Computational Finance

Lectured by Panos Parpas

Typed by Aris Zhu Yi Qing

December 7, 2022

Contents

1 Introduction

- **Forward contract**: a customized *binding* contract to buy/sell an asset at a specific price (**forward price**) on a *future* date.
- **American option**: gives the holder the *right* to exercise the option at any time before the date of expiration.
- **European option**: can only be exercised at the date of expiration.
- **call/put option**: gives the holder the *right* to buy/sell.
- **Derivative asset**: assets defined in terms of underlying financial asset.

2 The Binomial Model

2.1 The One Period Model

2.1.1 Definitions

- **Bond** price process is deterministic and given by

$$\begin{aligned} B_0 &= 1, \\ B_1 &= 1 + R. \end{aligned}$$

  where constant R is the spot rate for the period.

- **Stock** price process is stochastic and given by

$$\begin{aligned} S_0 &= s, \\ S_1 &= s \cdot Z \end{aligned}$$

  where Z is a stochastic variable defined as

$$Z = \begin{cases} u, & \text{with probability } p_u, \\ d, & \text{with probability } p_d. \end{cases}$$

  with $p_u + p_d = 1$ and the assumption that $d < u$.

- **Portfolio** on the (B, S) market is a vector

$$h = (x, y).$$

  with a deterministic market value at $t = 0$ and stochastic value at $t = 1$.

- **Value Process** of the portfolio h is defined as

$$V_t^h = xB_t + yS_t, \quad t = 0, 1,$$

or, in more detail,

$$V_0^h = x + ys,$$
$$V_1^h = x(1+R) + ysZ.$$

- **Arbitrage** portfolio is an h with the properties

$$V_0^h = 0,$$
$$V_1^h > 0, \quad \textit{with probability 1}$$

2.1.2 Contingent Claim Pricing

Proposition 1. The model h is free of arbitrage $\iff d \leq (1+R) \leq u$.

Comments The above proposition implies that $(1+R)$ is a convex combination of u and d, i.e.

$$1 + R = q_u \cdot u + q_d \cdot d,$$

where q_u and q_d can be interpreted as probabilities for a new probability measure Q with $P(Z = u) = q_u$ and $P(Z = d) = q_d$. Denoting expectation w.r.t. this measure by E^Q and with the following calculation

$$\frac{1}{1+R} E^Q[S_1] = \frac{1}{1+R}[q_u su + q_d sd] = \frac{1}{1+R} \cdot s(1+R) = s,$$

we arrive at the following relation

$$s = \frac{1}{1+R} E^Q[S_1].$$

Definition 2. A probability measure Q is **martingale** if the following condition holds:

$$S_0 = \frac{1}{1+R} E^Q[S_1].$$

Proposition 3. Arbitrage-free model $\iff \exists$ martingale measure Q.

Proposition 4. For the binomial model above, the martingale probabilities are

$$\begin{cases} q_u = \dfrac{(1+R) - d}{u - d}, \\ q_d = \dfrac{u - (1+R)}{u - d}. \end{cases}$$

Proof. Solve the system of equation

$$\begin{cases} 1 + R = q_u \cdot u + q_d \cdot d \\ q_u + q_d = 1. \end{cases}$$

$\square$

Definition 5. A **contingent claim** (financial derivative) is any stchastic variable X of the form $X = \Phi(Z)$, where Z is the stochastic variable driving the stock price process above, Φ is the **contract function**.

Example 6. European call option on the stock with strike price K. Assuming that $sd < K < su$, we have

$$X = \begin{cases} su - K, & \text{if } Z = u, \\ 0, & \text{if } Z = d, \end{cases}$$

so the contract function is given by

$$\Phi(u) = su - K,$$
$$\Phi(d) = 0.$$

Definition 7. A given contingent claim X can be **replicated** / is **reachable** if

$$\exists h \text{ s.t. } V_1^h = X, \text{ with probability 1.}$$

We call such an h a **hedging**/**replicating** portfolio. If all claims can be replicated, then the market is **complete**.

Proposition 8. If a claim X is reachable with replicating portfolio h, then the only reasonable price process for X is

$$\Pi_t[X] = V_t^h,\ t = 0, 1.$$

Here, "reasonable" means that $\Pi_0[X] \neq V_0^h \Rightarrow$ arbitrage possibility.

Proposition 9. The general binomial model is free of arbitrage $\Rightarrow$ it is complete.

Proof. Say a claim X has contract function Φ s.t.

$$V_1^h = \begin{cases} \Phi(u), & \text{if } Z = u \\ \Phi(d), & \text{if } Z = d, \end{cases}$$

we can obtain the following system of equations by expanding V_1^h

$$(1+R)x + suy = \Phi(u),$$
$$(1+R)x + sdy = \Phi(d),$$

and solve it to find out the replicating portfolio as

$$x = \frac{1}{1+R} \cdot \frac{u\Phi(d) - d\Phi(u)}{u-d},$$
$$y = \frac{1}{s} \cdot \frac{\Phi(u) - \Phi(d)}{u-d}.$$

□

Proposition 10. If the binomial model is free of arbitrage, then the arbitrage free price of a contingent claim X is

$$\Pi_0[X] = \frac{1}{1+R} E^Q[X],$$

where the martingale measure Q is uniquely determined by the relation

$$S_0 = \frac{1}{1+R} E^Q[S_1].$$

Proof. Using the results derived previously, we can find out that

$$\begin{aligned}\Pi_0[X] &= x + sy \\ &= \frac{1}{1+R}\left[\frac{(1+R)-d}{u-d} \cdot \Phi(u) + \frac{u-(1+R)}{u-d} \cdot \Phi(d)\right] \\ &= \frac{1}{1+R}\left[\Phi(u)q_u + \Phi(d)q_d\right] \\ &= \frac{1}{1+R} E^Q[X].\end{aligned}$$

□

2.2 The Multiperiod Model

2.2.1 Definitions

- Model
  - The bond price dynamics is given by

    $$B_{n+1} = (1+R)B_n, \quad B_0 = 1.$$

  - The stock price dynamics is given by

    $$S_{n+1} = S_n \cdot Z_n, \quad S_0 = s.$$

    where $Z_0, Z_1, \ldots, Z_{T-1}$ are assumed to be i.i.d. stochastic variables, with only two variables u and d and probabilities

    $$P(Z_n = u) = p_u, \quad P(Z_n = d) = p_d.$$

  - Illustrating the stock dynamics by means of a tree, we can see that it is **recombining**.
- A **portfolio strategy** is a stochastic process

  $$\{h_t = (x_t, y_t);\ t = 1, \ldots, T\}$$

  s.t. h_t is a function of $S_0, S_1, \ldots, S_{t-1}$. By convention, $h_0 = h_1$.
  - Interpretation of h_t: at $t-1$, x_t and y_t of bonds and stocks are bought and held until time t.

