FINANCIAL MATHEMATICS

I-Liang Chern

Department of Mathematics

National Taiwan University
and
Chinese University of Hong Kong

December 14, 2016

Contents

Chapter 1

Introduction

1.1 Assets

Assets represent value of ownership that can be converted into cash. There are two kinds of assets: tangible and intangible. Commodity, foreign currency, house, building, equipments are tangible, while copyright, trademarks, patterns, computer programs and financial assets are intangible.

The financial assets include bank deposits, debt instrument, stocks and derivatives. Debt instruments are issued by anyone who borrows money – firms, governments, and households. They include corporate bonds, government bonds, residential, commercial mortgages and consumer loans. These debt instruments are also called fixed-income instruments because they promise to pay fixed sums of cash in the future.

The stocks, shares and equities are all words used to describe what is essentially the same thing. They are the claim of the ownership of a firm. When someone buys a share, he is buying ownership of part of a company and becomes a shareholder in that company.

The prices of assets are settled through trading. Many assets are traded in markets. There are commodity markets as well as financial markets. Financial markets include bond markets, stock markets, currency markets, financial derivatives markets, etc. The financial market provides a link between saving and investment. Savers can earn high returns from their saving and borrowers can execute their investment plans to earn future profits.

The derivatives are contracts derived from some underlying assets. Usually, the prices of assets fluctuate time by time. In order to reduce the risk of price fluctuation, a corresponding contract is introduced to make such uncertainty more certain. For instance, suppose today's price of corn is US$3 per bushel. You want to buy 5,000 bushels of corn for delivery two months later. You can sign an agreement with someone who is willing to sell you this amount at such price at such future time. This agreement is called a forward contract. It certainly costs you some money but make your price uncertainty more certain. We call that such forward contact hedges the risk of price fluctuation. Forward contract is one kind of financial derivative. The corn is the underlying asset. More financial derivatives will be introduced later.

1.2 Financial Derivatives

1.2.1 Forwards contract

A forward contract is an agreement which allows the holder of the contract to buy or sell a certain asset at or by a certain day at a certain price. Here,

- the certain day—maturity or expiration date,
- the certain price—delivery price,
- the person who write the contract (has the asset) is called in short position,
- the person who holds the contract is called in long position.

Example 1. Suppose today's price of corn is US$3 per bushel. You want to buy 5,000 bushels of corn for delivery two months later. You can sign an agreement with someone who is willing to sell you this amount at price, say US$3.05, at such future time. Then you make your uncertain risk of price fluctuation more certain.

Example 2. (quoted from Chan) Suppose it is May 30 now and you need to pay your tuition to Oxford University £1,000 pound by September 1. Suppose you can earn HK$15,600 in the summer. The exchange rate now is £1.00 = HK$15.40. What should you do?

- 1. Try your luck: If the exchange rate is less than HK$15.60 by August 31, you earn some extra cash besides the registration fee. If the exchange rate is higher than HK$15.60, you ... ?
- 2. Buy a forward on British pounds: Look for a forward contract on British pounds that entitles you to use HK$15,600 to buy £1,000 on August 31. Then you have eliminated, or hedged, your risk. Thus, a forward contracts eliminate your risk.

1.2.2 Futures (futures contracts)

A futures contract is very similar to a forward contract. Futures contracts are usually traded through an exchange, or clearinghouse, which standardizes the terms of the contracts. The exchange helps to eliminate the risk of default of either party through a margin account. Each party has to pay an initial margin as deposit at the inception of the contract. The profit or loss from the futures position is calculated every day and the change in this value is paid from one party to the other. The maintenance margin is the minimum level below which the investor is required to deposit additional margin.

The difference between forward contract and futures contract are

	Forward Contract	Future Contract
nature	customized contract	standarized contract
Trading	over the counter	through exchanges
liquidity	less liquid	highly liquid
counter party risk	high	negligible
Settlement	delivery (at the end)	closed out prior to maturity
Margin	no margin	compulsorily needs to be paid by the parties

Example Mr. Chan takes a long (buy) position of one contract in corn (5,000 bushels) for March delivery at a price of US$3.682 per bushel at the Chicago Board of Trade (CBOT). See quotation at `http://www.cbot.com/.CBOT` It requires maintenance margin of US$700 with an initial margin markup of 135%, i.e. the initial margin is US$945 which Mr. Chan and the seller each has to deposit into the broker's account on the first day they enter the contract. The next day the price of this contract drops to US$3.652. This represents a loss of US$0.03 $\times$ 5,000 = US$150. The broker will take this amount from Mr. Chan's margin account and deposit it to the seller's margin account. It leaves Mr. Chan with a balance of US$795. The following day the price drops again to US$3.552. This represents an additional loss of US$500, which is again deducted from the margin account. As this point the margin account is US$295, which is below the maintenance level. The broker calls Mr. Chan and tells him that he must deposit at least US$405 in his margin account, or his position will be closed out, i.e. both sides agree to settle the contract at this point. Once the position is closed, Mr. Chan will not be able to earn back any money in the future even if the price rises above US$3.682.

1.2.3 Options

There are two kinds of options — call options and put options. A call (put) option is a contract between two parties, in which the holder has the **right** to buy (sell) and the writer has the **obligation** to sell (buy) an asset at certain time in the future at a certain price. The price is called the exercise price (or strike price). The holder is called in long position, while the writer is called in short position. The underlying assets of an option can be commodity, stocks, stock indices, foreign currencies, or future contracts.

There are two kinds of exercise features:

- European options : Options can only be exercised at the maturity date.
- American options : Options can be exercised any time up to the maturity date.

Notation

- t current time
- T maturity date
- S current asset price

- S_T asset price at time T
- E strike price
- c premium, the price of call option
- r bank interest rate

Examples An investor buys 100 European call options on XYZ stock with strike price \$140. Suppose

$$\begin{aligned} E &= 140, \\ S_t &= 138, \\ T &= 2 \text{ months}, \\ c &= 5 \text{ (the price of one call option)}. \end{aligned}$$

If at time T, $S_T > E$, then he should exercise this option. The payoff is $100 \times (S_T - E) = 100 \times (146 - 140) = 600$. The premium is $5 \times 100 = 500$. Hence, he earns \$100. If $S_T \leq T$, then he should not exercise his call contracts. The payoff is 0.

The payoff function for a call option is $\Lambda = max\{S_T - E, 0\}$. One needs to pay premium (c_t) to buy the options. Thus the net profit from buying this call is

$$\Lambda - c_t e^{r(T-t)}.$$

Example Suppose

$$\begin{aligned} \text{today is } t &= \text{03/09/2016}, \\ \text{expiration is } T &= \text{31/12/2016}, \\ \text{the strike price } E &= 250 \end{aligned}$$

for some stock. If $S_T = 270$ at expiration, which is bigger than the strike price, we should exercise this call option, then buy the share for 250, and sell it in the market immediately for 270. The payoff $\Lambda = 270 - 250 = 20$. If $S_T = 230$, we should give up our option, and the payoff is 0. Suppose the share take 230 or 270 with equal probability. Then the expected profit is

$$\frac{1}{2} \times 0 + \frac{1}{2} \times 20 = 10.$$

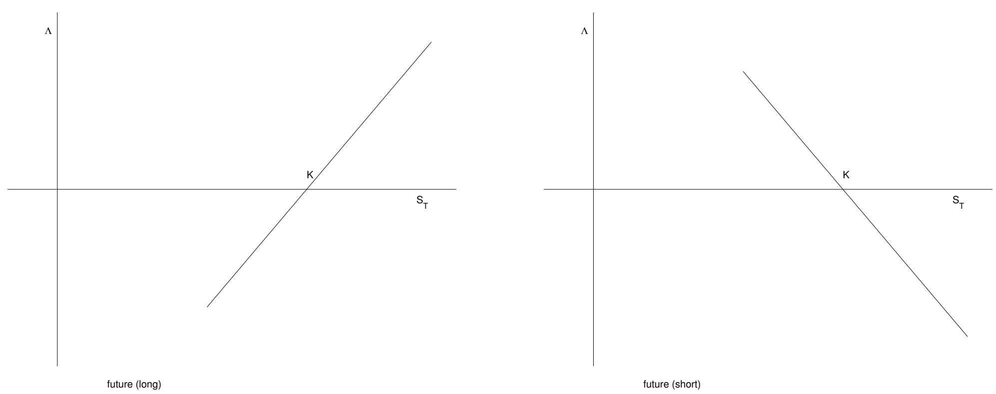
Ignoring the interest of bank, then a reasonable price for this call option should be 10. If $S_T = 270$, then the net profit$= 20 - 10 = 10$. This means that the profits is 100% (He paid 10 for the option). If $S_T = 230$ the loss is 10 for the premium. The loss is also 100%. On the other hand, if the investor had instead purchased the share for 250 at t, then the corresponding profit or loss at T is ± 20. Which is only $\pm 8\%$ of the original investment. Thus, option is of high risk and with high return.

1.3 Payoff functions

At the expiration day, the payoff of a future or an option is the follows.

Payoff of futures

- Payoff of a future in long position: At the expiration day, the price of the asset is S_T. The holder can buy the asset at price E. He has the obligation to buy it. Thus the payoff of the holder is $S_T - E$.
- Payoff of a future in short position: at expiration, the writer also has the obligation to sell the asset at price E. So, he needs to use S_T to buy the asset, then sell to the holder at price E. Thus his payoff is $E - S_T$.

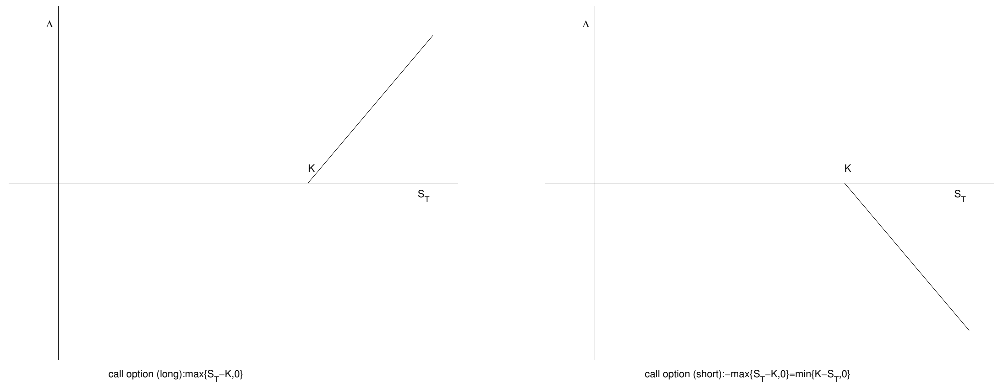


(a) left (b) right

Figure 1.1: Payoff of a future, long position (left) and short position (right)

Payoff of call options

- Long call: $\Lambda = \max\{S_T - E, 0\}$
  - If $S_T > E$, then the holder exercise the call at price E then sell at price S_T
  - If $S_T \leq E$, then he gives up the call option.
- Short call: $\Lambda = \min\{E - S_T, 0\}$
  - If $S_T > E$, he has the obligation to sell the asset at price E, thus he needs to buy the asset at price S_T then sell it at price E. He losses $S_T - E$.
  - If $S_T \leq E$, the holder will give up the call option. So the writer losses nothing.



(a) left (b) right

Figure 1.2: Payoff of a call, long position (left) and short position (right)

Payoff of put options

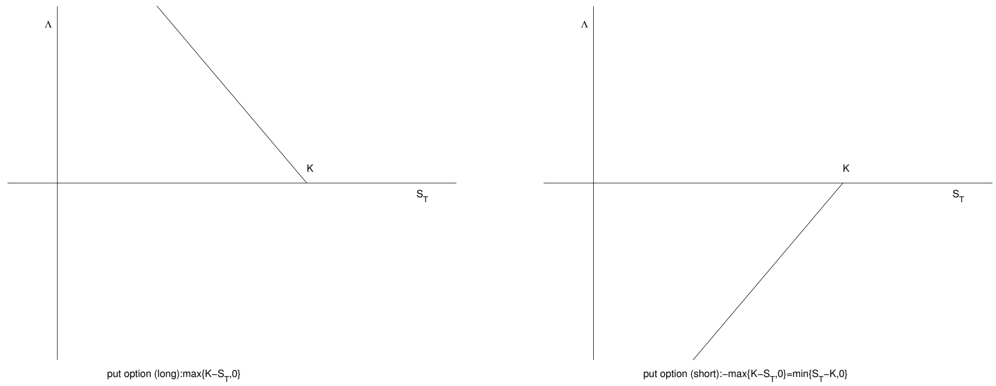
- Long put: $\Lambda = \max\{E - S_T, 0\}$
  - If $S_T < E$, the holder has the **right** to sell the asset at E, then buy it back at S_T. Thus the payoff is $E - S_T$.
  - If $S_T \geq E$, the holder just gives up the option.
- Short put: $\Lambda = \min\{S_T - E, 0\}$
  - If $S_T < E$, he has the **obligation** to buy the asset at E. He then sell it at S_T. Thus he losses $E - S_T$.
  - If $S_T \geq E$, the holder will give up the put option, So the writer losses nothing.

1.3.1 Premium of an option

The person who hold an option has the **right** and the counter party has the **obligation** to fulfill the contract. Thus, a premium should be paid by the holder to the writer. Below is a portion of a call option copied from the Financial Times.

$$\begin{aligned} \text{the current time } t &= \text{Feb 3} \\ \text{the expiration } T &= \text{end of Feb,} \\ T - t &\approx \text{10 days} \\ S_t &= 2872 \end{aligned}$$

E	2650	2700	2750	2800	2850	2900	2950	3000
c	233	183	135	89	50	24	9	3



(a) left (b) right

Figure 1.3: Payoff of a put, long position (left) and short position (right)

1.4 Other kinds of options

- **Barrier option**: The option only exists when the underlying asset price is in some prescribed value before expiry.
- **Asian option**: It is a contract giving the holder the right to buy or sell an asset for its average price over some prescribed period.
- **Look-back option**: The payoff depends not only on the asset price at expiry but also its maximum or minimum over some period price to expiry. For example, $\Lambda = \max\{J - S_0, 0\}$, $J = \max_{0 \le \tau \le T} S(\tau)$.

1.5 Types of traders

1. Speculators (high risk, high rewards)
2. Hedgers (to make the outcomes more certain)
3. Arbitrageurs (Working on more than one markets, p12, p13, p14, Hull).

1.6 Basic assumption

Arbitrage opportunities cannot last for long. Only small arbitrage opportunities are observed in financial markets. Our arguments concerning future prices and option prices will be based on the assumption that "there is no arbitrage opportunities".

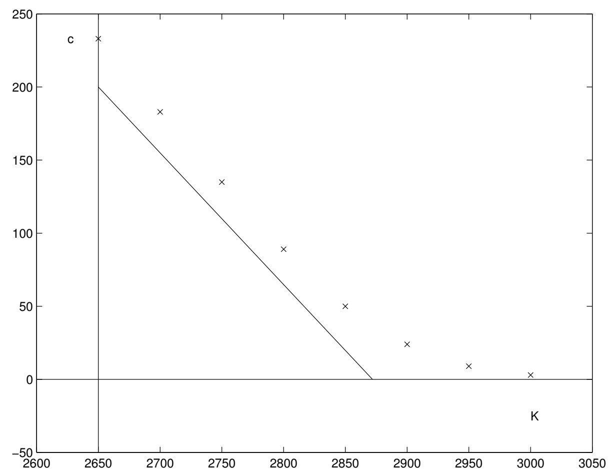


Figure 1.4: The FT-SE index call option values versus exercise price.

Chapter 2

Asset Price Model

2.1 Efficient market hypothesis

The asset prices move randomly because of the following efficient market hypothesis:

1. The past history is fully reflected in the present price, which does not hold any future information. This means the future price of the asset only depends on its current value and does not depends on its value one month ago, or one year ago. If this were not true, technical analysis could make above-average return by interpreting chart of the past history of the asset price. This contradicts to the hypothesis of no arbitrage opportunities. In fact, there is very little evidence that they are able to do so.

2. Market responds immediately to any new information about an asset.

2.2 The asset price model

We shall introduce a discrete model and a continuous model. We will show that the continuous model is the continuous limit of the discrete model.

2.2.1 The discrete asset price model

The time is discrete in this model. The time sequence is $n\Delta t$, $n \in \mathbb{N}$. Let us denote the asset price at time step n by S_n. We model the asset price by

$$\frac{S_{n+1}}{S_n} = \begin{cases} u & \text{with probability } p \\ d & \text{with probability } 1-p. \end{cases} \tag{2.1}$$

Here, $0 < d < 1 < u$. The information we are looking for is the following transition probability $P(S_n = S|S_0)$, the probability that the asset price is S at time step n with initial price S_0. We shall find this transition probability later.

2.2.2 The continuous asset price model

Let us denote the asset price at time t by $S(t)$. The meaningful quantity for the change of an asset price is its relative change

$$\frac{dS}{S},$$

which is called *the return.* The change $\frac{dS}{S}$ can be decomposed into two parts: one is deterministic, the other is random.

- Deterministic part: This can be modeled by

$$\frac{dS}{S} = \mu dt.$$

  Here, μ is a measure of the growth rate of the asset. We may think μ is a constant during the life of an option.

- Random part: this part is a random change in response to external effects, such as unexpected news. It is modeled by a Brownian motion

$$\sigma dz,$$

  the σ is the order of fluctuations or the variance of the return and is called *the volatility.* The quantity dz is sampled from a normal distribution which we shall discuss below.

The overall asset price model is then given by

$$\frac{dS}{S} = \mu dt + \sigma dz. \tag{2.2}$$

We shall look for the transition probability density function $\mathcal{P}(S(t) = S|S(0) = S_0)$. Or equivalently, the integral

$$\int_a^b \mathcal{P}(S(t) = S|S(0) = S_0)\, dS$$

is the probability that the asset price $S(t)$ lies in (a, b) at time t and is S_0 initially.

2.3 The solution of the discrete asset price model

Let us consider the discrete price model

$$\frac{S_{n+1}}{S_n} = \begin{cases} u & \text{with probability } 1/2 \\ d & \text{with probability } 1/2. \end{cases}$$

A sequence of movements $(S_0, S_1, ..., S_n)$ is called an n-step path. In such a path, it can consist of ℓ times up movements and $n - \ell$ times down movements, where $0 \le \ell \le n$. The corresponding values of S_n are $S_0 u^\ell d^{n-\ell}$. Since the probability of each movement is independent, the probability of an n-step path with ℓ up movements (and $n-\ell$ down movements) is

$$\left(\frac{1}{2}\right)^\ell \left(\frac{1}{2}\right)^{n-\ell} = \left(\frac{1}{2}\right)^n.$$

There are $\binom{n}{\ell}$ paths with ℓ up movements in n-step paths, we then obtain the transition probability of the asset price to be:

$$P(S_n = S|S_0) = \begin{cases} \binom{n}{\ell}\left(\frac{1}{2}\right)^n & \text{when } S = S_0 u^\ell d^{n-\ell}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

Such S_n is called a (discrete) random walk. It is hard to deal with the values $S_0 u^\ell d^{n-\ell}$. Instead, we take its logarithmic function. Let us denote $\ln S_n$ by X_n. This random walk X_n obeyes the rule

$$X_{n+1} - X_n = \begin{cases} \ln u & \text{with probability } 1/2 \\ \ln d & \text{with probability } 1/2. \end{cases}$$

Let us define

$$\Delta x = \frac{\ln u - \ln d}{2}, \quad \ln(1+r) = \frac{\ln u + \ln d}{2}.$$

Then the rule can be written as

$$X_{n+1} - X_n = \ln(1+r) + \begin{cases} \Delta x & \text{with probability } 1/2 \\ -\Delta x & \text{with probability } 1/2. \end{cases}$$

The term $\ln(1+r) \approx r$ is called a drift term. It measures the growth of S_n. It can be absorpt into X_n. Indeed, if we define $Z_n = \ln((1+r)^{-n} S_n)$, then Z_n satisfies

$$Z_{n+1} - Z_n = \begin{cases} \Delta x & \text{with probability } 1/2 \\ -\Delta x & \text{with probability } 1/2. \end{cases} \tag{2.4}$$

The advantage of this formulation (using Z_n now) is that this random walk has equal increment Δx, the possible values of the random variable Z_n are $m\Delta x$, $m \in \mathbb{Z}$. This is the random walk with binomial distribution.

2.4 Binomial distribution

Consider a particle moves randomly on an one dimensional lattice $\{m\Delta x | m \in \mathbb{Z}\}$ according to the rule (2.4), where Z_n denotes the location of this particle at time step n. Suppose the particle is located at 0 initially. Our goal is to find the transition probability

$$P(Z_n = m\Delta x | Z_0 = 0)$$

explicitly and also find its properties.

This random walk is equivalent to the following binomial trials: flipping a coin n times. Suppose we have equal probability to get Head or Tail in each toss. We move the particle to the right or left adjacent grid point according to whether we get Head or Tail. Let X_k be the result of the kth experiment. Namely,

$$X_k = \begin{cases} 1 & \text{if we get Head} \\ -1 & \text{if we get tail} \end{cases}$$

Then $Z_n/\Delta x = \sum_{k=1}^n X_k$. The random walk Z_n is equivalent to the random walk Y_n, the number of Heads we get in n trials. Namely,

$$P(Z_n = m\Delta x | Z_0 = 0) = P(Y_n = \ell)$$

with

$$m = \ell - (n - \ell) = 2\ell - n, \text{ or equivalently } \ell = \frac{1}{2}(n + m).$$

Notice that m is even (odd), when n is even (odd). There is a one-to-one correspondence between

$$\{\ell \,|\, 0 \le \ell \le n\} \leftrightarrow \{m \,|\, -n \le m \le n, m + n \text{ is even}\}.$$

Thus, we can also use ℓ instead m to label our particle. There are $\binom{n}{\ell}$ paths to reach $m\Delta x$ from 0 with $\ell = (n + m)/2$, we get

$$P(Z_n = m\Delta x | Z_0 = 0\} = \begin{cases} 0, & \text{if } m + n \text{ is odd,} \\ \binom{n}{(n+m)/2}(\frac{1}{2})^n, & \text{if } m + n \text{ is even.} \end{cases}$$

Or equivalently,

$$P(Y_n = \ell) = \binom{n}{\ell}\left(\frac{1}{2}\right)^\ell \left(\frac{1}{2}\right)^{n-\ell}.$$

It is clear that

$$\sum_{\ell=0}^n P(Y_n = \ell) = \left(\frac{1}{2} + \frac{1}{2}\right)^n = 1.$$

This probability distribution is called the binomial distribution. We denote them by p_{Z_n} and p_{Y_n}, respectively.

Properties of the binomial distribution We shall compute the moments of p_{Z_n} or p_{Y_n}. They are defined by

$$< \ell^k >:= \sum_{\ell=0}^n \ell^k p_{Y_n}(\ell), \quad < m^k >:= \sum_{\substack{m=-n \\ m+n \text{ even}}}^n m^k p_{Z_n}(m)$$

Since $m = 2\ell - n$, we can find the moments $< m^k > = < (2\ell - n)^k >$ by computing $< \ell^k >$, which in turn can be computed through the help of the following moments generating function:

$$G(s) := \sum_\ell s^\ell p_{Y_n}(\ell) = \sum_\ell s^\ell \left(\frac{1}{2}\right)^n \binom{n}{\ell} = \left(\frac{1+s}{2}\right)^n.$$

To compute moments, for instance

$$G'(1) = \sum_{\ell=0}^{n} \ell \cdot p_{Y_n}(\ell) =< \ell > .$$

$$G''(1) = \sum_{\ell=1}^{n} \ell(\ell-1) p_{Y_n}(\ell) =< \ell^2 > - < \ell > .$$

Hence, we obtain the mean

$$< \ell >= G'(1) = \frac{n}{2}.$$

The second moment

$$< \ell^2 >= G''(1) + < \ell >= \frac{n(n+1)}{4}.$$

We can also compute the mean and variance of p_{Z_n} as the follows. From $m = 2\ell - n$, we have

$$< m >= 2 < \ell > -n = 0.$$

To compute the second moment $< m^2 >$, from $m = 2\ell - n$, we have

$$< m^2 >= 4 < \ell^2 > -4n < \ell > +n^2 = n.$$

The mean of this random walk is $< m >= 0$, while its variance is $< (m- < m >)^2 >= n$.

Exercise

1. Find the transition probability, mean and variance for the case

$$Z_{n+1} - Z_n = \begin{cases} \Delta x & \text{with probability } p \\ -\Delta x & \text{with probability } 1-p \end{cases}$$

2.5 Normal distribution

A random variable Z is said to be distributed as standard normal if its probability density is

$$p_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

We denote it as $Z \sim \mathcal{N}(0,1)$. A random variable X is said to be distributed as normal with mean μ and variance σ^2 if its probability density function is

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}.$$

We denote it by $X \sim \mathcal{N}(\mu, \sigma^2)$. For such X, it is clearly that

$$Z := \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1).$$

For

$$P(a \le \frac{X-\mu}{\sigma} < b) = P(\mu+\sigma a \le X < \mu+\sigma b) = \int_{\mu+\sigma a}^{\mu+\sigma b} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)}\, dx = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

The moments of Z (i.e. $\mathbb{E}[Z^m]$) can be computed through the helps of moment generating function

$$G(s) := \mathbb{E}[e^{sZ}] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\mathbb{R}} s^m x^m p_Z(x)\, dx = \sum_{m=0}^{\infty} \frac{s^m}{m!} \mathbb{E}[Z^m].$$

We have

$$G(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sx} e^{-x^2/2}\, dx = e^{s^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-s)^2/2}\, dx = e^{s^2/2} = \sum_{k=0}^{\infty} \frac{s^{2k}}{2^k k!}.$$

By comparing the two expansions above, we obtain

$$\mathbb{E}[Z^m] = \begin{cases} 0 & \text{if } m \text{ is odd} \\ \frac{(2k)!}{2^k k!} & \text{if } m = 2k. \end{cases}$$

2.6 The Brownian motion

2.6.1 The definition of a Brownian motion

The definition of the Brownian motion (or the standard Wiener process) $z(t)$ is the following:

Definition 2.1. *A time dependent function $z(t)$, $t \in \mathbb{R}$ is said to be a Brown motion if*

(a) $\forall t$, $z(t)$ is a random variable.

(b) $z(t)$ is continuous in t.

(c) For any $u > 0$, $s > 0$, the increment $z(t+s) - z(t)$, $z(t) - z(t-u)$ are independent.

(d) $\forall s > 0$, the increment $z_{t+s} - z_t$ is normally distributed with mean zero and variance s, i.e.

$$z_{t+s} - z_t \sim \mathcal{N}(0, s).$$

2.6.2 The Brownian motion as a limit of a random walk

We may realize the Brownian motion as the limit of the standard random walk. Let us study Brownian motion in $[0, t]$. First, we partition the time interval $[0, t]$ into n subintervals evenly. Let us imagine a particle moves randomly on lattice points $\{m\Delta x | m \in \mathbb{Z}\}$ at discrete time $k\Delta t$, $k = 0, ..., n$. We choose $\Delta x = \sqrt{\Delta t}$. The motion of the particle is:

$$Z_0 = 0,$$

$$Z_k = Z_{k-1} + X_k \Delta x$$

where

$$X_k = \begin{cases} 1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2 \end{cases}$$

The location Z_n at time step n obeys the binomial distribution $p_{Z_n}(m) = P(Z_n = m\Delta x | Z_0 = 0)$. We can connect discrete path $(Z_0, Z_1, ..., Z_n)$ by linear function and form a piecewise linear path $Z^n(s)$, $0 \le s \le t$ with $Z^n(k\Delta t) = Z_k$. As $n \to \infty$, we expect that these kinds of paths tend to a class of zig-zag paths $z(\cdot)$, called Brownian motions. They are time-dependent random variable. Their properties are:

- $z(0) = 0$;
- $z(\cdot)$ is continuous;
- Any non-overlapping increments $z(t_4) - z(t_3)$ and $z(t_2) - z(t_1)$ ($t_1 < t_2 < t_3 < t_4$) are independent.

Let us denote the collection of these "zig-zag" paths by Ω, the sample space of the Brownian motions.

Next, we find the probability distribution of Brownian motions. The probability mass function of the random walk is $P(Z_n = m\Delta x | Z_0 = 0\}$. Since $n + m$ is even, the corresponding probability density function on the real line is $P_m^n := P(Z_n = m\Delta x | Z_0 = 0)/(2\Delta x)$. We claim that

Proposition 2.1. *The probability density function of the random walk has the limit*

$$P_m^n \to \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)}$$

as $n \to \infty$ with $m\Delta x \to x$ and $n\Delta t = t$.

This is the probability density of the standard normal distribution $\mathcal{N}(0, t)$. Thus, the 4th property of the Brownian motion is

- The probability distribution of $z(t)$ is $z(t) \sim \mathcal{N}(0, t)$.

We prove this proposition by using the Stirling formula:

$$n! \approx \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}.$$

Recall that the probability

$$P(Z_n = m\Delta x | Z_0 = 0) = \binom{n}{\frac{1}{2}(m+n)} \left(\frac{1}{2}\right)^n.$$

Using the Stirling formula, we have for n, m, $n - m >> 1$,

$$\begin{aligned} \binom{n}{\frac{1}{2}(m+n)} \left(\frac{1}{2}\right)^n &= \left(\frac{1}{2}\right)^n \frac{n!}{(\frac{1}{2}(n+m))!(\frac{1}{2}(n-m))!} \\ &\approx \left(\frac{1}{2}\right)^n \frac{\sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}}{\sqrt{2\pi}(\frac{1}{2}(n+m))^{\frac{1}{2}(n+m)+\frac{1}{2}} e^{-(n+m)/2} \sqrt{2\pi}(\frac{1}{2}(n-m))^{\frac{1}{2}(n-m)+\frac{1}{2}} e^{-(n-m)/2}} \end{aligned}$$

$$= \left(\frac{2}{\pi n}\right)^{1/2}\left(1+\frac{m}{n}\right)^{-\frac{1}{2}(n+m)-\frac{1}{2}}\left(1-\frac{m}{n}\right)^{-\frac{1}{2}(n-m)-\frac{1}{2}}$$

We shall use the formula

$$\left(1+\frac{x}{n}\right)^n \to e^x, \quad \text{as } n\to\infty.$$

We may treat m as a real number because the functions above are smooth functions. Now, we fix x, t, define $m = x/\Delta x = \sqrt{Cn}$, where $C = x^2/t$. We take $n \to \infty$. Then

$$\left(1+\frac{m}{n}\right)^{-n/2}\left(1-\frac{m}{n}\right)^{-n/2} = \left(1-\frac{m^2}{n^2}\right)^{-n/2} = \left(1-\frac{C}{n}\right)^{-n/2} \to e^{C/2}$$

$$\left(1+\frac{m}{n}\right)^{-m/2} = \left(1+\frac{\sqrt{C}}{\sqrt{n}}\right)^{-\sqrt{Cn}/2} \to e^{-C/2}$$

$$\left(1-\frac{m}{n}\right)^{m/2} = \left(1-\frac{\sqrt{C}}{\sqrt{n}}\right)^{\sqrt{Cn}/2} \to e^{-C/2}$$

$$\left(1+\frac{m}{n}\right)^{-1/2}\left(1-\frac{m}{n}\right)^{-1/2} = \left(1-\frac{m^2}{n^2}\right)^{-1/2} = \left(1-\frac{C}{n}\right)^{-1/2} \to 1$$

Thus, we obtain

$$\left(1+\frac{m}{n}\right)^{-\frac{1}{2}(n+m)-\frac{1}{2}}\left(1-\frac{m}{n}\right)^{-\frac{1}{2}(n-m)-\frac{1}{2}} \to e^{-C/2} = e^{-x^2/(2t)}.$$

We also have

$$\frac{\sqrt{\frac{2}{\pi n}}}{2\Delta x} = \frac{1}{\sqrt{2\pi n}\sqrt{\Delta t}} = \frac{1}{\sqrt{2\pi t}}.$$

Summarizing above, we obtain

$$P_m^n = \frac{1}{2\Delta x}\binom{n}{\frac{1}{2}(m+n)}\left(\frac{1}{2}\right)^n \to \frac{1}{\sqrt{2\pi t}}e^{-x^2/(2t)}.$$

Remarks.

1. We can also show this limiting process as a corollary of the central limit theorem. Recall that

$$Z_n = \sum_{k=1}^n X_k \Delta x,$$

where $X_k \sim X$ are iid and

$$X = \begin{cases} 1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2 \end{cases}$$

Since $\Delta x = \sqrt{\Delta t} = \sqrt{t/n}$, we get

$$Z_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \sqrt{t} X_k.$$

The mean $< \sqrt{t}X >= 0$, the variance $< (\sqrt{t}X)^2 >= t$. Thus, by the central limit theorem, the random variable $\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \sqrt{t} X_k$ converges in distribution to a normal $\mathcal{N}(0,t)$:

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \sqrt{t} X_k \to \mathcal{N}(0,t).$$

2. We can consider more general random walk

$$Z_n = \sum_{k=1}^{n} X_k \sigma \Delta x,$$

where $X_k \sim X$ are i.i.d. The parameter σ, called volatility, measures the variation of Z in each movement. In this case,

$$Z_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \sqrt{t} \sigma X_k \to \mathcal{N}(0, \sigma^2 t).$$

2.6.3 Properties of the Brownian motion

By (d) of the definition

$$\mathcal{P}(z(t) = x | z(0) = 0) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}.$$

Let us consider the Brownian motion starting from y at time s and reaches x at later time t. The probability density is denoted by $p(s,y,x,t)$, called the transition probability density of the stochastic process $z(t)$:

$$\begin{aligned} p(s,y,t,x) &= \mathcal{P}(z(t) = x | z(s) = y) = \mathcal{P}(z(t) = x - y | z(s) = 0) \\ &= \mathcal{P}(z(t-s) = x - y | z(0) = 0) = \frac{1}{\sqrt{2\pi(t-s)}} e^{-(x-y)^2/2(t-s)}. \end{aligned}$$

Let us simplify the notation $p(0,0,t,x)$ by $p(t,x)$. This transition probability density is the probability density function for Brownian motions $z(t)$ starting from 0 at time 0. Any function $f(z)$, its expectation is defined to be

$$\mathbb{E}[f(z(t))] := \int_{-\infty}^{\infty} f(x) p(t,x)\, dx.$$

This transition probability function has the following properties.

