CFRM 530: Fixed Income

Matthew Lorig [1]

This version: February 24, 2021

[1]Department of Applied Mathematics, University of Washington, Seattle, WA, USA. e-mail: mlorig@uw.edu

Contents

Contents

Preface

These notes are intended to give undergraduate- and masters-level students in computatinoal finance an introduction to fixed income markets. Because the focus of this course is on the *applied* aspects of this topic, many of the theorems will presented without proof. The aim is to provide the minimal level of rigor needed to complete fixed-income computations. The hope is that, what the notes lack in rigor, they make up in clarity.

These notes are a work in progress. Students are encouraged to e-mail the professor if (when) they find errors or typos.

Donations

If you find these notes useful and would like to make a donation to help me develop them further you can donate Bitcoin, Ethereum or Ethereum-based ERC20 tokens to the addresses below.

BTC address: 3QVEVC6QGJ8EmKsnpfXadkshwkDKzCrR4w

ETH address: 0x089d4d033F3E55B80Ac40D5da46812EFeFC8824b

Chapter 1

Basic Fixed Income Instruments

In this chapter, we will introduce several fixed income instruments.

1.1 Short-rate and money market account

Throughout this text, we will denote by $\mathrm{M} = (\mathrm{M}_t)_{t \geq 0}$ the value of a *money market* or *savings* account. The dynamics of M will always be given by

$$\mathrm{dM}_t = \mathrm{R}_t \mathrm{M}_t \mathrm{d}t \qquad \Rightarrow \qquad \mathrm{M}_t = \mathrm{M}_0 \exp\left(\int_0^t \mathrm{R}_s \mathrm{d}s \right),$$

where the process $\mathrm{R} = (\mathrm{R}_t)_{t \geq 0}$ is known as the *short rate*, *spot rate* or *instantaneous interest rate*. In general, the short rate R will be a stochastic (i.e., random) process. as such, the rate at which the money market grows at time t_1 will not be the same as the rate at which the money market account grows at time t_2. We will assume throughout this text that the spot rate is non-negative

$$\mathrm{R}_t \geq 0, \qquad \forall\, t \geq 0.$$

As a result, the money market account M will always be a non-decreasing process

$$\mathrm{M}_{t_2} \geq \mathrm{M}_{t_1}, \qquad \forall\, 0 \leq t_1 < t_2 < \infty.$$

1.2 Zero-coupon bonds and yields

A T-*maturity zero-coupon bond*, denoted $\mathrm{B}^{\mathrm{T}} = (\mathrm{B}^{\mathrm{T}}_t)_{0 \leq t \leq \mathrm{T}}$, is a financial instrument that pays 1 unit of currency at time T. Clearly, we must have

$$\mathrm{B}^{\mathrm{T}}_{\mathrm{T}} = 1, \qquad \forall\, \mathrm{T} \geq 0. \tag{1.1}$$

If we fix T, the map $t \mapsto B_t^T$ represents the evolution of the T-maturity bond price. In general, bond prices will evolve as stochastic processes. If we fix t, the map $T \mapsto B_t^T$ gives us the value of bonds with different maturity dates.

In financial markets, bonds only trade with certain maturity dates $T_1, T_2, \ldots, T_n$. However, it will be useful for us to imagine that bonds trade at every maturity date $T \geq t$.

The map $T \mapsto B_t^T$ will always be non-increaseing.

$$B_t^{T_1} \geq B_t^{T_2}, \qquad \forall\, t \leq T_1 < T_2 < \infty. \tag{1.2}$$

To see why this is the case, suppose that $T_2 > T_1$ and assume by contradiction that $B_t^{T_2} > B_t^{T_1}$. Then at time t an investor could do the following:

- buy a T_1-maturity bond for $B_t^{T_1}$,
- sell a T_2-maturity bond for $B_t^{T_2}$, and
- put $(B_t^{T_2} - B_t^{T_1})$ in the money market account.

The total initial cost of this strategy is zero. At time T_1 the investor could

- put the payment from the T_1 maturity bond $B_{T_1}^{T_1} = 1$ in the money market account.

The value of this trading strategy at time T_2 would be

$$(B_t^{T_2} - B_t^{T_1})\frac{M_{T_2}}{M_t} + \frac{M_{T_2}}{M_{T_1}} - B_{T_2}^{T_2} \geq (B_t^{T_2} - B_t^{T_1})\frac{M_{T_2}}{M_t} \geq 0,$$

where we have used the fact that M is non-decreasing and $B_{T_2}^{T_2} = 1$. With zero initial investment, the above strategy has generated a guaranteed profit. This is what is known as an *arbitrage* (we will give a precise definition for *arbitrage* later in this text). We generally accept that financial markets to not allow for arbitrage opportunities. As such, it follows that we must have $B_t^{T_2} \leq B_t^{T_1}$.

Note that, from (1.1) and (1.2), zero-coupon bond prices will always be worth less than one unit of currency

$$B_t^T \leq 1.$$

The *yield* of a T-maturity bond, denoted $Y^T = (Y_t^T)_{0 \leq t \leq T}$, is defined via the relation

$$B_t^T \exp\Big((T-t)Y_t^T\Big) = 1 \qquad \Rightarrow \qquad Y_t^T = \frac{-\log B_t^T}{T-t}.$$

Stated differently, if an investor were to buy a T-maturity bond at time t and hold it to maturity T, this would be equivalent to investing B_t^T units of currency in a savings account that pays a continuously compounded constant rate of interest Y_t^T.

We call the map $\mathrm{T} \mapsto \mathrm{Y}_t^\mathrm{T}$ the *yield curve*. Note that, as B_t^T is observable at time t for all $\mathrm{T} \geq t$ the yield curve is also observable at time t. As time t moves forward bond prices $(\mathrm{B}_t^\mathrm{T})_{\mathrm{T}\geq t}$ evolve stochastically and, as such, the entire yield curve evolves stochastically in time as well. Modeling the random movements of the yield curve is one of the main challenges of fixed income markets.

1.3 Forward rates

Fix two maturity dates $\mathrm{T}_1 \leq \mathrm{T}_2$ and consider the following investment strategy. At time $t \leq \mathrm{T}_1$ an investor

- sells N zero-coupon bonds with maturity T_1
- buys $\mathrm{NB}_t^{\mathrm{T}_1}/\mathrm{B}_t^{\mathrm{T}_2}$ zero-coupon bonds with mortuary T_2.

The total initial investment of this strategy is zero because

$$\mathrm{NB}_t^{\mathrm{T}_1} - \frac{\mathrm{NB}_t^{\mathrm{T}_1}}{\mathrm{B}_t^{\mathrm{T}_2}}\mathrm{B}_t^{\mathrm{T}_2} = 0.$$

At time T_1 the investor

- pays N units of currency for the T_1-maturity bonds he sold at time t.

At time T_2 the investor

- receives $\mathrm{NB}_t^{\mathrm{T}_1}/\mathrm{B}_t^{\mathrm{T}_2}$ for the T_2-maturity bonds he bought at time t.

Thus, by executing the above strategy at time t, the investor has guaranteed that an investment of N units of currency at time T_1 will grow to $\mathrm{NB}_t^{\mathrm{T}_1}/\mathrm{B}_t^{\mathrm{T}_2}$ at time T_2

$$\mathrm{N} \to \mathrm{N}\frac{\mathrm{B}_t^{\mathrm{T}_1}}{\mathrm{B}_t^{\mathrm{T}_2}}.$$

This motivates a the following definitions.

The *simple forward rate from* T_1 *to* T_2, denoted $\mathrm{F}^{\mathrm{T}_1,\mathrm{T}_2} = (\mathrm{F}_t^{\mathrm{T}_1,\mathrm{T}_2})_{0\leq t\leq \mathrm{T}_1}$ is defined through the relation

$$1 + (\mathrm{T}_2 - \mathrm{T}_1)\mathrm{F}_t^{\mathrm{T}_1,\mathrm{T}_2} = \frac{\mathrm{B}_t^{\mathrm{T}_1}}{\mathrm{B}_t^{\mathrm{T}_2}} \qquad \Rightarrow \qquad \mathrm{F}_t^{\mathrm{T}_1,\mathrm{T}_2} := \frac{1}{\mathrm{T}_2 - \mathrm{T}_1}\left(\frac{\mathrm{B}_t^{\mathrm{T}_1}}{\mathrm{B}_t^{\mathrm{T}_2}} - 1\right). \tag{1.3}$$

The *continuously compounded forward rate from* T_1 *to* T_2, denoted $f^{\mathrm{T}_1,\mathrm{T}_2} = (f_t^{\mathrm{T}_1,\mathrm{T}_2})_{0\leq t\leq \mathrm{T}_1}$ is defined through the relation

$$\exp\Big((\mathrm{T}_2-\mathrm{T}_1)f_t^{\mathrm{T}_1,\mathrm{T}_2}\Big) = \frac{\mathrm{B}_t^{\mathrm{T}_1}}{\mathrm{B}_t^{\mathrm{T}_2}} \qquad \Rightarrow \qquad f_t^{\mathrm{T}_1,\mathrm{T}_2} := \frac{1}{\mathrm{T}_2-\mathrm{T}_1}\log\left(\frac{\mathrm{B}_t^{\mathrm{T}_1}}{\mathrm{B}_t^{\mathrm{T}_2}}\right) = -\Big(\frac{\log \mathrm{B}_t^{\mathrm{T}_2} - \log \mathrm{B}_t^{\mathrm{T}_1}}{\mathrm{T}_2-\mathrm{T}_1}\Big).$$

Lastly, the *instantaneous forward rate with maturity* T, denoted $f^{\mathrm{T}} = (f_t^{\mathrm{T}})_{0\leq t\leq \mathrm{T}}$ is defined as

$$f_t^{\mathrm{T}} := \lim_{\mathrm{T}_2\to\mathrm{T}} f_t^{\mathrm{T},\mathrm{T}_2} = -\partial_{\mathrm{T}} \log \mathrm{B}_t^{\mathrm{T}}. \tag{1.4}$$

In words, f_t^{T} is the instantaneous rate of interest that one can lock in at time t for an investment made over the period T to $\mathrm{T}+\mathrm{d}t$. Noting that R_t is the instantaneous rate of interest one can lock in at time t by investing in the money market account M, to avoid arbitrage we must have

$$\mathrm{R}_t = f_t^t.$$

We call the map $\mathrm{T} \mapsto f_t^{\mathrm{T}}$ the *forward rate curve*. Note that, as $\mathrm{B}_t^{\mathrm{T}}$ is observable at time t for all $\mathrm{T} \geq t$ the forward rate curve is also observable at time t. As time t moves forward bond prices $(\mathrm{B}_t^{\mathrm{T}})_{\mathrm{T}\geq t}$ evolve stochastically and, as such, the entire forward rate curve evolves stochastically in time as well.

Observe that, if we integrate the equation (1.4) with respect the maturity date we have

$$\int_t^{\mathrm{T}} f_t^s \mathrm{d}s = -\int_t^{\mathrm{T}} \partial_s \log \mathrm{B}_t^s = -\Big(\log \mathrm{B}_t^{\mathrm{T}} - \log \mathrm{B}_t^t\Big).$$

Solving for $\mathrm{B}_t^{\mathrm{T}}$ and using the fact that $\mathrm{B}_t^t = 1$ we obtain

$$\mathrm{B}_t^{\mathrm{T}} = \exp\left(-\int_t^{\mathrm{T}} f_t^s \mathrm{d}s\right). \tag{1.5}$$

Thus, there is a one-to-one correspondence between bond prices $(\mathrm{B}_t^{\mathrm{T}})_{\mathrm{T}\geq t}$ and the forward rate curve $(f_t^{\mathrm{T}})_{\mathrm{T}\geq t}$. We can obtain $(f_t^{\mathrm{T}})_{\mathrm{T}\geq t}$ from $(\mathrm{B}_t^{\mathrm{T}})_{\mathrm{T}\geq t}$ using (1.4) and we can obtain $(\mathrm{B}_t^{\mathrm{T}})_{\mathrm{T}\geq t}$ from $(f_t^{\mathrm{T}})_{\mathrm{T}\geq t}$ using (1.5).

1.4 Coupon-bearing bonds

A typical *coupon-bearing bond* specifies a series of dates $\mathcal{T} = (\mathrm{T}_1, \mathrm{T}_2, \ldots, \mathrm{T}_n)$ at which deterministic (i.e., non-random) payments $c = (c_1, c_2, \ldots c_n)$ are made. Additionally, on the last date T_n and additional payment of one unit of currency is made. At time t, the payments of a coupon-bearing bond that have yet to be paid can be replicated by a static portfolio of zero-coupon bonds as follows

- for every i such that $\mathrm{T}_i > t$, purchase c_i zero-coupon bonds with maturity T_i.
- buy one additional T_n-maturity zero-coupon bond.

To avoid arbitrage, the price of the coupon-bearing bond, denoted $\mathrm{CB}(\mathcal{T}, c) = (\mathrm{CB}_t(\mathcal{T}, c))_{t \geq 0}$ must be equal to the value of the replicating portfolio

$$\mathrm{CB}_t(\mathcal{T}, c) = \sum_{\mathrm{T}_i > t}^{n} c_i \mathrm{B}_t^{\mathrm{T}_i} + \mathrm{B}_t^{\mathrm{T}_n}. \tag{1.6}$$

Thus, coupon-bearing bond prices can be determined from zero-coupon bond prices.

Alternatively, suppose we observe coupon-bearing bond prices $\mathrm{CB}_t(\mathcal{T}^1, c^1), \mathrm{CB}_t(\mathcal{T}^2, c^2), \ldots, \mathrm{CB}_t(\mathcal{T}^n, c^n)$. Then we have from (1.6) that

$$\underbrace{\begin{pmatrix} \mathrm{CB}_t(\mathcal{T}^1, c^1) \\ \mathrm{CB}_t(\mathcal{T}^2, c^2) \\ \vdots \\ \mathrm{CB}_t(\mathcal{T}^n, c^n) \end{pmatrix}}_{\mathbf{CB}_t} = \underbrace{\begin{pmatrix} c_1^1 + 1 & 0 & \ldots & 0 \\ c_1^2 & c_2^2 + 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ c_1^n & c_2^n & \ldots & c_n^n + 1 \end{pmatrix}}_{\mathbf{c}} \underbrace{\begin{pmatrix} \mathrm{B}_t^{\mathrm{T}_1} \\ \mathrm{B}_t^{\mathrm{T}_2} \\ \vdots \\ \mathrm{B}_t^{\mathrm{T}_n} \end{pmatrix}}_{\mathbf{B}_t},$$

where we have assumed that $\mathcal{T}^i = (\mathrm{T}_1, \mathrm{T}_2, \ldots, \mathrm{T}_i)$. Denoting the above matrix equation as $\mathbf{CB}_t = \mathbf{c}\mathbf{B}_t$ we have $\mathbf{c}^{-1}\mathbf{CB}_t = \mathbf{B}_t$.

1.5 Floating rate notes

A typical *floating rate note* specifies a series of dates $\mathcal{T} = (\mathrm{T}_0, \mathrm{T}_1, \mathrm{T}_2, \ldots, \mathrm{T}_n)$. At all dates $\mathrm{T}_i > \mathrm{T}_0$ the holder of the note receives a payment of

$$\text{Payment at time } \mathrm{T}_i := (\mathrm{T}_i - \mathrm{T}_{i-1}) \mathrm{F}_{\mathrm{T}_{i-1}}^{\mathrm{T}_{i-1}, \mathrm{T}_i} = \left(\frac{1}{\mathrm{B}_{\mathrm{T}_{i-1}}^{\mathrm{T}_i}} - 1 \right),$$

where we have used (1.3) and $\mathrm{B}_{\mathrm{T}_{i-1}}^{\mathrm{T}_{i-1}} = 1$. Additionally, at the final date T_n an additional payment of one unit of currency is made. Observe that the payment received at time T_i is random but known at time T_{i-1}. We can replicate the payments of a floating rate note as follows. Assume for simplicity that $t \leq \mathrm{T}_0$. At time t

- buy a T_0-maturity zero-coupon bond for total cost of $\mathrm{B}_t^{\mathrm{T}_0}$.

At time T_0

- receive one unit of currency from the T_0-maturity zero-coupon bond,

- use the payment buy a $1/B_{T_0}^{T_1}$ T_1-maturity bonds for exactly one unit of currency.

At time T_1

- receive $1/B_{T_0}^{T_1}$ units of currency from the T_1 maturity bonds,
- pay out the coupon payment of $(1/B_{T_0}^{T_1} - 1)$,
- use the remaining single unit of currency to purchase $1/B_{T_1}^{T_2}$ T_2-maturity bonds.

Repeat the procedure until at time $T_2, T_3, \ldots T_{n-1}$. At time T_n

- receive $1/B_{T_{n-1}}^{T_n}$ units of currency from the T_n maturity bonds,
- make a final coupon payment of $(1/B_{T_{n-1}}^{T_n} - 1)$,
- make a final payment of one unit of currency.

As the above strategy replicates the coupon payments, the value of the floating rate note must equal the initical cost of the investment strategy $B_t^{T_0}$. Thus

$$\text{Time } t \text{ value of floating rate note} = B_t^{T_0}. \tag{1.7}$$

1.6 Interest rate swaps

A typical *swap* is an agreement between two parties. The *long* side agrees to pay a fixed amount K at a future date (or dates) in exchange for a random quantity. The *short* side agrees to pay the random quantity in exchange for the fixed amount K. The constant K is determined at inception so that the initial value of the contract is zero.

A typical *interest rate swap* specifies a series of dates $\mathcal{T} = (T_0, T_1, T_2, \ldots, T_n)$. At all dates $T_i > T_0$ the long side of the swap receives

$$\text{Payment at time } T_i := \underbrace{(T_i - T_{i-1})F_{T_{i-1}}^{T_{i-1},T_i}}_{\text{floating leg}} - \underbrace{(T_i - T_{i-1})K}_{\text{fixed leg}}. \tag{1.8}$$

From equation (1.7), we have

$$\text{time } t \text{ value of floating leg payments} = B_t^{T_0} - B_t^{T_n}$$

where we have subtracted $B_t^{T_n}$ because, unlike the floating rate note, the floating leg of the interest rate swap does not make a final payment of one unit of currency at the maturity date T_n.

Next, noting that the payments of the fixed leg can be replicated by a static portfolio of bonds with different maturities, we have

$$\text{time } t \text{ value of fixed leg payments} = \sum_{i=1}^{n} (\mathrm{T}_i - \mathrm{T}_{i-1}) \mathrm{K} \mathrm{B}_t^{\mathrm{T}_i}$$

Thus, the total value of the long side of the interest rate swap at time t is

$$\text{time } t \text{ value of interest rate swap} = \mathrm{B}_t^{\mathrm{T}_0} - \mathrm{B}_t^{\mathrm{T}_n} - \mathrm{K} \sum_{i=1}^{n} (\mathrm{T}_i - \mathrm{T}_{i-1}) \mathrm{B}_t^{\mathrm{T}_i},$$

where we have pulled the constant K out of the sum. The *swap rate*, denoted $\mathrm{K}_t^{\text{swap}}$, is the value of K that makes the time t value of the swap equal to zero. Thus, setting the right-hand side above equal to zero and solving for K we obtain

$$\mathrm{K}_t^{\text{swap}} = \frac{\mathrm{B}_t^{\mathrm{T}_0} - \mathrm{B}_t^{\mathrm{T}_n}}{\sum_{i=1}^{n} (\mathrm{T}_i - \mathrm{T}_{i-1}) \mathrm{B}_t^{\mathrm{T}_i}}. \tag{1.9}$$

Observe that as time t moves forward and bond prices change, so with the swap rate $\mathrm{K}_t^{\text{swap}}$.

1.7 Forward Contracts and T-forward prices

Let $\mathrm{A} = (\mathrm{A}_t)_{t \geq 0}$ be the value some traded asset (e.g., stock, bond, derivative, etc.). A *forward contract* written on A is an agreement between two parties – *long* and *short*. The long side agrees to receive $\mathrm{A}_\mathrm{T} - \mathrm{K}$ at time T and the short side agrees to pay $\mathrm{A}_\mathrm{T} - \mathrm{K}$. The date T is called the *expiration* or *maturity date* and the constant K is called the *delivery price*. The T-forward price of A, denoted $\mathrm{A}^\mathrm{T} = (\mathrm{A}_t^\mathrm{T})_{0 \leq t \leq \mathrm{T}}$ is the value of K at time t that makes a forward contract have zero value.

We can determine the value of K through a replication argument. For simplicity, assume that A pays no dividends or coupon payments. Suppose that at time t and investor

- sells $\mathrm{A}_t / \mathrm{B}_t^\mathrm{T}$ zero-coupon bonds maturity at time T,
- buys the asset A for A_t.

The total cost of this strategy is zero. At time T the investor has a contract that is worth

$$\mathrm{A}_\mathrm{T} - \frac{\mathrm{A}_t}{\mathrm{B}_t^\mathrm{T}} \mathrm{B}_\mathrm{T}^\mathrm{T} = \mathrm{A}_\mathrm{T} - \frac{\mathrm{A}_t}{\mathrm{B}_t^\mathrm{T}}.$$

Thus, the investor has replicated the payoff of the long side of a forward contract with delivery price $\mathrm{K} = \mathrm{A}_t / \mathrm{B}_t^\mathrm{T}$. As the investor's strategy had zero initial cost, we conclude that the T-forward price of A at time t is

$$\mathrm{A}_t^\mathrm{T} = \frac{\mathrm{A}_t}{\mathrm{B}_t^\mathrm{T}}. \tag{1.10}$$

Note: when we refer to the T_1-forward price of a zero-coupon bond with maturity $\mathrm{T}_2 > \mathrm{T}_1$ we will use the notation $\mathrm{B}_t^{\mathrm{T}_1,\mathrm{T}_2}$

$$\mathrm{B}_t^{\mathrm{T}_1,\mathrm{T}_2} = \frac{\mathrm{B}_t^{\mathrm{T}_2}}{\mathrm{B}_t^{\mathrm{T}_1}}.$$

1.8 Caps and Caplets

A *caplet* with reset date T_1 and settlement date T_2 pays the holder the positive part of the difference between $\mathrm{F}_{\mathrm{T}_1}^{\mathrm{T}_1,\mathrm{T}_2}$ and the strike rate κ at time T_2. That is

$$\text{Caplet payoff at time } \mathrm{T}_2 = (\mathrm{F}_{\mathrm{T}_1}^{\mathrm{T}_1,\mathrm{T}_2} - \kappa)^+, \qquad\qquad x^+ := \max\{x, 0\}.$$

Thus, a caplet is essentially a call written on $\mathrm{F}_{\mathrm{T}_1}^{\mathrm{T}_1,\mathrm{T}_2}$. A cap is simply a strip of caplets with reset dates $(\mathrm{T}_0, \mathrm{T}_1, \ldots \mathrm{T}_{n-1}$ and settlement dates $(\mathrm{T}_1, \mathrm{T}_2, \ldots, \mathrm{T}_n)$. Specifically

$$\text{Cap payoff at time } \mathrm{T}_i = (\mathrm{F}_{\mathrm{T}_{i-1}}^{\mathrm{T}_{i-1},\mathrm{T}_i} - \kappa)^+, \qquad\qquad i = 1, 2, \ldots, n. \tag{1.11}$$

Clearly, the value of a cap is equal to the sum of the values of the individual caplets. Assuming $t < \mathrm{T}_0$ we have

$$\mathrm{V}_t^{\text{cap}} = \sum_{i=1}^{n} \mathrm{V}_t^{\text{caplet},i}.$$

We will discuss how to find the value of a caplet later in this course.

Caps and caplets provide to their holder protection against rising interest rates. Specifically, they guarantees that the interest to be paid on a floating rate loan never exceeds the predetermined cap rate κ. For example, suppose an investor must make a payment at time T_2 of $\mathrm{N}\mathrm{F}_{\mathrm{T}_1}^{\mathrm{T}_1,\mathrm{T}_2}$. If the investor purchases a caplet, then his cash flow at time T_2 will be

$$\mathrm{N}(\mathrm{F}_{\mathrm{T}_1}^{\mathrm{T}_1,\mathrm{T}_2} - \kappa)^+ - \mathrm{N}\mathrm{F}_{\mathrm{T}_1}^{\mathrm{T}_1,\mathrm{T}_2} = -\mathrm{N}\min\{\kappa, \mathrm{F}_{\mathrm{T}_1}^{\mathrm{T}_1,\mathrm{T}_2}\}.$$

Note that the cash flow will be bounded from below by $-\mathrm{N}\kappa$.

1.9 Floors and Floorlets

A *floorlet* with reset date T_1 and settlement date T_2 pays the holder the positive part of the difference between a strike rate κ and $\mathrm{F}_{\mathrm{T}_1}^{\mathrm{T}_1,\mathrm{T}_2}$ at time T_2. That is

$$\text{Floorlet payoff at time } \mathrm{T}_2 = (\kappa - \mathrm{F}_{\mathrm{T}_1}^{\mathrm{T}_1,\mathrm{T}_2})^+.$$

Thus, a floorlet is essentially a put written on $F_{T_1}^{T_1,T_2}$. A floor is simply a strip of floorlets with reset dates $(T_0, T_1, \dots T_{n-1}$ and settlement dates $(T_1, T_2, \dots, T_n)$. Specifically

$$\text{Floor payoff at time } T_i = (\kappa - F_{T_{i-1}}^{T_{i-1},T_i})^+, \qquad i = 1, 2, \dots, n.$$

Clearly, the value of a floor is equal to the sum of the values of the individual floorlets. Assuming $t < T_0$ we have

$$V_t^{\text{floor}} = \sum_{i=1}^{n} V_t^{\text{floorlet},i}.$$

We will discuss how to find the value of a floorlet later in this course.

Floors and floorlets provide to their holder protection against falling interest rates. Specifically, they guarantees that the interest to be received on a floating rate loan never falls below the predetermined cap rate κ. For example, suppose an investor will receive a payment at time T_2 of $NF_{T_1}^{T_1,T_2}$. If the investor purchases a floorlet, then his cash flow at time T_2 will be

$$N(\kappa - F_{T_1}^{T_1,T_2})^+ + NF_{T_1}^{T_1,T_2} = N \max\{\kappa, F_{T_1}^{T_1,T_2}\}.$$

Note that the cash flow will be bounded from below by $N\kappa$.

1.10 Exercises

EXERCISE 1.1. Throughout this exercise, we will suppose that the short-rate R is a deterministic function of time: $R_t = R(t)$.

(a) Show via a no-arbitrage argument that

$$B_t^T = \frac{M_t}{M_T}, \qquad \forall 0 \le t \le T < \infty.$$

(b) Suppose that

$$R_t = r\Big(1 + \cos(2\pi t)\Big), \qquad r > 0. \tag{1.12}$$

Compute B_t^T, Y_t^T. $F_t^{T_1,T_2}$, $f_t^{T_1,T_2}$ and f_t^T.

(c) Is it true that $f_t^T = R_t$? Is it true that $f_t^T = R_T$? Explain your answer.

(d) Let $r = 0.05$ in (1.12) and plot B_0^T, Y_0^T and f_0^T as functions of T over the interval $[0, 1]$. Also plot $B_{1/2}^T$, $Y_{1/2}^T$ and $f_{1/2}^T$ as functions of T over the interval $[1/2, 3/2]$.

(e) Let R be given by (1.12). For what values of $\delta > 0$ do we have $B_t^T = B_{t+\delta}^{T+\delta}$?

EXERCISE 1.2. Fix dates $T_j = j$ for $j = 0, 1, 2, \dots, n = 5$. and Consider a floating rate swap with payments given by (1.8). Suppose

$$R_t = a + bt, \qquad 0 \leq t \leq 5, \qquad a = 0.10, \qquad b = -0.01.$$

(a) Compute the swap rate K_0^{swap}.
(b) Compute the payments received by the long side at times T_j for $j = 1, 2, \dots, 5$ assuming $K = K_0^{\text{swap}}$.
(c) Suppose an investor enters the long side on an interest rate swap at time $t = 0$. Then he will receive cash flows at times T_j for $j = 1, 2, \dots, n$. If the payment is negative, the investor borrows the money to make the payment from the money market account M. If the payment is positive, the investor invests the money in the money market account. How much money does the investor have after the final payment at time T_n?

EXERCISE 1.3. Consider a floating rate swap with payments given by (1.8) Recall that the swap rate K_t^{swap} is given by (1.9). Show that K_t^{swap} can be written of the form

$$K_t^{\text{swap}} = \sum_{i=1}^{n} w_t^i F_t^{T_{i-1}, T_i},$$

and identify w_t^i.

EXERCISE 1.4. Firx series of dates $\mathcal{T} = (T_0, T_1, T_2, \dots, T_n)$. At all dates $T_i > T_0$ the long side of the swap receives

$$\text{Payment at time } T_i := (T_i - T_{i-1}) F_{T_{i-1}}^{T_{i-1}, T_i} - (T_i - T_{i-1}) K.$$

(a) Show that the time $t < T_0$ value of the long side, denoted V_t^{swap} is equal to

$$V_t^{\text{swap}} = \sum_{i=1}^{n} (T_i - T_{i-1}) B_t^{T_i} \left(K_t^{\text{swap}} - K \right)$$

where K_t^{swap} is given by (1.9).
(b) Consider a caplet that pays $(T_2 - T_1)(F_{T_1}^{T_1, T_2} - K)^+$ at time $T_2 > T_1$. Show that this is equal in value to a cash flow at time T_1 of

$$\left(1 + (T_2 - T_1)K\right) \left(\frac{1}{1 + (T_2 - T_1)K} - B_{T_1}^{T_2} \right)^+.$$

(c) Let V_t^{swap} be the time $t < T_0$ value of the long side of the swap described above. And let V_t^{cap} and V_t^{floor} denote the time $t < T_0$ values of a cap and floor respectively, where

$$\text{Cap payment at time } T_i := \left((T_i - T_{i-1}) F_{T_{i-1}}^{T_{i-1}, T_i} - (T_i - T_{i-1}) K \right)^+,$$

$$\text{Floor payment at time } T_i := \left((T_i - T_{i-1}) K - (T_i - T_{i-1}) F_{T_{i-1}}^{T_{i-1}, T_i} \right)^+.$$

Prove that $V_t^{\text{swap}} = V_t^{\text{cap}} - V_t^{\text{floor}}$.

Chapter 2

Review of probability

The notes from this chapter are taken primarily from (Shreve, 2004, Chapter 1) and (Grimmett and Stirzaker, 2001, Chapters 1–5).

2.1 Events as sets

Definition 2.1.1. The set of all possible outcomes of an experiment is called the *sample space*. We denote the sample space as Ω.

We will typically denote by ω a generic element of Ω.

Definition 2.1.2. An *event* is a subset of the sample space. We usually denote events by capital roman letters $A, B, C, \ldots$.

Example 2.1.3 (Toss two distinguishable coins). $\Omega = \{(HH), (HT), (HT), (TT)\}$. One element of Ω is, e.g., $\omega = (HT)$. Possible event: ''second toss a tail.'' $A = \{(HT), (TT)\}$.

Example 2.1.4 (Roll a die). $\Omega = \{1, 2, 3, 4, 5, 6\}$. One element of Ω is, e.g., $\omega = 2$. Possible event: "roll an odd number." $A = \{1, 3, 5\}$.

If A and B are subsets of Ω, we can reasonably concern ourselves with events such as "not A" (A^c), "A or B" ($A \cup B$), "A and B" ($A \cap B$), etc. A σ-algebra is a mathematical way to descibe all possible sets of interest for a given sample space Ω.

Definition 2.1.5. A collection $\mathcal{F}$ of subsets of Ω is called a *σ-algebra* if it satisfies

1. contains the empty set: $\emptyset \in \mathcal{F}$;
2. is closed under countable unions: $A_1, A_2, A_3, \ldots \in \mathcal{F} \Rightarrow \cup_i A_i \in \mathcal{F}$;

3. is closed under complements: $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$;

Alternatively, one can define of a σ-algebra $\mathcal{F}$ as a set of subsets of Ω that contains at least the empty set $\emptyset$ and is closed under countable set operations (though, *not* necessarily closed under *un*countable set operators).

EXAMPLE 2.1.6 (TRIVIAL σ-ALGEBRA). The set of subsets $\mathcal{F}_0 := \{\emptyset, \Omega\}$ of Ω is commonly referred to as the *trivial* σ-algebra.

EXAMPLE 2.1.7. If A is a subset of Ω then $\mathcal{F}_A := \{\emptyset, \Omega, A, A^c\}$ is a σ-algebra.

EXAMPLE 2.1.8. The *power set* of Ω, written 2^Ω is the collection of all subsets of Ω. The power set $\mathcal{F} = 2^\Omega$ is a σ-algebra.

DEFINITION 2.1.9. Let $\mathcal{G}$ be a collection of subsets of Ω. The σ-algebra *generated* by $\mathcal{G}$, written $\sigma(\mathcal{G})$, is the smallest σ-algebra that contains $\mathcal{G}$.

By "smallest" σ-algebra we mean the σ-algebra with the fewest sets. One can show (although we will not do so in these notes) that $\sigma(\mathcal{G})$ is equal to the intersection of all σ-algebras that contain $\mathcal{G}$.

EXAMPLE 2.1.10. The collection of sets $\mathcal{G} = \{\emptyset, A, \Omega\}$ is not a σ-algebra because it does not contain A^c. However, we could create a σ-algebra from $\mathcal{G}$ by simply adding the set A^c. Thus, we have $\sigma(\mathcal{G}) = \{\emptyset, \Omega, A, A^c\}$.

DEFINITION 2.1.11. Let $\mathcal{O}(\mathbb{R}^d)$ be the set of open sets in $\mathbb{R}^d$. The *Borel σ-algebra on* $\mathbb{R}^d$, denoted $\mathcal{B}(\mathbb{R}^d)$ is the σ-algebra generated by open sets $\mathcal{B}(\mathbb{R}^d) := \sigma(\mathcal{O}(\mathbb{R}^d))$.

REMARK 2.1.12. Do not worry too much about what exactly Borel σ-algebras are. Just think of them as "reasonable" sets in $\mathbb{R}^d$. In fact, you would have to think very hard to come up with a set that is not a Borel set.

DEFINITION 2.1.13. The pair $(\Omega, \mathcal{F})$ where Ω is a sample space and $\mathcal{F}$ is a σ-algebra of subsets of Ω is called a *measurable space*.

2.2 PROBABILITY

So far, we have not yet talked about probabilities at all – only outcomes of a random experiment (elements $\omega \in \Omega$) and events (subsets $A \subseteq \Omega$). A probability measure assigns probabilities to events.

DEFINITION 2.2.1. A *probability measure* defined on $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ that satisfies

1. $\mathbb{P}(\Omega) = 1$;
2. if $A_i \cap A_j = \emptyset$ for $i \neq j$ then $\mathbb{P}(\cup_i A_i) = \sum_i \mathbb{P}(A_i)$. (countable additivity)

A probability measure $\mathbb{P}$ does *not* need to correspond to empirically observed probabilities! For example, from experience, we know that if we toss a fair coin we have $\mathbb{P}(\mathrm{H}) = \mathbb{P}(\mathrm{T}) = 1/2$. However, we can always define a measure $\widetilde{\mathbb{P}}$ that assigns different probabilities $\widetilde{\mathbb{P}}(\mathrm{H}) = p$ and $\widetilde{\mathbb{P}}(\mathrm{T}) = 1 - p$. As long as $p \in [0, 1]$ the measure $\widetilde{\mathbb{P}}$ is a probability measure on $(\Omega, \mathcal{F})$ where $\Omega = \{\mathrm{H}, \mathrm{T}\}$ and $\mathcal{F} = \{\emptyset, \Omega, \mathrm{H}, \mathrm{T}\}$.

