Foundations of Mathematical Finance

Yuguang Yang

Notations

- $\mathbb{R}$: real numbers.
- $\mathbb{C}$: complex numbers.
- $\mathbb{F}$:real or complex numbers.
- $\mathbb{N}$:natural numbers.
- $\mathcal{R}(A)$: the range of matrix A.
- $\mathcal{N}(A)$: the null space of matrix A.
- V: vector space.
- $det(A)$:the determinant of matrix A.
- $rank(A)$:the rank of matrix A.
- $Tr(A)$:the trace of matrix A.
- $L^2[a,b]$:Lebesgue integrable function on $[a,b]$.
- $L^1[a,b]$:Lebesgue integrable function on $[a,b]$.
- $N(0,1)$: standard Gaussian distribution.
- $N(\mu,\sigma^2)$: Gaussian distribution with mean μ and variance σ^2.
- $MN(\mu,\Sigma)$: multivariate Gaussian distribution with mean vector μ and covariance matrix Σ.
- $E_Q[\cdot]$: expectation taken with respect to risk-neutral measure Q.
- $\mathcal{F}_t$: σ algebra geneated by a driving stochastic process, say Bronwian motion, up to time t.
- r: risk-free short rate.
- $B(t), M(t)$:stochastic bank/money market account; $B(t) = M(t) = \exp(\int_0^t r(s)ds)$.
- $DF(t_1,t_2)$: discount factor from t_2 to t_1; $DF(t_1,t_2) = E_Q[\exp(-\int_{t_1}^{t_2} r(s)ds)|\mathcal{F}_{t_1}]$.
- $P(t,T)$: time-t price of a zero coupon bond with a maturity at T; $P(t,T) = E_Q[\exp(-\int_t^T r(s)ds)|\mathcal{F}_t]$.
- $df(t_1,t_2)$: stochastic discount factor discounting cash flow from t_2 to t_1; $df(t_1,t_2) = \exp(-\int_{t_1}^{t_2} r(s)ds)$.
- $DF(t_1,t_2)$: discount factor (a deterministic quantity) discounting cash flow from t_2 to t_1; $DF(t_1,t_2) = E_Q[\exp(-\int_{t_1}^{t_2} r(s)ds)|\mathcal{F}_{t_1}]$.
- $L(S,T)$: LIBOR interest rate over period from S to T.
- $F(t,S,T)$: forward LIBOR interest rate over period from S to T observed at time t.
- $S(t,T_0,T_n)$ par swap rate for a fixed-floating swap starting from T_0 to T_n.
- $A(t,T_0,T_n)$ annuity with payments starting from T_0 to T_n.
- $EE(t)$ undiscounted expected exposure; $EE(t) = E_N[V(t)|\mathcal{F}_0]$, where E_N is the expectation with respect to an appropriate martingale measure.
- $EPE(t), ENE(t)$ undiscounted expected positive exposure and negative exposure; $EPE(t) = E_N[V(t)^+|\mathcal{F}_0]$, $ENE(t) = E_N[V(t)^-|\mathcal{F}_0]$, where $E_N[\cdot]$ is the expectation taken with respect to the equivalent martingale measure associated with numeriare $N(t)$.

- CCY1/CCY2 spot rate, where CCY2 is the quoting currency and CCY1 is the base currency, shows the amount of CCY1 needed to exchange for 1 unit of CCY2.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1

NO-ARBITRAGE PRICING PRINCIPLE

1.1 Introduction

1.1.0.1 *Overview*

The no-arbitrage pricing framework developed in this work is based on a fundamental assumption that market participants cannot gain riskless profits by buying and selling instruments traded in an active and integrated market, as we emphasize in Assumption 1.1.

> **Assumption 1.1 (the fundamental no-arbitrage assumption).** *There is no opportunities for market participants to gain riskless profits by buying and selling instruments traded in an active and integrated market*

If a real-world market or a market mathematical model constructed satisfies such assumption, we say the market (model) is no-arbitrage market.

Assuming markets are free of arbitrage is a reasonable assumption, particularly to active markets where any 'free lunch' will be quickly arbitraged away by arbitragers. Without constructing any sophisticated model, this no-arbitrage assumption can lead to Law of one price, which essentially states that financial instruments with same payoff at a specified future time should have same prices; otherwise arbitrage opportunities exist.

Law of One Price can be readily applied to derive no-arbitrage price for instruments and construct no-arbitrage market models. For example, linear pricing theorem, a direct consequence of Law of One Price, can be used to price complex instruments whose payoff can be linearly decomposed to payoff functions of simple and actively traded instruments. Law of One Price is also useful in deriving price bounds and relations among different financial instruments.

Despite all the merits of Law of One Price, more exotic instruments and portfolio risk management (where how price will change as response to market conditions) requires sophisticated mathematical market model to be constructed.

The breakthrough in option pricing came with the famous Black and Scholes paper in 1973[1], where they show that option could be priced by dynamically creating a risk-less portfolio via buy and sell the underlying stock continuously.

The other side of the idea of dynamical hedging is dynamical replication. Consider managing a portfolio based on a trading strategy such that the portfolio at the expiration date of an option exactly replicate the payoff, the Law of One Price gives that the initial capital of the replicating portfolio should be the same of the price of the option.

Such dynamical replication based pricing methodology has been formalized and greatly extended by risk-neutral pricing framework developed by M. Harrison, D. Kreps, and S.

Pliska [2]. The construction and application of this types of models usually require additional assumption, as we note in

Assumption 1.2 (additional assumptions for the market).

- *We always consider an economy with continuous and frictionless (no transaction cost) trading take place within a finite horizon* $[0, T]$.
- *We can borrow and lend at the same rate without limit.*
- *The market participants are subject to the same tax rate on all net trading profits.*

They pointed out that once a model is built, one can change the probability measure on the space on which the stock price is defined so that it has mean rate of return equal to the interest rate. This is akin to building a model based on tosses of a fair coin, and then pretending for computational purposes that the coin is biased. Under this socalled risk-neutral measure, both the stock and the money-market account held by the portfolio replicating a call have mean rate of return equal to the interest rate, and so the portfolio itself has this mean rate of return. Therefore, the initial value of the portfolio, which is the Black-Scholes price of the call, can be obtained by discounting the call payoff at the interest rate and taking the expected value under the risk-neutral measure.

In this chapter, instead of introducing concrete market models, we cover different types generic market model frameworks, including single-period finite-state model, multiple-period finite state-state model and continuous-time infinite-stat model driven by Brownian motion.

For these models framework, we demonstrate how the no-arbitrage conditions are ensured using the martingale theory. These no-arbitrage martingale pricing frameworks can then serve as model factories to enable us to construct different types of concrete model based on specific requirements of markets and instruments.

1.1.1 The concept of price

Definition 1.1.1 (different types of prices).

- ***Instrinsic price*** *of an instrument is the amount of money that rational buyers and rational sellers agree on in a liquid market. It is the result of market forces.*
- ***No-arbitrage price*** *of an instrument is the price that arbitrage opportunities cannot occur.*
- *In an efficient market where no arbitrage opportunities exist, the no-arbitrage price will equal the instrinsic price.*
- *In asset pricing, we use no-arbitrage price to approximate the instrinsic price in a no-arbitrage market model.*

1.1.2 The goal of asset price

Remark 1.1.1 (The goal of asset pricing).

- Price of an asset is the amount of money that rational buyers and rational sellers agree on in a liquid market.
- Some asset prices are already existing, such that bonds and stocks.
- Our goal is to
  - Determine prices for untraded instruments based on the prices of instruments currently traded.
  - Understand the relationship between prices and underlying risk factors, laying the foundation for risk managemenet.
- There are generally two types of methods for pricing assets: model-based pricing and model-free pricing.
  - Examples for model-free pricing include pricing via replication, linear pricing theorem.
  - Examples of model based pricing is: Assuming asset dynamics are given by SDE driven by Brownian motions. The model should be arbitrage free.

1.2 The Law of one price

1.2.1 The Law of One Price

Proposition 1.2.1 (The law of One Price). *[3] In an **arbitrage-free** market, consider the values at time t of two portfolio V_X, V_Y (they are random variables parametrized by time $t' > t$): if they are the 'same' random variable at a future time $\tau > t$, in the sense that, they have the equivalent mapping from the random sample space to price, then they must have the same value at time t. If one portfolio at a future time $\tau > t$ is more valuable(or less) regardless of the random outcomes, then one portfolio is more valuable (or less).*

Proof. The simplest proof is that: a portfolio is a good that can generate cash flow, therefore it has value and price. (The portfolio is a good can bring payoff under different state of the world). In a competitive market, same goods must have the same price otherwise there will be arbitrage opportunities. □

Remark 1.2.1 (Implicit assumptions of no-arbitrage condition). Implicit assumptions underlying the no-arbitrage condition:

- Markets are liquid: sufficient number of buyers and sellers

- Price information is available to all buyers and sellers
- Competitions in supply and demand will correct any deviation from no-arbitrage prices.
- Same borrow interest rate and lending rate.

Corollary 1.2.1.1 (valuation of zero-valued portfolio). *If the value of a portfolio is equal to 0 regardless of the random outcomes in future $\tau > t$, then the value of the portfolio is 0 at time t.*

Remark 1.2.2 (implications for pricing).

- The law of one price will give **two ways of pricing a portfolio: portfolio replication and construction of risk-free portfolio.**
- Note that these two method can be converted mutually. Suppose we can replicate a state claim using some portfolio, which has value V_0. Then we can short this portfolio at V_0, and use it to buy the state claim at p. At time T, we return the state claim. In the process, we are guaranteed/deterministically to get zero at T. Therefore, our initial cash flow must be $0 = V_0 - p$.

Corollary 1.2.1.2 (valuation of risk free portfolio). *Consider a no-arbitrage market with risk-free asset. If the value $V(T)$ of a portfolio at time T in the future is independent of the random outcome, then*

- *its current time t value is*
$$V(t) = e^{-r(T-t)}V(T), t < T.$$
where r is the constant risk free rate.
- *the growth rate of $V(t)$ is r.*

Remark 1.2.3. The implication of this corollary is that if we can construct a risk-free portfolio, i.e., a portfolio without randomness and get a deterministic payoff V_T at future time T. Then its initial price V_0 must satisfy $V_0 = \exp(-rT)V_T$. It has to be such, otherwise people can just long(short) infinite amount of this portfolio by borrowing/lending money at interest rate r.

1.2.2 Linear pricing theorem

Proposition 1.2.2 (linear pricing theorem). *Consider a **arbitrage-free** market with N finite states. Assume there are two assets with payoff vector $D_1, D_2 \in \mathbb{R}^N$ with price p_1, p_2. Then*

- *For an asset with payoff vector $D = D_1 + D_2$, its price is $p = p_1 + p_2$.*
- *For an asset with payoff vector $aD_1, (a \in \mathbb{R})$, its price is $p = ap_1$.*

- *For an asset with payoff vector $aD_1 + bD_2, (a, b \in \mathbb{R})$, its price is $p = ap_1 + bp_2$.*

Proof. (1) If $p < p_1 + p2$, then we can short $D_1 + D_2$ and long D to earn risk-free money. Similar operations can apply to $p > p_1 + p_2$. Therefore, $p = p_1 = p_2$. (2)(3) similar to (1). □

Corollary 1.2.2.1 (Linear pricing theorem with Arrow securities). *Consider a **arbitrage-free** market with N finite states, assume there are a set of N Arrow securities that produce a unit payoff in each market scenario. If each Arrow security has price P_i, then for any other securities that with payoff vector $D \in \mathbb{R}^N$, its price is $D^T P$.*

Proof. Because the payoff vector is in vector space; that is, a payoff vector can be decomposed as linear combination of Arrow securities payoff vectors. □

1.2.3 Forward contract pricing

Definition 1.2.1 (forward price). *When the two parties enter a contract that on a future date T, an asset is traded at the **forward price** F such that the two parties will not pay extra money to enter the contract.*

Lemma 1.2.1 (forward pricing using no-arbitrage argument). *The fair price should satisfy:*

$$F + \sum_{i=1}^{N} D_i e^{(r-q)(T-t)} = S_0 e^{(r-q)T}$$

where r is the risk free rate, D_i is the dividend guaranteed to pay at t_i, S_0 is the spot price of the underlying asset, q is the cost-of-carry. In particular, for the case of non-dividend and zero cost-of-carry, we have

$$F = S_0 e^{rT}.$$

Proof. The method to replicate a forward for the underlying asset is to buy the asset, and wait to contract expired date. The full cash flow generated within the process discounted to the payment date is the price of forward price. The idea is the asset hold till the expired date is the replication.

For example, If $F < S_0 e^{rT}$, we can short the asset S_0, and put the money into bank to get interest, and then at the expiry date we buy at $F < S_0 e^{rT}$ to return the asset; If $F > S_0 e^{rT}$, we will borrow money S_0 at a rate r to buy the asset, and then sell at $F > S_0 e^{rT}$ and clean the debt at T. □

1.3 Single period finite-state market model

1.3.1 No-arbitrage and asset pricing

1.3.1.1 *The setup and the fundamental theorem*

Setup of general finite-state market

- A probability space (Ω, F, P) describing the world.
- A market in which N assets, labeled $A_1, A_2, ..., A_N$, are freely traded.
- Uncertainty about the behavior of the market is encapsulated in a finite-sized sample space Ω of K possible market scenarios or sample points, labeled $\omega_1, \omega_2, ..., \omega_K$, and $P(\omega) > 0 \forall \omega \in \Omega$. There is an N by K payoff matrix $D \in \mathbb{R}^{N\times K}$ with entries $V_j(\omega_i)$ such that, in scenario i, the value of the asset A_j at after one-period is $V_j(\omega_i)$. The payoff matrix D is given by

$$\begin{pmatrix} V_1(\omega_1) & V_1(\omega_2) & \cdots & V_1(\omega_K) \\ V_2(\omega_1) & V_2(\omega_2) & \cdots & V_2(\omega_K) \\ \vdots & \vdots & \ddots & \vdots \\ V_N(\omega_1) & V_N(\omega_2) & \cdots & V_N(\omega_K) \end{pmatrix}$$

- We denote a portfolio vector by $\theta \in \mathbb{R}^N$
- The prices of the N assets at current time is represented by a price vector $p \in \mathbb{R}^N$.

Definition 1.3.1 (arbitrage portfolio). *A θ is an arbitrage portfolio is either*

1. *(type I arbitrage) initial price $p^T\theta \leq 0$ but final payoff vector $D^T\theta > 0, D^T\theta \in \mathbb{R}^K$.*
2. *(type II arbitrage) initial price $p^T\theta < 0$ but final payoff vector $D^T\theta \geq 0$.*

Remark 1.3.1. According to this, an arbitrary portfolio θ guarantees some positive return in every possible states, yet it costs nothing to purchase. Or it guarantees a non-negative return whiling have negative cost.

Proposition 1.3.1 (Fundamental theorem of asset pricing). *Consider a single-period finite state market model represented by payoff matrix $D \in \mathbb{R}^{N\times K}$. Further let the prices of the N assets at current time be represented by a price vector $p \in \mathbb{R}^N$. It follows that*

- *The market model has no arbitrage opportunities if and only there exists a nonnegative vector $\pi \in \mathbb{R}^K$ such that the price of each asset i is given as*

$$p_i = \sum_{j=0}^{K} \pi_j D_{ij};$$

or in matrix form

$$p = D\pi.$$

- *And π is also known as* ***state price****.*

Proof. (1)**Forward (if there is such nonnegative vector, then there is no arbitrage):** If there is a nonnegative vector $\pi \in \mathbb{R}^K$ such that the price of each asset i is given as

$$p_i = \sum_{j=0}^{K} \pi_j D_{ij}$$

Then for any portfolio θ, its initial value is given as

$$V_0 = \theta^T p = \sum_{i=1}^{N} \theta_i p_i = \sum_{i=1}^{N} \sum_{j=0}^{K} \theta_i \pi_j D_{ij} = \theta^T D\pi.$$

- For any final payoff $\theta^T D > 0$(positive payoff), its initial price $V_0 = \theta^T p = \theta^T D\pi > 0$. Therefore, the type I arbitrage opportunity will not exist.
- For any final payoff $\theta^T D \geq 0$(nonnegative payoff), its initial price $V_0 = \theta^T p = \theta^T D\pi \geq 0$. Therefore, the type II arbitrage opportunity will not exist.

(2) **Converse:** (use **??**)

- When there is no arbitrage type I, i.e. the situation that there exists some portfolio vector $\theta \in \mathbb{R}^N$ such that the payoff $V = D^T\theta \geq 0$ and $\theta^T p < 0$ cannot happen, then the situation that there exists some positive $\pi \in \mathbb{R}^K$ such that $D\pi = p$ will definitely happen.
- When there is no arbitrage type II, i.e. the situation that there exists some portfolio vector $\theta \in \mathbb{R}^N$ such that the payoff $V = D^T\theta > 0$ and $\theta^T p \leq 0$ cannot happen, then the situation that there exists some positive $\pi \in \mathbb{R}^K$ such that $D\pi = p$ will definitely happen.

□

Remark 1.3.2 (existence and uniqueness)**.** Note that this theorem only guarantees existence but not uniqueness.

Note 1.3.1 (detecting arbitrage conditions)**.** Given a payoff matrix D and the price vector p observed at the current market. There are following three possibilities:

1. There does not exist a π (no matter positive or negative) satisfying $p = D^T\pi$.

2. There exists a π satisfying $p = D^T\pi$, but π is not fully non-negative for all its components.
3. There exists one or infinite non-negative π satisfying $p = D\pi$.

Example 1.3.1. A simple application of the Arbitrage condition Given $D = \begin{pmatrix} 1+r & 1+r \\ s_{up} & s_{down} \end{pmatrix}$ and $S = (1, s)^T$. Using above theorem, we have $S = D\phi$, or

$$0 = ((1+r) - s_{up}/s)\phi_1 + ((1+r) - s_{down}/s)\phi_2$$

We requires $\phi_1, \phi_2 > 0$, then $s_{down}/s < 1 + r < s_{up}/s$

1.3.1.2 *No-arbitrage pricing*

Lemma 1.3.1 (no-arbitrage pricing using arbitrary state price vector). *Consider a single-period finite state market model ,consisting of N assets, represented by payoff matrix $D \in \mathbb{R}^{N\times K}$. Suppose the current prices of the N assets represented by a price vector $p \in \mathbb{R}^N$* ***is unknown.*** *Then the no-arbitrage prices for these assets are*

$$p = D^T\pi,$$

where π is ***arbitrary non-negative state vector****.*

Proof. From Theorem 1.3.1, we can see that p satisfying $p = D^T\pi$ will not give rise to arbitrage opportunities. □

Remark 1.3.3 (no-arbitrage price vs. instrinsic price)**.** As we discussed in Definition 1.1.1, no-arbitrage price is not necessarily the instrinsic price.

Definition 1.3.2 (replication portfolio). *Consider a market where there are freely trade asset B, and $A_1, A_2, ..., A_N$. The one-period payoff vectors of As are given by $V_1, V_2, ..., V_N \in \mathbb{R}^K$. A portfolio $\theta \in \mathbb{R}^N$ in the asset As is a replicating portfolio of asset B if there exists a portfolio vector $\theta = (\theta_1, \theta_2, ..., \theta_N)$ such that*

$$V^B = \sum_{i=1}^{N} \theta_i V_i.$$

Proposition 1.3.2 (pricing via replication). *indexreplation pricing Suppose that θ is a replicating portfolio of B in the asset As. Denote the price vector of asset As by p. If the market is* ***arbitrage-free****, then $t = 0$ the no-arbitrage price of B is given as*

$$p^B = \sum_{i=1}^{n} \theta_i p_i$$

Proof. Because the replication portfolio has the same payoff as the asset B, the law of one price (Theorem 1.2.1) requires that they have the same price. □

Remark 1.3.4. The replication pricing procedure only relies on law of one price and does not require additional conditions like existence of risk neutral measure, market completeness.

Lemma 1.3.2 (replication indicates correlation, necessary condition for replication). *Let X_B be the random variable representing the payoff of B, let $Y_1, ..., Y_N$ be the random variables representing the payoff of assets $A_1, ..., A_N$. Suppose an asset B is replicated by the* ***mutually independent*** *assets $A_1, ..., A_K$ such that the payoff X_B of B is*

$$X_B = \sum_{i=1}^{N} \theta_i Y_i, \theta \in \mathbb{R}^K$$

Then if** $\theta_i \neq 0$**, then** X_B **is correlated with** Y_i**.

Proof. WLOG, we assuming $\theta > 0$ for every component. Then

$$cov(Y_i, X_B) = cov(Y_i, \sum_{i=1}^{N} Y_i) = \theta cov(Y_i, Y_i) \geq 0$$

□

Remark 1.3.5 (implications).

- In general, only economically connected(or correlated) assets can be replicated. Different uncorrelated stocks, for example, can not be replicated. For example, derivatives can be replicated by using its underlying assets.
- **hedging via replicating** For every asset B sold, buy replicating portfolio θ such that the future payoff is neutral to the world state, that is, net gain = net loss = 0.

1.3.2 Risk-neutral measure and pricing

1.3.2.1 *Setup and the risk-neutral measure*

Setup of general finite-state market with risk-free asset

- A probability space (Ω, F, P) describing the world.
- A market in which N assets, labeled $A_1, A_2, ..., A_N$, are freely traded. **Assume that one of these, say A_1, is risk-free.**
- Uncertainty about the behavior of the market is encapsulated in a finite-sized sample space Ω of K possible market scenarios or sample points, labeled $\omega_1, \omega_2, ..., \omega_K$. There is an N by K payoff matrix $D \in \mathbb{R}^{N\times K}$ with entries $V_j(\omega_i)$ such that, in scenario i, the value of the asset A_j at after one-period is $V_j(\omega_i)$. The payoff matrix D is given by

$$\begin{pmatrix} V_1(\omega_1) & V_1(\omega_2) & \cdots & V_1(\omega_K) \\ V_2(\omega_1) & V_2(\omega_2) & \cdots & V_2(\omega_K) \\ \vdots & \vdots & \ddots & \vdots \\ V_N(\omega_1) & V_N(\omega_2) & \cdots & V_N(\omega_K) \end{pmatrix} = \begin{pmatrix} 1+r & 1+r & \cdots & 1+r \\ V_2(\omega_1) & V_2(\omega_2) & \cdots & V_2(\omega_K) \\ \vdots & \vdots & \ddots & \vdots \\ V_N(\omega_1) & V_N(\omega_2) & \cdots & V_N(\omega_K) \end{pmatrix},$$

  where r is the risk-free return.
- We denote a portfolio vector by $\theta \in \mathbb{R}^N$.
- The prices of the N assets at current time is represented by a price vector $p \in \mathbb{R}^N$. In particularly, $p_1 = 1$.

Proposition 1.3.3 (Fundamental theorem of asset pricing in a world with risk-free asset). *Consider a single-period finite state market model represented by payoff matrix $D \in \mathbb{R}^{N\times K}$. Further let the prices of N assets at current time be represented by a price vector $p \in \mathbb{R}^N$. It follows that*

- *The market model has no arbitrage opportunities if and only if there exists a nonnegative vector $\pi \in \mathbb{R}^K$ such that the price of each asset i is given as*

$$p_i = \frac{1}{1+r}\sum_{j=0}^{K} \pi_j D_{ij};$$

  or in matrix form

$$p = \frac{1}{1+r} D\pi.$$

- *π is known as **risk-neutral measure** Q since $\sum_{i=1}^{K} \pi_i = 1, \pi_i \geq 0 \forall i$.*
- *Define expectation operator $E_Q[V] : \mathbb{R}^K \to \mathbb{R}$ as*

$$E_Q[V] = \sum_{i=1}^{K} \pi_i V(\omega_i),$$

then for the existing assets $V_1, ..., V_2$,

$$p_i = \frac{1}{1+r} E_Q[V_i].$$

- *For any other assets with payoff vector $V_B \in \mathbb{R}^K$, its no-arbitrage price is*

$$p_B = \frac{1}{1+r} E_Q[V_i]$$

Proof. Theorem 1.3.1 □

Remark 1.3.6 (risk-neutral measure has to be equivalent to original real-world measure).

- We are implicitly requiring that $\pi(\omega_i) \geq 0, \forall \omega \in \Omega$. Moreover, let the real probability be P, the risk-netural measure π, should satisfy $P(\omega) = 0 \Rightarrow \pi(\omega) = 0; \pi(\omega) = 0 \to P(\omega) = 0, \forall \omega \in \Omega$.

Remark 1.3.7 (existence and uniqueness).

1. Under no-arbitrage condition, the risk-neutral measure must exist. However, it might not be unique.

Remark 1.3.8 (Relation with state price in Theorem 1.3.1).

- We can also apply *Theorem* 1.3.1 to a market with risk-free market, let π be the state price, then $1 = e^{-r} \sum_i \pi_i e^r \implies \sum_i \pi_i = 1$.
- Therefore, we can understand as **risk-neutral measure is the discounted state price**.

1.3.2.2 *Computing risk-netural measure*

Remark 1.3.9 (how to compute risk-neutral measure). Suppose currently thare N actively traded assets and the market has no-arbitrage opportunties. By Theorem 1.3.3, there exists an risk-neutral measure. The risk-netrual measure π is about solving the following linear equation:

$$p = \frac{1}{1+r} D\pi.$$

Since the no-arbitrage condition ensures that the linear equation is consistent(or equivalently, p is in the column space of D), then we have the following situations:

- D has full column rank, then the risk-neutral measure is unique
- D has no full column rank(one situation is $K > N$), then there are infinitely many risk-netural measure.

Example 1.3.2 (Martingale measure for two state spot market). Consider a discrete stock price process consists of S_0 and S_T. We denote the discounted stock price as:

$$S_0^* = S_0, S_T^* = S_T/(1+r)$$

In the martingale measure, we have

$$S_0^* = E_{\mathbb{P}^*}[S_T^*]$$

In the case that the sample space of S^T has only two states, we have

$$S_0 = (1+r)^{-1}(Q(S^u)S^u + Q(S^d)S^d)$$

Solve it and we have

$$Q(S^u) = \frac{(1+r)S_0 - S^d}{S^u - S^d}$$
$$Q(S^d) = \frac{S^u - (1+r)S_0}{S^u - S^d}$$

Now a call option on this stock on the expiry date T will have

$$C_0(1+r) = \mathbb{E}_Q(C^T)$$

Example 1.3.3. [4, p. 10] Consider a market with only one risky asset and one money account. The current discounted price is 5. The one period discounted payoff of the risky asset are 4 and 6. To solve the risk neutral measure, we have

$$5 = 6Q(\omega_1) + 4Q(\omega_2), 1 = Q(\omega_1) + Q(\omega_2).$$

We can get

$$Q(\omega_1) = Q(\omega_2) = 1/2.$$

Example 1.3.4 (infinitely many risk-neutral measure). [4, p. 12] Consider a market with one risky asset and one money account. The current discounted price is 5. The one period discounted payoff of the risky asset are 3, 4 and 6. To solve the risk neutral measure, we have

$$5 = 6Q(\omega_1) + 4Q(\omega_2) + 3Q(\omega_3), 1 = Q(\omega_1) + Q(\omega_2) + Q(\omega_3).$$

We can get the solution

$$Q(\omega_1) = \lambda, Q(\omega_2) = 2 - 3\lambda, Q(\omega_3) = -1 + 2\lambda, 1/2 < \lambda < 2/3.$$

Example 1.3.5 (non-existence of risk-neutral measure). [4, p. 12] Consider a market with two risky asset and one money account. The current discounted price is 5 and 10. The one period discounted payoff of the risky asset are (6,6,4) and (12,8,8). To solve the risk neutral measure, we have

$$\begin{aligned}
5 &= 6Q(\omega_1) + 6Q(\omega_2) + 4Q(\omega_3) \\
10 &= 12Q(\omega_1) + 8Q(\omega_2) + 8Q(\omega_3) \\
1 &= Q(\omega_1) + Q(\omega_2) + Q(\omega_3)
\end{aligned}$$

We can get the unique solution

$$Q(\omega_1) = 1/2, Q(\omega_2) = 0, Q(\omega_3) = 1/2.$$

We claim that there must exist an arbitrage opportunity.

1.3.3 Numeriares and other measures for pricing

1.3.3.1 *Pricing in arbitrage numeriare*

Setup of general finite-state market with risk-free asset

- A probability space (Ω, F, P) describing the world.
- A market in which N assets, labeled $A_1, A_2, ..., A_N$, are freely traded. **Assume that one of these, say A_1, is risk-free.**
- Uncertainty about the behavior of the market is encapsulated in a finite-sized sample space Ω of K possible market scenarios or sample points, labeled $\omega_1, \omega_2, ..., \omega_K$. There

is an N by K payoff matrix $D \in \mathbb{R}^{N\times K}$ with entries $V_j(\omega_i)$ such that, in scenario i, the value of the asset A_j at after one-period is $V_j(\omega_i)$. The payoff matrix D is given by

$$\begin{pmatrix} V_1(\omega_1) & V_1(\omega_2) & \cdots & V_1(\omega_K) \\ V_2(\omega_1) & V_2(\omega_2) & \cdots & V_2(\omega_K) \\ \vdots & \vdots & \ddots & \vdots \\ V_N(\omega_1) & V_N(\omega_2) & \cdots & V_N(\omega_K) \end{pmatrix} = \begin{pmatrix} 1+r & 1+r & \cdots & 1+r \\ V_2(\omega_1) & V_2(\omega_2) & \cdots & V_2(\omega_K) \\ \vdots & \vdots & \ddots & \vdots \\ V_N(\omega_1) & V_N(\omega_2) & \cdots & V_N(\omega_K) \end{pmatrix},$$

where r is the risk-free return.

-
- We denote a portfolio vector by $\theta \in \mathbb{R}^N$
- The prices of the N assets at current time is represented by a price vector $p \in \mathbb{R}^N$. In particularly, $p_1 = 1$.

Proposition 1.3.4 (Fundamental theorem of asset pricing in a world with risk-free asset). *Consider a single-period finite state market model represented by payoff matrix $D \in \mathbb{R}^{N\times K}$. Further let the prices of the N assets at current time be represented by a price vector $p \in \mathbb{R}^N$. It follows that*

- *The market model has no arbitrage opportunities if and only if there exists a nonnegative vector $\pi \in \mathbb{R}^K$ such that the price of each asset i is given as*

$$p_i = \frac{1}{1+r}\sum_{j=0}^{K} \pi_j D_{ij};$$

or in matrix form

$$p = \frac{1}{1+r} D^T \pi.$$

- *Define π^B as*

$$\pi_j^B \triangleq \frac{\pi_j}{p_B} V_{B,j};$$

*π is known as **martingale measure** associated with numeriare B since $\sum_{i=1}^K \pi_i = 1$.*
- *Define expectation operator $E_B[V] : \mathbb{R}^K \to \mathbb{R}$ as*

$$E_B[V] = \sum_{i=1}^{K} \pi_i^B V(\omega_i),$$

then for the existing assets $V_1, ..., V_2$,

$$\frac{p_i}{p_B} = E_Q[\frac{V_i}{V_B}].$$

- *For any other assets with payoff vector $V_S \in \mathbb{R}^K$, its no-arbitrage price is*

$$p_B = \frac{1}{1+r} E_Q[V_i]$$

Proof.

$$p_i = \sum_{i=1}^{K} \pi_j D_{ij}$$
$$p_i = \sum_{i=1}^{K} \pi_j \frac{D_{ij}}{V_{B,j}} \cdot V_{B,j}$$

For the numeriare B, we have

$$p_B = \sum_{j=1}^{K} \pi_j \frac{D_{B,j}}{V_{B,j}} \cdot V_{B,j}$$

therefore

$$p_B = \sum_{j=1}^{K} \pi_j V_{B,j}.$$

Define $\pi_j^B \triangleq \frac{\pi_j}{p_B} V_{B,j}$, then $\sum_{j=1}^{K} \pi_j^B = 1$.

To pricing, we have

$$\begin{aligned} p_i &= \sum_{i=1}^{K} \pi_j D_{ij} \\ &= \sum_{i=1}^{K} p_B \pi_j^B \frac{D_{ij}}{V_{B,j}} \\ &= p_B E_B[\frac{D_i}{V_B}] \end{aligned}$$

□

Remark 1.3.10. Risk-neutral measure is the martingale measure associated with money account.

1.3.4 Market completeness and pricing

1.3.4.1 *Market completeness*

Definition 1.3.3 (market completeness). *A finite-state market model is complete if an asset with arbitrary payoff vector $V^B \in \mathbb{R}^K$ can be replicated and price.*

Proposition 1.3.5 (Second fundamental theorem of asset pricing). *Consider a no-arbitrage finite-state market model. Suppose there exists a set of actively traded assets whose prices vector p and the associated payoff matrix D are known. The following are equivalent:*

- *the market is complete if and only if D has the row space of dimensionality of K (the number of linearly independent rows is K).*
- *the market is complete if and only if has it has a **unique** risk-neutral measure π such that*

$$p = \frac{1}{1+r} D\pi.$$

Proof. (1) If its row space is of full rank, then for an asset with payoff vector $V \in \mathbb{R}^K$ can be replicated and price because $V \in \mathcal{R}(D^T)$ (Theorem 1.3.2); therefore the linear equation solution must be unique.

(2) (a)When the row space dimensionality equals K, the matrix D must be full column rank(note that the column rank is the number of linearly independent columns, which equals the number of linearly independent rows, see **??**).(b) When the linear equation $p = \frac{1}{1+r} D\pi$ has a **unique** solution, D must have full column rank; because row rank equals column rank(**??**), we get (1). □

Remark 1.3.11 (quick-check necessary condition for completeness)**.** The number of assets N should be greater than the number of states K; otherwise the state space cannot be spanned(or more precisely, the row space dimensionality cannot be greater than K).

Example 1.3.6*.* Consider there are only two assets in the market, a bond B and a stock S. If there are three market scenarios in $T = 1$, then we cannot **uniquely price** a derivative given the price of B and S at time 0.

Example 1.3.7*.* Given the following one-period payoff matrix associated with a two-state market

$$D = \begin{bmatrix} 1 & 10.51.2 \end{bmatrix}.$$

D has two linearly independent rows, and the market is complete.

Example 1.3.8. Given the following one-period payoff matrix associated with a three-state market

$$D = \begin{bmatrix} 1 & 1 & 10.5 & 1.2 & 1.7 \end{bmatrix}.$$

D has two linearly independent rows, which is smaller than number of states; therefore the market is incomplete.

1.3.4.2 *Pricing in incomplete market*

Remark 1.3.12 (pricing in incomplete market)**.**

- A complete market gives a unique price for assets with all possible payoff vectors; however, the real-world market is often incomplete.
- In an incomplete market, we have the following pricing properties:
  - there are infinitely many risk-netural measures.
  - the no-arbitrage price for an asset with payoff vector $V \in \mathbb{R}^K$ is given by $p = \frac{1}{1+r}E_\pi[V]$.
  - for an asset that **can be replicated** by existing actively traded assets, its price is unique.
  - for an asset that **cannot be replicated** by existing actively traded assets, its price is not unique.

1.4 Multi-period finite-state market model

1.4.1 Market model setup

Definition 1.4.1 (multi-period finite-state market model).

- *A set of discrete time $t = 0, 1, 2, ..., T$.*
- *A finite sample space $\Omega = \{\omega_1, ..., \omega_K\}$.*
- *A probability measure P defined on Ω with $P(\omega) > 0, \forall \omega \in \Omega$.*
- *A filtration $\{cF_t, t = 0, ..., T\}$.*
- *A* ***stochastic money account process*** *$B(t), t = 0, ..., T$, where $B(0) = 1$ and $B(t) > 0$ almost surely.*

- *N risky stochastic security processes $S_i(t), t = 0, 1, ..., T, i = 1, 2, ..., N$, where $S_i(t) > 0, \forall i, t$ almost surely.*

1.4.2 Self-financing discrete-time trading strategies

Definition 1.4.2 (trading strategies).

- *A trading strategy $h = (h_0, h_1, ..., h_N)$ is a* ***vector of processes*** *where $h_i = \{h_i(t), i = 1, 2, ..., T\}$ representing the number of units of asset i carried from time $t-1$ to time t.*
- *$h_0(t)B(t-1)$ represents the amount of money invested in the money market account from time $t-1$ to time t.*
- *$h_i(t)S_i(t-1)$ represents the long position of $h_i(t)$ shares in asset S_i from time $t-1$ to time t.*
- *A negative value of $h_i(t)$ representing borrowing from the bank or short positions.*

Definition 1.4.3 (value process and gain process).

- *The value process $V = \{V(t), t = 0, 1, ..., T\}$ associated with the trading strategy h and the market model consists of*
  - *initial value*
  $$V(0) = h_0(1)B(0) + \sum_{i=1}^{N} h_i(1)S_i(0),$$
  - *the value at time $t = 1, 2, ..., T$*
  $$V(t) = h_0(t)B(t) + \sum_{i=1}^{N} h_i(t)S_i(t).$$
- *The gain process $G(t), t = 1, 2, ..., T$ is defined as*
  $$G(t) = \sum_{u=1}^{t} h_0(u)\Delta B(u) + \sum_{u=1}^{t}\sum_{i=1}^{N} h_i(u)\Delta S_i(u),$$
  where $\Delta B(u) = B(u) - B(u-1), \Delta S_i(u) = S_i(u) - S_i(u-1)$.
- *The single-period gain $\Delta G(t)$ is defined as*
  $$\Delta G(t) = h_0(t)(B(t) - B(t-1)) + \sum_{i=1}^{N} h_i(t)(S_i(t) - S_i(t-1)),$$
  such that $G(t) = \sum_{u=1}^{t} \Delta G(u)$.

Definition 1.4.4 (self-financing strategy). *A trading strategy h is called **self-financing** if for* $t = 1, 2, ..., T - 1$,

$$V(t) = h_0(t+1)B(t) + \sum_{i=1}^{N} h_i(t+1)S_i(t),$$

where the LHS represents the time t value of the portfolio just before any transactions take place at that time, while the RHS represents the time t value of the portfolio right after any transactions (i.e. before the portfolio is carried forward to $t+1$*).*

Lemma 1.4.1 (criterion for self-financing strategy). *The following are equivalent:*

- *The trading strategy h is self-financing.*
- *The trading strategy h satisfies*

$$(h_0(t+1) - h_0(t))B(t) + \sum_{i=1}^{N}(h_i(t+1) - h_i(t))S_i(t) = 0, t = 0, 2, ..., T-1.$$

- *The trading strategy h satisfies*

$$V(t) = V(t-1) + \Delta G(t), t = 1, 2, ..., T.$$

- *The trading strategy h satisfies*

$$V(t) = V(0) + G(t), t = 1, 2, ..., T.$$

That is, for a self-financing strategy, any change in the portfolio's value is due to a gain or loss in the investments.

Proof. (1) equivalent to (2): Based on definition, we have

$$\begin{aligned} V(t) &= h_0(t+1)B(t) + \sum_{i=1}^{N} h_i(t+1)S_i(t) \\ &= h_0(t)B(t) + \sum_{i=1}^{N} h_i(t)S_i(t) \end{aligned}$$

Simple algebra can lead to the result. (2) equivalent (3)

$$\begin{aligned}V(t-1)+\Delta G(t) =& h_0(t-1)B(t-1)+\sum_{i=1}^N h_i(t-1)S_i(t-1)\\ & h_0(t)(B(t)-B(t-1)+\sum_{i=1}^N h_i(t)(S_i(t)-S_i(t-1))\\ =& h_0(t)B(t-1)+\sum_{i=1}^N h_i(t)S_i(t-1)\\ & h_0(t)(B(t)-B(t-1)+\sum_{i=1}^N h_i(t)(S_i(t)-S_i(t-1))\\ =& h_0(t)B(t)+\sum_{i=1}^N h_i(t)S_i(t) = V(t)\end{aligned}$$

where we use (2) in the second line. (3) equivalent (4): Use the definition $G(t) = \sum_{i=1}^t \Delta G$.

Note that for $t = 1$, we have

$$\begin{aligned}V(0)+\Delta G(1) = & h_0(1)B(0)+\sum_{i=1}^N h_i(1)S_i(0)+\\ & h_0(1)(B(1)-B(0))+\sum_{i=1}^N h_i(1)(S_i(1)-S_i(0))\\ &= h_0(1)B(1)+\sum_{i=1}^N h_i(1)S_i(1)\\ &= V(1)\end{aligned}$$

□

Definition 1.4.5 (disocunted asset price process).

- *The* ***discounted price process*** $S_i^*(t), t = 0, ..., T, i = 1, ..., N$ *is defined by*

$$S_i^*(t) = \frac{S_i(t)}{B(t)}, t = 0, 1, ..., T.$$

- *The* ***discounted value process*** $V^*(t), t = 0, ..., T, i = 1, ..., N$ *is defined by*

$$V^*(0) = h_0(1) + \sum_{i=1}^N h_i(1)S_i^*(0),$$

and

$$V^*(t) = h_0(t) + \sum_{i=1}^{N} h_i(t) S_i^*(t).$$

- *The* ***discounted gain process*** $G^*(t), t = 1, ..., T, i = 1, ..., N$ *is defined by*

$$G^*(t) = \sum_{u=1}^{t} \sum_{i=1}^{N} h_i(u) \Delta S_i^*(u),$$

- *A strategy h is self-financing if and only if*

$$V^*(t) = V^*(0) + G^*(t), t = 1, ..., T.$$

- *A strategy h is self-financing if and only if*

$$h_0(t+1) - h_0(t) + \sum_{i=1}^{N} (h_i(t+1) - h_i(t)) S_i^*(t)$$

1.4.3 First fundamental theorem of asset pricing

1.4.3.1 *Arbitrage and self-financing*

Definition 1.4.6 (arbitrage). *In the multi-period finite-state market model, an arbitrage opportunity exists if there is a self-financing strategy h whose value satisfying*

1. $V(0) = 0$;
2. $V(T) \geq 0$;
3. $Pr(V(T) > 0) > 0$.

More generally, the arbitrage condition can be

1. $V(0) = V_0$;
2. $V(T) \geq V_0$;
3. $Pr(V(T) > V_0) > 0$.

where V_0 is a deterministic constant.

Proposition 1.4.1 (arbitrage opportunity, change of numeraire, and change of measure).

- *A self-financing strategy h is an arbitrage opportunity under a probability measure P* ***if and only if*** *its discounted value V^* satisfying*

1. $V^*(0) = 0$;
2. $V^*(T) \geq 0$;
3. $Pr(V^*(T) > 0) > 0$.

- *A self-financing strategy h is an arbitrage opportunity under a probability measure P* ***if and only if*** *h is an arbitrage opportunity under an* ***equivalent*** *probability measure Q.*

Proof. (1) Note that because $B(t) > 0$ almost surely, then

1. $V(0)/B(0) = 0 \Leftrightarrow V(0) = 0$;
2. $V(T)/B(T) \geq 0 \Leftrightarrow V(0) = 0$;
3. $Pr(V(T)/B(T) > 0) > 0 \Leftrightarrow Pr(V(T) > 0) > 0$ since $V(T)/B(T) > 0 \Leftrightarrow V(T) > 0$.

(2) Let h be an arbitrage opportunity under P. Let $\Omega' \subseteq \Omega$ be a set such that $V(T)(\omega) > 0, \omega \in \Omega'$. Then

$$P(\Omega') > 0 \Leftrightarrow Q(\Omega') > 0,$$

due to the equivalence of P and Q. □

Note 1.4.1 (implication).

- The first part of the theorem simply says that an arbitrage is always an arbitrage no matter which numeraire we are using.
- The second part of the theorem simply says that an arbitrage is always an arbitrage under all equivalent probability measure.
- This theorem gives us the latitude to find the most convenient numeriare and measure to pricing assets. As long as under the chosen measure and numeriare no arbitrage exists, then no arbitrage exists in the real world.

Lemma 1.4.2 (martingale value process is not an arbitrage). *Let $V(t)$ be a martingale under a probability measure P with $V(0) = V_0$, where V_0 is a deterministic constant. Then $V(t)$ is not an arbitrage under any equivalent probability measure.*

Proof. If $V'(T) \geq V_0, P(V(T) > V_0) > 0$, we cannot have $E_P[V(t)] = V_0$. Therefore V cannot be an arbitrage. □

1.4.3.2 *First fundamental theorem*

Lemma 1.4.3 (self-financing strategy of martingale price processes cannot be arbitrage). *Let $S_1(t), S_2(t), ..., S_N(t)$ be stochastic price process of assets.*

- *If $S_1, ..., S_N(t)$ are martingale under the probability measure P, then under any equivalent measure any self-financing strategy involving buying and selling these assets cannot be arbitrage.*
- *If the discounted price processes under some numeraire $B(t)$ such that $S_1 * (t), ..., S_N^*(t)$ ($S_i^*(t) = S_i(t)/B(t)$), then under any equivalent measure any self-financing strategy involving buying and selling these assets cannot be arbitrage.*

Proof. (1)We are trying to show that the value processes are martingales:

$$\begin{aligned}
E_P[V(0) + G(t)|\mathcal{F}_s] &= V(0) + E_P[\sum_{u=1}^{t}\sum_{i=1}^{N} h_i(u)\Delta S_i(u)|\mathcal{F}_s] \\
&= V(0) + E_P[\sum_{u=1}^{s}\sum_{i=1}^{N} h_i(u)\Delta S_i(u)|\mathcal{F}_s] \\
&+ E_P[\sum_{u=s+1}^{t}\sum_{i=1}^{N} h_i(u)\Delta S_i(u)|\mathcal{F}_s] \\
&= V(0) + E_P[\sum_{u=1}^{s}\sum_{i=1}^{N} h_i(u)\Delta S_i(u)|\mathcal{F}_s] \\
&= V(s)
\end{aligned}$$

where we use the fact that

$$E_P[\Delta S_i(u)|\mathcal{F}_s] = E_P[\Delta S_i(u+1) - S_i(u)|\mathcal{F}_s] = 0, \forall u \geq s.$$

We know that a martingale value process cannot be arbitrage under any equivalent meausre(Lemma 1.4.2). (2) From (1) we know that discounted value process $V^*(t)$ is a martingale that cannot satisfy

1. $V^*(0) = V_0^*$;
2. $V^*(T) \geq V_0^*$;
3. $Pr(V^*(T) > V_0^*) > 0$.

