Financial mathematics and risk theory

Péter Kevei

November 9, 2022

Contents

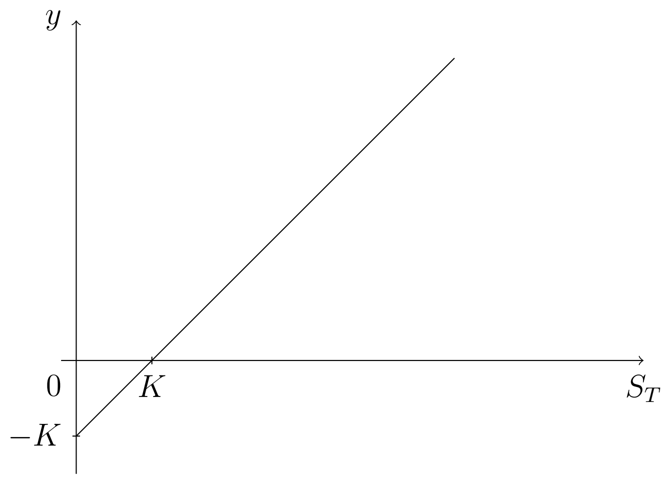

Figure 1: Payoff of a forward

1 Introduction

These notes are based on the Hungarian lecture notes by Gáll and Pap [5], on Shiryaev's monograph [7], and on Elliott and Kopp [2].

There are two type of financial instruments: the basic financial units and their derivatives.

Underlying:

- bond: risk-free asset, basically money. Its price is deterministic B_t;
- stock: risky asset. Its price is a random, modeled by a stochastic process S_t (or with d risky assets $= (S_t^1, \ldots, S_t^d)$).

Derivatives are bets on the underlying. They are used to share or reduce risk. Here we consider forward contracts and options.

1.1 Forward

A forward contract is an agreement to buy or sell an asset (stock) for a price previously agreed K in the future time T.

From the buyers point of view, at time T his wealth is $S_T - K$, that is the payoff function is $f(s) = s - K$.

We want to determine the fair price of this contract, and to understand the meaning of 'fair'. Assume $B_0 = 1$.

Seller's point of view: At time 0, we can buy a stock for S_0. Then at time T selling a stock for K and paying back the loan $S_0 \cdot B_T$, we have $K - S_0 B_T$. Therefore,

$$K \geq S_0 B_T.$$

Buyer's point of view: At time 0, we sell a stock for S_0. At time T we pay K for a stock, and the our wealth is $S_0 B_T - K$. Thus,

$$K \leq S_0 B_T.$$

We see that the fair price has to be $K = S_0 B_T$. Otherwise, either the seller or the buyer would have a strategy providing riskless profit (arbitrage).

Example 1. Let $S_0 = 40$, $B_t = e^{rt}$, $r = 0.1$ being the annual interest, $T = 1$ year. What is the fair price of this forward, and what is the value of the contract after half a year if $S_{0.5} = 45$?

The forward price at time 0 is

$$K = S_0 B_1 = 40 \cdot e^{0.1} = 44.2.$$

At time $t = 0.5$ the forward price

$$K_2 = S_{0.5} B_{0.5} = 45 \cdot e^{\frac{1}{2}0.1} = 47.3.$$

Thus the current value of the contract

$$e^{-\frac{1}{2}r}(47.3 - 44.2) = 2.9.$$

1.2 Options

An option is *right* to do something but not an obligation. European option can be executed only at the expiration date, while American options can be executed at any time.

The writer of a European call option agrees to sell a stock for a previously agreed price K. Clearly, the buyer of this option will not use his right if $S_T < K$. The payoff function for the buyer is $f(s) = (s - K)_+$

In case of a put option the writer agrees to buy a stock for K. The payoff function of the buyer

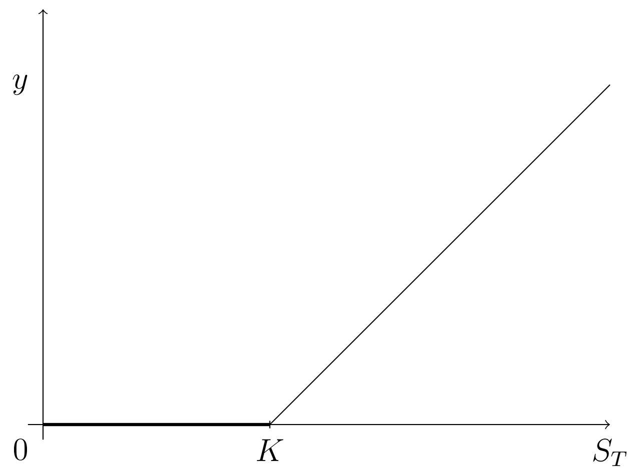

Figure 2: Payoff a call option

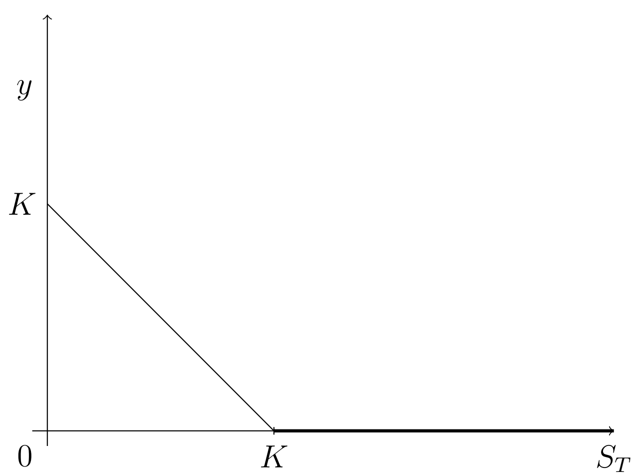

Figure 3: Payoff of a put option

1.3 Put–call parity

The aim of the course is to determine the fair price of an option, and understand the fairness. However, there is a simple relation between call and put prices regardless of the underlying market model.

Let C_K be the fair price of the call, and P_K be the fair price of the put, both with strike price K. Assume that $B_0 = 1$. Then, from the payoff functions it is easy to see that having put, a stock, and -1 call results at the expiration date (regardless of the stock price) a wealth K. That is, after discounting

$$\frac{K}{B_T} = P_K + S_0 - C_K.$$

This is the put-call parity.

2 Portfolio, claim, and hedging in discrete time

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. In the discrete time case we always assume (if not stated otherwise) that Ω is finite, and $\mathbf{P}(\{\omega\}) > 0$ for each $\omega \in \Omega$. We assume that transactions are made only at the time instants $0, 1, \ldots, N$. Let $(\mathcal{F}_n)_{n=0,1,\ldots,N}$ be a filtration, an increasing sequence of σ-algebras, such that $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_N = \mathcal{F}$. Assume that there are d risky assets and a bond. The price of the risky asset i at time is S_n^i, an $\mathcal{F}_n$-measurable random variable, and the bond price at time n is B_n.

2.1 Portfolio

An investment portfolio (strategy) is $\pi_n = (\beta_n, \gamma_n)$, where $\beta_n \in \mathbb{R}$ represents the amount of bonds in the portfolio at time n, while $\gamma_n = (\gamma_n^1, \ldots, \gamma_n^d) \in \mathbb{R}^d$, where γ_n^i represents the amount of type-i stock at time n. The random variables (β_n, γ_n) are $\mathcal{F}_{n-1}$-measurable, which means the investor has to decide at time $n-1$ how to invest on time n. That is the sequence (β_n, γ_n) is *predictable*. For simplicity

$$\gamma_n S_n = \sum_{i=1}^{d} \gamma_n^i S_n^i.$$

The wealth of the investor at time n under the strategy π is

$$X_n^\pi = \beta_n B_n + \gamma_n S_n.$$

This is the *value process* of the investment portfolio.

A strategy is *self-financing (SF)* if the investor does not take out money from, and does not invest money to the portfolio after time 0. That is π is self-financing if

$$X_{n-1}^\pi = \beta_n B_{n-1} + \gamma_n S_{n-1} \quad \text{for all } n.$$

For a sequence a_n put $\Delta a_n = a_n - a_{n-1}$.

{lemma:SF}

Lemma 1. *The following are equivalent:*

(i) π *is SF;*

(ii) $\Delta X_n^\pi = \beta_n \Delta B_n + \gamma_n \Delta S_n$*;*

(iii) $B_{n-1}\Delta\beta_n + S_{n-1}\Delta\gamma_n = 0$*.*

Proof. We have

$$\begin{aligned}
&\Delta X_n = X_n - X_{n-1} \\
&= \beta_n B_n - \beta_{n-1}B_{n-1} + \gamma_n S_n - \gamma_{n-1}S_{n-1} \\
&= \beta_n(B_n - B_{n-1}) + (\beta_n - \beta_{n-1})B_{n-1} + \gamma_n(S_n - S_{n-1}) + (\gamma_n - \gamma_{n-1})S_{n-1} \\
&= \beta_n\Delta B_n + \Delta\beta_n B_{n-1} + \gamma_n\Delta S_n + \Delta\gamma_n S_{n-1},
\end{aligned}$$

and the equivalence follows. □

In what follows, unless otherwise stated all the strategies are meant to be SF.

We can decompose the value process as

$$\begin{aligned}
X_n^\pi &= X_{n-1}^\pi + \Delta X_n^\pi = \dots \\
&= X_0^\pi + \sum_{i=1}^n (\beta_i \Delta B_i + \gamma_i \Delta S_i) \\
&=: X_0^\pi + G_n^\pi,
\end{aligned}$$

where G_n^π is the *gain process*. So the value of the strategy is the initial investment plus the gain.

2.2 Claim and hedging

Let f_N be a nonnegative random variable, which is the *payoff function*, or *obligation*, or *contingent claim*. A strategy π is an *upper* (x, f_N)*-hedge*, if $\mathbf{P}$-almost surely

$$X_0^\pi = x, \quad X_N^\pi \geq f_N.$$

It is a *lower* (x, f_N)*-hedge*, if a.s.

$$X_0^\pi = x, \quad X_N^\pi \leq f_N.$$

The hedge is perfect if $=$ holds a.s.

Put

$$C^*(f_N) = \inf\{x : \exists \text{ upper } (x, f_N)\text{-hedge } \},$$

and similarly

$$C_*(f_N) = \sup\{x : \exists \text{ lower } (x, f_N)\text{-hedge } \}.$$

For the class of upper (x, f_N)-hedge strategies put $H^*(x, f_N, \mathbf{P})$, and for the lower $H_*(x, f_N, \mathbf{P})$.

{lemma:hedge}

Lemma 2. *For any payoff function f_N there exists an x such that there is an upper (x, f_N)-hedge.*

Proof. Put

$$x = \frac{B_0}{B_N} \max_{\omega \in \Omega} |f_N(\omega)|.$$

Then the (trivial) strategy $\pi_n \equiv (\frac{x}{B_0}, 0)$ (start with enough money and don't do anything) is an upper hedge. □

2.3 Binomial market

{ss:bin}

2.3.1 One-step market

Consider a one-step binomial market with $d = 1$ stock. That is $\Omega = \{0, 1\}$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_1 = \mathcal{F} = 2^\Omega$. Assume that $\mathbf{P}(\{0\}) \in (0, 1)$. The bond price $B_1 = (1 + r)B_0$, that is $r > -1$ is the interest rate, and for some $a < b$, $S_1 = (1 + \rho)S_0$, $\rho \in \{a, b\}$. Say, $\rho(1) = b$, $\rho(0) = a$. Let f be a payoff, that is $f(0) = f_0$, $f(1) = f_1$. We construct a perfect hedge.

Using the strategy $\pi_1 = (\beta_1, \gamma_1)$ we want that

$$X_1^\pi = \beta_1 B_1 + \gamma_1 S_1 = f \quad \text{a.s.}$$

Since there are only two possibilities, a.s. means

$$\begin{aligned}\beta_1 B_0(1+r) + \gamma_1 S_0(1+a) &= f_0 \\ \beta_1 B_0(1+r) + \gamma_1 S_0(1+b) &= f_1.\end{aligned}$$

Solving the linear system

$$\gamma_1 = \frac{1}{S_0}\frac{f_1 - f_0}{b-a}, \quad \beta_1 = \frac{f_1 - (1+b)\frac{f_1-f_0}{b-a}}{B_0(1+r)}.$$

This is deterministic, so $\mathcal{F}_0$-measurable, as it should be. The initial cost of this strategy is

$$X_0^\pi = B_0\beta_1 + S_0\gamma_1 = \frac{1}{1+r}\left(\frac{r-a}{b-a}f_1 + \frac{b-r}{b-a}f_0.\right)$$

If $a < r < b$ this can be written as

$$X_0^\pi = \frac{1}{1+r}\mathbf{E}_{\mathbf{Q}} f,$$

with the probability measure $\mathbf{Q}(\{0\}) = (b-r)/(b-a)$, $\mathbf{Q}(\{1\}) = (r-a)/(b-a)$.

This shows that the 'fair' price of the payoff is $\mathbf{E}_{\mathbf{Q}} f/(1+r)$. Note that this does not depend on the probability measure $\mathbf{P}$.

2.3.2 N-step market

Assume we have only one stock, $d = 1$. For the bond $B_n = (1+r_n)B_{n-1}$, and for the share $S_n = (1+\rho_n)S_{n-1}$, where $\rho_n \in \{a_n, b_n\}$.

Exercise 1. Give a concrete construction of the probability space and the filtration!

Solution 1. Let

$$\Omega = \{0,1\}^N = \{\omega = (\omega_1, \dots, \omega_N) : \omega_i \in \{0,1\}\}.$$

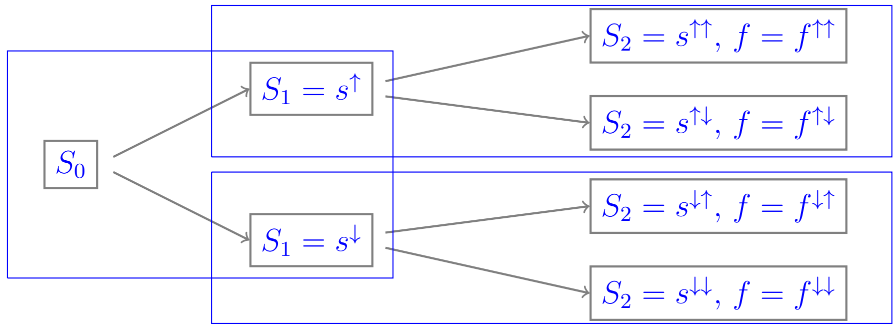

Figure 4: 2-step binary market as 3 1-step binary market

Define the random variables $\rho_n : \Omega \to \{a_n, b_n\}$ as

$$\rho_n(\omega) = \begin{cases} a_n, & \text{if } \omega_n = 0, \\ b_n, & \text{if } \omega_n = 1. \end{cases}$$

For the filtration let $\mathcal{F}_n = \sigma(\rho_1, \dots, \rho_n)$, i.e. the natural filtration generated by the variables $\rho_1, \dots, \rho_n$.

Consider any payoff function f_N. A perfect hedge can be constructed recursively, using the simple one-step market. Indeed, a two-step model can be seen as 3 one-step markets.

3 Arbitrage and pricing in discrete time

3.1 Arbitrage

A SF strategy π is an *arbitrage strategy* if

- $X_0^\pi = 0$;
- $X_n^\pi \geq 0$ for all $n = 0, 1, \dots, N$;
- $\mathbf{P}(X_N^\pi > 0) > 0$.

That is, using the strategy π with 0 money we have riskless profit.

If the second assumption only holds for $n = N$ then π is a *weak arbitrage strategy*. According to the following if weak arbitrage strategy exists, then also arbitrage strategy exists.

{lemma:arbitrage}

Lemma 3. *Assume that π is a weak arbitrage strategy. Then there exists an arbitrage strategy π'.*

Proof. If $X_n^\pi \geq 0$ a.s. for all n, we are ready. Otherwise, there exists $m < N$ such that $\mathbf{P}(X_m^\pi < 0) > 0$, and $X_n^\pi \geq 0$ for any $n \geq m+1$. Let

$$A_m = \{X_m < 0\} \in \mathcal{F}_m.$$

Consider the strategy

$$\beta_n' = \mathbf{I}(A_m)\mathbf{I}(n > m)\left(\beta_n - \frac{X_m}{B_m}\right), \quad \gamma_n' = \mathbf{I}(A_m)\mathbf{I}(n > m)\gamma_n.$$

It is easy to check that this strategy is predictable, SF, and arbitrage strategy. Indeed,

(i) predictable: for $n \leq m$ this is clear, since $\beta_n' = 0$ and $\gamma_n' = 0$, while for $n > m$ A_m is $\mathcal{F}_m$-measurable and thus $\mathcal{F}_{n-1}$-measurable as well, and β_n, γ_n are $\mathcal{F}_{n-1}$-measurable by the assumption.

(ii) SF: for $n \leq m$ this is again clear. For $n = m+1$

$$\begin{aligned} &B_m \Delta\beta_{m+1}' + S_m \Delta\gamma_{m+1}' \\ &= \mathbf{I}(A_m)\left(B_m\beta_{m+1}(\omega) - X_m^\pi(\omega) + S_m\gamma_{m+1}(\omega)\right) = 0, \end{aligned}$$

since π is SF. For $n > m+1$ we have $\Delta\beta_n' = I_{A_m}\Delta\beta_n$, and $\Delta\gamma_n' = I_{A_m}\Delta\gamma_n$, and the result follows, using again that π is SF.

(iii) arbitrage: we have

$$X_n^{\pi'} = \mathbf{I}(A_m)\mathbf{I}(n > m)\left(\beta_n B_n + \gamma_n S_n - \frac{X_m^\pi B_n}{B_m}\right),$$

where the sum of the first two terms in the bracket is nonnegative by the definition of m and the last is strictly negative on A_m, which proves the statement.

$\square$

Exercise 2. Assume that $a < b < r$ in the one-step binomial model. Give an arbitrage strategy.

Assume that $a_n < b_n < r_n$ for some n in the N-step binomial model. Give an arbitrage strategy.

3.2 Martingale measures

A probability measure $\mathbf{Q}$ is called *equivalent martingale measure* (EMM) if $\mathbf{P} \sim \mathbf{Q}$ and $(S_n^i/B_n, \mathcal{F}_n)$ is a $\mathbf{Q}$-martingale for each $i = 1, 2, \ldots, d$.

3.2.1 EMM in binomial markets

In a one-step binomial market the martingale property is easy to check. Indeed, $(S_i/B_i)_{i=0,1}$ is a martingale iff

$$\mathbf{E_Q}\left[\frac{S_1}{B_1}\middle|\mathcal{F}_0\right] = \frac{S_0}{B_0}.$$

We have

$$\begin{aligned}
\mathbf{E_Q}\left[\frac{S_1}{B_1}\middle|\mathcal{F}_0\right] &= \mathbf{E_Q}\frac{S_1}{B_1} \\
&= \mathbf{Q}(\rho = a)\frac{(1+a)S_0}{(1+r)B_0} + (1 - \mathbf{Q}(\rho = a))\frac{(1+b)S_0}{(1+r)B_0} \\
&= \frac{S_0}{B_0}.
\end{aligned}$$

Solving the equation we obtain that

$$\mathbf{Q}(\rho = a) = \frac{b-r}{b-a}, \quad \text{and} \quad \mathbf{Q}(\rho = b) = \frac{r-a}{b-a}.$$

That is $\mathbf{Q}(\{0\}) = (b-r)/(b-a)$, $\mathbf{Q}(\{1\}) = (r-a)/(b-a)$. This is the probability measure $\mathbf{Q}$ we obtained at pricing.

Let us see the general N-step model. Then

$$S_n = \prod_{i=1}^{n}(1+\rho_i)\, S_0,$$

thus the martingale property reads as

$$\mathbf{E_Q}\left[\frac{S_n}{B_n}\middle|\mathcal{F}_{n-1}\right] = \frac{S_{n-1}}{B_{n-1}} \quad n = 0, 1, \ldots N.$$

Using the properties of conditional expectation we have

$$\mathbf{E_Q}\left[\frac{S_n}{B_n}\middle|\mathcal{F}_{n-1}\right] = \frac{S_{n-1}}{B_{n-1}}\frac{1}{1+r_n}\mathbf{E_Q}[1+\rho_n|\mathcal{F}_{n-1}].$$

Therefore S_n/B_n is a $\mathbf{Q}$-martingale iff

$$\mathbf{E}_{\mathbf{Q}}[\rho_n|\mathcal{F}_{n-1}] = r_n.$$

This condition exactly tells that under the new measure $\mathbf{Q}$ the risky asset behaves as the bond on average. Using that $\rho_n \in \{a_n, b_n\}$, we obtain as above

$$\mathbf{Q}(\rho_n = a_n|\mathcal{F}_{n-1}) = \frac{b_n - r_n}{b_n - a_n}, \quad \text{and } \mathbf{Q}(\rho_n = b_n|\mathcal{F}_{n-1}) = \frac{r_n - a_n}{b_n - a_n}.$$

Note the conditioning on $\mathcal{F}_{n-1}$ gives a constant, meaning that ρ_n is independent of $\mathcal{F}_{n-1}$ under the measure $\mathbf{Q}$.

We obtained the following.

{thm:binom-EMM}

Theorem 1. *In the binomial market if $a_n < r_n < b_n$ for each n then there exists a unique EMM $\mathbf{Q}$ given by the formulas above. Moreover, under $\mathbf{Q}$ the random variables $\rho_1, \ldots, \rho_N$ are independent.*

In the proof we used the following simple result.

Exercise 3. Assume that $Y \in \{a, b\}$ and

$$\mathbf{P}(Y = a|\mathcal{F}) = p \text{ a.s.}$$

Show that Y is independent of $\mathcal{F}$.

Note that the original measure $\mathbf{P}$ is irrelevant.

In the special case of the homogeneous binomial market we get that

$$\mathbf{Q}(S_N = S_0(1+b)^k(1+a)^{N-k}) = \binom{N}{k} q^k (1-q)^{N-k}, \quad k = 0, 1, \ldots, N.$$

3.2.2 Pricing with EMM

{prop:Xbar-mtg}

Proposition 1. *If $\mathbf{Q}$ is an EMM then $(\overline{X}_n^\pi = X_n^\pi / B_n)_n$ is a $\mathbf{Q}$-martingale for any SF strategy π.*

Proof. Easily follows from the SF property. Indeed, using that β_n, γ_n are $\mathcal{F}_{n-1}$-measurable

$$\begin{aligned}
\mathbf{E}_\mathbf{Q}\left[\frac{X_n^\pi}{B_n}\middle|\mathcal{F}_{n-1}\right] &= \mathbf{E}_\mathbf{Q}\left[\beta_n + \gamma_n\frac{S_n}{B_n}\middle|\mathcal{F}_{n-1}\right] \\
&= \beta_n + \gamma_n \mathbf{E}_\mathbf{Q}\left[\frac{S_n}{B_n}\middle|\mathcal{F}_{n-1}\right] \\
&= \beta_n + \gamma_n\frac{S_{n-1}}{B_{n-1}} \\
&= \frac{\beta_n B_{n-1} + \gamma_n S_{n-1}}{B_{n-1}} \\
&= \frac{X_{n-1}^\pi}{B_{n-1}},
\end{aligned}$$

where the last equality follow from the self-financing property. □

The following main result is the *first fundamental theorem of asset pricing.*

{thm:emm-arb}

Theorem 2. *There exists an EMM if and only if the market is arbitrage-free.*

Proof. Let $\mathbf{Q}$ be an EMM and π be any strategy with $X_0^\pi = 0$. Then, by the previous statement

$$\mathbf{E}_\mathbf{Q}\frac{X_N^\pi}{B_N} = \mathbf{E}_\mathbf{Q}\frac{X_0^\pi}{B_0} = 0.$$

Thus $X_N \geq 0$ $\mathbf{P}$-a.s., then also $\mathbf{Q}$-a.s., which implies $X_N \equiv 0$ $\mathbf{Q}$-a.s., thus $\mathbf{P}$-a.s.