1. $\mathbb{E}[z(t)] = 0$
$$\int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)}\, dx = 0.$$

2. $\mathbb{E}[z(t)^2] = t$
$$\int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)}\, dx = t.$$

3. Let us consider an infinitesimal change $dz := z(t + dt) - z(t)$. Here, dt is an infinitesimal time increment. Then we have
$$(dz)^2 = dt \text{ with probability } 1. \tag{2.5}$$
This formula is very important for the stochastic calculus below. It is interpreted by integrating the above equation in t:
$$\int_0^t (dz)^2 = t.$$
The left-hand side is defined to be the limit of the Riemann sum $\sum_{k=1}^n (z(t_k) - z(t_{k-1}))^2$.

Proposition 2.2. *Let $z(\cdot)$ be the Brownian motion. Then*
$$\lim_{n\to\infty} \sum_{k=1}^{n} (z(t_k) - z(t_{k-1}))^2 = t \quad \textit{with probability } 1$$
as $n \to \infty$. Here $t_k := kt/n$.

Proof. 1. We shall apply the strong law of large numbers. Let us consider the random variables: $Y_k := n(z(t_k) - z(t_{k-1}))^2$. We see that
$$\sum_{k=1}^{n} (z(t_k) - z(t_{k-1}))^2 = \frac{1}{n} \sum_{k=1}^{n} Y_k.$$
The random variables Y_k are i.i.d. $Y_k \sim Y$, where
$$Y = n(z(t/n) - z(0))^2.$$

2. The mean of Y is
$$\mathbb{E}[Y] = \int n x^2 \frac{1}{\sqrt{2\pi(t/n)^2}} e^{-x^2/(2t/n)}\, dx = n(t/n) = t.$$
The variance of Y is also finite (check). That is
$$\mathbb{E}[(Y - t)^2] < \infty, \tag{2.6}$$
where
$$E[f(Y)] = \int_{-\infty}^{\infty} f(nx^2) \frac{1}{\sqrt{2\pi(t/n)}} e^{-x^2/(2t/n)}\, dx$$

3. By the strong law of large numbers

$$\frac{1}{n}\sum_{k=1}^{n} Y_k \to t \text{ with probability } 1.$$

□

Exercise

1. Find $\mathbb{E}[z(t)^m]$.
2. Check (2.6).

2.7 Itô's Lemma

In this section, we shall study differential equations which consist of deterministic part: $\dot{x} = b(x)$, and stochastic part $\sigma \dot{z}(t)$. Here, $z(t)$ is the Brownian motion. We call such an equation a stochastic differential equation (s.d.e.) and expressed as

$$dx(t) = b(x(t))dt + \sigma(x(t))dz(t). \tag{2.7}$$

An important lemma for finding their solution is the following Itô's lemma.

Lemma 2.1 (Itô). *Suppose $x(t)$ satisfies the stochastic differential equation (2.7), and $f(x,t)$ is a smooth function. Then $f(x(t),t)$ satisfies the following stochastic differential equation:*

$$df = \left(f_t + bf_x + \frac{1}{2}\sigma^2 f_{xx}\right) dt + \sigma f_x\, dz. \tag{2.8}$$

Proof. According to the Taylor expansion,

$$df = f_t\, dt + f_x\, dx + \frac{1}{2} f_{tt}\, (dt)^2 + f_{xt}\, dx\, dt + \frac{1}{2} f_{xx}\, (dx)^2 + \cdots .$$

Plug (2.7) into this equation. The term

$$(dx)^2 = b^2(dt)^2 + 2b\sigma dt \cdot dz + \sigma^2(dz)^2.$$

In the Taylor expansion of df, the terms $(dt)^2$, $dt \cdot dz$ are relative unimportant as comparing with the dt term and $(dz)^2$ term. Using (2.7) and noting $(dz)^2 = dt$ with probability 1, we obtain (2.8). □

Example 1 The stochastic process

$$x(t) := x_0 + at + \sigma z(t)$$

solves the s.d.e.

$$dx = adt + \sigma dz,$$

where a and σ are constants. By letting $y = x - at$, then y is a function of x and t. From Itô's lemma, $y(x(t), t)$ is also a stochastic process and y satisfies $dy = \sigma dz$. Since σ is a constant, we then have $y(t) = y_0 + \sigma z(t)$. The transition probability density function for y is

$$\mathcal{P}(y(t) = y | y(0) = y_0) = \frac{1}{\sqrt{2\pi\sigma^2 t}}\, e^{-(y-y_0)^2/2\sigma^2 t}.$$

Or equivalently, the transition probability density function for x is

$$\mathcal{P}(x(t) = x | x(0) = x_0) = \frac{1}{\sqrt{2\pi\sigma^2 t}}\, e^{-(x-at-x_0)^2/2\sigma^2 t}.$$

Example 2. Let $y(t) = y_0 + z(t)^2$. Then $y(t)$ solves

$$dy = dt + 2z(t)dz(t).$$

To show this, we apply Itô's lemma. Let $y = f(x, t) = x^2$, and $x = z$. Then

$$dy(t) = f_x dx + \frac{1}{2} f_{xx} dt = 2z(t)dz(t) + dt.$$

This means

$$d(z^2) = 2zdz + dt.$$

Example 3. Let $dx = dz$ and $y = \ln x$. Then $f_x = 1/x$ and $f_{xx} = -1/x^2$. Thus, we have

$$d \ln x = \frac{1}{x} dx - \frac{1}{2x^2}(dx)^2.$$

In terms of Brownian motion, it can be written as

$$d \ln z = \frac{dz}{z} - \frac{dt}{2z^2}.$$

Thus, $\ln z$ solves the above s.d.e.

Example 4. Let $y = e^{\sigma z}$, σ is a constant. We choose $dx = \sigma dz$ and $y = e^x$. Then

$$de^x = e^x dx + \frac{1}{2} e^x (dx)^2 = e^x \sigma dz + \frac{1}{2} e^x \sigma^2 dt.$$

Thus

$$de^{\sigma z} = e^{\sigma z} \sigma dz + \frac{1}{2} e^{\sigma z} \sigma^2 dt.$$

2.8 Geometric Brownian Motion: the solution of the continuous asset price model

In this section, we want to find the transition probability density function for the continuous asset price model:

$$dS = \mu S\,dt + \sigma S\,dz. \tag{2.9}$$

with initial data $S(0) = S_0$. We apply Itô's lemma with $x = f(S) = \ln S$. Then, $f_S = 1/S$ and $f_{SS} = -1/S^2$. By Itô's lemma, x satisfies the s.d.e.

$$dx = f_S dS + \frac{1}{2} f_{SS}(dS)^2 = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma\,dz.$$

and $x(0) = x_0 := \ln S_0$. From Example 1 of the previous section, we obtain

$$x(t) = x_0 + (\mu - \frac{\sigma^2}{2})t + \sigma z(t).$$

Since $S = e^x$, we obtain

$$S(t) = S_0 \exp\left((\mu - \frac{\sigma^2}{2})t + \sigma z(t)\right). \tag{2.10}$$

The probability density of $x(t)$ is

$$\mathcal{P}(x(t) = x | x(0) = x_0) = \frac{1}{\sqrt{2\pi\sigma^2 t}}\, e^{-(x - x_0 - (\mu - \frac{\sigma^2}{2})t)^2/2\sigma^2 t}.$$

That is,

$$x(t) \sim \mathcal{N}(x_0 + (\mu - \frac{\sigma^2}{2})t, \sigma^2 t).$$

From

$$\begin{aligned}\mathcal{P}(S(t) = S|S(0) = S_0)dS &= \mathcal{P}(x(t) = x|x(0) = x_0)dx \\ &= \mathcal{P}(x(t) = x|x(0) = x_0)dS/S\end{aligned}$$

we obtain that the transition probability density function for $S(t)$ is

$$\mathcal{P}(S(t) = S|S(0) = S_0) = \frac{1}{\sqrt{2\pi\sigma^2 t}S}\, e^{-(\ln\frac{S}{S_0} - (\mu - \frac{\sigma^2}{2})t)^2/2\sigma^2 t}. \tag{2.11}$$

This is called the *log-normal distribution*. The stochastic process $x(t)$ obeys the normal distribution, while $S(t) = e^{x(t)}$ is called to obey the geometric Brownian motion.

Remark. Alternatively, we may also use probability distribution function to derive the probability density function as below. The probability distribution function of $S(t)$ is

$$F_S(S) = \int_{-\infty}^{S} \mathcal{P}(S(t) = S_1 | S(0) = S_0)\, dS_1$$

On the other hand,

$$F_x(x) = \int_{-\infty}^{x} \mathcal{P}(x(t) = x_1 | x(0) = x_0)\, dx_1$$

We notice that

$$F_S(S) = F_x(x) \quad \text{for } x = \ln S,\ x_0 = \ln S_0.$$

Thus,

$$\begin{aligned} \mathcal{P}(S(t) = S | S(0) = S_0) &= \frac{d}{dS} F_S(S) = \frac{d}{dS} F_x(x) = \frac{dx}{dS}\frac{d}{dx} F_x(x) \\ &= \frac{1}{S}\mathcal{P}(x(t) = \ln S | x(0) = \ln S_0) = \frac{1}{\sqrt{2\pi\sigma^2 t} S}\, e^{-(\ln\frac{S}{S_0} - (\mu - \frac{\sigma^2}{2})t)^2/2\sigma^2 t}. \end{aligned}$$

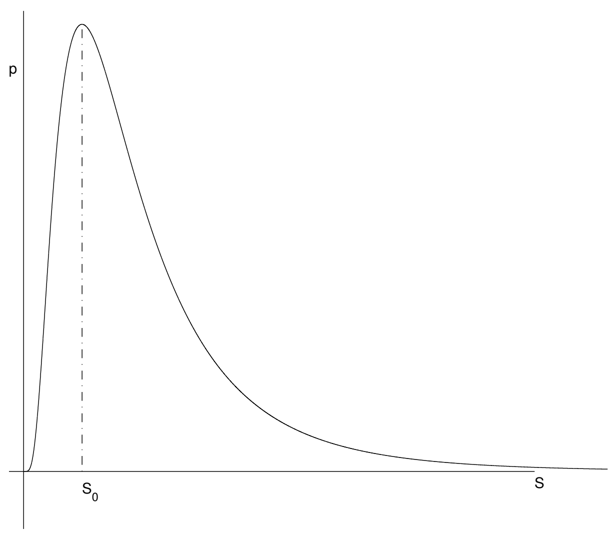


Figure 2.1: The log-normal distribution.

Properties of the geometric normal distribution $S(t)$ Let us denote the transition probability density $\mathcal{P}(S(t) = s | S(0) = S_0)$ by $p_{S(t)}(s)$ for short. That is

$$p_{S(t)}(s) = \frac{1}{\sqrt{2\pi\sigma^2 t s}}\, e^{-(\ln\frac{s}{S_0} - (\mu - \frac{\sigma^2}{2})t)^2/2\sigma^2 t}$$

We compute its mean and variance below.

Proposition 2.3. *The mean and variance of the geometric Brownian motion are*

(a) $\mathbb{E}[S(t)] = \int_{-\infty}^{\infty} s p_{S(t)}(s)\, ds = S_0 e^{\mu t}$,

(b) *Var*$[S(t)] = S_0^2 e^{2\mu t}[e^{\sigma^2 t} - 1]$.

Proof. (a)

$$\begin{aligned}
\mathbb{E}[S(t)] &= \int_0^{\infty} s p_{S(t)}(s)\, ds = \int_0^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}}\, e^{-(\ln \frac{s}{S_0} - (\mu - \frac{\sigma^2}{2})t)^2/2\sigma^2 t}\, ds \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^x e^{-(x - x_0 - (\mu - \frac{\sigma^2}{2})t)^2/2\sigma^2 t}\, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{x + x_0 + \mu t} e^{-(x + \frac{\sigma^2}{2}t)^2/2\sigma^2 t}\, dx \\
&= S_0 e^{\mu t} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{x - (x + \frac{\sigma^2}{2}t)^2/2\sigma^2 t}\, dx \\
&= S_0 e^{\mu t} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-(x - \frac{\sigma^2}{2}t)^2/2\sigma^2 t}\, dx \\
&= S_0 e^{\mu t}.
\end{aligned}$$

(b)

$$\begin{aligned}
\mathbb{E}[S(t)^2] &= \int_0^{\infty} s^2 p_{S(t)}(s)\, ds = \int_0^{\infty} s \frac{1}{\sqrt{2\pi\sigma^2 t}}\, e^{-(\ln \frac{s}{S_0} - (\mu - \frac{\sigma^2}{2})t)^2/2\sigma^2 t}\, ds \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{2x} e^{-(x - x_0 - (\mu - \frac{\sigma^2}{2})t)^2/2\sigma^2 t}\, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{2(x + x_0 + \mu t)} e^{-(x + \frac{\sigma^2}{2}t)^2/2\sigma^2 t}\, dx \\
&= S_0^2 e^{2\mu t} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{2x - (x + \frac{\sigma^2}{2}t)^2/2\sigma^2 t}\, dx \\
&= S_0 e^{\mu t} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-(x - \frac{3\sigma^2}{2}t)^2/2\sigma^2 t + \sigma^2 t}\, dx \\
&= S_0^2 e^{2\mu t + \sigma^2 t}.
\end{aligned}$$

$$\text{Var}[S(t)] = \mathbb{E}[S(t)^2] - \mathbb{E}[S(t)]^2 = S_0^2 e^{2\mu t}[e^{\sigma^2 t} - 1].$$

□

2.9 Calibrating geometric Brownian motion

How do we obtain μ and σ of an asset? We estimate them from the historical data S_i, at discrete time $t_i := i\Delta t$, $i = 0, ..., n$. The procedure is as below. First, let $x(t) = \ln S(t)$. Let $\Delta x_i :=$

$x(t_{i+1}) - x(t_i)$. The variables Δx_i are independent random variables with identical distribution:

$$\Delta x_i \sim \Delta x \sim \mathcal{N}((\mu - \frac{\sigma^2}{2})\Delta t, \sigma^2 \Delta t).$$

The historical data we collect are S_i which is a sample of $S(t_i)$. We define $x_i = \ln S_i$. We estimate the mean and variance of Δx by using these historical data x_i:

$$\left(\hat{\mu} - \frac{\hat{\sigma}^2}{2}\right)\Delta t = \frac{1}{n}\sum_{i=1}^{n}(x_{i+1} - x_i) := \hat{m},$$

$$\hat{\sigma}^2 \Delta t = \frac{1}{n-1}\sum_{i=1}^{n}\left((x_{i+1} - x_i) - \hat{m}\right)^2.$$

The $\hat{\mu}$ and $\hat{\sigma}$ are the estimators of μ and σ. [1]

[1]The variance estimator has $n-1$ in denominator instead of n. This is called unbiased estimator because the expectation of this estimator is the correct variance. This is called Bessel's correction.

Chapter 3

Black-Scholes Analysis

3.1 The hypothesis of no-arbitrage-opportunities

The option pricing theory was introduced by Black and Scholes. The fundamental hypothesis of their analysis is that ”there is no arbitrage opportunities in financial markets”.

For simplicity, we shall also assume

1. There exists a risk-free investment that gives a guaranteed return with interest rate r. (e.g. government bond, bank deposit.)

2. Borrowing or lending at such riskless interest rate is always possible.

3. There is no transaction costs.

4. All trading profits are subject to the same tax rate.

We will use the following notations:

S	current asset price
E	exercise price
T	expiry time
t	current time
μ	growth rate of an asset
σ	volatility of an asset
S_T	asset price at T
r	risk-free interest rate
c	value of European call option
C	value of American call option
p	value of European put option
P	value of American put option
Λ	the payoff function

Remark.

1. In assumption 1, if the annual interest rate is r and the compound interest is counted daily, then the principal plus the compound interest in one year are
$$A\left(1+\frac{r}{365}\right)^{365} \approx Ae^r.$$

2. Assumption 1 implies that an amount of money A at time t has value $Ae^{r(T-t)}$ at time T if it is deposited into bank. This is called the future value of A. Similarly, a strike price E at time T should be deducted to $Ee^{-r(T-t)}$ at time t. This is called its present value.

3. The option price c, p, C and P are functions of (S_t, t). Usually, we only write c, which means the current value $c(S_t, t)$. We use the same convention for S, an abbreviation of the current value S_t.

3.2 Basic properties of option prices

3.2.1 The relation between payoff and options

1. The payoff of a contract is the return from the deal. For European option, it can only be exercised on expiry date. Therefore, its payoff Λ is only meaningful at the expiry date T. However, for American option, the option can be exercised at anytime between t and the expiration date T. It is therefore meaningful to define the payoff function at t. For a person who longs an American call option, his payoff is $\Lambda(t) = \max(S_t - E, 0)$, while for American put option, $\Lambda(t) = \max(E - S_t, 0)$.

2. $c(S_T, T) = \Lambda(T) = \max(S_T - E, 0)$.
Otherwise, there is a chance of arbitrage. For instance, if $c(S_T, T) < \Lambda(T)$, then we can buy a call on price c, exercise it immediately. If $S_T > E$, then $\Lambda = S_T - E > 0$ and $c < \Lambda$ by our assumption. Hence we have an immediate net profit $S_T - E - c > 0$. This contradicts to our hypothesis. If $S_T \le E$, then $\Lambda(T) = 0$ and $c < \Lambda(T) = 0$ leading to c has negative value. This means that anyone who can buy this call option with negative amount of money. This again contradicts to our hypothesis.
If $c(S_T, T) > \Lambda(T)$, we can short a call and earn $c(S_T, T)$. If the person who buy the call does not claim, then we have net profit c. If he does exercise his call, then we can buy an asset from the market on price S_T and sell to that person with price E. The cost to us is $S_T - E$. By doing so, the net profit we get is $c(S_T, T) - (S_T - E) > 0$. Again, this is a contradiction.

3. $p(S_T, T) = \Lambda(T) = \max(E - S_T, 0)$.
If $p(S_T, T) < \Lambda(t)$, then an arbitrageur longs a put at T, exercises it right away, and get net profit $\Lambda(T) - p(S_T, T)$. If $p(S_T, T) > \Lambda(T)$, then an arbitrageur shorts a put. If the person who buys the put exercise it, then the arbitrageur earns $P(S_T, T) - \Lambda(T)$; if he gives up this put, then the arbitrageur earns $p(S_T, T)$.

4. We show $C(S_t, t) = \Lambda(t) := \max(S_t - E, 0)$.
If $C(S_t, t) > \Lambda(t)$, then an arbitrageur can short the call and earn $C(S_t, t)$. If the person who exercises it right away, the arbitrageur need to pay $\Lambda(t)$. He still has net profit $C(S_t, t) - \Lambda(t)$. If $C(S_t, t) < \Lambda(t)$. Then an arbitrageur buy C and exercises it right away and get net profit $\Lambda(t) - C(S_t, t)$.

5. $P(S_t, t) = \Lambda(t) := \max(E - S_t, 0)$. This can be shown similarly.

3.2.2 European options

Theorem 3.1. *The following statements hold for European options*

$$\max\{S - Ee^{-r(T-t)}, 0\} \leq c \leq S \tag{3.1}$$

$$\max\{Ee^{-r(T-t)} - S, 0\} \leq p \leq Ee^{-r(T-t)} \tag{3.2}$$

and the put-call parity

$$p + S = c + Ee^{-r(T-t)} \tag{3.3}$$

To show these, we need the following definition and lemmae.

Definition 3.2. *A portfolio is a collection of investments.*

For instance, a portfolio $I = c - \Delta S$ means that we long a call and short Δ amount of an asset S.

Remark. The value of a portfolio depends on time. Suppose a portfolio I has A amount of cash at time t, its value is $Ae^{r(T-t)}$ at time T. On the other hand, a portfolio involves E amount of cash, its present value is $Ee^{-r(T-t)}$. We should make such deduction, otherwise there is an arbitrage opportunity.

Lemma 3.2. *Suppose $I(t)$ and $J(t)$ are two portfolios containing no American options. Then under the hypothesis of no-arbitrage-opportunities, we can conclude that*

$$I(T) \leq J(T) \Rightarrow I(t) \leq J(t), \quad \forall t \leq T,$$

$$I(T) = J(T) \Rightarrow I(t) = J(t), \quad \forall t \leq T.$$

Proof. Suppose the conclusion is false, i.e., there exists a time $t \leq T$ such that $I(t) > J(t)$. An arbitrageur can buy (long) $J(t)$ and short $I(t)$ and immediately gain a profit $I(t) - J(t)$. Its value at time T is $(I(t) - J(t))e^{r(T-t)}$. Since I and J containing no American options, nothing can be exercised before T. At time T, since $I(T) \leq J(T)$, he can use $J(T)$ (what he has) to cover $I(T)$ (what he shorts) and gains another profit $J(T) - I(T)$. This contradicts to the hypothesis of no-arbitrage-opportunities.

The equality case can be proven similarly. □

Proof of Theorem 3.1.

1. Let $I = c$ and $J = S$. At T, we have

   $$I(T) = c_T = \max\{S_T - E, 0\} \leq \max\{S_T, 0\} = S_T = J(T).$$

   Hence, $I(t) \leq J(t)$ holds for all $t \leq T$.

   Remark. The equality holds when $E = 0$. In this case $c = S$

2. Consider $I = c$ and $J = 0$. At time T, $I(T) = \Lambda(T) \geq 0 = J(T)$, hence $I(t) \geq 0$ for all $t \leq T$. Similarly, we also get $p \geq 0$.

3. Consider $I = c + Ee^{-r(T-t)}$ and $J = S$. At time T,

   $$I(T) = \max\{S_T - E, 0\} + E = \max\{S_T, E\} \geq S_T = J(T).$$

   This implies $I(t) \geq J(t)$.

4. Let $I = p$ and $J = Ee^{-r(T-t)}$. At time T,

   $$I(T) = \max\{E - S_T, 0\} \leq E = J(T).$$

   Hence, $p(t) \leq Ee^{-r(T-t)}$. The equality holds only when $S_T = 0$.

5. Consider $I = p + S$ and $J = Ee^{-r(T-t)}$. At time T,

   $$I(T) = \max(E - S_T, 0) + S_T = \max\{S_T, E\} \geq E = J(T).$$

   Hence, $I(t) \geq J(t)$.

6. For put-call parity, we consider $I = c + Ee^{-r(T-t)}$ and $J = p + S$. At time T,

   $$\begin{aligned} I(T) &= c + E = \max\{S_T - E, 0\} + E = \max\{S_T, E\}, \\ J(T) &= p + S = \max\{E - S_T, 0\} + S_T = \max\{E, S_T\} \end{aligned}$$

   Hence, $I(t) = J(t)$. □

3.2.3 Basic properties of American options

Theorem 3.2. *For American options, we have*

(i) The optimal exercise time for American call option is T, and we have $C = c$.

(ii) The optimal exercise time for American put option is as earlier as possible, and we have $P(t) \geq p(t)$.

(iii) The put-call parity for American option is

$$C + Ee^{-r(T-t)} \leq S + P \leq C + E. \tag{3.4}$$

To prove these properties, we need the following lemma.

Lemma 3.3. *Let I or J be two portfolios that contain American options. Suppose $I(\tau) \le J(\tau)$ at some $\tau < T$. Then $I(t) \le J(t)$, for all $t \le \tau$.*

Proof. Suppose $I(t) > J(t)$ at some $t \le \tau$. An arbitrageur can long $J(t)$ and short $I(t)$ at time t to make profit $I(t) - J(t)$ immediately. At later time τ, he can use $J(\tau)$ to cover $I(\tau)$ with additional profit $J(\tau) - I(\tau)$, in case the person who owns I exercises his American option. □

Remark. The equality also holds , that is, if $I(\tau) = J(\tau)$ for $\tau < T$, then $I(t) = J(t)$ for $t \le \tau$.

Proof of Theorem 3.2.

1. Firstly, we show $C_t \ge c_t$ for all $t < T$. If not, then $c(\tau) > C(\tau)$ for some time $\tau < T$, we can buy C and sell c at time τ to make a profit $c(\tau) - C(\tau)$. The right of C is even more than that of c. This is an arbitrage opportunity which is a contradiction.

   Secondly, we show $c_t \ge C_t$. Consider two portfolios $I(t) = C_t + Ee^{-r(T-t)}$ and $J(t) = S_t$. The portfolio I is an American call plus an amount of cash $Ee^{-r(T-t)}$. Suppose we exercise C at some time $\tau < T$, then $I(\tau) = S_\tau - E + Ee^{-r(T-\tau)}$ and $J(\tau) = S_\tau$. This implies $I(\tau) < J(\tau)$. By our lemma, $I(t) \le J(t)$ for all $t \le \tau$. Since $\tau \le T$ arbitrary, we conclude $I(t) \le J(t)$ for all $t < T$. Combine this inequality with the inequality

   $$c_t + Ee^{-r(T-t)} \ge S_t$$

   in section 3.2,we conclude $c_t = C_t$ for all $t < T$. Further, early exercise results in $C(\tau) + Ee^{-r(T-\tau)} = S(\tau) - E + Ee^{-r(T-\tau)} < S(\tau)$. But $S(\tau) \le c_\tau + Ee^{-r(T-\tau)}$. Thus, early exercise leads to decrease the value of $C(\tau)$, if $\tau < T$. Hence, the optimal exercise time for American option is T.

2. We show $p(t) \le P(t)$. Suppose $p(t) > P(t)$. Then we can make an immediate profit by selling p and buying P. We earn $p - P$ and gain more right. This is a contradiction.

   Next, we show that we should exercise American puts as early as possible (the first time $E > S_\tau$). We consider a portfolio $I(t) = P_t + S_t$. This is called a covered put. That is, we use S to cover P when we exercise P. When we exercise P at $\tau < T$, the payoff of I is $E - S_\tau + S_\tau = E$. By the time T, its value is $Ee^{r(T-\tau)}$. Thus, in order to maximize the profit, we exercise time should be as earlier as possible.

3. The inequality $C + Ee^{-r(T-t)} \le S + P$ follows from the put-call parity (3.3) and the facts that $c = C$ and $P \ge p$. To show the inequality $S + P < C + E$, we consider two portfolios: $I(t) = c(t) + E$ and $J(t) = P(t) + S_t$. Suppose P is not exercised before T. Then $J(T) = \max(E - S_T, 0) + S_T = \max(E, S_T)$, and

   $$I(T) = c(T) + Ee^{r(T-t)} = \max(E, S_T) + E(e^{r(T-t)} - 1) > J(T).$$

Suppose P is exercised at some time τ, with $\tau < T$. Then we sell S_τ at price E. Thus $J(\tau) = E$. On the other hand,

$$I(\tau) = c(\tau) + Ee^{r(\tau-t)} \geq E = J(\tau).$$

From lemma, we have $I(t) > J(t)$. That is, $c(t) + E \geq P(t) + S_t$. Since $c(t) = C(t)$, we also have $C(t) + E \geq P(t) + S_t$. □

Remark. $P - p$ is called the time value of a put. The maximal time value is $E - Ee^{-r(T-t)}$.

Examples.

1. Suppose $S(t) = 31$, $E = 30$, $r = 10\%$, $T - t = 0.25$ year, $c = 3$, $p = 2.25$. Consider two portfolios:

$$\begin{aligned} I &= c + Ee^{-r(T-t)} = 3 + 30 \times e^{-0.1\times 0.25} = 32.26, \\ J &= p + S = 2.25 + 31 = 33.25. \end{aligned}$$

We find $J(t) > I(t)$. It violates the put-call parity. This means that there is an arbitrage opportunity.
Strategy : long the security in portfolio I and short the security in portfolio J. This results a cashflow: $-3+2.25+31 = 30.25$. Put this cash into a bank. We will get $30.25 \times e^{0.1\times 0.25} = 31.02$ at time T. Suppose at time T, $S_T > E$, we can exercise c, also we should buy a share for E to close our short position of the stock. Suppose $S_T < E$, the put option will be exercised. This means that we need to buy the share for E to close our short position. In both cases, we need to buy a share for E to close the short position. Thus, the net profit is

$$31.02 - 30 = 1.02.$$

2. Consider the same situation but $c = 3$ and $p = 1$. In this case

$$\begin{aligned} I &= c + Ee^{-r(T-t)} = 32.25 \\ J &= p + S = 1 + 31 = 32. \end{aligned}$$

and we see that J is cheaper.
Strategy: We long J and short I. To long J, we need an initial investment $31 + 1$, to short c, we gain 3. Thus, the net investment is $31 + 1 - 3 = 29$ initially. We can finance it from the bank, and we need to pay $29 \times e^{0.1\times 0.25} = 29.73$ to the bank at time T. Now, at T, we must have that either c or p be exercised. If $S_T > E$, then c is exercised. We need to sell the share for E to close our short position for c. If $S_T < E$, we exercise p. That is, we sell the share for E. In both cases, we sell the share for E. Thus, the net profit is $30 - 29.73 = 0.27$.

3.2.4 Dividend Case

Many stocks pay out dividends. These are payments to shareholders out of the profits made by the company. Since the company's wealth does not change after paying the dividends, the stock price, the strike prices fall as the dividends being paid. If a company declared a cash dividend, the strike price for options was reduced on the ex-dividend day by the amount of the dividend.

The inequalities below are all at the present time t. Thus, let D be the present value of the dividend, no matter when they are paid. Certainly its value at time T is $De^{r(T-t)}$. Thus, a company pays dividend before T, its asset at time T becomes $S_T + De^{r(T-t)}$.

Theorem 3.3. *Suppose a dividend will be paid during the life of an option. Let D denotes its present value. Then we have for European option*

$$S - D - Ee^{-r(T-t)} \leq c \leq S \tag{3.5}$$

$$-S + D + Ee^{-r(T-t)} \leq p \leq Ee^{-r(T-t)}. \tag{3.6}$$

and the put-call parity:

$$c + Ee^{-r(T-t)} = p + S - D. \tag{3.7}$$

For the American options, we have (i)

$$S - D - E \leq C - P \leq S - Ee^{-r(T-t)} \tag{3.8}$$

provided the dividend is paid before exercising the put option, or (ii)

$$S - E \leq C - P \leq S - Ee^{-r(T-t)} \tag{3.9}$$

if the put is exercised before the dividend being paid.

Proof. 1. Proof of $c \geq S - Ee^{-r(T-t)} - D$.

We consider two portfolios:

$$\begin{aligned} I &= c + D + Ee^{-r(T-t)}, \\ J &= S. \end{aligned}$$

Then at time T,

$$\begin{aligned} I(T) &= \max\{S_T - E, 0\} + D + E = \max\{S_T, E\} + De^{r(T-t)} \\ J(T) &= S_T + De^{r(T-t)}. \end{aligned}$$

Hence $I(T) \geq J(T)$. This yields $I(t) \geq J(t)$ for all $t \leq T$. This proves

$$c \geq S - D - Ee^{-r(T-t)}.$$

In other word, c is reduced by an amount D.

2. Proof of $p \geq D + Ee^{-r(T-t)} - S$. That is, p increases by an amount D.
We consider
$$I = p + S$$
$$J = Ee^{-r(T-t)} + D.$$
At time T, $I(T) = p(T) = \max(E - S_T, 0) + S_T + De^{r(T-t)}$, while $J(T) = E + De^{r(T-t)}$. Thus, $I(T) \geq J(T)$, we then obtain $p + S \geq Ee^{-r(T-t)} + D$.

3. For the put-call parity, we consider
$$\begin{array}{rcl} I &=& c + D + Ee^{-r(T-t)} \\ J &=& S + p. \end{array}$$
At time T,
$$I = J = \max\{S_T, E\} + De^{r(T-t)}.$$
This yields the put-call parity for all time.

4. When there is no dividend, we have shown that
$$C - P \leq S - Ee^{-r(T-t)}.$$
When there is dividend payment, we know that
$$C_D \leq C, \quad P_D \geq P$$
Hence,
$$C_D - P_D \leq C - P \leq S - Ee^{-r(T-t)}.$$
Here, C_D and P_D refer to Amperican call and put with paying dividends. For the American call option, we should not exercise it early, because the dividend will cause the stock price to jump down, making the option less attractive. We should exercise it immediately after the ex-dividend date.

5. For the American put option, we consider
$$I = C + D + E, \quad J = P + S.$$
If we exercise P at $\tau \leq T$ after the dividend being paid, then $S_\tau < E$ and
$$\begin{array}{rcl} I(\tau) &=& De^{r(\tau-t)} + Ee^{r(\tau-t)}, \\ J(\tau) &=& E + De^{r(\tau-t)}. \end{array}$$
We have $J(\tau) \leq I(\tau)$. Hence $J(t) \leq I(t)$ for all $t \leq \tau$.

6. If the put option is exercised before the dividend being paid, then we should consider $I = C + E$ and $J = P + S$. At τ,
$$\begin{array}{rcl} I(\tau) &=& Ee^{r(\tau-t)}, \\ J(\tau) &=& E. \end{array}$$
Again, we have $J(\tau) \leq I(\tau)$. Hence $J(t) \leq I(t)$ for all $t \leq \tau$.

□

3.3 The Black-Scholes Equation

3.3.1 Black-Scholes Equation

The fundamental hypothesis of the Black-Scholes analysis is that *there is no arbitrage opportunities.* Besides, we make the following additional assumptions:

(1) The asset price follows the log-normal distribution. That is, the asset price S satisfies the s.d.e.

$$dS = \mu S dt + \sigma S dz.$$

(2) There exists a risk-free interest rate r.

(3) No transaction costs.

(4) No dividend paid.

(5) Shorting selling is permitted.