Using Theorem 1.4.1, it is easy to see that $V(t) = B(t)V^*(t)$ cannot satisfy

1. $V(0) = V_0 B(0)$;
2. $V(T) \geq V_0 B(T)$;
3. $P(V(T)B(T) > V_0 B(T)) > 0$.

therefore h is not an arbitrage under P or any equivalent measure. □

Proposition 1.4.2 (first fundamental theorem of asset pricing). *In the multi-period finite-state market model, no arbitrage opportunity exists if and only if there exists a probability* *measure Q, known as risk-neutral measure, with $Q(\omega) > 0, \forall \omega$, such that every discounted price $S_i^*(t), t = 0, 1, ..., T$ is a martingale under measure Q.*

Proof. (1) (martingale measure implies no arbitrage) (2) (no arbitrage implies existence of martingale)

□

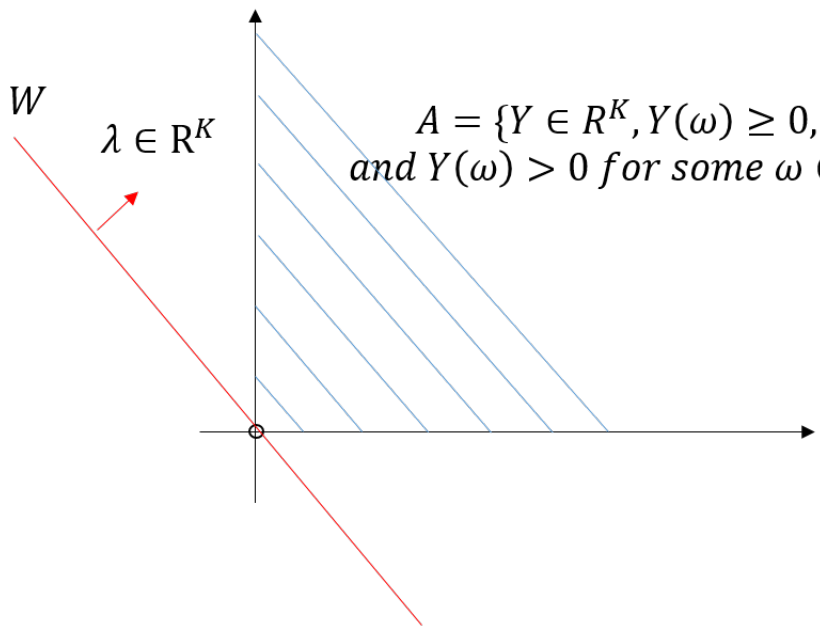


Figure 1.4.1: Demonstration of First Fundamental Theorem.

Lemma 1.4.4 (arbitrage connection between multiperiod and single period market model).

- *A multi-period finite-state market model is free of arbitrage if and only if all its constituent single-period finite-state market model is free of arbitrage.*
- *Under the equivalent martingale measure Q associated with a multi-period finite-state market model, for any single period n,*

$$S^*(n) = E_Q[S^*(n+1)|\mathcal{F}_n].$$

Proof. (1)(a) If single-period market model has arbitrage self-financing strategy, we can simply use it in the multi-period market model to get arbitrage. (b) If the □

1.4.3.3 *Assets with dividends*

Proposition 1.4.3 (first fundamental law for assets with dividends). *Let $S_1(t), S_2(t), ..., S_N(t), t = 0, 1, ..., T$ be stochastic price process of assets. Let $D_1(t), D_2(t), ..., D_N(t), t = 0, 1, ..., T$ be the dividend payment of asset i at time t. Then the market model has no arbitrage opportunities if and only if there exists a probability measure Q, known as risk-neutral measure, with $Q(\omega) > 0, \forall \omega$, such that for every i and t,*

$$S_i^*(t) = E_Q[\sum_{s=t+1}^{\tau} \frac{D(s)}{B(s)} + S_i^*(\tau)|\mathcal{F}_t], \forall \tau > t.$$

Explicitly,

$$S_i^*(t) - \sum_{s=t+1}^{\tau} E_Q[\frac{D(s)}{B(s)}|\mathcal{F}_t] = E_Q[S_i^*(\tau)|\mathcal{F}_t].$$

Proof. From the results for non-dividend paying assets(Theorem 1.4.2,Lemma 1.4.4) and $t = \tau$, we have

$$\frac{S(\tau-1)}{B(\tau-1)} = E_Q[\frac{S(\tau)}{B(\tau)} + \frac{D(\tau)}{B(\tau)}|\mathcal{F}_{\tau-1}].$$

Continuing forward, we have

$$\begin{aligned}
\frac{S(\tau-2)}{B(\tau-2)} &= E_Q[\frac{S(\tau-1)}{B(\tau-1)} + \frac{D(\tau-1)}{B(\tau-1)}|\mathcal{F}_{\tau-2}] \\
&= E_Q[E_Q[\frac{S(\tau)}{B(\tau)} + \frac{D(\tau)}{B(\tau)}|\mathcal{F}_{\tau-1}] + \frac{D(\tau-1)}{B(\tau-1)}|\mathcal{F}_{\tau-2}] \\
&= E_Q[\frac{S(\tau)}{B(\tau)} + \frac{D(\tau)}{B(\tau)} + \frac{D(\tau-1)}{B(\tau-1)}|\mathcal{F}_{\tau-2}]
\end{aligned}$$

Continuing forward to t and we will get the result. □

1.4.3.4 *Application*

Lemma 1.4.5 (martingale properties of self-financing strategies under risk-neutral measure). *Assume no arbitrage opportunity exists in the multi-period finite-state market model. Under the risk-neutral measure Q,*

- *the* ***discounted gain process*** *of every self-financing strategy h is martingale.*
- *the* ***discounted value process*** *of every self-financing strategy h is martingale.*
- *every self-financing strategy h admits no arbitrage.*

Proof. (1)

$$\begin{aligned} E_Q[G^*(t)|\mathcal{F}_s] &= E_Q[\sum_{u=1}^{t}\sum_{i=1}^{N} h_i(u)\Delta S_i^*(u)|\mathcal{F}_s] \\ &= E_Q[\sum_{u=1}^{s}\sum_{i=1}^{N} h_i(u)\Delta S_i^*(u)|\mathcal{F}_s] \\ &+ E_Q[\sum_{u=s+1}^{t}\sum_{i=1}^{N} h_i(u)\Delta S_i^*(u)|\mathcal{F}_s] \\ &= E_Q[\sum_{u=1}^{s}\sum_{i=1}^{N} h_i(u)\Delta S_i^*(u)|\mathcal{F}_s] = G^*(s) \end{aligned}$$

where we use the fact that

$$E_Q[\Delta S_i^*(u)|\mathcal{F}_s] = E_Q[\Delta S_i^*(u+1) - S_i^*(u)|\mathcal{F}_s] = 0, \forall u \geq s.$$

Therefore $G^*(t)$ is a Q-martingale. (2) Note that $V^*(t) = V^*(s) + G^*(t) - G^*(s)$, we have

$$E_Q[V^*(t)|\mathcal{F}_s] = E_Q[V^*(s) + G^*(t) - G^*(s)|\mathcal{F}_s] = V^*(s),$$

where we use $E_Q[G^*(t) - G^*(s)|\mathcal{F}_s] = 0$ from (1). (3) Because (2), h is not an abitrage due to Theorem 1.4.1Lemma 1.4.3. (Lemma 1.4.3 shows that the discounted version V^* satisfies the no-arbitrage condition)

□

Proposition 1.4.4 (no-arbitrage price of any attainable claim). *Assume no arbitrage opportunity exists in the multi-period finite-state market model. Let h be a self-financing strategy replicate an attainable claim Y. Let $V(t)$ be the value process associated with h. Then the* ***no-arbitrage price process*** *Y is given by*

$$V_Y(t) = B(t)E_Q[\frac{Y}{B(T)}|\mathcal{F}_t].$$

Proof. Our goal is to prove that when the price process of all attainable claims are defined in this way, **no self-financing strategies involving buying and selling these attainable claims and other assets are arbitrage opportunities.**

Becuase the discounted price proccesses of Y and other assets are martingales, then no arbitrage exists due to Lemma 1.4.3. □

Note 1.4.2 (uniqueness, no-arbitrage price and intrinsic price).

- In the theorem, we give one possible form of price process of attainable claims; we can only say these price processes admit no arbitrage; we cannot ensure its uniqueness.
- We should distinguish the no-arbitrage prices from intrinsic prices, which are the results of market forces equilibrium. See Definition 1.1.1 for more details.

Remark 1.4.1 (pitfalls in 'other possible forms' of no-arbitrage price). It is tempting to define the price process of any attainable claim Y by

$$V(t) = E_P[Y|\mathcal{F}_t],$$

such that $V(t)$ is a martingale because

$$E_P[V(t)|\mathcal{F}_s] = E_P[E_P[Y|\mathcal{F}_t]|\mathcal{F}_s] = V(s).$$

However, such price is not arbitrage free. Let h be the self-financing strategy such that $V_h(T) = Y$. In general,

$$V_h(0) = B(0)E_Q[V_h(T)/B(T)] \neq V_Y(0) = E_P[Y]$$

which violates law of one price and introduce arbitrage opportunities.

1.4.3.5 *Forwards*

Proposition 1.4.5 (forward price under risk-neutral measure). *Consider a security price process S with its dividend process D. Let $F(t,\tau)$ be the forward price process with delivery date τ.*

- *Then under risk-neutral measure Q,*

$$F(t,\tau) = E_Q[\frac{S(\tau)B(t)}{B(\tau)}|\mathcal{F}_t]/DF(t,\tau),$$

where

$$DF(t,\tau) = E_Q[\frac{B(t)}{B(\tau)}|\mathcal{F}_t].$$

- *Under the equivalent martingale measure Q_τ associated with numeraire $DF(t,\tau)$, we have*

$$F(t,\tau) = E_\tau[S(\tau)|\mathcal{F}_t] = E_\tau[F(\tau,\tau)|\mathcal{F}_t],$$

and

$$E_\tau[F(t,\tau)] = F(0,\tau).$$

That is $F(t,\tau)$ is the conditional expectation process of $F(\tau,\tau)$ under measure Q_τ

- *If the bank account process $B(t)$ is deterministic, then*

$$F(t,\tau) = E_Q[S(\tau)|\mathcal{F}_t] = \frac{B(\tau)}{B(t)}S_i(t) - \sum_{s=t+1}^{\tau} D(s)\frac{B(\tau)}{B(s)}|\mathcal{F}_t].$$

- *If the bank account process $B(t)$ is deterministic, then $F(t,\tau)$ is a martingale and*

$$F(t,\tau) = E_Q[S(\tau)|\mathcal{F}_t] = E_Q[F(\tau,\tau)|\mathcal{F}_t],$$

and

$$E_Q[F(t,\tau)] = E_Q[S(\tau)] = F(0,\tau),$$

That is $F(t,\tau)$ is the conditional expectation process of $F(\tau,\tau)$.

Proof. (1) Because the time t value of the forward contract is zero, using pricing formula Theorem 1.4.4, we have

$$E_Q[\frac{B(t)}{B(\tau)}(S(\tau) - F(t,\tau)|\mathcal{F}_t] = 0,$$

which gives

$$F(t)E_Q[\frac{B(t)}{B(\tau)}|\mathcal{F}_t] = E_Q[\frac{B(t)}{B(\tau)}S(\tau)|\mathcal{F}_t].$$

(2) We use the dynamics of $S(t)$ under risk-netrual measure given in Theorem 1.4.3. (3) We use the boundary condition $S(\tau) = F(\tau,\tau)$ and conditional expectation process is martingale(Theorem 1.6.3). □

Proposition 1.4.6 (value of forward contract after initiation). *Let $F(t,T)$ be the forward price of an asset at time t with delivery time T. Consider a forward/futures contract initiated at t_0 with forward price $F(t_0,T)$. Let $V(t)$ be the value of the contract at time t. Denote $DF(t,\tau) = E_Q[\frac{B(t)}{B(\tau)}|\mathcal{F}_t]$.*

- *The* ***stochastic*** *value of the forward contract at time t_0+1 is*

$$V(t_0+1) = (F(t_0+1,T) - F(t_0,T))DF(t+1,T)$$

- *The* ***stochastic*** *value of the forward contract at time t is*

$$V(t) = (F(t,T) - F(t_0,T))DF(t,T)$$

Proof. (1) Using pricing formula Theorem 1.4.4 for the value $t_0 + 1$, we have

$$\begin{aligned} V(t_0+1) &= E_Q[\frac{B(t_0+1)}{B(T)}(S(T) - F(t_0,T)|\mathcal{F}_{t+1}] \\ &= E_Q[\frac{B(t_0+1)}{B(T)}S(T)|\mathcal{F}_{t+1}] - F(t_0,T)DF(t_0+1,T) \\ &= F(t_0+1,F)DF(t_0+1,T) - F(t_0,T)DF(t_0+1,T) \end{aligned}$$

(2) We can similarly prove other cases. □

1.4.3.6 *Futures*

Proposition 1.4.7 (futures price under risk-neutral measure). *Consider a security price process S with its dividend process D. Let $Fur(t,\tau)$ be the further price process with delivery date τ. Assume marking to market at every discrete time step.*

- *Then under risk-neutral measure Q,*
$$E_Q[\frac{Fur(t,\tau) - Fur(t-1,\tau)}{B(t)}|\mathcal{F}_{t-1}], t = 1,...,T.$$
- *If $B(t)$ is predictable, i.e., $B(t)$ is measurable with respect to $\mathcal{F}_{t-1}$, then $Fur(t,\tau)$ is a martingale and given by*
$$Fur(t,\tau) = E_Q[S(\tau)|\mathcal{F}_t] = E_Q[Fur(\tau,\tau)|\mathcal{F}_t],$$
and
$$E_Q[Fur(t,\tau)] = Fur(0,\tau).$$
- *If $B(t)$ is deterministic,*
$$F(t,\tau) = Fur(t,\tau),$$
that is, forward and futures price process are equivalent.

Proof. (1) Because of the marking to market convention, the time t payoff or value of the futures contract is $Fur(t,\tau) - Fur(t-1,\tau)$. Using pricing formula Theorem 1.4.4, we have

$$V(t-1) = B(t-1)E_Q[\frac{Fur(t,\tau) - Fur(t-1,\tau)}{B(t)}|\mathcal{F}_{t-1}], t = 1,...,T.$$

Setting $V(t-1) = 0$ due to marking to market, we get the result.

$$F(t)E_Q[\frac{B(t)}{B(\tau)}|\mathcal{F}_t] = E_Q[\frac{B(t)}{B(\tau)}S(\tau)|\mathcal{F}_t].$$

(2) When $B(t)$ is predictable, we have

$$0 = \frac{1}{B(t)} E_Q[Fur(t,\tau) - Fur(t-1,\tau)|\mathcal{F}_{t-1}] = 0.$$

That is,

$$E_Q[Fur(t,\tau)|\mathcal{F}_{t-1}] = Fur(t-1,\tau)$$

Therefore,

$$\begin{aligned} Fur(t,\tau) &= E_Q[Fur(t+1,\tau)|\mathcal{F}_t] \\ &= E_Q[E_Q[Fur(t+2,\tau)|\mathcal{F}_{t+1}]|\mathcal{F}_t] \\ &= E_Q[Fur(t+2,\tau)|\mathcal{F}_t] \\ &\cdots \\ &= E_Q[Fur(\tau,\tau)|\mathcal{F}_t] \end{aligned}$$

(3) When $B(t)$ is deterministic, from Theorem 1.4.5, we have

$$F(t,\tau) = Fur(t,\tau) = E_Q[S(\tau)|\mathcal{F}_t].$$

□

1.4.4 Complete market and second fundamental theorem

1.4.4.1 *Complete market*

Definition 1.4.7 (complete market). *A market is said to be* ***complete*** *if every claim in the market is attainable.*

1.4.4.2 *Second fundamental theorem*

1.4.4.3 *Complete vs. incomplete market*

Note 1.4.3 (pricing in complete vs. incomplete market)**.**

- If the model is complete then we know how to price all the contingent claims.
- If the model is incomplete, then we **only know how to price some of the contingent claims**, namely, all the attainable ones. Let $\mathcal{M}$ denote the set of all the risk-neutral measures. Then
$$p = B(0)E_Q[Y(T)/B(T)], \forall Q \in \mathcal{M},$$

that is, we **get the same price under all risk-neutral measures.**
- For the pricing of non-attainable claims in the incomplete market, we use optimization method to provide lower and upper bounds.

Definition 1.4.8 (no-arbitrage price range for non-attainable claims). *Assume the market has no arbitrage but is incomplete.*

- *The upper bound price is*

$$V_+(X) \triangleq \inf\{B_0 E_Q[Y/B_T] : Y \geq X, Y \text{ attainable}\}.$$

- *The lower bound price is*

$$V_-(X) \triangleq \sup\{B_0 E_Q[Y/B_T] : Y \leq X, Y \text{ attainable}\}.$$

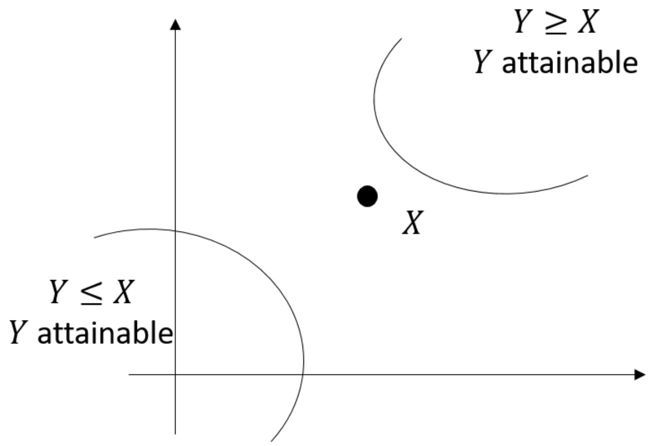


Figure 1.4.2: An illustration of the upper and lower bound price for non-attainable claims.

Remark 1.4.2 (arbitrage justification for lower and upper bound prices)**.**

- If X could be sold for a greater amount p than $V^+(X)$, then we should short X, and use the proceeds to buy Y, which has better payoff than X at time T, i.e., $Y(T) \geq X(T)$. Then at time T sell Y and buy X to cover the short position.
- If X could be bought for a smaller amount p than $V^-(X)$, then we should buy X and short Y. X has better payoff than Y at time T, i.e., $X(T) \geq Y(T)$. Then at time T sell X and buy Y to cover the short position.

Definition 1.4.9 (no-arbitrage price range for non-attainable claims, alternative). *Assume the market has no arbitrage but is incomplete. Let $\mathcal{M}$ denote the set of all the risk-neutral measures.*

- *The upper bound price is*

$$V_+(X) \triangleq \sup\{B_0 E_Q[X/B_T] : Q \in \mathcal{M}\}.$$

- *The lower bound price is*

$$V_-(X) \triangleq \inf\{B_0 E_Q[X/B_T] : Q \in \mathcal{M}\}.$$

- *In particular, if X is attainable, then $V_+(X) = V_-(X)$.*

Note 1.4.4 (equivalence of the two definitions). [4, p. 26]

- Note that the two definitions can be showed equivalent using linear programing duality theorem.
- The alternative definition is easier for calculating lower and upper bound prices.

Example 1.4.1. Consider a contingent claim X with payoff $(30, 20, 10)$. Consider the set of risk-neutral measure are parameterized by $Q = (q, 2-3q, -1+2q)$, where $1/2 < q < 2/3$. Assume zero interest rate. We have

- $E_Q[X] = 30q + 20(2-3q) + 10(-1+2q) = 30 - 10q$.
- $$V_+(X) = \sup_q 30 - 10q = 25.$$
- $$V_-(X) = \inf_q 30 - 10q = 70/3.$$

1.4.5 Extension to infinite states

1.5 Continuous-time market model I: hedging

Assumptions for this section:

- All assets are trade-able
- The risk free rate r are deterministic
- No transaction cost.
- The underlying asset pays no dividends.
- Short selling is permitted.

- The trading of assets can take place continuously in time and amount.

1.5.1 Single source of uncertainty

Lemma 1.5.1 (hedging one source of risk, constraints on parameters of dynamics with common source of risks). *[5, p. 656][6, p. 55] Consider a market with risk-free asset with short rate r. Consider values of two assets as stochastic processes given as:*

$$\frac{df_1}{f_1} = \mu_1 dt + \sigma_1 dz$$

and

$$\frac{df_2}{f_2} = \mu_2 dt + \sigma_2 dz$$

where z is the Wiener process,i.e., the only ***common source of uncertainty****. Under no-arbitrage condition, there exist a λ(called market price of risk) such that*

$$\frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2} = \lambda$$

And the stochastic processes f_1 and f_2 will be given as

$$\frac{df_1}{f_1} = (r + \sigma_1 \lambda)dt + \sigma_1 dz$$

and

$$\frac{df_2}{f_2} = (r + \sigma_2 \lambda)dt + \sigma_2 dz$$

Proof. If we construct the portfolio

$$\Pi = k_1 f_1 + k_2 f_2$$

where $k_1 = \sigma_2 f_2$ and $k_2 = -\sigma_1 f_1$. Then we have

$$d\Pi = k_1 df_1 - k_2 df_2 = (\mu_1 \sigma_2 f_1 f_2 - \mu_2 \sigma_1 f_1 f_2)dt = r\Pi dt.$$

That is, the portfolio Π is deterministic, and it must have growth rate equal to risk-free rate as required by no-arbitrary and law of one price(Corollary 1.2.1.2). Then, we have

$$\mu_1 \sigma_2 + \mu_2 \sigma_1 = \sigma_2 r + \sigma_1 r.$$

$\square$

Remark 1.5.1 (interpretation)**.**

- Assets have the common source of uncertainty have dependent growth rate. Because of the common uncertainty, they are correlated and they can be used to hedging each other to reduce risks.
- **If the asset prices do not evolve like this, there exists arbitrage opportunities**. That is, risk-free portfolio does not grow at risk-free rate.

Note 1.5.1 (constraints on parameters of dynamics with common source of risks)**.** Under the no-arbitrage assumption, law of one price(Theorem 1.2.1), (and assumptions about continuous trading, short selling, frictionless market etc,), two Ito processes with single common source of uncertainty have constraints on the drift parameters they have. This lemma says that, they have to satisfy

$$\frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2} = \lambda.$$

From mathematical point of view, any $\lambda \in \mathbb{R}$ is reasonable(no arbitrage exists); from economics point of view, any $\lambda > 0$ is reasonable, the large λ reflects investors are more risk-averse and require higher growth rate.

Example 1.5.1. Assume that there are two financial assets, both are dependent on a 90 day interest rate. The first instrument has an expected return 6% per year and a volatility of 20%. For the second instrument a volatility of 30% per year is assumed. Further more, the risk-free rate is 3% per year. The market price of risk for the first instrument is

$$\frac{0.06 - 0.03}{0.2} = 0.15.$$

Then for the second instrument, we estimate its expected return is

$$0.15 \times 0.3 + 0.03 = 0.075.$$

Note 1.5.2 (Derivation of Black-Scholes model and martingale pricing, alternative)**.** indexBlack-Scholes equation Let $V(S(t), t)$ be the value of the derivative as a function of the asset price $S(t)$ and time t. Assume $S(t)$ is governed by

$$dS(t) = \mu S(t)dt + \sigma S(t)dW_t$$

where W_t is the Brownian motion. Then $V(S_t, t)$ is governed by

$$dV = \sigma S \frac{\partial V}{\partial S} dW_t + (\mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t})dt$$

via Ito lemma(**??**)

Then from Lemma 1.5.1, there exists a λ such that

$$\mu = r + \lambda\sigma$$
$$(\mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t})/V = r + \lambda\sigma S \frac{\partial V}{\partial S}/V$$

Eliminate λ, we get the Black-Scholes equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

If we change our measure to a new measure Q(risk-neutral measure) such that $\lambda = 0$, then

$$dS(t) = rS(t)dt + \sigma S(t) dW_t$$
$$dV = rVdt + \sigma S \frac{\partial V}{\partial S} dW_t$$

That is, any derivative V will have drift rV no matter its functional form(assume V is a 'nice' function). **Therefore,**

$$e^{-rt}V_t$$

is a martingale, and justify the martingale pricing method

$$V(t) = E_Q[V(T)|\mathcal{F}_t].$$

Note 1.5.3 (When risks cannot be hedged). **Assume** an asset S governed by

$$dS = \mu S dt + \sigma S dW$$

where W is the Brownian motion. Assume the risk associated with W cannot be eliminated(for example, due to the lack of derivatives and assets sharing the same risks).

- If we can estimate μ and σ from the market data, then we can determine market price of risk $\lambda = (\mu - r)/\sigma$.

- In this single asset with unhedgeable risk, any μ value is reasonable, since the model itself will not admit arbitrage. However, for multiple assets with common risks, the drift parameters are constrained if no arbitrage is allowed(Lemma 1.5.1).

Lemma 1.5.2 (hedging one source of risk, constraints on parameters of dynamics with common source of risks, assets with dividends). *[5, p. 656][6, p. 55] Consider a market with risk-free asset with short rate r. Consider values of two assets as stochastic processes given as:*

$$\frac{df_1}{f_1} = \mu_1 dt + \sigma_1 dz$$

and

$$\frac{df_2}{f_2} = \mu_2 dt + \sigma_2 dz$$

where dz is the Wiener process,i.e., the only ***common source of uncertainty****. Further assume that the asset is paying continuous dividends with ratio q_1 and q_2. Under no-arbitrage condition, there exist a λ(called market price of risk) such that*

$$\frac{\mu_1 + q_1 - r}{\sigma_1} = \frac{\mu_2 + q_2 - r}{\sigma_2} = \lambda$$

And the stochastic processes f_1 and f_2 will be given as

$$\frac{df_1}{f_1} = (r - q_1 + \sigma_1\lambda)dt + \sigma_1 dz$$

and

$$\frac{df_2}{f_2} = (r - q_2 + \sigma_2\lambda)dt + \sigma_2 dz$$

Proof. (1) The most common proof is that we form a new security f^*, which the same as f except that all income produced by f is reinvested in f. f^* is related to f via

$$f^* = fe^{qt}, df^*/f^* = (\mu + q)dt + \sigma dz.$$

Then we can use analog of Lemma 1.5.1. (2) The most common proof above relies on a special portfolio. Now we can prove most general case: suppose we divide the dividend for two usage, ϕ for reinvestment and $q - \phi$ for money account m, then the total wealth $g = f^* + m$ at time t has dynamics:

$$dg = df^* + dm, df^* = (\mu + \phi q)f^* dt + \sigma f^* dz, dm = rmdt + f(1 - \phi)dt.$$

That is:

$$dg = (\mu + q)f^* + rmdt + \sigma f^* dz$$

From Lemma 1.5.1, we have

$$(\mu + q)f^* + rm = r(m + f^*) + \lambda\sigma f^* \implies (\mu + q - r) = \lambda\sigma.$$

□

Lemma 1.5.3 (linearity properties in no-arbitrage condition). *Consider a market with risk-free asset with short rate r. Consider values of two assets under* ***no arbitrage condition*** *having stochastic processes given by:*

$$\frac{df_1}{f_1} = r + \lambda\sigma_1 dt + \sigma_1 dz$$

and

$$\frac{df_2}{f_2} = r + \lambda\sigma_2 dt + \sigma_2 dz$$

where dz is the Wiener process,i.e., the only ***common source of uncertainty*** *and λ is the market price of risk. Then for arbitrary constant α and β,*

$$\frac{d(\alpha f_1 + \beta f_2)}{\alpha f_1 + \beta f_2} = (r + \lambda\sigma)dt + \sigma dz,$$

where

$$\sigma = \frac{\alpha\sigma_1 f_1 + \beta\sigma_2 f_2}{\alpha f_1 + \beta f_2}.$$

That is, the new asset $\alpha f_1 + \beta f_2$ also satisfies the no arbitrage condition.

Proof. We have

$$\begin{aligned} d\alpha f_1 &= (r + \lambda\sigma_1)\alpha f_1 dt + \sigma_1 \alpha f_1 dz \\ d\beta f_2 &= (r + \lambda\sigma_2)\beta f_2 dt + \sigma_2 \beta f_2 dz \end{aligned}$$

Add together and factor out the common factor $\alpha f_1 + \beta f_2$ can prove the result. □

Remark 1.5.2. This is the restatement of the linear pricing theorem(Theorem 1.2.2).

1.5.2 Multiple sources of uncertainty

Proposition 1.5.1 (hedging multiple sources of risks, constraints on parameters of dynamics). *[5, p. 659][6, p. 55] Consider a market with risk-free asset with short rate r. Consider n assets as stochastic processes given as:*

$$\frac{df_i}{f_i} = a_i dt + \sum_{j=1}^{m} b_{ij} dz_j, i = 1, 2, ..., n.$$

*where $z_j, j = 1, 2, ..., m$ are independent Wiener processes. If there is **no arbitrage**, then there exists (unnecessarily unique) $\lambda_1, ..., \lambda_m$*

$$a_i - r = \sum_{j=1}^{m} b_{ij} \lambda_j, j = 1, 2..., m.$$

Proof. Denote $A = (a_1, ..., a_n)^T$, $B_{ij} = b_{ij}$, and a portfolio vector $k = (k_1, ..., k_n)^T$. To construct a risk-free portfolio, we have

$$k^T B = 0.$$

This risk-free portfolio should grow at risk-free rate r_f, we have

$$k^T A = k^T r_f.$$

Combine $k^T B = 0, k^T(A - r_f) = 0$, we know that $A - r_f = B\lambda$. (from $k^T B = 0$, we know that $k \in \mathcal{N}(B)$; from $k^T(A - r_f)$, we know that $(A - r_f) \in \mathcal{N}(B)^{\perp} = \mathcal{R}(B)$). □

Remark 1.5.3 (trivial solution for portfolio vector k). In our proof, if B is a tall and thin matrix(no fewer sources of uncertainty than number of assets, for example, 2 stocks with 2 sources of uncertainty; note that when there are more stocks than uncertainty sources, it is guaranteed to have infinitely many solutions for k.), then k might only have the trivial solution of $k = 0$. However, this does not undermine validity of our conclusion that $(A - r_f) \in \mathcal{N}(B)^{\perp} = \mathcal{R}(B)$.

Note 1.5.4 (existence and uniqueness of λ, market completeness). We know that the A, B, λ are connected by

$$B\lambda = (A - r_f).$$

Assume A, r_f and B are known, the solutions to λ has the following situations.

- The linear equation is **inconsistent**, then λ **does not exist. Then there exists arbitrage in the model.**
- The linear equation has one unique solution, then the market is complete and the price is unique under no-arbitrage condition.
- The linear equation has infinitely many solutions, then the market is incomplete and the price is not unique under no-arbitrage condition. The non-unique prices

means that there exists other sets of model parameters$(A, B, A = r_f + B\lambda)$ such that no arbitrage can happen. The unique price situation means that only current model parameters can satisfy the no-arbitrage condition.
- The non-uniqueness situation usually occur when $m \geq n$; i.e. not all the uncertainty can be eliminated. The most simple case is that there is only 1 risky asset in market, then its drift parameter can be any value and the model will not admit arbitrage.

Remark 1.5.4 (interpretation of λ). If $\lambda_i \sigma_i > 0$, then the risk associated with z_i cannot be eliminated, and the investor require a positive excess return for the risk associated with z_i;If $\lambda_i \sigma_i = 0$, then the risk associated with z_i can be eliminated, and the investor would not require compensation for the risk associated with z_i; If $\lambda_i \sigma_i < 0$, then the investor is willing to accept lower return since the risk associated with z_i can help them reduce total risks;

Remark 1.5.5 (market price of risk in CAPM).

- In the CAPM model, where we can only hedge non-systematic risk by portfolio combination, the market price of the risk is

$$\frac{E[r_m] - r_f}{\sigma_m},$$

  and the return of **efficient** portfolio is

$$\frac{E[r] - r_f}{\sigma} = \frac{E[r_m] - r_f}{\sigma_m}$$

  Investors can only get compensated by taking systematic risks.
- More generally, consider a world of multiple assets described by SDE with multiple sources of uncertainties. **If there exists a portfolio having no risk, then its growth rate must be risk-free rate**.
- Every market price of risk is corresponding to some probability measure. Some value of the market price of risk corresponds to the real world probability measure and the growth rates of security prices that are observed in practice.

Proposition 1.5.2 (hedging multiple sources of risks, constraints on parameters of dynamics, assets with dividends). *[5, p. 659][6, p. 55] Consider a market with risk-free asset with short rate r. Consider n assets as stochastic processes given as:*

$$\frac{df_i}{f_i} = a_j dt + \sum_{i=1}^{n} b_{ij} dz_i, j = 1, 2, ..., n.$$

where $z_i, i = 1, 2, ..., m$ are independent Wiener processes. ***Further assume each asset has a continuous dividend*** $q_i, i = 1, ..., m$***.*** *If there is* ***no arbitrage****, then there exists (unnecessarily unique)* $\lambda_1, ..., \lambda_m$

$$a_i + q_i - r = \sum_{j=1}^{m} \lambda_i b_{ij}, j = 1, 2..., m.$$

Proof. See Lemma 1.5.2. □

1.5.3 Case study on market risk

Proposition 1.5.3 (Law of one price on market risks). *[7, p. 25] Identify unavoidable risks should have identical expected returns.*

Lemma 1.5.4 (asset dynamics in a world with a few uncorrelated stocks and a riskless bond). *Consider a market with risk-free asset with short rate r. Consider n assets characterized by stochastic processes given as:*

$$\frac{df_i}{f_i} = a_i dt + b_i dz_i, j = 1, 2, ..., n.$$

where $z_i, i = 1, 2, ..., n$ are independent Wiener processes. If there is ***no arbitrage****, then there exists (unnecessarily unique) $\lambda > 0$ such that*

$$a_i - r = b_i \lambda, \forall i.$$

Proof. From Theorem 1.5.1, we know that there exists $\lambda_1, \lambda_2, ..., \lambda_n$ such that

$$a_i - r = b_i \lambda_i, \forall i.$$

The law of one price on market risk(Theorem 1.5.3) requires these λ_i should be equal to each other. If one asset offer a smaller growth rate at the same risk level, the investors will view this asset is undervalued, then the buying force will push up the growth rate.

We further requires $\lambda > 0$ because investors are risk averse(no one is willing to take risks unless being rewarded.) □

Lemma 1.5.5 (asset dynamics in a world with infinitely many uncorrelated stocks and a riskless bond). *Consider a market with risk-free asset with short rate r. Consider infinitely many assets characterized by stochastic processes given by:*

$$\frac{df_i}{f_i} = a_i dt + b_i dz_i, j = 1, 2, ..., n.$$

where $z_i, i = 1, 2, ..., n$ are independent Wiener processes. If there is ***no arbitrage****, then*

$$a_i - r = 0, \forall i.$$

That is, the individual asset dynamics is

$$\frac{df_i}{f_i} = r dt + b_i dz_i, i = 1, 2, ..., n.$$

Or equivalently, ***the market price of idiosyncratic risk is zero.***

Proof. Via linear combination, we can create a portfolio consisting of all the assets such that the risk is zero(due to central limit theorem). This portfolio will have zero risk and therefore requires $\lambda = 0$. Here is the formal proof.

We know that

$$df_i = (r + \lambda b_i) f_i dt + b_i f_i dz_i, b_i > 0$$

then for this special portfolio, we have

$$\begin{aligned} d(\sum_{i=1}^{N} f_i) &= (r + \lambda \frac{\sum_{i=1}^{N} b_i f_i}{\sum_{[i=1]}^{N} f_i}) \sum_{[i=1]}^{N} f_i dt + \sum_{[i=1]}^{N} b_i f_i dz_i \\ &= (r + \lambda \frac{\sum_{i=1}^{N} b_i f_i}{\sum_{[i=1]}^{N} f_i}) \sum_{[i=1]}^{N} f_i dt + \frac{\sum_{i=1}^{N} b_i^2 f_i^2}{\sum_{i=1}^{N} f_i} \sum_{i=1}^{N} f_i dZ \end{aligned}$$

We can show that (assume $\{b_i\}, \{f_i\}$ is unformly bounded away from 0 and ∞)

$$\frac{\sum_{i=1}^{N} b_i^2 f_i^2}{\sum_{i=1}^{N} f_i} \leq \frac{\max\{b_i f_i\}\sqrt{N}}{\min\{f_i\} N} \to 0, \ as\ N \to \infty$$

and

$$\frac{\sum_{i=1}^{N} b_i f_i}{\sum_{[i=1]}^{N} f_i} \geq \min\{b_i\} > 0.$$

Therefore, as $N \to \infty$, the portfolio has no risk and its growth rate needs to be r. Therefore, $\lambda = 0$. $\square$

Remark 1.5.6 (intuition and interpretation)**.**

- Suppose $\lambda > 0$, then we will just buy the market portfolio, which has no risk but growth rate greater than bond. Then we will borrow infinite money to buy this portfolio and push up r.
- This example clearly shows hedgeable risks will not be rewarded.

Remark 1.5.7 (compared with finite asset case)**.** The finite asset case has a $\lambda > 0$ depending on N(which determines how much can be diversified) and investors risk aversion.

Lemma 1.5.6 (asset dynamics in a world with infinitely many all simultaneously correlated with the entire market, and a riskless bond). *Consider a market with risk-free asset with short rate r. Consider infinitely many assets characterized by stochastic processes given by:*

$$\frac{df_i}{f_i} = a_i dt + \rho_i b_i dW + \sqrt{1-\rho_i} b_i dz_i, b_i > 0 \forall i = 1, 2, ..., n.$$

where $z_i, i = 1, 2, ..., n$, and W are independent Wiener processes. Assume there is no arbitrage opportunities in the market. Suppose we form a market portfolio $f_M = \sum_{i=1}^{\infty} f_i$.

It follows that

- *the market portfolio has dynamics*

$$\frac{df_M}{f_M} = (r + \lambda b_M)dt + b_M dW, b_M = \frac{\sum_{i=1}^{\infty} b_i f_i}{\sum_{i=1}^{\infty} f_i}.$$

- *the individual asset has dynamics*

$$\frac{df_i}{f_i} = (r + \lambda \rho_i b_i)dt + \rho_i b_i dW + \sqrt{1-\rho_i} dz_i, b_M = \frac{\sum_{i=1}^{\infty} \rho_i b_i f_i}{\sum_{i=1}^{\infty} f_i}.$$

Proof. From above two lemmas, we can assume the individual asset dynamics is given by

$$\frac{df_i}{f_i} = (r + \lambda \rho b_i + \lambda' \sqrt{1-\rho_i} b_i)dt + \rho_i b_i dW + \sqrt{1-\rho_i} b_i dz_i, \forall i.$$

The market portfolio will require $\lambda' = 0$, and it is given by

$$\frac{df_M}{f_M} = (r + \lambda b_M)dt + b_M dW, b_M = \frac{\sum_{i=1}^{\infty} \rho_i b_i f_i}{\sum_{i=1}^{\infty} f_i}.$$

□

Remark 1.5.8 (alternative derivation of CAPM). Denote $\mu_m = r + \lambda b_M, \mu_i = r + \rho_i \lambda b_i$. Then we have

$$\frac{\mu_i - r}{b_i} = \rho_i \frac{\mu_M - r}{b_M}$$

or equivalently

$$(\mu_i - r) = \beta(\mu_M - r),$$

where

$$\beta = \frac{Cov(df_i, df_M)}{Var[df_M]} = \frac{\rho b_i b_M}{b_M^2} = \frac{\rho_i b_i}{b_M}.$$

1.5.4 Asset dynamics when all investors are risk-neutral

Proposition 1.5.4 (constraints on parameters of dynamics when all investors are risk-neutral). *Consider a market with risk-free asset with short rate r. Consider n assets described by stochastic processes given by:*

$$\frac{df_i}{f_i} = a_i dt + \sum_{i=1}^{n} b_{ij} dz_i, i = 1, 2, ..., n.$$

where $z_i, i = 1, 2, ..., m$ are independent Wiener processes. If there is ***no arbitrage*** *and* ***all investors are risk-neutral****, then the asset dynamics is given by*

$$\frac{df_i}{f_i} = r dt + \sum_{i=1}^{n} b_{ij} dz_i, i = 1, 2, ..., n.$$

Proof. In a risk-neutral world, all investors only concern the expected return

$$\frac{dE[f_i]}{E[f_i]} = a_i dt.$$

From the law of one price, $a_i, \forall i$ should equal to the risk-free rate r. Suppose there is an asset grows at a rate greater than r, then all investors will borrow at r and invest this asset, which will push up the risk-free rate. If there is an asset grows at a rate smaller than r, then we should short this asset and invest the money into money market, which has growth rate of r. □

Remark 1.5.9 (zero market price of risk when all investors are risk-neutral). In a risk-neural world, the market price of risk is zero since investors does not require risk premium.

1.6 Continuous-time market model II: pricing and martingale methods

1.6.1 Trading strategies and arbitrage

1.6.1.1 *Self-financing continuous-time trading strategies*

Definition 1.6.1 (continuous-time trading strategy). *[8, p. 25] A continuous-time trading strategies is a k-dimensional stochastic process θ_t representing the portfolio held. Each component $\theta_t^1, \theta_t^2, ..., \theta_t^n$ are predictable.*

Definition 1.6.2 (discrete-time trading strategy). *A discrete-time trading strategies is a k-dimensional stochastic process θ_t representing the portfolio held:*

- *immediately after trading at time $t-1$ so it is know at time $t-1$.*
- *and immediately before trading at time t.*

Remark 1.6.1. Consider the real axis as the time axis, trading operation only happen at these discrete time points $t_1, t_2, ...$ θ_n represents the portfolio held in the interval (t_{n-1}, t_n). θ_t is generally un-defined at time t since the trading is right occurring.

Definition 1.6.3 (value process and gains process associated with a trading strategy). *[8, p. 24]*

- *The* ***value process*** *$V_t(\theta)$, associated with a* ***discrete-time or continuous-time trading strategy*** *θ_t, with the price vector S_t, is defined by:*

$$V_t = \sum_{i=1}^{K} \theta_t^i S_t^i.$$

- *The gains process $G_t(\theta)$ associated with a* ***continuous-time trading strategy*** *θ_t, with the price vector S_t, is defined by*

$$G_t(\theta) = \sum_{i=1}^{K} \int_0^t \theta_u^i dS_u^i.$$

- *The gains process $G_t(\theta)$ associated with a* ***discrete-time trading strategy*** *θ_t, with the price vector S_t, is defined by*

$$G_t(\theta) = \sum_{i=1}^{K} \sum_{j=1}^{t} \theta_j^i (S_j - S_{j-1}).$$

Definition 1.6.4 (self-financing strategy). *[8, p. 25] A self-financing trading strategy is a trading strategy θ_t where its changes in $V_t(\theta)$ are due to entirely trading gains or loss rather than the injection or withdrawal of cash funds. In particular, a self-financing strategy satisfies:*

$$V_t(\theta) = V_0 + G_t(\theta).$$

Remark 1.6.2 (interpretation). Intuitively, a strategy is self-financing if its value changes only due to changes in the asset prices. In other words, no additional cash inflows or outflows occur after the initial time.

1.6.1.2 *Self-financing trading strategies on discounted assets*

1.6.1.3 *Attainable claims*

Definition 1.6.5 (attainable contingent claim).

- *A contingent claim Y can be defined as a random varible defined on Ω represents the payoff at time T.*
- *A contingent claim Y is said to be **attainable** if there exists a self-financing trading strategy h whose value at T satisfies $V(T) = Y$.*

Lemma 1.6.1 (vector space of attainable claims of zero initial values). *Consider a sample space $\Omega = \{\omega_1, ..., \omega_K\}$ and a contingent claim Y defined on Ω. The set of possible values $Y(\omega_1), ..., Y(\omega_K)$ of Y can be considered as an element in $\mathbb{R}^K$. Let*

$$\mathcal{G} = \{Y \in \mathbb{R}^K, Y = G(T) \ for\ some\ trading\ strategy\ h\}.$$

Then $\mathcal{G}$ is a vector space.