We prove the converse later. □

Assume that f_N is a replicable payoff, i.e. there is a prefect hedge π. This means that

$$X_N^\pi = f_N \quad \text{a.s.}$$

Then the fair price for f_N is the initial cost of the portfolio, $X_0^\pi = x$. By the martingale property

$$\mathbf{E}_\mathbf{Q}\frac{f_N}{B_N} = \mathbf{E}_\mathbf{Q}\frac{X_N^\pi}{B_N} \overset{\text{mtg}}{=} \mathbf{E}_\mathbf{Q}\frac{X_0^\pi}{B_0} = \frac{x}{B_0}.$$

That is, the fair price x for a replicable payoff f_N is

$$x = \frac{B_0}{B_N}\mathbf{E}_\mathbf{Q} f.$$

In particular, it also follows that for a replicable f, the value $\mathbf{E}_{\mathbf{Q}} f$ is the same for any EMM $\mathbf{Q}$.

Summarizing, we proved the following:

{thm:pricing}

Theorem 3. *Consider an arbitrage-free market and let f be a replicable payoff. Then the fair price of f is*

$$C(f) = C_* = C^* = \frac{B_0}{B_N} \mathbf{E}_{\mathbf{Q}} f,$$

where $\mathbf{Q}$ is any EMM.

3.3 Complete markets

We proved that if EMM exists then we have the fair price for any replicable payoff. A market is *complete* if any payoff is replicable.

We have seen in Theorem 3 that on a complete arbitrage-free market any payoff f has a unique well-defined fair price $B_0 \mathbf{E}_{\mathbf{Q}} f / B_N$.

In section 2.3 we showed that a binomial market is complete.

The second fundamental theorem of asset pricing is the following.

{thm:complete-mark

Theorem 4. *Consider an arbitrage-free market with EMM $\mathbf{Q}$. Then the following are equivalent:*

(i) the market is complete;

(ii) $\mathbf{Q}$ is the unique EMM;

(iii) for any $\mathbf{Q}$-martingale (M_n) there exists a predictable sequence γ_n such that M_n can be represented as

$$M_n = M_0 + \sum_{k=1}^{n} \gamma_k \left(\frac{S_k}{B_k} - \frac{S_{k-1}}{B_{k-1}} \right) = M_0 + \sum_{k=1}^{n} \sum_{i=1}^{d} \gamma_k^i \left(\frac{S_k^i}{B_k} - \frac{S_{k-1}^i}{B_{k-1}} \right).$$

Proof. We prove again the easy parts (i) $\Rightarrow$ (ii), and (iii) $\Leftrightarrow$ (i), and postpone the difficult (ii) $\Rightarrow$ (i) implication later.

(i) $\Rightarrow$ (ii): Assume that $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are EMM's. Consider any $A \in \mathcal{F}$. We show that $\mathbf{Q}_1(A) = \mathbf{Q}_2(A)$ implying the uniqueness. Let π be a perfect hedge to $f = I_A$. Then X_n^π / B_n is both $\mathbf{Q}_1$ and $\mathbf{Q}_2$ martingale, so

$$\mathbf{Q}_1(A) = \mathbf{E}_{\mathbf{Q}_1} f = \mathbf{E}_{\mathbf{Q}_1} X_N^\pi = B_N \mathbf{E}_{\mathbf{Q}_1} \frac{X_N^\pi}{B_N} = B_N \frac{X_0^\pi}{B_0} = \ldots = \mathbf{Q}_2(A).$$

(i) $\Rightarrow$ (iii): Consider a $\mathbf{Q}$-martingale M_n. There exists a strategy π_n such that a.s.

$$X_N^\pi = B_N M_N.$$

Using that both M_n and X_n^π/B_n are martingales

$$M_n = \mathbf{E_Q}[M_N|\mathcal{F}_n] = \mathbf{E_Q}\left[\frac{X_N^\pi}{B_N}|\mathcal{F}_n\right] = \frac{X_n^\pi}{B_n} = \beta_n + \gamma_n\frac{S_n}{B_n}.$$

Thus, using that π is SF

$$\begin{aligned}
M_n - M_{n-1} &= \Delta\beta_n + \gamma_n\frac{S_n}{B_n} - \gamma_{n-1}\frac{S_{n-1}}{B_{n-1}} \\
&= \gamma_n\left(\frac{S_n}{B_n} - \frac{S_{n-1}}{B_{n-1}}\right) + \frac{1}{B_{n-1}}\left(B_{n-1}\Delta\beta_n + S_{n-1}\Delta\gamma_n\right) \\
&= \gamma_n\left(\frac{S_n}{B_n} - \frac{S_{n-1}}{B_{n-1}}\right),
\end{aligned}$$

as claimed.

(iii) $\Rightarrow$ (i): Consider a payoff f. We are looking for a strategy π such that $X_N^\pi = f$ $\mathbf{Q}$-a.s. We know that $(X_n^\pi/B_n)_n$ is a martingale, so this should be (M_n). Now the following choice is clear: let

$$M_n = \mathbf{E_Q}\left[\frac{f}{B_N}|\mathcal{F}_n\right].$$

Then M_n is a martingale, therefore by the assumption

$$M_n = M_0 + \sum_{k=1}^{n}\gamma_k\Delta\frac{S_k}{B_k}.$$

Let

$$\beta_n = M_n - \gamma_n\frac{S_n}{B_n},$$

and consider the strategy $\pi_n = (\beta_n, \gamma_n)$. To see that this is indeed a strategy we have to show that it is predictable and SF. The sequence γ_n is predictable by the assumption (iii), and β_n is predictable because all the terms in M_n are $\mathcal{F}_{n-1}$-measurable except $\gamma_n S_n/B_n$, which is subtracted. To see that it is

SF note that

$$
\begin{aligned}
&B_{n-1}\Delta\beta_n + S_{n-1}\Delta\gamma_n \\
&= B_{n-1}\left(M_n - M_{n-1} - \gamma_n \frac{S_n}{B_n} + \gamma_{n-1}\frac{S_{n-1}}{B_{n-1}}\right) + S_{n-1}\Delta\gamma_n \\
&= B_{n-1}\left(\gamma_n \Delta\frac{S_n}{B_n} - \gamma_n \frac{S_n}{B_n} + \gamma_{n-1}\frac{S_{n-1}}{B_{n-1}}\right) + S_{n-1}\Delta\gamma_n = 0,
\end{aligned}
$$

showing that π is SF. It is clearly a perfect hedge since

$$X_N^\pi = \beta_N B_N + \gamma_N S_N = B_N M_N = f,$$

as claimed. □

3.4 Proof of the difficult part of Theorem 2

Here we use strongly that Ω is finite, and let $|\Omega| = k$.

Assume that there is no arbitrage strategy. Let

$$\mathcal{V}_0 = \{X : \Omega \to \mathbb{R} \text{ r.v. } | \exists \pi : X_0^\pi = 0 \text{ and } X_N^\pi = X\},$$

and

$$\mathcal{V}_1 = \{X : \Omega \to \mathbb{R} \text{ r.v. } | X \geq 0, \mathbf{E}X \geq 1\}.$$

We identify a random variable $X : \Omega \to \mathbb{R}$ with a vector in $\mathbb{R}^k$, as $X \leftrightarrow (X(\omega_1), \ldots, X(\omega_k))$. Clearly, $\mathcal{V}_0$ is a linear subspace and $\mathcal{V}_1$ is convex set in $\mathbb{R}^k$.

Since there is no arbitrage strategy, $\mathcal{V}_0 \cap \mathcal{V}_1 = \emptyset$. Therefore, by the Kreps–Yan theorem, there exists a linear functional $\ell : \mathbb{R}^k \to \mathbb{R}$ such that $\ell|_{\mathcal{V}_0} \equiv 0$ and $\ell(v_1) > 0$ for all $v_1 \in \mathcal{V}_1$. A linear function in $\mathbb{R}^k$ (in any Hilbert space) is a inner product, thus there exists $q \in \mathbb{R}^k$ such that

$$\ell(v) = \langle v, q \rangle.$$

Define the random variables

$$X_i(\omega_j) = \delta_{i,j}\frac{1}{\mathbf{P}(\{\omega_i\})}.$$

Then $X_i \geq 0$ and $\mathbf{E}X_i = 1$, so $X_i \in \mathcal{V}_1$. Furthermore

$$\ell(X_i) = \frac{q_i}{\mathbf{P}(\{\omega_i\})} > 0,$$

implying $q_i > 0$ for any i. Define the probability measure $\mathbf{Q}$ as

$$\mathbf{Q}(\{\omega_i\}) = \frac{q_i}{\sum_{i=1}^k q_i}.$$

It is clear that $\mathbf{Q} \sim \mathbf{P}$. We have to check that (S_n/B_n) is a $\mathbf{Q}$-martingale. First we need a lemma.

Lemma 4. *Let $(X_n)_{n=1}^N$ be an adapted process. If for any stopping time $\tau : \Omega \to \{0, \ldots, N\}$*

$$\mathbf{E}X_\tau = \mathbf{E}X_0,$$

then (X_n) is martingale.

Proof. We show that $X_n = \mathbf{E}[X_N|\mathcal{F}_n]$, which implies that X is martingale.

Let $A \in \mathcal{F}_n$ and consider the stopping time

$$\tau_A(\omega) = \begin{cases} n, & \omega \in A, \\ N, & \text{otherwise.} \end{cases}$$

This is indeed a stopping time, since $\{\tau_A \leq k\} = \emptyset$ for $k < n$, and A for $k \geq n$, which is $\mathcal{F}_k$-measurable. Then, by the assumption

$$\mathbf{E}X_0 = \mathbf{E}X_{\tau_A} = \mathbf{E}X_n I(A) + \mathbf{E}X_N I(A^c).$$

With $A = \emptyset$ we see that $\mathbf{E}X_0 = \mathbf{E}X_N$, implying

$$\mathbf{E}X_n I(A) = \mathbf{E}X_N I(A).$$

This exactly means that

$$X_n = \mathbf{E}[X_N|\mathcal{F}_n],$$

as claimed. □

We show that (S_n/B_n) satisfies the condition of the lemma above. Let τ be a stopping time and define the strategy

$$\beta_n = \frac{S_\tau}{B_\tau} I(\tau \leq n-1) - \frac{S_0}{B_0}, \quad \gamma_n = I(\tau > n-1).$$

Since $\{\tau < n\} = \{\tau \leq n-1\} \in \mathcal{F}_{n-1}$, the sequence (β_n, γ_n) is predictable. Furthermore,

$$B_{n-1}\Delta\beta_n + S_{n-1}\Delta\gamma_n = \frac{S_\tau}{B_\tau} B_{n-1} I(\tau = n-1) - S_{n-1} I(\tau = n-1) = 0,$$

so it is SF. Finally,

$$X_0^\pi = -\frac{S_0}{B_0} B_0 + S_0 = 0,$$

so $X_N^\pi \in \mathcal{V}_0$. Therefore

$$\begin{aligned}
0 = \mathbf{E}_{\mathbf{Q}} X_N^\pi &= \mathbf{E}_{\mathbf{Q}} \beta_N B_N + \gamma_N S_N \\
&= \mathbf{E}_{\mathbf{Q}} \left(\left(\frac{S_\tau}{B_\tau} I(\tau \le N-1) - \frac{S_0}{B_0} \right) B_N + \frac{S_\tau}{B_\tau} I(\tau = N) B_N \right) \\
&= B_N \mathbf{E}_{\mathbf{Q}} \left(\frac{S_\tau}{B_\tau} - \frac{S_0}{B_0} \right).
\end{aligned}$$

That is (S_n/B_n) is indeed a $\mathbf{Q}$-martingale.

3.5 Proof of the difficult part of Theorem 4

Here we prove the implication (ii) $\Rightarrow$ (i).

We use the notation of the previous proof. Let

$$\mathcal{V}_2 = \{X : \Omega \to \mathbb{R} \text{ r.v. } | \mathbf{E}_{\mathbf{Q}} X = 0\}.$$

Then $\mathcal{V}_2$ is a linear subspace in $\mathbb{R}^k$ and we have seen in the previous proof that $\mathcal{V}_0 \subset \mathcal{V}_2$. We claim that equality holds.

Assume first that this is indeed true. Then for any claim X the centered version $X - \mathbf{E}_{\mathbf{Q}} X \in \mathcal{V}_2 = \mathcal{V}_0$, meaning that there is a perfect hedge. Thus the market is complete. So we only have to show that $\mathcal{V}_0 = \mathcal{V}_2$.

Assume on the contrary that $\mathcal{V}_0 \neq \mathcal{V}_2$. Then there is an $y \in \mathcal{V}_2$, which is orthogonal to $\mathcal{V}_0$. Since $q_i > 0$ (see the previous proof) for all $i = 1, \ldots, k$, we may choose $\varepsilon > 0$ small enough such that

$$q_i' = q_i - \varepsilon y_i > 0 \quad \text{for all } i.$$

As both q and y are orthogonal to $\mathcal{V}_0$, q' is also orthogonal. Define the measure

$$\mathbf{Q}'(\{\omega_i\}) = \frac{q_i'}{\sum_{i=1}^k q_i'}.$$

Exactly as in the previous proof we can show that $\mathbf{Q}'$ is EMM. The uniqueness of the EMM implies

$$\frac{q_i'}{\sum_{i=1}^k q_i'} = \frac{q_i}{\sum_{i=1}^k q_i},$$

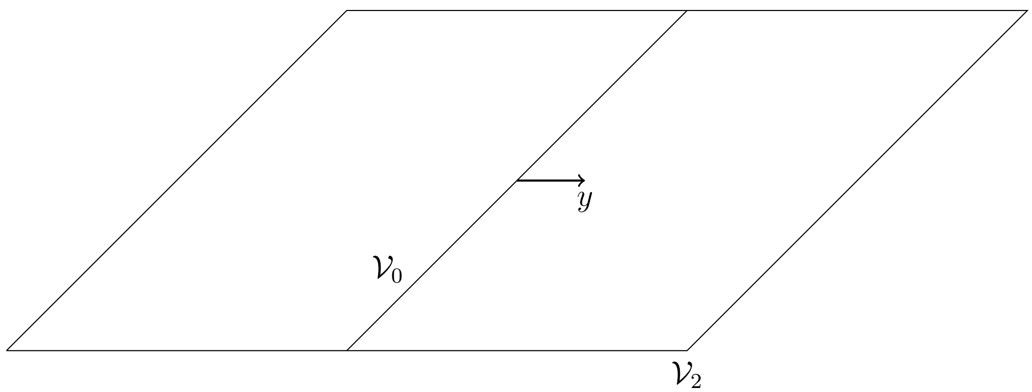

Figure 5: Choice of y

that is, using also the definition of q',

$$q = \alpha q' = \alpha q - \alpha\varepsilon y,$$

with $\alpha = \sum q_i / \sum q_i'$. Thus

$$(1-\alpha)q = -\alpha\varepsilon y.$$

But $y \in \mathcal{V}_2$ and q is chosen so that it is orthogonal to $\mathcal{V}_2$, so $y \perp q$, which is a contradiction. The proof is complete.

4 Pricing and hedging European options

In this section we summarize our findings on pricing and hedging, and consider some special cases in detail.

4.1 Complete markets

Consider an arbitrage-free complete market. The *fair price* of the contingent claim f_N is

$$C(f_N) = \inf\{x : \exists\pi, X_0^\pi = x,\, X_N^\pi = f_N\}.$$

Then, by Theorems 2 and 4 there exists a unique EMM $\mathbf{Q}$. Since (X_n^π/B_n) is $\mathbf{Q}$-martingale

$$\mathbf{E}_\mathbf{Q}\frac{f_N}{B_N} = \mathbf{E}_\mathbf{Q}\frac{X_N^\pi}{B_N} = \mathbf{E}_\mathbf{Q}\frac{x}{B_0} = \frac{x}{B_0},$$

therefore

$$C(f_N) = x = \frac{B_0}{B_N}\mathbf{E_Q} f_N.$$

Note that x is independent of the hedge π itself, that is for different hedges the initial value is the same.

For a hedge we need to know not only the fair price C, but also the strategy π itself. For the given claim f_N consider the martingale

$$M_n = \mathbf{E_Q}\left[\frac{f_N}{B_N}\middle|\mathcal{F}_n\right].$$

By Theorem 4 there exists a representation

$$M_n = M_0 + \sum_{k=1}^{n} \gamma_k \Delta \frac{S_k}{B_k},$$

with a predictable sequence (γ_n). Let

$$\beta_n = M_n - \frac{\gamma_n S_n}{B_n}.$$

We proved that $\pi = (\beta_n, \gamma_n)_n$ is an SF strategy and is a perfect hedge for f_N.

Summarizing, we obtained the following.

Theorem 5. *In an arbitrary arbitrage-free complete market the price of the contingent claim f_N is*

$$C(f_N) = B_0 \mathbf{E_Q} \frac{f_N}{B_N}.$$

Moreover, there exists a strategy π which is a perfect hedge of f_N, i.e.

$$X_N^\pi = f_N,$$

where (β_n, γ_n) are given above. The value process is determined by

$$X_n^\pi = B_n \mathbf{E_Q}\left[\frac{f_N}{B_N}\middle|\mathcal{F}_n\right].$$

4.2 Homogeneous binomial market – CRR formula

Consider a homogeneous binomial N-step market with $a < r < b$. That is

$$B_n = (1+r)^n, \quad S_n = S_0 \prod_{k=1}^{n} (1+\rho_k),$$

where $\rho_k \in \{a, b\}$. We proved that this market is arbitrage-free and complete, and the unique EMM is given by

$$\mathbf{Q}(\rho_i = a) = \frac{b-r}{b-a},$$

and ρ_i's are independent. If the claim f_N only depends on the final price S_N, and not on the whole trajectory, i.e.

$$f_N(\omega) = f_N(S_N(\omega)),$$

then the pricing formula simplifies, and we obtain the Cox–Ross-Rubinstein formula:

$$C(f_N) = \frac{1}{(1+r)^N} \sum_{k=0}^{N} f_N(S_0(1+b)^k(1+a)^{N-k}) \binom{N}{k} q^k (1-q)^{N-k},$$

where $q = \frac{r-a}{b-a}$.

5 American options

While European options can be exercised only at the terminal date N, American options can be exercised at any time. Formally, instead of a fixed random payoff function f_N, a sequence of payoffs $(f_n)_{n=0,1,\ldots,N}$ is given, where f_n is $\mathcal{F}_n$-measurable, i.e. $(f_n)_n$ is adapted to $(\mathcal{F}_n)_n$. So f_n is the random payoff if the option is exercised at time n. Clearly, the exercise time has to be a stopping time.

5.1 Optimal stopping problems

Consider a probability space with a filtration $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n=0,1,\ldots,N}, \mathbf{P})$, and let $\mathcal{M}$ denote the set of stopping times. Consider a sequence of nonnegative adapted random variables $(X_n)_n$, and define by backward induction its Snell-envelope $(Z_n)_n$ as follows. We are interested in the value

$$Z_N = X_N, \quad Z_n = \max\{X_n, \mathbf{E}[Z_{n+1}|\mathcal{F}_n]\},\ n < N.$$

For a stopping time τ the stopped process is denoted by Z^τ, i.e.

$$Z_n^\tau = Z_{\tau\wedge n},$$

where $a \wedge b = \min\{a, b\}$.

Proposition 2. *Let (Z_n) be the Snell-envelope of (X_n) with $X_n \geq 0$ a.s.*

(i) Z is the smallest supermartingale dominating X.

(ii) The random variable $\tau^* = \min\{n : Z_n = X_n\}$ is a stopping time and the stopped process $Z_{n\wedge\tau^*} = Z_n^{\tau^*}$ is martingale.

Proof. From the definition it is clear that Z is supermartingale and dominates X. Let Y be another supermartingale dominating X. Then $Y_N \geq X_N = Z_N$. Assuming that $Y_n \geq Z_n$ we have

$$Y_{n-1} \geq \max\{\mathbf{E}[Y_n|\mathcal{F}_{n-1}], X_{n-1}\} \geq \max\{\mathbf{E}[Z_n|\mathcal{F}_{n-1}], X_{n-1}\} = Z_{n-1}.$$

Thus the minimality follows.

To see that τ^* is stopping time note that

$$\{\tau^* = n\} = \cap_{k=0}^{n-1}\{Z_k > X_k\} \cap \{Z_n = X_n\}.$$

For the last assertion note that

$$Z_n^{\tau^*} - Z_{n-1}^{\tau^*} = \mathbf{I}(\tau^* \geq n)(Z_n - Z_{n-1}).$$

On the event $\{\tau^* \geq n\}$ we have $Z_{n-1} = \mathbf{E}[Z_n|\mathcal{F}_{n-1}]$ therefore

$$\mathbf{E}[\mathbf{I}(\tau^* \geq n)(Z_n - Z_{n-1})|\mathcal{F}_{n-1}] = 0.$$

□

A stopping time σ is optimal if

$$\mathbf{E}(X_\sigma) = \sup_{\tau\in\mathcal{M}} \mathbf{E}(X_\tau).$$

Proposition 3. *The stopping time τ^* is optimal for X, and*

$$Z_0 = \mathbf{E} X_{\tau^*} = \sup_{\tau \in \mathcal{M}} \mathbf{E} X_\tau.$$

Proof. Since Z^{τ^*} is martingale

$$Z_0 = Z_0^{\tau^*} = \mathbf{E} Z_N^{\tau^*} = \mathbf{E} Z_{\tau^*} = \mathbf{E} X_{\tau^*}.$$

On the other hand for any stopping time τ the process Z^τ is supermartingale (by Doob's optional sampling), thus

$$Z_0 = \mathbf{E} Z_0^\tau \geq \mathbf{E} Z_\tau \geq \mathbf{E} X_\tau.$$

□

5.2 Pricing American options

Let us return to our pricing problem. Assume that we have an arbitrage-free complete market, that is the EMM $\mathbf{Q}$ is unique. Let $(f_n)_{n=0,\dots,N}$ be the payoff of an American option. A hedging strategy now has to fulfill the conditions

$$X_n^\pi \geq f_n, \quad n = 0, 1, \dots, N,$$

as the option can be exercised at any time. A hedge is *minimal*, if for a stopping time τ^* we have $X_{\tau^*}^\pi = f_{\tau^*}$.