Our purpose is to value the price of an option (call or put). Let $V(S,t)$ denotes for the price of an option. The randomness of $V(S(t),t)$ would be fully correlated to $S(t)$. Thus, we consider a portfolio which contains only S and V, but in opposite position in order to cancel out the randomness. Then this portfolio becomes deterministic. To be more precise, let the portfolio be

$$\Pi = V - \Delta S,$$

that is, long an option and short Δ amount of the underlying asset. In a small time step dt, the change of this portfolio is

$$d\Pi = dV - \Delta dS.$$

Here Δ is held fixed during the time step. From Itô's lemma

$$\begin{aligned} d\Pi &= V_t dt + V_S dS + \frac{1}{2} V_{SS} \sigma^2 dt - \Delta dS \\ &= \sigma S \left(\frac{\partial V}{\partial S} - \Delta \right) dz \\ &\quad + \left(\mu S \frac{\partial V}{\partial S} - \mu S \Delta + \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt. \end{aligned}$$

Now, we can eliminate the randomness by choosing

$$\Delta = \frac{\partial V}{\partial S} \tag{3.10}$$

at the starting time of each time step. The resulting portfolio

$$\Pi = V - V_S S$$

has the property that its change

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt$$

is wholly deterministic. From the hypothesis of no arbitrage opportunities, the return, $\frac{d\Pi}{\Pi}$, should be the same as investing Π in a riskless bank with interest rate r, i.e.

$$\frac{d\Pi}{\Pi} = rdt.$$

Otherwise, there would be either a net loss or an arbitrage opportunity. Hence we must have

$$\begin{aligned} r\Pi dt &= \left(\mu S \frac{\partial V}{\partial S} - \mu S \Delta + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt \\ &= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt, \end{aligned}$$

or

$$\boxed{\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = r\left(V - S\frac{\partial V}{\partial S} \right).} \tag{3.11}$$

This is the Black-Scholes partial differential equation (P.D.E.) for option pricing.

- Its left-hand side $V_t + \frac{\sigma^2}{2} S^2 V_{SS}$ is the return from the hedged portfolio per unit time,
- while its right-hand side $r(V - SV_S)$ is the return from bank deposit per unit time.

Note that the equation is independent of μ.
Remark. Notice that the Black-Scholes equation is invariant under the change of variable $S \mapsto \lambda S$.

3.3.2 Boundary and Final condition for European options

- Final condition:

$$\begin{aligned} c(S,T) &= \max\{S - E, 0\} \\ p(S,T) &= \max\{E - S, 0\}. \end{aligned}$$

  In general, the final condition is

$$V(S,T) = \Lambda(S),$$

  where Λ is the payoff function. These final conditions are derived from the "no-arbitrage opportunities" assumption.

- Boundary conditions:

  (i) On $S = 0$:

$$c(0,\tau) = 0, \quad \forall t \le \tau \le T.$$

  This means that you wouldn't want to buy a right whose underlying asset costs nothing.

(ii) On $S = 0$:
$$p(0,\tau) = Ee^{-r(T-\tau)}.$$
This follows from the put-call parity and $c(0,t) = 0$.

(iii) For call option, at $S = \infty$:
$$c(S,t) \sim S - Ee^{-r(T-t)}, \text{ as } S \to \infty.$$
Since $S \to \infty$, the call option must be exercised, and the price of the option must be closed to $S - Ee^{-r(T-t)}$.

(iv) For put option, at $S = \infty$:
$$p(S,t) \to 0, \text{ as } S \to \infty$$
As $S \to \infty$, the payoff function $\Lambda = \max\{E - S, 0\}$ is zero. Thus, the put option is unlikely to be exercised. Hence $p(S,T) \to 0$ as $S \to \infty$.

3.4 Exact solution for the B-S equation for European options

3.4.1 Reduction to parabolic equation with constant coefficients

Let us recall the Black-Scholes equation
$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0, \quad (0,T] \times (0,\infty). \tag{3.12}$$
This P.D.E. is a parabolic equation with variable coefficients. Notice that this equation is invariant under $S \to \lambda S$. That is, it is homogeneous in S with degree 0. We therefore make the following change-of-variable:
$$x = \log\frac{S}{E}.$$
The fraction S/E makes x dimensionless. The domain $S \in (0,\infty)$ becomes $x \in (-\infty,\infty)$ and
$$\begin{aligned}
\frac{\partial V}{\partial x} &= \frac{\partial S}{\partial x}\frac{\partial V}{\partial S} = S\frac{\partial V}{\partial S}, \\
\frac{\partial^2 V}{\partial x^2} &= \frac{\partial}{\partial x}\left(S\frac{\partial V}{\partial S}\right) \\
&= \frac{\partial S}{\partial x}\frac{\partial V}{\partial S} + S\frac{\partial S}{\partial x}\frac{\partial^2 V}{\partial S^2} \\
&= S\frac{\partial V}{\partial S} + S^2\frac{\partial^2 V}{\partial S^2} \\
&= \frac{\partial V}{\partial x} + S^2\frac{\partial^2 V}{\partial S^2}.
\end{aligned}$$
Next, let us reverse the time by letting
$$\tau = T - t.$$

Then the Black-Scholes equation becomes

$$\frac{\partial V}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial V}{\partial x} - rV.$$

We can also make V dimensionless by setting $v = V/E$. Then v satisfies

$$\boxed{\frac{\partial v}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial x^2} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial v}{\partial x} - rv.} \tag{3.13}$$

The initial and boundary conditions for v becomes

$$v(x, 0) = \bar{\Lambda}(x) = \Lambda(Ee^x)/E,$$

- call option: $\bar{\Lambda}(x) = \max\{e^x - 1, 0\}$, $v(-\infty, \tau) = 0$, $v(x, \tau) \to e^x - e^{-r\tau}$ as $x \to \infty$;
- put option: $\bar{\Lambda}(x) = \max\{1 - e^x, 0\}$, $v(-\infty, \tau) = e^{-r\tau}$, $v(x, \tau) \to 0$ as $x \to \infty$.

Our goal is to solve v for $0 \le \tau \le T$.

3.4.2 Further reduction

Let us divide the v equation by $\sigma^2/2$:

$$\frac{2}{\sigma^2}\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2} + \left(\frac{2r}{\sigma^2} - 1\right)\frac{\partial v}{\partial x} - \frac{2r}{\sigma^2}v.$$

Let us make a change of variable and define the following new constants:

$$\begin{aligned} s &= \tau/(\frac{1}{2}\sigma^2) \\ a &= 1 - r/(\frac{1}{2}\sigma^2) \\ b &= r/(\frac{1}{2}\sigma^2) \end{aligned}$$

Then the equation becomes

$$v_s + av_x + bv = v_{xx}. \tag{3.14}$$

The part, $v_s + av_x$ is call the advection term. The term bv is called the source term, and the term v_{xx} is called the diffusion term. The advection part:

$$v_\tau + av_x = (\partial_\tau + a\partial_x)\, v$$

is a direction derivative along the curve (called characteristic curve)

$$\frac{dx}{d\tau} = a.$$

This suggests the following change-of-variable:

$$\begin{aligned} y &= x - as \\ s' &= s. \end{aligned}$$

Then the direction derivative become

$$\begin{aligned} \partial_{s'} &= \partial_s + a\partial_x \\ \partial_y &= \partial_x \end{aligned}$$

Hence the equation is reduced to

$$v_{s'} + bv = v_{yy}.$$

Since $s' = s$, we will still use s below instead of s'. Therefore, the equation becomes

$$v_s + bv = v_{yy}.$$

Next, the equation $v_s + bv$ suggests that v behaves like e^{bs} along the characteristic curves. Thus, it is natural to make the following change-of-variable

$$v = e^{-bs}u.$$

Then

$$\partial_s u = \partial_s(e^{bs}v) = e^{bs}\partial_s v + be^{bs}v = e^{bs}(\partial_s + b)v = e^{bs}\partial_y^2 v = \partial_y^2 u.$$

Thus, the equation is reduced to

$$u_s = u_{yy}, \quad y \in (-\infty, \infty), s > 0.$$

This is the standard heat equation on the real line. Its solution can be expressed as

$$u(y,s) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi s}} e^{-\frac{(y-z)^2}{4s}} f(z)\,dz, \quad s > 0,$$

where f is the initial data.

3.4.3 Black-Scholes formula

Let us return to the Black-Scholes equation (5.6). Thus, the solution v

$$\boxed{v(x,\tau) = e^{-r\tau} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2\tau}} e^{-\frac{(x-z+(r-\frac{1}{2}\sigma^2)\tau)^2}{2\sigma^2\tau}} \bar{\Lambda}(z)\,dz} \tag{3.15}$$

In terms of the original variables, with the change of variable $S' = Ee^z$, we have the following Black-Scholes formula:

$$V(S,t) = e^{-r(T-t)} \int_0^{\infty} \frac{1}{\sqrt{2\pi\sigma^2(T-t)}S'} e^{-\frac{(\ln(\frac{S}{S'})+(r-\frac{1}{2}\sigma^2)(T-t))^2}{2\sigma^2(T-t)}} \Lambda(S')\,dS' \tag{3.16}$$

We may express it as

$$\boxed{V(S,t) = e^{-r(T-t)} \int_0^\infty \mathcal{P}(S',T,S,t)\Lambda(S')\,dS'.} \tag{3.17}$$

Here,

$$\mathcal{P}(S',T,S,t) := \frac{1}{\sqrt{2\pi\sigma^2(T-t)}S'} \exp\left(-\frac{[\ln(\frac{S'}{S}) - (r-\frac{1}{2}\sigma^2)(T-t)]^2}{2\sigma^2(T-t)}\right). \tag{3.18}$$

This is the transition probability density of an asset price model with growth rate r and volatility σ. In other words, V is the present value (deducted by $e^{-r(T-t)}$) of the expectation of the payoff Λ under an asset price model with volatility σ and growth rate r. It is important to note that it depends on r, not on the growth rate of the underlying asset. We shall come back to this point later.

3.4.4 Special cases

1. **European call option.** The rescaled payoff function for a European call option is

$$\bar{\Lambda}(z) = \max\{e^z - 1, 0\}.$$

Then

$$v(x,\tau) = e^{-r\tau} \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2\tau}} e^{-(x-z-(\frac{1}{2}\sigma^2-r)\tau)^2/(2\sigma^2\tau)} (e^z-1)\,dz. \tag{3.19}$$

This can be integrated. Finally, we get the exact solution for the European call option

$$\boxed{c\,(S,t) = S\mathcal{N}(d_1) - Ee^{-r(T-t)}\mathcal{N}(d_2),} \tag{3.20}$$

where

$$\mathcal{N}(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{z^2}{2}}\,dz, \tag{3.21}$$

$$d_1 = \frac{\log(\frac{S}{E}) + (r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \tag{3.22}$$

$$d_2 = \frac{\log(\frac{S}{E}) + (r-\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}. \tag{3.23}$$

The calculation is done as below. Let us use the following notations for abbreviation.

$$a = r - \sigma^2/2, \quad D = \sigma^2\tau, \quad d_1\sqrt{D} = x + a\tau + D, \quad d_2\sqrt{D} = x + a\tau$$

We have

$$u(x,\tau) = \int_0^\infty \frac{1}{\sqrt{2\pi D}} \exp\left(-\frac{1}{2D}(z-x-a\tau)^2\right)(e^z-1)\,dz = I - II.$$

$$
\begin{aligned}
I &= \int_0^\infty \frac{1}{\sqrt{2\pi D}} \exp\left(-\frac{1}{2D}(z-x-a\tau)^2\right) e^z\, dz \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi D}} \exp\left(-\frac{1}{2D}(z-(x+a\tau+D))^2 + \frac{D^2+2D(x+a\tau)}{2D}\right) dz \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi D}} \exp\left(-\frac{1}{2D}(z-d_1\sqrt{D})^2 + (d_1+d_2)\sqrt{D}/2\right) dz \\
&= e^{(d_1+d_2)\sqrt{D}/2} \int_0^\infty \frac{1}{\sqrt{2\pi D}} \exp\left(-\frac{1}{2D}(z-d_1\sqrt{D})^2\right) dz \\
&= \frac{S}{E} e^{r\tau} \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(z/\sqrt{D}-d_1)^2\right) dz/\sqrt{D} \\
&= \frac{S}{E} e^{r\tau} \int_{-d_1}^\infty \frac{1}{2\pi} e^{-z^2/2}\, dz = \frac{S}{E} e^{r\tau} \mathcal{N}(d_1). \\
II &= \int_0^\infty \frac{1}{\sqrt{2\pi D}} \exp\left(-\frac{1}{2D}(z-x-a\tau)^2\right) dz \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi D}} \exp\left(-\frac{1}{2D}(z-d_2\sqrt{D})^2\right) dz \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(z/\sqrt{D}-d_2)^2\right) dz/\sqrt{D} \\
&= \int_{-d_2}^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = \mathcal{N}(d_2).
\end{aligned}
$$

$$
V(S,t) = Ee^{-r(T-t)}\,(I - II) = S\mathcal{N}(d_1) - Ee^{-r(T-t)}\mathcal{N}(d_2).
$$

Remarks

- We have seen that a lower bound of c is $c \geq S - Ee^{-r(T-t)}$.
- The formula is a modification of the previous lower bound formula. The function $\mathcal{N}$ is the error function. $\mathcal{N}(+\infty) = 1$. The two parameters $d_1 - d_2 = \sigma\sqrt{T-t}$, the standard deviation of the asset price model.

Exercise. Show that

(a) $\mathcal{N}(0) = 1/2, \mathcal{N}(x) \to 1$ as $x \to \infty$;

(b) $\mathcal{N}(x) + \mathcal{N}(-x) = 1$ for any $x \in \mathbb{R}$

Exercise. Prove the formula (3.20).

2. **European put option.** Recall the put-call parity

$$
c + Ee^{-r(T-t)} = p + S.
$$

We can obtain the price for p from c:

$$\boxed{p(S,t) = Ee^{-r(T-t)}\mathcal{N}(-d_2) - S\mathcal{N}(-d_1).} \tag{3.24}$$

Exercise. Prove (3.24).

3. **Forward contract** Recall that a forward contract is an agreement between two parties to buy or sell an asset at certain time in the future for certain price. Its value V also satisfies the B-S equation. The payoff function for such a forward contract is

$$\Lambda(S) = S - E.$$

In terms of the rescaled variable, $\bar{\Lambda}(z) := \Lambda(Ee^z)/E = e^z - 1$. Let $x = \ln(S/E)$, $\tau = T - t$) and $u(x,\tau) = e^{r\tau}V(Ee^x, t)/E$. Then u satisfies the advect heat equation. Its solution with initial data $\bar{\Lambda}(x)$. Its solution is given by

$$\begin{aligned} u(x,\tau) &= \frac{1}{\sqrt{2\pi\sigma^2\tau}}\int_{-\infty}^{\infty} e^{-\frac{(x-z-(r-\sigma^2/2)\tau)^2}{2\sigma^2\tau}}\,(e^z-1)\,dz \\ &= e^{x+r\tau} - 1. \end{aligned}$$

Hence,

$$\boxed{V(S,t) = S - Ee^{-r(T-t)}.} \tag{3.25}$$

This means that the current value of a forward contract is nothing but the difference of S and the discounted E. Notice that this value is independent of the volatility σ of the underlying asset.

Exercise. Show that the payoff function of a portfolio $c - p$ is $S - E$. From this and the Black-Scholes formula (3.16), show the formula of the put-call parity.

4. **Cash-or-nothing.** A contact with cash-or-nothing is just like a bet. If $S_T > E$, then the reward is B. Otherwise, you get nothing. The payoff function is

$$\Lambda(S) = \begin{cases} B & \text{if } S > E \\ 0 & \text{otherwise.} \end{cases}$$

Using the Black-Scholes formula (3.16), we obtain the value of a cash-or-nothing contract to be

$$V(S,t) = Be^{-r(T-t)}\mathcal{N}(d_2). \tag{3.26}$$

Exercise. Verify this formula.

5. **Supershare.** Supershare is a binary option whose payoff function is defined to be

$$\Lambda(S) = \begin{cases} B & \text{if } E_1 < S < E_2 \\ 0 & \text{otherwise.} \end{cases}$$

One can show that the value for this binary option is

$$V(S,t) = Be^{-r(T-t)}\left(\mathcal{N}(d_2(E_1)) - \mathcal{N}(d_2(E_2))\right),$$

where $d_2(E)$ is given by (3.23).

Exercise. Verify this formula.

6. **Deterministic case ($\sigma = 0$).** In this case, the Black-Scholes equation is reduced to

$$V_t + rSV_S - rV = 0.$$

Or in τ, x and u variables:

$$u_\tau - ru_x = 0$$

with initial data

$$u(x,0) = \Lambda(Ee^x)/E,$$

Thus, its solution is given by

$$u(x,\tau) = u(x + r\tau, 0) = \Lambda(Ee^{x+r\tau})/E = \Lambda(Se^{r\tau})/E.$$

Or

$$V(S,t) = e^{-r(T-t)}\Lambda(Se^{r(T-t)}).$$

This means that when the process is deterministic, the value of the option is the payoff function evaluated at the future price of S at T (that is $Se^{r(T-t)}$), and then discounted by the factor $e^{-r(T-t)}$.

3.5 Risk Neutrality

Notice that the growth rate μ of the underlying asset does not appear in the Black-Scholes equation. The option may be valued as if all random walks involved are risk neutral. This means that the drift term (growth rate) μ in the asset pricing model can be replaced by r. The option is then valued by calculating the present value of its expected return at expiry. Recall the lognormal probability density function with growth rate r, volatility σ is

$$\mathcal{P}(S', T, S, t) := \frac{1}{\sqrt{2\pi\sigma^2(T-t)}S'} \exp\left(-\frac{[\ln(\frac{S'}{S}) - (r - \frac{1}{2}\sigma^2)(T-t)]^2}{2\sigma^2(T-t)}\right). \tag{3.27}$$

This is the transition probability density of an asset price model in a *risk-neutral world*:

$$\frac{dS}{S} = rdt + \sigma dz. \tag{3.28}$$

The expected return at time T in this risk-neutral world is

$$\int \mathcal{P}(S', T, S, t)\Lambda(S')dS'.$$

At time t, this value should be discounted by $e^{-r(T-T)}$:

$$V(S,t) = e^{-r(T-t)} \int \mathcal{P}(S', T, S, t)\Lambda(S')dS'.$$

We may reinvestigate the function $\mathcal{N}$ and the parameters d_i in the Black-Scholes formula. After some calculation, we find

$$\mathcal{N}(d_2) = \int_E^\infty \mathcal{P}(S', T, S, t)dS'. \tag{3.29}$$

This is the probability of the event $\{\tilde{S} \geq E\}$, where $\tilde{S}$ obeys the risk-neutral pricing model:

$$\frac{d\tilde{S}}{\tilde{S}} = rdt + \sigma dz.$$

Similarly, one can show that

$$\mathcal{N}(d_1) = \frac{\int_E^\infty \mathcal{P}(S', T, S, t)S'dS'}{Se^{r(T-t)}}. \tag{3.30}$$

is the expectation of $\tilde{S}$ at T when $\tilde{S} = 1$ at t and under the condition that $\tilde{S} \geq E$ at T.
Exercise. Check formulae (3.29) and (3.30).

3.6 Hedging

Hedging is the reduction of sensitivity of a portfolio to the movement of the underlying of asset by taking opposite position in different financial instruments. The Black-Scholes analysis is a dynamical hedging strategy. The delta hedge is instantaneously risk free. It requires a continuous rebalancing of the portfolio and the ratio of the holdings in the asset and the derivative product. The delta (Δ) for a whole portfolio is $\Delta = \frac{\partial \Pi}{\partial S}$. This is the sensitivity of Π against the change of S. By taking $d\Pi - \Delta \cdot dS = 0$, the sensitivity of the portfolio to the asset price change is instantaneously zero.

Besides the delta hedge, there are more sophisticated hedging strategies such as:

$$\begin{aligned}
\text{Gamma: } \Gamma &= \frac{\partial^2 \Pi}{\partial S^2}, \\
\text{Theta: } \Theta &= \frac{\partial \Pi}{\partial t}, \\
\text{Vega: } &= \frac{\partial \Pi}{\partial \sigma}, \\
\text{rho: } \rho &= \frac{\partial \Pi}{\partial r}.
\end{aligned}$$

Hedging against any of these dependencies requires the use of another option as well as the asset itself. With a suitable balance of the underlying asset and other derivatives, hedgers can eliminate the short-term dependence of the portfolio on the movement in t, S, σ, r.

3.6.1 The delta hedging

Suppose a financial institution write $1,000$ European calls on 10,000 shares of a stock. How does this institution hedge the risk from the price fluctuation of the underlying stock? The Delta-hedge is such a strategy. It is a dynamic hedging strategy. The institution long Δ shares of the underlying stock. The Δ is chosen to be

$$\Delta = \frac{\partial c}{\partial S}.$$

If $\Delta = 0.5$, this means that it should long $0.5 \times 10,000 = 5,000$ shares. In this case, the total portfolio $\Pi = c - \Delta S$ is instantaneously neutral. That is $d\Pi = dc - \Delta dS = 0$ under a small change of dS. We call Π is delta neutral. However, this delta neutral holds only for a short period of time. It should be adjusted periodically. This is called a dynamic hedging.

For the Delta-hedge for the European call and put options, we have the following propositions.

Proposition 1. *For European call options, its Δ hedge is given by*

$$\Delta = \mathcal{N}(d_1).$$

Proof. By definition, $\Delta = \partial c/\partial S$. Since $c = S\mathcal{N}(d_1) - Ee^{-r(T-t)}\mathcal{N}(d_2)$, we get

$$\frac{\partial c}{\partial S} = \mathcal{N}(d_1) + S \cdot \mathcal{N}'(d_1) \cdot d_{1S} - Ee^{-r(T-t)}\mathcal{N}'(d_2)d_{2S}.$$

Since

$$d_1 = \frac{\log(\frac{S}{E}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = \frac{\log(\frac{S}{E}) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}},$$

we get

$$d_{1S} = d_{2S} = \frac{1}{S\sigma\sqrt{(T-t)}}, \quad \mathcal{N}'(d_i) = \frac{1}{\sqrt{2\pi}}e^{-d_i^2/2}.$$

Hence,

$$\begin{aligned} \frac{\partial c}{\partial S} &= \mathcal{N}(d_1) + \left(S\mathcal{N}'(d_1) - Ee^{-r(T-t)}\mathcal{N}'(d_2)\right)/(S\sigma\sqrt{T-t}) \\ &\equiv \mathcal{N}(d_1) + \mathcal{I}/(S\sigma\sqrt{T-t}). \end{aligned}$$

We claim that $\mathcal{I} = 0$. Or equivalently,

$$\frac{S}{E}\frac{\mathcal{N}'(d_1)}{\mathcal{N}'(d_2)} = e^{-r\tau}$$

This follows from the computation below.

$$\frac{S}{E}\frac{\mathcal{N}'(d_1)}{\mathcal{N}'(d_2)} = e^x \cdot e^{-(d_1^2-d_2^2)/2}.$$

From (3.22), (3.23),

$$\begin{aligned} d_1^2 - d_2^2 &= \frac{1}{\sigma^2\tau}\left((x + r\tau + \frac{\sigma}{2}\tau)^2 - (x + r\tau - \frac{\sigma}{2}\tau)^2\right) \\ &= 2(x + r\tau) \end{aligned}$$

Hence,

$$\frac{S}{E}\frac{\mathcal{N}'(d_1)}{\mathcal{N}'(d_2)} = e^x \cdot e^{-x-r\tau} = e^{-r\tau}.$$

□

Proposition 2. *For European put options, its Δ hedge is given by*

$$\Delta = \mathcal{N}(-d_1).$$

Proof. From the put-call parity,

$$\Delta = \frac{\partial p}{\partial S} = \frac{\partial c}{\partial S} - 1 = \mathcal{N}(d_1) - 1 = -\mathcal{N}(-d_1),$$

□

3.7 Time-Dependent r, σ and μ for Black-Sholes equation

We can extend Black-Schole analysis to the case when r, σ, μ are functions of t, but still deterministic. We use the change-of-variables:

$$S = Ee^x, \quad V = Ev, \quad \tau = T - t.$$

The Black-Scholes equation is converted to

$$v_\tau = \frac{\sigma^2(\tau)}{2}v_{xx} + (r(\tau) - \frac{\sigma^2(\tau)}{2})v_x - r(\tau)v \tag{3.31}$$

We look for a new time variable $\hat{\tau}$ such that

$$d\hat{\tau} = \sigma^2(\tau)d\tau$$

For instance, we can choose

$$\hat{\tau} = \int_0^\tau \sigma^2(\tau)\,d\tau.$$

Then the equation becomes

$$v_{\hat{\tau}} = \frac{1}{2}v_{xx} + a(\hat{\tau})v_x - b(\hat{\tau})v. \tag{3.32}$$

To handle the term $a(\hat{\tau})$, we consider the following characteristic equation:

$$\frac{dx}{d\hat{\tau}} = -a(\hat{\tau})$$

This can be integrated and yields

$$x = -\int_0^{\hat{\tau}} a(\tau')d\tau' + y,$$

where y is an integration constant. The lines with constant y's are called characteristic curves. Along the characteristic curves, the derivative $\partial_{\hat{\tau}} + a(\hat{\tau})\partial_x$ is just the derivative in time with fixed y. In fact, we introduce the following change-of-variable:

$$\begin{bmatrix} x \\ \hat{\tau} \end{bmatrix} \to \begin{bmatrix} y \\ \hat{\tau}_1 \end{bmatrix} \quad \text{with} \quad \left\{ \begin{array}{l} y = x + \int_0^{\hat{\tau}} a(\tau')d\tau' \equiv x + A(\hat{\tau}) \\ \hat{\tau}_1 = \hat{\tau} \end{array} \right.$$

Then,

$$\frac{\partial}{\partial x}|_{\hat{\tau}} = \frac{\partial y}{\partial x}|_{\hat{\tau}} \frac{\partial}{\partial y} = \frac{\partial}{\partial y},$$

and

$$\frac{\partial}{\partial \hat{\tau}_1}|_y = \frac{\partial \hat{\tau}}{\partial \hat{\tau}_1}\frac{\partial}{\partial \hat{\tau}} + \frac{\partial x}{\partial \hat{\tau}_1}\frac{\partial}{\partial x} = \frac{\partial}{\partial \hat{\tau}} - a(\hat{\tau})\frac{\partial}{\partial x}.$$

The equation (3.32)is transformed to

$$v_{\hat{\tau}_1} = \frac{1}{2}v_{yy} - b(\hat{\tau}_1)v.$$

Let $B(\hat{\tau}_1) = \int_0^{\hat{\tau}_1} b(\tau')d\tau'$ and $u = e^{B(\hat{\tau}_1)}v$, then $u_{\hat{\tau}_1} = \frac{1}{2}u_{yy}$. And we can solve this heat equation explicitly.

3.8 Trading strategy involving options

An investor can design his/her payoff by using options and underlying stocks. The options whose payoff are $\max\{S_T - E, 0\}$ or $\max\{E - S_T, 0\}$ are called vanilla option. They are the simplest payoff functions. In this section, we shall discuss how to design more general payoff functions to fit an investor's anticipation.

3.8.1 Strategies involving a single option and stock

There are four ways:

a. $\Pi = S - c$ (writing a covered call option). When an investor short a call, we say he write a nake call. The investor has cash inflow at the beginning. However, short selling is highly risky. If the stock rises $S_T > E$, then the investor needs to buy the share on S_T and sell to the call holder on E. The investor loses $S_T - E$. This loss can be unlimited because S_T may rise very high. To limit this risk, the investor can long a share S_t at time t. That is, his portfolio is $\Pi = -c + S$. At expiry, if $S_T > E$, then he can use this share to sell to the call owner on E, instead of S_T. His risk is limited. We say that this share covers the call, and say the writer writes a covered call option. The payoff of Π is $\Lambda = S_T - \max(S_T - E, 0) = \min\{S_T, E\}$.

b. $\Pi = c - S$ (reverse of a covered call). Suppose an investor anticipates the stock price will decrease. So he shorts a share and receive money S_t at time t. If S_T does decrease, he earns money. If unfortunately, the stock price increases, then he loses $S_T - S_t$ at time T. This risk is unlimited. Thus, short selling has unlimited risk. However, he can buy a call to limit this risk. That is, his portfolio $\Pi = -S + c$. Since $c < S$, he still receive some money from Π at beginning. At expiry, if the price goes beyond E, then he can exercise this call to cover the shorted share. The payoff now is $\Lambda = -\min\{S, E\} \ge -E$. Thus, his loss is at most E.

c. $\Pi = p + S$ (protective put). When an investor anticipates a stock will increase, he set up a portfolio $\Pi = S$. This costs him S_t. If the prices S_T goes up, he earns $S_T - S_t$. However, if the price goes down, he sells it and loses $S_t - S_T$. If he wants to limit his loss, he can set up a portfolio $\Pi = S + p$. This costs him more at beginning. But at expiry, if the price goes down, he can sell the share on E instead of a lower price S_T. Thus, he limits his loss. In this portfolio, the payoff is $\Lambda = S + \max\{E - S, 0\} = \max\{S, E\}$. The downside is limited (by E) and the upside is unlimited (S_T can be arbitrary large).

d. $\Pi = -p - S$ (reverse of a protective put). In this portfolio, the investor anticipates the stock price will decrease. He shorts a share S_t and also a put. He receives money from shorting these. At expiry, the payoff is $-\max\{S, E\}$. This means that he pays E if $S_T \le E$. However, if S_T is large, then his risk is unlimited.

Below are the payoff functions for the above four cases.

3.8.2 Bull spreads

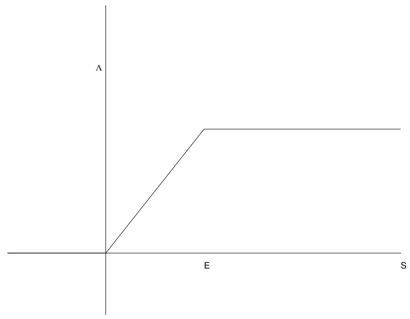
In this strategy, an investor anticipates the stock price will increase. However, he would like to give up some of his right if the price goes beyond certain price, say E_2. Indeed, he does not anticipate the stock price will increase beyond E_2. Therefore he does not want to own a right beyond E_2. Such a portfolio can be designed as

$$\Pi = C_{E_1} - C_{E_2}, \quad E_1 < E_2,$$

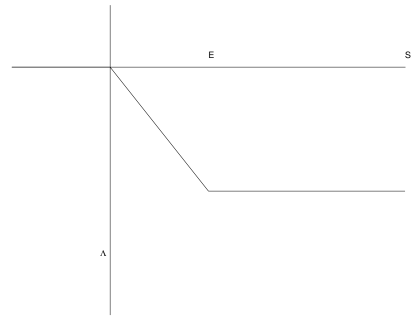
where C_{E_i} is a European call option with exercise price E_i and C_{E_1}, C_{E_2} have the same expiry. The payoff

$$\begin{aligned}\Lambda &= \max\{S_T - E_1, 0\} - \max\{S_T - E_2, 0\} \\ &= \begin{cases} 0 & \text{if } S_T < E_1 \\ S_T - E_1 & \text{if } E_1 < S_T < E_2 \\ E_2 - E_1 & \text{if } S_T > E_2 \end{cases}\end{aligned}$$

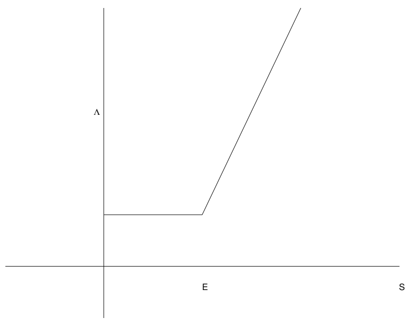
Since $E_1 < E_2$, we have $C_{E_1} > C_{E_2}$. A bull spread, when created from $C_{E_1} - C_{E_2}$, requires an initial investment. We can describe the strategy by saying that the investor has a call option with a strike price E_1 and has chosen to give up some upside potential by selling a call option with strike price $E_2 > E_1$. In return, the investor gets $E_2 - E_1$ if the price goes up beyond E_2.



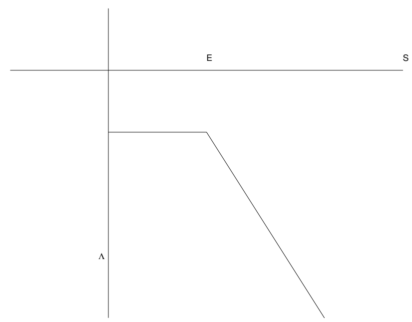
(a)



(b)



(c)



(d)

Example: $C_{E_1} = 3$, $C_{E_2} = 1$ and $E_1 = 30$, $E_2 = 35$. The cost of the strategy is 2. The payoff is

$$\begin{cases} 0 & \text{if } S_T \leq 30 \\ S_T - 30 & \text{if } 30 < S_T < 35 \\ 5 & \text{if } S_T \geq 35 \end{cases}$$

The bull spread can also be created by using put options

$$\Pi = P_{E_1} - P_{E_2}, \quad E_1 < E_2.$$

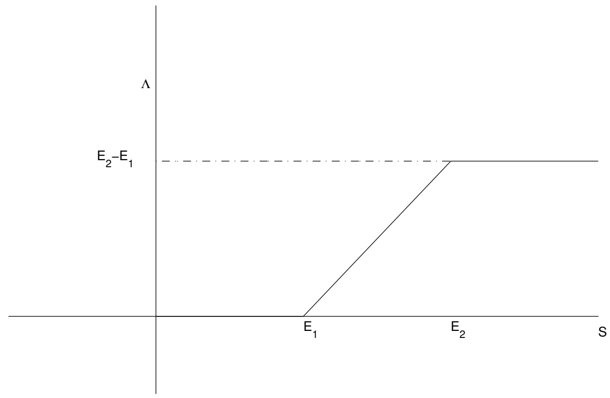
3.8.3 Bear spreads

An investor entering into a bull spread is hoping that the stock price will increase. By contrast, an investor entering into a bear spread is expecting the stock price will go down. The bear spread is

$$\Pi = C_{E_2} - C_{E_1}, \quad E_1 < E_2.$$

There is cash flow entered $(C_{E_2} - C_{E_1})$. The payoff is

$$\begin{aligned} \Lambda &= -\max\{S_T - E_1, 0\} + \max\{S_T - E_2, 0\} \\ &= \begin{cases} 0 & \text{if } S_T < E_1 \\ -S_T + E_1 & \text{if } E_1 < S_T < E_2 \\ -E_2 + E_1 & \text{if } S_T > E_2 \end{cases} \end{aligned}$$



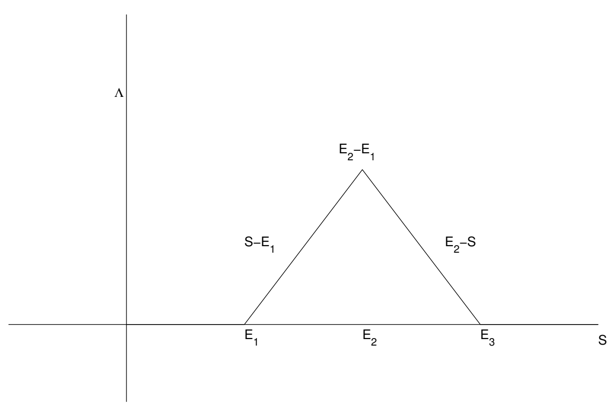
3.8.4 Butterfly spread

If an investor anticipate the stock price will stay in certain region, say, $E_1 < S_T < E_3$, he or she can have a butterfly spread such that the payoff function is positive in that region and he or she gives up the return outside that region.