Proof. Note that the linear structure

$$Y = \sum_{u=1}^{t} h_0(u)\Delta B(u) + \sum_{u=1}^{t}\sum_{i=1}^{N} h_i(u)\Delta S_i(u).$$

If $Y_1, Y_2 \in \mathcal{G}$ and the associated self-financing strategies are h_1, h_2, then $\alpha Y_1 + \beta Y_2$ will have self-financing strategy of $\alpha h_1 + \beta h_2$ □

1.6.1.4 *Arbitrage*

Definition 1.6.6 (arbitrage). *In the multi-period finite-state market model, an arbitrage opportunity exists if there is a self-financing strategy h whose value satisfying*

1. $V(0) = 0$;
2. $V(T) \geq 0$;
3. $Pr(V(T) > 0) > 0$.

More generally, the arbitrage condition can be

1. $V(0) = V_0$;
2. $V(T) \geq V_0$;
3. $Pr(V(T) > V_0) > 0$.

where V_0 is a deterministic constant.

Proposition 1.6.1 (arbitrage opportunity, change of numeraire, and change of measure, recap). *Theorem 1.4.1*

- *A self-financing strategy h is an arbitrage opportunity under a probability measure P* ***if and only if*** *its discounted value V^* satisfying*
  1. $V^*(0) = 0$;
  2. $V^*(T) \geq 0$;
  3. $Pr(V^*(T) > 0) > 0$.
- *A self-financing strategy h is an arbitrage opportunity under a probability measure P* ***if and only if*** *h is an arbitrage opportunity under an* ***equivalent*** *probability measure Q.*

Proof. (1) Note that because $B(t) > 0$ almost surely, then

1. $V(0)/B(0) = 0 \Leftrightarrow V(0) = 0$;
2. $V(T)/B(T) \geq 0 \Leftrightarrow V(0) = 0$;
3. $Pr(V(T)/B(T) > 0) > 0 \Leftrightarrow Pr(V(T) > 0) > 0$ since $V(T)/B(T) > 0 \Leftrightarrow V(T) > 0$.

(2) Let h be an arbitrage opportunity under P. Let $\Omega' \subseteq \Omega$ be a set such that $V(T)(\omega) > 0, \omega \in \Omega'$. Then

$$P(\Omega') > 0 \Leftrightarrow Q(\Omega') > 0,$$

due to the equivalence of P and Q. □

Note 1.6.1 (implication).

- The first part of the theorem simply says that an arbitrage is always an arbitrage no matter which numeraire we are using.
- The second part of the theorem simply says that an arbitrage is always an arbitrage under all equivalent probability measure.

- This theorem gives us the latitude to find the most convenient numeriare and measure to pricing assets. As long as under the chosen measure and numeriare no arbitrage exists, then no arbitrage exists in the real world.

Lemma 1.6.2 (martingale value process is not an arbitrage, recap). *Lemma 1.4.2 Let $V(t)$ be a martingale under a probability measure P with $V(0) = V_0$, where V_0 is a deterministic constant. Then $V(t)$ is not an arbitrage under any equivalent probability measure.*

Proof. If $V'(T) \geq V_0, P(V(T) > V_0) > 0$, we cannot have $E_P[V(t)] = V_0$. Therefore V cannot be an arbitrage. □

Proposition 1.6.2 (self-financing strategy on discounted assets that are martingale).

1.6.2 First fundamental theorems of asset pricing

1.6.2.1 *Harrison-Pliska martingale no-arbitrage theorem*

Proposition 1.6.3 (Harrison-Pliska martingale no-arbitrage theorem). *Consider a financial market with time horizon T and price processes of the risky assets $S_1(t), S_2(t), ..., S_N(t), 0 \leq t \leq T$ and riskless bond B. The riskless bond $B(t) \geq 0$ almost surely. The market model is aribtrage-free under the probability measure P if and only if there exists another probability measure Q such that*

- *P and Q are equivalent probability measures.*
- *the discounted price processes*

$$\frac{S_1(t)}{B(t)}, \frac{S_2(t)}{B(t)}, ..., \frac{S_N(t)}{B(t)},$$

are martingales under Q.

Proposition 1.6.4 (change of numeraire in conditional expectation, extension to other numeriares, recap). *Theorem 1.6.15 Assume there exists a numeraire N and a probability measure Q_N, equivalently to the original probability measure P, such that the price of any traded asset X(without intermediate payments) relative to N is a martingale under Q^N, i.e.,*

$$\frac{X_t}{N_t} = E_N[\frac{X_T}{N_T}|\mathcal{F}_t], 0 \leq t \leq T.$$

Let U be an arbitrary numeraire. Then ***there exists*** *a probability measure Q_U, equivalent to the measures P and Q_N, such that the price of any attainable claim*

$$\frac{Y_t}{U_t} = E_U[\frac{Y_T}{U_T}|\mathcal{F}_t], 0 \leq t \leq T.$$

Moreover, the Radon-Nikodyn derivative defining the measure Q_U is given by

$$\frac{dQ_U}{dQ_N} = \frac{U_T N_0}{U_0 N_T}.$$

Remark 1.6.3 (extension to other numeriares)**.**

Remark 1.6.4 (model free implications)**.** The Harrison-Pliska martingale no-arbitrage theorem and its extension to other numeraires does not specify any model dynamics associated with the assets. Therefore, any dynamic model (geometric Brownian motion, jump process, or more general levy process) satisfying the martingale relationship can be used as the model.

1.6.2.2 *First fundamental theorem of asset pricing*

Definition 1.6.7 (risk-neutral measure). *[9, p. 228] A probability measure Q is said to be risk-neutral if*

- *Q and P are equivalent(i.e., P and Q have the same zero probability subsets)*
- *under Q, the discounted stock process $S_i(t)/M(t)$ is a martingale for every $i = 1, 2..., m$*

Proposition 1.6.5 (existence of risk-neutral measure implies no arbitrage, first fundamental theorem of asset pricing). *[9, p. 231]*

- *Let a martingale $M(t)$ under equivalent risk-neutral probability measure Q representing the value of a portfolio with $M(0) = 0$, then it does not admit arbitrage.*
- *If a market admits no arbitrage, then it has an risk-neutral probability measure.*

Proof. (1) Consider the arbitrage ($M(0) = 0, P(M(T) \geq 0) = 1, P(M(T) > 0) > 0$ [10, p. 230]) exists, then $E_Q[M(t)] > 0$ due to the fact that probability measure Q and P are equivalent. This contradicts with the fact that $M(t)$ is a martingale $E[M(t)] = M(0) = 0$. (2) See Theorem 1.6.7. □

1.6.2.3 *Risk-neutral measure I: basic concepts*

Definition 1.6.8 (portfolio process and it value). *A portfolio process $\Delta_i(t)$ represents the position of stock i as a function of time. Its value $X(t)$ is governed by the following SDE as:*

$$dX(t) = \sum_{i=1}^{m} \Delta_i dS_i + r(t)(X(t) - \sum_{i=1}^{m} \Delta_i S_i(t))$$

Lemma 1.6.3 (discount portfolio value process is martingale under risk-neutral measure). *[9, p. 230] Let Q be a risk-neutral measure(that is, the discount stock process will be a martingale), then the discounted portfolio value process $D(t)X(t)$ will be a martingale.*

Proof.

$$\begin{aligned} dX(t) &= \sum_{i=1}^{m} \Delta_i dS_i + r(t)(X(t) - \sum_{i=1}^{m} \Delta_i S_i(t)) \\ &= r(t)X(t)dt + \sum_{i=1}^{m} \Delta_i (dS_i(t) - r(t)S_i(t)dt) \\ &= r(t)X(t)dt + \sum_{i=1}^{m} \frac{\Delta_i(t)}{D(t)} d(D(t)S_i(t)) \end{aligned}$$

Then

$$d(D(t)X(t)) = D(t)(dX(t) - R(t)X(t)dt) = \sum_{i=1}^{m} \Delta_i(t) d(D(t)S_i(t)).$$

Since under Q, $D(t)S_i(t)$ is a martingale, $D(t)X(t)$ will also be a martingale. □

Remark 1.6.5 (general pricing procedures). If we can use a portfolio process replicating the payoff of a derivative, then the initial value of the portfolio process is the initial value of the derivative.

Remark 1.6.6 (mathematically interpretation of risk-neutral measure).

- It is tempting to interpret risk-neutral measure from the economical point of view, which, however, can create confusion.
- Mathematically, risk-neutral measure is simply a equivalent measure with some special properties(the most important: the discount portfolio process is a martingale).
- There are infinitely many equivalent measures, but they do not have properties as nice as risk-neutral measure.
- Usually, it is not important to know the explicit form of the risk-neutral measure; however, knowing the dynamic model of assets(e.g., stocks) under such measure is essential.

Remark 1.6.7 (Implications for a martingale). From the properties of martingales(**??**), we have:

- A martingale is a zero-drift stochastic process. A stochastic process $\theta(t)$ is a martingale if it has the form
$$d\theta = \sigma dW$$
where the variable σ can itself be stochastic or as a function of θ and other stochastic processes, W is the Wiener process.
- The expected value at any future time is equal to its value today:
$$E[\theta(T)] = \theta_0$$
note that $E[]$ here does not condition on any thing.

Remark 1.6.8 (Why cares about martingale)**.**

- Consider the discounted value process $M(t)$. If there exists an **equivalent measure** Q such that under Q, $M(t)$ is a martingale, then $M(t)$ will not admit arbitrage in arbitrary equivalent measure(including the real-world measure).
- Suppose $M(0) = 0$, and let $M_T \triangleq M(T)$ be a random variable defined on $(\Omega, \mathcal{F})$. If there exists a measure P such that $P(M_T \geq 0) = 1$, then that mean $M_T(A), \forall \{A \in \mathcal{F}, P(A) > 0\}$. Then for any other probability measure Q equivalent to P, we must also have $Q(M_T \geq 0) = 1$, since the subsets $B \in \mathcal{F}$ such that $M_T < 0$ have measure 0 in both probability measure P and Q.
- Therefore, the no-arbitrage condition implies there exists a probability measure under which discounted value process is martingale. And if there exists

1.6.2.4 *Risk-neutral measure II: existence*

Proposition 1.6.6 (Sufficient condition (coefficient constraints) for existence of risk-neutral measure). *[9, p. 229] The SDE for the discount stock process $D(t)S_i(t)$ is*
$$d(D(t)S_i(t)) = D(t)S_i(t)[(\mu_i - r(t))dt + \sum_{j=1}^{d} \sigma_{ij}(t)dW_j(t)].$$
If there exists processes *$\theta_1(t), ..., \theta_d(t)$ satisfying the following,*
$$\mu_i(t) - r(t) = \sum_{j=1}^{d} \sigma_{ij}(t)\theta_j(t), i = 1, ..., m,$$

*then there exists a measure, called **risk-neutral** measure, such that the discount stock process under Q becomes:*

$$d(D(t)S_i(t)) = D(t)S_i(t)[\sum_{j=1}^{d} \sigma_{ij}(t)d\hat{W}_j(t)]$$

where $d\hat{W}$ are Brownian motions under Q.

More specifically, Let

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{11} & \dots & \sigma_{1d} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2d} \\ \vdots & & & \vdots \\ \sigma_{M1} & \sigma_{M2} & \dots & \sigma_{Md} \end{pmatrix}, x = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{pmatrix}, b = \begin{pmatrix} u_1 - r \\ u_2 - r \\ \vdots \\ u_M - r \end{pmatrix}.$$

If $Ax = b$ has a solution, then there exists a risk-neutral measure; If $Ax = b$ has a unique solution, then there exists a unique risk-neutral measure.

Proof. Use Girsanov theorem(**??**), which says there exists a probability Q such that

$$W(t) + \int_0^t \theta(u)du$$

can be made to be a driftless Brownian motion under Q. Therefore, as long as we can solve $\theta_1(t), ..., \theta_d(t)$ from the linear equation, we can use the Girsanov theorem to find such measure. □

Remark 1.6.9 (Some algebraic conditions)**.**

- First, we always assume $d \leq M$ without loss of generality, since $d > M$ case is equivalent to $d = M$ case(see **??**).
- If $d \leq M$, then there are exactly three possibilities: no solution, unique solution, and infinitely many solutions.

Proposition 1.6.7 (No arbitrage implies the existence of risk-neutral measure). *[9, p. 228]* *The SDE for the discount stock process $D(t)S_i(t)$ is*

$$d(D(t)S_i(t)) = D(t)S_i(t)[(\mu_i - r(t))dt + \sum_{j=1}^{d} \sigma_{ij}(t)dW_j(t)].$$

If there exists ***no arbitrages*** *in the market model, then* ***there exists processes*** $\theta_1(t), ..., \theta_d(t)$ *satisfying the following,*

$$\mu_i(t) - r(t) = \sum_{j=1}^{d} \sigma_{ij}(t)\theta_j(t), i = 1, ..., m,$$

or equivalently from (Theorem 1.6.6), there exists a measure, called ***risk-neutral*** *measure, such that the discount stock process under Q becomes:*

$$d(D(t)S_i(t)) = D(t)S_i(t)[\sum_{j=1}^{d} \sigma_{ij}(t)d\hat{W}_j(t)]$$

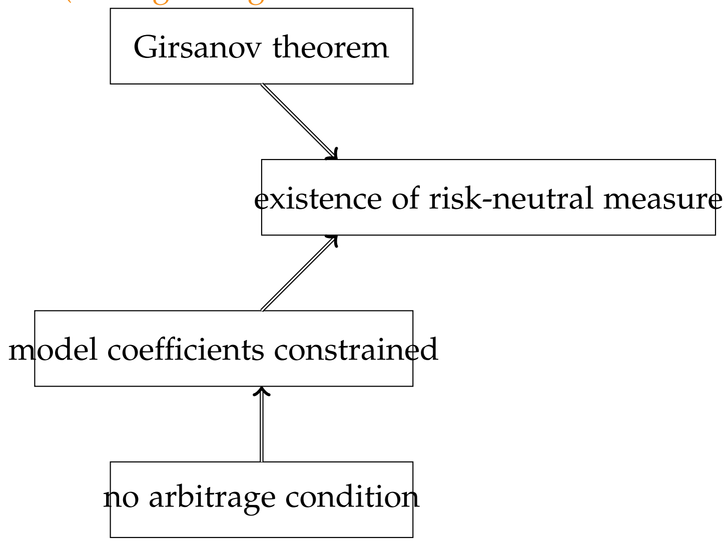
where $d\hat{W}$ are Brownian motions under Q.

Proof. See Theorem 1.5.1. □

Note 1.6.2 (continuous version of finite-state price)**.**

- Consider a special stock $S(t)$ that pays 1 if and only if an event $A(t) \in \mathcal{F}_t$ occurs, then its prices is
$$E_Q[D(t)S(t)] = \int_{\infty}^{\infty} D(t)I(A(t))dP = D(t)Q(A),$$
which equals the discounted probability measure.
- Suppose in the market, there are infinite number of such stocks $S_A$that pay 1 if and only if an event $A(t) \in \mathcal{F}(t)$ occurs. Then $Q(A(t)) = S_A(t)/D(t)$; that is, we can calculate the risk-neutral measure from the stock price.
- In both finite-state model and continuous-time version, the no-arbitrage condition guarantees the existence of the risk-neutral measure or state price. See (Theorem 1.3.1 and Theorem 1.6.7).
- For another discussion, see subsubsection 2.1.2.1.

Remark 1.6.10 (the logic diagram for existence of risk neutral measure)**.**



- In the Brownian asset dynamics model, the no-arbitrage condition implies that the model coefficients are constrained (Theorem 1.5.1).
- The constrained coefficients together with Girsanov theorem Theorem 1.6.6 can imply the existence of risk-neutral measure.
- Further, the existence of risk-neutral measure will imply there is no arbitrage opportunitiesTheorem 1.6.5.

1.6.2.5 *Risk-neutral pricing in no-arbitrage markets*

Proposition 1.6.8 (no-arbitrage pricing in no-arbitrage marketcomplete market). *[9, p. 218] Let $D(t)$ be a discount process, and let $V(t)$ be an asset value. Assume the market is free of abitrages. Denote the risk-neutral measure be Q. Then for* ***any attainable contingent claim*** *with payoff $V(T)$ at time T, its time-t* ***no-arbitrage price*** *is given by*

$$D(t)V(t) = \frac{1}{D(t)} E_Q[D(T)V(T)|\mathcal{F}_t].$$

In particular at current time $t = 0$,

$$V(0) = E_Q[D(T)V(T)].$$

Proof. Because the market is free of abitrages, then there exist a portfolio process $X(t)$ such that $X(T) = V(T)$. Let $X(T)$ be a replicate of asset $V(T)$. Based on the lemma (Lemma 1.6.3), then $D(t)X(t)$ is a martingale with Q. Therefore,

$$D(t)X(t) = E_Q[D(T)X(T)|\mathcal{F}_t]$$

If $X(T)$ replicates $V(T)$, then from the Law of one price(Theorem 1.2.1), $V(t)$ will be equal to $X(t)$. □

Example 1.6.1 (European call put price). The European call with strike price K and maturity time T has price C

$$C = E_Q[D(T)\max(S(T) - K, 0)];$$

And a put has price P

$$P = E_Q[D(T)\max(K - S(T), 0)].$$

And note that under risk-neutral measure, the S_t is following Lemma 1.8.1:

$$dS_t = rS_t dt + \sigma S_t dW_t$$

where W_t is a Wiener process.

Example 1.6.2 (pricing of a special call). Suppose we have a call on stock S_t and matures at T with payoff $\max(S_T^2 - K, 0)$. Then the value at $t = 0$ is

$$C = E_Q[D(T)\max(S_T^2 - K, 0)],$$

where $Y_t = S_t^2$ under risk-neutral measure Q will follow

$$dY_t = (2r + \sigma^2)Y_t dt + (2\sigma)Y_t dW_t.$$

1.6.2.6 *No-arbitrage pricing of futures and forwards*

Corollary 1.6.8.1 (Zero-Coupon Bond price). *[9, p. 240] In a complete market with risk-neutral measure Q, the price for the zero-coupon bond that pays 1 at time T has price at time t as*

$$B(t, T) = \frac{1}{D(t)} E_Q[D(T)|\mathcal{F}_t]$$

In particular at $t = 0$,

$$B(0, T) = E_Q[D(T)].$$

Proof. Use Theorem 1.6.13. Note that $B(T, T) = 1$. □

Remark 1.6.11 (risk-neutral measure is unknown). Note that in this lemma, we do not have the form of risk-neutral measure or how the short-rate dynamics will be under such measure. We just assume it exist, and its form will be discussed in the following sections.

Definition 1.6.9 (forward contract). *[9, p. 241] A forward contract is an agreement to pay a specified delivery price K at a delivery date T for an asset S whose price at t is $S(t)$. Let $F(t,T)$ denote the forward price of this asset at time t, then $F(t,T)$ is the value of K that makes the forward contract having no-arbitrage price zero at at time t.*

Corollary 1.6.8.2 (forward contract pricing). *[9, p. 241] In a complete market with risk-neutral measure Q, the forward price for $S(t)$ is*

$$F(t,T) = \frac{S(t)}{B(t,T)}$$

Proof. Use Theorem 1.6.13.

$$\begin{aligned} 0 &= \frac{1}{D(t)} E_Q[D(T)(S(T)-K)|\mathcal{F}_t] \\ &= \frac{1}{D(t)} E_Q[D(T)S(T)|\mathcal{F}_t] - \frac{K}{D(t)} E_Q[D(T)|\mathcal{F}_t] \\ &= S(t) - KB(t,T) \end{aligned}$$

Therefore,

$$F(t,T) = K = \frac{S(t)}{B(t,T)}$$

□

Remark 1.6.12 (forward pricing vs. future pricing)**.**

- When the interest rate is non-random(even as a function of time), the forward and the future has the same price.
- When the interest rate is random, then the forward and the future price will be different.[9, p. 244]

1.6.3 Second fundamental theorem of asset pricing

1.6.3.1 *Martingale representation theorem*

Proposition 1.6.9 (one dimensional martingale representation theorem). *[11, p. 193][12, p. 49]*

- *Let $W(t), 0 \leq t \leq T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ and with natrual filtration $\{\mathcal{F}_t\}$. If $M_t, 0 \leq t \leq T$, is a martingale with respect to the filtration $\{\mathcal{F}_t\}$, then there exists an predictable process [a] h_t such that*

$$M_t = M_0 + \int_0^t h(s)dW(s), 0 \leq t \leq T.$$

- *Let $W(t)$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ and with natrual filtration $\{\mathcal{F}_t\}$. Let X be a square integrable random variable measurable with respect to $\mathcal{F}_T$, then there exists a predictable process h which is adapted with respect to $\{\mathcal{F}_t\}$, such that*

$$X = E[X] + \int_0^T h(s)dW(s).$$

[a] Because Brownian motion has continuous sample path, with respect to the natural filtration of the Brownian motion, predictable process is the same as adapted process

Proof. (2) Define $X(t) = E[X]\mathcal{F}_t]$ and $X(T) = E[X]\mathcal{F}_T]$. From (1), we have

$$\begin{aligned} X_T &= X_0 + \int_0^T h_u dW_u \\ &= E[X]\mathcal{F}_0] + \int_0^T h_u dW_u \\ &= E[X] + \int_0^T h(s)dW(s) \end{aligned}$$

□

Example 1.6.3. [11, p. 193]

- If $X = W_T$, then $h_t = 1$.
- If $X = W_T^2$, then $h_t = 2W_t$.
- If $X = W_T^3$, then $h_t = 3(W_t^2 + T - t)$.
- If $X = \exp(\sigma W_T)$, then $h_t = \sigma \exp(\sigma W_t + \frac{1}{2}\sigma^2(T-t))$.

Proposition 1.6.10 (Multidimensional dimensional martingale representation theorem).

- *Let $W(t), 0 \leq t \leq T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ and with natrual filtration $\{\mathcal{F}_t\}$. If $M_t, 0 \leq t \leq T$, is a martingale with respect to the filtration $\{\mathcal{F}_t\}$, then there exists an predictable process [a] h_t such that*

$$M_t = M_0 + \sum_{i=1}^{M} \int_0^t h_i(s) dW_i(s), 0 \leq t \leq T.$$

- *Let $W(t)$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ and with natrual filtration $\{\mathcal{F}_t\}$. Let X be a square integrable random variable measurable with respect to $\mathcal{F}_T$, then there exists a predictable process h which is adapted with respect to $\{\mathcal{F}_t\}$, such that*

$$X = E[X] + \int_0^T h(s) dW(s).$$

a Because Brownian motion has continuous sample path, with respect to the natural filtration of the Brownian motion, predictable process is the same as adapted process

Proof. Similar to Theorem 1.6.9. □

1.6.3.2 *Second fundamental theorem of asset pricing*

Definition 1.6.10 (complete market). *A market model is complete if **every finite-variance derivative security measurable with respect to $\mathcal{F}_T$** can be hedged. Or equivalently, a market model is complete if the set of contingent claims measurable with respect to $\mathcal{F}_\infty$ are attainable.*

Proposition 1.6.11 (linear algebra condition for unique risk-neutral measure). ***Assume there exist no arbitrage in the market.*** *Let the SDE of N risky assets be*

$$\frac{df_i}{f_i} = (r_i + \sum_{j=1}^{M} \lambda_j \sigma_{ij})dt + \sum_{j=1}^{M} \sigma_{ij} dw_j$$

where $w_1, ..., w_M$ are M independent Brownian motions. Further, let

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \cdots & \sigma_{1M} \\ \sigma_{21} & \sigma_{22} & \cdots & \cdots & \sigma_{2M} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \cdots & \sigma_{NM} \end{pmatrix}, V = \begin{pmatrix} v_1 + \lambda_1 \\ v_2 + \lambda_2 \\ \vdots \\ \vdots \\ v_M + \lambda_M \end{pmatrix}$$

Then, ***the no arbitrage condition implies there always exists a risk-neutral measure****. Particularly,*

- *When $M > N$ (more risk factors than underlying assets excluding bonds), there exists infinitely many risk-neutral measure Q.*
- *When $M \leq N$ (less risk factors than underlying assets excluding bonds), there are infinitely many or one unique risk-neutral measure Q.*
- *Particularly, there are M independent columns in Σ if and only if there exists one unique risk-neutral measure.*

Under this measure Q, the SDE dynamics is

$$\frac{df_i}{f_i} = r_i dt + \sum_{j=1}^{M} \sigma_{ij} dw_j.$$

Proof. From Girsanov theorem(**??**), we can construct a new measure Q, such that $dw_i = d\hat{w}_i + v_i, \forall i$. The existence of a risk-neutral measure is equivalent to find a solution to

$$\Sigma V = 0.$$

We have exactly two situations,

- When $M > N$, there are infinitely many solution; that is, there exists infinitely many risk-neutral measure Q.
- When $M \leq N$, there are infinitely many solutions or one unique solution. if there are M independent columns in Σ, there exists one unique solution.

□

Remark 1.6.13 (interpretation).

- Under no arbitrage condition, there always exists risk-neutral measure, which is a restatement of first fundamental theorem of asset pricing(Theorem 1.3.1) that no-arbitrage implies the existence of risk-neutral measure.

- When there exists multiple risk-neutral measures, the pricing will not be unique(since price is obtained by taking expectation with respect to such measure).

Proposition 1.6.12 (Second fundamental theorem of asset pricing, uniqueness of risk-neural measure). *[9, p. 232] Consider a no-arbitrage market model consisting of a money market account $B(t)$ and N risky asset $S(t)$ such that*

$$dB(t) = r(t)B(t)dt$$
$$dS_i(t)/S_i(t) = (r_i + \sum_{j=1}^{M} \lambda_j \sigma_{ij})dt + \sum_{j=1}^{M} \sigma_{ij} dw_j.$$

Then the market is complete if and only if it has a unique risk-neutral measure.

Proof. First note that no-arbitrage condition ensures that the existence of risk-neutral measure(Theorem 1.6.5). (1) Assume the model is complete. Suppose we have two risk-neutral measures $Q_1, Q_2 : \mathcal{F} \to \mathbb{R}$. Consider a special indicator derivative with payoff $I_A, A \in \mathcal{F}$.

$$E_{Q_1}[D(T)I_A] = E_{Q_1}[D(0)I_A] = Q_1(A) = E_{Q_2}[D(T)I_A] = E_{Q_2}[D(0)I_A] = Q_2(A).$$

Since A is arbitrary, then this two measure are equal for all subsets in $\mathcal{F}$. (2) (uniqueness implies completeness) For any derivative with payoff $V(T)$ measurable to the natural filtration generated by $W_1, ..., W_M$ (which is the same natural filtration generated by $W_1^Q, W_2^Q, ..., W_M^Q$, Brownian motions under risk-neutral measure), the martingale representation theorem (Theorem 1.6.10) says that **under risk-neutral measure there exist adaptive processes** $\Lambda_1(t), \Lambda_2(t), ..., \Lambda_M(t)$ such that

$$\frac{V(t)}{B(t)} = \frac{V(0)}{B(0)} + \sum_{i=1}^{M} \int_0^t \Lambda_i(u) dW_i^Q.$$

On the other hand, **if** $V(T)$ **can be replicated via a self-financing strategy** such that

$$\begin{aligned}\frac{V(t)}{B(t)} &= \frac{V(0)}{B(0)} + \sum_{i=1}^{N} \int_0^t h_i(u) d(S_i(u)/B(u)) \\ &= \frac{V(0)}{B(0)} + \sum_{i=1}^{N} \int_0^t h_i(u) \sum_{j=1}^{M} \frac{\sigma_{ij}(u)}{B(u)} dW_j^Q.\end{aligned}$$

To solve for $h_1(t), ..., h_N(t)$ for all t, we need to solve

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \dots & \dots & \sigma_{2N} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ \sigma_{M1} & \sigma_{N2} & \dots & \dots & \sigma_{MN} \end{pmatrix} \begin{pmatrix} S_1(t)h_1(t)/B(t) \\ S_1(t)h_2(t)/B(t) \\ \vdots \\ \vdots \\ S_N(t)h_N(t)/B(t) \end{pmatrix} = \begin{pmatrix} \Lambda_1 \\ \Lambda_2 \\ \vdots \\ \vdots \\ \Lambda_M \end{pmatrix}$$

When there exists an unique risk-neutral measure, the linear algebra condition in gives that the matrix Σ has to be full row rank (Theorem 1.6.11) and therefore the linear equation is the consistent and will always have a solution, even though it might not be unique. □

Corollary 1.6.12.1 (completeness of two asset market model). *Consider a market model consisting of a money market account $M(t)$ and a risky asset $S(t)$ such that*

$$\begin{aligned} dB(t) &= r(t)B(t)dt \\ dS(t) &= \mu(t)S(t)dt + \sigma(t)dW(t) \end{aligned}$$

Then any derivative instrument with payoff $V(T)$ measurable to the natural filtration generated by $W(t)$ can be hedged.

Proof. For any derivative with payoff $V(T)$ measurable to the natural filtration generated by $W(t)$, the martingale representation theorem (Theorem 1.6.9) says that **under risk-neutral measure there exist an adaptive process** $\Lambda(t)$ such that

$$\frac{V(t)}{B(t)} = \frac{V(0)}{B(0)} + \int_0^t \Lambda(u)dW^Q.$$

On the other hand, **if** $V(T)$ **can be replicated via a self-financing strategy** such that

$$\begin{aligned} \frac{V(t)}{B(t)} &= \frac{V(0)}{B(0)} + \int_0^t h(u)d(S(u)/B(u)) \\ &= \frac{V(0)}{B(0)} + \int_0^t h(u)\frac{\sigma(u)}{B(u)}dW^Q. \end{aligned}$$

Therefore, we have

$$h(t) = \Lambda(t)B(t)/(S(t)\sigma(t));$$

that is, we can replicate any derivatives. □

Note 1.6.3 (completeness under other equivalent measures).

- Consider a no-arbitrage market model. If the market model is complete under risk-neutral measure, then it is also complete under other equivalent martingale measure.
- This is because when a derivative can be replicated using a self-financing strategy h, then random payoff of the derivative and the strategy is equal, independent of the measure used.

1.6.3.3 *Pricing in complete markets*

Proposition 1.6.13 (no-arbitrage pricing in complete market). *[9, p. 218] Let $D(t)$ be a discount process, let $V(t)$ be an asset value. Assume the market is complete and let the risk-neutral measure be Q. Then*

$$D(t)V(t) = E_Q[D(T)V(T)|\mathcal{F}_t].$$

In particular at $t = 0$,

$$D(0)V(0) = E_Q[D(T)V(T)]$$

which simplifies as

$$V(0) = E_Q[D(T)V(T)]$$

where $D(0) = 1$.

Remark 1.6.14 (Complete market vs. incomplete market).

1.6.4 Dividend paying underlying assets

1.6.4.1 *General remarks*

Remark 1.6.15 (special issues concerning dividend paying assets).

- If the stock is constantly paying dividends, its price will decrease. From the perspective of investors, its value is decreasing; However, from the perspective of stock holders, their total values are not because the dividends belong to them.
- The general model for a stock paying dividend is given as

$$dS_t = (\mu(t) - a(t))S_t dt + \sigma S_t dW_t$$

- Recall that for non-dividend paying stocks, we require the risk-neutral measure to be such that discount stock process $D(t)S_i(t)$ is a martingale(Definition 1.6.7), and as a consequence, the discount portfolio value process $D(t)X(t)$ is a martingale(Lemma 1.6.3). Therefore, we can use discount portfolio value process to replicate and price other derivatives.
- When the stock is paying dividend, we will just require the risk-neutral measure Q to be such that the discount portfolio value process $D(t)X(t)$ is a martingale. **Under Q, $D(t)S_i(t)$ is usually not a martingale([9, p. 248]); Instead, under Q, the discounted value of the portfolio and its dividend accumulating interest rate will be a martingale.**

Definition 1.6.11 (risk-neutral measure with asset paying dividend). *[9, p. 235] A probability measure Q is said to be risk-neutral if*

- *Q and P are equivalent(i.e., P and Q have the same zero probability subsets)*
- *under Q, the discounted portfolio value process $D(t)X(t)$ is a martingale.*

1.6.4.2 *Continuous-dividend paying stocks*

Lemma 1.6.4 (SDE for continuous dividend paying stocks under risk-neutral measure). *Assume the stock S_t in the real-world probability measure follows*

$$dS_t = (\mu - a)S_t dt + \sigma S_t dW_t$$

where W_t is a Wiener process and a is the dividend rate. Then under the risk-neutral measure, the stock dynamics will be

$$dS_t = [r(t) - a(t)]S_t dt + \sigma(t)S_t dW(t).$$

The solution to the SDE is

$$S(t) = S(0)\exp[\int_0^t \sigma(u)dW_u + \int_0^t (r(t) - a(t) - \frac{1}{2}\sigma^2(u)]du$$

Proof.

$$\begin{aligned} dX(t) &= \Delta(t)dS(t) + \Delta(t)a(t)S(t)dt + r(t)[X(t) - \Delta(t)S(t)]dt \\ &= r(t)X(t)dt + (a(t) - R(t))\Delta(t)S(t)dt + \sigma(t)\Delta(t)S(t)dW(t) \\ &= r(t)X(t)dt + \Delta(t)S(t)\sigma(t)[\theta(t)dt + dW(t)] \end{aligned}$$

where

$$\theta(t) = \frac{\mu(t) - r(t)}{\sigma(t)}.$$

If we define

$$\hat{W}(t) = W(t) + \int_0^t \theta(u)du,$$

then from **??**, there exist a measure Q such that $\hat{W}(t)$ will be a Brownian motion. Note that

$$dS_t = [\mu(t) - a(t)]S_t dt + \sigma(t)S_t dW(t).$$

Plug into

$$dW(t) = d\hat{W}(t) - \frac{\mu(t) - r(t)}{\sigma(t)}dt,$$

then we have

$$dS_t = [r(t) - a(t)]S_t dt + \sigma(t)S_t dW(t).$$

□

Remark 1.6.16 (implication for pricing).

- This distribution of S_t under risk-neutral measure Q is important in pricing options on stocks.
- Note the distribution under risk-neutral measure is different for dividend paying stocks and non-dividend paying stocks.

Remark 1.6.17 (risk-neutral dynamics for stocks and its accumated dividends). Consider we have a money account $M(t)$. Then our stock price S_t and the money account M_t dynamics is given by

$$\begin{aligned} dS_t &= (r(t) - a(t))S_t dt + \sigma(t)S_t dW(t) \\ dM_t &= r(t)M(t)dt + a(t)S_t dt \end{aligned}$$

The total value $S(t) + M(t)$ has the following dynamics given by

$$d(S_t + M_t) = r(t)(S_t + M_t)dt + \sigma(t)S_t dW_t.$$

And the discounted total value $D(t)(S_t + M(t))$ will be martingale under risk-neutral measure.

Proposition 1.6.14 (pricing derivatives of continuous dividend paying stocks). *Let $D(t)$ be a discount process, let $V(t)$ be an derivative of a dividend paying stock S_t. Assume the market is complete and let the risk-neutral measure be Q. Then,*

$$D(t)V(t) = E_Q[D(T)V(T)|\mathcal{F}_t].$$

In particular at $t = 0$,

$$D(0)V(0) = E_Q[D(T)V(T)]$$

which simplifies as

$$V(0) = E_Q[D(T)V(T)]$$

where $D(0) = 1$. ***Note that under the risk-neutral measure Q,***

$$dS_t = [r(t) - a(t)]S_t dt + \sigma(t)S_t dW(t).$$

Proof. Same as Theorem 1.6.13. □

1.6.4.3 *Discrete-dividend paying stocks*

Lemma 1.6.5 (SDE for discrete dividend paying stocks under risk-neutral measure). *[9, p. 240] Consider a stock S_t with discrete dividend payment D_i at $t_i, i = 1, 2, 3, ...$ Let current time be t, and $t < t_1 < t_2 < \cdots$. Then its risk-neutral dynamics will be*

$$dS_t = r(t)S_t dt - \sum_{i=1} H(t - t_i)D_i + \sigma(t)S_t dW_t.$$

Proof. Consider we have a money account $M(t)$. Then our stock price S_t and the money account M_t dynamics is given by

$$\begin{aligned} dS_t &= r(t)S_t dt - \sum_{i=1} H(t - t_i)D_i + \sigma(t)S_t dW(t) \\ dM_t &= r(t)M(t)dt + \sum_{i=1} H(t - t_i)D_i \end{aligned}$$

The total value $S(t) + M(t)$ has the following dynamics given by

$$d(S_t + M_t) = r(t)(S_t + M_t)dt + \sigma(t)S_t dW_t.$$

And the discounted total value $D(t)(S_t + M(t))$ will be martingale under risk-neutral measure. □

1.6.5 Numeraire and pricing

1.6.5.1 *No-arbitrage pricing under different numeraire*

Definition 1.6.12 (numeraire). *A numeraire is any positive non-dividend-paying asset.*

Proposition 1.6.15 (change of numeraire in conditional expectation). *[8, p. 27][13, p. 24]* *Assume there exists a numeraire N and its associated martingale probability measure Q_N, equivalently to the initial measure Q_0, such that the price of any traded asset X(without intermediate payments) relative to N is a martingale under Q^N, i.e.,*

$$\frac{X_t}{N_t} = E_N[\frac{X_T}{N_T}|\mathcal{F}_t], 0 \leq t \leq T.$$

Let U be an arbitrary numeraire. Then ***there exists*** *a probability measure Q_U, called* ***martingale measure associated with numeriare*** *U, equivalent to the initial Q_0, such that*

$$\frac{X_t}{U_t} = E_U[\frac{X_T}{U_T}|\mathcal{F}_t], 0 \leq t \leq T.$$

Moreover, the Radon-Nikodyn derivative defining the measure Q_U is given by

$$\frac{dQ_U}{dQ_N} = \frac{U_T N_0}{U_0 N_T},$$

such that

$$U_t E_U[\frac{X_T}{U_T}|\mathcal{F}_t] = U_t E_N[\frac{X_T}{U_T}\frac{dQ_U}{dQ_N}|\mathcal{F}_t] = E_U[\frac{X_T}{U_T}|\mathcal{F}_t].$$

Proof. From the assumption, we have

$$E_N[\frac{U_T}{N_T}|\mathcal{F}_t] = \frac{U_t}{N_t}.$$

We can define

$$Z = \frac{dQ_U}{dQ_N} = \frac{U_T N_0}{U_0 N_T}$$

such that Z is strictly positive and $E[Z] = 1$.

Then from Bayes theorem for conditional expectation(**??**), we have

$$E_U[\frac{Y_T}{U_T}|\mathcal{F}_t] = \frac{E_N[\frac{Y_T}{U_T}Z|\mathcal{F}_t]}{E_N[Z|\mathcal{F}_t]}.$$

Note that

$$E_N[\frac{Y_T}{U_T}Z|\mathcal{F}_t] = \frac{N_0}{U_0}E_N[\frac{Y_T}{N_T}|\mathcal{F}_t] = \frac{N_0}{U_0}\frac{Y_t}{N_t},$$

and

$$E_N[Z|\mathcal{F}_t] = \frac{N_t}{U_t}E_N[\frac{U_T}{N_T}|\mathcal{F}_t] = \frac{N_0}{U_0}\frac{U_t}{N_t}.$$

Then

$$E_U[\frac{Y_T}{U_T}|\mathcal{F}_t] = \frac{Y_t}{U_t}.$$

□

Note 1.6.4 (implication).

- The most important aspect of this theorem is that: combining the first fundamental theorem (Theorem 1.6.6) to ensure the existence of risk neutral measure, we can define no-arbitrage price for all attainable claims in a consistent way (i.e., price of a claim is independent of the numeriare used). Therefore, we can choose the type of numeriare most convinient for us.
- The specific consequences of the this new measure on asset dynamics is discussed at Theorem 1.6.18 and Theorem 1.6.19.

Example 1.6.4 (from risk-neutral measure to forward measure). Denote risk-neutral and forward measure by Q and Q_T, we have

$$\frac{dQ}{dQ_T} = \frac{B(T)}{B(t)}\frac{P(t,T)}{P(T,T)}.$$

For an asset X, under risk-neutral measure, we have

$$\begin{aligned}
X(t) &= B(t)E_Q[\frac{X(T)}{B(T)}|\mathcal{F}_t] \\
&= B(t)E_T[\frac{X(T)}{B(T)}\frac{dQ}{dQ_T}|\mathcal{F}_t] \\
&= B(t)E_T[\frac{X(T)}{B(T)}\frac{B(T)}{B(t)}\frac{P(t,T)}{P(T,T)}|\mathcal{F}_t] \\
&= P(t,T)E_T[\frac{X(T)}{P(T,T)}|\mathcal{F}_t]
\end{aligned}$$

Methodology 1.6.1 (pricing in different numeraire). *Let $N(t)$ be a numeraire. Let E_N denote the expectation with resepct to the martingale measure associated with $N(t)$.*

- *Consider a contingent claim with payoff C unit of domestic currency at future time T, where C is a random variable. Then its price is*

$$V(t) = P(t,T)E_T[\frac{V(T)}{P(T,T)}|\mathcal{F}_t] = P(t,T)\cdot E_T[C|\mathcal{F}_t],$$

where $P(t,T)$ is the zero-coupon bond numeriare, and E_T denote the expectation with resepct to the martingale measure associated with $P(t,T)$.

- *Consider a contingent claim with payoff $c \cdot N(T)$ at future time T, where c is a constant multiplier. Then its price is*

$$V(t) = N(t)E_N[\frac{C \cdot N(T)}{N(T)}|\mathcal{F}_t] = c \cdot N(t).$$

- *Consider a contingent claim with payoff $C \cdot N(T)$ at future time T, where c is a random variable. Then its price is*

$$V(t) = N(t) \cdot E_N[\frac{c \cdot N(T)}{N(T)}|\mathcal{F}_t] = N(t) \cdot E_N[C|\mathcal{F}_t].$$

Example 1.6.5. Consider a payer swaption (Lemma 3.3.15) with strike K and expiry T. Its payoff is either zero or entering a swap, which can be writen by

$$V(T) = N(T) \cdot (Swap(T) - K)^+,$$

where $Swap(T)$ is the swap rate at future time T, and $N(T)$ is the annuity.

By choosing the annuity as the numeriare, we have its current value given by

$$V(t) = N(t)E_N[\frac{N(T) \cdot (Swap(T) - K)^+}{N(T)}|\mathcal{F}_t] = N(t)E_N[(Swap(T) - K)^+|\mathcal{F}_t].$$

We can evaluate the swaption value if we know the distribution $Swap(T)$ under the martingale measure associated with $N(t)$.

1.6.5.2 *State price and risk-neutral pricing*

Lemma 1.6.6 (state price density and pricing in zero-interest world). *Consider a world with zero interest rate. Define $V(t,\omega_0) \triangleq PV[\delta(\omega - \omega_0)]$, where $PV[\cdot]$ is the present value operator taking an random variable as input and produce a real number as output. It follows that*

- *For an arbitrary payoff $f(\omega)$, a random variable $\Omega \to \mathbb{R}$, we have the decomposition*

$$f(\omega) = \int \delta(\omega - \omega_0) f(\omega_0) d\omega_0.$$

- *No arbitrage condition requires $1 = PV[1]$ and $PV[\cdot]$ is a linear operator.*

- *The present value of an arbitrary payoff* $f(\omega)$ *is given by*

$$V(t) = PV[f(\omega)] = \int V(t,\omega)f(\omega)d\omega.$$

- $$1 = \int V(t,\omega)d\omega.$$

- *If we define* $E_Q[f|\mathcal{F}_t] = \int f(\omega)V(\omega)d\omega$*, then*

$$V(t) = E_Q[f|\mathcal{F}_t]$$

.

Proof. (1) Directly from the property of delta function. (2) From law of one price and linear pricing theorem.Theorem 1.2.2Theorem 1.2.1(3)

$$\begin{aligned}
V(t) &\triangleq PV[f(\omega)] \\
&= PV[\int \delta(\omega - \omega_0)f(\omega_0)d\omega_0] \\
&= \int PV[\delta(\omega - \omega_0)]f(\omega_0)d\omega_0 \\
&= \int V(t,\omega_0)f(\omega_0)d\omega_0
\end{aligned}$$

□

Definition 1.6.13 (state price density and pricing using money account as numeraire). *A state price density using money account as numeraire is the time t price of the payoff denominated in the money account when the world is realized sample point* $\omega \in \Omega$ *at time* T*. We write the state price density as*

$$\frac{V(t,\omega_0)}{B_t} = PV[\frac{\delta(\omega - \omega_0)}{B_T(\omega)}].$$

Proposition 1.6.16. *Assume we have the state price density for all possible states. Given any payoff function* $f(S_T), S_T : \Omega \to \mathbb{R}$*. We have*

- *The present value of money account* B_T*(a random variable) is given by*

$$V(t) = B_t.$$

- $\int V(t,\omega)d\omega = 1.$

- $$V(t) = B_t PV[\int \frac{\delta(\omega-\omega_0)f(\omega_0)}{B_T(\omega_)}d\omega_0]$$
- $$V(t) = B_t E_Q[\frac{f(\omega)}{B_T(\omega)}|\mathcal{F}_t]$$

Note 1.6.5 (the fallacy of using real-world probability measure). Let S_T be a random variable. Let $f(S_T)$ be the payoff at time T. Let current time be 0. Further assume zero interest rate. One fallacy of pricing is given by

$$V(0) = E_P[f(S_T)],$$

where P is the real world measure.