- The **value process** corresponding to the portfolio h is defined by

$$V_t^h = x_t(1+R) + y_t S_t.$$

- A portfolio strategy h_t is said to be **self-financing** if $\forall t = 0, \dots, T-1$,

$$x_t(1+R) + y_t S_t = x_{t+1} + y_{t+1} S_t$$

which says that the market value of the "old" portfolio h_t equals to the purchase value of the "new" portfolio h_{t+1}.

- An **arbitrage** possibility is a self-financing portfolio h with the properties

$$V_0^h = 0,$$
$$P(V_T^h \geq 0) = 1,$$
$$P(V_T^h > 0) > 0.$$

- The **martingale** probabilities q_u and q_d are defined as the probabilities for which the following relation holds

$$s = \frac{1}{1+R} E^Q[S_{t+1} | S_t = s]$$

2.2.2 Binomial Model Pricing Algorithm

Proposition 11. The model is free of arbitrage $\Rightarrow d \leq (1+R) \leq u$.

Assumption 12. From now, assume that $d < u$ and $d \leq (1+R) \leq u$.

Definition 13. A **contingent claim** is a stochastic variable X of the form

$$X = \Phi(S_T),$$

where the **contract function** Φ is some given real valued function.

Definition 14. A given contingent claim X can be **replicated** / is **reachable** if

$$\exists \text{ self-financing } h \text{ s.t. } V_T^h = X, \text{ with probability } 1.$$

We call such an h a **hedging**/**replicating** portfolio. If all claims can be replicated, then the market is **(dynamically) complete**.

(Binomial Algorithm) Proposition 15. Consider a T-claim $X = \Phi(S_T)$, which could be replicated with a self-financing portfolio h. Let k be the number of up-moves occurred. So $V_t(k)$ can be computed recursively as

$$\begin{cases} V_t(k) &= \dfrac{1}{1+R}\left[q_u V_{t+1}(k+1) + q_d V_{t+1}(k)\right], \\ V_T(k) &= \Phi(s u^k d^{T-k}). \end{cases}$$

where the martingale probabilities q_u and q_d are

$$\begin{cases} q_u &= \dfrac{(1+R)-d}{u-d} \\ q_d &= \dfrac{u-(1+R)}{u-d}. \end{cases}$$

and the hedging portfolio is given by

$$\begin{cases} x_t(k) &= \dfrac{1}{1+R} \cdot \dfrac{uV_t(k) - dV_t(k+1)}{u-d}, \\ y_t(k) &= \dfrac{1}{S_{t-1}} \cdot \dfrac{V_t(k+1) - V_t(k)}{u-d}. \end{cases}$$

N.B. In Fig 2.9 in notes, (the figure might be confusing) $(-22.5, 5/8)$ is both h_0 and h_1, $(-42.5, 95/120)$ is $h_2(1)$, $(-2.5, 1/8)$ is $h_2(0)$, etc.

Proposition 16. The arbitrage free price at $t = 0$ of a T-claim X is given by

$$\Pi_0[X] = \frac{1}{(1+R)^T} \cdot E^Q[X],$$

where Q denotes the martingale measure, or more explicitly,

$$\Pi_0[X] = \frac{1}{(1+R)^T} \cdot \sum_{k=0}^{T} \binom{T}{k} q_u^k q_d^{T-k} \Phi(s u^k d^{T-k}).$$

Proposition 17. $d < (1+R) < u \iff$ free of arbitrage.

3 Brownian Motion

3.1 Definitions

- Let A, B be two random variables, we define **equal in distribution** as
$$A \stackrel{d}{=} B \iff P(A \in C) = P(B \in C)\ \forall \text{ possible sets } C.$$

- We say a random variable X has **stationary increments** if
$$X_t - X_s \stackrel{d}{=} X_{t+h} - X_{s+h} \quad \forall h > 0.$$

- A stochastic process B_t is called a **Brownian motion** if
  - $B_0 = 0$,
  - $B_t - B_s \sim N(0, t-s)\ \forall t \geq s \geq 0$, implying *stationary* increments,
  - it has *independent* increments.

  with
  - $\mathbb{E}[B_t] = 0$,
  - $\mathrm{Cov}(B_t, B_s) = \min\{t, s\}$. Show this using (Practice!)
    * $\mathrm{Cov}(B_t, B_s) = \mathbb{E}[(B_t - \mathbb{E}[B_t])(B_s - \mathbb{E}[B_s])]$
    * $\mathrm{Cov}(B_t, B_s) = \mathbb{E}[B_t B_s] - \mathbb{E}[B_t]\mathbb{E}[B_s]$.

3.2 Path Properties

- Paths are continuous (no jumps).
- Nowhere differentiable.
- Paths are **self-similar**, i.e.
$$(T^H B_{t_1}, \ldots, T^H B_{t_n}) \stackrel{d}{=} (B_{Tt_1}, \ldots, B_{Tt_n})$$
  with $H = 0.5$ is the Hurst coefficient.

3.3 Brownian Motion with Drift

- A **Gaussian Process** is a stochastic process s.t. every finite collection of those random variables has a multivariate normal distribution.

- The process is
$$X_t = \mu t + \sigma B_t, \quad t \geq 0.$$

- It is a Gaussian process with the following properties:
  - $\mathbb{E}[X_t] = \mu t$,
  - $\mathrm{Cov}(X_t, X_s) = \sigma^2 \min\{t, s\}$. (Practice!)

3.4 Geometric Brownian Motion with Drift

- The process is
$$X_t = e^{\mu t + \sigma B_t}, \quad t \geq 0.$$

- It is *not* Gaussian, but log-normal instead, with the following properties:
  - $\mathbb{E}[X_t] = e^{\mu t + \frac{1}{2}\sigma^2 t}$, (Practice!)
  - $\mathrm{Cov}(X_t, X_s) = e^{\left(\mu + \frac{1}{2}\sigma^2\right)(t+s)} \left(e^{\sigma^2 s} - 1\right)$. (Practice!)

4 Stochastic Integration

4.1 Time Value of Money

- £x today is worth more than £x in the future.
- Compensation for postponed consumption
- Inflation: prices may rise so £x in the future will not have the same purchasing power
- Risk – money in the future may never be received.

