Chapter 1

Financial Instruments

The Greek philosopher Thales of Miletus (circa 624-546 BC) is regarded the first “thinker” to use an option. In one version of the story, Thales is a frustrated philosopher who wanted to show his peers that he was poor by choice. That year, Thales predicted good weather and a bumper crop for the olive harvest. On the basis of this prediction, he paid a small non-refundable deposit to the owner of the olive press, to secure the right to rent the olive press at the normal rate. To the owner of the press this was a win-win: not only was he getting a deposit ahead of time, but he would also get his normal rate if Thales did indeed rent the press. The bumper olive crop did arrive, and Thales, the effective owner of the olive press, was able to make a small fortune renting out the press at a significant premium over the price he was paying.

Thales had invented the option. To study financial algorithms, we must first introduce the simple financial instruments that will be our nuts and bolts.

This first use of an option highlights the need for financial derivatives. Thales needed to exploit his prediction of good weather in Greece, but you cannot buy or sell weather as a good on a marketplace. Instead, Thales created a contract another quantity (olive presses) which can be bought and sold. His key observation is that the future value of the tradeable olive press depends indirectly on the non-tradeable quantity that he predicted.

Constructing the “derivative contract” is itself ingenious of Thales – he figured out a way to profit from a weather prediction. Now, put yourself in Thales’ shoes; having invented this derivative contract giving you the future right to rent the olive press. How much deposit will you pay? The deposit would naturally depend on the profit you stand to make, which depends on how good your weather prediction is and how big the bumper crop will be. But that is not all. The profit comes at the end of the summer and you must pay the deposit now. If you make $100 at the end of the summer, how much deposit would you pay now to realize that future profit? To address this question we need to study how the value of money grows with time.

Exercise 1.1
You predict a drought in Florida, adversely affecting the orange crop. How could you profit from this prediction?

1.1 Interest: How Money Grows

Roughly speaking, money can be equated to consumption, and it is more or less accepted that present consumption is worth more than an equal consumption at some future date.

Money today is more valuable than the same amount of money in the future.

You are certainly no worse off if you have $1 today instead of tomorrow, because you can always just hold on to it for a day. However, you could be better off with the money today because you might “invest” at a profit.
Example 1.1. Suppose that $1 today buys a corn seed. If you plant the seed, a year later you crop produces 1 corn cob (which you can eat) and 2 corn seeds. Therefore, 1 seed today is worth 2 seeds and a cob one year later (if we assume that labor, water, land, etc. would be otherwise idle).

To quantify our intuition, we say that

\[(\text{today}) \quad 1 \rightarrow (1 + R) \quad (\text{in a year})\]

The parameter $R > 0$ is the interest ‘accrued’ over the one year period starting today, and it represents how much more you should pay me one year from now to make me give up $1 today. We can ask a more general question. Suppose I were given $1 at a time $t_1$. How much more would you have to give me at a later time $t_2$ for me to be willing to give up the $1 at time $t_1$.

\[1 \rightarrow (1 + R(t_1, t_2)) \quad (at \ t_1) \quad (at \ t_2)\]

The interest $R(t_1, t_2)$ is the additional amount needed to equate the $1 at time $t_1$ with its equivalent amount at $t_2$. Instead of $1 at $t_1$, you could have an arbitrary amount $X$ at $t_1$:

\[X \rightarrow X(1 + R(t_1, t_2)) \quad (at \ t_1) \quad (at \ t_2)\]

At time $t_1$ if you give a bank $X$, you can expect the bank to give you back $X \cdot (1 + R(t_1, t_2))$ at time $t_2$. Thus, your $X$ has accrued an amount $X \cdot R(t_1, t_2)$ in interest. We will assume that you can borrow and lend any amount of money at the same interest rate, which is a slight approximation because typically you can lend to a bank at a lower interest rate than you can borrow from the bank. So, if you borrow $X$ at time $t_1$ from the bank, you are expected to pay $X \cdot (1 + R(t_1, t_2))$ back to the bank at time $t_2$. To simplify notation, we will write $R(0, t)$ as just $R(t)$, that is by default we will always assume $t_1 = 0$ unless we explicitly state otherwise.

1.1.1 The Interest Rate Law - Compounding

It turns out that knowing $R(t)$ is enough because there is a relationship between $R(t_1, t_2)$, $R(t_1)$ and $R(t_2)$.

Let’s track the growth of an amount $X$ as it grows from time $0$ to $t_1$, and then from time $t_1$ to time $t_2$.

\[
\begin{align*}
X &\rightarrow X' = X(1 + R(t_1)) \\
X' &\rightarrow X'(1 + R(t_1, t_2)) = X(1 + R(t_1))(1 + R(t_1, t_2))
\end{align*}
\]

Thus, $X$ at time $0$ grows to $X(1 + R(t_1))(1 + R(t_1, t_2))$ at time $t_2$. By definition of $R(t_2)$, $X$ at time $0$ grows to $X(1 + R(t_2))$ at time $t_2$. We therefore have that

\[
1 + R(t_2) = (1 + R(t_1))(1 + R(t_1, t_2)).
\]

(1.1)

The growth of money is multiplicative. The growth from $0$ to $t_2$ is the product of the growth from $0$ to $t_1$ and the growth from $t_1$ to $t_2$. This is known as compounding. The law in (1.1) also shows that $R(t_1, t_2)$ is determined once $R(t_1)$ and $R(t_2)$ are known,

\[
R(t_1, t_2) = \frac{R(t_2) - R(t_1)}{1 + R(t_1)}.
\]

(1.2)

Equation (1.1) places quite a strong constraint on what possibilities can be chosen for the interest rate function $R(t)$.

Exercise 1.2

Generalize the interest rate law to $N$ periods defined by the times $0 = t_0, t_1, t_2, \ldots, t_N$. Specifically, show that for all $2 \leq k \leq N$,

\[
1 + R(t_k) = \prod_{i=1}^{k}(1 + R(t_{i-1}, t_i)).
\]

This is called the multi-period compounding law.
In the above exercise, let \( r_i = R(t_{i-1}, t_i) \), and let \( X_0 \) be put in the bank at time 0. Then at time \( t_k \), the money has grown to
\[
X_k = X_0 \prod_{i=1}^{k} (1 + r_i).
\]
This is known as multi-period compounding. A common case is to choose the periods \([t_{i-1}, t_i]\) so that \( r_i \) are a constant equal to \( r \) (in practice, this is approximately so if all the periods are of the same length). Then, we get the familiar compounding formula
\[
X_k = X_0 (1 + r)^k.
\]

The future compounded value at time \( t_k \) is growing \textit{exponentially} with \( k \). We can run the entire process in reverse and ask, how much money should we put in the bank at time 0 to have \( X_k \) at time \( t_k \). This is called the present value (PV) of \( X_k \). From our compounding formula, \( X_0 = X_k / (1 + r)^k \), and so
\[
PV(X_k) = \frac{X_k}{(1 + r)^k}.
\]

**Exercise 1.3**
Show that, \( PV(X; t) = X / (1 + R(t)) \). Suppose \([0, t_k]\) is broken into periods defined by \( 0 = t_0, t_1, \ldots, t_k \) with \( r_i = R(t_{i-1}, t_i) \). Then,
\[
PV(X_k) = \frac{X_k}{\prod_{i=1}^{k} (1 + r_i)}.
\]
Computing the present value of \( X_k \) is also called discounting, where \( X_k / (1 + r)^k \) is the discounted value of \( X_k \) to time 0 and \( 1 / (1 + r)^k \) is the discount factor. Does money have to grow in this exponential way?

The short answer is yes, under the reasonable assumption that money grows by the same amount over the same-sized time period. That is, \( R(\tau, \tau + \Delta t) \) only depends on \( \Delta t \) and not on \( \tau \). This is known as a stationarity assumption. Note, we are not assuming any particular dependence on \( \Delta t \). Consider two times \( t_1 = t \) and \( t_2 = t + \Delta t \) in (1.1). Since \( R(t, t + \Delta t) = R(\Delta t) \), (1.1) becomes
\[
1 + R(t + \Delta t) = (1 + R(t))(1 + R(\Delta t)). \tag{1.3}
\]
If we now make a mild technical assumption, namely that the function \( R \) is continuous at some (any) point, then a startling conclusion arises. The only possible function \( R(t) \) that satisfies (1.3) is
\[
R(t) = e^{rt} - 1,
\]
for some constant parameter \( r \), which is known as the instantaneous rate of interest.\(^1\) Money grows according to \( 1 + R(t) = e^{rt} \), that is exponentially. This is the only possibility, given stationarity. So, if we invest \( X \) at time 0, then at time \( t \), we will have
\[
X(t) = X e^{rt},
\]
which corresponds to continuous compounding at interest rate \( r \). Using (1.2),
\[
R(t_1, t_2) = \frac{e^{r t_2} - e^{r t_1}}{e^{r t_1}} = e^{r(t_2 - t_1)} - 1 = R(t_2 - t_1).
\]

**Exercise 1.4**
Consider \( X_0 \) at time 0 growing to time \( t_k \) over periods defined by \( 0 = t_0, t_1, \ldots, t_k \). Recall that \( X_k = X_0 \prod_{i=1}^{k} (1 + r_i) \) where \( r_i = R(t_{i-1}, t_i) = e^{r(t_i - t_{i-1})} - 1 \). Verify that \( X_k = X_0 e^{r t_k} \), consistent with our exponential growth law.

---

\(^1\) Define the growth multiplier \( G(t) = 1 + R(t) \). Then (1.3) says that for every \( t \) and \( \Delta t \),
\[
G(t + \Delta t) = G(t) G(\Delta t).
\]
For the analysis inclined, this is one of the well known characterizations of the exponential function. Namely the unique function \( G \) which is continuous at some (any) point satisfying this equation is \( G(t) = e^{t \alpha} \) for a parameter \( \alpha \). Here is an informal argument assuming \( G \) is differentiable. \( \ln G(t + \Delta t) - \ln G(t) = \ln G(\Delta t) \). Taking the limit \( \Delta t \to 0 \) gives \( \frac{d \ln G}{d t} = (\ln G)'(0) \equiv \alpha \). Since \( G(0) = 1, \alpha = 0 \).
Money grows exponentially at the rate \( r \). In practice, a bank will not quote \( r \), but rather a rate of interest for a year. So for example, an interest rate of 5% from your bank means that in 1 year, $1.00 will grow to $1.05. The rate being quoted by the bank is the yearly interest rate, which is slightly larger than the instantaneous interest rate.