- this method is appealing because we can usually estimate the distribution of S_T in the real world measure via historical data.
- However, this method does not take into account the risk-aversion nature of the market, in which market price should be higher than $E_P[f(S_T)]$.

1.6.6 Market model under different martingale measure

1.6.6.1 *Brownian motion in different measures*

Proposition 1.6.17 (changing dynamics between measures). *[8, p. 32] Consider a numeraire U and a numeraire N with its associated measure Q_U and Q_N measure. Assume a n-dimensional diffusion process X_t,* ***under measure*** *Q_U, has dynamics*

$$dX_t = \mu_X^U(X_t,t)dt + \sigma_X(X_t,t)dW_t^U,$$

where $\mu_X^U \in \mathbb{R}^n$ and $\sigma_X \in \mathbb{R}^{n\times n}$, $W_t^U \in \mathbb{R}^n$ is a n-dimensional correlated Brownian motion under measure Q_U, and $dW_t^U[dW_t^U]^T = \rho \in \mathbb{R}^{n\times n}$, $\sigma_X = diag(\sigma_{X,1},\sigma_{X,2},...,\sigma_{X,n})$.[a]

Now assume the dynamics of X_t under measure Q_N is given by

$$dX_t = \mu_X^N(X_t,t)dt + \sigma_X(X_t,t)dW_t^N,$$

where $W_t^U \in \mathbb{R}^n$ is a n-dimensional correlated Brownian motion under measure Q_N.

Further assume the numeraire has dynamics, under measure Q_U, given by

$$dU_t = \mu_U^U dt + \sigma_{N,t} dW_t^U$$
$$dN_t = \mu_N^U dt + \sigma_{U,t} dW_t^U$$

where $\sigma_{N,t}, \sigma_{U,t} \in \mathbb{R}^{1\times n}$.

It follows that

- $$\mu_X^N(X_t,t) = \mu_X^U(X_t,t) + \sigma_X(X_t,t)\rho(\frac{\sigma_N}{N_t} - \frac{\sigma_U}{U_t})^T,$$

  or equivalently,

  $$\mu_X^N(X_t,t) = \mu_X^U(X_t,t) + dX_t d(\log(\frac{N_t}{U_t})),$$

- $$dW_t^N = dW_t^U - \rho(\frac{\sigma_N}{N_t} - \frac{\sigma_U}{U_t})^T dt.$$

- *Moreover, if for any X_i/U is a martingale under Q_U, then X_i/N is a martingale under Q_N.*

---

[a] Note that the superscript denotes different measures and volatilities(the coefficients before Brownian motion) will not change under changing measure.

Proof. Note that we know that $dE_U[X_t] = \mu_X^U dt$. Then we have

$$\begin{aligned}
\mu_X^U dt &= E_U[dX_t] \\
&= E_N[dX_t \frac{dQ_U}{dQ_N}] \\
&= E_N[dX_t \frac{U(t+dt)N(t)}{U(t)N(t+dt)}] \\
&= E_N[dX_t \frac{\frac{U}{N}(t+dt)}{\frac{U}{N}(t)}] \\
&= E_N[dX_t(1 + d(\frac{U}{N}))/(U/N)] \\
&= E_N[dX_t] + E_N[dX_t d(\frac{U}{N})/(U/N)] \\
&= E_N[dX_t] + E_N[dX_t d(\log(\frac{U}{N}))] \\
&= \mu_X^N dt + E[\sigma_X dW_t^N((\frac{\sigma_U}{U_t} - \frac{\sigma_N}{N_t})dW_t^N)^T] \\
&= \mu_X^N dt + \sigma_X \rho(\frac{\sigma_U}{U_t} - \frac{\sigma_N}{N_t})^T
\end{aligned}$$

Note that we will get the same result if we use

$$\frac{dQ_U}{dQ_N} = \frac{U(T)N(t)}{U(t)N(T)},$$

since dX_t only correlated with the $[t, t+dt]$ part in $\frac{U(T)}{N(T)}$ in the $O(t^{1/2})$ scale. □

Note 1.6.6 (parameter constraint on drifting parameters).

- Note that under no-arbitrage condition, the drifting parameters $\mu_i, i = 1, 2, ..., N$ under different measures are not unrelated. There relationship is given by Theorem 1.6.20.
- Also see the following example.

Example 1.6.6. Consider three assets X, U and N. Assume under the measure associated with numeraire associated with U, we have

$$\begin{aligned} dX_t/X_t &= \mu_X^U dt + \sigma_X dW_1^U \\ dU_t/U_t &= \mu_U^U dt + \sigma_U dW_2^U \\ dN_t/N_t &= \mu_N^U dt + \sigma_N dW_3^U \end{aligned}$$

where W_1^U, W_2^U, W_3^U are correlated Brownian motions with $dW_i dW_j = \rho_{ij} dt$.

Note that $\mu_X^U, \mu_U^U, \mu_N^U$ are satisfying

$$\begin{aligned} \mu_X^U - \mu_U^U &= \rho_{XU}\sigma_X\sigma_U - \sigma_U^2 \\ \mu_N^U - \mu_U^U &= \rho_{NU}\sigma_N\sigma_U - \sigma_U^2 \end{aligned}$$

such that X/U and N/U are martingales.

Under the measure Q_N associated with numeraire N, we assume

$$\begin{aligned} dX_t/X_t &= \mu_X^N dt + \sigma_X dW_1^N \\ dU_t/U_t &= \mu_U^N dt + \sigma_U dW_2^N \\ dN_t/N_t &= \mu_N^N dt + \sigma_N dW_3^N \end{aligned}$$

where W_1^N, W_2^N, W_3^N are correlated Brownian motions with $dW_i dW_j = \rho_{ij} dt$.

When changing to the measure Q_N associated with numeraire N, we have

$$\begin{aligned} \mu_X^N &= \mu_X^U + \sigma_X\rho_{XN}\sigma_N - \sigma_X\rho_{XU}\sigma_U \\ \mu_N^N &= \mu_N^U + \sigma_N^2 - \sigma_N\rho_{NU}\sigma_U \end{aligned}$$

Then it can be shown that

$$\begin{aligned}\mu_X^N - \mu_N^N &= \mu_X^U - \mu_N^U + \sigma_X\rho_{XN}\sigma_N - \sigma_X\rho_{XU}\sigma_U - \sigma_N^2 + \sigma_N\rho_{NU}\sigma_U \\ &= \sigma_X\rho_{XN}\sigma_N - \sigma_N^2;\end{aligned}$$

that is, X/N will be a martingale under Q_N.

Lemma 1.6.7. *Consider a SDE under real probability measure given by*

$$dX_t = \mu dt + \sigma dW_t.$$

Let the SDE of a zero-coupon bond under real probability measure be given by

$$dP(t,T)/P(t,T) = b(t)dt + \Sigma(t,T)dW_t.$$

Let the SDE of an arbitrary asset $S(t)$ under risk-neutral probability measure be given by

$$dS(t,T)/S(t,T) = r(t)dt + \sigma_S(t)dW_t.$$

It follows that

- *Under risk-neutral measure Q,*

$$dX_t = rdt + \sigma dW_t^Q.$$

  where W_t^Q is a Brownian motion under Q.
- *Under risk-neutral measure Q,*

$$dW_s = dW_s^Q - \frac{\mu(s) - r(s)}{\sigma(s)}ds.$$

- *Under forward measure Q_T with respect to $T-$maturity zero-coupon bond,*

$$dX_t = (r + \Sigma^2)dt + \sigma dW_t^T,$$

  where W_t^T is a Brownian motion under Q_T.
- *Under forward measure Q_T with respect to $T-$maturity zero-coupon bond,*

$$dW_s = dW_s^T - \frac{\mu - (r + \Sigma^2(t,T))}{\sigma}ds.$$

Lemma 1.6.8 (changing dynamics from risk-neutral measure to forward measure). *Consider a SDE under risk-neutral probability measure given by*

$$dX_t = \mu dt + \sigma dW_t.$$

Let the SDE of a zero-coupon bond under real probability measure be given by

$$dP(t,T)/P(t,T) = b(t)dt + \Sigma(t,T)dW_t.$$

It follows that

- *Under forward measure Q_T with respect to $T-$maturity zero-coupon bond,*

$$dX_t = (r+\sigma^2)dt + \sigma dW_t^T,$$

  where W_t^T is a Brownian motion under Q_T.
- *Under forward measure Q_T with respect to $T-$maturity zero-coupon bond,*

$$dW_s = dW_s^T - \frac{\mu - (r+\sigma^2)}{\sigma} ds.$$

- *Under forward measure,*

$$\frac{X(t)}{P(t,T)}$$

  is a martingale.

1.6.6.2 *SDE approach: Single source of uncertainty*

Proposition 1.6.18 (change of numeraire, single source uncertainty). *[5, p. 661] Consider two Ito process f and g of assets with common source of uncertainty, given by*

$$df = \mu_f f dt + \sigma_f f dz,$$

$$dg = \mu_g g dt + \sigma_g g dz,$$

where z is the Brownian motion under the original measure. Assume g is always positive. It follows that

- *Under no arbitrage condition, there exists an measure Q_g (i.e. set market price of risk to σ_g)such that under such measure, we have*

$$df = (r + \sigma_g \sigma_f) f dt + \sigma_f f d\hat{z},$$

and

$$dg = (r + \sigma_g^2)gdt + \sigma_g g d\hat{z},$$

where r is the risk-free rate and $\hat{z}$ is a Brownian motion under the new measure Q_g.

- *Under the new measure Q_g,*

$$d(\frac{f}{g}) = (\sigma_f - \sigma_g)\frac{f}{g}d\hat{z},$$

that is, the quantity

$$\frac{f}{g}$$

is a martingale under the this measure.

- *Under the new measure Q_g, the Brownian motion z in original measure is given by*

$$dz = d\hat{z} + (\sigma_g - \lambda)dt,$$

where λ is the market price of risk.

Note that if set market price of risk to σ_f, we have

$$d(\frac{f}{g}) = (\sigma_f - \sigma_g)^2\frac{f}{g}dt + (\sigma_f - \sigma_g)\frac{f}{g}dz.$$

Proof. From the lemma (Lemma 1.5.1), we know that the drift parameters of the two SDE with single common sources are constrained as:

$$\frac{\mu_g - r}{\sigma_g} = \frac{\mu_f - r}{\sigma_f} = \lambda,$$

where λ is unknown function resulting from market forces equilibrium. From Girsanov theorem(**??**), we can define random quantity

$$Z = \exp(\int_0^T (\sigma_g - \lambda)dz(s) - \frac{1}{2}\int_0^T (\sigma_g - \lambda)^2 ds)$$

to generate a new measure Q_g such that under Q_g

$$dz = d\hat{z} + (\sigma_g - \lambda)dt,$$

Then under Q_g,

$$df = (r + \lambda\sigma_f)fdt + \sigma_f f dz \Leftrightarrow df = (r + \lambda\sigma_f)fdt + \sigma_f f(d\hat{z} + (\sigma_g - \lambda)dt)$$

and

$$dg = (r + \lambda\sigma_g)gdt + \sigma_g g dz \Leftrightarrow dg = (r + \lambda\sigma_g)gdt + \sigma_g g(d\hat{z} + (\sigma_g - \lambda)dt),$$

Arrange terms, we have

$$df/f = (r + \sigma_f\sigma_g)dt + \sigma_f d\hat{z}$$

and

$$dg/g = (r + \sigma_g^2/2)dt + \sigma_g d\hat{z}$$

Or equivalently, we have

$$d(\ln f) = (r + \sigma_f\sigma_g - \sigma_f^2/2)dt + \sigma_f d\hat{z}$$

and

$$d(\ln g) = (r + \sigma_g^2/2)dt + \sigma_g d\hat{z}$$

Therefore, under Q_g

$$d(\ln f - \ln g) = -\frac{(\sigma_f - \sigma_g)^2}{2}dt + (\sigma_f - \sigma_g)d\hat{z}$$

and

$$d(\frac{f}{g}) = (\sigma_f - \sigma_g)\frac{f}{g}d\hat{z}.$$

□

Note 1.6.7. Note that under the measure Q_g, $1/g$ is not a martingale but f/g is.

Remark 1.6.18. It can be showed that Lemma 1.5.1, under the market price of risk σ_g, we have

$$\frac{r + \sigma_g^2 - r}{\sigma_g} = \frac{r + \sigma_g\sigma_f - r}{\sigma_f} = \sigma_g.$$

Example 1.6.7. [14, p. 169] A stock,S_t, follows the Black-Scholes model. A derivative D_t pays $S_T I_{S_T > K}$ at time T. Develop a formula for its current price at time t.

Solution: Use the stock as the numeraire, then under this new measure Q_S,

$$\frac{D(T)}{S(T)}$$

is a martingale. Therefore,

$$\frac{D(t)}{S(t)} = E_S[\frac{D(T)}{S(T)}|\mathcal{F}_t] = E_S[I_{S_T>K}|\mathcal{F}_t].$$

Note that under measure Q_S, the dynamics of $S(t)$ follows

$$dS(t) = (r+\sigma^2)S(t)dt + \sigma S(t)dW(t).$$

This allows us to to obtain

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

with Z a standard normal variable. Then

$$D(t) = S(t)P(S_T > K),$$

can be evaluated.

Example 1.6.8. [14, p. 170] A stock,S_t, follows the Black-Scholes model. A derivative D_t pays $S_T \log S_T$ at time T. Develop a formula for its current price at time t.

Solution: Use the stock as the numeraire, then under this new measure Q_S,

$$\frac{D(T)}{S(T)}$$

is a martingale. Therefore,

$$\frac{D(t)}{S(t)} = E_S[\frac{D(T)}{S(T)}|\mathcal{F}_t] = E_S[\log S_T|\mathcal{F}_t].$$

Note that under measure Q_S, the dynamics of $S(t)$ follows

$$dS(t) = (r+\sigma^2)S(t)dt + \sigma S(t)dW(t).$$

This allows us to to obtain

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

with Z a standard normal variable. Then

$$E_S[\log S(T)|\mathcal{F}_t] = \log S(t) + (r + \frac{1}{2}\sigma^2)(T-t),$$

and

$$D(t) = S(t)(\log S(t) + (r + \frac{1}{2}\sigma^2)(T-t)).$$

1.6.6.3 *SDE approach: Multiple sources of uncertainty*

Proposition 1.6.19 (change of numeraire, multiple sources of uncertainty). *[5, p. 668] Consider the SDE for assets $f_1, f_2, ..., f_M$ and g driven by n independent Brownian motions $z_1, z_2, ..., z_n$, given by,*

$$\frac{df_j}{f_j} = [r + \sum_{i=1}^n \lambda_i \sigma_{f_j,i}]dt + \sum_{i=1}^n \sigma_{f_j,i} dz_i, j = 1,2,...,M$$
$$\frac{dg}{g} = [r + \sum_{i=1}^n \lambda_i \sigma_{g,i}]dt + \sum_{i=1}^n \sigma_{g,i} dz_i$$

where r is the risk-free rate and $\lambda_1, \lambda_2, ..., \lambda_n$ are market prices of risks. Assume g is always positive. It follows that

- *There exists a new measure Q_g such that*

$$d\hat{z}_i = dz_i + \sigma_{g,i} dt, i = 1,2,...,n;$$

- *Under the new measure Q_g Then*

$$d(\frac{f_j}{g}) = (\frac{f_j}{g})(\sum_{i=1}^n (\sigma_{f_j,i} - \sigma_{g,i}))d\hat{z}_i, j = 1,2,...,M.$$

Proof. From Girsanov theorem(**??**), we can define random quantity

$$Z = \exp(\int_0^T \theta(s) \cdot dz(s) - \frac{1}{2}\int_0^T \|\theta(s)\|^2 ds),$$

where $\theta_i(s) = (\sigma_{g,i} - \lambda_i), i = 1,2,...,n.$ to generate a new measure Q_g such that under Q_g

$$dz_i = d\hat{z}_i + (\sigma_{g,i} - \lambda_i)dt, i = 1,2,...,n.$$

Therefore, under the measure Q_g, we have

$$d\log f_j = [r + \sum_{i=1}^{n}(\sigma_{g,i}\sigma_{f_j,i} - \sigma_{f_j,i}^2/2)]dt + \sum_{i=1}^{n}\sigma_{f_j,i}d\hat{z}_i, j = 1,2,...,M$$
$$d\log g = [r + \sum_{i=1}^{n}(\sigma_{g,i}\sigma_{g,i} - \sigma_{g,i}^2/2)]dt + \sum_{i=1}^{n}\sigma_{g,i}d\hat{z}_i$$

so that

$$d\log\frac{f_j}{g} = d(\log f_j - \log g) = [\sum_{i=1}^{n}(\sigma_{g,i}\sigma_{f_j,i} - \sigma_{g,i}^2/2 - \sigma_{f_j,i}^2/2)] + \sum_{i=1}^{n}(\sigma_{f_j,i} - \sigma_{g,i})dz_i, j = 1,2,...,M.$$

□

Example 1.6.9. [14, p. 272]Lemma 5.3.12[The Margrabe option] Two stocks,$S_1(t)$ and $S_2(t)$, follows the SDE in real-world measure.

$$dS_1(t) = (r + \lambda_1\sigma_1)S_1(t)dt + \sigma_1 S_1(t)dW_1(t)$$
$$dS_2(t) = (r + \rho\lambda_1\sigma_2 + \lambda_2\sqrt{1-\rho^2}\sigma_2)S_2(t)dt + \rho\sigma_2 S_2(t)dW_1(t) + \sqrt{1-\rho^2}\sigma_2 S_2(t)dW_2(t)$$

where W_1 and W_2 are independent Brownian motions.

A derivative D_t pays $\max(S_2(T) - S_1(T), 0)$ at time T. Develop a formula for its current price at time t.

Solution:

Use the stock S_1 as the numeraire, then under this new measure Q_S,

$$\frac{D(T)}{S_1(T)}$$

is a martingale. Therefore,

$$\frac{D(t)}{S_1(t)} = E_{S_1}[\frac{D(T)}{S(T)}|\mathcal{F}_t] = E_{S_1}[\max(\frac{S_2(T)}{S_1(T)} - 1)|\mathcal{F}_t].$$

Note that under measure Q_S (we take $\lambda_1 = \sigma_1, \lambda_2 = 0$ following Theorem 1.6.19), the dynamics of $S_1(t)$ and $S_2(t)$ follows

$$dS_1(t) = (r + \sigma_1^2)S_1(t)dt + \sigma_1 S_1(t)dW_1(t)$$
$$dS_2(t) = (r + \rho\sigma_1\sigma_2)S_2(t)dt + \ \rho\sigma_2 S_2(t)dW_1(t) + \sqrt{1-\rho^2}\sigma_2 S_2(t)dW_2(t)$$
$$d\frac{S_2}{S_1} = \frac{S_2}{S_1}((\rho\sigma_2 - \sigma_1)dW_1 + \sqrt{1-\rho^2}\sigma_2\sigma dW_2)$$

Denote $Y = \frac{S_2}{S_1}$, then Y is a geometric Brownian motion with volatility $\sigma_Y = \sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}$.

Then we have $D(t) = S_1(t)E_{S_1}[\max(Y(T) - 1, 0)]$, which can be evaluated.

1.6.6.4 *Market price of risk under different measures*

Proposition 1.6.20 (No arbitrage constraints on parameters of dynamics under different measures). *Consider a market with risk-free asset with short rate r. Consider n assets as stochastic processes given as:*

$$\frac{df_i}{f_i} = a_i dt + \sum_{j=1}^{m} b_{ij} dz_j, i = 1, 2, ..., n.$$

where $z_j, j = 1, 2, ..., m$ are independent Wiener processes.

*Assume there is **no arbitrage** in the market. It follows that*

- *there exists (unnecessarily unique) $\lambda_1, ..., \lambda_m$*

$$a_i - r = \sum_{j=1}^{m} b_{ij}\lambda_j, j = 1, 2..., m,$$

such that

$$\frac{df_i}{f_i} = (r + \sum_{j=1}^{m} \lambda_j b_{ij})dt + \sum_{j=1}^{m} b_{ij} dz_j, i = 1, 2, ..., n.$$

- *under any new equivalent measure Q such that*

$$dz_i = d\hat{z}_i + \theta_i dt, i = 1, 2, ..., n,$$

where $\hat{z}_i, i = 1, 2, ..., n$ are independent Brownian motions under Q, there exists $\lambda_1^Q, ..., \lambda_m^Q$

$$a_i - r = \sum_{j=1}^{m} b_{ij}\lambda_j^Q, j = 1, 2..., m.$$

- *under this measure Q*

$$\frac{df_i}{f_i} = (r + \sum_{j=1}^{m} \lambda_j^Q b_{ij})dt + \sum_{j=1}^{m} b_{ij} d\hat{z}_j, i = 1, 2, ..., n.$$

Proof. (1) See Theorem 1.5.1. (2)(3) Note that we can use Girsanov theorem(**??**)to generate a new equivalent measure Q via random quantity

$$Z = \exp(\int_0^T \theta(s) \cdot dz(s) - \frac{1}{2}\int_0^T \|\theta(s)\|^2 ds),$$

where $\theta_i(s) = (\lambda_i^Q - \lambda_i), i = 1, 2, ..., n$. Under Q, we have

$$dz_i = d\hat{z}_i + (\lambda_i^Q - \lambda_i)dt, i = 1, 2, ..., n.$$

plug into

$$\frac{df_i}{f_i} = (r + \sum_{j=1}^m \lambda_j b_{ij})dt + \sum_{j=1}^m b_{ij} dz_j, i = 1, 2, ..., n,$$

we have

$$\frac{df_i}{f_i} = (r + \sum_{j=1}^m \lambda_j^Q b_{ij})dt + \sum_{j=1}^m b_{ij} d\hat{z}_j, i = 1, 2, ..., n.$$

□

Note 1.6.8 (implications)**.**

- In Theorem 1.5.1, we prove that under the real world probability measure, the drifting parameters has to satisfy

$$a_i - r = \sum_{j=1}^m b_{ij}\lambda_j, j = 1, 2..., m.$$

  such that the model admits no arbitrage.
- In this theorem, we prove that under different equivalent measure, or under different numeraire associated measure, such parameter constraint relation always exist.

1.7 Black-Scholes framework

1.7.1 Canonical version of Black-Scholes model

Assumption 1.3 (Assumptions for canonical Black-Scholes model:). *[15, p. 41][7, p. 85][1]*

- *The asset price S follows geometric random walk.*
- *The risk free rate r and volatility σ are known.*
- *No transaction cost.*

- *The underlying asset pays no dividends.*
- ***No arbitrage****.*
- *Short selling is permitted.*
- *The trading of assets can take place continuously in time and amount.*
- *No bid-ask spreads.*

Proposition 1.7.1 (Black-Scholes equation). *[15, p. 41] Let $V(S(t), t)$ be the value of the derivative as a function of the asset price $S(t)$ and time t. Assume $S(t)$* ***under real world measure*** *is governed by*

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

where W_t is the Brownian motion. Then V is governed by

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

with final condition $V(S(T), T) = V_T(S(T))$ and boundary condition $V(S, t) = V_a(t)$ on $S = a$ and $V(S, t) = V_b(t)$ on $S = b$.

Proof. Use Ito's lemma(**??**), we have

$$dV = \sigma S\frac{\partial V}{\partial S}dW_t + (\mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t})dt$$

Construct a portfolio $\Pi = V - \Delta S$, then

$$d\Pi = dV - \Delta dS = \sigma S(\frac{\partial V}{\partial S} - \Delta)dW_t + (\mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} - \mu\Delta S)dt$$

If we set $\frac{\partial V}{\partial S} = \Delta$, then the value of the portfolio Π will evolutes deterministically, we have

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

as required by risk free portfolio valuation theorem(Corollary 1.2.1.2). Finally, we have

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

A simpler proof:

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(dS)^2$$

then

$$d\Pi = dV - \Delta dS = \frac{\partial V}{\partial t}dt + (\frac{\partial V}{PaS} - \Delta)dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(dS)^2$$

which implies $\Delta = \partial V/\partial S$. And

$$d\Pi = dV - \frac{\partial V}{\partial S}dS = \frac{\partial V}{\partial t}dt + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(dS)^2 = \frac{\partial V}{\partial t}dt + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2 dt = r(V - \frac{\partial V}{\partial S}S)dt.$$

□

Remark 1.7.1 (interpretation and applicability)**.**

- This formula applies to both European call and put.
- **The σ, r, μ can be a function of S_t and t, but they must be a deterministic function** .
- The growth rate of S does not enter into the equation.

Remark 1.7.2 (boundary and final conditions)**.**

- For an European call option with strike price E and expiry date T, the final condition is $C(S,T) = \max(S-E,0)$, and the boundary conditions are $C(0,t) = 0$(when $S = 0$, it will always stays at 0, and the call option is worthless.) and $C(\infty, t) \sim S$(when S is infinitely large, then C has infinite value for any finite E).
- For an European put option with strike price E and expiry date T, the final condition is $P(S,T) = \max(E-S,0)$, and the boundary conditions are $P(0,t) = Ee^{-r(T-t)}$(when $S = 0$, it will always stays at 0, and the put option is worthing the discounted payoff E.) and $P(\infty, t) \sim 0$(when P is infinitely large, then P has 0 value for any finite E. Since it is unwise to buy at an infinitely high price and sell at a finite price E).

Lemma 1.7.1 (martingale pricing under risk-neutral measure). *Consider a derivative V is governed by*

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

Then there exists a measure Q, called ***risk-neutral measure****, under which the underlying asset dynamics is given by*

$$dS = rSdt + \sigma S dW_t$$

where W_t is a Brownian motion.

Moreover, the quantity $e^{-rt}V(t)$ is a martingale under risk-neutral measure Q, which enables martingale pricing formula given by

$$V(t) = E_Q[\exp(-r(T-t))V(S_T)|\mathcal{F}_t].$$

Proof. (1) The existence of such risk-neutral measure is directly from Feyman-Kac theorem (Theorem 1.8.1).

(2) To show $e^{-rt}V(t)$ is a martingale, we want to show it is a driftless SDE under Q.

$$\begin{aligned}
&d(\exp(-rt)V(t,S_t,\sigma_t)) \\
&= -\exp(-rt)rVdt + \exp(-rt)dV \\
&= -\exp(-rt)rVdt + \exp(-rt)[(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})dt + \frac{\partial V}{\partial S}dS] \\
&= -\exp(-rt)rVdt + \exp(-rt)[rVdt + \sigma S\frac{\partial V}{\partial S}dW_t] \\
&= \exp(-rt)\sigma S\frac{\partial V}{\partial S}dW_t
\end{aligned}$$

where we use the Black-Scholes equation in the derivation. □

Note 1.7.1 (asset dynamics under real probability and risk-neutral probability). The asset dynamics under risk-neutral probability is given by

$$dS = rSdt + \sigma SdW_t$$

and

$$\begin{aligned}
dV &= (\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})dt + \frac{\partial V}{\partial S}dS \\
&= rVdt + \frac{\partial V}{\partial S}\sigma SdW_t
\end{aligned}$$

where we use the Black-Scholes equation in the derivation.

Under real probability measure, the dynamics are

$$dS = \mu Sdt + \sigma SdW_t$$

and

$$\begin{aligned}
dV &= (\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})dt + \frac{\partial V}{\partial S}dS \\
&= rV + \frac{\partial V}{\partial S}S(\mu - r)dt + \frac{\partial V}{\partial S}\sigma SdW_t
\end{aligned}$$

If we define $\lambda = (\mu - r)/\sigma$, we find that

$$\begin{aligned}
dS/S &= (r + \lambda\sigma)dt + \sigma dW_t \\
dV/V &= (r + \lambda\frac{\partial V}{\partial S}\frac{\sigma S}{V})dt + \frac{\partial V}{\partial S}\frac{\sigma S}{V}dW_t
\end{aligned}$$

which is consistent with the no-arbitrage condition for single source uncertainty dynamics(Lemma 1.5.1).

1.7.2 Fundamental solution and risk-neutral measure

Lemma 1.7.2 (fundamental solution to Black-Scholes equation). *[16, p. 35] Assume constant short rate r. Given Black-Scholes equation*

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

and let $V^\delta(t, S_t, s) > 1$ be the solution for final condition $V_T^\delta(S_T, s) = \delta(S_T - s)$. It follows that

- $$e^{-r(T-t)} = \frac{B(t)}{B(T)} = \int_{-\infty}^{\infty} V^\delta(t, S_t, s)ds$$
- *Define $p^\delta(t, S_t, s) \triangleq e^{r(T-t)} V^\delta(t, S_t, s)$, then*
  $$\int_{-\infty}^{\infty} p^\delta(t, S_t, s)ds = 1.$$
- *For any payoff function $V_T(S_T)$, we have its current value given by*
  $$\begin{aligned} V(t) &= \int_{-\infty}^{\infty} V_T(S_T = s)V^\delta(t, S_t, S_T = s)ds \\ &= e^{-r(T-t)} \int_{-\infty}^{\infty} V_T(S_T = s)p^\delta(t, S_t, S_T = s)ds \\ &= e^{-r(T-t)} E_Q[V_T(S_T)|\mathcal{F}_t] \end{aligned}$$
  where E_Q denotes taking expectation with respect to p^δ.

Proof. (1)(2) Use the fact that risk-free asset grows at rate r at all world states. (3) use linearity of pricing. □

Remark 1.7.3 (interpretation).

- This lemma shows that the nature of risk-neutral probability measure, which is the state price resulted from market equilibrium based on current market condition(time t and stock price S_t).
- The risk-neutral probability/state price therefore will depends on time t and stock price S_t.

1.7.3 Asset paying dividends

1.7.3.1 *Continuous proportional dividend*

Lemma 1.7.3 (martingale properties of asset paying continuous proportional dividend). *Consider an asset S_t following geometric Brownian motion with dividend rate D. Let Q denote the risk-neutral measure. Then,*

- $$e^{-(r-D)t}S_t = E_Q[e^{-(r-a)T}S_T|\mathcal{F}_t],$$
  that is, the random process $e^{-(r-a)T}S_T$ is a martingale with respect to Q.
- $$e^{-r(T-t)}E_Q[S_T|\mathcal{F}_t] = S_t e^{-D(T-t)}$$
- *(Put-call parity)*
  $$C(K,T,S_t) - P(K,T,S_t) = S_t e^{-D(T-t)} - Ke^{-r(T-t)}$$

Proof. (1)(2) recall that under risk-neutral measure, the asset dynamics is given by(Lemma 1.6.4)

$$dS_t/S_t = (r-D)dt + \sigma dW_t.$$

Then use the expectation property of geometric Brownian motion(**??**). (3)

$$\begin{aligned} C - P &= e^{-r(T-t)}E_Q[(S_T-K)^+ - (K-S_T)^+|\mathcal{F}_t] \\ &= e^{-r(T-t)}E_Q[(S_T-K)|\mathcal{F}_t] \\ &= S_t e^{-D(T-t)} - Ke^{-r(T-t)}. \end{aligned}$$

□

Remark 1.7.4 (only valid for deterministic short rate)**.** Only when r is deterministic, we can separate $E_Q[e^{-(r-a)T}S_T|\mathcal{F}_t] = e^{-(r-a)T}E_Q[S_T|\mathcal{F}_t]$; when r is the random, we have to use forward measure method.

Proposition 1.7.2 (Black-Scholes equation with constant dividends). *[15, p. 90] Let $V(S(t),t)$ be the value of the derivative as a function of the asset price $S(t)$ and time t. Assume $S(t)$ is governed by*

$$dS_t = (\mu S_t - D_0 S_t)dt + \sigma S_t dW_t$$

where W_t is the Brownian motion. Then V is governed by

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

with final condition $V(S(T),T) = V_T(S(T))$ and boundary condition $V(S,t) = V_a(t)$ on $S = a$ and $V(S,t) = V_b(t)$ on $S = b$.

Proof. Use Ito's lemma(**??**), we have

$$dV = \sigma S\frac{\partial V}{\partial S}dW_t + (\mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t})dt$$

Construct a portfolio $\Pi = V - \Delta S$, and **note that**

$$d\Pi = dV - \Delta(dS + \Delta S dt) = dV - \Delta dS - \Delta D_0 S dt$$

because the dividends are returned back to the stock holders(see Remark 1.6.15). Then we have

$$d\Pi = \sigma S(\frac{\partial V}{\partial S} - \Delta)dW_t + ((\mu - D_0 S)(\frac{\partial V}{\partial S} - \Delta) + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} - \Delta D_0 S)dt$$

If we set $\frac{\partial V}{\partial S} = \Delta$, then the value of the portfolio Π will evolutes deterministically, we have

$$d\Pi = (\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} - \Delta D_0 S)dt = r\Pi dt = r(V - \Delta S)dt$$

as required by risk free portfolio valuation theorem(Corollary 1.2.1.2). Finally, we have

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

□

Remark 1.7.5 (relation to non-dividend paying model). If we make the change of variable as

$$V(S,t) = e^{-D_0(T-t)}V_1(S,t)$$

then the function V_1 satisfies the canonical Black-Scholes equation with r replaced by $r - D_0$; That is, if V satisfies the canonical Black-Scholes equation, then V_1 satisfies the dividend paying Black-Scholes equation.

1.7.4 Forwards and futures

1.7.4.1 *Forwards contract*

Lemma 1.7.4 (forward price with constant interest rate). *[15, p. 98] Let current time be 0. Let F be the forward/future price of an asset at T. Under martingale/risk-neutral measure Q and constant short rate r,we have*

$$0 = E_Q[\exp(-rT)(S_T - F)] = S_0 - F\exp(-rT),$$

or equivalently,

$$F = S_0 \exp(rT).$$

If the asset is paying an dividend at rate q, then

$$F = S_0 \exp((r-q)T).$$

Proof. (1) Note that the payoff of the contract is $S_T - F$, and the initial price is zero. Using martingale pricing theorem(Theorem 1.6.13), we have $0 = E_Q[\exp -rT(S_T - F)]$. Also from the fact that $\exp(-rT)S_T$ is a martingale under risk-neutral measure , we have $E_Q[\exp -rTS_T] = S_0$.
(2) Similarly to (1). Note that $\exp(-(r-q)T)S_T$ is a martingale under risk-neutral measure , we have $E_Q[\exp -rTS_T] = S_0 \exp(qT)$. Therefore, $F = S_0 \exp((r-q)T)$. □

Proposition 1.7.3 (value of forward contract after initiation). *Let $F(t,T)$ be the forward/future price of an asset at time T. Consider a forward/futures contract initiated at t_0 with forward price $F(t_0,T)$. Let $V(t)$ be the value of the contract at time t. Assuming constant short rate r, we have*

- $$\begin{aligned} V(t) &= S(t) - F(t_0,T)\exp(-r(T-t)) \\ &= S(t) - S(t_0)\exp(r(t-t_0))) \end{aligned}$$
- *If the underlying is paying dividend at a rate q, we have*
$$\begin{aligned} V(t) &= S(t)\exp(-q(T-t)) - F(t_0,T)\exp(-r(T-t)) \\ &= S(t)\exp(-q(T-t)) - S(t_0,T)\exp(r(t-t_0))\exp(-q(T-t_0)) \end{aligned}$$

Proof. (1)

$$\exp(-rt)V(t) = E_Q[\exp(-rT)(S_T - F(t_0,T))|\mathcal{F}_t] \implies V(t) = S_t - F(t_0,T)\exp(-r(T-t)).$$

(2) Note that $\exp(-(r-q)T)S_T$ is a martingale under risk-neutral measure , we have $E_Q[\exp -r(T-t)S_T] = S_0 \exp(-q(T-t))$. Then

$$\exp(-rt)V(t) = E_Q[\exp(-rT)(S_T - F(t_0,T))|\mathcal{F}_t] \implies V(t) = S_t \exp(-q(T-t)) - F\exp(-r(T-t)).$$

□

Note 1.7.2 (simulating forward contract value). To generate the trajectories of the forward contract value evolution, we can simulate the underlying, S_t, and then use the relation between S_t and $V(t)$ in Theorem 1.7.3.

- As showed in Figure 1.7.1, we simulate the underlying in the real world via model

$$dS_t/S_t = \mu dt + \sigma dW_t,$$

and then evaluate the forward contract via

$$V(t) = S_t - S(t_0)\exp(r(t-t_0)).$$

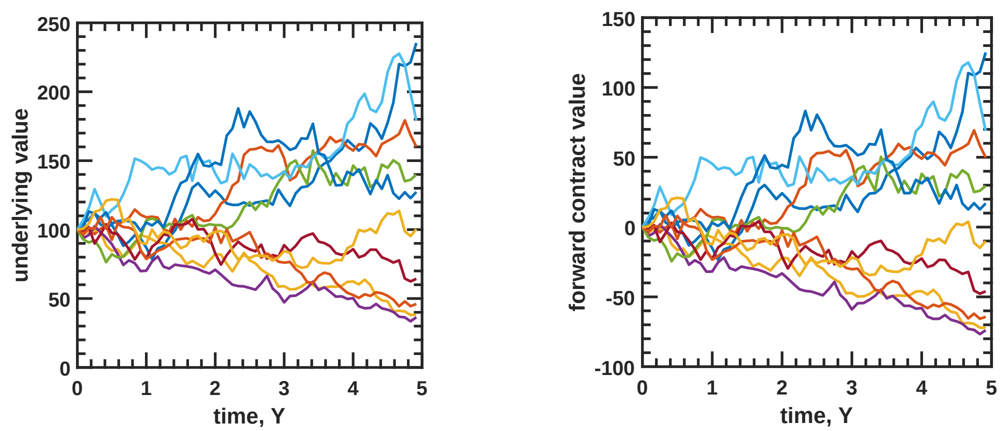
- Note that **we should not simulate the underlying in the risk-neutral measure.** This is because in the risk neutral measure $V(t)$ will have zero expected growth rate,i.e.,

$$V(t) = S(t_0)\exp(r(t-t_0))(M(t)-1),$$

where $M(t)$ is exponential martingale given by

$$M(t) = \exp(-\frac{1}{2}\sigma^2(t-t_0) + \sigma(W(t) - W(t_0))).$$

If a forward contract has zero expected return with risks, then no one wants to enter due to risk aversion even though it costs zero to enter.
- On the other hand, if we simulate the underlying in the real world measure where S_t will have a **different expected growth rate** from r and V_t will grow at a nonzero expected rate(could be positive or negative). Then the market participants will enter either short or long positions based on their estimation of the S_t dynamics.
- Also see [17, p. 145] for a discussion.



(a) Underlying price trajectories. (b) Forward contract value trajectories.

Figure 1.7.1: Demonstration of forward contract value evolution. Underlying simulation parameters: $\mu = 0.06, r = 0.02, \sigma = 0.2, q = 0.0, t_0 = 0.$

Lemma 1.7.5 (forward price with non-constant deterministic interest rate). *Let F be the forward price of an asset. Under the risk-neutral measure Q and non-constant deterministic short rate $r(t)$, we have*

$$0 = E_Q[D(T)(S_T - F)] = S_0 - FD(T),$$

or equivalently,

$$F = S_0/D(T),$$

where $D(t) = \exp(-\int_0^t r(s)ds)$

Proof. Same as the constant short rate case. □

Lemma 1.7.6 (forward price with non-constant deterministic interest rate). *Let F be the forward price of an asset. Under the risk-neutral measure Q and non-constant deterministic short rate $r(t)$, we have*

$$0 = E_Q[D(T)(S_T - F)] = S_0 - FD(T),$$

or equivalently,

$$F = S_0/D(T),$$

where $D(t) = \exp(-\int_0^t r(s)ds)$

Proof. Same as the constant short rate case. □

Lemma 1.7.7 (non-interest-rate asset forward price dynamics with stochastic interest rate). *Let $F(t,T)$ be the current forward price (with maturity T) of an asset S_t uncorrelated with interest rate. Under the risk-neutral measure Q and stochastic short rate $r(t)$, we have*

- $$F(t,T) = \frac{S_t}{P(t,T)}.$$
- ***If assuming independence between*** S_T ***and*** *r, then $F(t,T)$ is martingale under Q; that is*

$$F(t,T) = E_Q[F(T,T)|\mathcal{F}_t].$$

where $P(t,T)$ is the zero-coupon bond price.

Proof. (1)

$$\begin{aligned}
0 &= E_Q[\frac{1}{B(T)}(S_T - F(t,T))|\mathcal{F}_t] \\
&= \frac{S_t}{B(t)} - F(t,T)E_Q[\frac{1}{B(T)}|\mathcal{F}_t] \\
&= \frac{S_t}{B(t)} - F(t,T)P(t,T)
\end{aligned}$$

(2)

$$\begin{aligned}
0 &= E_Q[\frac{1}{B(T)}(S_T - F(t,T))|\mathcal{F}_t] \\
&= E_Q[\frac{1}{B(T)}\mathcal{F}_t]E_Q[(S_T - F(t,T))|\mathcal{F}_t] \\
\implies E_Q[(S_T - F(t,T))|\mathcal{F}_t] &= 0 \\
E_Q[F(T,T)|\mathcal{F}_t] &= F(t,T)
\end{aligned}$$

Note that we use the fact that $(S_T - F(t,T))$ is independent of r(which is usually highly unlikely in real world). □

Lemma 1.7.8 (evolution of forward rate agreement). *Consider a forward contract entered at t_0 with fair rate K. At future time $t_1, t_0 < t_1 < T$, the value of the forward contract is given by*

$$V(t_1) = P(t_1,T)(F(t_0,T) - F(t_1,T)),$$

where $F(t_1,T)$, $F(t_0,T)$ are the fair forward value at t_1 and t_0 and $P(t,T)$ is the T zero coupon bond price at t. ***Note that $P(t_1,T)$, $F(t_1,T)$ and $V(t_1)$ are all random quantities at time t_0.***

Proof. For a contract entered at time t_0, we agree to pay $F(t_0,T)$ at time T to exchange for the asset $S(T)$. At time t_1, a new forward agreement will allow us to pay $F(t_1,T)$ to exchange for the same asset $S(T)$ at time T. The payment $F(t_0,T)$ at time T is worth $P(t_1,T)$ at time t_1, the asset $S(T)$ is worth $P(t_1,T)S(T) = P(t_1,T)F(t_0,T)$. Therefore, the value of the contract at time t_1 is to add the gain of $P(t_1,T)F(t_0,T)$ and subtract the loss $P(t_1,T)F(t_1,T)$, given by

$$V(t_1) = P(t_1,T)(F(t_0,T) - F(t_1,T)).$$

□

1.7.4.2 *Futures contract*

Lemma 1.7.9 (futures price dynamics under risk-neutral measure). *[18, p. 29] Under the risk-neutral measure Q, the underlying asset dynamics is given by*

$$dS_t/S_t = rdt + \sigma dW_t.$$

Consider a futures price $Fur(t)$ on the underlying S_t. It follows that

- *$Fur(t)$ is governed by*
$$\frac{\partial Fur}{\partial t} + \sigma^2 S^2 \frac{\partial^2 Fur}{\partial S^2} + \frac{\partial Fur}{\partial S} rS = 0,$$
with terminal condition $Fur(T, S_T) = S_T$ at maturity T.
- *Assume deterministic short rate $r(t)$, then $Fur(t) = E_Q[S_T|\mathcal{F}_t] = S_t \exp(\int_t^T r(s)ds)$. Moreover, $Fur(t)$ is martingale under Q; that is,*
$$Fur(t;T) = E_Q[S_T|\mathcal{F}_t] = E_Q[Fur(T;T)|\mathcal{F}_t],$$
where T is the maturity time.
- *The dynamics of $Fur(t)$, under risk-neutral measure Q, is given by*
$$dFur(t) = \sigma(t) Fur(t) dW_t.$$

Proof. (1) (a)Consider an investor has a marginal money account M continuous reinvesting in money market and handling the mark-to-market requirement of the futures. Note that the futures contract has no value since the mark-to-market gain/loss will go to the marginal money account.

The dynamics of the money account M is given by

$$dM = rMdt + dFur = rMdt + (\frac{\partial Fur}{\partial t} + \sigma^2 S^2 \frac{\partial^2 Fur}{\partial S^2} + \frac{\partial Fur}{\partial S} rS)dt + \frac{\partial F}{\partial S}\sigma S dW_t.$$

Under risk-neutral measure, the growth rate of any portfolio should be r(Theorem 1.5.4), therefore

$$\frac{\partial Fur}{\partial t} + \sigma^2 S^2 \frac{\partial^2 Fur}{\partial S^2} + \frac{\partial Fur}{\partial S} rS = 0.$$

(b) Another way to derive is to construct riskless portfolio $\Pi = M - \Delta S$, where $\Delta = \partial F/\partial S$. Then

$$d\Pi = (M - \partial F/\partial S S)rdt \implies \frac{\partial Fur}{\partial t} + \sigma^2 S^2 \frac{\partial^2 Fur}{\partial S^2} + \frac{\partial Fur}{\partial S} rS = 0.$$

(2) From Feyman-Kac theorem (Theorem 1.8.1),

$$Fur(t) = E_Q[Fur(T)|\mathcal{F}_t] = E_Q[S_T|\mathcal{F}_t].$$

$$\begin{aligned}
Fur(t) &= E_Q[S_t \exp(\int_t^T r(s)ds) \exp(\int_t^T \sigma(s)ds - \frac{1}{2}\int_t^T \sigma(s)^2 ds)|\mathcal{F}_t] \\
&= S_t \exp(\int_t^T r(s)ds) E_Q[\exp(\int_t^T \sigma(s)ds - \frac{1}{2}\int_t^T \sigma(s)^2 ds)|\mathcal{F}_t]. \\
&= S_t \exp(\int_t^T r(s)ds)
\end{aligned}$$

where we use the fact that $\exp(\int_t^T \sigma(s)ds - \frac{1}{2}\int_t^T \sigma(s)^2 ds)$ is a martingale.

