FINANCIAL MODELLING AND DATA SCIENCE

BRAD BAXTER

Version: 202309271030

Abstract. These notes contain all examinable theoretical material for the first term of the Financial Modelling and Data Science course.

Contents

1. Introduction 2
1.1. Reading List 2
2. The Binomial Model Universe 4
2.1. The Binomial Model and Delta Hedging 5
3. Brownian Motion 8
3.1. Simple Random Walk 8
3.2. Discrete Brownian Motion 9
3.3. Basic Properties of Brownian Motion 9
3.4. Martingales 10
3.5. Brownian Motion and Martingales 10
3.6. The Black–Scholes Equation 11
3.7. Itô Calculus 13
3.8. Itô rules and SDEs 18
3.9. Multivariate Geometric Brownian Motion 19
3.10. The Ornstein–Uhlenbeck Process 23
3.11. Δ-Hedging for GBM 24
4. The Geometric Brownian Motion Universe 26
4.1. European Puts and Calls 30
4.2. Digital Options 32
5. Mathematical Background Material 34
5.1. Probability Theory 34
5.2. Differential Equations 36
5.3. Recurrence Relations 37
5.4. Mortgages – a once exotic instrument 40
5.5. Pricing Mortgages via lack of arbitrage 41
6. Numerical Linear Algebra 42
6.1. Orthogonal Matrices 42

1. Introduction

You can access these notes, and other material, via my office machine:

`http://econ109.econ.bbk.ac.uk/brad/FMDS/`

An earlier version of my lecture notes for both terms is available here:

`http://econ109.econ.bbk.ac.uk/brad/Methods/Methods.pdf`

These notes are fairly stable, having evolved while teaching multiple MSc programmes, including MSc Mathematical Finance at Imperial College, London, MSc Financial Engineering here, and MSc Mathematical Finance. I do still add new examples and make minor changes, so please check you have the latest version.

Students will also be taking my Matlab course and the notes are available on my server too:

`http://econ109.econ.bbk.ac.uk/brad/FDMS/matlab_intro_notes.pdf`

Past exams can also be downloaded from my server:

`http://econ109.econ.bbk.ac.uk/brad/FinEngExams/`

The first three questions are suitable in my exams set since 2020. The first four questions are suitable in earlier exams.

Many students will find my Numerical Analysis notes helpful too:

`http://econ109.econ.bbk.ac.uk/brad/FDMS/nabook.pdf`

I wrote these notes for an undergraduate course in Numerical Analysis when lecturing at Imperial College, London, from 1995–2001. However, they have often been found useful by MSc students who need to improve their general understanding of theoretical Numerical Analysis. The first section of the notes is on matrix algebra and contain many examples and exercises, together with solutions.

Finally, there is lots of interesting material, including extensive notes for several related courses (e.g. Analysis) available on my office Linux server, so please do explore:

`http://econ109.econ.bbk.ac.uk/brad/`

1.1. **Reading List.** Everything required for this term is in these notes or in my slides. For further reading, the following books are all useful.

(i) M. Baxter and A. Rennie, *Financial Calculus*, Cambridge University Press. This gives a fairly informal description of the mathematics of pricing, concentrating on martingales. It's not a source of information for efficient numerical methods.

(ii) A. Etheridge, *A Course in Financial Calculus*, Cambridge University Press. This does not focus on the algorithmic side but is very lucid for students with a strong mathematical background.

(iii) D. Higham, *An Introduction to Financial Option Valuation*, Cambridge University Press. This book provides many excellent Matlab examples, although its mathematical level is undergraduate.

(iv) J. Hull, *Options, Futures and Other Derivatives*, 6th edition. [Earlier editions are probably equally suitable for much of the course.] Fairly clear, with lots of background information on finance. The mathematical treatment is lower than the level of much of our course (and this is *not* a mathematically rigorous book), but it's still the market leader in many ways.

(v) J. Michael Steele, *Stochastic Calculus and Financial Applications*, Springer. This is an excellent book, but is one to read near the end of this term, once you are more comfortable with fundamentals.
(vi) P. Wilmott, S. Howison and J. Dewynne, *The Mathematics of Financial Derivatives*, Cambridge University Press. This book is very useful for its information on partial differential equations. If your first degree was in engineering, mathematics or physics, then you probably spent many happy hours learning about the diffusion equation. This book is very much mathematical finance from the perspective of a traditional applied mathematician. It places much less emphasis on probability theory than most books on finance.
(vii) P. Wilmott, *Paul Wilmott introduces Quantitative Finance*, 2nd edition, John Wiley. More chatty than his previous book. The author's ego grew enormously between the appearance of these texts, but there's some good material here.
(viii) Y.-K. Kwok, *Mathematical Models of Financial Derivatives*, Springer. A dry, but very detailed treatment of finite difference methods. If you need a single book for general reference work, then this is probably it.

There are lots of books suitable for mathematical revision. The **Schaum series** publishes many good inexpensive textbooks providing worked examples. The inexpensive paperback *Calculus*, by K. G. Binmore (Cambridge University Press) will also be useful to students wanting an introduction to, say, multiple integrals, as will *Mathematical Methods for Science Students*, by Geoff Stephenson. At a slightly higher level, *All you wanted to know about Mathematics but were afraid to ask*, by L. Lyons (Cambridge University Press, 2 vols), is useful and informal.

The ubiquitous *Numerical Recipes in C++*, by S. Teukolsky et al, is extremely useful. Its coverage of numerical methods is generally reliable and it's available online at `www.nr.com`. A good hard book on partial differential equations is that of A. Iserles (Cambridge University Press).

At the time of writing, finance is going through a turbulent period which began in 2007, in which politicians sometimes profess their longstanding doubts that the subject was well-founded – surprisingly, many omitted to voice such doubts earlier! It is good to know that we have been here before. The following books are included for general cultural interest. My recommendation for a single book is Lanchester.

(i) M. Balen, *A Very English Deceit: The Secret History of the South Sea Bubble and the First Great Financial Scandal.*
(ii) C. Eagleton and J. William (eds), *Money: A History.*
(iii) C. P. Kindleberger, R. Aliber and R. Solow, *Manias, Panics, and Crashes: A History of Financial Crises*, Wiley. This is still a classic.
(iv) J. Lanchester, *How to Speak Money*, Faber. This is an **excellent** introduction to finance and economics for all readers. Lanchester is a journalist and author, as well as being a gifted expositor.
(v) N. N. Taleb, *The Black Swan.* In my view, this is greatly over-rated, but you should still read it.

No text is perfect: please report any slips to `b.baxter@bbk.ac.uk`.

2. The Binomial Model Universe

We begin with a discrete model for asset prices which was first demonstrated by Cox, Ross and Rubinstein in the late 1970s. It is still of practical use and its limit, as the number of timesteps tends to infinity, will be the geometric Brownian motion (GBM) model for asset prices.

Our model will be entirely specified by two parameters, $\alpha > 0$ and $p \in [0,1]$. We choose $S_0 > 0$ and define

$$S_k = S_{k-1}\exp(\alpha X_k), \qquad k > 0, \tag{2.1}$$

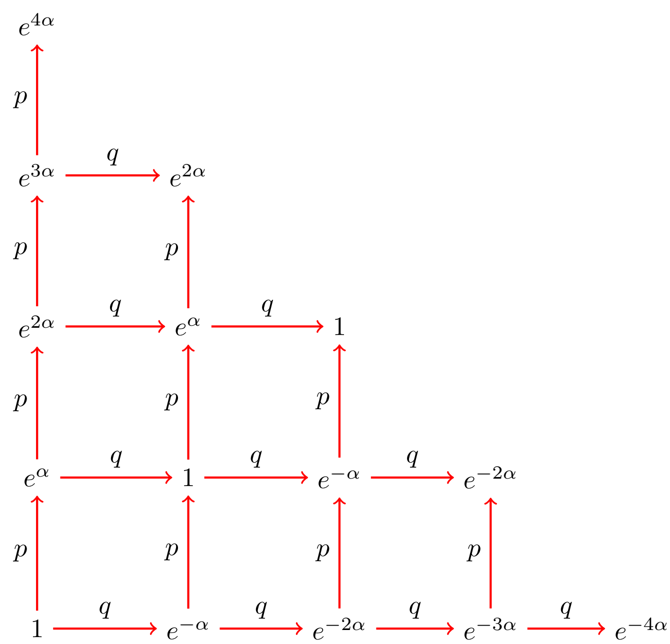
where the independent random variables $X_1, X_2, \ldots$ satisfy

$$\mathbb{P}\left(X_k = 1\right) = p, \qquad \mathbb{P}\left(X_k = -1\right) = 1 - p =: q. \tag{2.2}$$

Thus

$$S_m = S_0 e^{\alpha(X_1 + X_2 + \cdots + X_m)}, \qquad m > 0. \tag{2.3}$$

It is usual to display this random process graphically.

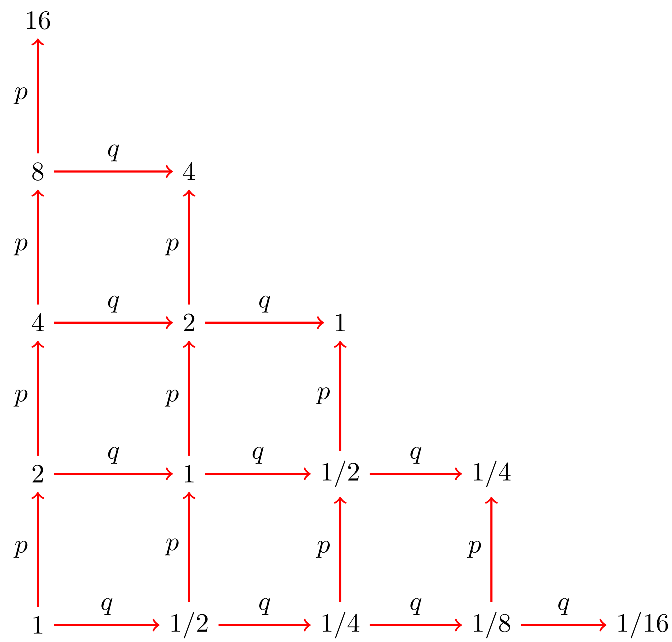


At this stage, we haven't specified p and α. In practice, these would be chosen numerically to fit data, but this calibration is not crucial here. Instead, we shall first learn how to price **functions** $f(S_n, t_n)$ of the asset price, which we shall call **options**. In the jargon of mathematical finance these are also called **derivatives** and **contingent claims**. The technique is fundamental and was discovered by Black and Scholes in the early 1970s.

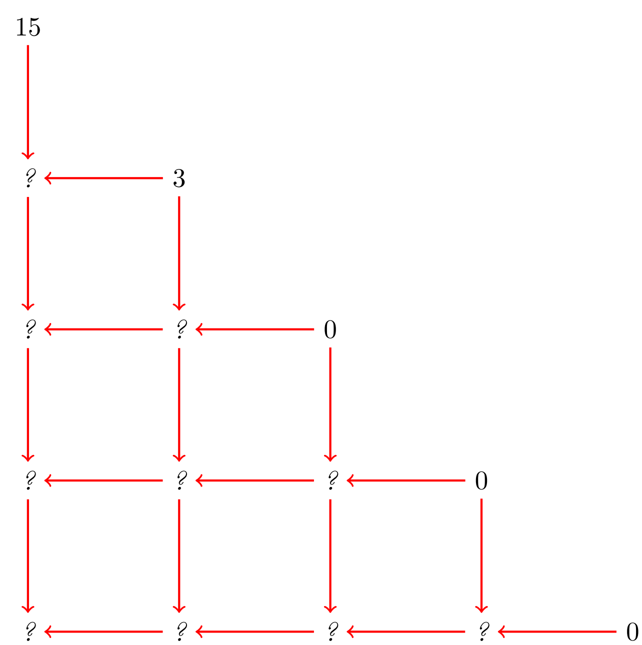
Example 2.1. *Suppose* $e^\alpha = 2$, $p = 1/2$. *Let* $m = 4$ *and define the* **call option**

$$f(S(mh), mh) = \left(S(mh) - 1\right)_+ .$$

Using (2.16)*, we obtain the following diagram for the asset prices.*



The corresponding diagram for the option prices is as follows. Our aim is to find all of the earlier values of the option price.



We shall see how to calculate these unknown values very soon.

2.1. **The Binomial Model and Delta Hedging.** In this section we learn how to calculate option prices using a **delta hedging** argument. You only need to accept one fundamental axiom from mathematical economics: every deterministic asset grows at the **risk-free rate** r. In other words, if we borrow or lend M, then at time t we have $M\exp(rt)$, where the risk-free rate r is set by the currency's central bank.

