# Advanced Mathematical Economics

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

Lecture 7
24.11.2021
(revised 15.12.2021)

# Contents

**12 Introduction to stochastic differential equations and stochastic calculus 3**

12.1 Introduction . . . 3

12.2 The Wiener process . . . 5

12.2.1 Stochastic processes: a brief description . . . 5

12.2.2 Wiener process: definition . . . 6

12.3 The Itô's processes . . . 11

12.3.1 The Itô's integral and stochastic calculus . . . 12

12.3.2 Itô's formula for a multi-dimensional process . . . 14

12.4 Characterizing Itô's processes . . . 15

12.4.1 Functions of the diffusion . . . 15

12.4.2 Dynamics of the density: the Kolmogorov forward equation . . . 16

12.4.3 Moment equations . . . 20

12.5 Backward distributions . . . 22

12.5.1 Generator of a diffusion . . . 22

12.5.2 Kolmogorov backward equation . . . 23

12.5.3 The Feynman-Kac formula . . . 24

12.6 References . . . 25

**13 Linear scalar stochastic differential equations 27**

13.1 Introduction . . . 27

13.2 Autonomous equations . . . 28

13.2.1 Brownian motion . . . 28

13.2.2 Geometric Brownian motion . . . 30

13.2.3 Ornstein-Uhlenback processes . . . 32

13.2.4 The linear autonomous SDE . . . 35

13.2.5 Stochastic dynamic properties of the linear autonomous SDE . . . 37

13.3 The general linear SDE: the non-autonomous case . . . 38

13.4 Economic applications . . . 38

13.4.1 The Solow stochastic growth model . . . 38

13.4.2 Derivation of the Black and Scholes (1973) equation . . . 40

13.5 References . . . 42

**14 Stochastic optimal control** **43**
14.1 Introduction . . . . . . . . . . 43
14.2 Stochastic dynamic programming . . . . . . . . . . 43
14.2.1 Finite horizon . . . . . . . . . . 43
14.2.2 Infinite horizon . . . . . . . . . . 45
14.2.3 Economic applications using stochastic dynamic programming . . . . . . . . 45
14.3 The stochastic PMP . . . . . . . . . . 50
14.3.1 Volatility function independent of the control variable . . . . . . . . . . 51
14.3.2 Volatility dependent on the control variable . . . . . . . . . . 51
14.3.3 Economic applications using stochastic maximum principle . . . . . . . . . . 52
14.4 References . . . . . . . . . . 56

# Chapter 12

# Introduction to stochastic differential equations and stochastic calculus

## 12.1 Introduction

If we consider again the ordinary differential equation

$$\dot{y} = f(y(t)) \tag{12.1}$$

we can extend it by introducing a random perturbation,

$$\dot{Y} = f(Y(t)) + \epsilon(t) \tag{12.2}$$

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

$$\dot{Y} = f(Y(t), \epsilon(t)). \tag{12.3}$$

While the solution of (12.1) is a mapping $y : \mathbb{R}_+ \to \mathbb{R}^n$, in the case of equations (12.2) or (12.3), 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. In continuous time models applied to economics and finance there are two main ways to introduce a stochastic component[1]:

- to model rare events with a local high impact uncertainty is introduced via a **Poisson process**, $(Q(t)\ )_{t\in \mathrm{T}}$,
$$dY(t) = f(Y(t), t)dt + v(Y(t^-, t^-)\, dQ(t) \tag{12.4}$$
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** $(W(t))_{t\in \mathrm{T}}$ is a **Wiener process**. The most common model is called **diffusion equation**
$$dY(t) = f(Y(t), t)dt + \sigma(Y(t), t)dW(t) \tag{12.5}$$
where $f(.)$ and $\sigma(.)$ are continuous and differentiable known functions.

The main reason for using the previous formalism is related to the fact that $Y(t)$ in both cases is not differentiable in the classic sense, and specific stochastic calculus rules have to be developed before solving those equations.

Therefore, in general, the term **stochastic differential equation with jumps** (SDEJ) is reserved to equations of the form (12.4) in the differential form or to in the integral form

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

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

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

Poisson process and of the Poisson integral. Likewise, the term **stochastic differential equation** (SDE) is reserved to equations as (12.5) 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 a **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.

However, there are SDE which have both jump and diffusion components

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

The most common approach to SDE's view "noise" as generated by a Wiener process and builds upon the Itô process. 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. We present a very brief introduction to the Itô's stochastic calculus applied to stochastic differential equation of type (12.5), following an applied and heuristic approach. In particular, we emphasise the connections with ordinary and partial differential equations.

In section 12.2 we define and describe the properties of the Wiener process. In section 12.3 we define and present the properties of the Itô's process and integral. In section 12.4 we present methods for characterizing diffusion processes.

## 12.2 The Wiener process

### 12.2.1 Stochastic processes: a brief description

A uni-dimensional **stochastic process** can be seen as a flow of random variables $\Big(X(t,\omega)\Big)_{t\in\mathrm{T}}$, where $X(t,\omega) : \mathrm{T}\times\Omega \to \mathbb{R}$, where $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ is a filtered probability space. A filtered probability space is defined by the sample space $\Omega$, by the set of events $\mathcal{F}$ (i.e., the set of all subsets of $\Omega$, 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 associated to 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 $t$, $X(t) = x(t)$ is a particular number which can be observed **after** nature (or the computer) makes a draw. Therefore, we consider non-anticipating processes, or **processes adapted** to a non-anticipating filtration.

### 12.2.2 Wiener process: definition

There are several ways of characterising the **Wiener process** also called **standard Brownian motion**. As we mentioned in the last subsection, a stochastic process can be defined by the way 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 $W : \Omega \times T \rightarrow \mathbb{R}$ 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. *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 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 last property implies that the Wiener process is a Markovian, or memory-less, process.

A **propagator** can be defined as

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

that is 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 $t$.

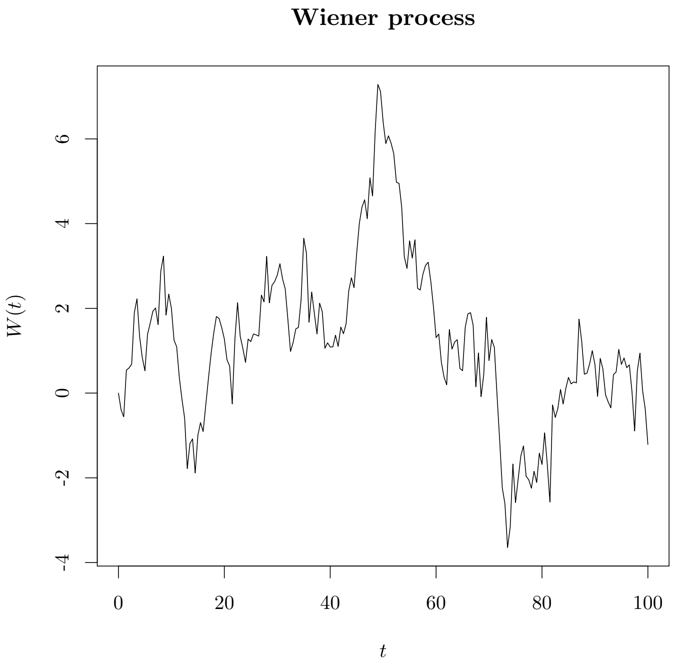
If we write $w' = w + dw$ then the **propagator of a Wiener process** is

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

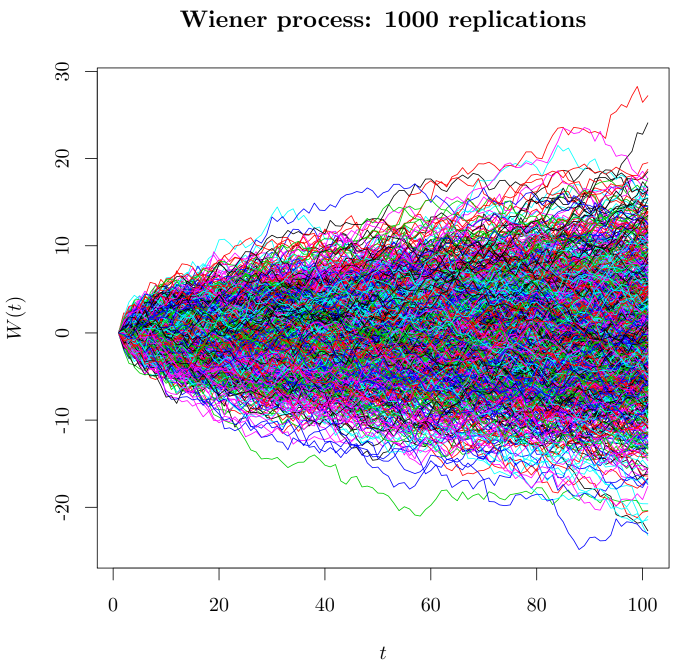
There are several ways of charaterizing stochastic processes. Next we characterize the Wiener process by its **sample path** and **statistic** (or stochastic) properties.