(3) (a)

$$\begin{aligned}
dFur(t) &= d(S_t \exp(\int_t^T r(s)ds)) \\
&= \exp(\int_t^T r(s)ds)dS_t - S_t \exp(\int_t^T r(s)ds)r(t)dt \\
&= \exp(\int_t^T r(s)ds)S_t\sigma(t)dW_t \\
&= Fur(t)\sigma(t)dW_t.
\end{aligned}$$

(b) From Ito lemma,

$$dFur(t) = (\frac{\partial Fur}{\partial t} + \frac{\partial^2 Fur}{\partial S^2}\sigma^2 S^2 + \frac{\partial Fur}{\partial S} rS_t)dt + \frac{\partial Fur}{\partial S}\sigma S_t dW_t = \frac{\partial Fur}{\partial S}\sigma S_t dW_t = Fur(t)\sigma(t)dW_t$$

where we used (1) to simplify and note that $\frac{\partial Fur}{\partial S} = \exp(\int_t^T r(s)ds)$. □

1.8 Connection between Black-Scholes framework and martingale pricing

1.8.1 Pricing theory via Feyman Kac theorem

Proposition 1.8.1 (Feyman Kac theorem). *Consider the 1D parabolic*

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(S,t)\frac{\partial^2 V}{\partial S^2} + m(S,t)\frac{\partial V}{\partial S} - rV = 0.$$

The solution is given as

$$V(s,t) = E_Q[e^{\int_s^t r(\tau)d\tau} V(S_t,t)|\mathcal{F}_s]$$

where S_t is a stochastic process, and Q is probability measure, under which the dynamics of S is given as

$$dS = mdt + \sigma dW_t$$

where W_t is the Wiener process. Moreover,

$$e^{\int_0^t r(\tau)d\tau} V(S_t, t)$$

is a martingale under the measure Q.

Proof.

$$\begin{aligned} d(e^{\int_0^t r(\tau)d\tau} V(S_t, t)) &= d(e^{\int_0^t r(\tau)d\tau}) V(t) + e^{\int_0^t r(\tau)d\tau} dV(t) + d(e^{\int_0^t r(\tau)d\tau}) dV \\ &= -e^{\int_0^t -r(\tau)d\tau} r(t) V dt + e^{-\int_0^t r(\tau)d\tau} dV. \end{aligned}$$

Use the fact that

$$dV = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{1}{2}(dS)^2 \frac{\partial^2 V}{\partial S^2} = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} mdt + \sigma \frac{\partial V}{\partial S} dW_t + \frac{1}{2}(\sigma)^2 \frac{\partial^2 V}{\partial S^2} dt.$$

Then plug in dV, we have

$$\begin{aligned} d(e^{\int_0^t r(\tau)d\tau} V(S_t, t)) &= e^{\int_0^t -r(\tau)d\tau} (-r(t) V dt + \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} mdt + \sigma \frac{\partial V}{\partial S} dW_t + \frac{1}{2}(\sigma)^2 \frac{\partial^2 V}{\partial S^2} dt) \\ &= e^{\int_0^t -r(\tau)d\tau} \sigma \frac{\partial V}{\partial S} dW_t \end{aligned}$$

Therefore, $e^{\int_0^t r(\tau)d\tau} V(S_t, t))$ is a martingale. Then we can easily show using martingale property:

$$V(s) = E_Q[e^{-\int_s^t r(\tau)d\tau} V(S_t, t) | \mathcal{F}_s].$$

□

Proposition 1.8.2 (Feyman Kac theorem, alternative). *[9, p. 268] Assume X_t is governed by the following SDE*

$$dX_t = \mu dt + \sigma dW_t.$$

Suppose we are given a payoff $V(X_T)$ at time T. Define

$$f(x, t) = E[e^{-\int_t^T r(u)du} V(X_T) | X_t = x]$$

for all $t \leq T$, where r is a non-random function. Then $f(x,t)$ is governed by

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

for $t \leq T$, subject to the final condition $f(x,T) = V(X_T)$.

Moreover, $E[e^{-\int_0^T r(u)du} V(X_T)|X_t = x]$ is a martingale.

Proof. $e^{-\int_0^t r(u)du} f(x,t)$ is a martingale. Therefore,

$$\begin{aligned} d(e^{-\int_0^t r(u)du} f(x,t)) &= e^{-\int_0^t r(u)du}(f_t + f_x dX + \frac{1}{2} f_{xx} dX dX - rf) \\ &= e^{-\int_0^t r(u)du}(f_t + \mu f_x dt + \frac{1}{2} f_{xx}\sigma^2 - rf)dt + e^{-\int_0^t r(u)du}\sigma f_x dW. \end{aligned}$$

Set the drift term to zero, we get the PDE. □

Remark 1.8.1 (interpretation and financial applications in path-independent derivatives)**.**

- Feyman Kac theorem shows that certain types of parabolic equation can be solve using stochastic differential equation method(by simulating trajectories and take expectations.) **Note that in parabolic differential equation S is not a random variable, S is simply a variable.**
- **Path independence**. Note that the function $V(S_t, t)$ can only be a function of instantaneous value S_t. If V is dependent on the history of S, then we cannot use Feyman Kac theorem.

Remark 1.8.2 (special case of $r = 0$)**.** When $r = 0$, the dynamics of S will not change(i.e. Q will not change), then $V(S_t, t)$ is a martingale. And the parabolic equation becomes Kologorov backward equation.

Proposition 1.8.3 (Kolmogorov backward equation, expectation pricing). *Assume X_t is governed by the following SDE*

$$dX_t = \mu dt + \sigma dW_t.$$

Suppose we are given a payoff $V(X_T)$ at time T. Define $f(x,t) = E[V(X_T)|X_t = x]$ for all $t \leq T$. Then $f(x,t)$ is governed by the ***Kolmogorov backward equation*** *given as*

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

for $t \leq T$, subject to the final condition $f(x,T) = V(X_T)$.

Proof. Note that

$$E[dX_t|cF_t] = E[dX_t|X_t = x] = a(X_t, t)dt$$

and

$$E[dX_t^2|cF_t] = E[dX_t^2|X_t = x] = b^2(X_t, t)dt.$$

then we have (use the notation $E[f(X_s, s)|X_t = t] = E_{x,t}[f(X_s, s)]$)

$$\begin{aligned}
E_{x,t}[f(X_{t+dt}, t+dt)] &= E_{x,t}[f(x + dX_t, t + dt)] \\
&\approx E_{x,t}[f(x,t) + \partial_x f(x,t)dX_t + \frac{1}{2}\partial_x^2 f(x,t)dX_t^2 + \partial_t f(x,t)dt] \\
&= f(x,t) + \partial_x f(x,t)E_{x,t}[dX_t] + \frac{1}{2}\partial_x^2 f(x,t)E_{x,t}[dX_t^2] + \partial_t f(x,t)dt \\
&= f(x,t) + dt(\partial_t f(x,t) + \mu(x,t)\partial_x f(x,t) + \frac{\sigma^2(x,t)}{2}\partial_x^2 f(x,t))
\end{aligned}$$

$$E_{x,t}[f(X_{t+dt}, t+dt)] = f(x,t) \implies \partial_t f(x,t) + \mu(x,t)\partial_x f(x,t) + \frac{\sigma^2(x,t)}{2}\partial_x^2 f(x,t) = 0$$

□

Remark 1.8.3 (interpretation)**.**

- We can interpret $f(x,t)$ as the price at t and the state is at x. When we cannot hedge the risk, we can price the asset by the expected value of payoff with respect to the real probability. Note that such pricing method does not take into account of the risk-aversion.
- The only difference to Black-Scholes is the extra source decreasing term.

Proposition 1.8.4 (Kolmogorov backward equation with discount). *Assume X_t is governed by the following SDE*

$$dX_t = \mu dt + \sigma dW_t.$$

Suppose we are given a payoff $V(X_T)$ at time T. Define $f(x,t) = E[e^{-\int_t^T r(tau)d\tau}V(X_T)|X_t = x]$ for all $t \leq T$. Then $f(x,t)$ is governed by the ***Kolmogorov backward equation*** *given as*

$$-\frac{\partial}{\partial t}f(x,t) = \mu\frac{\partial}{\partial x}f(x,t) + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}f(x,t) - rf$$

for $t \leq T$, subject to the final condition $f(x,T) = V(X_T)$.

Proof. Note that

$$E[dX_t|cF_t] = E[dX_t|X_t = x] = a(X_t, t)dt$$

and

$$E[dX_t^2|cF_t] = E[dX_t^2|X_t = x] = b^2(X_t, t)dt.$$

then we have (use the notation $E[f(X_s, s)|X_t = t] = E_{x,t}[f(X_s, s)]$)

$$\begin{aligned} E_{x,t}[(1-rdt)f(X_{t+dt}, t+dt)] &= E_{x,t}[(1-rdt)f(x+dX_t, t+dt)] \\ &\approx (1-rdt)E_{x,t}[f(x,t)+\partial_x f(x,t)dX_t + \frac{1}{2}\partial_x^2 f(x,t)dX_t^2 + \partial_t f(x,t)dt] \\ &= f(x,t) - rf(x,t)dt + \partial_x f(x,t)E_{x,t}[dX_t] + \frac{1}{2}\partial_x^2 f(x,t)E_{x,t}[dX_t^2] + \partial_t f(x,t)dt \\ &= f(x,t) - rf(x,t)dt + dt(\partial_t f(x,t) + \mu(x,t)\partial_x f(x,t) + \frac{\sigma^2(x,t)}{2}\partial_x^2 f(x,t)) \end{aligned}$$

$$\begin{aligned} &E_{x,t}[(1-rdt)f(X_{t+dt}, t+dt)] = f(x,t) \\ &\implies \partial_t f(x,t) + \mu(x,t)\partial_x f(x,t) + \frac{\sigma^2(x,t)}{2}\partial_x^2 f(x,t) - rf(x,t) = 0 \end{aligned}$$

□

Proposition 1.8.5 (Kolmogorov backward equation, multi-dimensional version). *Assume X_1, X_2 is governed by the following SDE*

$$dX_1(t) = \mu_1 dt + \sigma_1 dW_1(t).$$

and

$$dX_2(t) = \mu_2 dt + \sigma_2 dW_2(t).$$

Suppose we are given a payoff $V(X_1(T), X_2(T))$ at time T. Define

$$f(x_1, x_2, t) = E[e^{-\int_t^T r(\tau)d\tau} V(X_1(T), X_2(T)|X_1(t) = x_1, X_2(t) = x_2]$$

*for all $t \leq T$. Then $f(x_1, x_2, t)$ is governed by the **Kolmogorov backward equation** given as*

$$-\frac{\partial}{\partial t} f(x_1, x_2, t) = \mu_1 \frac{\partial}{\partial x_1} f + \mu_2 \frac{\partial}{\partial x_2} f + + \frac{1}{2}\sigma_1^2 \frac{\partial^2}{\partial x_2^2} f + \frac{1}{2}\sigma_2^2 \frac{\partial^2}{\partial x_2^2} f + \sigma_1\sigma_2 \frac{\partial^2}{\partial x_1 \partial x_2} f - rf$$

for $t \leq T$, subject to the final condition $f(x_1, x_2, T) = V(X_1(T), X_2(T))$.

Proof. Similar to 1D case. □

Proposition 1.8.6 (Feyman Kac theorem, multi-dimensional). *Consider the multidimensional parabolic*

$$\frac{\partial V}{\partial t} + \sum_{i=1}^N \mu_i \frac{\partial V}{\partial S} + \frac{1}{2}\sum_{i=1}^N \sum_{j=1}^N \gamma_{ij} \frac{\partial^2 V}{\partial S_i \partial S_j} + -rV = 0$$

where $\gamma_{ij} = \sum_{k=1}^{N} \sigma_{ik}\sigma_{jk}$. *The solution is given as*

$$V(s,t) = E_Q[e^{\int_s^t r(\tau)d\tau} V(S_t,t)|\mathcal{F}_s]$$

where S_t *is a N dimensional stochastic process, and Q is probability measure, under which the dynamics of S is given as*

$$dS_i = \mu_i dt + \sum_{j=1}^{N} \sigma_{i,j} dW_j(t), i = 1,2,...,N$$

where W_t *is the Wiener process. Moreover,*

$$e^{\int_0^t r(\tau)d\tau} V(S_t,t)$$

is a martingale under the measure Q.

Proof. Similar to 1D case. □

1.8.2 Model dynamics under risk-neutral measure

Lemma 1.8.1 (stock dynamics under risk-neutral measure). *Assume the stock* S_t *in the real-world probability measure follows*

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

where W_t *is a Wiener process. Then under the risk-neutral measure Q, the stock dynamics will be*

$$dS_t = rS_t dt + \sigma S_t d\hat{W}_t,$$

where $\hat{W}_t$ *is a Brownian motion under risk-neutral measure Q, no matter what original* μ *is.*

Proof. Based on the definition of risk-neutral measure(Definition 1.6.7), it can be show that

$$d(D(t)S_t) = D(t)dS_t + S_t dD(t) = D(t)(rS_t dt + \sigma S_t dW_t) - rD(t)S_t dt = D(t)\sigma S_t dW_t.$$

□

1.9 Notes on bibliography

For no-arbitrage theory, see [19].

For treatment from economical perspective, see [20] [21].

For martingale methods, see [22].

For PDE methods, see [22][23].

For incomplete markets, see [24].

For the differences between real world measure and risk-neutral measure, see [25].

An excellent book of "P" method for risks and asset allocations, see [26].

For a comprehensive discussion on risks, see [27].

For statistical arbitrage, see [28].

BIBLIOGRAPHY

1. Black, F. & Scholes, M. The pricing of options and corporate liabilities. *Journal of political economy* **81,** 637–654 (1973).
2. Harrison, J. M. & Pliska, S. R. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic processes and their applications* **11,** 215–260 (1981).
3. Stefanica, D. *A primer for the Mathematics of Financial Engineering* (Fe Press, 2008).
4. Pliska, S. *Introduction to Mathematical Finance: Discrete Time Models* (Blackwell: Malden, MA, 1997).
5. Hull, J. *Options, Futures and Other Derivatives 9 edition* (Pearson Education, 2009).
6. Primbs, J. A. *A Factor Model Approach to Derivative Pricing* (CRC Press, 2016).
7. Derman, E., Miller, M. & Park, D. *The Volatility Smile* ISBN: 9781118959169. https://books.google.com/books?id=WKfBDAAAQBAJ (Wiley, 2016).
8. Brigo, D. & Mercurio, F. *Interest rate models-theory and practice: with smile, inflation and credit* (Springer Science & Business Media, 2007).
9. Shreve, S. E. *Stochastic calculus for finance II: Continuous-time models* (Springer Science & Business Media, 2004).
10. Shreve, S. *Stochastic Calculus for Finance I: The Binomial Asset Pricing Model* (Springer Science & Business Media, 2012).
11. Chin, E., Nel, D. & Ãlafsson, S. *Problems and Solutions in Mathematical Finance: Equity Derivatives* (John Wiley & Sons, 2017).
12. Oksendal, B. *Stochastic differential equations: an introduction with applications* (Springer Science & Business Media, 2013).
13. McInerney, D. & Zastawniak, T. *Stochastic Interest Rates* ISBN: 9781107002579. https://books.google.com/books?id=8JTwCAAAQBAJ (Cambridge University Press, 2015).
14. Joshi, M. S. *The concepts and practice of mathematical finance* (Cambridge University Press, 2003).
15. Lalley, S. *The mathematics of financial derivatives: a student introduction* (Chicago University Statistics 390/ Mathematical Finance 345, 2001).
16. Pelsser, A. *Efficient methods for valuing interest rate derivatives* (Springer Science & Business Media, 2000).
17. Hull, J. *Risk management and financial institutions,+ Web Site* (John Wiley & Sons, 2012).

18. Clark, I. J. *Commodity Option Pricing: A Practitioner's Guide* (John Wiley & Sons, 2014).

19. Björk, T. *Arbitrage theory in continuous time* (Oxford university press, 2004).

20. LeRoy, S. F. & Werner, J. *Principles of financial economics* (Cambridge University Press, 2014).

21. Mishkin, F. S. *The economics of money, banking, and financial markets* (Pearson education, 2007).

22. Pascucci, A. *PDE and martingale methods in option pricing* (Springer Science & Business Media, 2011).

23. Wilmott, P. *Paul Wilmott on quantitative finance* (John Wiley & Sons, 2013).

24. Cernỳ, A. *Mathematical techniques in finance: tools for incomplete markets* (Princeton University Press, 2009).

25. Meucci, A. 'P'Versus' Q': Differences and Commonalities between the Two Areas of Quantitative Finance. *GARP Risk Professional*, 47–50 (2011).

26. Meucci, A. *Risk and asset allocation* (Springer Science & Business Media, 2009).

27. Jorion, P. *Financial Risk Manager Handbook: FRM Part I / Wiley Finance* **pt. 2.** ISBN: 9781118017913. https://books.google.com/books?id=4ceVmGJSNpcC (Wiley, 2010).

28. Leung, T. & Li, X. *Optimal Mean Reversion Trading: Mathematical Analysis and Practical Applications* (World Scientific, 2015).

2

MARKET INSTRUMENT, REPLICATION AND MODELS

2.1 Market implied distributions

2.1.1 Overview

Various kinds of financial instruments, including forwards, futures, swaps and options, are actively traded in the market. Their current prices and historical prices reveal critical information that can be employed to give no-arbitrage price for other financial instruments whose prices not observed in the market.

We use previously developed martingale pricing framework to interpret the price information and price other instruments via replication.

In general, forwards, futures and swaps reveal the market implied mean values of a financial random variable, such as interest rate, spot price, forward price, realized volatility, etc.

Vanilla option like European call and put with sufficient number strikes can enable reconstruction of distribution of the underlying at a specified time.

Forward starting products or options on options can further illuminate conditional probabiliy density of underlying asset stochastic process.

2.1.2 Implied distributions

2.1.2.1 *Implied distribution and risk-neutral measure*

Definition 2.1.1 (implied distribution of underlying asset). *Consider a no-arbitrage market probabilitic model $(\Omega, \mathcal{F}, P)$ such that there exists an equivalent martingale measure, called risk-neutral measure Q. Let current time be 0. Let the asset spot price S_0 be observed in the market.*

- *The implied distribution of an asset value S_t at future time T is the distribution under measure Q* ***conditioning on*** *S_0.*
- *More concretely, given the call/puts market prices $C(K, T, S_0)$ at maturity T. The* ***implied distribution*** *f_{S_T} of S_T is the distribution function such that*

$$C(K, T, S_0) = e^{-rT} \int_0^\infty f_{S_T}(s)(S_T - K)^+ ds, \forall K;$$

that is, implied distribution is the one that produces market prices of all calls/puts.

Remark 2.1.1 (uniqueness of implied distribution).

- Given a finite set of liquid instruments such as calls/puts, there might exist multiple implied distribution such that the distribution of the underlying asset reprice these liquid instruments.
- When calls/puts are actively traded for all strikes $K \in \mathbb{R}^+$, the implied distribution can be uniquely determined by methods in Lemma 2.1.2, Lemma 2.1.4.

Proposition 2.1.1 (no-arbitrage pricing of path-independent options using implied distribution). *Let current time be 0. Let the asset spot price S_0 be observed in the market. Assume we have the implied distribution f_{S_T} for the underlying S_T derived from vanilla option market* *prices. Denote constant risk-free rate by r. It follows that an option with arbitrary payoff $V(S_T)$ at maturity has current value*

$$V_0 = e^{-rT}E_Q[f(S_T)] = e^{-rT}\int_{-\infty}^{\infty} V(s)f_{S_T}(s)ds.$$

Note that in reality, the implied probability for $S_T < 0$ is zero.

Proof. Directly use the definition of implied distribution and expectation under risk-neutral measure. □

Definition 2.1.2 (risk-neutral probability,state price). *Consider a probabilistic model $(\Omega, \mathcal{F}, P)$ for the real world. Let Q be an equivalent measure to P and satisfy*

$$Q(A) = V_0, \forall A \in \mathcal{F},$$

where V_0 is the current price of an asset with payoff $1_A e^{rT}$ at maturity T, and r is the constant interest rate.

*Then, we call Q **risk-neutral probability measure**, or (discounted) state price.*

Note 2.1.1 (constraints on risk-neutral probability/state price)**.**

- The risk-neutral probability/state price can be any functional form as long as it is **equivalent**(Definition 1.6.7) to real-world probability; risk-neutral probability is the resulting equilibrium from market forces.
- **(why it should be equivalent?)** Suppose there is a positive price on a payoff event with zero probability to happen. Then the owner of this asset will try to sell it but no one wants to buy it since it cannot occur. As a result, the price will drop to zero. Similarly, suppose there is a zero price for a future payoff greater than 0. This is impossible since buyers will drive the price up.
- Each market participants have their own point of view of the market, their own risk-preference, and their own reserved prices for the assets. Their selling and buying behavior will determine the equilibrium price, or the state price. As long as there is few arbitrageurs, the market will be arbitrage free.
- When market reaches equilibrium and there is no arbitrage opportunities, we should be able to deduce the state price from the existing markets prices of all assets.

Note 2.1.2. [1, p. 177]

- (why is discounted state price) Q_S represents a 'true' state price, then a zero coupon bond has present value given by $e^{-rT} = \int_\Omega dQ_s$, which makes Q_s not satisfying the sum-to-one condition of probability measure. If we define $Q \triangleq e^{rT}Q$, then $\int_\Omega dQ = 1$.

- risk-neutral measure/probability si not the real probability describing likelihood of event occurrence; but it shares important common features of real probability, for example, non-negativeness, the sum of probability of all mutually exclusive event should equal 1.
- (pricing is model independent)We usually assume the real market has no arbitrage opportunities. When the market has prices along all strikes available, the linearity of pricing(Theorem 1.2.2) gives a no-arbitrage price on the spike(delta function) payoff the law of one price(Theorem 1.2.1) plus the convolution theorem enables us to derive a no-arbitrage price for options with arbitrage pay-off profile. **It does not rely on any extra model assumptions**(e.g., Brownian motion model).
- (constraints on risk-neutral probability/state price) The risk-neutral probability/state price can be any functional form as long as it is equivalent to real-world probability; risk-neutral probability is the resulting equilibrium from market forces.

Note 2.1.3 (implied distribution pricing is no-arbitrage pricing; relation to model-based pricing)**.**

- We usually assume the real market has no arbitrage opportunities. When the market has prices along all strikes available, the linearity of pricing(Theorem 1.2.2) gives a no-arbitrage price on the spike(delta function) payoff the law of one price(Theorem 1.2.1) plus the convolution theorem enables us to derive a no-arbitrage price for options with arbitrage pay-off profile. **It does not rely on any extra model assumptions**(e.g., Brownian motion model).
- Implied distribution is not real-world distribution, but is the state price. See 1.6.2,Theorem 1.3.1.
- When we assume an asset model(e.g., Brownian motion), we have to take great effort to find out the model dynamics under risk-neutral measure. Implied distribution pricing is theoretically more general than model based pricing.

Example 2.1.1. Assume that the risk neutral probability distribution for stock XYZ at the end of one year is uniform between 100 and 200. Assume the interest rate is 0.1. What is the value of a security that pays 1 if XYZ is between 140 and 151?

From the risk-neutral pricing theorem we have

$$V = e^{-0.1}\int_{140}^{151} f_{S_T}(x)dx = e^{-0.1}\frac{11}{100}$$

2.1.2.2 *Implied distribution from derivatives*

Lemma 2.1.1 (derivatives of calls as step and Dirac delta function). *[2, p. 34] Let current time be 0. Let the asset spot price S_0 be observed in the market. Let $C(K,T), P(K,T)$ be the call/put price with strike K and maturity T. It follows that*

- $$\frac{dC}{dK} = -H(S_T - K), \frac{dP}{dK} = H(K - S_T).$$
- $$\frac{d^2C}{dK^2} = \frac{d^2P}{dK^2} = \delta(S_T - K).$$
- *Let $D(K,T)$ be a digital call also on S_T with strike K. Then $D(K,T) = -\frac{dC}{dK}$.*

Proof. (informal)

- This is the definition of step function.
- Let $C(x;K)$ denote function $C(x;K) \triangleq (x-K)^+$. Then,
$$\lim_{\epsilon \to 0} \frac{C(x;K-\epsilon) - 2C(x;K) + C(x;K+\epsilon)}{\epsilon^2} = \frac{d^2C}{dK^2} = \delta(x-K),$$
where $\delta(x-K)$ is the Dirac delta function with spike located at K.
- $C(x;K+\epsilon) - C(x;K)$ is a call spread; long a call spread $C(x;K+\epsilon) - C(x;K)$ and short a call spread $C(x;K) - C(x;K-\epsilon)$ will generate a triangle spike with width 2ϵ and height ϵ and area ϵ^2.

□

Lemma 2.1.2 (implied distribution from digital call/put). *Let current time be 0. Let the asset spot price S_0 be observed in the market. Assume constant risk-free rate r. Consider market prices of digital call/puts available along a continuum of strikes K with the same maturity T. Then the implied distribution distribution*

$$f_{S_T}(S_T = K) = -e^{rT}\frac{dD(K)}{dK}.$$

where $D(K)$ is the current price of digital call at strike K.

Proof. From the martingale pricing theorem, we have

$$D(K) = e^{-rT}\int_0^\infty H(x-K)f_{S_T}(x)dx,$$

where $H(x-k)$ is the step function. Use the fact that derivative of a step function is Dirac delta, we have

$$D'(K) = e^{-rT}\int_0^\infty -\delta(x-K)f_{S_T}(x)dx = -e^{-rT}f_{S_T}(K).$$

□

Lemma 2.1.3 (implied distribution from Dirac-delta option). *Let current time be 0. Let the asset spot price S_0 be observed in the market. Consider a special call option on S_T with payoff $\delta(S_T-K)$. If the market prices of $Delta(K)$ are available along a continuum of strike K with same maturity T. Then the implied distribution is given by*

$$f_{S_T}(S_T-K) = e^{rT}Delta(K).$$

Proof.

$$Delta(K) = e^{-rT}E_Q[\delta(S_T-K)] = f_{S_T}(S_T=K).$$

□

Lemma 2.1.4 (implied distribution from European call/put). *Let current time be 0. Let the asset spot price S_0 be observed in the market. Assume constant risk-free rate r. Consider market prices of European call/put available along a continuum of strikes K with the same maturity T. Then the implied distribution distribution*

$$f_{S_T}(S_T=K) = e^{rT}\frac{d^2C(K)}{dK^2} = e^{rT}\frac{d^2P(K)}{dK^2},$$

where $C(K), P(K)$ are the prices of call/put at strike K.

Proof. From the martingale pricing formula, we have

$$\begin{aligned}
C(K) &= e^{-rT}\int_K^\infty (s-K)f_{S_T}(s)ds \\
&= e^{-rT}\int_K^\infty (s-K)f_{S_T}(s)ds \\
\implies e^{rT}\frac{dC(K)}{dK} &= \int_K^\infty \frac{d(s-K)}{dK}f_{S_T}(s)ds \\
e^{rT}\frac{d^2C(K)}{dK^2} &= f_{S_T}(K).
\end{aligned}$$

The equality for the put is directly from Lemma 5.2.9. □

Note 2.1.4 (implied distribution is state price or risk-neutral measure).

- We usually assume the real market has no arbitrage opportunities. Then no-arbitrage condition can ensure there existence a state price vector or risk-neutral measure. With the state price vector, we can price any payoff by

$$V = E_Q[V(T)],$$

which is from the linearity of pricing(Theorem 1.2.2) and the law of one price(Theorem 1.2.1), and **is irrelevant to any specific asset models**(e.g., Brownian motion model).
- The implied distribution is actually the state price or risk-neutral measure. See 1.6.2,Theorem 1.3.1.
- **(non-uniqueness)** Note that density function maybe only related to derivatives of option prices, denoted here by $C(K)$. Therefore, for different $C(K)$ profiles(e.g., $C(K)$ profiles differ only by a linear function), the density function will be the same.

Note 2.1.5 (features of implied distribution in practice). [2, p. 35]

- The implied distribution usually shows that there exists a long tail on the lower price side, whereas the lognormal distribution stipulates there is a long tail on the higher price side.
- That is, contradicting to the model assumptions(the underlying follows a geometric Brownian motion with constant volatility), the market believes that there is higher likelihood that the market will experience a crash than a large rise.

2.1.2.3 *Implied distribution from derivatives: discrete case*

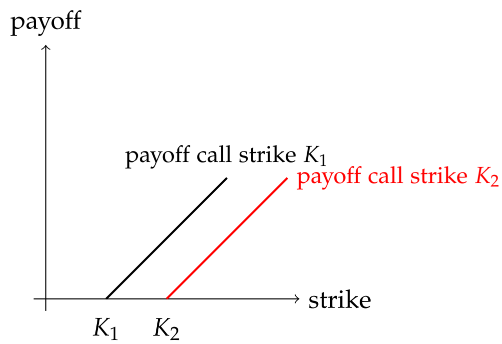
Lemma 2.1.5 (construct implied distribution from discrete call options). *Suppose we have observed market prices $V_0, V_1, V_2, ..., V_n$ for a series of call options on the same underlying S_T with the same expiry T but different strikes $K_0 < K_2 < K_3 < \cdots < K_n$.*

Suppose the implied density of S_T takes piecewise constant, that is, $Pr(S_T \in [K_{i-1}, K_i]) = p_i(K_i - K_{i-1}), i = 1, 2, ..., n$.

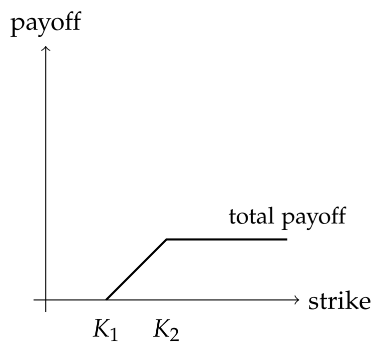
Then the following linear system

$$p_i \frac{K_i - K_{i-1}}{2} + \sum_{j>i}^{n} p_i(K_i - K_{i-1}) = V_i - V_{i-1}, i = 1, 2, ..., n$$

can be used to solve $p_i, i = 1, 2, ..., n$.



(a) payoff of calls at strike K_1, K_2



(b) payoff of $C_1 - C_2$

Figure 2.1.1: Use call option difference to find risk neutral density

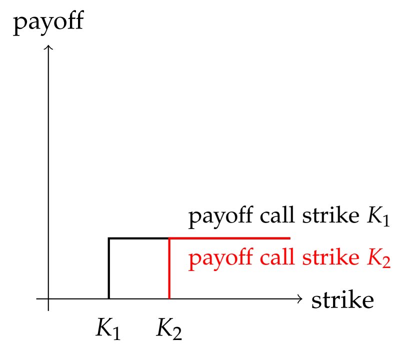
Lemma 2.1.6 (construct risk-neutral density from digital options). *Suppose we have observed market prices $V_0, V_1, V_2, ..., V_n$ for a series of call options on the same underlying S_T with the same expiry T but different strikes $K_0 < K_2 < K_3 < \cdots < K_n$.*

Suppose the implied density of S_T takes piecewise constant, that is, $Pr(S_T \in [K_{i-1}, K_i]) = p_i(K_i - K_{i-1}), i = 1, 2, ..., n$.

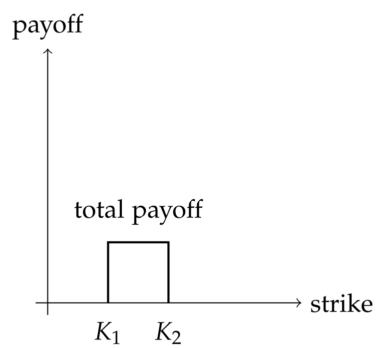
Then the following identities

$$p_i = \frac{V_i - V_{i-1}}{K_i - K_{i-1}}, i = 1, 2, ..., n$$

can be used to solve $p_i, i = 1, 2, ..., n$.



(a) payoff of calls at strike K_1, K_2



(b) payoff of $C_1 - C_2$

Figure 2.1.2: Use digital option difference to find risk neutral density

2.1.3 Application in replicating strategies

2.1.3.1 *Continuous static hedging path-independent derivatives*

Proposition 2.1.2 (payoff function decomposition lemma). *[3] Given a twice-continuously differentiable function $f(x)$(assume $f(x) = 0, \forall x < 0$), it has following decomposition:*

- $$f(x) = \int_0^\kappa f(K)\delta(x-K)dK + \int_\kappa^\infty f(K)\delta(x-K)dK;$$

- $$\begin{aligned} f(x) &= f(K)1(x<K)|_0^\kappa - \int_0^\kappa f'(K)1(x<K)dK \\ &\quad - f(K)1(x \geq K)|_\kappa^\infty + \int_\kappa^\infty f'(K)1(x \geq K)dK \\ &= f(\kappa) - \int_0^\kappa f'(K)1(x<K)dK + \int_\kappa^\infty f'(K)1(x \geq K)dK \end{aligned}$$

- $$f(x) = f(\kappa) + f'(\kappa)(x-\kappa) + \int_0^\kappa f''(K)(K-x)^+ dK + \int_\kappa^\infty f''(K)(x-K)^+ dK;$$

Proof. (1)straight forward; (2) use the properties of a step function that

$$\frac{d1(x<K)}{dK} = \delta(K-x) = \delta(x-K)$$

and

$$\frac{d1(x \geq K)}{dK} = -\delta(K-x) = -\delta(x-K).$$

(3) use the step function definition

$$\max\{K-x, 0\} = \frac{dH(K-x)}{dK}$$

□

Lemma 2.1.7 (Path-independent option hedging using digital calls). *Assume we have the implied distribution f_{S_T} for the underlying S_T derived from vanilla option market prices. It follows that an option with arbitrary payoff $V(S_T)$ at maturity has current value*

$$V_0 = \int_0^\infty V(s) f_{S_T}(s) ds = V(0) + \int_0^\infty V'(K) D(K) dK.$$

The hedging strategy is: long zero-coupon bonds in quantity $V(0)e^{-rT}$ and long all digital calls struck at $K > 0$ in quantities $V'(K)$.

Proof.

$$\begin{aligned} V_0 &= -\int_0^\infty V(K)D'(K)dK \\ &= -[V(K)D(K)]_0^\infty + \int_0^\infty V'(K)D(K)dK \\ &= V(0)D(0) - \int_0^\infty V'(K)D(K)dK \\ &= V(0) + \int_0^\infty V'(K)D(K)dK \end{aligned}$$

where we the fact of $D(0) = 1$.

(2) We can also directly use the decomposition lemma Theorem 2.1.2, and set $\kappa = 0$ such that

$$V(S_T) = V(0) + V'(0)S_T + \int_0^\infty V''(K)(S_T - K)^+ dK.$$

□

Lemma 2.1.8 (Path-independent option hedging using European calls). *[2, p. 36][1, p. 188] Assume we have the implied distribution f_{S_T} for the underlying S_T derived from vanilla option market prices. It follows that an option with arbitrary payoff $V(S_T)$ at maturity has current value*

$$V_0 = \int_{-\infty}^\infty V(s)f_{S_T}(s)ds = V(0)e^{-rT} + V'(0)S_0 + \int_0^\infty V''(K)c(K)dK.$$

The hedging strategy *is: long zero-coupon bonds in quantity $V(0)e^{-rT}$, long underlying stock in quantity $V'(0)$, and long all vanilla calls struck at $K > 0$ in quantities $V''(K)$.*

Proof. (1)Let $c(K)$ represents the market prices of calls with strike K and maturity date T.

$$\begin{aligned} V_0 &= \int_0^\infty V(K)c''(K)dK \\ &= [V(K)c'(K)]_0^\infty - \int_0^\infty V'(K)c'(K)dK \\ &= -V(0)c'(0) + [V'(K)c(K)]_0^\infty + \int_0^\infty V''(K)c(K)dK \\ &= -V(0)c'(0) + V'(0)c(0) + \int_0^\infty V''(K)c(K)dK \\ &= V(0)e^{-rT} + V'(0)S_0 + \int_0^\infty V''(K)c(K)dK \end{aligned}$$

where $c'(0) = -e^{-rT}$, $c(0) = S_0$.

Consider a portfolio contains a zero-coupon bond of quantity $V(0)$, stock of quantity $V'(0)$, and all vanilla calls struck at $K > 0$ in quantities $V''(K)$. Its current value is given by

$$\begin{aligned} V_0 &= e^{-rT} E_Q[V(0) + S_T + \int_0^\infty V''(K)(S_T - K)^+ dK] \\ &= V(0)e^{-rT} + V'(0)S_0 + \int_0^\infty V''(K)c(K)dK \end{aligned}$$

where E_Q is taken expectation with respect to the implied distribution such that

$$S_0 = e^{-rT} E_Q[S_T], c(K) = e^{-rT} E_Q[(S_T - K)^+].$$

(2) We can also directly use the decomposition lemma Theorem 2.1.2, and set $\kappa = 0$ such that

$$V(S_T) = V(0) + V'(0)S_T + \int_0^\infty V''(K)(S_T - K)^+ dK.$$

□

Lemma 2.1.9 (Path-independent option hedging using European calls and puts). *[2, p. 36][1, p. 188] Assume we have the implied distribution f_{S_T} for the underlying S_T derived from vanilla option market prices. It follows that an option with arbitrary payoff $V(S_T)$ at maturity has current value*

$$V_0 = \int_{-\infty}^\infty V(s) f_{S_T}(s)ds = (V(A) - V'(A)A)e^{-rT} + V'(A)S_0 + \int_0^A V''(K)p(K)dK + \int_A^\infty V''(K)c(K)dK.$$

where A is an arbitrary nonnegative number.

***The hedging strategy** is: long zero-coupon bonds in quantity $(V(A) - V'(A)A)e^{-rT}$, long the underlying stock in quantity $V'(A)$, long all vanilla puts struck at $K < A$ in quantities $V''(K)$, and long all vanilla calls struck at $K > A$ in quantities $V''(K)$.*

Proof. We can also directly use the decomposition lemma Theorem 2.1.2, and set $\kappa = A$. □

Remark 2.1.2 (practical importance in using both calls and puts). In reality, puts with small strikes are much more liquid than calls with small strikes, since these puts serve as insurance of market crash and are in high demand.

Example 2.1.2 (replicating log contract using vanilla options). Lemma 5.2.17

- The final payoff of a log contract, $L_T = \ln(S_T/K)$, can be decomposed as

$$\ln(S_T/K) = \frac{S_T - K}{K} - \int_K^\infty \frac{1}{v^2}(S_T - v)^+ dv - \int_0^K \frac{1}{v^2}(v - S_T)^+ dv.$$

- The replicating strategy is:
  - long $1/K$ unit of forward with strike price K.
  - short $1/v$ unit of put options at strike v where v ranges from 0 to K.
  - short $1/v$ unit of call options at strike v where v ranges from K to ∞.

2.1.3.2 *Discrete static hedging path-independent derivatives*

Lemma 2.1.10 (replicating positively-supported piece-wise linear continuous payoff function via calls). *[1, pp. 41, 78] Consider a piece-wise linear continuous payoff function $f(x)$ satisfying*

- $f(x) = 0, \forall x < 0$;
- *On x axis, $K_0 = 0$, $K_1 < K_2 < ...$ are successive kink points.*
- *$f(x)$ intercept at y axis I.*
- *$\lambda_1, \lambda_2, ...$ are slopes of successive linear pieces.*

Let $C(K)$ denote the current market price of a call option with strike K and same maturity. Then, the current value is given by

$$V(t) = Ie^{-r(T-t)} + \lambda_1(C(K_0) - C(K_1)) + \lambda_2(C(K_1) - C(K_2)) + \ldots.$$

Note that $S_t = C(K_0 = 0)$.

Proof. The final payoff can be constructed using the following functions

$$f(x) = I + \lambda_1((x - K_0)^+ - (x - K_1)^+) + \lambda_2((x - K_1)^+ - (x - K_2)^+) + \ldots,$$

since $f(x) \triangleq (x - K_0)^+ - (x - K_1)^+$ is a linear segment given by

$$f(x) = \begin{cases} 0, x \leq K_0 \\ x - K_0, K_0 < x < K_1 \\ K_1 - K_0, x \geq K_1 \end{cases}.$$

□

Lemma 2.1.11 (replicating positively-supported piece-wise linear continuous payoff function via puts). *Consider a piece-wise linear continuous payoff function $f(x)$ satisfying*

- $f(x) = 0, \forall x < 0$;
- *On x axis, $K_0 = 0$, $0 < K_1 < K_2 < \ldots$ are successive kink points from right to left.*
- *$f(x)$ intercept at y axis I.*
- *$\lambda_1, \lambda_2, \ldots$ are slopes of successive linear pieces.*

Let $C(K)$ denote the current market price of a call option with strike K and same maturity. Then, the current value is given by

$$\begin{aligned} V(t) = Ie^{-r(T-t)} - \lambda_1(-P(K_0) + P(K_1) - \\ (K_1 - K_0)e^{-r(T-t)}) - \lambda_2(P(K_1) - P(K_2) - (K_2 - K_1)e^{-r(T-t)}) + \ldots. \end{aligned}$$

Note that $S_t = C(K_0 = 0)$.

Proof. The final payoff can be constructed using the following functions

$$f(x) = I + \lambda_1(-(K_0 - x)^+ + (K_1 - x)^+) + \lambda_2(-(x - K_1)^+ + (x - K_2)^+) + \ldots,$$

since $g(x) \triangleq -(K_0 - x)^+ + (K_1 - x)^+$ is a linear segment given by

$$g(x) = \begin{cases} K_1 - K_0, x \leq K_0 \\ -(x - K_1), K_0 < x < K_1 \\ 0, x \geq K_2 \end{cases},$$

and $f(x) \triangleq -(K_0 - x)^+ + (K_1 - x)^+ - (K_1 - K_0)$ is a linear segment given by

$$f(x) = \begin{cases} 0, x \leq K_0 \\ -(x - K_0), K_0 < x < K_1 \\ -(K_1 - K_0), x \geq K_1 \end{cases}.$$

□

Remark 2.1.3 (practical implications)**.**

- In practice, we can use this method to approximate arbitrary payoff function using available call options in the market.
- We can further use put-call parity to replace illiquid out-of-money calls by more liquid out-of-money puts.

Proposition 2.1.3 (general discrete replication via Hilbert space approximation). *Suppose we have a continuous payoff function $f(x)$ with the support $[a,b]$. Suppose we have liquid products with continuous payoff function given by $g_1(x), g_2(x), ..., g_n(x)$(the support is also $[a,b]$). Define inner product between two payoff functions as*

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx.$$

It follows that

- *we can approximatively replicate the payoff function via*

$$\hat{f}(x) = \sum_{i=1}^{n} c_i g_i(x),$$

where the vector c is determined by

$$Gc = p,$$

where $G_{ij} = \left\langle g_i, g_j \right\rangle, p_i = \langle g_i, f \rangle$.

- *The price due to the approximate replication is given by*

$$NPV[\hat{f}] = \sum_{i=1}^{n} c_i NPV[g_i].$$

Proof. (1) See Hilbert space approximation theorem(**??**). (2) See linear pricing theorem(Theorem 1.2.2). □

Note 2.1.6 (path-dependent derivative pricing using implied distributions)**.** Suppose we have the market prices of European calls for all strikes K and all maturities T, then we are able to obtain the implied distribution of $S(t)$ for all future t, which will enables us to price path-dependent derivatives(e.g., Asian options).

2.1.4 Implied forward distribution

2.1.4.1 *Forward starting option basics*

Definition 2.1.3 (forward starting call/put option). *[4, p. 602]*

- *Consider two dates $0 < T_0 < T$. A **forward starting call/put option** allows the holder to receive, at time T_0, a call/put option expiring at T, with strike set equal to $S(T_0)K$, for some*

$K > 0$. That is, the option's life starts at T_0, but the holder pays at time 0 the premium of the option.

- *We can view a call option has payoff at time T given by*

$$(S(T) - KS(T_0))^+.$$

Lemma 2.1.12 (forward starting option with general distributions). *Consider two dates $0 < T_1 < T_2$ and a forward starting call option payoff at time T given by*

$$(S(T_2) - KS(T_1))^+.$$

Under risk-neutral measure Q, the time 0 value of the forward starting call option is given by

$$\begin{aligned} V(0) &= e^{-rT_2} E_Q[(S(T_2) - KS(T_1))^+] \\ &= e^{-rT_1} E_Q[g(S_1)] \end{aligned}$$

where

$$\begin{aligned} g(s_1) &= e^{-r(T_2-T_1)} E_Q[(S_2 - KS_1)^+ | S_1 = s_1] \\ &= e^{-r(T_2-T_1)} \int_{\mathbb{R}} (s_2 - Ks_1)^+ f_{S_2|S_1}(s_2|s_1) ds_2 \end{aligned}$$

and $f_{S_2|S_1}$ is the conditional pdf of S_2 given S_1.

