in \mathrm{T}.$$

  (see the proof next).

13.3.1 The Itô's integral and stochastic calculus

Itô's integral, in equation (13.8), is written in the integral form **differential form** as

$$dI(t) = f(t)dW(t) \tag{13.10}$$

where $dW(t)$ is a variation of the Wiener process. Even if $f(.)$ 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

Lemma 1 (Itô's formula)**.** *Assume that $X(t)$ is an Itô's integral and assume that $f(x) : \mathbb{R} \to \mathbb{R}$ is a C^2 function. Then the integral $Y(t) = g(t, X(t))$ satisfies, in its differential form, the* ***Itô's formula***

$$dY(t) = g_t(t, X(t))dt + g_x(t, X(t))dX(t) + \frac{1}{2}g_{xx}(t, X(t))(dX(t))^2. \tag{13.11}$$

where

$$g_t(t,x) \equiv \frac{\partial g(t,x)}{\partial t}, \ g_x(t,x) \equiv \frac{\partial g(t,x)}{\partial x}, \ \textit{and} \ \ g_{xx}(t,x) \equiv \frac{\partial^2 g(t,x)}{\partial x^2}.$$

In its application the following ***Itô's rules*** *are used*

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

Proof. Itô (1951) □

We can use the previous formulas for show that

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

where $f(.)$ is deterministic, and $(I(t))$ is an Itô integral. To prove this: first, let $dI(t) = f(t)\,dW(t)$ and determine $I(t)^2$: applying the Itôs formula yields

$$\begin{aligned} d(I(t))^2) &= 2\,I(t)\,dI(t) + \frac{2}{2}\,(dI(t))^2 \\ &= 2\,f(t)\,dW(t) + \big(f(t)\,dW(t)\big)^2 \\ &= f(t)^2\,dt + 2\,f(t)\,dW(t); \end{aligned}$$

second, integrating

$$I(t)^2 = I(0)^2 + \int_0^t \; dI(s)^2 = \int_0^t \; dI(s)^2 = \int_0^t f(s)^2\,ds + 2\int_0^t f(s)\,dW(s),$$

at last, $\mathbb{E}[I(t)^2] \; = \int_0^t f(s)^2\,ds$ because $\mathbb{E}\Big[\int_0^t f(s)\,dW(s)\Big] \; = 0$.

In applications the following results are usefull:

- Let $dX(t) = dW(t)$ be an Itô's SDE, and $Y(t) = g(t, X(t))$, where $g(\cdot)$ is a $C^{1,2}$ function. Tthen $Y(t)$ is follows

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

  is also an Itô's SDE

- Let $dX(t) = f(t)\,dW(t)$, then $Y(t) = g(t, X(t))$ satisfies

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

In particular, if $dX(t) = dW(t)$ and $Y(t) = g(X(t))$, we have several particular cases:

- for a linear function: $g(x) = a\,x + b$, as $g_t(x) = 0$, $g_x(x) = a$ and $g_{xx}(x) = 0$, then

$$dY(t) = a\,dX(t) = a\,dW(t)$$

- for a power function: $g(x) = x^a$, for $a \neq 0$, as $g_t(x) = 0$, $g_x(x) = a\,x^{a-1}$ and $g_{xx}(x) = a\,(a-1)\,x^{a-2}$, then

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

- for an exponential function: $g(x) = e^{\lambda x}$, for $\lambda \neq 0$, as $g_t(x) = 0$, $g_x(x) = \lambda\, e^{\lambda x}$ and $g_{xx}(x) = \lambda^2\, e^{\lambda x}$, then

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

- for a logarithmic function: $g(x) = \ln(x)$, then

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

13.3.2 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,

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

where $\nabla_x f(\cdot, \mathbf{x})$ is the Jacobian and $\nabla_{xx} f(\cdot, \mathbf{x})$ is the Hessian on function $f(\cdot, \mathbf{x})$,

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

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

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

Example 1: product rule Let $Y(t) = f(X_1(t), X_2(t)) = X_1(t)X_2(t)$. Then

$$dY(t) = X_1(t)dX_2(t) + X_2(t)dX_1(t) + dX_1(t)dX_2(t),\ \text{for each}\ \ t \in \mathrm{T}$$

that is equivalent to

$$\frac{dY(t)}{Y(t)} = \frac{dX_1(t)}{X_1(t)} + \frac{dX_2(t)}{X_2(t)} + \frac{dX_1(t)}{X_1(t)}\,\frac{dX_2(t)}{X_2(t)},\ \text{for each}\ \ t \in \mathrm{T},$$

where the presence of the last term distinguishes the Itô's stochastic calculus from product rule of elementary calculus.

To prove this, apply the Itô rule observing that we have the following derivatives of $f(x_1, x_2)$ [3]:

$$\nabla_x f(x_1, x_2) = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}, \; \nabla_{xx}\; f(x_1, x_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then

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

Example 2: quocient rule Let $Y(t) = f(X_1(t), X_2(t)) = X_1(t)/X_2(t)$. Then

$$\frac{dY(t)}{Y(t)} = \frac{dX_1(t)}{X_1(t)} - \frac{dX_2(t)}{X_2(t)} - \frac{dX_2(t)}{X_2(t)} \left(\frac{dX_1(t)}{X_1(t)} - \frac{dX_2(t)}{X_2(t)} \right), \text{ for each } \; t \in \mathrm{T},$$

where, again, the presence of the last term distinguishes the Itô's stochastic calculus from the quotient rule of elementary calculus.

Exercise: prove this.

13.4 Diffusion processes

Let us consider the stochastic differential equation in the Itô interpretation, which is also called **diffusion equation**, in differential form

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dW(t), \text{ for each } \; t \in [0, \infty) \tag{13.12}$$

or in the integral form

$$X(t) = X(0) + \int_0 t \; \mu(X(s))ds + \int_0^t \; \sigma(X(s))dW(s), \text{ for each } \; t \in [0, \infty) \tag{13.13}$$

where the solution $(X(t))_{t \in \mathrm{T}}$ 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 us to characterise the properties of the diffusion process. In the next chapter we will apply them to obtain explicit solutions of linear SDE's.

[3]We write the function as $f(x_1, x_2)$ and not $f(X_1, X_2)$ because this function is the same independently from the realizations of the two stochastic processes. That is, it is state-independent. This would not be the case if the function is state dependent, as $f(X_1, X_2, \omega)$ in which ω is a function of past values of the Wiener process $(W(t))$.

13.4.1 Functions of the diffusion

Before determining the statistics for the process $\Big(X(t)\Big)_{t\in \mathrm{T}}$ it is useful to apply the Itô's formula to a function of the diffusion.