By Doob's optional stopping $(X_0^\pi / B_0, X_\tau^\pi / B_\tau)$ is martingale for any stopping time τ, i.e.

$$\frac{x}{B_0} = \mathbf{E}_\mathbf{Q} \frac{X_0^\pi}{B_0} = \mathbf{E}_\mathbf{Q} \frac{X_\tau^\pi}{B_\tau} \geq \mathbf{E}_\mathbf{Q} \frac{f_\tau}{B_\tau}.$$

Therefore the initial cost of the hedge is at least

$$x \geq B_0 \sup_{\tau \in \mathcal{M}_0^N} \mathbf{E}_\mathbf{Q} \frac{f_\tau}{B_\tau}. \tag{1}$$ {eq:am-opt-ineq}

For a hedging strategy π we have that

(i) $(X_n^\pi / B_n)_n$ is a $\mathbf{Q}$-martingale (since $\mathbf{Q}$ is EMM and π is SF), and

(ii) (X_n^π / B_n) dominates (f_n / B_n) (since π is a hedge).

Therefore, the value process of a hedge is larger than the Snell-envelope of (f_n/B_n), i.e.

$$\frac{X_n^\pi}{B_n} \geq Z_n, \quad n = 0, 1, \ldots, N, \tag{2}$$

{eq:di-american-1}

where (Z_n) is the Snell-envelope of (f_n/B_n). The Snell-envelope (Z_n) is a supermartingale, therefore by the Doob-decomposition (that's stated for submartingale, but multiply by -1) we have

$$Z_n = M_n - A_n, \quad n = 0, 1, \ldots, N, \tag{3}$$

{eq:di-american-2}

where M_n is a $\mathbf{Q}$-martingale, and (A_n) is an increasing predictable sequence, $A_0 = 0$.

The market is complete, therefore (see the easy parts of the proof of Theorem 4) there exists a strategy π such that

$$\frac{X_n^\pi}{B_n} = M_n, \quad n = 0, 1, \ldots, N.$$

This is a hedging strategy with initial cost

$$\frac{x}{B_0} = \frac{X_0^\pi}{B_0} = M_0 = Z_0.$$

Comparing to (1), we see that π is optimal.

{thm:price-di-amer

Theorem 6. *Consider an arbitrage-free complete market with unique EMM $\mathbf{Q}$. Let (f_n) be the nonnegative payoff sequence of an American option. Let (Z_n) be the Snell-envelope of the discounted payoff sequence (f_n/B_n). The fair price for this option is*

$$C = B_0 Z_0 = B_0 \sup_{\tau \in \mathcal{M}_0^N} \mathbf{E_Q} \frac{f_\tau}{B_\tau} = B_0 \mathbf{E_Q} \frac{f_{\tau^*}}{B_{\tau^*}},$$

where τ^* is an (not unique in general) optimal exercise time given by

$$\tau^* = \min \left\{ n : \frac{f_n}{B_n} = Z_n \right\}.$$

Furthermore, there exists a SF strategy π which is an optimal hedge with initial cost C and

$$X_{\tau^*}^\pi = \frac{f_{\tau^*}}{B_{\tau^*}}.$$

5.3 American vs. European options

Clearly, an American option with payoff sequence $(f_n)_{n=0,1,\dots,N}$ worth at least as a European option with payoff f_N. However, in some cases the fair prices are equal.

Consider an American call option with strike price K, that is

$$f_n = f(S_n) = (S_n - K)_+.$$

Assume that the deterministic sequence (B_n) is nondecreasing (i.e. the interest rate is nonnegative). Let (Z_n) denote the Snell envelope of (f_n/B_n), that is

$$Z_N = \frac{f_N}{B_N}, \quad Z_n = \max\left\{ \frac{f_n}{B_n}, \mathbf{E}\left[Z_{n+1}|\mathcal{F}_n\right] \right\}, \quad n = 0, 1, \dots, N-1.$$

Using that (S_n/B_n) is a $\mathbf{Q}$-martingale, by Jensen's inequality

$$\begin{aligned}
\frac{f_{N-1}}{B_{N-1}} &= \frac{(S_{N-1} - K)_+}{B_{N-1}} \\
&= \left(\frac{S_{N-1}}{B_{N-1}} - \frac{K}{B_{N-1}} \right)_+ \\
&\leq \mathbf{E}_\mathbf{Q}\left[\left(\frac{S_N}{B_N} - \frac{K}{B_{N-1}} \right)_+ \middle| \mathcal{F}_{N-1} \right] && \text{Jensen's inequality} \\
&\leq \mathbf{E}_\mathbf{Q}\left[\left(\frac{S_N}{B_N} - \frac{K}{B_N} \right)_+ \middle| \mathcal{F}_{N-1} \right] && \text{by } B_N \geq B_{N-1} \\
&= \mathbf{E}_\mathbf{Q}\left[\frac{(S_N - K)_+}{B_N} \middle| \mathcal{F}_{N-1} \right] \\
&= \mathbf{E}_\mathbf{Q}[Z_N|\mathcal{F}_{N-1}].
\end{aligned}$$

This means that at time $N-1$ it is always good to hold the option and continue to step N.

An induction argument shows that at any time it is better to hold the option. Indeed, assume for some n

$$\frac{f_n}{B_n} \leq \mathbf{E}_\mathbf{Q}[Z_{n+1}|\mathcal{F}_n].$$

We just proved this for $n = N - 1$. The same way as above we have

$$\begin{aligned}
\frac{f_{n-1}}{B_{n-1}} &= \frac{(S_{n-1} - K)_+}{B_{n-1}} \\
&= \left(\frac{S_{n-1}}{B_{n-1}} - \frac{K}{B_{n-1}} \right)_+ \\
&\leq \mathbf{E_Q} \left[\left(\frac{S_n}{B_n} - \frac{K}{B_{n-1}} \right)_+ \middle| \mathcal{F}_{n-1} \right] && \text{Jensen's inequality} \\
&\leq \mathbf{E_Q} \left[\left(\frac{S_n}{B_n} - \frac{K}{B_n} \right)_+ \middle| \mathcal{F}_{n-1} \right] && \text{by } B_n \geq B_{n-1} \\
&= \mathbf{E_Q} \left[\frac{(S_n - K)_+}{B_n} \middle| \mathcal{F}_{n-1} \right] \\
&= \mathbf{E_Q} \left[\frac{f_n}{B_n} \middle| \mathcal{F}_{n-1} \right] \\
&\leq \mathbf{E_Q} \left[\mathbf{E_Q}[Z_{n+1} | \mathcal{F}_n] \middle| \mathcal{F}_{n-1} \right] && \text{induction} \\
&\leq \mathbf{E_Q}[Z_n | \mathcal{F}_{n-1}] && Z \text{ supermartingale}
\end{aligned}$$

Thus $\tau^* \equiv N$ is an optimal stopping time, which means that no matter what happens, we wait until the end. Then the American option behaves as the European, so the prices are equal.

Theorem 7. *Assume that the market is arbitrage free and complete, and the interest rate is nonnegative. Then the price of a European call option equals to the price of the American call option.*

6 Stochastic integral

This part is from Karatzas and Shreve [6], in a rather simplified way. Stochastic integration is only worked out in detail with respect to SBM, and not with respect to a continuous martingales. A lot of technical details are omitted.

Here we define the integration with respect to the Brownian motion. Note that SBM is not of bounded variation, therefore we cannot define the integral pathwise. This is the major difficulty in the theory.

6.1 Integration of simple processes

In what follows we work on $[0, T]$, for $T < \infty$. Let $(W_t, \mathcal{F}_t)$ be SBM.

The process (X_t) is a *simple process*, if

$$X_t(\omega) = \xi_0(\omega)\mathbf{I}_{\{0\}}(t) + \sum_{i=1}^{n-1} \xi_i(\omega)\mathbf{I}_{(t_i,t_{i+1}]}(t),$$

where $0 = t_0 < t_1 < \ldots < t_n = T$ is a partition of $[0,T]$, and ξ_i is $\mathcal{F}_{t_i}$-measurable.

That is $(X_t(\omega))$ is a step function for each $\omega \in \Omega$, where the step sizes are random. Note that ξ_i is measurable with respect to the σ-algebra corresponding to the left end point of the interval.

Exercise 4. Show that a simple process is adapted.

The definition of the integral of simple processes is straightforward. Let k be such that $t \in (t_k, t_{k+1}]$. Then

$$I_t(X) = \int_0^t X_s \mathrm{d}W_s = \sum_{i=0}^{k-1} \xi_i(W_{t_{i+1}} - W_{t_i}) + \xi_k(W_t - W_{t_k}), \quad t \in [0,T].$$

Note that we defined the process for each $t \in [0,T]$.

{thm:stint-prop}

Theorem 8. *Let X, Y be simple processes with square integrable coefficients.*

(i) $I_t(X)$ is a continuous martingale, $I_0(X) = 0$ a.s.

(ii) For $t > s$

$$\mathbf{E}\left[\left(\int_s^t X_u \mathrm{d}W_u\right)^2 \Big| \mathcal{F}_s\right] = \mathbf{E}\left[\int_s^t X_u^2 \mathrm{d}u \Big| \mathcal{F}_s\right];$$

in particular $\mathbf{E} I_t(X)^2 = \mathbf{E}\int_0^t X_u^2 \mathrm{d}u$.

(iii) The integral is linear, that is

$$I(\alpha X + \beta Y) = \alpha I(X) + \beta I(Y), \quad \alpha, \beta \in \mathbb{R}.$$

(iv) $\mathbf{E}\sup_{0 \le t \le T}\left(\int_0^t X_u \mathrm{d}W_u\right)^2 \le 4\mathbf{E}\int_0^T X_u^2 \mathrm{d}u$.

Proof. (iii) is clear. (iv) follows from Doob's maximal inequality.

(i) The continuity is obvious and $I_0(X) = 0$. We prove that (I_t) is martingale. Let $s < t$ and $s \in (t_k, t_{k+1}], t \in (t_m, t_{m+1}]$. Then

$$\begin{aligned}\int_0^t X_u \mathrm{d}W_u = & \sum_{i=0}^{k-1} \xi_i (W_{t_{i+1}} - W_{t_i}) + \xi_k (W_s - W_{t_k}) \\ & + \xi_k (W_{t_{k+1}} - W_s) + \sum_{i=k+1}^{m-1} \xi_i (W_{t_{i+1}} - W_{t_i}) + \xi_m (W_t - W_{t_m}).\end{aligned}$$

By the tower rule

$$\begin{aligned}\mathbf{E}[\xi_i (W_{t_{i+1}} - W_{t_i}) | \mathcal{F}_s] &= \mathbf{E}\left[\mathbf{E}[\xi_i (W_{t_{i+1}} - W_{t_i}) | \mathcal{F}_{t_i}] | \mathcal{F}_s\right] \\ &= \mathbf{E}\left[\xi_i \mathbf{E}[W_{t_{i+1}} - W_{t_i} | \mathcal{F}_{t_i}] | \mathcal{F}_s\right] \\ &= \mathbf{E}[\xi_i \cdot 0 | \mathcal{F}_s] = 0.\end{aligned}$$

The first and last term can be handled similarly.

(ii) We showed that

$$\int_s^t X_u \mathrm{d}W_u = \xi_k (W_{t_{k+1}} - W_s) + \sum_{i=k+1}^{m-1} \xi_i (W_{t_{i+1}} - W_{t_i}) + \xi_m (W_t - W_{t_m}).$$

Taking square and conditional expectation we end up with sum of terms

$$\mathbf{E}[\xi_i (W_{t_{i+1}} - W_{t_i}) \xi_j (W_{t_{j+1}} - W_{t_j}) | \mathcal{F}_s]$$

We show that this equals 0, whenever $i \neq j$. Indeed,

$$\begin{aligned}&\mathbf{E}[\xi_i (W_{t_{i+1}} - W_{t_i}) \xi_j (W_{t_{j+1}} - W_{t_j}) | \mathcal{F}_s] \\ &= \mathbf{E}\left[\mathbf{E}[\xi_i (W_{t_{i+1}} - W_{t_i}) \xi_j (W_{t_{j+1}} - W_{t_j}) | \mathcal{F}_{t_j}] | \mathcal{F}_s\right] = 0.\end{aligned}$$

Therefore

$$\begin{aligned}&\mathbf{E}\left[\left(\int_s^t X_u \mathrm{d}W_u\right)^2 | \mathcal{F}_s\right] \\ &= \mathbf{E}\left[\xi_k^2 (W_{t_{k+1}} - W_s)^2 + \sum_{i=k+1}^{m-1} \xi_i^2 (W_{t_{i+1}} - W_{t_i})^2 + \xi_m^2 (W_t - W_{t_m})^2 | \mathcal{F}_s\right].\end{aligned}$$

By the tower rule again

$$\begin{aligned}\mathbf{E}[\xi_i^2(W_{t_{i+1}}-W_{t_i})^2|\mathcal{F}_s] &= \mathbf{E}\left[\mathbf{E}[\xi_i^2(W_{t_{i+1}}-W_{t_i})^2|\mathcal{F}_{t_i}]\mathcal{F}_s\right] \\ &= \mathbf{E}[\xi_i^2(t_{i+1}-t_i)|\mathcal{F}_s] \\ &= \mathbf{E}\left[\int_{t_i}^{t_{i+1}} X_u^2 \mathrm{d}u|\mathcal{F}_s\right].\end{aligned}$$

Summing we obtain the result. □

6.2 Extending the definition

The idea is the following. We defined the integral for simple processes. Adapted processes can be approximated by simple processes, so we can define the integral of adapted process as a limit and hope for the best. This was the method at the definition of both Riemann and Lebesgue integral.

Let

$$\mathcal{H} = \left\{(X_t)\ : \mathcal{F}_t\text{-adapted and } \mathbf{E}\int_0^T X_u^2 \mathrm{d}u < \infty\right\}.$$

We extend the definition to the class $\mathcal{H}$.

Lemma 5. *Let $(X_t) \in \mathcal{H}$. There exists a sequence of simple processes $\{(X_t^n)\}_n$ such that*

$$\lim_{n\to\infty} \mathbf{E}\int_0^T (X_s - X_s^n)^2 \,\mathrm{d}s = 0.$$

Proof. We only prove in the special case when X is bounded and continuous. Let

$$X_t^n(\omega) = X_0(\omega)\mathbf{I}_{\{0\}}(t) + \sum_{k=0}^{2^n-1} X_{\frac{kT}{2^n}}(\omega)\mathbf{I}_{(\frac{kT}{2^n},\frac{(k+1)T}{2^n}]}(t).$$

These are simple processes. Since continuous function is uniformly continuous on compacts, almost surely

$$\int_0^T |X_u^n - X_u|^2 \,\mathrm{d}t \to 0.$$

Lebesgue's dominated convergence gives the proof. □

Let $X \in \mathcal{H}$ and $\{X^n\}_n$ given in the lemma. By Theorem 8 (iv)

$$\mathbf{E} \sup_{t\in[0,T]} \left(\int_0^t (X_u^n - X_u^m) \mathrm{d}W_u \right)^2 \leq 4\mathbf{E} \int_0^T (X_u^n - X_u^m)^2 \mathrm{d}u. \tag{4}$$

{eq:ito-int-lim1}

The right-hand side tends to 0 by the lemma above, therefore the left-hand side too. Thus there exists a sequence $\{n_k\}$ such that

$$\mathbf{E} \sup_{t\in[0,T]} \left(\int_0^t (X_u^{n_{k+1}} - X_u^{n_k}) \mathrm{d}W_u \right)^2 \leq 2^{-k}. \tag{5}$$

{eq:unif-conv-ineq

Then by Chebyshev

$$\mathbf{P} \left(\sup_{t\in[0,T]} |I_t^{n_{k+1}} - I_t^{n_k}| > k^{-2} \right) \leq k^4 2^{-k},$$

which is summable. Therefore, the first Borel–Cantelli lemma implies that $I(X^{n_k})$ converges uniformly on $[0,T]$-n a.s. Let define the stochastic integral $I(X)$ as the limit

$$I_t(X) := \lim_{k\to\infty} I_t(X^{n_k}).$$

As $I(X^{n_k})$ is continuous, so is $I(X)$. We have to show that $I(X)$ does not depend on the subsequence. In (4) letting $m \to \infty$

$$\mathbf{E} \sup_{t\in[0,T]} (I_t(X) - I_t(X^n))^2 \leq 4\mathbf{E} \int_0^T (X_u - X_u^n)^2 \mathrm{d}u,$$

so $I(X)$ does not depend on the subsequence.

Next we show that $I(X)$ is martingale, i.e. for any $s < t$

$$\mathbf{E}[I_t(X)|\mathcal{F}_s] = I_s(X).$$

For any n

$$\begin{aligned} \|\mathbf{E}[I_t(X)|\mathcal{F}_s] - I_s(X)\|_{L^2} &\leq \|\mathbf{E}[I_t(X) - I_t(X^n)|\mathcal{F}_s]\|_{L^2} \\ &+ \|\mathbf{E}[I_t(X^n) - I_s(X^n)|\mathcal{F}_s]\|_{L^2} + \|I_s(X^n) - I_s(X)\|_{L^2}, \end{aligned}$$

where $\|X\|_{L^2} = \sqrt{\mathbf{E}X^2}$. The second term on the RHS equals 0, since $I(X^n)$ is martingale, while the first and third term can be arbitrarily small. So $I(X)$ is indeed a martingale.

Summarizing, for $X \in \mathcal{H}$ we defined the stochastic integral

$$I_t(X) = \int_0^t X_u \mathrm{d}W_u$$

and showed that it satisfies the properties of Theorem 8.

We note that the definition of the integral can be further extended from $\mathcal{H}$ to the larger class

$$\mathcal{H}' = \left\{ (X_t) \ : \mathcal{F}_t\text{-adapted and } \int_0^T X_u^2 \, \mathrm{d}u < \infty \text{ a.s.} \right\}$$

such that Theorem 8 remains true.

{example:W-appr}

Example 2 (Approximation of $\int_0^t W_s \mathrm{d}W_s$)**.** Fix $\varepsilon \in [0,1]$ and consider

$$S_\varepsilon(\Pi) = \sum_{i=0}^{n-1} \left(\varepsilon W_{t_{i+1}} + (1-\varepsilon) W_{t_i} \right) \left(W_{t_{i+1}} - W_{t_i} \right).$$

We prove that

$$\lim_{\|\Pi\| \to 0} S_\varepsilon(\Pi) \stackrel{L^2}{=} \frac{1}{2} W_t^2 + \left(\varepsilon - \frac{1}{2} \right) t. \tag{6}$$ {eq:W-int}

We know that $(W_t^2 - t)$ is martingale, thus the limit above is martingale iff $\varepsilon = 0$, which corresponds to the definition of Itô stochastic integral. There are other stochastic integrals: $\varepsilon = 1/2$ corresponds to the *Fisk–Stratonovich integral*, and $\varepsilon = 1$ corresponds to the *backward Itô integral.*

By (6)

$$\int_0^t W_s \mathrm{d}W_s = \frac{W_t^2 - t}{2}.$$

Next we prove (6). Since

$$\varepsilon W_{t_{i+1}} + (1-\varepsilon) W_{t_i} = \frac{W_{t_{i+1}} + W_{t_i}}{2} + \left(\varepsilon - \frac{1}{2} \right) \left(W_{t_{i+1}} - W_{t_i} \right),$$

we have to determine the limits

$$\sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2, \quad \sum_{i=0}^{n-1} (W_{t_{i+1}}^2 - W_{t_i}^2).$$

The first is exactly the quadratic variation of SBM, therefore converges to t in L^2, while the second is a telescopic sum, giving W_t^2.

{example:exp}

Example 3. Let X be simple process and W SBM. Let

$$\zeta_t^s(X) = \int_s^t X_u \mathrm{d}W_u - \frac{1}{2}\int_s^t X_u^2 \mathrm{d}u, \quad \zeta_t = \zeta_t^0.$$

We show that $(Y_t = e^{\zeta_t})$ is martingale.

Since X is simple, we have

$$X_t = \xi_0 \mathbf{I}_{\{0\}}(t) + \sum_{i=0}^{n-1} \xi_i \mathbf{I}_{(t_i, t_{i+1}]}(t),$$

where ξ_i is $\mathcal{F}_{t_i}$-measurable. Thus if $s \in (t_k, t_{k+1}]$, $t \in (t_m, t_{m+1}]$, then

$$\begin{aligned}\zeta_t^s = & \xi_k (W_{t_{k+1}} - W_s) - \frac{\xi_k^2}{2}(t_{k+1} - s) + \sum_{i=k+1}^{m-1} \left[\xi_i (W_{t_{i+1}} - W_{t_i}) - \frac{\xi_i^2}{2}(t_{i+1} - t_i) \right] \\ & + \xi_m (W_t - W_{t_m}) - \frac{\xi_m^2}{2}(t - t_m).\end{aligned} \tag{7}$$

{eq:zeta-felbontas

Since ζ_s is $\mathcal{F}_s$-measurable we obtain

$$\mathbf{E}[e^{\zeta_t} | \mathcal{F}_s] = e^{\zeta_s} \mathbf{E}[e^{\zeta_t^s} | \mathcal{F}_s].$$

We only have to show that

$$\mathbf{E}[e^{\zeta_t^s} | \mathcal{F}_s] = 1.$$

This can be done by a repeated application of the tower rule. In (7) all terms but the last are $\mathcal{F}_{t_m}$-measurable and

$$\begin{aligned}&\mathbf{E}\left[\exp\left\{ \xi_m (W_t - W_{t_m}) - \frac{\xi_m^2}{2}(t - t_m) \right\} | \mathcal{F}_{t_m} \right] \\ &= e^{-\frac{\xi_m^2}{2}(t - t_m)} \mathbf{E}\left[\exp\{ \xi_m (W_t - W_{t_m}) \} | \mathcal{F}_{t_m} \right].\end{aligned}$$

In the exponent of the RHS ξ_m is $\mathcal{F}_{t_m}$-measurable and $W_t - W_{t_m}$ is independent of $\mathcal{F}_{t_m}$, therefore (by the next exercise) ξ_m can be handled as a constant. We have

$$\mathbf{E} e^{\lambda Z} = e^{\frac{\lambda^2}{2}},$$

therefore

$$\mathbf{E}\left[\exp\{\xi_m(W_t - W_{t_m})\}|\mathcal{F}_{t_m}\right] = e^{\frac{\xi_m^2}{2}(t-t_m)}.$$

Summarizing

$$\mathbf{E}\left[\exp\{\xi_m(W_t - W_{t_m}) - \frac{\xi_m^2}{2}(t-t_m)\}|\mathcal{F}_{t_m}\right] = 1.$$

Applying repeatedly the tower rule first to the σ-algebra $\mathcal{F}_{t_{m-1}}$, then to $\mathcal{F}_{t_{m-2}}$, ..., we obtain that each factor equals 1.