At time $t_{n-1} = (n-1)h$, we construct a new portfolio

$$\Pi_{n-1} = f(S_{n-1}, t_{n-1}) - \Delta_{n-1} S_{n-1}. \tag{2.4}$$

and we choose Δ_{n-1} so that the evolution of Π_{n-1} is deterministic. Now, at time $t_n = nh$, the portfolio Π_{n-1} has the new value

$$\Pi_n = f(S_{n-1}e^{\alpha X_n}, t_n) - \Delta_{n-1} S_{n-1} e^{\alpha X_n}. \tag{2.5}$$

Thus Π_n is deterministic if the two possible values of (2.5) are equal, that is,

$$f(S_{n-1}e^{\alpha}, t_n) - \Delta_{n-1} S_{n-1} e^{\alpha} = f(S_{n-1}e^{-\alpha}, t_n) - \Delta_{n-1} S_{n-1} e^{-\alpha}. \tag{2.6}$$

It is useful to introduce the notation

$$f_{\pm} = f(S_{n-1}e^{\pm\alpha}, t_n). \tag{2.7}$$

Then (2.6) and (2.7) imply that

$$\Delta_{n-1} S_{n-1} = \frac{f_+ - f_-}{e^{\alpha} - e^{-\alpha}}. \tag{2.8}$$

Thus the resulting portfolio values are given by

$$\Pi_{n-1} = f(S_{n-1}, t_{n-1}) - \frac{f_+ - f_-}{e^{\alpha} - e^{-\alpha}} \tag{2.9}$$

and

$$\begin{aligned} \Pi_n &= f(S_{n-1}e^{\alpha}, t_n) - \frac{f_+ - f_-}{e^{\alpha} - e^{-\alpha}} e^{\alpha} \\ &= \frac{f_- e^{\alpha} - f_+ e^{-\alpha}}{e^{\alpha} - e^{-\alpha}}. \end{aligned} \tag{2.10}$$

Now that the portfolio's evolution from Π_{n-1} to Π_n is deterministic, we must have $\Pi_n = \exp(rh)\Pi_{n-1}$, i.e.

$$\frac{f_- e^{\alpha} - f_+ e^{-\alpha}}{e^{\alpha} - e^{-\alpha}} = e^{rh} \left(f(S_{n-1}, t_{n-1}) - \frac{f_+ - f_-}{e^{\alpha} - e^{-\alpha}} \right). \tag{2.11}$$

The key point here is that $f(S_{n-1}, t_{n-1})$ is a linear combination of f_+ and f_-. Specifically, if we introduce

$$P = \frac{e^{rh} - e^{-\alpha}}{e^{\alpha} - e^{-\alpha}}, \tag{2.12}$$

then (2.11) becomes

$$f(S_{n-1}, t^{n-1}) = e^{-rh} \left(P f_+ + (1-P) f_- \right). \tag{2.13}$$

Note that original model probability p does not occur in this formula: instead, it is as if we had begun with the alternative binomial model

$$S_n = S_{n-1} e^{\alpha Y_n}, \tag{2.14}$$

where the independent Bernoulli random variables $Y_1, Y_2, \ldots, Y_n$ satisfy $\mathbb{P}(Y_k = 1) = P$ and $\mathbb{P}(Y_k = -1) = 1 - P$, where P is given by (2.12). Indeed, we have

$$\mathbb{E} S_n | S_{n-1} = S_{n-1} \mathbb{E} e^{\alpha Y_n} = S_{n-1} e^{rh}. \tag{2.15}$$

Exercise 2.1. *Prove that* $\mathbb{E} S_n | S_{n-1} = S_{n-1} e^{rh}$.

We can now easily price a European option given these parameters. If S_k denotes our Binomial Model asset price at time kh, for some positive time interval h, then the Binomial Model European option requirement is given by

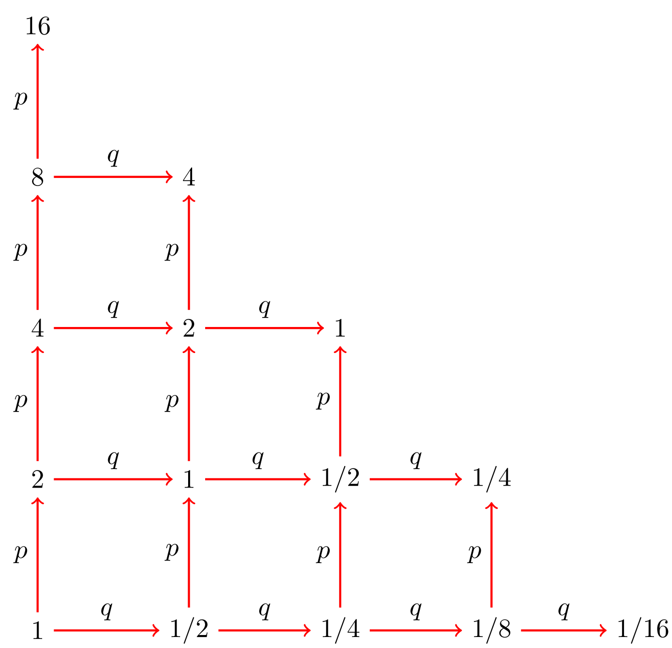
$$\begin{aligned} f(S_{k-1},(k-1)h) &= e^{-rh}\mathbb{E}f(S_{k-1}e^{\alpha X_k},kh) \\ &= e^{-rh}\left(pf(S_{k-1}e^{\alpha},kh)+(1-p)f(S_{k-1}e^{-\alpha},kh)\right). \end{aligned} \tag{2.16}$$

Thus, given the $m+1$ possible asset prices at expiry time mh, and their corresponding option prices, we use (2.16) to calculate the m possible values of the option at time $(m-1)h$. Recurring this calculation provides the value of the option at time 0. Let's illustrate this by solving Example 2.1.

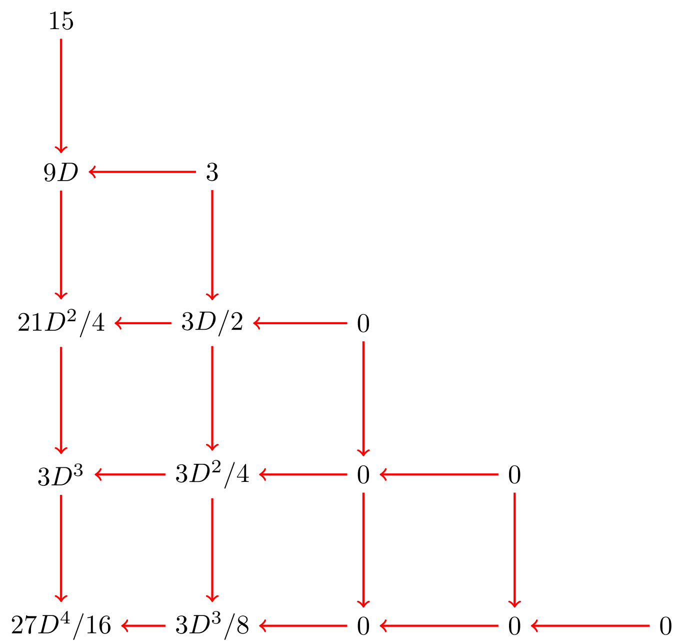
Example 2.2. *Suppose $e^{\alpha}=2$, $p=1/2$ and $D=e^{-rh}$. Let $m=4$ and let's use the Binomial Model to calculate all earlier values of the call option whose expiry value is*

$$f(S(mh),mh)=\left(S(mh)-1\right)_{+}.$$

Using (2.16), we obtain the following diagram for the asset prices.



The corresponding diagram for the option prices is as follows.



3. Brownian Motion

3.1. **Simple Random Walk.** Let $X_1, X_2, \ldots$ be a sequence of independent random variables all of which satisfy

$$\mathbb{P}\left(X_i = \pm 1\right) = 1/2 \tag{3.1}$$

and define

$$S_n = X_1 + X_2 + \cdots + X_n. \tag{3.2}$$

We can represent this graphically by plotting the points $\{(n, S_n) : n = 1, 2, \ldots\}$, and one way to imagine this is as a *random walk*, in which the walker takes identical steps forwards or backwards, each with probability 1/2. This model is called *simple random walk* and, whilst easy to define, is a useful laboratory in which to improve probabilistic intuition.

Another way to imagine S_n is to consider a game in which a fair coin is tossed repeatedly. If I win the toss, then I win £1; losing the toss implies a loss of £1. Thus S_n is my fortune at time n.

Firstly note that

$$\mathbb{E}S_n = \mathbb{E}X_1 + \cdots + \mathbb{E}X_n = 0.$$

Further, $\mathbb{E}X_i^2 = 1$, for all i, so that $\operatorname{var} X_i = 1$. Hence

$$\operatorname{var} S_n = \operatorname{var} X_1 + \operatorname{var} X_2 + \cdots + \operatorname{var} X_n = n,$$

since $X_1, \ldots, X_n$ are independent random variables.

3.2. **Discrete Brownian Motion.** We begin with a slightly more complicated random walk this time. We choose a timestep $h > 0$ and let $Z_1, Z_2, \ldots$ be independent $N(0, h)$ Gaussian random variables. We then define a curve ${B^{(h)}}_t$ by defining ${B^{(h)}}_0 = 0$ and

$$B^{(h)}(kh) = Z_1 + Z_2 + \cdots + Z_k, \tag{3.3}$$

for positive integer k. We then join the dots to obtain a piecewise linear function. More precisely, we define

$${B^{(h)}}_t = {B^{(h)}}_{kh} + (t - kh)\left(\frac{{B^{(h)}}_{(k+1)h} - {B^{(h)}}_{kh}}{h}\right), \qquad \text{for } t \in (kh, (k+1)h).$$

The resultant random walk is called discrete Brownian motion.

Proposition 3.1. *If $0 \le a \le b \le c$ and $a, b, c \in h\mathbb{Z}$, then the discrete Brownian motion increments ${B^{(h)}}_c - {B^{(h)}}_b$ and ${B^{(h)}}_b - {B^{(h)}}_a$ are independent random variables. Further, ${B^{(h)}}_c - {B^{(h)}}_b \sim N(0, c - b)$ and ${B^{(h)}}_b - {B^{(h)}}_a \sim N(0, b - a)$.*

Proof. Exercise. □

3.3. **Basic Properties of Brownian Motion.** It's not obvious that discrete Brownian motion has a limit, in some sense, when we allow the timestep h to converge to zero. However, it can be shown that this is indeed the case (and will see the salient features of the *Lévy–Cieselski* construction of this limit later). For the moment, we shall state the defining properties of Brownian motion.

Definition 3.1. *There exists a stochastic process W_t, called Brownian motion, which satisfies the following conditions:*

(i) $W_0 = 0$;
(ii) *If $0 \le a \le b \le c$, then the Brownian increments $W_c - W_b$ and $W_b - W_a$ are independent random variables. Further, $W_c - W_b \sim N(0, c - b)$ and $W_b - W_a \sim N(0, b - a)$;*
(iii) *W_t is continuous almost surely.*

Proposition 3.2. *$W_t \sim N(0, t)$ for all $t > 0$.*

Proof. Just set $a = 0$ and $b = t$ in (ii) of Definition 3.1. □

The increments of Brownian motion are independent Gaussian random variables, but the actual values W_a and W_b are **not** independent random variables, as we shall now see.

Proposition 3.3. *If $a, b \in [0, \infty)$, then $\mathbb{E}(W_a W_b) = \min\{a, b\}$.*

Proof. We assume $0 < a < b$, the remaining cases being easily checked. Then

$$\begin{aligned}
\mathbb{E}(W_a W_b) &= \mathbb{E}\left(W_a [W_b - W_a] + W_a^2\right) \\
&= \mathbb{E}(W_a [W_b - W_a]) + \mathbb{E}\left(W_a^2\right) \\
&= 0 + a \\
&= a.
\end{aligned}$$

□

Brownian motion is continuous almost surely but it is easy to see that it cannot be differentiable. The key observation is that

$$\frac{W_{t+h} - W_t}{h} \sim N(0, \frac{1}{h}). \tag{3.4}$$

In other words, instead of converging to some limiting value, the variance of the random variable $(W_{t+h} - W_t)/h$ tends to infinity, as $h \to 0$.

3.4. **Martingales.** A martingale is a mathematical version of a fair game, as we shall first illustrate for simple random walk.

Proposition 3.4. *We have*

$$\mathbb{E}\left(S_{n+k} | S_n\right) = S_n.$$

Proof. The key observation is that

$$S_{n+k} = S_n + X_{n+1} + X_{n+2} + \cdots + X_{n+k}$$

and $X_{n+1}, \ldots, X_{n+k}$ are all independent of $S_n = X_1 + \cdots + X_n$. Thus

$$\mathbb{E}\left(S_{n+k} | S_n\right) = S_n + \mathbb{E}X_{n+1} + \mathbb{E}X_{n+2} + \cdots + \mathbb{E}X_{n+k} = S_n.$$

□

To see why this encodes the concept of a fair game, let us consider a biased coin with the property that

$$\mathbb{E}\left(S_{n+10} | S_n\right) = 1.1 S_n.$$

Hence

$$\mathbb{E}\left(S_{n+10\ell} | S_n\right) = 1.1^{\ell} S_n.$$

In other words, the expected fortune $S_{n+10\ell}$ grows exponentially with ℓ. For example, if we ensure that $S_4 = 3$, by fixing the first four coin tosses in some fashion, then our expected fortune will grow by 10% every 10 tosses thereafter.

3.5. **Brownian Motion and Martingales.**

Proposition 3.5. *Brownian motion is a martingale, that is,* $\mathbb{E}\left(W_{t+h} | W_t\right) = W_t$, *for any* $h > 0$.

Proof.

$$\begin{aligned}
\mathbb{E}\left(W_{t+h} | W_t\right) &= \mathbb{E}\left([W_{t+h} - W_t] + W_t | W_t\right) \\
&= \mathbb{E}\left([W_{t+h} - W_t] | W_t\right) + W_t \\
&= \mathbb{E}\left([W_{t+h} - W_t]\right) + W_t \\
&= 0 + W_t \\
&= W_t.
\end{aligned}$$

□

We can sometimes use a similar argument to prove that functionals of Brownian motion are martingales.

Proposition 3.6. *The stochastic process* $X_t = W_t^2 - t$ *is a martingale, that is,* $\mathbb{E}\left(X_{t+h} | X_t\right) = X_t$, *for any* $h > 0$.

Proof.

$$\begin{aligned}\mathbb{E}\left(X_{t+h}|X_t\right) &= \mathbb{E}\left([W_{t+h} - W_t + W_t]^2 - [t+h]|W_t\right)\\ &= \mathbb{E}\left([W_{t+h} - W_t]^2 + 2W_t\left[W_{t+h} - W_t\right] + W_t^2 - t - h|W_t\right)\\ &= \mathbb{E}[W_{t+h} - W_t]^2 + W_t^2 - t - h\\ &= h + W_t^2 - t - h\\ &= X_t.\end{aligned}$$

□

The following example will be crucial.

Proposition 3.7. *Geometric Brownian motion*

$$Y_t = e^{\alpha + \beta t + \sigma W_t} \tag{3.5}$$

is a martingale, that is, $\mathbb{E}\left(Y_{t+h}|Y_t\right) = Y_t$*, for any* $h > 0$*, if and only if* $\beta = -\sigma^2/2$.