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

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

where $X(t)$ *is a diffusion process given by equation* (13.12)*, and* $f(\cdot)$ *is a point-wise mapping* $f(x) : \mathbb{R} \to \mathbb{R}$ *which is at least* $C^2(\mathbb{R})$*, and it is invertible, such that* $x = f^{-1}(y) = g(y)$*, where* $g(\cdot)$ *is continuous. Then* $Y(t)$ ***is also a diffusion process*** *such that*

$$dY(t) = \mu_y(Y(t))dt + \sigma_y(Y(t))dW(t). \tag{13.14}$$

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 $dY(t) = d(f(X(t)),$

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

If the function $f(.)$ is invertible then we substitute, for every realization, $x = f^{-1}(y) = g(y)$ 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.

13.4.2 Moment equations

In this subsection we derive the first and the second moments for the diffusion process $(X(t))_{t\in\mathbb{R}_+}$ generated by the one-dimensional diffusion equation in integral form in equation (13.13).

Proposition 4. *Consider the diffusion in integral form* (13.13)*, 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,

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

yields

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

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

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

□

Taking time-derivatives, the previous properties have differential equivalents for the variation of the first moment

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

the variation of the second moment,

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

and the variation of the centered second moment

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

These equations allows to find a relationship between ordinary differential equations (ODEs) and the moments of the processes represented by stochastic differential equations (SDE). If this is the case, we can obtain an asymptotic characterization of the moments of the solution to SDEs.

For particular cases, depending on the functional form of $\mu(\cdot)$ and $\sigma(\cdot)$, this allows to obtain ordinary linear differential equations over moments of $X(t)$, and, after solving them obtain time varying solutions for the moments.

Example Consider the linear diffusion equation

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

where $X(0) = x_0$ is known, and γ and σ are two constants.

The first moment of $X(t)$ 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

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

Therefore, the dynamic properties of the moments can be the following, depending on the value of γ:

1. if $\gamma > 0$ then
$$\lim_{t\to\infty} \mathbb{E}[X(t)] = 0, \text{ and } \lim_{t\to\infty} \mathbb{V}[X(t)] = \lim_{t\to\infty} \mathbb{E}[X(t)^2] = \frac{\sigma^2}{2\gamma}.$$
then the process tends asymptotically to a bounded limit distribution $N\left(0, \frac{\sigma^2}{2\gamma}\right)$. In this case we say we have an ergodic process;

2. if $\gamma = 0$ then
$$\lim_{t\to\infty} \mathbb{E}[X(t)] = x_0, \lim_{t\to\infty} \mathbb{V}[X(t)] = \frac{\sigma^2}{2\gamma}, \text{ and } \lim_{t\to\infty} \mathbb{E}[X(t)^2] = x_0^2$$
then the process is path dependent as it tends to a limit distribution $N\left(x_0, \frac{\sigma^2}{2\gamma}\right)$. In this case we say we have an ergodic process;

3. if $\gamma < 0$ then
$$\lim_{t\to\infty} \mathbb{E}[X(t)] = \lim_{t\to\infty} \mathbb{V}[X(t)] = \lim_{t\to\infty} \mathbb{E}[X(t)^2] = \infty$$
then the process unbounded.

13.4.3 Forward density dynamics: the Fokker-Planck-Kolmogorov equation

An alternative method to determine the dynamics of moments involves finding the dynamics of the probability distribution. This approach allows for finding a relationship between SDE and partial differential equations (PDE).

Consider again the diffusion process specified in equation (13.12).

We define the **unconditional probability**
$$p(t, x) = \mathbb{P}[X(t) = x|X(0) = x_0],$$
as a mapping $p : \mathrm{T} \times \mathrm{X} \to (0, 1)$. It is the probability, determined with the information at time $t = 0$, that the realization of the process at time $t > 0$ will be equal to x (a scalar), i.e. $X(t) = x$, when we observe that at time $t = 0$ the process it is equal to x_0, i.e. $X(0) = x_0$. The initial state can be made consistent with the unconditional probability if we let the probability at time $t = 0$ to be a Dirac-delta distribution centered at x_0, $p(0, x) = \delta(x - x_0)$.

We introduce the following assumption: the support of x is the whole set of real numbers $\mathrm{X} = \mathbb{R}$, it satisfies $\lim_{x\to\pm\infty} p(t, x) = 0$, and a normalization condition holds
$$\int_{-\infty}^{\infty} p(t, x)\, dx = 1, \text{for every}\, t \geq 0.$$

Then the mean value of the process $(X(t))_{t\in\mathbb{R}_+}$ is the function of time
$$\mathbb{E}[X(t)] \;= \int_{-\infty}^{\infty} x\, p(t, x)\, dx, \text{ for every}\, t \in \mathbb{R}_+.$$

Definition 3. *Let $X(t) = x$ be the realization of the diffusion process* (13.12) *at time $t > 0$ and let $p(t,x)$ be its unconditional probability. Then the following operator*

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

is called ***adjoint operator****.*

The following (forward) partial differential equation

$$\frac{\partial p(t,x)}{\partial t} = G^*[p](t,x) = -\frac{\partial(\mu(x)\,p(t,x))}{\partial x} + \frac{1}{2}\frac{\partial^2(\sigma(x)^2\,p(t,x))}{\partial x^2} \tag{13.15}$$

is called **Kolmogorov-Fokker-Planck equation**.

Proposition 5 (Forward density dynamics)**.**
Let the stochastic process $(X(t))_{t\geq 0}$ be generated by the SDE (13.12)*, and let the initial state be $X(0) = x_0$. Then the probability distribution of $X(t)$ at time $t > 0$, is the solution of the forward partial differential equation (PDE)*

$$\begin{cases} p_t(t,x) = G^*[p](t,x) & \text{for } \ t > 0 \\ p(0,x) = \delta(x - x_0) & \text{for } \ t = 0 \end{cases} \tag{13.16}$$

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

Proof. (Heuristic) We introduce a test function. Let $t \in \mathrm{T} = [0,T]$ and $x \in \mathrm{X} = (-\infty,\infty)$ and consider an arbitrary stationary and bounded function $f(t,x)$ such that $f(0,X(0)) = f(T,X(T)) = 0$, for $X(0) = x_0$ and any realization of $X(T)$, and $\lim_{x\to\pm\infty} f(t,x) = 0$.

Therefore,

$$f(t,X(t)) = f(0,x_0) + \int_0^t df(s,X(s)) = \int_0^t df(s,X(s)),$$

and we expect

$$\mathbb{E}[f(T,X(T))] = \mathbb{E}\Big[\int_0^T df(t,X(t))\Big] = 0.$$

Using the Itô's Lemma we find, for any realization of the process at time t, $X(t) = x$, we have

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

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

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

Assume there is an unconditional probability distribution $p(t,x)$. The unconditional expected value of the variation of f from $t=0$ to $t=T$ is

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

Because function $f(\cdot)$ is arbitrary, but has 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

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

With the boundary conditions introduced then

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

Therefore, for an arbitrary stationary process $\mathbb{E}\Big[\int_0^T df(t)\Big]=0$ if equation (13.15) 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

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

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 $p(x,0) = \mathbb{P}[X(0)] = \delta(x - x_0)$ 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 10 that the solution to this problem is

$$p(t,x) = \frac{1}{\sigma\sqrt{2\pi t}}\, e^{-\frac{(x-x_0)^2}{2\sigma^2 t}}, \text{ for } t > 0.$$

Therefore, the expected value and the variance are

$$\mathbb{E}[X(t)] = \int_{-\infty}^{\infty} x\, p(t,x)\, dx = x_0$$

and

$$\mathbb{V}[X(t)] = \int_{-\infty}^{\infty} (x - \mathbb{E}[X(t)]\ )^2\, p(t,x)\, dx = \int_{-\infty}^{\infty} (x - x_0\ )^2\, p(t,x)\, dx = \sigma^2\, t$$

and the process is not ergodic.

