Advanced Mathematical Economics

Paulo B. Brito
PhD in Economics: 2019-2020
ISEG
Universidade de Lisboa
pbrito@iseg.ulisboa.pt

Lecture 7
13.2.2020

Contents

Part VII

Parabolic partial differential equations

Chapter 9

Stochastic differential equations

9.1 Introduction

If we consider again the ordinary differential equation

$$\dot{y} = f(y(t)) \tag{9.1}$$

we can extend it by introducing a random perturbation,

$$\dot{Y} = f(Y(t)) + \epsilon(t) \tag{9.2}$$

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

$$\dot{Y} = f(Y(t), \epsilon(t)). \tag{9.3}$$

While the solution of (9.1) is a mapping $y : \mathbb{R}_+ \to \mathbb{R}^n$, in the cases of equations (9.2) and (9.3) 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 stochastic in dynamic models. However, the most common model is called **diffusion equation**

$$dY(t) = f(Y(t), t)dt + \sigma(Y(t), t)dW(t) \tag{9.4}$$

where $(W(t))_{t\geq 0}$ is a Wiener process and $f(.)$ and $\sigma(.)$ are continuous and differentiable known functions. The main reason for this formalism is related to the fact that $Y(t)$ is not differentiable in the classic way and there are stochastic calculus rules provided by the Itô's Lemma.

Therefore, in general, the term *stochastic differential equation* (SDE) is reserved to equations as (9.4) in the differential form or in the integral form

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

where the first integral in the right-hand-side is a Riemmann integral, but the second is 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.

Next we present a very brief introduction to stochastic differential equations following a heuristic approach and with a view to characterizing the properties of the solutions.

9.2 Introduction to stochastic calculus

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

9.2.1 Wiener process

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

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

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

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

The conditional probability (or propagator) is

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

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

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

Sample path properties

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

Proof. (Heuristic) Let

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

for a given $0 < t < \infty$.

Then

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

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

$$\begin{aligned}
\mathbb{E}[|x|] &= \int_{-\infty}^{\infty} \frac{|x|}{\sqrt{2\pi dt}} e^{-\dfrac{x^2}{2dt}} dx = \\
&= \frac{\sqrt{2dt}}{\sqrt{\pi}} \int_{-\infty}^{\infty} \sqrt{\frac{x^2}{2dt}} e^{-\dfrac{x^2}{2dt}} \frac{dx}{\sqrt{2dt}} = \\
&= \sqrt{\frac{2dt}{\pi}}
\end{aligned}$$

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

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

which is of order $dt^{-1/2}$ when dt tends to zero, meaning that it is not first-order differentiable. □

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

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

in the integral form, but

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

is not well defined.

Figure 9.1 presents one sample path and 100 sample path replications of a Wiener process

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

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

- *The Wiener process is stationary in expected value*

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

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

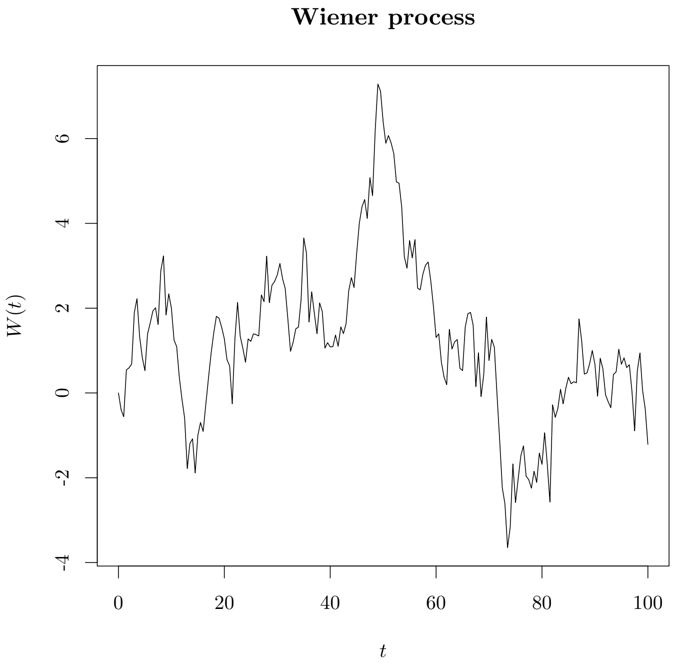
$$\mathbb{E}[(dW(t))^2] = dt$$

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

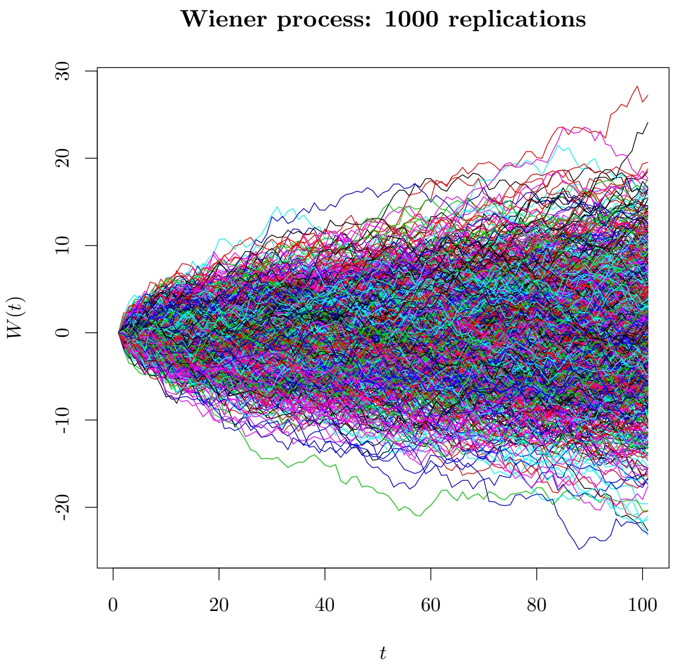
$$\mathbb{V}[dW(t)] = \mathbb{E}[dW(t)^2] - \mathbb{E}[dW(t)]^2 = dt$$