4.2 Interest

4.2.1 Definitions

- $V(t)$: value of investment at time t
- r: interest rate $r \geq 0$
- t: time measured in years
- P: principal/initial investment (£)
- simple/annual compounding:

$$\begin{cases} V(0) &= P \\ V(t) &= (1+tr)P \end{cases}$$

4.2.2 Compounding

- Periodic, e.g. monthly compounding:

$$V(t) = \left(1 + \frac{r}{12}\right)^t P$$

- Continuous:

$$\begin{aligned} V(t) = e^{tr}P \quad &\implies \quad \frac{\mathrm{d}V(t)}{\mathrm{d}t} = rV(t) \\ &\implies \quad V(t) = V(0) + \int_0^t rV(s)\mathrm{d}s. \end{aligned}$$

4.2.3 Stochastic Integral

Example:

$$I(T) = \int_0^T B(t)\mathrm{d}B(t)$$

Using left-hand point approximation, we can get

$$\begin{aligned} I(T) &= \lim_{L\to\infty} \sum_{i=0}^{L-1} B(t_i)\left[B(t_{i+1}) - B(t_i)\right] \\ &= \lim_{L\to\infty} \sum_{i=0}^{L-1} -\frac{1}{2}B(t_i)^2 - \frac{1}{2}[B(t_{i+1}) - B(t_i)]^2 + \frac{1}{2}B(t_{i+1})^2 \\ &= \frac{1}{2}\lim_{L\to\infty} \sum_{i=0}^{L-1} [B(t_{i+1})^2 - B(t_i)^2] - \delta B_i^2 \end{aligned}$$

where $\delta B_i = B(t_{i+1}) - B(t_i)$. Analyzing the above expression (exercise!), we can obtain

$$I(T) = \frac{1}{2}B(T)^2 - \frac{1}{2}T.$$

We can also show the following: (exercise!)

- $\mathbb{E}[\delta B_i] = 0$
- $\mathrm{Var}(\delta B_i) = \delta t$
- $\mathbb{E}[\delta B_i^2] = \delta t$
- $\mathrm{Var}(\delta B_i^2) = 2(\delta t)^2$ (hint: $\mathbb{E}[Z^4] = 3$ for $Z \sim N(0,1)$)

5 Stochastic Calculus

5.1 Ito's Formula

5.1.1 Ito's Multiplication Rules

$$\begin{aligned} (\mathrm{d}t)^2 &= 0, \\ (\mathrm{d}t)(\mathrm{d}B(t)) &= 0, \\ (\mathrm{d}B(t))^2 &= \mathrm{d}t. \end{aligned}$$

5.1.2 Ito's Lemma

Consider an **Ito's process**

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}B_t,$$

where

- μ_t (or just μ): drift process,
- σ_t (or just σ): diffusion process,

both can be either stochastic or deterministic. Using the multiplication rules, we can derive that

$$(\mathrm{d}X_t)^2 = \mu^2(\mathrm{d}t)^2 + \sigma^2(\mathrm{d}B_t)^2 + 2\mu\sigma(\mathrm{d}t)(\mathrm{d}B_t) = \sigma^2\mathrm{d}t.$$

Let $Z(t) = f(t, X_t)$, then

$$\begin{aligned}\mathrm{d}Z(t) &= \frac{\partial f}{\partial X_t}\mathrm{d}X_t + \frac{\partial f}{\partial t}\mathrm{d}t + \underbrace{\frac{1}{2}\frac{\partial^2 f}{\partial X_t^2}(\mathrm{d}X_t)^2}_{\text{Itô correction}} \\ &= \left(\frac{\partial f}{\partial t} + \mu\frac{\partial f}{\partial X_t} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial X_t^2}\right)\mathrm{d}t + \frac{\partial f}{\partial X_t}\sigma\mathrm{d}B(t).\end{aligned}$$

5.1.3 Examples

1. $\mathrm{d}X_t = \mathrm{d}W_t$, $f(t,x) = te^{\alpha x}$. Since

$$\frac{\partial f}{\partial t} = e^{\alpha x}, \quad \frac{\partial f}{\partial x} = \alpha te^{\alpha x}, \quad \frac{\partial^2 f}{\partial x^2} = \alpha^2 te^{\alpha x},$$

and in this case $\mu = 0$, $\sigma = 1$, we have

$$\mathrm{d}f(t, W_t) = \left(e^{\alpha x} + +\frac{1}{2}\alpha^2 te^{\alpha W_t}\right)\mathrm{d}t + \alpha te^{\alpha W_t}\mathrm{d}W_t$$

2. $\mathrm{d}X = \mathrm{d}W$, $f(x) = x^2$.

$$\begin{aligned}\mathrm{d}f(t, W_t) &= \frac{\partial f}{\partial t}\mathrm{d}t + \frac{\partial f}{\partial x}\mathrm{d}x + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\mathrm{d}t \\ &= 0 + 2W\mathrm{d}W + \frac{1}{2}2\mathrm{d}t \\ &= \mathrm{d}t + 2W_t\mathrm{d}W_t\end{aligned}$$

3. Compute $\mathbb{E}\left[B_t^4\right]$. This could be transformed to the question

$$\mathrm{d}X_t = \mathrm{d}B_t, \quad Z_t = X_t^4.$$

Thus,

$$\mathrm{d}Z_t = 6B_t^2\mathrm{d}t + 4B_t^3\mathrm{d}B_t,$$

i.e.

$$Z_t = Z_0 + 6\int_0^t B_s^2\mathrm{d}s + 4\int_0^t B_s^3\mathrm{d}B_s$$

so

$$\begin{aligned}\mathbb{E}[B_t^4] = \mathbb{E}Z_t &= 0 + 6\int_0^t \mathbb{E}[B_s^2]\mathrm{d}s + 4\int_0^t \mathbb{E}[B_s^3]\mathrm{d}B_s \\ &= 6\int_0^t s\mathrm{d}s + 0 \\ &= 3t^2\end{aligned}$$

5.2 Stochastic Differential Equation

Example Given that

$$\mathrm{d}S(t) = \mathrm{d}B(t), \quad u(t,x) = y_0 e^{\mu t + \sigma x},$$

and

$$Y(t) = u(t, S), \quad Y(0) = y_0,$$

we have

$$\begin{aligned}\mathrm{d}Y_t &= \mu u(t,S)\mathrm{d}t + \sigma u(t,s)\mathrm{d}B + \frac{1}{2}\sigma^2 u(t,s)\mathrm{d}t \\ &= \left(\mu + \frac{1}{2}\sigma^2\right)Y(t)\mathrm{d}t + \sigma Y(t)\mathrm{d}B_t.\end{aligned}$$