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 the Gaussian kernel is

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

Therefore the unconditional probability distribution for an arbitrary realization $X(t) = x$ is

$$p(t,x) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\dfrac{(x - x_0 - \mu t)^2}{2\sigma^2 t}}. \tag{13.17}$$

and the expected value and variance are

$$\mathbb{E}[X(t)] \ = \int_{-\infty}^{\infty} x\, p(t,x)\, dx = x_0 + \mu\, t$$

and

$$\mathbb{V}[X(t)] \ = \int_{-\infty}^{\infty} (x - \mathbb{E}[X(t)]\ )^2\, p(t,x)\, dx = \int_{-\infty}^{\infty} (x - x_0 - \mu\, t\ )^2\, p(t,x)\, dx = \sigma^2\, t$$

13.5 Backward probability distribution

In some problems, particularly in finance applications, we may be interested in determining the distribution dynamics such that a terminal condition is observed. Assume again we have a stochastic process generated by the diffusion SDE (13.12).

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

13.5.1 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 $G[f](t,x)$,

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

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

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

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.**

13.5.2 Kolmogorov backward equation

The Kolmogorov backward equation allows for the determination of the probability that the state of the process is $x \in \mathrm{X}$, $X(t) = x$, if the value of the process belongs to a target set ϕ_T at time $T > t$.

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

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

when $X(t)$ follows a diffusion process (13.12). The probability distribution $q(\cdot)$ satisfies the **Kolmogorov backward** (KB) equation

$$q_t(t,x) = -G[q](t,x) = -q_x(t,x)\mu(x) - \frac{1}{2}\sigma(x)^2 q_{xx}(t,x), \text{ for } \ t \in (0,T), \tag{13.18}$$

which we want to solve together with with the terminal condition specifying the conditional terminal probability distribution

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

where $q_T(\cdot)$ is a known mapping $q_T : \mathrm{X} \to (0,1)$ such that $\int_{\mathbb{X}} q_T(y)\, dy = 1$.

Example Let $dX(t) = \sigma dW(t)$. The backward FPK equation is

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

In order to determine the backward probability we consider two different terminal conditions.

First let $X(T) = x_T$ where x_T is a known number. To obtain the probability distribution $q(t,x)$ we have to solve the well posed backward parabolic PDE problem

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

The solution is

$$q(t,x) = \frac{e^{-\dfrac{-(x-x_T)^2}{2\,\sigma^2\,(T-t)}}}{\sigma\,\sqrt{2\,\pi\,(T-t)}}, \text{ for } 0 < t < T,$$

and is illustrated in Figure 13.2 panel (a).

Second, let the terminal distribution be $q_T(x) = e^{-x^2}/\sqrt{\pi}$. Then $q(t,x)$ is the solution we have to solve the problem

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

The solution is

$$q(t,x) = \frac{e^{-\dfrac{-x^2}{1+2\,\sigma^2\,(T-t)}}}{\sigma\,\sqrt{1+2\,\pi\,(T-t)}}, \text{ for } 0 < t < T$$

and is illustrated in Figure 13.2 panel (b).

13.5.3 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 $(X(t))_{t\in[0,T]}$, solving equation (13.12), when there is a discount factor with discount rate $f(X(t))$.

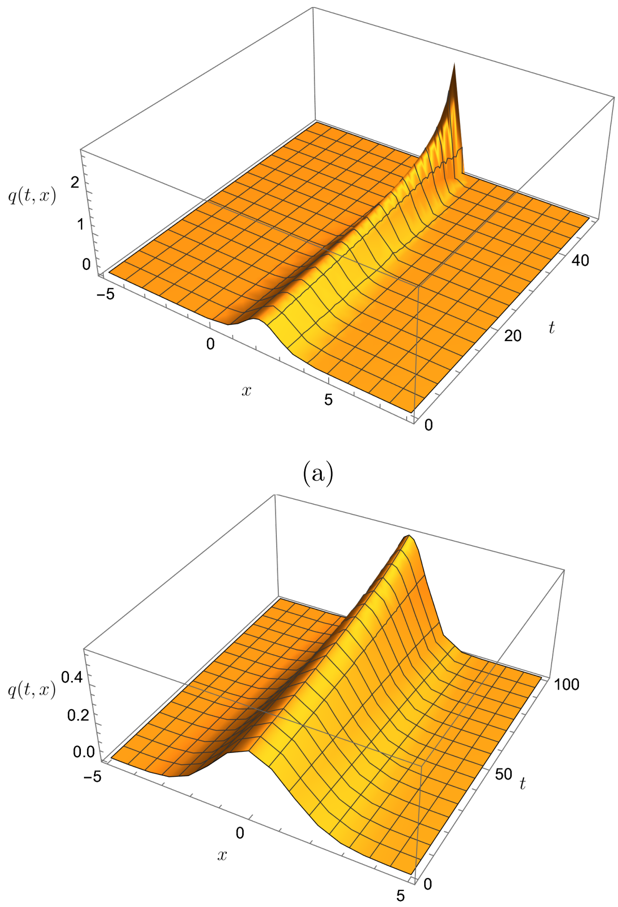
Let $v(t,x)$ be the probability at time t for a realization $X(t) = x$. Assume that the function $v(t,x)$ is the solution for the partial differential equation boundary value problem

$$\begin{cases} v_t(t,x) = -G[v](t,x) + f(x)\,v(t,x), & 0 < t \leq T \\ v(T, X(T)), & T \end{cases} \tag{13.19}$$

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

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

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



(a)

(b)

Figure 13.2: Backward FPK for $\sigma = 0.1$

Proposition 6. *The solution to the PDE problem* (13.19) *is the* ***Feynman-Kac*** *formula:*

$$v(t,x) = \mathbb{E}\left[\ v(T, X(T))e^{-\int_t^T f(X(s))ds} | X(t) = x\right]. \tag{13.20}$$

Then $v(t,x)$ is the present value of a terminal value $v(T, X(T))$, discount at rate $f(X(t))$, conditional on having $X(t) = x$.