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

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



(a) One replication



(b) 100 replications

Figure 9.1: Sample paths for the Wiener process

- *The correlation coefficient is*

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

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

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

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

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

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

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

For the covariance

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

$\square$

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

9.2.2 The Itô's process

In the definition of the stochastic differential equation, in its integral form, we had the expression

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

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

Let $f(t)$ be a bounded function of time. We write

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

This definition can be extended to functions of type $f(t, w)$.

Definition We call **Itô's integral** to

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

where w is the outcome of a non-anticipating Wiener process, i.e, $w = W(s)$ for $s \leq t$. Also assume that the function is bounded in the sense $\mathbb{E}[\int_0^t f(t)^2 dt] < \infty$.

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

Properties of the Itô's integral

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

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

- The variance variance of the Itô's integral is

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

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

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

- The Itô integral is additive as regards the time integrand

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

  for $0 < t < T$.

9.2.3 The Itô's integral and stochastic calculus

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

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

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

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

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

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

$$dY(t) = f_t(t, X(t))dt + f_x(t, X(t))dX(t) + \frac{1}{2}f_{xx}(t, X(t))(dX(t))^2$$

where the following **Itô's rules** are used

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

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

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

verifies the variation, in its differential form,

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

where

$$\nabla_x f(\mathbf{X}(t), t) = \begin{pmatrix} f_{x_1}(\mathbf{X}(t), t) \\ \vdots \\ f_{x_n}(\mathbf{X}(t), t) \end{pmatrix},\ \nabla_x^2 f(\mathbf{X}(t), t) = \begin{pmatrix} f_{x_1x_1}(\mathbf{X}(t), t) & \dots & f_{x_1x_n}(\mathbf{X}(t), t) \\ \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{X}(t), t) & \dots & f_{x_nx_n}(\mathbf{X}(t), t) \end{pmatrix}$$

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

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

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

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

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

$$\nabla f(x_1, x_2) = \begin{pmatrix} f_{x_1}(x_1, x_2) \\ f_{x_2}(x_1, x_2) \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}, \nabla^2 f(x_1, x_2) = \begin{pmatrix} f_{x_1x_1}(x_1, x_2) & f_{x_1x_2}(x_1, x_2) \\ f_{x_2x_1}(x_1, x_2) & f_{x_2x_2}(x_1, x_2) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Then

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

9.3 The diffusion equation

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

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dW(t) \tag{9.5}$$

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

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

9.3.1 Functions of the diffusion

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

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

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

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

where

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

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

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

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

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

9.3.2 Moment equations

The one-dimensional diffusion equation in integral form is

$$X(t) = X(0) + \int_0^t \mu(X(s))ds + \int_0^t \sigma(X(s))dW(s). \tag{9.6}$$

Proposition 4. *Consider the diffusion integral form in equation* (9.6) *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

$$Y(t) = X(t)^2$$

then, using the Itô's formula

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

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

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

$\square$

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 verifies 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 verifies

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

also verifies 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 $N\left(0, \dfrac{\sigma^2}{2\gamma}\right)$. It is an ergodic process.

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

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

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

9.3.3 Generator of a diffusion

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

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

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

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

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

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

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

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

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

9.3.4 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

$$\begin{cases} v_t(t,x) = -G(t,x)[v] + v(t,x)f(x), & 0 < t \leq T \\ v(T, X(T)), & T \end{cases} \tag{9.7}$$

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 5. *The solution to the PDE problem* (9.7) *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 (9.7). 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. □

9.3.5 Kolmogorov backward equation

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

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

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

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

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

The equation is called **Kolmogorov backward equation**

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

which we want to solve together with with the terminal condition

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

Using the Feynman-Kac the probability verifies

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

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

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

From the Feynman-Kac formula

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

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

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

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

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

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

Then

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

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

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

we would reach the same solution.

9.3.6 Kolmogorov forward equation

The Kolmogorov forward equation, also called the **Fokker-Planck equation**, gives the probability of a somewhat inverse reasoning: if the initial state is x_0 at $t = 0$, that is $X(0) = x_0$, what is the density distribution of $X(t)$ at time $t > 0$, when $X(t)$ follows a diffusion process ?

If $p(t,x)$ represents now the probability, as off time $t = 0$, of $X(t) = x$ at $t > 0$, that is

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

and is given by the equation

$$p_t(t,x) = G^*[p](t,x)$$

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

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

where $p(x_0, 0) = 1$.

Example Let $dX(t) = \sigma dW(t)$ and let $X(0) = x_0$ and let the initial distribution be a Dirac delta with the distribution mass concentrated at x_0.