This gives hint on how to solve

$$\mathrm{d}Y(t) = \mu Y(t)\mathrm{d}t + \frac{1}{2}\sigma Y(t)\mathrm{d}B_t,$$

whose solution is

$$Y(t) = u(t, B_t) = y(0)e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma B_t}.$$

This can be solved with a change of variable

$$Z(t) = \ln Y(t).$$

Applying Ito's lemma, we can derive that

$$\begin{aligned}\mathrm{d}Z(t) &= \frac{1}{Y(t)}\mathrm{d}Y(t) - \frac{1}{2}\frac{1}{Y(t)^2}\sigma^2 Y(t)^2\mathrm{d}t \\ &= \mu\mathrm{d}t + \sigma\mathrm{d}B(t) - \frac{1}{2}\sigma^2\mathrm{d}t\end{aligned}$$

$$= \left(\mu - \frac{1}{2}\sigma^2\right) \mathrm{d}t + \sigma \mathrm{d}B_t$$

Integrating both sides from 0 to t, we get

$$Z(t) - Z(0) = \ln \frac{Y(t)}{Y(0)} = \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma B_t$$

leading to the solution.

5.3 SDE & PDE

Consider the SDE

$$\begin{aligned} \mathrm{d}X_s &= \mu(s, X_s)\mathrm{d}s + \sigma(s, B_s)\mathrm{d}B_s, \\ X_t &= x, \end{aligned}$$

which starts at x at time t and evolves in the interval $[t, T]$. Now the parabolic PDE

$$\begin{aligned} \frac{\partial F}{\partial t} + \mu(t,x)\frac{\partial F}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 F}{\partial x^2} = \frac{\partial F}{\partial t} + \mathcal{A}F = 0 \\ F(T, x) = \phi(x), \end{aligned}$$

where

$$\mathcal{A} = \mu(t,x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2}{\partial x^2}$$

is called the *infinitesimal operator*. Applying Ito's Lemma to $F(s, X(s))$, we have

$$\mathrm{d}F = \frac{\partial F}{\partial t}\mathrm{d}t + \frac{\partial F}{\partial X}\mathrm{d}X + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial X^2}\mathrm{d}t.$$

Integrate from t to T, we get

$$\begin{aligned} &F(T, X_T) - F(t, X_t) \\ &= \int_t^T \underbrace{\left[\frac{\partial F(s,X_s)}{\partial t} + \frac{\partial F(s,X_s)}{\partial X}\mu(s,X_s) + \frac{1}{2}\sigma^2\frac{\partial^2 F(s,X_s)}{\partial X^2}\right]}_{\text{exactly the PDE, } = 0} \mathrm{d}s \\ &\quad + \int_t^T \sigma(s,X_s)\frac{\partial F(s,X_s)}{\partial X}\mathrm{d}W_s \\ &= \int_t^T \left[\frac{\partial F(s,X_s)}{\partial t} + \mathcal{A}F(s,X_s)\right]\mathrm{d}s + \int_t^T \sigma(s,X_s)\frac{\partial F(s,X_s)}{\partial X}\mathrm{d}W_s \\ &= \int_t^T \sigma(s,X_s)\frac{\partial F(s,X_s)}{\partial X}\mathrm{d}W_s, \\ &\implies \phi(X_T) - F(t,X) = \int_t^T \sigma(s,X_s)\frac{\partial F(s,X_s)}{\partial X}\mathrm{d}W_s \end{aligned}$$

Taking expectation on both sides (conditional on t),

$$\mathbb{E}\phi(X_T) - \mathbb{E}[F(t,X(t))|X(t) = x] = \mathbb{E}\left[\int_t^T \sigma(s,X_s)\frac{\partial F(s,X_s)}{\partial X}\mathrm{d}W_s | X(t) = x\right].$$

Since stochastic integral has 0 expectation, i.e.

$$F(t,X) = \mathbb{E}\left[\phi(X_T)|X(t) = x\right]$$

The above conclusion is the **Feynman Kac Theorem**

Proposition 18. If F is a solution to

$$\frac{\partial F}{\partial t} + \mathcal{A}F = 0, \quad F(T,X) = \phi(X),$$

where $\mathcal{A}$ is the infinitesimal operator associated with the SDE

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

then

$$F(t,x) = \mathbb{E}\left[\phi(X_T)|X_t = x\right]$$

Note By "associated" means that the quantity μ and σ are the same in both SDE and PDE.

Example 19. Solve the PDE

$$\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 F(t,x)}{\partial x^2} = 0, \quad F(T,x) = x^2$$

using the Feynman Kac formula. Let

$$\mathrm{d}X_s = \sigma \mathrm{d}W_s, \quad X_t = x$$

The solution to the SDE is

$$X_t = x + \sigma\left[W_T - W_t\right]$$

implying that $X_T \sim N(x, \sigma^2(T-t))$. So

$$F(t,x) = \mathbb{E}X_T^2 = \sigma^2(T-t) + x^2.$$

6 Arbitrage Pricing

6.1 Definitions

- **out-of-the-money (OTM)** for a call means $S(t) - K < 0$.
- **in-the-money** for a call means $S(t) - K > 0$.
- **at-the-money** for a call means $S(t) = K$.
- **Exotic options**: more complicated products that have "exotic" features, e.g. early exercise, multiple strikes, etc.

6.2 The Black & Scholes Model and Arbitrage

6.2.1 Model Introduction

The model is based on:

1. risk-free bank (letter B stands for bank):

$$\mathrm{d}B(t) = rB(t)\mathrm{d}t \quad \Rightarrow \quad B(t) = B(0)e^{rt}$$

2. stock:

$$\mathrm{d}S(t) = \mu S(t)\mathrm{d}t + \sigma S(t)\mathrm{d}W(t), \quad S(0) = s(0)$$
$$\Rightarrow \quad S(t) = S(0)e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t}$$

3. option, that depends on stock price and the current time:

$$V(t, S).$$

6.2.2 The PDE approach of pricing

The model:

$$\mathrm{d}B_t = rB_t\mathrm{d}t, \quad \mathrm{d}S_t = \mu S_t\mathrm{d}t + \sigma S_t\mathrm{d}W_t$$