Proof.

$$\begin{aligned}\mathbb{E}\left(Y_{t+h}|Y_t\right) &= \mathbb{E}\left(e^{\alpha+\beta(t+h)+\sigma W_{t+h}}|Y_t\right)\\ &= \mathbb{E}\left(Y_t e^{\beta h + \sigma(W_{t+h} - W_t)}|Y_t\right)\\ &= Y_t \mathbb{E} e^{\beta h + \sigma(W_{t+h} - W_t)}\\ &= Y_t e^{\left(\beta + \sigma^2/2\right)h}.\end{aligned}$$

□

In this course, the mathematical model chosen for option pricing is risk-neutral geometric Brownian motion: we choose a geometric Brownian motion S_t with the property that $Y_t = e^{-rt} S_t$ is a martingale, where Y_t is given by (3.5). Thus we have

$$Y_t = e^{\alpha + (\beta - r)t + \sigma W_t}$$

and Proposition 3.7 implies that $\beta - r = -\sigma^2/2$, i.e.

$$S_t = e^{\alpha + (r - \sigma^2/2)t + \sigma W_t} = S_0 e^{(r - \sigma^2/2)t + \sigma W_t}.$$

3.6. **The Black–Scholes Equation.** We can also use (4.6) to derive the famous *Black–Scholes partial differential equation*, which is satisfied by any European option. The key is to choose a *small* positive h in (4.6) and expand. We shall need Taylor's theorem for functions of two variables, which states that

$$\begin{aligned}G(x + \delta x, y + \delta y) \quad &= \quad G(x, y) + \left(\frac{\partial G}{\partial x}\delta x + \frac{\partial G}{\partial y}\delta y\right)\\ &\quad + \frac{1}{2}\left(\frac{\partial^2 G}{\partial x^2}(\delta x)^2 + 2\frac{\partial^2 G}{\partial x \partial y}(\delta x)(\delta y) + \frac{\partial^2 G}{\partial y^2}(\delta y)^2\right) + \cdots .\end{aligned}$$

(3.6)

Further, it simplifies matters to use "log-space": we introduce $u(t) := \log S(t)$, where $\log \equiv \log_e$ in these notes (*not* logarithms to base 10). In log-space, (4.3) becomes

$$u(t+h) = u(t) + (r - \sigma^2/2)h + \sigma\delta W_t, \tag{3.7}$$

where

$$\delta W_t = W_{t+h} - W_t \sim N(0,h). \tag{3.8}$$

We also introduce

$$g(u(t),t) := f(\exp(u(t),t)), \tag{3.9}$$

so that (4.6) takes the form

$$g(u(t),t) = e^{-rh}\mathbb{E}g(u(t+h),t+h). \tag{3.10}$$

Now Taylor expansion yields the (initially daunting)

$$\begin{aligned} g(u(t+h),t+h) &= g(u(t) + (r-\sigma^2/2)h + \sigma\delta W_t, t+h) \\ &= g(u(t),t) + \frac{\partial g}{\partial u}\left((r-\sigma^2/2)h + \sigma\delta W_t\right) + \\ &\quad \frac{1}{2}\frac{\partial^2 g}{\partial u^2}\sigma^2\left(\delta W_t\right)^2 + h\frac{\partial g}{\partial t} + \cdots, \end{aligned} \tag{3.11}$$

ignoring all terms of higher order than h. Further, since $\delta W_t \sim N(0,h)$, i.e. $\mathbb{E}\delta W_t = 0$ and $\mathbb{E}[(\delta W_t)^2] = h$, we obtain

$$\mathbb{E}g(u(t+h),t+h) = g(u(t),t) + h\left(\frac{\partial g}{\partial u}(r-\sigma^2/2) + \frac{1}{2}\frac{\partial^2 g}{\partial u^2}\sigma^2 + \frac{\partial g}{\partial t}\right) + \cdots. \tag{3.12}$$

Recalling that

$$e^{-rh} = 1 - rh + \frac{1}{2}(rh)^2 + \cdots,$$

we find

$$\begin{aligned} g &= (1 - rh + \cdots)\Big(g + h\left[\frac{\partial g}{\partial u}(r-\sigma^2/2) + \frac{1}{2}\frac{\partial^2 g}{\partial u^2}\sigma^2 + \frac{\partial g}{\partial t}\right] + \cdots\Big) \\ &= g + h\Big(-rg + \frac{\partial g}{\partial u}(r-\sigma^2/2) + \frac{1}{2}\frac{\partial^2 g}{\partial u^2}\sigma^2 + \frac{\partial g}{\partial t}\Big) + \cdots. \end{aligned} \tag{3.13}$$

For this to be true for all $h > 0$, we must have

$$-rg + \frac{\partial g}{\partial u}(r-\sigma^2/2) + \frac{1}{2}\frac{\partial^2 g}{\partial u^2}\sigma^2 + \frac{\partial g}{\partial t} = 0, \tag{3.14}$$

and this is the celebrated Black–Scholes partial differential equation (PDE). Thus, instead of computing an expected future value, we can calculate the solution of the Black–Scholes PDE (3.14). The great advantage gained is that there are highly efficient numerical methods for solving PDEs numerically. The disadvantages are complexity of code and learning the mathematics needed to exploit these methods effectively.

Exercise 3.1. *Use the substitution* $S = \exp(u)$ *to transform* (3.14) *into the non-linear form of the Black–Scholes PDE.*

3.7. **Itô Calculus.** Equation (3.12) is really quite surprising, because the second derivative contributes to the $O(h)$ term. This observation is at the root of the Itô rules. We begin by considering the *quadratic variation* $I_n[a,b]$ of Brownian motion on the interval $[a,b]$. Specifically, we choose a positive integer n and let $nh = b-a$. We then define

$$I_n[a,b] = \sum_{k=1}^{n} \left(W_{a+kh} - W_{a+(k-1)h}\right)^2. \tag{3.15}$$

We shall prove that $\mathbb{E}I_n[a,b] = b-a$, for every positive integer n, but that $\operatorname{var} I_n[a,b] \to 0$, as $n \to \infty$. We shall need the following simple property of Gaussian random variables.

Lemma 3.8. *Let $Z \sim N(0,1)$. Then $\mathbb{E}Z^4 = 3$.*

Proof. Integrating by parts, we obtain

$$\begin{aligned}
\mathbb{E}Z^4 &= \int_{-\infty}^{\infty} s^4 (2\pi)^{-1/2} e^{-s^2/2}\, ds \\
&= (2\pi)^{-1/2} \int_{-\infty}^{\infty} s^3 \frac{d}{ds}\left(-e^{-s^2/2}\right) ds \\
&= (2\pi)^{-1/2} \left\{ \left[-s^3 e^{-s^2/2}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} 3s^2 \left(-e^{-s^2/2}\right) ds \right\} \\
&= 3.
\end{aligned}$$

□

Exercise 3.2. *Calculate $\mathbb{E}Z^6$ when $Z \sim N(0,1)$. More generally, calculate $\mathbb{E}Z^{2m}$ for any positive integer m.*

Proposition 3.9. *We have $\mathbb{E}I_n[a,b] = b-a$ and $\operatorname{var} I_n[a,b] = 2(b-a)^2/n$.*

Proof. Firstly,

$$\mathbb{E}I_n[a,b] = \sum_{k=1}^{n} \mathbb{E}\left(W_{a+kh} - W_{a+(k-1)h}\right)^2 = \sum_{k=1}^{n} h = nh = b-a.$$

Further, the Brownian increments $W_{a+kh} - W_{a+(k-1)h}$ are independent $N(0,h)$ random variables. We shall define independent $N(0,1)$ random variables $Z_1, Z_2, \ldots, Z_n$ by

$$W_{a+kh} - W_{a+(k-1)h} = \sqrt{h} Z_k, \qquad 1 \le k \le n.$$

Hence

$$\begin{aligned}
\operatorname{var} I_n[a,b] &= \sum_{k=1}^{n} \operatorname{var}\left(\sqrt{h} Z_k\right)^2 \\
&= \sum_{k=1}^{n} \operatorname{var}\left[h Z_k^2\right] \\
&= \sum_{k=1}^{n} h^2 \operatorname{var}\left[Z_k^2\right] \\
&= \sum_{k=1}^{n} h^2 \left(\mathbb{E} Z_k^4 - \left[\mathbb{E} Z_k^2\right]^2\right) \\
&= \sum_{k=1}^{n} 2h^2 \\
&= 2nh^2 \\
&= 2(b-a)^2/n.
\end{aligned}$$

□

With this in mind, we *define*

$$\int_a^b (dW_t)^2 = b - a$$

and observe that we have shown that

$$\int_a^b (dW_t)^2 = \int_a^b dt,$$

for any $0 \le a < b$. Thus we have really shown the Itô rule

$$dW_t^2 = dt.$$

Using a very similar technique, we can also prove that

$$dt dW_t = 0.$$

We first define

$$J_n[a,b] = \sum_{k=1}^{n} h \left(W_{a+kh} - W_{a+(k-1)h}\right),$$

where $nh = b - a$, as before.

Proposition 3.10. *We have* $\mathbb{E} J_n[a,b] = 0$ *and* $\operatorname{var} J_n[a,b] = (b-a)^3/n^2$.

Proof. Firstly,

$$\mathbb{E} J_n[a,b] = \sum_{k=1}^{n} \mathbb{E} h \left(W_{a+kh} - W_{a+(k-1)h}\right) = 0.$$

The variance satisfies

$$\begin{aligned}
\operatorname{var} J_n[a,b] &= \sum_{k=1}^{n} \operatorname{var} h \left(W_{a+kh} - W_{a+(k-1)h}\right) \\
&= \sum_{k=1}^{n} h^2 \operatorname{var} \left(W_{a+kh} - W_{a+(k-1)h}\right) \\
&= \sum_{k=1}^{n} h^3 \\
&= nh^3 \\
&= (b-a)^3/n^2.
\end{aligned}$$

□

With this in mind, we *define*

$$\int_a^b dt dW_t = 0, \qquad \text{for any } 0 \le a < b,$$

and observe that we have shown that

$$dt dW_t = 0.$$

Exercise 3.3. *Setting $nh = b - a$, define*

$$K_n[a,b] = \sum_{k=1}^{n} h^2.$$

Prove that $K_n[a,b] = (b-a)^2/n \to 0$, as $n \to \infty$. Thus

$$\int_a^b (dt)^2 = 0,$$

for any $0 \le a < b$. Hence we have $(dt)^2 = 0$.

Proposition 3.11 (Itô Rules). *We have $dW_t^2 = dt$ and $dW_t dt = dt^2 = 0$.*

Proof. See Propositions 3.9, 3.10 and Exercise 3.3 □

The techniques used in Propositions 3.9 and 3.10 are crucial examples of the basics of stochastic integration. We can generalize this technique to compute other useful stochastic integrals, as we shall now see. However, computing these stochastic integrals directly from limits of stochastic sums is cumbersome compared to direct use of the Itô rules: compare the proof of Proposition 3.12 to the simplicity of Example 3.3.

Proposition 3.12. *We have*

$$\int_0^t W_s dW_s = \frac{1}{2}\left(W_t^2 - t\right).$$

Proof. We have already seen that, when $h = t/n$,

$$\sum_{k=1}^{n} \left(W_{kh} - W_{(k-1)h}\right)^2 \to t, \tag{3.16}$$

as $n \to \infty$. Further, we shall use the telescoping sum

$$\sum_{k=1}^{n} \left(W_{kh}^2 - W_{(k-1)h}^2 \right) = W_{nh}^2 - W_0^2 = W_t^2. \tag{3.17}$$

Subtracting (3.16) from (3.17), we obtain

$$\sum_{k=1}^{n} \left[\left(W_{kh}^2 - W_{(k-1)h}^2 \right) - \left(W_{kh} - W_{(k-1)h} \right)^2 \right] = 2 \sum_{k=1}^{n} W_{(k-1)h} \left(W_{kh} - W_{(k-1)h} \right). \tag{3.18}$$

Now the LHS converges to $W_t^2 - t$, whilst the RHS converges to

$$2 \int_0^t W_s dW_s,$$

whence (3.12). □

Example 3.1. *Here we shall derive a useful formula for*

$$\int_0^t f(s) dW_s, \tag{3.19}$$

where f is continuously differentiable. The corresponding discrete stochastic sum is

$$S_n = \sum_{k=1}^{n} f(kh) \left(W_{kh} - W_{(k-1)h} \right) \tag{3.20}$$

where $nh = t$, as usual. The key trick is to introduce another telescoping sum:

$$\sum_{k=1}^{n} \left(f(kh) W_{kh} - f((k-1)h) W_{(k-1)h} \right) = f(t) W_t. \tag{3.21}$$

Subtracting (3.21) *from* (3.20) *we find*

$$\begin{aligned} S_n - f(t) W_t &= - \sum_{k=1}^{n} \left(f(kh) - f((k-1)h) \right) W_{(k-1)h} \\ &= - \sum_{k=1}^{n} \left(h f'(kh) + O(h^2) \right) W_{(k-1)h} \\ &\to - \int_0^t f'(s) W_s ds, \end{aligned} \tag{3.22}$$

as $n \to \infty$. Hence

$$\int_0^t f(s) dW_s = f(t) W_t - \int_0^t f'(s) W_s \, ds. \tag{3.23}$$

Exercise 3.4. *Modify the technique of Example 3.1 to prove that*

$$\mathbb{E} \left[\left(\int_0^t h(s) dW_s \right)^2 \right] = \int_0^t h(s)^2 \, ds. \tag{3.24}$$

This is the Itô isometry property*.*

We now come to Itô's lemma itself.