As we have $\mu(x) = 0$ and $\sigma(x) = \sigma$ then

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

To find the probability distribution for $X(t) = x$ we solve the problem with a forward parabolic PDE and an initial condition:

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

where $\int_{-\infty}^{\infty} p(x_0)dx = \int_{-\infty}^{\infty} \delta(x - x_0)dx = 1$. The solution to this problem is

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

9.4 The linear diffusion equation

We apply some of the previous results to obtain explicit solutions of some simple stochastic differential equation.

9.4.1 Brownian motion

The Brownian motion $B(t)$ is a process generated by the SDE

$$dB = \mu dt + \sigma dW(t)$$

where $B(0) = \phi$

The solution is

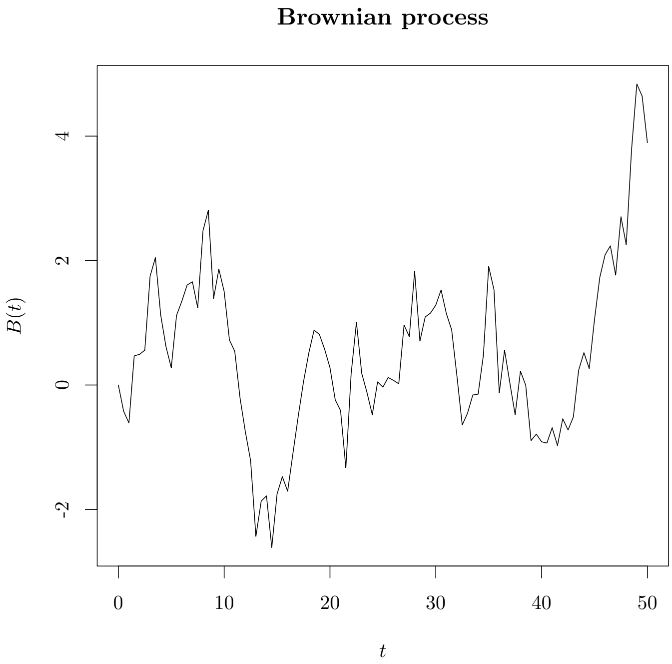
$$B(t) = \phi + \mu t + \sigma \int_0^t dW(t)$$

It has moments

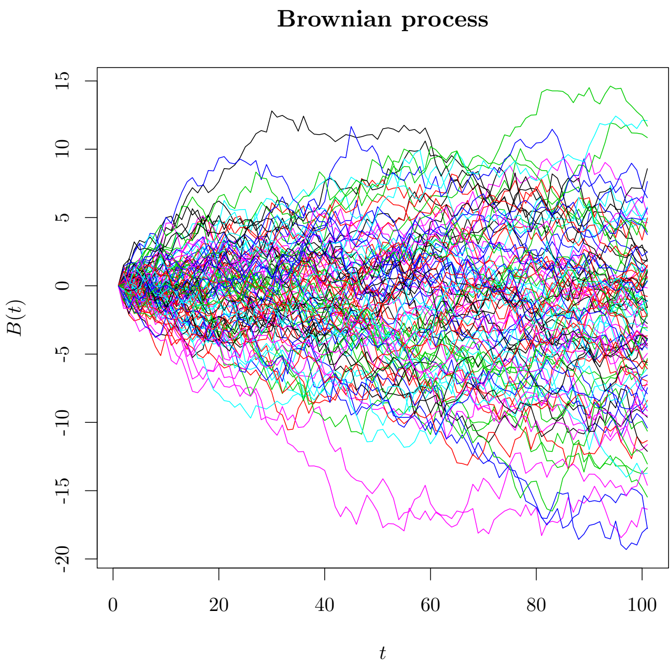
$$\mathbb{E}[B(t)] = B(0) + \mu t, \ \mathbb{V}[B(t)] = \sigma^2 t$$

The probability distribution is

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



(a) One replication



(b) 100 replications

Figure 9.2: Sample path for the Brownian process for $\mu = -0.5$ and $\sigma = 1$.

9.4.2 Linear diffusion or geometric Brownian motion

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

where $X(0) = x_0$ with $\mathbb{P}[X(0) = x_0] = 1$.

The explicit solution is

$$X(t) = x_0 e^{(\mu-\sigma^2/2)t+\sigma W(t)}. \tag{9.8}$$

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

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

Then, because $Y(t) = Y(0) + \int_0^t dY(s)$ we get

$$\ln X(t) = \ln X(0) + \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t)$$

and taking exponential we arrive at equation (9.8).

Properties In figure 9.2 we plot one sample path and several sample paths for the linear diffusion equation where $\mu < 0$ and $\sigma > 0$.