Applying Ito's lemma on $V(t, S)$, we have

$$\begin{aligned}\mathrm{d}V &= \frac{\partial V}{\partial t}\mathrm{d}t + \frac{\partial V}{\partial S}\mathrm{d}S + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(\mathrm{d}S)^2 \\ &= \left[\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right]\mathrm{d}t + \sigma S\frac{\partial V}{\partial S}\mathrm{d}W_t.\end{aligned}$$

Consider a portfolio $P(t)$ with Δ stocks and one short option (**Delta hedging**). The value of the portfolio is

$$P(t) = \underbrace{\Delta S(t)}_{\text{long the stock}} - \underbrace{V(t)}_{\text{short the option}}$$

The portfolio evolves according to

$$\begin{aligned}\mathrm{d}P(t) &= \Delta\mathrm{d}S(t) - \mathrm{d}V \\ &= \left[\Delta\mu S - \frac{\partial V}{\partial t} - \mu S\frac{\partial V}{\partial S} - \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right]\mathrm{d}t + \left[\Delta\sigma S - \sigma S\frac{\partial V}{\partial S}\right]\mathrm{d}W_t.\end{aligned}$$

If we let $\Delta = \frac{\partial V}{\partial S}$, because if the market has no arbitrage then its return must equal $rP\mathrm{d}t$, we are left with

$$\mathrm{d}P = \left[\Delta\mu S - \frac{\partial V}{\partial t} - \mu S\frac{\partial V}{\partial S} - \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right]\mathrm{d}t = rP\mathrm{d}t$$

$$\Rightarrow \quad \Delta\mu S - \frac{\partial V}{\partial t} - \mu S\frac{\partial V}{\partial S} - \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} = rP$$

Since $P = \frac{\partial V}{\partial S}S - V, \Delta = \frac{\partial V}{\partial S}$, after these substitution, we obtain the following second-order parabolic PDE called the **Black & Scholes Equation (BSE)**:

$$\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} = rV, \quad V(T, S) = \max(S - K, 0),$$

where the second equation is the boundary condition for European call option.

Hedge Parameters The following parameters (appearing in the PDE) are the sensitivities of the option value w.r.t. small changes in the problem.

- $$\Delta = \frac{\partial V}{\partial S}$$

  It tells the trader how to balance the portfolio so that it is always equal to the option.

- $$\Gamma = \frac{\partial^2 V}{\partial S^2}$$
  It gives an indication of how stable the hedging portfolio is. If Γ is large, the trader needs to rebalance more often.

- $$\theta = \frac{\partial V}{\partial t}$$
  If S stays constant then the value of the option will change by θ

- $$\zeta = \frac{\partial V}{\partial \sigma}$$
  It measures the change in price w.r.t. volatility

- $$\rho = \frac{\partial V}{\partial r}$$
  It measures sensitivity w.r.t. interest rate.

6.2.3 The Martingale Method of Pricing

Option valuation could be reduced to the calculation of the following:

$$V(S_0) = \mathbb{E}^Q \left[e^{-rT} V(S_T) | S_0 \right].$$

In the case of the Black-Scholes model, Q is the probability distribution of the following SDE

$$\mathrm{d}S_t = rS_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t$$

or

$$S(t) = S(0)e^{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t}$$

where

- r: risk-free rate
- σ: volatility
- w: standard brownian motion

This says that the assumption of the Black Scholes model is that the underlying asset follows a geometric Brownian motion i.e. lognormal random walk.

7 Monte-Carlo methods

7.1 Strong Law of Large Numbers

Let $\xi^{(i)}$, $i = 1, 2, \dots, N$ be i.i.d. random variables with values in $\mathbb{R}^d$, and mean $\mathbb{E}|\xi| < \infty$. Let $\hat{S}_N$ denote the empirical mean $\hat{S}_N = \frac{1}{N}\sum_{i=1}^N \xi^{(i)}$. Then the SLLN holds true:

$$\lim_{N \to \infty} \hat{S}_N = \mathbb{E}(\xi)$$

7.2 Central Limit Theorem

Let $\xi^{(i)}$, $i = 1, 2, \dots, N$ be i.i.d. random variables with mean $\mathbb{E}\xi$ and $\mathrm{Var}(\xi) < \infty$. Then

$$Y_N = \frac{\sqrt{N}}{\sigma} \left(\frac{1}{N} \sum_{i=1}^N \xi^i - \mathbb{E}\xi \right) \longrightarrow_d Z \sim N(0,1)$$

as $N \to \infty$. In practice, we use empirical variance to compute σ (unbiased estimate):

$$\sigma_N^2 = \frac{1}{N-1} \sum_{i=1}^N \left(\xi^i - \bar{\xi} \right)^2.$$

In other words,

$$\mathbb{P}\left(\left| \frac{\sqrt{N}}{\sigma} \left(\frac{1}{N} \sum_{i=1}^N \xi^i - \mathbb{E}\xi \right) \right| < R \right) \longrightarrow \int_{-R}^{R} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathrm{d}x$$

7.3 Monte Carlo for Option Valuation

Since

$$S(T) = S(0)e^{\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z}, \quad Z \sim N(0,1),$$

if we can generate i.i.d. normal random variables $\left\{Z^i\right\}$, $i = 1, 2, \dots, N$, we have

$$V(S_0) \approx \frac{1}{N} \sum_{i=1}^N V(S_T^i) e^{-rT} = V_N(S_0),$$

where

$$S_T^i = e^{\left(r - \frac{1}{2}\sigma^2\right)T + \sqrt{T}Z^i}$$

7.4 Variance Reduction methods for MC

7.4.1 Antithetic Variables

If X_1 and X_2 are i.i.d. RVs,

$$\mathrm{Var}\left(\frac{X_1+X_2}{2}\right) = \frac{1}{4}\left(\mathrm{Var}(X_1)+\mathrm{Var}(X_2)+2\mathrm{Cov}(X_1,X_2)\right)$$

so variance is reduced if $\mathrm{Cov}(X_1, X_2) < 0$.

Let $f(x)$ be monotonic. For instance, in the case of <u>Gaussian variables</u> X, $f(x)$ and $f(-x)$ are negatively correlated, i.e.

$$\mathrm{Cov}(f(X), f(-X)) < 0.$$

In the case of <u>uniform distribution variables</u> U,

$$\mathrm{Cov}(f(U), f(1-U)) < 0.$$

7.4.2 Control Variates

Suppose we want to estimate $\mathbb{E}X$ and know the mean of another random variable Y, $\mathbb{E}Y$, then

$$\mathbb{E}X = \mathbb{E}[X - Y + \mathbb{E}Y].$$

Y is called a control variate. While the expectation of

$$Z = X - Y + \mathbb{E}Y$$

is $\mathbb{E}X$ its variance will be

$$\mathrm{Var}(X-Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X,Y).$$

This idea works when X and Y are close.