Proof. Let us write the discounted probability distribution $V(t) = V(t, X(t))$ where

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

where $H(t) \equiv e^{-Z(t)} = e^{-\int_s^t f(X(\tau))d\tau}$ and $X(t)$ is driven by the diffusion process . Next, we want to obtain the terminal value $V(T)$. First, because $dZ(t) = f(X(t))dt$ then

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

Second, as $v(t, X(t))$ depends on the diffusion equation (13.12)

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

from the PDE in (13.19). Third, using the Itô's 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 Feynman-Kac formula (13.20). □

Observe that $v(\cdot)$ does not need to be a distribution. If there is no discounting and $v(T)$ is a distribution the Feynman-Kac formula is equivalent to backward Kolmogorov equation.

13.6 References

- Mathematics of SDE's: Karatzas and Shreve (1991), Øksendal (2003), Pavliotis (2014)
- Very useful hands-on introduction to SDE: Särkkä and Solin (2019)
- 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), Dixit and Pindyck (1994), Cvitanić and Zapatero (2004) , Stokey (2009)

Chapter 14

Linear scalar stochastic differential equations

14.1 Introduction

Forward scalar linear stochastic differential equations (SDE) are equations of the form

$$dX(t) = \Big(\mu_0(t) + \mu_1(t)\,X(t)\Big)\,dt + \Big(\sigma_0(t) + \sigma_1(t)\,X(t)\Big)\,dW(t). \tag{14.1}$$

where $X(0) = x_0$ is a known constant, and $\mu_0(\cdot)$, $\mu_1(\cdot)$, $\sigma_0(\cdot)$ and $\sigma_1(\cdot)$ are known functions and $(W(t))_{t\in\mathbb{R}_+}$ is a standard one-dimensional Wiener process, and therefore, a non-anticipating process.

In this chapter we provide explicit solutions, and their properties, to autonomous equations, which are the simplest linear SDE's. 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.

We can compare those solutions with the analogous (deterministic) autonomous ODEs

$$dy(t) = (\mu_0 + \mu_1\,y(t))\,dt$$

we saw that the solution is

$$y(t) = \begin{cases} -\frac{\mu_0}{\mu_1} + (y(0) + \frac{\mu_0}{\mu_1})\,e^{\mu_1\,t}, & \text{if}\,\mu_0 \neq 0 \text{ and } \ \mu_1 \neq 0 \\ y(0)\,e^{\mu_1\,t}, & \text{if}\,\mu_0 = 0 \text{ and } \ \mu_1 \neq 0 \\ y(0) + \mu_0\,t, & \text{if}\,\mu_0 \neq 0 \text{ and } \ \mu_1 = 0 \\ y(0), & \text{if}\,\mu_0 = \mu_1 = 0, \end{cases} \tag{14.2}$$

for every $t \in \mathrm{T}$. We saw that: (1) if $\mu_1 < 0$ the solution is asymptotically stable, such that $\lim_{t\to\infty}\ y(t) = -\frac{\mu_0}{\mu_1}$; (2) if $\mu_1 > 0$ or if $\mu_1 = 0$ and $\mu_0 \neq 0$ the solution is unstable; (4) the solution is stationary if $\mu_0 = \mu_1 = 0$.

We can compare those results with the solutions of a linear SDE. In section 14.2 we present the solutions to several linear equations. In section **??** we present the general solution to non-autonomous equations. Section 14.4 provides some applications to economics and finance.

14.2 Autonomous equations

In this section we consider the linear autonomous forward SDE,

$$dX(t) = \Big(\mu_0 + \mu_1\, X(t)\Big)\, dt + \Big(\sigma_0 + \sigma_1\, X(t)\Big)\, dW(t). \tag{14.3}$$

in which the coefficients are known constants. We assume a known initial value $X(0) = x_0$.

Next we present the explicit solutions, and the first and second moments of the solutions. With a view to comparing with the deterministic ODE, we discuss in the stochastic dynamic properties, that is, the asymptotic statistic properties of the solutions.

14.2.1 Brownian motion

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

$$dX = \mu dt + \sigma dW(t),\ t \in \mathbb{R}_+ \tag{14.4}$$

with $X(0) = x_0 \in \mathbb{R}$ and $\sigma > 0$. This is a special case of equation (14.3) with $\mu_1 = \sigma_1 = 0$ and $\mu_0 = \mu$ and $\sigma_0 = \sigma$.

Proposition 1. *The solution of equation* (14.4)*, given* $X(0) = x_0$ *is*

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

Proof. Writing $X(t)$ in the integral form, yields

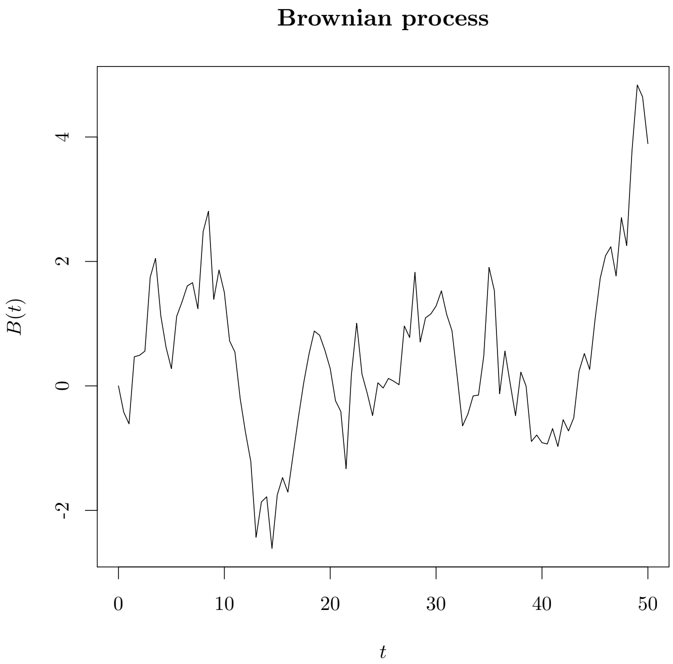
$$\begin{aligned}
X(t) &= X(0) + \int_0^t dX(s) \\
&= x_0 + \int_0^t \mu\, ds + \int_0^t \sigma dW(s) \\
&= \phi + \mu\, t + \sigma\left(W(t) - W(0)\right) \\
&= \phi + \mu\, t + \sigma\, W(t)
\end{aligned}$$

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

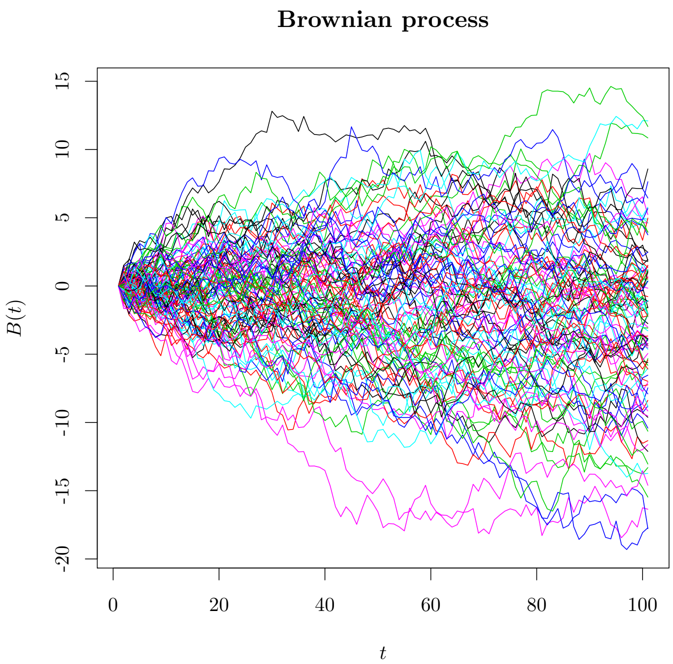
Figure 14.1 presents one sample path, in subfigure (a), and 100 sample paths, in subfigure (b), for the case in which $\mu = -0.5$ and $\sigma = 1$.