Using the Itô formula we show that Y is martingale for more general processes and it satisfies a certain stochastic differential equation.

Exercise 5. Let X, Y be random variables, X is $\mathcal{G}$-measurable, and Y is independent of $\mathcal{G}$. Then

$$\mathbf{E}[h(X,Y)|\mathcal{G}] = \int h(X,y)\mathrm{d}F(y),$$

where $F(y) = \mathbf{P}(Y \le y)$ is the distribution function of Y.

6.3 Itô's formula

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, $(\mathcal{F}_t)$ a filtration, and (W_t) SBM for this filtration. Then (X_t) *is Itô process* if

$$X_t = X_0 + \int_0^t K_s \mathrm{d}s + \int_0^t H_s \mathrm{d}W_s, \tag{8}$$ {eq:ito-proc}

where

- X_0 $\mathcal{F}_0$-measurable;
- K, H are $\mathcal{F}_t$-adapted processes;
- $\int_0^T |K_u|\mathrm{d}u < \infty$, $\int_0^T H_s^2 \mathrm{d}s < \infty$ a.s.

The part $\int_0^t K_s \mathrm{d}s$ is the bounded variation part of the process, while $\int_0^t H_s \mathrm{d}W_s$ is the martingale part.

{lemma:korlatosval

Lemma 6. *If $M_t = \int_0^t K_s \mathrm{d}s$ is a continuous martingale and $\int_0^T |K_s|\mathrm{d}s < \infty$ almost surely then $M_t \equiv 0$.*

Proof. Assume that $\int_0^T |K_s| \mathrm{d}s \leq C$ for some $C < \infty$. Then for a sequence of partitions $(\Pi_n = \{0 = t_0 < t_1 < \ldots < t_n = T\})$ of $[0, T]$

$$\begin{aligned}
\mathbf{E} \sum_{i=0}^{n-1} (M_{t_{i+1}} - M_{t_i})^2 &\leq \mathbf{E} \sup_{0 \leq i \leq n-1} |M_{t_{i+1}} - M_{t_i}| \int_0^T |K_s| \mathrm{d}s \\
&\leq C \mathbf{E} \sup_{0 \leq i \leq n-1} |M_{t_{i+1}} - M_{t_i}| \to 0,
\end{aligned}$$

as $\|\Pi_n\| \to 0$. We used that continuous function is uniformly continuous on compacts and Lebesgue's dominated convergence can be used because of the boundedness.

Furthermore,

$$\begin{aligned}
\mathbf{E}(M_t - M_s)^2 &= \mathbf{E} M_t^2 + \mathbf{E} M_s^2 - 2\mathbf{E} \left(\mathbf{E}[M_t M_s | \mathcal{F}_s] \right) \\
&= \mathbf{E} M_t^2 - \mathbf{E} M_s^2,
\end{aligned}$$

for $s < t$, thus

$$\mathbf{E} \sum_{i=0}^{n-1} (M_{t_{i+1}} - M_{t_i})^2 = \mathbf{E}(M_t^2 - M_0^2) = \mathbf{E} M_t^2.$$

Therefore $\mathbf{E} M_t^2 = 0$ for all t, and the statement follows. □

Corollary 1. *Representation* (8) *is unique.*

Proof. Indeed, if

$$\int_0^t K_s \mathrm{d}s + \int_0^t H_s \mathrm{d}W_s = \int_0^t L_s \mathrm{d}s + \int_0^t G_s \mathrm{d}W_s,$$

then

$$\int_0^t (K_s - L_s) \mathrm{d}s = \int_0^t (G_s - H_s) \mathrm{d}W_s.$$

The RHS is a continuous martingale, therefore by the previous lemma it has to be constant 0. □

In what follows we use the *notation*

$$\mathrm{d}X_t = K_t \mathrm{d}t + H_t \mathrm{d}W_t.$$

Theorem 9 (Itô formula (1944)). *Let $X_t = X_0 + \int_0^t K_s \mathrm{d}s + \int_0^t H_s \mathrm{d}W_s$ be an Itô process and $f \in C^2$. Then*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)H_s^2 \mathrm{d}s.$$

That is $(f(X_t))$ is an Itô process too, with representation (8)

$$f(X_t) = f(X_0) + \int_0^t \left(f'(X_s)K_s + \frac{1}{2} f''(X_s)H_s^2 \right) \mathrm{d}s + \int_0^t f'(X_s)H_s \mathrm{d}W_s.$$

Example 4. We already calculated the stochastic integral $\int W_s \mathrm{d}W_s$ in Example 2. Now we determine it again.

The SBM as an Itô process can be represented with $K_s \equiv 0$, $H_s \equiv 1$. Let $f(x) = x^2$. Then

$$W_t^2 = W_0^2 + \int_0^t 2W_s \mathrm{d}W_s + \frac{1}{2}\int_0^t 2\mathrm{d}s.$$

From this we obtain

$$\int_0^t W_s \mathrm{d}W_s = \frac{W_t^2 - t}{2}.$$

We see immediately that $W_t^2 - t$ is martingale.

Proof. We only prove under the following extra assumptions: f is compactly supported; $\sup_{s,\omega} |K_s(\omega)| < K$, $\sup_{s,\omega} |H_s(\omega)| < K$ for some $K < \infty$. (This is not an essential restriction.)

Take $\Pi = \{0 = t_0 < t_1 < \ldots < t_m = T\}$. Using the Taylor formula

$$\begin{aligned}
f(X_t) - f(X_0) &= \sum_{k=1}^m \left[f(X_{t_k}) - f(X_{t_{k-1}}) \right] \\
&= \sum_{k=1}^m f'(X_{t_{k-1}})(X_{t_k} - X_{t_{k-1}}) + \frac{1}{2}\sum_{k=1}^m f''(\eta_k)(X_{t_k} - X_{t_{k-1}})^2 \\
&= \sum_{k=1}^m f'(X_{t_{k-1}}) \int_{t_{k-1}}^{t_k} K_s \mathrm{d}s + \sum_{k=1}^m f'(X_{t_{k-1}}) \int_{t_{k-1}}^{t_k} H_s \mathrm{d}W_s \\
&\quad + \frac{1}{2}\sum_{k=1}^m f''(\eta_k)(X_{t_k} - X_{t_{k-1}})^2 \\
&= I_1 + I_2 + I_3,
\end{aligned}$$

where $\eta_k(\omega)$ is between $X_{t_{k-1}}(\omega)$ and $X_{t_k}(\omega)$.

It is easy to handle I_1. As f' and X_t are continuous

$$I_1 = \sum_{k=1}^{m} f'(X_{t_{k-1}}) \int_{t_{k-1}}^{t_k} K_s \mathrm{d}s \longrightarrow \int_0^t f'(X_s) K_s \mathrm{d}s \quad \text{a.s.}, \tag{9}$$ {eq:i1}

as $\|\Pi\| \to 0$.

Rewrite I_2 as

$$I_2 = \sum_{k=1}^{m} f'(X_{t_{k-1}}) \int_{t_{k-1}}^{t_k} H_s \mathrm{d}W_s = \int_0^t \sum_{k=1}^{m} f'(X_{t_{k-1}}) \mathbf{I}_{(t_{k-1},t_k]}(s) H_s \mathrm{d}W_s.$$

As $\|\Pi\| \to 0$

$$\mathbf{E} \int_0^t \left(f'(X_s) H_s - \sum_{k=1}^{m} f'(X_{t_{k-1}}) \mathbf{I}_{(t_{k-1},t_k]}(s) H_s \right)^2 \mathrm{d}s \to 0.$$

Indeed, for any $\omega \in \Omega$ fix the integrand is bounded and by continuity goes to 0, therefore the dominated Lebesgue convergence theorem applies. Theorem 8 (ii) implies

$$I_2 = \int_0^t \sum_{k=1}^{m} f'(X_{t_{k-1}}) I_{(t_{k-1},t_k]}(s) H_s \mathrm{d}W_s \xrightarrow{L^2} \int_0^t f'(X_s) H_s \mathrm{d}W_s. \tag{10}$$ {eq:i2-konv}

Next comes I_3, the difficult part. We have to show that

$$I_3 \to \frac{1}{2} \int_0^t f''(X_s) H_s^2 \mathrm{d}s.$$

Write

$$\begin{aligned}
(X_{t_k} - X_{t_{k-1}})^2 &= \left(\int_{t_{k-1}}^{t_k} K_s \mathrm{d}s + \int_{t_{k-1}}^{t_k} H_s \mathrm{d}W_s \right)^2 \\
&= \left(\int_{t_{k-1}}^{t_k} K_s \mathrm{d}s \right)^2 + 2 \int_{t_{k-1}}^{t_k} K_s \mathrm{d}s \cdot \int_{t_{k-1}}^{t_k} H_s \mathrm{d}W_s \\
&\quad + \left(\int_{t_{k-1}}^{t_k} H_s \mathrm{d}W_s \right)^2.
\end{aligned}$$

We show that the contribution of the first two terms is negligible to the whole sum. For the first

$$\left| \sum_{k=1}^m f''(\eta_k) \left(\int_{t_{k-1}}^{t_k} K_s \mathrm{d}s \right)^2 \right| \leq \|f''\|_\infty \cdot K^2 \sum_{k=1}^m (t_k - t_{k-1})^2 \to 0 \quad \text{a.s.} \tag{11}$$ {eq:i3-1}

To handle the second introduce $M_t = \int_0^t H_s \mathrm{d}W_s$. Then

$$\begin{aligned} &\left| \sum_{k=1}^m f''(\eta_k) \int_{t_{k-1}}^{t_k} K_s \mathrm{d}s \cdot \int_{t_{k-1}}^{t_k} H_s \mathrm{d}W_s \right| \\ &\leq \|f''\|_\infty \cdot K \sup_{1 \leq k \leq m} |M_{t_k} - M_{t_{k-1}}| \cdot \sum_{k=1}^m (t_k - t_{k-1}) \\ &= \|f''\|_\infty \cdot K\, t \sup_{1 \leq k \leq m} |M_{t_k} - M_{t_{k-1}}| \to 0, \quad \text{a.s.,} \end{aligned} \tag{12}$$ {eq:i3-2}

since $M_t = \int_0^t H_s \mathrm{d}W_s$ is a continuous martingale.

We have to deal with the sum

$$\sum_{k=1}^m f''(\eta_k) \left(\int_{t_{k-1}}^{t_k} H_s \mathrm{d}W_s \right)^2.$$

First we change η_k to $X_{t_{k-1}}$. Taking the difference

$$\begin{aligned} &\sum_{k=1}^m [f''(\eta_k) - f''(X_{t_{k-1}})](M_{t_k} - M_{t_{k-1}})^2 \\ &\leq \sup_{1 \leq k \leq m} |f''(\eta_k) - f''(X_{t_{k-1}})| \cdot \sum_{k=1}^m (M_{t_k} - M_{t_{k-1}})^2. \end{aligned}$$

By the Cauchy–Schwarz inequality

$$\begin{aligned} &\left| \mathbf{E} \sum_{k=1}^m [f''(\eta_k) - f''(X_{t_{k-1}})](M_{t_k} - M_{t_{k-1}})^2 \right| \\ &\leq \sqrt{\mathbf{E} \sup_{1 \leq k \leq m} (f''(\eta_k) - f''(X_{t_{k-1}}))^2} \sqrt{\mathbf{E} \left(\sum_{k=1}^m (M_{t_k} - M_{t_{k-1}})^2 \right)^2}. \end{aligned} \tag{13}$$ {eq:i3-3}

The first term tends to 0 because (X_t) is continuous and f'' is bounded. The second is bounded by the following lemma.

{lemma:Ito-aux}

Lemma 7. *Let* (M_t) *be a continuous bounded martingale on* $[0,t]$, *that is* $\sup_{s,\omega} |M_s(\omega)| \le K$, *and let* $\Pi = \{0 = t_0 < t_1 < \ldots < t_m = t\}$ *be a partition. Then*

$$\mathbf{E}\left(\sum_{i=1}^{m}(M_{t_i} - M_{t_{i-1}})^2\right)^2 \le 6K^4.$$

Proof. Expanding the square

$$\begin{aligned}&\mathbf{E}\left(\sum_{i=1}^{m}(M_{t_i} - M_{t_{i-1}})^2\right)^2 \\ &= \sum_{i=1}^{m}\mathbf{E}(M_{t_i} - M_{t_{i-1}})^4 + \sum_{i\neq j}\mathbf{E}(M_{t_i} - M_{t_{i-1}})^2(M_{t_j} - M_{t_{j-1}})^2.\end{aligned}$$

Using several times that

$$\mathbf{E}[(M_t - M_s)^2|\mathcal{F}_s] = \mathbf{E}[M_t^2 - M_s^2|\mathcal{F}_s], \quad s < t,$$

we obtain

$$\begin{aligned}&\sum_{i\neq j}\mathbf{E}(M_{t_i} - M_{t_{i-1}})^2(M_{t_j} - M_{t_{j-1}})^2 \\ &= 2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\mathbf{E}(M_{t_i} - M_{t_{i-1}})^2(M_{t_j} - M_{t_{j-1}})^2 \\ &= 2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\mathbf{E}\left[\mathbf{E}[(M_{t_i} - M_{t_{i-1}})^2(M_{t_j} - M_{t_{j-1}})^2|\mathcal{F}_{t_{j-1}}]\right] \\ &= 2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\mathbf{E}(M_{t_i} - M_{t_{i-1}})^2(M_{t_j}^2 - M_{t_{j-1}}^2) \\ &= 2\sum_{i=1}^{m-1}\mathbf{E}(M_{t_i} - M_{t_{i-1}})^2(M_t^2 - M_{t_i}^2) \\ &\le 2K^2\sum_{i=1}^{m-1}\mathbf{E}(M_{t_i} - M_{t_{i-1}})^2 \\ &= 2K^2\sum_{i=1}^{m-1}\mathbf{E}(M_{t_i}^2 - M_{t_{i-1}}^2) \le 2K^4.\end{aligned}$$

While, for the sum of 4th powers

$$\sum_{i=1}^{m}\mathbf{E}(M_{t_i}-M_{t_{i-1}})^4 \le 4K^2\mathbf{E}\sum_{i=1}^{m}\mathbf{E}(M_{t_i}-M_{t_{i-1}})^2$$
$$= 4K^2\mathbf{E}(M_t^2-M_0^2) \le 4K^4.$$

□

Summarizing from I_3 we have the sum

$$\sum_{k=1}^{m} f''(X_{t_{k-1}})(M_{t_k}-M_{t_{k-1}})^2.$$

We claim that

$$\sum_{k=1}^{m} f''(X_{t_{k-1}})(M_{t_k}-M_{t_{k-1}})^2 \xrightarrow{L^1} \int_0^t f''(X_s)H_s^2\mathrm{d}s. \tag{14}$$ {eq:i3-negyzetesva

Since X and f'' are continuous

$$\sum_{k=1}^{m} f''(X_{t_{k-1}})\int_{t_{k-1}}^{t_k} H_s^2\mathrm{d}s \to \int_0^t f''(X_s)H_s^2\mathrm{d}s \quad \text{a.s.}$$

Since almost sure convergence and boundedness implies L^1 convergence, and L^2 convergence implies L^1 convergence, it is enough to show that

$$\sum_{k=1}^{m} f''(X_{t_{k-1}})\left((M_{t_k}-M_{t_{k-1}})^2-\int_{t_{k-1}}^{t_k} H_s^2\mathrm{d}s\right) \xrightarrow{L^2} 0.$$

Theorem 8 (ii) implies

$$\mathbf{E}\left[(M_{t_k}-M_{t_{k-1}})^2|\mathcal{F}_{t_{k-1}}\right] = \mathbf{E}\left[\left(\int_{t_{k-1}}^{t_k} H_s\,\mathrm{d}W_s\right)^2|\mathcal{F}_{t_{k-1}}\right]$$
$$= \mathbf{E}\left[\int_{t_{k-1}}^{t_k} H_s^2\,\mathrm{d}s|\mathcal{F}_{t_{k-1}}\right],$$

so in

$$\mathbf{E}\left(\sum_{k=1}^{m} f''(X_{t_{k-1}})\left((M_{t_k}-M_{t_{k-1}})^2-\int_{t_{k-1}}^{t_k} H_s^2\mathrm{d}s\right)\right)^2$$

the expectation of the mixed term is 0. Thus this equals

$$
\begin{aligned}
&= \mathbf{E} \sum_{k=1}^{m} f''(X_{t_{k-1}})^2 \left((M_{t_k} - M_{t_{k-1}})^2 - \int_{t_{k-1}}^{t_k} H_s^2 \mathrm{d}s \right)^2 \\
&\leq \|f''\|_\infty^2 \Bigg[\mathbf{E} \sum_{k=1}^{m} (M_{t_k} - M_{t_{k-1}})^4 + 2\mathbf{E} \sum_{k=1}^{m} (M_{t_k} - M_{t_{k-1}})^2 \int_{t_{k-1}}^{t_k} H_s^2 \mathrm{d}s \\
&\qquad\qquad + \mathbf{E} \sum_{k=1}^{m} \left(\int_{t_{k-1}}^{t_k} H_s^2 \mathrm{d}s \right)^2 \Bigg] \\
&\leq \|f''\|_\infty^2 \left[\mathbf{E} \sum_{k=1}^{m} (M_{t_k} - M_{t_{k-1}})^4 + 2K^2 t \mathbf{E} \sup_{1 \leq k \leq m} (M_{t_k} - M_{t_{k-1}})^2 + K^4 t \|\Pi\| \right].
\end{aligned}
$$

The second and third term tend to 0, and for the first

$$
\begin{aligned}
\mathbf{E} \sum_{k=1}^{m} (M_{t_k} - M_{t_{k-1}})^4 &\leq \mathbf{E} \left[\sum_{k=1}^{m} (M_{t_k} - M_{t_{k-1}})^2 \cdot \sup_{1 \leq k \leq m} |M_{t_k} - M_{t_{k-1}}|^2 \right] \\
&\leq \sqrt{\mathbf{E} \left[\sum_{k=1}^{m} (M_{t_k} - M_{t_{k-1}})^2 \right]^2} \sqrt{\mathbf{E} \sup_{1 \leq k \leq m} |M_{t_k} - M_{t_{k-1}}|^4} \\
&\leq \sqrt{6} K^2 \sqrt{\mathbf{E} \sup_{1 \leq k \leq m} |M_{t_k} - M_{t_{k-1}}|^4} \to 0.
\end{aligned}
$$

Summarizing we obtained L^1, L^2 and almost sure convergence in (9)–(14). Since everything is bounded, L^1 convergence follows in each case, that is

$$
\begin{aligned}
f(X_t) - f(X_0) &= \sum_{k=1}^{m} [f(X_{t_k}) - f(X_{t_{k-1}})] \\
&\xrightarrow{L^1} \int_0^t f'(X_s) \mathrm{d}X_s + \frac{1}{2} \int_0^t f''(X_s) H_s^2 \mathrm{d}s.
\end{aligned}
$$

Convergence in L^1 implies a.s. convergence on a subsequence. As both sides are continuous we obtained that the two process are indistinguishable. □

{example:exp-2}

Example 5 (Continuation of Example 3)**.** Let

$$\zeta_t^s = \int_s^t X_u \mathrm{d}W_u - \frac{1}{2}\int_s^t X_u^2 \mathrm{d}u, \quad \zeta_t = \zeta_t^0,$$

where X_t is an adapted process. Then $Z_t = e^{\zeta_t}$ satisfies the stochastic differential equation

$$Z_t = 1 + \int_0^t Z_s X_s \mathrm{d}W_s,$$

or with a common notation

$$\mathrm{d}Z_t = Z_t X_t \mathrm{d}W_t, \quad Z_0 = 1.$$

Writing ζ as an Itô process

$$\zeta_t = \int_0^t -\frac{1}{2}X_u^2 \mathrm{d}u + \int_0^t X_u \mathrm{d}W_u.$$

Using Itô's formula with $f(x) = e^x$

$$\begin{aligned}
Z_t = e^{\zeta_t} &= 1 + \int_0^t e^{\zeta_s}\mathrm{d}\zeta_s + \frac{1}{2}\int_0^t e^{\zeta_s} X_s^2 \mathrm{d}s \\
&= 1 + \int_0^t e^{\zeta_s}\left(-\frac{1}{2}X_s^2 \mathrm{d}s + X_s \mathrm{d}W_s\right) + \frac{1}{2}\int_0^t e^{\zeta_s} X_s^2 \mathrm{d}s \\
&= 1 + \int_0^t e^{\zeta_s} X_s \mathrm{d}W_s \\
&= 1 + \int_0^t Z_s X_s \mathrm{d}W_s,
\end{aligned}$$

as claimed. We see that Z_t is martingale.

Exercise 6. Let ζ_t be as above. Show that $Y_t = e^{-\zeta_t}$ satisfies the SDE

$$\mathrm{d}Y_t = Y_t X_t^2 \mathrm{d}t - X_t Y_t \mathrm{d}W_t, \quad Y_0 = 1.$$

Similarly, one can show a more general version, where f depends on the time variable t.

Theorem 10 (More general Itô formula)**.** *Let X_t be an Itô process and $f \in C^{1,2}$. Then*

$$\begin{aligned}
f(t, X_t) = f(0, X_0) &+ \int_0^t \frac{\partial}{\partial s} f(s, X_s)\mathrm{d}s + \int_0^t \frac{\partial}{\partial x} f(s, X_s)\mathrm{d}X_s \\
&+ \frac{1}{2}\int_0^t \frac{\partial^2}{\partial x^2} f(s, X_s) H_s^2 \mathrm{d}s.
\end{aligned}$$

6.4 Multidimensional Itô processes

Let $W = (W^1, W^2, \dots, W^r)$ be an r-dimensional SBM, that is its component are iid SBM's. Then (X_t) is a *d-dimensional Itô process*, if

$$X_t^i = X_0^i + \int_0^t K_s^i \mathrm{d}s + \sum_{j=1}^r \int_0^t H_s^{i,j} \mathrm{d}W_s^j, \qquad (15)$$ {eq:multid-ito}

where $\int_0^T |K_s^i|\mathrm{d}s < \infty$, $\int_0^T (H_s^{i,j})^2 \mathrm{d}s < \infty$ a.s., and $K^i, H^{i,j}$ are $\mathcal{F}_t$-adapted, $i = 1, 2, \dots, d$, $j = 1, 2, \dots, r$.