7.4.3 Importance Sampling

We can write the expectation with respect to the probability density p as

$$\mathbb{E}^p g(X) = \int g(x)p(x)\mathrm{d}x.$$

Suppose that we had another density q that had the same support i.e. $p(x) > 0 \iff q(x) > 0$, then

$$\mathbb{E}^p g(x) = \int g(x)p(x)\mathrm{d}x = \int g(x)\frac{p(x)}{q(x)}q(x)\mathrm{d}x = \mathbb{E}^q g(x)\frac{p(x)}{q(x)}$$

Thus instead of calculating $\mathbb{E}^p g(x)$ using i.i.d. samples from p we calculate

$$\mathbb{E}^q g(x)\Lambda(x)$$

using samples from q, where $\Lambda(x)$ is the likelihood ratio.

7.5 Computation of Monte-Carlo Greeks

Monitoring the P&L of the hedging portfolio

$$\pi(t,s) = V(t,s) - \Delta S(t).$$

If $\pi(t,S)$ is not small, then it means that some element of the model is not performing well i.e. the value of the option is changing well beyond the Δ parameter predicts. There will be several parameters deployed: $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$, so

$$V(t,s,\boldsymbol{\theta})$$

With Monte Carlo, we can use *finite difference* to calculate the Greeks.

7.5.1 Finite Difference

For convenience, we consider a single parameter θ. We know that by a Taylor series expansion,

$$V(\theta+\Delta\theta) = V(\theta) + \frac{\partial V}{\partial \theta}\Delta\theta + \frac{1}{2}\frac{\partial^2 V}{\partial \theta^2}\Delta\theta^2 + \cdots$$

therefore

$$\begin{aligned}\frac{\partial V}{\partial \theta} &= \frac{V(\theta+\Delta\theta)-V(\theta)}{\Delta\theta} - \frac{1}{2}\frac{\partial^2 V}{\partial\theta^2}\Delta\theta + O(\Delta\theta^2)\\ &\approx \frac{\overline{V}_N(\theta+\Delta\theta)-\overline{V}_N(\theta)}{\Delta\theta},\end{aligned}$$

where

$$\overline{V}_N(\theta) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{M(T)}V(T,S^i,\theta),$$

and $M(T)$ is the discount factor, e.g. $M(T) = e^{r(T-t_0)}$. Since $\mathbb{E}\overline{V}_N(\theta) = V(\theta)$, the error will be $O(\Delta\theta)$.

The above is <u>**forward finite difference**</u>. We could also consider <u>**central finite difference**</u>:

$$\frac{\partial V}{\partial\theta} = \frac{V(\theta+\Delta\theta)-V(\theta-\Delta\theta)}{2\Delta\theta} + O(\Delta\theta^2),$$

where the second derivative error in the Taylor expansion of $V(\theta+\Delta\theta)$ and $V(\theta-\Delta\theta)$ cancels out. Thus the error is much less than the forward finite difference.

Central F.D. needs one more simulation $V(\theta-\Delta\theta)$ than forward F.D.

8 Numerical Methods for SDEs

8.1 Deterministic methods

We are looking at

$$\dot{x} = f(t,x), \quad t \geq t_0,\ x(t_0) = x_0,$$

and assume that f satisfies the **Lipscitz condition**:

$$|f(t,x) - f(t,y)| \leq L|x-y|, \quad \forall x, y \in \mathbb{R}, t > t_0,$$

where L is the **Lipscitz constant**.

8.1.1 Euler's Method

Given that we know $t_0, f, x(t_0)$, the most elementary approximation is

$$x(t) = x(t_0) + \int_{t_0}^{t} f(s, x(s))\mathrm{d}s \approx x(t_0) + f(t_0, x(t_0))(t-t_0).$$

Breaking the time interval $[t_0, \ldots, T]$ into $t_0 < t_1 < t_2 < \cdots < t_{N-1} = T$ s.t.

$$\delta = t_{i+1} - t_i, \quad i = 0, 1, \ldots, N-1,$$

we arrive at the recursive scheme

$$x_{n+1} = x_n + f(t_n, x_n)\delta.$$

The following theorem provides the condition where the above recursive scheme converges.

Theorem 20. Euler's method is convergent if

$$\delta < \frac{1}{L}.$$

Regarding **order of convergence**, we re-write the Euler's method as

$$\begin{aligned}
&x(t_{n+1}) - [x(t_n) + \delta f(t_n, x(t_n))] \\
=&[x(t_n) + \dot{x}(t_n)\delta + O(\delta^2)] - [x(t_n) + \delta f(t_n, x(t_n))] \\
=&O(\delta^2),
\end{aligned}$$

and we say that the Euler method converges with order 1, i.e. it recovers exactly every polynomial solution of degree 1 or less.

8.1.2 The Trapezoidal Rule

Instead of approximating with the left end-point we could just average the two:

$$\begin{aligned}
x(t) &= x(t_n) + \int_{t_n}^{t} f(s, x(s))\mathrm{d}s \\
&\approx x(t_n) + \frac{1}{2}(t-t_n)\left[f(t_n, x(t_n)) + f(t, x(t))\right],
\end{aligned}$$

giving the trapezoidal rule

$$x_{n+1} = x_n + \frac{1}{2}\delta\left[f(t_n, x_n) + f(t_{n+1}, x_{n+1})\right]$$

similarly, using Taylor's expansion and several substituions, we could get

$$x(t_{n+1}) - \left[x(t_n) + \frac{1}{2}\delta f(t_n, x_n(t_n)) + f(t_{n+1}, x_{n+1})\right] = O(\delta^3),$$

suggesting that the trapezoidal method has a convergence of order 2. We can also show that it is convergent. Comparing with the Euler method, it has a superior convergence rate but it is implicit, whereas Euler method is explicit.

8.1.3 The Theta method

Both the Euler and trapezoidal rule fit in the general scheme

$$x_{n+1} = x_n + \delta\left[\theta f(t_n, x_n) + (1-\theta) f(t_{n+1}, x_{n+1})\right],$$

where $\theta = 1$ is the Euler's scheme, $\theta = \frac{1}{2}$ is the trapezoidal scheme.