The probability distribution $p(t, x) = \mathbb{P}[X(t) = x | X(0) = x_0]$ was already derived in chapter **??**, see equation (13.17),

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



(a) One replication



(b) 100 replications

Figure 14.1: Sample path for the Brownian process for $\mu = -0.5$ and $\sigma = 1$.

Properties The first and second moments are linear functions of time

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

We observe that the process is not ergodic, because

$$\lim_{t\to\infty} \mathbb{E}^{x_0}[X(t)] = \lim_{t\to\infty} \mathbb{V}^{x_0}[X(t)] = \pm\infty$$

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

Observe that the solution of the skeleton $\frac{dx(t)}{dt} = \mu$, given x_0 is $x(t) = x_0 + \mu\, t$ (see equation (14.2)).

14.2.2 Geometric Brownian motion

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

$$dX(t) = \mu\, X(t)dt + \sigma\, X(t)dW(t),\ t \in \mathbb{R}_+. \tag{14.5}$$

This is a special case of equation (14.3) with $\mu_0 = \sigma_0 = 0$ and $\mu_1 = \mu$ and $\sigma_1 = \sigma$.

Proposition 2. *The explicit solution to equation* (14.5) *where* $\mathbb{P}[X(0) = x_0]\ = 1$ *is*

$$X(t) = x_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma\, W(t)},\ \textit{a.s. for}\ \ t \in \mathbb{R}_+. \tag{14.6}$$

Proof. Let $Y(t) = \ln X(t)$. 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, because $X(t) = e^{Y(t)}$, equation (14.6) is obtained. □

To find the probability distribution $p(t,x) = \mathbb{P}[X(t) = x | X(0) = x_0]$ we solve the Fokker-Planck-Kolmogorov (FPK) equation

$$\begin{cases} \frac{\partial}{\partial t} p(t,x) = -\frac{\partial}{\partial x}\left(\mu\, x\, p(t,x)\right) + \frac{1}{2}\, \frac{\partial^2}{\partial x^2}\left(\sigma\, x\, p(t,x)\right) & \text{for } \ t > 0 \\ p(0, x_0) = \delta(x - x_0), \ \text{where } \ x_0 > 0 & \text{for } \ t = 0. \end{cases}$$

The solution to this problem only exists for positive values of the state variable $x \in \mathrm{X} = \mathbb{R}_{++}$ and it is

$$p(t,x) = \frac{1}{x\sigma\sqrt{2\pi t}} e^{-\frac{\left(\ln(x/x0) - \left(\mu - \frac{1}{2}\sigma^2\right) t\ \right)^2}{2\sigma^2 t}}, \text{for } (t,x) \in \mathbb{R}^2_+. \tag{14.7}$$

To prove this result, expanding the FPK equation yields

$$\partial_t p(t,x) = \frac{\sigma^2}{2}\ \ x^2\, \partial_{xx} p(t,x) + (2\sigma^2 - \mu)\, \partial_x p(t,x) + (\sigma^2 - \mu)\, p(t,x).$$

We write the independent variable as a mapping $x : \mathbb{R}_+ \to \mathbb{R}$, as $x = e^z$ and write a transformed probability distribution $u(t,z) = p(t, x(z))$. As $\partial_t u(t,z) = \partial_t p(t, x(z))$, $\partial_z u(t,z) = \partial_x p(t, x(z))\, x(z)$ and $\partial_{zz} u(t,z) = \partial_{xx} p(t, x(z))\, x(z)^2 + \partial_x p(t, x(z))\, x(z)$ the FPK equation, over $p(t,x)$, is equivalent to the linear parabolic autonomous and homogenous parabolic PDE over $u(t,z)$

$$\begin{cases} \partial_t u(t,z) = \frac{\sigma^2}{2}\, \partial_{zz} u(t,z) + \left(\frac{3}{2}\sigma^2 - \mu\right) \partial_z u(t,z) + (\sigma^2 - \mu)\, u(t,z), & (t,z) \in \mathbb{R}_+ \times \mathbb{R} \\ u(0,z) = \delta(z - \ln(x_0)) & (t,z) \in \{t = 0\}\ \times \mathbb{R}. \end{cases}$$

Using the results for the linear parabolic PDE (in the unbounded spatial domain) the solution is

$$u(t,z) = \int_{-\infty}^{\infty}\ \delta(s - \ln(x_0))\, g(t, z - s)\, ds = g(t, z - \ln(x_0))$$

where

$$g(t,\xi) = \frac{1}{\sqrt{2\pi\sigma^2 t}}\ \exp\left\{\ -\ \frac{\left(\xi - (\mu - \frac{\sigma^2}{2}\ )\right)^2}{2\sigma^2 t} - \xi\right\}.$$

Transforming back to the original variable we have $p(t,x) = u(t, \ln(x) - \ln(x_0)) = g(t, \ln(x/x_0))$ yields equation (14.7).

Therefore, the geometric Brownian motion has the moments

$$\mathbb{E}^{x_0}[X(t)]\ = x_0 e^{\mu t},\ t \in \mathbb{R}_+,$$

$$\mathbb{V}^{x_0}[X(t)]\ = x_0^2 e^{2\mu t}\ \left(e^{\sigma^2 t} - 1\right),\ t \in \mathbb{R}_+.$$

Properties In Figure 14.2 we plot one sample path and several sample paths for the linear diffusion equation where $\mu < 0$ and $\sigma > 0$ and in Figure 14.3 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. This is the case because the variance tends to zero asymptotically, which means that all the mass of the distribution will be concentrated at point $x = 0 \in \mathrm{X}$.

Intuitively, we can say that the steady state is a degenerate distribution.

14.2.3 Ornstein-Uhlenback processes

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

$$dX = \theta\,(\mu - X)\,dt + \sigma dW(t). \text{ for } \ t \in \mathbb{R}_+ \tag{14.8}$$

This is a special case of equation (14.3) with $\mu_0 = \theta\,\mu$, $\mu_1 = -\theta$, $\sigma_0 = \sigma$ and $\sigma_1 = 0$.