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is often referred to as a *probability space* or *probability triple*. To review, the sample space Ω is the collection of all possible outcomes of an experiment. The σ-algebra $\mathcal{F}$ is all sets of interest of an experiment. And the probability measure $\mathbb{P}$ assigns probabilities to these sets.

When a sample space is countable $\Omega = \{\omega_1, \omega_2, \ldots\}$, we can always take the σ-algebra as the power set $\mathcal{F} = 2^\Omega$ and construct a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ by specifying the probabilities of each individual outcome $\mathbb{P}(\omega_i) = p_i$. However, when the sample space Ω is uncountable, choosing an appropriate σ-algebra $\mathcal{F}$, and constructing a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a more delicate procedure.

2.3 Infinite probability spaces

In this section we consider an infinite sequence of coin tosses. We define

$$\Omega := \text{ the set of infinite sequences of Hs and Ts .}$$

Note that this set is uncountable because there is a one-to-one correspondence between Ω and the set of reals in $[0, 1]$. We will denote a generic element of Ω as follows:

$$\omega = \omega_1 \omega_2 \omega_3 \ldots$$

where ω_i is the result of the ith coin toss. We want to construct a σ-algebra for this experiment.

Let us define some σ-algebras. First, consider the trivial σ-algebra

$$\mathcal{F}_0 = \{\emptyset, \Omega\}.$$

Given no information, I can tell if ω is in the sets in $\mathcal{F}_0$ because we know $\omega \in \Omega$ and $\omega \notin \emptyset$. Next, define two sets

$$\mathrm{A_H} = \{\omega \in \Omega : \omega_1 = \mathrm{H}\}, \qquad \mathrm{A_T} = \{\omega \in \Omega : \omega_1 = \mathrm{T}\}.$$

Noting that $\mathrm{A_H} = \mathrm{A_T^c}$ we see that

$$\mathcal{F}_1 := \{\emptyset, \Omega, \mathrm{A_H}, \mathrm{A_T}\},$$

satisfies the conditions of σ-algebra. Given ω_1 it is possible to say whether or not ω is in each of the sets in $\mathcal{F}_1$. For example, if $\omega_1 = \mathrm{H}$ then $\omega \in \mathrm{A_H}$ and $\omega \in \Omega$, but $\omega \notin \mathrm{A_T}$ and $\omega \notin \emptyset$. Next define four sets

$$\begin{aligned} \mathrm{A_{HH}} &:= \{\omega \in \Omega : \omega_1 = \mathrm{H}, \omega_2 = \mathrm{H}\}, & \mathrm{A_{HT}} &:= \{\omega \in \Omega : \omega_1 = \mathrm{H}, \omega_2 = \mathrm{T}\}, \\ \mathrm{A_{TT}} &:= \{\omega \in \Omega : \omega_1 = \mathrm{T}, \omega_2 = \mathrm{T}\}, & \mathrm{A_{TH}} &:= \{\omega \in \Omega : \omega_1 = \mathrm{T}, \omega_2 = \mathrm{H}\}. \end{aligned}$$

We wish to construct a σ-algebra that contains these sets and the sets in $\mathcal{F}_1$. The smallest such σ-algebra is

$$\mathcal{F}_2 = \left\{ \begin{matrix} \emptyset, \Omega, \mathrm{A_H}, \mathrm{A_T}, \mathrm{A_{HH}}, \mathrm{A_{HT}}, \mathrm{A_{TT}}, \mathrm{A_{TH}}, \mathrm{A^c_{HH}}, \mathrm{A^c_{HT}}, \mathrm{A^c_{TT}}, \mathrm{A^c_{TH}} \\ \mathrm{A_{HH}} \cup \mathrm{A_{TH}}, \mathrm{A_{HH}} \cup \mathrm{A_{TT}}, \mathrm{A_{HT}} \cup \mathrm{A_{TH}}, \mathrm{A_{HT}} \cup \mathrm{A_{TT}} \end{matrix} \right\}.$$

Given ω_1 and ω_2, we can say if ω belongs to each of the sets in $\mathcal{F}_2$. Continuing in this way, we can define a σ-algebra $\mathcal{F}_n$ for every $n \in \mathbb{N}$. Finally, we take

$$\mathcal{F} := \sigma(\mathcal{F}_\infty), \qquad \mathcal{F}_\infty = \cup_n \mathcal{F}_n.$$

One might ask if we could have simply taken $\mathcal{F} = \mathcal{F}_\infty$? Well, $\mathcal{F}_\infty$ contains every set that can be described in terms of *finitely many* coin tosses. However, we may be interested in sets such as "sequences for which x percent of coin tosses are heads," and these sets are not in $\mathcal{F}_\infty$. It turns out such sets are in $\mathcal{F}$.

Now, we want to construct a probability measure on $\mathcal{F}$. Let us assume the coin tosses are independent (a term we will describe rigorously later on) and that the probability of a head is p. Setting $q = 1 - p$, it should be obvious that

$$\begin{aligned} &\mathbb{P}(\emptyset) = 0, & &\mathbb{P}(\Omega) = 1, & &\mathbb{P}(\mathrm{A_H}) = p, & &\mathbb{P}(\mathrm{A_T}) = q, \\ &\mathbb{P}(\mathrm{A_{HH}}) = p^2, & &\mathbb{P}(\mathrm{A_{HT}}) = pq, & &\mathbb{P}(\mathrm{A_{TH}}) = pq, & &\mathbb{P}(\mathrm{A_{TT}}) = q^2, \ldots \end{aligned}$$

Continuing in this way, we can define $\mathbb{P}(\mathrm{A})$ for every $\mathrm{A} \in \mathcal{F}_\infty$. What about the sets that are in $\mathcal{F}$ but not in $\mathcal{F}_\infty$? It turns out that once we have defined $\mathbb{P}$ for sets in $\mathcal{F}_\infty$ there is only one way to assign probabilities to those sets that are in $\mathcal{F}$ but not in $\mathcal{F}_\infty$. We refer the interested reader to *Carathéodory's Extension Theorem* for details.

Now, let us define

$$\mathrm{A} = \left\{ \omega : \lim_{n \to \infty} \frac{\#\mathrm{H} \text{ in first } n \text{ coin tosses}}{n} = \frac{1}{2} \right\}.$$

The strong law of large numbers (SLLN) tells us that $\mathbb{P}(\mathrm{A}) = 1$ if $p = 1/2$ and $\mathbb{P}(\mathrm{A}) = 0$ if $p \neq 1/2$ (if you have not yet seen the SLLN, you should be able to see this from intuition). Now it should be clear why *uncountable* additivity does *not* hold for probability measures. The probability of any given sequence of infinite coin tosses is zero: $\mathbb{P}(\omega) = 0$. If we were to attempt to compute $\mathbb{P}(A)$ by adding up the probabilities $\mathbb{P}(\omega)$ of all elements $\omega \in A$ we would find

$$\sum_{\omega \in A} \mathbb{P}(\omega) = \sum_{\omega \in A} 0 = 0 \neq 1 = \mathbb{P}(A), \qquad \text{(when } p = 1/2\text{)}.$$

Thus, uncountable additivity clearly does *not* hold.

We finish this example (we will come back to it!) with the following definition

<u>Definition</u> 2.3.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. If a set $A \in \mathcal{F}$ satisfies $\mathbb{P}(A) = 1$, we say that the event A occurs $\mathbb{P}$ *almost surely* (written, $\mathbb{P}$-a.s.).

Note in the example above that, when $p = 1/2$ we have $\mathbb{P}(A) = 1$ and thus A occurs almost surely. But it is important to recognize that $A \neq \Omega$ and $A^c \neq \emptyset$. The elements of A^c are part of the sample space Ω, but they have zero probability of occurring.

2.4 Random variables and distributions

A random variable maps the outcome of an experiment to $\mathbb{R}$. We capture this idea with the following definition.

<u>Definition</u> 2.4.1. A *random variable* defined on $(\Omega, \mathcal{F})$ is a function $X : \Omega \to \mathbb{R}$ with the property that

$$\{X \in A\} := \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F},$$

for all $A \in \mathcal{B}(\mathbb{R})$.

Observe all any random variables must be defined on a measurable space $(\Omega, \mathcal{F})$, as these appear in the definition. Note, however, that the probability meausre $\mathbb{P}$ does *not* appear in the definition. Random variables are defined *independent* of a probability measure $\mathbb{P}$.

What does Definition 2.4.1 mean? Recall that a probability measure $\mathbb{P}$ defined on $(\Omega, \mathcal{F})$ maps $\mathcal{F} \to [0, 1]$. In order for us to answer the question: "what is the probability that $X \in A$?" we need for the set $\{X \in A\} \in \mathcal{F}$. And this is precisely what Definition 2.4.1 requires. Why do we only consider sets $A \in \mathcal{B}(\mathbb{R})$ rather than any set $A \subset \mathbb{R}$? The answer is rather technical and, frankly, not worth exploring at the moment.

A word on notation: the standard convention is to use capital Roman letters (typically, X, Y, Z) for random variables and lower case Roman letters (x, y, z) for real numbers.

Let us us look at some random variables.

Example 2.4.2 (Discrete time model for stock prices). Consider the infinite sequence of coin tosses in Section 2.3. We Define a sequence of random variables $(\mathrm{S}_n)_{n\geq 0}$ via

$$\mathrm{S}_0(\omega) = 1, \qquad \mathrm{S}_{n+1}(\omega) = \begin{cases} u\mathrm{S}_n & \text{if } \omega_n = \mathrm{H} \\ d\mathrm{S}_n & \text{if } \omega_n = \mathrm{T} \end{cases} \tag{2.1}$$

Here, S_n represents the value of a stock at time n. Note that $\mathbb{P}(\mathrm{S}_1 = u) = \mathbb{P}(\mathrm{A}_\mathrm{H}) = p$. Likewise $\mathbb{P}(\mathrm{S}_2 = ud) = \mathbb{P}(\mathrm{A}_{\mathrm{HT}} \cup \mathrm{A}_{\mathrm{TH}}) = 2pq$. More generally, one can show that

$$\mathbb{P}(\mathrm{S}_n = u^k d^{n-k}) = \binom{n}{k} p^k q^{n-k}. \tag{2.2}$$

Note if we had simply defined the random variables $(\mathrm{S}_n)_{n\geq 1}$ as having probabilities given by (2.2) we would have no information about how, e.g., S_n relates to S_{n-1}. From the above construction (2.1), however, we know that if $\mathrm{S}_n = u^n$ then $\mathrm{S}_{n-1} = u^{n-1}$. Thus, the structure of a given probability space, not just the probabilities of events, is very important.

Example 2.4.3. Let $(\Omega, \mathcal{F}) = ((0,1), \mathcal{B}((0,1))$ Define random variables $\mathrm{X}(\omega) = \omega$ and $\mathrm{Y}(\omega) = 1 - \omega$. Clearly, we have $\mathrm{X} = 1 - \mathrm{Y}$. Now, suppose we defined $\mathbb{P}(\mathrm{d}\omega) := \mathrm{d}\omega$. Then X and Y have the same distribution. For $x \in [0,1]$ we have

$$\begin{aligned} \mathbb{P}(\mathrm{X} \leq x) &= \mathbb{P}(\omega \leq x) = \int_0^x \mathbb{P}(\mathrm{d}\omega) = \int_0^x \mathrm{d}\omega = x, \\ \mathbb{P}(\mathrm{Y} \leq x) &= \mathbb{P}(1 - \omega \leq x) = \int_{1-x}^1 \mathbb{P}(\mathrm{d}\omega) = \int_{1-x}^1 \mathrm{d}\omega = x. \end{aligned}$$

However, if we defined a new probability measure via $\widetilde{\mathbb{P}}(\mathrm{d}\omega) = 2\omega\mathrm{d}\omega$ then X and Y have different distributions. For $x \in [0,1]$ we have

$$\begin{aligned} \mathbb{P}(\mathrm{X} \leq x) &= \mathbb{P}(\omega \leq x) = \int_0^x \widetilde{\mathbb{P}}(\mathrm{d}\omega) = \int_0^x 2\omega\mathrm{d}\omega = x^2, \\ \mathbb{P}(\mathrm{Y} \leq x) &= \mathbb{P}(1 - \omega \leq x) = \int_{1-x}^1 \widetilde{\mathbb{P}}(\mathrm{d}\omega) = \int_{1-x}^1 2\omega\mathrm{d}\omega = 1 - (1-x)^2. \end{aligned}$$

The distribution of a random variable X is most easily described through its cumulative distribution function.

Definition 2.4.4. The *distribution function* $\mathrm{F}_\mathrm{X} : \mathbb{R} \to [0,1]$ of a random variable X defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is given by

$$\mathrm{F}_\mathrm{X}(x) := \mathbb{P}(\mathrm{X} \leq x).$$

Observe that, while a random variable X is defined with respect to $(\Omega, \mathcal{F})$ (with no reference to $\mathbb{P}$), the distribution F_X is specific to a probability measure $\mathbb{P}$.

Note that we put the random variable X in the subscript of F_X to remind us that F_X is the distribution function corresponding to the random variable X (and not, e.g., Y). It is a good idea to do this.

Many (but not all) random variables fall in to one of two categories: discrete and continuous. We describe these two categories below.

<u>Definition</u> 2.4.5. A random variable X is called *discrete* if it takes values in some countable set $A := \{x_1, x_2, \ldots\} \subset \mathbb{R}$. We associate is a discrete random variable a *probability mass function* $f_X : A \to \mathbb{R}$, defined by $f_X(x_i) := \mathbb{P}(X = x_i)$.

<u>Definition</u> 2.4.6. A random variable X is called *continuous* if its distribution function F_X can be written as

$$F_X(x) = \int_{-\infty}^{x} \mathrm{d}u \, f_X(u), \qquad x \in \mathbb{R},$$

for some $f_X : \mathbb{R} \to [0, \infty)$ called the *probability density function.*

It may help to think of the density function f_X as $f_X(x)\mathrm{d}x = \mathbb{P}(X \in \mathrm{d}x)$.

Note that for a continuous random variable X we have $f_X = F'_X$.

If X is either discrete or continuous, it is easy to compute $\mathbb{P}(X \in A)$ for any $A \in \mathcal{B}(\mathbb{R})$. We have

$$\text{discrete} : \qquad \mathbb{P}(X \in A) = \sum_{\{i : x_i \in A\}} f_X(x_i),$$

$$\text{continuous} : \qquad \mathbb{P}(X \in A) = \int_A \mathrm{d}x \, f_X(x).$$

<u>Remark</u> 2.4.7. Although we have defined $F_X : \mathbb{R} \to [0, 1]$ by $F_X(x) := \mathbb{P}(X \le x)$, it is common to also to utilize F_X as a set function $F_X : \mathcal{B}(\mathbb{R}) \to [0, 1]$, which means $F_X(B) := \mathbb{P}(X \in B)$. It should always be clear from the argument of F_X, which of the two meanings we intend.

Examples of discrete random variables

The following discrete random variables frequently arise in applications in nature and social sciences.

<u>Example</u> 2.4.8. If X is distributed as a *Bernoulli random variable* with parameter $p \in [0, 1]$, written $X \sim \mathrm{Ber}(p)$, then

$$X \in \{0, 1\}, \qquad f_X(k) = \begin{cases} 1 - p & k = 0, \\ p & k = 1. \end{cases}$$

<u>Example</u> 2.4.9. If X is distributed as a *Binomial random variable* with parameters $n \in \mathbb{N}$ and $p \in [0,1]$, written $X \sim \text{Bin}(n,p)$, then

$$X \in \{0,1,2,\ldots,n\}, \qquad f_X(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Note that if $X_i \sim \text{Ber}(p)$ and independent of each other then $Y := \sum_{i=1}^n X_i \sim \text{Bin}(n,p)$.

Examples of continuous random variables

Before introducing some common continuous random variables, let us introduce a useful function.

<u>Definition</u> 2.4.10. Let A be a set in some topological space Ω (e.g., $\Omega = \mathbb{R}^d$). The *indicator function* $\mathbb{1}_A : \Omega \to \{0,1\}$ is defined as follows

$$\mathbb{1}_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

We now introduce some continuous random variables that frequently arise in applications.

<u>Example</u> 2.4.11. If X is distributed as a *Exponential random variable* with mean $\lambda > 0$, written $X \sim \mathcal{E}(\lambda)$, then

$$X \in [0,\infty), \qquad f_X(x) = \mathbb{1}_{[0,\infty)}(x)\lambda e^{-\lambda x}.$$

Note that $f_X(x) = 0$ if $x < 0$ due to the presence of the indicator function.

<u>Example</u> 2.4.12. If X is distributed as a *Gaussian* or *Normal random variable* with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$ (we will give a meaning for "mean" and "variance" below), written $X \sim \mathcal{N}(\mu, \sigma^2)$, then

$$X \in \mathbb{R}, \qquad f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

A random variable $Z \sim \mathcal{N}(0,1)$ is referred to as *standard normal.*

2.5 Stochastic Processes

Intuitively, we think of a stochastic process as a process that evolves randomly in time. Now that we understand what a random variable is, we can define rigourously what we mean when we say *stochastic process.*

Definition 2.5.1. A *Stochastic process* is a collection of random variables $X = (X_t)_{t \in \mathbb{T}}$ where $\mathbb{T}$ is some index set. If the index set $\mathbb{T}$ is countable (e.g., $\mathbb{T} = \mathbb{N}_0$) we say that X is a *discrete time* process. If the index set $\mathbb{T}$ is uncountable (e.g., $\mathbb{T} = \mathbb{R}_+$) we say that X is a *continuous time* process. The *State Space* S of a stochastic process X is union of the state spaces of $(X_t)_{t \in \mathbb{T}}$.

We can think of a stochastic process $X : \mathbb{T} \times \Omega \to \mathbb{R}$ in (at least) two ways. First, for any $t \in \mathbb{T}$ we have that $X_t : \Omega \to \mathbb{R}$ is random variable. Second, for any $\omega \in \Omega$, we have that $X_\cdot(\omega) : \mathbb{T} \to \mathbb{R}$ is a function of time. Both interpretations can be useful.

2.6 Expectation

When we think of averaging we think of weighting outcomes by their probabilities. The mathematical way to encode this is via the expectation.

Definition 2.6.1. Let X be a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$. The *expectation* of X, writtien $\mathbb{E}X$, is defined as

$$\mathbb{E}X := \int_\Omega X(\omega)\mathbb{P}(d\omega),$$

where the integral is understood in the Lebesgue sense.

2.6.1 Integration in the Lebesgue sense

For those who have not previously encountered Lesbesgue integration, we now give a brief (*very brief*!) overview of this concept.

Definition 2.6.2. Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $A \in \mathcal{F}$. The *indicator random variable*, denoted $\mathbb{1}_A$, is defined by

$$\mathbb{1}_A(\omega) := \begin{cases} 1 & \omega \in A, \\ 0 & \omega \notin A. \end{cases}$$

Observe that $\mathbb{1}_A \sim \mathrm{Ber}(p)$ with $p = \mathbb{P}(A)$. For disjoint sets A and B we have

$$\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B, \qquad\qquad A \cap B = \emptyset.$$

And, for any two sets A and B we have

$$\mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B.$$

<u>Definition</u> 2.6.3. A collection of non-empty sets (A_i) is said to be a *partition* of Ω if $A_i \cap A_j \neq \emptyset$ for all i and j and $\cup_i A_i = \Omega$.

<u>Definition</u> 2.6.4. Let (A_i) be a finite partition of Ω. A non-negative random variable X, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which is of the form

$$X(\omega) = \sum_{i=1}^{n} x_i \mathbb{1}_{A_i}(\omega), \qquad x_i \geq 0, \qquad A_i \in \mathcal{F},$$

is called *simple*.

Let X be a simple random variable. We define the expectation of X as follows

$$\mathbb{E}X := \sum_{i=1}^{n} x_i \mathbb{P}(A_i). \qquad \text{(if X is simple)}$$

Note that, from this definition, we have

$$\mathbb{E}\mathbb{1}_A = \mathbb{P}(A).$$

Thus, *we can always represent probabilities of sets as expectations of indicator random variables.*

Now, consider a non-negative random variable X, which is not necessarily simple. Let $(X_n)_{n\geq 0}$ be an increasing sequence of simple random variables that converges almost surely to X. That is

$$X_i \leq X_{i+1}, \qquad \lim_{i\to\infty} X_i \to X, \qquad \mathbb{P}\text{-a.s.}.$$

We define the expectation of a non-negative random variable X as the following limit

$$\mathbb{E}X := \lim_{i\to\infty} \mathbb{E}X_i, \qquad \text{(if X is non-negative)} \tag{2.3}$$

where each of the expectations on the right-hand side are well-defined because all of the X_i are simple by construction. Finally, consider a general random variable X that could take either positive or negative values. Define

$$X^+ = \max\{X, 0\}, \qquad X^- = \max\{-X, 0\}.$$

Note that X^+ and X^- are non-negative and $X = X^+ - X^-$. With this in mind, we define

$$\mathbb{E}X := \mathbb{E}X^+ - \mathbb{E}X^-,$$

where the expectations of X^+ and X^- are defined via (2.3).

Definition 2.6.1 of $\mathbb{E}X$ makes sense if $\mathbb{E}|X| < \infty$ or if $\mathbb{E}X^\pm = \infty$ and $\mathbb{E}X^\mp < \infty$. In the latter case, we have $\mathbb{E}X = \pm\infty$. If both $\mathbb{E}X^+ = \infty$ and $\mathbb{E}X^- = \infty$, then we find ourselves in an $\infty - \infty$ situation and, in this case, $\mathbb{E}X$ is undefined.

2.6.2 Computing expectations

If X is either discrete or continuous Definition 2.6.1 reduces to the formulas one learns as an undergraduate.

$$\begin{aligned} \text{discrete}: && \mathbb{E}X &= \sum_i x_i f_X(x_i), \\ \text{continuous}: && \mathbb{E}X &= \int_{\mathbb{R}} dx\, x\, f_X(x). \end{aligned}$$

In the discrete case, the sum runs over all possible values of x.

Note that $\mathbb{E}$ is a linear operator. If X and Y are random variables and a and b are constants, then

$$\mathbb{E}(aX + bY) = a\mathbb{E}X + b\mathbb{E}Y.$$

How does one compute $\mathbb{E}g(X)$ where $g : \mathbb{R} \to \mathbb{R}$? Although we have not stated it explicitly, it should be obvious that if X is a random variable, then $Y := g(X)$ is also a random variable. [1] Thus, we have

$$\mathbb{E}Y = \mathbb{E}g(X) = \int_\Omega g(X(\omega))\mathbb{P}(d\omega),$$

which in the discrete and continuous cases become

$$\begin{aligned} \text{discrete}: && \mathbb{E}g(X) &= \sum_i g(x_i) f_X(x_i), \\ \text{continuous}: && \mathbb{E}g(X) &= \int_{\mathbb{R}} dx\, g(x) f_X(x). \end{aligned}$$

2.7 Change of measure

Consider two probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ defined on a measurable space $(\Omega, \mathcal{F})$. What is the relation between $\mathbb{P}$ and $\widetilde{\mathbb{P}}$? The following theorem answers this question.

Theorem 2.7.1. *Fix a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and let* $Z \geq 0$ *be a random variable satisfying* $\mathbb{E}Z = 1$. *Define a* $\widetilde{\mathbb{P}} : \mathcal{F} \to [0, 1]$ *by*

$$\widetilde{\mathbb{P}}(A) := \mathbb{E}Z\mathbb{1}_A. \tag{2.4}$$

Then $\widetilde{\mathbb{P}}$ *is a probability measure on* $(\Omega, \mathcal{F})$. *Denote by* $\widetilde{\mathbb{E}}$ *the expectation taken with respect to* $\widetilde{\mathbb{P}}$. *Then*

$$\widetilde{\mathbb{E}}X = \mathbb{E}ZX, \qquad \textit{and if} \qquad Z > 0, \qquad \textit{then} \qquad \mathbb{E}X = \widetilde{\mathbb{E}}\frac{1}{Z}X. \tag{2.5}$$

where X *is a random variable defined on* $(\Omega, \mathcal{F})$.

[1] Rigorously, g should be a *measurable function*, meaning $g^{-1}(A) \in \mathcal{B}(\mathbb{R})$ for all $A \in \mathcal{B}(\mathbb{R})$. Do not concern yourself too much with this.

<u>Definition</u> 2.7.2. We call the random variable Z in Theorem 2.7.1 the *Radon-Nikodým derivative* of $\widetilde{\mathbb{P}}$ with respect to $\mathbb{P}$.

<u>Definition</u> 2.7.3. Two probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ on $(\Omega, \mathcal{F})$ are *equivalent*, written $\mathbb{P} \sim \widetilde{\mathbb{P}}$, if

$$\mathbb{P}(A) = 0 \qquad \Leftrightarrow \qquad \widetilde{\mathbb{P}}(A) = 0.$$

Two probability measures are equivalent $\mathbb{P} \sim \widetilde{\mathbb{P}}$ if and only if the Radon-Nikodým Derivative that relates them is strictly positive $Z > 0$. Equivalent measures agree on which events will happen with probability zero (and thus, they agree on which events will happen with probability one).

<u>Example</u> 2.7.4. Set $(\Omega, \mathcal{F}) = ((0,1), \mathcal{B}((0,1))$. On this measure space, we define two probability measures $\mathbb{P}(d\omega) = d\omega$ and $\widetilde{\mathbb{P}}(d\omega) = 2\omega d\omega$. Note that we have

$$\widetilde{\mathbb{P}}(A) = \widetilde{\mathbb{E}}\mathbb{1}_A = \int_\Omega \mathbb{1}_A(\omega)\widetilde{\mathbb{P}}(d\omega) = \int_\Omega \mathbb{1}_A(\omega)2\omega d\omega = \int_\Omega \mathbb{1}_A 2\omega \mathbb{P}(d\omega) = \mathbb{E}\mathbb{1}_A Z, \qquad Z(\omega) := 2\omega.$$

One can easily check that $\mathbb{E}Z = 1$ and $Z > 0$. Defining $\widetilde{\mathbb{P}}$ by (2.4), one can easily check that (2.5) holds true.

It is quite common to use the notation

$$Z(\omega) = \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}(\omega), \qquad \widetilde{\mathbb{P}}(d\omega) = \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}(\omega)\mathbb{P}(d\omega),$$

as a reminder of how the Radon-Nikodým Derivative Z relates $\widetilde{\mathbb{P}}$ to $\mathbb{P}$. For a finite probability space, it is true that $Z(\omega) = \widetilde{\mathbb{P}}(\omega)/\mathbb{P}(\omega)$. However, for an infinite probability space, it makes no sense in general to define $Z(\omega) = \widetilde{\mathbb{P}}(\omega)/\mathbb{P}(\omega)$ since it may be that $\mathbb{P}(\omega) = 0$. Nevertheless, the heuristic $Z(\omega) = \widetilde{\mathbb{P}}(\omega)/\mathbb{P}(\omega)$ gives the correct intuition. In particular, for the special case of an infinite probability space in which $\mathbb{P}(d\omega) = p(\omega)d\omega$ and $\widetilde{\mathbb{P}}(d\omega) = \widetilde{p}(\omega)d\omega$ and $\mathbb{P} \sim \widetilde{\mathbb{P}}$, we have $Z(\omega) = \widetilde{p}(\omega)/p(\omega)$.

<u>Example</u> 2.7.5 (Change of measure Normal random variable). On $(\Omega, \mathcal{F}, \mathbb{P})$ let $X \sim \mathcal{N}(0,1)$ and define $Y = X + \theta$. Clearly, we have $Y \sim \mathcal{N}(\theta, 1)$. Now, define a random variable Z by

$$Z = e^{-\theta X - \frac{1}{2}\theta^2}.$$

Clearly $Z > 0$. We also have $\mathbb{E}Z = 1$. To see this, simply compute

$$\begin{aligned}
\mathbb{E}Z &= \int_{\mathbb{R}} dx\, e^{-\theta x - \frac{1}{2}\theta^2} f_X(x) \\
&= \int_{\mathbb{R}} dx\, e^{-\theta x - \frac{1}{2}\theta^2} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \\
&= \int_{\mathbb{R}} dx\, \frac{1}{\sqrt{2\pi}} e^{-(x+\theta)^2/2} = 1.
\end{aligned}$$

Since $Z > 0$ and $\mathbb{E}Z = 1$, we can define a new probability measure $\widetilde{\mathbb{P}}$ with $Z = d\widetilde{\mathbb{P}}/d\mathbb{P}$ as the Radon-Nikodým derivative. Let us compute the distribution of Y under $\widetilde{\mathbb{P}}$. We have

$$\begin{aligned}\widetilde{\mathbb{P}}(Y \leq b) &= \mathbb{E}Z\mathbb{1}_{\{Y \leq b\}} = \mathbb{E}e^{-\theta X - \frac{1}{2}\theta^2}\mathbb{1}_{\{X \leq b-\theta\}} \\ &= \int_{-\infty}^{b-\theta} dx\, e^{-\theta x - \frac{1}{2}\theta^2} f_X(x) \\ &= \int_{-\infty}^{b-\theta} dx\, \frac{1}{\sqrt{2\pi}} e^{-(x+\theta)^2/2} \\ &= \int_{-\infty}^{b} dz\, \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\end{aligned}$$

Thus, under $\widetilde{\mathbb{P}}$ we see that $Y \sim \mathcal{N}(0,1)$. The Radon-Nikodým derivative Z changes the mean of Y from θ to 0, but it does not affect the variance of Y.

2.8 Information and σ-algebras

Let us return to the coin-toss example of Section 2.3. If we are given no information about ω what can we say about ω? In other words, what are the subsets of Ω for which we can say: "ω is in this set" or "ω is not in this set"? The answer is $\emptyset$ and Ω, which, together, form the trivial σ-algebra $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

Now suppose we are given the value of ω_1. What are the subsets of Ω for which we can say: "ω is in this set" or "ω is not in this set"? The answer is the sets in $\mathcal{F}_0$ as well as A_H and A_T. Together, these sets form the σ-algebra $\mathcal{F}_1 = \{\emptyset, \Omega, A_H, A_T\}$. We say the sets in $\mathcal{F}_1$ are *resolved* by the first coin toss.

Now suppose we are given the value of ω_1 and ω_2. What are the subsets of Ω for which we can say: "ω is in this set" or "ω is not in this set"? The answer is the sets in $\mathcal{F}_2$, given by

$$\mathcal{F}_2 = \left\{\begin{matrix} \emptyset, \Omega, A_H, A_T, A_{HH}, A_{HT}, A_{TT}, A_{TH}, A_{HH}^c, A_{HT}^c, A_{TT}^c, A_{TH}^c \\ A_{HH} \cup A_{TH}, A_{HH} \cup A_{TT}, A_{HT} \cup A_{TH}, A_{HT} \cup A_{TT} \end{matrix}\right\}.$$

The sets in $\mathcal{F}_2$ are resolved by the first two coin tosses.

Continuing in this way, for each $n \in \mathbb{N}$ we can define $\mathcal{F}_n$ as the σ-algebra containing the sets that are resolved by the first n coin tosses. Note that if a set $A \in \mathcal{F}_n$ then $A \in \mathcal{F}_{n+1}$. Thus, $\mathcal{F}_n \subset \mathcal{F}_{n+1}$. In other words, $\mathcal{F}_{n+1}$ contains more "information" than $\mathcal{F}_n$. This kind of structure is encapsulated in the following definition.

Definition 2.8.1. Let Ω be a nonempty set. Let T be a fixed positive number, and assume that for each $t \in [0,T]$ there is a σ-algebra $\mathcal{F}_t$. Assume further that if $0 \leq s \leq t \leq T$, then $\mathcal{F}_s \subseteq \mathcal{F}_t$. Then we call the sequence of σ-algebras $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]}$ a *continuous time filtration.*

A *discrete time filtration* is a sequence of σ-algebras $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ that satisfies $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ for all n.

In the above example we generated a sequence of σ-algebras by observing directly an element $\omega \in \Omega$. Suppose that, instead of observing ω we can observe only a random variable $X(\omega)$. We can use this information to generate a σ-algebra as well.

<u>Definition</u> 2.8.2. Let X be a random variable defined on a nonempty sample space Ω. The *σ-algebra generated by* X, denoted $\sigma(X)$, is the collection of all subsets of Ω of the form $\{X \in A\}$ where $A \in \mathcal{B}(\mathbb{R})$.

<u>Example</u> 2.8.3. Let us return to Example 2.4.2. What is $\sigma(S_2)$? From the definition, we need to ask, which sets are of the form $\{S_2 \in A\}$? Since S_2 can only take three values, u^2, ud and d^2 we check the following sets

$$\{S_2 = u^2\} = A_{HH}, \qquad \{S_2 = ud\} = A_{HT} \cup A_{TH}, \qquad \{S_2 = d^2\} = A_{TT}.$$

We add to these sets the sets that are necessary to form a σ-algebra (i.e., $\emptyset$, Ω and unions and complements of the above sets) to obtain

$$\sigma(S_2) = \sigma(\{A_{HH}, A_{TT}, A_{HT} \cup A_{TH}\}).$$

Note, that $\sigma(S_2) \subset \mathcal{F}_2$ since $A_{HT}, A_{TH} \in \mathcal{F}_2$ but $A_{HT}, A_{TH} \notin \sigma(S_2)$. The reason is that, if $S_2 = ud$ we cannot say if $\omega_1 = T$ or $\omega_1 = H$.

<u>Definition</u> 2.8.4. Let X be a random variable defined on a nonempty sample space Ω. Let $\mathcal{G}$ be a σ-algebra of subsets of Ω. If $\sigma(X) \subset \mathcal{G}$ we say that X *is $\mathcal{G}$-measurable*, and we write $X \in \mathcal{G}$.

A random variable X is $\mathcal{G}$-measurable if and only if the information in $\mathcal{G}$ is sufficient to determine the value of X. Obviously, if $X \in \mathcal{G}$ then $g(X) \in \mathcal{G}$ (assuming is g is a measurable map from $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$).

Eventually, we will want to consider stochastic processes $X = (X_t)_{t\in[0,T]}$ and we will want to know at each time t if X_t is measureable with respect to σ-algebra $\mathcal{F}_t$.

<u>Definition</u> 2.8.5. Let Ω be a nonempty sample space equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$. Let $X = (X_t)_{t\in[0,T]}$ be a collection of random variables indexed by $t \in [0, T]$. We say this collection of random variables is $\mathbb{F}$-*adapted* if $X_t \in \mathcal{F}_t$ for all $t \in [0, T]$.

2.9 Independence

When $X \in \mathcal{G}$ this means that the information in $\mathcal{G}$ is sufficient to determine the value of X. On the other extreme, if X is independent (a term we will define soon) of $\mathcal{G}$ this means that the informtion in $\mathcal{G}$ tells us nothing about the value of X.

Definition 2.9.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We say that two sets two sets A and B in $\mathcal{F}$ are *independent*, written $A \perp\!\!\!\perp B$, if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

Having defined independent sets, we can now extend to independent σ-algebras and random variables.

Definition 2.9.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$ (i.e., $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$). We say these *two σ-algebras are independent*, written $\mathcal{G} \perp\!\!\!\perp \mathcal{H}$, if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B), \qquad \forall\, A \in \mathcal{G}, \qquad \forall\, B \in \mathcal{H}.$$

Let X and Y be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. We say these *two random variables are independent*, written $X \perp\!\!\!\perp Y$, if $\sigma(X) \perp\!\!\!\perp \sigma(Y)$. Lastly, we say the random variable X is *independent* of the σ-algebra $\mathcal{G}$, written $X \perp\!\!\!\perp \mathcal{G}$, if $\sigma(X) \perp\!\!\!\perp \mathcal{G}$.