**Exercise 1.5**

With time \( t \) measured in years, what is the instantaneous interest rate that is equivalent to a yearly interest rate of 5%? [Hint: $1.00 should grow to $1.05 in 1 year.]

When putting money in the bank, you get exponential growth. When computing present values, you have exponential discounting,

\[
PV(X; t) = X e^{-rt}.
\]

This present value is the money needed today (in the bank) to generate \( X \) at time \( t \).

### 1.1.2 A Stream of Cash Flows

Suppose you are about to retire and would like to ensure a retirement income of $50,000 per year every year for the next 25 years (if you retired at age 65, then you are expecting to live to 90). How much money do you need in the bank now? Your goal is to have a sufficient amount at time \( 0 \) to ensure your cash flow stream if you have its present value at time \( 0 \), the sum of the present values is sufficient to ensure your cash flow stream. Thus, the sum of the present values is equivalent to the cash flow stream. Therefore, this present value is the price you should be willing to pay for the cashflow stream.

For each individual cashflow \( X_i \) in the stream (occurring at time \( t_i \)), we can compute the present value \( PV(X_i; t_i) = X_i e^{-rt_i} \). This is the amount you need today to ensure \( X_i \) at time \( t_i \). If you imagine putting each of the \( PV(X_i; t_i) \) in the bank and allowing each to grow in parallel, then you will achieve the desired cashflow stream. Thus, the sum of the present values is sufficient to ensure your cashflow stream. The present value of a cashflow stream is the sum of the present values of the individual payments.

\[
PV(X_1, \ldots, X_N; t_1, \ldots, t_N) = \sum_{i=1}^{N} PV(X_i; t_i) = \sum_{i=1}^{N} X_i e^{-rt_i}.
\]

**Exercise 1.6**

Take any non-negative cashflow stream (that is \( X_i \geq 0 \)) and let its present value be \( X_0 = PV(X_1, \ldots, X_N; t_1, \ldots, t_N) \) as given in the formula above. Put \( X_0 \) in the bank at time \( 0 \). Money in the bank grows exponentially at interest rate \( r \). At time \( t_1 \) you withdraw \( X_1 \) to make your first payment. Similarly, at time \( t_i \) you withdraw \( X_i \) to make your \( i \)th payment for \( i = 1, \ldots, N \). Show:

(a) Your bank balance will never be negative (you can successfully make all payments).

(b) At the end of your last payment, your bank balance is zero.

(c) If you put in less than \( X_0 \), your balance will be negative at some time.

Not only is \( X_0 \) sufficient, but it is the minimum sufficient amount at time 0 to be able to produce this cashflow stream.

The exercise shows that you can replicate a positive cashflow stream if you have its present value at time 0, using the bank as the keeper of your money until the payments are due. In this sense, the present value is equivalent to the cashflow stream. Therefore, this present value is the price you should be willing to pay for the cashflow stream.

You can still compute a present value even if some cashflows are negative. For example, you can compute \( PV(e^{rt_1}, -e^{rt_2}; t_1, t_2) \) for \( t_2 > t_1 \), and verify that it is zero. However, if you try to replicate these payments by putting zero in the bank, at \( t_1 \) when the first payment comes along you will have a problem. You will need to incur a negative bank balance. Nevertheless, the present value of 0 is also the price you should pay for this cashflow stream. To explain why, we need to use the very powerful concept of arbitrage.
1.1.3 Arbitrage, Borrowing and Lending

Arbitrage is the mathematical formalization of “free money”. There is no free money. We need a more formal notion of free money, but the intuition is clear. If there was some free money somewhere, someone would have immediately grabbed it; every bit of it.\(^2\) Let’s first define free money, or arbitrage.

**Arbitrage Opportunity.** A cashflow stream \((X_1, \ldots, X_N; t_1, \ldots, t_N)\) is an arbitrage opportunity if \(X_i \geq 0\) for \(i = 1, \ldots, N\) and at least one of the cashflows is strictly positive, that is for some \(i \in [1, N]\), \(X_i > 0\).

The definition makes sense as a quantification of ‘free’ money, since every payment is non-negative with at least one positive payment. No one would refuse such an arbitrage cashflow stream. Strictly speaking, we have defined a deterministic arbitrage opportunity which corresponds to the setting of deterministic cashflows. Later on, we will generalize this concept to the case when cashflows can be non-deterministic.

**Example 1.2.** A deal which offers you $1 at times \(t = 1\) and \(t = 2\) is the cashflow stream \((1, 1, 1, 2)\). This is an arbitrage opportunity and you should desire infinite amounts of it. Suppose, instead, that the deal offers you $1 at time \(t = 1\) but you must pay $1 at time \(t = 2\). The cashflow stream is \((1, -1, 1, 2)\). This looks like a desirable stream that you should want in infinite amounts of (since you can put the $1 you get in the bank for a year), but it is not an arbitrage opportunity. Even if we changed the cashflow stream to \((100, -1, 1, 2)\), intuitively this is very desirable, but it is not an arbitrage opportunity. \(\square\)

Any financial dealing can be reduced to a stream of cashflows promised by one party to another. Our setting is deterministic, so the stream of cashflows is deterministic. There is a market place (for example the Chicago Board of Exchange (CBOE)) where such deals can be bought or sold by anyone, for a price that is determined in the market. Assume that anyone can borrow or lend money to a bank at an interest rate \(r\). This bank is exogenous to the market. We summarize our assumptions below. The first assumption is benign. The second has to be taken with a grain of salt because typically the rate at which you can borrow money (from a bank) is higher than the rate at which you can lend money.

1. Any financial deal (cashflow stream) can be bought or sold at a price.
2. Anyone can borrow or lend any amount of money at interest rate \(r\).

Though you can borrow or lend at interest rate \(r\), ultimately your account with the bank must end at zero. That is, if you borrowed (lent) money, ultimately it must be paid back (retrieved) with interest. One of our main tasks is to figure out how to compute the price of financial deals that can be bought or sold in the market. Our main tool is the no-arbitrage axiom.

**Axiom 1.3 (No Arbitrage).** Nobody has access to an arbitrage cashflow stream.

Let’s now see how the no-arbitrage axiom helps us to price financial deals. The only types of financial deals available to us at the moment involve the exchange of cashflows between two parties. So let’s take an example financial deal.

You pay $100 at time \(t_1\), you get $200 at time \(t_2\).

The price to buy this deal is \(P\). Do you want it? The deal, is equivalent to the cashflow stream \((-100, 200; t_1, t_2)\). The present value of this cashflow stream is

\[
PV(-100, 200; t_1, t_2) = -100 \cdot e^{-rt_1} + 200 \cdot e^{-rt_2}.
\]

\(^2\)An old joke has an economist walking in the park. He comes across a $20 bill lying on the path and walks right buy it. A child walking behind sees the money and picks it up. The child later catches up to the economist and asks why he (the economist) did not pick up the free $20 bill. “Son, you’re a very silly boy. Don’t you know that free money cannot possibly exist.”
1. Financial Instruments

1.1. Interest: How Money Grows

This present value is how much you should pay for this deal, and we will now prove it using the no-arbitrage concept. We will show that if the price is not the present value, then you can construct an arbitrage opportunity for yourself which violates the no-arbitrage axiom. The arbitrage argument has two steps.

(i) Suppose $P < -100 \cdot e^{-rt_1} + 200 \cdot e^{-rt_2}$. (The deal is cheap.) Buy the deal. To do so, borrow $P$ from the bank. Your cashflow at time 0 is $P$ (from bank) − $P$ (to buy deal) = 0. At time $t_1$ borrow $100$ to pay the $100$ that the deal requires you to pay; again you experience a nett zero cashflow. At time $t_2$ you get $200$. You must also pay off the $P$ you borrowed at time zero and the $100$ you borrowed at time $t_1$ so that your account with the bank is squared away to zero. The amount you owe at time $t_2$ is $Pe^{rt_2} + 100e^{r(t_2-t_1)}$. So the cashflow stream you have created for yourself is $(0, 0, 200 - Pe^{rt_2} - 100e^{r(t_2-t_1)}; 0, t_1, t_2)$.

The cashflow at $t_2$ is positive because

$$200 - Pe^{rt_2} - 100e^{r(t_2-t_1)} = e^{rt_2}(200e^{-rt_2} - 100e^{-rt_1} - P) > 0,$$

where the inequality follows by our assumption on $P$ and so you have an arbitrage cashflow stream. This contradicts the no-arbitrage axiom.

(ii) Suppose $P > -100 \cdot e^{-rt_1} + 200 \cdot e^{-rt_2}$. (The deal is expensive.) Sell the deal for $P$ and put in the bank to create an arbitrage opportunity (see next exercise).

**Exercise 1.7**

Show that if $P > -100 \cdot e^{-rt_1} + 200 \cdot e^{-rt_2}$, you can get an arbitrage opportunity by selling the deal at time 0 and putting the $P$ in the bank. When you sell the deal, your cashflows will be $+100$ at time $t_1$ and $-200$ at time $t_2$.

(a) What is your cashflow at time 0.

(b) What will you do at times $t_1$ and $t_2$.

(c) What is the entire cashflow stream you have constructed for yourself.

(d) Show that this cashflow stream is an arbitrage opportunity.

If there are no arbitrage opportunities, then cases (i) and (ii) cannot occur. It means that

$$P = -100 \cdot e^{-rt_1} + 200 \cdot e^{-rt_2} = PV(-100, 200; t_1, t_2)$$

is the only possibility, and at this price there will indeed be no arbitrage opportunities.

We have determined the no-arbitrage price of this ‘complex’ financial deal. Why do we care about pricing such complex deals in this way. The no-arbitrage price is sacred knowledge. If you know this price, and the deal is being quoted to you at some other price, then you know you can enter the deal and make sure money somehow. It turns out that large financial institutions spend a lot of energy trying to determine the profitability of complex deals that arrive at their doorstep. If the price is not the no-arbitrage price, you can see how a financial institution would love to take on that deal and make free money.