Lemma 3.13 (Itô's Lemma for univariate functions)**.** *If f is any infinitely differentiable univariate function and $X_t = f(W_t)$, then*

$$dX_t = f'(W_t)dW_t + \frac{1}{2}f^{(2)}(W_t)dt. \tag{3.25}$$

Proof. We have

$$\begin{aligned} X_{t+dt} &= f(W_{t+dt}) \\ &= f(W_t + dW_t) \\ &= f(W_t) + f'(W_t)dW_t + \frac{1}{2}f^{(2)}(W_t)dW_t^2 \\ &= X_t + f'(W_t)dW_t + \frac{1}{2}f^{(2)}(W_t)dt. \end{aligned}$$

Subtracting X_t from both sides gives (3.25). □

Example 3.2. *Let $X_t = e^{cW_t}$, where $c \in \mathbb{C}$. Then, setting $f(x) = \exp(cx)$ in Lemma 3.13, we obtain*

$$dX_t = X_t\left(cdW_t + \frac{1}{2}c^2dt\right).$$

Example 3.3. *Let $X_t = W_t^2$. Then, setting $f(x) = x^2$ in Lemma 3.13, we obtain*

$$dX_t = 2W_tdW_t + dt.$$

If we integrate this from 0 to T, say, then we obtain

$$X_T - X_0 = 2\int_0^T W_tdW_t + \int_0^T dt,$$

or

$$W_T^2 = 2\int_0^T W_tdW_t + T,$$

that is

$$\int_0^T W_tdW_t = \frac{1}{2}\left(W_T^2 - T\right).$$

This is an excellent example of the Itô rules greatly simplifying direct calculation with stochastic sums, because it is much easier than the direct proof of Proposition 3.12.

Example 3.4. *Let $X_t = W_t^n$, where n can be any positive integer. Then, setting $f(x) = x^n$ in Lemma 3.13, we obtain*

$$dX_t = nW_t^{n-1}dW_t + \frac{1}{2}n(n-1)W_t^{n-2}dt.$$

We can also integrate Itô's Lemma, as follows.

Example 3.5. *Integrating* (3.25) *from a to b, we obtain*

$$\int_a^b dX_t = \int_a^b f'(W_t)dW_t + \frac{1}{2}\int_a^b f^{(2)}(W_t)dt,$$

i.e.

$$X_b - X_a = \int_a^b f'(W_t)dW_t + \frac{1}{2}\int_a^b f^{(2)}(W_t)dt.$$

Lemma 3.13 is not quite sufficient to deal with geometric Brownian motion, hence the following bivariate variant.

Lemma 3.14 (Itô's Lemma for bivariate functions)**.** *If $g(x_1, t)$, for $x_1, t \in \mathbb{R}$, is any infinitely differentiable function and $Y_t = g(W_t, t)$, then*

$$dY_t = \frac{\partial g}{\partial x_1}(W_t, t)dW_t + \left(\frac{1}{2}\frac{\partial^2 g}{\partial x_1^2}(W_t, t) + \frac{\partial g}{\partial t}(W_t, t)\right) dt. \tag{3.26}$$

Proof. We have

$$\begin{aligned} Y_{t+dt} &= g(W_{t+dt}, t + dt) \\ &= g(W_t + dW_t, t + dt) \\ &= g(W_t, t) + \frac{\partial g}{\partial x_1}(W_t, t)dW_t + \frac{1}{2}\frac{\partial^2 g}{\partial x_1^2}(W_t, t)dW_t^2 + \frac{\partial g}{\partial t}(W_t, t)dt \\ &= g(W_t, t) + \frac{\partial g}{\partial x_1}(W_t, t)dW_t + \left(\frac{1}{2}\frac{\partial^2 g}{\partial x_1^2}(W_t, t) + \frac{\partial g}{\partial t}(W_t, t)\right) dt \end{aligned}$$

Subtracting Y_t from both sides gives (3.26).

□

Example 3.6. *Let $X_t = e^{\alpha+\beta t+\sigma W_t}$. Then, setting $f(x_1, t) = \exp(\alpha + \beta t + \sigma x_1)$ in Lemma 3.13, we obtain*

$$dX_t = X_t \left(\sigma dW_t + \left(\frac{1}{2}\sigma^2 + \beta\right) dt\right).$$

Example 3.7. *Let $X_t = e^{\alpha+(r-\sigma^2/2)t+\sigma W_t}$. Then, setting $\beta = r-\sigma^2/2$ in Example 3.6, we find*

$$dX_t = X_t (\sigma dW_t + rdt).$$

Exercise 3.5. *Let $X_t = W_t^2 - t$. Find dX_t.*

3.8. **Itô rules and SDEs.** Suppose now that the asset price S_t is given by the SDE

$$dS_t = S_t (\mu dt + \sigma dW_t), \tag{3.27}$$

that is, S_t is a geometric Brownian motion. Then the Itô rules imply that

$$(dS_t)^2 = \sigma^2 S_t^2 dt. \tag{3.28}$$

Hence, if we define $X_t = f(S_t)$, then

$$dX_t = f'(S_t)dS_t + \frac{1}{2}f^{(2)}(S_t)(dS_t)^2 = \sigma f'(S_t)S_t dW_t + dt\left(\mu f'(S_t)S_t + \frac{1}{2}\sigma^2 S_t^2 f^{(2)}(S_t)\right). \tag{3.29}$$

We illustrate this with the particularly important example of solving the SDE for geometric Brownian motion.

Example 3.8. *If $f(x) = \log x$, then $f'(x) = 1/x$, $f^{(2)}(x) = -1/x^2$ and (3.29) becomes*

$$dX_t = \sigma\frac{1}{S_t}S_t dW_t + dt\left(\mu\frac{1}{S_t}S_t + \frac{1}{2}\sigma^2 S_t^2\left(\frac{-1}{S_t^2}\right)\right) = \sigma dW_t + dt\left(\mu - \sigma^2/2\right).$$

Integrating from t_0 to t_1, say, we obtain

$$X_{t_1} - X_{t_0} = \sigma\left(W_{t_1} - W_{t_0}\right) + \left(\mu - \sigma^2/2\right)\left(t_1 - t_0\right),$$

or

$$\log \frac{S_{t_1}}{S_{t_0}} = \sigma\left(W_{t_1} - W_{t_0}\right) + \left(\mu - \sigma^2/2\right)\left(t_1 - t_0\right).$$

Taking the exponential of both sides, we obtain

$$S_{t_1} = S_{t_0} e^{\left(r-\sigma^2/2\right)(t_1-t_0)+\sigma\left(W_{t_1}-W_{t_0}\right)}.$$

3.9. **Multivariate Geometric Brownian Motion.** So far we have considered one asset only. In practice, we need to construct a multivariate GBM model that allows us to incorporate dependencies between assets via a covariance matrix. To do this, we first take a vector Brownian motion $\mathbf{W}_t \in \mathbb{R}^n$: its components are independent Brownian motions. Its covariance matrix C_t at time t is simply a multiple of the identity matrix:

$$C_t = \mathbb{E}\mathbf{W}_t\mathbf{W}_t^T = tI.$$

Now take any real, invertible, symmetric $n \times n$ matrix A and define

$$\mathbf{Z}_t = A\mathbf{W}_t.$$

The covariance matrix D_t for this new stochastic process is given by

$$D_t = \mathbb{E}\mathbf{Z}_t\mathbf{Z}_t^T = \mathbb{E}A\mathbf{W}_t\mathbf{W}_t^T A = A\left(\mathbb{E}\mathbf{W}_t\mathbf{W}_t^T\right)A = tA^2,$$

and A^2 is a symmetric positive definite matrix.

Exercise 3.6. *Prove that A^2 is symmetric positive definite if A is real, symmetric and invertible.*

In practice, we calculate the covariance matrix M from historical data, hence must construct a symmetric A satisfying $A^2 = M$. Now a covariance matrix is precisely a symmetric positive definite matrix, so that the following linear algebra is vital. We shall use $\|\mathbf{x}\|$ to denote the Euclidean norm of the vector $\mathbf{x} \in \mathbb{R}^n$, that is

$$\|\mathbf{x}\| = \left(\sum_{k=1}^{n} x_k^2\right)^{1/2}, \qquad \mathbf{x} \in \mathbb{R}^n. \tag{3.30}$$

Further, great algorithmic and theoretical importance attaches to those $n \times n$ matrices which preserve the Euclidean norm. More formally, an $n \times n$ matrix Q is called *orthogonal* if $\|Q\mathbf{x}\| = \|\mathbf{x}\|$, for all $\mathbf{x} \in \mathbb{R}^n$. It turns out that Q is an orthogonal matrix if and only if $Q^T Q = I$, which is equivalent to stating that its columns are orthonormal vectors. See Section 6 for further details.

Theorem 3.15. *Let $M \in \mathbb{R}^{n \times n}$ be symmetric. Then it can be written as $M = QDQ^T$, where Q is an orthogonal matrix and D is a diagonal matrix. The elements of D are the eigenvalues of M, while the columns of Q are the eigenvectors. Further, if M is positive definite, then its eigenvalues are all positive.*

Proof. Any good linear algebra textbook should include a proof of this fact; a proof is given in my numerical linear algebra notes. □

Given the *spectral decomposition* $M = QDQ^T$, with $D = \text{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$, we define

$$D^{1/2} = \text{diag}\,(\lambda_1^{1/2}, \lambda_2^{1/2}, \ldots, \lambda_n^{1/2})$$

when M is positive definite. We can now define the *matrix square-root* $M^{1/2}$ by

$$M^{1/2} = QD^{1/2}Q^T. \tag{3.31}$$

Exercise 3.7. *Prove that* $(M^{1/2})^2 = M$ *directly from* (3.31).

Given the matrix square-root $M^{1/2}$ for a chosen symmetric. positive definite matrix M, we now define the assets

$$S_k(t) = e^{(r - M_{kk}/2)t + \left(M^{1/2}\mathbf{W}_t\right)_k}, \qquad k = 1, 2, \ldots, n, \tag{3.32}$$

where $\left(M^{1/2}\mathbf{W}_t\right)_k$ denotes the kth element of the vector $M^{1/2}\mathbf{W}_t$. We now need to check that our assets remain risk-neutral.

Proposition 3.16. *Let the assets' stochastic processes be defined by* (3.32). *Then*

$$\mathbb{E}S_k(t) = e^{rt},$$

for all $k \in \{1, 2, \ldots, n\}$.

Proof. The key calculation is

$$\begin{aligned}
\mathbb{E}e^{\left(M^{1/2}\mathbf{W}_t\right)_k} &= \mathbb{E}e^{\sum_{\ell=1}^n (M^{1/2})_{k\ell} W_t(\ell)} \\
&= \mathbb{E}\prod_{\ell=1}^n e^{(M^{1/2})_{k\ell} W_t(\ell)} \\
&= \prod_{\ell=1}^n \mathbb{E}e^{(M^{1/2})_{k\ell} W_t(\ell)} \\
&= \prod_{\ell=1}^n e^{(M^{1/2})_{k\ell}^2 t/2} \\
&= e^{(t/2)\sum_{\ell=1}^n (M^{1/2})_{k\ell}^2} \\
&= e^{(t/2)M_{kk}},
\end{aligned}$$

using the independence of the components of $\mathbf{W}_t$. □

Exercise 3.8. *Compute* $\mathbb{E}[S_k(t)^2]$.

Exercise 3.9. *What's the covariance matrix for the assets* $S_1(t), \ldots, S_n(t)$*?*

In practice, it is usually easier to describe the covariance structure of multivariate Brownian motion via the Itô rules, which take the simple form

$$d\mathbf{W}_t d\mathbf{W}_t^T = M\ dt, \tag{3.33}$$

where $M \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix and $\mathbf{W}_t$ is an n-dimensional Brownian motion.

Proposition 3.17. *If* $X_t = f(\mathbf{W}_t)$, *then*

$$dX_t = \nabla f(\mathbf{W}_t)^T d\mathbf{W}_t + \frac{1}{2} d\mathbf{W}_t^T D^2 f(\mathbf{W}_t) d\mathbf{W}_t \tag{3.34}$$

or

$$dX_t = \sum_{j=1}^n \frac{\partial f}{\partial x_j} dW_{j,t} + \frac{dt}{2}\sum_{j=1}^n\sum_{k=1}^n \frac{\partial^2 f}{\partial x_j \partial x_k} M_{jk}. \tag{3.35}$$

Proof. This is left as an exercise. □

Example 3.9. *If $n = 2$ and*

$$f(x_1, x_2) = e^{a_1 x_1 + a_2 x_2},$$

and the correlated Brownian motions $W_{1,t}$ and $W_{2,t}$ satisfy

$$dW_{1,t} dW_{2,t} = \rho dt,$$

for some constant correlation coefficient $\rho \in [-1, 1]$, then $X_t = f(W_{1,t}, W_{2,t})$ satisfies

$$dX_t = \left(a_1 dW_{1,t} + a_2 dW_{2,t} + \frac{1}{2} dt \left(a_1^2 + 2\rho a_1 a_2 + a_2^2 \right) \right) X_t.$$

Example 3.10. *If $n = 3$ and*

$$f(x_1, x_2, x_3) = e^{a_1 x_1 + a_2 x_2 + a_3 x_3},$$

and the correlated Brownian motions $W_{1,t}, W_{2,t}, W_{3,t}$ satisfy

$$dW_{1,t} dW_{2,t} = M_{12} dt, \quad dW_{2,t} dW_{3,t} = M_{23} dt, \quad dW_{3,t} dW_{1,t} = M_{31} dt,$$

where $M \in \mathbb{R}^{3\times 3}$ is a symmetric positive definite matrix which also satisfies

$$M_{11} = M_{22} = M_{33} = 1,$$

then $X_t = f(W_{1,t}, W_{2,t}, W_{3,t})$ satisfies

$$dX_t = \left(a_1 dW_{1,t} + a_2 dW_{2,t} + a_3 dW_{3,t} + \frac{1}{2} dt \left(a_1^2 + a_2^2 + a_3^2 + 2a_2 a_3 M_{23} + 2a_3 a_1 M_{31} + 2a_1 a_2 M_{12} \right) \right) X_t.$$