Figure 12.1 depicts a sample path (upper diagram) and 100 draws of the process which allows for an illustration of its stochastic properties. In the upper diagram we see that every sample path is strongly jagged: although it looks like being continuous, it is not smooth enough to be differentiable. In the lower diagram we see that the distribution seems to change with time although it tends to be located most of the time close to $W = 0$.

Next we will make a heuristic confirmation of those perceptions.




(a) One replication




(b) 100 replications

Figure 12.1: Sample paths for the Wiener process

**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

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

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

$$\mathbb{E}\left[\left|\frac{dW(t)}{dt}\right|\right] = \frac{d}{dt}\Big(\sqrt{\frac{2dt}{\pi}}\Big) = \sqrt{\frac{2}{\pi dt}} = o\Big(\frac{1}{\sqrt{dt}}\Big).$$

The time derivative is of order $dt^{-1/2}$ meaning that as $\lim_{dt\to 0} \mathbb{E}\left[\left|\frac{dW(t)}{dt}\right|\right] = \infty$ which means that the sample path of $(W(t))$ is not first-order differentiable in time. □

Therefore, although we can write the process in the integral form

$$W(t) = W(0) + \int_0^t dW(t) = \int_0^t dW(s),$$

from the fundamental theorem of calculus, but the derivative

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

is not well defined.

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

**Statistic properties**

Looking again to Figure 12.1 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 $dt > 0$.*

*1 The Wiener process is **stationary** in expected value terms*

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

*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

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

from the properties of the Gaussian integral (see the Appendix); (2) the quadratic variation $(dW(t))^2 = (dW)^2$ has the expected value

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

(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 equal to zero*

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

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

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

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

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

*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 is*

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

*Proof.* (1) As $W(t) = \int_0^t dW(t)$ then $\mathbb{E}[W(t)] \;= \mathbb{E}\Big[\int_0^t dW(t)\Big] = \int_0^t \mathbb{E}[dW(s)] \;= 0$, for ant $t \in \mathrm{T}$; (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 ant $t \in \mathrm{T}$; (4) for the covariance

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

□

## 12.3 The Itô's processes

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 2.** *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 $f(t, \omega)$,where $\omega$ is adapted to the Wiener process, that is $\omega$ is a function of $W(s)$ with $s \le t$ (I.e, past realizations of $(W(t))$). 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 $s \le t$.

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

#### 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

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

- The variance of the Itô's integral is

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

  (see the proof next).

### 12.3.1 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

**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 $Y(t)$*

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

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

$$dY(t) = g_t(t, X(t))dt + g_x(t, X(t))dX(t) + \frac{1}{2}g_{xx}(t, X(t))(dX(t))^2. \tag{12.6}$$

*where*

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

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

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

*Proof.* Itô (1951) □

Examples

- Let $dX(t) = dW(t)$ and $Y(t) = g(t, X(t))$ then

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

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

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

We can use the previous formulas for proving that

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

where $f(.)$ is deterministic. and $(T(t))$ is an Itô integral.

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

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

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

**Examples** Let $dX(t) = dW(t)$ and $Y(t) = g(X(t))$. We can determine $dY(t)$ for 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

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

### 12.3.2 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 $f(x_1, x_2)$ [2]:

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

Then

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

[2] 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 $\omega$ is a function of past values of the Wiener process $(W(t))$.

**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.

## 12.4 Characterizing Itô's processes

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

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dW(t), \text{ for each } t \in [0, \infty) \tag{12.7}$$

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 to characterise the properties of the diffusion process. In the next chapter we will apply them to the solutions of linear SDE's.

### 12.4.1 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 (**??**), 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*

$$dY(t) = \mu_y(Y(t))dt + \sigma_y(Y(t))dW(t). \tag{12.8}$$

*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, 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.

### 12.4.2 Dynamics of the density: the Kolmogorov forward equation

Consider again the diffusion process specified in equation (12.7).

We write **unconditional probability** as

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

This 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), that is $X(t) = x$, when we observe that at time $t = 0$ it is equal to $x_0$. We can see the initial state as a Dirac-delta distribution $p(0,x) = \delta(x - x_0)$.

We introduce the following assumption: the support of $x$ is the whole set of real numbers $\mathbb{R}$, 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 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** [**adjoint operator**] Let $X(t) = x$ be the realization of the diffusion process (12.7) at time $t > 0$ and let $p(t,x)$ be its unconditional probability. Then the following operator

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

is called adjoint operator.

The following (forward) partial differential equation

$$\frac{\partial p(t,x)}{\partial t} = -\frac{\partial(\mu(x)\,p(t,x))}{\partial x} + \frac{1}{2}\frac{\partial^2(\sigma(x)^2\,p(t,x))}{\partial x^2} \tag{12.9}$$

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

**Proposition 4** (Forward density dynamics)**.**
*Assume the initial state is $x_0$ at $t=0$, that is $X(0)=x_0$. Then the density distribution of $X(t)$ at time $t>0$, when $X(t)$ is the solution to the problem*

$$\begin{cases} p_t(t,x) = \mathcal{G}^*[p](t,x) & \text{for } t>0 \\ p(0,x) = \delta(x-x_0) & \text{for } t=0 \end{cases} \tag{12.10}$$

*where $\mathcal{G}^*[(.)$ is the adjoint operator, together with the initial condition ).*

*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

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

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

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

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

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

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 (12.9) 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 $p(x,0) = \mathbb{P}[X(0)] = \delta(x - x_0)$ 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 9 that the solution to this problem is

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

**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

$$p(t,x) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\dfrac{(x - x_0 - \mu t)^2}{2\sigma^2 t}}. \tag{12.11}$$

### 12.4.3 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

$$X(t) = X(0) + \int_0^t \mu(X(s))ds + \int_0^t \sigma(X(s))dW(s). \tag{12.12}$$

**Proposition 5.** *Consider the diffusion integral form in equation* (12.12) *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

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

□

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.

## 12.5 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

### 12.5.1 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 $G(t, x)[f]$ ,

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

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

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

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.

### 12.5.2 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 $\phi_T$ 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 (see next subsection) the probability satisfies

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

**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.

### 12.5.3 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 $(X(t))_{t\in[0,T]}$, 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

$$\begin{cases} v_t(t, x) = -G(t, x)[v] + v(t, x)f(x), & 0 < t \leq T \\ v(T, X(T)), & T \end{cases} \tag{12.13}$$

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* (12.13) *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

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

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

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

if we use the PDE in problem (12.13). 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. □

## 12.6 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)

## 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

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

# Chapter 13

# Linear scalar stochastic differential equations

## 13.1 Introduction

In this chapter we provide explicit solutions and the statistics for forward scalar linear stochastic differential equations (SDE), which are the simplest SDE's.

These equations have the general form

$$dX(t) = \Big(\mu_0(t) + \mu_1(t)\, X(t)\Big)\, dt + \Big(\sigma_0(t) + \sigma_1(t)\, X(t)\Big)\, dW(t). \tag{13.1}$$

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. We also

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) ODE

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

we saw that the solution is

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

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 sotutions of a linear SDE.

## 13.2 Autonomous equations

In this section we consider the linear autonomous forward SDE,

$$dX(t) = \big(\mu_0 + \mu_1\, X(t)\big)\, dt + \big(\sigma_0 + \sigma_1\, X(t)\big)\, dW(t). \tag{13.2}$$

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.

### 13.2.1 Brownian motion

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

$$dX = \mu dt + \sigma dW(t),\ t \in \mathbb{R}_+ \tag{13.3}$$

with $X(0) = x_0 \in \mathbb{R}$ and $\sigma > 0$. This is a special case of equation (13.2) with $\mu_1 = \sigma_1 = 0$ and $\mu_0 = \mu$ and $\sigma_0 = \sigma$.

The solution of equation (13.3), 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, $W(0) = 0$.

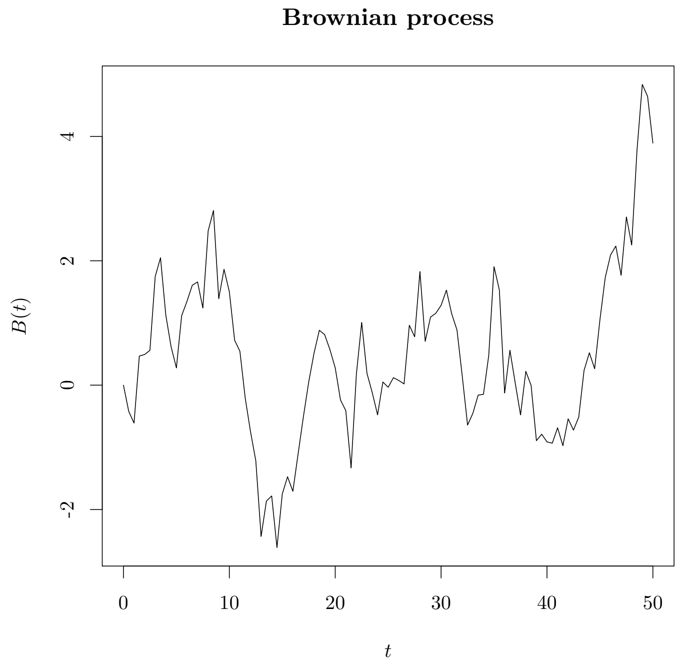
Figure 13.1 presents one sample path in panel (a) and 100 sample paths for the case in which $\mu = -0.5$ and $\sigma = 1$.