Do we need to go through such a complex argument to price every new deal? No. Life gets simpler because what we saw in our example can be generalized. The present value of an arbitrary deal’s (deterministic) cashflow stream is the no-arbitrage price of the deal. The no-arbitrage axiom implies the following theorem.

**Theorem 1.4.** The no-arbitrage price of any deterministic cashflow stream is the present value of the cashflow stream.

The proof of Theorem 1.4 follows exactly the same two steps in our example deal above, and we leave it as an exercise.
1.2 The Bond

Now that we are comfortable with the interest rate and the way money grows, here is the bad news: there is no stationary fixed interest rate $r$ in the market, at which we can borrow and lend money. The closest we can get to risk free growth is when we lend our money to the government who promises us interest in return. This is done with the treasury bond.

The Zero Coupon Bond. A contract that “guarantees” you $1 at a time $T$ in the future. The price of the bond today is $B(T)$.

The time $T$ is the maturity (or term) of the bond and the amount it pays ($1$) is the face value. In the market, the face value is typically $100$ so the price is $100B(T)$, and such bonds are issued every quarter at maturities $0.25, 0.5, 1, 2, 5, 10, 30$ (in years). The bond is like the bank account, except you put $B(T)$ in the “bank” and you withdraw it at time $T$ after it grows to $1$; you cannot withdraw your money earlier or later. The interest you make over this period $T$ corresponds to an interest rate which we can deduce from the formula

$$1 = B(T)e^{\bar{r}T}$$

because $B(T)$ invested in the bond at time zero grows to $1$ at time $T$. The parameter $\bar{r}$ is an effective (average) interest rate being paid by the bond over the period zero to $T$. This effective interest rate is called the yield,

$$\bar{r}(T) = \frac{1}{T} \ln \frac{1}{B(T)}.$$ 

If we were in the stationary setting, bonds with different maturities would give the same yield. To be more general we allow the yield to depend on the maturity $T$. How the yield depends on $T$ is called the term structure of the interest rate, and in the stationary setting, the term structure will be flat. Figure 1.1 shows how the 30 year, 10 year and 5 year yields have varied over the recent past. The three curves have similar behavior, but they are not identical. The yield does indeed depend on the maturity.
1. Financial Instruments

1.2. The Bond

Figure 1.1: The US treasury yield. The yield (interest rate) changes with time; this is also reflected in the fact that the yield to different maturity (5,10,30 years) are not the same. In particular, this implies that $R(0,5)$ and $R(5,10)$ are not equal, so the market is not paying equal interest over equal sized periods. However, the yields to different maturities are not far off from each other.

Exercise 1.9
From Figure 1.1, we observe that during periods of very high interest rate (e.g. early 1980s) the 30 year yield tends to be below the 5 year yield. During periods of rapidly dropping or very low interest rates, (e.g. around 2012), the reverse is true. During somewhat stable, normal interest rates around 6-8%, the three yields are about equal. If we plot the term structure of the yield (yield versus time to maturity) during these times we get three different looking curves.

Explain these different types of term structure during different interest rate regimes.
From Equation (1.4), $B(T) = e^{-rT}$, so, over time zero to $T$, the bond price is the discount factor with effective interest rate $\bar{r}$. We can redefine present value using the bond price,
\[
PV(X; T) = B(T)X.
\]
If you want to generate $X$ at time $T$, then you would buy $X$ bonds maturing at time $T$. The total price you pay is $X \times B(T)$ which by definition of the present value – the money you need now to generate the $X$ at time $T$. It remains true that the present value, when defined this way, is the price $P$ of a financial deal which delivers a cashflow stream $(X; T)$. This follows from a no-arbitrage argument as follows. Suppose that $P < B(T)X$ (the deal is cheap). Here is a way to construct an arbitrage opportunity.

At time 0. Sell $P/B(T)$ bonds at price $B(T)$ (to generate a cash inflow of $P$) and buy the deal at price $P$.
Your nett cashflow is 0.

At time $T$. You owe $P/B(T)$ to the person you sold the bond (you sold $P/B(T)$ bonds and each delivers $\$1$).
You get $X$ from the deal. So, your nett cashflow is
\[
X - \frac{P}{B(T)} = \frac{B(T)X - P}{B(T)} > 0.
\]
Thus, you have an arbitrage opportunity, contradicting the no-arbitrage axiom. Therefore our assumption that $P < B(T)X$ must be false. In a similar way, if $P > B(T)X$, then by selling the deal and buying bonds, you can construct an arbitrage opportunity. The conclusion is that $P = B(T)X$.

There is something subtle about the strategy when the deal is cheap. At time zero, you are to sell $P/B(T)$ bonds. How do you sell something you do not have? The market has a mechanism for this, called selling short. We will not worry about the mechanics. For our purposes, the bond is a financial deal and as such is a cashflow stream. Buying the bond means you will get that cashflow stream, and selling the bond means you will get the negative of the cashflows in the stream.

The bond plays the role of our old friend interest and determines both the rate at which money grows to time $T$, as well as the discount factor for computing the present value of a cashflow at time $T$. If we have bonds maturing at different times, we can compute the present value of arbitrary cashflow streams. Let $(X_1, \ldots, X_N; t_1, \ldots, t_N)$ be a cashflow stream, and let $B(t_i)$ be the price of the bond with maturity $t_i$. The present value of cashflow $X_i$ at time $t_i$ is $B(t_i)X_i$, and the present value of the entire cashflow stream is the sum of the present values,
\[
PV(X_1, \ldots, X_N; t_1, \ldots, t_N) = \sum_{i=1}^{N} X_i B(t_i).
\] (1.5)

Theorem 1.4 still holds. That is, the no-arbitrage price of any deterministic cashflow stream is its present value, where the present value is defined using (1.5).

**Exercise 1.10**
Rewrite the formula for the present value using $\tilde{r}(t_i)$, the yield of the bond whose maturity is $t_i$. (You should get a formula similar to the one in Exercise 1.8.)

**Coupon Paying Bonds.** A zero coupon bond with face value $\$1$ has a payment of $\$1$ at maturity. A bond can have intermediate payments called coupons. Such a bond has a face value as well as an interest rate $r$. Typically, the coupon is $r/2 \times$ (face value) paid every half year up to and including maturity. Such bonds behave like a bank paying an annual interest rate of $r$ (compounded every half year). A coupon paying bond is a stream of cashflows, rather than just a single cashflow at maturity. Here is an example obtained from Treasury Direct, where one can purchase US-Treasury securities.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Issue</th>
<th>Maturity</th>
<th>Interest</th>
<th>Yield</th>
<th>Price (per $100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>Date</td>
<td>Date</td>
<td>Rate ($r$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year</td>
<td>09-16-2013</td>
<td>09-15-2016</td>
<td>0.875%</td>
<td>0.913%</td>
<td>99.887890</td>
</tr>
</tbody>
</table>

Setting 09-16-2013 as time $t = 0$, this 3-year treasury bill corresponds to the following stream of cash flows
simplicity, your strategy will be to buy \( \alpha \) units \( \alpha \) of bonds, with maturities \( \alpha \) and \( \alpha \). Thus, the prices of the two bonds are \( \alpha \) and \( \alpha \). We obtain this by multiplying their value today by \( \alpha \) and \( \alpha \). Hence, you would like to set aside some money \( \alpha \) now, in order to “guarantee” that you will be able to meet your commitment at time \( \alpha \). Hence, you would like to have as little uncertainty as possible at time \( \alpha \).

Exercise 1.12
Show that for any choice of \( \alpha_1, \alpha_2 \) satisfying the settlement constraint, the present value of the bond portfolio is the same. Why must this be the case?

1.2.1 Bond Portfolio Immunization

Bond portfolios alone present several computational tasks that are generally non-trivial. We will now consider one such task, which is that of immunizing a bond portfolio. To understand the need for this, we need to cast the problem in a practical setting. Specifically, suppose that you have a payment \( \alpha \) which is due at time \( \alpha \), and you would like to have as little uncertainty as possible at time \( \alpha \). Hence, you would like to set aside some money now, in order to “guarantee” that you will be able to meet your commitment at time \( \alpha \).

The simplest thing to do is to purchase \( \alpha \) zero coupon bonds with maturity \( \alpha \) which will cost you \( \alpha \cdot B(t) \) now. You now have no need to worry about meeting your commitment because at time \( \alpha \) you will owe \( \alpha \) which is exactly the amount that your bonds will furnish you with at that time. The problem that you will face when you attempt to implement this strategy is that you will generally not find zero coupon bonds available on the market with maturity exactly at time \( \alpha \). Typically, the bonds are issued at discrete times, with discrete time to maturity (usually bonds are issued quarterly and at time of issue, the time to maturity is usually either 3 months, 1, 5, 10, 30 years). As time passes, the time to maturity for a given bond will decrease, however there will still only be a discrete set of maturities available to you. All of a sudden, your task has taken a non trivial twist.