Example 3.11. *If*

$$f(\mathbf{x}) = e^{\mathbf{a}^T \mathbf{x}}, \qquad \mathbf{x} \in \mathbb{R}^n,$$

then

$$\nabla f(\mathbf{x}) = \mathbf{a} f(\mathbf{x})$$

and

$$D^2 f(\mathbf{x}) = \mathbf{a}\mathbf{a}^T f(\mathbf{x}).$$

Let $\mathbf{W}_t$ be any n-dimensional Brownian motion satisfying

$$d\mathbf{W}_t d\mathbf{W}_t^T = M dt,$$

where $M \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix. Then $X_t = f(\mathbf{W}_t)$ satisfies

$$dX_t = \left(\mathbf{a}^T d\mathbf{W}_t + \frac{1}{2} d\mathbf{W}_t^T M d\mathbf{W}_t \right) f(\mathbf{x}),$$

or, in coordinate form,

$$dX_t = \left(\sum_{j=1}^n a_j dW_{j,t} + \frac{1}{2} dt \sum_{j=1}^n \sum_{k=1}^n a_j a_k M_{jk} \right) f(\mathbf{x}).$$

Proposition 3.18. *If* $Y_t = g(\mathbf{W}_t, t)$, *then*

$$dY_t = \nabla g(\mathbf{W}_t)^T d\mathbf{W}_t + \frac{\partial g}{\partial t} dt + \frac{1}{2} d\mathbf{W}_t^T D^2 g(\mathbf{W}_t) d\mathbf{W}_t \tag{3.36}$$

or

$$dY_t = \sum_{j=1}^n \frac{\partial g}{\partial x_j} dW_{j,t} + \frac{\partial g}{\partial t} dt + \frac{dt}{2} \sum_{j=1}^n \sum_{k=1}^n \frac{\partial^2 g}{\partial x_j \partial x_k} M_{jk}. \tag{3.37}$$

Proof. Exercise. □

Example 3.12. *If* $n = 2$ *and*

$$g(x_1, x_2) = e^{a_1 x_1 + a_2 x_2 + bt},$$

and the correlated Brownian motions $W_{1,t}$ *and* $W_{2,t}$ *satisfy*

$$dW_{1,t} dW_{2,t} = \rho dt,$$

for some constant correlation coefficient $\rho \in [-1, 1]$, *then* $X_t = f(W_{1,t}, W_{2,t})$ *satisfies*

$$dX_t = \left(a_1 dW_{1,t} + a_2 dW_{2,t} + b dt + \frac{1}{2} dt \left(a_1^2 + 2\rho a_1 a_2 + a_2^2 \right) \right) X_t.$$

Example 3.13. *If*

$$g(\mathbf{x}, t) = e^{\mathbf{a}^T \mathbf{x} + bt}, \qquad \mathbf{x} \in \mathbb{R}^n,$$

then

$$\nabla g(\mathbf{x}, t) = \mathbf{a} g(\mathbf{x}, t), \quad \frac{\partial g}{\partial t} = b g(\mathbf{x}, t),$$

and

$$D^2 g(\mathbf{x}, t) = \mathbf{a}\mathbf{a}^T g(\mathbf{x}, t).$$

Let W_t *be any* n*-dimensional Brownian motion satisfying*

$$dW_t dW_t^T = M dt,$$

where $M \in \mathbb{R}^{n \times n}$ *is a symmetric positive definite matrix. Then* $Y_t = g(\mathbf{W}_t, t)$ *satisfies*

$$dY_t = \left(\mathbf{a}^T d\mathbf{W}_t + b dt + \frac{1}{2} d\mathbf{W}_t^T M d\mathbf{W}_t \right) Y_t,$$

or, in coordinate form

$$dY_t = \left(\sum_{j=1}^n a_j dW_{j,t} + b dt + \frac{1}{2} dt \sum_{j=1}^n \sum_{k=1}^n a_j a_k M_{jk} \right) Y_t.$$

3.10. **The Ornstein–Uhlenbeck Process.** This interesting SDE displays mean-reversion and requires a slightly more advanced technique. We consider the SDE

$$dX_t = -\alpha X_t dt + \sigma dW_t, \qquad t \geq 0, \tag{3.38}$$

where $\alpha > 0$ and $\sigma \geq 0$ are constants and $X_0 = x_0$.

It's very useful to consider the special case $\sigma = 0$ first, in which case the SDE (3.38) becomes the ODE

$$\frac{dX_t}{dt} + \alpha X_t = 0. \tag{3.39}$$

There is a standard method for solving (3.39) using an *integrating factor*. Specifically, if we multiply (3.39) by $\exp(\alpha t)$, then we obtain

$$\frac{d}{dt}\left(X_t e^{\alpha t}\right) = 0, \tag{3.40}$$

so that $X_t \exp(\alpha t)$ is constant. Hence, recalling the initial condition $X_0 = x_0$, we must have

$$X_t = x_0 e^{-\alpha t}. \tag{3.41}$$

Thus the solution decays exponentially to zero, at a rate determined by the positive constant α, for any initial value x_0.

Fortunately the integrating factor method also applies to the $\sigma > 0$ case, with a little more work. Multiplying (3.38) by $\exp \alpha t$, we obtain

$$e^{\alpha t}\left(dX_t + \alpha X_t dt\right) = \sigma e^{\alpha t} dW_t,$$

or

$$d\left(X_t e^{\alpha t}\right) = \sigma e^{\alpha t} dW_t, \tag{3.42}$$

using the infinitesimal increments variant on the product rule for differentiation. Integrating (3.42) from 0 to s, we find

$$X_s e^{\alpha s} - x_0 = \int_0^s d\left(X_t e^{\alpha t}\right) = \sigma \int_0^s e^{\alpha t} dW_t, \tag{3.43}$$

or

$$X_s = x_0 e^{-\alpha s} + e^{-\alpha s} \int_0^s e^{\alpha t} dW_t. \tag{3.44}$$

We can say more using the following important property of stochastic integrals.

Proposition 3.19. *Let $f : [0, \infty) \to \mathbb{R}$ be any infinitely differentiable function and define the stochastic process*

$$F_s = \int_0^s f(t) dW_t. \tag{3.45}$$

Then $\mathbb{E} F_s = 0$ and

$$\operatorname{var} F_s = \int_0^s f(t)^2 dt. \tag{3.46}$$

Proof. The key point is that (3.45) is the limit of the stochastic sum

$$S_n = \sum_{k=1}^{n} f(kh)\left(W_{kh} - W_{(k-1)h}\right),$$

where $h > 0$ and $nh = s$. Now the increments $W_{kh} - W_{(k-1)h}$ are independent and satisfy

$$W_{kh} - W_{(k-1)h} \sim N(0, h),$$

by the axioms of Brownian motion, so

$$\mathbb{E}S_n = 0,$$

for all n. By independence of the terms in the sum, we see that

$$\begin{aligned}
\operatorname{var} S_n &= \sum_{k=1}^{n} \operatorname{var}\left(f(kh)\left(W_{kh} - W_{(k-1)h}\right)\right) \\
&= \sum_{k=1}^{n} f(kh)^2 \operatorname{var}\left(W_{kh} - W_{(k-1)h}\right) \\
&= h \sum_{k=1}^{n} f(kh)^2 \\
&\to \int_0^s f(t)^2\, dt,
\end{aligned}$$

as $n \to \infty$. □

Applying Proposition 3.19 to the Ornstein–Uhlenbeck process solution (3.44), we obtain $\mathbb{E}X_s = x_0 \exp(-\alpha s)$ and

$$\begin{aligned}
\operatorname{var} X_s &= \operatorname{var} \sigma e^{-\alpha s} \int_0^s e^{\alpha t} dW_t \\
&= \sigma^2 e^{-2\alpha s} \int_0^s e^{2\alpha t} dt \\
&= \sigma^2 e^{-2\alpha s} \left(\frac{e^{2\alpha s} - 1}{2\alpha}\right) \\
&= \sigma^2 \left(\frac{1 - e^{-2\alpha s}}{2\alpha}\right).
\end{aligned}$$

3.11. **Δ-Hedging for GBM.** We begin with the real world asset price

$$S_t = e^{\alpha + \beta t + \sigma W_t}, \tag{3.47}$$

where we do **not** assume there is any connection between the parameters α, β and σ: this is **not** risk-neutral GBM. It is a simple exercise in Itô calculus (see Example 3.6) to prove that

$$dS_t = S_t\left(\sigma dW_t + \left(\beta + \sigma^2/2\right) dt\right) \tag{3.48}$$

and

$$\left(dS_t\right)^2 = \sigma^2 S_t^2 dt. \tag{3.49}$$

By analogy with delta hedging in the Binomial Model (2.4), let us assume that $S_t = S$ and define the portfolio

$$\Pi_t = f(S_t, t) - \Delta S_t, \tag{3.50}$$

where Δ is a constant. Then

$$\begin{aligned}
\Pi_{t+dt} &= f(S_{t+dt}, t+dt) - \Delta S_{t+dt} \\
&= f(S + dS_t, t+dt) - \Delta S - \Delta dS_t \\
&= f(S,t) + dS_t f_S + \frac{1}{2} dS_t^2 f_{SS} + dt f_t - \Delta S - \Delta dS_t \\
&= \Pi_t + dS_t \left(f_S - \Delta\right) + \frac{1}{2} dS_t^2 f_{SS} + dt f_t \\
&= \Pi_t + dS_t \left(f_S - \Delta\right) + \left(\frac{1}{2}\sigma^2 S^2 f_{SS} + f_t\right) dt. \qquad (3.51)
\end{aligned}$$

In other words, we have the infinitesimal increment

$$\Pi_{t+dt} - \Pi_t = d\Pi_t = dS_t \left(f_S - \Delta\right) + \left(\frac{1}{2}\sigma^2 S^2 f_{SS} + f_t\right) dt. \qquad (3.52)$$

Thus we eliminate the stochastic dS_t component by setting

$$\Delta = f_S \qquad (3.53)$$

and (3.51) then becomes

$$d\Pi_t = \left(f_t + \frac{1}{2}\sigma^2 S^2 f_{SS}\right) dt, \qquad (3.54)$$

or

$$\frac{d\Pi_t}{dt} = f_t + \frac{1}{2}\sigma^2 S^2 f_{SS}. \qquad (3.55)$$

Now there is no stochastic component in (3.54), so we must also have

$$\frac{d\Pi_t}{dt} = r\Pi_t = r\left(f - f_S S\right), \qquad (3.56)$$

because all deterministic assets must grow at the risk-free rate. Equating (3.55) and (3.56) yields

$$f_t + \frac{1}{2}\sigma^2 S^2 f_{SS} = r\left(f - f_S S\right), \qquad (3.57)$$

or

$$f_t - rf + rf_S S + \frac{1}{2}\sigma^2 S^2 f_{SS} = 0, \qquad (3.58)$$

which is the Black–Scholes PDE.

It is often useful to restate the Black–Scholes PDE in terms of the logarithm of the asset price, i.e. via $S = e^x$. Thus

$$\partial_S \frac{dS}{dx} = \partial_x,$$

or

$$S\partial_S = \partial_x. \qquad (3.59)$$

Hence

$$\begin{aligned}
\partial_{xx} &= S\partial_S \left(S\partial_S\right) \\
&= S\left(\partial_S + S\partial_{SS}\right) \\
&= S\partial_S + S^2\partial_{SS}, \qquad (3.60)
\end{aligned}$$

or, using (3.59),

$$S^2\partial_{SS} = \partial_{xx} - \partial_x. \tag{3.61}$$

Therefore substituting (3.59) and (3.61) in (3.58) yields

$$\begin{aligned} 0 &= f_t - rf + rf_x + \frac{1}{2}\sigma^2\left(f_{xx} - f_x\right) & (3.62) \\ &= f_t - rf + \left(r - \sigma^2/2\right)f_x + \frac{1}{2}\sigma^2 f_{xx}. & (3.63) \end{aligned}$$

4. The Geometric Brownian Motion Universe

We shall begin with a brisk introduction to the main topics, filling in the details later. The real economy is vastly complex, so mathematical finance begins with vast oversimplification.

Let r be the *risk-free* interest rate, which we shall assume constant. This is really the interest paid by the state when borrowing money via selling bonds, and it is **nominally** risk-free in any state that issues its own currency, although the real value of that currency can greatly decrease. We assume that everyone in our mathematical economy can borrow and lend at this rate, so that such debts (or investments, if lent) satisfy $B_t = B_0 \exp(rt)$. In reality, banks and companies borrow and lend at a higher rate $r + \delta$, where δ increases with the perceived risk of the lender, but this complication is ignored here.

Notation: In most (but not all) areas of mathematics, a function B depending on time t would be denoted $B(t)$, but mathematical finance often uses the alternative notation B_t which is very common in probability theory, statistics and economics. I shall be consistent in using $S(t)$ to denote the share price in Section 2, but we shall move to S_t in Section 3.

We shall assume that every **risky** asset (such as a share) is described by a random process called *geometric Brownian motion* (GBM):

$$S(t) = S(0)e^{\beta t + \sigma W_t}, \qquad t > 0, \tag{4.1}$$

where W_t denotes Brownian motion, $\beta \in \mathbb{R}$ and σ is a non-negative parameter called the *volatility* of the asset. You can think of Brownian motion as an important generalization of random walk, but we shall postpone its detailed definition and properties until Section 3. Fortunately all we need for now is the fundamental property that W_T is a normal (or Gaussian) random variable with mean zero and variance T, that is,

$$W_t \sim N(0,t), \qquad \text{for all } t > 0. \tag{4.2}$$

As we shall see later, option pricing requires us to use $\beta = r - \sigma^2/2$, that is,

$$S(t) = S(0)e^{(r-\sigma^2/2)t + \sigma W_t}, \qquad t > 0, \tag{4.3}$$

and this is usually called *risk neutral* GBM. The reason for the disappearance of the parameter β when pricing options is rather deep and extremely important, but will be explained later. All you need to know at present is that (4.1) is the mathematical model for share prices in the real world, but the risk neutral variant (4.3) is used when pricing options, i.e. contracts whose value depends on the asset price.

Thus, when pricing options, to generate sample prices $S(T)$ at some future time T given the initial price $S(0)$, we use

$$S(T) = S(0)\exp\Big((r - \sigma^2/2)T + \sigma\sqrt{T}Z_T\Big), \qquad \text{where } Z_T \sim N(0,1), \tag{4.4}$$

because $W_T \sim N(0,T)$, so that we can write $W_T = T^{1/2}Z_T$.