Theorem 11 (Multidimensional Itô formula)**.** *Let (X_t) be a multidimensional Itô process and $f : \mathbb{R}^{1+d} \to \mathbb{R}$, $f \in C^{1,2}$. Then*

$$\begin{aligned} f(t, X_t^1, \dots, X_t^d) = f(0, X_0^1, \dots, X_0^d) &+ \int_0^t \frac{\partial}{\partial s} f(s, X_s^1, \dots, X_s^d)\, \mathrm{d}s \\ &+ \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} f(s, X_s^1, \dots, X_s^d)\, \mathrm{d}X_s^i \\ &+ \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} f(s, X_s^1, \dots, X_s^d) \sum_{k=1}^r H_s^{i,k} H_s^{j,k}\, \mathrm{d}s. \end{aligned}$$

6.5 Applications

Example 6 (Integration by parts I)**.** Let (X, Y) be a two-dimensional Itô process with representation

$$\begin{aligned} X_t &= X_0 + \int_0^t K_s\, \mathrm{d}s + \int_0^t H_s\, \mathrm{d}W_s \\ Y_t &= Y_0 + \int_0^t L_s\, \mathrm{d}s + \int_0^t G_s\, \mathrm{d}W_s, \end{aligned}$$

where K, L, H, G are as usual. Then

$$\int_0^t X_s \mathrm{d}Y_s = X_t Y_t - X_0 Y_0 - \int_0^t Y_s \mathrm{d}X_s - \int_0^t H_s G_s \mathrm{d}s.$$

Note that in the deterministic integration by parts formula the last term is missing.

For the proof apply Itô's formula for (X,Y) and $f(x,y)=xy$. Then

$$r=1,\ d=2,\ K_s^1=K_s,\ K_s^2=L_s,\ H_s^{1,1}=H_s,\ H_s^{2,1}=G_s.$$

Since $\frac{\partial f}{\partial x}=y$, $\frac{\partial f}{\partial y}=x$, $\frac{\partial^2 f}{\partial^2 x}=\frac{\partial^2 f}{\partial^2 y}=0$, and $\frac{\partial^2 f}{\partial x \partial y}=\frac{\partial^2 f}{\partial y \partial x}=1$, we obtain

$$X_tY_t=X_0Y_0+\int_0^t Y_s \mathrm{d}X_s+\int_0^t X_s \mathrm{d}Y_s+\frac{1}{2}2\int_0^t H_sG_s\mathrm{d}s,$$

as claimed.

Example 7 (Integration by parts II)**.** To change a bit let $\widetilde{W}$ be another SBM independent of W and (X,Y)

$$X_t=X_0+\int_0^t K_s\,\mathrm{d}s+\int_0^t H_s\,\mathrm{d}W_s$$
$$Y_t=Y_0+\int_0^t L_s\,\mathrm{d}s+\int_0^t G_s\,\mathrm{d}\widetilde{W}_s.$$

Then

$$\int_0^t X_s\mathrm{d}Y_s=X_tY_t-X_0Y_0-\int_0^t Y_s\mathrm{d}X_s.$$

The proof is the same but here $d=r=2$, and no extra term appears.

{pelda:exp-BM}

Example 8 (Geometric Brownian motion)**.** Let $\mu\in\mathbb{R}$, $\sigma>0$. Solve the SDE

$$\mathrm{d}X_t=\mu X_t\mathrm{d}t+\sigma X_t\mathrm{d}W_t. \tag{16}$$ {eq:exp-BM-sde}

We have

$$X_t=X_0+\int_0^t \mu X_s\mathrm{d}s+\int_0^t \sigma X_s\mathrm{d}W_s.$$

Applying Itô's formula with $f(x)=\log x$

$$\begin{aligned}\log X_t&=\log X_0+\int_0^t \frac{1}{X_s}\left(\mu X_s\mathrm{d}s+\sigma X_s\mathrm{d}W_s\right)+\frac{1}{2}\int_0^t -\frac{1}{X_s^2}\sigma^2X_s^2\mathrm{d}s\\&=\log X_0+\sigma W_t+\left(\mu-\frac{\sigma^2}{2}\right)t.\end{aligned}$$

Thus

$$X_t=X_0\cdot e^{\sigma W_t+\left(\mu-\frac{\sigma^2}{2}\right)t}. \tag{17}$$ {eq:exp-BM}

This is martingale iff $\mu=0$.

Note that $\log x$ is not defined at 0, so the proof is not complete. It only gives us a potential solution.

Exercise 7. Show that X_t in (17) is indeed a solution to the SDE (16).

A more constructive solution is to apply Itô's formula with a general f, and then choose f to obtain a simple equation. With $f(x) = \log x$ the integrand in the martingale part is constant.

Exercise 8. Show that $Y(t) = e^{t/2} \cos W_t$ is martingale.

Exercise 9. Show that

$$\int_0^t W_s^2 \mathrm{d}W_s = \frac{1}{3} W_t^3 - \int_0^t W_s \mathrm{d}s,$$

and

$$\int_0^t W_s^3 \mathrm{d}W_s = \frac{1}{4} W_t^4 - \frac{3}{2} \int_0^t W_s^2 \mathrm{d}s.$$

Exercise 10. Let $\mathbf{W} = (W^1, \ldots, W^r)$ be an r-dimensional SBM, $r \geq 2$, and let

$$R_t = \sqrt{\sum_{i=1}^r (W_t^i)^2}.$$

Show that R satisfies the SDE

$$\mathrm{d}R_t = \frac{r-1}{2R_t} \mathrm{d}t + \sum_{i=1}^r \frac{W_t^i}{R_t} \mathrm{d}W_t^i.$$

This is the Bessel equation and R is the *Bessel process*.

6.6 Quadratic variation and the Doob–Meyer decomposition

We proved that

$$\mathbf{E}\left[\left(\int_s^t X_u \mathrm{d}W_u \right)^2 | \mathcal{F}_s \right] = \mathbf{E}\left[\int_s^t X_u^2 \, \mathrm{d}u | \mathcal{F}_s \right],$$

which means that the process

$$\left(\int_0^t X_u \, \mathrm{d}W_u \right)^2 - \int_0^t X_u^2 \, \mathrm{d}u \qquad (18) \quad \{\texttt{eq:doob-meyer}\}$$

is a continuous martingale. In the decomposition

$$\left(\int_0^t X_u \mathrm{d}W_u\right)^2 = \int_0^t X_u^2 \,\mathrm{d}u + \left(\int_0^t X_u \mathrm{d}W_u\right)^2 - \int_0^t X_u^2 \,\mathrm{d}u$$

the first term is an increasing process and the second term is a martingale, that is we obtained the Doob–Meyer decomposition of $I_t(X)^2$.

On the other hand, at the proof of Itô's formula we showed (see (14)) that

$$\sum_{i=1}^n \left(\int_{t_{i-1}}^{t_i} X_u \mathrm{d}W_u\right)^2 \xrightarrow{L^1} \int_0^t X_u^2 \mathrm{d}u, \quad \text{as} \quad \|\Pi_n\| \to 0.$$

The left-hand side is exactly the *quadratic variation process* of the martingale $I_t(X)$.

Summarizing, we proved the following.

{thm:quad-DM}

Theorem 12. *For any Itô process X_t, the quadratic variation of $I_t(X)$ and the increasing process in the Doob–Meyer decomposition of $I_t(X)^2$ are the same.*

This result holds in a more general setup.

Let (X_t) be a (continuous) square integrable martingale, $X \in \mathcal{M}_2$ (or $X \in \mathcal{M}_2^c$). Then X_t^2 is a submartingale, so by the Doob–Meyer decomposition there exists a unique (up to indistinguishability) adapted increasing process A_t, such that $A_0 = 0$ a.s. and $X_t^2 - A_t$ is a martingale. The process $\langle X \rangle_t = A_t$ is the *quadratic variation of* X.

With this notation, Theorem 12 states that

$$\left\langle \int_0^\cdot X_u \mathrm{d}W_u \right\rangle_t = \langle I(X) \rangle_t = \int_0^t X_u^2 \,\mathrm{d}u.$$

Without proof we mention that Theorem 12 holds not only for Itô processes but for *continuous square integrable martingales.*

Theorem 13. *Let $X \in \mathcal{M}_2^c$. For partition Π of $[0,t]$ we have*

$$V_t^{(2)}(\Pi) := \sum_{k=1}^n (X_{t_k} - X_{t_{k-1}})^2 \xrightarrow{\mathbf{P}} \langle X \rangle_t \quad \text{as} \quad \|\Pi\| \to 0.$$

For square integrable martingales X, Y the *crossvariation process of X and Y* is

$$\langle X, Y \rangle_t = \frac{1}{4} \left(\langle X + Y \rangle_t - \langle X - Y \rangle_t \right).$$

The processes X and Y are *orthogonal* if $\langle X, Y \rangle_t = 0$ a.s. for any t.

Exercise 11. Show that if $X, Y \in \mathcal{M}_2$, then $XY - \langle X, Y \rangle$ is a martingale.

One can define stochastic integral with respect to more general processes. The process (X_t) is a continuous *semimartingale* if

$$X_t = M_t + A_t,$$

where M_t is a continuous martingale and A_t is of bounded variation, and both are adapted. As in Lemma 6 it can be shown that this decomposition is essentially unique.

We can define stochastic integral with respect to semimartingales. Indeed, integral with respect to A_t can be defined pathwise, since A is of bounded variation, and integration with respect to continuous M_t can be defined similarly as for SBM.

The following version of Itô's formula holds.

Theorem 14 (Itô formula for semimartingales)**.** *Let $X_t = M_t + A_t$ be a continuous semimartingale, and let $f \in C^2$. Then*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) \mathrm{d}X_s + \frac{1}{2} \int_0^t f''(X_s) \mathrm{d}\langle M \rangle_s.$$

7 Stochastic differential equations

7.1 Existence and uniqueness

We define the strong solution of SDEs and obtain existence and uniqueness results.

The followings are given:

- probability space $(\Omega, \mathcal{A}, \mathbf{P})$;
- with a filtration $(\mathcal{F}_t)_{t \in [0,T]}$;
- a d-dimensional SBM $W_t = (W_t^1, \ldots, W_t^r)$ with respect to the filtration $(\mathcal{F}_t)$;
- measurable functions $f : \mathbb{R}^d \times [0,T] \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{d \times r}$;
- $\mathcal{F}_0$-measurable rv $\xi : \Omega \to \mathbb{R}^d$.

The (d-dimensional) process (X_t) is *strong solution to the SDE*

$$\begin{aligned} \mathrm{d}X_t &= f(X_t,t)\,\mathrm{d}t + \sigma(X_t,t)\,\mathrm{d}W_t, \\ X_0 &= \xi, \end{aligned} \tag{19}$$

{eq:sde}

if $\int_0^t f(X_s,s)\mathrm{d}s$ are $\int_0^t \sigma(X_s,s)\mathrm{d}W_s$ well-defined for all $t \in [0,T]$ and the integral version of (19) holds, i.e.

$$X_t = \xi + \int_0^t f(X_s,s)\,\mathrm{d}s + \int_0^t \sigma(X_s,s)\,\mathrm{d}W_s, \quad \text{for all } t \in [0,T] \text{ a.s.}$$

Written coordinatewise

$$X_t^i = \xi^i + \int_0^t f^i(X_s,s)\,\mathrm{d}s + \int_0^t \sum_{j=1}^r \sigma_{i,j}(X_s,s)\,\mathrm{d}W_s^j, \quad i = 1,2,\ldots,d.$$

It is important to emphasize that with strong solutions not only the SDE (19) is given, but the driving SBM, the initial condition (not just distribution!) ξ and the filtration.

For d-dimensional vectors $|x| = \sqrt{x_1^2 + \ldots + x_d^2}$ stands for the usual Euclidean norm, and for a matrix $\sigma \in \mathbb{R}^{d\times r}$, define $|\sigma| = \sqrt{\sum_{i,j} \sigma_{ij}^2}$,

{thm:sde-exuni}

Theorem 15. *Assume that for the functions in* (19) *the following hold:*

$$\begin{aligned} &|f(x,t) - f(y,t)| + |\sigma(x,t) - \sigma(y,t)| \le K|x-y|, \\ &|f(x,t)|^2 + |\sigma(x,t)|^2 \le K_0(1+|x|^2), \\ &\mathbf{E}|\xi|^2 < \infty. \end{aligned}$$

Then (19) *has a unique strong solution X, and*

$$\mathbf{E} \sup_{0\le t\le T} |X_t|^2 \le C(1 + \mathbf{E}|\xi|^2).$$

Proof. We only prove for $d = r = 1$. The general case is similar, but notationally messy.

Recall the following statement from the theory of ordinary differential equations.

Lemma 8 (Gronwall–Bellman)**.** *Let α, β be integrable functions for which*

$$\alpha(t) \le \beta(t) + H \int_a^t \alpha(s)\,\mathrm{d}s, \quad t \in [a,b],$$

for some $H \ge 0$. Then

$$\alpha(t) \le \beta(t) + H \int_a^t e^{H(t-s)} \beta(s)\,\mathrm{d}s.$$

Uniqueness. Let X_t, Y_t be solutions. Then

$$X_t - Y_t = \int_0^t (f(X_s,s) - f(Y_s,s))\,\mathrm{d}s + \int_0^t (\sigma(X_s,s) - \sigma(Y_s,s))\,\mathrm{d}W_s.$$

Since $(a+b)^2 \le 2a^2 + 2b^2$, by Theorem 8 (ii) and the Cauchy–Schwarz inequality

$$\begin{aligned} \mathbf{E}(X_t - Y_t)^2 &\le 2\,\mathbf{E}\left(\int_0^t (f(X_s,s) - f(Y_s,s))\mathrm{d}s \right)^2 \\ &\quad + 2\,\mathbf{E} \int_0^t (\sigma(X_s,s) - \sigma(Y_s,s))^2 \mathrm{d}s \\ &\le 2(T+1)K^2 \int_0^t \mathbf{E}(X_s - Y_s)^2\,\mathrm{d}s. \end{aligned}$$

With the notation $\varphi(t) = \mathbf{E}(X_t - Y_t)^2$ we obtained

$$\varphi(t) \le 2(T+1)K^2 \int_0^t \varphi(s)\,\mathrm{d}s.$$

By the Gronwall–Bellman lemma $\varphi(t) \equiv 0$, i.e. $X_t = Y_t$ a.s. Since $X_t - Y_t$ is continuous, the two processes are indistinguishable, meaning

$$\mathbf{P}(X_t = Y_t,\ \forall t \in [0,T]) = 1.$$

Thus the uniqueness is proved.
Existence. Sketch. The proof goes similarly as the proof of the Picard–Lindelf theorem for ODEs. We do Picard iteration. Let $X_t^{(0)} \equiv \xi$, and if $X_t^{(n)}$ is given, let

$$X_t^{(n+1)} = \xi + \int_0^t f(X_s^{(n)}, s)\mathrm{d}s + \int_0^t \sigma(X_s^{(n)}, s)\mathrm{d}W_s.$$

Write

$$\begin{aligned} X_t^{(n+1)} - X_t^{(n)} &= \int_0^t \left(f(X_s^{(n)}, s) - f(X_s^{(n-1)}, s) \right) \mathrm{d}s \\ &\quad + \int_0^t \left(\sigma(X_s^{(n)}, s) - \sigma(X_s^{(n-1)}, s) \right) \mathrm{d}W_s \\ &=: B_t^{(n)} + M_t^{(n)}. \end{aligned}$$

By Doob's maximal inequality, as in the proof of uniqueness

$$\begin{aligned} \mathbf{E}\left(\sup_{s\in[0,t]} (M_s^{(n)})^2 \right) &\le 4\mathbf{E} \int_0^t \left(\sigma(X_s^{(n)}, s) - \sigma(X_s^{(n-1)}, s) \right)^2 \mathrm{d}s \\ &\le 4K^2 \int_0^t \mathbf{E}(X_s^{(n)} - X_s^{(n-1)})^2\,\mathrm{d}s. \end{aligned}$$

On the other hand, by Cauchy–Schwarz

$$\mathbf{E}\left(\sup_{s\in[0,t]} (B_s^{(n)})^2 \right) \le tK^2\,\mathbf{E} \int_0^t \left(X_s^{(n)} - X_s^{(n-1)} \right)^2 \mathrm{d}s.$$

This implies

$$\mathbf{E}\left(\sup_{s\in[0,t]} (X_s^{(n+1)} - X_s^{(n)})^2 \right) \le L \int_0^t \mathbf{E}(X_s^{(n)} - X_s^{(n-1)})^2 \mathrm{d}s,$$

with $L = 2(T+4)K^2$. Iterating and changing the order of integration

$$\begin{aligned} \mathbf{E}\left(\sup_{s\in[0,t]} (X_s^{(n+1)} - X_s^{(n)})^2 \right) &\le L \int_0^t \mathbf{E}(X_s^{(n)} - X_s^{(n-1)})^2\,\mathrm{d}s \\ &\le L^2 \int_0^t \int_0^s \mathbf{E}(X_u^{(n-1)} - X_u^{(n-2)})^2\,\mathrm{d}u\,\mathrm{d}s \\ &\le L^2 \int_0^t (t-s)\mathbf{E}(X_s^{(n-1)} - X_s^{(n-2)})^2\,\mathrm{d}s. \end{aligned}$$

Continuing, and using the assumption on ξ we obtain

$$\begin{aligned}&\mathbf{E}\left(\sup_{s\in[0,t]}(X_s^{(n+1)}-X_s^{(n)})^2\right)\\ &\leq L^n\int_0^t\frac{(t-s)^{n-1}}{(n-1)!}\mathbf{E}(X_s^1-\xi)^2\,\mathrm{d}s\leq C\frac{(LT)^n}{n!}.\end{aligned}$$

By Chebyshev

$$\sum_{n=1}^{\infty}\mathbf{P}\left(\sup_{0\leq t\leq T}|X_t^{(n+1)}-X_t^n|>n^{-2}\right)\leq\sum_{n=1}^{\infty}C'n^4\frac{(LT)^n}{n!}<\infty.$$

Therefore, applying the first Borel–Cantelli lemma the infinite sum

$$\sum_{n=0}^{\infty}(X_t^{(n+1)}-X_t^n)$$

converges a.s. Clearly the sum is a solution to the SDE (19). □

7.2 Examples

Most of the examples and exercises are from Evans [3].

Example 9. Let g be a continuous function, and consider the SDE

$$\begin{cases}\mathrm{d}X_t=g(t)X_t\mathrm{d}W_t\\ X_0=1.\end{cases}$$

Show that the unique solution is

$$X_t=\exp\left\{-\frac{1}{2}\int_0^t g(s)^2\mathrm{d}s+\int_0^t g(s)\mathrm{d}W_s\right\}.$$

The uniqueness follows from Theorem 15, assuming g is nice enough. To check that X_t is indeed a solution, we use Itô's formula. Let

$$Y_t=-\frac{1}{2}\int_0^t g(s)^2\mathrm{d}s+\int_0^t g(s)\mathrm{d}W_s.$$

With $f(x)=e^x$, we have

$$\begin{aligned}X_t=e^{Y_t}&=1+\int_0^t e^{Y_s}\mathrm{d}Y_s+\frac{1}{2}\int_0^t e^{Y_s}g^2(s)\mathrm{d}s\\ &=1+\int_0^t X_s g(s)\mathrm{d}W_s,\end{aligned}$$

as claimed.

Exercise 12. Let f and g be continuous functions, and consider the SDE

$$\begin{cases} \mathrm{d}X_t = f(t)X_t\mathrm{d}t + g(t)X_t\mathrm{d}W_t \\ X_0 = 1. \end{cases}$$

Show that the unique solution is

$$X_t = \exp\left\{\int_0^t \left[f(s) - \frac{1}{2}g(s)^2\right]\mathrm{d}s + \int_0^t g(s)\mathrm{d}W_s\right\}.$$

Exercise 13 (Brownian bridge)**.** Show that

$$B_t = (1-t)\int_0^t \frac{1}{1-s}\mathrm{d}W_s$$

is the unique solution of the SDE

$$\begin{cases} \mathrm{d}B_t = -\frac{B_t}{1-t}\mathrm{d}t + \mathrm{d}W_t \\ B_0 = 0. \end{cases}$$

Calculate the mean and covariance function of B.

A mean zero Gaussian process B_t on $[0,1]$ is called *Brownian bridge* if its covariance function is

$$\mathbf{Cov}(B_s, B_t) = \min(s,t) - st.$$

Exercise 14. Show that if W is SBM then $B_t = W_t - tW_1$ is Brownian bridge.

Exercise 15. Solve the SDE

$$\begin{cases} \mathrm{d}X_t = -\frac{1}{2}e^{-2X_t}\mathrm{d}t + e^{-X_t}\mathrm{d}W_t \\ X(0) = 0 \end{cases}$$

and show that it explodes in a finite random time. *Hint: Look for a solution* $X_t = u(W_t)$.

Exercise 16. Solve the SDE

$$\mathrm{d}X_t = -X_t\mathrm{d}t + e^{-t}\mathrm{d}W_t.$$

Exercise 17. Show that $(X_t, Y_t) = (\cos W_t, \sin W_t)$ is a solution to the SDE

$$\begin{cases} \mathrm{d}X_t = -\frac{1}{2}X_t\mathrm{d}t - Y_t\mathrm{d}W_t \\ \mathrm{d}Y_t = -\frac{1}{2}Y_t\mathrm{d}t + X_t\mathrm{d}W_t. \end{cases}$$

Show that $\sqrt{X_t^2 + Y_t^2}$ is a constant for any solution (X, Y)!

Exercise 18. Solve the SDE

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}t + \mathrm{d}W_t^{(1)} \\ \mathrm{d}Y_t = X_t\mathrm{d}W_t^{(2)}, \end{cases}$$

where $W^{(1)}$ and $W^{(2)}$ are independent SBMs.

Exercise 19. Solve the SDE

$$\begin{cases} \mathrm{d}X_t = Y_t\mathrm{d}t + \mathrm{d}W_t^{(1)} \\ \mathrm{d}Y_t = X_t\mathrm{d}t + \mathrm{d}W_t^{(2)}, \end{cases}$$

where $W^{(1)}$ and $W^{(2)}$ are independent SBMs.

7.3 Lévy characterization

One can define stochastic integral with respect to more general processes. The process (X_t) is a continuous *semimartingale* if

$$X_t = M_t + A_t,$$

where M_t is a continuous martingale and A_t is of bounded variation, and both are adapted.

We can define stochastic integral with respect to semimartingales. Indeed, integral with respect to A_t can be defined pathwise, since A is of bounded variation, and integration with respect to continuous M_t can be defined similarly as for SBM.

We recall Itô's formula.

Theorem 16 (Itô formula for semimartingales)**.** *Let $X_t = M_t + A_t$ be a continuous semimartingale, and let $f \in C^2$. Then*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)\mathrm{d}\langle M\rangle_s.$$

We have seen that if W_t is SBM, then it W_t is a continuous martingale, and $W_t^2 - t$ is a martingale. It turns out that this characterizes SBM.