Proof. Based on the definition, we have

$$\begin{aligned} V(0) &= e^{-rT_2} E_Q[(S(T_2) - KS(T_1))^+] \\ &= e^{-rT_2} \int_{\mathbb{R}^2} (s_2 - Ks_1)^+ f_{S_1,S_2}(s_1, s_2) ds_1 ds_2 \\ &= e^{-rT_1} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-r(T_2-T_1)} (s_2 - Ks_1)^+ f_{S_2|S_1}(s_2|s_1) ds_2 f_{S_1}(s_1) ds_1 \\ &= e^{-rT} \int_{\mathbb{R}} g(s_1) f_{S_1}(s_1) ds_1 \\ &= e^{-rT} E_Q[g(S_1)] \end{aligned}$$

□

Lemma 2.1.13 (forward starting option pricing in Geometric Browian motion model). *Consider two dates $0 < T_0 < T$. A forward starting call option allows the holder to receive, at time*

T_0, a call option expiring at T, with strike set equal to $S(T_0)K$, for some $K > 0$. Assuming the underlying asset S_t has dynamics given by

$$dS_t/S_t = rdt + \sigma dW_t.$$

The value at time 0 is given by

$$V(0) = S(0) \cdot c(1, T - T_0, K),$$

where $c(1, T - T_0, K)$ is the call option price with spot 1, strike K, and tenor $T - T_0$ in the Black model given by

$$c(1, T - T_0, K) = N(d_1) - K\exp(-r(T - T_0))N(d_2),$$

where

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T - T_0}}[\ln(\frac{1}{K}) + (r + \sigma^2/2)(T - T_0)] \\ d_2 &= \frac{1}{\sigma\sqrt{T - T_0}}[\ln(\frac{1}{K}) + (r - \sigma^2/2)(T - T_0)]. \end{aligned}$$

Proof. Note that the payoff at terminal time T is given by

$$FS(T) = (S(T) - KS(T_0))^+$$

then its value at T_0 is given by(Lemma 5.2.1)

$$c(S(T_0), T - T_0, KS(T_0)) = N(d_1)S(T_0) - N(d_2)KS(T_0)e^{-r(T-t)},$$

where

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T - T_0}}[\ln(\frac{S(T_0)}{KS(T_0)}) + (r + \sigma^2/2)(T - T_0)] \\ d_2 &= \frac{1}{\sigma\sqrt{T - T_0}}[\ln(\frac{S(T_0)}{KS(T_0)}) + (r - \sigma^2/2)(T - T_0)]. \end{aligned}$$

Note that we can also write

$$c(S(T_0), T - T_0, KS(T_0)) = S(T_0)c(1, T - T_0, K),$$

where $c(1, T - T_0, K)$ is a deterministic quantity.

Then,

$$\begin{aligned} V(0) &= E_Q[\exp(-rT_0)S(T_0) \cdot c(1, T - T_0, K)] \\ &= E_Q[\exp(-rT_0)S(T_0)]c(1, T - T_0, K) \\ &= S(0)c(1, T - T_0, K) \end{aligned}$$

where we use the fact that $\exp(-rt)S(t)$ is a martingale under risk-neutral measure. □

Remark 2.1.4 (strike is stochastic and unknown). Note that usually a forward starting call option has strike set equal to $S(T_0)K$, for some $K > 0$. And $S(T_0)$ is a stochastic unknown quantity.

Remark 2.1.5 (the situation of general strike). For the case of general strike K instead of $KS(T_0)$, we have

$$V(0) = E_Q[\exp(-rT_0)c(S(T_0), T - T_0, KS(T_0))]$$

where

$$\begin{aligned} c(S(T_0), T - T_0, K) &= N(d_1)S(T_0) - N(d_2)Ke^{-r(T-t)} \\ d_1 &= \frac{1}{\sigma\sqrt{T - T_0}}[\ln(\frac{S(T_0)}{K}) + (r + \sigma^2/2)(T - T_0)] \\ d_2 &= \frac{1}{\sigma\sqrt{T - T_0}}[\ln(\frac{S(T_0)}{K}) + (r - \sigma^2/2)(T - T_0)]. \end{aligned}$$

Note that we need the distribution of $S(T_0)$ in order to evaluate the expectation; and the integral will require numerical integration.

Remark 2.1.6 (how to extract complete implied transitioning probability from the market).

- If in the market there are actively traded double Dirac delta option with payoff $\delta(S_1 - K_1) \times \delta(S_2 - K)$, where $S_1 \triangleq S(T_1), S_2 \triangleq S(T_2)$. Then we have the joint pdf given by
$$f_{S_1,S_2}(K_1, K_2) = V_0(K_1, K_2)e^{rT_2}.$$
- Such joint pdf can be integrated to get marginal pdf and conditional pdf (i.e., transition pdf $f_{S_2|S_1}$).

Proposition 2.1.4 (pricing arbitrary forward starting derivatives). *Consider two dates $0 < T_1 < T_2$. Consider a derivative on the asset process S_t with payoff at T_2 given by $S_1 \cdot h(\frac{S_2}{S_1})$, where $S_1 \triangleq S(T_1), S_2 \triangleq S(T_2)$, and h is function of the ration S_2/S_1.*

Then the price of this derivative is given by

$$V_0 = E_Q[e^{-rT_2}S_1 \cdot h(\frac{S_2}{S_1})] = \int_{\mathbb{R}} V_F(k)h(k)dk,$$

where $V_F(k)$ is the second derivative of a forward starting call option price at strike k, that is

$$V_F(k) = \frac{\partial}{\partial k^2}e^{-rT_2}E_Q[(S_2 - S_1 \cdot k)^+].$$

2.1.4.2 *Implied forward volatility*

Definition 2.1.4 (Implied forward volatility defined via European and digital option).

- *Consider two dates* $0 < T_0 < T$ *and a forward starting call option allows the holder to receive, at time* T_0*, a call option expiring at* T*, whose value is given by*

$$(\frac{S_T}{S_{T_0}} - K)^+$$

Let the market price of the forward starting option be C^{mkt}*. Then, the implied forward volatility, denoted by* $\sigma(T_0, T, K)$*, is the value that the following equation hold; that is*

$$N(d_1) - \exp(-r(T-T_0))KN(d_2) = C^{mkt},$$

where

$$d_1 = \frac{1}{\sigma\sqrt{T-T_0}}[\ln(\frac{1}{K}) + (r + \sigma^2/2)(T-T_0)]$$
$$d_2 = \frac{1}{\sigma\sqrt{T-T_0}}[\ln(\frac{1}{K}) + (r - \sigma^2/2)(T-T_0)].$$

- *Consider two dates* $0 < T_0 < T$ *and a forward starting digital call allows the holder to receive, at time* T_0*, a digital call option expiring at* T*, whose value is given by*

$$\mathbf{1}_{(\frac{S_T}{S_{T_0}} - K)}$$

Let the market price of the forward starting option be C^{mkt}*. Then, the implied forward volatility, denoted by* $\sigma(T_0, T, K)$*, is the value that the following equation hold; that is*

$$\exp(-r(T-T_0))KN(d_2) = C^{mkt},$$

where

$$d_2 = \frac{1}{\sigma\sqrt{T-T_0}}[\ln(\frac{1}{K}) + (r - \sigma^2/2)(T-T_0)].$$

Note 2.1.7 (interpretation of implied forward volatility in the Black-Scholes setting)**.**

- The implied forward volatility is parameterized by three parameters, denoted by $\sigma(T_0, T, K)$.

- Under the geometric Brownian motion, the implied forward volatility determines the implied distribution of S_T/S_{T_0} conditioning on the information at time 0; that is

$$\frac{S_T}{S_{T_0}} = \exp(r(T-T_0) - \frac{1}{2}\sigma(T-T_0)^2 + \sigma\sqrt{T-T_0}Z), Z \sim N(0,1),$$

where the σ depends on the strike.
- Conditioning on the information at time T_0(that is, S_{T_0} is known), then

Lemma 2.1.14 (application examples for forward implied volatility). *Consider two dates $0 < T_0 < T$. Suppose we know the implied volatility σ_0 and forward implied volatility $\sigma_{0,1}$. Then we can price the following instruments with market consistence.*

- $$(S_T - S_{T_0} - K)^+$$
$$E[E[(S_T - S_{T_0} - K)^+|cF_{T_0}]|\mathcal{F}_0]$$

Proof. (1) The current value of the instrument can be written by

$$E[E[(S_T - S_{T_0} - K)^+|\mathcal{F}_{T_0}]|\mathcal{F}_0].$$

With forward volatility, we can evaluate

$$\begin{aligned} &E[(S_T - S_{T_0} - K)^+|cF_{T_0}] \\ &= S_{T_0}E[(\frac{S_T}{S_{T_0}} - 1 - K/S_{T_0})|\mathcal{F}_{T_0}] \\ &= f(S_{T_0}) \end{aligned}$$

since for

$$E[(S_T - S_{T_0} - K)^+|cF_{T_0}]$$

□

2.2 Implied volatility surface

2.2.1 Fundamentals and facts

Definition 2.2.1 (implied volatility). *[4, p. 341] Let $C(K,T;S_t)$ be the current time t market value of a call option with strike K, maturity T and underlying S_t. The implied volatility for European options on the stock S_t is the $\sigma^*(K,T;S_t)$ that the following equality hold:*

$$C_{BS}(S_t, K, r, T, \sigma^*) = C(K, T, S_t)$$

.

Remark 2.2.1 (interpretation).

- The implied volatility reflects the market's view on the volatility of the stock.
- The implied volatility obtained from actively traded options are usually used to price other options.

Definition 2.2.2 (volatility smile, term structure, and surface). *[4, p. 431]*

- *The implied volatility of an option with a certain life as a function of its strike price is known as a volatility smile.*
- *The implied volatility of an option with a certain life as a function of its maturity is known as a volatility term structure*
- *The implied volatility of an option with a certain life as a function of its strike price and its maturity is known as a volatility surface.*

Remark 2.2.2 (interpretation).

- The implied volatility as a function of strike price for equity option market is usually down-sloping and it is also known as **volatility skew**.
- The implied volatility vs the strike price is usually valley-shaped for currency markets.

Lemma 2.2.1 (No arbitrage bounds on European call/put price profile and smiles). *[1, p. 153][2, p. 22] Let $C(K,T;S_t)$ and $P(K,T;S_t)$ denote the current market prices of European calls/puts with strike K and maturity T.*

The bounds on price profile is given by

- $C \geq Se^{-d\tau} - Ke^{-r\tau}$, *where d is the dividend rate, $\tau = T - t$, and r the short rate.*
- $$\frac{\partial C}{\partial K} \leq 0, \frac{\partial P}{\partial K} \geq 0;$$
- $$\frac{\partial^2 C}{\partial K^2} \geq 0, \frac{\partial^2 P}{\partial K^2} \geq 0;$$

- $$\int_0^\infty \frac{\partial^2 C}{\partial K^2} dK = \int_0^\infty \frac{\partial^2 P}{\partial K^2} dK = e^{-rT}.$$

Proof. (1) use the call-put parity(Lemma 1.7.3) and the fact that a put price $P \geq 0$. (2)(3) use the relations between derivatives of call/puts and delta/step function. See Lemma 2.1.1. (4) Use the fact that $\frac{\partial^2 C}{\partial K^2}$ will give a discounted risk-neutral measure that integrate to 1. See Lemma 2.1.4. □

Remark 2.2.3 (bounds are model independent).

- Note that bounds on prices only relies on the assumption that there exists no arbitrage opportunities in the market rather than any model.
- Any market model that violates above should not be used.

2.2.2 Stylized market facts about volatility

Note 2.2.1 (equity market implied volatility feature). [5, pp. 233, 239][1, p. 5][6, p. 486] Consider European call/put options having same maturity dates but different strikes. The volatility skew (Figure 2.2.1) is often observed in the market due to the following reasons:

- Stock investors, as a whole, are worrying more about the falling stock price than the rising stock price; therefore, the protective put strategy(**??**) with out-of-money put is in high demand and push the out-of-money put price up.
- Out-of-money puts have higher implied volatility(that is, higher price than Black model predicts) than in-the-money puts because out-of-money puts resemble an insurance against market crash and risk-averse buyers are willing to pay extra money for insurance.
- Out-of-money puts have higher implied volatility(that is, higher price) than out-of-money calls because buyers/market worry more about the falling prices than the rising prices.

Consider European call/put options having same strike but different maturity dates.

- At volatile period, options of smaller maturities tend to have higher implied volatility.
- At tranquil period, options of larger maturities tend to have higher implied volatility.

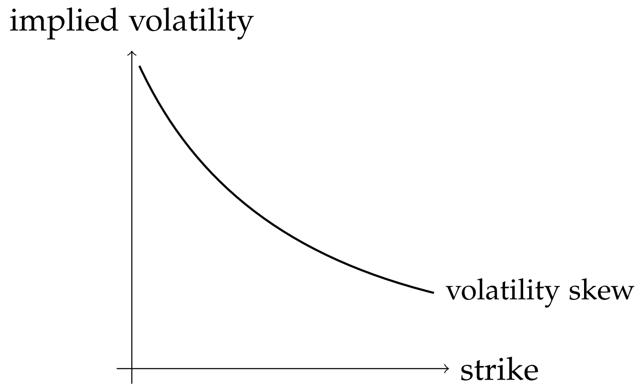


Figure 2.2.1: volatility skew for stock index

Note 2.2.2 (commodity market implied volatility feature)**.** [6, p. 487][5, p. 236] Consider the commodity market. The volatility skew, as showed in Figure 2.2.2, is often observed in the market.

- In commodity markets where end users try to protect themselves against rising prices by either buying **protective calls**(**??**) at higher exercise prices(out-of-money calls) or selling **covered puts**(**??**) at lower exercise prices(out-of-money put).
- As a result, in the commodity market, lower exercise prices have lower implied volatilities, and higher exercise prices have higher implied volatilities.

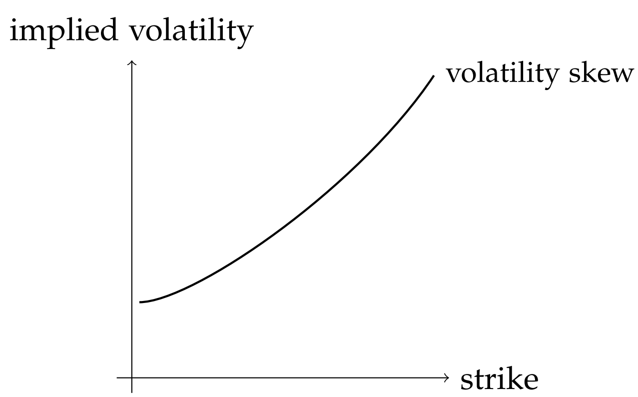


Figure 2.2.2: volatility skew for commodity market

Note 2.2.3 (FX market implied volatility feature)**.** [6, p. 488][5, p. 236] Consider the FX market. The volatility skew, as showed in Figure 2.2.3, is often observed in the market.

- For FX market market participants, when one party is worrying about the rising exchange rate(buy protective call or covered put), there is also an counterparty worrying about the falling exchange rate(buy protective put or covered call).
- As a result, in the FX market, lower strike and higher strike prices tend to have higher implied volatilities, and intermediate strike have lower implied volatility.

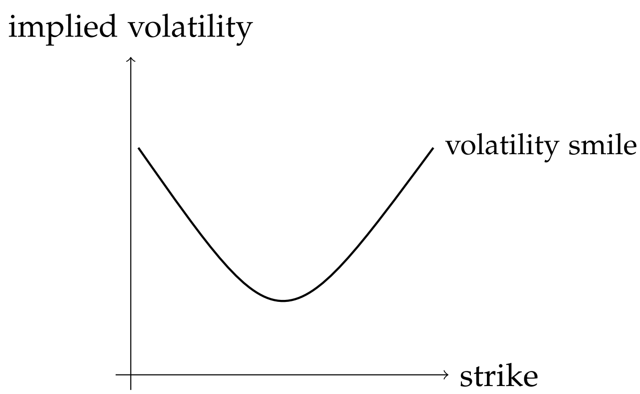


Figure 2.2.3: volatility smile for FX market

Remark 2.2.4 (discussion on smiles in different markets). For smiles in equity, interest rate, foreign exchange, see [1, p. 143].

Note 2.2.4 (geometric Brownian motion with varying volatility can generate arbitrage distributions).

2.2.3 Implied distribution from implied volatility

Proposition 2.2.1 (deriving implied distribution from implied volatility). [7] *Let $\sigma^*(K,T;S_t)$ be the implied volatility such that the market prices of calls $C(K,T;S_t)$ can be parameterized by*

$$C(K,T;S_t) = N(d_1)S_t - N(d_2)Ke^{-r(T-t)}$$

where

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r + [\sigma^*]^2/2)(T-t)]$$
$$d_2 = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r - [\sigma^*]^2/2)(T-t)].$$

Then, the implied distribution parametrized in log moneyness $\kappa = \ln(K/F(t,T))$ is given by

$$f_\kappa(\kappa = k) = e^{rT}\frac{g(k)}{\sqrt{2\pi w(k,T)}}\exp(-\frac{d_-(k)^2}{2}),$$

where

$$\begin{aligned} w(k,T) &= \sigma^2(k,T)(T-t) \\ d_\pm(k) &= -\frac{k}{\sqrt{w(k,T)}} \pm \frac{\sqrt{w(k,T)}}{2} \\ g(k) &= 1 - \frac{\kappa}{w}\frac{\partial w}{\partial \kappa} + \frac{1}{4}(\frac{\partial w}{\partial \kappa})^2(\frac{\kappa^2}{w^2} - \frac{1}{w} - \frac{1}{4}) + \frac{1}{2}\frac{\partial^2 w}{\partial \kappa^2} \end{aligned}$$

The implied distribution parameterized in strike K is given by

$$f_{S_T}(S_T = K) = e^{rT}\frac{g_K(K)}{K\sqrt{2\pi w(k,T)}}\exp(-\frac{d_2(K)^2}{2}).$$

where $g_K(K)$ is the composite function $g_K(K) = g(\ln K/F(t,T))$.

Proof. See Lemma 2.1.4. Note that σ^* also has dependence on K. □

2.3 Static distribution model

Various distribution for ST can be selected as long as this condition is satisfied: for example, normal, lognormal. One distribution is preferred over other distribution based on the analytical pricing possibility and calibration to market.

Parameters in the assumed distribution were tuned so that the model prices fit the market prices for contracts whose market prices could be observed.

2.3.1 Principles

Remark 2.3.1 (consistence with the no-arbitrage pricing framework)**.** We have showed in (), to have no-arbitrage pricing, the spot price should satisfy

$$\frac{S_t}{M(t)} = E_Q[\frac{S(T)}{M(T)}|\mathcal{F}_t]$$

under risk-neutral measure.

And the forward price should satisfy

$$F(t, T_M) = E_Q[F(T, T_M)|\mathcal{F}_t]$$

under risk-neutral measure.

Remark 2.3.2 (achieve market consistence via calibration)**.** The goal of calibration is to find a set of parameters such that the static distribution is close to implied distribution.

2.3.2 Black lognormal model

Definition 2.3.1. *Let current time be t. In a Black lognormal model, we assume the spot price S_t of an asset has a lognormal distribution under risk-neutral measure at time T, given by*

$$S_T = S_t \exp((r - q - \frac{1}{2}\sigma^2)(T - t) + \sigma\sqrt{T - t}Z), Z \in N(0, 1).$$

Similarly, we assume the forward price $F(t, T_M)$ of an asset has a lognormal distribution under risk-neutral measure at time T, given by

$$F(T, T_M) = F(t, T_M) \exp((-\frac{1}{2}\sigma^2)(T - t) + \sigma\sqrt{T - t}Z), Z \in N(0, 1).$$

Proposition 2.3.1 (Black model European call/put pricing for assets without dividends). *[8, p. 219] Let current time t. Denote the spot by S_t and forward by $F(t, T)$. In the Black-Scholes model, an European call option with strike price K and expiry T is given as:*

$$C(S_t, t) = N(d_1)S_t - N(d_2)Ke^{-r(T-t)}$$

where

$$\begin{aligned}
d_1 &= \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r + \sigma^2/2)(T - t)] = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{F(t, T)}{K}) + \sigma^2/2(T - t)] \\
d_2 &= \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r - \sigma^2/2)(T - t)] = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{F(t, T)}{K}) - \sigma^2/2(T - t)] \\
d_2 &= d_1 - \sigma\sqrt{T - t} \\
N(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \\
F(t, T) &= S_t \exp(r(T - t))
\end{aligned}$$

We further have

- *The price of a zero strike is S_t.*
- *The price of the put can be derived based on put-call parity $P_t + S_t = C_t + Ke^{-r(T-t)}$ (**??**) as*

$$P(S_t, t) = N(-d_2)Ke^{-r(T-t)} - N(-d_1)S_t.$$

Proof. We use martingale method.

$$C(S_T, T) = E_Q[e^{-r(T-t)} \max(S_T - K, 0) | \mathcal{F}_t].$$

where

$$S(T) = S(t)\exp((r - \frac{1}{2}\sigma(s)^2)(T-t) + \sigma(W(T) - W(t))).$$

Let $Y = (W(T) - W(t))/\sqrt{T-t}$, then $Y \sim N(0,1)$, and we can write

$$S(T) = S(t)\exp((r - \frac{1}{2}\sigma(s)^2)(T-t) + \sigma\tau Y)).$$

Let

$$d_2 = \frac{\ln\frac{S(t)}{K} + (r - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}$$

such that

$$S(t)\exp((r - \frac{1}{2}\sigma^2)\tau + \sigma\sqrt{\tau}(-d_2)) = K.$$

Then the for the standard normal Y, We have

$$Pr(S(t)\exp((r - \frac{1}{2}\sigma(s)^2)(T-t) + \sigma\tau Y)) > K) = Pr(Y > -d_2) = Pr(Y < d_2).$$

$$\begin{aligned}
C(S_T, T) &= \frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} e^{-r\tau}(S(t)\exp(+\sigma\tau y + (r - \frac{1}{2}\sigma^2)\tau) - K)e^{-\frac{1}{2}y^2} dy \\
&= [\frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} S(t)\exp(-\sigma\tau y - \frac{1}{2}\sigma^2\tau)e^{-\frac{1}{2}y^2} dy] - [\frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} e^{-r\tau}e^{-\frac{1}{2}y^2} dy] \\
&= [\frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} S(t)\exp(-\frac{1}{2}(y - \sigma\sqrt{\tau})^2) dy] - e^{-r\tau}KN(d_2) \\
&= [\frac{1}{\sqrt{2\pi}} \int_{-d_2-\sigma\sqrt{\tau}}^{\infty} S(t)\exp(-\frac{1}{2}z^2) dz] - e^{-r\tau}KN(d_2) \\
&= S(t)N(d_1) - e^{-r\tau}KN(d_2)
\end{aligned}$$

where we used the fact that $-d_1 = -d_2 - \sigma\sqrt{\tau}$. □

Proposition 2.3.2 (Black model European call and put for assets with dividends). *[8, p. 236][4, p. 373] Let current time t. Denote the spot by S_t and forward by $F(t,T)$. In the Black-Scholes model, an European call option with strike price K for an asset paying continuous dividends a is given as:*

$$C(S_t,t) = N(d_1)S_te^{-a(T-t)} - N(d_2)Ke^{-r(T-t)}$$

where

$$\begin{aligned}
d_1 &= \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r-a+\sigma^2/2)(T-t)] = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{F(t,T)}{K}) + \sigma^2/2(T-t)] \\
d_2 &= \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r-a-\sigma^2/2)(T-t)] = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{F(t,T)}{K}) - \sigma^2/2(T-t)] \\
d_2 &= d_1 - \sigma\sqrt{T-t} \\
N(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \\
F(t,T) &= S_t \exp((r-a)(T-t))
\end{aligned}$$

The price of the put can be derived based on put-call parity

$$P_t + S_te^{-a(T-t)} = C_t + Ke^{-r(T-t)}$$

*(**??**) as*

$$P(S_t,t) = N(-d_2)Ke^{-r(T-t)} - N(-d)S_te^{-a(T-t)}.$$

Proof. We use martingale method.

$$C(S_T,T) = E_Q[e^{-r(T-t)}\max(S_T-K,0)|\mathcal{F}_t].$$

where

$$S(T) = S(t)\exp((r-a-\frac{1}{2}\sigma(s)^2)(T-t) + \sigma(W(T)-W(t))).$$

Let $Y = (W(T)-W(t))/\sqrt{T-t}$, then $Y \sim N(0,1)$, and we can write

$$S(T) = S(t)\exp((r-a-\frac{1}{2}\sigma(s)^2)(T-t) + \sigma\tau Y)).$$

Let

$$d_2 = \frac{\ln\frac{S(t)}{K} + (r-a-\frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}$$

such that

$$S(t)\exp((r-a-\frac{1}{2}\sigma^2)\tau + \sigma\sqrt{\tau}(-d_2)) = K.$$

Then the for the standard normal Y, We have

$$Pr(S(t)\exp((r-a-\frac{1}{2}\sigma(s)^2)(T-t)+\sigma\tau Y))>K)=Pr(Y>-d_2)=Pr(Y<d_2).$$

$$\begin{aligned}
C(S_T,T) &= \frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty} e^{-r\tau}(S(t)\exp(+\sigma\tau y+(r-a-\frac{1}{2}\sigma^2)\tau)-K)e^{-\frac{1}{2}y^2}dy \\
&= [\frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty} S(t)\exp(-\tau a)\exp(-\sigma\tau y-\frac{1}{2}\sigma^2\tau)e^{-\frac{1}{2}y^2}dy]-[\frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty} e^{-r\tau}e^{-\frac{1}{2}y^2}dy] \\
&= [\frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty} S(t)\exp(-\tau a)\exp(-\frac{1}{2}(y-\sigma\sqrt{\tau})^2)dy]-e^{-r\tau}KN(d_2) \\
&= [\frac{1}{\sqrt{2\pi}}\int_{-d_2-\sigma\sqrt{\tau}}^{\infty} S(t)\exp(-\tau a)\exp(-\frac{1}{2}z^2)dz]-e^{-r\tau}KN(d_2) \\
&= S(t)\exp(-\tau a)N(d_1)-e^{-r\tau}KN(d_2)
\end{aligned}$$

where we used the fact that $-d_1=-d_2-\sigma\sqrt{\tau}$. □

2.3.3 Extended lognormal model

Lemma 2.3.1 (black formula with extended lognormal distribution family). *[9]*

- $$c=\exp(-rT)[M_1(T)N(d_1)-KN(d_2)],$$

  where

  $$d_{1,2}=\frac{\log M_1(T)-\log K}{\sigma}\pm\frac{1}{2}\sigma$$

  and

  $$\sigma=\sqrt{\log(\frac{M_2(T)}{M_1(T)^2})}.$$

- $$c=\exp(-rT)[(M_1(T)-\tau)N(d_1)-(K-\tau)N(d_2)],$$

  where

  $$d_{1,2}=\frac{\log(M_1(T)-\tau)-\log(K-\tau)}{\sigma}\pm\frac{1}{2}\sigma$$

  and

  $$\sigma=\sqrt{\log(\frac{M_2(T)}{M_1(T)^2})}.$$

- $$c=\exp(-rT)[M_1(T)N(d_1)-KN(d_2)],$$

where

$$d_{1,2} = \frac{\log -M_1(T) - \log -K}{\sigma} \pm \frac{1}{2}\sigma$$

and

$$\sigma = \sqrt{\log(\frac{M_2(T)}{M_1(T)^2})}.$$

- $$c = \exp(-rT)[M_1(T)N(d_1) - KN(d_2)],$$

  where

$$d_{1,2} = \frac{\log M_1(T) - \log K}{\sigma} \pm \frac{1}{2}\sigma$$

and

$$\sigma = \sqrt{\log(\frac{M_2(T)}{M_1(T)^2})}.$$

2.3.4 Bachelier normal model

Definition 2.3.2 (Bachelier normal model). *Let current time be t. In a Bachelier normal model, we assume the spot price S_t of an asset has a normal distribution under risk-neutral measure at time T, given by*

$$S_T = S_t + (r-q)(T-t) + \sigma\sqrt{T-t}Z, Z \in N(0,1).$$

Similarly, we assume the forward price $F(t, T_M)$ of an asset has a normal distribution under risk-neutral measure at time T, given by

$$F(T, T_M) = F(t, T_M) + \sigma\sqrt{T-t}Z, Z \in N(0,1).$$

Lemma 2.3.2 (European call/put pricing under normal model). *[10] Under the normal forward price model, the current values of European call and put are given by*

- $$\begin{aligned} C(t) &= \exp(-r(T-t))E_Q[(F(T,T)-K)^+|\mathcal{F}_t] \\ &= \exp(-r(T-t))((F(t,T)-K)N(d_1) + \frac{\sigma\sqrt{T-t}}{\sqrt{2\pi}}e^{-d_1^2/2}); \end{aligned}$$

•

$$P(t) = \exp(-r(T-t))E_Q[(K - F(T,T))^+|\mathcal{F}_t]$$
$$= \exp(-r(T-t))((K - F(t,T))N(-d_1) + \frac{\sigma\sqrt{T-t}}{\sqrt{2\pi}}e^{-d_1^2/2});$$

• ***(put call parity)***

$$C(t) - P(t) = \exp(-r(T-t))(F(t,T) - K)$$

where $d_1 = \frac{F-K}{\sigma\sqrt{T-t}}$

Proof. From Feyman-Kac theorem(Theorem 1.8.1), the forward price under risk-neutral measure has dynamics

$$dF = \sigma dW_t.$$

That is $F(T) - F(t) \sim N(0, \sigma^2(T-t))$

$$C(t) = e^{-r(T-t)}E_Q[(F_T - K)^+|\mathcal{F}_t]$$
$$= \frac{e^{-r(T-t)}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(F_t + \sigma\sqrt{T-t}x - K)^+e^{-x^2/2}dx$$
$$= \frac{e^{-r(T-t)}}{\sqrt{2\pi}}\int_{\frac{K-F}{\sigma\sqrt{T-t}}}^{\infty}(F_t + \sigma\sqrt{T-t}x - K)e^{-x^2/2}dx$$

The rest is straight forward. □

2.3.5 Multivariate lognormal model

Definition 2.3.3 (multivariate lognormal model). *Let current time be t. In a Black lognormal model, we assume the spot prices of assets* $S_1, S_2, ..., S_N$ *have a multivariate lognormal distribution under risk-neutral measure at time* T*, given by*

$$\ln S_i(T) = \ln S_t + (r - q - \frac{1}{2}\sigma_i^2)(T-t) + \sigma_i\sqrt{T-t}Z, Z \in N(0,1);$$

$$corr(\ln S_i, \ln S_j) = \rho_{ij}.$$

Similarly, we assume the forward prices of assets $F_1(t, T_M), F_2(t, T_M), ..., F_N(t, T_M)$ *have a multivariate lognormal distribution under risk-neutral measure at time* T*, given by*

$$\ln F_i(T, T_M) = \ln F_i(T, T_M) + (-\frac{1}{2}\sigma_i^2)(T-t) + \sigma_i\sqrt{T-t}Z, Z \in N(0,1).$$

$$corr(\ln F_i(T, T_M), \ln F_i(T, T_M)) = \rho_{ij}.$$

2.3.5.1 *Different assets*

2.3.5.2 *One asset at different future times*

2.3.5.3 *Moments*

Lemma 2.3.3 (moments of basket underlying dynamics under risk-neutral measure). *Assume under risk-neutral measure Q that the underlying dynamics are given by*

$$dS_t^{(i)}/S_t^{(i)} = (r-q)dt + \sigma_i dW_i(t), i = 1, 2, ..., n,$$

and $dW_i dW_j = \rho_{ij} dt$. *Denote* $B(T) = \sum_{i=1}^n w_i S_T^{(i)}$. *It following that*

- $$B(T) = \sum_{i=1}^n F_i w_i \exp(\sigma_i W_i(T) - \frac{1}{2}\sigma_i^2 T),$$

  where $F_i = S^{(i)}(0)\exp((r-q)T)$.

- $$E[B(T)] = \sum_{i=1}^n F_i w_i.$$

- $$E[B(T)^2] = \sum_{i=1}^n \sum_{j=1}^n F_i F_j w_i w_j \exp(\rho_{ij}\sigma_i\sigma_j T).$$

- $$E[B(T)^3] = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n F_i F_j F_k w_i w_j w_k (\exp((\rho_{ij}\sigma_i\sigma_j + \rho_{ik}\sigma_i\sigma_k + \rho_{jk}\sigma_j\sigma_k)T)).$$

Proof. (1) Note that

$$S_T^{(i)} = S_0^{(i)} \exp((r - q - \frac{1}{2}\sigma_i^2 T) + \sigma_i W_i(T)).$$

(2) Note that

$$E[\exp(\sigma_i W_i(T) - \frac{1}{2}\sigma_i^2 T)] = 1.$$

(3)

$$
\begin{aligned}
&E[\exp(\sigma_i W_i(T) - \frac{1}{2}\sigma_i^2 T)\exp(\sigma_j W_j(T) - \frac{1}{2}\sigma_j^2 T)] \\
&= \exp(-\frac{1}{2}\sigma_i^2 T)\exp(-\frac{1}{2}\sigma_j^2 T) M(\sigma_i, \sigma_j) \\
&= \exp(-\frac{1}{2}\sigma_i^2 T)\exp(-\frac{1}{2}\sigma_j^2 T)\exp(\frac{1}{2}(\sigma_i, \sigma_j)\begin{bmatrix} T & \rho T \\ \rho T & T \end{bmatrix}(\sigma_i, \sigma_j)^T) \\
&= \exp(-\frac{1}{2}\sigma_i^2 T)\exp(-\frac{1}{2}\sigma_j^2 T)\exp(\frac{1}{2}(\sigma_i^2 T + \sigma_j^2 T + 2\rho_{ij}\sigma_i\sigma_j T) \\
&= \exp(-\frac{1}{2}\sigma_i^2 T)\exp(-\frac{1}{2}\sigma_j^2 T)\exp(\frac{1}{2}(\sigma_i^2 T + \sigma_j^2 T + 2\rho_{ij}\sigma_i\sigma_j T) \\
&= \exp(\rho_{ij}\sigma_i\sigma_j)
\end{aligned}
$$

where $M(t_1, t_2)$ is the mgf for the random vector $(W_i(T), W_j(T))$. (4)

$$
\begin{aligned}
&E[\exp(\sigma_i W_i(T) - \frac{1}{2}\sigma_i^2 T)\exp(\sigma_j W_j(T) - \frac{1}{2}\sigma_j^2 T)\exp(\sigma_k W_k(T) - \frac{1}{2}\sigma_k^2 T)] \\
&= \exp(-\frac{1}{2}\sigma_i^2 T)\exp(-\frac{1}{2}\sigma_j^2 T)\exp(-\frac{1}{2}\sigma_k^2 T) M(\sigma_i, \sigma_j, \sigma_k) \\
&= \exp(-\frac{1}{2}\sigma_i^2 T)\exp(-\frac{1}{2}\sigma_j^2 T)\exp(-\frac{1}{2}\sigma_k^2 T)\exp(\frac{1}{2}(\sigma_i, \sigma_j, \sigma_k)\begin{bmatrix} T & \rho_{ij} T & \rho_{ik} T \\ \rho_{ij} T & T & \rho_{jk} T \\ \rho_{ik} T & \rho_{jk} T & T \end{bmatrix}(\sigma_i, \sigma_j, \sigma_k)^T) \\
&= \exp(-\frac{1}{2}\sigma_i^2 T)\exp(-\frac{1}{2}\sigma_j^2 T)\exp(-\frac{1}{2}\sigma_k^2 T)\exp(\frac{1}{2}(\sigma_i^2 T + \sigma_j^2 T + \sigma_k^2 T + 2\rho_{ij}\sigma_i\sigma_j T + 2\rho_{ik}\sigma_i\sigma_k T + 2\rho_{jk}\sigma_k\sigma_j T) \\
&= \exp(\rho_{ij}\sigma_i\sigma_j)
\end{aligned}
$$

where $M(t_1, t_2, t_3)$ is the mgf for the random vector $(W_i(T), W_j(T), W_k(T))$. □

2.3.5.4 *Moment matching methods*

Lemma 2.3.4 (two-parameter lognormal moment matching). *[11][12, p. 232] Let $T_1, T_2, ..., T_n$ be a set of dates. Consider a set of random variables given by*

$$X_j = S_0 \exp((r - \frac{1}{2}\sigma^2)T_j + \sigma W(T_j)).$$

and the averaging random variable

$$X = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

It follows that

- $$M_1 = E[X] = \frac{1}{n}S_0\sum_{i=1}^{n}\exp(rT_i)$$
$$M_2 = E[X^2] = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E[X_iX_j]$$
$$= \frac{1}{n^2}S_0^2\sum_{i=1}^{n}\exp(2rT_i+\sigma^2T_i) + \frac{2}{n^2}S_0^2\sum_{i=1}^{n}\sum_{j>i}\exp(r(T_j-T_i)+2rT_i+\sigma^2T_i)$$
- *The random variable* $Y = \mu\exp(-\frac{1}{2}\nu^2+\nu Z)$ *has first moment* μ *and second moment is* $\mu^2\exp(\nu^2)$.
- *Let* $\mu = M_1, \nu^2 = \log(M_2/M_1^2)$, *we have*

$$E[Y] = M_1, E[Y^2] = M_2.$$

Proof. (1)(a) Use linearity of expectation. Note that

$$E[\exp(\sigma\sqrt{T_j}Z)] = \exp(\frac{1}{2}\sigma^2T_j), Z\sim N(0,1).$$

(b) Note that for $j>i$, we have

$$X_iX_j = X_i^2(X_j/X_i) = X_i^2\exp((r-\frac{1}{2}\sigma^2)(T_j-T_i)+\sigma\sqrt{T_j-T_i}Z), Z\sim N(0,1).$$

Then due to independence

$$\begin{aligned}E[X_iX_j] &= E[X_i^2]E[\exp((r-\frac{1}{2}\sigma^2)(T_j-T_i)+\sigma\sqrt{T_j-T_i}Z)]\\ &= S_0^2\exp((2r+\sigma^2)T_j)\exp(r(T_j-T_i))\end{aligned}$$

(2)

$$E[Y] = \mu\exp(-\frac{1}{2}\nu^2)E[\exp(\nu Z)] = \mu\exp(-\frac{1}{2}\nu^2)\exp(\frac{1}{2}\nu^2) = \mu.$$

and

$$E[Y^2] = \mu^2\exp(-\nu^2)E[\exp(2\nu Z)] = \mu^2\exp(-\nu^2)\exp(\frac{1}{2}4\nu^2) = \mu^2\exp(\nu^2).$$

(3) Straight forward. □

Lemma 2.3.5 (European call price using moment matching approximation). *Let $0 < T_1 < T_2 < ... < T_n \le T$ be a set of dates. Consider a set of random variables given by*

$$X_j = F_0 \exp(-\frac{1}{2}\sigma^2 T_j + \sigma W(T_j)).$$

and the averaging random variable

$$X = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Consider a call option with strike K and maturity T such that its final payoff at T is given by

$$V(T) = (X - K)^+$$

. It follows that its value at $t = 0, t < T_1$, approximated by log-normal distribution, is given by

$$V(0) = P(t,T)(M_1 N(d_1) - KN(d_2),$$

where

$$d_{\pm} = \frac{\ln(M_1/K) \pm \frac{1}{2}\nu^2}{\nu},$$

and

$$M_1 = F_0, M_2 = \frac{2}{n^2}\sum_{i=1}^{n}(n-i+1)F_0^2 \exp(\sigma^2 T_i), \nu^2 = \log(M_2/M_1^2).$$

Proof. Using lognormal moment matching(Lemma 2.3.4), we know that the distribution of X can be approximated by $Y = M_1 \exp(-\frac{1}{2}\nu^2 + \nu Z), Z \sim N(0,1)$.Then we follow the standard procedure of deriving vanilla call option. □

Lemma 2.3.6 (two-parameter lognormal moment matching for weighted averaging log-normal). *[11][12, p. 232] Let $T_1, T_2, ..., T_n$ be a set of dates. Consider a set of random variables given by*

$$X_j = S_0 \exp((r - \frac{1}{2}\sigma^2)T_j + \sigma W(T_j)).$$

and the averaging random variable

$$X = \sum_{i=1}^{n} w_i X_i.$$

It follows that

- $$M_1 = E[X] = S_0 \sum_{i=1}^{n} w_i \exp(rT_i)$$
$$M_2 = E[X^2] = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j E[X_i X_j]$$
$$= S_0^2 \sum_{i=1}^{n} w_i^2 \exp(2rT_i + \sigma^2 T_i) + S_0^2 \sum_{i=1}^{n}\sum_{j>i} 2w_i w_j \exp(r(T_j - T_i) + 2rT_i + \sigma^2 T_i)$$
- *The random variable* $Y = \mu \exp(-\frac{1}{2}\nu^2 + \nu Z)$ *has first moment* μ *and second moment is* $\mu^2 \exp(\nu^2)$.
- *Let* $\mu = M_1, \nu^2 = \log(M_2/M_1^2)$, *we have*
$$E[Y] = M_1, E[Y^2] = M_2.$$

Proof. See Lemma 2.3.4. □

2.3.6 Multivariate normal model

Definition 2.3.4 (multivariate normal model). *Let current time be t. In a Black lognormal model, we assume the spot prices of assets* $S_1, S_2, ..., S_N$ *have a multivariate lognormal distribution under risk-neutral measure at time* T, *given by*

$$S_i(T) = S_t + (r-q)(T-t) + \sigma_i \sqrt{T-t} Z, Z \in N(0,1);$$

$$corr(S_i, S_j) = \rho_{ij}.$$

Similarly, we assume the forward prices of assets $F_1(t, T_M), F_2(t, T_M), ..., F_N(t, T_M)$ *have a multivariate lognormal distribution under risk-neutral measure at time* T, *given by*

$$F_i(T, T_M) = F_i(T, T_M) + \sigma_i \sqrt{T-t} Z, Z \in N(0,1).$$

$$corr(F_i(T, T_M), F_i(T, T_M)) = \rho_{ij}.$$

Remark 2.3.3 (applications). Multivariate normal model can be used to model the static distribution of averages of interest rate.

2.4 Local volatility method

2.4.1 Fundamentals

Proposition 2.4.1 (Dupire's equation). *[5, p. 245][2, p. 57] Let $C(K,T;S_0)$ denote the* ***current market call prices*** *at different strikes K and different maturities T. Further assume an asset S_t has risk-neutral SDE given by*

$$dS_t = (r-d)S_t dt + \sigma(S_t,t)S_t dW_t.$$

If the local volatility $\sigma(S,T)$ satisfies

$$C_T = \frac{1}{2}\sigma^2 K^2 C_{KK} - (r-d)KC_K - dC,$$

or equivalently

$$\sigma^2(K,T;S_0) = \frac{C_T + (r-d)KC_K + dC}{\frac{1}{2}K^2 C_{KK}},$$

where

$$C_T = \frac{C(K,T)}{\partial T}, C_K = \frac{C(K,T)}{\partial K}, C_{KK} = \frac{C(K,T)}{\partial K^2},$$

then

$$E_Q[(S_T-K)^+|S_0] = C(K,T|S_0), \forall K,T.$$

That is, the Black model prediction matches with the observed market price exactly.

Proof. Note that the Fokker-Planck equation(**??**) governing probability evolution p of S_t under risk-neutral measure is given by

$$p_t = -[(r-d)sp]_s + \frac{1}{2}[(\sigma s)^2 p]_s s.$$

The martingale pricing method gives the call option price with maturity T and strike K given by

$$C = e^{-rT}E_Q[(S_T-K)^+] = e^{-rT}\int_K^\infty (s-K)pds = e^{-rT}\int_K^\infty spds - Ke^{-rT}\int_K^\infty pds.$$

Differentiating C with respect to K gives $C_K = -e^{-rT}\int_K^\infty pds$, then we have $C - KC_K = e^{-rT}\int_K^\infty spds$. Differentiating C with respect to K twice $C_{KK} = e^{-rT}p$. Differentiating C with respect to T gives

$$\begin{aligned}
C_T &= -rC - e^{-rT}\int_K^\infty (s-K)[(r-d)sp]_s ds + \frac{1}{2}e^{-rT}\int_K^\infty (s-K)[(\sigma s)^2 p]_{ss} \\
&= -rC + (r-d)e^{-rT}\int_K^\infty spds + \frac{1}{2}e^{-rT}(\sigma K)^2 p \\
&= -rC + (r-d)(C - KC_K) + \frac{1}{2}\sigma^2 K^2 C_{KK} \\
&= -(r-d)KC_K - dC + \frac{1}{2}\sigma^2 K^2 C_{KK}
\end{aligned}$$

□

Remark 2.4.1 (interpretation)**.**

- Using the local volatility calculated from Dupire's equation and the martingale pricing formula, we have able to reproduce the market prices for calls with different strikes and maturities.
- Note that what we get from Dupire's equation is $\sigma(S_T, T)$. When we want to use $\sigma(S_T, T)$ in Monte Carlo simulation, we replace T by t and K by S_t. The validity of this replacement is the proof(the Fokker-Planck equation part).