Example 4.1. *Generating sample prices at a fixed time T using* (4.4) *is particularly easy in Matlab and Octave:*

```
S = S0*exp((r-sigma^2/2)*T + sigma*sqrt(T)*randn(m,1)
```

will construct a column vector of m sample prices once you've defined $S0$, r, σ and T. To calculate the sample average price, we type `sum(S)/m`.

To analytically calculate $\mathbb{E}S(T)$ we need the following simple, yet crucial, lemma.

Lemma 4.1. *If $W \sim N(0,1)$, then $\mathbb{E}\exp(\lambda W) = \exp(\lambda^2/2)$.*

Proof. We have

$$\mathbb{E}e^{\lambda W} = \int_{-\infty}^{\infty} e^{\lambda t}(2\pi)^{-1/2}e^{-t^2/2}\,dt = (2\pi)^{-1/2}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(t^2-2\lambda t)}\,dt.$$

The trick now is to *complete the square* in the exponent, that is,

$$t^2 - 2\lambda t = (t-\lambda)^2 - \lambda^2.$$

Thus

$$\mathbb{E}e^{\lambda W} = (2\pi)^{-1/2}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}([t-\lambda]^2 - \lambda^2)\right)\,dt = e^{\lambda^2/2}.$$

This is also described in detail in Example 5.3. □

Lemma 4.2. *For every $\sigma \geq 0$, we have the expected growth*

$$\mathbb{E}S(t) = S(0)e^{rt}, \qquad t \geq 0. \tag{4.5}$$

Proof. This is an easy consequence of Lemma 4.1. □

The option pricing risk-neutral geometric Brownian motion universe might therefore seem rather strange, because *every* asset has the same expected growth e^{rt} as the risk-free interest rate. Thus our universe of all possible assets in a risk-neutral world is specified by one parameter only: the volatility σ. Please do remember that this is *not* the asset price in the market, but simply a mathematical device required for pricing options based on the asset.

A *financial derivative* is any function $f(S,t)$. We shall concentrate on the following particular class of derivatives.

Definition 4.1. *A* European option *is any function $f \equiv f(S,t)$ that satisfies the conditional expectation equation*

$$f(S(t),t) = e^{-rh}\mathbb{E}\Big(f(S(t+h),t+h)|S(t)\Big), \qquad \textit{for any } h > 0. \tag{4.6}$$

We shall often simply write this as

$$f(S(t),t) = e^{-rh}\mathbb{E}f(S(t+h),t+h)$$

but you should take care to remember that this is an expected future value given *the asset's current value $S(t)$. We see that (4.6) describes a contract $f(S,t)$ whose* *current value is the* discounted value *of its expected future value in the risk-neutral GBM universe.*

We can learn a great deal by studying the mathematical consequences of (4.6) and (4.3).

Example 4.2. *A* plain vanilla European put option *is simply an insurance contract that allows us to sell one unit of the asset, for exercise price K, at time T in the future. If the asset's price $S(T)$ is less than K at this expiry time, then the option is worth $K - S(T)$, otherwise it's worthless. Such contracts protect us if we're worried that the asset's price might drop. The pricing problem here to calculate the value of the contract at time zero given its value at expiry, namely*

$$f_P(S(T), T) = (K - S(T))_+ \,, \tag{4.7}$$

where $(z)_+ := max\{z, 0\}$.

Typically, we know the value of the option $f(S(T), T)$ for all values of the asset $S(T)$ at some future time T. Our problem is to compute its value at some earlier time, because we're buying or selling this option.

Example 4.3. *A* plain vanilla European call option *gives us the right to buy one unit of the asset at the exercise price K at time T. If the asset's price $S(T)$ exceeds K at this expiry time, then the option is worth $S(T) - K$, otherwise it's worthless, implying the expiry value*

$$f_C(S(T), T) = (S(T) - K)_+ \,, \tag{4.8}$$

using the same notation as Example 4.2. Such contracts protect us if we're worried that the asset's price might rise.

How do we compute $f(S(0), 0)$? The difficult part is computing the expected future value $\mathbb{E}f(S(T), T)$. This can be done analytically for a tiny number of options, including the European Put and Call (see Theorem 4.5), but usually we must resort to a numerical calculation. This leads us to our first algorithm: *Monte Carlo simulation*. Here we choose a large integer N and generate N pseudo-random numbers $Z_1, Z_2, \ldots, Z_N$ that have the normalized Gaussian distribution; in Matlab, we simply write `Z = randn(N,1)`. Using (4.3), these generate the future asset prices

$$S_k = S(0) \exp\Big((r - \frac{\sigma^2}{2})T + \sigma\sqrt{T} Z_k \Big), \qquad k = 1, \ldots, N. \tag{4.9}$$

We then approximate the future expected value by an average, that is, we take

$$f(S(0), 0) \approx \frac{e^{-rT}}{N} \sum_{k=1}^{N} f(S_k, T). \tag{4.10}$$

Monte Carlo simulation has the great advantage that it is extremely simple to program. Its disadvantage is that the error is usually a multiple of $1/\sqrt{N}$, so that very large N is needed for high accuracy (each decimal place of accuracy requires about a hundred times more work). We note that (4.10) will compute the value of *any* European option that is completely defined by a known final value $f(S(T), T)$.

We shall now use Monte Carlo to approximately evaluate the European Call and Put contracts. In fact, Put-Call parity, described below in Theorem 4.3, implies that we only need a program to calculate one of these, because they are related by the simple formula

$$f_C(S(0), 0) - f_P(S(0), 0) = S(0) - Ke^{-rT}. \tag{4.11}$$

Here's the Matlab program for the European Put.

```
%
% These are the parameters chosen in Example 11.6 of
% OPTIONS, FUTURES AND OTHER DERIVATIVES,
% by John C. Hull (Prentice Hall, 4th edn, 2000)
%
%% initial stock price
S0 = 42;
% unit of time = year
% 250 working days per year
% continuous compounding risk-free rate
r = 0.1;
% exercise price
K = 40;
% time to expiration in years
T = 0.5;
% volatility of 20 per cent annually
sigma = 0.2;
% generate asset prices at expiry
Z = randn(N,1);
ST = S0*exp( (r-(sigma^2)/2)*T + sigma*sqrt(T)*Z );
% calculate put contract values at expiry
fput = max(K - ST,0.0);
% average put values at expiry and discount to present
mc_put = exp(-r*T)*sum(fput)/N
% calculate analytic value of put contract
wK = (log(K/S0) - (r - (sigma^2)/2)*T)/(sigma*sqrt(T));
a_put = K*exp(-r*T)*Phi(wK) - S0*Phi(wK - sigma*sqrt(T))
```

The function `Phi` denotes the cumulative distribution function for the normalized Gaussian distribution, that is,

$$\Phi(x) = \mathbb{P}(Z \le x) = \int_{-\infty}^{x} (2\pi)^{-1/2} e^{-s^2/2}\, ds, \qquad \text{for } x \in \mathbb{R}, \tag{4.12}$$

where $Z \sim N(0, 1)$.

Unfortunately, Matlab only provides the very similar *error function*, defined by

$$\operatorname{erf}(y) = \frac{2}{\sqrt{\pi}} \int_0^y \exp(-s^2)\, ds, \qquad y \in \mathbb{R}.$$

It's not hard to prove that

$$\Phi(t) = \frac{1}{2}\Big(1 + \operatorname{erf}(t/\sqrt{2})\Big), \qquad t \in \mathbb{R}.$$

We can add this to Matlab using the following function.

```
function Y = Phi(t)
Y = 0.5*(1.0 + erf(t/sqrt(2)));
```

We have only revealed the tip of a massive iceberg in this brief introduction. Firstly, the Black–Scholes model, where asset prices evolve according to (4.3), is rather poor: reality is far messier. Further, there are many types of option which are *path-dependent*: the value of the option at expiry depends not only on the final price $S(T)$, but on its previous values $\{S(t) : 0 \le t \le T\}$. In particular, there are *American options*, where the contract can be exercised at any time before its expiry. All of these points will be addressed in our course, but you should find that Hull's book provides excellent background reading (although his mathematical treatment is often sketchy). Higham provides a clear Matlab-based exposition.

Although the future expected value usually requires numerical computation, there are some simple cases that are analytically tractable. These are particularly important because they often arise in examinations!

4.1. **European Puts and Calls.** It's not too hard to calculate the values of these options analytically. Further, the next theorem gives an important relation between the prices of call and put options.

Theorem 4.3 (Put-Call parity)**.** *European Put and Call options, each with exercise price K and expiry time T, satisfy*

$$f_C(S,t) - f_P(S,t) = S - Ke^{-r\tau}, \qquad \text{for } S \in \mathbb{R} \text{ and } 0 \le t \le T, \tag{4.13}$$

where $\tau = T - t$, the time-to-expiry.

Proof. The trick is the observation that

$$y = y_+ - (-y)_+,$$

for any $y \in \mathbb{R}$. Thus

$$\begin{aligned} S(T) - K &= (S(T) - K)_+ - (K - S(T))_+ \\ &= f_C(S(T),T) - f_P(S(T),T), \end{aligned}$$

which implies

$$e^{-r\tau}\mathbb{E}\left(S(T) - K | S(t) = S\right) = f_C(S,t) - f_P(S,t).$$

Now

$$\begin{aligned} \mathbb{E}\left(S(T)|S(t) = S\right) &= (2\pi)^{-1/2}\int_{-\infty}^{\infty} Se^{(r-\sigma^2/2)\tau+\sigma\sqrt{\tau}w}e^{-w^2/2}\,dw \\ &= Se^{(r-\sigma^2/2)\tau}(2\pi)^{-1/2}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(w^2-2\sigma\sqrt{\tau}w\right)}\,dw \\ &= Se^{r\tau}, \end{aligned}$$

and some simple algebraic manipulation completes the proof. □

This is a useful check on the Monte Carlo approximations of the options' values. To derive their analytic values, we shall need the cumulative distribution function

$$\Phi(y) = (2\pi)^{-1/2}\int_{-\infty}^{y} e^{-z^2/2}\,dz, \qquad y \in \mathbb{R}, \tag{4.14}$$

for the Gaussian probability density, that is, $\mathbb{P}(Z \le y) = \Phi(y)$ and $\mathbb{P}(a \le Z \le b) = \Phi(b) - \Phi(a)$, for any normalized Gaussian random variable Z. Further, we have the following relation which will be of use in subsequent formulae.

Lemma 4.4. *We have* $1 - \Phi(a) = \Phi(-a)$, *for any* $a \in \mathbb{R}$.

Proof. Observe that

$$\begin{aligned} 1 - \Phi(a) &= \int_a^\infty (2\pi)^{-1/2} e^{-s^2/2}\, ds \\ &= \int_{-\infty}^{-a} (2\pi)^{-1/2} e^{-u^2/2}\, du \\ &= \Phi(-a), \end{aligned}$$

where we have made the substitution $u = -s$. □

Theorem 4.5. *A European Put option satisfies*

$$f_P(S,t) = Ke^{-r\tau}\Phi(w(K)) - S\Phi(w(K) - \sigma\sqrt{\tau}), \quad \text{for } S \in \mathbb{R}, \tag{4.15}$$

where $\tau = T - t$, *i.e. the time-to-expiry, and* $w(K)$ *is defined by the equation*

$$K = Se^{(r-\sigma^2/2)\tau + \sigma\sqrt{\tau}w(K)},$$

that is

$$w(K) = \frac{\log(K/S) - (r - \sigma^2/2)\tau}{\sigma\sqrt{\tau}}. \tag{4.16}$$

Proof. We have

$$\mathbb{E}\left(f_P(S(T),T)|S(t) = S\right) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} \left(K - Se^{(r-\sigma^2/2)\tau + \sigma\sqrt{\tau}w}\right)_+ e^{-w^2/2}\, dw.$$

Now the function

$$w \mapsto K - S\exp((r - \sigma^2/2)\tau + \sigma\sqrt{\tau}w)$$

is strictly decreasing, so that

$$K - Se^{(r-\sigma^2/2)\tau + \sigma\sqrt{\tau}w} \ge 0$$

if and only if $w \le w(K)$, where $w(K)$ is given by (4.16). Hence

$$\begin{aligned} \mathbb{E}\left(f_P(S(T),T)|S(t) = S\right) &= (2\pi)^{-1/2} \int_{-\infty}^{w(K)} \left(K - Se^{(r-\sigma^2/2)\tau + \sigma\sqrt{\tau}w}\right) e^{-w^2/2}\, dw \\ &= K\Phi(w(K)) - Se^{(r-\sigma^2/2)\tau}(2\pi)^{-1/2} \int_{-\infty}^{w(K)} e^{-\frac{1}{2}\left(w^2 - 2\sigma\sqrt{\tau}w\right)}\, dw \\ &= K\Phi(w(K)) - Se^{r\tau}\Phi(w(K) - \sigma\sqrt{\tau}). \end{aligned}$$

Thus

$$\begin{aligned} f_P(S,t) &= e^{-r\tau}\mathbb{E}\left(f_P(S(T),T)|S(t) = S\right) \\ &= Ke^{-r\tau}\Phi(w(K)) - S\Phi(w(K) - \sigma\sqrt{\tau}). \end{aligned}$$

□

There is an almost standard notation for Theorem 4.5, which is contained in the following corollary.

Corollary 4.6. *A European Put option satisfies*

$$f_P(S,t) = Ke^{-r\tau}\Phi(-d_-) - S\Phi(-d_+), \qquad \text{for } S \in \mathbb{R}, \tag{4.17}$$

where $\tau = T - t$, *i.e. the time-to-expiry, and*

$$d_\pm = \frac{\log(S/K) + (r \pm \sigma^2/2)\tau}{\sigma\sqrt{\tau}}. \tag{4.18}$$

Proof. This is simply rewriting Theorem 4.5 in terms of (4.18). □

We can now calculate the price of a European call using Corollary 4.6 and the Put-Call parity Theorem 4.3.

Corollary 4.7. *A European Call option satisfies*

$$f_C(S,t) = S\Phi(d_+) - Ke^{-r\tau}\Phi(d_-), \qquad \text{for } S \in \mathbb{R}, \tag{4.19}$$

where $\tau = T - t$, *i.e. the time-to-expiry, and* $d_\pm$ *is given by* (4.18).