Let’s try to formalize the problem we would like to solve, and at the same time simplify the setting to illustrate the essential issues. Your commitment is \( \alpha \) at time \( \alpha \); in the market place are two bonds, with maturities \( \alpha < \alpha \). Let’s for simplicity assume that the market perceives a stationary interest rate, and so there is a well defined instantaneous interest rate \( \alpha \). Thus, the prices of the two bonds are \( \alpha \) and \( \alpha \). For simplicity, your strategy will be to buy \( \alpha_1 \) and \( \alpha_2 \) units of each bond (you will not hold cash to settle your debt – this task is already completed). What are \( \alpha_1 \) and \( \alpha_2 \) units of each bond worth at time \( \alpha \)? We obtain this by multiplying their value today by \( \alpha \). Since we need to settle our debt of \( \alpha \), we see that \( \alpha, \alpha_2 \) must satisfy the settlement constraint

\[
\alpha_1 e^{-\alpha(t_1 - \alpha)} + \alpha_2 e^{-\alpha(t_2 - \alpha)} = \alpha.
\]
How should we select \( \alpha_1, \alpha_2 \). The market being stationary means that there is a constant rate of interest \( r \). This can be viewed as the perceived rate of interest. After some time \( \Delta t \), this instantaneous interest rate may change to \( r + \Delta r \). The settlement condition for the new interest rate \( r + \Delta r \) at time \( \Delta t \) may not be satisfied any more. In fact the degree to which it is not satisfied can be viewed as an instability in our bond portfolio, or as a sensitivity to the interest rate. We can measure the degree to which we are immunized to changes in the interest rate by how much the settlement condition is violated. Thus, we consider

\[
\epsilon(\Delta r; \alpha_1, \alpha_2) = \alpha_1 e^{-(r+\Delta r)(t_1-t)} + \alpha_2 e^{-(r+\Delta r)(t_2-t)} - X
\]

as a measure of how sensitive we are to the interest rate. \( \epsilon(\Delta r) \) measures how secure I am in the knowledge that I will be able to meet my debt with minimal additional cash flows at time \( t \). We emphasize that if \( \epsilon(\Delta r) > 0 \) then it is certainly a good thing for you, as this means you expect to not only cover your commitment but also get cash flow back. However we view the goal as not to maximize profit, but rather to reduce uncertainty. We wish to make \( \epsilon(\Delta r) \) a small, if not zero function of fluctuations in the interest rate. Lets formalize this notion. Expanding \( \epsilon(\Delta r) \) as a power series, we obtain

\[
\epsilon(\Delta r; \alpha_1, \alpha_2) = a_0(\alpha_1, \alpha_2) + a_1(\alpha_1, \alpha_2) \Delta r + a_2(\alpha_1, \alpha_2) \Delta r^2 + \cdots ,
\]

\[
= \sum_{k=0}^{\infty} a_k(\alpha_1, \alpha_2) \Delta r^k.
\]

**Exercise 1.13**

Show that

\[
\begin{align*}
a_0(\alpha_1, \alpha_2) &= 0; \\
a_1(\alpha_1, \alpha_2) &= \alpha_1 (t-t_1)e^{-r(t_1-t)} + \alpha_2 (t-t_2)e^{-r(t_2-t)}; \\
a_2(\alpha_1, \alpha_2) &= \alpha_1 (t-t_1)^2 e^{-r(t_1-t)} + \alpha_2 (t-t_2)^2 e^{-r(t_2-t)}.
\end{align*}
\]

Obtain a general formula for \( a_k(\alpha_1, \alpha_2) \).

A natural approach to minimizing the sensitivity to interest rate fluctuations is for \( \epsilon \) to depend only on high order powers of \( \Delta r \). Since the settlement constraint provides only one constraint, and we have two parameters, we can choose the parameters \( \alpha_1, \alpha_2 \) so that the first order dependence disappears and \( \epsilon \) will fluctuate only in proportion to \( \Delta r^2 \). Thus we minimize the sensitivity to the interest rate by setting

\[
\alpha_1(\alpha_1, \alpha_2) = 0.
\]

This constraint, together with the settlement constraint gives two equations in two unknowns. Solving gives the bond portfolio that optimally immunizes you to interest rate risk. We leave the details as an exercise.

**Exercise 1.14**

Let \( \rho = (t-t_2)/(t-t_1) \). Show that the optimally immunized bond portfolio is obtained by setting

\[
\alpha_1 = -\frac{\rho}{1-\rho} X \cdot e^{-r(t-t_1)}, \quad \alpha_2 = \frac{1}{1-\rho} X \cdot e^{-r(t-t_2)}.
\]

Thus, while we do not know exactly what the interest rate fluctuation will be, we can attempt to minimize our exposure to it. An alternative way of writing the settlement constraint is to use present values, present values,

\[
\alpha_1 e^{-rt_1} + \alpha_2 e^{-rt_2} = X e^{-rt}.
\]

Correspondingly, one could define the sensitivity \( \epsilon \) by

\[
\epsilon(\Delta r; \alpha_1, \alpha_2) = \alpha_1 e^{-(r+\Delta r)(t_1-\Delta t)} + \alpha_2 e^{-(r+\Delta r)(t_2-\Delta t)} - X e^{-(r+\Delta r)(t-\Delta t)}
\]

**Exercise 1.15**

Show that using sensitivities for the present value constraint leads to exactly the same optimal portfolio.
The next exercise generalizes exactly this task to a more general setting. Specifically, assume that your commitments are \( X_1, \ldots, X_m \) at the times \( \tau_1, \ldots, \tau_m \), and in the market place you have access to zero coupon bonds with maturities \( t_1, \ldots, t_n \). Assume that the current instantaneous interest rate is \( r \).

Exercise 1.16
We discuss one approach to immunizing the bond portfolio that meets the commitments \((X_1, \ldots, X_m; \tau_1, \ldots, \tau_m)\).

Suppose that the bond portfolio which you choose contains \( \alpha_i \) units of the bond with maturity \( t_i \).

(a) Show that the settlement constraint is given by
\[
\sum_{i=1}^{n} \alpha_i e^{-rt_i} - \sum_{j=1}^{m} X_j e^{-r\tau_j} = 0.
\]

(b) At time \( \Delta t \), assume that the interest rate changes to \( r + \Delta r \). We define the sensitivity with respect to fluctuation in the interest rate \( \Delta r \) over the time \( \Delta t \), denoted \( \epsilon(\Delta r; \alpha_1, \ldots, \alpha_n) \), to be the amount by which the settlement condition is violated. Show that
\[
\epsilon(\Delta r; \alpha_1, \ldots, \alpha_n) = \sum_{i=1}^{n} \alpha_i e^{-(r+\Delta r)(\tau_i - \Delta t)} - \sum_{j=1}^{m} X_j e^{-(r+\Delta r)(\tau_j - \Delta t)}.
\]

(c) Write \( \epsilon(\Delta r; \alpha_1, \ldots, \alpha_n) \) as a Taylor series in \( \Delta r \),
\[
\epsilon(\Delta r; \alpha_1, \ldots, \alpha_n) = \sum_{k=1}^{\infty} a_k(\alpha_1, \ldots, \alpha_n)\Delta r^k.
\]
What is \( a_0 \)?

(d) Argue that since there are \( n \) parameters, \( \alpha_1, \ldots, \alpha_n \) (one for each bond), and since the settlement constraint provides one constraint, we should be able to set \( n - 1 \) of the coefficients to 0. Which coefficients do you choose to set to 0 and why?

(e) The natural choice which gives optimal immunization is to set
\[
a_1 = a_2 = \cdots = a_{n-1} = 0.
\]
Define \( a_k'(\alpha_1, \ldots, \alpha_n) \) by
\[
a_k'(\alpha_1, \ldots, \alpha_n) = \sum_{i=1}^{n} \alpha_i t_i^k e^{-rt_i} - \sum_{j=1}^{m} X_j \tau_j^k e^{-r\tau_j}.
\]
Show that \( a_1 = a_2 = \cdots = a_{n-1} = 0 \) if and only if \( a_1' = a_2' = \cdots = a_{n-1}' = 0 \). Note that this set of equations is independent of \( \Delta t \).

(f) Suppose one also held cash. How does our general solution apply even in this setting?
Exercise 1.17

Show that this set of equations, in Exercise 1.16(e) together with the settlement constraint has a unique solution.

Here is an outline of the argument. Let \( \beta_k = \sum_{j=1}^{n} X_j \tau_j^k e^{-r \tau_j} \), and define \( \alpha, \beta \) as follows.

\[
\alpha = \begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_n
\end{bmatrix}, \quad \beta = \begin{bmatrix}
\beta_1 \\
\beta_2 \\
\vdots \\
\beta_{n-1}
\end{bmatrix}.
\]

Show that the set of constraints can be written in the matrix form

\[
\begin{bmatrix}
\lambda_1 & \lambda_2 & \cdots & \lambda_n \\
\lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_1 x_1^{n-1} & \lambda_2 x_2^{n-1} & \cdots & \lambda_n x_n^{n-1}
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_n
\end{bmatrix}
= \begin{bmatrix}
\beta_1 \\
\beta_2 \\
\vdots \\
\beta_n
\end{bmatrix},
\]

where \( \lambda_i = e^{-r t_i} > 0 \) and \( x_i = t_i \). More compactly, this can be written as

\[ \Lambda \alpha = \beta, \]

for an appropriately defined \( \Lambda \). The task is to now show that \( \Lambda \) is invertible. This can be done by showing that

\[ \Lambda = VL, \]

where \( V \) is a Vandermonde matrix and \( L \) is a positive diagonal matrix,

\[
V = \begin{bmatrix}
1 & 1 & \cdots & 1 \\
x_1 & x_2 & \cdots & x_n \\
x_1^2 & x_2^2 & \cdots & x_n^2 \\
x_1^3 & x_2^3 & \cdots & x_n^3 \\
\vdots & \vdots & \ddots & \vdots \\
x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1}
\end{bmatrix}, \quad L = \begin{bmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & 0 & \cdots \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n
\end{bmatrix}.
\]

Clearly \( L \) is invertible, and it is well known that the Vandermonde matrix is invertible if and only if all the \( x_i \) are distinct. (The energetic may wish to prove this.) Show that in our case, all the \( x_i \) are distinct to conclude the proof.

This exercise shows that the immunization problem reduces to solving a linear system involving a Vandermonde matrix. Clearly this system can be solved by inverting the Vandermonde matrix. We refer the interested reader to texts on matrix analysis (e.g., \( \beta \)) for efficient algorithms to invert Vandermonde matrices. When the bond maturities have the form \( t_i = i \cdot t \), then the Vandermonde matrix \( V \) has a very special form, and can be inverted in closed form, \( \beta \).

Exercise 1.18

Consider a stream of commitments of $100 at each of the times 1, 2, 3. Let the interest rate be 0.05, and assume that there are 10 bonds in the market place with maturities \( \frac{1}{2}, 1 \frac{1}{2}, 2 \frac{1}{2}, \ldots, 9 \frac{1}{2} \). You would like to build an immunized bond portfolio to guarantee this cashflow stream.

(a) What is the maximum value of \( k \) for which you can be immunized to fluctuations up to order \( \Delta r^k \).

(b) Construct the bond portfolio which is immunized up to order \( \Delta r^k \) for the value of \( k \) in the previous part.