Recall from Definition 2.8.2 that $\sigma(X)$ contains all sets of the form $\{X \in A\}$, where $A \in \mathcal{B}(\mathbb{R})$. Combining this with Definition 2.9.2 we see that

$$X \perp\!\!\!\perp Y \qquad \Leftrightarrow \qquad \mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \cdot \mathbb{P}(Y \in B), \qquad \forall\, A, B \in \mathcal{B}(\mathbb{R}). \tag{2.6}$$

It follows from (2.6) that

$$X \perp\!\!\!\perp Y \qquad \Rightarrow \qquad \mathbb{E}XY = \mathbb{E}X \cdot \mathbb{E}Y.$$

Note that $\mathbb{E}XY = \mathbb{E}X \cdot \mathbb{E}Y$ does *not* imply $X \perp\!\!\!\perp Y$.

The above notion of independence is called *pairwise* independence. If $X \perp\!\!\!\perp Y$ and $Y \perp\!\!\!\perp Z$, this notion of independence *not* imply $X \perp\!\!\!\perp Z$ (for example, what if $Z = X$?). Thus, at times, we may need a stronger notion of independence.

Definition 2.9.3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n$ be sub-σ-algebras of $\mathcal{F}$. We say the *sequence of σ-algebras are independent*, if

$$\mathbb{P}(\cap_{i=1}^n A_i) = \prod_{i=1}^n \mathbb{P}(A_i), \qquad \forall\, A_1 \in \mathcal{G}_1, \forall\, A_2 \in \mathcal{G}_2, \ldots, \forall A_n \in \mathcal{G}_n.$$

Let $X_1, X_2, \ldots, X_n$ be a sequence of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. We say the *sequence of random variables are independent* if the σ-algebras $\sigma(X_1), \sigma(X_2), \ldots, \sigma(X_n)$ are independent.

As with with a pair of random variables, a sequence of random variables $(X_i)_{i \geq 1}$ is independent if and only if

$$\mathbb{P}\left(\cap_{i=1}^n \{X_i \in A_i\}\right) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i), \qquad \forall\, A_1 \in \mathcal{B}(\mathbb{R}), \forall\, A_2 \in \mathcal{B}(\mathbb{R}), \ldots, \forall\, A_n \in \mathcal{B}(\mathbb{R}).$$

We will often say that a sequence of random variables $(X_i)_{i\geq 0}$ is *independent and identically distributed* (iid), by which me mean all X_i have the same distribution and $(X_i)_{1\leq i\leq n}$ are independent for every $n \in \mathbb{N}$.

It is not easy to verify if two random variables X and Y are independent using Expression (2.6), since the equation must be verified for *all* Borel sets $A, B \in \mathcal{B}(\mathbb{R})$. In fact, there is an easier way to check independence.

Definition 2.9.4. The *joint distribution function* $F_{X,Y} : \mathbb{R}^2 \to [0,1]$ of two random variables X and Y defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is given by

$$F_{X,Y}(x,y) := \mathbb{P}(X \leq x, Y \leq y).$$

Again, we have two special cases for jointly discrete and jointly continuous random variables.

Definition 2.9.5. Two random variables X and Y are called *jointly discrete* if the pair (X,Y) takes values in some countable set $A = \{x_1, x_2, \ldots\} \times \{y_1, y_2, \ldots\} \subset \mathbb{R}^2$. We associate is a discrete random variable a *probability mass function* $f_{X,Y} : A \to \mathbb{R}$, defined by $f_{X,Y}(x_i, y_j) := \mathbb{P}(X = x_i, Y = y_j)$.

Definition 2.9.6. A pair of random variables X and Y is called *jointly continuous* if its joint distribution function $F_{X,Y}$ can be written as

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} \mathrm{d}u\mathrm{d}v\, f_{X,Y}(u,v), \qquad (x,y) \in \mathbb{R}^2,$$

for some $f_{X,Y} : \mathbb{R}^2 \to [0,\infty)$ called the *joint probability density function*.

As in the one-dimensional case, it may help to think of the joint density function $f_{X,Y}$ as $f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = \mathbb{P}(X \in \mathrm{d}x, Y \in \mathrm{d}y)$.

Note that for jointly continuous random variables X and Y we have $f_{X,Y}(x,y) = \partial_x \partial_y F_X(x,y)$.

If the pair (X,Y) is either jointly discrete or jointly continuous, it is easy to compute $\mathbb{P}((X,Y) \in A)$ for any $A \in \mathcal{B}(\mathbb{R}^2)$. We have

$$\begin{aligned}
\text{discrete :} &\qquad \mathbb{P}((X,Y) \in A) = \sum_{\{i,j:(x_i,y_j)\in A\}} f_{X,Y}(x_i,y_j), \\
\text{continuous :} &\qquad \mathbb{P}((X,Y) \in A) = \int_A \mathrm{d}x\mathrm{d}y\, f_{X,Y}(x,y).
\end{aligned}$$

To recover the *marginal* distribution F_X from $F_{X,Y}$, simply note that

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X \leq x, Y \leq \infty) = F_{X,Y}(x,\infty).$$

It follows that for the discrete and continuous cases, we have, respectively

$$\text{discrete :} \qquad f_X(x_i) = \sum_j f_{X,Y}(x_i, y_j),$$

$$\text{continuous :} \qquad f_X(x) = \int_{\mathbb{R}} dy\, f_{X,Y}(x, y).$$

The following theorem gives some easy-to-check conditinos for independence.

THEOREM 2.9.7. *Let X and Y be random variables definied on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. *The following conditions are equivalent (that is, if one of them holds, all of them hold)*

1. $X \perp\!\!\!\perp Y$.
2. $F_{X,Y}(x, y) = F_X(x)F_Y(y)$ *for every* $(x, y) \in \mathbb{R}^2$.
3. *Discrete case:* $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ *for every* $(x, y) \in \mathbb{R}^2$.
   Continuous case: $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ *for 'almost' every* $(x, y) \in \mathbb{R}^2$.

Together with expectation, the most important statistical properties of a random variable (or pair) are the variance and co-variance.

DEFINITION 2.9.8. The *variance* of a random variable X, written $\mathbb{V}X$ is defined by

$$\mathbb{V}X = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2,$$

whenever the expectation exists.

DEFINITION 2.9.9. The *co-variance* of two random variables X and Y, written $\mathrm{Co}\mathbb{V}[X, Y]$ is defined by

$$\mathrm{Co}\mathbb{V}[X, Y] = \mathbb{E}(X - \mathbb{E}X)(X - \mathbb{E}Y) = \mathbb{E}XY - \mathbb{E}X \cdot \mathbb{E}Y,$$

whenever the expectation exists.

Note that $\mathrm{Co}\mathbb{V}[X, X] = \mathbb{V}X$.
Note $\mathbb{V}\cdot$ is *not* a linear operator, since

$$\mathbb{V}[aX + bY] = a^2\mathbb{V}X + b^2\mathbb{V}Y + 2ab\,\mathrm{Co}\mathbb{V}[X, Y].$$

where a and b are constants.

DEFINITION 2.9.10. We say two random variables are *un-correlated* if $\mathrm{Co}\mathbb{V}[X, Y] = 0$.

Note that $X \perp\!\!\!\perp Y$ implies X and Y are uncorrelated. However, the converse is *not* true.

2.10 Conditional expectation

Let X be a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G}$ be a sub-σ algebra of $\mathcal{F}$. When $X \in \mathcal{G}$ this means that the information in $\mathcal{G}$ is sufficient to determin the value of X. When $X \perp\!\!\!\perp \mathcal{G}$, this means that the information in $\mathcal{G}$ gives us no information at all about X. Usually, however, the information in $\mathcal{G}$ gives us some information about X, but not enough to determine X exactly. And this brings us to the notion of conditioning.

Presumably, you have run across the following formula for the conditional probability of a set A given B

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \qquad \mathbb{P}(B) > 0.$$

When (X, Y) have are jointly discrete or jointly continuous, this readily leads to *conditional probability mass function*

$$\begin{aligned}
\text{discrete :} \quad & f_{X|Y}(x_i, y_j) := \mathbb{P}(X = x_i | Y = y_j) = \frac{\mathbb{P}(X = x_i \cap Y = y_j)}{\mathbb{P}(Y = y_j)} = \frac{f_{X,Y}(x_i, y_j)}{f_Y(y_j)}, \\
\text{continuous :} \quad & f_{X|Y}(x, y)\mathrm{d}x := \mathbb{P}(X \in \mathrm{d}x | Y = y) = \frac{\mathbb{P}(X \in \mathrm{d}x \cap Y \in \mathrm{d}y)}{\mathbb{P}(Y \in \mathrm{d}y)} = \frac{f_{X,Y}(x, y)}{f_Y(y)}\mathrm{d}x.
\end{aligned}$$

And from this, we can define $\mathbb{E}[X|Y = y]$, the *conditional expectation of* X *given* $Y = y$

$$\begin{aligned}
\text{discrete :} \quad & \mathbb{E}[X|Y = y_j] := \sum_i x_i \, f_{X|Y}(x_i, y_j), \\
\text{continuous :} \quad & \mathbb{E}[X|Y = y] := \int_{\mathbb{R}} \mathrm{d}x \, x \, f_{X|Y}(x, y)
\end{aligned}$$

Note that $\mathbb{E}[X|Y = y]$ is simply a function of y – there is nothing random about it.

Unfortunately, there are cases for which the pair (X, Y) are neither jointly discrete nor jointly continuous. And, for these cases we need a more general notion of conditional expectation. Here we will make two conceptual leaps:

1. We will condition with respect to a σ-algebra rather than conditioning on an event.
2. The conditional expectation will be a random variable.

We will just hop in with our new definition of conditional expectation and then we will see, through an example, that his new definition makes sense.

Definition 2.10.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$, and let X be a random variable that is either nonnegative or integrable. The conditional expectation of X given $\mathcal{G}$, denoted $\mathbb{E}[X|\mathcal{G}]$, is any random variable that satisfies

1. Measurability: $\mathbb{E}[X|\mathcal{G}] \in \mathcal{G}$.
2. Partial averaging: $\mathbb{E}[\mathbb{1}_A \mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[\mathbb{1}_A X]$ for all $A \in \mathcal{G}$.
   Alternatively, $\mathbb{E}[Z\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[ZX]$ for all $Z \in \mathcal{G}$.

When $\mathcal{G} = \sigma(Y)$ we shall often use the short-hand notation $\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)]$.

Conditional probabilities are defined from conditional expectations using

$$\mathbb{P}(A|\mathcal{G}) = \mathbb{E}[\mathbb{1}_A|\mathcal{G}].$$

Admittedly, Definition 2.10.1 is rather abstract (and, for the purposes of computation, useless). In fact, it is not at all clear from Definition 2.10.1 that $\mathbb{E}[X|\mathcal{G}]$ even exists! It does exist, though we will not prove this here.

When conditioning on the σ-algebra generated by a random variable, it is easiest to use the following formula

$$\mathbb{E}[X|Y] = \psi(Y), \qquad\qquad \psi(y) := \mathbb{E}[X|Y = y].$$

The following properties are arguably more important to remember than the definition of conditional expectation. Memorize them!

<u>Theorem</u> 2.10.2. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Conditional expectations satisfy the following properties.*

1. *Linearity:* $\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$.
2. *Taking out what is known: if* $X \in \mathcal{G}$ *then* $\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$.
3. *Iterated conditioning: if* $\mathcal{H}$ *is a sub-σ-algebra of* $\mathcal{G}$ *then* $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$.
4. *Independence: if* $X \perp\!\!\!\perp \mathcal{G}$ *then* $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}X$.

Theorem 2.10.2 can be proved directly from Definition 2.10.1, though we will not do so here.

<u>Definition</u> 2.10.3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number, and let $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ be a filtration of sub-σ-algebras of $\mathcal{F}$. Consider an $\mathbb{F}$-adapted stochastic process $M = (M_t)_{t\in[0,T]}$. We say that M is a *martingale* if

$$\mathbb{E}[M_t|\mathcal{F}_s] = M_s, \qquad\qquad \forall\, 0 \le s \le t \le T.$$

We have given above the definition of a *continuous time* martingale. We can also define *discrete-time* martingales by making the obvious modifications. Note: when we say that a process M is a martingale this is with respect to a fixed probability measure and filtration. If $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ are two probability measures and $\mathbb{F}$ and $\mathbb{G}$ are two filtrations, it is entirely possible that a process M may be a martingale with respect to $(\mathbb{P}, \mathbb{F})$ and may not be a martingale with respect to $(\widetilde{\mathbb{P}}, \mathbb{F})$, $(\mathbb{P}, \mathbb{G})$ or $(\widetilde{\mathbb{P}}, \mathbb{G})$.

DEFINITION 2.10.4. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number, and let $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]}$ be a filtration of sub-σ-algebras of $\mathcal{F}$. Consider an $\mathbb{F}$-adapted stochastic process $X = (X_t)_{t\in[0,T]}$. Assume that for all $0 \le s \le t \le T$ and for every nonnegative, Borel-measurable function f, there is another Borel-measurable function g (which depends on s, t, and f) such that

$$\mathbb{E}[f(X_t)|\mathcal{F}_s] = g(X_s).$$

Then we say X is a *Markov process* or simply "X is Markov."

Identifying $g(X_s) \equiv \mathbb{E}[f(X_t)|X_s]$ we can write the Markov property as follows

$$\mathbb{E}[f(X_t)|\mathcal{F}_s] = \mathbb{E}[f(X_t)|X_s]$$

A Markov process is a process for which the following holds: *given the present (i.e., X_s), the future (i.e, X_t, $t \ge s$) is independent of the past (i.e, $\mathcal{F}_s$).* What this means in practice is that

$$\mathbb{P}(X_t \in A|\mathcal{F}_s) = \mathbb{P}(X_t \in A|X_s), \qquad \forall s \le t, \qquad \forall\, A \in \mathcal{B}(\mathbb{R}). \qquad \text{(if X is Markov)}$$

If X_t is a discrete or continuous random variable for every t then we have a *transition kernel*, written as P in the discrete case and Γ in the continuous case.

$$\begin{aligned} \text{discrete}: &\qquad P(s,x;t,y) := \mathbb{P}(X_t = y|X_s = x), \\ \text{continuous}: &\qquad \Gamma(s,x;t,y)\mathrm{d}y = \mathbb{P}(X_t \in \mathrm{d}y|X_s = x). \end{aligned}$$

If you can write the transition kernel of a process explicitly, then you have essentially proved that the process is Markov.

Note that *any process* that has independent increments is Markov since, if $X_t - X_s \perp\!\!\!\perp X_s$ for $t \ge s$, then

$$\mathbb{P}(X_t \in A|\mathcal{F}_s) = \mathbb{P}(X_t - X_s + X_s \in A|X_s),$$

Where $\mathcal{F}_s$ is the filtration generated by observing X up to time s.

Markov processes and Martingales are *entirely separate* concepts. A process X can be both a martingale and a Markov process, it can be a martingale but not a Markov process, it can be a Markov process but not a martingale, and it can be neither a Markov process nor a martingale. We illustrate the difference with an example.

EXAMPLE 2.10.5. Let us return to the stock price Example 2.4.2. Let us show that $S = (S_n)_{0\le n}$ is a Markov process. Recall that $\mathcal{F}_m$ is the σ-algebra generated by observing $\omega_1, \omega_2, \ldots, \omega_m$. Observe that $S_m \in \mathcal{F}_m$. Next, note that

$$\mathbb{P}(S_{n+m} = S_m u^k d^{n-k}|S_m) = \binom{n}{k} p^k q^{n-k}.$$

Since we have written the transition kernel explicitly, we have established that S is Markov. Let us also find the function g in Definition 2.10.4. For any $f : \mathbb{R} \to \mathbb{R}$ we have

$$\mathbb{E}[f(S_{n+m})|\mathcal{F}_m] = \sum_{k=0}^{n} f(S_m u^k d^{n-k}) \cdot \binom{n}{k} p^k q^{n-q} =: g(S_m).$$

Thus, we have found g. Now, to see if S is a martingale note that

$$\mathbb{E}[S_{n+1}|\mathcal{F}_n] = \mathbb{E}[S_{n+1}|S_n] = p \cdot uS_n + q \cdot dS_n = (p \cdot u + q \cdot d)S_n.$$

Thus, if $(p \cdot u + q \cdot d) = 1$, then $\mathbb{E}[S_{n+1}|\mathcal{F}_n] = S_n$. Let us assume that $(p \cdot u + q \cdot d) = 1$. Then we have

$$\begin{aligned}
\mathbb{E}[S_{n+m}|\mathcal{F}_n] &= \mathbb{E}[\mathbb{E}[S_{n+m}|\mathcal{F}_{n+m-1}]|\mathcal{F}_n] = \mathbb{E}[S_{n+m-1}|\mathcal{F}_n] \\
&= \mathbb{E}[\mathbb{E}[S_{n+m-1}|\mathcal{F}_{n+m-2}]|\mathcal{F}_n] = \mathbb{E}[S_{n+m-2}|\mathcal{F}_n] \\
&= \dots \\
&= \mathbb{E}[\mathbb{E}[S_{n+1}|\mathcal{F}_n]|\mathcal{F}_n] = \mathbb{E}[S_n|\mathcal{F}_n] = S_n,
\end{aligned}$$

Therefore, the process S is a martingale.

2.11 Exercises

Exercise 2.1. Let $\mathcal{F}$ be a σ-algebra of Ω. Suppose $B \in \mathcal{F}$. Show that $\mathcal{G} := \{A \cap B : A \in \mathcal{F}\}$ is a σ-algebra of B.

Exercise 2.2. Let $\mathcal{F}$ and $\mathcal{G}$ be σ-algebras of Ω. (a) Show that $\mathcal{F} \cap \mathcal{G}$ is a σ-algebra of Ω. (b) Show that $\mathcal{F} \cup \mathcal{G}$ is not necessarily a σ-algebra of Ω.

Exercise 2.3. Describe the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for the following three experiments: (a) a biased coin is tossed three times; (b) two balls are drawn without replacement from an urn which originally contained two blue and two red balls; (c) a biased coin is tossed repeatedly until a head turns up.

Exercise 2.4. Suppose X is a continuous random variable with distribution F_X. Let g be a strictly increasing continuous function. Define $Y = g(X)$. (a) What if F_Y, the distribution of Y? (b) What is f_Y, the density of Y?

Exercise 2.5. Suppose X is a continuous random variable with distribution F_X. Find F_Y where Y is given by (a) X^2 (b) $\sqrt{|X|}$ (c) $\sin X$ (d) $F_X(X)$.

Exercise 2.6. Suppose X is a continuous random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let f be the density of X under $\mathbb{P}$ and assume $f > 0$. Let g be the density function of a random variable. Define $Z := g(X)/f(X)$. (a) Show that $Z \equiv d\widetilde{\mathbb{P}}/d\mathbb{P}$ defines a Radon-Nikodým derivative. (b) What is the density of X under $\widetilde{\mathbb{P}}$?

<u>Exercise</u> 2.7. Let X be uniformly distributed on $[0,1]$. For what function g is the random variable $g(X)$ exponentially distributed with parameter 1 (i.e. $g(X) \sim \mathcal{E}(1)$)?

<u>Exercise</u> 2.8. Let $\Omega = \{a, b, c, d\}$ and let $\mathcal{F} = 2^{\Omega}$ (the set of all subsets of Ω). We define a probability measure $\mathbb{P}$ as follows

$$\mathbb{P}(a) = 1/6, \qquad \mathbb{P}(b) = 1/3, \qquad \mathbb{P}(c) = 1/4, \qquad \mathbb{P}(d) = 1/4,$$

Next, define three random variables

$$\begin{aligned} &X(a) = 1, && X(b) = 1, && X(c) = -1, && X(d) = -1, \\ &Y(a) = 1, && Y(b) = -1, && Y(c) = 1, && Y(d) = -1, \end{aligned}$$

and $Z = X + Y$. (a) List the sets in $\sigma(X)$. (b) What are the values of $\mathbb{E}[Y|X]$ for $\{a, b, c, d\}$? Verify the partial averaging property: $\mathbb{E}[\mathbb{1}_A \mathbb{E}[Y|X]] = \mathbb{E}[\mathbb{1}_A Y]$ for all $A \in \sigma(X)$. (c) What are the values of $\mathbb{E}[Z|X]$ for $\{a, b, c, d\}$? Verify the partial averaging property.

<u>Exercise</u> 2.9. Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let Y be a square integrable random variable: $\mathbb{E}Y^2 < \infty$ and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Show that

$$\mathbb{V}(Y - \mathbb{E}[Y|\mathcal{G}]) \leq \mathbb{V}(Y - X), \qquad \forall X \in \mathcal{G}.$$

<u>Exercise</u> 2.10. Give an example of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a random variable X and a function f such that $\sigma(f(X))$ is strictly smaller than $\sigma(X)$ but $\sigma(f(X)) \neq \{\emptyset, \Omega\}$. Give a function g such that $\sigma(g(X)) = \{\emptyset, \Omega\}$.

<u>Exercise</u> 2.11. On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ define random variables X and $Y_0, Y_1, Y_2, \ldots$ and suppose $\mathbb{E}|X| < \infty$. Define $\mathcal{F}_n := \sigma(Y_0, Y_1, \ldots, Y_n)$ and $X_n = \mathbb{E}[X|\mathcal{F}_n]$. Show that the sequence $X_0, X_1, X_2, \ldots$ is a martingale under $\mathbb{P}$ with respect to the filtration $(\mathcal{F}_n)_{n \geq 0}$.

<u>Exercise</u> 2.12. Let $X_0, X_1, \ldots$ be i.i.d Bernoulli random variables with parameter p (i.e., $\mathbb{P}(X_i = 1) = p$). Define $S_n = \sum_{i=1}^{n} X_i$ where $S_0 = 0$. Define

$$Z_n := \left(\frac{1-p}{p}\right)^{2S_n - n}, \qquad n = 0, 1, 2, \ldots.$$

Let $\mathcal{F}_n := \sigma(X_0, X_1, \ldots, X_n)$. Show that Z_n is a martingale with respect to this filtration.

Chapter 3

Brownian motion and Stochastic Calculus

3.1 Brownian motion

<u>Definition</u> 3.1.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A *Brownian motion* is a stochastic process $W = (W_t)_{t\geq 0}$ that satisfies:

1. $W_0 = 0$.
2. If $0 \leq r < s < t < u < \infty$ then $(W_u - W_t) \perp\!\!\!\perp (W_s - W_r)$.
3. If $0 \leq r < s$ then $W_s - W_r \sim \mathcal{N}(0, s-r)$.
4. The map $t \to W_t$ is continuous for every ω.

<u>Definition</u> 3.1.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be on probability space on which a Brownian motion $W = (W_t)_{t\geq 0}$ is defined. A *filtration for the Brownian motion* W is a collection of σ-algebras $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying:

1. Information accumulates: if $0 \leq s < t$ then $\mathcal{F}_s \subset \mathcal{F}_t$.
2. Adaptivity: for all $t \geq 0$, we have $W_t \in \mathcal{F}_t$.
3. Independence of future increments: if $u > t \geq 0$ then $(W_u - W_t) \perp\!\!\!\perp \mathcal{F}_t$.

The most natural choice for this filtration $\mathbb{F}$ is the *natural filtration* for W. That is $\mathcal{F}_t = \sigma(W_u, 0 \leq u \leq t)$. In principle the filtration $(\mathcal{F}_t)_{t\geq 0}$ could contain more than the information obtained by observing W. However, the information in the filtration is not allowed to destroy the independence of future increments of Brownian motion.

Not surprisingly, if $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is a filtration for a Brownian motion W then W is a martingale with respect to this filtration. We see this, let $0 \leq s < t$ and observe that

$$\mathbb{E}[W_t|\mathcal{F}_s] = \mathbb{E}[W_t - W_s + W_s|\mathcal{F}_s] = \mathbb{E}[W_t - W_s|\mathcal{F}_s] + \mathbb{E}[W_s|\mathcal{F}_s] = \mathbb{E}[W_t - W_s] + W_s = W_s.$$

3.2 Quadratic variation

In this Section we will define what we mean by "quadratic variation" and we will compute this quantity for a Brownian motion W.

Definition 3.2.1. Let $f : [0, \mathrm{T}] \to \mathbb{R}$. We define the *quadratic variation of f up to time* T, denoted $[f, f]_\mathrm{T}$ as

$$[f, f]_\mathrm{T} := \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \Big[f(t_{j+1}) - f(t_j) \Big]^2 ,$$

where Π and $\|\Pi\|$ are as defined as follows

$$\Pi = \{t_0, t_1, \ldots, t_n\}, \qquad 0 = t_0 < t_1 < \ldots < t_n = \mathrm{T}, \qquad \|\Pi\| = \max_i (t_{i+1} - t_i). \tag{3.1}$$

Theorem 3.2.2. *Let* W *be a Brownian motion. Then, for all* $\mathrm{T} \geq 0$ *we have* $[\mathrm{W}, \mathrm{W}]_\mathrm{T} = \mathrm{T}$ *almost surely.*

The above Theorem can roughly be understood as follows. Suppose $\mathrm{d}t \ll 1$ and define $\mathrm{dW}_t := \mathrm{W}_{t+\mathrm{d}t} - \mathrm{W}_t$. Because $\mathrm{dW}_t \sim \mathcal{N}(0, \mathrm{d}t)$ we have $\mathbb{E}((\mathrm{dW}_t)^2) = \mathrm{d}t$ and $\mathbb{V}((\mathrm{dW}_t)^2) = 2\mathrm{d}t^2$. As $\mathrm{d}t^2$ is practically zero for $\mathrm{d}t \ll 1$, one can imagine that $(\mathrm{dW}_t)^2$ is *almost* equal to a constant $\mathrm{d}t$. *Informally*, we write this as

$$\mathrm{dW}_t \mathrm{dW}_t = \mathrm{d}t. \tag{3.2}$$

This informal statement, while not rigorously correct, captures the spirit of the quadratic variation computation for W.

Definition 3.2.3. Let $f, g : [0, \mathrm{T}] \to \mathbb{R}$. We define the *covaration of f and g up to time* T, denoted $[f, g]_\mathrm{T}$ as

$$[f, g]_\mathrm{T} := \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \Big[f(t_{j+1}) - f(t_j) \Big] \Big[g(t_{j+1}) - g(t_j) \Big] ,$$

where Π and $\|\Pi\|$ are as defined in (3.1).

Theorem 3.2.4. *Let* W *be a Brownian motion and let* Id *be the identity function:* $\mathrm{Id}(t) = t$. *Then, for all* $\mathrm{T} \geq 0$ *we have* $[\mathrm{W}, \mathrm{Id}]_\mathrm{T} = 0$ *almost surely and* $[\mathrm{Id}, \mathrm{Id}]_\mathrm{T} = 0$.

Just as (3.2) captures the spirit of the computation of $[\mathrm{W}, \mathrm{W}]_\mathrm{T}$, the following equations

$$\mathrm{dW}_t \mathrm{d}t = 0, \qquad \mathrm{d}t \mathrm{d}t = 0,$$

informally capture the spirit of the $[\mathrm{W}, \mathrm{Id}]_\mathrm{T}$ and $[\mathrm{Id}, \mathrm{Id}]_\mathrm{T}$ computations.

3.3 Markov property of Brownian motion

Theorem 3.3.1. *Let* $W = (W_t)_{t \geq 0}$ *be a Brownian motion and let* $(\mathcal{F}_t)_{t \geq 0}$ *be a filtration for this Brownian motion. Then* W *is a Markov process.*

Proof. According to Definition 2.10.4, we must show that there exists a function g such that

$$\mathbb{E}[f(W_T)|\mathcal{F}_t] = g(W_t),$$

where $T \geq t \geq 0$. Noting that $W_t \in \mathcal{F}_t$ and $W_T - W_t \perp\!\!\!\perp \mathcal{F}_t$, we have

$$\mathbb{E}[f(W_T)|\mathcal{F}_t] = \mathbb{E}[f(W_T - W_t + W_t)|W_t] = \int_{\mathbb{R}} dy\, f(y + W_t)\Gamma(t, 0; T, y) =: g(W_t),$$

where $\Gamma(t, x; T, \cdot)$ is the density of a normal random variable with mean x and variance $T - t$. □

3.4 Itô Integrals

Assumption 3.4.1. In what follows $W = (W_t)_{t \geq 0}$ will always represent a Brownian motion and $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ will always be a filtration for this Brownian motion. We shall assume the integrand $\Delta = (\Delta_t)_{t \geq 0}$ is adapted to $\mathbb{F}$, meaning $\Delta_t \in \mathcal{F}_t$ for all t.

Note that the process Δ *can* and, in many cases, *will* be random. However, the information available in $\mathcal{F}_t$ will always be sufficient to determine the value of Δ_t at time t. Also note, since $(W_T - W_t) \perp\!\!\!\perp \mathcal{F}_t$ for $T > t$, it follows that $(W_T - W_t) \perp\!\!\!\perp \Delta_t$. In other words, future increments of Brownian motion are independent of the Δ process.

Itô integrals for simple integrands

To begin let us assume that Δ is a *simple process*, meaning Δ is of the form

$$\Delta_t = \sum_{j=0}^{n-1} \Delta_{t_j} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}}, \qquad 0 = t_0 < t_1 < \dots t_n = T, \qquad \Delta_{t_j} \in \mathcal{F}_{t_j}.$$

Since the process Δ is constant over intervals of the form $[t_j, t_{j+1})$, it makes sense to define

$$I_T = \int_0^T \Delta_t dW_t := \sum_{j=0}^{n-1} \Delta_{t_j}(W_{t_{j+1}} - W_{t_j}). \qquad \text{(for } \Delta \text{ a simple process)}$$

Itô integrals for general integrands

Clearly, it is rather restrictive to limit ourselves to integrands Δ that are simple processes. We now allow the process Δ to be any process that is adapted to $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ and which satisfies the following integrability condition

$$\mathbb{E} \int_0^{\mathrm{T}} \Delta_t^2 \mathrm{d}t < \infty. \tag{3.3}$$

To construct an Itô integral with Δ as the integrand, we first approximate Δ by a simple process

$$\Delta_t \approx \Delta_t^{(n)} := \sum_{j=0}^{n-1} \Delta_{t_j} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}}, \qquad 0 \leq t_0 < t_1 < \ldots < t_n = \mathrm{T}.$$

As $n \to \infty$ the process $\Delta^{(n)}$ converges to Δ in the sense that

$$\lim_{n \to \infty} \mathbb{E} \int_0^{\mathrm{T}} \left(\Delta_t - \Delta_t^{(n)} \right)^2 \mathrm{d}t = 0. \tag{3.4}$$

We now *define* the Itô integral for a general integrand Δ by

$$\mathrm{I}_{\mathrm{T}} \equiv \int_0^{\mathrm{T}} \Delta_t \mathrm{dW}_t := \lim_{n \to \infty} \int_0^{\mathrm{T}} \Delta_t^{(n)} \mathrm{dW}_t. \tag{3.5}$$

Note that the integrals $\int_0^{\mathrm{T}} \Delta_t^{(n)} \mathrm{d}t$ are well-defined for every n, since $\Delta^{(n)}$ is a simple process. Furthermore, the condition (3.4) ensures that the limit exists in $\mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. The following Theorem lists some important properties of Itô integrals.

<u>Theorem</u> 3.4.2. *Let* W *be a Brownian motion and let* $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ *be a filtration for this Brownian motion. Let* $\Delta = (\Delta_t)_{0 \leq t \leq \mathrm{T}}$ *be adapted to the filtration* $\mathbb{F}$ *and satisfy* (3.4). *Let* $\mathrm{I} = (\mathrm{I}_t)_{0 \leq t \leq \mathrm{T}}$ *be given by* $\mathrm{I}_t = \int_0^t \Delta_s \mathrm{dW}_s$*, where the integral is defined as in* (3.5). *Then the process* I *has the following properties.*

1. *The sample paths of* I *are continuous.*
2. *The process* I *is adapted to the filtration* $\mathbb{F}$. *That is,* $\mathrm{I}_t \in \mathcal{F}_t$ *for all* t.
3. *If* $\Gamma = (\Gamma_t)_{0 \leq t \leq \mathrm{T}}$ *satisfies the same conditions as* Δ*, then*

$$\int_0^{\mathrm{T}} (a\Delta_t + b\Gamma_t) \mathrm{dW}_t = a \int_0^{\mathrm{T}} \Delta_t \mathrm{dW}_t + b \int_0^{\mathrm{T}} \Gamma_t \mathrm{dW}_t,$$

   where a *and* b *are constants.*
4. *The process* I *is a martingale with respect to the filtration* $\mathbb{F}$.
5. *We have the Itô isometry* $\mathbb{E}\mathrm{I}_t^2 = \mathbb{E} \int_0^{\mathrm{T}} \Delta_t^2 \mathrm{d}t$.
6. *The quadratic variation of* I *is given by* $[\mathrm{I}, \mathrm{I}]_{\mathrm{T}} = \int_0^{\mathrm{T}} \Delta_t^2 \mathrm{d}t$.

3.5 Itô processes and the Itô formula

Definition 3.5.1. Let $W = (W_t)_{t\geq 0}$ be a Brownian motion and let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be a filtration for this Brownian motion. An *Itô process* is any process $X = (X_t)_{t\geq 0}$ of the form

$$X_t = X_0 + \int_0^t \Theta_s ds + \int_0^t \Delta_s dW_s, \tag{3.6}$$

where $\Theta = (\Theta_t)_{t\geq 0}$ and $\Delta = (\Delta_t)_{t\geq 0}$ are adapted to the filtration $\mathbb{F}$ and satisfy

$$\int_0^T |\Theta_t| dt < \infty, \qquad \mathbb{E}\int_0^T \Delta_t^2 dt < \infty, \qquad \forall\, T \geq 0,$$

and X_0 is not random.

We sometimes write an Itô process in *differential form*

$$dX_t = \Theta_t dt + \Delta_t dW_t. \tag{3.7}$$

Expression (3.7) literally means that X satisfies (3.6). Informally, the differential form can be understood as follows: in a small interval of time δt , the process X changes according to

$$X_{t+\delta t} - X_t \approx \Theta_t \delta t + \Delta_t \left(W_{t+\delta t} - W_t\right). \tag{3.8}$$

In fact, noting that $W_{t+\delta t} - W_t \sim \mathcal{N}(0, \delta t)$ and $W_{t+\delta t} - W_t \perp\!\!\!\perp \mathcal{F}_t$, one can use expression (3.8) to simulate the increment $X_{t+\delta t} - X_t$. This way of simulating X is called the *Euler scheme* and is the workhorse of many Monte Carlo methods.