The probability distribution is given by equation (12.11)

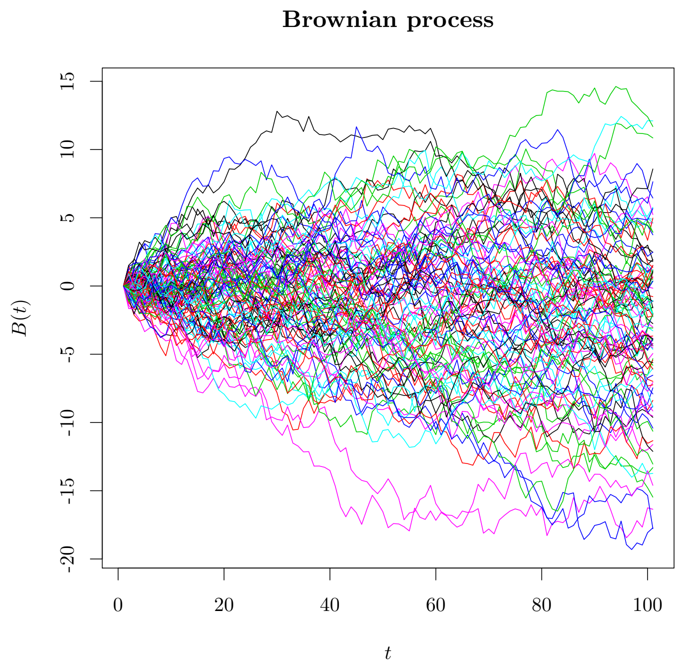
$$p(t, x) = \frac{1}{\sigma\sqrt{2\pi t}}\, e^{-\frac{(x - x_0 - \mu t)^2}{2\sigma^2 t}},\ (t, x) \in \mathbb{R}_+ \times \mathbb{R}.$$

**Properties** The first and second moments are functions of time

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




(a) One replication




(b) 100 replications

Figure 13.1: Sample path for the Brownian process for $\mu = -0.5$ and $\sigma = 1$.

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 $\frac{dx(t)}{dt} = \mu$, given $x_0$ is $x(t) = x_0 + \mu\, t$.

### 13.2.2 Geometric Brownian motion

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

$$dX(t) = \mu\, X(t)dt + \sigma\, X(t)dW(t),\ t \in \mathbb{R}_+, \tag{13.4}$$

where $X(0) = x_0$ with $\mathbb{P}[X(0) = x_0] = 1$. This is a special case of equation (13.2) with $\mu_0 = \sigma_0 = 0$ and $\mu_1 = \mu$ and $\sigma_1 = \sigma$.

The explicit solution is

$$X(t) = x_0 e^{\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma\, W(t)},\ t \in \mathbb{R}_+. \tag{13.5}$$

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, because $x = e^y$, equation (13.5) results.

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

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

Next we assume that $x \in \mathbb{R}_+$, which means that $x_0 > 0$. The solution to this problem is

$$p(t,x) = \frac{x_0}{x\sigma\sqrt{2\pi t}} e^{-\frac{\left(\ln(x/x0) - \left(\mu - \frac{1}{2}\sigma^2\right) t\ \right)^2}{2\sigma^2 t}}, \text{for } (t,x) \in \mathbb{R}^2_+. \tag{13.6}$$

To prove this result, we derive the Fokker-Planck equation and the associated initial condition

$$\begin{cases} \partial_t p(t,x) = \frac{\sigma^2}{2}\ x^2\, \partial_{xx} p(t,x) + (2\sigma^2 - \mu)\, \partial_x p(t,x) + (\sigma^2 - \mu)\, p(t,x) & t \geq 0 \\ p(0,x) = \delta(x - x_0) & t = 0 \end{cases}$$

Performing a transformation of variables $x = e^z$ mapping $x : \mathbb{R}_+ \to \mathbb{R}$, we write $u(t,z) = p(t, x(z))$. 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 Fokker-Planck equation is equivalent to the linear parabolic, with constant coefficients, parabolic PDE,

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

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 $p(t,x) = u(t, \ln(x) - \ln(x_0)) = g(t, \ln(x/x_0))$ as in equation (13.6).

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 13.2 we plot one sample path and several sample paths for the linear diffusion equation where $\mu < 0$ and $\sigma > 0$ and in Figure 13.3 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.

### 13.2.3 Ornstein-Uhlenback processes

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

$$dX = \theta\,(\mu - X)\,dt + \sigma dW(t) \tag{13.7}$$

where $X(0) = x_0$. This is a special case of equation (13.2) with $\mu_0 = \theta\,\mu$, $\mu_1 = -\theta$, $\sigma_0 = \sigma$ and $\sigma_1 = 0$.

The solution is

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

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

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

Integrating on time we have

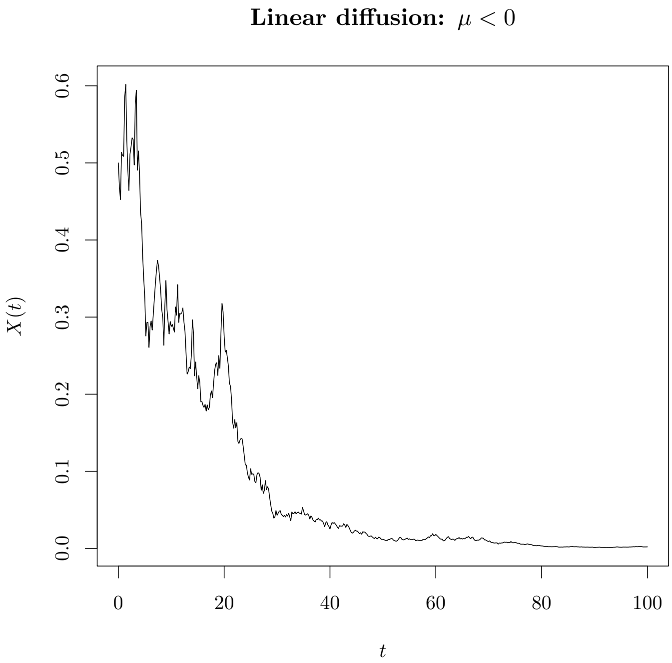
$$Y(t) = y_0 + y_0 + \theta\mu\int_0^t e^{\theta s}ds + \sigma\int_0^t e^{\theta s}dW(s).$$

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 (13.7)

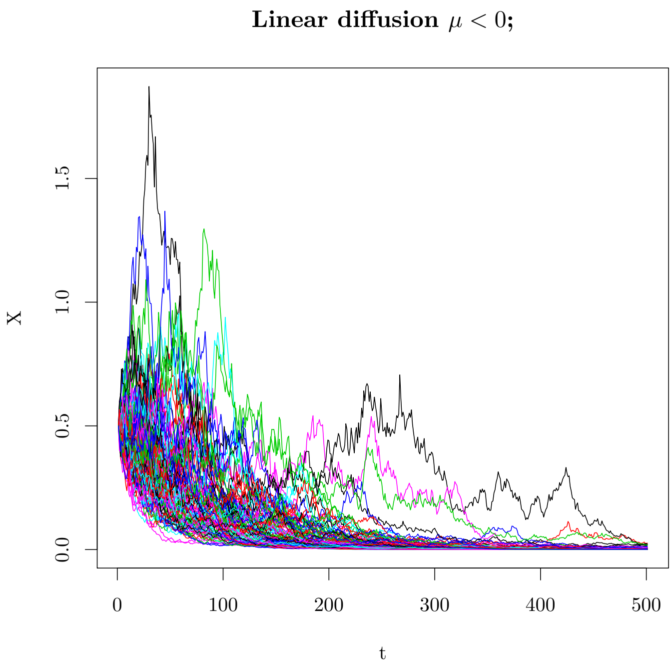
$$\begin{aligned} X(t) &= e^{-\theta t}\left(y_0 + \theta\mu\int_0^t e^{\theta s}ds + \sigma\int_0^t e^{\theta s}dW(s)\right) \\ &= x_0 e^{-\theta t} + \mu e^{-\theta t}\left(e^{\theta t} - 1\right) + \sigma\int_0^t e^{-\theta(t-s)}dW(s). \end{aligned}$$