8.2 The Euler-Maruyama (EM) Method

Given that
$$\mathrm{d}X(t) = f(X(t))\mathrm{d}t + g(X(t))\mathrm{d}W(t),$$
with $X(0)$ given, $0 \leq t \leq T$, we define a step-size $\delta = \frac{T}{N}$. The exact solution is
$$X(t_{n+1}) = X(t_n) + \int_{t_n}^{t_{n+1}} f(X(s))\mathrm{d}s + \int_{t_n}^{t_{n+1}} g(X(s))\mathrm{d}W(s).$$
Assuming f and g are constant in $[t_n, t_{n+1}]$, then
$$X_{n+1} = X_n + f(X_n)\Delta t + g(X_n)\Delta W_n$$
where $\Delta W_n = W(t_{n+1}) - W(t_n)$ is a Brownian motion increment:
$$\Delta W_n \sim N(0, t_{n+1} - t_n) \sim \sqrt{\delta}N(0,1).$$
The EM algorithm is then like the following:

1. Fix δ, the stepsize.
2. compute iteratively
$$X_{n+1} = X_n + f(X_n)\delta + \sqrt{\delta}g(X_n)\xi_n,$$
where $\xi_n \sim N(0,1)$.

8.2.1 Weak v.s. Strong Convergence

Weak error associated with a discretization scheme is
$$e_\delta^{\text{weak}} = \max_{n=0,\dots,N-1} |\underbrace{\mathbb{E}[\phi(X_n)]}_{\text{EM approx.}} - \underbrace{\mathbb{E}[\phi(X_{t_n})]}_{\text{exact}}|$$
where ϕ is a class of functinos, e.g. polynomials/Lipscitz. We say that a method **converges weakly** if
$$e_\delta^{\text{weak}} \longrightarrow 0 \quad \text{as} \quad \delta \longrightarrow 0.$$
We say that the method **converges with weak order** p if
$$e_\delta^{\text{weak}} \leq K\delta^p \quad \forall 0 < \delta \leq \delta^* \text{ for some } \delta^*.$$

Strong error is given by
$$e_\delta^{\text{strong}} = \max_{n=0,\dots,N-1} \mathbb{E}[|\underbrace{X_n}_{\text{approx.}} - \underbrace{X(t_n)}_{\text{exact}}|]$$
A method **converges strongly** if
$$e_\delta^{\text{strong}} \longrightarrow 0 \quad \text{as} \quad \delta \longrightarrow 0$$
We say that the method **converges with strong order** p if
$$e_\delta^{\text{strong}} \leq K\Delta t^p \quad \forall 0 < \Delta t \leq \Delta t^*.$$

8.3 Implicit Methods and Numerical Stability

In stability analysis, we are interested in long-time behaviourof the numerical scheme e.g. if we make small errors then will this error propagate?

8.3.1 Stochastic θ-method

$$X_{n+1} = X_n + (1-\theta)f(X_n)\delta + \theta f(X_{n+1})\Delta t + g(X_n)\Delta W_n,$$
with

- $\theta = 0 \Rightarrow$ EM,
- $\theta = \frac{1}{2} \Rightarrow$ stochastic trapezoidal,
- $\theta = 1 \Rightarrow$ backward (implicit) Euler.

8.3.2 The Log-Normal Distribution

We say that X has a log-normal distribution, or $X \sim \mathrm{LN}(\mu, \sigma^2)$, if
$$\log(X) \sim N(\mu, \sigma^2).$$
The mean and variance of the log-normal distribution satisfy
$$\mathbb{E}[X] = \exp\left(\mu + \frac{\sigma^2}{2}\right),$$
$$\mathrm{Var}(X) = \exp\left(2\mu + \sigma^2\right)\left(\exp(\sigma^2) - 1\right).$$

8.3.3 Stability Analysis

Given that

$$\mathrm{d}X(t) = \mu X(t)\mathrm{d}t + \sigma X(t)\mathrm{d}W(t), \quad X(0) = x_0,$$

we know that

$$X(t) = X(0)e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t},$$

we can derive using log-normal distribution that

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

Definition 21. **Mean-square stability** is defined as

$$\lim_{t\to\infty} \mathbb{E}\left[X(t)^2\right] = 0 \iff 2\mu + \sigma^2 < 0.$$

Definition 22. **Asymptotic stability** is defined as

$$\lim_{t\to\infty} |X(t)| = 0 \text{ with probability } 1 \iff (\mu - \frac{1}{2}\sigma^2)t < 0.$$

Since $(\mu + \frac{1}{2}\sigma^2)t < 0 \Rightarrow (\mu - \frac{1}{2}\sigma^2)t < 0$, we can deduce that

$$\text{mean-square stability} \implies \text{asymptotic stability}.$$

8.3.4 Mean-square Stability of the θ-method

Definition 23. Let

$$X_{n+1} = X_n + \left[(1-\theta)\mu X_n + \theta\mu X_{n+1}\right]\delta + \sqrt{\delta}\sigma X_n \xi_n,$$

where $\xi_n \sim N(0,1)$. We have

$$\lim_{n\to 0} \mathbb{E}X_n^2 = 0 \iff (1-2\theta)\mu\delta < -2\left(\mu + \frac{\sigma^2}{2}\right)$$

Comments We can show that

- $0 \le \theta < \frac{1}{2}$ for a stable SDE the method has a finite region of stability,
- $\theta = \frac{1}{2}$ SDE stable $\iff$ mean square stable $\forall\delta$,
- $\frac{1}{2} < \theta \le 1$ the method is overstable – it is stable for a stable SDE and stable even for an unstable SDE.

9 Asset and Portfolio

9.1 Asset Returns

- **Asset**: investment instrument that can be bought/sold.
- If you buy an asset today at price X_0 and sell it in one year at price X_1, then the **total return** R on the investment is defined as

$$R = \frac{X_1}{X_0}$$

- **Rate of return** is defined as

$$r = \frac{X_1 - X_0}{X_0} = R - 1$$

- The rate of return acts much like an interest rate:

$$X_1 = (1+r)X_0.$$

9.2 Portfolio Returns

- We form a **master asset** or **portfolio** by apportioning an amount X_0 among n assets, where each amount X_{0i} is such that

$$\sum_{i=1}^{n} X_{0i} = X_0.$$

- Alternatively, we could use w_i as the **weight** of asset i in the portfolio

$$X_{0i} = w_i X_0$$

  such that $\sum_{i=1}^{n} w_i = 1$.
- Using the above concepts, the total return and rate of return are

$$R = \sum_{i=1}^{n} w_i R_i \quad \text{and} \quad r = \sum_{i=1}^{n} w_i r_i$$

9.3 Variance as a Risk Measure

- Suppose there are n assets with random rates of return r_i, and expected values $\mathbb{E}(r_i) = \overline{r}_i$, where $i \in \{1, 2, \ldots, n\}$.
- The variance of r_i is denoted as σ_i^2 and the covariance of r_i and r_j is σ_{ij} ($\sigma_{ii} = \sigma_i^2$).
- The expected return of a portfolio is given by

$$\overline{r} = \mathbb{E}(r) = \mathbb{E}\left(\sum_{i=1}^{n} w_i r_i\right) = \sum_{i=1}^{n} w_i \mathbb{E}(r_i) = \sum_{i=1}^{n} w_i \overline{r}_i.$$