Proposition 3. *The solution to equation* (14.8) *such that* $\mathbb{P}[X(0) = x_0] \ = 1$ *is*

$$X(t) = \mu + e^{-\theta t}\Big(x_0 - \mu + \sigma \int_0^t e^{\theta\, s}\ dW(s)\Big),\ a.s\ \ for\ \ t \in \mathbb{R}_+$$

Proof. Introduce the change in variables $Y(t) = X(t)\,e^{\theta t}$. Itô's formula yields

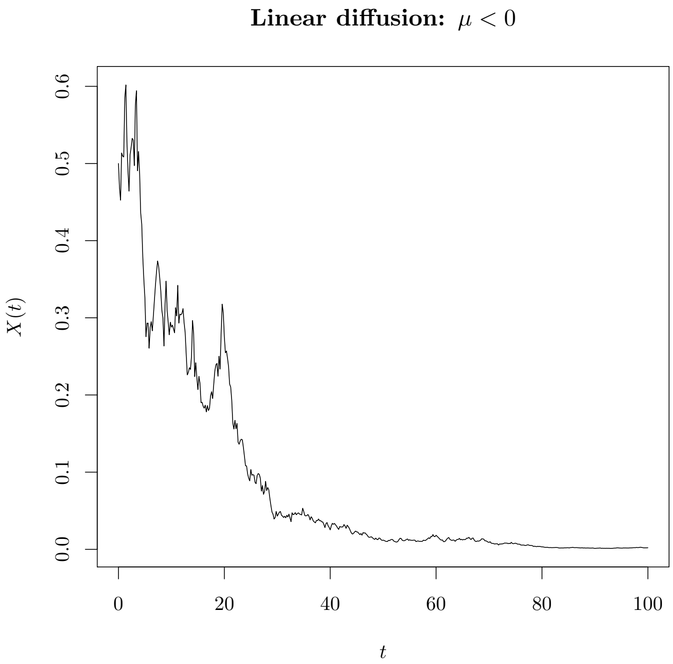
$$\begin{aligned} dY(t) &= \theta\, X(t)\, e^{\theta t}\, dt + e^{\theta t}\, dX(t) \\ &= \theta\, X(t)\, e^{\theta t}\, dt + e^{\theta t}\ (\theta\,(\mu - X(t))\ dt + \sigma\, dW(t)) \\ &= e^{\theta t}\,(\theta\mu dt + \sigma dW(t))\,. \end{aligned}$$

Integrating over time we have

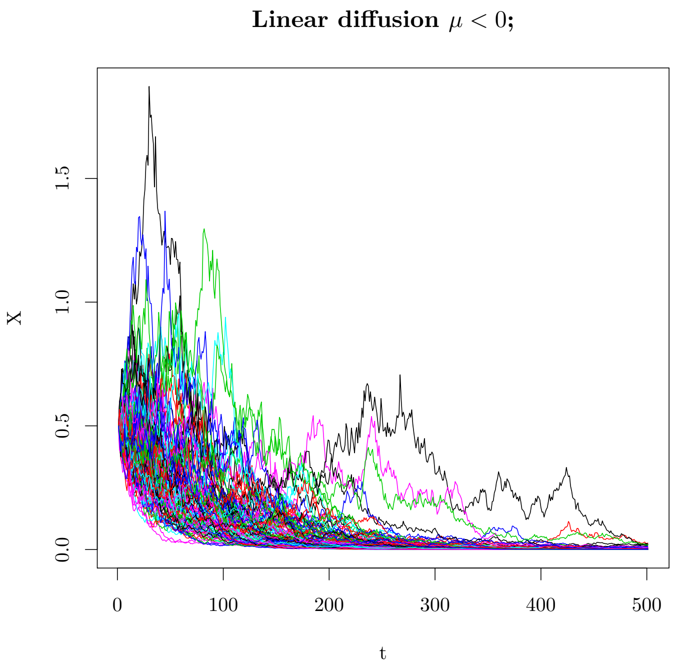
$$Y(t) = y_0 + \theta\mu \int_0^t e^{\theta s} ds + \sigma \int_0^t e^{\theta s} dW(s).$$

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 (14.8)

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

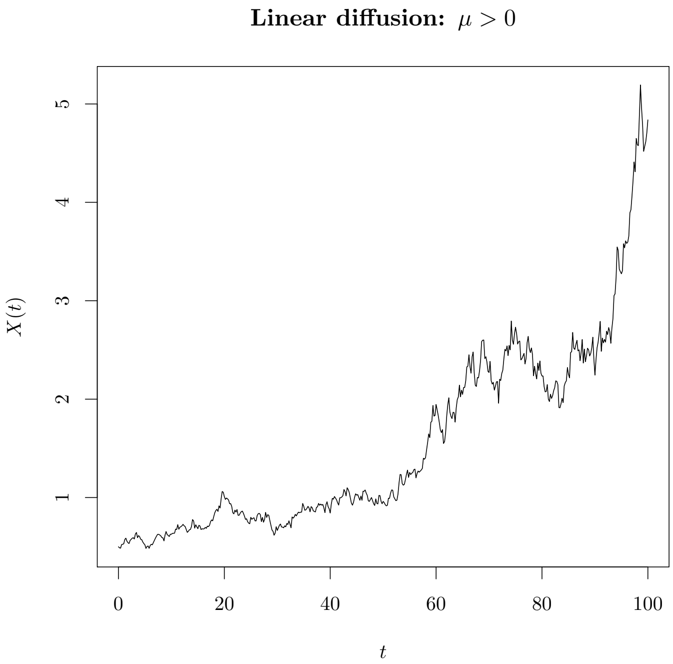


(a) One replication

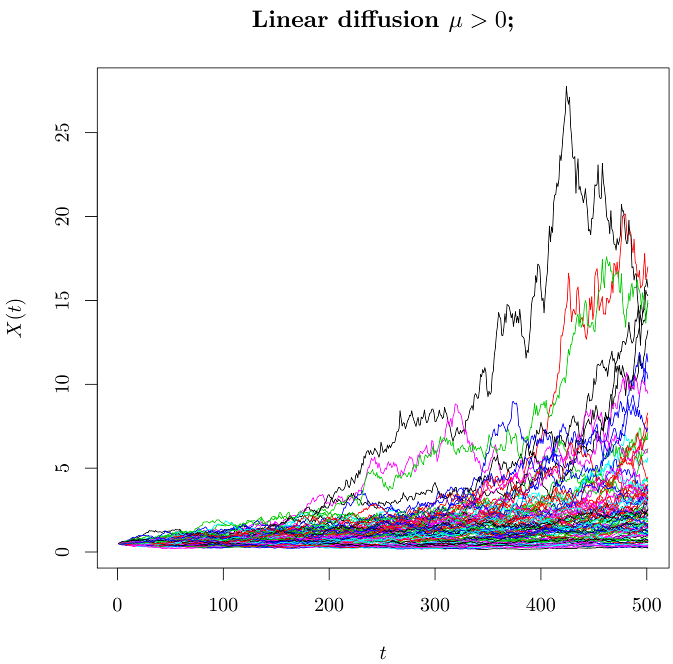


(b) 100 replications

Figure 14.2: Sample paths for the linear diffusion process with $\mu < 0$



(a) One replication



(b) 100 replications

Figure 14.3: Sample paths for the linear diffusion process with $\mu > 0$

□

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

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

The solution to this problem is

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

Therefore, the conditional expected value and variance, for $\mathbb{P}[X(0) = x_0] = 1$ are