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

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

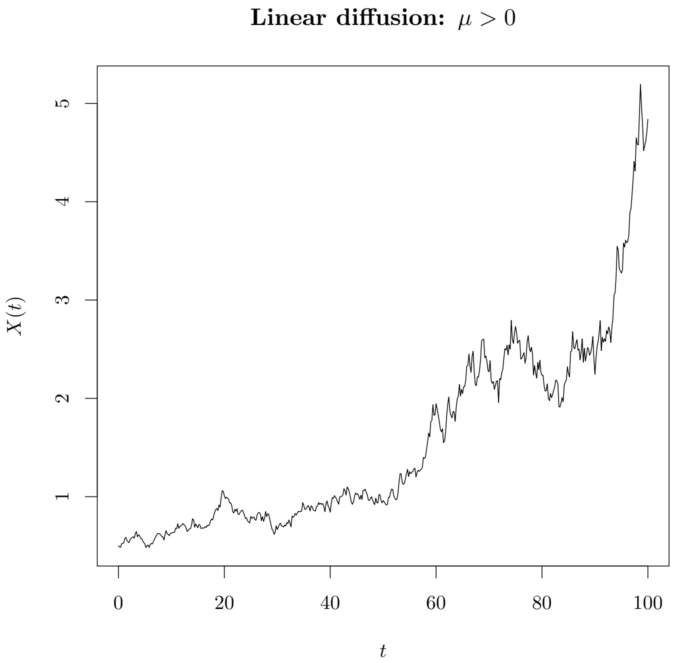



(a) One replication

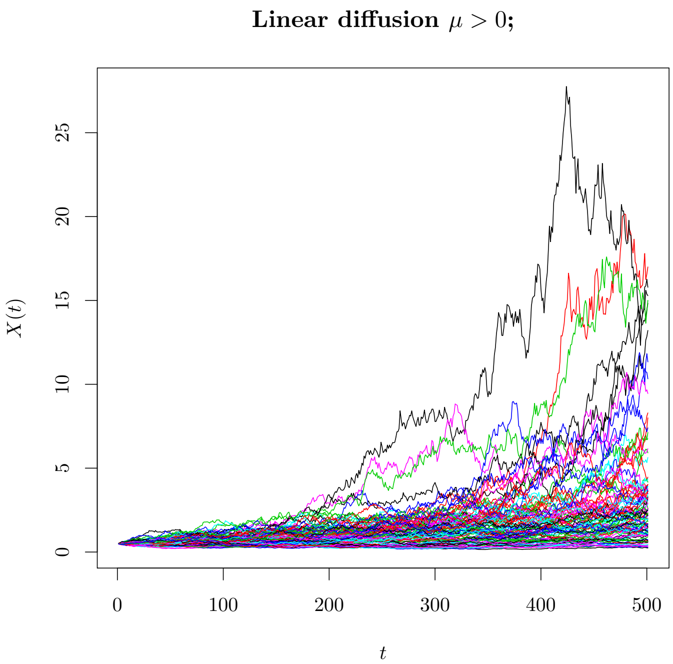



(b) 100 replications

Figure 13.2: Sample paths for the linear diffusion process with $\mu < 0$




(a) One replication




(b) 100 replications

Figure 13.3: Sample paths for the linear diffusion process with $\mu > 0$

The solution to this problem is

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

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 $\theta$. 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

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

has the solution

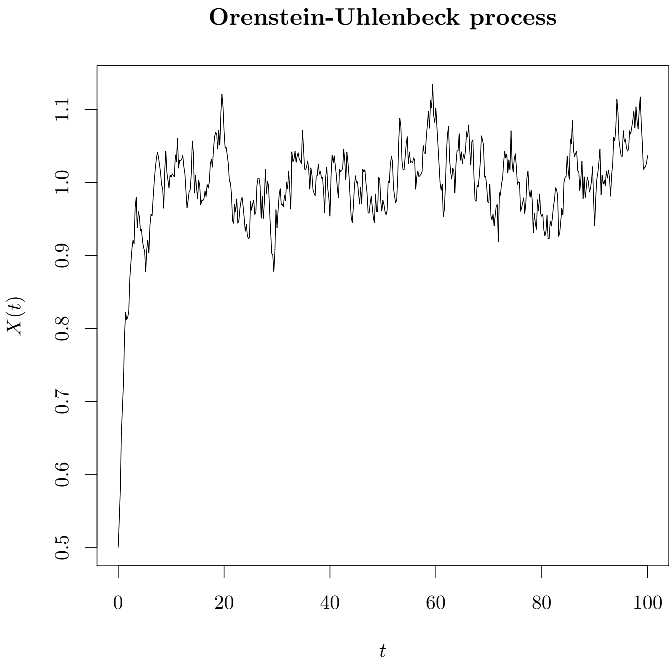
$$x(t) = \mu + (x_0-\mu)\,e^{-\theta\, t}$$

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

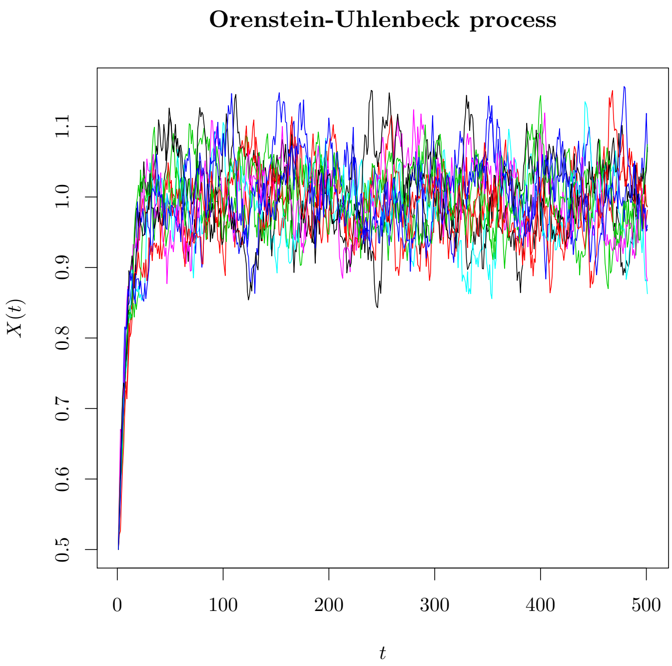
The sample paths for the case $\theta > 0$ are illustrated in figure 13.4: we see that they converge in average to $X(t)=\mu$, however this value is not an attractor, that is the solution although not stationary is ergodic.

### 13.2.4 The linear autonomous SDE

Now consider equation the general linear Itô-SDE (13.2) with $X(0)=x_0$.




(a) One replication




(b) 100 replications

Figure 13.4: Sample paths for Ornstein-Uhlenbeck process for $\theta > 0$ and $\mu = 1$

It can be proved that it has an explicit solution which is

$$X(t) = \Phi(t)\left(x_0 + (\mu_0 - \sigma_0\,\sigma_1)\int_0^t \Phi(s)^{-1}\,ds + \sigma_0\int_0^t \Phi(s)^{-1}\,dW(s)\right) \tag{13.8}$$

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 (13.2) 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 $\mu_1 < 0$ 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

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

### 13.2.5 Stochastic dynamic properties of the linear autonomous SDE

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

1. if $\mu_1 + \sigma_1^2 < 0$ 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 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 $x = -\dfrac{\mu_0}{\mu_1}$ 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.

## 13.3 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

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

and $\Phi(0) = 1$

## 13.4 Economic applications

### 13.4.1 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 $\mu$ is the rate mean rate of growth of population and $\sigma$ 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

$$dK = sLf(k)dt$$

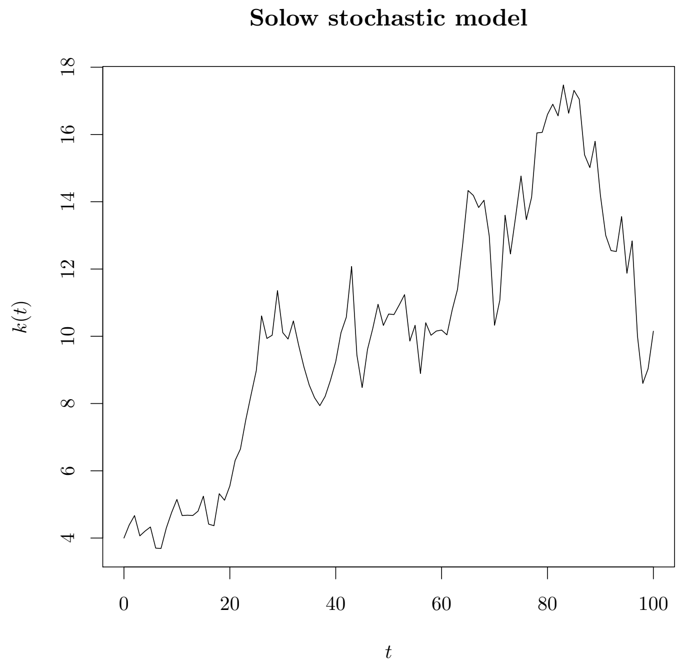



Figure 13.5: Sample path for the capital intensity: $s = 0.1$, $\alpha = 0.3$, $\mu = 0.01$, $\sigma = 0.1$

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

$$dk = \left(sf(k) - (\mu - \sigma^2)k\right) dt - k\sigma dW(t) \tag{13.9}$$

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 13.5 and 13.6 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)