1. Butterfly spread using calls: Define the portfolio:

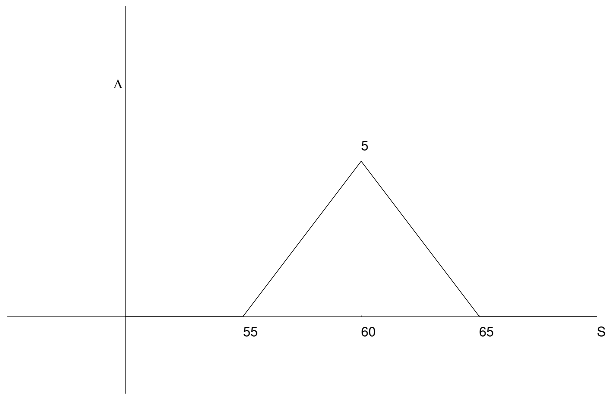
$$\Pi = C_{E_1} - 2C_{E_2} + C_{E_3}, \text{ with } E_1 < E_2 < E_3.$$

where $E_3 = E_2 + (E_2 - E_1)$. Its payoff function is a piecewise linear function and is determined by $\Lambda(E_1) = \Lambda(E_3) = 0$, $\Lambda(E_2) = E_2 - E_1$. Below is the graph of its payoff function.



Example: Suppose a certain stock is currently worth 61. A investor who feels that it is unlikely that there will be significant price move in the next 6 month. Suppose the market of 6 month calls are

E	C
55	10
60	7
65	5



The investor creates a butterfly spread by

$$\Pi = C_{E_1} - 2C_{E_2} + C_{E_3}.$$

The cost is $10 + 5 - 2 \times 7 = 1$. The payoff is the figure below (Figure (1)).

2. Butterfly spread using puts.

$$P_{E_1} + P_{E_3} - 2P_{E_2}, \quad E_1 < E_3, \quad E_2 = \frac{E_1 + E_3}{2}.$$

$$\varphi_{E_2} = \begin{cases} \text{linear} & \\ \Delta E & \text{if } S = E_2 \\ 0 & \text{if } S < E_2 - \Delta E, \text{ or } S > E_2 - \Delta E \end{cases}$$

Remark 1. Suppose European options were available for every possible strike price E, then any payoff function could be created theoretically:

$$\Lambda(S) = \sum \Lambda_i \varphi(S - E_i)$$

where $E_i = i\Delta E$, Λ_i is constant and

$$\varphi(S - E_i) = \begin{cases} 1 & \text{if } S - E_i = 0 \\ 0 & \text{if } |S - E_i| \geq \Delta E \\ \text{linear} & \text{for } |S - E_i| \leq \Delta E. \end{cases}$$

Then $\Lambda(E_i) = \Lambda_i$ and Λ is linear on every interval (E_i, E_{i+1}) and Λ is continuous. As $\Delta E \to 0$, we can approximate any payoff function by using butterfly spreads.

Remark 2. One can also use cash-or-nothing to create any payoff function:

$$\Lambda(S) = \sum \Lambda_i \psi(S - E_i),$$

where

$$\psi(S) := H(S) - H(S - \Delta E) = \begin{cases} 1 & \text{if } 0 \leq S < \Delta E \\ 0 & \text{otherwise.} \end{cases}$$

and H is called the Heaviside function. The value for such a portfolio is

$$\begin{aligned} V &= e^{-r(T-t)} \int \mathcal{P}(S', T, S, t)\Lambda(S')dS', \\ &= e^{-r(T-t)} \Sigma \Lambda_i \mathcal{P}(E_i \le S \le E_{i+1}). \end{aligned}$$

3.9 Derivation of heat equation and its exact solution

3.9.1 Derivation of the heat equation on $\mathbb{R}$

The heat equation on $\mathbb{R}$ reads

$$\begin{cases} u_t = u_{xx}, \quad x \in \mathbb{R}, t > 0, \\ u(x,0) = f(x). \end{cases}$$

This equation models heat propagation on the line. Here, $u(\cdot, t)$ represents the temperature distribution at time t and f the initial temperature distribution. The derivation of this equation is based on a physical law – the conservation of energy, and the Fourier law of heat flux. Given a temperature distribution $u(x,t)$, the Fourier law describes that the temperature flows from high temperature to low temperature at a rate q called heat flux. It is the energy passing through a point (per unit area) per unit time. The Fourier law states that

$$q = -\kappa u_x.$$

Here, the constant κ is called thermal conductivity. It is a positive constant. It differs for different material. For instant, κ of cooper is much larger than that of woods. According to thermal dynamics, the energy density e is related to temperature by $e = c_v u$. The constant c_v is called specific heat (at constant volume). The conservation of energy states that the change of energy in an interval (a,b) per unit time is the same as the heat flux flows into (a,b) from boundaries. Mathematically, it is

$$\partial_t \int_a^b e\, dx = q(a) - q(b) = -\int_a^b q_x\, dx.$$

This gives

$$\int_a^b c_v \partial_t u\, dx = \int_a^b (\kappa u_x)_x\, dx.$$

This holds for any interval (a,b). Thus, the integrands much be the same. This gives

$$\partial_t u = D u_{xx}.$$

where $D = \kappa / c_v > 0$ is the heat conductivity.

3.9.2 Exact solution of heat equation on $\mathbb{R}$.

If we rescale time by $t' = Dt$, then the heat equation becomes

$$\partial_{t'} u = u_{xx}.$$

Thus, without loss of generality, we consider the heat equation with $D = 1$. We shall show that its solution is given by

$$u(x,t) = \int_{-\infty}^{\infty} G(x-y,t) f(y)\, dy, \quad G(x,t) := \frac{1}{\sqrt{4\pi t}} e^{-x^2/(4t)}.$$

The function $G(\cdot, t)$ is the Gaussian distribution. It corresponds to the temperature distribution at time t when the initial temperature distribution is a hot spot at $x = 0$. The derivation of this formulation is decomposed into the following steps.

1. **Superposition principle**: If u and v are solutions, so is $au + bv$, where a and b are constants. This is called superposition principle for linear partial differential equations.
   Now, we imagine f can be approximated by piecewise step function: let h be a small mesh size, $x_j = jh$ and f can be approximated by f_h defined by

   $$f \approx f_h, \quad f_h(x) := \sum_{j=-\infty}^{\infty} f(x_j)\delta_h(x - x_j)h.$$

   Here,

   $$\delta_h(x) := \frac{1}{h}\chi_h(x), \quad \chi_h(x) := \begin{cases} 1 & \text{for } -h/2 < x < h/2 \\ 0 & \text{otherwise.} \end{cases}$$

   Suppose each solution corresponding to $f(x_j)\delta_h(x)$ is $f(x_j)G_h(x - x_j)$. Because the equation is linear, then the general solution is just the linear superposition of these solutions:

   $$u(x,t) \approx \sum_j G_h(x - x_j, t) f(x_j) \cdot h \approx \int_{-\infty}^{\infty} G(x-y,t) f(y)\, dy.$$

   Here, $G_h(x - x_j, t)$ is the solution of the heat equation corresponding to the initial data $\delta_h(x - x_j)$. Notice that we have used another important of heat equation, namely, the translation invariance: if $u(x,t)$ is a solution of the heat equation with initial data $u(x,0)$, then $u(x-a,t)$ is a solution of heat equation with initial data $u(x-a,0)$. This is because the equation is unchanged under the translation transformation $x \mapsto x - a$. Now if $G_h(x,t)$ is the solution with initial data $\delta_h(x)$, then $G_h(x - x_j, t)$ is the solution corresponding to the initial data $\delta_h(x - x_j)$.

2. **Translational invariance.** Because the equation is translationally invariant, we can solve the equation with initial data to be $\chi_h(x)/h$. This initial function tends so-called δ-function as

$h \to 0$. Thus, what we need is to solve the case when the initial function is a δ function. A δ-function is a generalized function having the property:

$$\delta(x) = \begin{cases} 0 & x \neq 0 \\ \infty & x = 0 \end{cases}$$

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1.$$

The δ-function is considered to be the limit of $\delta_h(x)$ as $h \to 0+$ in the sense that

$$\int_{-\infty}^{\infty} f(x)\delta(x)\, dx = \lim_{h\to 0} \int_{-\infty}^{\infty} f(x)\delta_h(x)\, dx = \frac{1}{h}\int_{-h/2}^{h/2} f(x)\, dx = f(0),$$

for any smooth function f.

3. **Parabolic scaling.** To solve the heat equation with an initial hot spot at $x = 0$, we use the parabolic invariance property of the heat equation. Namely, the equation and the initial condition are both invariant under the transformation

$$x \mapsto \lambda x, \quad t \mapsto \lambda^2 t, \quad \lambda > 0.$$

This suggests that the solution is a function of $\xi = x/\sqrt{t}$. So, let us look a solution of the form $u(x,t) = U(x/\sqrt{t})$. Plug it into the equation, we obtain

$$u_t = U'(\xi)\left(-\frac{1}{2}\frac{x}{t^{3/2}}\right), \quad u_{xx} = U''(\xi)\frac{1}{t}.$$

This leads to

$$-\frac{\xi}{2}U'(\xi) = U''(\xi).$$

$$\left(\ln U'(\xi)\right)' = -\frac{\xi}{2}.$$

Integrate it,

$$\ln U'(\xi) = -\frac{\xi^2}{4} + Const., \quad U'(\xi) = Ce^{-\xi^2/4}.$$

Hence

$$U(\xi) = \int_{-\infty}^{\xi} Ce^{-\eta^2/4}\, d\eta.$$

Such a solution has the property: $U(-\infty) = 0$ and $U(+\infty) = C\sqrt{4\pi}$. We choose $C = 1/\sqrt{4\pi}$ for normalization. However, this is still not what we want. But we notice that if $v(x,t) = U(x/\sqrt{t})$ is a solution, so is v_x. Thus, we let

$$G(x,t) = \partial_x U\left(\frac{x}{\sqrt{t}}\right) = \frac{1}{\sqrt{4\pi t}} e^{x^2/(4t)}$$

Then G is a solution and satisfies

$$\int_{-\infty}^{\infty} G(x,t)\, dx = 1.$$

4. **Verification.** Finally, one check

   - $u(x,t) := \int_{-\infty}^{\infty} G(x-y,t)f(y)\,dy$ is a solution of the heat equation
   - $u(x,t) \to f(x)$ as $t \to 0+$ for any bounded continuous function f.

   To show the first one, we check that $G_t = G_{xx}$. This is left to you.
   For the second one, the integral

$$\int_{-\infty}^{\infty} G(x-y,t)f(y)\,dy = \int_{-\infty}^{\infty} G(-y,t)f(x+y)\,dy = \int_{-\infty}^{\infty} G(y,t)f(x+y)\,dy$$

represents the average of f about x with weight $G(y,t)$. We notice that $G(y,t)$ is a Gaussian distribution with standard deviation $\sqrt{t}$. As $t \to 0$, this weight is more and more concentrated at 0. Therefore, this average becomes the average of f roughly in a $\sqrt{t}$-neighborhood of x. Its limit is $f(x)$ for any bounded smooth function f.

We can also give a more rigorous proof. We want to show

$$\lim_{t\to 0} |u(x,t) - f(x)| = \left| \int_{-\infty}^{\infty} G(x-y,t)f(y)\,dy - f(x) \right| = 0.$$

Notice that

$$\int_{-\infty}^{\infty} G(x-y,t)\,dy = 1 \text{ for all } x \in \mathbb{R}, t > 0.$$

Hence

$$\int_{-\infty}^{\infty} G(x-y)f(y)\,dy - f(x) = \int_{-\infty}^{\infty} G(x-y)\left(f(y) - f(x)\right)\,dy.$$

Therefore,

$$|u(x,t) - f(x)| \le \int_{-\infty}^{\infty} G(x-y,t)\left|f(y) - f(x)\right|\,dy.$$

Here, we have used $G(x,t) \ge 0$. To estimate this integral, we break it into two parts

$$\int_{-\infty}^{\infty} G(x-y,t)\left|f(y)-f(x)\right|\,dy = \left(\int_{|x-y|<\delta} + \int_{|x-y|\ge\delta}\right) G(x-y,t)\left|f(y)-f(x)\right|\,dy = I + II,$$

where $\delta > 0$ is a small parameter to be chosen later. Since f is continuous, we have for any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|f(y) - f(x)| < \epsilon, \text{ whenever } |x-y| < \delta.$$

Thus,

$$I = \int_{|x-y|<\delta} G(x-y,t)|f(y)-f(x)|\,dy \le \epsilon \int_{|x-y|<\delta} G(x-y,t)\,dy < \epsilon.$$

For II, since we assume f is bounded, there exists a constant $M > 0$ such that $|f(x)-f(y)| \leq 2M$. Thus,

$$II \leq 2M \int_{|x-y|\geq\delta} G(x-y,t)\,dy = 2M \int_{|y|\geq\delta} G(y,t)\,dy.$$

This integral is the tail estimate of a Gaussian distribution:

$$\int_{|y|\geq\delta} G(y,t)\,dy = 2\int_{y\geq\delta} \frac{1}{\sqrt{4\pi t}} e^{-y^2/(4t)} dy = \int_{z\geq\delta/(2\sqrt{t})} \frac{1}{\sqrt{\pi}} e^{-z^2} dz.$$

As $t \to 0+$, this integral goes to 0. Therefore, we can choose a small $t_1 > 0$ such that

$$II \leq \epsilon \text{ for all } 0 < t < t_1.$$

Combining I and II, we have for any $\epsilon > 0$, there exists $t_1 > 0$ such that for any $0 < t < t_1$ we have $I + II < 2\epsilon$. This shows $I + II \to 0$ as $t \to 0$.

Chapter 4

Variations on Black-Scholes models

4.1 Options on dividend-paying assets

Dividends are payments to the shareholders out of the profits made by the company. We will consider two "deterministic" models for dividend. One has constant dividend yield. The other has discrete dividend payments.

4.1.1 Constant dividend yield

Suppose that in a short time dt, the underlying asset pays out a dividend $D_0 S dt$, where D_0 is a constant, called the dividend yield. This continuous dividend structure is a good model for index options and for short-dated currency options. In the latter case, $D_0 = r_f$, the foreign interest rate.

As the dividend is paid, the return $\frac{dS}{S}$ must fall by the amount of the dividend payment $D_0 dt$. It follows the s.d.e. for the asset price is

$$\frac{dS}{S} = (\mu - D_0)dt + \sigma dz.$$

To value the corresponding option V, we consider the portfolio : $\Pi = V - \Delta S$ as before. In one time step, the change of the portfolio is

$$d\Pi = dV - \Delta dS - \Delta D_0 S dt,$$

where the last term $-\Delta D_0 S dt$ is the dividend our assets received.[1] Thus

$$\begin{aligned} d\Pi &= dV - \Delta(dS + D_0 S dt) \\ &= \sigma S \left(\frac{\partial V}{\partial S} - \Delta \right) dz \\ &\quad + \left((\mu - D_0) S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + V_t - (\mu - D_0)\Delta S - \Delta D_0 S \right) dt \end{aligned}$$

[1] If you own a share, after dt, you will receive $SD_0 dt$ dividend.

$$= \left(V_t - D_0 S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt,$$

Here, we have chosen $\Delta = \frac{\partial V}{\partial S}$ to eliminate the random term. From the absence of arbitrage opportunities, we must have

$$d\Pi = r\Pi dt.$$

Thus,

$$V_t - D_0 S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = r\left(V - S\frac{\partial V}{\partial S} \right).$$

i.e.,

$$\boxed{V_t + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - D_0) S \frac{\partial V}{\partial S} - rV = 0}$$

This is the Black-Scholes equation when there is a continuous dividend payment.

The boundary conditions are:

$$\begin{aligned} c(0,t) &= 0, \\ c(S,t) &\sim Se^{-D_0(T-t)} \end{aligned}$$

The latter is the asset price S discounted by $e^{-D_0(T-t)}$ from the payment of the dividend. The payoff function $c(S,T) = \Lambda(S) = \max\{S - E, 0\}$.

To find the solution, let us consider

$$c(S,t) = e^{-D_0(T-t)} c_1(S,t).$$

Then c_1 satisfies the original Black-Scholes equation with r replaced by $r - D_0$ and the same final condition. The boundary conditions for c_1 are

$$\begin{aligned} c_1(0,t) &= 0, \\ c_1(S,t) &\sim S \text{ as } S \to \infty \end{aligned}$$

Hence,

$$c_1(S,t) = S\mathcal{N}(d_{1,0}) - Ee^{-(r-D_0)(T-t)}\mathcal{N}(d_{2,0})$$

where

$$\begin{aligned} d_{1,0} &= \frac{\ln \frac{S}{E} + (r - D_0 + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \\ d_{2,0} &= d_{1,0} - \sigma\sqrt{T-t}. \end{aligned}$$

The original

$$c(S,t) = Se^{-D_0(T-t)}\mathcal{N}(d_{1,0}) - Ee^{-r(T-t)}\mathcal{N}(d_{2,0}).$$

This means that the call option on a dividend-paying asset is just like a call option in risk-free environment with rate $r - D_0$ and discounted by $e^{-r(T-t)}$.

Remark. We notice that $c \searrow$ as $D_0 \nearrow$. This is natural because paying dividend will lower the value of S, hence c. If the dividend is paid more, the value of call becomes lower.

Exercise. Derive the put-call parity for the European options on dividend-paying assets.

4.1.2 Discrete dividend payments

Suppose our asset pays just one dividend during the life time of the option, say at time t_d. At t_d+, the asset holder receiver a payment $d_y S(t_d-)$. Hence,

$$S(t_d+) = S(t_d-) - d_y S(t_d-) = (1 - d_y)S(t_d-).$$

We claim that across the jumps, V should be continuous, i.e.,

$$V(S(t_d-), t_d-) = V(S(t_d+), t_d+).$$

Otherwise, there is a net loss or gain from buying V before t_d then sell it right after t_d. To find $V(S,t)$, here is a procedure.

1. Solve the Black-Scholes from T to t_d+ to obtain $V(S, t_d+)$ (using the payoff function Λ)
2. Adjusting V by
$$V(S, t_d-) = V((1 - d_y)S, t_d+)$$
3. Solve Black-Scholes equation from t_d to t with the final condition $V((1 - d_y)S, t_d+)$.

For the case of call option, we can find its explicit expression. Let c_d be the European option for this dividend-paying asset. Then

$$c_d(S,t) = c(S,t,E) \text{ for } t_d+ \le t \le T$$

$$c_d(S, t_d-) = c_d(S(1 - d_y), t_d+) = c(S(1 - d_y), t_d, E)$$

Note that

$$c(S(1 - d_y), T, E) = \max\{S(1 - d_y) - E, 0\} = (1 - d_y)\max\{S - (1 - d_y)^{-1}E, 0\}$$

and the linearity of the Black-Scholes equation, we obtain

$$c(S(1 - d_y), t_d, E) = (1 - d_y)c(S, t_d, (1 - d_y)^{-1}E).$$

Therefore,

$$c_d(S,t) = (1 - d_y)c(S, t, (1 - d_y)^{-1}E).$$

That is, the call option price is reduced to its original value with strike price E replaced by $E/(1 - d_y)$.

4.2 Futures and futures options

4.2.1 Forward contracts

Recall that a forward contract is an agreement between two parties to buy or sell an underlying asset on a certain price E at a certain future time T. Here, E is called the delivery price. The payoff function for this forward contract is $\Lambda = S_T - E$. Based on the no arbitrage opportunity, we have seen that the price for this forward contract is (3.25)

$$f = S - Ee^{-r(T-t)}.$$

Definition 4.3. *The forward price F for a forward contract is defined to be the delivery price E which would make that contract have zero value, i.e.,*

$$f_t = 0, \quad \textit{or} \quad F = S_t e^{r(T-t)}.$$

One can take another point of view. Consider a party who shorts the contract. He can borrow an amount of money S_t at time t to buy an asset and use it to close his short position at T. The money he received at expiry, F, is used to pay the loan. If no arbitrage opportunities, then

$$F = S_t e^{r(T-t)}.$$

If the delivery price E of a forward contract is not equal to the forward price F, then the value of this forward contract is $f = S_t - Ee^{-r(T-t)}$. This is the result we obtained in the last chapter. Thus, in order to have $f = 0$, we should take the delivery price to be $S_t e^{r(T-t)}$.

4.2.2 Futures

Futures are very similar to the forward contracts, except they are traded in an exchange, thus, they are required to be standardized. This includes size, quality, price, expiry, ... etc. Let us explain the characters of a future by the following example.

1. Trading future contracts
   - Suppose you call your broker to buy one July corn futures contract (5,000 bushels) on the Chicago Board of Trade (CBOT) at current market price.
   - The broker send this signal to traders on the floor of the exchange.
   - The trader signal this to ask other traders to sell, if no one wants to sell, the trader who represents you will raise the price and eventually find someone to sell
   - Confirmation: Price obtained are sent back to you.
2. Specification of the futures: In the above example, the specification of this future is
   - Asset : quality
   - Contract size: 5,000 bushel

- Delivery arrangement: delivery month is on December
- price quotes
- Daily price movement limits: these are specified by the exchange.
- Position limits: the maximum number of contracts that a speculator may hold.

3. Operation of margins

   - Marking to market: Suppose an investor who contacts his or her broker on June 1, 2016, to buy two December 2016 gold futures contracts on New York Commodity Exchange. We suppose that the current future price is \$400 per ounce. The contract size is \$100 ounces, the investor want to buy \$200 ounces at this price. The broker will require the investor to deposit funds in a "margin account". The initial margin, say is \$2,000 per contract. As the futures prices move everyday, the amount of money in the margin account also changes. Suppose, for example, by the end of June 1, the futures price has dropped from \$400 to \$397. The investor has a loss of \$200×3=600. This balance in the margin account would therefore be reduced by \$600. Maintaining margin needs to deposit. Certain account of money to keep that futures contract.
   - Maintenance margin: To insure the balance in the margin account never becomes negative, a maintenance margin, which is usually lower than the initial margin, is set.

   We tabulate the futures prices, the main/loss in each day. Suppose there are n days to the maturity date. Let $\delta = r(T-t)/n$ be the riskless interest rate per date. Let F_i be the future price at the end of day i. Suppose the margin has money M initially. The gain/loss in each day is the following table.

Day	0	1	2	$\cdots$	$n-1$	n
futures price	F_0	F_1	F_2	$\cdots$	F_{n-1}	$F_n = S_T$
Margin (compounded)	M	Me^{δ}	$Me^{2\delta}$	$\cdots$	$\cdots$	$Me^{n\delta}$
gain/loss	0	$(F_1 - F_0)$	$(F_2 - F_1)$	$\cdots$	$\cdots$	$(F_n - F_{n-1})$
gain/loss compounded to day n	0	$(F_1 - F_0)e^{(n-1)\delta}$	$(F_2 - F_1)e^{(n-2)\delta}$	$\cdots$	$\cdots$	$(F_n - F_{n-1})e^{n\delta}$

   At the end of day i, the margin account should have

$$Me^{i\delta} + \sum_{k=1}^{i}(F_k - F_{k-1})e^{(i-k)\delta}.$$

   This account is required to be kept an amount money greater than certain maintenance fee. If it is lower than this maintenance fee, it will be closed.

Here, the futures prices F_i are the market future prices. What are the futures prices if there is no arbitrage opportunities? We show below that under the no-arbitrage assumption, the future price at time t is indeed identical to the forward price. That is $F_t = S_t e^{r(T-t)}$.

Theorem 4.4. *Forward price and futures price are equal when the interest rates are constant.*

Proof. Suppose a futures contract lasts for n days. Let the future prices are

$$F_0, \cdots, F_n$$

at the end of each business day. Let δ be the risk-free interest rate per day. That is, $\delta = r(T-t)/n$. Consider the following two strategies:

- Strategy 1: Suppose the forward price is G_0.
  - We invest G_0 in a risk-free bond;
  - take a long position of amount $e^{n\delta}$ forward contract.

  At day n, $G_0 e^{n\delta}$ is used to buy the underlying asset at price $S_T e^{n\delta}$. The payoff is $S_T e^{n\delta}$.
- Strategy 2: Suppose the futures price at day 0 is F_0.
  - We invest F_0 amount of money in a risk-free bond;
  - take a long position of $e^{i\delta}$ amount of future at the end of day $i-1$, $i = 1, ..., n$.

Day	0	1	2	$\cdots$	$n-1$	n
futures price	F_0	F_1	F_2	$\cdots$	F_{n-1}	F_n
position	e^{δ}	$e^{2\delta}$	$e^{3\delta}$	$\cdots$	$e^{n\delta}$	0
gain/loss	0	$e^{\delta}(F_1-F_0)$	$e^{2\delta}(F_2-F_1)$	$\cdots$	$\cdots$	$e^{n\delta}(F_n-F_{n-1})$
compound	0	$e^{\delta}(F_1-F_0)e^{(n-1)\delta}$	$e^{2\delta}(F_2-F_1)e^{(n-2)\delta}$	$\cdots$	$\cdots$	$e^{n\delta}(F_n-F_{n-1})$

The total gain/loss from the long position of the futures is

$$\sum_{i=1}^{n}(F_i - F_{i-1})e^{i\delta} \cdot e^{(n-i)\delta} = (F_n - F_0)e^{n\delta} = (S_T - F_0)e^{n\delta}.$$

At time T, we receive $F_0 e^{n\delta}$ from the initial investment F_0. Thus, the total payoff of strategy 2 is

$$F_0 e^{n\delta} + (S_T - F_0)e^{n\delta} = S_T e^{n\delta}.$$

Since both strategies have the same payoff, we conclude their initial investments must be the same, i.e., $F_0 = G_0 = S_T e^{r(T-t)}$. □

4.2.3 Futures options

Options on futures are traded in many different exchanges. They require the delivery of an underlying futures contract when exercised. When a call futures option is exercised, the holder acquires a long position in the underlying futures contract plus a cash amount equal to the current futures price minus the exercise price. Suppose the expiry date of the call is T_1, which is $T_1 \leq T$, the expiration of the underlying future. The payoff of the call option on future is thus

$$\Lambda = \max(F_{T_1} - E, 0).$$

In addition of this payoff, the holder also receive a future contract, yet it costs zero to enter the long position of a future.

Example An investor who has a September futures call option on 25,000 pounds of copper with exercise price $E = 70$ cents/pound. Suppose the current future price of copper for delivery in September is 80 cents/pound. If the option is exercised, the investor received 10 cents $\times 25,000$+long position in futures contract to buy 25,000 pound of copper in September at price 80 cents/pound.

The maturity date of futures option is generally on, or a few days before, the earlist delivery date of the underlying futures contract.

Futures options are more attractive to investors than options on the underlying assets when it is cheaper or more convenient to deliver futures contracts rather than the asset itself. Futures options are usually more liquid and involved lower transaction costs.

4.2.4 Black-Scholes analysis on futures options

Future Price Model As we have seen that the futures price is identical to the forward price when the interest rate is a constant, i.e., $F = Se^{r(T-t)}$. From Itô's lemma, we obtain a pricing model for F:

$$\begin{aligned} dF &= \left(\frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S^2\right) dt + \frac{\partial F}{\partial S} dS \\ &= (-re^{r(T-t)}S)dt + Se^{r(T-t)}\frac{dS}{S} \\ &= (-rF)dt + F(\mu dt + \sigma dz) \\ &= ((\mu - r)F)dt + F\sigma dz. \end{aligned}$$

Hence,

$$\frac{dF}{F} = (\mu - r)dt + \sigma dz. \tag{4.1}$$

We conclude this discussion by the following proposition.

Proposition 3. *The futures price is the same as a stock paying a dividend yield at rate r.*

Next, we study the value V of a futures option. For a call option on a future, it is the right to buy the future on strike price E. Its value V is a function of F, t and E. We apply Black-Sholes analysis to value V.

Futures Option Price Consider a portfolio

$$\Pi = V - \Delta F.$$

As discussed before, we choose $\Delta = \frac{\partial V}{\partial F}$ to eliminate randomness of $d\Pi$. Then

$$\begin{aligned} d\Pi &= dV - \Delta dF \\ &= \left(\frac{\partial V}{\partial F}\mu_F F + \frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial F^2}\sigma^2 F^2\right)dt + \frac{\partial V}{\partial F}\sigma F dz - \frac{\partial V}{\partial F}(\mu_F F dt + \sigma F dz) \\ &= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial F^2}\sigma^2 F^2\right)dt. \end{aligned}$$

Here, $\mu_F := \mu - r$, is the growth rate of the undelying future. Since it costs nothing to enter into a future contract, the cost of setting up the above portfolio is just V. [2] Thus, based on the no arbitrage opportunity,

$$d\Pi = rV dt,$$

and we obtain

$$V_t + \frac{1}{2}\sigma^2 F^2 \frac{\partial^2 V}{\partial F^2} = rV.$$

The payoff function for a call option is $\Lambda = \max\{F - E, 0\}$.

To solve this equation, we recall the option price equation for stock paying dividend is

$$V_t + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - D_0)S\frac{\partial V}{\partial S} - rV = 0$$

In our case, $D_0 = r$, so the futures call option

$$c(F,t) = e^{-r(T-t)} c_1(F,t)$$

where c_1 satisfies Black-Scholes equation with r replaced by $r - r = 0$. This gives

$$\begin{aligned} c_1(F,t) &= F\mathcal{N}(d_1) - E\mathcal{N}(d_2), \\ d_2 &= \frac{\ln\frac{F}{E} - \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}}, \\ d_1 &= d_2 + \sigma\sqrt{T-t} \end{aligned}$$

Notice that this option price V (i.e. c in the present case) is the same as $\tilde{V}(S(F,t),t)$, where $\tilde{V}$ is the solution of the option corresponding to the underlying asset S. That is

$$\begin{aligned} \tilde{V}(S,t) &= S\mathcal{N}(\tilde{d}_1) - Ee^{-r(T-t)}\mathcal{N}(\tilde{d}_2) \\ \tilde{d}_2 &= \frac{\ln\frac{S}{E} + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \\ \tilde{d}_1 &= \tilde{d}_2 + \sigma\sqrt{T-t} \end{aligned}$$

[2] When the futures $F_t = e^{r(T-t)}S_t$, it is the same as the forward contract price. The price f of corresponding contract is 0.

We can write this $\tilde{V}$ in terms of F by $S = Fe^{-r(T-t)}$. Plug this into the above equation to obtain

$$\begin{aligned}\tilde{d}_2 &= \frac{\ln \frac{Fe^{-r(T-t)}}{E} + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ &= \frac{\ln \frac{F}{E} - \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}} = d_2,\end{aligned}$$

Similarly, $\tilde{d}_1 = d_1$. Thus

$$\begin{aligned}\tilde{V}(S(F,t),t) &= S\mathcal{N}(d_1) - Ee^{-r(T-t)}\mathcal{N}(d_2) \\ &= Fe^{-r(T-t)}\mathcal{N}(d_1) - Ee^{-r(T-t)}\mathcal{N}(d_2) \\ &= V(F,t)\end{aligned}$$

We conclude the above discussion by the following proposition.

Proposition 4. *The price for future options is the same as the price for options on the underlying assets.*

Finally, let us find the put-call parity for futures options.

Proposition 5. *Let c and p be the call and put options on a future. Then*

$$c + Ee^{-r(T-t)} = p + Fe^{-r(T-t)}. \tag{4.2}$$

Proof. Consider two portfolios:

$$\begin{aligned}A &= c + Ee^{-r(T-t)} \\ B &= p + Fe^{-r(T-t)} + \text{ a futures contract}\end{aligned}$$

At time T,

$$\begin{aligned}\Lambda_A &= \max\{F_T - E, 0\} + E = \max\{F_T, E\}, \\ \Lambda_B &= \max\{E - F_T, 0\} + F + (F_T - F) = \max\{E, F_T\}.\end{aligned}$$

Hence we obtain $A = B$. $\square$

Chapter 5

Numerical Methods

To evaluate the option price numerically, we need to discretize our option price model. We recall that there are two equivalent versions of option price models:

- The Black-Scholes equation with initial and boundary conditions:

$$V_t + \frac{\sigma^2}{2} S^2 V_{SS} = r(V - SV_S), \quad S > 0,\ 0 \le t \le T,$$
$$V(S,T) = \Lambda(S).$$

- The exact expression in terms of the transition probability of the risk-free asset price model:

$$V(S,0) = e^{-rT} \int_0^\infty \mathcal{P}(S', T, S, 0)\Lambda(S')dS', \tag{5.1}$$

  where $\mathcal{P}$ is the transition probability for the risk-free asset price model

$$d\tilde{S} = r\tilde{S}dt + \sigma\tilde{S}dz, \quad 0 \le t \le T$$
$$\tilde{S}(t) = S.$$

  That is, $\mathcal{P}(S', T, S, 0)$ is the probability density of $\tilde{S}(T)$ with $\tilde{S}(0) = S$.