Remark 2.4.2 (static smile property of local volatility surface)**.** If the market option price set are the same for different spot price S_0, the local volatility $\sigma(S, T)$ will be the same.

Lemma 2.4.1 (deriving implied distribution from local volatility). *Let $\sigma_{loc}(K, t)$ be the local volatility satisfying Dupire's equation. Then the implied distribution for S_t at a future time t is given by*

$$S(t) = S(0)\exp\left(\int_0^t [\mu(S_u, u) - \frac{1}{2}\sigma(S_u, u)^2_{loc}]du + \int_0^t \sigma(S_u, u)^*_{loc} dB(u)\right).$$

where $B(t)$ is a Brownian motion, and μ is the drift at risk-neutral measure.

Proof. The solution to the SDE

$$dS(t) = \mu(t)S(t)dt + \sigma^*(t)S(t)dB(t)$$

is at **??**. □

Lemma 2.4.2 (connection between local volatilty and implied volatility when there is no strike dependence). *[1, p. 279] If both implied volatility σ^* and local volatility σ_{loc} have no dependence on the strike K. Then,*

$$\frac{1}{T}\int_0^T \sigma_{loc}^2(u)du = \sigma^{*2}(T),$$

that is, implied volatility square is the time-average value of local volatility square.

Proof. When there is no strike dependence, the implied distribution from the local volatility is given by

$$\begin{aligned} S(T) &= S(0)\exp(\int_0^T [\mu(S_u,u) - \frac{1}{2}\sigma(S_u,u)_{loc}^2]du + \int_0^T \sigma(S_u,u)_{loc}^* dB(u)) \\ &= S(0)\exp(\int_0^T \mu(S_u,u)du - \frac{1}{2}\sigma_m^2 T + \sigma_m B(T)) \end{aligned}$$

where we use result in **??** to derive

$$\sigma_m^2 = \frac{1}{T}\int_0^T \sigma_{loc}^2(u)du.$$

The distribution of

$$S(0)\exp(\int_0^T \mu(S_u,u)du - \frac{1}{2}\sigma_m^2 T + \sigma_m B(T))$$

with σ_m replaced by σ_{loc} is exactly what we use to derive Black-Scholes equation. □

Lemma 2.4.3 (deriving local volatility from implied volatility). *[1, p. 278][13, p. 13] The local volatility can be derived from implied volatility, given by,*

$$\sigma^2(K,T;S_0) = \frac{C_T + (r-d)KC_K + dC}{\frac{1}{2}K^2 C_{KK}},$$

where

$$\begin{aligned} C_K &= \frac{\partial C_{BS}}{\partial K} + \frac{\partial C_{BS}}{\partial \sigma^*}\frac{\partial \sigma^*}{\partial K} \\ C_T &= \frac{\partial C_{BS}}{\partial T} + \frac{\partial C_{BS}}{\partial \sigma^*}\frac{\partial \sigma^*}{\partial T} \\ C_{KK} &= \frac{\partial^2 C_{BS}}{\partial K^2} + \frac{\partial^2 C_{BS}}{\partial \sigma^* \partial K}\frac{\partial \sigma^*}{\partial K} + (\frac{\partial}{\partial K}(\frac{\partial C_{BS}}{\partial \sigma^*})) + \frac{\partial C_{BS}}{\partial \sigma^*}\frac{\partial^2 \sigma^*}{\partial K^2} \end{aligned}$$

$$\sigma^2(K,T;S_0) = \frac{\frac{\partial\sigma}{\partial T} + \frac{\sigma}{T} + 2K(r-d)\frac{\partial\sigma}{\partial K}}{K^2(\frac{\partial^2\sigma}{\partial K^2} - D\sqrt{T}(\frac{\partial\sigma}{\partial K})^2 + \frac{1}{\sigma}(\frac{1}{K\sqrt{T}} + D\frac{\partial\sigma}{\partial K})^2)},$$

where $D = \frac{1}{\sigma\sqrt{T}}(\ln\frac{S_0}{K} + (r - d + \frac{1}{2}\sigma^2)T)$.

Denote $w = w(T,K) = \sigma^2(T,K)T$*, then*

$$\sigma^2(K,T;S_0) = \frac{\frac{\partial w}{\partial T}}{1 - \frac{\kappa}{w}\frac{\partial w}{\partial\kappa} + \frac{1}{4}(\frac{\partial w}{\partial\kappa})^2(\frac{\kappa^2}{w^2} - \frac{1}{w} - \frac{1}{4}) + \frac{1}{2}\frac{\partial^2 w}{\partial\kappa^2}}$$

where $\kappa = \ln(\frac{K}{F})$*, F is the forward price.*

Proof. Note that the market prices of calls, $C(K,T)$, can be written by

$$C(K,T) = C_{BS}(S_t, K, T, \sigma^*(K,T)),$$

where C_{BS} is the Black-Scholes formula for calls and σ^* is the implied volatility.

Then, we use chain rule. □

2.5 Stochastic volatility model

2.5.1 Black-Scholes model for stochastic volatility

Lemma 2.5.1 (Black-Scholes model with stochastic volatility). *[14, p. 300][1, p. 345][15, p. 882] Assume **under real-world probability measure**, we have dynamics*

$$\begin{aligned} dS &= \mu S dt + \sigma S dW_1 \\ d\sigma &= p(S,\sigma,t)dt + q(S,\sigma,t)dW_2 \end{aligned}$$

where W_1 and W_2 are two correlated Brownian motions with $dW_1 dW_2 = \rho$.

Then the value of an option on S, denoted by $V(S,\sigma,t)$ is governed by

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + \rho\sigma S q\frac{\partial^2 V}{\partial S\partial\sigma} + \frac{1}{2}q^2\frac{\partial^2 V}{\partial\sigma^2} + rS\frac{\partial V}{\partial S} + (p - \lambda q)\frac{\partial V}{\partial\sigma} - rV = 0.$$

Proof. Consider a portfolio

$$\Pi = V - \Delta S - \Delta_1 V_1,$$

where V_1 is another tradable option to hedge volatility risks.

$$\begin{aligned}
d\Pi = (&\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma S q \frac{\partial^2 V}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V}{\partial \sigma^2})dt \\
-\Delta_1(&\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho\sigma S q \frac{\partial^2 V_1}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V_1}{\partial \sigma^2})dt \\
&+(\frac{\partial V}{\partial S} - \Delta - \Delta_1 \frac{\partial V_1}{\partial S})dS \\
&+(\frac{\partial V}{\partial \sigma} - \Delta_1 \frac{\partial V_1}{\partial \sigma})d\sigma
\end{aligned}$$

To eliminate all randomness, we choose Δ, Δ_1 such that

$$\frac{\partial V}{\partial S} - \Delta - \Delta_1 \frac{\partial V_1}{\partial S} = 0, \frac{\partial V}{\partial \sigma} - \Delta_1 \frac{\partial V_1}{\partial \sigma} = 0.$$

This gives us

$$\Delta_1 = \frac{\partial V / \partial \sigma}{\partial V_1 / \partial \sigma}, \Delta = \frac{\partial V}{\partial S} + \frac{\partial V / \partial \sigma}{\partial V_1 / \partial \sigma} \frac{\partial V_1}{\partial S}.$$

Then, we have

$$\begin{aligned}
d\Pi = (&\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma S q \frac{\partial^2 V}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V}{\partial \sigma^2})dt \\
& - \Delta_1(\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho\sigma S q \frac{\partial^2 V_1}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V_1}{\partial \sigma^2})dt \\
& = r\Pi dt = r(V - \Delta S - \Delta_1 V_1)dt
\end{aligned}$$

Rearrange and remove cancel dt, we have

$$\begin{aligned}
&\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma S q \frac{\partial^2 V}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V}{\partial \sigma^2} + rS\frac{\partial V}{\partial S} - rV / \frac{\partial V}{\partial \sigma} \\
= &\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho\sigma S q \frac{\partial^2 V_1}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V_1}{\partial \sigma^2} + rS\frac{\partial V_1}{\partial S} - rV_1 / \frac{\partial V_1}{\partial \sigma}
\end{aligned}$$

Note that since V_1 is an arbitrary option, therefore, the ratio can only be a function, denoted by $-(p - \lambda q)$, of S, σ, t (without V, V_1). Thus, we have

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma S q \frac{\partial^2 V}{\partial S \partial \sigma} + \frac{1}{2} q^2 \frac{\partial^2 V}{\partial \sigma^2} + rS\frac{\partial V}{\partial S} - rV / \frac{\partial V}{\partial \sigma} = -(p - \lambda q).$$

□

Lemma 2.5.2 (risk-neutral pricing in stochastic volatility model). *Assume there exists a risk-neutral measure Q such that the dynamics of the asset and the volatility is given by*

$$\begin{aligned} dS &= rSdt + \sigma SdW_1 \\ d\sigma &= m(S,\sigma,t)dt + q(S,\sigma,t)dW_2 \end{aligned}$$

where W_1 and W_2 are two correlated Brownian motions with $dW_1 dW_2 = \rho$.

Then for a derivative with payoff $V_T(S_T)$, its current value is given by

$$V(S_t,t) = E_Q[e^{-r(T-t)}V_T(S_T)|\mathcal{F}_t]$$

Proof. We want to show that the dynamics of $e^{-rt}V(t,S_t,\sigma_t)$ **under risk-neutral measure** is driftless SDE.

$$\begin{aligned} &d(\exp(-rt)V(t,S_t,\sigma_t) \\ &= -\exp(-rt)rV + \exp(-rt)dV \\ &= -\exp(-rt)rV + \exp(-rt)[(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + \rho\sigma Sq\frac{\partial^2 V}{\partial S\partial\sigma} + \frac{1}{2}q^2\frac{\partial^2 V}{\partial\sigma^2})dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial\sigma}d\sigma] \\ &= \frac{\partial V}{\partial S}\sigma SdW_1 + \frac{\partial V}{\partial\sigma}qdW_2 \\ &= \sqrt{(\frac{\partial V}{\partial S}\sigma S)^2 + (\frac{\partial V}{\partial\sigma}q)^2 + (2\rho\frac{\partial V}{\partial\sigma}q)\frac{\partial V}{\partial S}\sigma S}dW_3 \end{aligned}$$

where W_3 is a Brownian motion and we use the Black-Scholes equation in the derivation. □

Remark 2.5.1 (risk-neutral interpretation using Feyman Kac theorem)**.** In Lemma 2.5.1, we show that the value of the derivative is governed by

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + \rho\sigma Sq\frac{\partial^2 V}{\partial S\partial\sigma} + \frac{1}{2}q^2\frac{\partial^2 V}{\partial\sigma^2} + rS\frac{\partial V}{\partial S} + (p-\lambda q)\frac{\partial V}{\partial\sigma} - rV = 0.$$

The Feyman-Kac theorem enables us to interpret the underlying dynamics, where we replace $(p-\lambda q)$ by m.

Remark 2.5.2 (how to implement the pricing model)**.**

- We can use Monte Carlo simulation to evaluate the expectation as long as we have the model parameter under risk-neutral measure.
- Even we know the model parameter in the real-probabiliy measure(which can be estimated from the historical data), we cannot use them. The risk-neutral parameter has to be estimated from prices of other derivatives.

Note 2.5.1 (dynamics in real probability measure and risk-neutral measure). The asset dynamics under risk-neutral probability is given by

$$dS = rSdt + \sigma SdW_t$$

and

$$\begin{aligned} dV &= (\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})dt + \frac{\partial V}{\partial S}dS \\ &= rVdt + \frac{\partial V}{\partial S}\sigma SdW_1 + \frac{\partial V}{\partial \sigma}qdW_2 \end{aligned}$$

where we use the Black-Scholes equation in the derivation.

Under real probability measure, the dynamics are

$$dS = \mu Sdt + \sigma SdW_t$$

and

$$\begin{aligned} dV &= (\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{2}q^2 \frac{\partial^2 V}{\partial \sigma^2} + \rho q \sigma \frac{\partial^2 V}{\partial \sigma^2})dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial \sigma}d\sigma \\ &= rV + \frac{\partial V}{\partial S}S(\mu - r)dt + \frac{\partial V}{\partial \sigma}(p - m) + \frac{\partial V}{\partial S}\sigma SdW_1 + \frac{\partial V}{\partial \sigma}qdW_2 \end{aligned}$$

If we define $\lambda_1 = (\mu - r)/\sigma$ and $\lambda_2 = (p - m)/q$, we find that

$$\begin{aligned} dS/S &= (r + \lambda\sigma)dt + \sigma dW_t \\ dV/V &= (r + \lambda_1 \frac{\partial V}{\partial S}\frac{\sigma S}{V} + \lambda_2 \frac{\partial V}{\partial \sigma}\frac{q}{V})dt + \frac{\partial V}{\partial S}\frac{\sigma S}{V}dW_1 + \frac{\partial V}{\partial \sigma}\frac{q}{V}dW_2 \end{aligned}$$

which is consistent with the no-arbitrage condition for multiple source uncertainty dynamics(Theorem 1.5.1).

Example 2.5.1 (Heston stochastic volatility model). [8, p. 288] Suppose that under a risk-neutral measure Q, a stock price is governed by

$$dS(t) = rS(t)dt + \sqrt{V(t)}S(t)dW(t),$$

where the interest rate r is constant and volatility $\sqrt{V(t)}$ is governed by

$$dV(t) = (a - bV(t))dt + \sigma\sqrt{V(t)}dW_2,$$

where W_1 and w_2 are correlated Brownian motions with $dW_1dW_2 = \rho dt$.

Consider a call maturing at $T \geq t$ has price at t given by

$$c(t, S(t), V(t)) = E_Q[e^{-r(T-t)} \max(S(T) - K, 0)|\mathcal{F}_t],$$

then $c(t, s, v)$ will satisfy the following PDE

$$c_t + rsc_s + (a - bv)c_v + \frac{1}{2}s^2 vc_{ss} + \rho\sigma svc_{sv} + \frac{1}{2}\sigma^2 vc_{vv} = rc$$

in the region $0 \leq t \leq T, s \geq 0$ and $v \geq 0$. The function $c(t, s, v)$ also satisfies the boundary condition

$$\begin{aligned}
c(T, s, v) &= \max(s - K, 0), \forall s \geq 0, v \geq 0 \\
c(t, 0, v) &= 0, \forall 0 \leq t \leq T, v \geq 0 \\
c(t, s, 0) &= \max(s - e^{-r(T-t)}K, 0), \forall 0 \leq t \leq T, s \geq 0 \\
\lim_{s\to\infty} \frac{c(t, s, v)}{S - K} &= 1, \forall 0 \leq t \leq T, v \geq 0 \\
\lim_{v\to\infty} c(t, s, v) &= s, \forall 0 \leq t \leq T, s \geq 0
\end{aligned}$$

Remark:
It has practical difficulties in using the pricing formula since $V(t)$ cannot be directly observed from the market.

2.5.2 SARB model

2.5.2.1 *The model*

Definition 2.5.1 (SABR stochastic volatility model). *The futures prices $F(t, T_m)$ with maturity T_m is given by*

$$\begin{aligned}
dF(t, T_m) &= \alpha(t)F^\beta(t, T_m)dW_1 \\
d\alpha(t, T_m) &= \nu\alpha(t, T_m)dW_2 \\
\alpha(0) &= \alpha \\
dW_1dW_2 &= \rho dt
\end{aligned}$$

where $\alpha(t, T_m)$ is the stochastic volatility associated with maturity T_m, β is the CEV constant, ν is the volatility of volatility, α is the initial value of the volatility, and ρ is the correlation coefficient.

Lemma 2.5.3 (European call/option pricing). *[16]*

- *The European call price with strike K and maturity T is given by*

$$V_c(t) = P(t,T)(F(t,T)N(d_1) - KN(d_2))$$

where $F(t,T)$ is the forward price the underlying, and

$$d_{1,2} = \frac{\log(F(t,T)/K) \pm \frac{1}{2}\sigma_B^2(T-t)}{\sigma_B^2\sqrt{T-t}},$$

and the implied volatility $\sigma(F,K)$ is given by

$$\begin{aligned}
\sigma_B(F,K) &= \frac{\alpha}{(FK)^{(1-\beta)/2}[1 + \frac{(1-\beta)^2}{24}\log^2(F/K) + \frac{(1-\beta)^4}{1920}\log^4(F/K) + \cdots]} \\
&\cdot \frac{z}{\chi(z)} \cdot \{1 + [\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(FK)^{1-\beta}} + \frac{1}{4}\frac{\rho\beta\nu\alpha}{(FK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24}\nu^2](T-t) + \cdots\} \\
z &= \frac{\nu}{\alpha}(FK)^{(1-\beta)/2}\log(F/K) \\
\chi(z) &= \log\frac{\sqrt{1-2\rho z + z^2} + z - \rho}{1-\rho}
\end{aligned}$$

- *At-the-money volatility*

$$\sigma_B(F,F) = \frac{\alpha}{(FK)^{(1-\beta)/2}}\{1 + [\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(FK)^{1-\beta}} + \frac{1}{4}\frac{\rho\beta\nu\alpha}{(FK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24}\nu^2](T-t) + \cdots\}$$

- *The put price given by*

$$V_p(t) = V_c(t) + P(t,T)(K - F(t,T)).$$

Note 2.5.2 (choices of parameters for the equity volatility surface)**.**

- At every maturity T, we use four parameter α_0, β, ν and ρ to characterize the volatility smile; usually, β is pre-selected, and the rest is calibrated to implied volatility surface using non-linear optimization.
- For a set of maturity dates $T_1, T_2, ..., T_n$, we use a set of $4n$ parameters to characterize the volatility surface. Usually, β is chosen to be constant independent of maturity date.

Note 2.5.3 (choices of parameters for the interest rate volatility surface). subsection 3.8.1

- At every maturity T and the interest rate or swap rate tenors Δ, we use four parameter α_0, β, ν and ρ to characterize the volatility smile; usually, β is pre-selected, and the rest is calibrated to implied volatility surface using non-linear optimization.
- For a set of maturity dates $T_1, T_2, ..., T_n$ and tenors $\Delta_1, \Delta_2, ..., \Delta_m$, we use a set of $4mn$ parameters to characterize the volatility surface/cube. Usually, β is chosen to be constant independent of maturity date and tenors.

2.5.2.2 *Managing smile risk*

Note 2.5.4 (SABR model and the risks associated with vanilla options).

- Let $BS(f, K, \sigma, T)$ denote the call option prices at time 0. Then the call option price under SABR model can be written as
$$V(f, K, \sigma_B(f, K; \alpha, \beta, \rho, \nu), T).$$
Therefore, market risk is associated with parameters $f, \alpha, \beta, \rho, \mu$(usually β is assumed to be a constant and therefore will be the source of market risk).
- The eliminate the market risks, we want to maintain a portfolio, whose value is denoted by V_P, such that
$$\frac{\partial V_P}{\partial f} = \frac{\partial V_P}{\partial \alpha} = \frac{\partial V_P}{\partial \beta} = \frac{\partial V_P}{\partial \nu} = 0,$$
where
$$\begin{aligned} \frac{\partial V}{\partial f} &= \frac{\partial BS}{\partial f} + \frac{\partial BS}{\partial \sigma_B} \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, \mu)}{\partial f} \\ \frac{\partial V}{\partial \alpha} &= \frac{\partial BS}{\partial \sigma_B} \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, \mu)}{\partial f} \\ \frac{\partial V}{\partial \nu} &= \frac{\partial BS}{\partial \sigma_B} \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, \mu)}{\partial \nu} \\ \frac{\partial V}{\partial \rho} &= \frac{\partial BS}{\partial \sigma_B} \frac{\partial \sigma_B(K, f; \alpha, \beta, \rho, \mu)}{\partial \rho} \end{aligned}$$
- To achieve this, we can buy/sell underlying, and other options. Note that α, β, ν are non-tradables; therefore, we need to buy and sell other options to make the portfolio neutral.

2.6 Multi-factor model

2.6.1 Primer on factor model dynamics

Definition 2.6.1 (one factor model for a curve). *Let $F(t,T)$ denote a forward curve. A factor model is given by*

$$dF(t,T) = \sigma(t,T)dW_t.$$

We can interpret this SDE as inifite number SDE parameterized by parameter T sharing the same risk driver W_t.

Remark 2.6.1 (about factor model).

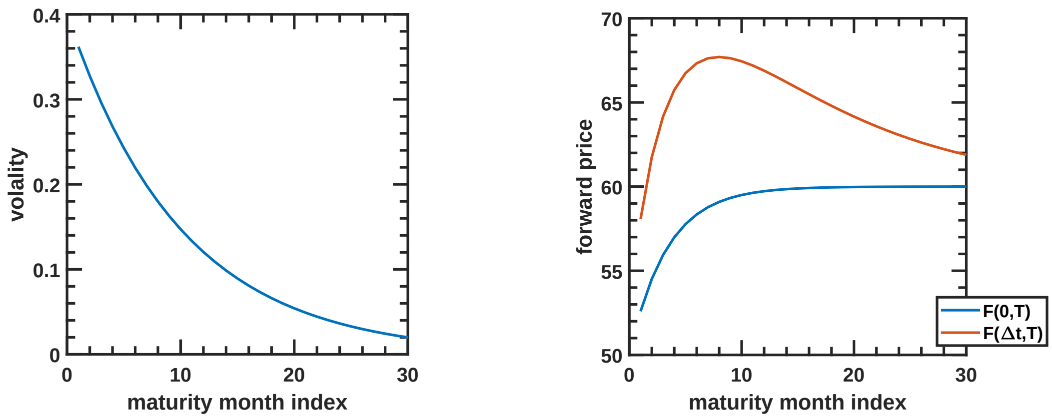
- **Static factor model** represent a large number N of random variables in terms of a small number K of difference random variables called factors.
- **Dynamic factor models** represent a large number N of time series in terms of a small number K of different time series called dynamic factors.

Note 2.6.1 (moving mode with different volatilty structures). The first-order approximation discrete-time form of the one factor model is given by

$$F(t+\Delta t, T) = F(t,T) + \sigma(t,T)\sqrt{\Delta t}Z, Z \in T.$$

By choosing different volatility structure $\sigma(t,T)$, the one-factor model can produce different moving mode in the forward curve.

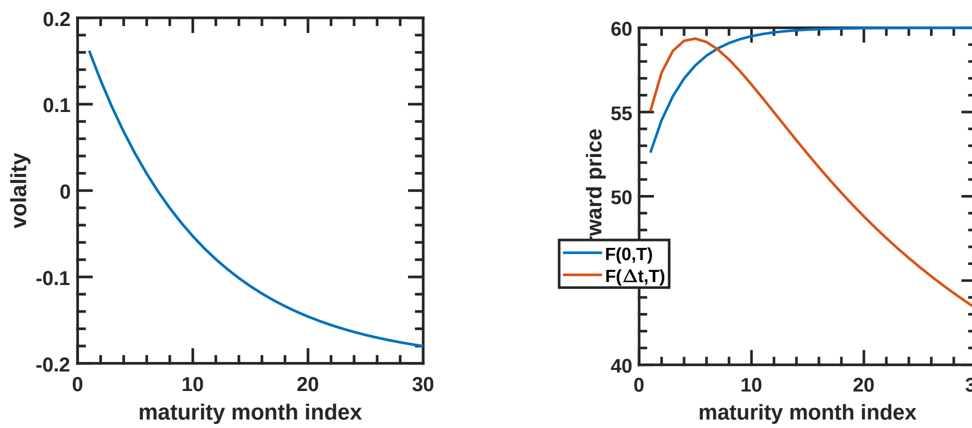
- (**parallel movment**) If we choose $\sigma(t,T) > 0$ for all T, then the curve $F(t+\Delta t, T)$ will be overall above(or below, depends on the realization of Z) the initial curve $F(t,T)$, as showed in Figure 2.6.1.
- (**twisting movement**) If we choose $\sigma(t,T) > 0$ for small T and $\sigma(t,T) < 0$ for large T, then the curve $F(t+\Delta t, T)$ will be above(or below, depends on the realization of Z) the initial curve $F(t,T)$ at the near end and below the initial curve $F(t,T)$ at the far end, as showed in Figure 2.6.2.
- (**bending movement**) If we choose $\sigma(t,T) > 0$ for small and large T, then the curve $F(t+\Delta t, T)$ will be overall above(or below, depends on the realization of Z) the initial curve $F(t,T)$, as showed in Figure 2.6.3.



(a) Volatility term structure $\sigma(t, T)$ as a function of T at time t.

(b) Demonstration of forward curve movement.

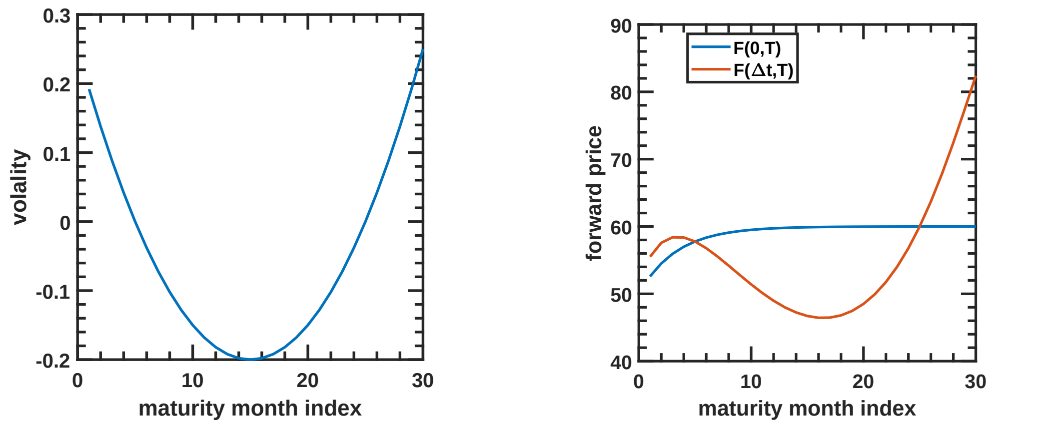
Figure 2.6.1: One factor model producing parallel movement of forward curve



(a) Volatility term structure $\sigma(t, T)$ as a function of T at time t.

(b) Demonstration of forward curve movement.

Figure 2.6.2: One factor model producing twisting movement of forward curve



(a) Volatility term structure $\sigma(t, T)$ as a function of T at time t.

(b) Demonstration of forward curve movement.

Figure 2.6.3: One factor model producing bending movement of forward curve

2.7 Binomial tree market model

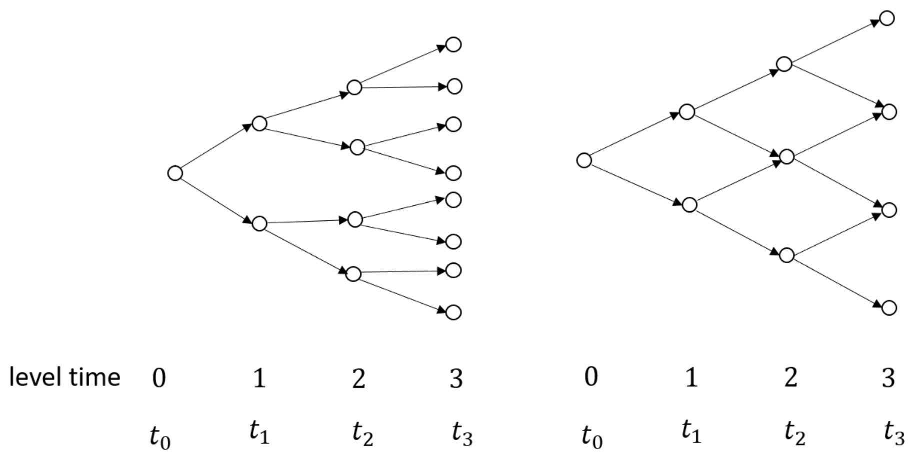


Figure 2.7.1: non-recombining tree (left) vs. recombining tree (right).

Definition 2.7.1 (Binomial tree market model).

- *A risky stock price process $S(t)$, is represented by $S = \{S(t), t = 0, 1, ..., T\}$.*
- *In each time period the stock price either goes up by a factor $u > 1$ with probability p, or goes down by a factor $0 < d < 1$ with probability $1 - p$. For each t, $S(t) = S(0)u^{n_t}d^{t-n_t}$, where n_t represents the number of up moves up to t.*
- *The bank account process B is deterministic with $B(0) = 1$ and a constant interest rate $0 < r < 1$. Hence $B(t) = (1 + r)^t$.*

- *The sample space Ω contains $K = 2^T$ different paths. Each sample point ω is a path with $U(\omega)$ up moves, and $T - U(\omega)$ down moves. The underlying probability P is defined by*

$$P(\omega) = p^U(\omega)(1-p)^{T-U(\omega)}, 0 < p < 1.$$

Lemma 2.7.1. *There exists a **unique** martingale equivalent measure Q such that*

$$Q(\omega) = q^U(\omega)(1-q)^{T-U(\omega)}, q = \frac{1+r-d}{u-d}$$

if and only if $d < 1 + r < u$.

Moreover, the binomial tree model is complete.

Proof. Let $\theta_t = n_t - n_{t-1}$ (i.e., $\theta_t = 1$ is a Bernoulli random variable and n_t is the Binomial random variable) Then for every t, the martingale condition gives

$$\frac{S(t)}{B(t)} = S^*(t) = S^*(t-1)(1+r)^{-1}u^{\theta_t}d^{1-\theta_t}.$$

Therefore,

$$\begin{aligned}
&E_Q[S^*(t)|\mathcal{F}_{t-1}] = S^*(t-1) \\
\Leftrightarrow & uQ(\theta_t = 1|n_{t-1}) + d(1 - Q(\theta_t = 1|n_{t-1})) = 1 + r \\
\Leftrightarrow & q = Q(\theta_t = 1|n_{t-1}) = \frac{1+r-d}{u-d}
\end{aligned}$$

Only when $d < 1 + r < u$, we can guarantee $0 < q < 1$.

The tree model is complete because the martingale measure is unique. □

Remark 2.7.1 (interpretation on parameterizing probability measure function)**.**

- For a multi-period finite state model, it may require $K^T - 1$ number of fully parameterize the probability measure funciton. Then we might need multiple existing assets to find an unique martingale measure.
- The tree model using only one parameter, i.e., jump probability to specify the dynamics of the model and only one parameter to parameterize the probability measure function.

Example 2.7.1 (pricing a call using binomial tree). Consider a call with payoff $g(S(T)) = (S(T) - K)^+$. Note that

$$S(t)u^n d^{T-t-n} - K > 0 \Leftrightarrow n > \frac{\log(K/S(t)d^{T-t})}{\log(u/d)}.$$

Let n^* be the smallest n such that the inequality holds.

Then the value of the call at time t is given by

$$V(t) = \underbrace{\frac{S(t)B(t)}{B(T)} \sum_{n=n^*}^{T-t} \binom{T-t}{n} (uq)^n ((1-q)d)^{T-t-n}}_{B(t)E_Q[S(T)/B(T)\mathbf{1}(S(T)>K)|\mathcal{F}_t]}$$
$$- \underbrace{\frac{KB(t)}{B(T)} \sum_{n=n^*}^{T-t} \binom{T-t}{n} (q)^n ((1-q))^{T-t-n}}_{B(t)E_Q[K/B(T)\mathbf{1}(S(T)>K)|\mathcal{F}_t]}$$

2.7.1 Basic constructions

Definition 2.7.2 (binomial tree model). *A binomial tree model for an asset (e.g. stock) dynamics consists of*

- *nodes (i,j): representing date $i = 0,1,...,N$ and states $j = 0,1,...,N$*
- *values on node (i,j) represent the asset price at i and state j.*
- *the binomial tree model is a representation of discrete-time asset stochastic process $S(1), S(2), ..., S(N)$, where $S(i)$ can take values $s_{i,1}, ..., s_{i,i}$.*
- *state transition probability:*

$$P((i+1,s)|(i,j)) = \begin{cases} q, s = j+1 \\ 1-q, s = j \end{cases}$$

- *the sample space Ω consists of all possible 2^N sample paths; We use $\omega_{i,j}$ to denote the set of sample paths that pass through state j at time i.*

2.7.2 State price and general pricing

Definition 2.7.3 (Arrow-Debreu security). *An **Arrow-Debreu security** is a security that has a payoff that solely pays 1 at time n if world state $\omega_{n,j}$ is realized. Its price, called, **state price** at $t = 0$ is denoted as $\lambda_{n,j}$*

With the existence of risk-neutral measure Q, we have

$$\lambda_{n,j} = \frac{1}{(1+r\Delta t)^n} E_Q[\mathbf{1}(S_n = s_{n,j})] = \frac{1}{(1+r\Delta t)^n} E_Q[\mathbf{1}(\omega_{n,j})].$$

Further, the probability of reaching state $s_{n,j}$ from state $s_{0,0}$ is given by

$$Pr_Q(S_n = s_{n,j}) = (1+r\Delta t)^n \lambda_{n,j}.$$

Proposition 2.7.1 (state price recursive relation). *In a N period binomial tree model, where there will be $K+1$ states, then state prices satisfying* ***forward equations****:*

$$\begin{aligned}
\lambda_{n+1,j} &= \frac{1}{1+r\Delta t}(Q_n(j;j-1)\lambda_{n,j-1} + Q_n(j;j)\lambda_{n,j}), 0 < j \leq n+1 \ (interior) \\
\lambda_{n+1,0} &= \frac{1}{1+r\Delta t} Q_n(0;0)\lambda_{n,0} \ (upper\ boundary) \\
\lambda_{n+1,n+2} &= \frac{1}{1+r\Delta t} Q_n(n+2;n+1)\lambda_{n,n+1} \ (lower\ boundary)
\end{aligned}$$

where we have ***boundary condition*** $\lambda_{0,0} = 1$*,* $Q_n(j;i)$ *denotes the equilibrium measure for transitioning from state i at period n to state j at period $n+1$, and r is the interest rate.*

More over, we have

$$\sum_{j=1}^{n+2} \lambda_{n+1,j} = \frac{1}{1+r\Delta t} \sum_{j=1}^{n+1} \lambda_{n,j}.$$

Proof. (1)Consider an interior node (n,i). By definition,

$$\begin{aligned}
\lambda_{n,i} &= \frac{1}{(1+r\Delta t)^n} E_Q[\mathbf{1}(S_n = s_{n,i})] \\
&= \frac{1}{(1+r\Delta t)^n} Pr_Q(S_n = s_{n,i}) \ (Pr_Q \ is\ the\ measure) \\
&= \frac{1}{(1+r\Delta t)^n} [Pr_Q((S_n = s_{n,i}) \cap (S_{n-1} = s_{n-1,i}) + Pr_Q((S_n = s_{n,i}) \cap (S_{n-1} = s_{n-1,i-1})] \\
&= \frac{1}{(1+r\Delta t)^n} [Q_{n-1}(i,i) Pr_Q(S_{n-1} = s_{n-1,i}) + Q_{n-1}(i,i-1) Pr_Q(S_{n-1} = s_{n-1,i-1})] \\
&= \frac{1}{(1+r\Delta t)} [\frac{1}{(1+r\Delta t)^{n-1}} Q_{n-1}(i,i) E_Q[\mathbf{1}(S_{n-1} = s_{n-1,i})] + Q_{n-1}(i,i-1) E_Q[\mathbf{1}(S_{n-1} = s_{n-1,i-1})]] \\
&= \frac{1}{1+r\Delta t}(Q_n(i;i)\lambda_{n-1,i} + Q_n(i;i-1)\lambda_{n-1,i-1})
\end{aligned}$$

(2)

$$\sum_{j=1}^{n+2} \lambda_{n+1,j} = \frac{1}{(1+r\Delta t)}(\lambda_{n,0}(Q_n(0,0)+Q_n(1,0)) + \lambda_{n,1}(Q_n(1,1)+Q_n(2,1)) + \cdots + \lambda_{n,n+1}(Q_n(n+1,n+$$
$$= \frac{1}{(1+r\Delta t)}(\lambda_{n,0} + \lambda_{n,1} + \cdots + \lambda_{n,n+1})$$

□

Proposition 2.7.2 (pricing simple derivatives using state prices). *Suppose $\lambda_{i,j}$ is known for all possible i, j.*

- *for a zero coupon bond that has a payoff 1 at period n, its no-arbitrage price at time 0 is*

$$V(0) = \sum_{j=1}^{n+1} \lambda_{n,j} = \frac{1}{(1+r\Delta t)^n}.$$

- *for any derivative that has a payoff vector $D(S_n)$ at period n, its arbitrage price at time 0is*

$$V(0) = \sum_{j=1}^{n+1} \lambda_{n,j} D(s_{n,j}).$$

Proof. (1) Using relationship

$$\lambda_{i,j} = \frac{1}{(1+r\Delta t)^n} E_Q[\mathbf{1}(S_i = s_{i,j})] = \frac{1}{(1+r\Delta t)^n} E_Q[\mathbf{1}(\omega_{i,j})]$$

(2) Using either linear pricing theorem(Theorem 1.2.2). Or using definition

$$V(0) = \frac{1}{(1+r\Delta t)^n} E_Q[D(S_n)] = \sum_{j=1}^{n+1} \frac{1}{(1+r\Delta t)^n} D(s_{n,j}) E_Q[\mathbf{1}(S_n = s_{n,j})].$$

□

Remark 2.7.2 (pricing more complex derivative using state price).

- Pricing using state price usually limit to European style derivatives or American style derivatives with addition dynamical programming.
- For more complex derivative that involves the joint distribution of S at multiple time points (e.g., spread option), the state price method cannot help.

2.7.3 GBM to a Binomial Model

Given the continuous geometric Brownian motion, we can convert them into equivalent binomial model parameters:

- $R_n = \exp(rT/n)$, where n is the number of periods in binomial model
- $R_n - c_n = \exp((r-c)T/n) \approx 1 + rT/n - cT/n$, r is the short rate and c is the constant dividend paying rate.
- $u_n = \exp(\sigma\sqrt{T/n})$
- $d_n = 1/u_n$
- the risk-neutral probability is calculated as $q = \frac{\exp((r-c)T/n)-d}{u-d}$

Definition 2.7.4 (binomial model). *A binomial lattice model for the stock consists of*

- *nodes (i,j): representing date $i = 0,1,...,n$ and states $j = 0,1,...,n$*
- *values on node (i,j) represent the stock price at i and state j.*
- *state transition probability:*

$$P((i+1,s)|(i,j)) = \begin{cases} q, s = j+1 \\ 1-q, s = j \end{cases}$$

2.7.4 Convergence of Binomial model

Proposition 2.7.3 (convergence in distribution of the Binomial model). *[17, p. 169] Let $r_n = e^{r/n} - 1, d_n = e^{-\sigma/\sqrt{n}}, u_n = e^{\sigma/\sqrt{n}}$ such that the equivalent equivalent martingale measure of the nth binomial model for simulating S is given as*

$$q = \frac{r_n + 1 - d_n}{u_n - d_n}$$

Then the multi-period binomial model for $S_n(t)$ under risk-neutral probability converges in distribution (as $n \to \infty$) to

$$S_t = S_0 \exp(\sigma W_t + (r - \frac{1}{2}\sigma^2)t)$$

where S_t is the solution to the Ito process

$$dS = rSdt + \sigma SdW$$

Remark 2.7.3 (implications). We know that under martingale measure, the geometric Brownian motion of the stock will be

$$dS = rSdt + \sigma SdW.$$

And this theorem shows that if we increase n, the binomial tree simulation is approximating the stock price dynamics under martingale measure.

2.7.5 Two factor trees

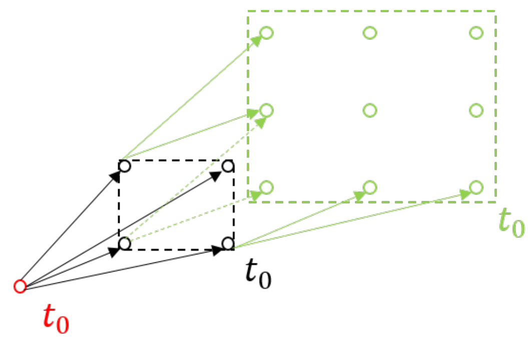


Figure 2.7.2: A three-dimensional tree representation of two-asset dynamics

Remark 2.7.4. A three-dimensional tree might be used to approximate the two-asset dynamics given by

$$\begin{aligned} dS_1/S_1 &= (r - q_1)dt + \sigma_1 dW_1 \\ dS_2/S_2 &= (r - q_2)dt + \sigma_2 dW_2 \\ dW_1 dW_2 &= \rho dt \end{aligned}$$

We can use transformation

$$x_1 = \sigma_2 \ln S_1 + \sigma_1 \ln S_2, x_2 = \sigma_2 \ln S_1 - \sigma_1 \ln S_2,$$
$$S_1 = \exp(\frac{x_1 + x_2}{2\sigma_2}), S_2 = \exp(\frac{x_1 - x_2}{2\sigma_1}),$$

and use the following independent process

$$dx_1 = (\sigma_2(r - q_1 - \sigma_1^2/2) + \sigma_1(r - q_2 - \sigma_2^2/2))dt + \sigma_1\sigma_2\sqrt{2(1+\rho)}dZ_A$$
$$dx_1 = (\sigma_2(r - q_1 - \sigma_1^2/2) - \sigma_1(r - q_2 - \sigma_2^2/2))dt + \sigma_1\sigma_2\sqrt{2(1+\rho)}dZ_B$$

where Z_A, Z_B are indepndent Brownian motions.

2.8 Notes on bibliography

For no-arbitrage theory, see [18].

For treatment from economical perspective, see [19] [20].

For martingale methods, see [21].

For PDE methods, see [21][15].

For incomplete markets, see [22].

For the differences between real world measure and risk-neutral measure, see [23].

An excellent book of "P" method for risks and asset allocations, see [24].

For a comprehensive discussion on risks, see [25].

For statistical arbitrage, see [26].

For stochastic volatility model, see [27], [28], [29].

For empirical evidence of volatility surface, see [30]. For no-arbitrage theory, see [18].

For treatment from economical perspective, see [19] [20].

BIBLIOGRAPHY

1. Derman, E., Miller, M. & Park, D. *The Volatility Smile* ISBN: 9781118959169. https://books.google.com/books?id=WKfBDAAAQBAJ (Wiley, 2016).
2. Bossu, S. *Advanced Equity Derivatives: Volatility and Correlation* (John Wiley & Sons, 2014).
3. Carr, P. & Madan, D. Towards a theory of volatility trading. *Volatility: New estimation techniques for pricing derivatives,* 417–427 (1998).
4. Hull, J. *Options, Futures and Other Derivatives 9 edition* (Pearson Education, 2009).
5. Alexander, C. *Market Risk Analysis, Pricing, Hedging and Trading Financial Instruments* ISBN: 9780470772812. https://books.google.com/books?id=M-6UEM18UHkC (Wiley, 2008).
6. Natenberg, S. *Option Volatility and Pricing Workbook, Second Edition* ISBN: 9780071819053. https://books.google.com/books?id=5calngEACAAJ (McGraw-Hill Education, 2014).
7. Gatheral, J. & Jacquier, A. Arbitrage-free SVI volatility surfaces. *Quantitative Finance* **14,** 59–71 (2014).
8. Shreve, S. E. *Stochastic calculus for finance II: Continuous-time models* (Springer Science & Business Media, 2004).
9. Borovkova, S., Permana, F. J. & Weide, H. V. A closed form approach to the valuation and hedging of basket and spread options. *Journal of Derivatives* **14,** 8 (2007).
10. Iwasawa, K. Analytical formula for the European normal Black Scholes formula. *Unpublished working paper, viewed* **27** (2013).
11. Brigo, D., Mercurio, F., Rapisarda, F. & Scotti, R. Approximated moment-matching dynamics for basket-options simulation (2001).
12. Joshi, M. S. *The concepts and practice of mathematical finance* (Cambridge University Press, 2003).
13. Gatheral, J. *The volatility surface: a practitioner's guide* (John Wiley & Sons, 2011).
14. Wilmott, P. *Derivatives: The Theory and Practice of Financial Engineering* ISBN: 9780471983897. https://books.google.com/books?id=9NPuAAAAMAAJ (Wiley, 1998).
15. Wilmott, P. *Paul Wilmott on quantitative finance* (John Wiley & Sons, 2013).
16. Hagan, P. K. & Kumar, D. Managing smile risk. *Wilmott Magazine,* 84–108.
17. Steland, A. *Financial statistics and mathematical finance: methods, models and applications* (John Wiley & Sons, 2012).