Lemma 3.5.2. *The quadratic variation* $[X, X]_T$ *of an Itô process* (3.6) *is given by*

$$[X, X]_T = \int_0^T \Delta_t^2 dt.$$

Definition 3.5.3. Let $X = (X_t)_{t\geq 0}$ be an Itô process, as described in Definition 3.5.1. Let $\Gamma = (\Gamma_t)_{t\geq 0}$ be adapted to the filtration of the Brownian motion $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$. We define

$$\int_0^T \Gamma_t dX_t := \int_0^T \Gamma_t \Theta_t dt + \int_0^T \Gamma_t \Delta_t dW_t,$$

where we assume

$$\int_0^T |\Gamma_t \Theta_t| dt < \infty, \qquad \mathbb{E}\int_0^T (\Gamma_t \Delta_t)^2\, dt < \infty, \qquad \forall\, T \geq 0.$$

Theorem 3.5.4 (Itô formula in one dimension). *Let* $X = (X_t)_{t\geq 0}$ *be an Itô process and suppose* $f : \mathbb{R} \to \mathbb{R}$ *satisfies* $f \in C^2(\mathbb{R})$. *Then, for any* $T \geq 0$ *we have*

$$f(X_T) - f(X_0) = \int_0^T f'(X_t) dX_t + \tfrac{1}{2}\int_0^T f''(X_t) d[X, X]_t.$$

In differential form, with X given by (3.7), Itô's formula becomes

$$\begin{aligned} \mathrm{d}f(\mathrm{X}_t) &= f'(\mathrm{X}_t)\mathrm{d}\mathrm{X}_t + \tfrac{1}{2}f''(\mathrm{X}_t)\mathrm{d}[\mathrm{X},\mathrm{X}]_t \\ &= f'(\mathrm{X}_t)\left(\Theta_t \mathrm{d}t + \Delta_t \mathrm{d}\mathrm{W}_t\right) + \tfrac{1}{2}f''(\mathrm{X}_t)\Delta_t^2 \mathrm{d}t, \end{aligned} \tag{3.9}$$

where we have used $\mathrm{d}[\mathrm{X},\mathrm{X}]_t = \Delta_t^2 \mathrm{d}t$. Perhaps the easiest way to remember (3.9) is to use the following two-step procedure:

1. Expand $f(\mathrm{X}_t + \mathrm{d}\mathrm{X}_t) - f(\mathrm{X}_t)$ to second order about the point X_t

$$\mathrm{d}f(\mathrm{X}_t) = f(\mathrm{X}_t + \mathrm{d}\mathrm{X}_t) - f(\mathrm{X}_t) = f'(\mathrm{X}_t)\mathrm{d}\mathrm{X}_t + \tfrac{1}{2}f''(\mathrm{X}_t)(\mathrm{d}\mathrm{X}_t)^2, \tag{3.10}$$

2. Insert the differential $\mathrm{d}\mathrm{X}_t = \Theta_t \mathrm{d}t + \Delta_t \mathrm{d}\mathrm{W}_t$ into (3.10), expand $(\mathrm{d}\mathrm{X}_t)^2$ and use the rules

$$\mathrm{d}\mathrm{W}_t \mathrm{d}\mathrm{W}_t = \mathrm{d}t, \qquad \mathrm{d}\mathrm{W}_t \mathrm{d}t = 0, \qquad \mathrm{d}t\mathrm{d}t = 0.$$

The resulting formula gives the correct expression for $\mathrm{d}f(\mathrm{X}_t)$.

Example 3.5.5. Let X be an Itô process with the following dynamics

$$\mathrm{d}\mathrm{X}_t = \mu_t \mathrm{X}_t \mathrm{d}t + \sigma_t \mathrm{X}_t \mathrm{d}\mathrm{W}_t. \tag{3.11}$$

Assuming $\mu = (\mu_t)_{t\geq 0}$ and $\sigma = (\sigma_t)_{t \geq 0}$ are bounded above and below and $\mathrm{X}_0 > 0$, the process X remains strictly positive. We call X a *generalized geometric Brownian motion*. The "geometric" part refers to the fact that the relative step size $\mathrm{d}\mathrm{X}_t/\mathrm{X}_t$ has dyanmics $\mu_t \mathrm{d}t + \sigma_t \mathrm{d}\mathrm{W}_t$. The "generalized" part refers to the fact that the processes σ and μ are stochastic rather than constant. Define $\mathrm{Y}_t = \mathrm{X}_t^p$. What is $\mathrm{d}\mathrm{Y}_t$? Let $f(x) = x^p$. Then $f'(x) = px^{p-1}$ and $f''(x) = p(p-1)x^{p-2}$. Thus, we have

$$\begin{aligned} \mathrm{d}\mathrm{Y}_t = \mathrm{d}f(\mathrm{X}_t) &= p\mathrm{X}_t^{p-1}\mathrm{d}\mathrm{X}_t + \tfrac{1}{2}p(p-1)\mathrm{X}_t^{p-2}(\mathrm{d}\mathrm{X}_t)^2 \\ &= p\mathrm{X}_t^{p-1}\left(\mu_t \mathrm{X}_t \mathrm{d}t + \sigma_t \mathrm{X}_t \mathrm{d}\mathrm{W}_t\right) + \tfrac{1}{2}p(p-1)\mathrm{X}_t^{p-2}\left(\mu_t \mathrm{X}_t \mathrm{d}t + \sigma_t \mathrm{X}_t \mathrm{d}\mathrm{W}_t\right)^2 \\ &= p\mathrm{X}_t^{p-1}\left(\mu_t \mathrm{X}_t \mathrm{d}t + \sigma_t \mathrm{X}_t \mathrm{d}\mathrm{W}_t\right) + \tfrac{1}{2}p(p-1)\mathrm{X}_t^{p-2}\sigma_t^2 \mathrm{X}_t^2 \mathrm{d}t \\ &= \left(p\mu_t + \tfrac{1}{2}p(p-1)\sigma_t^2\right)\mathrm{X}_t^p \mathrm{d}t + p\sigma_t \mathrm{X}_t^p \mathrm{d}\mathrm{W}_t \\ &= \left(p\mu_t + \tfrac{1}{2}p(p-1)\sigma_t^2\right)\mathrm{Y}_t \mathrm{d}t + p\sigma_t \mathrm{Y}_t \mathrm{d}\mathrm{W}_t. \end{aligned}$$

We see from the last line that $\mathrm{Y} = (\mathrm{Y}_t)_{t\geq 0}$ is also a generalized geometric Brownian motion.

Example 3.5.6. Let X have generalized geometric Brownian motion dynamics as in (3.11). We would like to find an explicit expression for X_t (i.e., an expression of the form $\mathrm{X}_t = \ldots$ where $\ldots$ does not contain X). To this end,we let $Y_t = \log X_t$. With $f(x) = \log x$ we have $f'(x) = 1/x$ and $f''(x) = -1/x^2$. Thus, we have

$$dY_t = \frac{1}{X_t} dX_t + \frac{1}{2}\frac{-1}{X_t^2}(dX_t)^2 = \left(\mu_t - \tfrac{1}{2}\sigma_t^2\right) dt + \sigma_t dW_t.$$

Thus, we have

$$\begin{aligned} X_T = \exp(Y_T) &= \exp\left(Y_0 + \int_0^T \left(\mu_t - \tfrac{1}{2}\sigma_t^2\right) dt + \int_0^T \sigma_t dW_t\right) \\ &= X_0 \exp\left(\int_0^T \left(\mu_t - \tfrac{1}{2}\sigma_t^2\right) dt + \int_0^T \sigma_t dW_t\right), \end{aligned}$$

where we have used $Y_0 = \log X_0$.

PROPOSITION 3.5.7. *Let* $W = (W_t)_{t\geq 0}$ *be a Brownian motion. Suppose* $g : \mathbb{R}_+ \to \mathbb{R}_+$ *is a deterministic function. Then*

$$I_T := \int_0^T g(t) dW_t \sim \mathcal{N}(0, v(T)), \qquad v(T) = \int_0^T g^2(t) dt.$$

3.6 Multivariate stochastic calculus

DEFINITION 3.6.1. A *d-dimensional Brownian motion* is a process

$$W = (W_t^1, W_t^2, \dots, W_t^d)_{t\geq 0}$$

with the the following properties.

1. Each $W^i = (W_t^i)_{t\geq 0}$, $i = 1, 2, \dots d$, is a one-dimensional Brownian motion.
2. The processes $(W^i)_{1\leq i\leq d}$ are independent.

A *filtration for* W is a collection of σ-algebras $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ such that

1. Information accumulates: $\mathcal{F}_s \subset \mathcal{F}_t$ for all $0 \leq s < t$.
2. Adaptivity: $W_t \in \mathcal{F}_t$ for all $t \geq 0$.
3. Independent increments: for $0 \leq s < t$ we have $W_t - W_s \perp\!\!\!\perp \mathcal{F}_s$.

THEOREM 3.6.2. *Let* $X^i = (X_t^i)_{t\geq 0}$ $i = 1, 2, \dots, n$ *be the Itô processes given by*

$$dX_t^i = \Theta_t^i dt + \sum_{j=1}^d \sigma_t^{ij} dW_t^j, \qquad i = 1, 2, \dots, n, \tag{3.12}$$

where $W = (W_t^1, W_t^2, \dots, W_t^d)_{t\geq 0}$ *is a d-dimensional Brownian motion. Then*

$$d[X^i, X^j]_t = \sum_{k=1}^d \sigma_t^{ik} \sigma_t^{jk} dt.$$

We will not prove Theorem 3.6.2. Rather, we simply remark that it can be obtained *informally* by writing

$$\mathrm{d}[\mathrm{X}^i, \mathrm{X}^j]_t = \mathrm{d}\mathrm{X}^i_t \mathrm{d}\mathrm{X}^j_t, \tag{3.13}$$

inserting expression (3.12) into (3.13) and using the multiplication rules

$$\mathrm{d}\mathrm{W}^i_t \mathrm{d}\mathrm{W}^j_t = \delta_{ij}\mathrm{d}t, \qquad \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \qquad \mathrm{d}\mathrm{W}^j_t \mathrm{d}t = 0, \qquad \mathrm{d}t\mathrm{d}t = 0. \tag{3.14}$$

Note that $\mathrm{d}[\mathrm{X}^i, \mathrm{X}^j]_t = 0$ unless X^j and X^j are driven by at least one common one-dimensional Brownian motion.

We can now give a n-dimensional version of Itô's Lemma. We present the formula in differential form, as it is written more compactly in this way.

<u>Theorem</u> 3.6.3 (Itô formula in two dimensions). *Let* $\mathrm{X} = (\mathrm{X}^1_t, \mathrm{X}^2_t, \ldots, \mathrm{X}^n_t)_{t \geq 0}$ *be an* n*-dimensional Itô process and suppose* $f : \mathbb{R}^n \to \mathbb{R}$ *satisfies* $f \in \mathrm{C}^2(\mathbb{R}^n)$*. Then, for any* $\mathrm{T} \geq 0$ *we have*

$$\mathrm{d}f(\mathrm{X}_t) = \sum_{i=1}^n \frac{\partial f(\mathrm{X}_t)}{\partial x_i}\mathrm{d}\mathrm{X}^i_t + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(\mathrm{X}_t)}{\partial x_i \partial x_j}\mathrm{d}[\mathrm{X}^i, \mathrm{X}^j]_t.$$

The proof of Theorem 3.6.3 is a straightforward extension of Theorem 3.5.4 to the n-dimensional case and will not be presented here.

To obtain an explicit expression for $\mathrm{d}f(\mathrm{X}_t)$ in terms of $\mathrm{d}\mathrm{W}^1_t, \mathrm{d}\mathrm{W}^2_t, \ldots, \mathrm{d}\mathrm{W}^d_t$ and $\mathrm{d}t$ we can repeat the same informal procedure we used in the one-dimensional case.

1. Expand $\mathrm{d}f(\mathrm{X}_t) = f(\mathrm{X}_t + \mathrm{d}\mathrm{X}_t) - f(\mathrm{X}_t)$ about the point X_t to second order

$$\mathrm{d}f(\mathrm{X}_t) = \sum_{i=1}^n \frac{\partial f(\mathrm{X}_t)}{\partial x_i}\mathrm{d}\mathrm{X}^i_t + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(\mathrm{X}_t)}{\partial x_i \partial x_j}\mathrm{d}\mathrm{X}^i_t \mathrm{d}\mathrm{X}^j_t. \tag{3.15}$$

2. Insert expression for $\mathrm{d}\mathrm{X}^i_t$ into (3.15) and use the multiplication rules given in (3.14).

<u>Example</u> 3.6.4 (Product rule). To compute $\mathrm{d}(\mathrm{X}_t\mathrm{Y}_t)$ wherer X and Y are one-dimensional Itô processes, we define $f(x, y) = xy$ and use $f_x = y$, $f_y = x$, $f_{xy} = 1$ and $f_{xx} = f_{yy} = 0$ to compute

$$\mathrm{d}(\mathrm{X}_t\mathrm{Y}_t) = \mathrm{Y}_t\mathrm{d}\mathrm{X}_t + \mathrm{X}_t\mathrm{d}\mathrm{Y}_t + \mathrm{d}[\mathrm{X}, \mathrm{Y}]_t.$$

<u>Example</u> 3.6.5 (OU process). An *Ornstein-Uhlenbeck process* (OU process, for short) is an Itô process $\mathrm{X} = (\mathrm{X}_t)_{t \geq 0}$ that satisfies

$$\mathrm{d}\mathrm{X}_t = \kappa(\theta - \mathrm{X}_t)\mathrm{d}t + \sigma\mathrm{d}\mathrm{W}_t, \tag{3.16}$$

where $W = (W_t)_{t \geq 0}$ is a one-dimensional Brownian motion and $\kappa, \theta > 0$. The OU process is *mean-reverting* in the following sense. If $X_t > \theta$ then $\kappa(\theta - X_t) < 0$ and the deterministic part of (3.16) (i.e., the dt-term) pushes the process down towards θ. If $X_t < \theta$ then $\kappa(\theta - X_t) > 0$ and the deterministic part of (3.16) pushes the process up towards θ. The OU process mean-reverts to the *long-run mean* θ. We often call κ the *rate of mean reversion*, though this nomenclature is somewhat misleading since the instantaneous rate of mean reversion is actually $\kappa(\theta - X_t)$.

We will find an explicit expression for X_t and also compute $\mathbb{E}X_t$ and $\mathbb{V}X_t$. To this end, let us define $Y_t = X_t - \theta$ so that

$$dY_t = -\kappa Y_t dt + \sigma dW_t$$

Note that Y is an OU process that mean-reverts to zero. Next, we define $Z_t = f(t, Y_t) = e^{\kappa t} Y_t$. We can use the two-dimensional Itô formula to compute dZ_t. Using $f_{yy} = 0$ and the heuristic rules $dt dW_t = 0$ and $dt dt = 0$ we have

$$\begin{aligned} dZ_t &= f_t dt + f_y dY_t + \tfrac{1}{2} f_{yy} d[Y, Y]_t \\ &= \kappa e^{\kappa t} Y_t dt + e^{\kappa t} dY_t \\ &= \kappa e^{\kappa t} Y_t dt + e^{\kappa t} \left(-\kappa Y_t dt + \sigma dW_t\right) \\ &= e^{\kappa t} \sigma dW_t. \end{aligned}$$

Thus, we have obtained an expression for Z_t:

$$Z_t = Z_0 + \int_0^t e^{\kappa s} \sigma dW_s.$$

Next, we use $Y_t = e^{-\kappa t} Z_t$ and $X_t = Y_t + \theta$ to obtain

$$\begin{aligned} Y_t &= e^{-\kappa t} Y_0 + \int_0^t e^{-\kappa(t-s)} \sigma dW_s, \\ X_t &= \theta + e^{-\kappa t}(X_0 - \theta) + \int_0^t e^{-\kappa(t-s)} \sigma dW_s. \end{aligned}$$

Note that X_t has a normal distribution at every time $t > 0$ because Itô integrals with deterministic integrands are normally distributed random variables; see Proposition 3.5.7. Thus, the distribution of X_t is completely determined by its mean and variance. We have

$$\begin{aligned} \mathbb{E}X_t &= \theta + e^{-\kappa t}(X_0 - \theta) + \mathbb{E} \int_0^t e^{-\kappa(t-s)} \sigma dW_s \\ &= \theta + e^{-\kappa t}(X_0 - \theta), \\ \mathbb{V}X_t &= \mathbb{V} \int_0^t e^{-\kappa(t-s)} \sigma dW_s \\ &= \int_0^t \left(e^{-\kappa(t-s)} \sigma\right)^2 ds = \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa t}\right). \end{aligned}$$

where we have used Proposition 3.5.7.

3.7 Girsanov's Theorem for a single Brownian motion

We briefly recall some results from Section 2.7. Suppose that, on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we have a random variable $Z \geq 0$ that has expectation $\mathbb{E}Z = 1$. Then we can define a new probability measure $\widetilde{\mathbb{P}}$ via

$$\widetilde{\mathbb{P}}(A) = \mathbb{E}Z\mathbb{1}_A, \qquad A \in \mathcal{F},$$

and we call $Z = \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}$ the Radon-Nikodým derivative of $\widetilde{\mathbb{P}}$ with respect to $\mathbb{P}$. If Z is strictly positive $Z > 0$, then we also have

$$\mathbb{P}(A) = \widetilde{\mathbb{E}}\frac{1}{Z}\mathbb{1}_A, \qquad A \in \mathcal{F},$$

and we call $\frac{1}{Z} = \frac{d\mathbb{P}}{d\widetilde{\mathbb{P}}}$ the Radon-Nikodým derivative of $\mathbb{P}$ with respect to $\widetilde{\mathbb{P}}$.

In Example 2.7.5, on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we defined $X \sim \mathcal{N}(0, 1)$ and a Radon-Nikodým derivative $Z = e^{-\theta X - \frac{1}{2}\theta^2}$. We showed that $Y := X + \theta$ was $\mathcal{N}(\theta, 1)$ under $\mathbb{P}$ and $\mathcal{N}(0, 1)$ under $\widetilde{\mathbb{P}}$. Thus, Z had the effect of changing the mean of Y.

We would like to extend this idea from a static to a dynamics setting. Specifically, we would like to find a measure change that modifies the dynamics of a stochastic process $X = (X_t)_{t \geq 0}$.

<u>Definition</u> 3.7.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ be a filtration on this space. A *Radon-Nikodým derivative process* $(Z_t)_{0 \leq t \leq T}$ is any process of the form

$$Z_t := \mathbb{E}[Z|\mathcal{F}_t]$$

where Z is a random variable satisfying $\mathbb{E}Z = 1$ and $Z > 0$.

Note that Z in Definition 3.7.1 satisfies the conditions of a Radon-Nikodým derivative. As such, one can define a measure change $\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}$ from Z.

<u>Theorem</u> 3.7.2 (Girsanov). *Let $W = (W_t)_{0 \leq t \leq T}$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ be a filtration for W. Suppose $\Theta = (\Theta_t)_{0 \leq t \leq T}$ is adapted to the filtration $\mathbb{F}$. Define $(Z_t)_{0 \leq t \leq T}$ and $\widetilde{W} = (\widetilde{W}_t)_{0 \leq t \leq T}$ by*

$$Z_t = \exp\left(-\int_0^t \tfrac{1}{2}\Theta_s^2 ds - \int_0^t \Theta_s dW_s\right), \qquad d\widetilde{W}_t = \Theta_t dt + dW_t, \qquad \widetilde{W}_0 = 0.$$

Assume that

$$\mathbb{E}\int_0^T \Theta_t^2 Z_t^2 dt < \infty.$$

Define a Radon-Nikodým derivative $Z \equiv \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} := Z_T$. Then the process $\widetilde{W}$ is a Brownian motion under $\widetilde{\mathbb{P}}$.

EXAMPLE 3.7.3. On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$, consider a process $X = (X_t)_{0 \le t \le T}$ which is defined by the following Itô process

$$dX_t = \mu_t dt + \sigma_t dW_t,$$

where μ and σ are $\mathbb{F}$-adapted. Let us define a change of measure as follows

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} := \exp\Big(-\frac{1}{2}\int_0^T \gamma_t^2 dt - \int_0^T \gamma_t dW_t \Big),$$

where γ is $\mathbb{F}$-adapted. What are the dynamics of X under $\widetilde{\mathbb{P}}$? We know that the process $\widetilde{W} = (\widetilde{W}_t)_{0 \le t \le T}$ defined by

$$d\widetilde{W}_t := \gamma_t dt + dW_t, \qquad \widetilde{W}_0 := 0,$$

is a Brownian motion under $\widetilde{\mathbb{P}}$. Thus, we have

$$\begin{aligned} dX_t &= \mu_t dt + \sigma_t (d\widetilde{W}_t - \gamma_t dt) \\ &= (\mu_t - \sigma_t \gamma_t) dt + \sigma_t d\widetilde{W}_t. \end{aligned}$$

Thus, while X has a drift of μ_t under $\mathbb{P}$, it has a drift of $\mu_t - \sigma_t\gamma_t$ under $\widetilde{\mathbb{P}}$.

3.8 Girsanov's Theorem for d-dimensional Brownian motion

We conclude this chapter by stating (wihtout proof) Girsanov's Theorem and the martingale representation theorem for multi-dimensional Brownian motions.

THEOREM 3.8.1 (GIRSANOV). *Let* $W = (W_t^1, W_t^2, \dots, W_t^d)_{0 \le t \le T}$ *be a d-dimensional Brownian motion on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and let* $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ *be a filtration for* W. *Suppose* $\Theta = (\Theta_t^1, \Theta_t^2, \dots, \Theta_t^d)_{0 \le t \le T}$ *is adapted to the filtration* $\mathbb{F}$. *Define* $(Z_t)_{0 \le t \le T}$ *and* $\widetilde{W} = (\widetilde{W}_t^1, \widetilde{W}_t^2, \dots, \widetilde{W}_t^d)_{0 \le t \le T}$ *by*

$$Z_t = \exp\Big(-\int_0^t \tfrac{1}{2}\langle \Theta_s, \Theta_s \rangle ds - \int_0^t \langle \Theta_s, dW_s \rangle \Big), \qquad d\widetilde{W}_t = \Theta_t dt + dW_t, \qquad \widetilde{W}_0 = 0,$$

where $\langle \cdot, \cdot \rangle$ *denotes a d-dimensional Euclidean inner product. Assume that*

$$\mathbb{E}\int_0^T \langle \Theta_t, \Theta_t \rangle Z_t^2 dt < \infty.$$

Define a Radon-Nikodým derivative $Z \equiv \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} := Z_T$. *Then the process* $\widetilde{W}$ *is a d-dimensional Brownian motion under* $\widetilde{\mathbb{P}}$.

It is interesting to note that the components of $\widetilde{W}$ in Theorem 3.8.1 could be co-dependent under $\mathbb{P}$ (as Θ_j could depend on any of $W^1, W^2, \dots, W^d$). Nevertheless, under $\widetilde{\mathbb{P}}$, the components of $\widetilde{W}$ are independent of each other.

3.9 Stochastic differential equations

Definition 3.9.1. A *stochastic differential equation* (SDE) is an equation of the form

$$\mathrm{d}X_s = \mu(s, X_s)\mathrm{d}s + \sigma(s, X_s)\mathrm{d}W_s, \qquad X_t = x, \tag{3.17}$$

where $X = (X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(d)})_{t\geq 0}$ lives in $\mathbb{R}^d$, $W = (W_t^{(1)}, W_t^{(2)}, \dots, W_t^{(m)})_{t\geq 0}$ is an m-dimensional Brownian motion, $\mu : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$. We call functions μ and σ the *drift* and *diffusion*, respectively, and we call $X_t = x$ the *initial condition*. A *(strong) solution* of an SDE is a stochastic process $X = (X_s)_{s\geq t}$ such that

$$X_T = x + \int_t^T \mu(s, X_s)\mathrm{d}s + \int_t^T \sigma(s, X_s)\mathrm{d}W_s, \tag{3.18}$$

for all $T \geq t$.

One way to envision a strong solution of an SDE is as follows: think of a sample path $W_\cdot(\omega) : [t, \infty) \to \mathbb{R}$ as input. From this input, we can construct a unique sample path $X_\cdot(\omega) : [t, \infty) \to \mathbb{R}$.

Ideally, we would like to write X_T as an *explicit* functional of the Brownian path $(W_s)_{s\geq t}$. Unfortunately, this is typically not possible. Still, it will help to build intuition if we see some explicitly solvable examples.

Example 3.9.2 (Geometric Brownian motion). A *geometric Brownian motion* is a process $Z = (Z)_{t\geq 0}$ that satisfies

$$\mathrm{d}Z_t = \mu(t)Z_t\mathrm{d}t + \sigma(t)Z_t\mathrm{d}W_t, \qquad Z_0 = z,$$

where μ and σ are *deterministic* functions of t. To solve this SDE, we consider $X_t = \log Z_t$. Using the Itô formula, we obtain

$$\begin{aligned} \mathrm{d}X_t = \mathrm{d}\log Z_t &= \frac{1}{Z_t}\mathrm{d}Z_t + \frac{1}{2}\left(\frac{-1}{Z_t^2}\right)\mathrm{d}[Z, Z]_t \\ &= \left(\mu(t) - \tfrac{1}{2}\sigma^2(t)\right)\mathrm{d}t + \sigma(t)\mathrm{d}W_t. \end{aligned}$$

Integrating from 0 to T, we obtain

$$X_T = x + \int_0^T \left(\mu(t) - \tfrac{1}{2}\sigma^2(t)\right)\mathrm{d}t + \int_0^T \sigma(t)\mathrm{d}W_t, \qquad x = \log z.$$

Finally, we obtain our expression for Z_T.

$$Z_T = \exp(X_T) = z\exp\left(\int_0^T \left(\mu(t) - \tfrac{1}{2}\sigma^2(t)\right)\mathrm{d}t + \int_0^T \sigma(t)\mathrm{d}W_t\right).$$

THEOREM 3.9.3 (MARKOV PROPERTY OF SOLUTIONS OF AN SDE). *Let* $X = (X_t)_{t \geq 0}$ *be the solution of an SDE of the form* (3.17). *The* X *is a Markov process. That is, for* $t \leq T$ *and for some suitable function* φ*, there exists a function* g *(which depends on* t*,* T *and* φ*) such that*

$$\mathbb{E}[\varphi(X_T)|\mathcal{F}_t] = g(X_t),$$

where $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ *is any filtration to which* X *is adapted.*

The proof of Theorem 3.9.3 is somewhat technical and will not be given here. But, the intuitive idea for why the theorem is true is rather simple. From (3.18), we see that the value of X_T depends only on the path of the Brownian motion over the interval $[t, T]$ and the initial value $X_t = x$. The path that X took to arrive at $X_t = x$ plays no role. In other words, *given the present* $X_t = x$*, the future* $(X_T)_{T > t}$ *is independent of the past* $\mathcal{F}_t$. With this in mind, the process X should admit a transition density

$$\mathbb{P}(X_T \in dy | X_t = x) = \Gamma(t, x; T, y) dy,$$

and thus, the function g should be given by

$$g(X_t) = \mathbb{E}[\varphi(X_T)|\mathcal{F}_t] = \mathbb{E}[\varphi(X_T)|X_t] = \int dy \, \Gamma(t, X_t; T, y) \varphi(y).$$

Of course, finding an explicit representation of the transition density Γ may not be possible.

3.10 EXERCISES

EXERCISE 3.1. Let W be a Brownian motion and let $\mathbb{F}$ be a filtration for W. Show that $W_t^2 - t$ is a martingale with respect to the filtration $\mathbb{F}$.

EXERCISE 3.2. Define

$$X_t = \mu t + W_t,$$

where $W = (W_t)_{t \geq 0}$ is a Brownian motion. Let $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ be a filtration for W. Show that Z is a martingale with respect to $\mathbb{F}$ where

$$Z_t = \exp\left(\sigma X_t - (\sigma\mu + \sigma^2/2)t\right).$$

EXERCISE 3.3. Compute $d(W_t^4)$. Write W_T^4 as an integral with respect to W plus an integral with respect to t. Use this representation of W_T^4 to show that $\mathbb{E}W_T^4 = 3T^2$. Compute $\mathbb{E}W_T^6$ using the same technique.

EXERCISE 3.4. Find an explicit expression for Y_T where

$$dY_t = r dt + \alpha Y_t dW_t.$$

Hint: compute $d(Y_t Z_t)$ where $Z_t := \exp(-\alpha W_t + \frac{1}{2}\alpha^2 t)$.

EXERCISE 3.5. Suppose X, Δ and Π are given by

$$dX_t = \sigma X_t dW_t, \qquad \Delta_t = \frac{\partial f}{\partial x}(t, X_t), \qquad \Pi_t = X_t \Delta_t$$

where f is some smooth function. Show that if f satisfies

$$\left(\frac{\partial}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2}\right) f(t,x) = 0,$$

for all (t, x), then Π is a martingale with respect to a filtration $\mathcal{F}_t$ for W.

EXERCISE 3.6. Suppose X is given by

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t.$$

For any smooth function f define

$$M_t^f := f(t, X_t) - f(0, X_0) - \int_0^t \left(\frac{\partial}{\partial s} + \mu(s, X_s)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(s, X_s)\frac{\partial^2}{\partial x^2}\right) f(s, X_s) ds.$$

Show that M^f is a martingale with respect to a filtration $\mathcal{F}_t$ for W.

EXERCISE 3.7. Let $X = (X_t)_{0 \le t \le T}$ be an OU process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$

$$dX_t = K(\theta - X_t)dt + \sigma dW_t.$$

Where $W = (W_t)_{0 \le t \le T}$ is a *Brownian motion* under probability measure $\mathbb{P}$. Then we can define a new probability measure $\widetilde{\mathbb{P}}$ such that the process $\widetilde{W} = (\widetilde{W}_t)_{0 \le t \le T}$ is a *Brownian motion* under $\widetilde{\mathbb{P}}$. Then the OU process $X = (X_t)_{0 \le t \le T}$ on the new probability space $(\Omega, \mathcal{F}, \widetilde{\mathbb{P}})$ will be

$$dX_t = K(\theta^* - X_t)dt + \sigma d\widetilde{W}_t.$$

Find the *Radon-Nikodým* derivative $\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}$.

EXERCISE 3.8. For $i = 1, 2, \dots, d$, let $X^{(i)}$ satisfy

$$dX_t^{(i)} = -\frac{b}{2}X_t^{(i)}dt + \frac{1}{2}\sigma dW_t^{(i)},$$

where the $(W^{(i)})_{i=1}^d$ are independent Brownian motions. Define

$$R_t := \sum_{i=1}^d (X_t^{(i)})^2, \qquad B_t := \sum_{i=1}^d \int_0^t \frac{1}{\sqrt{R_s}} X_s^{(i)} dW_s^{(i)}.$$

Lévy's Theorem says that, if a process $M = (M_t)_{t \ge 0}$ is a martingale and $[M, M]_t = t$ for all $t \ge 0$, then M is a Brownian motion. Use this information to show that B is a Brownian motion. Derive an SDE for R that involves only dt and dB_t terms (i.e., no $dW_t^{(i)}$ terms should appear).

Chapter 4

No-arbitrage pricing

In this chapter, we will give a precise definition of arbitrage and present the fundamental theorem of asset pricing.

4.1 Arbitrage

We have previously characterized *arbitrage* as an opportunity to make a guaranteed profit with zero initial investment. In this section, we will provide a more precise definition. First, let us define what we mean by *self-financing portfolio.*

Definition 4.1.1. Consider a financial market with assets $(A_t^1, A_t^2, \ldots, A_t^n)_{t \geq 0}$. A portfolio is *self-financing* if its value $X = (X_t)_{t \geq 0}$ at all times is given by

$$X_t = \sum_{i=1}^{n} \Delta_t^i A_t^i,$$

where Δ_t^i represents the number of shares of A^i held at time t, and changes to the value of the portfolio are due only to changes in the value the assets

$$dX_t = \sum_{i=1}^{n} \Delta_t^i dA_t^i.$$

Gains and/or losses of a self-financing portfolio are due only to changes in the values of the assets in the portfolio. A portfolio would not be self-financing if, for example, and investor added cash to the portfolio at different times. Now that we understand what a self-financing portfolio is, we can define what we mean by arbitrage.

Definition 4.1.2. An *arbitrage* is any self-financing portfolio whose value $X = (X_t)_{t \geq 0}$ satisfies

1. $X_0 = 0$,
2. $\mathbb{P}(X_T \geq 0) = 1$,
3. $\mathbb{P}(X_T > 0) > 0$,

for some $T > 0$.

From the above definition, we see that an arbitrage is a trading strategy that can be financed with zero initial investment, has no probability of losing money, and has some strictly positive probability of making money. Let us take a looke at a few examples of arbitrage portfolios.

<u>Example</u> 4.1.3. Recall that the time t value of a zero-coupon bond with maturity $T > t$ is given by

$$B_t^T = \exp\left(- \int_t^T f_t^s \mathrm{d}s \right)$$

Now, suppose that the forward rate-curve has the following dynamics

$$f_t^T = \begin{cases} f_0^T & t < 1, \\ f_0^T + \varepsilon & t \geq 1, \end{cases}$$

where $\varepsilon \in \mathbb{R}$ is a random variable to be realized at time $t = 1$. From the above dynamics, we see that the entire forward rate curve experiences a jump at time t. The time $t = 1$ value a bond with maturity $T_i > 1$ is

$$\begin{aligned} B_1^{T_i} &= \exp\left(- \int_1^{T_i} f_t^s \mathrm{d}s \right) \\ &= \exp\left(- \int_1^{T_i} (f_0^s + \varepsilon) \mathrm{d}s \right) \\ &= \exp\left(\int_0^1 f_0^s \mathrm{d}s - \int_0^{T_i} f_0^s \mathrm{d}s - \varepsilon(T_i - 1) \right) \\ &= \frac{B_0^{T_i}}{B_0^1} e^{-\varepsilon(T_i - 1)}. \end{aligned}$$

We wish to see if this market has an arbitrage. Consider a potfolio $X = (X_t)_{t \geq 0}$ consisting of three bonds with matrities $1 < T_1 < T_2 < T_3$. The initial value of this portflolio is

$$X_0 = \sum_{i=1}^{3} \Delta_0^i B_0^{T_i}.$$

If we do not adjust the porfolio weights, then the value of the portfolio at time $t = 1$ is

$$X_1 = \sum_{i=1}^{3} \Delta_0^i B_1^{T_i}$$

$$
\begin{aligned}
&= \sum_{i=1}^{3} \Delta_0^i \frac{B_0^{T_i}}{B_0^1} e^{-\varepsilon(T_i-1)} \\
&= \frac{e^{-\varepsilon(T_2-1)}}{B_0^1} g(\varepsilon), \qquad\qquad g(\varepsilon) := \sum_{i=1}^{3} \Delta_0^i B_0^{T_i} e^{-\varepsilon(T_i-T_2)}.
\end{aligned}
$$

Observe that $X_1 > 0 \Leftrightarrow g(\varepsilon) > 0$ and also $g(0) = X_0$. As such, there will be an arbitrage if

$$
g(0) = 0, \qquad \text{and} \qquad g(\varepsilon) > 0, \qquad \forall\, \varepsilon \neq 0.
$$

Observe that $g(\varepsilon) > 0$ for all $\varepsilon \neq 0$ if $g(0) = g'(0) = 0$ and $g''(\varepsilon) > 0$ for all ε (draw a graphs of $g(\varepsilon)$ as a function of ε). Thus, we check if it is possible to have

$$
\begin{aligned}
0 &= g(0) = \Delta_0^1 B_0^{T_1} + \Delta_0^2 B_0^{T_2} + \Delta_0^3 B_0^{T_3}, && (4.1)\\
0 &= g'(0) = -(T_1 - T_2)\Delta_0^1 B_0^{T_1} - (T_3 - T_2)\Delta_0^3 B_0^{T_3}, && (4.2)\\
0 &< g''(\varepsilon) = (T_1 - T_2)^2 \Delta_0^1 B_0^{T_1} e^{-\varepsilon(T_1-T_2)} + (T_3 - T_2)^2 \Delta_0^3 B_0^{T_3} e^{-\varepsilon(T_3-T_2)}. && (4.3)
\end{aligned}
$$

If we choose $\Delta_0^1 > 0$ and $\Delta_0^3 > 0$ then (4.3) will be satisfied. Furthermore, with $\Delta_0^1 > 0$ and $\Delta_0^3 > 0$ we have

$$
(T_1 - T_2)\Delta_0^1 B_0^{T_1} < 0, \qquad\qquad (T_3 - T_1)\Delta_0^3 B_0^{T_3} > 0.
$$

So, we can choose $\Delta_0^1 > 0$ and $\Delta_0^3 > 0$ such that (4.2) is satisfied. Finally, we can choose $\Delta_0^2 < 0$ such that (4.1) is satisfied. As such, there exists an arbitrage.