The first one is a P.D.E. model. The second one is the exact solution of the P.D.E. It can also be viewed as a solution of a stochastic model in a risk-free world. The finite difference method is a numerical procedure for solving the Black-Scholes P.D.E., whereas the Monte-Carlo method is a numerical procedure for the second expression. The binomial method can be derived from both versions.

5.1 Monte Carlo method

We recall that the value of a European option is given by

$$V(S,0) = e^{-rT} \int \mathcal{P}(\tilde{S}, T, S, 0)\Lambda(\tilde{S})d\tilde{S} \tag{5.2}$$

where Λ is the payoff function, $\mathcal{P}$ is the transition probability density function of $\tilde{S}$ which satisfies

$$\frac{d\tilde{S}}{\tilde{S}} = rdt + \sigma dz, \quad \text{(initial state } S(0) = S) \tag{5.3}$$

i.e., it is the asset price model in the risk-neutral world. The Monte Carlo simulation is a numerical procedure to evaluate V numerically based on the random simulation of the distribution of $\tilde{S}(T)$. There are two ways to find an approximation of probability distribution of $\tilde{S}(T)$.

- Use the exact solution formula for $\tilde{S}$. Recall (2.10)

$$\tilde{S}(T) = S\exp\left((r - \frac{\sigma^2}{2})T + \sigma z(T)\right), \tag{5.4}$$

  where $z(T)$ is the Brownian motion. We have seen that

$$z(T) = \sqrt{T}\epsilon, \quad \epsilon \sim \mathcal{N}(0,1).$$

  We can simulate the distribution of $\tilde{S}(T)$ by using this formula. We generate M random numbers ϵ_i with distribution $\mathcal{N}(0,1)$. We approximate

$$\tilde{S}_i(T) := S\exp\left((r - \frac{\sigma^2}{2})T + \sigma\sqrt{T}\epsilon_i\right).$$

- Solving the s.d.e. numerically. We simulate M paths (for example, $M = 10,000$) from (5.3) to obtain $\tilde{S}_i(T), i = 1, ..., M$. To sample M paths from (5.3), we divide the interval $[0,T]$ into N subintervals with equal length $\Delta t = \frac{T}{N}$. We sample MN random numbers ϵ_k^i, $i = 1, ..., M$, $k = 1, \ldots, N$ with distribution $\mathcal{N}(0,1)$ We then discretize (5.3) by using the forward Euler method:

$$\frac{\tilde{S}_i(k\Delta t) - \tilde{S}_i((k-1)\Delta t)}{\tilde{S}_i((k-1)\Delta t)} = r\Delta t + \sigma\epsilon_k^i\sqrt{\Delta t}, \quad k = 1, ..., N, i = 1, ..., M$$

$$\tilde{S}_i(0) = S.$$

  This approach certainly costs more. But it is for general s.d.e.

The option price at time 0 is the expectation of the payoff according the probability distribution of $\tilde{S}_i(T)$, then discounted by e^{-rT}. Thus,

$$V(S,0) \approx e^{-rT}\frac{1}{M}\sum_{i=1}^{M}\Lambda(\tilde{S}_i).$$

Remarks.

1. The error from the numerical integration by Monte-Carlo method is $O\left(\frac{1}{\sqrt{M}}\right)$. This can be proved by the central limit theorem. The error from the numerical integration of the s.d.e. for a sample path by using forward Euler method is $O(\frac{1}{N})$. This can be proved by standard error analysis of numerical ordinary differential equation. Thus, the total error for Monte-Carlo method by using forward Euler integrator for s.d.e. is $O\left(\frac{1}{\sqrt{M}}+\frac{1}{N}\right)$. If there is only one underlying asset, the Monte Carlo does not have any advantage. However, if there are many underlying assets, say more than three, the corresponding Black Scholes equation is a diffusion equation in high dimensions. In this case, finite difference method is very difficult and the Monte Carlo method wins.

2. The calibration of the volatility σ of an asset from market data has been demonstrated in Section 2.9.

Below we provide a C code for Monte-Carlo method by using the exact formula of $\tilde{S}(T)$.

```
Monte_Carlo_European_option(V,S,E,r,sigma,T,M)
{
    double St, sd, R;
    double payoff(), randn();   //randn() is the random number generator obeying normal
    distribution
    int i;
    R = (r-sigma^2 /2)*T;
    sd = sigma * sqrt(T);
    V = 0.;
    for (i=1; i <=  M; i++)
    {
        St = S*exp(R+sd*randn(0,1));
        V += payoff(St,E);
    }
    V *= exp(-r*T)/M;
}
double payoff(S,E)
{
    return max(S-E,0);
}
```

5.2 Binomial Methods

The binomial model is to approximate the risk-neutral asset price model $d\tilde{S}/\tilde{S} = rdt + \sigma dz$ by a binomial model, then find the option price by taking expectation of this binomial model. We first recall the binomial asset model. Then we illustrate how to find option price from this binomial model.

5.2.1 Binomial method for asset price model

We consider the underlying asset is risk-neutral, i,e.,

$$\frac{d\tilde{S}}{\tilde{S}} = rdt + \sigma dz \tag{5.5}$$

We shall approximate this continuous model by the following binomial model.

- We choose an N, define time step size $\Delta t = (T-t)/N$. Define

$$\Delta x = \sigma\sqrt{\Delta t},$$
$$u = e^{\Delta x}, \quad d = e^{-\Delta x}.$$

- Define a random walk: $\{S^0, ...S^N\}$ by

$$\frac{S^{n+1}}{S^n} = \begin{cases} u & \text{with probability } p \\ d & \text{with probability } 1-p \end{cases} \qquad n = 0, ...N-1$$

  with initial price

$$S^0 = S,$$

  where the probability p is determined by

$$p = \frac{e^{r\Delta t} - d}{u - d} = \frac{e^{r\Delta t} - e^{-\Delta x}}{e^{\Delta x} - e^{-\Delta x}}.$$

  The reason how u, d and p are chosen in this form will be illustrated later.

This process can be illustrated by a binomial tree. We define

$$S_j = Se^{j\Delta x}, \quad -N \le j \le N.$$

These are possible prices in every time step up to N. We start from $S^0 = S_0 = S$. The price goes up to $S_1 = Su = Se^{\Delta x}$, or down to $S_{-1} = Sd = Se^{-\Delta x}$. In the second step, the possible prices are

$$Su^2, \quad Sud, \quad Sdu, \quad Sd^2,$$

which are

$$S_{-2}, S_0, S_2.$$

After N steps, the possible values of the asset prices are

$$\{S_j \mid |j| \le N \text{ and } j + N = \text{ even}\}.$$

Each S^n is a random variable. Let P_j^n be the probability of the event that $S^n = S_j$:

$$P_j^n := P(S^n = S_j), \quad |j| \le n, j + n = \text{ even}.$$

Because the initial position is S_0, we have $P_0^0 = 1$. After n step random walks, the probability distribution of $S^n = S_j$ is the binomial distribution:

$$P(S^n = S_j) = P_j^n = \begin{cases} \binom{n}{\ell} p^\ell (1-p)^{n-\ell} & n + j = 2\ell \\ 0 & \text{otherwise.} \end{cases}$$

Here, ℓ is the number of steps that the price goes up in n steps.

This binomial model depends on two parameters: u and p. (The down ratio $d = 1/u$.) They are determined by the conditions so that the discrete model and the continuous model have the same mean and variance in one time step Δt. We recall that these conditions are

$$\begin{aligned} pu + (1-p)d &= e^{r\Delta t} \\ pu^2 + (1-p)d^2 &= e^{(2r+\sigma^2)\Delta t}. \end{aligned}$$

From these two equations, we get respectively

$$p = \frac{e^{r\Delta t} - d}{u - d}$$

$$p = \frac{e^{(2r+\sigma^2)\Delta t} - d^2}{u^2 - d^2}.$$

From these two equations, we get

$$u + d = \frac{e^{(2r+\sigma^2)\Delta t} - d^2}{e^{r\Delta t} - d}.$$

Since $d = u^{-1}$, we get a quadratic equation for u:

$$u^2 - \alpha u + 1 = 0, \quad \text{where} \quad \alpha = \frac{1 + e^{(2r+\sigma^2)\Delta t}}{e^{r\Delta t}}.$$

Solving this equation, we get

$$u = \frac{1}{2}\left[\alpha + \sqrt{\alpha^2 - 4}\right].$$

From this, we recall that d and p can be expressed in terms of u:

$$d = \frac{1}{u}, \quad p = \frac{e^{r\Delta t} - d}{u - d}.$$

Thus, p, u and d can be expressed in terms of r, σ and Δt.

By Taylor expansion, a simple calculation gives

$$u = 1 + \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t + O((\Delta t)^{3/2}) \approx e^{\sigma\sqrt{\Delta t}}$$

Thus, up to $O((\Delta t)^{-3/2})$ error, we can choose

$$u = e^{\sigma\sqrt{\Delta t}},$$

This is why we choose

$$\Delta x = \sigma\sqrt{\Delta t}$$

at the beginning. The probability p is thus given by

$$p = \frac{e^{r\Delta t} - d}{u - d} = \frac{e^{r\Delta t} - e^{-\Delta x}}{e^{\Delta x} - e^{-\Delta x}}.$$

A simple Taylor expansion gives

$$p = \frac{1}{2} + \frac{1}{2}\frac{r\Delta t}{\Delta x} - \frac{1}{4}\Delta x + O(\Delta t + (\Delta x)^2)$$

We may also choose

$$p = \frac{1}{2} + \frac{1}{2}\frac{r\Delta t}{\Delta x} - \frac{1}{4}\Delta x = \frac{1}{2} + \left(r - \frac{\sigma^2}{2}\right)\frac{\Delta t}{2\Delta x}.$$

We should require Δt so that $0 \le p \le 1$. This leads to require

$$\left|r - \frac{\sigma^2}{2}\right| \frac{\Delta t}{\Delta x} \le 1.$$

5.2.2 Binomial method for option

We recall that the option price is the expectation of the payoff $\Lambda(S)$ under the risk neutral probability distribution $\mathcal{P}(\tilde{S}(T), T, S_0, t)$, then deducted by $e^{-r(T-t)}$. This transition probability distribution can be approximated by a binomial distribution P_j^n as shown in the last section. In the discrete model, the risk neutral asset price only takes discrete values S_j, we shall approximate $V(S_j, t + n\Delta t)$ by V_j^n. We notice that: if the random variable $S^n = S_j$, then

$$S^{n+1} = \begin{cases} S_{j+1} & \text{with probability } p \\ S_{j-1} & \text{with probability } 1-p \end{cases}$$

Therefore, the expected value of V at time step n at S_j should satisfy

$$V_j^n = e^{-r\Delta t}\left(pV_{j+1}^{n+1} + (1-p)V_{j-1}^{n+1}\right),$$

the expectation of the binomial distribution discounted by $e^{-r\Delta t}$. Thus, the binomial method for option pricing is

(i) Choose N and define $\Delta t = (T-t)/N$. N is chosen such that $\Delta t \le \sigma^2/r^2$.

(ii) Define $\Delta x = \sigma\sqrt{\Delta t}$, and for each $n = N, ..., 0$, define $S_j = S_0 e^{j\Delta x}$ with $|j| \le N$.

(iii) Define $V_j^N = \Lambda(S_j^N)$, and define V_j^n for $n = N-1, N-2, ..., 0$, by

$$V_j^n = e^{-r\Delta t}\left(pV_{j+1}^{n+1} + (1-p)V_{j-1}^{n+1}\right), \text{ for } |j| \le n,\ j+n = \text{ even.}$$

(iv) Output: V_0^0 is the option price at time t for asset price S_0.

Example. For put option,

$$
\begin{aligned}
T &= 5 \text{ months } = 0.4167 \text{ year} \\
\Delta t &= 1 \text{ months } = 0.0833 \text{ year} \\
r &= 0.1, \quad \sigma = 0.4 \\
S_0 &= \$50, \quad E = \$50 \\
u &= e^{\sigma\sqrt{\Delta t}} = 1.1224 \\
d &= 0.8909 \\
p &= 0.5076 \\
e^{r\Delta t} &= 1.0084
\end{aligned}
$$

We begin to generate a binomial tree from $S_0 = 50$ consisting of $S_j^n = S_0 u^\ell d^{n-\ell}$, where $n+j = 2\ell$, $-n \le j \le n$. Then we compute V_j^n inductively from $n = N-1$ to $n = 0$ by

$$V_j^n = e^{-r\Delta t}\left(pV_{j+1}^{n+1} + (1-p)V_{j-1}^{n+1}\right).$$

with V_j^N being the payoff function $\Lambda(S_j^N)$. The value V_0^0 is our answer.
A C code for European option is given below. In computer code, we use the data structure

- $S[n][m]$, $V[n][m]$, $m = 0, ..., n$, $n = 0, ..., N$. Before, we use S_j^n, where $n + j =$ even and $|j| \le n$. The relation between j and m is $m = (j+n)/2$.
- The relation
$$V_j^n = e^{-r\Delta t}\left(pV_{j+1}^{n+1} + (1-p)V_{j-1}^{n+1}\right)$$
is replaced by
$$V_m^n = e^{-r\Delta t}\left(pV_{m+1}^n + (1-p)V_m^n\right).$$

```
Binomial_European_option(V,S,S0,E,r,sigma,dt,N)
{
    double u,d,p;
    double discount, tmp;
    int m,n;
    discount = exp(-r*dt);
    u = exp(sigma*sqrt(dt));
    d = 1/u;
    p = (1/discount - d)/(u-d);

    //Binomial tree for asset price: at level n.
    S[0][0]=S0;
    for (n = 1; n <= N; ++n)
    {
        for (m=n; m > 0; --m)
            S[n][m] = u*S[n-1][m-1];
        S[n][0] = d*S[n-1][0];
    }

    // Binomial tree for option.
    for (m=0; m <=N; ++m)
```

```
        V[N][m] = payoff(S[N][m],E);
    for ( n = N-1; n >= 0; --n)
    {
        for (m=0; m < n; ++m)
            V[n][m] = discount*(p*V[n+1][m+1] + (1-p)*V[n+1][m]);
    }
}
```

5.3 Finite difference methods (for the modified B-S eq.)

In this section, we shall solve the Black-Scholes equation by finite difference methods. Recall that the Black-Scholes equation is

$$V_t + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = r(V - S\frac{\partial V}{\partial S}).$$

We set up the initial time is 0 and reserve the symbol t for time parameter. Using dimensionless variables $S = Ee^x$, $V = Ev$, $\tau = T - t$, we have

$$v_\tau = \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial x^2} + (r - \frac{1}{2}\sigma^2)\frac{\partial v}{\partial x} - rv, \quad \tau \in [0, T].$$

Let $v = e^{-r\tau}u$, then u satisfies

$$u_\tau = \frac{1}{2}\sigma^2 \frac{\partial^2 u}{\partial x^2} + (r - \frac{1}{2}\sigma^2)\frac{\partial u}{\partial x}, \quad \tau \in [0, T]. \tag{5.6}$$

The initial condition for u is

$$u(x, 0) = \begin{cases} \max\{e^x - 1, 0\} & \text{for call option} \\ \max\{1 - e^x, 0\} & \text{for put option} \end{cases}$$

The far field boundary condition for u is

$$u(-\infty, t) = 0, \quad u(x, t) = e^x e^{r\tau} \text{ as } x \to \infty$$

for a call option, and

$$u(-\infty, t) = 1, \qquad u(x, t) = 0 \text{ as } x \to \infty$$

for a put option.

5.3.1 Discretization methods

To solve (5.6) numerically, we follow the following procedure:

- Discretize space and time. We discretize $[0, T]$ into N steps, $\Delta\tau = T/N$. We choose a small proper Δx (to be determined later) and define $x_j = j\Delta x, j = -N, ..., N$. We shall approximate $u(x_j, n\Delta\tau)$ by U_j^n.

- Spatial discretization. We shall approximate the spatial derivatives by finite differences:
  - u_x is approximated by
$$u_x \leftarrow \frac{u_{j+1} - u_{j-1}}{2\Delta x}.$$
  - u_{xx} is approximated by
$$u_{xx} \leftarrow \frac{u_{j+1} - 2u_j + u_{j-1}}{(\Delta x)^2}.$$

  Then the right-hand-side of (5.6) is discretized into
$$(QU)_j \equiv (\frac{\sigma^2}{2})\frac{U_{j+1} - 2U_j + U_{j-1}}{(\Delta x)^2} + (r - \frac{\sigma^2}{2})\frac{U_{j+1} - U_{j-1}}{2\Delta x}$$

- Temporal discretization. For the temporal discretization, we introduce the following three methods:
  - Forward Euler method:
$$\frac{U_j^{n+1} - U_j^n}{\Delta\tau} = (QU^n)_j.$$
  - Backward Euler method:
$$\frac{U_j^{n+1} - U_j^n}{\Delta\tau} = (QU^{n+1})_j.$$
  - Crank-Nicolson method:
$$\frac{U_j^{n+1} - U_j^n}{\Delta\tau} = \frac{1}{2}[(QU^{n+1})_j + (QU^n)_j].$$

- Boundary conditions: we impose
$$U_{-N}^n = 0, \quad U_N^n = e^{N\Delta x}e^{rn\Delta t}$$
  for call option.

For forward Euler method, we find U_j^n inductively in n for $n = 1, ..., N$. We start from $n = 0$, which is the initial condition. We are given U_j^0, $j = -N, ..., N$. From n to $n+1$, the boundary values U_{-N}^{n+1} and U_N^{n+1} are assigned. For $-N < j < N$, we find U_j^{n+1} by
$$U_j^{n+1} := U_j^n + \Delta\tau Q(U^n)_j.$$
We stop this recursive process until all U_j^N are found.

For other methods, which involves $Q(U^{n+1})$ in the equation of U^{n+1}. They can be expressed as a linear system for U^{n+1}. There are standard method to solve such linear system. We shall discuss this issue later.

5.3.2 Binomial method is a forward Euler finite difference method

We claim that the forward Euler method for u in backward time τ is indeed the binomial method for option price V if we choose Δx properly. Let us use V_j^n to approximate the solution $V(x_j, n\Delta t)$, where $\Delta t = \Delta\tau = T/N$. We notice that the time variable for V is the forward time t and the time variable for u and v is the backward time τ. Thus, $V(x_j, n\Delta t) = v(x_j, (N-n)\Delta\tau)$ and $v = e^{-r\tau}u$. We then have

$$V_j^n = e^{-r(N-n)\Delta t} U_j^{N-n}.$$

For the forward Euler method, we rewrite it in terms of V and also

$$U_j^{N-n} = U_j^{N-n-1} + \Delta\tau(QU^{N-n-1})_j$$

In terms of V_j^n, we have

$$\begin{aligned} e^{r\Delta t} V_j^n &= V_j^{n+1} + \Delta t Q(V^{n+1})_j \\ &= V_j^{n+1} + \frac{1}{2}\frac{\Delta t}{(\Delta x)^2}\sigma^2(V_{j+1}^{n+1} - 2V_j^{n+1} + V_{j-1}^{n+1}) + (r - \frac{\sigma^2}{2})\frac{\Delta t}{2\Delta x}(V_{j+1}^{n+1} - V_{j-1}^{n+1}) \\ &= aV_{j+1}^{n+1} + bV_j^{n+1} + cV_{j-1}^{n+1} \end{aligned}$$

where

$$\begin{aligned} a &= \frac{1}{2}\frac{\Delta t}{(\Delta x)^2}\sigma^2 + (r - \frac{\sigma^2}{2})\frac{\Delta t}{2\Delta x} \\ b &= (1 - \frac{\Delta t}{(\Delta x)^2}\sigma^2)V_j^{n+1}, \\ c &= \frac{1}{2}\frac{\Delta t}{(\Delta x)^2}\sigma^2 - (r - \frac{\sigma^2}{2})\frac{\Delta t}{2\Delta x}. \end{aligned}$$

Notice that

$$a + b + c = 1.$$

We should require

$$a, b, c \geq 0,$$

for stability reason to be explained later. Thus V_j^n is the "average" of V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1} with weights a, b, c, then discounted by $e^{-r\Delta t}$.

The stability condition $a, b, c \geq 0$ reads

$$\left| r - \frac{\sigma^2}{2} \right| \frac{\Delta t}{\Delta x} \leq \frac{\Delta t}{(\Delta x)^2}\sigma^2 \leq 1$$

These give constraints on Δx and Δt:

$$\Delta x \leq \frac{1}{\left|\frac{r}{\sigma^2} - \frac{1}{2}\right|}, \quad \Delta t \leq \frac{(\Delta x)^2}{\sigma^2}.$$

Binomial case: $b = 0$ When we choose

$$\Delta x = \sigma\sqrt{\Delta t},$$

then $b = 0$. In this case,

$$e^{r\Delta t}V_j^{n+1} = pV_{j+1}^{n+1} + (1-p)V_{j-1}^{n+1}$$

where

$$p = a = \frac{1}{2} + (r - \frac{\sigma^2}{2})\frac{\Delta t}{2\Delta x}.$$

The stability condition is satisfied if and only if

$$0 \le p \le 1. \tag{5.7}$$

We see that this finite difference is identical to the binomial method in the previous section.

5.3.3 Stability

Definition 5.4. *A finite difference method is called consistent to the corresponding P.D.E. if for any solution of the corresponding P.D.E., it satisfies*

$$\textit{F.D.E (finite difference equation)} \; + \epsilon(\Delta x, \Delta t)$$

and $\epsilon \to 0$ *as* $\Delta t, \Delta x \to 0$

Definition 5.5. *The truncation error of a finite difference method is defined to be the function* $\epsilon(\Delta x, \Delta t)$ *in the previous definition.*

For instance, the truncation error for central difference is

$$Qu = \frac{\sigma^2}{2}u_{xx} + (r - \frac{\sigma^2}{2})u_x + O((\Delta x)^2).$$

And the truncation for various temporal discretizations are

1. Forward Euler:
$$\frac{u(j\Delta x, (n+1)\Delta t) - u(j\Delta x, n\Delta t)}{\Delta t} - (Qu)(j\Delta x, n\Delta t) = O((\Delta x)^2) + O(\Delta t).$$

2. Backward Euler:
$$\frac{u(j\Delta x, (n+1)\Delta t) - u(j\Delta x, n\Delta t)}{\Delta t} - (Qu)(j\Delta x, (n+1)\Delta t) = O((\Delta x)^2) + O(\Delta t).$$

3. Crank-Nicolson method
$$\begin{aligned} & \frac{u(j\Delta x, (n+1)\Delta t) - u(j\Delta x, n\Delta t)}{\Delta t} - \frac{1}{2}\left[(Qu)(j\Delta x, (n+1)\Delta t) + (Qu)(j\Delta x, n\Delta t)\right] \\ = \; & O((\Delta x)^2) + O((\Delta t)^2). \end{aligned}$$

The true error $U_j^n - u(j\Delta x, n\Delta t)$ is usually estimated in terms of the truncation error.

Definition 5.6. *A finite difference equation is said to be (L^2-)stable if the norm*

$$\|U^n\|^2 := \sum_j |U_j^n|^2 \Delta x$$

is bounded for all $n \geq 0$.

Definition 5.7. *A finite difference method for a P.D.E. is convergent if its solution U_j^n converges to the solution $u(j\Delta x, n\Delta t)$ of the corresponding P.D.E..*

Theorem 5.5 (Lax)**.** *For linear partial differential equations, a finite difference method is convergent if and only if it is consistent and stable.*

This theorem is standard and its proof can be found in most numerical analysis text book. We therefore omit it here.

Since the consistency is easily to achieve, we shall focus on the stability issue. A standard method to analyze stability issue is the von Neumann stability analysis. It works for P.D.E. with constant coefficients. It also works "locally" and serves as a necessary condition for linear P.D.E. with variable coefficients and nonlinear P.D.E.. We describe his method below.

We take Fourier transform of $\{U_j\}_{j=-\infty}^{\infty}$ by defining

$$\hat{U}(\xi) = \sum_{j=-\infty}^{\infty} U_j e^{-ij\xi}$$

It is a well-known fact that

$$\begin{aligned} \sum_j |U_j|^2 &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} |\hat{U}(\xi)|^2 \, d\xi \\ &\equiv \|\hat{U}\|^2 \end{aligned}$$

Thus, the boundedness of $\sum_j |U_j|^2$ can be estimated by using $\|\hat{U}\|^2$. The advantage of using $\hat{U}$ is that the finite difference operation becomes a multiplier in terms of $\hat{U}$. Namely,

$$\begin{aligned} \widehat{DU}(\xi) &= \sum_j \left(\frac{U_{j+1} - U_{j-1}}{2} \right) e^{-ij\xi} \\ &= \sum_j \left(\frac{U_j e^{i(j+1)\xi} - U_j e^{i(j-1)\xi}}{2} \right) \\ &= \left(\frac{e^{i\xi} - e^{-i\xi}}{2} \right) \sum_j U_j e^{-ij\xi} \\ &= (2i \sin \xi) \hat{U}(\xi) \end{aligned}$$

For the finite difference operator QU,m we have

$$\begin{aligned}\widehat{(QU)} &= \sum_j (Qu)_j e^{-ij\xi} \\ &= [\frac{\sigma^2}{2}\frac{1}{(\Delta x)^2}(2cos\xi - 2) + (r - \frac{\sigma^2}{2})\frac{1}{\Delta x}(2isin\xi)]\hat{U} \\ &\equiv \widehat{Q}(\xi)\hat{U}(\xi).\end{aligned}$$

For forward Euler method,

$$\begin{aligned}\widehat{U^{n+1}}(\xi) &= (1 + \Delta t\widehat{Q}(\xi))\widehat{U^n} \\ &= G(\xi)\widehat{U^n} \\ &= G(\xi)^{n+1}\widehat{U^n}\end{aligned}$$

We observe that

$$\begin{aligned}\int_{-pi}^{\pi} |\widehat{U^n}(\xi)|^2\, d\xi &= \int |G(\xi)|^{2n}\widehat{U^0}(\xi)|^2\, d\xi \\ &\leq \max_{\xi\in(-\pi,\pi)} |G(\xi)|^{2n} \int |\widehat{U^0}(\xi)|^2\, d\xi\end{aligned}$$

If $|G(\xi)| \leq 1, \rangle\forall\xi \in (-\pi, \pi)$, then stability condition holds. On the other hand, if $|G(\xi)| > 1$ at some point ξ_0, then by the continuity of G, we have that

$$|G(\xi)| \geq 1 + \epsilon$$

for some small $\epsilon > 0$ and for all ξ with $|\xi - \xi_0| \leq \delta$ for some $\delta > 0$. Let consider an initial condition such that

$$\widehat{U^0}(\xi) = \begin{cases} 1 & |\xi - \xi_0| \leq \delta \\ 0 & \text{otherwise.} \end{cases}$$

Then the corresponding $\widehat{U^n}$ will have

$$\begin{aligned}\int_{-pi}^{\pi} |\widehat{U^n}(\xi)|^2\, d\xi &= \int |G(\xi)|^{2n}\widehat{U^0}(\xi)|^2\, d\xi \\ &\to \infty\end{aligned}$$

as $n \to \infty$. We conclude the above discussion by the following theorem.

Theorem 5.6. *For a finite difference equation with constant coefficients, suppose its fourier transform satisfies'*

$$\widehat{U^{n+1}}(\xi) = G(\xi)\widehat{U^n}(\xi)$$

Then the finite difference equation is stable if and only if

$$|G(\xi)| \leq 1\ \forall\xi \in (-\pi, \pi].$$

Example: Let apply the forward Euler method for the heat equation: $u_t = u_{xx}$. Then

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \frac{1}{(\Delta x)^2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n), \quad \Longrightarrow U^{n+1} = U^n + \frac{\Delta t}{(\Delta x)^2} D^2 U^n$$

From von Neumann analysis:

$$\begin{aligned} \widehat{U^{n+1}} &= [1 + \frac{\Delta t}{(\Delta x)^2}(2\cos\xi - 2)]\widehat{U^n} \\ &= (1 - 4\frac{\Delta t}{(\Delta x)^2}\sin^2\frac{\xi}{2})\widehat{U^n} \\ &\equiv G(\xi)\widehat{U^n}. \end{aligned}$$

Hence

$$|G(\xi)| \leq 1 \Longrightarrow \frac{\Delta t}{(\Delta x)^2}\sin^2\frac{\xi}{2} \leq \frac{1}{2} \Longrightarrow \frac{\Delta t}{(\Delta x)^2} \leq \frac{1}{2} \text{ (stability condition)}$$

If we rewrite the finite difference scheme by

$$\begin{aligned} U_j^{n+1} &= \frac{\Delta t}{(\Delta x)^2}U_{j+1}^n + (1 - 2\frac{\Delta t}{(\Delta x)^2})U_j^n + \frac{\Delta t}{(\Delta x)^2}U_{j-1}^n \\ &\equiv aU_{j+1}^n + bU_j^n + cU_{j-1}^n. \end{aligned}$$

Then the stability condition is equivalent to

$$a, b, c \geq 0.$$

Since we have $a + b + c = 1$ from the definition, thus we see that the finite difference scheme is nothing but saying U_j^{n+1} is the average of U_{j+1}^n, U_j^n and U_{j-1}^n with weights a, b, c. In particular, if we choose $\frac{\Delta t}{(\Delta x)^2} = \frac{1}{2}$, then $b = 0$. If we rename $a = p$, $c = 1 - p$, then $U_j^{n+1} = pU_{j+1}^n + (1 - p)U_{j-1}^n$. This can be related to the random walk as the follows.

Consider a particle move randomly on the grid points $j\Delta x$. In one time step, the particle moves toward right with probability p and left with probability $1 - p$. Let U_j^n be the probability of the particle at $j\Delta x$ at time step n for a random walk.

$$U_j^{n+1} = pU_{j-1}^n + (1 - p)U_{j+1}^n.$$

We can also apply the above stability analysis to backward Euler method and the Crank-Nicolson method. Let us only demonstrate the analysis for the heat equation. We left the analysis for the Black-Scholes equations as exercises.

1. For backward Euler method,

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \frac{1}{(\Delta x)^2}(U_{j-1}^{n+1} - 2U_j^{n+1} + U_{j+1}^{n+1}).$$

Then

$$(1 + (4\sin^2\frac{\xi}{2})\frac{\Delta t}{(\Delta x)^2})\widehat{U^{n+1}} = \widehat{U^n}\rangle \Longrightarrow\rangle \widehat{U^{n+1}} = G(\xi)\widehat{U^n},$$

where

$$G(\xi) = \frac{1}{1 + 4\frac{\Delta t}{(\Delta x)^2}(\sin^2\frac{\xi}{2})}.$$

We find that $|G(\xi)| \le 1$, for all ξ. Hence, the backward Euler method is always stable.

2. For Crank-Nicolson method,

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \frac{1}{2(\Delta x)^2}\left[(U_{j+1}^{n+1} - 2U_j^{n+1} + U_{j-1}^{n+1}) + (U_{j+1}^n - 2U_j^n + U_{j-1}^n)\right].$$

Its Fourier transform satisfies

$$\frac{\widehat{U^{n+1}} - \widehat{U^n}}{\Delta t} = \frac{1}{2(\Delta x)^2}\left[-(4\sin^2\frac{\xi}{2})\widehat{U^{n+1}} - (4\sin^2\frac{\xi}{2})\widehat{U^n}\right].$$

We have

$$(1 + 2\frac{\Delta t}{(\Delta x)^2}(\sin^2\frac{\xi}{2}))\widehat{U^{n+1}} = (1 - 2\frac{\Delta t}{(\Delta x)^2}(\sin^2\frac{\xi}{2}))\widehat{U^n}$$

and hence

$$\widehat{U^{n+1}} = \frac{1 - 2\frac{\Delta t}{(\Delta x)^2}\sin^2\frac{\xi}{2}}{1 + 2\frac{\Delta t}{(\Delta x)^2}\sin^2\frac{\xi}{2}}\widehat{U^n}.$$

Let $\alpha = 2\frac{\Delta t}{(\Delta x)^2}\sin^2\frac{\xi}{2}$, then $G(\xi) = \frac{1-\alpha}{1+\alpha}$. We find that for all $\alpha \ge 0$, $|G(\xi)| \le 1$, hence Crank-Nicolson method is always stable for all Δt, $\Delta x > 0$.

Exercise study the stability criterion for the modified Black-Scholes equation

$$u_\tau = \frac{\sigma^2}{2}u_{xx} + (r - \frac{\sigma^2}{2})u_x,$$

for the forward Euler method, back Euler method and Crank-Nicolson method.

5.3.4 Convergence

Let us study the convergence for finite difference schemes for the modified Black-Scholes equation. Let us take the forward Euler scheme as our example. The method below can also be applied to other scheme.