(c) Plot the amount by which the settlement constraint is violated as a function of the fluctuation in the interest rate for interest rate fluctuations \( \Delta r \in [-0.02, 0.02] \). The settlement constraint violation is given by

\[
\epsilon(\Delta r) = e^{(r + \Delta r) \Delta t} \left( \sum_{i=1}^{m} \alpha_i e^{-(r + \Delta r) t_i} \sum_{j=1}^{m} X_j e^{-(r + \Delta r) t_j} \right).
\]

Select a value of \( \Delta t \) for your plot, for example \( \Delta t = \frac{1}{12} \) corresponds to 1 month.
Exercise 1.19
For the general case of one commitment and \( n \) bonds, one could define the settlement constraint and sensitivity \( \varepsilon \) either using the present value or the value at the time of the commitment. Show that both ways lead to the same optimally immunized portfolio.

Exercise 1.20
A much harder bond portfolio immunization problem arises if there are a large number (hundreds) of bonds with different maturities available, and there are constraints on the size of the bond portfolio. For example, imagine there are 1000 bond maturities, but your bond portfolio is limited to size 3.

Naturally, any 3 bond maturities will serve to immunize the settlement constraint violation against fluctuations of order 2, but what is the best set of 3 maturities to select. One criterion is that the coefficient of the 3rd order fluctuation should be minimized. More generally, suppose that there are \( n \) maturities available, but the portfolio is limited to size \( \ell \).

(a) Formulate the problem as one of finding the optimal set of bonds which immunize the portfolio to fluctuations of order \( \ell - 1 \) and minimizes the magnitude of the order \( \ell \) coefficient of the fluctuation. Specifically show that the problem reduces to the following mixed integer-linear program (MILP),

\[
\min_z, \quad \text{s.t.} \\
\alpha_i \in \{0, 1\}, \\
\sum_{i=1}^{n} \alpha_i \leq k, \\
\sum_{i=1}^{n} \alpha_i t_i e^{-rt_i} - \sum_{j=1}^{m} X_j \tau_j e^{-r\tau_j} = 0, \quad k \in \{0, \ldots, \ell - 1\}, \\
\sum_{i=1}^{n} \alpha_i t_i e^{-rt_i} - \sum_{j=1}^{m} X_j \tau_j e^{-r\tau_j} \leq z, \\
- \sum_{i=1}^{n} \alpha_i t_i e^{-rt_i} + \sum_{j=1}^{m} X_j \tau_j e^{-r\tau_j} \leq z.
\]

(b) Formulate a brute force strategy for solving this problem, and determine its running time complexity.

(c) Formulate a greedy heuristic strategy for solving this problem and determine its runtime complexity. Run some experiments to compare the performance of the greedy with the brute force.

(d) [hard, open-ended] Can you formulate a heuristic with a provable approximation guarantee on the quality of the solution with respect to the optimal solution. One approach would be to relax this MILP to an LP and then use a randomized rounding technique to obtain a valid solution. Then show that the expected value of the objective after randomized rounding is within some approximation factor of the objective in the relaxed LP solution. Since the solution to the relaxed LP gives an upper bound on the solution to the MILP, you have your result.

Exercise 1.21
Other important complex contracts derived from bonds are interest rate caps and interest rate swaptions. We define these complex contracts here, and the energetic reader may want to ponder how to price these contracts.

1.2.2 Yield Curve Interpolation from Coupon Bond Data

Just as a reminder, the zero coupon price for the bond with maturity \( t \) is \( B(t) \), and the yield \( Y(t) \) is defined by \( Y(t) = -\frac{\log B(t)}{t} \). Typically, bonds are only available at discrete maturities. A sample of the daily yields as published by the US treasury department (http://www.ustreas.gov/) is given in the table below.
1. Financial Instruments

1.2. The Bond

<table>
<thead>
<tr>
<th>Date</th>
<th>1 mo</th>
<th>3 mo</th>
<th>6 mo</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>10</th>
<th>20</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/02/90</td>
<td>7.83</td>
<td>7.89</td>
<td>7.81</td>
<td>7.87</td>
<td>7.90</td>
<td>7.87</td>
<td>7.98</td>
<td>7.94</td>
<td>8.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/03/90</td>
<td>7.89</td>
<td>7.94</td>
<td>7.85</td>
<td>7.94</td>
<td>7.96</td>
<td>7.92</td>
<td>8.04</td>
<td>7.99</td>
<td>8.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/04/90</td>
<td>7.84</td>
<td>7.90</td>
<td>7.82</td>
<td>7.92</td>
<td>7.93</td>
<td>7.91</td>
<td>8.02</td>
<td>7.98</td>
<td>8.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01/05/90</td>
<td>7.79</td>
<td>7.85</td>
<td>7.79</td>
<td>7.90</td>
<td>7.92</td>
<td>7.92</td>
<td>8.03</td>
<td>7.99</td>
<td>8.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>08/26/08</td>
<td>1.67</td>
<td>1.71</td>
<td>1.95</td>
<td>2.19</td>
<td>2.35</td>
<td>2.64</td>
<td>3.06</td>
<td>3.37</td>
<td>3.79</td>
<td>4.43</td>
<td>4.40</td>
</tr>
<tr>
<td>08/27/08</td>
<td>1.58</td>
<td>1.67</td>
<td>1.93</td>
<td>2.16</td>
<td>2.31</td>
<td>2.58</td>
<td>3.02</td>
<td>3.34</td>
<td>3.77</td>
<td>4.41</td>
<td>4.38</td>
</tr>
<tr>
<td>08/28/08</td>
<td>1.63</td>
<td>1.74</td>
<td>1.98</td>
<td>2.19</td>
<td>2.37</td>
<td>2.62</td>
<td>3.09</td>
<td>3.42</td>
<td>3.79</td>
<td>4.41</td>
<td>4.38</td>
</tr>
<tr>
<td>08/29/08</td>
<td>1.63</td>
<td>1.72</td>
<td>1.97</td>
<td>2.17</td>
<td>2.36</td>
<td>2.60</td>
<td>3.10</td>
<td>3.45</td>
<td>3.83</td>
<td>4.47</td>
<td>4.43</td>
</tr>
</tbody>
</table>

Notice that in the early days, some maturities were not available since the US government was not issuing those bonds. The immediate problem is one of obtaining the present value of a cashflow at a maturity which is not one of the discrete maturities available. One needs to interpolate bond prices in order to obtain prices at other maturities.

The typical properties one would like of any interpolation method are that it be continuous, and that it be local. Local means that the value of the interpolated price at some maturity \( \tau \) depends only on the available maturities close to \( \tau \). An additional bonus is that the interpolated function has a continuous derivative. There are many approaches for interpolation of yield curves \( \? \). The simplest is the linear interpolation which is local, continuous, though not continuously differentiable.

Even with linear interpolation, the results differ depending on whether one interpolates the bond prices \( B(t) \) or whether one interpolates the yields \( Y(t) \), and obtain the bond prices from the resulting interpolated yield.

For a set of maturities \( t_1 \leq t_2 \leq \cdots \leq t_n \) and corresponding values \( y_1, y_2, \ldots, y_n \), for a given \( t \) such that \( t_i \leq t < t_{i+1}, \) the interpolated value is given by

\[
I(t) = y_i \frac{t_{i+1} - t}{t_{i+1} - t_i} + y_{i+1} \frac{t - t_i}{t_{i+1} - t_i},
\]

Linear interpolation of the yields gives smoother looking bond prices. One can now use the interpolated bond prices for computing present values.

1.2.3 Yield Curve Interpolation from Swap Data

Exercise ?? gives a relationship between the par swap rates and the bond prices. Swap rates are also periodically published by the Federal reserve, again at discrete maturities. The swap rates could also be used to infer the yields, which can then be interpolated.
1.2.4 Data Driven Yield Curve Dynamics

If we look at the yield curve over time, it is definitely not constant. The daily evolution of the yield curve over a short period of time is shown in the next figure. Each line is the yield curve for a particular day.

While the yield curve is not constant, it is believed that the yield curve exhibits only a few main degrees of freedom which capture most of its dynamics. Typically three degrees of freedom are used, often denoted by the shift (the whole curve shifts up or down), the rotation (the whole yield curve rotates) and the bend (the center of the yield curve moves in one direction while the extremes move in the other).

The fact that a small number of degrees of freedom should suffice to describe yield curve dynamics suggests that it should be possible to effectively model yield curve dynamics using some kind of Markov Models.

Modeling Yield Curve Dynamics Using Hidden Markov Models

Hidden Markov Models are very useful for modeling processes in computational finance, and we will use the context of yield curve modeling to introduce them. A Hidden Markov Model (HMM) is a state based model. The market exists in various states, for example the high interest rate state, the low interest rate state, the upward sloping yield curve state, the inverted yield curve state (downward sloping), etc. The actual yield curve depends on the state the market is in. Finally, the market may transition from one state to the next, according to some transition probability distribution.

Suppose that there are $K$ state $S = \{s_1, \ldots, s_K\}$. If the market is in state $s_i$, it will transition to state $s_j \in S$ with probability $P_{ij} \geq 0$, where $\sum_{j=1}^{K} P_{ij} = 1$. The matrix $P$ is called the state transition matrix.

Given that the state of the market on a particular day is $s_i$, the observed yield curve $x$ will be drawn from some (state dependent) distribution $P_i(x)$. The Hidden Markov Model is then fully specified by the initial state (or initial probability distribution for the initial state), the transition matrix $P$ and the state dependent densities $P_i$. Let $\pi_i(t)$ be the probability that the state at time $t$ is $s_i$. Then the probability distribution for
the yield curve at time \( t + 1 \) is given by

\[
P_{t+1}(x) = \sum_{i=1}^{K} \pi_i(t) \sum_{j=1}^{K} P_{ij} P_i(x).
\]

\( P_{t+1}(x) \) is the probability distribution for the yield curve at time \( t + 1 \). The first sum is over the possible states at time \( t \), weighting each state by the probability of being in that state. The second sum considers all possible states that one can transition to from the state at time \( t \), weighting by the probability of transitioning to that state and the probability of observing \( x \) from the state transitioned into. Note that we can obtain a relationship between \( \pi_j(t) \) and \( \pi_i(t+1) \):

\[
\pi_i(t+1) = \sum_j \pi_j(t) P_{ji},
\]
or in more compact vector notation,

\[
\pi(t+1) = P^T \pi(t).
\]

**Exercise 1.22**

Show that if \( \sum_i \pi_i(t) = 1 \), then \( \sum_i \pi_i(t+1) = 1 \).