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

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

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

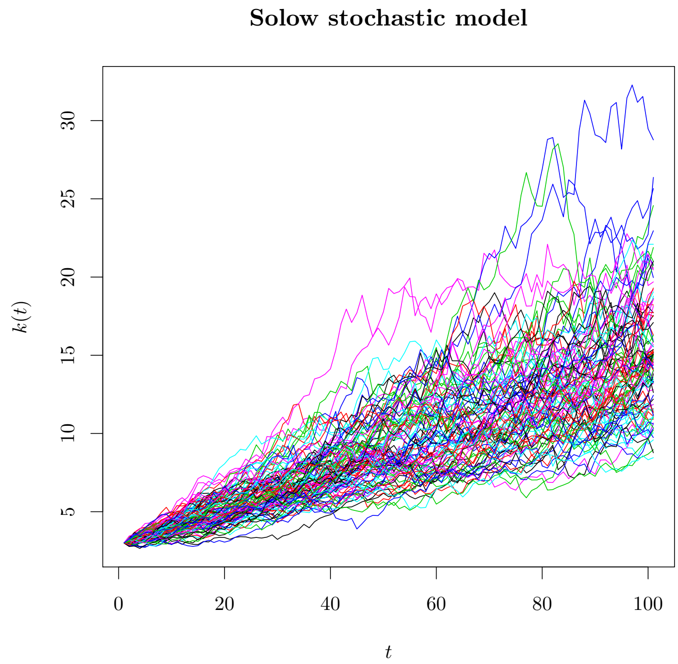



Figure 13.6: Sample paths for the capital intensity: $s = 0.1$, $\alpha = 0.3$, $\mu = 0.01$, $\sigma = 0.1$, 100 replications

### 13.4.2 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 $0 \leq t \leq T$. 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 $\theta$ 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:

$$\begin{cases} V_t(S,t) = -\frac{\sigma^2}{2}S^2V_{ss}(S,t) - rSV_s(S,t) + rV(S,t), & (S,t) \in (0,\infty) \times [0,T] \\ V(S,T) = \max\{S-K,0\}, & (S,t) \in (0,\infty) \times \{\ t=T\} \end{cases} \tag{13.10}$$

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 $\sigma$ are estimated or conjectured.

## 13.5 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)

# Chapter 14

# Stochastic optimal control

## 14.1 Introduction

In this chapter we identify the stochastic optimal control problem as an optimal control problem of an Itô forward stochastic differential equation (FSDE) together with an initial condition on the state variable and some cases in which there are terminal conditions. We deal with both the finite and the infinite horizon cases. We, again, present the simplest problems, present heuristic proofs, and are mostly concerned with characterizing solutions.

There are three approaches to solving the stochastic optimal control problem: (1) using the principle of dynamic programming (DP); (2) using the Pontriyagin maximum principle (PM); and (3) the convex duality method (see Pham (2009)).

The first method is the most well known (see Fleming and Rishel (1975) or Malliaris and Brock (1982) for applications in economics and finance) and leads to the solution of a parabolic PDE, or a second order ODE for infinite horizon problems. The second method is less well known and leads directly to a system of forward-backward stochastic differential equations (FBSDE). The third method is used in association to the Malliavin calculus and is still new. It is not presented in the following notes.

## 14.2 Stochastic dynamic programming

### 14.2.1 Finite horizon

Again we assume the filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, \mathbb{P}\right)$, where a non-anticipating filtration is generated by a Wiener process $\{W(t) : t \in \mathbb{R}_+\}$. This means that all the information is given by the past.

We consider the stochastic optimal control problem, that consists in determining the value function, $V(.)$,

$$V(x_0) = \max_{(U(t))_{t\in[0,T]}} \mathbb{E}_0\left[\int_0^T f(t, X(t), U(t))dt\right] \tag{14.1}$$

subjected to

$$dX(t) = g(t, X(t), U(t))dt + \sigma(t, X(t), U(t))dW(t) \tag{14.2}$$

given the initial distribution for the state variable $X(0) = x_0$. We call $U(.)$ the control variable and assume that the objective, the drift and the volatility functions, $f(.)$, $g(.)$ and $\sigma(.)$. Function $g(.)$ is assumed to be of class $H$ and functions $f(.)$ and $\sigma(.)$ are of class N.

One important difference as regards deterministic optimal control is that while in this case the control variable, together with the transversality condition can be seen as a backward looking variable, in the stochastic case it should be a $\mathcal{F}_t$-adapted process. Therefore, some type of terminal condition should be imposed.

The stochastic dynamic programming principle is the analogue to the dynamic programming principle for the optimal control of ODE's. It gives a local necessary condition for optimality.

**Proposition 1.** ***Stochastic dynamic programming*** *Let the processes* $(X^*(t), U^*(t))_{t\in[0,T]}$ *be solution to the SOC problem* (14.1)-(14.2)*. Then, at time t, the realizations of the state and control variables,* $X^*(t) = x$ *and* $U^*(t) = u$*, satisfy the* ***Hamilton-Jacobi-Bellman*** *equation*

$$-\frac{\partial V(t,x)}{\partial t} = \max_u \left( f(t,x,u) + g(t,x,u)\frac{\partial V(t,x)}{\partial x} + \frac{1}{2}\sigma(t,x,u)^2 \frac{\partial^2 V(t,x)}{\partial x^2} \right). \tag{14.3}$$

*Proof.* (Heuristic) Observe that a solution of the problem satisfies

$$\begin{aligned} V(0,x_0) &= \max_{(u(t))_{t\in[0,T]}} \mathbb{E}_0 \left( \int_0^T f(t,X(t),U(t))dt \right) = \\ &= \max_{(u(t))_{t\in[0,T]}} \mathbb{E}_0 \left( \int_0^{\Delta t} f(t,X(t),U(t))dt + \int_{\Delta t}^T f(t,X(t),U(t))dt \right) \end{aligned}$$

by the principle of the dynamic programming and the law of iterated expectations we have

$$\begin{aligned} V(x_0) &= \max_{(u(t))_{t\in[0,\Delta t]}} \mathbb{E}_0 \left[ \int_0^{\Delta t} f(t,X(t),U(t))dt + \max_{(u(t))_{t\in[\Delta t,T]}} \mathbb{E}_{\Delta t} \left[ \int_{\Delta t}^T f(t,X(t),U(t))dt \right] \right] \\ &= \max_{(u(t))_{t\in[0,\Delta t]}} \mathbb{E}_0 \left[ f(t,X(t),U(t))\Delta t + V(\Delta t, x(\Delta x)) \right] \end{aligned}$$

if we write $x(\Delta t) = x_0 + \Delta x$. If $V$ is continuously differentiable of the second order, the Itô's lemma may be applied to get, for pair $(t, x(t)) = (t, x)$

$$V(t+dt, x+dx) = V(t,x) + V_t(t,x)dt + V_x(t,x)dx + \frac{1}{2}V_{xx}(t,x)(dx)^2 + h.o.t$$

where

$$\begin{aligned} dx &= g(.)dt + \sigma(.)dW \\ (dx)^2 &= g(.)^2(dt)^2 + 2g(.)\sigma(.)(dt)(dW) + (\sigma(.))^2(dW)^2 = (\sigma(.))^2 dt. \end{aligned}$$

Then,

$$\begin{aligned} V &= \max_u \mathbb{E}\left[ f dt + V + V_t dt + V_x g dt + V_x \sigma dW + \frac{1}{2}\sigma^2 V_{xx} dt \right] \\ &= \max_u \left[ f + V_t + V_x g + \frac{1}{2}\sigma^2 V_{xx} \right] dt + V \end{aligned}$$

because $\mathbb{E}_0(dW) = 0$. The equation is only true if and only if the HJB equation holds.

□

### 14.2.2 Infinite horizon

The autonomous discounted infinite horizon problem is

$$V(x_0) = \max_u \mathbb{E}_0 \left[ \int_0^\infty f\left(X(t), U(t)\right)\, e^{-\rho t} dt \right] \tag{14.4}$$

where $\rho > 0$, subject to

$$dX(t) = g\left(X(t), U(t)\right) dt + \sigma\left(X(t), U(t)\right) dW(t) \tag{14.5}$$

given the initial distribution of the state variable $X(0) = x_0$, and assuming the same properties for functions $f(.)$, $g(.)$ and $\sigma(.)$.

Applying, again, the Bellman's principle, now the HJB equation is the nonlinear second order ODE of the form

$$\rho V(x) = \max_u \left( f(x,u) + g(t,x,u)V'(x) + \frac{1}{2}\sigma(x,u)^2 V''(x) \right). \tag{14.6}$$

**References** (Kamien and Schwartz, 1991, cap. 22).

### 14.2.3 Economic applications using stochastic dynamic programming

#### The representative agent problem

The Merton (1971) model is the standard micro model for the simultaneous determination of the strategies of consumption and portfolio investment. Next, we present a simplified version with one risky and one risk-free asset.

Assume that an agent can invest in two types of assets, a risk-free and a risky asset, whose prices are denoted by $B$ and $S$, respectively. We denote by $\theta_0(t)$ and $\theta_1(t)$ the number of risk free and risky assets in the portfolio, and by $A(t)$ net financial wealth of the agent at time $t$, we have $A(t) = \theta_0(t)B(t) + \theta_1(t)S(t)$, for any $t \in [0, \infty)$. The agent can have a short or a long position on any asset: if $\theta_j(t) < 0$ ($\theta_j(t) > 0$) this means that the agent has a short (long) position in asset $j$ at time $t$.