Proof. Theorem 4.3 implies that

$$\begin{aligned} f_C(S,t) &= f_P(S,t) + S - Ke^{-r\tau} \\ &= Ke^{-r\tau}\Phi(-d_-) - S\Phi(-d_+) + S - Ke^{-r\tau} \\ &= S\left(1 - \Phi(-d_+)\right) - Ke^{-r\tau}\left(1 - \Phi(-d_-)\right) \\ &= S\Phi(d_+) - Ke^{-r\tau}\Phi(d_-), \end{aligned}$$

using Lemma 4.4. □

Exercise 4.1. *Modify the proof of Theorem 4.5 to derive the analytic price of a European Call option. Check that your price agrees Corollary 4.7.*

4.2. **Digital Options.** A digital option is simply an option that only takes the values 0 and 1, that is, it is the indicator function for some event. Recall that, for any indicator function I_A, we have

$$\mathbb{E}I_A = \mathbb{P}(A).$$

Our first example is the digital call option with exercise price K and expiry time T is defined by

$$f_{DC}(S(T),T) = \begin{cases} 1 & \text{if } S(T) \ge K, \\ 0 & \text{otherwise.} \end{cases} \tag{4.20}$$

Theorem 4.8. *The digital call option* f_{DC} *satisfies*

$$f_{DC}(S,t) = e^{-r\tau}\Phi(d_-), \tag{4.21}$$

where $\tau = T - t$ *and* d_- *is defined by* (4.18).

Proof. Its price at any earlier time $t \in [0,t)$ is therefore given by

$$f_{DC}(S,t) = e^{-r\tau}\mathbb{E}\left(f_{DC}(S(T),T) | S(t) = S\right)$$

$$= e^{-r\tau}\mathbb{E}f_{DC}(Se^{(r-\sigma^2/2)\tau + \sigma\tau^{1/2}Z}, \tag{4.22}$$

(4.23)

where $Z \sim N(0,1)$.

Now

$$Se^{(r-\sigma^2/2)\tau+\sigma\tau^{1/2}Z} \geq K$$

if and only if

$$\log S + (r-\sigma^2/2)\tau + \sigma\tau^{1/2}Z \geq \log K,$$

because the logarithm is an increasing function. Rearranging this inequality, we find

$$Z \geq -\frac{\log(S/K) - (r-\sigma^2/2)\tau}{\sigma\tau^{1/2}} = -d_-.$$

Thus

$$\begin{aligned} f_{DC}(S,t) &= e^{-r\tau}\mathbb{P}\left(Z \geq -d_-\right) \\ &= e^{-r\tau}\left(1-\Phi(-d_-)\right) \\ &= e^{-r\tau}\Phi(d_-), \end{aligned}$$

by Lemma 4.4. □

It is now simple to define and price the digital put option f_{DP}, which is defined by

$$f_{DP}(S(T),T) = \begin{cases} 1 & \text{if } S(T) < K, \\ 0 & \text{otherwise.} \end{cases} \tag{4.24}$$

A pair of digital put and call options with the same exercise price K and expiry time T satisfy a digital put-call parity relation, specifically

$$f_{DC}(S(T),T) + f_{DP}(S(T),T) \equiv 1,$$

at expiry, which implies

$$f_{DC}(S,t) + f_{DP}(S,t) \equiv e^{-r\tau}, \qquad \text{for } S \in \mathbb{R}, \tag{4.25}$$

where $\tau = T - t$.

Theorem 4.9. *The digital put option f_{DP} satisfies*

$$f_{DP}(S,t) = e^{-r\tau}\Phi(-d_-), \tag{4.26}$$

where $\tau = T - t$ and d_- is defined by (4.18)*.*

Proof. We use (4.25) and Lemma 4.4:

$$f_{DP}(S,t) = e^{-r\tau} - f_{DC}(S,t) = e^{-r\tau}\left(1-\Phi(d_-)\right) = e^{-r\tau}\Phi(-d_-).$$

□

Another way to express digital calls and puts is as follows. Observe that

$$f_{DC}(S(T),T) = (S(T)-K)_+^0 \quad \text{and} \quad f_{DP}(S(T),T) = (K-S(T))_+^0. \tag{4.27}$$

Thus we have shown that

$$\begin{aligned} \mathbb{E}\left((S(T)-K)_+ | S(t)=S\right) &= Se^{r\tau}\Phi(d_+) - K\Phi(d_-), \\ \mathbb{E}\left((S(T)-K)_+^0 | S(t)=S\right) &= \Phi(d_-), \\ \mathbb{E}\left((K-S(T))_+ | S(t)=S\right) &= K\Phi(-d_-) - Se^{r\tau}\Phi(-d_+), \\ \mathbb{E}\left((K-S(T))_+^0 | S(t)=S\right) &= \Phi(-d_-), \end{aligned}$$

5. Mathematical Background Material

I've collected here a miscellany of mathematical methods used (or reviewed) during the course.

5.1. **Probability Theory.**

You may find my more extensive notes on Probability Theory useful:

`http://econ109.econ.bbk.ac.uk/brad/Probability_Course/probnotes.pdf`

A *random variable* X is said to have (continuous) *probability density function* $p(t)$ if

$$\mathbb{P}(a < X < b) = \int_a^b p(t)\, dt. \tag{5.1}$$

We shall assume that $p(t)$ is a continuous function (no jumps in value). In particular, we have

$$1 = \mathbb{P}(X \in \mathbb{R}) = \int_{-\infty}^{\infty} p(t)\, dt.$$

Further, because

$$0 \le \mathbb{P}(a < X < a + \delta a) = \int_a^{a+\delta a} p(t)\, dt \approx p(a)\delta a,$$

for small δa, we conclude that $p(t) \ge 0$, for all $t \ge 0$. In other words, a probability density function is simply a non-negative function $p(t)$ whose integral is one. Here are two fundamental examples.

Example 5.1. *The Gaussian probability density function, with mean μ and variance σ^2, is defined by*

$$p(t) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right). \tag{5.2}$$

We say that the Gaussian is normalized *if $\mu = 0$ and $\sigma = 1$.*

To prove that this is truly a probability density function, we require the important identity

$$\int_{-\infty}^{\infty} e^{-Cx^2}\, dx = \sqrt{\pi/C}, \tag{5.3}$$

which is valid for any $C > 0$. [In fact it's valid for any complex number C whose real part is positive.]

Example 5.2. *The Cauchy probability density function is defined by*

$$p(t) = \frac{1}{\pi(1+t^2)}. \tag{5.4}$$

This distribution might also be called the Mad Machine Gunner *distribution; imagine our killer sitting at the origin of the (x, y) plane. He*[1] *is firing (at a constant rate) at the infinite line $y = 1$, his angle θ (with the x-axis) of fire being uniformly distributed in the interval $(0, \pi)$. Then the bullets have the Cauchy density.*

---

[1]The sexism is quite accurate, since males produce vastly more violent psychopaths than females.

If you draw some graphs of these probability densities, you should find that, for small σ, the graph is concentrated around the value μ. For large σ, the graph is rather flat. There are two important definitions that capture this behaviour mathematically.

Definition 5.1. *The* mean, *or* expected value, *of a random variable X with p.d.f $p(t)$ is defined by*

$$\mathbb{E}X := \int_{-\infty}^{\infty} t p(t)\, dt. \tag{5.5}$$

It's very common to write μ instead $\mathbb{E}X$ when no ambiguity can arise. Its variance *$Var\ X$ is given by*

$$Var\ X := \int_{-\infty}^{\infty} (t-\mu)^2\, p(t)\, dt. \tag{5.6}$$

Exercise 5.1. *Show that the Gaussian p.d.f. really does have mean μ and variance σ^2.*

Exercise 5.2. *What happens when we try to determine the mean and variance of the Cauchy probability density defined in Example 5.4?*

Exercise 5.3. *Prove that $Var\ X = \mathbb{E}(X^2) - (\mathbb{E}X)^2$.*

We shall frequently have to calculate the expected value of *functions* of random variables.

Theorem 5.1. *If*

$$\int_{-\infty}^{\infty} |f(t)| p(t)\, dt$$

is finite, then

$$\mathbb{E}\left(f(X)\right) = \int_{-\infty}^{\infty} f(t)p(t)\, dt. \tag{5.7}$$

Example 5.3. *Let X denote a normalized Gaussian random variable. We shall show that*

$$\mathbb{E}e^{\lambda X} = e^{\lambda^2/2}, \tag{5.8}$$

Indeed, applying (5.7), we have

$$\mathbb{E}e^{\lambda X} = \int_{-\infty}^{\infty} e^{\lambda t}(2\pi)^{-1/2}e^{-t^2/2}\, dt = (2\pi)^{-1/2}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(t^2-2\lambda t)}\, dt.$$

The trick now is to complete the square *in the exponent, that is,*

$$t^2 - 2\lambda t = (t-\lambda)^2 - \lambda^2.$$

Thus

$$\mathbb{E}e^{\lambda X} = (2\pi)^{-1/2}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}([t-\lambda]^2 - \lambda^2)\right)\, dt = e^{\lambda^2/2}.$$

Exercise 5.4. *Let W be any Gaussian random variable with mean zero. Prove that*

$$\mathbb{E}\left(e^{W}\right) = e^{\frac{1}{2}\mathbb{E}\left(W^2\right)}. \tag{5.9}$$

5.2. **Differential Equations.** A *differential equation*, or *ordinary differential equation* (ODE), is simply a functional relationship specifying first, or higher derivatives, of a function; the *order* of the equation is just the degree of its highest derivatives. For example,

$$y'(t) = 4t^3 + y(t)^2$$

is a univariate first-order differential equation, whilst or

$$\mathbf{y}'(t) = A\mathbf{y}(t),$$

where $\mathbf{y}(t) \in \mathbb{R}^d$ and $A \in \mathbb{R}^{d\times d}$ is a first-order differential equation in d-variables. A tiny class of differential equations can be solved analytically, but numerical methods are required for the vast majority. The numerical analysis of differential equations has been one of the most active areas of research in computational mathematics since the 1960s and excellent free software exists. It is extremely unlikely that any individual can better this software without years of effort and scholarship, so you should use this software for any practical problem. You can find lots of information at `www.netlib.org` and `www.nr.org`. This section contains the minimum relevant theory required to make use of this software.

You should commit to memory one crucial first-order ODE:

Proposition 5.2. *The general solution to*

$$y'(t) = \lambda y(t), \qquad t \in \mathbb{R}, \tag{5.10}$$

where λ can be any complex number, is given by

$$y(t) = c\exp(\lambda t), \qquad t \in \mathbb{R}. \tag{5.11}$$

Here $c \in \mathbb{C}$ is a constant. Note that $c = y(0)$, so we can also write the equation as $y(t) = y(0)\exp(\lambda t)$.

Proof. If we multiply the equation $y' - \lambda y = 0$ by the *integrating factor* $\exp(-\lambda t)$, then we obtain

$$0 = \frac{d}{dt}\left(y(t)\exp(-\lambda t)\right),$$

that is

$$y(t)\exp(-\lambda t) = c,$$

for all $t \in \mathbb{R}$. □

In fact, there's a useful **slogan** for ODEs: try an exponential $\exp(\lambda t)$ or use reliable numerical software.

Example 5.4. *If we try $y(t) = \exp(\lambda t)$ as a trial solution in*

$$y'' + 2y' - 3y = 0,$$

then we obtain

$$0 = \exp(\lambda t)\left(\lambda^2 + 2\lambda - 3\right).$$

Since $\exp(\lambda t) \neq 0$*, for any t, we deduce the* associated equation

$$\lambda^2 + 2\lambda - 3 = 0.$$

The roots of this quadratic are 1 and -3, which is left as an easy exercise. Now this ODE is linear: any linear combination of solutions is still a solution. Thus we have a general family of solutions

$$\alpha\exp(t) + \beta\exp(-3t),$$

for any complex numbers α and β. We need two pieces of information to solve for these constants, such as $y(t_1)$ and $y(t_2)$, or, more usually, $y(t_1)$ and $y'(t_1)$. In fact this is the general solution of the equation.

In fact, we can always change an mth order equation is one variable into an equivalent first order equation in m variables, a technique that I shall call *vectorizing* (some books prefer the more pompous phrase "reduction of order"). Most ODE software packages are designed for first order systems, so vectorizing has both practical and theoretical importance.

For example, given

$$y''(t) = \sin(t) + (y'(t))^3 - 2(y(t))^2,$$

we introduce the *vector function*

$$\mathbf{z}(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix},$$

Then

$$\mathbf{z}'(t) = \begin{pmatrix} y' \\ y'' \end{pmatrix} = \begin{pmatrix} y' \\ \\ \sin(t) + (y')^3 - 2(y)^2 \end{pmatrix}.$$

In other words, writing

$$\mathbf{z}(t) = \begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} \equiv \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix},$$

we have derived

$$\mathbf{z}' = \begin{pmatrix} z_2 \\ \sin(t) + z_2^3 - 2z_1^2 \end{pmatrix},$$

which we can write as

$$\mathbf{z}' = \mathbf{f}(\mathbf{z}, t).$$

Exercise 5.5. *You probably won't need to consider ODEs of order exceeding two very often in finance, but the same trick works. Given*

$$y^{(n)}(t) = \sum_{k=0}^{n-1} a_k(t) y^{(k)}(t),$$

we define the vector function $\mathbf{z}(t) \in \mathbb{R}^{n-1}$ by

$$z_k(t) = y^{(k)}(t), \qquad k = 0, 1, \ldots, n-1.$$

Then $\mathbf{z}'(t) = M\mathbf{z}(t)$. Find the matrix M.

5.3. **Recurrence Relations.** In its most general form, a *recurrence relation* is simply a sequence of vectors $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \ldots$ for which some functional relation generates $\mathbf{v}^{(n)}$ given the earlier iterates $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(n-1)}$. At this level of generality, very little more can be said. However, the theory of linear recurrence relations is simple and very similar to the techniques of differential equations.