The forward Euler scheme is given by:

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = (QU^n)_j$$

We have known that it has first-order truncation error, namely, suppose $u_j^n := u(j\Delta x, n\Delta t)$, where u is the solution of the modified Black-Scholes equation, then

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = (Qu^n)_j + O(\Delta t) + O((\Delta x)^2).$$

We subtract the above two equations, and let e_j^n denotes for $u_j^n - U_j^n$ and ϵ_j^n denotes for the truncation error. Then we obtain

$$\frac{e_j^{n+1} - e_j^n}{\Delta t} = (Qe^n)_j + \epsilon_j^n$$

Or equivalently,

$$e_j^{n+1} = ae_{j+1}^n + be_j^n + ce_{j-1}^n + \Delta t \epsilon_j^n \tag{5.8}$$

Here, $a, b, c \geq 0$ and $a + b + c = 1$. We can take Fourier transformation $\widehat{e^n}$ of e^n. It satisfies

$$\widehat{e^{n+1}}(\xi) = G(\xi)\widehat{e^n}(\xi) + \Delta t \widehat{\epsilon^n}(\xi)$$

where

$$G(\xi) = ae^{i\xi} + b + ce^{-i\xi}.$$

Recall that the stability $|G(\xi)| \leq 1$ is equivalent to $a, b, c \geq 0$. Thus, by applying the above recursive formula, we obtain

$$\begin{aligned}
\|\widehat{e^n}\| &\leq \|\widehat{e^{n-1}}\| + \Delta t \|\widehat{\epsilon^{n-1}}\| \\
&\leq \|\widehat{e^{n-2}}\| + \Delta t \left(\|G\widehat{\epsilon^{n-2}}\| + \|\widehat{\epsilon^{n-1}}\| \right) \\
&\leq \|\widehat{e^{n-2}}\| + \Delta t \left(\|\widehat{\epsilon^{n-2}}\| + \|\widehat{\epsilon^{n-1}}\| \right) \\
&\leq \|\widehat{e^0}\| + \Delta t \sum_{k=0}^{n-1} \|\widehat{\epsilon^k}\| \\
&\leq O(\Delta t) + O((\Delta x)^2).
\end{aligned}$$

Here, we have used the estimate for the truncation error

$$\|\epsilon^n\| = O(\Delta t) + O((\Delta x)^2,$$

and that $n\Delta t = O(1)$. We conclude the error analysis as the following theorem.

Theorem 5.7. *The error $e_j^n := u(j\Delta x, n\Delta t) - U_j^n$ for the Euler method has the following convergence rate estimate:*

$$(\sum_j |e_j^n|^2 \Delta x)^{1/2} \leq O(\Delta t) + O((\Delta x)^2), \textit{ for all } n.$$

It is simpler to ontain the maximum norm estimate. Let $E(n) := \max_j |e_j^n|$ be the maximum error. From (5.8), we have

$$\begin{aligned} |e_j^{n+1}| &\le a|e_{j+1}^n| + b|e_j^n| + c|e_{j-1}^n| + \Delta t|\epsilon_j^n| \\ &\le aE(n) + bE(n) + cE(n) + \Delta t\epsilon \\ &= E(n) + \Delta t\epsilon \end{aligned}$$

where

$$\epsilon := \max_{j,n} |\epsilon_j^n| = O(\Delta t) + O((\Delta x)^2).$$

Hence,

$$E(n+1) \le E(n) + \Delta t\epsilon.$$

Since we take $U_j^0 = u_j^0$, there is no error initially. Hence, we have

$$\begin{aligned} E(n) &\le \sum_{k=0}^{n-1} \Delta t\epsilon \\ &\le n\Delta t\epsilon \end{aligned}$$

Since $n\Delta t$ is a fixed number, as we take the limit $n \to \infty$, we obtain the error is bounded by the truncation error. We summarize the above discussion as the following theorem.

Theorem 5.8. *The error $e_j^n := u(j\Delta x, n\Delta t) - U_j^n$ for the Euler method has the following convergence rate estimate:*

$$\max_j |e_j^n| \le O(\Delta t) + O((\Delta x)^2).$$

Exercise. Prove that the true error of the Crank-Nicolson scheme is $O((\Delta t)^2) + O((\Delta x)^2)$.

5.3.5 Boundary condition

For the modified Black-Scholes equation, we have

$$u(-\infty, t) = 0, \quad u(x,t) = e^x e^{r\tau} \text{ as } x \to \infty$$

for a call option, and

$$u(-\infty, t) = 1, \qquad u(x,t) = 0 \text{ as } x \to \infty$$

In computation, we can choose a finite domain (x_L, x_R) with $x_L << -1$ and $x_R >> 1$. The boundary condition at the boundary points are an approximation to the above far field boundary condition.

In practice, we don't even use this boundary condition. Indeed, if we want to know $u(x_k, n\Delta t)$, we can find the numerical domain of this quantity, which is the triangle

$$\{(j\Delta x, m\Delta t) \mid |j-k| \le n-m\}$$

We only need to compute u in this domain, which needs no boundary data.

5.4 Converting the B-S equation to finite domain

The transformation $x = \log(S/E)$ converts the B-S equation to a heat equation. However, the domain of x is the whole real line. For numerical computation, it is desirable to have a finite computation domain. The transformation in this section converts S to ξ with $\xi \in (0,1)$. The price is that the resulting equation has variable coefficients. But this is not a problem for numerical computation.

We define the transformation:

$$\xi = \frac{S}{S+E} \tag{5.9}$$

$$\overline{V} = \frac{V(S,t)}{S+E} \tag{5.10}$$

$$\tau = T - t. \tag{5.11}$$

Notice that ξ is dimensionless and important values of ξ are near $1/2$. With this, the inverse transformation is

$$S = \frac{E\xi}{1-\xi}, \ \frac{d\xi}{dS} = \frac{(1-\xi)^2}{E}$$

We plug this transformation to the B-S equation. We allow σ depend on S. Define $\bar{\sigma}(\xi) = \sigma(E\xi/1-\xi)$. Then the resulting equation is

$$\frac{\partial \overline{V}}{\partial \tau} = \frac{1}{2}\bar{\sigma}^2(\xi)\xi^2(1-\xi)^2\frac{\partial^2 \overline{V}}{\partial \xi^2} + r\xi(1-\xi)\frac{\partial \overline{V}}{\partial \xi} - r(1-\xi)\overline{V}, \tag{5.12}$$

for $0 \le \xi \le 1$ and $\tau > 0$. The initial data reads

$$\overline{V}(\xi,0) = \frac{1-\xi}{E}\Lambda\left(\frac{E\xi}{1-\xi}\right). \tag{5.13}$$

For a call option, the payoff is $\Lambda(S) = \max(S-E,0)$. The corresponding

$$\begin{aligned}\overline{V}(\xi,0) &= \max(S-E,0)(1-\xi)/E \\ &= \max(2\xi-1,0).\end{aligned}$$

Similarly, $\overline{V}(\xi,0) = \max(1-2\xi,0)$ for a put option.

On the boundaries $\xi = 0$ and $\xi = 1$, the diffusion coefficients are degenerate. If the solution is smooth up to the boundaries, then on the boundary, the equation is degenerate to the following ordinary differential equations:

$$\frac{\partial \overline{V}(0,\tau)}{\partial \tau} = -r\overline{V}(0,\tau)$$

$$\frac{\partial \overline{V}(1,\tau)}{\partial \tau} = 0.$$

The corresponding solutions are

$$\overline{V}(0,\tau) = \overline{V}(0,0)e^{r\tau} \tag{5.14}$$
$$\overline{V}(1,\tau) = \overline{V}(1,0). \tag{5.15}$$

We can discretize equation (5.12) by finite difference method. Let $\Delta\xi$ and $\Delta\tau$ are the spatial and temporal mesh sizes, respectively. Let $\xi_j = j\Delta\xi$, $\tau^n = n\Delta\tau$. The boundaries points are ξ_0 and ξ_M. We use central difference for $\partial^2\overline{V}/\partial\xi^2$ and $\partial\overline{V}/\partial\xi$. The resulting finite difference equation reads

$$\begin{aligned}\frac{dv_j}{d\tau} &= \frac{1}{2}\sigma_j^2\xi_j^2(1-\xi_j)^2\frac{v_{j+1}-2v_j+v_{j-1}}{\Delta\xi^2}\\ &\quad + r\xi_j(1-\xi_j)\frac{v_{j+1}-v_{j-1}}{2\Delta\xi} - r(1-\xi_j)v_j\end{aligned}$$

We can discretize this equation in the time direction by forward Euler method. The stability constraint is

$$\left| r\xi_j(1-\xi_j) - \frac{1}{2}\sigma_j^2\xi_j^2(1-\xi_j)^2 \right| \frac{\Delta\tau}{\Delta\xi} \le \sigma_j^2\xi_j^2(1-\xi_j)^2\frac{\Delta\tau}{2\Delta\xi^2} \le 1. \tag{5.16}$$

Remark. Many options have non-smooth payoff functions. This causes low order accuracy for finite difference scheme. Fortunately, many simple payoff function has exact solution. For instance, the European call option. For general payoff function, we may subtract its non-smooth part for which an exact solution is available. The remainder is smooth, and a finite difference scheme can yield high-order accuracy.

5.5 Fast algorithms for solving linear systems

In the backward Euler method and the Crank-Nicolson method, we need to solve linear systems of the form

$$AU = F.$$

For the backward Euler scheme,

$$\begin{aligned} A &= \text{diag}\,(-a, 1+a+c, -c)\\ &:= \begin{bmatrix} 1+a+c & -c & 0 & \cdots & & & \\ -a & 1+a+c & -c & 0 & \cdots & & \\ 0 & -a & 1+a+c & -c & 0 & \cdots & \\ & \ddots & \ddots & \ddots & \ddots & & \\ & \cdots & 0 & -a & 1+a+c & -c & 0 \\ & & \cdots & 0 & -a & 1+a+c & -c \\ & & & \cdots & 0 & -a & 1+a+c \end{bmatrix}\end{aligned}$$

and

$$U = \begin{bmatrix} U^n_{j_L+1} \\ \vdots \\ U^n_{j_R-1} \end{bmatrix}, \quad F = \begin{bmatrix} aU^{n+1}_{j_L} \\ \vdots \\ cU^{n+1}_{j_R} \end{bmatrix}.$$

For the Crank-Nicolson scheme We have

$$AU^{n+1} = BU^n + b^{n+1/2}$$

where $A =$ diag $(-\frac{a}{2}, 1 + \frac{a}{2} + \frac{c}{2}, -\frac{c}{2})$ $B =$ diag $(\frac{a}{2}, 1 - \frac{a}{2} - \frac{c}{2}, \frac{c}{2})$, and

$$b^{n+1/2} = \begin{bmatrix} a(U^{n+1}_{j_L} + U^n_{j_L})/2 \\ 0 \\ \vdots \\ 0 \\ c(U^{n+1}_{j_R} + U^n_{j_R})/2 \end{bmatrix}.$$

Now, we concentrate on solving the linear system

$$Ax = f.$$

The matrix A is tridiagonal and diagonally dominant. Let us rewrite $A =$ diag (a, b, c). Here, the constants a, b, c are different from the average weights we had before. We may assume $b > 0$. We say that A is diagonally dominant if $b > |a| + |c|$. More generally, A may takes the form $A =$ diag (a_j, b_j, c_j). and $|b_j| > |a| + |c_j|$. Without loss of generality, we may normalize the $j - th$ so that $b_j = 1$.

There are two classes of methods to solve the above linear systems. One is called direct methods, the other is called iterative methods. For one-dimensional case as we have here, direct method is usually better. However, for high-dimensional cases, iterative methods are better.

5.5.1 Direct methods

Gaussian elimination

Let us illustrate this method by the simple example: $A =$ diag $(a, 1, c)$. We multiple the first equation by $-a$ and add it into the second equation to eliminate the term x_{j_L+1} in the second equation. Then the resulting equation becomes

$$\begin{bmatrix} 1 & c & 0 & 0 & \cdots & 0 \\ 0 & 1-ac & c & 0 & \cdots & 0 \\ 0 & a & 1 & c & \cdots & 0 \\ & & & & \vdots & 0 \\ & & & & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_{j_L+1} \\ x_{j_L+2} \\ x_{j_L+3} \\ \vdots \\ x_{j_R-1} \end{bmatrix} = \begin{bmatrix} b_{j_L+1} \\ -ab_{j_L+1} + b_{j_L+2} \\ b_{j_L+3} \\ \vdots \\ b_{j_R-1} \end{bmatrix}$$

We continue to eliminate the term a in the third equation, and so on. Finally, we arrive

$$\begin{bmatrix} 1 & c & 0 & 0 & \cdots & 0 \\ 0 & 1-ac & c & 0 & \cdots & 0 \\ 0 & 0 & 1-c/(1-ac) & c & \cdots & 0 \\ 0 & 0 & 0 & & \vdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_{j_L+1} \\ x_{j_L+2} \\ x_{j_L+3} \\ \vdots \\ x_{j_R-1} \end{bmatrix} = \begin{bmatrix} b'_{j_L+1} \\ b'_{j_L+2} \\ b'_{j_L+3} \\ \vdots \\ b'_{j_R-1} \end{bmatrix}$$

Then x_j can be solved easily. The diagonal dominance condition guarantee that the reduced matrix is also diagonally dominant. Thus, this scheme is numerical stable.

LU decomposition

We decompose $A = LU$, where

$$L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \ell_{j_L+2} & 1 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots 0 & \\ 0 & \cdots 0 & \ell_{j_R-1} & 1 & \end{bmatrix}, U = \begin{bmatrix} u_{j_L+1} & v_{j_L+1} & 0 & \cdots & 0 \\ 0 & u_{j_L+2} & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots v_{j_R-2} & \\ 0 & \cdots 0 & 0 & u_{j_R-1} & \end{bmatrix}$$

It is easy to find a recursion formula to find the coefficients ℓ, u and v's. Once these are found, we can find x by solving

$$Ly = b, Ux = y.$$

These two equations are easy to solve. One can show that both L and U are diagonally dominant if A is.

If we watch carefully, LU-decomposition is equivalent to the Gaussian elimination.

Cyclic reduction method

Let us take the case $A =$ diag $(a, 1, c)$ to illustrate this method. Consider three consecutive equations

$$\begin{aligned} ax_{2j-2} + x_{2j-1} + cx_{2j} &= b_{2j-1} \\ ax_{2j-1} + x_{2j} + cx_{2j+1} &= b_{2j} \\ ax_{2j} + x_{2j+1} + cx_{2j+2} &= b_{2j+1} \end{aligned}$$

We can eliminate the odd-index terms x_{2j-1} and x_{2j+1}. Namely, $-a \times (2j-1)-$eq $+(2j)$-eq $-c \times (2j+1)$-eq: After normalization, we obtain

$$a'x_{2j-2} + x_{2j} + c'x_{2j+2} = b'_j$$

Here,

$$a' = -\frac{a^2}{1-2ac},\ c' = -\frac{c^2}{1-2ac},$$

$$b'_j = (b_{2j} - ab_{2j-1} - cb_{2j+1})/(1-2ac).$$

If we rename $x'_j = x_{2j}$. Then we have $A'x' = b'$, where $A' =$ diag $(a', 1, c')$. Notice that the system is reduced to half and with the same form. One can show that the iterative mapping

$$\begin{bmatrix} a \\ c \end{bmatrix} \mapsto \begin{bmatrix} a' \\ c' \end{bmatrix}$$

converges to $(0,0)^t$ quadratically fast, provided $|a| + |c| < 1$ initially. Thus, for few iteration, the matrix A is almost an identity matrix. We can invert it trivially. Once x_{2j} are found, the odd-index $x + 2j + 1$ can be found from the equation:

$$ax_{2j} + x_{2j+1} + cx_{2j+2} = b_{2j+1}.$$

A careful reader should find that the cyclic reduction is also a version of the Gaussian elimination method.

5.5.2 Iterative methods

Most iterative methods can be viewed as a proper decomposition of A, then solve an important and treat the rest as a perturbation term.

Jacob method

In Jacobi method, wer decompose

$$A = D + B$$

where D is the diagonal part and B is the off diagonal part. Since A is diagonally dominant, we may approximate x by the sequence x^n, where x^n is defined by the following iteration scheme:

$$Dx^{n+1} + Bx^n = b.$$

Let the error $e^n := x^{n+1} - x^n$. Then

$$De^n = -Be^{n-1}$$

Or

$$e^n = -D^{-1}Be^{n-1}.$$

Let us define the maximum norm

$$\|e^n\| := \max_j |e^n_j|$$

Then

$$\begin{aligned}
|e_j^n| &= |-\frac{a}{b}e_{j-1}^{n-1} - \frac{c}{b}e_{j+1}^{n-1}| \\
&\le |\frac{|a|}{|b|}\|e^{n-1}\| + \frac{|c|}{|b|}\|e^{n-1}\| \\
&= \frac{|a|+|c|}{|b|}\|e^{n-1}\|
\end{aligned}$$

Hence,

$$\|e^n\| \le \rho\|e^{n-1}\|$$

where $\rho = \frac{|a|+|c|}{|b|} < 1$ from the fact that A is diagonally dominant. This yields the convergence of the sequence x_n. The limit x satisfies the equation $Ax = b$.

Gauss-Seidel method

In Gauss-Seidel method, A is decomposed into $A = (D + L) + U$, where D is the diagonal part, L, the lower triangular part, and U, the upper triangular part of A. The approximate solution sequence is given by

$$(D + L)x^{n+1} + Ux^n = b.$$

As before, the error $e^n := x^{n+1} - x^n$ satisfies

$$e^n = -(D + L)^{-1}Ue^{n-1}$$

To analyze the decay of e^n, we use Fourier method. Let

$$\widehat{e^n}(\xi) := \sum_j e_j^n e^{-ij\xi}.$$

Then we have

$$\begin{aligned}
\widehat{e^n}(\xi) &= G(\xi)\widehat{e^{n-1}}, \\
G(\xi) &= -\frac{ce^{i\xi}}{b + ae^{-i\xi}}
\end{aligned}$$

It is easy to see that the amplification matrix G satisfies

$$\max_\xi |G(\xi)| := \rho < 1, \text{ provided } |b| > |a| + |c|. \tag{5.17}$$

This shows that the Gauss-Seidel method also converges for diagonally dominant matrix.
Exercise. Show the above statement (5.17).

Successive over-relaxation method (SOR)

In the methods of Jacobi and Gauss-Seidel, the approximate sequence x^n is usually convergent monotonely. We therefore have a chance to speed them up by an extrapolation procedure described below.

$$\begin{aligned} y^{n+1} &= -D^{-1}(L+U)x^n + b \\ x^{n+1} &= x^n + \omega(y^{n+1} - x^n). \end{aligned}$$

Here, ω is a parameter. In order to speed up, we require $\omega > 1$. We also need to require $\omega < 2$ for stability. The optimal ω is chosen to minimize the amplification matrix $G_\omega(\xi)$.

Exercise. Find the amplification matrix G_ω and the optimal ω for the matrix $A = \text{ diag } (a, b, c)$. Also, determine the rate

$$\rho := \min_\omega \max_\xi |G_\omega(\xi)|.$$

Chapter 6

American Option

6.1 Introduction

An American option has the right to exercise any time during the life of the option. We denote the American call by C and American put by P. The first important thing we should note is that the value of an American option is greater than or equal to the payoff function: $V(S,t) \geq \Lambda(S,t)$. Otherwise, there is an arbitrage opportunity because we can buy the American option then exercise it immediately to gain a net profit $V - \Lambda$.

We recall that the value of an American call option is equal to that of a European call option, see Theorem 3.2. However, for other cases like the the American put option or the American call option on dividend-paying asset, the American options do cost more. We explain why it is so below.

American put First, we notice that there must be some region in $0 < S < E$ in the S-P plane where the European put $p(S,t) < E - S$. In fact, we know that $p(0,t) = Ee^{-r(T-t)} < E - 0$, which is below the line $p = E - S$ in a neighborhood of $S = 0$, see Figure 6.1. On the other hand, the corresponding American option $P(S,t)$ must satisfy

$$P(S,t) \geq \Lambda(S) = E - S.$$

Otherwise, we can buy an asset S and a put $P(S,t)$, then exercise it immediately. We have the right to sell S on E. This gives a risk-free profit: $E - P - S > 0$. Thus, in the region where $p(S,t) < E - S \leq P(S,t)$, we must have $p(S,t) < P(S,t)$.

American call on a dividend-paying asset For an European call on dividend-paying asset, its value must lie below the payoff function $\Lambda(S) = S - E$ in some region $S > E$, see Figure 6.2. In fact, $c(S,t) \sim Se^{-D_0(T-t)}$ for large S, which is below the line $c = S - E$. Therefore, there must be a region in $S > E$ in (S,t)-plane where $c(S,t) < \Lambda(S) := \max\{S - E, 0\}$. In this region, if we could exercise the corresponding American call option, then based on $V(S,t) \geq \Lambda(S,t)$ (i.e. $C(S,t) \geq S - E$), the corresponding American call option C must also satisfy

$$C(S,t) \geq \Lambda(S) = \max\{S - E, 0\} > c(S,t).$$

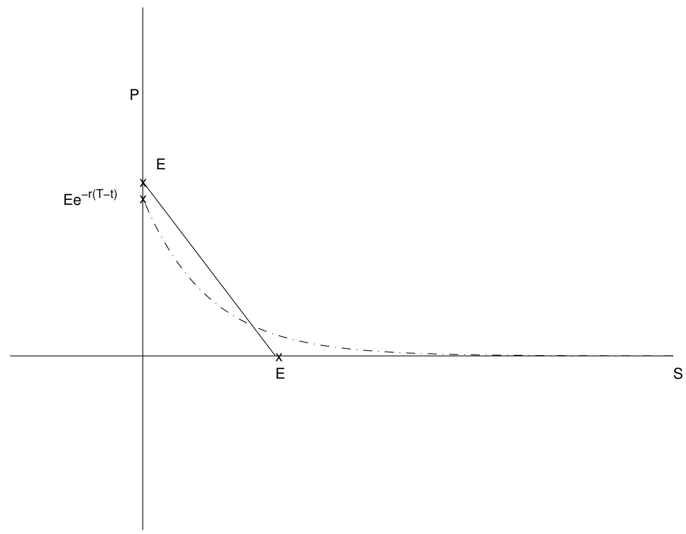


Figure 6.1: The dash line is the price of an European put option. There is a region where $p(S,t) < \max\{E-S,0\}$. This can be seen from the boundary condition $P(0,t) = Ee^{-r(T-t)}$, which is below the payoff function $p = E - S$.

We summarize the discussion as below.

Proposition 6. *American option V always satisfies $V(S,t) \geq \Lambda(S)$. Furthermore,*

- *for American put option, $P(S,t) > p(S,t)$ in some region where $p(S,t) < \Lambda_p(S)$;*
- *for American call option on dividend-paying asset, $C(S,t) > c(S,t)$ in some region where $c < \Lambda_c(S)$.*

6.2 American options as a free boundary value problem

6.2.1 American put option

Existence of optimal exercise price $S_f(t)$. First, there must be some value of S for which it is optimal from the holder's point of view to exercise the American option. Otherwise, we should hold the option for all possible S. Then this option is identical to a European option. But we have seen that this is not the case. Hence, for each time t, there is an optimal exercise price $S_f(t)$ such that, when $S < S_f(t)$ one should exercise the put option, and when $S > S_f(t)$, we should hold the option. In other words, we should have:

- $P(S,t) \equiv \max(E-S, 0)$ for $S < S_f(t)$,
- $P(S,t)$ satisfies the Black-Scholes equation for $S > S_f(t)$.

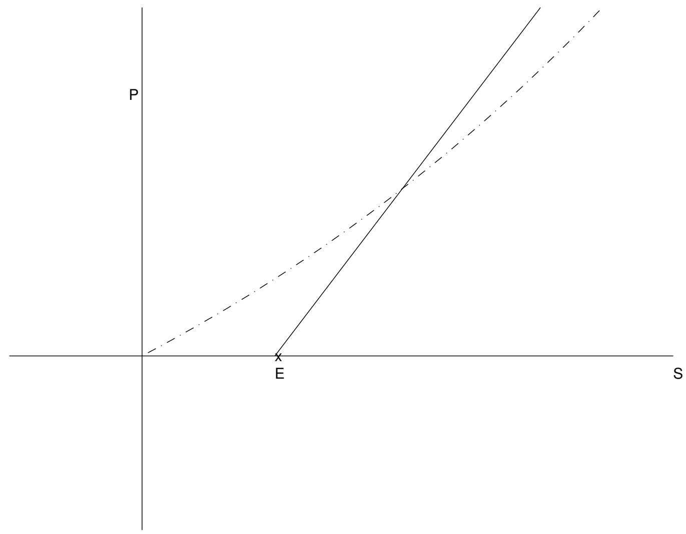


Figure 6.2: Dash line is the European call option on a dividend-paying asset. Notice that $c(S,t) \sim Se^{-D_0(T-t)}$ for $S >> 1$, hence $C(S,t) < \Lambda(S) = \max(S-E,0)$ for some $S_f(t)$.

The first statement can be viewed as the definition $S_f(t)$. That is,

$$S_f(t) := \max\{S | S \le E \text{ and } P(S,t) \equiv \max(E-S,0)\}. \tag{6.1}$$

The optimal exercise price can also be viewed as the minimal exercise price that one should hold the put option:

$$S_f(\cdot) = \min\{b(\cdot) | P(S,t) \text{ is a B-S solution in } S > b(t) \text{ satisfying } P(b(t),t) = E - b(t)\}.$$

This means that, given a exercise price curve $b(\cdot)$, the holder holds the option if $S > b(t)$ and exercise it if $S < b(t)$. When it is exercised, the payoff is $E - b(t)$. The value of $P(b(t),t)$ must be equal to $E - b(t)$ by the no-arbitrage assumption, otherwise we can buy or short P then exercise it right away to gain a net profit. Thus, given an exercise price curve $b(\cdot)$, we can find a solution $P(S,t)$ with boundary condition $P(b(t),t) = E - b(t)$. The optimal exercise price $S_f(\cdot)$ is the smallest such exercise price curve.

Notice that

$$S_f(t) \le E.$$

If S (or $S_f(t)$) goes up beyond E, then the right of a put to sell a higher price S with a lower price E is meaningless. Thus, such put $P(S,t)$ has no value at all. This means that the optimal exercise price cannot exceed E. Furthermore, at time T, we should have

$$S_f(T) = E.$$

This follows from the Definition (6.1).

However, we don't know the optimal exercise price $S_f(t)$, $t < T$ in advance. It is a new unknown, called free boundary. On this free boundary, there will be two boundary conditions for P. One boundary condition is used to determine the B-S solution P_{BS} uniquely in the region $S > S_f(t)$, $0 \le t \le T$. The other boundary condition is used to determine the free boundary itself. The following proposition gives the proper boundary conditions on $S_f(t)$.

Proposition 7. *Based on the no-arbitrage hypothesis, the price of an American put should satisfy the following boundary conditions on the free boundary $S_f(t)$:*

- $P(S_f(t), t) = E - S_f(t)$,
- $\frac{\partial P}{\partial S}(S_f(t), t) = -1$.

Remark. Notice that in the region $S < S_f(t)$, $P(S,t) \equiv E - S$, where we have $\partial P/\partial S(S,t) \equiv -1$. Therefore, the above boundary conditions are equivalent to

- $P(S,t)$ is continuous across the free boundary $S = S_f(t)$.
- $\frac{\partial P}{\partial S}(S,t)$ is continuous across the free boundary $S = S_f(t)$.

Proof.

1. Given the optimal price $S_f(t)$, it means that it is the price to the holder to excise it, the return is $E - S_f(t)$. The corresponding option price must equal $E - S_f(t)$. Otherwise, there is an arbitrage opportunity. Thus, we get $P(S_f(t), t) = E - S_f(t)$ for $0 \le t \le T$.

2. The free boundary $S = S_f(t)$ is an unknown. We want to derive a constraint to determine it. Suppose we are given a boundary $S = b(t)$, $0 \le t \le T$. We solve the B-S equation in the region $S > b(t)$, $0 \le t \le T$ with boundary condition

$$P(b(t), t) = E - b(t)$$

and final condition $P(S,T) = \max(E - S, 0)$. Such a solution depends on the path $b(\cdot)$. We denote it by $P(S,t;b(\cdot))$. The optimal exercise price $S_f(t)$ is the minimal such curve $b(\cdot)$. Below $S_f(t)$, then one should exercise the put option. Therefore, a small variation of P with respective to the change of $b(\cdot)$ is zero at the optimal exercise price $S_f(\cdot)$. This means

$$\frac{\delta P(S_f(t), t; S_f(\cdot))}{\delta b(\cdot)} = 0.$$

Here, $\delta P/\delta b$ means the variation of P w.r.t. the path $b(\cdot)$. [1]

[1] The variation of P w.r.t. a path $b(\cdot)$ is defined as below. Consider a small change of the path, say $b(\cdot) + \epsilon r(\cdot)$. The function r is called the direction of the variation. Then, just like the definition of gradient in n dimensions, we define the variation of P w.r.t. $b(\cdot)$ at a path $b(\cdot)$ by

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0} P(S,t;b+\epsilon r) = \int_0^T \frac{\delta P(S,t;b(s))}{\delta b(\cdot)} r(s)\, ds.$$

3. On the boundary $S = b(t)$, P satisfies the boundary condition

$$P(b(t), t; b(\cdot)) = E - b(t).$$

We can take variation of this formula w.r.t. $b(\cdot)$, then evaluate at the optimal exercise price $S_f(\cdot)$. This gives

$$\frac{\partial P}{\partial S}(S_f(t), t; S_f(\cdot))\frac{\delta b(\cdot)}{\delta b(\cdot)}\Big|_{b(\cdot)=S_f(\cdot)} + \frac{\delta P(S_f(t), t; S_f(\cdot))}{\delta b(\cdot)} = -1.$$

Since the second term on the left-hand side is zero. We get

$$\frac{\partial P}{\partial S}(S_f(t), t; S_f(\cdot)) = -1.$$

□

Thus, we treat the American put option as the following free boundary value problem. There exists an optimal exercise price $S_f(t)$ such that

1. for $S < S_f(t)$, early exercise is optimal, and $P(S, t) = E - S$;
2. for $S > S_f(t)$, one should hold the put option and P satisfies the Black-Scholes equation:

$$\frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP = 0;$$

3. across the free boundary $(S_f(t), t)$, both P and $\frac{\partial P}{\partial S}$ are continuous.

We will solve this free boundary value problem in the next section.

6.2.2 American call option on a dividend-paying asset

As we have seen in the introduction of this chapter that an American call option $C(S, t)$ on a dividend-paying asset has asymptotic value $C(S, t) \sim Se^{-D_0(T-t)}$ for large S. This value is below the payoff function $\Lambda \equiv \max(S - E, 0)$. Therefore, there must an optimal $S_f(t)$ such that we should exercise this call option when $S > S_f(t)$ and hold it when $S < S_f(t)$. On the free boundary $S = S_f(t)$, based on the no-arbitragy hypothesis, we should require that

- both C and $\partial C/\partial S$ are continuous across the free boundary $(S_f(t), t)$ for $0 < t < T$.

We summarize this by the following equations

$$C_t + \frac{\sigma^2}{2}S^2\frac{\partial^2 C}{\partial S^2} + (r - D_0)S\frac{\partial C}{\partial S} - rC = 0,\ 0 < S < S_f(t)$$
$$C(S, t) \equiv \Lambda(S) \equiv \max\{S - E, 0\},\ S > S_f(t).$$

On the free boundary $S = S_f(t)$, the boundary condition is required

$$C(S_f(t), t) = S_f(t) - E,$$
$$\frac{\partial C}{\partial S}(S_f(t), t) = 1,\ 0 \le t \le T.$$

6.3 American option as a linear complementary problem

The American option can also be formulated as a linear complementary problem, where the free boundary is treated *implicitly* and numerical schemes can be easily designed. To illustrate this linear complementary problem, first we notice that an American option V should satisfy the following conditions:

(i) $V(S,t) \geq \Lambda(S)$,

(ii) $V_t + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \leq r(V - S\frac{\partial V}{\partial S})$,

(iii) either $V(S,t) = \Lambda(S)$ or $V_t + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = r(V - S\frac{\partial V}{\partial S})$ should hold,

(iv) both V and $\frac{\partial V}{\partial S}$ are continuous.

Here, $\Lambda(S) = \max(E - S, 0)$ is the corresponding payoff function.

We have seen the reasons for (i), and (iv). We explain the reasons of (ii) below. Let us consider the portfolio $\Pi = V - \Delta S$. As we have seen that the Delta hedge eliminate the randomness of Π and yields

$$d\Pi = V_t + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}.$$

When it is optimal to hold the option, then

$$d\Pi = r(V - \frac{\partial V}{\partial S} S).$$

Otherwise, we should have

$$d\Pi \leq r(V - \frac{\partial V}{\partial S} S),$$

based on no arbitrage opportunities. Thus, the Black-Scholes is replaced by the Black-Scholes inequality.

To show (iii), we know that if we exercise the option, then $V = \Lambda$, otherwise, we hold the option and its value should satisfy the Black-Scholes equation.

Properties (i)-(iii) can be formulated as the following linear complementary equations:

$$\begin{cases} V - \Lambda \geq 0 \\ V_t + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - r(V - S\frac{\partial V}{\partial S}) \leq 0, \\ (V - \Lambda)\left(V_t + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - r(V - S\frac{\partial V}{\partial S})\right) = 0 \end{cases} \tag{6.2}$$

We look for C^1 solution (i.e. both V and $\frac{\partial V}{\partial S}$ are continuous) in the domain $S \in (0, \infty)$ and $t \geq 0$ with final condition

$$P(S,T) = \Lambda(S).$$

Such a problem is called a linear complementary problem. The advantage of this formulation is that the free boundary is treated implicitly, which can be solved by some penalty method or projection method to be explained below.

Remark For American put, in the region $S < S_f(t)$, $P(S,t) = E - S$. We plug it into the Black-Schole equation

$$\frac{\partial P}{\partial t} + \frac{\sigma^2}{2} S^2 \frac{p^2 P}{\partial S^2} + r\left(S\frac{\partial P}{\partial S} - P\right) = -rE < 0.$$

Thus, the B-S inequality holds.

6.4 *Penalty method for linear complementary problem

We can reformulate this problem in terms of x variable. As before, we use the following change of variables: $V = Ev$, $S = Ee^x$, $\tau = T - t$. The free boundary now in x-variable is $x_f(t)$. The free boundary value problem is formulated as

(i) for $-\infty < x < x_f(t)$, $v = 1 - e^x$, and

$$v_\tau - \frac{\sigma^2}{2} v_{xx} - (r - \frac{\sigma^2}{2}) v_x + rv \geq 0.$$

(ii) for $x_f(t) < x < \infty$, $v > 1 - e^x$, and

$$v_\tau - \frac{\sigma^2}{2} v_{xx} - (r - \frac{\sigma^2}{2}) v_x + rv = 0,$$

(iii) both v and $\frac{\partial v}{\partial x}$ are continuous.