The prices of the assets are given to the agent and are assume to follow the exogenous processes

$$\begin{aligned} dB(t) &= rB(t)dt \\ dS(t) &= \mu S(t)dt + \sigma S(t)dW(t) \end{aligned}$$

where $r$ is the risk-free interest rate, $\mu$ and $\sigma$ are the constant rates of return and volatility for the risky asset. The change in financial income in the time interval $dt$, starting at time $t$, is therefore,

$$\theta_0(t)\, r\, B(t)\, dt + \theta_1(t)(\mu S(t)dt + \sigma S(t)dW(t)).$$

Assume that the agent is entitled to a deterministic endowment $\{y(t), t \in \mathbb{R}\}$ which adds to the financial income. Then the value of financial wealth at time $t$ is

$$A(t) = A(0) + \int_0^t (r\theta_0(s)B(s) + \mu\theta_1(s)S(s) + y(s) - c(s))\, ds + \int_0^t \sigma\mu\theta_1(s)S(s)dW(s),$$

where the process for consumption $\{c(t), t \in \mathbb{R}\}$ is endogenous. Denoting the shares of the equity and of the risk-free asset by $w = \frac{\theta_1 S}{A}$ and $1-w = \frac{\theta_0 B}{A}$, the budget constraint is the Itô's stochastic differential equation

$$dA(t) = \Big[\Big(r\,(1-w(t)) + \mu\, w(t)\Big)\, A(t) + y(t) - c(t)\Big]\, dt + \sigma\, w(t)\, A(t)\, dW(t), \text{ for } t \geq 0 \tag{14.7}$$

and the initial net wealth $A(0) = \theta_0(0)B(0) + \theta_1(0)S(0)$ is known. The rate of return on the total asset position $r^a(t) = r\,(1-w(t)) + \mu\, w(t)$ is a weighted sum of the rates of return of the risk-free and the risky asset, and there is time-varying.

The problem for the consumer-investor is

$$\max_{c,w} \mathbb{E}_0 \left[\int_0^\infty u(c(t))e^{-\rho t} dt\right] \tag{14.8}$$

subject to the instantaneous budget constraint (14.7), given $A(0) = a_0$ and assuming that the utility function is increasing and concave.

This is a stochastic optimal control problem with infinite horizon, and has two control variables, $c$ and $w$. We solve it by using proposition 1.

The Hamilton-Jacobi-Bellman equation (14.6) is

$$\rho\, V(A) = \max_{c,w} \left\{u(c) + V'(A)[(r(1-w) + \mu w)A + y - c] + \frac{1}{2}w^2\sigma^2 A^2 V''(A)\right\}.$$

The first order necessary conditions allows us to get the optimal controls, i.e. the optimal policies for consumption and portfolio composition

$$u'(c^*) = V'(A), \tag{14.9}$$

$$w^* = W(A) = \frac{(\mu - r)}{\varepsilon_v(A)\, \sigma^2} \tag{14.10}$$

where the $\frac{(\mu - r)}{\sigma}$ is the Sharpe index and $\varepsilon_v(A) \equiv -\frac{V'(A)}{AV''(A)}$ is the inverse of the elasticity of the value function.

If $u''(.) < 0$ then the optimal policy function for consumption may be written as $c^* = C(A) \equiv (u')^{-1}(V'(A))$. Substituting the policy functions into the HJB equation, we get the differential equation over $V(A)$

$$\rho V(A) = u(C(A)) + V'(A)(y + rA - C(A)) + \frac{1}{2}\Big(\frac{r-\mu}{\sigma}\Big)^2 \frac{(V'(A))^2}{V''(A)}. \tag{14.11}$$

In some cases, in particular when the utility function is a generalized mean and the constraint is a linear SDE, the HJB equation can be solved explicitly.

**Example: the CRRA case** In particular, let the utility function display constant relative risk aversion (CRRA)

$$u(c) = \frac{c^{1-\eta} - 1}{1-\eta}, \text{ for } \eta > 0,$$

and define total net wealth

$$N = N(A) = \frac{y}{r} + A,$$

as the sum of human wealth $(\frac{y}{r})$ and net financial wealth.

We can solve equation solve equation (14.11) by using the method of undetermined coefficients.

Conjecture that the solution for equation (14.11) is of type

$$V(A) = \alpha + \theta\, N(A)^{1-\eta}$$

where $\alpha$ and $\theta$ are arbitrary constants to be determined. If the functional form of this function is correct, by substituting in equation (14.11) the state variable, we obtain the HJB equation, at the optimum, containing only the unknowns $\alpha$ and $\theta$. By finding a particular solution of that equation we find particular values for those two coefficients.

First, as

$$V^{'}(A) = \theta\,(1-\eta)\,N^{-\eta}, \text{ and } V^{''}(A) = -\theta\,\eta\,(1-\eta)\,N^{-\eta-1}$$

then the optimal policy functions are: for consumption is

$$C(A) = \left(\theta(1-\eta)\right)^{-\frac{1}{\eta}} N(A)$$

which requires that $\theta\,(1-\eta) > 0$ to be a real number, and for the portfolio composition is

$$W(A) = \frac{(\mu - r)}{\sigma^2}\ \frac{N}{\eta\, A}.$$

Substituting in (14.11), we obtain

$$\begin{aligned} \rho\Big(\alpha + \theta N^{1-\eta}\Big) \ &= \ \frac{1}{1-\eta}\Big(\left(\theta\,(1-\eta)\right)^{\frac{\eta-1}{\eta}} N^{1-\eta} - 1\Big) + \\ &\quad + \Big(\theta\,(1-\eta)\,N^{1-\eta}\Big)\Big(r - \left(\theta\,(1-\eta)\right)^{\frac{-1}{\eta}} - \frac{1}{2\,\eta}\Big(\frac{\mu - r}{\sigma}\Big)^2\Big). \end{aligned}$$

If we set $\alpha\,\rho\,(1-\eta) + 1 = 0$, we can eliminate $N^{1-\eta}$ and obtain an equation in $\theta$. Solving it, yields

$$\theta = \theta^* \ \equiv \frac{1}{1-\eta}\left[\frac{\rho + r(1-\eta)}{\eta} + \frac{(1-\eta)}{2\eta^2}\Big(\frac{\mu - r}{\sigma}\Big)^2\right]^{-\eta}$$

Then

$$V(A) = \frac{1}{1-\eta}\left\{\ \left[\frac{\rho - r(1-\eta)}{\eta} + \frac{(1-\eta)}{2\eta^2}\Big(\frac{\mu - r}{\sigma}\Big)^2\right]^{-\eta} N^{1-\eta} - \frac{1}{\rho}\right\}.$$

Then the optimal consumption is

$$c^* = \left(\frac{\rho + r(\eta - 1)}{\eta} + \frac{(1-\eta)}{2\eta^2}\left(\frac{\mu - r}{\sigma}\right)^2\right) N,$$

and the share of the risky asset in the portfolio is again

$$w^* = -\frac{(r-\mu)}{\sigma^2} \frac{N}{\eta A}.$$

In the deterministic analogue, with only the risk-free asset, optimal consumption would be

$$c^* = \frac{\rho + r(\eta - 1)}{\eta} N,$$

which means that if $\eta > 1$ consumption will be smaller in the stochastic environment than in the stochastic one.

We see that the consumer cannot eliminate risk, in general. If we write $c^* = \chi N$, where $\chi \equiv \frac{\rho - r(1-\eta)}{\eta} + \frac{(1-\eta)}{2\eta^2}\left(\frac{\mu - r}{\sigma}\right)^2$, then the optimal net wealth is stochastic and follows a geometric Brownian motion

$$dN(t) = \Big[\ \mu_n dt + \sigma_n dW(t)\Big]\ N(t)$$

where

$$\begin{aligned} \mu_n &= r + \left(\frac{\mu - r}{\sigma}\right)^2 \left(\frac{1-\eta}{\eta}\right) - \chi \\ \sigma_n &= \frac{\mu - r}{\sigma}\frac{1-\eta}{\eta}. \end{aligned}$$

Given the initial wealth $n(0) = \dfrac{y}{r} + a_0$, and using the results in the previous chapter, we find that the probability density of a realization $A(t) = a/a_0$ follows a log-normal distribution.

As $c^* = c(N)$, the optimal consumption is also stochastic. Iff we apply Itô's lemma,

$$dC = \chi dN = C\left(\mu_c dt + \sigma_c dW(t)\right)$$

where

$$\begin{aligned} \mu_c &= \frac{r - \rho}{\eta} \\ \sigma_c &= \frac{r - \eta\rho}{\eta} + .\frac{1}{2}\left(\frac{\mu - r}{\sigma}\frac{1-\eta}{\eta}\right)^2 \end{aligned}$$

The sde has the solution

$$C(t) = c(0)\exp\left\{\left(\mu_c - \frac{\sigma_c^2}{2}\right)t + \sigma_c W(t)\right\}$$

where

$$c(0) = (1-\eta)(\theta^*)^{\frac{1}{\eta}} n(0) = (1-\eta)(\theta^*)^{\frac{1}{\eta}}\left(\frac{y + ra_0}{r}\right).$$

The unconditional expected value for consumption at time $t$

$$\mathbb{E}_0[C(t)] = c(0)\, e^{\mu_c t}.$$

The value function follows a stochastic process which is a monotonous function for wealth. The optimal strategy for consumption follows a stochastic process which is a linear function of the process for wealth and the fraction of the risky asset in the optimal portfolio is a direct function of the premium of the risky asset relative to the riskless asset and is a inverse function of the volatility.