Theorem 17 (Lévy's characterization of SBM)**.** *Let M_t be a continuous martingale, such that $M_0 = 0$, and $M_t^2 - t$ is martingale. Then M_t is SBM.*

Proof. We determine the conditional characteristic function of M_t with respect to $\mathcal{F}_s$, $t > s$. Apply Itô with $f(x) = e^{\mathrm{i}ux}$, where $u \in \mathbb{R}$ is arbitrary but fixed. Since $f'(x) = \mathrm{i}ue^{\mathrm{i}ux}$, $f''(x) = -u^2 e^{\mathrm{i}ux}$, and by assumption $\langle M \rangle_t = t$, therefore

$$e^{\mathrm{i}uM_t} - e^{\mathrm{i}uM_s} = \int_s^t \mathrm{i}ue^{\mathrm{i}uM_v} \mathrm{d}M_v + \frac{1}{2} \int_s^t (-u^2) e^{\mathrm{i}uM_v} \mathrm{d}v.$$

Let $A \in \mathcal{F}_s$ arbitrary. Multiplying by $e^{-\mathrm{i}uM_s}$, and integrating on A we get

$$\mathbf{E}\left[e^{\mathrm{i}u(M_t - M_s)} I_A\right] = \mathbf{P}(A) - \frac{u^2}{2} \int_s^t \mathbf{E}\left[e^{\mathrm{i}u(M_v - M_s)} I_A\right] \mathrm{d}v.$$

With A and s fixed, define

$$g_{A,s}(t) = g(t) = \mathbf{E}\left[e^{\mathrm{i}u(M_t - M_s)} I_A\right].$$

With this notation

$$g(t) = \mathbf{P}(A) - \frac{u^2}{2} \int_s^t g(v) \mathrm{d}v.$$

Differentiating we obtain

$$g'(t) = -\frac{u^2}{2} g(t), \quad g(s) = \mathbf{P}(A).$$

Therefore, the solution

$$g(t) = \mathbf{P}(A) \cdot e^{-\frac{u^2}{2}(t-s)}.$$

This holds for any $A \in \mathcal{F}_s$, which means that

$$\mathbf{E}\left[e^{\mathrm{i}u(M_t - M_s)} | \mathcal{F}_s\right] = e^{-\frac{u^2}{2}(t-s)}$$

for $u \in \mathbb{R}$. That is the increment $M_t - M_s$ is independent of $\mathcal{F}_s$, and it is Gaussian with mean 0 and variance $(t - s)$. Since it is continuous, it is SBM. $\square$

Note that the continuity assumption is important. Indeed, if N_t is a Poisson process with intensity 1, then both $(N_t - t)$ and $(N_t - t)^2 - t$ are martingales.

7.4 Girsanov's theorem

Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space, and $(\mathcal{F}_t)$ a filtration. Let $\mathbf{Q}$ be another probability measure on $(\Omega, \mathcal{A})$, which is absolute continuous with respect to $\mathbf{P}$, i.e. $\mathbf{Q} \ll \mathbf{P}$. Let M_∞ denote the Radon–Nikodym-derivative,

$$M_\infty = \frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}},$$

that is

$$\mathbf{Q}(A) = \int_A M_\infty \mathrm{d}\mathbf{P}.$$

In what follows, we have more (usually 2) probability measures, therefore we put in the lower index of $\mathbf{E}$ the corresponding measure. That is $\mathbf{E}_\mathbf{P} X = \int_\Omega X \mathrm{d}\mathbf{P}$, and $\mathbf{E}_\mathbf{Q} X = \int_\Omega X \mathrm{d}\mathbf{Q}$. Note that the notion of martingale does depend on the underlying measure. Therefore, we have $\mathbf{P}$-martingale, and $\mathbf{Q}$-martingale.

Define the $\mathbf{P}$-martingale

$$M_t = \mathbf{E}_\mathbf{P}[M_\infty | \mathcal{F}_t].$$

{lemma:p-q-mtg}

Lemma 9. *The adapted process (X_t) is $\mathbf{Q}$-martingale if and only if $(M_t X_t)$ is $\mathbf{P}$-martingale.*

Proof. Since

$$\mathbf{E}_\mathbf{P}[M_\infty X_t | \mathcal{F}_t] = X_t M_t,$$

for each $A \in \mathcal{F}_t$

$$\int_A X_t M_\infty \mathrm{d}\mathbf{P} = \int_A X_t M_t \mathrm{d}\mathbf{P}.$$

Therefore, if $A \in \mathcal{F}_s \subset \mathcal{F}_t$, then

$$\int_A X_t \mathrm{d}\mathbf{Q} = \int_A X_t M_\infty \mathrm{d}\mathbf{P} = \int_A X_t M_t \mathrm{d}\mathbf{P}$$
$$\int_A X_s \mathrm{d}\mathbf{Q} = \int_A X_s M_\infty \mathrm{d}\mathbf{P} = \int_A X_s M_s \mathrm{d}\mathbf{P}.$$

Then (X_t) is $\mathbf{Q}$-martingale if the left-hand sides are equal for each $A \in \mathcal{F}_s$, $s < t$, which is obviously equivalent to the equality of the right-hand sides, which means that $(M_t X_t)$ is $\mathbf{P}$-martingale. □

Let

$$\zeta_t^s = \int_s^t \theta_u \mathrm{d}W_u - \frac{1}{2}\int_s^t \theta_u^2 \mathrm{d}u, \quad \zeta_t = \zeta_t^0,$$

where θ_t is adapted. Then $Z_t = e^{\zeta_t}$ satisfies the SDE

$$Z_t = 1 + \int_0^t Z_s X_s \mathrm{d}W_s. \tag{20}$$ {eq:Gir-sde}

We use this formula in the proof of Girsanov's theorem. We can write the SDE above as

$$\mathrm{d}Z_t = Z_t X_t \mathrm{d}W_t, \quad Z_0 = 1.$$

Indeed, rewriting ζ as an Itô process

$$\zeta_t = \int_0^t -\frac{1}{2}\theta_u^2 \mathrm{d}u + \int_0^t \theta_u \mathrm{d}W_u.$$

Using the Itô formula with $f(x) = e^x$, we obtain

$$\begin{aligned} Z_t = e^{\zeta_t} &= 1 + \int_0^t e^{\zeta_s}\mathrm{d}\zeta_s + \frac{1}{2}\int_0^t e^{\zeta_s}\theta_s^2 \mathrm{d}s \\ &= 1 + \int_0^t e^{\zeta_s}\left(-\frac{1}{2}\theta_s^2 \mathrm{d}s + X_s \mathrm{d}W_s\right) + \frac{1}{2}\int_0^t e^{\zeta_s}\theta_s^2 \mathrm{d}s \\ &= 1 + \int_0^t e^{\zeta_s}\theta_s \mathrm{d}W_s \\ &= 1 + \int_0^t Z_s \theta_s \mathrm{d}W_s, \end{aligned}$$

as claimed. We also see that Z_t is a martingale.

Exercise 20. Let ζ_t as above. Prove that $Y_t = e^{-\zeta_t}$ satisfies the SDE

$$\mathrm{d}Y_t = Y_t \theta_t^2 \mathrm{d}t - \theta_t Y_t \mathrm{d}W_t, \quad Y_0 = 1.$$

{thm:Girsanov}

Theorem 18 (Girsanov's theorem)**.** *Let (θ_t) be an adapted process, such that $\int_0^T \theta_s^2 \mathrm{d}s < \infty$ a.s., and assume that*

$$\Lambda_t = \exp\left\{-\int_0^t \theta_s \mathrm{d}W_s - \frac{1}{2}\int_0^t \theta_s^2 \mathrm{d}s\right\} \tag{21}$$ {eq:Lambda}

is $\mathbf{P}$*-martingale, where* (W_t) *is* $\mathbf{P}$*-SBM. Define* $\mathbf{Q}_\theta = \mathbf{Q}$

$$\left.\frac{\mathrm{d}\mathbf{Q}_\theta}{\mathrm{d}\mathbf{P}}\right|_{\mathcal{F}_T} = \Lambda_T.$$

Then $\widetilde{W}_t = W_t + \int_0^t \theta_s \mathrm{d}s$ *is* $\mathbf{Q}$*-SBM.*

Remark 1. We have seen above that Λ_t is martingale. In fact, in general it is only local martingale, and we need integrability conditions. These technical assumptions are omitted.

Proof. First we show that $\mathbf{Q}$ is indeed a probability measure. By (20)

$$\Lambda_t = 1 - \int_0^t \Lambda_s \theta_s \mathrm{d}W_s,$$

which is martingale, so

$$\mathbf{E}_\mathbf{P} \Lambda_T = \mathbf{E}_\mathbf{P} \Lambda_0 = 1.$$

Since $\Lambda_T > 0$ we see that $\mathbf{Q}$ is probability measure.

Next we show that $\widetilde{W}$ satisfies the conditions of the Lévy characterization.

The continuity is clear, since W is SBM and $\mathbf{Q} \ll \mathbf{P}$. By Lemma 9 $(\widetilde{W}_t)$ is $\mathbf{Q}$-martingale iff $(\widetilde{W}_t \Lambda_t)$ is $\mathbf{P}$-martingale. We apply the Itô formula with $f(x,y) = xy$ and the Itô process

$$\begin{aligned}
\widetilde{W}_t &= \int_0^t \theta_s \mathrm{d}s + \int_0^t 1 \mathrm{d}W_s \\
\Lambda_t &= 1 - \int_0^t \Lambda_s \theta_s \mathrm{d}W_s.
\end{aligned}$$

Then

$$\begin{aligned}
\Lambda_t \widetilde{W}_t &= \int_0^t \widetilde{W}_s \mathrm{d}\Lambda_s + \int_0^t \Lambda_s \mathrm{d}\widetilde{W}_s + \int_0^t -\Lambda_s \theta_s \mathrm{d}s \\
&= -\int_0^t \widetilde{W}_s \Lambda_s \theta_s \mathrm{d}W_s + \int_0^t \Lambda_s \left(\theta_s \mathrm{d}s + \mathrm{d}W_s\right) - \int_0^t \Lambda_s \theta_s \mathrm{d}s \\
&= \int_0^t \Lambda_s (1 - \theta_s \widetilde{W}_s) \mathrm{d}W_s,
\end{aligned}$$

which is $\mathbf{P}$-martingale. Thus $(\widetilde{W}_t)$ is $\mathbf{Q}$-martingale.

Next we show that $(\widetilde{W}_t^2 - t)$ is $\mathbf{Q}$-martingale. Again, by Itô's formula with $f(x) = x^2$

$$\widetilde{W}_t^2 = 2\int_0^t \widetilde{W}_s \mathrm{d}\widetilde{W}_s + \frac{1}{2}\int_0^t 2\mathrm{d}t,$$

from which

$$\widetilde{W}_t^2 - t = 2\int_0^t \widetilde{W}_s \left(\theta_s \mathrm{d}s + \mathrm{d}W_s\right).$$

Therefore

$$\begin{aligned}
\Lambda_t(\widetilde{W}_t^2 - t) &= \int_0^t \Lambda_s 2\widetilde{W}_s \left(\theta_s \mathrm{d}s + \mathrm{d}W_s\right) + \int_0^t (\widetilde{W}_s^2 - s)\mathrm{d}\Lambda_s - \int_0^t \Lambda_s \theta_s 2\widetilde{W}_s \mathrm{d}s \\
&= \int_0^t \left[2\Lambda_s \widetilde{W}_s - (\widetilde{W}_s^2 - s)\Lambda_s \theta_s\right] \mathrm{d}W_s,
\end{aligned}$$

which is $\mathbf{P}$-martingale. Thus $(\widetilde{W}_t^2 - t)$ is $\mathbf{Q}$-martingale, and the proof is complete. □

Finally, we state without proof (and precise statement) the martingale representation theorem.

{thm:martingale-re

Theorem 19 (Martingale representation)**.** *Let (W_t) SBM on $(\Omega, \mathcal{A}, \mathbf{P})$, and let $(\mathcal{F}_t)$ the generated filtration, together with the $\mathbf{P}$-zero sets. If (M_t) is continuous square integrable martingale with $M_0 = 0$ a.s., then there exists an adapted (Y_t) such that*

$$M_t = \int_0^t Y_s \mathrm{d}W_s.$$

8 Continuous time markets

8.1 General setup

The general notations are the same as in the discrete time setup.

In what follows, we work on the finite time horizon $[0, T]$, $T < \infty$. Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space, and $(\mathcal{F}_t)$ a filtration. There are two financial instruments on the market, the bond, which is the riskless asset, and the stock, which is the risky asset. The price process of the bond is given by the deterministic process $(B_t = e^{rt})$, $r \in \mathbb{R}$ being the continuous interest rate, while the price process of the stock is (S_t), which is nonnegative, adapted to $(\mathcal{F}_t)$. Furthermore, we assume that (S_t) is an Itô process.

A *strategy / portfolio* is a process $(\pi_t = (\beta_t, \gamma_t))$, where the components are adapted and

$$\int_0^T |\beta_t| \mathrm{d}t < \infty, \quad \int_0^T \gamma_t^2 \mathrm{d}t < \infty, \text{ a.s.}$$

The process β_t represents the amount of bonds at time t, while γ_t is the amount of stock. Both processes are real valued (short selling is possible).

The value of the portfolio (π) at t is

$$X_t^\pi = \beta_t B_t + \gamma_t S_t. \tag{22}$$ {eq:ertekfoly}

Recall that in discrete time an equivalent formulation of self-financing portfolio is

$$X_{n+1} - X_n = \beta_{n+1}(B_{n+1} - B_n) + \gamma_{n+1}(S_{n+1} - S_n).$$

The continuous time analogue of the above is the SDE

$$\mathrm{d}X_t^\pi = \beta_t \mathrm{d}B_t + \gamma_t \mathrm{d}S_t.$$

The strategy $(\pi_t = (\beta_t, \gamma_t))$ is *self-financing* (SF) if it satisfies the SDE

$$\mathrm{d}X_t^\pi = \beta_t \mathrm{d}B_t + \gamma_t \mathrm{d}S_t. \tag{23}$$ {eq:onfin}

In what follows all strategies are SF unless otherwise stated.

The discounted processes are defined as $(\overline{S}_t = S_t B_0 / B_t)$ and $(\overline{X}_t^\pi = X_t^\pi B_0 / B_t)$.

{all:onfin-ekv}

Proposition 4. *A strategy $(\pi_t = (\beta_t, \gamma_t))$ is SF iff*

$$\overline{X}_t^\pi = X_0^\pi + \int_0^t \gamma_s \mathrm{d}\overline{S}_s, \quad t \in [0, T].$$

Proof. Assume that π is SF. Then, by Itô's formula

$$\begin{aligned}
\mathrm{d}\overline{X}_t^\pi &= \mathrm{d}\left(e^{-rt} X_t^\pi\right) = -re^{-rt} X_t^\pi \mathrm{d}t + e^{-rt} \mathrm{d}X_t^\pi \\
&= -re^{-rt}(\beta_t e^{rt} + \gamma_t S_t)\mathrm{d}t + e^{-rt}\left(\beta_t \mathrm{d}e^{rt} + \gamma_t \mathrm{d}S_t\right) \\
&= -re^{-rt}\gamma_t S_t \mathrm{d}t + e^{-rt}\gamma_t \mathrm{d}S_t \\
&= \gamma_t \mathrm{d}\left(e^{-rt} S_t\right),
\end{aligned}$$

as claimed.

For the reverse direction, we have

$$\mathrm{d}\overline{X}_t^\pi = \gamma_t \mathrm{d}\overline{S}_t.$$

Since $X_t^\pi = \beta_t e^{rt} + \gamma_t S_t$, so

$$\mathrm{d}\overline{X}_t^\pi = -re^{-rt} X_t^\pi \mathrm{d}t + e^{-rt}\mathrm{d}X_t^\pi = -e^{-rt}\beta_t \mathrm{d}B_t - re^{-rt}\gamma_t S_t \mathrm{d}t + e^{-rt}\mathrm{d}X_t^\pi.$$

The right-hand side

$$\gamma_t \mathrm{d}\overline{S}_t = -re^{-rt}\gamma_t S_t \mathrm{d}t + \gamma_t e^{-rt}\mathrm{d}S_t.$$

The equality of the sides gives

$$\mathrm{d}X_t^\pi = \beta_t \mathrm{d}B_t + \gamma_t \mathrm{d}S_t,$$

which is the definition of SF. □

An SF strategy π is *arbitrage*, if $X_0^\pi = 0$ a.s., $X_T \geq 0$ a.s., and $\mathbf{P}(X_T^\pi > 0) > 0$. The market is *arbitrage free* if there exists no arbitrage strategy.

A probability measure $\mathbf{Q}$ is *equivalent martingale measure* (EMM) if $\mathbf{P} \sim \mathbf{Q}$ (that is $\mathbf{P} \ll \mathbf{Q}$ and $\mathbf{Q} \ll \mathbf{P}$), and $(\overline{S}_t)$ is $\mathbf{Q}$-martingale.

We have seen in the discrete time setup that the existence of EMM is equivalent to the arbitrage free property. One of the implications is rather simple in the continuous time setup. Assume that $\mathbf{Q}$ is EMM, and let π be an (SF) strategy. By Proposition 4 the discounted value process has the representation

$$\overline{X}_t^\pi = X_0^\pi + \int_0^t \gamma_s \mathrm{d}\overline{S}_s.$$

Since $(\overline{S}_t)$ is $\mathbf{Q}$-martingale, and $\overline{X}_t^\pi$ is a stochastic integral with respect to $\overline{S}$, we see that $(\overline{X}_t^\pi)$ is $\mathbf{Q}$-martingale. (Recall the discrete time analogue of this statement.) Therefore

$$\mathbf{E}_{\mathbf{Q}} \overline{X}_T^\pi = \mathbf{E}_{\mathbf{Q}} X_0^\pi.$$

Since $\mathbf{P} \sim \mathbf{Q}$, $X_0^\pi = 0$, $X_T^\pi \geq 0$ $\mathbf{P}$-a.s., implies $\mathbf{Q}$-a.s. Then $\mathbf{E}_{\mathbf{Q}} \overline{X}_T^\pi = \mathbf{E}_{\mathbf{Q}} X_0^\pi = 0$, from which $X_T^\pi \equiv 0$ $\mathbf{Q}$-a.s., and so $\mathbf{P}$-a.s.

We proved the following.

Theorem 20. *Assume that on the market* $(\Omega, \mathcal{A}, \mathbf{P}, (S_t), (B_t = e^{rt}), (\mathcal{F}_t))$ *there exists EMM. Then the market is arbitrage free.*

8.2 Black–Scholes model

In a special model we explicitly construct the EMM via Girsanov's theorem, and compute the fair price of a payoff. In particular, we prove the Nobel-prize winner Black–Scholes pricing formula, which gives the fair price of a European call option.

Fix $r > 0, \mu \in \mathbb{R}$ and $\sigma > 0$. Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space, (W_t) SBM on $[0, T]$, $T < \infty$, and $\mathcal{F}_t$ be the generated filtration. The bond and stock price in the *Black–Scholes-model* is given by

$$\begin{aligned} \mathrm{d}B_t &= rB_t \,\mathrm{d}t, \quad B_0 = 1, \\ \mathrm{d}S_t &= \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t, \quad S_0 = S_0., \end{aligned} \tag{24}$$ {eq:black-shcholes

From the form of S_t we immediately see that S_t is a martingale if and only if $\mu = 0$.

The bond price is simply $B_t = e^{rt}$.

Writing S_t as an Itô process

$$S_t = S_0 + \int_0^t \mu S_s \mathrm{d}s + \int_0^t \sigma S_s \mathrm{d}W_s.$$

Applying Itô with $f(x) = \log x$

$$\begin{aligned} \log S_t &= \log S_0 + \int_0^t \frac{1}{S_s} \left(\mu S_s \mathrm{d}s + \sigma S_s \mathrm{d}W_s \right) + \frac{1}{2} \int_0^t -\frac{1}{S_s^2} \sigma^2 S_s^2 \mathrm{d}s \\ &= \log S_0 + \sigma W_t + \left(\mu - \frac{\sigma^2}{2} \right) t. \end{aligned}$$

From which

$$S_t = S_0 \cdot e^{\sigma W_t + \left(\mu - \frac{\sigma^2}{2} \right) t}. \tag{25}$$ {eq:exp-BM2}

This is the *exponential Brownian motion.*

Note that the proof is not complete, because the logarithm is not smooth at 0. The argument above only helps to find out the solution. (A more constructive approach is to apply Itô with a general f, and then choose f to obtain a solvable equation.)

Exercise 21. Prove that (25) is indeed a solution.

8.2.1 Equivalent martingale measure and the fair price

As an application of Girsanov's theorem, we construct a new measure, such that $\overline{S}_t$ is a martingale under this measure.

By (24)

$$\mathrm{d}\overline{S}_t = \overline{S}_t\left((\mu - r)\mathrm{d}t + \sigma \mathrm{d}W_t\right) = \overline{S}_t \sigma \mathrm{d}\widetilde{W}_t^{\mu}, \tag{26}$$ {eq:tildeS}

where

$$\widetilde{W}_t^{\mu} = W_t + \frac{\mu - r}{\sigma} t. \tag{27}$$ {eq:tildeW}

Therefore, we need a measure $\mathbf{Q}$ such that the process $\widetilde{W}_t^{\mu}$ is $\mathbf{Q}$-SBM. Then, by (26) $(\overline{S}_t)$ is $\mathbf{Q}$-martingale.

Let $\theta_t \equiv \theta = \frac{\mu - r}{\sigma}$, and

$$\left.\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}}\right|_{\mathcal{F}_T} = \Lambda_T = \exp\left\{-\int_0^T \theta \mathrm{d}W_s - \frac{1}{2}\int_0^T \theta^2 \mathrm{d}s\right\} = e^{-\theta W_T - \frac{\theta^2 T}{2}}.$$

By Girsanov's theorem (Theorem 18) the shifted process $(\widetilde{W}_t^{\mu})$ is $\mathbf{Q}$-SBM, thus $(\overline{S}_t)$ is $\mathbf{Q}$-martingale. Since $\Lambda_T > 0$ a.s., $\mathbf{P} \sim \mathbf{Q}$, therefore $\mathbf{Q}$ is EMM. By (26)

$$\overline{S}_t = S_0 \cdot e^{\sigma \widetilde{W}_t^{\mu} - \frac{\sigma^2}{2} t}. \tag{28}$$ {eq:S-mu}

Next, we determine the fair price of a claim f_T, for which $\mathbf{E} f_T^2 < \infty$. Let

$$N_t = \mathbf{E}_{\mathbf{Q}}\left[e^{-rT} f_T | \mathcal{F}_t\right],\ 0 \le t \le T.$$

By the martingale representation theorem (Theorem 19) there exists an adapted process Y_t, such that

$$N_t = N_0 + \int_0^t Y_s \mathrm{d}\widetilde{W}_s^{\mu}, \tag{29}$$ {eq:N-def}

where $N_0 = \mathbf{E}_{\mathbf{Q}} e^{-rT} f_T$. Define the strategy $\pi_t = (\beta_t, \gamma_t)$ as

$$\beta_t = N_t - \frac{Y_t}{\sigma}, \quad \gamma_t = \frac{Y_t e^{rt}}{\sigma S_t}.$$

Lemma 10. *The strategy $(\pi_t = (\beta_t, \gamma_t))$ is self-financing and $\overline{X}_t^{\pi} = N_t$.*

Proof. By the definition

$$X_t^\pi = \beta_t B_t + \gamma_t S_t = \left(N_t - \frac{Y_t}{\sigma}\right) e^{rt} + \frac{Y_t}{\sigma} e^{rt} = e^{rt} N_t,$$

i.e. $\overline{X}_t^\pi = N_t$.