The linear complementary problem is formulated as:

(i) $v - (1 - e^x) \geq 0$,

(ii) $v_\tau - \frac{\sigma^2}{2} v_{xx} - (r - \frac{\sigma^2}{2}) v_x + rv \geq 0$,

(iii) $(v_\tau - \frac{\sigma^2}{2} v_{xx} - (r - \frac{\sigma^2}{2}) v_x + rv)(v - (1 - e^x)) = 0$,

(iv) v and v_x are continuous.

with initial condition $v(x, 0) = \Lambda(x) = 1 - e^x$.

We may replace v by $ue^{-r\tau}$ to eliminate the term rv. Then we have

$$\begin{aligned}
\left(u_\tau - \frac{\sigma^2}{2} u_{xx} - (r - \frac{\sigma^2}{2}) u_x\right)(u - g) &= 0 \\
u_\tau - \frac{\sigma^2}{2} u_{xx} - (r - \frac{\sigma^2}{2}) u_x &\geq 0 \\
u - g &\geq 0, \\
u(x, 0) &= g(x, 0),
\end{aligned}$$

with u, u_x being continuous. Here $g(x,\tau) = \max\{e^{r\tau}(1-e^x), 0\}$. The far field boundary conditions are

$$u(x,\tau) \to 0, \text{ as } x \to \infty, \quad u(x,\tau) \to e^{r\tau}, \text{ as } x \to -\infty.$$

A mathematical theory called parabolic variational inequality gives *construction, existence, uniqueness* of the solution. (see reference: A. Friedman, *Variational Inequality*). Let us demonstrate this theory briefly. The method we shall use is called the penalty method. Let us consider the following penalty function:

$$\phi_N(v) = -e^{-Nv}.$$

It has the properties: (i) $\phi_N > 0$, (ii) $\phi_N(v) \to 0$ whenever $v > 0$.. We consider the following penalized P.D.E.:

$$\begin{aligned} u_\tau - \frac{1}{2}\sigma^2 u_{xx} - (r - \frac{\sigma^2}{2})u_x + \phi_N(u-g) &= 0 \\ u(x,0) &= g(x,0), \end{aligned}$$

with $u \to 0$ as $x \to +\infty$, $\quad u \to e^{r\tau}$ as $x \to -\infty$.

From a standard theory of nonlinear P.D.E. (by monotone method, for instance), one can show that the solution u_N exists for all $N > 0$. We then need an estimate for u_N and $\frac{\partial u_N}{\partial x}$. The boundedness of these two gives that u_N has a convergent subsequence, say u_{N_i} such that

$$u_{N_i} \to u,$$

with u_{N_i}, u, $\frac{\partial}{\partial x}u_{N_i}$, u_x being continuous. Moreover,

$$|\phi_{N_i}(u_{N_i} - g)| \leq \text{ constant}$$

As $N_i \to \infty$, we conclude $u - g \geq 0$. Further, on the set $\{u - g > 0\}$, $\phi_{N_i}(u_{N_i} - g) \to 0$. Hence we have

$$u_\tau - \frac{1}{2}u_{xx} - (r - \frac{\sigma^2}{2})u_x = 0 \text{ on } \{u - g > 0\}.$$

6.5 Numerical Methods

6.5.1 Binomial method for American puts

We recall that the solution of the Black-Scholes equation can be discretized by the following binomial method (or the forward Euler method):

- Choose a large N, define $\Delta t = T/N$, $\Delta x = \sigma\sqrt{\Delta t}$, $S_j^n = S_0 e^{j\Delta x}$, $|j| \leq n$, $j + n =$ even.
- Define
$$e^{r\Delta t}V_j^n = pV_{j+1}^{n+1} + qV_{j-1}^{n+1},$$
where
$$p = \frac{e^{r\Delta t} - u}{u - d}, u = e^{\sigma\sqrt{\Delta}}, d = 1/u, q = 1 - p.$$

Similarly, the linear complementary problem can be discretized as

$$\begin{aligned}\left(e^{r\Delta t}V_j^n - pV_{j+1}^{n+1} - qV_{j-1}^{n+1}\right)\left(V_j^n - \Lambda_j^n\right) &= 0,\\ e^{r\Delta t}V_j^n - pV_{j+1}^{n+1} - qV_{j-1}^{n+1} &\geq 0,\\ V_j^n - \Lambda_j^n &\geq 0\end{aligned}$$

Here, $\Lambda_j^n := E - S_j^n$. This recursive procedure finding V^n from V^{n+1} for $n = N-1, N-2, ..., 0$ with initial condition

$$V_j^N = \Lambda_j^N.$$

This discretized linear complementary problem can be solved by the following projected forward Euler method:

$$\boxed{V_j^n = \max\{e^{-r\Delta t}(pV_{j+1}^{n+1} + qV_{j-1}^{n+1}), \Lambda_j^n\}.} \tag{6.3}$$

Here, $\Lambda_j^n := \Lambda(S_j^n)$. This is called projected binomial method.

Theorem 6.9. *The projected binomial method for American option is stable and converges. Namely, $V_j^n \to V(S,t)$ with $S_j := S_0 e^{j\Delta x} \to S$ and $n\Delta t \to t$ as $n \to \infty$. Furthermore, $V(S,t)$ satisfies the linear complementary equations (6.2).*

Sketch the idea of the proof.

1. One can show that this method is stable and converges in $C_b^1(\mathbb{R}^+)$, the space of C^1-continuous and bounded functions, by using the monotone operator theory. (Ref. Majda & Crandell, Math. Comp.) An operator $\Phi : C_b(\mathbb{R}) \to C_b(\mathbb{R})$ is called monotone operator if $f \geq g$ implies $\Phi(f) \geq \Phi(g)$. For continuous case, the solution operator $\Phi(t)$ of the Black-Scholes equation is a monotone operator. Their discrete version, the one-step solution operator which maps $V^{n+1} \mapsto V^n$ is also a monotone operator. The projection operator is a monotone operator, which can be proven easily by the fact that $\max(a,c) \leq \max(b,c)$ if $a \leq b$.

2. Furthermore, from the projection $V_j^n = \max\{e^{-r\Delta t}(pV_{j+1}^{n+1} + qV_{j-1}^{n+1}), \Lambda_j^n\}$, we obtain

$$\begin{aligned}&V_j^n \geq \Lambda_j^n\\ &V_j^n \geq e^{-r\Delta t}\left(pV_{j+1}^{n+1} + qV_{j-1}^{n+1}\right).\end{aligned}$$

   Thus the limiting function $V(S,t)$ also satisfies $V(S,t) \geq \Lambda(S,t)$ and the Black-Scholes inequality.

3. From construction, the projection: $V_j^n = \max\{e^{-r\Delta t}(pV_{j+1}^{n+1} + qV_{j-1}^{n+1}), \Lambda_j^n\}$ satisfies either $V_j^n - \Lambda_j^n = 0$, or $V_j^n - e^{-r\Delta t}(pV_{j+1}^{n+1} + qV_{j-1}^{n+1}) = 0$. Thus, it satisfies the complementary condition:

$$\left(e^{r\Delta t}V_j^n - pV_{j+1}^{n+1} - qV_{j-1}^{n+1}\right)\left(V_j^n - \Lambda_j^n\right) = 0.$$

   From convergence result, their limiting function also satisfies the continuous complementary condition.

Below is a C code for binomial method for American put.

```
Binomial_American_option(V,S,S0,E,r,sigma,dt,N)
{
    double u,d,p;
    double discount, tmp;
    int m,n;
    discount = exp(-r*dt);
    u = exp(sigma*sqrt(dt));
    d = 1/u;
    p = (1/discount - d)/(u-d);

    //Binomial tree for asset price: at level n.
    S[0][0]=S0;
    for (n = 1; n <= N; ++n)
    {
        for (m=n; m > 0; --m)
            S[n][m] = u*S[n-1][m-1];
        S[n][0] = d*S[n-1][0];
    }

    // Binomial tree for option.
    for (m=0; m <=N; ++m)
        V[N][m] = payoff_put(S[N][m],E);
    for ( n = N-1; n >= 0; --n)
    {
        for (m=0; m < n; ++m)
        {
            tmp = p*V[n+1][m+1] + (1-p)*V[n+1][m];
            tmp *= discount;
            V[n][m] = max(tmp,payoff_put(S[n][m],E));
        }
    }
}

double payoff_put(S,E)
{
    return max(E-S,0);
}
```

6.5.2 Binomial method for American call on dividend-paying asset

The linear complementary problem for this American call option is

(i) $V(S,t) \geq \Lambda(S) \equiv \max\{S - E, 0\}$

(ii) $V_t + \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2} + (r - D_0)S\frac{\partial V}{\partial S} - rV \leq 0,$

(iii) $\left(V_t + \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2} + (r - D_0)S\frac{\partial V}{\partial S} - rV\right)(V - \Lambda) = 0.$

(iv) V and V_S are continuous.

The binomial approximation for the B-S equation is

$$e^{r\Delta t}V_j^n = pV_{j+1}^{n+1} + qV_{j-1}^{n+1},$$

where

$$p = \frac{\sigma^2}{2}\frac{\Delta t}{(\Delta x)^2} + (r - D_0 - \frac{\sigma^2}{2})\frac{\Delta t}{2\Delta x}, \quad q = \frac{\sigma^2}{2}\frac{\Delta t}{(\Delta x)^2} - (r - D_0 - \frac{\sigma^2}{2})\frac{\Delta t}{2\Delta x},$$

and $p + q = 1$. We choose Δt and Δx so that $p > 0$ and $q > 0$. For American option, V has to be greater than $\Lambda(S, t)$, the payoff function at time t. Hence, we should require

$$V_j^n = \max\{e^{-r\Delta t}(pV_{j+1}^{n+1} + qV_{j-1}^{n+1}), \Lambda_j^n\}.$$

The above is the projective forward Euler method.

For the corresponding binomial model, first we determine the up/down ratios u and d for the riskless asset price by

$$pu + (1 - p)d = e^{(r-D_0)\Delta t}, \quad pu^2 + (1 - p)d^2 = e^{(2(r-D_0)+\sigma^2)\Delta t},$$

or equivalently,

$$\begin{aligned} u &= A + \sqrt{A^2 - 1}, \quad d = 1/u, \\ A &= \frac{1}{2}(e^{-(r-D_0)\Delta t} + e^{(r-D_0+\sigma^2)\Delta t}), \\ p &= \frac{e^{(r-D_0)\Delta t} - d}{u - d}, \rangle q = 1 - p. \end{aligned}$$

Then the binomial model is given by

$$\begin{aligned} S_j^n &= \begin{cases} uS_{j-1}^{n-1}, & \text{with probability } p \\ dS_{j+1}^{n-1}, & \text{with probability } q \end{cases} \\ S_0^0 &= S. \end{aligned}$$

and

$$V_j^n = \max\{e^{-r\Delta t}(pV_{j+1}^{n+1} + qV_{j-1}^{n+1}), S_j^n - E\}.$$

6.5.3 *Implicit method

For implicit method like backward Euler or Crank-Nicolson method, we need to add the constraints $u^{n+1} \geq g^{n+1}$ for American option. It is important to know that if an iterative method is used, then we should require this condition hold in each iteration steps. For instance, in the SOR iteration method,

$$\begin{aligned} V^{n,(k+1)} &= \max\{V^{n,(k)} + \omega(y^{n,(k+1)} - V^{n,(k)}), \Lambda^n\}, \\ y^{n,(k+1)} &= (D + L)^{-1}(-UV^{n,(k)} + f^{n+1}), \ k = 0, \cdots K \\ V^n &\equiv V^{n,(K)} \end{aligned}$$

This guarantees that $V^{n+1} \geq \Lambda^{n+1}$.

6.6 Converting American option to a fixed domain problem

6.6.1 American call option with dividend paying asset

We consider the American call option on a dividend paying asset:

$$V_t + \frac{\sigma^2}{2} S^2 \frac{\partial^2 V}{\partial S^2} + (r - D_0) S \frac{\partial V}{\partial S} - rV = 0,$$
$$V(S,t) \geq \Lambda(S) \equiv \max\{S - E, 0\},$$
$$0 \leq S \leq S_f(t), 0 \leq t \leq T.$$

where $S_f(t)$ is the free boundary. On this free boundary, the boundary condition is required

$$V(S_f(t),t) = S_f(t) - E,$$
$$\frac{\partial V}{\partial S}(S_f(t),t) = 1,\ 0 \leq t \leq T.$$

We also need a condition for S_f at final time

$$S_f(T) = \max(E, rE/D_0)$$

Before converting the problem, we first remove the singularity of the final data (i.e. non-smoothness of the payoff function) as the follows. We may substract V by an European call option c with the same payoff data. Notice that $c(S,t)$ has exact solution. The new variable $V - c$ satisfies the same equation, yet it has smooth final data.

To convert the free boundary problem to a fixed domain problem, we introduce the following change-of-variables:

$$\begin{cases} \xi &= S/S_f(t) \\ \tau &= T - t \\ u(\xi,\tau) &= (V(S,t) - c(S,t))/E \\ s_f(\tau) &= S_f(t)/E \end{cases}$$

The new equations for these new variables are

$$\begin{cases} \frac{\partial u}{\partial \tau} = \frac{\sigma^2}{2}\xi^2 \frac{\partial^2 u}{\partial \xi^2} + \left((r - D_0) + \frac{1}{s_f}\frac{ds_f}{d\tau}\right)\xi \frac{\partial u}{\partial \xi} - ru, & 0 \leq \xi \leq 1, \\ & 0 \leq \tau \leq T, \\ u(\xi, 0) = 0, & 0 \leq \xi \leq 1, \\ u(1,\tau) = g(s_f(\tau),\tau), & 0 \leq \tau \leq T \\ \frac{\partial u}{\partial \xi}(1,\tau) = h(s_f(\tau),\tau), & 0 \leq \tau \leq T \\ s_f(0) = \max(1, r/D_0), & \end{cases} \tag{6.4}$$

where

$$\begin{aligned} g(s_f(\tau),\tau) &= s_f(\tau) - 1 - c(Es_f(\tau), T - \tau), \\ h(s_f(\tau),\tau) &= s_f(\tau)\left[1 - \frac{\partial c}{\partial S}(Es_f(\tau), T - \tau)\right]. \end{aligned}$$

At the boundary $\xi = 0$, the Black-Sholes equation is degenerate to

$$\frac{\partial u}{\partial \tau} = -ru.$$

With the trivial initial condition yields

$$u(0,\tau) = 0,\ 0 \le \tau \le T.$$

In practice, we can solve the modified Black-Sholes equation (6.4) with the boundary conditions

$$\begin{cases} u(0,\tau) = 0 & 0 \le \tau \le T \\ \frac{\partial u}{\partial \xi}(1,\tau) = h(s_f(\tau),\tau), & 0 \le \tau \le T \end{cases} \tag{6.5}$$

We can differentiate the other boundary condition in τ and yield an ODE for the free boundary:

$$\frac{\partial u}{\partial \tau}(1,\tau) = \frac{\partial g}{\partial \xi}\frac{ds_f}{d\tau}.$$

with $s_f(0) = \max(1, r/D_0)$.

6.6.2 American put option

For American put option P, the B-S equation is on the infinite domain $S > S_f(t)$, $0 \le t \le T$. Through the change-of-variable

$$\begin{cases} \eta = \frac{E^2}{S} \\ u(\eta,t) = \frac{EP(S,t)}{S} \end{cases}$$

the infinity domain problem is converted to a finite domain problem:

$$\begin{cases} \frac{\partial u}{\partial t} + \widehat{f}\sigma^2\eta^2\frac{\partial^2 u}{\partial \eta^2} - r\eta\frac{\partial u}{\partial \eta} = 0, & 0 \le \eta \le \eta_f(t), \\ & 0 \le t \le T \\ u(\eta,T) = \max(\eta - E, 0), & 0 \le \eta \le \eta_f(T), \\ u(\eta_f(t),t) = \eta_f(t) - E, & 0 \le t \le T, \\ \frac{\partial u}{\partial \eta}(\eta_f(t),t) = 1, & 0 \le t \le T, \\ \eta_f(T) = \max(E,0) & \end{cases}$$

We can further convert it to a fixed domain problem as that in the last section.

Chapter 7

Exotic Options

Option with more complicated payoff then the standard European or American calls and puts are called exotic options. They are usually traded over the counter. Their prices are usually not quoted on an exchange. We list some common exotic options below.

1. Binary options
2. compound options
3. chooser options
4. barrier options
5. Asian options
6. Lookback options

In the last two, the payoff depends on the history of the asset prices, for instance, the averages, the maximum, etc., we shall call these kinds of options, the path-dependent options, and will be discussed in the next Chapter.

7.1 Binaries

The payoff function $\Lambda(S)$ is an arbitrary function. One particular binary option is the cash-or-nothing call, whose payoff is

$$\Lambda(S) = BH(S - E).$$

This option can be interpreted as a simple bet on an asset price: if $S > E$ at expiry the payoff is B, otherwise zero. We have seen its value is

$$V = e^{-r(T-t)} \int_0^\infty \mathcal{P}(S', T, S, t)\Lambda(S')dS' = e^{-r(T-t)} B\mathcal{N}(d_2),$$

where

$$\mathcal{P}(S', T, S, t) = \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} e^{-\frac{\log(\frac{S'}{S})-(r-\frac{\sigma^2}{2})(T-t)}{2\sigma^2(T-t)}}$$

is the transition probability density for asset price in risk-neutral world.

7.2 Compounds

A compound option may be described as an option on an option. We consider the case where the underlying option is a vanilla put or call and the compound option is vanilla put or call on the underlying option. The extension to more complicated option on more complicated option is relatively straightforward. There are four different classes of basic compound options:

1. call-on-call,
2. call-on-put,
3. put-on-call,
4. put-on-put.

Let us investigate the case *call-on-call*. Other cases can be treated similarly. The underlying option is

$$\text{Expiry} : T_2, \quad \text{Strike price} : E_2.$$

The compound option on this option

$$\text{Expiry} : T_1 < T_2, \quad \text{Strike price} : E_1.$$

The underlying option has value $C(S, t, T_2, E_2)$. At time T_1, its value $C(S, T_1, T_2, E_2)$. The payoff for the compound call option is $\max\{C(S, T_1, T_2, E_2) - E_1, 0\}$. Because the compound options value is governed only by the randomness of S, according to the Black-Scholes analysis, it also must satisfy the same Black-Scholes equation. We then solve the Black-Scholes equation with payoff

$$\max\{C(S, T_1, T_2, E_2) - E_1, 0\}.$$

7.3 Chooser options

A regular chooser option gives its owner the right to purchase, for an amount E_1 at time T_1, either a call or a put with exercise price E_2 at time T_2. Thus, it is a "call on a call or put". Certainly, we have $T_1 < T_2$. The payoff at T_1 for this call-on-a-call-or-put" is

$$\Lambda = \max\{C(S, T_1) - E_1, P(S, T_1) - E_1, 0\}.$$

The compound option also satisfies the Black-Scholes equation for the same reason as above. From this and payoff function at T_1, we can value V at t. The contract can be made more general by having the underlying call and put with different exercise prices and expiry dates, or by allowing the right to sell the vanilla put or call. By using the Black-Scholes formula for vanilla option, there is no difficulty to value these complex chooser options.

7.4 Barrier option

Barrier options differ from vanilla options in that part of the option contract is triggered if the asset price hits some barrier, $S = X$, say at some time prior to T. As well as being either calls or puts, barrier options are categorized as follows.

1. **up-and-in**: the option expires worthless *unless* S reaches X from *below* before expiry.
2. **down-and-in**: the option expires worthless *unless* S reaches X from *above* before expiry.
3. **up-and-out**: the option expires worthless *if* S reaches X from *below* before expiry.
4. **down-and-out**: the option expires worthless *if* S reaches X from *above* before expiry.

7.4.1 down-and-out call(knockout)

A European option whose value becomes zero if S ever goes as low as $S = X$. Sometimes in the knockout options, one can have boundary of time, or one can have rebate if the barrier is crossed. In the latter case, the option holder receives a specific amount Z for compensation.

Let us consider the case of a European style down-and-out option without relate. We assume $X < E$. The boundary conditions are

$$V(X,t) = 0, \text{ (boundary condition)}, \quad V(S,t) \sim S \text{ as } S \to \infty.$$

The final condition, $V(S,T) = \max\{S - E, 0\}$. For $S > X$, the option becomes a vanilla call, it satisfies the Black-Scholes equation.

Let us find its explicit solution. Let $S = Ee^x$, $t = T - \frac{\tau}{(\sigma^2/2)}$, $V = Ev$. The Black-Scholes equation is transformed into

$$v_\tau = v_{xx} + (k-1)v_x - kv,$$

where $k = \frac{r}{\sigma^2/2}$. We make another change of variable :

$$v = e^{\alpha x + \beta\tau} u.$$

We choose α, β to eliminate the lower order terms in the derivatives of x:

$$\begin{aligned} \beta e^{\alpha x+\beta\tau}u + e^{\alpha x+\beta\tau}u_\tau &= \alpha^2 e^{\alpha x+\beta\tau}u + 2\alpha e^{\alpha x+\beta\tau}u_x + e^{\alpha x+\beta\tau}u_{xx} \\ &\quad +(k-1)(\alpha e^{\alpha x+\beta\tau}u + e^{\alpha x+\beta\tau}u_x) - ke^{\alpha x+\beta\tau}u. \end{aligned}$$

This implies that $\alpha = -\frac{1}{2}(k-1)$ and $\beta = -\frac{1}{4}(k+1)^2$ and equation becomes $u_\tau = u_{xx}$.

Let $x_0 = \log(\frac{x}{E})$, or $X = Ee^{x_0}$. The boundary condition becomes $u(x_0, \tau) = 0$ and $u(x,\tau) \sim e^{(1-\alpha)x - \beta\tau}$ as $x \to \infty$. The initial condition becomes

$$u(x,0) = u_0(x) = \max\{e^{\frac{1}{2}(k+1)x} - e^{\frac{1}{2}(k-1)x}, 0\}.$$

This follows from the payoff function being

$$\Lambda = \max\{S - E, 0\} = E\max\{\frac{S}{E} - 1, 0\} = E\max\{e^x - 1, 0\},$$

and $V = e^{\alpha x+\beta\tau}u(x, T-\frac{\tau}{(\sigma^2/2)})$, with $u(x,0) = u_0(x) = e^{-\alpha x}\max\{e^x-1, 0\} = \max\{e^{(-\alpha+1)x}-e^{-\alpha x}, 0\}$. Notice that because $X < E$, we have $x_0 < 0$, and

$$u_0(x) = \begin{cases} 0 & \text{for } x_0 < x < 0, \\ \max\{e^{(-\alpha+1)x} - e^{-\alpha x}, 0\} & \text{otherwise.} \end{cases}$$

We use method-of-reflection to solve above heat equation with zero boundary condition. We reflect the initial condition about x_0 as

$$u(x,0) = \begin{cases} u_0(x), & \text{for } x_0 < x < \infty \\ -u_0(2x_0 - x), & \text{for } -\infty < x < x_0. \end{cases}$$

The equation and the initial condition are unchanged under the change-of-variable: $x \to 2x_0 - x$, $u \to -u$. From the uniqueness of the solution, the solution has the property:

$$u(2x_0 - x, t) = -u(x,t).$$

From this, we can obtain that $u(x_0, t) = -u(x_0, t) = 0$.

Since $C = Ee^{\alpha x+\beta\tau}u_1$ is the vanilla call, where u_1 satisfies the heat equation with the initial condition:

$$u_1(x,0) = \begin{cases} e^{\frac{1}{2}(k+1)x} - e^{\frac{1}{2}(k-1)x} & \text{for } x > 0 \\ 0, & \text{for } x \le 0 \end{cases}$$

Using this and the method of reflection, we may express V in terms of C as the follows. First, we may write $V = Ee^{\alpha x+\beta\tau}(u_1 + u_2)$, where the initial condition for u_2 is the reflected condition from u_1:

$$u_2(x,0) = \begin{cases} u_0(2x_0 - x, 0) & \text{for } x \le 0 \\ 0 & \text{for } x > 0 \end{cases}$$

The solution u_1 corresponds to $C(S,t)$. The solution u_2 is corresponds to $e^{2\alpha(x-x_0)}C(x^2/S, t)$. We conclude

$$V = C(S,t) - (\frac{S}{X})^{-(k-1)}C(X^2/S, t).$$

7.4.2 down-and-in(knock-in) option

An "in" option becomes worthless unless the asset price reaches the barrier before expiry. If S crosses the line $S = X$ at some time prior to expiry, then the option becomes a vanilla option. It is common for in-type barrier option to give a rebate, usually a fixed amount, if the barrier is not hit. This compensates the holder for the loss of the option.

The boundary condition for an "in" option is the follows. The option is worthless as $S \to \infty$, i.e., $V(S,t) \to 0$ as $S \to \infty$. At T, if $S > X$, then $V(S,T) = 0$. For $t < T$, $V(X,t) = C(X,t)$. Since the option immediately turns into a vanilla call and must have the same value of this vanilla call. For $S \le X$, $V(S,t) = C(X,t)$. We only need to solve V for $S > X$. V still satisfies the same Black-Scholes equation for all S, t, because its randomness is fully correlated to the randomness of S.

We may write $V = c - \overline{V}$, where c is the value of a vanilla call. Then the boundary condition for $\overline{V}$ is $\overline{V}(S,t) = c - V \sim S - 0 = S$ as $S \to \infty$. And

$$\overline{V}(X,t) = c(X,t) - V(X,t) = 0, \quad \overline{V}(S,T) = c(S,T) - V(S,T) = c(S,T) = \Lambda(S).$$

We observe that $\overline{V}$ is indeed a "down-and-out" barrier option. In other words, 1(down-and-in) plus 1(down-and-out) equal to 1 vanilla call. This is because one and only one of the two barrier options can be active at expiry and whichever it is, its value is the value of a vanilla call.

7.5 Asian options and lookback options

In Asian options and lookback options, their payoff functions depend on the history of the underlying asset. For example,

1. a European-type average strike option has the following payoff function

$$\max\{S_T - \frac{1}{T}\int_0^T S(\tau)d\tau, 0\}.$$

2. an American-type average strike option,

$$\Lambda(S,t) = \max\{S - \frac{1}{t}\int_0^t S(\tau)d\tau, 0\}.$$

3. geometric mean,

$$\Lambda(S,T) = \max\{S - e^{\int_0^T \log S(\tau)d\tau}, 0\}.$$

4. Lookback call,

$$\Lambda(S,T) = \max\{S - J, 0\}, \qquad J = \max_{0\le\tau\le T} S(\tau)$$

In general, the payoff depends on I, which is defined by

$$I = \int_0^t f(S(\tau),\tau)d\tau,$$

where f is a smooth function. The payoff function is $\Lambda(S,I)$. It is important to notice that $I(t)$ is independent of $S(t)$. The value of an asian option should depend on S, t s well as I. Indeed, we shall see in the next chapter that $dI = f dt$. The only randomness is through S, therefore V can be valued through a delta hedge.

For the lookback option, it will be treated as a limiting case of an asian option. We shall discuss this in the next chapter.

Chapter 8

Path-Dependent Options

8.1 Introduction

If the payoff depends on the history of the underlying asset price, such an option is called a path-dependent option. The Asian options and the Russian options (Lookback options) are the typical examples. The payoff functions for these options are, for example,

1. average strike call option: $\Lambda = \max\{S - \frac{1}{T}\int_0^T S(\tau)d\tau, 0\}$,
2. average rate call option: $\Lambda = \max\{\frac{1}{T}\int_0^T S(\tau)d\tau - E, 0\}$
3. geometric mean: the arithmetic mean $\frac{1}{T}\int_0^T S(\tau)d\tau$ above is replaced by $e^{\int_0^T \log(S(\tau))d\tau}$.
4. lookback strike put: $\Lambda = \max\{\max_{0\le\tau\le T} S(\tau) - S, 0\}$.
5. lookback rate put: $\Lambda = \max\{E - \max_{0\le\tau\le T} S(\tau), 0\}$.

8.2 General Method

Let f be a smooth function, define

$$I(t) = \int_0^t f(S(\tau), \tau)d\tau.$$

In previous examples, $f(S(\tau), \tau) = S(\tau)$ for arithmetic mean and $f(S(\tau), \tau) = \log S(\tau)$ for geometric mean.

Notice that $I(t)$ is a random variable and is independent of $S(t)$. (This is because $I(t)$ is the sum of increment of functions of S before time t, and each increment of $S(\tau)$, $\tau < t$, is independent of $S(t)$.) Therefore, we should introduce another independent variable I besides S to value the derivative $V(S, I, t)$.

The stochastic differential equations governed by S and I are

$$\begin{aligned}\frac{dS}{S} &= \mu dt + \sigma dz,\\ dI(t) &= f(S(t),t)dt.\end{aligned}$$

Notice that there is no noize term in dI. The only randomness is from dS. Therefore, we can use delta hedge to eliminate this randomness. Namely, we consider the portfolio

$$\Pi = V - \Delta S,$$

as before. We have

$$d\Pi = dV - \Delta dS = (V_t + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 S}{\partial V^2})dt + V_I dI + V_S dS - \Delta dS.$$

We choose $\Delta = \frac{\partial V}{\partial S}$ to eliminate the randomness in $d\Pi$. From the arbitrage assumption, we arrive

$$V_t + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + f\frac{\partial V}{\partial I} = r(V - S\frac{\partial V}{\partial S}).$$

8.3 Average strike options

8.3.1 European calls

Let us consider an average strike call option with European exercise feature. Its payoff function is defined by

$$\max\{S - \frac{1}{T}\int_0^T S(\tau)d\tau, 0).$$

Or, in terms of I, $\Lambda(S, I, T) = \max\{S - \frac{I}{T}, 0\}$.

We notice that the modified Black-Scholes equation

$$V_t + S\frac{\partial V}{\partial I} + \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

and the initial data for the average strike options are invariant under the transformation: $(S, I) \to \lambda(S, I)$. Therefore, we expect that its solution is a function of the scale-invariant variable $R = I/S$. Notice that this is also reflected in that

$$dR = (1 + (\sigma^2 - \mu)R)dt - \sigma R dz,$$

depends on R only. Since the initial data can be expressed as $\Lambda = S\max\{1 - R/T, 0\}$, we may also expect that $V = SH(R, t)$. This reduces one independent variable.

Plug $V = SH$ into the above modified Black-Scholes equation:

$$SH_t + \frac{1}{2}\sigma^2 S^2(2\frac{\partial H}{\partial S} + S\frac{\partial^2 H}{\partial S^2}) + S\cdot S\frac{\partial H}{\partial I} + rS\frac{\partial}{\partial S}(SH) - r(SH) = 0.$$

From

$$\begin{aligned} \frac{\partial}{\partial S} &= \frac{\partial R}{\partial S}, \quad \frac{\partial}{\partial R} = -\frac{R}{S}\frac{\partial}{\partial R} \\ \frac{\partial^2}{\partial S^2} &= \frac{2I}{S^3}\frac{\partial}{\partial R} + \frac{I^2}{S^4}\frac{\partial^2}{\partial R^2} = \frac{1}{S^2}(2R\frac{\partial}{\partial R} + R^2\frac{\partial^2}{\partial R^2}), \end{aligned}$$

we obtain

$$\begin{aligned} & SH_t + \frac{1}{2}\sigma^2 S^2(2(-\frac{R}{S}\frac{\partial H}{\partial R}) + S\cdot\frac{1}{S^2}(2R\frac{\partial H}{\partial R} + R^2\frac{\partial^2 H}{\partial R^2})) \\ & +S^2\frac{1}{S}H_R + rSH + rS^2(-\frac{R}{S}H_R - rSH) = 0 \end{aligned}$$

Finally, we arrive

$$H_t + \frac{\sigma^2}{2}R^2\frac{\partial^2 H}{\partial R^2} + (1 - rR)\frac{\partial H}{\partial R} = 0.$$

The payoff function

$$\Lambda(R, T) = \max\{1 - \frac{R}{T}, 0\} = H(R, T), \text{ (final condition) }.$$

we should require the boundary conditions.

- $H(\infty, t) = 0$. Since as $R \to \infty$ implies $S \to 0$, then $V \to 0$, then $H \to 0$.
- $H(0, t)$ is finite. When $R = 0$, we have $S(\tau) = 0$, for all τ with probability 1. That implies that $\Lambda = 0$, and consequently, H is finte at $(0, t)$.

Next, we expect that the solution is smooth up to $R =$. This implies that H_R, H_{RR} are finite at $(0, t)$. We have the following two cases: (i) If $R^2\frac{\partial^2 H}{\partial R^2} = O(1) \neq 0$ as $R \to 0$ then $H = O(\log R)$ for $R \to 0$. Or (ii) if $R\frac{\partial H}{\partial R} = O(1) \neq 0$ as $R \to 0$, then $H = O(\log R)$. Both cases contradict to $H(0, t)$ being finite. Hence, we have $RH_R(0, t)$, $R^2 H_{RR}(0, t)$ are zeros as $R \to 0$. Hence the boundary condition $H(0, t)$ is finite is equivalent to $H_t(0, t) + H_R(0, t) = 0$.

This equation with boundary condition can be solved by using the hypergeometric functions. However, in practice, we solve it by numerical method.

8.3.2 American call options

We consider the average strike call option with American exercise feature. In this case,

$$\begin{aligned} H_t + \frac{1}{2}\sigma^2 R^2\frac{\partial^2 H}{\partial R^2} + (1 - rR)\frac{\partial H}{\partial R} &\leq 0 \\ H - \Lambda &\geq 0 \\ (H_t + \frac{1}{2}\sigma^2 R^2\frac{\partial^2 H}{\partial R^2} + (1 - rR)\frac{\partial H}{\partial R})(H - \Lambda) &= 0, \end{aligned}$$

where $\Lambda(R, t) = \max\{1 - \frac{R}{t}, 0\}$, $R(t) = I(t)/S(t)$, $I(t) = \int_0^t S(\tau)d\tau$.