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

and

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

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

$$\lim_{t\to\infty}\ \mathbb{E}^{x_0}\left[\ X(t)\right] = \mu$$

$$\lim_{t\to\infty}\ \mathbb{V}^{x_0}\left[\ X(t)\right] = \frac{\sigma^2}{2\theta}$$

  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 $\lim_{t\to\infty}\ \mathbb{V}^{x_0}\left[\ X(t)\right] = \infty$

Observe that the skeleton

$$\frac{dx(t)}{dt} = \theta\,(\mu - x(t))$$

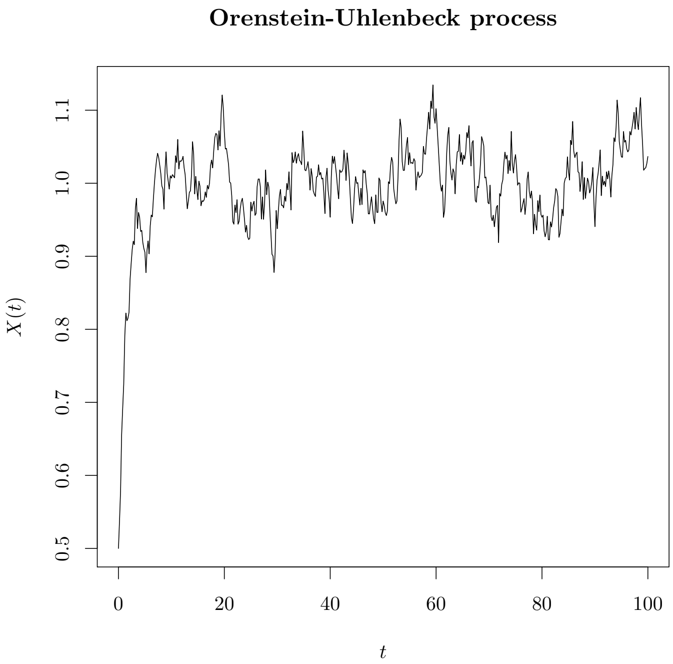
has the solution

$$x(t) = \mu + (x_0 - \mu)\,e^{-\theta\,t}$$

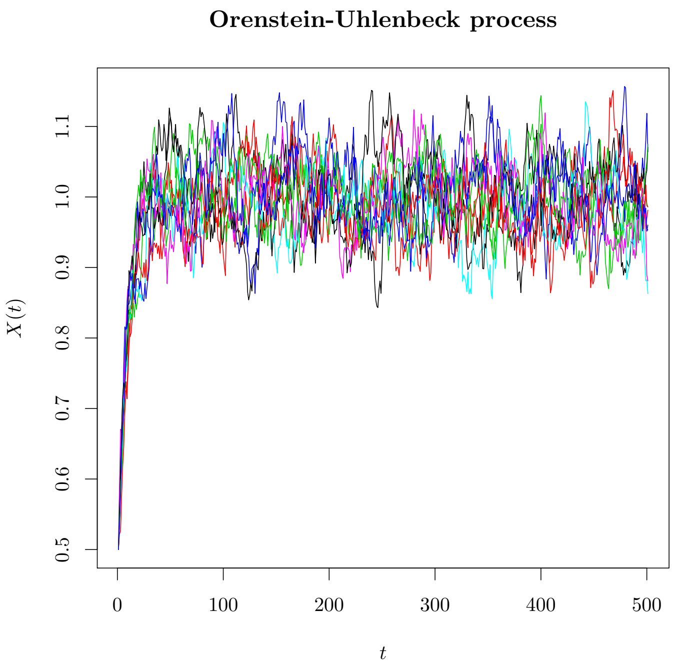
which is asymptotically stable if $\theta > 0$.

The sample paths for the case $\theta > 0$ are illustrated in figure 14.4: we see that they converge in average to $X(t) = \mu$, however this value is not an absorbing state. This can be understood from the fact the variance tends asymptotically to a positive finite value. Therefore, the solution of the OU process is ergodic: it will not be attracted to an absorbing state but will stay close to a finite value.

Intuitively, we can say that the steady state is a non-degenerate distribution.



(a) One replication



(b) 100 replications

Figure 14.4: Sample paths for Ornstein-Uhlenbeck process for $\theta > 0$ and $\mu = 1$

14.2.4 The linear autonomous SDE

Now consider equation the general linear Itô-SDE (14.3). If $\mathbb{P}[X(0) = x_0] = 1$ the explicit solution is

$$X(t) = \Phi(t)\left(x_0 + (\mu_0 - \sigma_0\,\sigma_1)\int_0^t \Phi(s)^{-1}\,ds + \sigma_0\int_0^t \Phi(s)^{-1}\,dW(s)\right), \text{ a.s. } t \in \mathbb{R}_+ \tag{14.9}$$

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 (14.3) to find a linear SDE for the process $dY(t)$. Solve this equation with the initial condition $Y(0) = x_0$. Use again the transformation $X(t) = \Phi(t)\,Y(t)$ to find equation (14.9).

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

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

and

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

If $\mu_1 < 0$ then the first moment is asymptotically finite:

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

However, if $\mu_1 < 0$ is sufficiently large in absolute value, such that $\mu_1 + \sigma_1^2 < 0$, which implies $2\,\mu_1 + \sigma_1^2 < 0$, and then the process is ergodic because in this case

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

14.2.5 Stochastic dynamic properties of the linear autonomous SDE

From the perspective of the asymptotic dynamics, the following qualitative stochastic dynamics can be expected from a general linear autonomous Itô-SDE, equation (14.3).

1. if $\mu_1 + \sigma_1^2 < 0$ and $\mu_1\,\sigma_0 - \mu_1\sigma_1 \neq 0$ the process is ergodic and tends asymptotically to a Gaussian distribution with positive variance $N\Big(-\dfrac{\mu_0}{\mu_1}, -\dfrac{(\mu_1\,\sigma_0 - \mu_1\sigma_1)^2}{\mu_1^2\,(2\,\mu_1 + \sigma_1^2)}\Big)$, which is a steady state non-degenerate 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 $x = -\dfrac{\mu_0}{\mu_1}$ 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. However, the expected value will converge asymptotically to an exponential with growth rate μ_1

$$\lim_{t\to\infty} \quad \mathbb{E}[X(t)] \propto \lim_{t\to\infty} \quad e^{\mu_1\,t}$$

and the valiance will also converge asymptotically to an exponential with growth rate $2\,\mu_1 + \sigma_1^2$

$$\lim_{t\to\infty} \quad \mathbb{V}[X(t)] \propto \lim_{t\to\infty} \quad e^{2\,\mu_1 + \sigma_1^2\,t}.$$

14.3 The general linear SDE: the non-autonomous case

The general linear SDE has the form

$$dX = (\mu_0(t) + \mu_1(t)X(t))\,dt + (\sigma_0(t) + \sigma_1(t)\,X)dW(t)$$

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

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

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

$$d\Phi(t) = \mu_1(t)\Phi(t)dt + \sigma_1(t)\,\Phi(t)dW(t),\ \text{for}\ \ t \in [0,\infty)$$

and $\Phi(0) = 1$.