In order to show that π is SF, by Proposition 4 we need that $\mathrm{d}\overline{X}_t^\pi = \gamma_t \mathrm{d}\overline{S}_t$. By (29)

$$\mathrm{d}\overline{X}_t^\pi = \mathrm{d}N_t = Y_t \mathrm{d}\widetilde{W}_t^\mu,$$

while (26) gives

$$\gamma_t \mathrm{d}\overline{S}_t = \gamma_t \overline{S}_t \sigma \mathrm{d}\widetilde{W}_t^\mu = Y_t \mathrm{d}\widetilde{W}_t^\mu.$$

□

Since

$$X_T^\pi = e^{rT} N_T = e^{rT} \mathbf{E}_{\mathbf{P}_\mu}\left[e^{-rT} f_T | \mathcal{F}_T\right] = f_T,$$

π is a perfect hedge for f_T, and $X_0^\pi = N_0 = \mathbf{E}_{\mathbf{Q}} e^{-rT} f_T$. Therefore, we proved the following.

{thm:bs-price}

Theorem 21. *In the Black–Scholes model the fair price of the contingent claim f_T is*

$$C_T(f_T) = \mathbf{E}_{\mathbf{Q}} e^{-rT} f_T.$$

Furthermore, $\pi_t = (\beta_t, \gamma_t)$,

$$\beta_t = N_t - \frac{Y_t}{\sigma}, \quad \gamma_t = \frac{Y_t e^{rt}}{\sigma S_t},$$

is a perfect hedge, where $N_t = \mathbf{E}_{\mathbf{Q}}[e^{-rT} f_T | \mathcal{F}_t]$, and $N_t = N_0 + \int_0^t Y_s \mathrm{d}\widetilde{W}_s^\mu$.

8.2.2 Black–Scholes formula

The famous Black–Scholes formula gives the fair price of a European call option. In this case the payoff function is $f_T = (S_T - K)_+$, where K is the strike price. By Theorem 21, the fair price is

$$C_T(K) = \mathbf{E}_{\mathbf{Q}}\left(e^{-rT}(S_T - K)_+\right).$$

By (28)

$$S_T = S_0 e^{rT} e^{\sigma \widetilde{W}_T^\mu - \frac{\sigma^2}{2} T},$$

where $\widetilde{W}_T^\mu \sim \mathrm{N}(0,T)$ under $\mathbf{Q}$. Therefore, writing Z for a standard normal

$$\begin{aligned}
C_T(K) &= \mathbf{E}_{\mathbf{Q}}\left(e^{-rT}(S_T-K)_+\right) \\
&= \mathbf{E}_{\mathbf{Q}}\left(S_0 e^{\sigma \widetilde{W}_T^\mu - \frac{\sigma^2}{2}T} - e^{-rT}K\right)_+ \\
&= \mathbf{E}\left(S_0 e^{\sigma\sqrt{T}Z - \frac{\sigma^2}{2}T} - e^{-rT}K\right)_+ \\
&= \frac{1}{\sqrt{2\pi}}\int_\gamma^\infty \left(S_0 e^{\sigma\sqrt{T}x - \frac{\sigma^2}{2}T} - e^{-rT}K\right) e^{-\frac{x^2}{2}}\,\mathrm{d}x \qquad (30) \quad \{\texttt{eq:BS-calc}\} \\
&= S_0 \frac{1}{\sqrt{2\pi}}\int_\gamma^\infty e^{-\frac{(x-\sigma\sqrt{T})^2}{2}}\,\mathrm{d}x - e^{-rT}K(1-\Phi(\gamma)) \\
&= S_0\left(1-\Phi(\gamma-\sigma\sqrt{T})\right) - e^{-rT}K(1-\Phi(\gamma)),
\end{aligned}$$

where

$$\gamma = \frac{1}{\sigma\sqrt{T}}\left[\log\frac{K}{S_0} + \left(\frac{\sigma^2}{2}-r\right)T\right].$$

The pricing formula

$$C_T(K) = S_0\left(1-\Phi(\gamma-\sigma\sqrt{T})\right) - e^{-rT}K(1-\Phi(\gamma))$$

is the *Black–Scholes formula*, which was published by Fischer Black and Myron Scholes in 1973. The underlying theory was generalized later by Merton. In 1997, Scholes and Merton received the Nobel prize for this (Black died in 1995).

8.2.3 From CRR to Black–Scholes

Here we derive the Black–Scholes formula as the limit of the discrete CRR pricing formula. This part is based on [2], section 2.6.

Consider the continuous model on $[0,T]$. Let $r>0$ be the continuous interest rate and $\sigma>0$ the volatility. In the approximating discrete model choose, for N fixed

$$0=\tau_0<\tau_1<\ldots<\tau_N=T, \quad \tau_i=\frac{i}{N}T$$

Put $h=T/N$. The parameters of the N-step homogeneous binomial market are r_N, a_N, and b_N. The price of the bond and stock is denoted by $B_{\tau_i}^N$ and $S_{\tau_i}^N$, respectively.

Choose

$$r_N = r\frac{T}{N} = rh, \quad \log\frac{1+b_N}{1+r_N} = \sigma\sqrt{h}, \quad \log\frac{1+a_N}{1+r_N} = -\sigma\sqrt{h}. \tag{31}$$ {eq:rab-choice}

It is easy to show that this implies

$$B^N_{\tau_{\frac{tN}{T}}} = (1+r_N)^{\lfloor \frac{tN}{T} \rfloor} \to e^{rt} = B_t,$$

which in fact suggests the choice of r_N. Similar, but more complicated calculations gives that with the choice above $\mathbf{Var} S^N_{\tau_N}$ converges.

In the homogeneous binomial model the EMM was given by the upwards step probability

$$p^*_N = \frac{r_N - a_N}{b_N - a_N}.$$

Under the EMM

$$S^N_{\tau_N} = S_0(1+b_N)^{Y_N}(1+a_N)^{N-Y_N} = S_0\left(\frac{1+b_N}{1+a_N}\right)^{Y_N}(1+a_N)^N, \tag{32}$$ {eq:crrS_N}

where $Y_N \sim \text{Binomial}(N, p^*_N)$.

The CRR pricing formula gives

$$C_N(K) = \mathbf{E}^*_N \frac{(S^N_{\tau_N} - K)_+}{B^N_{\tau_N}}. \tag{33}$$ {eq:crr-ar}

By the central limit theorem (Lindeberg–Feller theorem)

$$\frac{Y_N - Np^*_N}{\sqrt{Np^*_N(1-p^*_N)}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1), \;\; N \to \infty, \tag{34}$$ {eq:Y_N-conv}

whenever $0 < \liminf_{N\to\infty} p^*_N \le \limsup_{N\to\infty} p^*_N < 1$. Simple calculation gives that $\lim_{N\to\infty} p^*_N = 1/2$, so (34) holds. Rewriting (32)

$$\begin{aligned}\left(\frac{1+b_N}{1+a_N}\right)^{Y_N}(1+a_N)^N &= \exp\left\{Y_N \log\frac{1+b_N}{1+a_N} + N\log(1+a_N)\right\} \\ &= \exp\left\{\frac{Y_N - Np^*_N}{\sqrt{Np^*_N(1-p^*_N)}}\sqrt{Np^*_N(1-p^*_N)}\log\frac{1+b_N}{1+a_N}\right. \\ &\qquad \left. + N\left(p^*_N \log\frac{1+b_N}{1+a_N} + \log(1+a_N)\right)\right\}.\end{aligned}$$

By (34) we need to determine the limits

$$\lim_{N\to\infty} \sqrt{Np_N^*(1-p_N^*)} \log \frac{1+b_N}{1+a_N}, \quad \text{and}$$
$$\lim_{N\to\infty} N\left(p_N^* \log \frac{1+b_N}{1+a_N} + \log(1+a_N)\right).$$

Taylor expansion and (31) gives

$$\begin{aligned}1+b_N = e^{\sigma\sqrt{h}}(1+r_N) &= \left(1+\sigma\sqrt{h}+\frac{\sigma^2}{2}h+O(h^{3/2})\right)(1+rh)\\ &= 1+\sigma\sqrt{h}+\left(\frac{\sigma^2}{2}+r\right)h+O(h^{3/2}),\end{aligned}$$

thus

$$b_N = \sigma\sqrt{h}+\left(\frac{\sigma^2}{2}+r\right)h+O(h^{3/2}).$$

Similarly,

$$a_N = -\sigma\sqrt{h}+\left(\frac{\sigma^2}{2}+r\right)h+O(h^{3/2}).$$

From this

$$\begin{aligned}p_N^* = \frac{r_N-a_N}{b_N-a_N} &= \frac{\sigma\sqrt{h}-\frac{\sigma^2}{2}h+O(h^{3/2})}{2\sigma\sqrt{h}+O(h^{3/2})}\\ &= \frac{1}{2+O(h)} - \frac{\sigma\sqrt{h}+O(h)}{4+O(h)}\\ &= \frac{1}{2}-\frac{\sigma}{4}\sqrt{h}+O(h).\end{aligned}$$

Substituting back, and using the second order expansion $\log(1+x) = x - x^2/2 + O(x^3)$, $x \to 0$, we obtain

$$\lim_{N\to\infty} \sqrt{Np_N^*(1-p_N^*)} \log \frac{1+b_N}{1+a_N} = \lim_{N\to\infty} \sqrt{p_N^*(1-p_N^*)}2\sigma\sqrt{T} = \sigma\sqrt{T},$$

and

$$\begin{aligned}
&\lim_{N\to\infty} N\left(p_N^* \log\frac{1+b_N}{1+a_N} + \log(1+a_N)\right) \\
&= \lim_{N\to\infty} N\left(\left[\frac{1}{2} - \frac{\sigma}{4}\sqrt{\frac{T}{N}} + O(N^{-1})\right] 2\sigma\sqrt{\frac{T}{N}} - \sigma\sqrt{\frac{T}{N}} + r\frac{T}{N} + O(N^{-3/2})\right) \\
&= \left(r - \frac{\sigma^2}{2}\right) T.
\end{aligned}$$

Substituting back into (33)

$$\begin{aligned}
\lim_{N\to\infty} C_N(K) &= e^{-rT}\mathbf{E}^*\left(S_0 e^{\sigma\sqrt{T}Z + T\left(r-\frac{\sigma^2}{2}\right)} - K\right)_+ \\
&= \mathbf{E}^*\left(S_0 e^{\sigma\sqrt{T}Z - \frac{\sigma^2}{2}T} - e^{-rT}K\right)_+,
\end{aligned}$$

which is exactly the second line in (30).

In fact, we need the convergence of moments, that is, uniform integrability. That can be done with a little more work, but the details are skipped. It is important to note that not only the price convergence, but the whole process $(S_{\tau_i}^N)$ converges to the exponential Brownian motion. This follows from Donsker's theorem.

9 Ruin theory

This part is based on Asmussen, Steffensen: Risk and Insurance [1].

9.1 Risk processes

The *classical risk process* or *Cramér–Lundberg model* is the following. Let $(N_t)_{t\geq 0}$ be a Poisson process with intensity $\lambda > 0$, independently let Z, Z_1, Z_2, ... nonnegative iid random variables with distribution G. In terms of the insurance, the kth jump time of the Poisson process represents that claim arrives of size Z_k. Then the reserve of the insurance company at time t is given by

$$U_t = u + ct - \sum_{k=1}^{N_t} Z_k, \qquad (35) \quad \texttt{\{eq:def-R\}}$$

where $u > 0$ is the initial capital, c is the rate of premium inflow.

For more general risk processes, called *renewal model*, we only assume that N is a renewal process (not necessarily Poisson). The interarrival times are $X, X_1, X_2, \ldots$, (which in case of Poisson process are iid $\mathrm{Exp}(\lambda)$), and we assume that $\mathbf{E}X =: 1/\lambda < \infty$.

The general risk process has the following ingredients:

- claims: $Z, Z_1, Z_2, \ldots$ nonnegative iid with df G, $\mathbf{E}Z = \mu$;
- time between claims / interarrival times: $X, X_1, X_2, \ldots$ iid, $\mathbf{E}X = \frac{1}{\lambda}$; this sequence defines the renewal process N_t;
- rate of premium inflow: $c > 0$;
- initial capital $u \geq 0$.

Ruin occurs if $U_t < 0$ for some t, and the time of ruin is

$$\tau(u) = \inf\{t > 0 : U_t < 0\}.$$

The ruin probability is the probability that ruin ever occurs

$$\psi(u) = \mathbf{P}(\tau(u) < \infty).$$

Proposition 5. *In the general renewal model,* $\lim_{t\to\infty} \frac{U_t}{t} = c - \lambda\mu$. *Furthermore, if* $\lambda\mu \geq c$ *then* $\psi(u) = 1$ *for all* $u \geq 0$, *while if* $\lambda\mu < c$ *then* $\psi(u) < 1$ *for all* $u \geq 0$.

Proof. By the SLLN

$$\frac{N_t}{t} \to \lambda, \quad \frac{Z_1 + \ldots + Z_n}{n} \to \mu,$$

implying

$$\frac{\sum_{k=1}^{N_t} Z_k}{t} = \frac{\sum_{k=1}^{N_t} Z_k}{N_t} \frac{N_t}{t} \to \lambda\mu \quad \text{a.s.},$$

thus the first claim follows.

If $c < \lambda\mu$, then $U_t \to -\infty$ meaning that ruin occurs a.s. If $c > \lambda\mu$ then $U_t \to \infty$, and $\mathbf{P}(\inf U_t \geq 0) > 0$ for $u = 0$ (thus for any $u \geq 0$).

If $c = \lambda\mu$ then the corresponding random walk oscillates between $+\infty$ and $-\infty$. Details are omitted. □

Therefore, in what follows we always assume the *net-profit condition*

$$c > \lambda\mu.$$

From the structure of U in (35) it is clear that ruin occurs when claim arrives. Recall that $X, X_1, \ldots$ are the interarrival times of N (the times between two consecutive claims). Define the variables $Y_k := Z_k - cX_k$. The net profit condition is exactly that $\mathbf{E}Y < 0$, which assures that $\sum_{k=1}^n Y_k \to -\infty$, so

$$M := \sup_{n \geq 0} \sum_{k=1}^n Y_k < \infty \quad \text{a.s.}$$

Then it is clear that

$$\psi(u) = \mathbf{P}(M > u),$$

that is the ruin probability is the tail of the maximum of a random walk with negative drift.

9.2 Maximum of a random walk

Here we follow Foss, Korshunov, and Zachary [4].

Let $Y, Y_1, Y_2, \ldots$ be iid random variables with $\mathbf{E}(Y) < 0$ and $S_n = \sum_{i=1}^n Y_i$ their partial sum. Define

$$M := \sup_{n \geq 0} Y_n < \infty.$$

We want to determine the tail probabilities $\mathbf{P}(M > u)$.

Let

$$\tau_+ = \tau_+(1) = \min\{k > 0 : S_k > 0\}$$

denote the *first strict ascending ladder epoch*, and S_{τ_+} the *first strict ascending ladder height*. Since $S_n \to -\infty$, these are defective random variables, that is

$$\mathbf{P}(\tau_+ = \infty) = \mathbf{P}(S_k \leq 0, \ \forall k) = \mathbf{P}(M = 0) =: p > 0. \tag{36}$$ {eq:maxrw-def-p}

If $\tau_+(1) < \infty$ we can define the second ladder epoch

$$\tau_+(2) = \min\{k > \tau_+(1) : S_k > S_{\tau_+(1)}\},$$

and inductively, if $\tau_+(n) < \infty$

$$\tau_+(n+1) = \min\{k > \tau_+(n) : S_k > S_{\tau_+(n)}\}.$$

Since $Y_1, Y_2, \ldots$ are independent, $\tau_+(2) - \tau_+(1)$ given $\tau_+(1) < \infty$ has the same distribution as $\tau_+(1)$, and in general if $\tau_+(n) < \infty$ then $\tau_+(1)$, $\tau_+(2) - \tau_+(1)$,

…, $\tau_+(n) - \tau_+(n-1)$ are iid. Furthermore, $\mathbf{P}(\tau_+(n) < \infty) = (1-p)^n$, and it is defined (possibly infinite) with probability $(1-p)^{n-1}$. The same independence shows that given $\tau_+(n) < \infty$ the variables

$$S_{\tau_+(1)}, S_{\tau_+(2)} - S_{\tau_+(1)}, \ldots, S_{\tau_+(n)} - S_{\tau_+(n-1)}$$

are independent and identically distributed. Put

$$H_d(x) = \mathbf{P}(S_{\tau_+} \le x), \quad H(x) = \mathbf{P}(S_{\tau_+} \le x | \tau_+ < \infty) = \frac{1}{1-p} H_d(x). \quad (37)$$ {eq:ladderheight}

Then H_d is a *defective* distribution function, that is $H_d(\infty) = 1 - p < 1$. Therefore,

$$\begin{aligned} \mathbf{P}(M > x) &= \sum_{k=1}^{\infty} \mathbf{P}(S_{\tau_+(k)} > x, \tau_+(k) < \infty, \tau_+(k+1) = \infty) \\ &= \sum_{k=1}^{\infty} \mathbf{P}(S_{\tau_+(k)} > x | \tau_+(k) < \infty)(1-p)^k p. \end{aligned} \quad (38)$$ {eq:M-expr}

Or, what is the same, if $\varepsilon_+(1), \varepsilon_+(2), \ldots$ are iid with df H, and independently N has geometric distribution $\mathbf{P}(N = k) = p(1-p)^k$, $k = 0, 1, \ldots$, then

$$M \stackrel{\mathcal{D}}{=} \sum_{i=1}^{N} \varepsilon_+(i).$$

Define the taboo renewal measure on $[0, \infty)$

$$R_+(B) = \mathbf{I}(0 \in B) + \sum_{n=1}^{\infty} \mathbf{P}(S_1 > 0, S_2 > 0, \ldots, S_n > 0, S_n \in B).$$

The random vector $(Y_1, \ldots, Y_n)$ has the same distribution as $(Y_n, \ldots, Y_1)$, thus

$$\begin{aligned} &\mathbf{P}(S_1 > 0, S_2 > 0, \ldots, S_n > 0, S_n \in B) \\ &= \mathbf{P}(Y_n > 0, Y_n + Y_{n-1} > 0, \ldots, S_n > 0, S_n \in B) \\ &= \mathbf{P}(S_n > S_{n-1}, S_n > S_{n-2}, \ldots, S_n > 0, S_n \in B) \\ &= \mathbf{P}(S_n \text{ is a strict ascending ladder height in } B\ ) \\ &= \mathbf{E}\mathbf{I}(S_n \text{ is a strict ascending ladder height in } B\ ). \end{aligned}$$

Therefore, summing over n we obtain the following.

{lemma:R+}

Lemma 11. *For any $B \subset (0,\infty)$, $R_+(B)$ is the expected number of ladder heights in B. In particular,*

$$R_+((0,\infty)) = \sum_{k=1}^{\infty} kp(1-p)^k = \frac{1-p}{p},$$

and

$$R_+([0,\infty)) = 1 + R_+((0,\infty)) = \frac{1}{p}.$$

Similarly, define the *weak descending ladder epochs* and *ladder heights* as follows. Let

$$\tau_- = \tau_-(1) = \min\{k > 0 : S_k \leq 0\},$$

the first weak descending ladder epoch, and S_{τ_-} is the ladder height. and inductively, for $\tau_-(n)$ defined let

$$\tau_-(n+1) = \min\{k > \tau_-(n) : S_k \leq S_{\tau_-(n)}\}.$$

Furthermore, let $\varepsilon_-(1), \varepsilon_-(2), \ldots$ iid with the same distribution as S_{τ_-}. Since $S_n \to -\infty$ a.s., these random variables are all proper (finite with probability 1). Moreover, by Lemma 11

$$\begin{aligned}
\mathbf{E}\tau_- &= \sum_{k=0}^{\infty} \mathbf{P}(\tau_- > k) = 1 + \sum_{k=1}^{\infty} \mathbf{P}(S_1 > 0, \ldots, S_k > 0) \\
&= R_+([0,\infty)) = \frac{1}{p}.
\end{aligned}$$

Since $\tau_-(1)$ is a stopping time for (S_n), by Wald's identity

$$\mathbf{E}\varepsilon_- = \mathbf{E}S_{\tau_-} = \mathbf{E}(\tau_-)\mathbf{E}(Y) = \frac{\mathbf{E}(Y)}{p}.$$

Define for $B \subset (-\infty, 0]$

$$R_-(B) = \sum_{n=0}^{\infty} \mathbf{P}(S_0 \leq 0, \ldots, S_n \leq 0, S_n \in B).$$

As in Lemma 11 we obtain

{lemma:R-}

Lemma 12.

$$R_-(B) = \mathbf{I}(0 \in B) + \mathbf{E}|\{\textit{weak descending ladder heights in } B\}|.$$

The distribution of the ascending ladder height can be expressed for $x > 0$ as

$$\begin{aligned}
\mathbf{P}(S_{\tau_+} > x) &= \sum_{n=1}^{\infty} \mathbf{P}(S_n > x, \tau_+ = n) \\
&= \sum_{n=1}^{\infty} \mathbf{P}(S_1 \le 0, \ldots, S_{n-1} \le 0, S_n > x) \\
&= \sum_{n=1}^{\infty} \int_{(-\infty,0]} \mathbf{P}(Y_n > x - y) \mathbf{P}(S_1 \le 0, \ldots, S_{n-1} \le 0, S_{n-1} \in \mathrm{d}y) \\
&= \int_{(-\infty,0]} \mathbf{P}(Y > x - y) R_-(\mathrm{d}y).
\end{aligned} \tag{39}$$

{eq:Stau+}

Similarly, for the distribution of the descending ladder height

$$\mathbf{P}(S_{\tau_-} \le -x) = \int_{[0,\infty)} \mathbf{P}(Y \le -x - y) R_+(\mathrm{d}y), \quad x \ge 0. \tag{40}$$

{eq:Stau-}

9.3 The Cramér–Lundberg model

In what follows we mainly consider the *classical risk process*, or *Cramér–Lundberg process*, where (N_t) is a Poisson process with intensity λ, that is the interarrival times are exponential(λ). Here we specialize the formulas obtained in the previous section.