**References** Merton (1971), Merton (1990), Duffie (1996) Cvitanić and Zapatero (2004)

### The stochastic Ramsey model

Let $K$ denote the stock of physical capital and $L$ the labor input which is equal to the population (no unemployment, diseases, etc). The economy is represented by the the differential equations

$$\begin{aligned} dK(t) &= (F(K(t), L(t)) - C(t))dt \\ dL(t) &= \mu L(t)dt + \sigma L(t)dW(t) \end{aligned}$$

where we assume that $F(K, L)$ is linearly homogeneous, given the (deterministic) initial stock of capital and labor $K(0) = K_0$ and $L(0) = L_0$. The growth of the labor input (or its productivity) is stochastic.

If we define the variables in intensity terms,

$$k(t) \equiv \frac{K(t)}{L(t)}, \;\; c(t) \equiv \frac{C(t)}{L(t)},$$

we can get an equilvalent representation of the economy by a single stochastic differential equation over $k$. Using the Itô's lemma yields

$$dk = \left(f(k) - c - (\mu - \sigma^2)k\right) dt - \sigma^2 k dW(t) \tag{14.12}$$

where the production function in intensity terms is $f(k) = F\left(\frac{K}{L}, 1\right)$.

There is a central who wants to find the optimal path of the economy maximizing the intertemporal utility functional

$$\mathbb{E}_0\left[\int_0^\infty u(c(t))e^{-\rho t}dt\right]$$

subject to the budget constraint (14.12).

We use the stochastic dynamic programming principle to solve the problem. The HJB equation, (14.6), is

$$\rho V(k) = \max_c \left\{ u(c) + V^{'}(k)\left(f(k) - c - (\mu - \sigma^2)k\right) + \frac{1}{2}(k\sigma)^2 V^{"}(k) \right\}$$

the optimality condition is again

$$u^{'}(c) = V^{'}(k)$$

and, substituting in the HJB equation yields an implicit second-order ODE

$$\rho V(k) = u(h(k)) + V^{'}(k)\left(f(k) - h(k) - (\mu - \sigma^2)k\right) + \frac{1}{2}(k\sigma)^2 V^{"}(k).$$

Again, we assume the benchmark **particular case**: $u(c) = \dfrac{c^{1-\theta}}{1-\theta}$ and $f(k) = k^{\alpha}$. Then the optimal policy function becomes

$$c^* = V^{'}(k)^{-\frac{1}{\theta}}$$

and the HJB becomes

$$\rho V(k) = \frac{\theta}{1-\theta} V^{'}(k)^{\frac{\theta-1}{\theta}} + V^{'}(k)\left(k^{\alpha} - (\mu - \sigma^2)k\right) + \frac{1}{2}(k\sigma)^2 V^{"}(k).$$

This equation does not seem to have a closed form solution.

However, to illustrate how a solution would be obtained in the case in which a closed-form solution would be obtained, we consider the (unrealistic) case $\theta = \alpha$. Again we conjecture that the solution if of the form

$$V(k) = B_0 + B_1 k^{\alpha}$$

Using the same methods as before we get

$$\begin{aligned} B_0 &= (1-\alpha)\frac{B_1}{\rho} \\ B_1 &= \frac{1}{1-\alpha}\left[\frac{(1-\alpha)\theta}{(1-\theta)(\rho - (1-\alpha)^2\sigma^2)}\right]^{\alpha}. \end{aligned}$$

Then

$$V(k) = B_1\left(\frac{1-\alpha}{\rho} + k^{1-\alpha}\right)$$

and

$$c^* = c(k) = \left(\frac{(1-\theta)(\rho - (1-\alpha)^2\sigma^2)}{(1-\alpha)\theta}\right) k \equiv \varrho k$$

as we see an increase in volatility decreases consumption for every level of the capital stock.

Then the optimal dynamics of the per capita capital stock is the SDE

$$dk^*(t) = \left(f(k^*(t)) - (\mu + \varrho - \sigma^2)k^*(t)\right)dt - \sigma^2 k^*(t) dW(t).$$

In this case we can not solve it explicitly as in the deterministic case.

References: Brock and Mirman (1972), Merton (1975), Merton (1990)

## 14.3 The stochastic PMP

Consider again the optimal control problem with value function (14.1).

In order to find the necessary optimality conditions by using the stochastic version of the Pontriyagin maximum principle (SPMP) it is useful to distinguish the case in which the volatility component depends on the control variable, as in equation (14.2), from the case in which it does not, as in equation

$$dX(t) = g(t, X(t), U(t))dt + \sigma(t, X(t))dW(t). \tag{14.13}$$

The reason for this is, again, related to the fact that the control variable should be $\mathcal{F}_t$ adapted.

### 14.3.1 Volatility function independent of the control variable

**Proposition 2.** ***Stochastic PMP*** *Let the processes* $(X^*(t), U^*(t))_{t\in[0,T]}$ *be solution to the SOC problem* (14.1)-(14.13)*. Then, there are two processes* $(p(t), q(t))_{t\in[0,T]}$ *satisfying the adjoint equation and a terminal condition*

$$\begin{cases} dp(t) = -\Big\{ f_x(t, X^*(t), U^*(t)) + p(t)g_x(t, X^*(t), U^*(t)) + q(t)\sigma_x(t, X^*(t))\Big\}\, dt + q(t)dW(t) \\ p(T) = 0 \end{cases}$$

*and, defining the Hamiltonian function by*

$$H(t, x, u, p, q) = f(t, x, u) + pg(t, x, u) + q\sigma(t, x),$$

*the optimal control satisfies for the realizations of the state and the control variables* $X^*(t) = x$ *and* $U^*(t) = u$,

$$H(t, x^*, u^*, p, q) = \max_u H(t, x^*, u, p, q)$$

The proof is in (Yong and Zhou, 1999, p.123-137)

### 14.3.2 Volatility dependent on the control variable

**Proposition 3.** ***Stochastic PMP*** *Let the processes* $(X^*(t), U^*(t))_{t\in[0,T]}$ *be solution to the SOC problem* (14.1)-(14.2)*. Then, there are four processes* $(p(t), q(t), P(t), Q(t))_{t\in[0,T]}$ *satisfying the two adjoint equations and associated terminal conditions*

$$\begin{cases} dp(t) = -\Big\{ f_x(t, X^*(t), U^*(t)) + p(t)g_x(t, X^*(t), U^*(t)) + q(t)\sigma_x(t, X^*(t), U^*(t))\Big\}\, dt + q(t)dW(t) \\ p(T) = 0 \end{cases}$$

*and*

$$\begin{cases} dP(t) & = -\Big\{ f_{xx}(t, X^*(t), U^*(t)) + 2P(t)g_x\ (t, X^*(t), U^*(t)) + P(t)\left(g_x(t, X^*(t), U^*(t))\right)^2 + \\ & +2Q(t)\sigma_x(t, X^*(t), U^*(t))\Big\}\, dt + Q(t)dW(t) \\ P(T) & = 0 \end{cases}$$

*and, defining the Hamiltonian function,*

$$H(t, x, u, p) = f(t, x, u) + pg(t, x, u)$$

*the Generalized Hamiltonian function*

$$G(t, x, u, p, P) = f(t, x, u) + pg(t, x, u) + \frac{1}{2}\sigma^2(t, x, u)P$$

*the optimal control satisfies locally* $X^*(t) = x^*$ *and* $U^*(t) = u^*$ *such that defining*

$$\mathscr{H}(t, x^*, u) = G(t, x^*, u, p, P) + \sigma(t, x^*, u)\left(q - P\sigma(t, x^*, u^*)\right)$$

*it satisfies*

$$\mathcal{H}(t, x^*, u^*) = \max_u \mathcal{H}(t, x^*, u)$$

The proof is in (Yong and Zhou, 1999, p.123-137)

### 14.3.3 Economic applications using stochastic maximum principle

We present next two applications of the stochastic PMP: a stochastic endogenous growth model and, again, the Merton model. In the first case the control variable does not affect the volatility term and in the second it does. This means that we use Proposition 2 in the first case and Proposition 3 in the second.

**Application: the stochastic $AK$ model**

This is a stochastic version of the simplest endogenous growth model:

$$\max_{C(.)} \int_0^T \ln(C(t)) e^{-\rho t} dt$$

subject to

$$dK(t) = (\mu K(t) - C(t))\, dt + \sigma K(t) dW(t) \tag{14.14}$$
$$K(0) = k_0$$

Observe that, as in this case the volatility term is independent of the control variable, $C$, we use proposition 2.