The linear diffusion has the moments

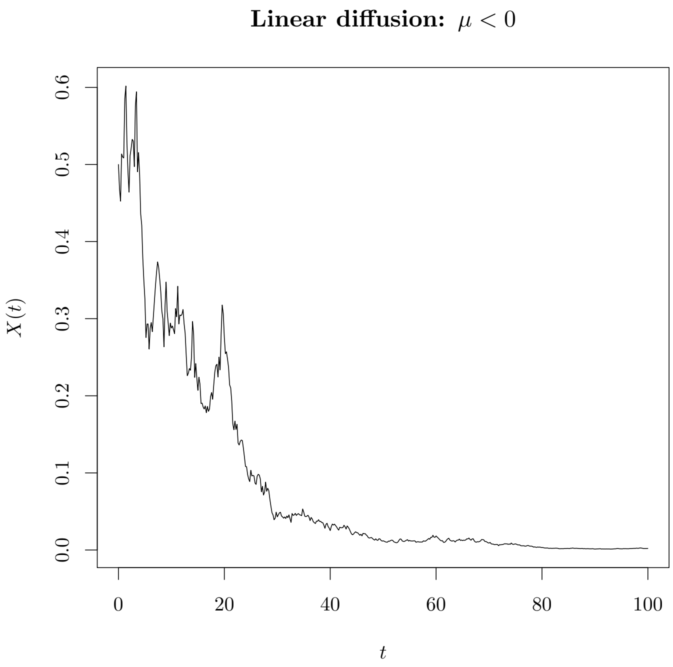
$$\begin{aligned} \mathbb{E}[X(t)] &= x_0 e^{\mu t},\ t \in [0, \infty) \\ \mathbb{V}[X(t)] &= x_0^2 e^{2\mu t}\left(e^{\sigma^2 t} - 1\right),\ t \in [0, \infty) \end{aligned}$$

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

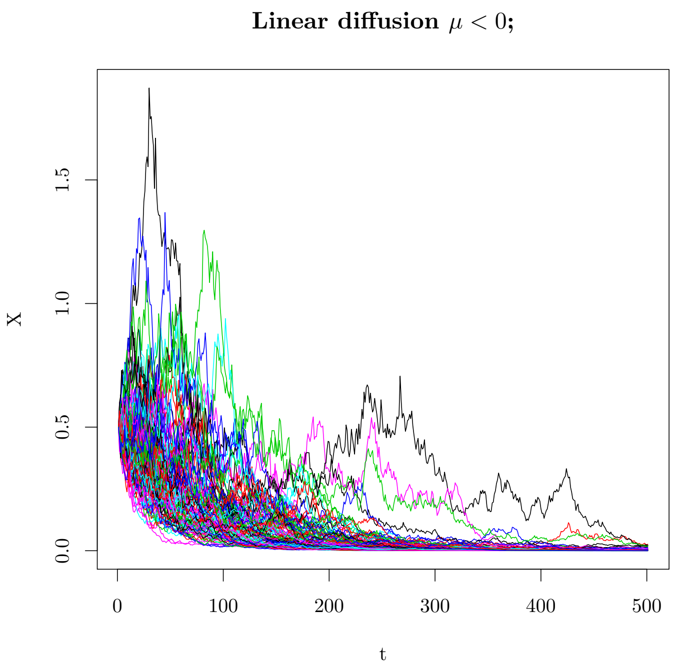
$$\begin{cases} p_t(t,x) = -G(t,x)[p] = -\dfrac{\partial}{\partial x}\left(\mu x p(t,x)\right) + \dfrac{1}{2}\dfrac{\partial^2}{\partial x^2}\left(\sigma x p(t,x)\right) \\ p(x_0, 0) = \delta(x - x_0) \end{cases}$$