By induction one then has that

\[
\pi(t) = (P^T)^t \pi(0).
\]

We will assume that the state dependent yield curve distribution \( P_i(x) \) is a Normal distribution. The number of dimensions in the vector \( x \) is equal to the number of maturities for which the yield curve data exists. Thus,

\[
P_i(x) = N(x; \mu_i, \Sigma_i),
\]

\[
= \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp \left( -\frac{1}{2} (x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i) \right).
\]

The Hidden Markov Model is therefore fully specified by \( \pi(0), P, \{\mu_i, \Sigma_i\}_{i=1}^{K} \). The first task is to estimate, or *calibrate* these parameters of the HMM to observed market yield curves.

**Immunization of Bond Portfolios to Data Driven Yield Curve Dynamics**

### 1.3 The Stock

A share of stock in a company entitles the owner to a part ownership in the company. The payoff for owning stock in a company typically comes in two forms: dividends (cash flows) that are distributed by the company to share holders and capital gain of the company – the company becomes worth more, hence the value of your part ownership in the company has increased. We will not spend much time discussing the ins and outs of stock valuation, and how one would analyze the financial books of a company in order to determine its value. Rather, here we will give a very simple idea of how much one should be willing to pay for a share of stock in a company.

Assume that cash flows in the form of dividends are expected to be distributed yearly in the amounts of \( X_1, X_2, \ldots \), and assume that the 1 year interest rate \( R(1) \) is known. In this case let \( P_0 \) be the be the price of the share now, then using the tools of the previous section, we know that the price \( P_0 \) should be the present value of the future stream of cash flows,

\[
P_0 = PV(X_1, X_2, X_3 \ldots; 1, 2, 3 \ldots),
\]

\[
= \frac{X_1}{1 + R(1)} + \frac{X_2}{(1 + R(1))^2} + \frac{X_3}{(1 + R(1))^3} + \cdots.
\]

Similarly, the price of the share at time 1, denoted \( P_1 \), would be the present value computed at time 1,

\[
P_1 = \frac{X_2}{1 + R(1)} + \frac{X_3}{(1 + R(1))^2} + \cdots.
\]
Multiplying $P_0$ by $1 + R(1)$, we see that

$$R(1)P_0 = X_1 + P_1 - P_0.$$ 

The RHS is the cash flow after 1 year plus the gain in the stock price (the capital gain). This sum is typically called the economic earnings per share (EPS), which includes the dividend plus any capital gain. Thus we see that

$$\frac{P_0}{EPS} = \frac{1}{R(1)}.$$ 

The LHS is typically called the price to earnings ratio (P to E ratio). Typically the yearly interest rate $R(1)$ is around 0.05, in which case the P to E ratio of a company should be about 20. Companies with a significantly lower P to E ratio could generally be considered a good buy, of course, assuming that the previous P to E ratio is a good indicator of the future P to E ratio.

The formula above explains a well known behavior in the markets. As the interest rate rises, so the P to E ratio will fall, and since typically the EPS will not rise, this means that the price will fall. Thus it is often the case the the Federal Reserve uses the yearly interest as a way to stimulate or hold back the stock markets.

### 1.3.1 Predicting Stock Prices

Of course, if we (the author) could do this, then we would stop writing immediately, get rich and go home. Thus, it is likely that your author cannot predict stock prices, however many people believe that this it is possible to predict stock prices. For those who are in this category, we will briefly discuss some approaches that one might take. Specifically, on the basis of what indicators should one try to do prediction. We briefly discuss some approaches that people have found useful. Are they profitable? well many of these techniques have been profitable when applied to historical data. The real question is whether they will remain profitable on future data. An alternative way to look at this problem is that the systems that have survived are precisely the ones that are profitable on historical data. If you try many systems, one of them will be profitable, by chance. There is no reason to believe that this system will be profitable in the future. These issues fall into the realm of inference. When can one conclude that the best system you have picked with respect to performance on historical data is really a system which is above random.

To drive the point further, suppose that all the systems you try are really just random. Inevitably some of these systems will do better than others on historical data. Picking one of these good performers has still left you with a random system. What has happened? This type of issue is sometimes called *data snooping* or *over fitting*. We will not address these issues here. Rather, we will discuss some indicators that you may want to use in developing a predictor. How you will determine whether your predictor performs well in practice is an issue you undertake at your own risk; in fact this task is an entire area of research itself, within the realm of statistical learning theory.

### Plotting or Charting Methods

The basic idea is to plot the price $P(t)$ and try to observe/detect trends. An example is shown below for a small sample of the price data from IBM, a fairly heavily traded stock.
The band seems to indicate both the trending behavior of the price of IBM stock, in addition to defining its region of activity. Extrapolating the upper and lower boundaries into the future in the natural way suggested by the picture, one might expect that a trading strategy which sells when the price approaches the upper boundary and buys when it approaches the lower boundary should be profitable. Along these lines, several tactics are used.

Support and Resistance Levels. Support and resistance levels are “psychological” barriers that the price appears not to break. A resistance is an upper barrier. Typically, the price will approach this upper barrier and then get reflected down. If one could determine a resistance, then it might be possible to make money by shorting the stock when the price approaches the resistance. A support is a lower barrier below which the price is not expected to fall.

Typical supports might be the previously attained high, or a round number that is above the current trading price. For example if the current trading price is $34.50, then a support might be $35.00. Similarly, a resistance tends to be the previous low, or a round number below the price. In the previous example, a resistance might be $34.00.

The reason for the round numbers is a purely psychological one, because most people will place limit buy orders at a round number below the trading price and similarly for limit sell orders above the trading price.

Trade Lines. Trade lines are similar to support and resistance in that they try to determine and extrapolate the trading region of a stock price by joining together successive maxima (to obtain the upper trade line) and successive minima (to obtain the lower trade line). This is illustrated with the small sampling from IBM stock price in the picture.

Bollinger Bands. Bollinger bands define the upper support curve and the lower resistance curve as some number of standard deviations above and below a specified moving average indicator (see below for the definition of moving average indicators). The standard deviation is typically computed on the same window length as the moving average.
1. Financial Instruments

1.3. The Stock

Exercise 1.23

In this example we will explore an algorithm to construct the lower support curve and the upper resistance curve as illustrated in the above example. As we see, the algorithm reasonably captures what we would intuitively expect of these lower and upper envelopes.

Let $S_0 = \{s_0, s_1, \ldots, s_T\}$ be the time series of prices at the times $0, 1, \ldots, T$. For simplicity, assume that all the $s_i$ are distinct. A local maximum of $S_0$ occurs at time $t$ if $s_{t-1} < s_t$ and $s_{t+1} < s_t$. Suppose that the local maxima of $S_0$ occur at the times $t_1, t_2, \ldots, t_k$. We define the level-1 upper envelope of $U_1$ as the time series $u_0, u_1, \ldots, u_T$ which linearly interpolates between the points $\{(0, s_0), (t_1, s_{t_1}), (t_2, s_{t_2}), \ldots, (t_k, s_{t_k}), (T, s_T)\}$, which is the sequence of local maxima together with the first and last point.

We define the level-2 upper envelope $U_2$ as the level-1 upper envelope of the time series $U_1$. Similarly we define $U_k$ recursively as the level-1 upper envelope of $U_{k-1}$.

Analogously, the level-1 lower envelope $L_1$ is defined as the linear interpolator of the local minima of $S_0$, and the level-$k$ lower envelope $L_k$ is the level-1 lower envelope of $L_{k-1}$.

The higher level envelopes are successively smoothed versions of the lower level envelopes.

(a) Implement this algorithm and show its result on a randomly generated time series. The example shown in this exercise is the level-2 envelope of the time series $U_1$.

(b) Show that the envelopes are invariant to scaling of the price or time axes. This is a useful property of this approach, i.e., it does not depend on any absolute thresholds or scales.

Quantitative Indicators

Moving Averages. A moving average is an average (over some probability distribution) of the previous prices of the stock. The resolution of the moving average is related to the time interval separating the points at which the stock price is sampled (the smaller the time interval, the higher the resolution). To be more specific, let $\tau$ be the time interval, and let $p_0, p_1, p_2, \ldots$ be a probability distribution over $0, 1, 2, \ldots$. Then a moving average has the form

$$MV(t) = \sum_{i=0}^{\infty} p_i S(t - i\tau),$$

$$= p_0 S(t) + p_1 S(t - \tau) + p_2 S(t - 2\tau) + p_2 S(t - 3\tau) + \cdots.$$

Note that this sum typically gets truncated because the time series of stock prices does not go back to $-\infty$. Some typical choices for the probability distributions are

Uniform Fixed Window.

$$p_i = \frac{1}{T+1}.$$
for $i = 1 \ldots T$. This choice for $p_i$ corresponds to a arithmetic average over the previous $T + 1$ time steps,

$$MV(t) = \frac{1}{T+1} \sum_{i=0}^{T} S(t - i\tau).$$

Typically, one may use such moving averages with different choices of the window size $T$. For example, if $\tau = 1$ trading day, then some useful choices of $T$ are $T = 10$ trading days (2 weeks), $T = 60$ trading days (3 months), $T = 120$ trading days (6 months). A good indication of the market trend might be obtained by looking at (for example) when the 10 day moving average crosses the 60 day moving average to indicate a change in the trend.

**Exercise 1.24**

You are given a time series for the stock price, $S_1, S_2, \ldots, S_N$. Assume that $\tau = 1$ and that the window size is $T$.

(a) For what values of $t$ is $MV(t)$ defined?

(b) Give a linear time algorithm (linear in $N$ and independent of $T$) to compute the time series $MV(t)$ for all values of $t$ on which it is defined.

Exponentially Weighted Moving Average.

$$p_i = (1 - e^{-\lambda})e^{-\lambda i}.$$  

This form of a moving average gives more weight to recent values of the stock price, and may be useful when the market dynamics is changing often. $\lambda$ is often called the exponential decay factor in the weighting. One can roughly equate $\frac{1}{\lambda}$ to an effective fixed window width, since once $i$ becomes greater than $\frac{1}{\lambda}$, the probabilities drop very sharply.

**Exercise 1.25**

You are given a time series for the stock price, $S_1, S_2, \ldots, S_N$. Assume that $\tau = 1$ and that the exponential decay factor is $\lambda$. You may assume that $S_i = 0$ for $i < 1$.

(a) Give an algebraic expression (a summation formula) for $MV(t)$.