8.3.3 Put-call parity for average strike option

We study the put-call parity for average strike options with European exercise feature. Consider a portfolio is $C - P$. The corresponding payoff function is

$$S\max\{1-\frac{R}{T},0\} - S\max\{\frac{R}{T}-1,0\} = S(1-\frac{R}{T}) \equiv S\cdot H(R,T).$$

Since the Black-Scholes equation in linear in R, we only need to solve the equation with final condition (i) $H(R,T) = 1$, (ii) $H(R,T) = -\frac{R}{T}$. For (i), $H(R,t) \equiv 1$. For (ii), since the final condition and the P.D.E. is linear in R, we expect that the solution is also linear in R. Thus, we consider H is of the following form

$$H(t,R) = a(t) + b(t)R$$

Plug this into the equation, we obtain

$$\frac{d}{dt}a + \frac{d}{dt}bR + (1-rR)b = 0.$$

and

$$\frac{d}{dt}a + b = 0, \quad \frac{d}{dt}b - rb = 0.$$

with $a(T) = 0$, $b(T) = -\frac{1}{T}$. This differential equation can be solved easily:

$$b(t) = -\frac{1}{T}e^{-r(T-t)}, \quad a(t) = -\frac{1}{rT}(1-e^{-r(T-t)}).$$

Consequently, we obtain the put-call parity:

$$\begin{aligned} C - P &= S(1-\frac{1}{rT}(1-e^{-r(T-t)} - \frac{1}{T}e^{-r(T-t)}\frac{1}{S}\int_0^t S(\tau)d\tau) \\ &= S - \frac{S}{rT}(1-e^{-r(T-t)}) - e^{-r(T-t)}\frac{1}{T}\int_0^t S(\tau)d\tau \end{aligned}$$

8.4 Lookback Option

A lookback option is a derivate product whose payoff depends on the maximum or minimum of its underlying asset price. For instance, the payoff function for a lookback option with European exercise feature is

$$\Lambda = \max_{0\le\tau\le T} S(\tau) - S(T).$$

Such an option is relatively expansive because it gives the holder an extremely advantageous payoff.

As before, let us introduce $J(t) = \max_{0\le\tau<t} S(\tau)$. Since $S(\tau), \tau < t$ are independent of $S(t)$, we see that $J(t)$ is independent of $S(t)$. This suggest that we should introduce another independent variable J to value the lookback option in addition to S and t. We can derive a stochastic differential equation for J as before. Indeed, it is $dJ = 0$. However, we shall give a more careful approach. We shall use the fact that for a continuous function $S(\cdot)$,

$$\max_{0\le\tau\le t} |S(\tau)| = \lim_{n\to\infty} \left(\int_0^t |S(\tau)|^n d\tau \right)^{\frac{1}{n}}.$$

We leave its proof as an exercise.

Remark. For the minimum, we have

$$\lim_{n\to-\infty} \left(\int_0^t (S(\tau))^n d\tau \right)^{\frac{1}{n}} = \min_{0\le\tau\le t} S(\tau).$$

Let us introduce

$$I_n = \int_0^t (S(\tau))^n d\tau, \quad J_n = I_n^{\frac{1}{n}}.$$

The s.d.e. for J_n,

$$\begin{aligned} dJ_n &= (\int_0^{t+dt} (S(\tau))^n d\tau)^{\frac{1}{n}} - (\int_0^t (S(\tau))^n d\tau)^{\frac{1}{n}} \\ &= \frac{1}{n} \frac{S(t)^n}{J_n^{n-1}} dt. \end{aligned}$$

Now, as before we consider the delta hedge:

$$\Pi = P - \Delta S.$$

From the arbitrage assumption, we can derive the equation for $P(S, J_n, t)$:

$$\begin{aligned} d\Pi &= P_t dt + \frac{1}{n} \frac{S^n}{J_n^{n-1}} \frac{\partial P}{\partial J_n} dt + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} dt \\ &= r(P - \frac{\partial P}{\partial S} S) dt \end{aligned}$$

Taking $n \to \infty$, using the facts that $J_n \to J$ and $\frac{S}{J} \le 1$, we arrive

$$P_t + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS \frac{\partial P}{\partial S} - rP = 0.$$

This is the usual Black-Scholes equation. The role of J here is only a parameter. This is consistent to the fact that

$$dJ = 0.$$

8.4.1 A lookback put with European exercise feature

The range for S is $0 \le S \le J$. This is because $S \le J$, for $0 \le t \le T$. We claim that

$$P(0, J, t) = Je^{-r(T-t)}.$$

Firstly, we have that $\Lambda(0, J, T) = \max\{J - S, 0\} = J$. Secondly, if $S(t) = 0$, then $S(\tau) = 0$ for $t \le \tau \le T$. The asset price process becomes deterministic. Therefore, the value of P is the discounted payoff: $P(0, J, t) = Je^{-r(T-t)}$.

Next, we claim that

$$\frac{\partial P}{\partial J}(J, J, t) = 0.$$

From $\mu > 0$, the current maximum cannot be the final maximum with probability 1. The value of P must be insensitive to a small change of J.

We can use method of image to solve this problem. Its solution is given by

$$P = S(-1 + N(d_7)(1 + k^{-1})) + Je^{-r(T-t)}N(d_5) - k^{-1}(\frac{S}{J})^{1-k}N(d_6),$$

where

$$\begin{aligned} d_5 &= [\ln(\frac{J}{S}) - (r - \frac{\sigma^2}{2})(T-t)]/\sigma\sqrt{T-t} \\ d_6 &= [\ln(\frac{S}{J}) - (r - \frac{\sigma^2}{2})(T-t)]/\sigma\sqrt{T-t} \\ d_7 &= [\ln(\frac{J}{S}) - (r + \frac{\sigma^2}{2})(T-t)]/\sigma\sqrt{T-t} \\ k &= \frac{r}{\sigma^2/2} \end{aligned}$$

8.4.2 Lookback put option with American exercise feature

We have the following linear complementary equation,

$$L_{BS}P \le 0, \quad P - \Lambda \ge 0, \quad (L_{BS}P)(P - \Lambda) = 0,$$

where

$$L_{BS} = \frac{\partial}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2}{\partial S^2} + rS\frac{\partial}{\partial S} - r.$$

The final condition

$$P(S, J, T) = \Lambda(S, J, T) = J - S.$$

The boundary condition

$$\frac{\partial P}{\partial J}(J, J, t) = 0.$$

We require P, $\frac{\partial P}{\partial S}$ $\frac{\partial P}{\partial J}$ are continuous.

For lookback call option, we simply replace $\max_{0\le\tau\le t} S(\tau)$ by $\min_{0\le\tau\le t} S(\tau)$. For instance, its payoff is $\Lambda = S(T) - \min_{0\le\tau\le T} S(\tau)$.

Chapter 9

Bonds and Interest Rate Derivatives

9.1 Bond Models

A bond is a long-term contract under which the issuer promises to pay the bondholder coupon payment (usually periodically) and principal (at the maturity dates). If there is no coupon payment, the bond is called a zero-coupon bond. The principal of a bond is called its face value.

9.1.1 Deterministic bond model

The value of a bond certainly depends on the interest rate. Let us first assume that the interest rate is deterministic temporarily, say $r(\tau)$, $t \leq \tau \leq T$, is known. Let $B(t,T)$ be the bond value at t with maturity date T, $k(t)$ be its coupon rate. This means that in a small dt, the holder receives coupon payment $k(t)\,dt$. From the no-arbitrage argument,

$$dB + k(t)\,dt = r(t)Bdt.$$

together with the final condition:

$$B(T,T) = F(\text{ face value}),$$

the bond value can be solved and has the following expression:

$$B(t,T) = e^{-\int_t^T r(\tau)\,d\tau}\left[F + \int_t^T k(\tau)e^{\int_\tau^T r(s)\,ds}\,d\tau\right].$$

Thus, the bond value is the sum of the present face value and the coupon stream.

However, the life span of a bond is long (usually 10 years or longer), it is unrealistic to assume that the interest rate is deterministic. In the next subsection, we shall provide a stochastic model.

9.1.2 Stochastic bond model

Let us assume that the interest rate satisfies the following s.d.e.:

$$dr = u(r,t)dt + w(r,t)dz,$$

where dz is the standard Wiener process. The drift u and the variance w^2 are proposed by many researchers. We shall discuss these issues later. To find the equation for B with stochastic property of r, we consider a portfolio containing bonds with different maturity dates:

$$\Pi = V(t, r, T_1) - \Delta V(t, r, T_2) \equiv V_1 - \Delta V_2.$$

The change $d\Pi$ in a small time step dt is

$$d\Pi = dV_1 - \Delta dV_2,$$

where

$$\begin{aligned} \frac{dV_i}{V_i} &= \mu_i dt + \sigma_i dz, \\ \mu_i &= \frac{1}{V_i}\left(V_{i,t} + uV_{i,r} + \widehat{f} w^2 V_{i,rr}\right) \\ \sigma_i &= \frac{1}{V_i} w V_{i,r}. \end{aligned}$$

We we choose $\Delta = V_{1,r}/V_{2,r}$, then the random term is canceled in $d\Pi$. From the no-arbitrage argument, $d\Pi = r\Pi dt$. We obtain

$$\mu_1 V_1 dt - \Delta \mu_2 V_2 dt = r(V_1 - \Delta V_2)\, dt.$$

This yields

$$(\mu_1 - r)V_1/V_{1,r} = (\mu_2 - r)V_2/V_{2,r},$$

or equivalently

$$\frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2}.$$

Since the left-hand side is a function of T_1, while the right-hand side is a function of T_2. Therefore, it is independent of T. Let us express it as a known function $\lambda(r, t)$:

$$\frac{\mu - r}{\sigma} = \lambda(r, t).$$

Plug μ_i and σ_i back to this equation, and drop the index i, we obtain

$$V_t + \frac{1}{2} w^2 V_{rr} + (u - \lambda w)V_r - rV = 0.$$

The function $\lambda(r, t) = \frac{\mu - r}{\sigma}$ is called the market price, since it gives the extra increase in expected instantaneous rate of return on a bond per an additional unit of risk.

This stochastic bond model depends on three parameter functions $u(r, t)$, $w(r, t)$ and $\lambda(r, t)$. In the next section, we shall provide some model to determine them.

9.2 Interest models

There are many interest rate models. We list some of them below.

- Merton (1973): $dr = \alpha dt + \sigma dz$.
- Vasicek (1977): $dr = \beta(\alpha - r)dt + \sigma dz$.
- Dothan (1978): $dr = \sigma r dz$.
- Marsh-Rosenfeld (1983): $dr = (\alpha r^{\delta-1} + \beta r)dt + \sigma r^{\delta/2} dz$.
- Cox-Ingersoll-Ross (1985) $dr = \beta(\alpha - r)dt + \sigma r^{\widehat{f}} dz$.
- Ho-Lee (1986): $dr = \alpha(t)dt + \sigma dz$.
- Black-Karasinski (1991): $d\ln r = (a(t) - b(t)\ln r)\, dt + \sigma dz$.

The main requirements for an interest rate model are

- positivity: $r(t) \geq 0$ almost surely,
- mean reversion: r should tends to increase (or to decrease) and toward a mean.

The C-I-R and B-K models have these properties.

Below, we shall illustrate a unified approach proposed by Luo, Yen and Zhang.

9.2.1 A functional approach for interest rate model

The idea is to design r to be a function of $x(t)$ and t, (i.e. $r = r(x(t), t)$) with $x(t)$ governed by a simple stochastic process. We notice that the Ornstein-Uhlenbeck process $dx = -\eta x dt + \sigma dz$ has the property to tend to its mean (which is 0) time asymptotically. While the Bessel process $dx = \epsilon/x dt + \sigma dz$ has positive property. We then design the underlying basic process is the sum of these two processes:

$$dx = \left(-\eta x + \frac{\epsilon}{x}\right) dt + \sigma dz.$$

In general, we allow η, ϵ, and σ are given functions of t. With this simple process, we can choose $r = r(x(t), t)$. Then all interest models mentioned above correspond to different choices of $r(x, t)$, η, ϵ and ϵ.

- Merton's model: we choose $\epsilon = \eta = 0$, $r = x + \alpha t$.
- Dothan's model: $\epsilon = \eta = 0$, $r = e^{x - \sigma^2/2t}$.
- Ho-Lee: $\epsilon = \eta = 0$, $r = x + \int_0^t \alpha(s)\, ds$.
- Vasicek: $\epsilon = 0$, $\eta = \beta$, $r = x + \alpha$.
- C-I-R: $\beta = 2\eta$, $\alpha = (\sigma^2 + 2\epsilon)/(8\eta)$, $r = \frac{1}{4}x^2$.

- Black-Karasinski: $\epsilon = 0$, $r = \exp(g(t,x))$, where $g = x + \int_0^t a(s)\,ds$, $\eta = b(t)$.

With the interest rate model, the zero-coupon bond price V is given by

$$V(x,t) = E\left(e^{\int_t^T r(s,x(s))\,ds} \mid x_t = x\right), t < T.$$

From the Feymann-Kac formula, V satisfies

$$\left[\frac{\partial}{\partial t} + \frac{1}{\sigma^2}\frac{\partial^2}{\partial x^2} + \left(-\eta x + \frac{\epsilon}{x}\right)\frac{\partial}{\partial x} - r\right] V = 0.$$

This model depends three parameter functions $\epsilon(t), \eta(t)$, $\sigma(t)$, and $r(x,t)$. There is no unified theory available yet with this approach and the approach of the previous subsection.

9.3 Convertible Bonds

A convertible bond is a bond plus a call option under which the bond holder has the right to convert the bond into a common shares. Thus, it is a function of r, S, t and T. Let the stochastic processes governed by S and r are

$$\begin{aligned} \frac{dS}{S} &= \mu dt + \sigma dz_S, \\ dr &= u dt + w dz_r. \end{aligned}$$

Suppose the correlation between dz_S and dz_R is

$$dz_S\, dz_r = \rho(S,r,t)dt.$$

The final value of the convertible bond $V(r,S,T) = F$, the face value of the bond. Suppose the bond can be converted to nS at any time priori to T. Then we have

$$V(r,S,t) \geq nS.$$

We also have the boundary conditions:

$$\begin{aligned} \lim_{S\to\infty} V(r,S,t) &= nS \\ \lim_{r\to\infty} V(r,S,t) &= 0. \end{aligned}$$

At $S = 0$ or $r = 0$, we should require $V(r,0,t)$ or $V(0,S,t)$ to be finite.

The Black-Scholes analysis for a convertible bond is similar to the analysis for a bond. Let $V_i = V(r,S,t,T_i)$, $i = 1,2$. Consider a portfolio

$$\Pi = \Delta_1 V_1 + \Delta_2 V_2 + \Delta_S S.$$

In a small time step dt, the change of $d\Pi$ is

$$d\Pi = \Delta_1 dV_1 + \Delta_2 dV_2 + \Delta_S dS,$$

where

$$\frac{dV_i}{V_i} = \mu_i dt + \sigma_{r,i} dz_r + \sigma_{S,i} dz_S,$$

$$\begin{aligned}
\mu_i &= \frac{1}{V_i}\left(V_{i,t} + \frac{\sigma^2}{2}S^2 V_{i,SS} + \rho S w V_{i,Sr} + \frac{w^2}{2}V_{i,rr} + \mu S V_{i,S} + u V_{i,r}\right), \\
\sigma_{S,i} &= \frac{1}{V_i}\sigma S V_{i,S} \\
\sigma_{r,i} &= \frac{1}{V_i} w V_{i,r}.
\end{aligned}$$

This implies

$$\begin{aligned}
d\Pi &= (\Delta_1 \mu_1 V_1 + \Delta_2 \mu_2 V_2 + \Delta \mu S)\, dt \\
&\quad + (\Delta_1 \sigma_{S,1} V_1 + \Delta_2 \sigma_{S,2} V_2 + \Delta \sigma S)\, dz_S \\
&\quad + (\Delta_1 \sigma_{r,1} V_1 + \Delta_2 \sigma_{r,2} V_2)\, dz_r.
\end{aligned}$$

We choose Δ_1, Δ_2 and Δ_S to cancel the randomness terms dz_r and dz_S. This means that

$$\begin{aligned}
(\Delta_1 \sigma_{S,1} V_1 + \Delta_2 \sigma_{S,2} V_2 + \Delta \sigma S)\, dz_S &= 0 \\
(\Delta_1 \sigma_{r,1} V_1 + \Delta_2 \sigma_{r,2} V_2)\, dz_r &= 0.
\end{aligned}$$

And it yields

$$\begin{aligned}
d\Pi &= (\Delta_1 \mu_1 V_1 + \Delta_2 \mu_2 V_2 + \Delta \mu S)\, dt \\
&= r(\Delta_1 V_1 + \Delta_2 V_2 + \Delta S)\, dt.
\end{aligned}$$

Or equivalently,

$$\Delta_1(\mu_1 - r)V_1 + \Delta_2(\mu_2 - r)V_2 + \Delta(\mu - r)S = 0.$$

This equality together with the previous two give that there exist λ_r and λ_S such that

$$\begin{aligned}
(\mu_1 - r) &= \lambda_S \sigma_{S,1} + \lambda_r \sigma_{r,1} \\
(\mu_2 - r) &= \lambda_S \sigma_{S,2} + \lambda_r \sigma_{r,2} \\
(\mu - r) &= \lambda_S \sigma
\end{aligned}$$

The functions λ_r and λ_S are called the market prices of risk with respect to r and S, respectively. Plug the formulae for μ_i, $\sigma_{r,i}$ and $\sigma_{S,i}$, we obtain the Black-Scholes equation for a convertible bond:

$$\left[\frac{\partial}{\partial t} + \frac{\sigma^2}{2}S^2\frac{\partial^2}{\partial S^2} + \rho S w\frac{\partial^2}{\partial S \partial r} + \frac{w^2}{2}\frac{\partial^2}{\partial r^2} + rS\frac{\partial}{\partial S} + (u - \lambda_r w)\frac{\partial}{\partial r} - r\right]V = 0.$$

Appendix A

Basic theory of stochastic calculus

A.1 From random walk to Brownian motion

This section gives a more rigorous approach to construct Brownian motion from random walk. Without loss of generality, we shall construct motion $B(t)$ for $t \in [0,1]$. Let us consider a hierarchical grids: for $\ell \geq 1$ integer, define

$$t_j^\ell := j2^{-\ell}, j = 1, ..., 2^\ell,$$

A hierarchical independent random variables Δ_j^ℓ are i.i.d.: $\Delta_j^\ell \sim \Delta$,

$$\Delta := \begin{cases} 1 & \text{prob. } 1/2 \\ -1 & \text{prob. } 1/2. \end{cases}$$

This corresponds to a binomial trial with sample space

$$\Omega^\ell := \{(a_1^\ell, ..., a_{2^\ell}^\ell) | a_j^\ell = 1 \text{ or } -1\}$$

The σ-algebra $\mathcal{F}^\ell$ is defined to be $\mathcal{F}^\ell := 2^{\Omega^\ell}$. There are $2^{(2^\ell)}$ elements in Ω^ℓ. The Probability P^ℓ at each sample in Ω^ℓ is defined to be 2^{-2^ℓ}.

We define the random variables

$$x_j^\ell := \frac{1}{\sqrt{2^\ell}} \sum_{k=1}^{j} \Delta_k^\ell, \quad j = 1, ..., 2^\ell,$$

and define a path x^ℓ which is piecewise linear and its values at grid points are defined to be

$$x^\ell(t_j^\ell) := x_j^\ell.$$

Each sample $\omega \in \Omega^\ell$ corresponds to a unique piecewise linear path. We may also imagine the sample space Ω^ℓ is the collection of such paths.

The grids $G^\ell := \{t_j^\ell\}_{j=0}^{2^\ell}$, the path x^ℓ, the sample space Ω^ℓ and the σ-algebra $\mathcal{F}^\ell$ have the nested structure:

(1) $G^{\ell-1} \subset G^{\ell}$ with $t_j^{\ell-1} = t_{2j}^{\ell}$;

(2) $\Omega^{\ell-1} \subset \Omega^{\ell}$ with $a_{2j-1}^{\ell} = a_{2j}^{\ell} = a_j^{\ell-1}$;

(3) $\mathcal{F}^{\ell-1} \subset \mathcal{F}^{\ell}$.

(4) The probability P^{ℓ} on $\mathcal{F}^{\ell}$, as restricted to $\mathcal{F}^{\ell-1}$ is identical to $P^{\ell-1}$

With this nested structure, we can define the Brownian motion to be

$$\Omega = \bigcup_{\ell} \Omega^{\ell}, \quad \mathcal{F} = \bigcup_{\ell} \mathcal{F}^{\ell},$$

with probability $P := P^{\ell}$ on F^{ℓ}.

The Brownian motion $B(t)$ has the following properties:

1. $B(0) = 0$ and $B(t)$ is continuous. Indeed, we should expect that $B(\cdot)$ is Hölder continuous continuous of order $1/2$;

2. If $t_1 < t_2 \le t_3 < t_4$, then $B(t_4) - B(t_3)$ and $B(t_2) - B(t_1)$ are independent;

3. $B(s+t) - B(s) \sim \mathcal{N}(0, t)$. This is mainly due to the central limit theorem.

A.2 Brownian motion

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A process is a function $X : [0, \infty) \times (\Omega, \mathcal{F}) \mapsto (E, \mathcal{B})$, such that for each $t \ge 0$, $X(t)$ is a random variable. Here, $E = R^d$, $\mathcal{B}$ is the set of all Borel sets in E. Let

$$\begin{aligned} \mathcal{F}_t &= \sigma\{X(s), s \le t\} \\ \mathcal{F}^t &= \sigma\{X(s), s \ge t\} \end{aligned}$$

A process is called Markov if

$$P(A \mid \mathcal{F}_t) = P(A \mid X(t)), \forall A \in \mathcal{F}^t.$$

This is equivalent to

$$P\{X(r) \in B | \mathcal{F}_t\} = P\{X(r) \in B | X(t)\} \quad \forall r > t,$$

A Markov process is characterized by its transition probability:

$$P(t, x, s, B) := P\{X(s) \in B \mid X(t) = x\}$$

with initial distribution

$$P\{X(0) \in B\} = \nu(B).$$

Theorem 1.10. *If X is a Markov process, then the corresponding transition probability P satisfies Chapman-Kolmogorov equation:*

$$\int P(t_0, x_0, t_1, dx_1) P(t_1, x_1, t_2, B)\, dx_1 = P(t_0, x_0, t_2, B).$$

Conversely, if P is a function satisfies Chapman-Kolmogorov equation, then there is a Markov process whose transition probability is P.

Two standard Markov processes are the Wiener process and the Poisson process. The Wiener process has the transition probability density function

$$p(t, x, s, y) = \frac{1}{(2\pi(t-s))^{d/2}} \exp(-\frac{|x-y|^2}{2(s-t)}).$$

Such a distribution is called a normal distribution with mean x and variance $s - t$.

Definition 1.8. *Brownian motion: A process is called a Brownian motion(or a Wiener process) if*

1. *$B_t - B_s$ has normal distribution with mean 0 and variance $t - s$,*
2. *it has independent increments: that is $B_t - B_s$ is independent of B_u for all $u \leq s < t$,*
3. *B_t is continuous in t.*

It is easy to see that B_t is markovian and its transition probability is

$$p(t, x, s, y) = \frac{1}{(2\pi(t-s))^{d/2}} \exp(-\frac{|x-y|^2}{s-t}).$$

Definition 1.9. *A process $\{X_t \mid t \geq 0\}$ is called martingale if*

(i) $EX_t < \infty$,

(ii) $E(X_u \mid X_s, 0 \leq s \leq t\} = X_t$.

This means that if we know the value of the process up to time t and $X_t = x$, then the future expectation of X_u is x.

Theorem 1.11. *1. B_t is a martingale.*

2. $B_t^2 - t$ is a martingale.

Proof. 1. $E(B_t) = 0$. From the fact that $B_{t+s} - B_t$ is independent of B_t, for all $s > 0$, we obtain $E(B_{t+s} - B_t \mid B_t) = 0$, for all $s > 0$. Hence,

$$\begin{aligned} E(B_{t+s} \mid B_t) &= E((B_{t+s} - B_t) + B_t \mid B_t) \\ &= E((B_{t+s} - B_t) \mid B_t) + E(B_t \mid B_t) \\ &= B_t \end{aligned}$$

2. $E(B_t^2) = t < \infty$.

3. Use

$$\begin{aligned} B_{t+s}^2 &= ((B_{t+s} - B_t) + B_t)^2 \\ &= (B_{t+s} - B_t)^2 + 2B_t(B_{t+s} - B_t) + B_t^2 \end{aligned}$$

and the fact that $B_{t+s} - B_t$ is independent of B_t, we obtain

$$E(B_{t+s}^2 \mid B_t) = s + B_t^2.$$

Hence,

$$E(B_{t+s}^2 - (t+s) \mid B_t) = B_t^2 - t.$$

□

We can show that the Brownian motion has infinite total variation in any interval. This means that

$$\lim \sum |B(t_i) - B(t_{i-1})| = \infty.$$

However, its quadratic variation, defined by

$$[B,B](a,b) := \lim \sum |B(t_i) - B(t_{i-1})|^2$$

is finite. Here, $a = t_0 < t_1 < \cdots < t_n = b$ is a partition of (a,b), and the limit is taken to be $\max(t_i - t_{i-1}) \to 0$.

Theorem 1.12. *We have $[B,B](0,t) = t$ almost surely.*

Proof. Let us partition $(0,t)$ evenly into 2^n subintervals. Let $T_n = \sum_{i=1}^{2^n} |B(t_i) - B(t_{i-1})|^2$. We see that

$$ET_n = \sum E(|B(t_i) - B(t_{i-1})|^2)) = \sum |t_i - t_{i-1}| = t.$$

$$\begin{aligned} Var(T_n) &= Var(\sum |B(t_i) - B(t_{i-1})|^2) \\ &= \sum Var((B(t_i) - B(t_{i-1}))^2) \\ &= 2\sum (t_i - t_{i-1})^2 \\ &= 2t^2 2^{-n} \end{aligned}$$

Hence, $\sum_{n=1}^{\infty} Var(T_n) < \infty$. From Fubini theorem,

$$E\left(\sum_{n=1}^{\infty} (T_n - ET_n)^2\right) < \infty.$$

This implies $E((T_n - ET_n)^2) \to 0$ and hence, $T_n \to ET_n$ almost surely. □

A.3 Stochastic integral

We shall define the integral

$$\int_0^t f(s)dB(s).$$

The method can be applied with B replaced by a martingale, or a martingale plus a function with finite total variation. We shall require that $f \in H^2$, which means:

(i) $f(t)$ depends only on the history $\mathcal{F}_t$ of B_s, for $s \le t$,

(ii) $\int_0^t E|f|^2 < \infty$.

For this kind of functions, we can define its Itô's integral as the follows.

1. We define Itô's integral for $f \in H^2$ and being step functions. Its Itô's integral is define by

$$\int_0^t f(s)\,dB(s) = \sum_{i=0}^n f(t_{i-1})(B(t_i) - B(t_{i-1})).$$

2. We use above step functions f_n to approximate general function $f \in H^2$. Using the fact that, for step functions $g \in H^2$,

$$E\left(\left(\int_0^t g(s)\,dB(s)\right)^2\right) = \int_0^t E|g(s)|^2\,ds,$$

one can show that

$$\lim_{n\to\infty}\int_0^t f_n(s)\,dB(s)$$

almost surely.

An typical example is

$$\int_0^t B(s)\,dB(s) = \frac{1}{2}B(t)^2 - \frac{t}{2}.$$

From the definition, the integral can be approximated by

$$I_n = \sum_{i=1}^n B(t_{i-1})(B(t_i) - B(t_{i-1}).$$

We have

$$\begin{aligned} I_n &= \frac{1}{2}\sum_{i=1}^n \left[B^2(t_i) - B^2(t_{i-1}) - (B(t_i) - B(t_{i-1}))^2\right] \\ &= \frac{1}{2}B^2(t) - \frac{1}{2}\sum_{i=1}^n (B(t_i) - B(t_{i-1}))^2 \end{aligned}$$

We have seen that the second on the right-hand side tends to $\frac{1}{2}t$ almost surely.

A.4 Stochastic differential equation

A stochastic differential equation has the form

$$dX_t = a(X_t,t)dt + b(X_t,t)dB(t) \tag{1.1}$$

A Markov process X is said to be a strong solution of this s.d.e. if it satisfies

$$X_t - X_0 = \int_0^t a(X_s,s)\,ds + \int_0^t b(X_s,s)dB(s).$$

Theorem 1.13 (Itô's formula)**.** *If X satisfies the s.d.e. $dX = adt + bdB$, then*

$$df(X(t)) = (af'(X(t)) + \frac{1}{2}b^2 f''(X(t)))dt + f'(X(t))bdB(t).$$

We shall give a brief idea of the proof. In a small time step (t_{i-1},t_i), let $\Delta B = B(t_i) - B(t_{i-1})$, $\Delta t = t_i - t_{i-1}$. We have

$$f(X(t_{i-1}) + \Delta X) - f(X(t_{i-1})) = f'\Delta X + \frac{1}{2}f''(\Delta X)^2 + o((\Delta X)^2).$$

We notice that

$$\begin{aligned}(\Delta X)^2 &= (a\Delta t + b\Delta B)^2\\ &= a^2(\Delta t)^2 + 2ab\Delta t + b^2(\Delta B)^2\\ &\approx b^2\Delta t + o(\Delta t).\end{aligned}$$

Plug ΔX and $(\Delta X)^2$ into the Taylor expansion formula for f. This yields the Itô's formula.

A.5 Diffusion process

For a s.d.e.(1.1), we define the operator T_t on functions $f \in L^2(\mathbb{R})$ by

$$T_t f = E_{x,t}(f(X(s)), s > t.$$

The operator T_t maps $L^2(\mathbb{R})$ to itself. It is a semigroup, that is

$$T_{t_1} \circ T_{t_2} = T_{t_1+t_2}.$$

This can be proven from the Markovian property of $X(\cdot)$. Indeed, in terms of the transition probability density function $p(t,x,s,y)$,

$$T_t f = \int p(t,x,s,y)f(y)\,dy.$$

For a semigroup T_t, we define its generator as

$$Lf := \lim_{t\to 0}\frac{T_t f - f}{t}.$$

From Itô's formula,

$$Lf := af' + \frac{1}{2}b^2 f''.$$

This follows from Itô's formula and $E(\int_0^t g(s)\,dB(s) = 0)$.

With a fixed $s > t$ and f, we define

$$u(x,t) := (T_t f)(x,t) = E_{x,t}(f(X(s)).$$

Theorem 1.14. *If X satisfies the s.d.e. (1.1), then the associate $u(x,t) := E_{x,t}(f(X(s))$, $s > t$, satisfies the backward diffusion equation:*

$$u_t + Lu = 0,$$

and has final condition $u(s,x) = f(x)$.

Proof. First, we notice that

$$\begin{aligned} u(x,t-h) &= E_{x,t-h}(f(X(s)) \\ &= E_{x,t-h}E_{X(t),t}(f(X(s)) \\ &= E_{x,t-h}(u(X(t),t)) \end{aligned}$$

This shifts final time from s to t. Now, we consider

$$\begin{aligned} \frac{u(x,t-h)-u(x,t)}{h} &= \frac{1}{h}E_{x,t-h}\left(u(X(t),t) - u(X(t-h),t)\right) \\ &= \frac{1}{h}\int_{t-h}^{t} Lu(X(s),t)\,ds \\ &\to Lu(x,t) \end{aligned}$$

as $h \to 0+$. Next, $u(x,s-h) = E_{x,s-h}f(X(s))$, we have $X(s) \to x$ as $h \to 0+$ almost surely. Hence $u(x,s-h) \to f(x)$ as $h \to 0+$. □

Since u can be represent as

$$u(x,t) = \int p(x,t,s,y)f(y)\,dy,$$

we obtain that p satisfies

$$p_t + L_x p = 0,$$

and

$$p(s,x,s,y) = \delta(x-y).$$

For diffusion equation with source term, its solution can be represented by the following Feymann-Kac formula.

Theorem 1.15 (Feymann-Kac). *If X satisfies the s.d.e. (1.1), then the associate*

$$u(x,t) := E_{x,t}\left[f(X(s))\exp\int_t^s g(X(\tau),\tau)\,d\tau\right],\ s>t \tag{1.2}$$

solves the backward diffusion equation:

$$u_t + Lu + gu = 0,$$

with final condition $u(s,x) = f(x)$.

Proof. As before, we shift final time from s to t:

$$\begin{aligned}
u(x,t-h) &= E_{x,t-h}\left[f(X(s))\exp\int_{t-h}^s g(X(\tau),\tau)\,d\tau\right]\\
&= E_{x,t-h}E_{X(t),t}\left[f(X(s))\exp\int_{t-h}^s g(X(\tau),\tau)\,d\tau\right]\\
&= E_{x,t-h}\left[E_{X(t),t}\left[f(X(s))\exp\int_t^s g(X(\tau),\tau)\,d\tau\right]\cdot E_{X(t),t}\left[\exp\int_{t-h}^t g(X(\tau),\tau)\,d\tau\right]\right]\\
&= E_{x,t-h}\left[u(X(t),t)\exp\int_{t-h}^t g(X(\tau),\tau)\,d\tau\right].
\end{aligned}$$

Here, we have used independence of X in the regions $(t-h,t)$ and (t,s). Now, from Itô's formula, we have

$$\begin{aligned}
u(x,t-h)-u(x,t) &= E_{x,t-h}\left[u(X(t),t)\exp\int_{t-h}^t g(X(\tau),\tau)\,d\tau - u(X(t-h),t)\right]\\
&= E_{x,t-h}\int_{t-h}^t d\left(u(X(s),t)\exp\int_{t-h}^s g(X(\tau),\tau)\,d\tau\right)\\
&= E_{x,t-h}\int_{t-h}^t \left(Lu(X(s),t)+u(X(s),t)g(X(s),s)\right)\,ds.
\end{aligned}$$

From this, it is easy to see that

$$\lim_{h\to 0+}\frac{u(x,t-h)-u(x,t)}{h} = Lu(x,t)+g(x,t)u(x,t).$$

□