The solution to this problem is

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

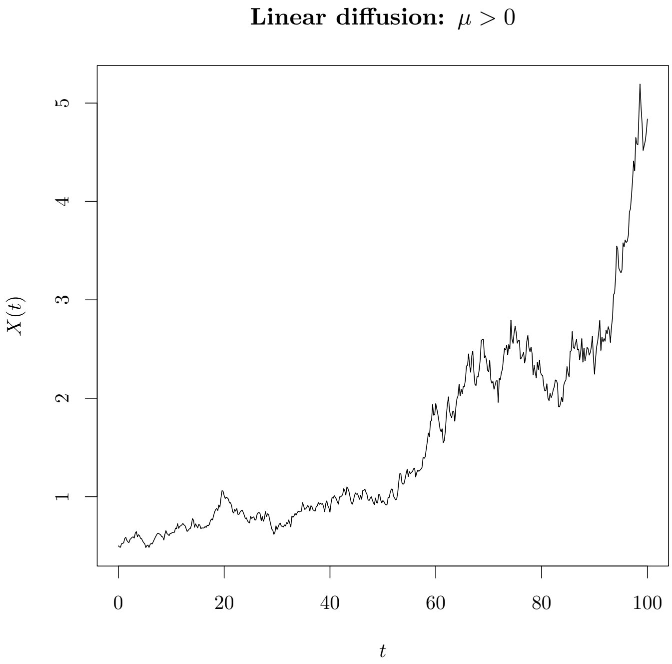


(a) One replication

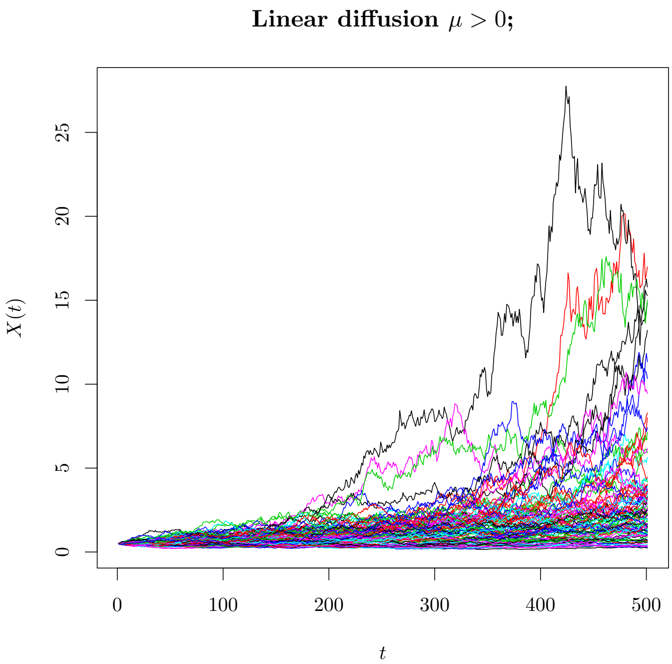


(b) 100 replications

Figure 9.3: Sample paths for the linear diffusion process with $\mu < 0$



(a) One replication



(b) 100 replications

Figure 9.4: Sample paths for the linear diffusion process with $\mu > 0$

9.4.3 Ornstein-Uhlenback process

It is the solution of the SDE

$$dX = \theta\left(\mu - X\right)dt + \sigma dW(t)$$

where $X(0) = x_0$.

The solution is

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

To prove this, let us consider a change in variables $Y(t) = X(t)e^{\theta t}$. If we apply Itô's formula , we have

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

Then, because, $X(t) = e^{-\theta t}Y(t)$ and $x_0 = y_0$,

$$\begin{aligned} X(t) &= e^{-\theta t}\left(y_0 + \theta\mu\int_0^t e^{\theta s}ds + \sigma\int_0^t e^{\theta s}dW(s)\right) \\ &= x_0e^{-\theta t} + \mu e^{-\theta t}\left(e^{\theta t} - 1\right) + \sigma\int_0^t e^{-\theta(t-s)}dW(s). \end{aligned}$$

Therefore, the conditional expected value and variance 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

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

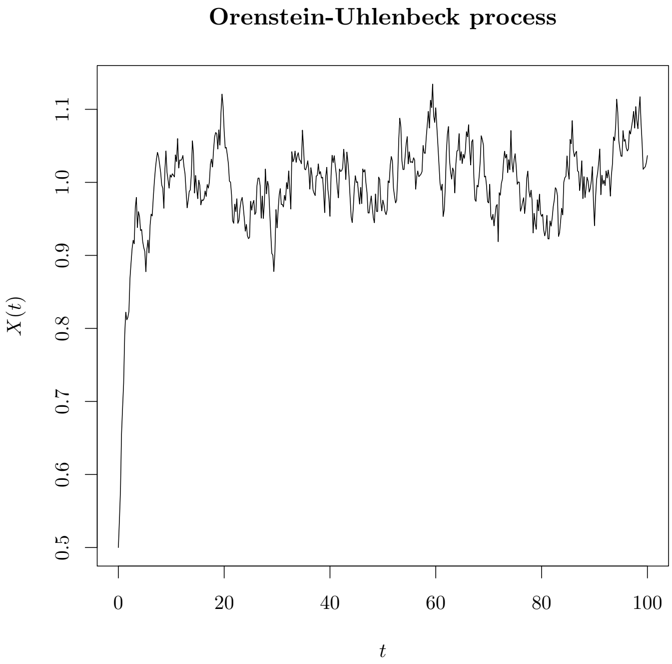
This means that the process is ergodic, because,

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

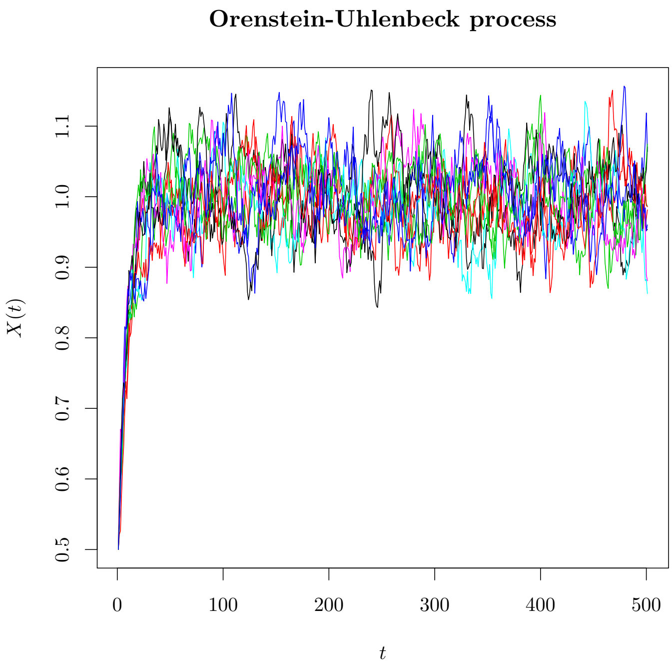
and it is asymptotically Gaussian, because

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

Sample paths are illustrated in figure 9.5.