(b) For what values of $t$ is $MV(t)$ defined?

(c) Give a linear time algorithm (linear in $N$) to compute the time series $MV(t)$ for all values of $t$ on which it is defined.

(Hint: First show that $MV(t + 1) = e^{-\lambda}MV(t) + (1 - e^{-\lambda})S_{t+1}$.)

(d) How would your answers above change if you did not assume that $S_i = 0$ for $i < 1$, i.e., you computed the weighted sum only back till $i = 1$.

(Hint: you need to renormalize the $p_i$’s at each time. Try to do this efficiently to still maintain a linear time algorithm. One approach is to define $A(t) = \sum_{i=1}^{t} e^{-\lambda(t-i)}S_i$, and $B(t) = \sum_{i=1}^{t} e^{-\lambda(t-i)}$. Now relate $MV(t)$ to $A(t)$ and $B(t)$, and show how to efficiently update $A(t)$ and $B(t)$.)

Moving averages are methods for attempting to “smooth” out the noise (random fluctuations), and extracting the true trend. Many people argue that there is predictive power in moving averages, however, our position here is a more neutral one of simply providing the reader with a possible indicator for predictive use.

The Relative Strength Index. The relative strength index $R_N$ is the percentage of up moves over the last $N$ moves. The relative strength index also gives some indication of the trend. When $R \geq 70\%$ it seems to indicate that the market has been in an up trend. How to use this predictively may of course depend on the market. Will the market continue to trend up or revert back down? Similarly, when $R \leq 30\%$ it seems to indicate that the market has been in a down trend.
Oscillators. An oscillator is a form of indicator that is commonly known as a stochastic. We describe a very simple form of oscillator, $\kappa_N$ which is computed over the previous $N$ time periods,

$$\kappa_N = \frac{\text{Current Price} - \text{Low}}{\text{High} - \text{Low}}.$$ 

The high and the low above are computed over the previous $N$ periods.

**Exercise 1.26**

Assume that $N$ is given. Give linear time algorithms to compute the relative strength index and the oscillator as functions of $t$ for a given stock price time series.

Specific Patterns.

Over time, qualitative traders have recognized the presence of certain patterns. Having established the existence of these patterns, one might argue that if the beginning of such a pattern is observed, one should be able to predict the behavior by using the remaining part of the pattern as a guide.

Some common patterns are: the **head and shoulders pattern**, and its reverse;

the **saucer top and bottom**;

etc.

**Exercise 1.27**

Given a database of patterns and a a stock price time series, develop efficient algorithms to extract the set of patterns from the database that are $K_1, K_2$-active at time $t$.

We define the active patterns as follows. View each pattern $P$ as a string $P = p_1 p_2, \ldots, p_{|P|}$. Also view the stock price time series up to time $t$, $S(t) = S_0 S_1 \ldots S_t$, as a string. A pattern $P$ is $K_1, K_2$-active at time $t$ if some prefix of length $\ell$, $K_1 \leq \ell \leq K_2$, of the corresponding pattern string $P$ matches the corresponding suffix of length $\ell$ of the string $S(t)$.

Define the database size $D$ as the total length of all the patterns in the database. Let $M_t$ be the total length of the patterns that are $K_1, K_2$-active at time $t$, and let $M = \sum_t M_t$ be the total length of the active patterns for this stock time series ($M$ is the size of the output).

Suppose that you are allowed to preprocess the database of patterns, and that the stock price time series and the value of $K_1, K_2$ are inputs to the algorithm. With $O(D)$ preprocessing, your algorithm should run in $O((K_2 - K_1 + 1)NK + M)$.

(Hint: You may want to preprocess your database into a suffix tree.)

A typical application of such algorithms would be to define the stock price series as a string over a three letter alphabet: down (-1), no significant move (0), up (+1). The pattern strings are also similarly defined. One then takes all the $K$-active patterns at a time $t$. Those patterns whose matching prefix is a proper prefix give a prediction of the future. These patterns can be “voted together” somehow to give a prediction of the future prices of the stock. If, for example, the prediction is sufficiently positive, one might then consider a buy trade.

[Extension, Open Ended] Suppose now that exact matches are not required, but certain transformations of a pattern, including noise in the data (approximate matching) should be allowed for. Formalize this problem and try to develop efficient algorithms or heuristics to solve the problem.
The Japanese Candlestick. An interesting pictorial representation of a short period of the stock price series is given by the Japanese candlestick. There are two kinds of candlestick, the solid and the clear.

Typically the period represented by the candlestick may be one day. The clear candlestick indicates that the close \((C)\) was above the open \((O)\). The height of the bar indicates the magnitude of the difference \(|C - O|\). The high \((H)\) is represented by upper tip of the top line and the low \((L)\) by the lower end of the bottom line. All height differences are proportional to their true values. An example of a market behavior and the corresponding candlestick are given below.

Note that a given market behavior can easily be converted to a candlestick, but the candlestick cannot uniquely be converted into a market behavior. However it gives a good idea, for example if the close is below the open, it probably means that the low is after the high, as this picture would result in smaller overall price movements.

**Exercise 1.28**
Give the market behaviors that would correspond to the following candlestick sequences,

Recap. The important point we wish to emphasize is the we have presented a set of indicators which are very good at giving us a picture of how the market has been behaving in the past (for example has it been recently trending up, down or neither). The real goal, however would be to predict how the market will behave in the future,

Prediction, however is not the content of our discussion. You may, of course at your own risk, take for granted that the market will behave in the future, as it did in the past. On a lucky day you may make some money. Some believe that it should be possible to predict based on such indicators. Some believe you are better off trying to predict the prices tomorrow based on the length of mini-skirts being sported by women today – these
people essentially believe that the only way to go is to do due diligence research on a company, its management, its products, its competition, etc, and obtain a better estimate of its EPS than the market has. We wash ourselves of these tasks of prediction because there are plenty of other computational tasks that are eagerly awaiting our solution.

**Data Snooping (Overfitting)** Before we go on, we will address one important issue, namely that of data snooping. Specifically, how to evaluate a predictor.

Suppose that John TinyTrader comes to you with a predictor and shows you that over the last 10 months, it predicted the market movement (up or down) correctly. This person then offers you this system for a price. Do you want it? The problem you face is one of trying to evaluate a predictor without knowing the process that created the predictor. The warning we give to you is that this is an extremely dangerous task. Your evaluation of the predictor depends heavily on the process that created the predictor. In the absence of any knowledge regarding the process that created the predictor, a prudent way to proceed would be to assume the worst. To illustrate the dilemma, consider the following three distinct processes that could have resulted in John TinyTrader approaching you:

1. John TinyTrader is the only trader in the market. One day he dreams up a system, and immediately gives you the system. You test it on the next 10 months. It works! Mr. TinyTrader now approaches you, and asks how well did my system do, and are you willing to pay for it?

2. There are a large number of TinyTraders on the market, John, Jack, Ann, . . . . For concreteness suppose that there are about 1024 such TinyTraders, and they all have a different set of predictions for the next 10 months (if any two have the same sets of predictions, we can ignore one of them). Clearly no one will approach you if they made mistakes. John happens to be the TinyTrader who gets it all right. Mr. TinyTrader now approaches you.

3. John TinyTrader is the only trader in the market. One day he dreams up a system and tests it on the next 10 months. It failed to predict the 8th month correctly. John continues testing his system, and lo and behold, it correctly predicts months 9-18 correctly. Mr. TinyTrader now approaches you with his data on the most recent ten months.

Under which process of creating the trading system would you pay the most. If your intuition said system 1 appears worth the most, followed by 3, then 2, then your intuition is not bad. If we rephrase each of these creation processes, perhaps the choice will become clearer,

1. I have a system which I use in the market for the next 10 months. It makes money, this is real money!

2. I have 1024 different systems, I need to pick one. So I look at the performance over 10 months, and luckily for me, one of them did perfectly. Was I really lucky? No! One system is guaranteed to perform well, as there are only 1024 possible outcomes over the next 10 months, and so one of my systems must be perfect. What can I conclude about this system? not much.

3. I have a system. I will continue testing until it performs perfectly for 10 consecutive months. Then I will approach you with my system. You should convince your self that a random system will eventually perform perfectly for 10 consecutive months. Don’t be fooled by the argument “. . .but over 19 months, the system predicted 18 months correct.”

It just got lucky. The fact is this system is guaranteed to be presented to you eventually, and it could be no better than random.

The underlying difference between system 1 and the other two is the notion of choice. In a (not so vague way) systems 2 and 3 are the result of someone being able to choose a system from among a set to present. Naturally TinyTrader will select the best. Such issues of how the availability of choices affects your belief in the quality of the system is a topic beyond the scope of our discussion, and falls in the arena of Machine Learning, specifically the sub field of **generalization** – how does the performance of the system you have seen so far generalize to the
future. The generalization ability is intimately related to the degree of choice one has in choosing which system to show.

To drive the point home, consider a uniform random variable over the interval $[0, 1]$. If you randomly pick one and show it to me, it will have an expected value of $\frac{1}{2}$ which one might argue is a fair representation of the distribution. If, however, one picks 10 such random numbers independently from the same distribution, and then chooses to show me the largest of these, this is no longer a fair representation of the distribution. In fact the number you show me will on the average be much large than $\frac{1}{2}$. The next exercise will convince you of this fact.

**Exercise 1.29**
Show that the expected value of the maximum of $n$ independent samples of a uniform random variable over $[0, 1]$ is $\frac{n}{n+1}$.

**Exercise 1.30**
[The Postal Scam.] You are a gambling man and every week you place a small bet on the Monday night (American) football game. One Friday you got a postcard with only the following on it

Ravens  Cardinals

You watch the game on Monday and the Cardinals win. For the next 4 consecutive Fridays, you get similar postcards with the following:

Bills  Falcons
Panthers  Bengals
Bears  Browns
Broncos  Cowboys

On each corresponding Monday, you watch the game and to your surprise, the sequence of winners is Falcons, Panthers, Bears and Cowboys.

You eagerly await the next postcard on the next Friday, which to your surprise arrives, with the following on it:

Colts  Packers

You are confused; you turn over the postcard and find a message. Call 1-900-XXX-XXXX to find out which team is underlined. When you call, they ask you for $10 for the answer. You need to make a quick decision. Do you bite (assuming that you can place any sized bet on the Colts-Packers game in Las Vegas).