EXAMPLE 4.1.4. Suppose at that zero-coupon bond prices are given by

$$
B_t^T = \begin{cases} e^{-R_0(T-t)} & t < 1, \\ e^{-(R_0+\varepsilon)(T-t)} & t \geq 1. \end{cases}
$$

where $\varepsilon \in \mathbb{R}$ is a random variable to be realized at time $t = 1$. The forward rate curve, given the above bond-prices, is given by

$$
f_t^T = -\partial_T \log B_t^T = \begin{cases} R_0 & t < 1, \\ R_0 + \varepsilon & t \geq 1. \end{cases}
$$

Noting that $f_0^T = R_0$, we see that the bond price dynamics are a special case of Example 4.1.3 with $f_0^T = R_0$. As such, there is an arbitrage in this market.

4.2 Fundamental Theorem of asset pricing

Theorem 4.2.1 (1st fundamental theorem of asset pricing). *Consider a financial market, defined under a probability measure* $\mathbb{P}$*. Let* $\mathrm{N} = (\mathrm{N}_t)_{t \geq 0}$ *be any stricly positive self-financing portfolio. A market is free of arbitrage if and only if there exists a probability measure* $\widetilde{\mathbb{P}}$*, equivalent to* $\mathbb{P}$*, under which* X/N *is martingale for all self-financing portfolios* X*.*

Proof. If $\widetilde{\mathbb{P}}$ exists, then we have

$$\widetilde{\mathbb{E}} \frac{\mathrm{X}_\mathrm{T}}{\mathrm{N}_\mathrm{T}} = \frac{\mathrm{X}_0}{\mathrm{N}_0},$$

for every self-financing portfolio X. In particlar, if $\mathrm{X}_0 = 0$, then we have

$$\widetilde{\mathbb{E}} \frac{\mathrm{X}_\mathrm{T}}{\mathrm{N}_\mathrm{T}} = 0. \tag{4.4}$$

Now, observe that

$$\mathbb{P}(\mathrm{X}_\mathrm{T} \geq 0) = 1 \quad \Leftrightarrow \quad \mathbb{P}(\mathrm{X}_\mathrm{T} < 0) = 0 \quad \Leftrightarrow \quad \widetilde{\mathbb{P}}(\mathrm{X}_\mathrm{T} < 0) = 0 \quad \Leftrightarrow \quad \widetilde{\mathbb{P}}(\mathrm{X}_\mathrm{T}/\mathrm{N}_\mathrm{T} < 0) = 0$$

where we have used the fact that $\widetilde{\mathbb{P}}$ is equivalent to $\mathbb{P}$. But

$$\widetilde{\mathbb{P}}(\mathrm{X}_\mathrm{T}/\mathrm{N}_\mathrm{T} < 0) = 0 \text{ and Equation (4.4)} \quad \Leftrightarrow \quad \widetilde{\mathbb{P}}(\mathrm{X}_\mathrm{T}/\mathrm{N}_\mathrm{T} > 0) = 0.$$

Lastly

$$\widetilde{\mathbb{P}}(\mathrm{X}_\mathrm{T}/\mathrm{N}_\mathrm{T} > 0) = 0 \quad \Leftrightarrow \quad \mathbb{P}(\mathrm{X}_\mathrm{T}/\mathrm{N}_\mathrm{T} > 0) = 0 \quad \Leftrightarrow \quad \mathbb{P}(\mathrm{X}_\mathrm{T} > 0) = 0,$$

where we have once again used the fact that $\widetilde{\mathbb{P}}$ is equivalent to $\mathbb{P}$. Thus, if $\widetilde{\mathbb{P}}$ exists, the portfolio has zero initial value $\mathrm{X}_0 = 0$ and $\mathbb{P}(\mathrm{X}_\mathrm{T} \geq 0) = 1$, then we also have $\mathbb{P}(\mathrm{X}_\mathrm{T} > 0) = 0$ (i.e., we cannot have $\mathbb{P}(\mathrm{X}_\mathrm{T} > 0) > 0$). As such, there is no arbitrage. □

In the above Theorem, we call the portfolio N the *numéraire*. Typically, we will choose the money-market account M as numéraire. However, for some applications, it will be easier to choose a T-maturity zero coupon bond B^T as numéraire. The probability measure $\widetilde{\mathbb{P}}$ in the above Theorem is referred to as a *risk-neutral*, *martingale* or *pricing measure*. Note that the measure $\widetilde{\mathbb{P}}$ depends on the choice of numéraire. That is, If $\widetilde{\mathbb{P}}^1$ is a risk-neutral measure with N^1 as numéraire and $\widetilde{\mathbb{P}}^2$ is a risk-neutral measure with N^2 as numéraire then, in general, $\widetilde{\mathbb{P}}^1$ and $\widetilde{\mathbb{P}}^2$ will be different

Theorem 4.2.1 is important for two reasons

1. It gives us a simple way to check if a model for the financial market contains an arbitrage. We should *never* price assets using models that admit arbitrage.

2. It provides a way for us to price derivative assets in such a way that there is no arbitrage. Specifically, the value of any asset $A = (A_t)_{t \geq 0}$ must satisfy

$$\frac{A_t}{N_t} = \widetilde{\mathbb{E}}\left(\frac{A_T}{N_T}\middle|\mathcal{F}_t\right), \tag{4.5}$$

where $\widetilde{\mathbb{P}}$ is a martingale measure for numérair $N = (N_t)_{t \geq 0}$.

You can think of (4.5) as a pricing equation.

4.3 Examples

Let us take a look at a few examples.

<u>Example</u> 4.3.1 (Black-Scholes). Consider a market with a stock $S = (S_t)_{0 \leq t \leq T}$ and a money market account, whose dynamics are of the form

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \qquad\qquad dM_t = r M_t dt,$$

where W is a Brownian motion under $\mathbb{P}$. Let us see if this model has an arbitrage. First we need to choose a numéraire portfolio. We will take the money market account M as numéraire, as this is typically the easiest choice. Next, we need to see if there exists a probability measure $\widetilde{\mathbb{P}}$, equivalent to $\mathbb{P}$, under which X/M is a martingale for all portfolios X. As there are only two assets in our market S and M, all porfolios must have dynamics of the form

$$\begin{aligned} dX_t &= \Delta_t dS_t + (X_t - \Delta_t S_t)\frac{1}{M_t}dM_t \\ &= \Delta_t S_t(\mu - r)dt + \Delta_t \sigma S_t dW_t + X_t r dt. \end{aligned}$$

The dynamics of X/M are then

$$\begin{aligned} d\frac{X_t}{M_t} &= X_t d\left(\frac{1}{M_t}\right) + \frac{1}{M_t}dX_t + d\left[X, \frac{1}{M}\right]_t \\ &= -r\frac{X_t}{M_t}dt + \Delta_t(\mu - r)\frac{S_t}{M_t}dt + \Delta_t \sigma \frac{S_t}{M_t}dW_t + \frac{X_t}{M_t}r dt \\ &= \Delta_t(\mu - r)\frac{S_t}{M_t}dt + \Delta_t \sigma \frac{S_t}{M_t}dW_t. \end{aligned}$$

From Girsanov's Theorem, we know that the process $\widetilde{W}$ defined by

$$d\widetilde{W}_t = \gamma_t dt + dW_t$$

is a Brownian motion under $\widetilde{\mathbb{P}}$ where

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = \exp\Big(- \int_0^T \gamma_t dW_t - \tfrac{1}{2}\int_0^T \gamma_t^2 dt \Big).$$

The process $\gamma = (\gamma_t)_{0 \le t \le T}$ is arbitrary. Let us write the dynamics of X/M in terms of $\widetilde{W}$. We have

$$\begin{aligned} d\frac{X_t}{M_t} &= \Delta_t(\mu - r)\frac{S_t}{M_t}dt + \Delta_t\sigma\frac{S_t}{M_t}(d\widetilde{W}_t - \gamma_t dt) \\ &= \Delta_t(\mu - r - \sigma\gamma_t)\frac{S_t}{M_t}dt + \Delta_t\sigma\frac{S_t}{M_t}d\widetilde{W}_t. \end{aligned}$$

If we choose $\gamma_t := (\mu - r)/\sigma$ then the dt-term will disappear in the dynamics of X/M and we have

$$d\frac{X_t}{M_t} = \Delta_t\sigma\frac{S_t}{M_t}d\widetilde{W}_t.$$

Thus, X/M will be a martingale under $\widetilde{\mathbb{P}}$ for all portfolios X. The market therefore does not have any arbitrage. If we were to consider a derivative asset that pays $\varphi(S_T)$ at time T, the value $V = (V_t)_{0 \le t \le T}$ would satisfy

$$\begin{aligned} \frac{V_t}{M_t} &= \widetilde{\mathbb{E}}\Big(\frac{V_T}{M_T}\Big|\mathcal{F}_t\Big) = \widetilde{\mathbb{E}}\Big(\frac{\varphi(S_T)}{M_t}\Big|\mathcal{F}_t\Big), \\ V_t &= e^{-r(T-t)}\widetilde{\mathbb{E}}\Big(\varphi(S_T)\Big|\mathcal{F}_t\Big), \end{aligned}$$

where we have used $M_t = M_0 e^{rt}$. Note that we can comput the expectatio on the last line, as the dynamics of S, under $\widetilde{\mathbb{P}}$ are

$$dS_t = rS_t dt + \sigma S_t d\widetilde{W}_t \qquad \Rightarrow \qquad S_T = S_t \exp\Big((r - \sigma^2/2)(T - t) + \sigma(\widetilde{W}_T - \widetilde{W}_t)\Big),$$

and $\widetilde{W}_T - \widetilde{W}_t \sim \mathcal{N}(0, T - t)$.

<u>Example</u> 4.3.2 (Zero-coupon bond prices). Consider a market consisting of a money market account, whose dynamics under the physical probability measrue $\mathbb{P}$ are of the form

$$\begin{aligned} dM_t &= R_t M_t dt & &\Rightarrow & M_t &= M_0 \exp\Big(\int_0^t R_s ds\Big), \\ dR_t &= b_t dt + a_t dW_t. \end{aligned}$$

As there is only one asset in this market, the only portfolio X one can hold is the money market accout

$$dX_t = \frac{X_t}{M_t}dM_t \qquad X_t = X_0 \exp\Big(\int_0^t R_s ds\Big) = X_0\frac{M_t}{M_0}$$

Noting that $X_t/M_t = X_0.M_0$, we see that X/M is automatically a martingale under $\widetilde{\mathbb{P}}$ for *any* choice of a and b. In fact, X/M is a martingale under *all* probability measures $\widetilde{\mathbb{P}}$ that are equivalent to $\mathbb{P}$. What would the dynamics of R look like under a different probability measure? We know from Girsanov's theorem that, for any process $\gamma = (\gamma_t)_{0 \leq t \leq T}$ a probability measure $\widetilde{\mathbb{P}}$ defined by

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = \exp\left(- \int_0^T \gamma_t dW_t - \tfrac{1}{2} \int_0^T \gamma_t^2 dt \right),$$

is equivalent to $\mathbb{P}$ and that the process $\widetilde{W}$ defined by

$$d\widetilde{W}_t = \gamma_t dt + dW_t$$

is a Brownian motion under $\widetilde{\mathbb{P}}$. The dynamics of R under $\mathbb{P}$ are

$$\begin{aligned} dR_t &= b_t dt + a_t (d\widetilde{W}_t - \gamma_t) dt \\ &= (b_t - a_t \gamma_t) dt + a_t d\widetilde{W}_t. \end{aligned}$$

Observe that the dyanmics of R under $\widetilde{\mathbb{P}}$ retain the form of the dynamics of R under $\mathbb{P}$; we have only made the replacement $b_t \to b_t - a_t \gamma_t$. Because the dynamics the processes a, b, and γ were arbitrary, we could have simply chosen to specify the dynamics of R directly under a risk-neutral measure $\widetilde{\mathbb{P}}$. In fact, this is what is typically done in fixed-income markets. Now, what would be the price of a zero-coupon bond in this market? Using the pricing formula (4.5) we have

$$\frac{B_t^T}{M_t} = \widetilde{\mathbb{E}}\left(\frac{B_T^T}{M_T} \middle| \mathcal{F}_t \right) \qquad \Rightarrow \qquad B_t^T = \widetilde{\mathbb{E}}\left(\frac{M_t}{M_T} \middle| \mathcal{F}_t \right) = \widetilde{\mathbb{E}}\left(\exp\left(- \int_t^T R_s ds \right) \middle| \mathcal{F}_t \right), \tag{4.6}$$

where we have used the fact that $B_T^T = 1$. We can think of (4.6) as a pricing formula for bonds. Once we have specified the dyanmics of the short rate R, under a risk-neutral measure $\widetilde{\mathbb{P}}$, then we can compute bond prices using (4.6). We will see some specific examples in the following chapters.

<u>Example</u> 4.3.3 (T-forward prices). Recall from Section 1.7 that the T-forward price of an asset $A = (A_t)_{t \geq 0}$ at time t is the value of K that makes a contract that pays $A_T - K$ at time T have zero value. We previously showed via a replication argument that the T-forward price, denoted $A^T = (A_t^T)$ is given by $A_t^T = A_t / B_t^T$. Let us derive this result once again using risk-neutral pricing. Using the risk-neutral pricing formula (4.5) with the money market M as numéraire, the value $V = (V_t)_{0 \leq t \leq T}$ of an asset that pays $A_T - K$ at time T satisfies

$$\begin{aligned} \frac{V_t}{M_t} &= \widetilde{\mathbb{E}}\left(\frac{V_T}{M_T} \middle| \mathcal{F}_t \right) = \widetilde{\mathbb{E}}\left(\frac{A_T - K}{M_T} \middle| \mathcal{F}_t \right) \\ &= \frac{A_t}{M_t} - K \widetilde{\mathbb{E}}\left(\frac{1}{M_T} \middle| \mathcal{F}_t \right), \end{aligned}$$

where we have used the fact the A/M is a martingale under $\widetilde{\mathbb{P}}$ (otherwise, there would be arbitrage). Multiplying through by M_t we obtain

$$V_t = A_t - K \widetilde{\mathbb{E}}\left(\frac{M_t}{M_T} \middle| \mathcal{F}_t \right) = A_t - K B_t^T.$$

The T-forward price A_t^T is the value of K that makes $V_t = 0$. Thus, we have

$$A_t^T = \frac{A_t}{B_t^T},$$

which agrees with the result we derived in Section 1.7.

4.4 Exercises

<u>Exercise</u> 4.1. Suppose that $f_0^T = 0.08$ for all $T \geq 0$. Three zero-coupon bonds trade with maturities at 5, 10 and 15 years. At time $t = 1$ the yield curve will jump to $f_1^T = f_0^T + \xi$ and $\mathbb{P}(\xi = 0.02) = 1/2$ and $\mathbb{P}(\xi = -0.02) = 1/2$. Construct an arbitrage using the three zero-coupon bonds.

<u>Exercise</u> 4.2. Consider a market consisting of a stock S and a money market account M, whose dynamics are given by

$$dM_t = rM_t dt, \qquad dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where W is a Brownian motion under the physical probability measure $\mathbb{P}$. Suppose that the owner of a share of S receives a dividend of $qS_t dt$ units of currency over the infinitesimally small time interval $[t, t + dt)$. Suppose that $\widetilde{\mathbb{P}}$ is a martingale measure with M as numéraire. How does $\widetilde{\mathbb{P}}$ relate to $\mathbb{P}$? i.e., what is $d\widetilde{\mathbb{P}}/d\mathbb{P}$? What are the dynamics of S under $\widetilde{\mathbb{P}}$? Is S/M a martingale under $\widetilde{\mathbb{P}}$? Why or why not.

Chapter 5

Short-rate modeling

In this Chapter, we will look at two equivalent ways to zero-coupon bonds and interest rate derivatives.

5.1 Pricing zero-coupon bonds by replication

In this section, we suppose that the dynamics of a money-market account M and and short-rate are of the form

$$\begin{aligned}
\mathrm{dM}_t &= \mathrm{R}_t \mathrm{M}_t \mathrm{d}t, \\
\mathrm{dR}_t &= b(t, \mathrm{R}_t)\mathrm{d}t + a(t, \mathrm{R}_t)\mathrm{dW}_t,
\end{aligned}$$

where W is a Brownian motion under the physics probability measure $\mathbb{P}$. We will attemp to replicate the payoff of a bond maturing at time T_1 by trading the money market account M and a bond maturity at time $\mathrm{T}_2 > \mathrm{T}_1$. The short-rate R is the solution of an SDE and is therefore a Markov process. It follows that the time t price of a bond B_t^T are deterministic function of t and R_t. That is

$$\mathrm{B}_t^\mathrm{T} = \mathrm{B}(t, \mathrm{R}_t; \mathrm{T}),$$

where the function $\mathrm{B} : [0, \mathrm{T}] \times \mathbb{R}^+ \mapsto [0, 1]$ is to be determined. Using Itô's formula, the dynamics of B^T are given by

$$\begin{aligned}
\mathrm{dB}_t^\mathrm{T} &= \mathrm{dB}(t, \mathrm{R}_t; \mathrm{T}) \\
&= \partial_t \mathrm{B}(t, \mathrm{R}_t; \mathrm{T})\mathrm{d}t + \partial_r \mathrm{B}(t, \mathrm{R}_t; \mathrm{T})\mathrm{dR}_t + \tfrac{1}{2}\partial_r^2 \mathrm{B}(t, \mathrm{R}_t; \mathrm{T})\mathrm{d}[\mathrm{R}, \mathrm{R}]_t \\
&= \left(\partial_t + b(t, \mathrm{R}_t)\partial_r + \tfrac{1}{2}a^2(t, \mathrm{R}_t)\partial_r^2\right)\mathrm{B}(t, \mathrm{R}_t; \mathrm{T})\mathrm{d}t + a(t, \mathrm{R}_t)\partial_r \mathrm{B}(t, \mathrm{R}_t; \mathrm{T})\mathrm{dW}_t \\
&= \mu_t^\mathrm{T} \mathrm{B}_t^\mathrm{T} \mathrm{d}t + \nu_t^\mathrm{T} \mathrm{B}_t^\mathrm{T} \mathrm{dW}_t,
\end{aligned}$$

where we have defined

$$\mu_t^{\mathrm{T}} := \frac{1}{\mathrm{B}(t, \mathrm{R}_t; \mathrm{T})}\Big(\partial_t + b(t, \mathrm{R}_t)\partial_r + \tfrac{1}{2}a^2(t, \mathrm{R}_t)\partial_r^2\Big)\mathrm{B}(t, \mathrm{R}_t; \mathrm{T}), \tag{5.1}$$

$$\nu_t^{\mathrm{T}} := \frac{1}{\mathrm{B}(t, \mathrm{R}_t; \mathrm{T})} a(t, \mathrm{R}_t)\partial_r \mathrm{B}(t, \mathrm{R}_t; \mathrm{T}). \tag{5.2}$$

Now, consider a portfolio X with two assets (i) the money market account M, and (ii) a T_2-maturity bond $\mathrm{B}^{\mathrm{T}_2}$. The dynamics of such a porfolio are given by

$$\mathrm{dX}_t = \Delta_t \mathrm{dB}_t^{\mathrm{T}_2} + (\mathrm{X}_t - \Delta_t \mathrm{B}_t^{\mathrm{T}_2})\frac{1}{\mathrm{M}_t}\mathrm{dM}_t,$$

where Δ_t is the number of T_2-maturity bonds held at time t. Comparing with Example 4.3.1 we have

$$\mathrm{d}\Big(\frac{\mathrm{X}_t}{\mathrm{M}_t}\Big) = \Delta_t(\mu_t^{\mathrm{T}_2} - \mathrm{R}_t)\frac{\mathrm{B}_t^{\mathrm{T}_2}}{\mathrm{M}_t}\mathrm{d}t + \Delta_t \nu_t^{\mathrm{T}_2}\frac{\mathrm{B}_t^{\mathrm{T}_2}}{\mathrm{M}_t}\mathrm{dW}_t. \tag{5.3}$$

Similarly, the dynamics of $\mathrm{B}^{\mathrm{T}_1}/\mathrm{M}$ are

$$\mathrm{d}\Big(\frac{\mathrm{B}_t^{\mathrm{T}_1}}{\mathrm{M}_t}\Big) = (\mu_t^{\mathrm{T}_1} - \mathrm{R}_t)\frac{\mathrm{B}_t^{\mathrm{T}_1}}{\mathrm{M}_t}\mathrm{d}t + \nu_t^{\mathrm{T}_1}\frac{\mathrm{B}_t^{\mathrm{T}_1}}{\mathrm{M}_t}\mathrm{dW}_t. \tag{5.4}$$

In order to replication the bond $\mathrm{B}^{\mathrm{T}_1}$ with X, we must match the right-hand sides of (5.3) and (5.4). Comparing the dW_t-terms, we see that we must have

$$\Delta_t = \frac{\nu_t^{\mathrm{T}_1}\mathrm{B}_t^{\mathrm{T}_1}}{\nu_t^{\mathrm{T}_2}\mathrm{B}_t^{\mathrm{T}_2}}.$$

Next, setting the $\mathrm{d}t$-terms equal to eqch other, and using the above expression for Δ we obtain

$$\frac{\nu_t^{\mathrm{T}_1}\mathrm{B}_t^{\mathrm{T}_1}}{\nu_t^{\mathrm{T}_2}\mathrm{B}_t^{\mathrm{T}_2}}(\mu_t^{\mathrm{T}_2} - \mathrm{R}_t)\frac{\mathrm{B}_t^{\mathrm{T}_2}}{\mathrm{M}_t} = (\mu_t^{\mathrm{T}_1} - \mathrm{R}_t)\frac{\mathrm{B}_t^{\mathrm{T}_1}}{\mathrm{M}_t}.$$

Multiplying both sides by $\mathrm{M}_t/(\nu_t^{\mathrm{T}_1}\mathrm{B}_t^{\mathrm{T}_1})$ we obtain

$$\frac{1}{\nu_t^{\mathrm{T}_2}}(\mu_t^{\mathrm{T}_2} - \mathrm{R}_t) = \frac{1}{\nu_t^{\mathrm{T}_1}}(\mu_t^{\mathrm{T}_1} - \mathrm{R}_t).$$

Noting that the left-hand side depends only on T_2 and the right-hand side depends only on T_1, we conclude that both sides must be equal to a function $\gamma(t, \mathrm{R}_t)$ that does not depend on T_1 or T_2. Thus, we must have

$$\frac{1}{\nu_t^{\mathrm{T}}}(\mu_t^{\mathrm{T}} - \mathrm{R}_t) = \gamma(t, \mathrm{R}_t).$$

Mutiplying both sides by ν_t^{T} and using expressions (5.2) and (5.1), we obtain

$$\begin{aligned}&\frac{1}{\mathrm{B}(t,\mathrm{R}_t;\mathrm{T})}\left(\partial_t + b(t,\mathrm{R}_t)\partial_r + \tfrac{1}{2}a^2(t,\mathrm{R}_t)\partial_r^2\right)\mathrm{B}(t,\mathrm{R}_t;\mathrm{T}) - \mathrm{R}_t \\ &= \gamma(t,\mathrm{R}_t)\frac{1}{\mathrm{B}(t,\mathrm{R}_t;\mathrm{T})}a(t,\mathrm{R}_t)\partial_r\mathrm{B}(t,\mathrm{R}_t;\mathrm{T}).\end{aligned}$$

Lastly, multiplying through by $\mathrm{B}(t,\mathrm{R}_t;\mathrm{T})$ and moving all terms to one side, we find

$$0 = \left(\partial_t - \mathrm{R}_t + (b(t,\mathrm{R}_t) - \gamma(t,\mathrm{R}_t)a(t,\mathrm{R}_t))\partial_r + \tfrac{1}{2}a^2(t,\mathrm{R}_t)\partial_r^2\right)\mathrm{B}(t,\mathrm{R}_t;\mathrm{T}).$$

We have derived a pricing PDE for the price of a bond. The function B must satisfy

$$0 = \left(\partial_t - r + (b(t,r) - \gamma(t,r)a(t,r))\partial_r + \tfrac{1}{2}a^2(t,r)\partial_r^2\right)\mathrm{B}(t,r;\mathrm{T}), \qquad \mathrm{B}(\mathrm{T},r;\mathrm{T}) = 1, \tag{5.5}$$

for some function γ, where the terminal condition follows from the fact that $\mathrm{B}_\mathrm{T}^\mathrm{T} = 1$. We call the function γ the *market price of risk*. We will see it arise in a different context in the next section.

5.2 Risk-neutral pricing of zero-coupon bonds

In this Section we will derive the bond pricing PDE (5.5) using risk-neutral pricing. As in Section 5.1, we suppose that the dynamics of a money-market account M and and short-rate are of the form

$$\begin{aligned}\mathrm{dM}_t &= \mathrm{R}_t\mathrm{M}_t\mathrm{d}t,\\ \mathrm{dR}_t &= b(t,\mathrm{R}_t)\mathrm{d}t + a(t,\mathrm{R}_t)\mathrm{dW}_t,\end{aligned}$$

where W is a Brownian motion under the the physical probability measure $\mathbb{P}$. Let us define a change probability measure

$$\frac{\mathrm{d}\widetilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}} = \exp\left(-\tfrac{1}{2}\int_0^\mathrm{T}\gamma^2(t,\mathrm{R}_t)\mathrm{d}t - \int_0^\mathrm{T}\gamma(t,\mathrm{R}_t)\mathrm{dW}_t\right).$$

By Girsanov's theorem 3.7.2, the process $\widetilde{\mathrm{W}}$ defined by

$$\mathrm{d}\widetilde{\mathrm{W}}_t = \gamma(t,\mathrm{R}_t)\mathrm{d}t + \mathrm{dW}_t, \qquad \widetilde{\mathrm{W}}_0 = 0,$$

is a Brownian motion under $\widetilde{\mathbb{P}}$. The dynamics of R under $\widetilde{\mathbb{P}}$ are as follows

$$\begin{aligned}\mathrm{dR}_t &= b(t,\mathrm{R}_t)\mathrm{d}t + a(t,\mathrm{R}_t)(\mathrm{d}\widetilde{\mathrm{W}}_t - \gamma(t,\mathrm{R}_t)\mathrm{d}t)\\ &= \Big(b(t,\mathrm{R}_t) - \gamma(t,\mathrm{R}_t)a(t,\mathrm{R}_t)\Big)\mathrm{d}t + a(t,\mathrm{R}_t)\mathrm{d}\widetilde{\mathrm{W}}_t.\end{aligned} \tag{5.6}$$

As there is only one asset in our market at this point – the money market account M – a self-financing portfolio X must be of the form

$$dX_t = \frac{X_t}{M_t} dM_t = X_t R_t dt \qquad \Rightarrow \qquad X_t = X_0 \exp\Big(\int_0^T R_s ds \Big).$$

Thus, the process X/M is a constant

$$\frac{X_t}{M_t} = \frac{X_0}{M_0}.$$

and therefore trivially a martingale. Thus, $\widetilde{\mathbb{P}}$ is a martingale measure (with M as numeraire) for *any* choice of γ. Now, to price a zero-coupon bond, we use the risk-neutral pricing formula (4.5). We have

$$\frac{B_t^T}{M_t} = \widetilde{\mathbb{E}}\Big(\frac{B_T^T}{M_T} \Big| \mathcal{F}_t \Big) = \widetilde{\mathbb{E}}\Big(\frac{1}{M_T} \Big| \mathcal{F}_t \Big).$$

Solving for B_t^T we obtain

$$\begin{aligned} B_t^T &= \widetilde{\mathbb{E}}\Big(\frac{M_t}{M_T} \Big| \mathcal{F}_t \Big) \\ &= \widetilde{\mathbb{E}}\Big(\exp\Big(- \int_t^T R_s ds \Big) \Big| \mathcal{F}_t \Big). \end{aligned}$$

Now, because R, as the solution of an SDE (5.6), is a Markov process, it follows that there exists a function $B(\cdot, \cdot; T) : [0, T] \times \mathbb{R}_+ \to [0, 1]$ such that

$$B_t^T = B(t, R_t; T),$$

where the function B is yet to be determined. In order to derive a PDE for the function B, we recall that B^T/M is a martingale under $\widetilde{\mathbb{P}}$. As such, the dt-term in $d(B^T/M)$ must equal zero. Noting that

$$\begin{aligned} dB_t^T &= dB(t, R_t; T) \\ &= \Big(\partial_t + (b(t, R_t) - \gamma(t, R_t) a(t, R_t)) \partial_r + \tfrac{1}{2} a^2(t, R_t) \partial_r^2 \Big) B(t, R_t; T) dt \\ &\quad + a(t, R_t) \partial_r B(t, R_t; T) d\widetilde{W}_t, \\ d\Big(\frac{1}{M_t} \Big) &= \frac{-R_t}{M_t} dt, \end{aligned}$$

the dynamics of B^T/M are given by

$$\begin{aligned} d\Big(\frac{B_t^T}{M_t} \Big) &= \frac{1}{M_t} dB_t^T + B_t^T d\Big(\frac{1}{M_t} \Big) + d\Big[B^T, \frac{1}{M} \Big]_t \\ &= \frac{1}{M_t} \Big(\partial_t - R_t + (b(t, R_t) - \gamma(t, R_t) a(t, R_t)) \partial_r + \tfrac{1}{2} a^2(t, R_t) \partial_r^2 \Big) B(t, R_t; T) dt \end{aligned}$$

$$+ \frac{1}{\mathrm{M}_t} a(t, \mathrm{R}_t)\partial_r \mathrm{B}(t, \mathrm{R}_t; \mathrm{T})\mathrm{d}\widetilde{\mathrm{W}}_t.$$

As the $\mathrm{d}t$-term must equal zero for all paths of R it must be the case that the function B satisfies the following PDE

$$0 = \Big(\partial_t - r + (b(t,r) - \gamma(t,r)a(t,r))\partial_r + \tfrac{1}{2}a^2(t,r)\partial_r^2\Big)\mathrm{B}(t,r;\mathrm{T}), \qquad \mathrm{B}(\mathrm{T},r;\mathrm{T}) = 1, \quad (5.7)$$

where the terminal condition $\mathrm{B}(\mathrm{T}, \mathrm{R}_\mathrm{T}; \mathrm{T}) = 1$ follows from the fact that $\mathrm{B}_\mathrm{T}^\mathrm{T} = 1$ by definition. Observe that the above PDE is *exactly* that same as the PDE (5.5) we derived in the previous section.

5.3 Pricing and hedging interest rate derivatives

Now, let us consider an option that pays $g(\mathrm{R}_{\mathrm{T}_1})$ at time T_1. We will denote by $\mathrm{V} = (\mathrm{V}_t)_{0 \le t \le \mathrm{T}_1}$ the value of this option. As in the previous sections, we will assume that the dyamics of R under $\mathbb{P}$ are of the form

$$\mathrm{dR}_t = b(t, \mathrm{R}_t)\mathrm{d}t + a(t, \mathrm{R}_t)\mathrm{dW}_t.$$

To avoid arbitrage, the dynamics of R under the pricing measure $\widetilde{\mathbb{P}}$ with M as numéraire must be of the form

$$\mathrm{dR}_t = \Big(b(t,\mathrm{R}_t) - \gamma(t,\mathrm{R}_t)a(t,\mathrm{R}_t)\Big)\mathrm{d}t + a(t,\mathrm{R}_t)\mathrm{d}\widetilde{\mathrm{W}}_t, \quad (5.8)$$

for some function γ. By the first fundamental Theorem of asset pricing we have

$$\frac{\mathrm{V}_t}{\mathrm{M}_t} = \widetilde{\mathbb{E}}\Big(\frac{\mathrm{V}_{\mathrm{T}_1}}{\mathrm{M}_{\mathrm{T}_1}}\Big|\mathcal{F}_t\Big) \qquad \Rightarrow \qquad \mathrm{V}_t = \widetilde{\mathbb{E}}\Big(\mathrm{e}^{-\int_t^{\mathrm{T}_1}\mathrm{R}_s\mathrm{d}s} g(\mathrm{R}_{\mathrm{T}_1})\Big|\mathcal{F}_t\Big).$$

Because R is a Markov process, there exists a function $\mathrm{V}(\cdot,\cdot;\mathrm{T}_1) : [0, \mathrm{T}_1] \times \mathbb{R}_+ \to \mathbb{R}$ such that

$$\mathrm{V}_t = \mathrm{V}(t, \mathrm{R}_t; \mathrm{T}_1).$$

To derive a PDE for the function V, we recall that V/M is a $\widetilde{\mathbb{P}}$ martingale. Thus, the $\mathrm{d}t$-term in $\mathrm{d}(\mathrm{V}_t/\mathrm{M}_t)$ must equal zero. Additionally, the value of the option at maturity must equal the option payoff: $\mathrm{V}_{\mathrm{T}_1} = g(\mathrm{R}_{\mathrm{T}_1})$. These two facts leads to the following PDE and terminal condition for the function V

$$0 = \Big(\partial_t - r + (b(t,r) - \gamma(t,r)a(t,r))\partial_r + \tfrac{1}{2}a^2(t,r)\partial_r^2\Big)\mathrm{V}(t,r;\mathrm{T}_1), \qquad \mathrm{V}(\mathrm{T}_1,r;\mathrm{T}_1) = g(r). \ (5.9)$$

Comparing (5.9) with (5.7), we observe that the function V satisfies the same PDE as the function B – only the terminal condition has changed. Assuming we can solve the PDE for V, the value of the derivative is $V_t = V(t, R_t; T_1)$.

Now, assume we can solve the PDE for V. How can we replicate the claim that pays $g(R_{T_1})$? We will construct a replicating portfolio X by trading the money market account M and a bond maturity at time $T_2 \geq T_1$. The dynamics of this portfolio are of the form

$$dX_t = \Delta_t dB_t^{T_2} + (X_t - \Delta_t B_t^{T_2}) \frac{1}{M_t} dM_t.$$

If we set $X_0 = V_0$ and choose Δ so that $d(X_t/M_t) = d(V_t/M_t)$ then we will have $X_{T_1} = V_{T_1}$. Using the fact that V satisfies (5.9), we find that

$$d\left(\frac{V_t}{M_t}\right) = \frac{1}{M_t} a(t, R_t) \partial_r V(t, R_t; T_1) d\widetilde{W}_t.$$

Similarly, using the fact that B satisfies (5.7), we find that

$$d\left(\frac{X_t}{M_t}\right) = \frac{\Delta_t}{M_t} a(t, R_t) \partial_r B(t, R_t; T_2) d\widetilde{W}_t.$$

Comparing the above equations, we see that, in order for $d(X_t/M_t) = d(V_t/M_t)$, we must have

$$\Delta_t = \frac{\partial_r V(t, R_t; T_1)}{\partial_r B(t, R_t; T_2)}. \tag{5.10}$$

Note that, for the special case $g(r) = 1$, we have $V(t, R_t; T_1) = B(t, R_t, T_1)$. Thus, (5.10) gives us a way to hedge a T_1 maturity bond by trading the money-market account M and a T_2-maturity bond (in addition to other more complicated interest rate derivatives that mature at time T_1). Of course, perfect replication requires knowledge of the $\widetilde{\mathbb{P}}$ dynamics of R.