(a) One replication



(b) 100 replications

Figure 9.5: Sample paths for Ornstein-Uhlenbeck process for $\theta > 0$ and $\mu = 1$

9.4.4 The general linear SDE

The general linear SDE has the form

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

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

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

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

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

and $\Phi(0) = 1$

9.5 Economic applications

9.5.1 The Solow stochastic growth model

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

Assume that population follows the SDE

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

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

The equilibrium equation for the product market is

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

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

$$dK = sLf(k)dt$$

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

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

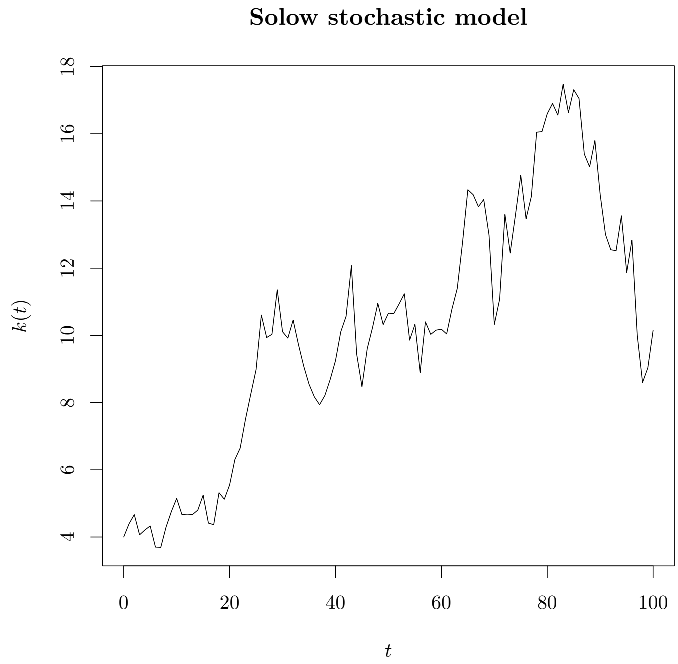


Figure 9.6: Sample path for the capital intensity: $s = 0.1$, $\alpha = 0.3$, $\mu = 0.01$, $\sigma = 0.1$

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

$$dk = \big(sf(k) - (\mu - \sigma^2)k\big)\, dt - k\sigma dW(t) \tag{9.9}$$

For a Cobb-Douglas function we have

$$dk = \big(sk^{\alpha} - (\mu - \sigma^2)k\big)\, dt - k\sigma dW(t)$$

where $0 < \alpha < 1$. Figures 9.6 and 9.7 present one replication and 100 replications for this equation for a deterministic initial value $k(0) = k_0$

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

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

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

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

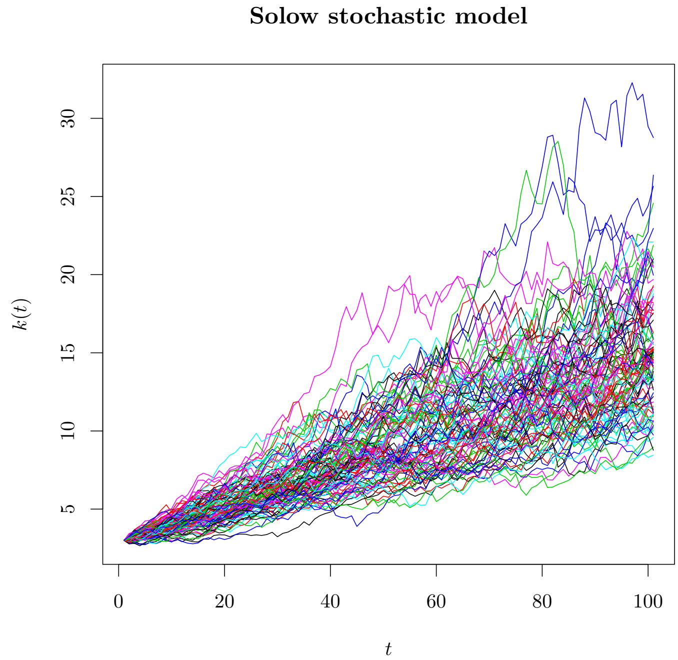


Figure 9.7: Sample paths for the capital intensity: $s = 0.1$, $\alpha = 0.3$, $\mu = 0.01$, $\sigma = 0.1$, 100 replications

9.5.2 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.

This implies that the value of the option at time 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 $t = 0$?

To answer the question, we assume there are **no arbitrage opportunities**, that is, the yields generated by the option should be equal to the yields of a portfolio composed by the available assets with the same value. We call this portfolio the replicating portfolio.

Let $V(S,t)$ be the value of the option on the risky asset at time t, for $0 \le t \le T$. The replication portfolio is composed of θ units of the risky asset and $(1-\theta)$ units of the risk free asset such that

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

Let us call $V^r(B,S) = (1-\theta)B + \theta S$.

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

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

The value of the replicating portfolio is

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

For ruling out arbitrage opportunities we should have $dV(S(t),t) = dV(B(t),S(t)$. To fulfil this condition we have to math the diffusion and the dispersion components of the processes for the option and the replicating portfolio values,

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

Therefore, the portfolio composition should be

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

and the options value is the solution of the following backward parabolic PDE

$$\begin{cases} V_t(S,t) = -\dfrac{\sigma^2}{2}S^2V_{ss}(S,t) - rSV_s(S,t) + rV(S,t), & (S,t) \in (0,\infty) \times [0,T] \\ V(S,T) = \max\{S(T) - K, 0\}, & (S,t) \in (0,\infty) \times \{t = T\} \end{cases} \tag{9.10}$$

We show in the PDE chapter that the solution is

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

where

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

Off course, if we find $V(S,t)$ then the cost of the contract would be $V(S,0)$, that is

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

where

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

which are all observable at time $t = 0$.