More specifically, what is the maximum you would be willing to pay for this information. You may make some reasonable assumptions on the price of a postcard stamp.

Does your answer to the above problem change if there is a bet-limit of $100 that you can make on the game in Las Vegas?

The moral is that while we do not engage in the task of prediction, we warn that to evaluate a predictor carefully, it is imperative that one knows the process that created the predictor. Trying only to evaluate it based on the resulting performance is flawed. In a conservative world, one should assume the worst for the process that created the predictor.
1.4 Problems

Problem 1.1. In reality, the situation in Exercise 1.5 is a simplification. When a bank quotes you an interest rate of 5% per year, there is also a compounding period (1 day or 1 month). If there are \( n \) compounding periods in a year, then an interest of \( 5%/n \) is applied for each period. A typical choice is \( n = 12 \), which corresponds to monthly compounding.

(a) Show that 1 dollar grows to \((1 + \frac{0.05}{n})^n\) dollars over the year.
(b) When \( n = 12 \), what is the equivalent yearly interest rate? (This is called the annualized percentage rate, APR)
(c) As a consumer, do you prefer \( n \) to be large or small?
(d) When \( n \to \infty \), the compounding is over smaller and smaller periods, but there are more of them. What does $1 grow to after a year and what is the equivalent yearly interest rate. \[ \text{[Hint: } (1 + a/n)^n \to e^a \text{]} \]

(When a bank uses more compounding periods the interest rate being quoted is closer to the rate that should be used in the exponential growth formula. It is better for the consumer if there are more compounding periods, especially when the interest rate is large.)

Problem 1.2. [The Mortgage Calculation]

(a) A mortgage with principle \( P \) and \( N \) payments of size \( X \) at the times \( t_1, \ldots, t_N \) is fair if you are indifferent about being the lender or borrower in the mortgage. For a given \( r, P, \{t_1, \ldots, t_N\} \), determine the fair payment value \( X \).

(Answer: \( X = P/\sum_{i=1}^{N} e^{-r t_i} \).)

Specialize your answer to the case where \( t_i = 1, \ldots, N \) (i.e., \( t_i = i \)), and obtain an answer that does not have a summation – i.e., get a closed form.

(b) Suppose that you can only afford to pay $\text{500}$ at the times \( 1, 2, 3, \ldots, 360 \) (i.e., \( t_i = i \)). Compute the maximum mortgage that you can afford to take out. Assume that \( e^r = 1.005 \). (This roughly corresponds to a 30 year mortgage with a monthly payment of $500 and a yearly interest rate of about 6%).

Suppose that you need a mortgage amount of $\text{100,000}$. For how many months will you be paying off the mortgage.

Problem 1.3. Show that the special case of \( D(t) = e^{rt} \) which we derived in the stationary setting gives a constant (instantaneous) forward interest rate \( r \).

Show that analogous to the formula \( D(t) = e^{rt} \) (in the stationary case), in the non-stationary case, one has that

\[ D(t) = e^{\int_0^t ds \ r(s)} \]

Problem 1.4. We give an alternate derivation of the formulas \( D(t) = e^{rt} \) \( (D(t) = e^{\int_0^t ds \ r(s)} \) in the non-stationary case.) by taking the continuous limit of compounding.

Divide the time period \([0, t]\) into \( n \) time periods each of length \( \Delta t = t/n \). Let \( r_i(\Delta t) \) be the interest rate function over the \( i \)th period. Assume that you invest \( X_0 \) in a bank at time 0 and let \( X_t \) be the amount accrued at time \( t \).

Show that

\[ X_t = X_0 \cdot \exp \left( \sum_{i=1}^{n} \log(1 + r_i(\Delta t)) \right) \]

We can assume that \( r_i(\Delta t) \) must go to zero as \( \Delta t \) goes to zero, as one should not be able to gain any interest over a period of time tending to zero. Hence, using the Taylor expansion for the logarithm, show that

\[ X_t = X_0 \cdot \exp \left( \lim_{n \to \infty} \sum_{i=1}^{n} r_i(\Delta t) \right) \]

The exponent has exactly the form of an integral, providing \( r_i(\Delta t) \) satisfies some regularity conditions. In particular if \( r_i(\Delta t) = r(\frac{1}{n} \Delta t) \Delta t \) then we recover the Riemann limit for the integral form we expect.

Problem 1.5. In Section 1.1.3 we assumed a single interest rate at which you can borrow and lend. Suppose you can lend at interest rate \( r_x \) and borrow at interest rate \( r_b \), with \( r_x < r_b \). The arbitrage argument is not as simple. Now you can only deduce a range for the arbitrage price.

Consider the deal \((X_1, X_2; t_1, t_2)\) whose no-arbitrage price is \( P \). Assume \( X_1 < 0 \) and \( X_2 > 0 \). So if you own the deal, at \( t_1 \) you pay \( X_1 \) and at \( t_2 \) you get \( X_2 \).
1. Financial Instruments

1.4. Problems

(a) Show that if \( P < X_1e^{-r_1t_1} + X_2e^{-r_2t_2} \), you can buy the deal and construct an arbitrage opportunity by borrowing money from the bank.

(b) Show that if \( P > X_1e^{-r_1t_1} + X_2e^{-r_2t_2} \), you can sell the deal and construct an arbitrage opportunity by lending money to the bank.

(c) What can you say about the no-arbitrage price \( P \)?

(d) Suppose now you reverse the payment and the receipt of cashflow. So, \( X_1 > 0 \) and \( X_2 < 0 \). Now derive upper and lower bounds on the no-arbitrage price \( P \). \([\text{Hint: It is more complicated than the previous case.}]\)

Problem 1.6. The Interest Rate Swap. Agreeing that the markets are not stationary, the \([t_1, t_2]\) period interest rate \( R(t_1, t_2) \) for a fixed period in the future may change with time.

An interest rate swap is the following somewhat bizarre contract.

A set of times \( t_1, \ldots, t_N \) are specified. The fixed payment or par value is \( X \). The contract is between two parties. The receiver receives the fixed payment \( X \) at the times \( t_1, \ldots, t_N \) from the payer. In return, the receiver pays the payer an amount \( R(t_{i-1}, t_i; t_{i-1}) \) at time \( t_i \), for \( i = 1, \ldots, N \). \( R(t_{i-1}, t_i; t_{i-1}) \) is the \([t_{i-1}, t_i]\) interest rate determined at time \( t_{i-1} \). For convenience set \( t_0 = 0 \). The third argument of \( R(\cdot) \) indicates that the value of the interest rate for the period \([t_{i-1}, t_i] \) is determined at the time \( t_{i-1} \). We will use the notation \( R(t_{i-1}) \) to denote the payment \( R(t_{i-1}, t_i; t_{i-1}) \).

The reason for calling it a swap should be clear – one is swapping a fixed payment \( X \) for a variable payment whose exact value will depend on a particular periods interest rate when that period arrives. So, right now, this interest rate is not known. Hence, from the present point of view, the variable payments are non-deterministic. This is a fairly complicated contract and it may be worth spending some time to understand it, before proceeding. In particular, it can be represented by the following picture.

![Interest Rate Swap Diagram](image)

The blue arrows indicate received payments, and the red arrows indicate payments out. The complication in this contract is that the payments out are not known at time 0. For example the payment \( R_{12} \) is only known at time \( t_1 \). Nevertheless, in this problem, we will investigate how to price this contract (i.e., determine the equilibrium par value for \( X \)) using an arbitrage argument. The \( R(t_{i-1}) \) are typically called the floating payments, since they are not determined a priori, and \( X \) is typically called the fixed payment. Notice that the payments occur at time \( t_i \), and that the exact value of the payments at time \( t_i \) are only known at \( t_{i-1} \), which is when the floating payment gets determined.

(a) Consider one of the floating payment \( R(t_{i-1}) \). Construct the following "portfolio" which buys a zero coupon bond with maturity \( t_{i-1} \) and sells one with maturity \( t_i \).

(i) Show that the cash flow at time 0 is \( B(t_i) - B(t_{i-1}) \).

(ii) At time \( t_{i-1} \), show that the value of this portfolio is \( 1 - B(t_i - t_{i-1}; t_{i-1}) \), where the zero coupon value is the value determined at time \( t_{i-1} \).

(iii) Use the relationship between \( B(t) \) and \( R(t) \) to conclude that the value of this portfolio at time \( t_{i-1} \) is

\[
\frac{R(t_{i-1})}{1 + R(t_{i-1})}. 
\]

Note that this value is the value at time \( t_{i-1} \), which is not known at time 0.
(iv) Now consider the floating payment \( R_{(i-1)i} \) which is made at time \( t_i \). This payment is determined precisely at time \( t_{i-1} \). Argue (for example using an arbitrage argument) that the value of this cash flow at time \( t_{i-1} \) is

\[
\frac{R_{(i-1)i}}{1 + R_{(i-1)i}}.
\]

(Remember that this cash flow only occurs at time \( t_i \)).

(v) Use an arbitrage argument to conclude that the value of the floating payment \( R_{(i-1)i} \) today must be equal to \( B(t_i) - B(t_{i-1}) \).

(vi) Use an arbitrage argument to conclude that the total value today of all the floating payments is

\[
1 - B(t_N).
\]

(vii) Using the result in the previous exercise, show that the total value of all the fixed payments is

\[
X \sum_{i=1}^{N} B(t_i).
\]

(viii) Use an arbitrage argument to show that the value of \( X \) must be

\[
\bar{X} = \frac{1 - B(t_N)}{\sum_{i=1}^{N} B(t_i)}
\]

if there is to be no arbitrage opportunity. \( \bar{X} \) is usually referred to as the *equilibrium par swap rate*. It is the rate at which one is indifferent between taking on the fixed or floating end of the contract.

(b) Typically in the market place, there are published par swap rates and zero coupon rates, and there may be a difference (usually denoted the spread) between the par rate and the formula above that gives the par rate in terms of the zero coupon rates. What are possible explanations for this?

We obtained the “price” of a fairly complicated bond portfolio involving cash flows that were not determined at time 0 using a fairly sophisticated arbitrage argument. One might wonder how one comes up with such arbitrage arguments, it almost looks like magic. Later we will set up a systematic framework using risk neutral (martingale) measures for addressing such topics.