The adjoint equation is

$$\begin{cases} dp(t) = -\left(\mu p(t) + \sigma q(t)\right) dt + q(t) dW(t), \quad t \in (0, T) \\ p(T) = 0 \end{cases}$$

and the Hamiltonian is

$$H(t, c, k, p, q) = \ln(c) e^{-\rho t} + p(\mu k - c) + q\sigma k.$$

We determine optimal consumption such that $C^* = c^*$ by making $\dfrac{\partial H}{\partial c} = 0$. Therefore,

$$C^*(t) = \left(p(t) e^{\rho t}\right)^{-1}.$$

Consumption is a stochastic process, depending on $p$. Using Itô's lemma yields

$$\begin{aligned} dC^*(t) &= -\rho \frac{e^{-\rho t}}{p(t)} dt - \frac{e^{-\rho t}}{p(t)^2} dp(t) + \frac{e^{-\rho t}}{p(t)^3} (dp(t))^2 \\ &= C^*(t) \left( -\rho dt - \frac{dp(t)}{p(t)} + \left( \frac{dp(t)}{p(t)} \right)^2 \right) \\ &= C^*(t) \left[ \left( \mu - \rho + \sigma \frac{q(t)}{p(t)} + \left( \frac{q(t)}{p(t)} \right)^2 \right) dt - \frac{q(t)}{p(t)} dW(t) \right] \end{aligned}$$

We have a stochastic differential equation for $p(.)$ but we do not have one equation allowing for the determination of $q(.)$. Based on our knowledge of the related deterministic model, we introduce a trial relationship

$$C(t) = \phi\, K(t)$$

where $\phi$ is a constant to be determined. Applying the Itô's lemma we have

$$\begin{aligned} dC(t) &= \phi dK(t) \\ &= \phi\left(\left(\mu K(t) - C(t)\right) dt + \sigma K(t) dW(t)\right) \end{aligned}$$

If we match the deterministic and the stochastic components of the two equations for $C$, we have, for any realization of $C(t) = c$, $K(t) = k$, $p(t) = p$, and $q(t) = q$

$$\begin{cases} c\left(A - \rho + \sigma \dfrac{q}{p} + \left(\dfrac{q}{p}\right)^2\right) = \phi(\mu k - c) \\ -c\dfrac{q}{p} = \phi\sigma k \end{cases}$$

that would hopefully allow for the determination of the two unknowns, the realization $q$ and the parameter $\phi$. Solving the system we get $q = -\sigma p$ and $\phi = \rho$. Therefore,

$$C^*(t) = \rho K^*(t)$$

substituting in equation (14.14) yields

$$dK^*(t) = K^*(t)\left((\mu - \rho)dt + \sigma dW(t)\right)$$

Therefore

$$K^*(t) = k_0 e^{(\mu - \rho - \frac{1}{2}\sigma^2)t + \sigma W(t)}$$

and

$$C^*(t) = \rho k_0 e^{(\mu - \rho - \frac{1}{2}\sigma^2)t + \sigma W(t)}$$

meaning that:

1. consumption and capital accumulation are perfectly correlated;

2. they both follow a log-normal process with mean, where

$$\mathbb{E}[K(t)] \; = k_0 e^{(\mu - \rho - \frac{1}{2}\sigma^2)t}$$

3. meaning that there wil be long-run growth if $\mu - \rho - \frac{1}{2}\sigma^2 > 0$ that is if volatility does not affect much total factor productivity.

**The Merton (1990) model**

Next we consider again the problem of maximizing the intertemporal utility functional (14.8) subject to the stochastic differential equation (14.7). Differently from the previous presentation of the Merton's model, we now assume that there is no non-financial income, that is $y = 0$ and the utility function is logarithmic.

We consider the problem

$$\max_{C.w} \mathbb{E}_0 \left[ \int_0^\infty \ln (C(t)) e^{-\rho t} dt \right]$$

subject to budget constraint, represents the dynamics of financial net wealth $N$,

$$dN(t) = [\ (r + (\mu - r)w)\, N - C]\ dt + \sigma w N dW(t)$$

and $N(0) = n_0$ is given and perfectly observed.

In this case there are two control variables, $C$ and $w$, but one control variable, $w$, affects the volatility term. Therefore, we have to apply Proposition 3.

The adjoint equations are

$$\begin{cases} dp(t) = -\left[(r + (\mu - r)w(t))\, p(t) + \sigma w(t) q(t)\right] dt + q(t) dW(t) \\ \lim_{t\to\infty} p(t) = 0 \end{cases}$$

and

$$\begin{cases} dP(t) = -\left[2\Big(r + (\mu - r)w(t)\Big)P(t) + \Big(r + (\mu - r)w(t)\Big)^2 P(t) + 2\sigma w(t) Q(t)\right] dt + Q(t) dW(t) \\ \lim_{t\to\infty} P(t) = 0. \end{cases}$$

To find the optimal controls we write the generalized Hamiltonian

$$G(t, N, C, w, p, P) = e^{-\rho t} \ln (C) + p\, [\ (r + (\mu - r)w)\, N - C]\ \ + \frac{1}{2}\sigma^2 w^2 N^2 P$$

and

$$\mathcal{H}(t, N, C, w) = G(t, N, C, w, p, P) + \sigma w N \left(q - P\sigma w^* N\right).$$

The optimal controls, $C^*$ and $w^*$ are found by maximizing function $\mathcal{H}(t, N, C, w)$ for $C$ and $w$. Therefore, we find

$$C^*(t) = e^{-\rho t} p(t)^{-1} \tag{14.15}$$

and the condition

$$p(t)(\mu - r)N^*(t) + w^*(t)\sigma^2 N^*(t)^2 P(t) + \sigma N^*(t) \left(q(t) - \sigma w^*(t) N^*(t) P(t)\right) = 0$$

which is equivalent to $p(t)(\mu - r)N^*(t) + \sigma q(t) N^*(t) = 0$. Therefore we find

$$q(t) = -p(t) \left(\frac{\mu - r}{\sigma}\right),$$

and, substituting in the adjoint equation,

$$dp(t) = -p(t)\left(rdt + \left(\frac{\mu - r}{\sigma}\right) dW(t)\right).$$

Observe that the structure of the model is such that the shadow value of volatility functions $P$ and $Q$ have no effect in the shadow value functions associated with the drift component $p$ and $q$, which simplifies the solution.

Applying the Itô's formula to consumption (14.15), and using this expression for the adjoint variable $q$, we find

$$\begin{aligned} dC(t) &= -\rho C(t)dt - \frac{C(t)}{p(t)}dp(t) + \frac{C(t)}{p^2(t)}(dp(t))^2 = \\ &= -\rho C(t)dt + C(t)\left(rdt + \left(\frac{\mu - r}{\sigma}\right) dW(t)\right) + C(t)\ \left(\frac{\mu - r}{\sigma}\right)^2 dt = \\ &= C(t)\left\{\left(r - \rho + \left(\frac{\mu - r}{\sigma}\right)^2\right) dt + \left(\frac{\mu - r}{\sigma}\right) dW(t)\right\}. \end{aligned}$$

Now, we **conjecture** that consumption is a linear function of net wealth $C = \xi N$. If this is the case this would allow us to obtain the optimal portfolio composition $w^*$. If the conjecture is right then we will also have

$$\begin{aligned} dC(t) &= \xi dN(t) \\ &= \xi N(t)\left[\ (r + (\mu - r)w - \xi)\ dt + \sigma w dW(t)\right] \\ &= C(t)\left[\ (r + (\mu - r)w - \xi)\ dt + \sigma w dW(t)\right] \end{aligned}$$

This is only consistent with the previous derivation if

$$\begin{cases} r - \rho + \left(\frac{\mu - r}{\sigma}\right)^2 = r + (\mu - r)w - \xi \\ \frac{\mu - r}{\sigma} = \sigma w \end{cases}$$

Solving for $\xi$ and $w$ we obtain the optimal controls

$$C^*(t) = \rho N^*(t) \tag{14.16}$$

$$w^*(t) = \frac{\mu - r}{\sigma^2} \tag{14.17}$$

Substituting in the budget constraint we have the optimal net wealth process

$$\frac{dN^*(t)}{N^*(t)} = \mu_n\ dt + \sigma_n dW(t)$$

where

$$\mu_n = r - \rho + \left(\frac{\mu - r}{\sigma}\right)^2 \tag{14.18}$$

$$\sigma_n = \frac{\mu - r}{\sigma} \tag{14.19}$$

which can be explicitly solved with the initial condition $N^*(0) = n_0$. We also find that

$$\frac{dC^*(t)}{C^*(t)} = \mu_n\ dt + \sigma_n dW(t)$$

the rates of return for consumption and wealth are perfectly correlated.

## 14.4 References

- Application to economics: Malliaris and Brock (1982), Chang (2004), Stokey (2009)
- Applications to finance: asset pricing Björk (2004) and Cvitanić and Zapatero (2004) , credit risk Bielecki and Rutkowski (2004). Advanced Pham (2009).
- Solution by DP methods: Fleming and Rishel (1975) and Seierstad (2009)
- Pontryiagin's principle for SDE: Bensoussan (1988), (Yong and Zhou, 1999, chap. 3)
- A survey on stochastic control: Kushner (2014),

# 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.