5.4 A note on γ

Suppose that the dynamics of R under the physical measure $\mathbb{P}$ are

$$dR_t = b(t, R_t)dt + a(t, R_t)dW_t,$$

Then, as mentioned above, to avoid arbitrage, the dynamics of R under the pricing measure $\widetilde{\mathbb{P}}$ with M as numéraire must be of the form

$$dR_t = \Big(b(t, R_t) - \gamma(t, R_t) a(t, R_t)\Big)dt + a(t, R_t) d\widetilde{W}_t,$$

for some function γ. Suppose we would like the dynamics of R under $\widetilde{\mathbb{P}}$ to be of the form

$$dR_t = \widetilde{b}(t, R_t)dt + a(t, R_t)d\widetilde{W}_t,$$

for some function $\widetilde{b}$. Then we can always achieve this by choosing

$$\gamma(t,r) = -\Big(\widetilde{b}(t,r) - b(t,r)\Big)/a(t,r).$$

Because of this, when one prices bonds and/or other interest rate derivatives, it is common to simply make the change $\widetilde{b} = (b - \gamma a) \to b$ and write the dynamics of R under $\widetilde{\mathbb{P}}$ as follows

$$\mathrm{dR}_t = b(t, \mathrm{R}_t)\mathrm{d}t + a(t, \mathrm{R}_t)\mathrm{d}\widetilde{\mathrm{W}}_t.$$

5.5 Affine models

As we have seen in Sections 5.1 and 5.2, when the short rate R is modeled as the solution of an SDE, then zero-coupon bond prices are given by $\mathrm{B}_t^{\mathrm{T}} = \mathrm{B}(t, \mathrm{R}_t; \mathrm{T})$ where the function B satisfies PDE (5.5). Of course, whether or not PDE (5.5) can be solved explicitly depends on the coefficients a, b, and γ appearing in SDE (5.6). In this section, we describe a large class of models for which PDE (5.5) can be solved explicitly.

<u>Definition</u> 5.5.1. A model for the short rate R is said to be an *Affine Term Structure* (ATS) model if zero-coupon bond prices are of the form

$$\mathrm{B}_t^{\mathrm{T}} \equiv \mathrm{B}(t, \mathrm{R}_t; \mathrm{T}) = \exp\Big(\mathrm{G}(t;\mathrm{T}) + \mathrm{H}(t;\mathrm{T})\mathrm{R}_t\Big), \tag{5.11}$$

for some deterministic functions of time G and H.

As we must have $\mathrm{B}_{\mathrm{T}}^{\mathrm{T}} = 1$ for any value of R_{T} it follows that $\mathrm{G}(\mathrm{T};\mathrm{T}) = \mathrm{H}(\mathrm{T};\mathrm{T}) = 0$.

<u>Theorem</u> 5.5.2. *A short-rate model described under the pricing measure $\widetilde{\mathbb{P}}$ of the form*

$$\mathrm{dR}_t = b(t, \mathrm{R}_t)\mathrm{d}t + a(t, \mathrm{R}_t)\mathrm{d}\widetilde{\mathrm{W}}_t, \tag{5.12}$$

produces bond prices in the affine form (5.11) if and only if

$$b(t, \mathrm{R}_t) = b_1(t) + b_2(t)\mathrm{R}_t, \qquad a^2(t, \mathrm{R}_t) = a_1(t) + a_2(t)\mathrm{R}_t., \tag{5.13}$$

where a_1, a_2, b_1 and b_2 are deterministic functions of time and $\widetilde{\mathrm{W}}$ is a Brownian motion under the pricing measure $\widetilde{\mathbb{P}}$ (with M as numéraire). Moreover, the functions G and H in (5.11) satisfy a pair of coupled ODEs

$$0 = \mathrm{H}' + b_2\mathrm{H} + \tfrac{1}{2}a_2\mathrm{H}^2 - 1, \qquad \mathrm{H}(\mathrm{T};\mathrm{T}) = 0, \tag{5.14}$$

$$0 = \mathrm{G}' + b_1\mathrm{H} + \tfrac{1}{2}a_1\mathrm{H}^2, \qquad \mathrm{G}(\mathrm{T};\mathrm{T}) = 0. \tag{5.15}$$

PROOF. First, we show that (5.11) implies (5.13). Assume (5.11) holds. With the dynamics of R given by (5.12), bond prices must be given by $B_t^T = B(t, R_t; T)$ where B satisfies

$$0 = \Big(\partial_t - r + b(t,r)\partial_r + \tfrac{1}{2}a^2(t,r)\partial_r^2\Big)B(t,r;T).$$

To see this, simply set $\gamma = 0$ in (5.5). Inserting the expression $B(t,r;T) = \exp(G(t;T) + H(t;T)r)$ into the above PDE and using

$$\begin{aligned}
\partial_t B(t,r;T) &= B(t,r;T)\Big(\partial_t G(t;T) + \partial_t H(t,T)r\Big),\\
\partial_r B(t,r;T) &= B(t,r;T)H(t;T),\\
\partial_r^2 B(t,r;T) &= B(t,r;T)H^2(t;T)
\end{aligned}$$

we find after multiplying through by $1/B(t,r;T)$ that

$$0 = \partial_t G(t;T) + \partial_t H(t,T)r - r + b(t,r)H(t;T) + \tfrac{1}{2}a^2(t,r)H^2(t;T). \tag{5.16}$$

Differentiating the above PDE twice with respect to r yields

$$0 = \partial_r^2 b(t,r)H(t;T) + \tfrac{1}{2}\partial_r^2 a^2(t,r)H^2(t;T)$$

which implies that $\partial_r^2 b(t,r) = 0$ and $\partial_r^2 a^2(t,r) = 0$. As such b and a^2 must be of the form (5.13), as claimed.

Now we show that (5.13) implies (5.11). Assuming b and a^2 are given by (5.13), if we guess that $B(t,r;T) = \exp(G(t,T) + H(t,T)r)$, then it follows from (5.16) that

$$0 = \Big(H' + b_2 H + \tfrac{1}{2}a_2 H^2 - 1\Big)r + \Big(G' + b_1 H + \tfrac{1}{2}a_1 H^2\Big),$$

where we have grouped terms of like order in r. In order for the above equations to hold for all r, it must be the case that G and H satisfy (5.15) and (5.14), respectively. □

When bond prices are of the affine form (5.11), the yield curve is given by

$$Y_t^T = \frac{-\log B_t^T}{T-t} = \frac{-1}{T-t}\Big(G(t;T) + R_t H(t;T)\Big).$$

and the instantaneous T-maturity forward rate satisfies

$$f_t^T = -\partial_T \log B_t^T = -\Big(\partial_T G(t;T) + R_t \partial_T H(t;T)\Big).$$

5.6 Examples of Affine models

Let us take a look at some affine term-structure models.

5.6.1 Vasicek

In the Vasicek model, the dynamics of the short-rate R has the dynamics of an Ornstein-Uhlenbeck (OU) process (see Example 3.6.5). More specifically, we have

$$\mathrm{d}R_t = \kappa(\theta - R_t)\mathrm{d}t + \sigma \mathrm{d}\widetilde{W}_t, \tag{5.17}$$

where $\widetilde{W}$ is a Brownian motion under the pricing measure $\widetilde{\mathbb{P}}$ with M as numéraire. We previously established that R_t is normally distributed with a mean and variance given by

$$\widetilde{\mathbb{E}}R_t = \theta + e^{-\kappa t}(R_0 - \theta), \qquad \widetilde{\mathbb{V}}R_t = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t}).$$

Because R_t is normally distributed, there is a non-zero probability that interest rates are negative

$$\widetilde{\mathbb{P}}(R_t < 0) > 0.$$

This os one of the main criticisms of the Vasicek model.

Comparing (5.17) with (5.12)-(5.13), we identify

$$b_1 = \kappa\theta, \qquad b_2 = -\kappa, \qquad a_1 = \sigma^2, \qquad a_2 = 0.$$

As such, the ODEs for H and G, given by (5.14) and (5.15) become

$$0 = H' - \kappa H - 1, \qquad H(T;T) = 0, \tag{5.18}$$

$$0 = G' + \kappa\theta H + \tfrac{1}{2}\sigma^2 H^2, \qquad G(T;T) = 0. \tag{5.19}$$

The solution to (5.18) is

$$H(t;T) = \frac{1}{\kappa}(e^{-\kappa(T-t)} - 1).$$

Inserting the above expression for H into (5.19) and integrating yields

$$\begin{aligned} G(t;T) &= \kappa\theta \int_t^T H(s;T)\mathrm{d}s + \frac{\sigma^2}{2}\int_t^T H^2(s;t)\mathrm{d}s \\ &= \left(\frac{\theta}{\kappa} - \frac{\sigma^2}{\kappa^3}\right)(1 - e^{-\kappa(T-t)}) + \left(\frac{\sigma^2}{2\kappa^2} - \theta\right)(T-t) + \frac{\sigma^2}{4\kappa^3}(1 - e^{-2\kappa(T-t)}). \end{aligned}$$

With G and H as given above, bond prices can now be computed using (5.11).

5.6.2 Cox-Ingersoll-Ross

In the Cox-Ingresoll-Ross (CIR) model, the dynamics of the short-rate R are of the form

$$\mathrm{dR}_t = \kappa(\theta - \mathrm{R}_t)\mathrm{d}t + \sigma\sqrt{\mathrm{R}_t}\mathrm{d}\widetilde{\mathrm{W}}_t, \tag{5.20}$$

where $\widetilde{\mathrm{W}}$ is a Brownian motion under the pricing measure $\widetilde{\mathbb{P}}$ with M as numéraire. Unlike the Vasicek model, interest rates are non-negative in the CIR model

$$\widetilde{\mathbb{P}}(\mathrm{R}_t < 0) = 0.$$

This is one of the main features of the CIR model.

Comparing (5.20) with (5.12)-(5.13), we identify

$$b_1 = \kappa\theta, \qquad b_2 = -\kappa, \qquad a_1 = 0, \qquad a_2 = \sigma^2.$$

As such, the ODEs for H and G, given by (5.14) and (5.15) become

$$\begin{aligned} 0 &= \mathrm{H}' - \kappa\mathrm{H} + \tfrac{1}{2}\sigma^2\mathrm{H}^2 - 1, & \mathrm{H}(\mathrm{T};\mathrm{T}) &= 0, \\ 0 &= \mathrm{G}' + \kappa\theta\mathrm{H}, & \mathrm{G}(\mathrm{T};\mathrm{T}) &= 0. \end{aligned} \tag{5.21}$$

Equation (5.21) is known as a *Riccati equation*. Its solution is

$$\mathrm{H}(t;\mathrm{T}) = \frac{2(1-\mathrm{e}^{\gamma(\mathrm{T}-t)})}{(\gamma+\kappa)(1-\mathrm{e}^{\gamma(\mathrm{T}-t)}) + 2\gamma}, \qquad \gamma := \sqrt{\kappa^2 + 2\sigma^2}.$$

Inserting the above expression for H into (5.19) and integrating yields

$$\mathrm{G}(t;\mathrm{T}) = \frac{2\kappa\theta}{\sigma^2}\log\left(\frac{2\gamma\mathrm{e}^{(\gamma+\kappa)(\mathrm{T}-t)/2}}{(\gamma+\kappa)(\mathrm{e}^{\gamma(\mathrm{T}-t)}-1)+2\gamma}\right).$$

With G and H as given above, bond prices can now be computed using (5.11).

5.6.3 Ho-Lee

In the Ho-Lee model, the dynamics of the short-rate R are of the form

$$\mathrm{dR}_t = \theta(t)\mathrm{d}t + \sigma\mathrm{d}\widetilde{\mathrm{W}}_t, \tag{5.22}$$

where $\widetilde{\mathrm{W}}$ is a Brownian motion under the pricing measure $\widetilde{\mathbb{P}}$ with M as numéraire. Clearly, we have

$$\mathrm{R}_t = \mathrm{R}_0 + \int_0^t \theta(s)\mathrm{d}s + \sigma\widetilde{\mathrm{W}}_t.$$

It follows that the mean and variance of R_t are

$$\begin{aligned}
\widetilde{\mathbb{E}}\mathrm{R}_t &= \mathrm{R}_0 + \int_0^t \theta(s)\mathrm{d}s, \\
\widetilde{\mathbb{V}}\mathrm{R}_t &= \widetilde{\mathbb{E}}(\mathrm{R}_t - \mathbb{E}\mathrm{R}_t)^2 = \widetilde{\mathbb{E}}\sigma^2\widetilde{\mathrm{W}}_t^2 = \sigma^2 t.
\end{aligned}$$

Moreover, R_t is normally distributed. As such, similar to the Vasicek model, interest rates in the Ho-Lee model may become negative

$$\widetilde{\mathbb{P}}(\mathrm{R}_t < 0) > 0.$$

Comparing (5.22) with (5.12)-(5.13), we identify

$$b_1 = \theta(t), \qquad b_2 = 0, \qquad a_1 = 0, \qquad a_2 = \sigma^2.$$

As such, the ODEs for H and G, given by (5.14) and (5.15) become

$$\begin{aligned}
0 &= \mathrm{H}' - 1, & \mathrm{H}(\mathrm{T};\mathrm{T}) &= 0, \\
0 &= \mathrm{G}' + \theta(t)\mathrm{H} + \tfrac{1}{2}\sigma^2\mathrm{H}^2, & \mathrm{G}(\mathrm{T};\mathrm{T}) &= 0.
\end{aligned}$$

The ODEs can be solved explicitly. We have

$$\begin{aligned}
\mathrm{H}(t;\mathrm{T}) &= -(\mathrm{T}-t), \\
\mathrm{G}(t;\mathrm{T}) &= \tfrac{1}{6}\sigma^2(\mathrm{T}-t)^3 - \int_t^{\mathrm{T}} \theta(s)(\mathrm{T}-s)\mathrm{d}s.
\end{aligned}$$

One of the features of the Ho-Lee model is that it can fit the observed instantaneous forward rate curve *exactly*. To see this, observe that

$$\begin{aligned}
f_t^{\mathrm{T}} &= -\partial_{\mathrm{T}} \log \mathrm{B}_t^{\mathrm{T}} = -\partial_{\mathrm{T}}\Big(\mathrm{G}(t;\mathrm{T}) + \mathrm{H}(t;\mathrm{T})\mathrm{R}_t\Big) \\
&= -\tfrac{1}{2}\sigma^2(\mathrm{T}-t)^2 + \int_t^{\mathrm{T}} \theta(s)\mathrm{d}s + \mathrm{R}_t.
\end{aligned}$$

Differentiating the above expession with respect to T we obtain

$$\partial_{\mathrm{T}} f_t^{\mathrm{T}} = -\sigma^2(\mathrm{T}-t) + \theta(\mathrm{T}) \qquad \Rightarrow \qquad \theta(\mathrm{T}) = \partial_{\mathrm{T}} f_t^{\mathrm{T}} + \sigma^2(\mathrm{T}-t). \tag{5.23}$$

By choosing θ as in (5.23), the forward rate curve will match the observed forward rate curve exactly (though, not necessarily for at future times).

5.6.4 Hull-White

The Hull-White model is an extension of the Vasicek model with a time-dependent mean θ. Specifically, the dynamics of R in the Hull-White model are given by

$$\mathrm{dR}_t = \kappa(\theta(t)/\kappa - \mathrm{R}_t)\mathrm{d}t + \sigma \mathrm{d}\widetilde{\mathrm{W}}_t, \tag{5.24}$$

where $\widetilde{\mathrm{W}}$ is a Brownian motion under the pricing measure $\widetilde{\mathbb{P}}$ with M as numéraire. Comparing (5.24) with (5.12)-(5.13), we identify

$$b_1(t) = \theta(t), \qquad b_2 = -\kappa, \qquad a_1 = \sigma^2, \qquad a_2 = 0.$$

As such, the ODEs for H and G, given by (5.14) and (5.15) become

$$0 = \mathrm{H}' - \kappa \mathrm{H} - 1, \qquad \mathrm{H}(\mathrm{T};\mathrm{T}) = 0, \tag{5.25}$$

$$0 = \mathrm{G}' + \theta(t)\mathrm{H} + \tfrac{1}{2}\sigma^2\mathrm{H}^2, \qquad \mathrm{G}(\mathrm{T};\mathrm{T}) = 0. \tag{5.26}$$

The solution to (5.25) is

$$\mathrm{H}(t;\mathrm{T}) = \frac{1}{\kappa}(\mathrm{e}^{-\kappa(\mathrm{T}-t)} - 1).$$

Inserting the above expression for H into (5.26) and integrating yields

$$\mathrm{G}(t;\mathrm{T}) = \int_t^{\mathrm{T}} \theta(s)\mathrm{H}(s;\mathrm{T})\mathrm{d}s + \tfrac{1}{2}\sigma^2 \int_t^{\mathrm{T}} \mathrm{H}^2(s;\mathrm{T})\mathrm{d}s.$$

With G and H as given above, instantaneous forward rate curve becomes

$$\begin{aligned} f_t^{\mathrm{T}} &= -\partial_{\mathrm{T}} \log \mathrm{B}_t^{\mathrm{T}} = -\partial_{\mathrm{T}}\Big(\mathrm{G}(t;\mathrm{T}) + \mathrm{R}_t\mathrm{H}(t;\mathrm{T})\Big) \\ &= -\int_t^{\mathrm{T}} \theta(s)\partial_{\mathrm{T}}\mathrm{H}(s;\mathrm{T})\mathrm{d}s + \tfrac{1}{2}\sigma^2 \int_t^{\mathrm{T}} -\partial_{\mathrm{T}}\mathrm{H}^2(s;\mathrm{T})\mathrm{d}s - \mathrm{R}_t\partial_{\mathrm{T}}\mathrm{H}(t;\mathrm{T}) \\ &= \underbrace{-\frac{\sigma^2}{2\kappa^2}\Big(\mathrm{e}^{-\kappa(\mathrm{T}-t)} - 1\Big)^2}_{=:g(t;\mathrm{T})} - \underbrace{\int_t^{\mathrm{T}} \theta(s)\mathrm{e}^{\kappa(\mathrm{T}-s)}\mathrm{d}s + \mathrm{R}_t\mathrm{e}^{-\kappa(\mathrm{T}-t)}}_{=:\phi(t;\mathrm{T})}. \end{aligned}$$

The function ϕ defined above satisfies

$$\partial_{\mathrm{T}}\phi(t;\mathrm{T}) = -\kappa\phi(t;\mathrm{T}) + \theta(\mathrm{T}), \qquad \phi(t;t) = \mathrm{R}_t.$$

Re-arranging the terms above, we obtain

$$\begin{aligned} \theta(\mathrm{T}) &= \partial_{\mathrm{T}}\phi(t;\mathrm{T}) + \kappa\phi(t;\mathrm{T}) \\ &= -\partial_{\mathrm{T}}\Big(f_t^{\mathrm{T}} + g(t;\mathrm{T})\Big) - \kappa\Big(f_t^{\mathrm{T}} + g(t;\mathrm{T})\Big). \end{aligned}$$

By setting θ equal to the right-hand side above, the Hull-While model will produce a forward rate curve that exactly matches the market's obseived forward rate curve.

5.6.5 Multi-factor models

Although a few of the affine models described above can be calibrated to exactly fit the observed forward rate curve, the dynamics of the forward rate-curve generally are not consistent with what we observe in the market. This has led to the development of multi-factor affine short-rate models. In a general multi-factor setting, the short-rate R is given by

$$\mathrm{R}_t = \sum_{i=1}^{n} \mathrm{X}_t^{(i)},$$

where each of the individual factors satisfies

$$\mathrm{d}\mathrm{X}_t^{(i)} = b_i(t, \mathrm{X}_t^{(i)})\mathrm{d}t + a_i(t, \mathrm{X}_t^{(i)})\mathrm{d}\widetilde{\mathrm{W}}_t^{(i)}$$

with $\widetilde{\mathrm{W}}^{(i)} \perp\!\!\!\perp \widetilde{\mathrm{W}}^{(j)}$ for all $i \neq j$ and

$$b_i(t, \mathrm{X}_t^{(i)}) = b_1^{(i)} + b_2^{(i)}\mathrm{X}_t^{(i)}, \qquad a_i^2(t, \mathrm{X}_t^{(i)}) = a_1^{(i)} + a_2^{(i)}\mathrm{X}_t^{(i)}, \qquad i = 1, 2, \ldots, n.$$

Because $\widetilde{\mathrm{W}}^{(i)} \perp\!\!\!\perp \widetilde{\mathrm{W}}^{(j)}$ it follows that $\mathrm{X}^{(i)} \perp\!\!\!\perp \mathrm{X}^{(j)}$. As such, bond prices are given by

$$\begin{aligned}
\mathrm{B}_t^{\mathrm{T}} &= \widetilde{\mathbb{E}}\Big(\exp\Big(-\int_t^{\mathrm{T}} \mathrm{R}_s \mathrm{d}s \Big) \Big| \mathcal{F}_t \Big) \\
&= \widetilde{\mathbb{E}}\Big(\exp\Big(-\sum_{i=1}^{n} \int_t^{\mathrm{T}} \mathrm{X}_s^{(i)} \mathrm{d}s \Big) \Big| \mathcal{F}_t \Big) \\
&= \prod_{i=1}^{n} \widetilde{\mathbb{E}}\Big(\exp\Big(\int_t^{\mathrm{T}} \mathrm{X}_s^{(i)} \mathrm{d}s \Big) \Big| \mathrm{X}_t^{(i)} \Big) \\
&= \prod_{i=1}^{n} \exp\Big(\mathrm{G}_i(t;\mathrm{T}) + \mathrm{H}_i(t;\mathrm{T})\mathrm{X}_t^{(i)} \Big) =: \mathrm{B}(t, \mathrm{X}_t; \mathrm{T}),
\end{aligned}$$

where $\mathrm{X} = (\mathrm{X}^{(1)}, \mathrm{X}^{(2)}, \ldots, \mathrm{X}^{(n)})$ and G_i and H_i satisfy the following ODEs

$$\begin{aligned}
0 &= \mathrm{H}_i' + b_2^{(i)}\mathrm{H}_i + \tfrac{1}{2}a_2^{(i)}\mathrm{H}_i^2 - 1, && \mathrm{H}_i(\mathrm{T};\mathrm{T}) = 0, \\
0 &= \mathrm{G}_i' + b_1^{(i)}\mathrm{H}_i + \tfrac{1}{2}a_1^{(i)}\mathrm{H}_i^2, && \mathrm{G}_i(\mathrm{T};\mathrm{T}) = 0.
\end{aligned}$$

5.7 Exercises

Exercise 5.1. In the Vasicek model described in Section 5.6.1 the short rate R has dynamics

$$\mathrm{d}\mathrm{R}_t = \kappa(\theta - \mathrm{R}_t)\mathrm{d}t + \sigma\mathrm{d}\widetilde{\mathrm{W}}_t,$$

where $\widetilde{W}$ is a Brownian motion under the pricing measure $\widetilde{\mathbb{P}}$ with M as numéraire.

(a) Show that $\int_0^T R_t dt$ is a Gaussian random variable and find its mean and variance.

(b) Find B_0^T using your answer from part (a) as well as the following fact

$$\widetilde{\mathbb{E}} e^{tZ} = e^{\mu t + \sigma^2 t^2/2}, \qquad Z \sim \mathcal{N}(\mu, \sigma^2).$$

(c) Show that your answer is consistent with the bond prices derived in Section 5.6.1.

<u>Exercise</u> 5.2. In the two-factor Vasicek model, the short-rate R is give by

$$\begin{aligned} dY_t^{(1)} &= -\lambda_1 Y_t^{(1)} dt + d\widetilde{W}_t^{(1)}, \\ dY_t^{(2)} &= -\lambda_{21} Y_t^{(1)} dt - \lambda_2 Y_t^{(2)} dt + d\widetilde{W}_t^{(2)}, \\ R_t &= \delta_0 + \delta_1 Y_t^{(1)} + \delta_2 Y_t^{(2)}, \end{aligned}$$

where $\widetilde{W}^{(1)}$ and $\widetilde{W}^{(2)}$ are independent Brownian motions under the pricing measure $\widetilde{\mathbb{P}}$ with M as numéraire.

(a) As the process $(Y^{(1)}, Y^{(2)})$ is a Markov process, there exists a function $B(\cdot, \cdot, \cdot; T) : [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ such that

$$B_t^T = B(t, Y_t^{(1)}, Y_t^{(2)}; T).$$

Derive a PDE for the function B as well as an appropriate boundary condition at $t = T$.

(b) In order to solve the PDE derived in part (a), assume that the function B is of the form

$$B(t, y_1, y_2; T) = \exp\Big(G(t; T) + y_1 H_1(t; T) + y_2 H_2(t; T)\Big).$$

The functions G, H_1 and H_2 satisfy a system of coupled ODEs. Derive these ODEs as well as the terminal conditions at time $t = T$. You do *not* need to solve the ODEs.

(c) Recall that the yield of a bond is given by $Y_t^T = -(\log B_t^T)/(T - t)$. Show that

$$\begin{pmatrix} R_t \\ Y_t^{t+\Delta} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} Y_t^{(1)} \\ Y_t^{(2)} \end{pmatrix} + \begin{pmatrix} C_1 \\ C_2 \end{pmatrix},$$

and state specifically what the coefficients A_{ij} and $C_i(t; T)$ are. Your answer should be written in terms of G, H_1, H_2 and other model parameters.

(d) Because the process $(Y^{(1)}, Y^{(2)})$ is time-homogeneous (i.e., the coefficients in the SDE for $(Y^{(1)}, Y^{(2)})$ do not depend on t explicitly), it follows that, for any $t_1, t_2 \geq 0$, we have $G(t_1; t_1 + \Delta) = G(t_2; t_2 + \Delta)$ and likewise H_1 and H_2. Use this information, as well as the result from part (c) to derive an system of SDEs for R and $Y_\cdot^{\cdot+\Delta}$. That is, find SDEs $dR_t = \ldots$ and $dY_t^{t+\Delta} = \ldots$ where the right-hand sides for these equations depend on R_t and $Y_t^{t+\Delta}$, but not on $Y_t^{(1)}$ or $Y_t^{(2)}$. You may assume the 2x2 matrix A you found in part (c) is invertible.

<u>Exercise</u> 5.3. In the CIR model discussed in Section 5.6.2, the short-rate R is non-negative (i.e., $R_t \geq 0$ for all $t \geq 0$). Because the CIR model is an affine model, bond prices are given by $B_t^T = \exp(G(t;T) + R_t H(t;T))$ for some functions G and H. Use the above information to prove that $G(t;T) \leq 0$ and $H(t;T) \leq 0$ for all $t \in [0,T]$.

<u>Exercise</u> 5.4. Consider the Vasicek model described in Section 5.6.1. Let X be the value of a self-financing portfolio that replicates a T_1 maturity bond by investing in a money market account M and a T_2-maturity bond

$$dX_t = \Delta_t dB_t^{T_2} + (X_t - \Delta_t B_t^{T_2})\frac{1}{M_t}dM_t.$$

(a) Give expressions for X_0 and Δ_t. You may leave your answer in terms of G and H.
(b) Using the following parameters

$$\kappa = 0.1, \qquad \theta = 0.05, \qquad \sigma = 0.1, \qquad R_0 = 0.07. \qquad T_1 = 1.00, \qquad T_2 = 2.00,$$

write a program in the language of your choice (e.g., Matlab, Mathematica, R, Python, etc.) to simulate a path of R, B^{T_1}, and X under $\widetilde{\mathbb{P}}$ over the interval $[0, T_1]$. On a single axis, plot a path of B^{T_1}, and X. On a seperate axis, plot a path of R. Remember, to simulate a path of and SDE of the form

$$dZ_t = \mu(t, Z_t)dt + \sigma(t, Z_t)d\widetilde{W}_t,$$

simply use the Euler approximation

$$Z_{t+\delta} \approx Z_t + \mu(t, Z_t)\delta + \sigma(t, Z_t)(\widetilde{W}_{t+\delta} - \widetilde{W}_t),$$

where $\widetilde{W}_{t+\delta} - \widetilde{W}_t \sim \mathcal{N}(0, \delta)$. Take $\delta = T_1/N$ with $N = 10,000$. Be sure to use the same Brownian motion $\widetilde{W}$ for all three processes when making your plots.
(c) Run 1000 paths of the above processes (with different realizations of $\widetilde{W}$). For each path, compute $X_{T_1} - B_{T_1}^{T_1}$ and plot a histogram of the results.
INSTRUCTIONS: turn in *both* a PDF file with your plots *as well as* the file with your source code.

<u>Exercise</u> 5.5. Suppose the short rate R under $\widetilde{\mathbb{P}}$ is of the form (5.8). Consider an option that pays $g(B_{T_1}^{T_2})$ at time T_1. Denote by $V = (V_t)_{0 \leq t \leq T_1}$ the value of this option. Because the process R is a Markov process, we know that there exists a function V such that $V_t = V(t, R_t)$. What PDE and terminal condition does V satisfy?

Chapter 6

Generating short-rates from bond prices

In Chapter 5 we specified a model for the short rate R and used this to derive bond prices. An alternative approach is to specify a model for bond prices B^{T} and derive the short-rate R using the relation between the short rate, the instantaneous forward rate f^{T} and bond prices

$$\mathrm{R}_t = f_t^t, \qquad\qquad f_t^{\mathrm{T}} = -\partial_{\mathrm{T}} \log \mathrm{B}_t^{\mathrm{T}}.$$

Let us see how this can be done.

6.1 General framework

As a starting point, consider a strictly positive diffusion process A. The process A must have dynamics of the form

$$\mathrm{d}\mathrm{A}_t = \mu_t \mathrm{A}_t \mathrm{d}t + \sigma_t \mathrm{A}_t \mathrm{d}\widehat{\mathrm{W}}_t, \tag{6.1}$$

where $\widehat{\mathrm{W}}$ is a Brownian motion under some probability measure $\widehat{\mathbb{P}}$. Now, we simply *define* bond prices as follows

$$\mathrm{B}_t^{\mathrm{T}} := \frac{\mathrm{D}_t^{\mathrm{T}}}{\mathrm{A}_t}, \qquad\qquad \mathrm{D}_t^{\mathrm{T}} := \widehat{\mathbb{E}}(\mathrm{A}_{\mathrm{T}} | \mathcal{F}_t).$$

Observe that $\mathrm{B}_{\mathrm{T}}^{\mathrm{T}} = 1$ by construction. Also, as A is strictly positive, so is B^{T}. However, we do not know at this point if defining bond prices, as described above, leads to arbitrage. To find out if thee market exhibits arbitrage, we must see if there exists a probability measure $\widetilde{\mathbb{P}}$, equivalent to $\widehat{\mathbb{P}}$, under which $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$ is a martingale for every T. If $\widetilde{\mathbb{P}}$ exists, then there is no arbitrage. To this end, observe that D^{T} is a strictly positive martingale under $\widehat{\mathbb{P}}$ by construction. As such, there exists a process π^{T} such that

$$\mathrm{d}\mathrm{D}_t^{\mathrm{T}} = \pi_t^{\mathrm{T}} \mathrm{D}_t^{\mathrm{T}} \mathrm{d}\widehat{\mathrm{W}}_t.$$

Now, let us compute the dynamics of $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$. Using

$$\mathrm{d}\left(\frac{1}{\mathrm{A}_t}\right) = (\sigma_t^2 - \mu_t)\frac{1}{\mathrm{A}_t}\mathrm{d}t - \sigma_t\frac{1}{\mathrm{A}_t}\mathrm{d}\widehat{\mathrm{W}}_t, \qquad \mathrm{d}\left(\frac{1}{\mathrm{M}_t}\right) = \frac{-\mathrm{R}_t}{\mathrm{M}_t}\mathrm{d}t,$$

we find that

$$\begin{aligned}
\mathrm{d}\left(\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\right) &= \mathrm{d}\left(\frac{\mathrm{D}_t^{\mathrm{T}}}{\mathrm{A}_t\mathrm{M}_t}\right) \\
&= \frac{1}{\mathrm{A}_t\mathrm{M}_t}\mathrm{d}\mathrm{D}_t^{\mathrm{T}} + \frac{\mathrm{D}_t^{\mathrm{T}}}{\mathrm{M}_t}\mathrm{d}\left(\frac{1}{\mathrm{A}_t}\right) + \frac{\mathrm{D}_t^{\mathrm{T}}}{\mathrm{A}_t}\mathrm{d}\left(\frac{1}{\mathrm{M}_t}\right) \\
&\quad + \frac{1}{\mathrm{M}_t}\mathrm{d}\left[\mathrm{D}^{\mathrm{T}}, \frac{1}{\mathrm{A}}\right]_t + \mathrm{D}_t^{\mathrm{T}}\underbrace{\mathrm{d}\left[\frac{1}{\mathrm{A}}, \frac{1}{\mathrm{M}}\right]_t}_{=0} + \frac{1}{\mathrm{A}_t}\underbrace{\mathrm{d}\left[\mathrm{D}^{\mathrm{T}}, \frac{1}{\mathrm{M}}\right]_t}_{=0} \\
&= \frac{1}{\mathrm{A}_t\mathrm{M}_t}\pi_t^{\mathrm{T}}\mathrm{D}_t^{\mathrm{T}}\mathrm{d}\widehat{\mathrm{W}}_t + \frac{\mathrm{D}_t^{\mathrm{T}}}{\mathrm{M}_t}\left((\sigma_t^2 - \mu_t)\frac{1}{\mathrm{A}_t}\mathrm{d}t - \sigma_t\frac{1}{\mathrm{A}_t}\mathrm{d}\widehat{\mathrm{W}}_t\right) + \frac{\mathrm{D}_t^{\mathrm{T}}}{\mathrm{A}_t}\left(\frac{-\mathrm{R}_t}{\mathrm{M}_t}\mathrm{d}t\right) \\
&\quad + \frac{1}{\mathrm{M}_t}\left(\frac{-\mathrm{D}_t^{\mathrm{T}}}{\mathrm{A}_t}\sigma_t\pi_t^{\mathrm{T}}\mathrm{d}t\right) \\
&= \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\left(\sigma_t^2 - \mu_t - \mathrm{R}_t - \sigma_t\pi_t^{\mathrm{T}}\right)\mathrm{d}t + \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\left(\pi_t^{\mathrm{T}} - \sigma_t\right)\mathrm{d}\widehat{\mathrm{W}}_t
\end{aligned}$$

Now, by Girsanov's Theorem 3.7.2, if $\widetilde{\mathbb{P}}$ is equivalent to $\widehat{\mathbb{P}}$, then there exists a process γ such that the process $\widetilde{\mathrm{W}}$, defined by

$$\widetilde{\mathrm{W}}_t := \widehat{\mathrm{W}}_t + \int_0^t \gamma_s \mathrm{d}s,$$

is a Brownian motion under $\widetilde{\mathbb{P}}$. The dynamics of $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$ under $\widetilde{\mathbb{P}}$ are

$$\begin{aligned}
\mathrm{d}\left(\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\right) &= \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\left(\sigma_t^2 - \mu_t - \mathrm{R}_t - \sigma_t\pi_t^{\mathrm{T}}\right)\mathrm{d}t + \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\left(\pi_t^{\mathrm{T}} - \sigma_t\right)\left(\mathrm{d}\widetilde{\mathrm{W}}_t - \gamma_t\mathrm{d}t\right) \\
&= \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\left((\sigma_t + \gamma_t)(\sigma_t - \pi_t^{\mathrm{T}}) - \mu_t - \mathrm{R}_t\right)\mathrm{d}t + \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\left(\pi_t^{\mathrm{T}} - \sigma_t\right)\mathrm{d}\widetilde{\mathrm{W}}_t
\end{aligned}$$

Again, in order for the market to be arbitrage free, we must have that $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$ is a $\widetilde{\mathbb{P}}$-martingale for every T. This means that the $\mathrm{d}t$-term must equal zero for every T. The only way for this to happen is for

$$\mathrm{R}_t = -\mu_t, \qquad \gamma_t = -\sigma_t. \tag{6.2}$$

With R and γ as described above $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$ is a $\widetilde{\mathbb{P}}$ martingale for every T (i.e., there is no arbitrage in the market), and the dynamics of $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$ are

$$\mathrm{d}\left(\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\right) = \left(\pi_t^{\mathrm{T}} - \sigma_t\right)\left(\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\right)\mathrm{d}\widetilde{\mathrm{W}}_t.$$

Now, observe from (6.2) that, if we want interest rates to be non-negative $\mathrm{R} \geq 0$ we must have that $\mu \leq 0$. Let us now see how the above framework can be applied in a few examples.