9.6 References

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

Appendix: The Gaussian integral

The gaussian kernel is a function

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

which has the well known bell shape.

A Gaussian integral is an integral of type

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

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

Some properties of the Gaussian integral are:

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

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

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

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

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

If we introduce a parameter $a > 0$

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

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

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

Chapter 12

Stochastic optimal control

12.1 Introduction

In this chapter we identify the stochastic optimal control problem as an optimal control problem of a 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)) 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 still new.

12.2 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 defined, and 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, $V(.)$,

$$V(x_0) = \max_{(U(t))_{t\in[0,T]}} \mathbb{E}_0\left[\int_0^T f(t, X(t), U(t))dt\right] \tag{12.1}$$

subjected to

$$dX(t) = g(t, X(t), U(t))dt + \sigma(t, X(t), U(t))dW(t) \tag{12.2}$$

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 with 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.

12.2.1 The stochastic DP principle

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* (12.1)-(12.2). *Then, at time* t, $X^*(t) = x$ *and* $U^*(t) = u$ *satisfy the* ***Hamilton-Jacobi-Bellman*** *equation*

$$-\frac{\partial V(t,x)}{\partial t} = \max_u \left(f(t,x,u) + g(t,x,u)\frac{\partial V(t,x)}{\partial x} + \frac{1}{2}\sigma(t,x,u)^2\frac{\partial^2 V(t,x)}{\partial x^2} \right). \tag{12.3}$$

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

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

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

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

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.

□

12.2.2 Infinite horizon

The autonomous discounted infinite horizon problem is

$$V(x_0) = \max_u \mathbb{E}_0\left[\int_0^\infty f\left(X(t), U(t)\right)\, e^{-\rho t} dt\right] \tag{12.4}$$

where $\rho > 0$, subject to

$$(\mathbf{??}) dX(t) = g\left(X(t), U(t)\right) dt + \sigma\left(X(t), U(t)\right) dW(t) \tag{12.5}$$

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

$$\rho V(x) = \max_u \left(f(x,u) + g(t,x,u)V'(x) + \frac{1}{2}\sigma(x,u)^2 V''(x)\right). \tag{12.6}$$

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

12.3 The stochastic PMP

Consider again the optimal control problem with value function (12.1).

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 (12.2), from the case in which it does not, as in equation

$$dX(t) = g(t, X(t), U(t))dt + \sigma(t, X(t))dW(t). \tag{12.7}$$

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

12.3.1 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* (12.1)-(12.7)*. Then, there are two processes* $(p(t), q(t))_{t\in[0,T]}$ *satisfying the adjoint equation and a terminal condition*

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

and, defining the Hamiltonian function

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

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

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

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

12.3.2 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* (12.1)-(12.2)*. 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*

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

and

$$\begin{cases} dP(t) = -\left\{f_x x(t, X(t), U(t)) + 2P(t)g_x(t, X(t), U(t)) + P(t)\left(g_x(t, X(t), U(t))\right)^2 + 2Q(t)\sigma_x(t, X(t), U(t))\right\} dt \\ P(T) = 0 \end{cases}$$

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

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

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)

12.3.3 Economic applications using stochastic dynamic programming

The representative agent problem

Here we present essentially the Merton (1971) model, which is a micro model for the simultaneous determination of the strategies of consumption and portfolio investment. We next present a simplified version with one risky and one riskless asset.

Let the exogenous processes be given to the representative consumer

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

where B and S are respectively the prices of the risky and the riskless assets, r is the interest rate, μ and σ are the constant rates of return and volatility for the equity.

The stock of financial wealth is denoted by $A(t) = \theta_0(t)B(t) + \theta_1(t)S(t)$, for any $t \in [0, \infty)$. Assume that $A(0) = \theta_0(0)B(0) + \theta_1(0)S(0)$ is known.

Assume that the agent also gets an endowment $\{y(t), t \in \mathbb{R}\}$ which adds to the incomes from financial investments and that the consumer uses the proceeds for consumption. Then the value of financial wealth at time t is

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

If the weight of the equity in total wealth is denoted by $w = \frac{\theta_1 S}{A}$ then $1 - w = \frac{\theta_0 B}{A}$. Then, we get the differential representation of the instantaneous budget constraint comes

$$dA(t) = [r(1 - w(t))A(t) + \mu w(t)A(t) + y(t) - c(t)]dt + w(t)\sigma A(t)dW(t). \tag{12.8}$$

The problem for the consumer-investor is

$$\max_{c,w} \mathbb{E}_0 \left[\int_0^\infty u(c(t))e^{-\rho t}dt\right] \tag{12.9}$$

subject to the instantaneous budget constraint (12.8), given $A(0)$ 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. We solve it by using proposition 1. The Hamilton-Jacobi-Bellman equation (12.6) 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

$$u^{'}(c^*) = V^{'}(A), \tag{12.10}$$

$$w^* = \frac{(r-\mu)V^{'}(A)}{\sigma^2 A V^{''}(A)} \tag{12.11}$$

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

$$\rho V(A) = u\left(h(V^{'}(A)`)\right) - h(V^{'}(A))V^{'}(A)(y+rA)V^{'}(A) - \frac{(r-\mu)^2(V^{'}(A))^2}{2\sigma^2 V^{''}(A)}. \tag{12.12}$$

In some cases the equation may be solve explicitly. In particular, let the utility function be CRRA as

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

and conjecture that the solution for equation (12.12) is of the type

$$V(A) = x(y+rA)^{1-\eta}$$

for x an unknow constant. If it is indeed a solution, there should be a constant, dependent upon the parameters of the model, such that equation (12.12) holds.