Since $Y = Z - cX$, where $Z \ge 0$ is the claim size, X is exponential(λ), and Z, X are independent, for $y \ge 0$

$$\begin{aligned}
\mathbf{P}(Y \le -y) &= \mathbf{P}(X \ge c^{-1}(Z + y)) \\
&= \int_{(0,\infty)} e^{-\lambda(c^{-1}(z+y)} \mathbf{P}(Z \in \mathrm{d}z) \\
&= e^{-\frac{\lambda}{c} y} \int_{(0,\infty)} e^{-\frac{\lambda}{c} z} \mathbf{P}(Z \in \mathrm{d}z) \\
&= e^{-\frac{\lambda}{c} y} C_1.
\end{aligned}$$

So, by (40)

$$\begin{aligned}
\mathbf{P}(S_{\tau_-} \leq -x) &= \int_{[0,\infty)} \mathbf{P}(Y \leq -x - y) R_+(\mathrm{d}y) \\
&= \int_{[0,\infty)} e^{-\frac{\lambda}{c}(x+y)} C_1 R_+(\mathrm{d}y) \\
&= e^{-\frac{\lambda}{c}x} \int_{[0,\infty)} e^{-\frac{\lambda}{c}(y)} C_1 R_+(\mathrm{d}y) \\
&= e^{-\frac{\lambda}{c}x},
\end{aligned}$$

where the last equality follows from $\mathbf{P}(S_{\tau_-} \leq 0) = 1$. That is, the distribution of $-S_{\tau_-}$ is exponential(λ/c). This means that the weak descending ladder heights form a Poisson process on $(-\infty, 0)$, so Lemma 12 gives that the measure R_- is λ/c times the Lebesgue measure together with a unit mass at 0, or more formally

$$R_-(\mathrm{d}t) = \frac{\lambda}{c}\mathrm{d}t + \delta_0.$$

Therefore, by (39)

$$\begin{aligned}
\mathbf{P}(S_{\tau_+} > x) &= \int_{(-\infty,0]} \mathbf{P}(Y > x - y) R_-(\mathrm{d}y) \\
&= \int_{-\infty}^{0} \mathbf{P}(Y > x - y) \frac{\lambda}{c} \mathrm{d}y + \mathbf{P}(Y > x) \qquad (41) \quad \{\texttt{eq:Stau+aux1}\} \\
&= \frac{\lambda}{c} \int_{x}^{\infty} \mathbf{P}(Y > y) \mathrm{d}y + \mathbf{P}(Y > x).
\end{aligned}$$

For $x > 0$

$$\begin{aligned}
\mathbf{P}(Y > x) = \mathbf{P}(Z - cX > x) &= \int_{(x,\infty)} \left(1 - e^{-\lambda \frac{z-x}{c}}\right) \mathbf{P}(Z \in \mathrm{d}z) \\
&= \mathbf{P}(Z > x) - \int_{(x,\infty)} e^{-\frac{\lambda}{c}(z-x)} \mathbf{P}(Z \in \mathrm{d}z).
\end{aligned}$$

Substituting back, using Fubini

$$\begin{aligned}
&\frac{\lambda}{c}\int_x^\infty \mathbf{P}(Y>y)\mathrm{d}y \\
&= \frac{\lambda}{c}\int_x^\infty \mathbf{P}(Z>y)\mathrm{d}y - \frac{\lambda}{c}\int_x^\infty \int_{(x,\infty)} \mathbf{I}(z>y) e^{-\frac{\lambda}{c}(z-x)} \mathbf{P}(Z\in \mathrm{d}z)\mathrm{d}y \\
&= \frac{\lambda}{c}\int_x^\infty \mathbf{P}(Z>y)\mathrm{d}y - \int_{(x,\infty)} \left(1-e^{-\frac{\lambda}{c}(z-x)}\right) \mathbf{P}(Z\in \mathrm{d}z) \\
&= \frac{\lambda}{c}\int_x^\infty \mathbf{P}(Z>y)\mathrm{d}y - \mathbf{P}(Z>x) + \int_{(x,\infty)} e^{-\frac{\lambda}{c}(z-x)} \mathbf{P}(Z\in \mathrm{d}z) \\
&= \frac{\lambda}{c}\int_x^\infty \mathbf{P}(Z>y)\mathrm{d}y - \mathbf{P}(Y>x).
\end{aligned}$$

Substituting back into (41)

$$\mathbf{P}(S_{\tau_+}>x) = \frac{\lambda}{c}\int_x^\infty \mathbf{P}(Z>y)\mathrm{d}y.$$

Thus we obtained that S_{τ_+} has a density, and substituting $x=0$ we also see that $\mathbf{P}(\tau_+<\infty)=\lambda\mu/c<1$. Using the notation

$$G_I(x) = \frac{1}{\mu}\int_0^x \mathbf{P}(Z>y)\mathrm{d}y,$$

we see that

$$H(x) = \mathbf{P}(S_{\tau_+}\le x | \tau_+<\infty) = G_I(x).$$

Recalling also (36) and (38), and that $\psi(u)=\mathbf{P}(M>u)$ we obtain the following.

{thm:Stau+}

Theorem 22. *In the Cramér–Lundberg model assume the net-profit condition $c>\lambda\mu$. Then for the strict ascending ladder epoch τ_+ and height S_{τ_+} we have*

$$\begin{aligned}
\mathbf{P}(\tau_+<\infty) &= \frac{\lambda\mu}{c}<1, \\
\frac{\mathrm{d}}{\mathrm{d}x}\mathbf{P}(S_{\tau_+}\le x) &= \frac{\lambda}{c}\mathbf{P}(Z>x).
\end{aligned}$$

For the maximum of the random walk $M=\sup_{n\ge 0} S_n$

$$\psi(u) = \mathbf{P}(M>u) = \sum_{k=1}^\infty \left(\frac{\lambda\mu}{c}\right)^k \left(1-\frac{\lambda\mu}{c}\right)(1-G_I^{*k}(u)).$$

Exercise 22. Assume that $Z \sim$Exp(δ). Compute explicitly $\psi(u)$ using Theorem 22.

9.4 Lundberg inequality

Introduce the notation

$$S_t = \sum_{k=1}^{N_t} Z_k - ct.$$

Next we calculate the moment generating function of S.

Proposition 6. *Let N be a nonnegative integer valued random variable, independent of the iid sequence $Y, Y_1, \ldots$. Then*

$$\mathbf{E} \exp\left\{ s \sum_{k=1}^{N} Y_k \right\} = \mathbf{E}\left[\left(\mathbf{E} e^{sY} \right)^N \right].$$

Proof. Simply,

$$\begin{aligned} \mathbf{E} \exp\left\{ s \sum_{k=1}^{N} Y_k \right\} &= \sum_{i=0}^{\infty} \mathbf{P}(N = i) (\mathbf{E} e^{sY})^i \\ &= \mathbf{E}\left[\left(\mathbf{E} e^{sY} \right)^N \right]. \end{aligned}$$

□

Introduce the notation

$$\widehat{G}(s) = \mathbf{E} e^{sZ} = \int_{(0,\infty)} e^{sz} G(\mathrm{d}z).$$

Recalling that if $N \sim$ Poisson(λ) then $\mathbf{E} s^N = \exp\{\lambda(s-1)\}$, we obtain

$$\kappa(\alpha) := \log \mathbf{E} e^{\alpha S(1)} = \lambda(\widehat{G}(\alpha) - 1) - \alpha c. \tag{42}$$ {eq:def-kappa}

Similarly,

$$\log \mathbf{E} e^{\alpha S(t)} = t\kappa(\alpha).$$

Corollary 2. *For $\alpha \in \mathbb{R}$ assume that $\widehat{G}(\alpha) < \infty$. Let $M(t) = e^{\alpha S(t) - t\kappa(\alpha)}$. Then $(M(t)_{t \geq 0})$ is a martingale with respect to the filtration $\mathcal{F}_t = \sigma(S(s) : s \in [0,t])$.*

Proof. Let $t > s > 0$. Since $S(t) - S(s)$ is independent of $\mathcal{F}_s$ we have

$$\begin{aligned}\mathbf{E}[M(t)|\mathcal{F}_s] &= \mathbf{E}\left[\exp\{\alpha S(s) - s\kappa(\alpha) + \alpha(S(t) - S(s)) - (t-s)\kappa(\alpha)|\mathcal{F}_s\right] \\ &= M(s)\end{aligned}$$

as claimed. □

We need that the function κ is convex.

Lemma 13. *For any random variable X, the function $\kappa(\alpha) = \log \mathbf{E} e^{\alpha X}$ is convex.*

Proof. Differentiate twice, and use Cauchy–Schwarz. Indeed,

$$\begin{aligned}\kappa'(\alpha) &= \frac{\mathbf{E} X e^{sX}}{\mathbf{E} e^{sX}}, \\ \kappa''(\alpha) &= \frac{\mathbf{E}(X^2 e^{sX})\mathbf{E} e^{sX} - (\mathbf{E} X e^{sX})^2}{(\mathbf{E} e^{sX})^2}.\end{aligned}$$

Cauchy–Schwarz inequality with $X e^{sX/2}$, and $e^{sX/2}$ shows that $\kappa''(\alpha) > 0$, that is κ is convex. □

Note that $\kappa(0) = 0$, and $\kappa'(0) = \lambda\mu - c < 0$, since the net profit condition hold. *Cramér's condition* is that

$$\exists \gamma > 0 \text{ such that } \kappa(\gamma) = 0, \qquad \kappa'(\gamma) < \infty. \tag{43}$$ {eq:cramer}

Convexity implies that there is at most 1 solution. Then γ is the *Lundberg exponent.*

At the time of ruin $\tau(u)$ the process S upcrosses u by making a jump. Let $\xi(u) = S(\tau(u)) - u$ denote the overshoot (defined on the event $\tau(u) < \infty$).

Proposition 7. *Under the Cramér condition*

$$\psi(u) = \frac{e^{-\gamma u}}{\mathbf{E}[e^{\gamma \xi(u)} | \tau(u) < \infty]}.$$

Proof. Clearly, $\tau(u)$ is a stopping time, thus by the optional sampling theorem

$$\mathbf{E} M_{\tau \wedge t} = \mathbf{E} M_0 = 1.$$

Taking $\alpha = \gamma$, we have

$$1 = \mathbf{E}e^{\gamma S(\tau(u)\wedge t)} = \mathbf{E}\left[e^{\gamma S(\tau(u)\wedge t)}\mathbf{I}(\tau(u) \le t)\right] + \mathbf{E}\left[e^{\gamma S(\tau(u)\wedge t)}\mathbf{I}(\tau(u) > t)\right].$$

As $t \to \infty$ the second term tends to 0, as the integrand is bounded by $e^{\gamma u}$, and tends to 0 pointwise since $S(t) \to -\infty$. Thus

$$\begin{aligned}1 = \mathbf{E}\left[e^{\gamma S(\tau(u)\wedge t)}\mathbf{I}(\tau(u) \le t)\right] &= e^{\gamma u}\mathbf{E}\left[e^{\gamma \xi(u)}\mathbf{I}(\tau(u) < \infty)\right] \\ &= e^{\gamma u}\psi(u)\mathbf{E}\left[e^{\gamma\xi(u)}|\tau(u) < \infty\right].\end{aligned}$$

□

Noting that $\xi(u) \ge 0$ Lundberg's inequality is an immediate corollary.

Corollary 3 (Lundberg's inequality)**.** *Under the Cramér condition for all* $u \ge 0$

$$\psi(u) \le e^{-\gamma u}.$$

Corollary 4. *Assume that* $Z \sim$ *Exp*(δ)*. Then*

$$\psi(u) = \frac{\lambda}{c\delta}e^{-(\delta-\frac{\lambda}{c})u}.$$

Proof. Since

$$\kappa(\alpha) = \lambda(\widehat{G}(\alpha) - 1) - \alpha c = \lambda\left(\frac{\delta}{\delta-\alpha} - 1\right) - \alpha c,$$

the unique solution to $\kappa(\alpha) = 0$ is $\gamma = \delta - \lambda/c$.

Given $\tau(u) = t$ and $S(t-) = x \le u$ we know that the claim size $V > u - x$. The overshoot $\xi(u) = V - u + x$ given that $V > u - x$ is again exponential(δ). Thus

$$\mathbf{E}\left[e^{\gamma\xi(u)}|\tau(u) < \infty\right] = \int_0^\infty e^{\gamma y}\delta e^{-\delta y}\mathrm{d}y = \frac{c\delta}{\lambda}.$$

□

9.5 Cramér–Lundberg theorem

{thm:CramLund}

Theorem 23 (Cramér–Lundberg approximation)**.** *Assume that Cramér's condition hold, i.e.* $\kappa(\gamma) = 0$ *for some* $\gamma > 0$*. Then*

$$\psi(u) \sim Ce^{-\gamma u} \quad \text{as } u \to \infty,$$

where

$$C = \frac{c - \lambda\mu}{\kappa'(\gamma)} = \frac{c - \lambda\mu}{\lambda\widehat{G}(\gamma)' - c}.$$

This means that the ruin probability decreases exponentially with the initial capital u. These are good news for the insurance company.

Proof. Conditioning on the first ascending ladder height we have

$$\psi(u) = \mathbf{P}(S_{\tau_+} > u) + \int_{(0,u]} \psi(u-y)\mathbf{P}(S_{\tau_+} \in \mathrm{d}y).$$

Using the explicit expressions in Theorem 22

$$\psi(u) = \frac{\lambda}{c} \int_u^\infty \mathbf{P}(Z > y)\mathrm{d}y + \int_0^u \psi(u-y)\frac{\lambda}{c}\mathbf{P}(Z > y)\mathrm{d}y.$$

This is a *defective renewal equation*, since the total mass on the integral involving ψ is

$$\int_0^\infty \frac{\lambda}{c}\mathbf{P}(Z > y)\mathrm{d}y = \frac{\lambda}{c}\mu < 1.$$

Put $\widetilde{\psi}(u) = e^{\gamma u}\psi(u)$. Multiplying both sides by $e^{\gamma u}$, we have

$$\widetilde{\psi}(u) = \frac{\lambda}{c}e^{\gamma u} \int_u^\infty \mathbf{P}(Z > y)\mathrm{d}y + \int_0^u \widetilde{\psi}(u-y)\frac{\lambda}{c}\mathbf{P}(Z > y)e^{\gamma y}\mathrm{d}y. \qquad (44) \quad \texttt{\{eq:reneq-psi\}}$$

This is proper renewal equation, as

$$\int_0^\infty \frac{\lambda}{c}\mathbf{P}(Z > y)e^{\gamma y}\mathrm{d}y = 1. \qquad (45) \quad \texttt{\{eq:reneq?\}}$$

Indeed, by Fubini

$$\begin{aligned}
\int_0^\infty \mathbf{P}(Z > y)e^{\gamma y}\mathrm{d}y &= \int_0^\infty \int_{(0,\infty)} \mathbf{I}(z > y)e^{\gamma y}G(\mathrm{d}z)\mathrm{d}y \\
&= \gamma^{-1}\left(\int_{(0,\infty)} e^{\gamma z}G(\mathrm{d}z) - 1\right) \\
&= \gamma^{-1}\left(\widehat{G}(\gamma) - 1\right),
\end{aligned}$$

and by the definition of the Lundberg exponent (see (43) and (42))

$$\widehat{G}(\gamma) = 1 + \frac{\gamma c}{\lambda}.$$

Substituting everything back we obtain (45).

Introducing the notation

$$a(u) = \frac{\lambda}{c} e^{\gamma u} \int_u^\infty \mathbf{P}(Z > y) \mathrm{d}y,$$

and

$$F(u) = \int_0^u \frac{\lambda}{c} \mathbf{P}(Z > y) e^{\gamma y} \mathrm{d}y,$$

we have the usual form

$$\widetilde{\psi}(u) = a(u) + \int_0^u \widetilde{\psi}(u - y) F(\mathrm{d}y).$$

Therefore, we can apply the key renewal theorem[1] to equation (44). We obtain

$$\lim_{u \to \infty} \widetilde{\psi}(u) = \frac{\int_0^\infty a(y)\mathrm{d}y}{\int_{(0,\infty)} yF(\mathrm{d}y)} = \frac{\int_0^\infty \frac{\lambda}{c} e^{\gamma u} \int_u^\infty \mathbf{P}(Z > y)\mathrm{d}y\mathrm{d}u}{\frac{\lambda}{c}\int_0^\infty y\mathbf{P}(Z > y)e^{\gamma y}\mathrm{d}y}.$$

It remains to evaluate the constant above. The numerator

$$\begin{aligned}
\int_0^\infty e^{\gamma u} \int_u^\infty \mathbf{P}(Z > y)\mathrm{d}y\mathrm{d}u &= \frac{1}{\gamma}\int_0^\infty (e^{\gamma y} - 1)(1 - G(y))\mathrm{d}y \\
&= \gamma^{-1}\left(\frac{c}{\lambda} - \mu\right).
\end{aligned}$$

For the denominator

$$\begin{aligned}
\int_0^\infty y\mathbf{P}(Z > y)e^{\gamma y}\mathrm{d}y &= \int_0^\infty \int_{(0,\infty)} \mathbf{I}(z > y)G(\mathrm{d}z)ye^{\gamma y}\mathrm{d}y \\
&= \gamma^{-1}\int_{(0,\infty)} \left(e^{\gamma z}z - \gamma^{-1}(e^{\gamma z} - 1)\right)G(\mathrm{d}z) \\
&= \gamma^{-1}\mathbf{E}(Ze^{\gamma Z}) - \gamma^{-2}\frac{\gamma c}{\lambda} = \frac{1}{\lambda\gamma}\kappa'(\gamma),
\end{aligned}$$

since

$$\kappa'(\gamma) = \lambda\widehat{G}(\gamma)' - c = \lambda\mathbf{E}(Ze^{\gamma Z}) - c.$$

□

[1] Here we have to check that a is directly Riemann integrable. In fact, it is of bounded variation, which is not difficult to show.

9.6 Heavy tails

We learned that if the claim sizes have exponential moments, more precisely Cramér's condition holds, then the ruin probability decreases exponentially with the capital. This is good for the insurance company, since in order to decrease the ruin probability by a factor 0.5, only a constant amount of money needed. The bad news is that in practice, claims are heavy-tailed, no exponential moments exists. Assume that the claim sizes have Pareto distribution,

$$\mathbf{P}(Z > x) = 1 - G(x) = x^{-\alpha}, \quad x \geq 1,$$

for some $\alpha > 1$. Then $\mathbf{E}Z = \mu = \frac{\alpha}{\alpha - 1} < \infty$. The integrated tail distribution

$$G_I(x) = \frac{1}{\mu} \int_0^x (1 - G(y)) \mathrm{d}y = \frac{\alpha - 1}{\alpha} \left(1 + \int_1^x y^{-\alpha} \mathrm{d}y \right) = 1 - \alpha^{-1} x^{1-\alpha}.$$

Then by Theorem 22 we see

$$\psi(u) = \sum_{k=1}^{\infty} \left(\frac{\lambda\mu}{c} \right)^k \left(1 - \frac{\lambda\mu}{c} \right) (1 - G_I^{*k}(u)) \geq \frac{\lambda\mu}{c} \left(1 - \frac{\lambda\mu}{c} \right) \alpha^{-1} u^{1-\alpha},$$

that is

$$\liminf_{u \to \infty} u^{\alpha - 1} \psi(u) > 0.$$

In this case the ruin probability is much larger, decreases only as a power of u.

In what follows, we are dealing with distributions on $(0, \infty)$ with unbounded support. In terms of random variables, $Z, Z_1, \ldots$ are nonnegative iid random variables, $G(x) = \mathbf{P}(Z \leq x)$, $G(x) < 1$ for all x. Let $S_n = Z_1 + \ldots + Z_n$ denote the partial sum. Then $\mathbf{P}(S_n \leq x) = G^{*n}(x)$.

The distribution G is *subexponential*, if

$$\lim_{x \to \infty} \frac{1 - G^{*2}(x)}{1 - G(x)} = 2. \qquad (46) \quad \{\texttt{eq:def-subexp}\}$$

Exercise 23. Determine the limit above if

- $G(x) = 1 - e^{-\lambda x}$;
- $G(x) = \Phi(x)$ standard normal distribution;
- $G(x) = 1 - x^{-\alpha}$.

Note that

$$1 - G^{*2}(x) = 1 - \mathbf{P}(Z_1 + Z_2 \leq x) = \mathbf{P}(Z_1 + Z_2 > x).$$

For any distribution with unbounded support

$$\begin{aligned}\mathbf{P}(\max(Z_1, Z_2) > x) &= 1 - \mathbf{P}(\max(Z_1, Z_2) \leq x) \\ &= 1 - G(x)^2 \sim 2(1 - G(x)),\end{aligned}$$

where the last asymptotic equality holds as $x \to \infty$. Thus definition (46) is equivalent to

$$\mathbf{P}(Z_1 + Z_2 > x) \sim \mathbf{P}(\max(Z_1, Z_2) > x) \quad \text{as} \quad x \to \infty.$$

That is, the sum is large if and only if one term is large. This is the *one big jump* property.

{prop:subexp}

Proposition 8. *Let G be subexponential. Then*

$$\lim_{x \to \infty} \frac{1 - G^{*n}(x)}{1 - G(x)} = n,$$

or equivalently

$$\mathbf{P}(S_n > x) \sim \mathbf{P}(\max\{Z_1, \ldots, Z_n\} > x) \quad \textit{as} \quad x \to \infty.$$

To ease notation write

$$\rho = \frac{\lambda \mu}{c}.$$

Theorem 22 states that

$$\psi(u) = \sum_{k=1}^{\infty} \rho^k (1 - \rho)(1 - G_I^{*k}(u)).$$

If G_I is subexponential, then by Proposition 8[2]

$$\begin{aligned}\psi(u) &= \sum_{k=1}^{\infty} \rho^k (1 - \rho)(1 - G_I^{*k}(u)) \\ &\sim \sum_{k=1}^{\infty} \rho^k (1 - \rho) k (1 - G_I(u)) = \frac{\rho}{1 - \rho}(1 - G_I(u)).\end{aligned}$$

Corollary 5. *Consider the Cramér–Lundberg risk process with the net-profit condition, and assume that G_I is subexponential. Then as $u \to \infty$*

$$\psi(u) \sim \frac{\rho}{1 - \rho}(1 - G_I(u)).$$

[2] In fact, here we need a bit more, a kind of uniform bound to ensure that we can interchange summation and limit.

References

[1] S. Asmussen and M. Steffensen. *Risk and insurance—a graduate text*, volume 96 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2020.

[2] R. J. Elliott and P. E. Kopp. *Mathematics of financial markets*. Springer Finance. Springer-Verlag, New York, second edition, 2005.

[3] L. C. Evans. *An introduction to stochastic differential equations*. American Mathematical Society, Providence, RI, 2013.

[4] S. Foss, D. Korshunov, and S. Zachary. *An introduction to heavy-tailed and subexponential distributions*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2013.

[5] J. Gáll and G. Pap. *Bevezetés a pénzügyi matematikába*. Polygon, 2010.

[6] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[7] A. N. Shiryaev. *Essentials of stochastic finance*, volume 3 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co., Inc., River Edge, NJ, 1999. Facts, models, theory, Translated from the Russian manuscript by N. Kruzhilin.