- The variance of the return of the portfolio is given by

$$\begin{aligned}
\sigma^2 &= \mathrm{Var}(r) \\
&= \mathbb{E}[(r - \overline{r})^2] \\
&= \mathbb{E}\left[\left(\sum_{i=1}^{n} w_i r_i - \sum_{i=1}^{n} w_i \overline{r}_i\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_{i=1}^{n} w_i (r_i - \overline{r}_i)\right)\left(\sum_{j=1}^{n} w_j (r_j - \overline{r}_j)\right)\right] \\
&= \mathbb{E}\left[\sum_{i,j=1}^{n} w_i w_j (r_i - \overline{r}_i)(r_j - \overline{r}_j)\right] \\
&= \sum_{i,j=1}^{n} w_i \sigma_{ij} w_j.
\end{aligned}$$

- **Example**: If the portfolio has $w_1 = \alpha$ and $w_2 = 1 - \alpha$,

$$\overline{r}_p = \alpha \overline{r}_1 + (1 - \alpha)\overline{r}_2,$$

$$\sigma_p^2 = \alpha^2 \sigma_1^2 + 2\alpha(1 - \alpha)\sigma_{12} + (1 - \alpha)^2 \sigma_2^2.$$

9.4 Mean-variance diagrams

- **Minimum variance set**: the left boundary of the feasible set.
- **Minimum variance point**: The point with lowest possible variance. It is obtained by minimizing the risk/variance for any given mean return.
- **Efficient frontier**: the upper half of the minimum variance set, of investors' interest.

9.5 The Markowitz Model

The **Markowitz Model** is defined as an optimization problem as

$$\begin{aligned}
\text{minimize} \quad & \frac{1}{2} \sum_{i,j=1}^{n} w_i \sigma_{ij} w_j \\
\text{subject to} \quad & \sum_{i=1}^{n} w_i \overline{r}_i = \overline{r}_p, \\
& \sum_{i=1}^{n} w_i = 1.
\end{aligned}$$

To solve this, we can form the **Lagrangian function** L given by

$$L = \frac{1}{2} \sum_{i,j=1}^{n} w_i \sigma_{ij} w_j - \lambda \left(\sum_{i=1}^{n} w_i \overline{r}_i - \overline{r}_p\right) - \mu \left(\sum_{i=1}^{n} w_i - 1\right),$$

and form the following system of equations:

$$\begin{cases}
\nabla_w L = \sum_{j=1}^{n} \sigma_{ij} w_j - \lambda \overline{r}_i - \mu = 0, \\
\nabla_\lambda L = \sum_{i=1}^{n} w_i \overline{r}_i - \overline{r}_p = 0 \\
\nabla_\mu L = \sum_{i=1}^{n} w_i - 1 = 0.
\end{cases}$$

Equivalently, using the vector notation, we have the problem as

$$\begin{aligned}
\text{minimize} \quad & \frac{1}{2} \mathbf{w}^T \Sigma \mathbf{w} \\
\text{subject to} \quad & \mathbf{w}^T \overline{\mathbf{r}} - \overline{r}_p = 0,
\end{aligned}$$

$$\mathbf{w}^T\mathbf{e} - 1 = 0.$$

The associated Lagrangian function is

$$L(\mathbf{w}, \lambda, \mu) = \frac{1}{2}\mathbf{w}^T\Sigma\mathbf{w} - \lambda(\mathbf{w}^T\overline{\mathbf{r}} - \overline{r}_p) - \mu(\mathbf{w}^T\mathbf{e} - 1),$$

and the optimality conditions become

$$\begin{cases} \nabla_w L = \Sigma\mathbf{w} - \lambda\overline{\mathbf{r}} - \mu\mathbf{e} = \mathbf{0}, \\ \nabla_\lambda L = \mathbf{w}^T\overline{\mathbf{r}} - \overline{r}_p = 0, \\ \nabla_\mu L = \mathbf{w}^T\mathbf{e} - 1 = 0, \end{cases}$$

which could be written in one vector equation

$$\begin{pmatrix} \Sigma & -\overline{\mathbf{r}} & -\mathbf{e} \\ -\overline{\mathbf{r}}^T & 0 & 0 \\ -\mathbf{e}^T & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{w} \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ -\overline{r}_p \\ -1 \end{pmatrix},$$

which then the solution can be derived as

$$\begin{pmatrix} \mathbf{w} \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \Sigma & -\overline{\mathbf{r}} & -\mathbf{e} \\ -\overline{\mathbf{r}}^T & 0 & 0 \\ -\mathbf{e}^T & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ -\overline{r}_p \\ -1 \end{pmatrix}$$

if Σ has *full rank* and $\overline{\mathbf{r}}$ is not a multiple of $\mathbf{e}$.

9.6 Estimation of mean and variance

We define the estimator of $\overline{r}$ as

$$\hat{\overline{r}} = \frac{1}{n}\sum_{i=1}^{n} r_i,$$

and we can see that

$$\mathbb{E}(\hat{\overline{r}}) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} r_i\right) = \overline{r},$$

$$\mathrm{var}(\hat{\overline{r}}) = \mathbb{E}\left[(\hat{\overline{r}} - \overline{r})^2\right] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}(r_i - \overline{r})\right]^2 = \frac{1}{n}\sigma^2.$$

so $\hat{\overline{r}}$ is an unbiased estimator of $\overline{r}$. With some calculation, we can see that

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(r_i - \hat{\overline{r}})^2$$

is an unbiased estimator of σ^2, whose variance is

$$\mathrm{var}(\hat{\sigma}^2) = \frac{2\sigma^4}{n-1}.$$

Similarly, the unbiased estimator of covariance is

$$\hat{\sigma}_{AB} = \frac{1}{n-1}\sum_{i=1}^{n}(r_{A,i} - \hat{\overline{r}}_A)(r_{B,i} - \hat{\overline{r}}_B).$$