14.4 Economic applications

14.4.1 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

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

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

$$dk = \left(sf(k) - (\mu - \sigma^2)k\right) dt - k\sigma dW(t) \tag{14.10}$$

For a Cobb-Douglas function we have

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

where $0 < \alpha < 1$. Figure 14.5 present one replication, in subfigure (a), and 100 replications, subfigure (b), 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)

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

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

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

if $2\,\mu > \sigma^2$. The asymptotic distribution is depicted in Figure 14.5 subfigure (c).

[1]For this case

$$m^{-1} = \frac{1}{1-\alpha}\left(\frac{(1-\alpha)\,\sigma^2}{2\,s}\right)^\zeta \Gamma(\zeta)$$

where $\zeta \equiv \dfrac{2\mu - \sigma^2}{(1-\alpha)\,\sigma^2}$ and $\Gamma(\cdot)$ is the Gamma distribution.

14.4.2 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

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

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:

$$\begin{cases} V_t(S,t) = -\frac{\sigma^2}{2} S^2 V_{ss}(S,t) - rSV_s(S,t) + rV(S,t), & (S,t) \in (0,\infty) \times [0,T] \\ V(S,T) = \max\{S-K, 0\}, & (S,t) \in (0,\infty) \times \{\ t = T\} \end{cases} \tag{14.11}$$

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

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

The price of the option is

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

with

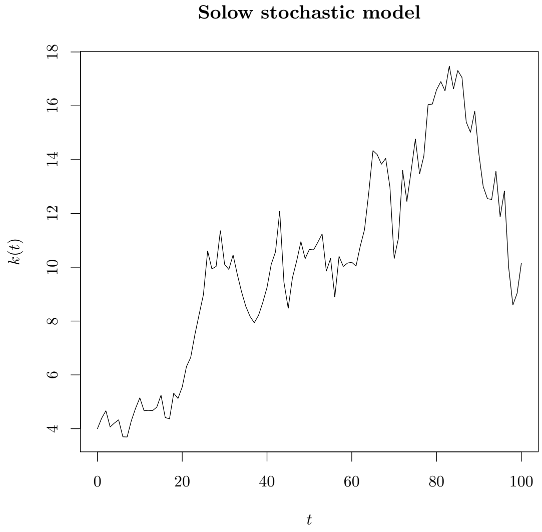
$$d_{\mp}(0) = \frac{\ln\left(\frac{S(0)}{K}\right) + T\left(r \mp \frac{\sigma^2}{2}\right)}{\sigma\sqrt{T}}.$$

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.

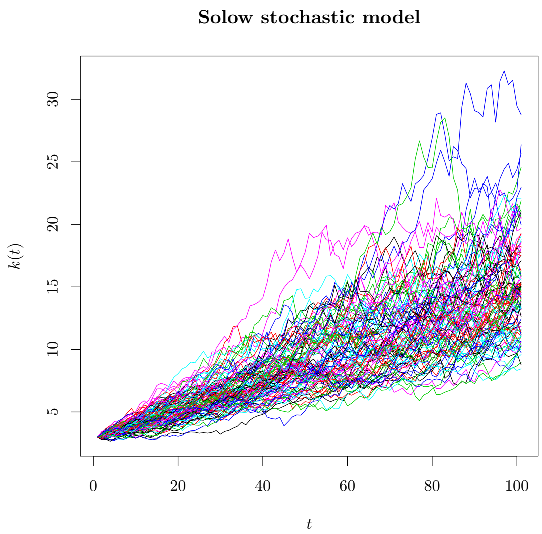
14.5 References

- Mathematics of SDE's: Karatzas and Shreve (1991), Øksendal (2003), Pavliotis (2014)

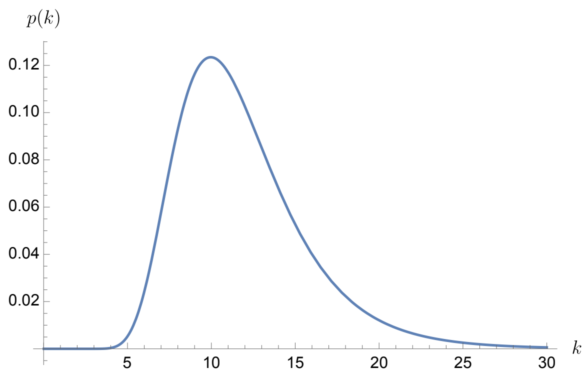
- Very useful hands-on introduction to SDE: Särkkä and Solin (2019). Explicit solutions: Kloeden and Platen (1992) and Gardiner (2009)
- Dynamic systems theory and SDE's: Cai and Zhu (2017)
- Numerical analysis of SDE Iacus (2010)
- Application to economics and finance: Malliaris and Brock (1982), Dixit and Pindyck (1994), Cvitanić and Zapatero (2004) , Stokey (2009)



(a) One sample path for $k(t)$



(b) 100 sample paths for the$k(t)$



(c) Asymptotic probability distribution

Figure 14.5: Solow model with $s = 0.1$, $\alpha = 0.3$, $\mu = 0.01$, $\sigma = 0.1$

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.

Chang, F.-R. (2004). *Stochastic Optimization in Continuous Time*. Cambridge University Press.

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

Dixit, A. K. and Pindyck, R. S. (1994). *Investment under Uncertainty*. Princeton University 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.

Gardiner, C. (2009). *Stochastic Methods. Handbook of stochastic methods for physics, chemistry, and the natural sciences*. Springer series in synergetics. Springer-Verlag, 4th ed edition.

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.

Karatzas, I. and Shreve, S. (1991). *Brownian Motion and Stochastic Calculus, 2nd ed.* Springer-Verlag.

Kloeden, P. E. and Platen, E. (1992). *Numerical Solutions of Stochastic Differential Equations*, volume 23 of *Applications of Mathematics*. Springer-Verlag.

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(3):375–393. Reproduced in R. C. Merton "Continuous-Time Finance", Blackwell, pp. 579-605.

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

Merton, R. C. (1982). On the mathematics and economics assumptions of continuous-time models. In *Essays in Honor of Paul Cootner*. Prentice-Hall. Reproduced in R. C. Merton "Continuous-Time Finance", Blackwell, pp. 57-93.

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.

Särkkä, S. and Solin, A. (2019). *Applied Stochastic Differential Equations*. Cambridge University Press.

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.