6.2 Example 1

Suppose A is given by

$$\mathrm{A}_t = f(t) + g(t)\mathrm{M}_t, \qquad \mathrm{dM}_t = h(t)\mathrm{M}_t\mathrm{d}\widehat{\mathrm{W}}_t, \qquad \mathrm{M}_0 = 1,$$

where f, g, and h are positive deterministic functions and f and g are strictly decreasing. The dynamics of A are

$$\begin{aligned}
\mathrm{dA}_t &= \big(f'(t) + g'(t)\mathrm{M}_t\big)\mathrm{d}t + g(t)\mathrm{dM}_t \\
&= \Big(\frac{f'(t) + g'(t)\mathrm{M}_t}{f(t) + g(t)\mathrm{M}_t}\Big)\mathrm{A}_t\mathrm{d}t + g(t)h(t)\mathrm{M}_t\mathrm{d}\widehat{\mathrm{W}}_t \\
&= \Big(\frac{f'(t) + g'(t)\mathrm{M}_t}{f(t) + g(t)\mathrm{M}_t}\Big)\mathrm{A}_t\mathrm{d}t + \Big(\frac{g(t)h(t)\mathrm{M}_t}{f(t) + g(t)\mathrm{M}_t}\Big)\mathrm{A}_t\mathrm{d}\widehat{\mathrm{W}}_t. \qquad (6.3)
\end{aligned}$$

Comparing (6.3) with (6.1) we identify

$$\mu_t = \frac{f'(t) + g'(t)\mathrm{M}_t}{f(t) + g(t)\mathrm{M}_t}.$$

Noting that

$$\mathrm{M}_t = \exp\Big(-\tfrac{1}{2}\int_0^t h^2(s)\mathrm{d}s + \int_0^t h(s)\mathrm{d}\widehat{\mathrm{W}}_s\Big) > 0.$$

we have

$$\mathrm{R}_t = -\mu_t = \frac{-f'(t) - g'(t)\mathrm{M}_t}{f(t) + g(t)\mathrm{M}_t} > 0, \qquad (6.4)$$

where we have used the fact that $f' < 0$ and $g' < 0$ because f and g are decreasing by assumption. Bond prices are given by

$$\begin{aligned}
\mathrm{B}_t^{\mathrm{T}} = \frac{\widehat{\mathbb{E}}(\mathrm{A}_{\mathrm{T}}|\mathcal{F}_t)}{\mathrm{A}_t} &= \frac{f(\mathrm{T}) + g(\mathrm{T})\widehat{\mathbb{E}}(\mathrm{M}_{\mathrm{T}}|\mathcal{F}_t)}{f(t) + g(t)\mathrm{M}_t} \\
&= \frac{f(\mathrm{T}) + g(\mathrm{T})\mathrm{M}_t}{f(t) + g(t)\mathrm{M}_t},
\end{aligned}$$

where we have used the fact that M is a martingale under $\widehat{\mathbb{P}}$. As a sanity check, we note that

$$\begin{aligned}
\mathrm{R}_t = f_t^t = -\partial_{\mathrm{T}} \log \mathrm{B}_t^{\mathrm{T}}\Big|_{\mathrm{T}=t} &= -\partial_{\mathrm{T}} \log \frac{f(\mathrm{T}) + g(\mathrm{T})\mathrm{M}_t}{f(t) + g(t)\mathrm{M}_t}\Big|_{\mathrm{T}=t} \\
&= \frac{-f'(\mathrm{T}) - g'(\mathrm{T})\mathrm{M}_t}{f(\mathrm{T}) + g(\mathrm{T})\mathrm{M}_t}\Big|_{\mathrm{T}=t} = \frac{-f'(t) - g'(t)\mathrm{M}_t}{f(t) + g(t)\mathrm{M}_t},
\end{aligned}$$

which agrees with (6.4).

6.3 Example 2

Suppose A is given by

$$\mathrm{A}_t = \mathrm{e}^{-at} h(\mathrm{X}_t), \qquad \mathrm{d}\mathrm{X}_t = b(t,\mathrm{X}_t)\mathrm{d}t + c(t,\mathrm{X}_t)\mathrm{d}\widehat{\mathrm{W}}_t,$$

where $a > 0$. The dynamics of A are can be computed as follows

$$\begin{aligned}
\mathrm{d}\mathrm{A}_t &= -a\mathrm{e}^{-at} h(\mathrm{X}_t)\mathrm{d}t + \mathrm{e}^{-at}\Big(b(t,\mathrm{X}_t)h'(\mathrm{X}_t) + \tfrac{1}{2}c^2(t,\mathrm{X}_t)h''(\mathrm{X}_t)\Big)\mathrm{d}t \\
&\quad + \mathrm{e}^{-at} c(t,\mathrm{X}_t)h'(\mathrm{X}_t)\mathrm{d}\widetilde{\mathrm{W}}_t \\
&= \Big(-a + \frac{b(t,\mathrm{X}_t)h'(\mathrm{X}_t) + \tfrac{1}{2}c^2(t,\mathrm{X}_t)h''(\mathrm{X}_t)}{h(\mathrm{X}_t)}\Big)\mathrm{A}_t\mathrm{d}t \\
&\quad + \frac{c(t,\mathrm{X}_t)h'(\mathrm{X}_t)}{h(\mathrm{X}_t)}\mathrm{A}_t\mathrm{d}\widetilde{\mathrm{W}}_t. \qquad (6.5)
\end{aligned}$$

Comparing (6.5) with (6.1), we identify

$$\mu_t = -a + \frac{b(t,\mathrm{X}_t)h'(\mathrm{X}_t) + \tfrac{1}{2}c^2(t,\mathrm{X}_t)h''(\mathrm{X}_t)}{h(\mathrm{X}_t)}.$$

Now, suppose we choose

$$b(t,x) = -\kappa x, \qquad c(t,x) = 1, \qquad h(x) = \cosh(\gamma x),$$

where $\kappa > 0$ and $\gamma > 0$. Using the fact that $h'(x)/h(x) = \gamma\tanh(\gamma x)$ and $h''(x)/h(x) = \gamma^2$, we have

$$\mathrm{R}_t = -\mu_t = a + \kappa\gamma\mathrm{X}_t\tanh(\gamma\mathrm{X}_t) - \tfrac{1}{2}\gamma^2. \qquad (6.6)$$

Noting that $\gamma x\tanh(\gamma x) \geq 0$ for all $x \in \mathbb{R}$, we see that if $\gamma^2/2 \leq a$ then $\mathrm{R} \geq 0$. Bond prices can be computed by noting that X is an OU process

$$\mathrm{d}\mathrm{X}_t = -\kappa\mathrm{X}_t\mathrm{d}t + \mathrm{d}\widehat{\mathrm{W}}_t.$$

We have from Example 3.6.5 that

$$\mathrm{X}_\mathrm{T} = \mathrm{e}^{-\kappa(\mathrm{T}-t)}\mathrm{X}_t + \int_t^\mathrm{T} \mathrm{e}^{-\kappa(\mathrm{T}-s)}\mathrm{d}\widehat{\mathrm{W}}_s.$$

from which it follows that

$$\mathrm{X}_\mathrm{T}|\mathcal{F}_t \sim \mathcal{N}(m_t, v_t^2), \qquad m_t = \mathrm{e}^{-\kappa(\mathrm{T}-t)}\mathrm{X}_t, \qquad v_t^2 = \frac{1}{2\kappa}\Big(\mathrm{e}^{-2\kappa(\mathrm{T}-t)} - 1\Big). \qquad (6.7)$$

Using the fact that

$$\widehat{\mathbb{E}}(\mathrm{e}^{\gamma\mathrm{X}_\mathrm{T}}|\mathcal{F}_t) = \mathrm{e}^{\gamma m_t + \frac{1}{2}\gamma^2 v_t^2}, \qquad \cosh(\gamma\mathrm{X}_\mathrm{T}) = \frac{1}{2}\Big(\mathrm{e}^{\gamma\mathrm{X}_\mathrm{T}} + \mathrm{e}^{-\gamma\mathrm{X}_\mathrm{T}}\Big),$$

we find that bond prices are given by

$$\begin{aligned} \mathrm{B}_t^{\mathrm{T}} &= \frac{\widehat{\mathbb{E}}(\mathrm{A}_{\mathrm{T}}|\mathcal{F}_t)}{\mathrm{A}_t} = \frac{\widehat{\mathbb{E}}(\mathrm{e}^{-a\mathrm{T}}\cosh(\gamma \mathrm{X}_{\mathrm{T}})|\mathcal{F}_t)}{\mathrm{e}^{-at}\cosh(\gamma \mathrm{X}_t)} \\ &= \frac{\mathrm{e}^{-a(\mathrm{T}-t)}}{2\cosh(\gamma \mathrm{X}_t)}\left(\mathrm{e}^{\gamma m_t + \frac{1}{2}\gamma^2 v_t^2} + \mathrm{e}^{-\gamma m_t + \frac{1}{2}\gamma^2 v_t^2}\right) \end{aligned}$$

where m_t and v_t^2 are given by (6.7). We leave it as an exercise for the reader to compute R_t from $\mathrm{B}_t^{\mathrm{T}}$ and verify that the result agrees with (6.6).

6.4 Exercises

Exercise 6.1. Derive R_t from $\mathrm{B}_t^{\mathrm{T}}$ when bond prices are as described in Section 6.3.

Exercise 6.2. Suppose that, for the class of models described in Section 6.3 we have

$$h(x) = \mathrm{e}^{\gamma x}, \qquad b(t,x) = \kappa(\theta - x), \qquad c(t,x) = \delta\sqrt{x}.$$

where γ, κ, θ, and δ are all positive. Observe that X is a CIR process, as discussed in Section 5.6.2. (a) Derive conditions under which this model process non-negative interest rates.

(b) If $\gamma < 0$, can we still guarantee non-negative interest rates?

(c) Suppose that $h(x) = \mathrm{e}^{\gamma/x}$. Is it still possible for interest rates to be non-negative? If so, under what conditions? You may assume $2\kappa\theta > \delta^2$, which guarantees that X remains strictly positive.

Exercise 6.3. A *Brownian bridge from x_0 to $\bar{x}$ on the time inteval* $[0, \overline{\mathrm{T}}]$, is a process $\mathrm{X} = (\mathrm{X}_t)_{[0,\overline{\mathrm{T}}]}$ given by

$$\mathrm{X}_t = x_0 + \frac{t}{\overline{\mathrm{T}}}(\bar{x} - x_0) + (\overline{\mathrm{T}} - t)\int_0^t \frac{1}{\overline{\mathrm{T}} - s}\mathrm{d}\widehat{\mathrm{W}}_s, \qquad 0 \le t \le \overline{\mathrm{T}}.$$

Suppose that, in the setting of Section 6.3, the process X is a Brownian Bridge as described above and

$$h(x) = \cosh(\gamma x).$$

(a) Compute the interest rate R_t in this setting where $t \le \overline{\mathrm{T}}$.

(b) Are any restrictions needed to guarantee that R is non-negative? If so, what are the restrictions?

(b) Compute the bond price $\mathrm{B}_t^{\mathrm{T}}$ where $\mathrm{T} \le \overline{\mathrm{T}}$.

Hint: To start this exercise, compute $\mathrm{dX}_t = \ldots$.

Chapter 7

Heath-Jarrow-Morton framework

In the Chapter 5 , we deduced bond prices by modeling the short-rate R directly under the risk-neutral probability measure $\widetilde{\mathbb{P}}$ (with M as numéraire), and deduced bond prices using risk-neutral pricing

$$\frac{B_t^T}{M_T} = \widetilde{\mathbb{E}}\Big(\frac{B_T^T}{M_T}\Big|\mathcal{F}_t\Big) \qquad \Rightarrow \qquad B_t^T = \widetilde{\mathbb{E}}\Big(\exp\Big(-\int_t^T R_s \mathrm{d}s\Big)\Big|\mathcal{F}_t\Big).$$

But, recall that bond prices can alternatively be deduced from the instantaneous forward rate cure using

$$B_t^T = \exp\Big(-\int_t^T f_t^s \mathrm{d}s\Big).$$

This begs the question: rather than deduce bond prices from the short rate R, why not model the forward rate curve f^T directly? This is precisely the approach taken by Heath, Jarrow and Morton (HJM).

In the HJM setting, the dynamics of the forward rate curve f^T are given by

$$\mathrm{d}f_t^T = \theta_t^T \mathrm{d}t + \sigma_t^T \mathrm{d}W_t, \tag{7.1}$$

where W is a Brownian motion under the real-world probability measure $\mathbb{P}$. Note that the above equation is actually infinitely many equations – one for each maturity $T \geq 0$. Let us work out the dynamics of the bond price B^T in this setting. First, noting that

$$B_t^T = \exp\Big(-J_t^T\Big), \qquad J_t^T := \int_t^T f_t^s \mathrm{d}s,$$

we compute

$$\begin{aligned}
\mathrm{d}J_t^T &= \mathrm{d}\int_t^T f_t^s \mathrm{d}s \\
&= -f_t^t \mathrm{d}t + \int_t^T \mathrm{d}f_t^s \mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&= -f_t^t \mathrm{d}t + \int_t^{\mathrm{T}} \Big(\theta_t^s \mathrm{d}t + \sigma_t^s \mathrm{dW}_t\Big)\mathrm{d}s \\
&= -\mathrm{R}_t \mathrm{d}t + \Big(\int_t^{\mathrm{T}} \theta_t^s \mathrm{d}s\Big)\mathrm{d}t + \Big(\int_t^{\mathrm{T}} \sigma_t^s \mathrm{d}s\Big)\mathrm{dW}_t \\
&= -\mathrm{R}_t \mathrm{d}t + \Theta_t^{\mathrm{T}} \mathrm{d}t + \Sigma_t^{\mathrm{T}} \mathrm{dW}_t,
\end{aligned}
$$

where we have defined

$$
\Theta_t^{\mathrm{T}} := \int_t^{\mathrm{T}} \theta_t^s \mathrm{d}s, \qquad \Sigma_t^{\mathrm{T}} := \int_t^{\mathrm{T}} \sigma_t^s \mathrm{d}s.
$$

Note that we have also used the fact that $f_t^t = \mathrm{R}_t$. Next, the dynamics of B^{T} are given by

$$
\begin{aligned}
\mathrm{dB}_t^{\mathrm{T}} &= \mathrm{d}\exp\Big(-\mathrm{J}_t^{\mathrm{T}}\Big) \\
&= -\mathrm{B}_t^{\mathrm{T}} \mathrm{dJ}_t^{\mathrm{T}} + \tfrac{1}{2}\mathrm{B}_t^{\mathrm{T}} \mathrm{d}[\mathrm{J}^{\mathrm{T}}, \mathrm{J}^{\mathrm{T}}]_t \\
&= -\mathrm{B}_t^{\mathrm{T}}\Big(-\mathrm{R}_t + \Theta_t^{\mathrm{T}} \mathrm{d}t + \Sigma_t^{\mathrm{T}} \mathrm{dW}_t\Big) + \tfrac{1}{2}\mathrm{B}_t^{\mathrm{T}}(\Sigma_t^{\mathrm{T}})^2 \mathrm{d}t \\
&= \Big(\mathrm{R}_t - \Theta_t^{\mathrm{T}} + \tfrac{1}{2}(\Sigma_t^{\mathrm{T}})^2\Big)\mathrm{B}_t^{\mathrm{T}} \mathrm{d}t - \Sigma_t^{\mathrm{T}} \mathrm{B}_t^{\mathrm{T}} \mathrm{dW}_t. \qquad (7.2)
\end{aligned}
$$

Thus, we have deduced the dynamics of B^{T} under the real-world probability measure $\mathbb{P}$.

7.1 No arbitrage condition

A natural question that comes to mind at this point is: does this market allow for arbitrage? To answer this question, we fix the money market account M as numéraire and we ask: is there are probability measure $\widetilde{\mathbb{P}}$, equivalent to $\mathbb{P}$, under which $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$ is a martingale? First we compute the dynamics of $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$ under $\mathbb{P}$. We have

$$
\begin{aligned}
\mathrm{d}\Big(\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\Big) &= \frac{1}{\mathrm{M}_t}\mathrm{dB}_t^{\mathrm{T}} + \mathrm{B}_t^{\mathrm{T}} \mathrm{d}\Big(\frac{1}{\mathrm{M}_t}\Big) + \mathrm{d}\Big[\mathrm{B}^{\mathrm{T}}, \frac{1}{\mathrm{M}}\Big]_t \\
&= \Big(-\Theta_t^{\mathrm{T}} + \tfrac{1}{2}(\Sigma_t^{\mathrm{T}})^2\Big)\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\mathrm{d}t - \Sigma_t^{\mathrm{T}} \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\mathrm{dW}_t. \qquad (7.3)
\end{aligned}
$$

Now, recall from Girsanov's theorem that, under the following change of measure

$$
\frac{\mathrm{d}\widetilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}} := \exp\Big(-\tfrac{1}{2}\int_0^{\mathrm{T}} \gamma_t^2 \mathrm{d}t - \int_0^{\mathrm{T}} \gamma_t \mathrm{dW}_t\Big),
$$

the process $\widetilde{\mathrm{W}}$ defined by

$$
\widetilde{\mathrm{W}}_t = \mathrm{W}_t + \int_0^t \gamma_s \mathrm{d}s, \qquad (7.4)
$$

is a $\widetilde{\mathbb{P}}$ Brownian motion. The dynamics of $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$ under $\widetilde{\mathbb{P}}$ are

$$\begin{aligned} \mathrm{d}\Big(\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\Big) &= \Big(-\Theta_t^{\mathrm{T}} + \tfrac{1}{2}(\Sigma_t^{\mathrm{T}})^2\Big)\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\mathrm{d}t - \Sigma_t^{\mathrm{T}}\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\Big(\mathrm{d}\widetilde{\mathrm{W}}_t - \gamma_t \mathrm{d}t\Big) \\ &= \Big(-\Theta_t^{\mathrm{T}} + \tfrac{1}{2}(\Sigma_t^{\mathrm{T}})^2 + \gamma_t\Sigma_t^{\mathrm{T}}\Big)\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\mathrm{d}t - \Sigma_t^{\mathrm{T}}\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\mathrm{d}\widetilde{\mathrm{W}}_t. \end{aligned}$$

In order for $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$ to be a martingale, the $\mathrm{d}t$-term must be zero. As such, we must have

$$0 = -\Theta_t^{\mathrm{T}} + \tfrac{1}{2}(\Sigma_t^{\mathrm{T}})^2 + \gamma_t\Sigma_t^{\mathrm{T}}. \tag{7.5}$$

Note that the above equation must hold for *every* maturity T, but the process γ cannot depend on T. Differentiating the above equation with respect to T, we find

$$0 = -\theta_t^{\mathrm{T}} + \sigma_t^{\mathrm{T}}\Sigma_t^{\mathrm{T}} + \gamma_t\sigma_t^{\mathrm{T}}. \tag{7.6}$$

Note that if γ satisfies (7.6) then it also solves (7.5). So see this, we simply integrate (7.6) with respect to T from t to T' which yields

$$0 = -\Theta_t^{\mathrm{T}}\Big|_{\mathrm{T}=t}^{\mathrm{T}=\mathrm{T}'} + \tfrac{1}{2}(\Sigma_t^{\mathrm{T}})^2\Big|_{\mathrm{T}=t}^{\mathrm{T}=\mathrm{T}'} + \gamma_t\Sigma_t^{\mathrm{T}}\Big|_{\mathrm{T}=t}^{\mathrm{T}=\mathrm{T}'} = -\Theta_t^{\mathrm{T}'} + \tfrac{1}{2}(\Sigma_t^{\mathrm{T}'})^2 + \gamma_t\Sigma_t^{\mathrm{T}'},$$

where we have used the fact that $\Theta_t^t = \Sigma_t^t = 0$. We have derived the following result.

<u>Theorem</u> 7.1.1. *Assume forward rates are given by* (7.1)*. If there exists a process* $\gamma = (\gamma_t)_{t\geq 0}$ *such that* (7.6) *for all* T*, then there is no arbitrage.*

Now, let us assume that there is no arbitrage. What are the dynamics of f^{T} under $\widetilde{\mathbb{P}}$? From (7.1) and (7.4), we have

$$\begin{aligned} \mathrm{d}f_t^{\mathrm{T}} &= \theta_t^{\mathrm{T}}\mathrm{d}t + \sigma_t^{\mathrm{T}}\Big(\mathrm{d}\widetilde{\mathrm{W}}_t - \gamma_t\mathrm{d}t\Big) \\ &= \Big(\theta_t^{\mathrm{T}} - \gamma_t\sigma_t^{\mathrm{T}}\Big)\mathrm{d}t + \sigma_t^{\mathrm{T}}\mathrm{d}\widetilde{\mathrm{W}}_t \\ &= \sigma_t^{\mathrm{T}}\Sigma_t^{\mathrm{T}}\mathrm{d}t + \sigma_t^{\mathrm{T}}\mathrm{d}\widetilde{\mathrm{W}}_t, \end{aligned} \tag{7.7}$$

where, in the last line, we have used the no-arbitrage condition (7.6). Note that under $\widetilde{\mathbb{P}}$, the drift (i.e., the coefficient of the $\mathrm{d}t$-term) is fixed by the volatility (i.e., the coefficient of the $\mathrm{d}\widetilde{\mathrm{W}}_t$-term). This is analagous to the situation in equity modeling, in which the drift of a stock with volatility σ_t must be $\mathrm{R}_t - \frac{1}{2}\sigma_t^2$ under $\widetilde{\mathbb{P}}$.

We can also derive the dynamics of B^{T} under $\widetilde{\mathbb{P}}$. From (7.8) we have

$$\mathrm{d}\mathrm{B}_t^{\mathrm{T}} = \Big(\mathrm{R}_t - \Theta_t^{\mathrm{T}} + \tfrac{1}{2}(\Sigma_t^{\mathrm{T}})^2\Big)\mathrm{B}_t^{\mathrm{T}}\mathrm{d}t - \Sigma_t^{\mathrm{T}}\mathrm{B}_t^{\mathrm{T}}\Big(\mathrm{d}\widetilde{\mathrm{W}}_t - \gamma_t\mathrm{d}t\Big)$$

$$
\begin{aligned}
&= \Big(\mathrm{R}_t - \Theta_t^{\mathrm{T}} + \tfrac{1}{2}(\Sigma_t^{\mathrm{T}})^2 + \gamma_t \Sigma_t^{\mathrm{T}} \Big) \mathrm{B}_t^{\mathrm{T}} \mathrm{d}t - \Sigma_t^{\mathrm{T}} \mathrm{B}_t^{\mathrm{T}} \mathrm{d}\widetilde{\mathrm{W}}_t \\
&= \mathrm{R}_t \mathrm{B}_t^{\mathrm{T}} \mathrm{d}t - \Sigma_t^{\mathrm{T}} \mathrm{B}_t^{\mathrm{T}} \mathrm{d}\widetilde{\mathrm{W}}_t, \qquad (7.8)
\end{aligned}
$$

where in the last line, we have used (7.5). Lastly, using (7.3) the dynamics of $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$ under $\widetilde{\mathbb{P}}$ are

$$
\begin{aligned}
\mathrm{d}\Big(\frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t}\Big) &= \Big(-\Theta_t^{\mathrm{T}} + \tfrac{1}{2}(\Sigma_t^{\mathrm{T}})^2 \Big) \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t} \mathrm{d}t - \Sigma_t^{\mathrm{T}} \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t} \Big(\mathrm{d}\widetilde{\mathrm{W}}_t - \gamma_t \mathrm{d}t \Big) \\
&= \Big(-\Theta_t^{\mathrm{T}} + \tfrac{1}{2}(\Sigma_t^{\mathrm{T}})^2 + \gamma_t \Sigma_t^{\mathrm{T}} \Big) \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t} \mathrm{d}t - \Sigma_t^{\mathrm{T}} \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t} \mathrm{d}\widetilde{\mathrm{W}}_t \\
&= -\Sigma_t^{\mathrm{T}} \frac{\mathrm{B}_t^{\mathrm{T}}}{\mathrm{M}_t} \mathrm{d}\widetilde{\mathrm{W}}_t, \qquad (7.9)
\end{aligned}
$$

where we have once again used (7.5). Observe that $\mathrm{B}^{\mathrm{T}}/\mathrm{M}$ is a $\widetilde{\mathbb{P}}$ martingale, as it must be to exclude arbitrage.

7.2 Relation of HJM to short-rate models

It turns out that every short-rate model driven by a single Brownian motion is, in fact, an HJM model. Consider the class of affine short-rate models described in Section 5.5. Forward rates, in this setting, are given by

$$
f_t^{\mathrm{T}} = -\partial_{\mathrm{T}} \mathrm{G}(t;\mathrm{T}) - \mathrm{R}_t \partial_{\mathrm{T}} \mathrm{H}(t;\mathrm{T}),
$$

where G and H solve (5.15) and (5.14), respectively. The dynamics of f_t^{T} are given by

$$
\begin{aligned}
\mathrm{d}f_t^{\mathrm{T}} &= -\partial_{\mathrm{T}}\partial_t \mathrm{G}(t;\mathrm{T})\mathrm{d}t - \mathrm{R}_t \partial_{\mathrm{T}}\partial_t \mathrm{H}(t;\mathrm{T})\mathrm{d}t - \partial_{\mathrm{T}}\mathrm{H}(t;\mathrm{T})\mathrm{d}\mathrm{R}_t \\
&= -\partial_{\mathrm{T}}\partial_t \mathrm{G}(t;\mathrm{T})\mathrm{d}t - \mathrm{R}_t \partial_{\mathrm{T}}\partial_t \mathrm{H}(t;\mathrm{T})\mathrm{d}t - \partial_{\mathrm{T}}\mathrm{H}(t;\mathrm{T})\Big(b(t,\mathrm{R}_t)\mathrm{d}t + a(t,\mathrm{R}_t)\mathrm{d}\widetilde{\mathrm{W}}_t \Big) \\
&= -\Big(\partial_{\mathrm{T}}\partial_t \mathrm{G}(t;\mathrm{T}) + \mathrm{R}_t \partial_{\mathrm{T}}\partial_t \mathrm{H}(t;\mathrm{T}) + b(t,\mathrm{R}_t)\partial_{\mathrm{T}}\mathrm{H}(t;\mathrm{T}) \Big)\mathrm{d}t - a(t,\mathrm{R}_t)\partial_{\mathrm{T}}\mathrm{H}(t;\mathrm{T})\mathrm{d}\widetilde{\mathrm{W}}_t. \qquad (7.10)
\end{aligned}
$$

Comparing (7.10) with (7.7), we identify

$$
\begin{aligned}
\sigma_t^{\mathrm{T}} &= -a(t,\mathrm{R}_t)\partial_{\mathrm{T}}\mathrm{H}(t;\mathrm{T}), \qquad (7.11) \\
\sigma_t^{\mathrm{T}}\Sigma_t^{\mathrm{T}} &= -\Big(\partial_{\mathrm{T}}\partial_t \mathrm{G}(t;\mathrm{T}) + \mathrm{R}_t \partial_{\mathrm{T}}\partial_t \mathrm{H}(t;\mathrm{T}) + b(t,\mathrm{R}_t)\partial_{\mathrm{T}}\mathrm{H}(t;\mathrm{T}) \Big). \qquad (7.12)
\end{aligned}
$$

Let us verify in a simple example that, when σ_t^{T} is given by the right-hand side of (7.11), then (7.12) holds.

7.3 Example: Vasicek model

Recall the Vasicek model from Section 5.6.1. We have

$$\begin{aligned} b(t,\mathrm{R}_t) &= b_1(t) + b_2(t)\mathrm{R}_t, & b_1(t) &= \kappa\theta, & b_2(t) &= -\kappa, \\ a^2(t,\mathrm{R}_t) &= a_1(t) + a_2(t)\mathrm{R}_t, & a_1(t) &= \sigma^2, & a_2(t) &= 0, \end{aligned} \tag{7.13}$$

and the functions G and H are given by

$$\begin{aligned} \mathrm{H}(t;\mathrm{T}) &= \frac{1}{\kappa}(\mathrm{e}^{-\kappa(\mathrm{T}-t)} - 1), \\ \mathrm{G}(t;\mathrm{T}) &= \kappa\theta \int_t^{\mathrm{T}} \mathrm{H}(s;\mathrm{T})\mathrm{d}s + \frac{\sigma^2}{2}\int_t^{\mathrm{T}} \mathrm{H}^2(s;\mathrm{T})\mathrm{d}s. \end{aligned}$$

Thus, we have

$$\sigma_t^{\mathrm{T}} = -a(t,\mathrm{R}_t)\partial_{\mathrm{T}}\mathrm{H}(t;\mathrm{T}) = -\sigma\partial_{\mathrm{T}}\Big(\frac{1}{\kappa}(\mathrm{e}^{-\kappa(\mathrm{T}-t)} - 1)\Big) = \sigma\mathrm{e}^{-\kappa(\mathrm{T}-t)}.$$

It follows that

$$\sigma_t^{\mathrm{T}}\Sigma_t^{\mathrm{T}} = \sigma_t^{\mathrm{T}}\int_t^{\mathrm{T}} \sigma_t^s\mathrm{d}s = \sigma\mathrm{e}^{-\kappa(\mathrm{T}-t)}\int_t^{\mathrm{T}} \sigma\mathrm{e}^{-\kappa(s-t)}\mathrm{d}s = \frac{\sigma^2}{\kappa}\Big(\mathrm{e}^{-\kappa(\mathrm{T}-t)} - \mathrm{e}^{-2\kappa(\mathrm{T}-t)}\Big). \tag{7.14}$$

Now, observe that

$$\begin{aligned} \partial_{\mathrm{T}}\mathrm{H}(t;\mathrm{T}) &= -\mathrm{e}^{-\kappa(\mathrm{T}-t)}, \\ \partial_{\mathrm{T}}\partial_t\mathrm{H}(t;\mathrm{T}) &= -\kappa\mathrm{e}^{-\kappa(\mathrm{T}-t)}, \\ \partial_{\mathrm{T}}\partial_t\mathrm{G}(t;\mathrm{T}) &= -\partial_{\mathrm{T}}\Big(\kappa\theta\mathrm{H}(t;\mathrm{T}) + \frac{\sigma^2}{2}\mathrm{H}^2(t;\mathrm{T})\Big) \\ &= \kappa\theta\mathrm{e}^{-\kappa(\mathrm{T}-t)} - \frac{\sigma^2}{\kappa}\Big(\mathrm{e}^{-2\kappa(\mathrm{T}-t)} - \mathrm{e}^{-\kappa(\mathrm{T}-t)}\Big). \end{aligned}$$

Inserting the above expressions into the right-hand side of (7.12) and using (7.13) we find

$$\text{R.H.S of (7.12)} = \frac{\sigma^2}{\kappa}\Big(\mathrm{e}^{-\kappa(\mathrm{T}-t)} - \mathrm{e}^{-2\kappa(\mathrm{T}-t)}\Big),$$

which agrees with the right-hand side of (7.14).

7.4 Exercises

Exercise 7.1. Suppose the dynamics of the forward rate curve f^{T} are given by

$$\mathrm{d}f_t^{\mathrm{T}} = \theta_t^{\mathrm{T}}\mathrm{d}t + \sum_{i=1}^{d}\sigma_t^{\mathrm{T},(i)}\mathrm{dW}_t^{(i)},$$

where $W = (W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(d)})_{t\geq 0}$ is a d-dimension Brownian motion under the real-world probability measure $\mathbb{P}$ with independent components. Consider the following change of measure

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = \prod_{i=1}^{d} \exp\Big(-\frac{1}{2}\int_0^T (\gamma_t^{(i)})^2 dt - \int_0^T \gamma_t^{(i)} dW_t \Big),$$

Under $\widetilde{\mathbb{P}}$ the process $\widetilde{W} = (\widetilde{W}_t^{(1)}, \widetilde{W}_t^{(2)}, \ldots, \widetilde{W}_t^{(d)})_{t\geq 0}$ defined by

$$\widetilde{W}_t^{(i)} = \int_0^t \gamma_s^{(i)} ds + W_t, \qquad i = 1, 2, \ldots, d,$$

is a d-dimension Brownian motion.

(a) Derive the dynamics of B^T/M under $\widetilde{\mathbb{P}}$ (i.e., $d(B_t^T/M_t) = \ldots$). Please give you answer in terms of processes

$$\Theta_t^T := \int_t^T \theta_t^s ds, \qquad \Sigma_t^{T,(i)} := \int_t^T \sigma_t^{s,(i)} ds.$$

(b) Show that the no-arbitrage condition is the existence of a process $\gamma = (\gamma_t^{(1)}, \gamma_t^{(2)}, \ldots, \gamma_t^{(d)})_{t\geq 0}$ such that

$$0 = -\theta_t^T + \sum_{i=1}^{d} \sigma_t^{T,(i)} (\Sigma_t^{T,(i)} + \gamma_t^{(i)}).$$

<u>Exercise</u> 7.2. Suppose the dynamics of the forward rate are given by (7.7) with

$$f_0^T = \lambda_0 + \lambda_1 e^{-\gamma T} - \frac{\eta^2}{2\gamma^2}(1 - e^{-\gamma T})^2,$$
$$\sigma_t^T = \eta e^{-\gamma (T-t)}.$$

Show that the short rate R admits the form

$$R_t = g(t, R_0) + \int_0^t h(s; t) d\widetilde{W}_s,$$

and identify the functions g and h.

<u>Exercise</u> 7.3. Suppose the dynamics of the forward rate are given by (7.7) with $\sigma_t^T = \sigma(t; T)$ a deterministic function of time. Show that $\log B_t^T$ is normally distributed (i.e., B_t^T is log-normal)

<u>Exercise</u> 7.4. Suppose the initial forward rate curve is give by

$$f_0^T = b_0 + b_1 e^{-a_1 T} + b_2 a_1 T e^{-a_1 T} + b_3 a_2 T e^{-a_2 T}.$$

(a) What is the bond price B_0^T?

(b) What is the short rate R_0?

(c) What is the yield Y_0^T?

Chapter 8

Forward measure

Throughout these notes, the numéraire we have used most often is the money market account M. The reason we have been using M as numéraire is that this choice often results in simple computations (or at least, relatively simple computations) for bond prices. There are, however, some financial assets, whose values can be more easily computed using a different numéraire – namely, the zero-coupon bond B^T.