The first order linear recurrence relation is simply the sequence $\{a_n : n = 0, 1, \ldots\}$ of complex numbers defined by

$$a_n = c a_{n-1}.$$

Thus

$$a_n = c a_{n-1} = c^2 a_{n-2} = c^3 a_{n-3} = \cdots = c^n a_0$$

and the solution is complete.

The second order linear recurrence relation is slightly more demanding. Here

$$a_{n+1} + pa_n + qa_{n-1} = 0$$

and, inspired by the solution for the first order recurrence, we try $a_n = c^n$, for some $c \neq 0$. Then

$$0 = c^{n-1}\left(c^2 + pc + q\right),$$

or

$$0 = c^2 + pc + q.$$

If this has two distinct roots c_1 and c_2, then one possible solution to the second order recurrence is

$$u_n = p_1 c_1^n + p_2 c_2^n,$$

for constants p_1 and p_2. However, is this the full set of solutions? What happens if the quadratic has only one root?

Proposition 5.3. *Let $\{a_n : n \in \mathbb{Z}\}$ be the sequence of complex numbers satisfying the recurrence relation*

$$a_{n+1} + pa_n + qa_{n-1} = 0, \qquad n \in \mathbb{Z}.$$

If α_1 and α_2 are the roots of the associated quadratic

$$t^2 + pt + q = 0,$$

then the general solution is

$$a_n = c_1\alpha_1^n + c_2\alpha_2^n$$

when $\alpha_1 \neq \alpha_2$. If $\alpha_1 = \alpha_2$, then the general solution is

$$a_n = (v_1 n + v_2)\,\alpha_1^n.$$

Proof. The same vectorizing trick used to change second order differential equations in one variable into first order differential equations in two variables can also be used here. We define a new sequence $\{\mathbf{b}^{(n)} : n \in \mathbb{Z}\}$ by

$$\mathbf{b}^{(n)} = \begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix}.$$

Thus

$$\mathbf{b}^{(n)} = \begin{pmatrix} a_{n-1} \\ -pa_{n-1} - qa_{n-2} \end{pmatrix},$$

that is,

$$\mathbf{b}^{(n)} = A\mathbf{b}^{(n-1)}, \tag{5.12}$$

where

$$A = \begin{pmatrix} 0 & 1 \\ -q & -p \end{pmatrix}. \tag{5.13}$$

This first order recurrence has the simple solution

$$\mathbf{b}^{(n)} = A^n\mathbf{b}^{(0)}, \tag{5.14}$$

so our analytic solution reduces to calculation of the matrix power A^n. Now let us begin with the case when the eigenvalues λ_1 and λ_2 are distinct. Then the corresponding eigenvectors $\mathbf{w}^{(1)}$ and $\mathbf{w}^{(2)}$ are linearly independent. Hence we can write our initial vector $\mathbf{b}^{(0)}$ as a unique linear combination of these eigenvectors:

$$\mathbf{b}^{(0)} = b_1\mathbf{w}^{(1)} + b_2\mathbf{w}^{(2)}.$$

Thus

$$\mathbf{b}^{(n)} = b_1 A^n \mathbf{w}^{(1)} + b_2 A^n \mathbf{w}^{(2)} = b_1 \lambda_1^n \mathbf{w}^{(1)} + b_1 \lambda_2^n \mathbf{w}^{(2)}.$$

Looking at the second component of the vector, we obtain

$$a_n = c_1 \lambda_1^n + c_2 \lambda_2^n.$$

Now the eigenvalues of A are the roots of the quadratic equation

$$\det(A - \lambda I) = \det \begin{pmatrix} -\lambda & 1 \\ -q & -p-\lambda \end{pmatrix},$$

in other words the roots of the quadratic

$$\lambda^2 + p\lambda + q = 0.$$

Thus the associated equation is precisely the characteristic equation of the matrix A in the vectorized problem. Hence $a_n = c_1 \alpha_1^n + c_2 \alpha_2^n$.

We only need this case in the course, but I shall lead you through a careful analysis of the case of coincident roots. It's a good exercise for your matrix skills.

First note that the roots are coincident if and only if $p^2 = 4q$, in which case

$$A = \begin{pmatrix} 0 & 1 \\ -p^2/4 & -p \end{pmatrix},$$

and the eigenvalue is $-p/2$. In fact, subsequent algebra is simplified if we substitute $\alpha = -p/2$, obtaining

$$A = \begin{pmatrix} 0 & 1 \\ -\alpha^2 & 2\alpha \end{pmatrix}.$$

The remainder of the proof is left as the following exercise. □

Exercise 5.6. *Show that*

$$A = \alpha I + \mathbf{u}\mathbf{v}^T,$$

where

$$\mathbf{u} = \begin{pmatrix} 1 \\ \alpha \end{pmatrix}, \qquad \mathbf{v} = \begin{pmatrix} -\alpha \\ 1 \end{pmatrix}$$

and note that $\mathbf{v}^T\mathbf{u} = 0$. *Show also that*

$$A^2 = \alpha^2 I + 2\alpha \mathbf{u}\mathbf{v}^T, \qquad A^3 = \alpha^3 I + 3\alpha^2 \mathbf{u}\mathbf{v}^T,$$

and use proof by induction to demonstrate that

$$A^n = \alpha^n I + n\alpha^{n-1}\mathbf{u}\mathbf{v}^T.$$

Hence find the general solution for a_n.

5.4. **Mortgages – a once exotic instrument.** The objective of this section is to illustrate some of the above techniques for analysing difference and differential equations via mortgage pricing. You are presumably all too familiar with a repayment mortgage: we borrow a large sum M for a fairly large slice T of our lifespan, repaying capital and interest using N regular payments. The interest rate is assumed to be constant and it's a secured loan: our homes are forfeit on default. How do we calculate our repayments?

Let $h = T/N$ be the interval between payments, let $D_h : [0, T] \to \mathbb{R}$ be our debt as a function of time, and let $A(h)$ be our payment. We shall assume that our initial debt is $D_h(0) = 1$, because we can always multiply by the true initial cost M of our house after the calculation. Thus D must satisfy the equations

$$D_h(0) = 1, \quad D_h(T) = 0 \quad \text{and} \quad D_h(\ell h) = D_h((\ell-1)h)e^{rh} - A(h). \tag{5.15}$$

We see that $D_h(h) = e^{rh} - A(h)$, while

$$D_h(2h) = D_h(h)e^{rh} - A(h) = e^{2rh} - A(h)\left(1 + e^{rh}\right).$$

The pattern is now fairly obvious:

$$D_h(\ell h) = e^{\ell rh} - A(h)\sum_{k=0}^{\ell-1} e^{krh}, \tag{5.16}$$

and summing the geometric series[2]

$$D_h(\ell h) = e^{\ell rh} - A(h)\left(\frac{e^{\ell rh} - 1}{e^{rh} - 1}\right). \tag{5.17}$$

In order to achieve $D(T) = 0$, we choose

$$A(h) = \frac{e^{rh} - 1}{1 - e^{-rT}}. \tag{5.18}$$

Exercise 5.7. *What happens if $T \to \infty$?*

Exercise 5.8. *Prove that*

$$D_h(\ell h) = \frac{1 - e^{-r(T-\ell h)}}{1 - e^{-rT}}. \tag{5.19}$$

Thus, if $t = \ell h$ is constant (so we increase ℓ as we reduce h), then

$$D_h(t) = \frac{1 - e^{-r(T-t)}}{1 - e^{-rT}}. \tag{5.20}$$

Almost all mortgages are repaid by 300 monthly payments for 25 years. However, until recently, many mortgages calculated interest *yearly*, which means that we choose $h = 1$ in Exercise 5.7 and then divide $A(1)$ by 12 to obtain the monthly payment.

Exercise 5.9. *Calculate the monthly repayment $A(1)$ when $M = 10^5$, $T = 25$, $r = 0.05$ and $h = 1$. Now repeat the calculation using $h = 1/12$. Interpret your result.*

---

[2]Many students forget the simple formula. If $S = 1 + a + a^2 + \cdots + a^{m-2} + a^{m-1}$, then $aS = a + a^2 + \cdots + a^{m-1} + a^m$. Subtracting these expressions implies $(a-1)S = a^m - 1$, all other terms cancelling.

In principle, there's no reason why our repayment could not be continuous, with interest being recalculated on our constantly decreasing debt. For continuous repayment, our debt $D : [0,T] \to \mathbb{R}$ satisfies the relations

$$D(0) = 1, \quad D(T) = 0 \quad \text{and} \quad D(t+h) = D(t)e^{rh} - hA. \tag{5.21}$$

Exercise 5.10. *Prove that*

$$D'(t) - rD(t) = -A, \tag{5.22}$$

where, in particular, you should prove that (5.21) implies the differentiability of $D(t)$. Solve this differential equation using the integrating factor e^{-rt}. You should find the solution

$$D(t)e^{-rt} - 1 = A \int_0^t \left(-e^{-r\tau}\right) d\tau = A\left(\frac{e^{-rt} - 1}{r}\right). \tag{5.23}$$

Hence show that

$$A = \frac{r}{1 - e^{-rT}} \tag{5.24}$$

and

$$D(t) = \frac{1 - e^{-r(T-t)}}{1 - e^{-rT}}, \tag{5.25}$$

agreeing with (5.20)*, i.e. $D_h(kh) = D(kh)$, for all k. Prove that* $\lim_{r\to\infty} D(t) = 1$ *for $0 < t < T$ and interpret.*

Observe that

$$\frac{A(h)}{Ah} = \frac{e^{rh} - 1}{rh} \approx 1 + (rh/2), \tag{5.26}$$

so that continuous repayment is optimal for the borrower, but that the mortgage provider is making a substantial profit. Greater competition has made yearly recalculations much rarer, and interest is often paid daily, i.e. $h = 1/250$, which is rather close to continuous repayment.

Exercise 5.11. *Construct graphs of $D(t)$ for various values of r. Calculate the time $t_0(r)$ at which half of the debt has been paid.*

5.5. **Pricing Mortgages via lack of arbitrage.** There is a very slick arbitrage argument to deduce the continuous repayment mortgage debt formula (5.25). Specifically, the simple fact that $D(t)$ is a deterministic financial instrument implies, via arbitrage, that $D(t) = a + b\exp(rt)$, so we need only choose the constants a and b to satisfy $D(0) = 1$ and $D(T) = 1$, which imply $a + b = 1$ and $a + b\exp(rT) = 0$. Solving these provides $a = \exp(rT)/(\exp(rT) - 1)$ and $b = -1/(\exp(rT) - 1)$, and equivalence to (5.25) is easily checked.

6. Numerical Linear Algebra

I shall not include much explicitly here, because you have my longer lecture notes on numerical linear algebra:

http://econ109.econ.bbk.ac.uk/brad/Methods/nabook.pdf

Please do revise the first long chapter of those notes if need to brush up on matrix algebra.

You will also find my Matlab notes useful:

http://econ109.econ.bbk.ac.uk/brad/Methods/matlab_intro_notes.pdf

6.1. Orthogonal Matrices.

Modern numerical linear algebra began with the computer during the Second World War, its progress accelerating enormously as computers became faster and more convenient in the 1960s. One of the most vital conclusions of this research field is the enormous practical importance of matrices which leave Euclidean length invariant. More formally:

Definition 6.1. *We shall say that $Q \in \mathbb{R}^{n\times n}$ is* distance-preserving *if $\|Q\mathbf{x}\| = \|\mathbf{x}\|$, for all $\mathbf{x} \in \mathbb{R}^n$.*

The following simple result is very useful.

Lemma 6.1. *Let $M \in \mathbb{R}^{n\times n}$ be any symmetric matrix for which $\mathbf{x}^T M\mathbf{x} = 0$, for every $\mathbf{x} \in \mathbb{R}^n$. Then M is the zero matrix.*

Proof. Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \in \mathbb{R}^n$ be the usual coordinate vectors. Then

$$M_{jk} = \mathbf{e}_j^T M\mathbf{e}_k = \frac{1}{2}\left(\mathbf{e}_j + \mathbf{e}_k\right)^T M\left(\mathbf{e}_j + \mathbf{e}_k\right) = 0, \qquad 1 \le j, k \le n.$$

□

Theorem 6.2. *The matrix $Q \in \mathbb{R}^n$ is distance-preserving if and only if $Q^T Q = I$.*

Proof. If $Q^T Q = I$, then

$$\|Q\mathbf{x}\|^2 = (Q\mathbf{x})^T(Q\mathbf{x}) = \mathbf{x}^T Q^T Q\mathbf{x} = \mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2,$$

and Q is distance-preserving. Conversely, if $\|Q\mathbf{x}\|^2 = \|\mathbf{x}\|^2$, for all $\mathbf{x} \in \mathbb{R}^n$, then

$$\mathbf{x}^T\left(Q^T Q - I\right)\mathbf{x} = 0, \qquad \mathbf{x} \in \mathbb{R}^n.$$

Since $Q^T Q - I$ is a symmetric matrix, Lemma 6.1 implies $Q^T Q - I = 0$, i.e. $Q^T Q = I$. □

The condition $Q^T Q = I$ simply states that the columns of Q are orthonormal vectors, that is, if the columns of Q are $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$, then $\|\mathbf{q}_1\| = \cdots = \|\mathbf{q}_n\| = 1$ and $\mathbf{q}_j^T\mathbf{q}_k = 0$ when $j \ne k$. For this reason, Q is also called an *orthogonal matrix*. We shall let $O(n)$ denote the set of all (real) $n \times n$ orthogonal matrices.

Exercise 6.1. *Let $Q \in O(n)$. Prove that $Q^{-1} = Q^T$. Further, prove that $O(n)$ is closed under matrix multiplication, that is, $Q_1Q_2 \in O(n)$ when $Q_1, Q_2 \in O(n)$. (In other words, $O(n)$ forms a group under matrix multiplication. This observation is important, and $O(n)$ is often called the* Orthogonal Group.*)*

School of Computing and Mathematical Sciences, Birkbeck College, University of London, Malet Street, London WC1E 7HX, England

Email address: b.baxter@bbk.ac.uk