First note that

$$V^{'}(A) = (1-\eta)rx(y+rA)^{-\eta}$$

$$V^{''}(A) = -\eta(1-\eta)r^2 x(y+rA)^{-\eta-1}$$

then: the optimal consumption policy is

$$c^* = (xr(1-\eta))^{-\frac{1}{\eta}}(y+rA)$$

and the optimal portfolio composition is

$$w^* = \left(\frac{\mu - r}{\sigma^2}\right)\frac{y+rA}{\eta r A}$$

Interestingly it is a linear function of the ratio of total (human plus financial wealth $\frac{y}{r} + a$) over financial wealth.

After some algebra, we get

$$V(A) = \Theta\left(\frac{y + rA}{r}\right)^{1-\eta}$$

where [1]

$$\Theta \equiv \frac{1}{1-\eta}\left[\frac{\rho r(1-\eta)}{\eta} - \frac{(1-\eta)}{2\eta^2}\left(\frac{\mu - r}{\sigma}\right)^2\right]^{-\eta}$$

Then the optimal consumption is

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

If we set the total wealth as $W = \frac{y}{r} + A$, we may write the value function and the policy functions for consumption and portfolio investment

$$\begin{aligned} V(W) &= \Theta W^{1-\eta} \\ c^*(W) &= (1-\eta)\Theta^{-\frac{1}{\eta}} W \\ w^*(W) &= \left(\frac{\mu - r}{\eta\sigma^2}\right)\frac{W}{A}. \end{aligned}$$

Remark 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.

We see that the consumer cannot eliminate risk, in general. If we write $c^* = \chi A$, where $\chi \equiv (1-\eta)\Theta^{-\frac{1}{\eta}}\frac{W}{A}$, then the optimal process for wealth is

$$dA(t) = [r^* + (\mu - r)w^* - \chi]A(t)dt + \sigma w^* A(t)dW(t)$$

where $r^* = r\frac{W}{A}$, which is a linear SDE. Then as $c^* = c(A)$, if we apply the Itô's lemma we get

$$dc = \chi dA = c\left(\mu_c dt + \sigma_c dW(t)\right)$$

where

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

[1] Of course, $x = r^{-(1-\eta)}\Theta$.

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 $\frac{\mu - r}{\sigma}$ is the Sharpe index, and the unconditional expected value for consumption at time t

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

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

$$dk = \left(f(k) - c - (\mu - \sigma^2)k\right)dt - \sigma^2 k dW(t) \tag{12.13}$$

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 (12.13).

We solve it by using proposition 1. The HJB equation, (12.6), 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, we get again a 2nd 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**: $u(c) = \frac{c^{1-\theta}}{1-\theta}$ and $f(k) = k^\alpha$. Then the optimal policy function becomes

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

and the HJB becomes

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

We can get, again, a closed form solution if we assume further that $\theta = \alpha$. 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)

12.3.4 Economic applications using stochastic maximum principle

Next we solve a stochastic optimal economic growth model by applying the stochastic Pontriyagin maximum principle. In the first case the control variable is not in the volatility term and in the second it is. This means that we use proposition 2 in the first case and 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

$$dK(t) = (\mu K(t) - C(t))\, dt + \sigma K(t) dW(t) \tag{12.14}$$

$$K(0) = k_0$$

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 $C^* = c$ is

$$\max_c \mathcal{H}(t, c, k, p, q) \equiv H(t, c, k, p, q) + \sigma k q$$

Therefore,

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

Using Itô's lemma

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

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

$$c = \phi k$$

where ϕ 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

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

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

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

substituting in equation (12.14) yields

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

Therefore

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

and

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

meaning that consumption and capital accumulation are perfectly correlated.

The Merton (1990)

Next we consider again the problem of maximizing the intertemporal utility functional (12.9) subject to the stochastic differential equation (12.8) where we assume there is no non-financial income, that is $y = 0$.

We consider the problem

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

subject to

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

where $N(0) = n_0$ is given.

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 **??**.

The adjoint equations are

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

and

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

To find the optimal controls we write the generalized Hamiltonian

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

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

$$C^*(t) = e^{-\rho t}p(t)^{-1} \tag{12.15}$$

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 (12.15), 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

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

This is only consistent with the previous derivation if

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

Solving for ξ and w we obtain the optimal controls

$$C^*(t) = \rho N^*(t) \tag{12.16}$$

$$w^*(t) = \frac{\mu - r}{\sigma^2} \tag{12.17}$$

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

$$\mu_n = r - \rho + \left(\frac{\mu - r}{\sigma}\right)^2 \tag{12.18}$$

$$\sigma_n = \frac{\mu - r}{\sigma} \tag{12.19}$$

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.

12.4 References

- Application to economics and finance: Malliaris and Brock (1982), Stokey (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.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–59.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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