Recall from (7.9) that the dynamics of B/M, are given by

$$d\left(\frac{B_t^T}{M_t}\right) = -\Sigma_t^T \frac{B_t^T}{M_t} d\widetilde{W}_t,$$

Where $\widetilde{W}$ is a Brownian motion under $\widetilde{\mathbb{P}}$, which is a risk-neutral measure with M as numéraire. From the above equation we have

$$\frac{B_T^T}{M_T} = \frac{B_0^T}{M_0} \exp\left(-\tfrac{1}{2}\int_0^T (-\Sigma_t^T)^2 dt - \int_0^T \Sigma_t^T d\widetilde{W}_t \right).$$

Recalling Girsanov's Theorem 3.7.2, we therefore can define a change of probability measure

$$\frac{d\widehat{\mathbb{P}}^T}{d\widetilde{\mathbb{P}}} := \frac{M_0}{B_0^T}\frac{B_T^T}{M_T} = \exp\left(-\tfrac{1}{2}\int_0^T (-\Sigma_t^T)^2 dt - \int_0^T \Sigma_t^T d\widetilde{W}_t \right).$$

The process $\widehat{W}^T$, defined by

$$\widehat{W}_t^T := \widetilde{W}_t + \int_0^t \Sigma_s^T ds, \tag{8.1}$$

is a Brownian motion under the probability measure $\widehat{\mathbb{P}}^T$. Now, consider an asset $A = (A_t)_{t\geq 0}$ (could be a stock, bond, or derivative). Using risk-neutral pricing, the value of A at time $t = 0$ can be computed as follows

$$\frac{A_0}{M_0} = \widetilde{\mathbb{E}}\left(\frac{A_T}{M_T}\right) = \frac{B_0^T}{M_0}\widetilde{\mathbb{E}}\left(\frac{M_0}{B_0^T}\frac{B_T^T}{M_T}\frac{A_T}{B_T^T}\right) = \frac{B_0^T}{M_0}\widetilde{\mathbb{E}}\left(\frac{d\widehat{\mathbb{P}}^T}{d\widetilde{\mathbb{P}}}\frac{A_T}{B_T^T}\right) = \frac{B_0^T}{M_0}\widehat{\mathbb{E}}^T\left(\frac{A_T}{B_T^T}\right),$$

where $\widehat{\mathbb{E}}^{T}$ indicates expectation under $\widehat{\mathbb{P}}^{T}$. Multiplying both sides of the above equation by M_0/B_0^T we obtain

$$\frac{A_0}{B_0^T} = \widehat{\mathbb{E}}^T\left(\frac{A_T}{B_T^T}\right) = \widehat{\mathbb{E}}^T A_T$$

where we have used $B_T^T = 1$. More generally, we have

$$\frac{A_t}{B_t^T} = \widehat{\mathbb{E}}^T\left(\frac{A_T}{B_T^T}\middle|\mathcal{F}_t\right) = \widehat{\mathbb{E}}^T(A_T|\mathcal{F}_t). \tag{8.2}$$

The above equation is simply the risk-neutral pricing formula (4.5) with the T-maturity bond B^T as numéraire. We call the probability measure $\widehat{\mathbb{P}}^T$ the T-*forward measure*. The reason for this name is that the T-forward price $A^T = A/B^T$ is a martingale under the T-forward measure $\widehat{\mathbb{P}}^T$ (see Section 1.7 for a description of T-forward prices). Note that, in order to use (8.2), we need to know the dynamics of A under $\widehat{\mathbb{P}}^T$. If we know the dynamics of A under $\widetilde{\mathbb{P}}$, then we can use (8.1) to determine the dynamics of A under $\widehat{\mathbb{P}}^T$.

8.1 Derivatives written on Bonds

Consider a derivative that pays $f(B_{T_1}^{T_2})$ at time $T_1 < T_2$, where $f : [0,1] \to \mathbb{R}$. Let $V = (V_t)_{0\leq t\leq T}$ denote the value of this derivative. If we try to price this derivative using M as numéraire

$$\frac{V_t}{M_t} = \widetilde{\mathbb{E}}\left(\frac{f(B_{T_1}^{T_2})}{M_{T_1}}\middle|\mathcal{F}_t\right) \qquad \Rightarrow \qquad V_t = \widetilde{\mathbb{E}}\left(\exp\left(-\int_t^T R_s ds\right)f(B_{T_1}^{T_2})\middle|\mathcal{F}_t\right),$$

we would need to know the joint density of $\exp\left(-\int_t^T R_s ds\right)$ and $B_{T_1}^{T_2}$ under $\widetilde{\mathbb{P}}$. On the other hand, if we price this derivative using B^{T_1} as numéraire

$$\frac{V_t}{B_t^{T_1}} = \widehat{\mathbb{E}}^{T_1}\left(\frac{f(B_{T_1}^{T_2})}{B_{T_1}^{T_1}}\middle|\mathcal{F}_t\right) \qquad \Rightarrow \qquad V_t = B_t^{T_1}\widehat{\mathbb{E}}^{T_1}\left(f(B_{T_1}^{T_2})\middle|\mathcal{F}_t\right),$$

we need only the density of $B_{T_1}^{T_2}$ under $\widehat{\mathbb{P}}^{T_1}$. Noting that $B_{T_1}^{T_2} = B_{T_1}^{T_2}/B_{T_1}^{T_1}$, rather than determine the dynamics of B^{T_2} under $\widehat{\mathbb{P}}^{T_1}$, we will determine the dynamics of B^{T_2}/B^{T_1}, as we know this process is a $\widehat{\mathbb{P}}^{T_1}$-martingale. Using (7.8), we compute

$$\begin{aligned} dB_t^{T_2} &= R_t B_t^{T_2} dt - \Sigma_t^{T_2} B_t^{T_2} d\widetilde{W}_t, \\ d\left(\frac{1}{B_t^{T_1}}\right) &= \frac{-1}{(B_t^{T_1})^2} dB_t^{T_1} + \frac{1}{(B_t^{T_1})^3} d[B^{T_1}, B^{T_1}]_t \end{aligned}$$

$$= \frac{1}{B_t^{T_1}}\Big(- R_t + (\Sigma_t^{T_1})^2\Big)dt + \frac{1}{B_t^{T_1}}\Sigma_t^{T_1}d\widetilde{W}_t.$$

And thus, we have

$$\begin{aligned}
d\Big(\frac{B_t^{T_2}}{B_t^{T_1}}\Big) &= \frac{1}{B_t^{T_1}}dB_t^{T_2} + B_t^{T_2}d\Big(\frac{1}{B_t^{T_1}}\Big) + d\Big[B^{T_2}, \frac{1}{B^{T_1}}\Big]_t \\
&= \frac{B_t^{T_2}}{B_t^{T_1}}\Big((\Sigma_t^{T_1})^2 - \Sigma_t^{T_1}\Sigma_t^{T_2}\Big)dt + \frac{B_t^{T_2}}{B_t^{T_1}}\Big(\Sigma_t^{T_1} - \Sigma_t^{T_2}\Big)d\widetilde{W}_t \\
&= \frac{B_t^{T_2}}{B_t^{T_1}}\Big((\Sigma_t^{T_1})^2 - \Sigma_t^{T_1}\Sigma_t^{T_2}\Big)dt + \frac{B_t^{T_2}}{B_t^{T_1}}\Big(\Sigma_t^{T_1} - \Sigma_t^{T_2}\Big)(d\widehat{W}_t^{T_1} - \Sigma_t^{T_1}dt) \\
&= \frac{B_t^{T_2}}{B_t^{T_1}}\Big(\Sigma_t^{T_1} - \Sigma_t^{T_2}\Big)d\widehat{W}_t^{T_1}.
\end{aligned}$$

Observe the B^{T_2}/B^{T_1} is a $\widehat{\mathbb{P}}^{T_1}$-martingale, as it must be. From the above, we have

$$B_{T_1}^{T_2} = \frac{B_{T_1}^{T_2}}{B_{T_1}^{T_1}} = \frac{B_t^{T_2}}{B_t^{T_1}} \exp\Big(- \frac{1}{2}\int_t^{T_1} \big(\Sigma_s^{T_1} - \Sigma_s^{T_2}\big)^2 ds + \int_t^{T_1} \big(\Sigma_s^{T_1} - \Sigma_s^{T_2}\big)d\widehat{W}_s^{T_1}\Big). \tag{8.3}$$

To go further, we must specify a particular model.

8.2 Example: Vasicek model

Recall the Vasicek model from Section 5.6.1 (and again in Section 7.3). We have

$$\sigma_t^T = \sigma e^{-\kappa(T-t)}.$$

It follows that

$$\Sigma_t^T = \int_t^T \sigma_t^s ds = \frac{\sigma}{\kappa}\Big(1 - e^{-\kappa(T-t)}\Big).$$

Because Σ_t^T is a deterministic function of time, it follows from (8.3) that

$$\begin{aligned}
\log B_{T_1}^{T_2}|\mathcal{F}_t = \log\Big(\frac{B_{T_1}^{T_2}}{B_{T_1}^{T_1}}\Big)|\mathcal{F}_t &\sim \mathcal{N}(m_t, v_t^2) \\
m_t &= \log\frac{B_t^{T_2}}{B_t^{T_1}} - \frac{1}{2}\int_t^{T_1} \big(\Sigma_s^{T_1} - \Sigma_s^{T_2}\big)^2 ds, \\
v_t^2 &= \int_t^{T_1} \big(\Sigma_s^{T_1} - \Sigma_s^{T_2}\big)^2 ds.
\end{aligned}$$

The above integrals can be computed explicitly, although we will not do so here. We can now compute the derivative price V_t as follows

$$V_t = B_t^{T_1} Q(t, Y_t), \qquad Q(t, Y_t) := \widehat{\mathbb{E}}^{T_1}\Big(f(Y_{T_1})\Big|Y_t\Big), \qquad Y_t := \frac{B_t^{T_2}}{B_t^{T_1}}, \tag{8.4}$$

where the conditional expectation can be computed using the $\mathcal{F}_t$-conditional density of $\log B_{T_1}^{T_2}$.

Just as important as pricing is the ability to replicate the derivative. To this end, observe that

$$\begin{aligned} d\Big(\frac{V_t}{B_t^{T_1}}\Big) &= dQ(t, Y_t) = \underbrace{(\ldots)}_{=0} dt + \partial_y Q(t, Y_t) dY_t \\ &= \partial_y Q(t, Y_t) \frac{B_t^{T_2}}{B_t^{T_1}} \Big(\Sigma_t^{T_1} - \Sigma_t^{T_2}\Big) d\widehat{W}_t^{T_1}, \end{aligned} \tag{8.5}$$

where the dt term must equal zero because Q is defined as a conditional expectation and is therefore a martingale. Now, consider a self-financing porfolio X, whose dynamics are of the form

$$dX_t = \Delta_t dB_t^{T_2} + (X_t - \Delta_t B_t^{T_2}) \frac{1}{B_t^{T_1}} dB_t^{T_1}. \tag{8.6}$$

Noting the X/B^{T_1} is a $\widehat{\mathbb{P}}^{T_1}$-martingale (and as such, the dt-terms sum to zero) we have

$$\begin{aligned} d\Big(\frac{X_t}{B_t^{T_1}}\Big) &= \frac{1}{B_t^{T_1}} dX_t + X_t d\Big(\frac{1}{B_t^{T_1}}\Big) + (\ldots)dt \\ &= \frac{1}{B_t^{T_1}}\Big(\Delta_t dB_t^{T_2} + (X_t - \Delta_t B_t^{T_2})\frac{1}{B_t^{T_1}} dB_t^{T_1}\Big) \\ &\quad + X_t\Big((\ldots)dt + \frac{1}{B_t^{T_1}} \Sigma_t^{T_1} d\widehat{W}_t^{T_1}\Big) + (\ldots)dt \\ &= \frac{1}{B_t^{T_1}}\Big(\Delta_t(\ldots dt - \Sigma_t^{T_2} B_t^{T_2} d\widehat{W}_t^{T_1}) + (X_t - \Delta_t B_t^{T_2})\frac{1}{B_t^{T_1}}(\ldots dt - \Sigma_t^{T_1} B_t^{T_1} d\widehat{W}_t^{T_1})\Big) \\ &\quad + X_t\Big((\ldots)dt + \frac{1}{B_t^{T_1}} \Sigma_t^{T_1} d\widehat{W}_t^{T_1}\Big) + (\ldots)dt \\ &= \Delta_t \frac{B_t^{T_2}}{B_t^{T_1}}\Big(\Sigma_t^{T_1} - \Sigma_t^{T_2}\Big) d\widehat{W}_t^{T_1} + \underbrace{(\ldots)}_{=0} dt. \end{aligned} \tag{8.7}$$

Comparing (8.5) and (8.7), we see that

$$d\Big(\frac{V_t}{B_t^{T_1}}\Big) = d\Big(\frac{X_t}{B_t^{T_1}}\Big) \qquad \Leftrightarrow \qquad \Delta_t = \partial_y Q(t, Y_t).$$

Thus, we have derived a replication strategy.

8.3 Exercises

<u>Exercise</u> 8.1. Consider the Vasicek model for R, described in Sections 5.6.1, 7.3 and 8.2.
(a) Compute $Q(t,y)$ and $\partial_y Q(t,y)$ where Q is defined in (8.4) and $f(y) = \log^2 y$.
(b) Compute V_t, the time t value of an option that pays pays $f(B^{T_2}_{T_1})$ at time T_1.
(c) Consider the portfolio X in (8.6), which replicates the payoff of an option that pays $f(B^{T_2}_{T_1})$ at time T_1. Fix the following parameters

$$\kappa = 0.01, \qquad \theta - 0.05, \qquad \sigma = 0.01,$$
$$R_0 = 0.07, \qquad T_1 = 1.00, \qquad T_2 = 2.00.$$

Write a program in the language of your choice (e.g., Matlab, Mathematica, R, Python, etc.) to simulate under $\widetilde{\mathbb{P}}$ paths of R, B^{T_1}, B^{T_2}, V and X over the interval $[0, T_1]$. On a single axis, plot a path of R, B^{T_1}, B^{T_2}, V and X. Remember, to simulate the path of an SDE of the form

$$dZ_t = \mu(t, Z_t)dt + \sigma(t, Z_t)d\widetilde{W}_t,$$

simply use the Euler approximation

$$Z_{t+\delta} \approx Z_t + \mu(t, Z_t)\delta + \sigma(t, Z_t)(\widetilde{W}_{t+\delta} - \widetilde{W}_t),$$

where $\widetilde{W}_{t+\delta} - \widetilde{W}_t \sim \mathcal{N}(0, \delta)$. Take $\delta = T_1/N$ with $N = 1000$. Make sure to use the same path of $\widetilde{W}$ for all processes when making the plot.
(d) Run $M = 1000$ paths of the above processes (with different realizations of $\widetilde{W}$). For each path, compute $X_{T_1} - f(B^{T_2}_{T_1})$ and plot a histogram of the results.

Chapter 9

LIBOR

Recall from (1.3) that the simple forward rate at time t over the interval $[T_1, T_2]$ is given by

$$F_t^{T_1,T_2} := \frac{1}{T_2 - T_1}\left(\frac{B_t^{T_1}}{B_t^{T_2}} - 1\right).$$

The *London Inter-Bank Offered Rate* (or LIBOR), denoted L_t^T is a simple forward rate with a fixed tenor $T_2 - T_1 = \delta$

$$L_t^T := F_t^{T,T+\delta} = \frac{1}{\delta}\left(\frac{B_t^T}{B_t^{T+\delta}} - 1\right). \tag{9.1}$$

If we set $T = t$ we obtain the *spot* LIBOR rate

$$L_t^t := \frac{1}{\delta}\left(\frac{B_t^t}{B_t^{t+\delta}} - 1\right) = \frac{1}{\delta}\left(\frac{1}{B_t^{t+\delta}} - 1\right).$$

9.1 Pricing a Backset LIBOR contract

Definition 9.1.1. A *backset LIBOR contract* pays its holder L_T^T at time $T + \delta$.

Theorem 9.1.2. *Let $S = (S_t)_{0\le t\le T+\delta}$ denote the value of backset LIBOR contract, described in Definition 9.1.1. We have*

$$S_t = \begin{cases} B_t^{T+\delta} L_t^T, & 0 \le t \le T, \\ B_t^{T+\delta} L_T^T, & T \le t \le T+\delta. \end{cases} \tag{9.2}$$

Proof. Let us first consider the case $T \le t \le T + \delta$. In this case the payment at time $T + \delta$ is known to be L_T^T. In this case the payment can be replicated by purchasing L_T^T bonds with maturity $T + \delta$. The cost of this purchase is $L_T^T B_t^{T+\delta}$, in agreement (9.2).

Now, consider the case $0 \le t \le T$. Consider the following investment strategy: at time t purchase $1/\delta$ bonds with maturity T and sell $1/\delta$ bonds with maturity $T + \delta$. The cost of this strategy is

$$\frac{1}{\delta}\left(B_t^T - B_t^{T+\delta}\right) = B_t^{T+\delta} L_t^T, \tag{9.3}$$

where we have used (9.1). At time $t = T$, the value of hits investment strategy is

$$\frac{1}{\delta}\left(B_T^T - B_T^{T+\delta}\right) = B_T^{T+\delta} L_T^T = S_T$$

The investment strategy replicates the value of the contract at time T. Thus, the value of the contract at time $t < T$ must equal the initial cost of the strategy (9.3), in agreement with (9.2). □

9.2 Black-Caplet Formula

Recall that a holder of a cap receives a series of payments called caplets where, from (1.11) we have

$$\text{Caplet payoff at time } T_i = (F_{T_{i-1}}^{T_{i-1},T_i} - \kappa)^+, \qquad i = 1, 2, \ldots, n.$$

If the dates are equally spaced $T_i - T_{i-1} = \delta$, then we have

$$\begin{aligned} \text{Caplet payoff at time } T_i &= (F_{T_{i-1}}^{T_{i-1},T_{i-1}+\delta} - \kappa)^+ \\ &= (L_{T_{i-1}}^{T_{i-1}} - \kappa)^+, \qquad i = 1, 2, \ldots, n, \end{aligned}$$

where we have used (9.1). The value of a cap is equal to the sum of the values of the caplets. Thus, we concentrate on valuing the following caplet

$$\text{Payoff at time } T + \delta = (L_T^T - \kappa)^+.$$

To begin, we note from (9.2)

$$\frac{S_t}{B_t^{T+\delta}} = L_t^T, \qquad 0 \le t \le T.$$

Recalling the definition of the forward price (1.10), the above equation tells us that, for $0 \le t \le T$, the $T + \delta$-forward price of S is L^T.

Suppose we had constructed a HJM model for forward rates driven by a single Brownian motion under the physical (i.e., real-world) probablity measure $\mathbb{P}$. Then, from (7.7) the no-arbitrage dynamics of forward rates under a risk-neutral measure $\widetilde{\mathbb{P}}$ with M as numéraire must be given by

$$df_t^T = \sigma_t^T \Sigma_t^T dt + \sigma_t^T d\widetilde{W}_t,$$

where $\widetilde{\mathrm{W}}$ is a $\widetilde{\mathbb{P}}$-Brownian motion. Moreover, from (7.9), the dynamics of $\mathrm{B}^{\mathrm{T}+\delta}/\mathrm{M}$ are given by

$$\mathrm{d}\Big(\frac{\mathrm{B}_t^{\mathrm{T}+\delta}}{\mathrm{M}_t}\Big) = -\Sigma_t^{\mathrm{T}+\delta}\frac{\mathrm{B}_t^{\mathrm{T}+\delta}}{\mathrm{M}_t}\mathrm{d}\widetilde{\mathrm{W}}_t,$$

Let us define the following change of measure

$$\frac{\mathrm{d}\widehat{\mathbb{P}}^{\mathrm{T}+\delta}}{\mathrm{d}\widetilde{\mathbb{P}}} = \frac{\mathrm{M}_0}{\mathrm{B}_0^{\mathrm{T}+\delta}}\frac{\mathrm{B}_{\mathrm{T}+\delta}^{\mathrm{T}+\delta}}{\mathrm{M}_{\mathrm{T}+\delta}} = \exp\Big(-\tfrac{1}{2}\int_0^{\mathrm{T}+\delta}(\Sigma_t^{\mathrm{T}+\delta})^2\mathrm{d}t - \int_0^{\mathrm{T}+\delta}\Sigma_t^{\mathrm{T}+\delta}\mathrm{d}\widetilde{\mathrm{W}}_t\Big)$$

From Theorem 3.7.2, the process $\widehat{\mathrm{W}}^{\mathrm{T}+\delta}$, defined by

$$\widehat{\mathrm{W}}_t^{\mathrm{T}+\delta} = \widetilde{\mathrm{W}}_t + \int_0^t \Sigma_s^{\mathrm{T}+\delta}\mathrm{d}s, \tag{9.4}$$

is a $\widehat{\mathbb{P}}^{\mathrm{T}+\delta}$ Brownian motion. Moreover, $\mathrm{S}/\mathrm{B}^{\mathrm{T}+\delta} = \mathrm{L}^{\mathrm{T}}$ is a martingale under $\widehat{\mathbb{P}}^{\mathrm{T}+\delta}$. If follows that there exists a process γ^{T} such that

$$\mathrm{dL}_t^{\mathrm{T}} = \gamma_t^{\mathrm{T}}\mathrm{L}_t^{\mathrm{T}}\mathrm{d}\widehat{\mathrm{W}}_t^{\mathrm{T}+\delta}, \qquad 0 \le t \le \mathrm{T}. \tag{9.5}$$

We can now value a derivative that pays $g(\mathrm{L}_{\mathrm{T}}^{\mathrm{T}})$ at time $\mathrm{T}+\delta$ using the following theorem.

<u>Theorem</u> 9.2.1. *Let $\mathrm{V} = (\mathrm{V}_t)_{0\le t\le \mathrm{T}+\delta}$ the the value of a derivative that pays $g(\mathrm{L}_{\mathrm{T}}^{\mathrm{T}})$ at time $\mathrm{T}+\delta$. Then we have*

$$\mathrm{V}_t = \mathrm{B}_t^{\mathrm{T}+\delta}\widehat{\mathbb{E}}^{\mathrm{T}+\delta}(g(\mathrm{L}_{\mathrm{T}}^{\mathrm{T}})|\mathcal{F}_t), \tag{9.6}$$

where the dynamics of L^{T} under $\widehat{\mathbb{P}}^{\mathrm{T}+\delta}$ are given by (9.5). In particular

$$\mathrm{V}_t = \mathrm{B}_t^{\mathrm{T}+\delta}g(\mathrm{L}_{\mathrm{T}}^{\mathrm{T}}), \qquad \mathrm{T} \le t \le \mathrm{T}+\delta. \tag{9.7}$$

Moreover, if $\gamma_t^{\mathrm{T}} = \gamma(t;\mathrm{T})$ is a deterministic function of time then $\log \mathrm{L}_{\mathrm{T}}|\mathcal{F}_t$ is noramally distributed

$$\log \mathrm{L}_{\mathrm{T}}^{\mathrm{T}}|\mathcal{F}_t \sim \mathcal{N}(m_t, v_t^2), \qquad 0 \le t \le \mathrm{T}, \tag{9.8}$$

with and mean and variance given by

$$m_t = \log \mathrm{L}_t^{\mathrm{T}} - \tfrac{1}{2}\int_t^{\mathrm{T}}\gamma^2(s;\mathrm{T})\mathrm{d}s, \qquad v_t^2 = \int_t^{\mathrm{T}}\gamma^2(s;\mathrm{T})\mathrm{d}s.$$

<u>Proof</u>. We have from risk-neutral pricing that

$$\frac{\mathrm{V}_t}{\mathrm{B}_t^{\mathrm{T}+\delta}} = \widehat{\mathbb{E}}^{\mathrm{T}+\delta}\Big(\frac{\mathrm{V}_{\mathrm{T}+\delta}}{\mathrm{B}_{\mathrm{T}+\delta}^{\mathrm{T}+\delta}}\Big|\mathcal{F}_t\Big) = \widehat{\mathbb{E}}^{\mathrm{T}+\delta}(g(\mathrm{L}_{\mathrm{T}}^{\mathrm{T}})|\mathcal{F}_t).$$

Equation (9.6) follows by mutiplying through by $B_t^{T+\delta}$ and equation (9.7) follows by noting that $g(L_T^T)$ is a known constant for any $t \geq T$. In the case that $\gamma_t^T = \gamma(t;T)$ is a deterministic function of time, we have from (9.5) and Itô's Lemma that

$$\log L_T^T = \log L_t^T - \tfrac{1}{2}\int_t^T \gamma^2(s;T)ds + \int_t^T \gamma(s;T)d\widehat{W}_s^{T+\delta}$$

from which (9.8) follows. □

Note that, if the payoff function is a call payoff $g(L) = (L-\kappa)^+$ then we obtain

$$\widehat{\mathbb{E}}^{T+\delta}\Big((L_T^T-\kappa)^+|\mathcal{F}_t\Big) = C^{BS}(t, L_t^T; T, \kappa, \overline{\gamma}(t;T))$$

where $C^{BS}(t, L_t^T; T, \kappa, \overline{\gamma}(t;T))$ denotes the time t Black-Scholes price of a call written on L^T with maturity T, strike κ and volatility $\overline{\gamma}(t,T)$, which is defined as follows

$$\overline{\gamma}^2(t;T) := \frac{1}{T-t}\int_t^T \gamma^2(s;T)ds.$$

9.3 Relation of LIBOR dynamics to bond volatilities

Let us see if we can relate the process γ^T in (9.5) to the process Γ^T in the HJM framework. In the computation that follows, we will ignore the dt-terms because under $\widehat{\mathbb{P}}^{T+\delta}$ the process L^T is a martingale. We have from (7.8) and (9.4) that

$$\begin{aligned}
dB_t^T &= (\dots)dt - \Sigma_t^T B_t^T d\widetilde{W}_t \\
&= (\dots)dt - \Sigma_t^T B_t^T d\widehat{W}_t^{T+\delta},
\end{aligned}$$

from which it follows that

$$\begin{aligned}
d\Big(\frac{1}{B_t^{T+\delta}}\Big) &= \frac{-1}{(B_t^{T+\delta})^2}dB_t^{T+\delta} + \frac{1}{(B_t^{T+\delta})^3}d[B^{T+\delta}, B^{T+\delta}]_t \\
&= (\dots)dt + \Sigma_t^{T+\delta}\frac{1}{B_t^{T+\delta}}d\widehat{W}_t^{T+\delta}.
\end{aligned}$$

We therefore have from from (9.1) that

$$\begin{aligned}
\delta dL_t^T &= d\Big(\frac{B_t^T}{B_t^{T+\delta}}\Big) \\
&= \frac{1}{B_t^{T+\delta}}dB_t^T + B_t^T d\Big(\frac{1}{B_t^{T+\delta}}\Big) + d\Big[B^T, \frac{1}{B^{T+\delta}}\Big]_t
\end{aligned}$$

$$
\begin{aligned}
&= \left(\Sigma_t^{T+\delta} - \Sigma_t^T\right)\frac{B_t^T}{B_t^{T+\delta}}d\widehat{W}_t^{T+\delta} \\
&= \left(\Sigma_t^{T+\delta} - \Sigma_t^T\right)(\delta L_t^T + 1)d\widehat{W}_t^{T+\delta}.
\end{aligned}
$$

where the dt-terms must cancel because L^T is a martingale. Dividing both sides by δ we obtain

$$
\begin{aligned}
dL_t^T &= \left(\Sigma_t^{T+\delta} - \Sigma_t^T\right)(L_t^T + \frac{1}{\delta})d\widehat{W}_t^{T+\delta} \\
&= \left(\Sigma_t^{T+\delta} - \Sigma_t^T\right)\left(1 + \frac{1}{\delta L_t^T}\right)L_t^T d\widehat{W}_t^{T+\delta}. \qquad (9.9)
\end{aligned}
$$

Comparing (9.5) with (9.9) we see that

$$
\gamma_t^T = \left(\Sigma_t^{T+\delta} - \Sigma_t^T\right)\left(1 + \frac{1}{\delta L_t^T}\right).
$$

9.4 A LIBOR MARKET MODEL

In this Section, we will show one method of constructing a LIBOR term-structure model that is consistent with observable market data. There are a number of practical reasons to consider modeling the dynamics of the LIBOR rate.

1. LIBOR is directly used as the underlying of many contracts, e.g. caplets, caps, swap, swaptions.
2. LIBOR is an observable and is routinely used as the input for loan calculations.
3. LIBOR is often considered as the fundamental interest rate that gives rise to other rates.

The LIBOR rates are the building blocks of the so-called market models. These models allow us to derive the bond prices and other quantities using the LIBORs. Let us see how this can be done

To begin, let us review some of the key equations we have developed thus far

$$
L_t^T = \frac{1}{\delta}\left(\frac{B_t^T}{B_t^{T+\delta}} - 1\right), \qquad (9.10)
$$

$$
d\widehat{W}_t^{T+\delta} = \Sigma_t^{T+\delta}dt + d\widetilde{W}_t, \qquad (9.11)
$$

$$
dL_t^T = \gamma_t^T L_t^T d\widehat{W}_t^{T+\delta}, \qquad (9.12)
$$

$$
\gamma_t^T = \left(\Sigma_t^{T+\delta} - \Sigma_t^T\right)\left(1 + \frac{1}{\delta L_t^T}\right), \qquad (9.13)
$$

Suppose now that, at time $t = 0$ we can observe at-the-money (i.e., $\kappa = L_0^{T_j}$) forward caplet prices at maturity dates $T_j = \delta j$ for $j = 1, 2, \dots n$. That is,

$$
\text{We observe:} \qquad \frac{V_0^{\text{caplet},T_{j+1}}}{B_0^{T_{j+1}}} = C^{BS}(0, L_0^{T_j}; T_j, L_0^{T_j}, \overline{\gamma}_j), \qquad j = 1, 2, \dots, n.
$$

We can then choose deterministic functions

$$\gamma(\cdot;\mathrm{T}_j):[0,\mathrm{T}_j]\to\mathbb{R}_+, \qquad \text{such that} \qquad \frac{1}{\mathrm{T}_j}\int_0^{\mathrm{T}_j}\gamma^2(t;\mathrm{T}_j)\mathrm{d}t=\overline{\gamma}_j^2. \tag{9.14}$$

For example, we could simply choose $\gamma(t;\mathrm{T}_j)=\overline{\gamma}_j$. Setting $\gamma_t^{\mathrm{T}}=\gamma(t;\mathrm{T}_j)$ in (9.12), where $\gamma(t;\mathrm{T}_j)$ satisfies (9.14), gives us a model for LIBOR that matches observed at-the-money caplet prices.

Now, observe from (9.11) that

$$\begin{aligned}\mathrm{d}\widehat{\mathrm{W}}_t^{\mathrm{T}_j}&=\left(\Sigma_t^{\mathrm{T}_j}-\Sigma_t^{\mathrm{T}_{j+1}}\right)\mathrm{d}t+\mathrm{d}\widehat{\mathrm{W}}_t^{\mathrm{T}_{j+1}}\\&=\frac{-\delta\gamma(t;\mathrm{T}_j)\mathrm{L}_t^{\mathrm{T}_j}}{1+\delta\mathrm{L}_t^{\mathrm{T}_j}}\mathrm{d}t+\mathrm{d}\widehat{\mathrm{W}}_t^{\mathrm{T}_{j+1}},\end{aligned}$$

where, in the second line, we have used (9.13). Using the above relation, one can easily derive that

$$\mathrm{d}\widehat{\mathrm{W}}_t^{\mathrm{T}_{j+1}}=\sum_{i=j+1}^{n}\frac{-\delta\gamma(t;\mathrm{T}_i)\mathrm{L}_t^{\mathrm{T}_i}}{1+\delta\mathrm{L}_t^{\mathrm{T}_i}}\mathrm{d}t+\mathrm{d}\widehat{\mathrm{W}}_t^{\mathrm{T}_{n+1}},$$

which holds for any $j=0,1,\ldots,n$, provided we interpret $\sum_{i=n+1}^{n}(\ldots)=0$. We therefore have that

$$\mathrm{d}\mathrm{L}_t^{\mathrm{T}_j}=\gamma(t;\mathrm{T}_j)\mathrm{L}_t^{\mathrm{T}_j}\left(\sum_{i=j+1}^{n}\frac{-\delta\gamma(t;\mathrm{T}_i)\mathrm{L}_t^{\mathrm{T}_i}}{1+\delta\mathrm{L}_t^{\mathrm{T}_i}}\mathrm{d}t+\mathrm{d}\widehat{\mathrm{W}}_t^{\mathrm{T}_{n+1}}\right). \tag{9.15}$$

The above equation gives the dynamics of $\mathrm{L}^{\mathrm{T}_j}$ for $j=1,2,\ldots,n$ in terms of a single Brownian motion $\widehat{\mathrm{W}}^{\mathrm{T}_{j+1}}$. More specifically, if we observe LIBOR rates at time zero $\mathrm{L}_0^{\mathrm{T}_j}$, $j=1,2,\ldots,n$, then (9.15) gives us the dynamics of $\mathrm{L}^{\mathrm{T}_j}$, $j=1,2,\ldots,n$ under $\widehat{\mathbb{P}}^{\mathrm{T}_{n+1}}$.

Now, let us see how we can construct the dynamics of discounted bond price $\mathrm{B}^{\mathrm{T}_j}/\mathrm{M}$. Using (7.9) and (9.11) we have

$$\begin{aligned}\mathrm{d}\left(\frac{\mathrm{B}_t^{\mathrm{T}_j}}{\mathrm{M}_t}\right)&=-\Sigma_t^{\mathrm{T}_j}\frac{\mathrm{B}_t^{\mathrm{T}_j}}{\mathrm{M}_t}\mathrm{d}\widetilde{\mathrm{W}}_t\\&=-\Sigma_t^{\mathrm{T}_j}\frac{\mathrm{B}_t^{\mathrm{T}_j}}{\mathrm{M}_t}\left(-\Sigma_t^{\mathrm{T}_{n+1}}\mathrm{d}t+\mathrm{d}\widehat{\mathrm{W}}_t^{\mathrm{T}_{n+1}}\right)\\&=\Sigma_t^{\mathrm{T}_j}\Sigma_t^{\mathrm{T}_{n+1}}\frac{\mathrm{B}_t^{\mathrm{T}_j}}{\mathrm{M}_t}\mathrm{d}t-\Sigma_t^{\mathrm{T}_j}\frac{\mathrm{B}_t^{\mathrm{T}_j}}{\mathrm{M}_t}\mathrm{d}\widehat{\mathrm{W}}_t^{\mathrm{T}_{n+1}}.\end{aligned} \tag{9.16}$$

The initial condition can be determined from observed LIBOR rates and the initial value of a money market account. Using (9.10) we find that

$$\frac{\mathrm{B}_0^{\mathrm{T}_j}}{\mathrm{M}_0}=\frac{1}{\mathrm{M}_0}\prod_{i=0}^{j-1}\frac{\mathrm{B}_0^{\mathrm{T}_{i+1}}}{\mathrm{B}_0^{\mathrm{T}_i}}=\frac{1}{\mathrm{M}_0}\prod_{i=0}^{j-1}\frac{1}{1+\delta\mathrm{L}_0^{\mathrm{T}_i}},$$

where we have used $B_0^{T_0} = B_0^0 = 1$. We have some freedom – but not complete freedom – regarding how we choose $\Sigma_t^{T_j}$. For example, recalling that

$$\Sigma_t^T = \int_t^T \sigma_t^s ds, \qquad \text{we must have} \qquad \Sigma_{T_j}^{T_j} = 0, \qquad \forall\, j = 1, 2, \ldots n.$$

Moreover, if we choose

$$\Sigma_t^{T_j}, \qquad T_{j-1} \leq t \leq T_j \qquad \text{for } j = 1, 2, \ldots, n+1$$

then we have, in fact, fixed $\Sigma_t^{T_j}$ for $0 \leq t \leq T_j$ for all j. The reason is that, from (9.13), we have

$$\Sigma_t^{T_j} = \Sigma_t^{T_{j-1}} + \frac{\delta\gamma(t; T_{j-1}) L_t^{T_{j-1}}}{1 + \delta L_t^{T_{j-1}}}, \qquad 0 \leq t \leq T_{j-1}.$$

Once we have chosen some processes $\Sigma_t^{T_j}$, $T_{j-1} \leq t \leq T_j$ for $j = 1, 2, \ldots, n+1$, we can solve (9.16) to construct the evolution of discounted bond prices

$$\frac{B_t^{T_j}}{M_t} = \frac{B_0^{T_j}}{M_0} \exp\left(\int_0^t \left(\Sigma_s^{T_j} \Sigma_s^{T_{n+1}} - \tfrac{1}{2} (\Sigma_s^{T_j})^2 \right) ds - \int_0^t \Sigma_s^{T_j} d\widehat{W}_s^{T_{n+1}} \right).$$

Bibliography

Grimmett, G. and D. Stirzaker (2001). *Probability and random processes.* Oxford university press.

Shreve, S. E. (2004). *Stochastic calculus for finance II: Continuous-time models*, Volume 11. Springer Science & Business Media.