18. Björk, T. *Arbitrage theory in continuous time* (Oxford university press, 2004).
19. LeRoy, S. F. & Werner, J. *Principles of financial economics* (Cambridge University Press, 2014).
20. Mishkin, F. S. *The economics of money, banking, and financial markets* (Pearson education, 2007).
21. Pascucci, A. *PDE and martingale methods in option pricing* (Springer Science & Business Media, 2011).
22. Cernỳ, A. *Mathematical techniques in finance: tools for incomplete markets* (Princeton University Press, 2009).
23. Meucci, A. 'P'Versus' Q': Differences and Commonalities between the Two Areas of Quantitative Finance. *GARP Risk Professional*, 47–50 (2011).
24. Meucci, A. *Risk and asset allocation* (Springer Science & Business Media, 2009).
25. Jorion, P. *Financial Risk Manager Handbook: FRM Part I / Wiley Finance* **pt. 2.** ISBN: 9781118017913. https://books.google.com/books?id=4ceVmGJSNpcC (Wiley, 2010).
26. Leung, T. & Li, X. *Optimal Mean Reversion Trading: Mathematical Analysis and Practical Applications* (World Scientific, 2015).
27. Lewis, A. L. *Option Valuation Under Stochastic Volatility II* (Finance Press, Newport Beach, CA, 2009).
28. Zhu, J. *Applications of Fourier transform to smile modeling: Theory and implementation* (Springer Science & Business Media, 2009).
29. Mikhailov, S. & Nögel, U. *Heston's stochastic volatility model: Implementation, calibration and some extensions* (John Wiley and Sons, 2004).
30. Daglish, T., Hull, J. & Suo, W. Volatility surfaces: theory, rules of thumb, and empirical evidence. *Quantitative Finance* **7,** 507–524 (2007).

3

INTEREST RATE MODELS & DERIVATIVES

3.1 Interest rate concepts

3.1.1 Basic definitions

3.1.1.1 *Zero-coupon bond*

Definition 3.1.1 (zero-coupon bond). *A **zero-coupon bond** is a bond that pays its face value (say, $1)at the maturity time T. The set of prices of zero-coupon bonds for various time-horizons/maturities is known as the **zero-coupon curve**.*

Remark 3.1.1 (importance of zero-coupon bond)**.**

- Zero-coupon bonds are not actively traded within the interbank market; however, they are important because interbank interest rates such as LIBOR and swap rates can be defined in terms of zero-coupon bonds.

Definition 3.1.2 (yield of zero-coupon bond). *The yield of the zero-coupon bond with maturity T is the equivalent **constant interest rate** such that the price of the zero-coupon bond invested at time 0 and accumulated at this interest would grow to the face value at T.*

Remark 3.1.2. We can view zero-coupon bond as a derivative on the interest rate that its value today depends on its face value and the interest rate dynamics.

Definition 3.1.3 (price of zero-coupon bonds). *We denote $P(t,T)$ as the price at time t of a zero-coupon bond that matures at time T(pays $1). where we have some practical constraints and issues:*

1. *$t \leq T$(before maturity)*
2. *t is not necessarily the time the bond is issuing.*
3. *$P(t,t) = 1, P(t,T) \leq 1, \forall T \geq t$, assuming that the interest rate is always non-negative.*

Remark 3.1.3 (interpretation)**.** The bond price is a function of two parameters t and T.

Remark 3.1.4 (caution!)**.** [1, p. 5] Let $t < S < T$. It is tempting to use no-arbitrage argument to show

$$P(t,T) = P(t,S)P(S,T).$$

However, such equality will not hold since $P(S,T)$ is a stochastic quantity whereas $P(t,T)$ and $P(t,S)$ are deterministic quantities.

3.1.1.2 *Spot rates*

Definition 3.1.4 (simply compounded spot rate). *[1, p. 2]*

- *The **simply compounded spot rate** at time t for maturity T is defined as the annualized rate of return from holding a zero-coupon bond from time t until maturity T. It is denoted by $L(t,T)$.*
- *The simply compounded spot rate for tenor 1M, 3M, 6M, 12M usually can be directly observed from the market, e.g., annual rate for CDs.*

Proposition 3.1.1 (no-arbitrage relation between spot rate and zero-coupon bond). *Let current time be t. Assume the market is free ofarbitrage opportunities.*

- *At current time t, the (default-free) zero-coupon bond price then can be expressed in terms of the spot rate as*
$$P(t,T) = \frac{1}{1+(T-t)L(t,T)}.$$
Or equivalently,
$$L(t,T) = \frac{1-P(t,T)}{(T-t)P(t,T)}.$$
- *At future time S, no arbitrage condition requires that*
$$P(S,T) = \frac{1}{1+(T-S)L(S,T)}.$$
- *At future time S, no arbitrage condition requires that*
$$1+(T-S)L(S,T) = \frac{M(T)}{M(S)}.$$

Proof. (1) If zero-coupon bond is not priced this way, then we can long or short the zero-coupon bond at current time t to make arbitrages. (2)(3) If zero-coupon bond or money market account are not priced this way, then we can long or short the zero-coupon bond at time S to make arbitrages. □

Remark 3.1.5 (interpretation of equality for stochastic quantities/processes).

- Let current time be t. For future time $S > t$, we can view $L(S,T)$ and $P(S,T)$ and $M(S)/M(T)$ as stochastic process indexed by S.

- The arbitrage strategy is applied to each realization path, therefore, on each sample path $\omega \in \Omega$, we have

$$\frac{1}{1+(T-S)L(S,T)(\omega)} = P(S,T)(\omega) = \frac{M(T)(\omega)}{M(S)(\omega)}.$$

Definition 3.1.5 (continuously compounded spot rate, yield, zero curve). *Because of the no-arbitrage relationship between sport rate and zero-coupon bonds (Theorem 3.1.1), we can similarly define*

- *The* ***continuously compounded spot rate*** *is the annualized logarithmic rate of return from holding the bond from time t until maturity T. It is denoted by $R(t,T)$, and related to zero-coupon bond as*

$$R(t,T) = -\frac{\ln P(t,T)}{T-t}.$$

- *$R(t,T)$ is also called* ***yield to maturity****. The graph of $R(t,T)$ versus maturity T is known as the* ***yield curve or zero curve****.*
- *The zero-coupon bond price then can be expressed in terms of the spot rate as*

$$P(t,T) = \exp(-R(t,T)(T-t)).$$

Remark 3.1.6 (the parameter of yield curve). We usually denote yield curve by $Y(T) = R(t = 0, T)$. Note that in yield curve, we can only vary the maturity rather than the current time t. Yield curves can be directly observed from the market

Remark 3.1.7 (relations of simply and continuously compounded spot rate).

$$\lim_{T\to t} R(t,T) = \lim_{T\to t} L(t,T).$$

That is, a short holding period, the two spot rates are the same.

3.1.1.3 *LIBOR payment*

Proposition 3.1.2 (current no-arbitrage value of LIBOR payment). *[1, p. 6] Let $S < T$. Consider the future LIBOR-based payment $(T-S)L(S,T)$ at time T. Its arbitrage-free value at time $t < T$ is $P(t,S) - P(t,T), T > S$. That is*

$$P(t,S) = P(t,T) = E_Q[\frac{M(t)(T-S)L(S,T)}{M(T)}|\mathcal{F}_t].$$

Proof. (replicating method) We use the following strategy to replicate the payoff at time T.

- At time t, we buy an $S-$bond and sell a $T-$bond.
- At time S, the long position in S-bond matures to yield one dollar. Then use this income to buy an amount of $1/P(S,T)$ of $T-$bonds.
- At time T, the net position is $1/P(S-T)-1$, which equals $(T-S)L(S,T)$ based on the definition of $L(S,T)$

□

3.1.1.4 *Forward zero-coupon bonds*

Definition 3.1.6 (forward (zero-coupon) bond contract, forward (zero-coupon)bond price). *[1, p. 7] Let $t<S<T$. Let the current time be t.*

- *A **forward (zero-coupon) bond contract** is a contract between two parties to buy or sell a zero-coupon bond, which matures at time T, at a future time S with a price agreed upon at time t today.*
- *The **forward bond price** is the agreed price such that the contract has zero value.*
- *The forward bond price*
$$FB(t;S,T)=\frac{P(t,T)}{P(t,S)},$$
which is the price we agree on time t that we need to pay at time S to get the $T-$bond.

Lemma 3.1.1 (forward zero-coupon bond price). *Let $t<S<T$. Let the current time be t. In a no-arbitrage market, the forward zero-coupon bond is given by*

- $$FB(t;S,T)=P(t,T)(1+(S-t)L(t,S))$$
- $$FB(t;S,T)=\frac{P(t,T)}{P(t,S)}$$

Proof. (1) We can replicate the forward contract to deliver a $T-$maturity bond at time S in the following way:

- Borrow $P(t,T)$ at time t; buy zero-coupon bond at time t with price $P(t,T)$.
- At time S, we delivery the bond and get payment $FB(t;S,T)$. Also, we pay the loan with a total of $P(t,T)(1+(S-t)L(t,S))$.

Because our initial capital is zero, then our final payoff $+FB(t;S,T-P(t,T)(1+(S-t)L(t,S))$ should also be zero. (2) Use the relation between spot rate and zero-coupon bond (Theorem 3.1.1). □

Remark 3.1.8. Let A be the amount of money we pay at time S to get the $T-$bond, the case flow at time 0 is

$$V(0) = -AP(t,S) + P(t,T).$$

where $AP(t,S)$ is the present value of the future cash A at time S, and $P(t,T)$ is the present value of the future cash 1 at time T. Set $V(0) = 0$, we have

$$A = \frac{P(t,T)}{P(t,S)}.$$

For positive interest rate, we have $FP(t;S,T) \leq 1$.

Remark 3.1.9 (forward rate vs. forward bond price)**.**

- $$FP(t;S,T) = \frac{P(t,T)}{P(t,S)} = \frac{1}{1+(T-S)F(t;S,T)}.$$

  We can think of the forward rate as the simply compounded rate of return over the time interval $[S,T]$ implied by the forward bond price.
- Recall that the zero-coupon bond price can be expressed as

$$P(t,T) = \frac{1}{1+(T-t)L(t,T)}.$$

Lemma 3.1.2 (zero-coupon bond decomposition). *Let $T_0 < T_1 < \cdots < T_n$ be a set of dates. Then*

$$\begin{aligned} P(T_0,T_n) &= FP(T_0;T_0,T_1)FP(T_0;T_1,T_2)\cdots FP(T_0;T_{n-1},T_n) \\ &= \prod_{i=1}^{n} \frac{1}{1+(T_i-T_{i-1})F(T_0,T_{i-1},T_i)} \\ &= \prod_{i=1}^{n} \exp(-(T_i-T_{i-1})R(T_0;T_{i-1},T_i)) \end{aligned}$$

Proof. Note the definitions:

$$FP(t;S,T) = \frac{P(t,T)}{P(t,S)} = \exp(-R(t;S,T)(T-S)),$$

and

$$FP(t;S,T) = \frac{P(t,T)}{P(t,S)} = \frac{1}{1+(T-S)F(t;S,T)}.$$

□

3.1.1.5 *Forward rates*

Definition 3.1.7 (forward rate agreement). *[1, p. 5]*

- *Let $t < S < T$. A unit notional **forward rate agreement(FRA)** is a contract entered into at time t, when the issuer agrees to pay the holder at time T the LIBOR $L(S,T)$ in exchange for a fixed rate K applied to unit notional amount. The value of the payoff at time T is given by*
$$(T-S)(K-L(S,T)).$$
- *The value of K is selected such that the value of FRA at time t is zero. Such K is called **forward rate**, denoted by $F(t,S,T)$.*

Remark 3.1.10 (using forward rate agreement to lock future rates).

- Suppose that party A use a forward rate agreement to lock the future rate on period $[S,T]$ at current time t.
- As showed in Figure 3.1.1, party A can enter the forward rate agreement with party B at time t, and then invest in the money market at time S.
- At time T, party A can exchange with party B the cash flow to receive fixed rate payments.

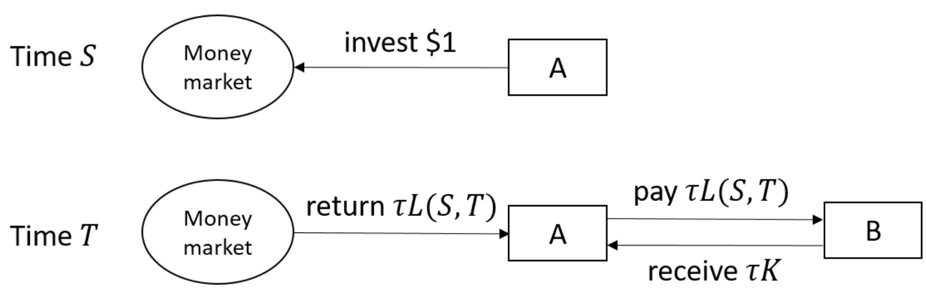


Figure 3.1.1: Using forward rate contract to lock future rate.

Remark 3.1.11 (interpretation, spot vs forward rate).

- Spot rate is the current market rate of return if we invest in money market for a period of time.
- Forward rate implied from forward rate agreement is future market rate of return if we invest in money market for a future period of time.

Proposition 3.1.3 (no-arbitrage forward rate). *Let $t < T_1 < T_2$. Let current time be t. In a no-arbitrage market, it is required that*

$$F(t, T_1, T_2) = \frac{1}{T_2 - T_1}\left(\frac{P(t, T_2)}{P(t, T_1)} - 1\right),$$

Further more, the time t value of LIBOR payment $(T_2 - T_1)L(T_1, T_2)$ is given by $(T_2 - T_1)F(t_1; T_1, T_2)P(t_1, T_2)$.

Proof. (1)(2)Note that the current value of a LIBOR payment via replication method (Theorem 3.1.2) is given by

$$P(t, T_1) - P(t, T_2),$$

which should equal the current value of the fixed rate payment, i.e.,

$$(T_2 - T_1)F(t, T_2, T_1)P(t, T_2).$$

□

Example 3.1.1. For two annually compounded spot rate $L(0, 1) = 4\%$ and $L(0, 2) = 5\%$, maturing in one and two years respectively, then the one-year to two-year forward rate $F(0; 1, 2)$ can be computed as follows:

$$F(0; 1, 2) = (P(0, 1)/P(0, 2) - 1)$$

where

$$P(0, 1)(1 + L(0, 1)) = 1, P(0, 2)(1 + 2L(0, 2)) = 1.$$

Proposition 3.1.4 (forward rate agreement value). *[1, p. 6] Consider a forward rate agreement(FRA), whose interest accrual period is $[S, T]$, has payoff at time T given by*

$$(T - S)(K - L(S, T)).$$

- ***(fair value)*** *It value at current time t is given as*

$$V(t) = P(t, T)(T - S)K - P(t, S) + P(t, T).$$

- ***(fair rate)*** *If we set*

$$K \triangleq F(t; S, T) = \frac{P(t, S) - P(t, T)}{(T - S)P(t, T)} = \frac{1}{T - S}\left(\frac{P(t, S)}{P(t, T)} - 1\right),$$

the FRA has value 0 at time t. Or equivalently,

$$\frac{P(t,T)}{P(t,S)} = \frac{1}{1+(T-S)F(t,S,T)}.$$

Note that by definition $F(S;S,T) = L(S,T)$.

- ***(value evolution)*** *Consider a forward contract entered at t_0 with fair rate K. At future time $t_1, t_0 < t_1 < T$, the value of the forward contract is given by*

$$V(t_1) = P(t_1,T)\tau(F(t_0;S,T) - F(t_1;S,T)),$$

where $F(t_1;S,T)$ is the fair forward rate at t_1, $\tau = T - S$. ***Note that** $P(t_1,T)$**,** $F(t_1;S,T)$ **and** $V(t_1)$ **are all random quantities at time** t_0**.***

Proof. (1)(2)Use the current value of LIBOR payment formula Theorem 3.1.2. (3) (a) (replication method) Because time t_1 of the floating payment $(T-S)F(t_1;S,T)P(t_1,T)$

(b) (martingale method) Note that at any time $t < T$, the fair value forward rate is related to the LIBOR payment via

$$\begin{aligned} &V(t) = P(t,T)E_T[\tau F(t;S,T) - \tau L(S,T)|\mathcal{F}_t] = 0 \\ &\implies P(t,T)\tau F(t;S,T) = P(t,T)E_T[\tau L(S,T)|\mathcal{F}_t] = PV_t[\tau L(S,T)] \end{aligned}$$

where $PV_t[\tau L(S,T)]$ denotes the t-value of the future T-LIBOR payment $\tau L(S,T)$.

For a contract entered at time t_0, the fixed payment has value $P(t_1,T)\tau F(t_0;S,T)$, the LIBOR payment has value

$$PV_{t_1}[\tau L(S,T)] = P(t_1,T)E_{t_1}[\tau L(S,T)|\mathcal{F}_{t_1}] = P(t_1,T)\tau F(t_1;S,T).$$

Therefore, the value of the forward contract is given by

$$V(t_1) = P(t_1,T)\tau F(t_0;S,T) - P(t_1,T)\tau F(t_1;S,T).$$

□

Definition 3.1.8 (continuously compounded forward rate). *The continuously compounded forward rate at time t, which applies between future times T and $S(t \leq S < T)$ is defined as*

$$R(t;S,T) \triangleq -\frac{1}{T-S}\log\frac{P(t,T)}{P(t,S)} = -\frac{1}{T-S}\log(FP(t;S,T)),$$

such that

$$FP(t;S,T) = \frac{P(t,T)}{P(t,S)} = \exp(-R(t;S,T)(T-S)).$$

Lemma 3.1.3 (no-arbitrage relationship for forward rates decomposition). *Let $t \leq T_1 < T_2 < T_3$. Then in a no-arbitrage market, it is required that*

$$(T_3 - T_1)F(t; T_1, T_3) + 1 = [(T_2 - T_1)F(t; T_1, T_2) + 1][(T_3 - T_2)F(t; T_2, T_3) + 1].$$

Proof. A forward contract locking the rate in period $[T_1, T_3]$ can be exactly replicated by two forward contracts locking the rate in the period $[T_1, T_2]$ and $[T_2, T_3]$. □

3.1.2 Instantaneous rates

Definition 3.1.9 (instantaneous forward rate).

- *The instantaneous forward rate at time t is*

$$f(t, T) \triangleq \lim_{S \to T} R(t; T, S) = \lim_{S \to T} F(t; T, S) = -\frac{\partial}{\partial T} \log P(t, T)$$

- *We also have (integration based on (1))*

$$P(t, T) = \exp(-\int_t^T f(t, u)du).$$

The dependence of $f(t, T)$ on the maturity T is known as the ***term structure of forward rate curve*** *or* ***forward rate curve***[a] *at time t.*

[a] Note that forward rate curve and forward curve usually refers to different curves. Forward curve is associated with a tenor.

Definition 3.1.10 (short rate, risk-free rate). *The* ***short rate*** *is defined as*

$$r(t) = \lim_{T \to t} R(t, T) = \lim_{T \to t} L(t, T) = R(t, t) = f(t, t).$$

Short rate is what we usually refer to as risk-free rate.

Note 3.1.1 (Caution! non-constant short rate cannot determine zero-coupon bond price). Note that

$$\begin{aligned} P(t,T) &= \exp(-R(t,T)(T-t)) = \exp(-\int_t^T f(t,s)ds) \\ &= E_Q[\exp(-\int_t^T r(s)ds)|\mathcal{F}_t] \neq \exp(-\int_t^T r(s)ds) \end{aligned}$$

Note that $\exp(-\int_t^T f(t,s)ds)$ is a scalar whereas $\exp(-\int_t^T r(s)ds)$ is a random variable.

However, under risk-neutral measure, we have

$$P(t,T) = g(t,r(t),T) = E_Q[\exp(-\int_t^T r(s)ds)|\mathcal{F}_t].$$

Lemma 3.1.4 (relationship between yield, zero curve and forward curve). *For a given current time t, each of the curves $P(t,T), f(t,T), R(t,T), F(t,T_1,T_2)$ have the following relationships:*

- $$P(t,T) = \exp(-R(t,T)(T-t)) = \exp(-\int_t^T f(t,s)ds).$$
- $$F(t,T_1,T_2) = \frac{1}{T_2-T_1}(\frac{P(t,T_2)}{P(t,T_1)} - 1) = \frac{1}{T_2-T_1}(\exp(\int_{T_1}^{T_2} f(t,s)ds) - 1).$$

Proof. (1) Based on the definition of yield, zero curve and instantaneous forward rate. (2) Note that from Theorem 3.1.4, we have

$$F(t,T_1,T_2) = \frac{1}{T_2-T_1}(\frac{P(t,T_2)}{P(t,T_1)} - 1).$$

□

Definition 3.1.11 (cash account, money market account). *The **cash account** is a special security that earns short rate interest, we denote the value by $B(t)$ with $B(0) = 1$. In continuous time model, we have*

$$B(t) = \int_0^t e^{r(s)}ds$$

In discrete time model, we have

$$B(n) = \prod_{i=1}^{n}(1 + r_i)$$

Remark 3.1.12 (interpretation)**.**

- Note that $B(t)$ is a **random variable** because the short rate $r(t)$ is random.
- The cash account/money market account can be thought of as the amount earned by starting with a unit dollar at time 0 and continuously reinvesting it at the short rate over infinitesimal time interval $[t, t + \delta t]$

Remark 3.1.13 (short rate contains less information than zero-coupon curve)**.**

- Short rate alone cannot determine the price of the zero-coupon bond.
- Money account and zero-coupon bond price have non-trivial relationship.

3.1.3 Coupon-bearing bonds

Definition 3.1.12 (fixed-coupon bond). *[1, p. 10] A **fixed-coupon bond** is a financial instrument that pays the holder deterministic(known at time $t \leq T_0$) amounts $c_1, ..., c_n$, referred to as **coupon rate**, at times $T_1, ..., T_n, T_0 < T_1 < \cdots < T_n$. At maturity time T_n, the holder receives the face value N and the final coupon c_n.*

Lemma 3.1.5 (present value of fixed coupon bonds). *The present value for a coupon-bearing bond is given by*

$$B_{fixed}(t) = \sum_{i=1}^{n} c_i P(t, T_i) + NP(t, T_n).$$

where c_i is the coupon rate and N is the face value.

Proof. Note that the present no-arbitrage price for a fixed payment c_i at time T_i is $c_i P(t, T_i)$. $\square$

Definition 3.1.13 (floating rate note). *[1, p. 10] Let $0 \leq T_0 < T_1 < \cdots < T_n$ be a set of dates. A **floating rate note** is a financial instrument that pays the holder stochastic(unknown at time $t \leq T_0$) amounts $\tau_1 L(T_0, T_1), ..., \tau_n L(T_{n-1}, T_n) \times N$ at times $T_1, ..., T_n, T_0 < T_1 < \cdots < T_n$. At maturity time T_n, the holder receives the face value N and the final coupon $\tau_n L(T_{n-1}, T_n) \times N$.*

Remark 3.1.14 (difference between fixed coupon bonds and floating rate note)**.** At time $t \leq T_0$, the coupon payment at future times is known for coupon-bearing bonds, whereas unknown for floating rate note(since $L(T_{i-1}, T_i)$ is unknown at t).

Lemma 3.1.6 (valuation of a floating-rate note). *Let $0 \leq T_0 < T_1 < \cdots < T_n$ be a set of dates. Consider a flating rate note that pays float coupon $\tau_i L(T_{i-1}, T_i) \times N$ at $T_i, i = 1, 2, ..., n$ and pays face value N at T_n. Let current time be $t_0, t_0 \leq T_0$. It follows that*

- *The present value for a floating-rate note is given by*

$$B_{floating}(t_0) = NP(t, T_0).$$

  Particularly if $t_0 = T_0$, then

$$B_{floating}(T_0) = N.$$

- *(value evolution)* ***Assume interest rate curve remain constant.***
  - *Consider a future time t. Immediately after coupon payment date $T_1, T_2, ..., T_{n-1}$, the value of the note is given by*

$$B_{floating}(T_i) = N.$$

  - *For $T_{i-1} < t < T_i$, the value is*

$$B_{floating}(t) = N \times DF(t, T_i)(1 + \tau_i F(t_0, T_{i-1}, T_i)).$$

Proof. (1)Use the result that the present value of $\tau L(T_{i-1}, T_i) = P(t, T_{i-1}) - P(t, T_i)$.

$$B_{floating}(t) = N \sum_{i=1}^{n} (P(t, T_{i-1}) - P(t, T_i)) + NP(t, T_n) = NP(t, T_0)$$

(2) Straight forward. □

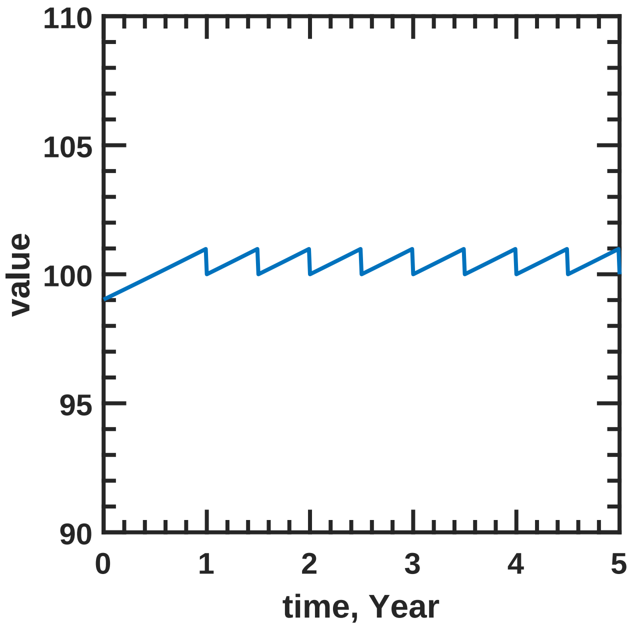


Figure 3.1.2: Value evoluation of the floating rate note with maturity 5Y. Coupon payment dates are 1Y, 1.5Y, ..., 5Y. The interest rate curve is flat with effective annual rate of 2%.

3.1.4 Interest rate swap

3.1.4.1 *The swap contract*

Definition 3.1.14 (interest rate swap). *[1, p. 12]*

- *An* ***interest rate swap*** *is an OTC instrument in which two counterparties exchange a set of payments at a fixed rate of interest for a set of payments at a floating rate, typically the LIBOR.*
- *If the holder is paying a floating rate and receiving the fixed rate, the swap is said to be a* ***receiver swap****. If the holder is receiving a floating rate and paying the fixed rate, the swap is said to be a* ***payer swap****.*
- *Consider a unit notional amount* $N = 1$ *and a set of dates* $T_0 < T_1 < ... < T_n$ *with* $\tau_i = T_i - T_{i-1}, i = 1, ..., n$*. At time* T_i*, the holder of a payer swap pays a fixed amount* $\tau_i K$*, where* K *is the* ***swap rate****, in exchange for a floating payment of* $\tau_i L(T_{i-1}, T_i)$*, where* $L(T_{i-1}, T_i)$ *is the spot LIBOR rate(a random quantity) in future time interval* T_{i-1} *and* T_i*.*
- *The payment dates* $T_1, ..., T_n$ *are* ***settlement dates****, and dates* $T_0, ..., T_{n-1}$ *are* ***reset dates.*** *The first reset data* T_0 *is called the* ***start date*** *of the swap.*
- *If current time* $t < T_0$*, the swap agreement is called a* ***forward-starting*** *swap; If current time* $t = T_0$*, the swap agreement is called* ***spot-starting*** *swap.*

Example 3.1.2 (application of forward starting swap). For example(see Figure 3.1.3a), company B takes a loan for $400 million at a fixed interest rate and company A takes a loan for $400 million at a floating interest rate. Company B expects that the rate six months in future will decline and therefore wants to convert its fixed rate into a floating one. But because the rate changes are not expected to happen right away, the company just wants to lock in the swap rate for later. On the other hand, company A believes that interest rates will increase six months in the future. It does not want to convert into a fixed rate loan right away but wants to protect itself by locking in the rate now. The two companies may enter into forward starting interest rate swap to hedge their risk.

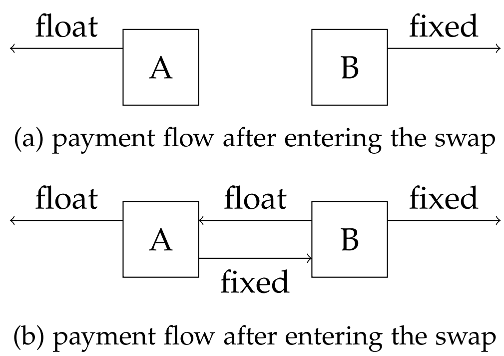


Figure 3.1.3: Use swap to exchange payment flow

Remark 3.1.15 (application of swaps in debt issuance). During a public offering of corporate bonds to the public, investors usually prefers floating rate bonds, whereas corporates prefer issuing fixed rate bonds. The corporate will enter a swap with a bank, in which the corporate will pay fixed coupon to the bond and pay investors the float rate coupon received from the bank.

Remark 3.1.16 (application of swaps). link

- **Portfolio management.** Interest rate swaps allow portfolio managers to adjust interest rate exposure to the yield curve shape change (parallel move, twisting, steepening, etc, but not the implied volatility of yield curves). By increasing or decreasing interest rate exposure in various parts of the yield curve using swaps, managers can either ramp-up or neutralize their exposure to changes in the shape of the curve, and can also express views on credit spreads. Swaps can also act as substitutes for other, less liquid fixed income instruments. Moreover, long-dated interest rate swaps can increase the duration of a portfolio, making them an effective tool in Liability Driven Investing, where managers aim to match the duration of assets with that of long-term liabilities.
- **Speculation.** Because swaps require little capital up front, they give fixed income traders a way to speculate on movements in interest rates while potentially avoiding the cost of

long and short positions in Treasuries. For example, to speculate that five-year rates will fall using cash in the Treasury market, a trader must invest cash or borrowed capital to buy a five-year Treasury note. Instead, the trader could 'receive' fixed in a five-year swap transaction, which offers a similar speculative bet on falling rates, but does not require significant capital up front.

- **Corporate debt issuance.** During a public offering of corporate/municipal bonds to the public, investors usually prefers floating rate bonds, whereas corporates/municipalities prefer issuing fixed rate bonds. The corporate/municipalities will enter a swap with a bank, in which the corporate/municipalities will pay fixed coupon to the bond and pay investors the float rate coupon received from the bank.
- **Corporate finance.** Firms with floating rate liabilities, such as loans linked to LIBOR, can enter into swaps where they pay fixed and receive floating, as noted earlier. Companies might also set up swaps to pay floating and receive fixed as a hedge against falling interest rates. Companies might also set up swaps pay floating and receive fixed to reduce interest rate exposure if floating rates more closely match their assets or income stream.
- **Risk management.** Banks and other financial institutions are involved in a huge number of transactions involving loans, derivatives contracts and other investments. The bulk of fixed and floating interest rate exposures typically cancel each other out, but any remaining interest rate risk can be offset with interest rate swaps.
- **Rate-locks on bond issuance.** When corporations decide to issue fixed-rate bonds, they usually lock in the current interest rate by entering into swap contracts. That gives them time to go out and find investors for the bonds. Once they actually sell the bonds, they exit the swap contracts. If rates have gone up since the decision to sell bonds, the swap contracts will be worth more, offsetting the increased financing cost.

Remark 3.1.17 (relation to coupon-bearing bonds). An interest rate swap can be viewed as longing a fixed coupon bond and shorting a floating-rate note.

3.1.4.2 *Valuation of a swap and swap rate*

Lemma 3.1.7 (value of an interest rate swap). *Let the current time be $t \leq T_0$. consider a swap with unit notional.*

- *The current value of the fixed leg is*

$$V_{fixed}(t) = K \sum_{i=1}^{n} \tau_i P(t, T_i).$$

- *The current value of the floating leg is*

$$V_{float}(t) = \sum_{i=1}^{n}(P(t,T_{i-1}) - P(t,T_i)) = P(t,T_0) - P(t,T_n).$$

Particularly for spot-starting swaps such that $t = T_0$,

$$V_{float}(t) = 1 - P(t,T_n).$$

- *The current value of a payer swap contract is*

$$V(t,K) = V_{float}(t) - V_{fixed}(t) = P(t,T_0) - P(t,T_n) - K\sum_{i=1}^{n}\tau_i P(t,T_i).$$

Proof. (1) Straight forward from cash flow discounting. (2) Note that from Theorem 3.1.2, we have $\tau_i L(T_{i-1},T_i) = P(t,T_{i-1}) - P(t,T_i)$. For future payment $K\tau_i$, its present value at time t is $K\tau_i P(t,T_i)$. □

Definition 3.1.15 (forward swap rate). *[1, p. 13] Let the current time be t. The **forward swap rate**, denote by $S(t;T_0,T_n)$, associated with a swap exchanging cash flows with specified frequency from T_0 to T_n ($t \leq T_0$), is given by*

$$S(t;T_0,T_n) = \frac{P(t,T_0) - P(t,T_n)}{\sum_{i=1}^{n}\tau_i P(t,T_i)},$$

such that the time t value of the forward swap zero.

Note that

- *The quantity $\Delta = T_n - T_0$ is known as **swap tenor**.*
- *The forward swap rate $S(t;T_0,T_n)$ is a known number since we know the current term structure $P(t,T), T \in \mathbb{R}^+$. For any future time $s > t$, $S(s;T_0,\Delta)$ is an unknown random variable since $P(s,T), T \in \mathbb{R}^+$ is random.*

Remark 3.1.18 (application of forward swap rate in calibration). In reality, swap rates are traded benchmark securities. Since we have expressed the swap rate in terms of the zero-coupon bond prices, we can determine the zero-coupon curve from market swap rates.

Lemma 3.1.8 (swap rate as the sum of weighted forward rate). *Consider a swap with underlying tenor $\Delta = T_n - T_0$. Let the current time be t.*

- *The forward swap rate given by*

$$S(t;T_0,T_n)(t) = \frac{P(t,T_0)-P(t,T_n)}{\sum_{i=1}^n \tau_i P(t,T_i)},$$

can also be written as

$$S(t;T_0,T_n) = \sum_{i=1}^n w_i(t)F(t;T_{i-1},T_i)$$

where

$$w_i(t) = \frac{\tau_i P(t,T_i)}{\sum_{j=1}^n P(t,T_j)}.$$

- *If this is a* ***continuous swap*** *where the exchange of cash flow occurs within* $[T_0,T_1]$*. Then*

$$S_c(t;T_0,T_n) = \frac{\int_{T_0}^{T_n} f(t,s)P(t,s)ds}{\int_{T_0}^{T_n} P(t,s)ds},$$

where $f(t,s)$ *is the instantaneous forward rate observed on current time* t*.*

Proof. (1) Note that the present value of a interest rate swap is given by

$$\begin{aligned} V(t,K) &= \sum_{i=1}^n PV(\tau_i L(T_{i-1},T_i)) - K\sum_{i=1}^n \tau_i P(t,T_i) \\ &= \sum_{i=1}^n \tau_i F(t;T_{i-1},T_i)P(t,T_i) - K\sum_{i=1}^n \tau_i P(t,T_i), \end{aligned}$$

where we use the result in Theorem 3.1.2. Set the present value to zero and we also get the result. (2) For a continuous swap, the fixed leg has value

$$K\int_{T_0}^{T_n} P(t,s)ds,$$

and the floating leg has

$$\int_{T_0}^{T_n} f(t,s)P(t,s)ds,$$

where we can view the floating leg is paying $f(t,s)$ in every time point. □

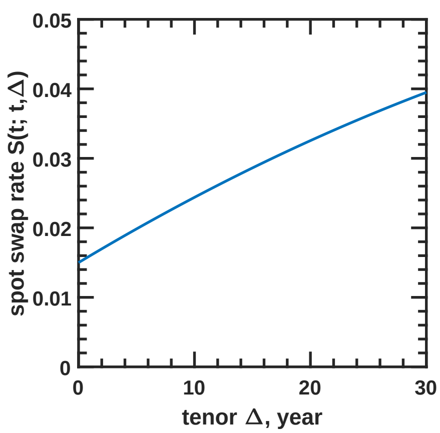
Lemma 3.1.9 (mark-to-market value of a swap contract). *Consider an interest rate payer swap starting at* T_0 *with tenor* Δ *entering at* $t_0 \leq T_0$*.*

- *During future time* $t, t_0 \leq t \leq T_0$*, the value of the swap contract* $V(t)$

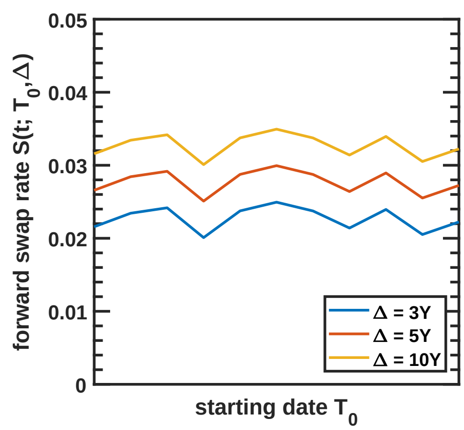
3.1.4.3 *Understand swap curves*

Note 3.1.2 (different swap rate curve). The swap rate $S(t; T_0, T_N), \Delta = T_N - T_0$ is characterized by three parameters. Varying these parameters can yield different curves, as showed in Figure 3.1.4.

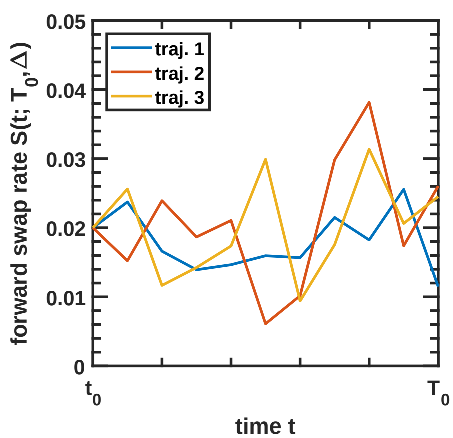
- Let the current time be t. The "**spot swap curve**" refers to an x-y chart of par swap rates $S(t; t, t+\Delta)$ plotted against their tenor Δ.
- If we fix current time t and tenor Δ, we can get a swap rate $S(t; T_0, T_0 + \Delta)$ vs. starting date T_0. This curve is like the forward curve of libor rates that can be observed at the current market.
- Consider a swap contract with fixed starting date T_0 and tenor Δ. Let current time be t_0. For any time $t_0 \leq t \leq T_0$, $S(t; T_0, T_0 + \Delta)$ is an unknown random quantity.



(a) Spot swap rate vs. tenor.



(b) Forward swap rate vs tenor and forward starting date T_0.



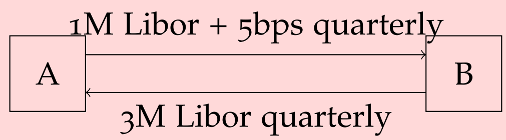
(c) The evolution of forward swap rate $S(t; T_0, \Delta)$ in different realizations.

Figure 3.1.4: Swap rate concept demo.

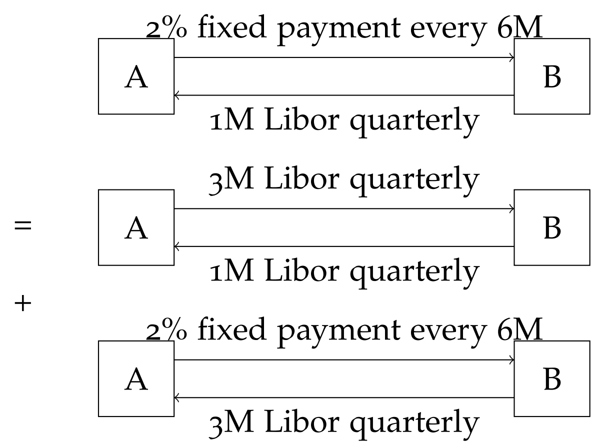
3.1.4.4 *Basis swap*

Definition 3.1.16 (basis swap). *[2, p. 28] A basis swap is an interest rate swap consisting of two floating legs, which are tied to different indices. To make a swap have zero value at starting date, a spread is added to one of the floating leg.*

Example 3.1.3. Consider a 2-year basis swap in the following figure that one counterparty pays 1-month LIBOR plus 5 basis point, and the other pays 3-month LIBOR. The two party will exchange the cash flow quarterly money for two years.



Remark 3.1.19 (use basis swap to construct other fixed-float swaps). We can use basis swap to construct new fixed-float swaps that are not available on the market. As we show in the following figure, we can use a 3M-1M basis swap and a regular 3M-fixed swap to construct a 1M-fixed swap.



3.1.5 Other common swaps and indices

3.1.5.1 *CMS and CMS index*

Definition 3.1.17 (constant-maturity swap, CMS). *[1, p. 133][3, p. 557] Take a unit notional amount and a set of dates $t < T_0 < T_1 < \cdots < T_n < \cdots < T_{m+n}$ with accrual periods $\tau_i = T_i - T_{i-1}$. A **constant-maturity swap(CMS)** involves a series of payments of an amount*

$\tau_{i+1}S_{i,i+m}(T_i)$, where $S_{i,i+m}(T_i)$ is the forward swap rate setting at at T_i and maturing at T_{i+m}, made in exchange for a fixed amount $\tau_{i+1}K$. These payments occur at times $T_{i+1}, i = 0, ..., n-1$.

Remark 3.1.20 (CMS swap vs. ordinary interest rate swap).

- In the ordinary interest rate swap, the future payment for a period is a quantity related to LIBOR rate $L(T_i, T_{i+1})$(which is a random quantity, but its expected value under the T_{i+1} forward measure is $F(t, T_i, T_{i+1})$).
- In the CMS swap,the future payment for a period is a quantity related to the swap rate of a constant tenor swap(which is a random quantity, but its expected value under the annuity measure is the current swap rate.)
- A CMS swap is like a swap on swap rate, while an ordinary interest rate swap is a swap on LIBOR interest rate.

Remark 3.1.21 (comparison between cap, swap, and CMS swap in hedging). Cap and swaps are usualy indexed to LIBOR 3M, LIBOR 1M, BMA etc., i.e., short-term interest rates.

- A cap with long life (e.g., the cap consisting of 40 caplets with a total coverage of 10 Year) can hedge long-term exposure to short-term interest rate risk. A 10Y cap indexed to LIBOR 3M can hedge 3M interest rate risk for 10 years.
- Similarly, a swap with long life can also hedge long-term exposure to short-term interst rate risk. A 10Y swap indexed to LIBOR 3M can hedge 3M interest rate risk for 10 years.
- A N-year-life CMS swap indexed to M-year can hedge N year exposure to M-year interest rate risk.

3.1.5.2 *BMA swap and BMA index*

Definition 3.1.18 (BMA swap and BMA index).

- *A Bond Market Association (BMA) Swap is a type of swap arrangement in which two parties agree to exchange interest rates on debt obligations, where the floating rate is based on the U.S. SIFMA Municipal Swap Index. One of the parties involved will swap a fixed interest rate for a floating rate, while the other party will swap a floating rate for a fixed rate.*
- *US BOND MARKET ASSOCIATION MUNICIPAL SWAP INDEX (BMA INDEX) This index is produced weekly, reflecting the average rate of issues of tax-exempt variable-rate debt, and serves as a benchmark floating rate in municipal swap transactions. The BMA index is usually 65%-70% of its taxable equivalent, the 1-month LIBOR.*

3.1.6 Negative interest rates

Remark 3.1.22 (why central bank set interest rate). link

- Central banks in many developed countries have set key interest rates below zero, and as a result, portions of the yield curves in these countries have dropped to negative levels.
- Although individuals are not paying banks to hold their money, negative interest rates imposed by a central bank effectively mean commercial banks are required to pay for holding excess reserves with the central bank. For example, if the deposit rate were –1%, for every $10 million held with the central bank, the commercial bank would have a balance of around $9.9 million at the end of a year.
- In general, the purpose of negative interest rate policy is to boost the economy. The theory is that commercial banks will be dissuaded from maintaining large balances with the central bank and will instead lend money to businesses and consumers who will, in turn, spend the money. The increase in lending and spending is likely to boost economic activity, leading to growth and inflation.
- For Switzerland, Denmark and Sweden, the rationale for lowering policy rates below zero had more to do with their currencies and the associated exchange rates than credit creation. The objective was to put downward pressure on the currency in order to stimulate trade by making exports cheaper and imports more expensive.

Remark 3.1.23 (why investors invest in negative yield bonds).

- **Currency speculation**: An investor may make an investment into a negative-yield bond as a proxy for a currency investment if they think the currency they will be repaid in will appreciate in value in an amount that exceeds the negative yield of the bond.
- **Speculation rolling yield:** Even though the yield curve is partial negative, or completely negative, as long as the yield curve has upward sloping shape, the investors can generate positive gain by buy long-dated bonds, and sell them before maturity.
- **Speculating further decreasing yield:** Investors can also generate positive return by speculating when the yield further decreases.
- **Regulatory Requirement:** Some investors have to compulsorily hold Govt bonds. These would be banks, central banks, insurance cos, etc. They do not have an option. They are required to purchase certain types of relatively safe assets, regardless of their yield.
- **Risk-averse institutional investosr:** Then there are investors who think the world will fall apart and feel the safest to keep their money with the Govt. They are ok to lose 0.5% of their capital, then to lose a lot more elsewhere. Consider an institution that has to be risk averse - when they have a lot of money they cannot stuff cash under a mattress and have rats eat it, nor you trust a bank deposit enough (banks default) and other bonds are too risky.

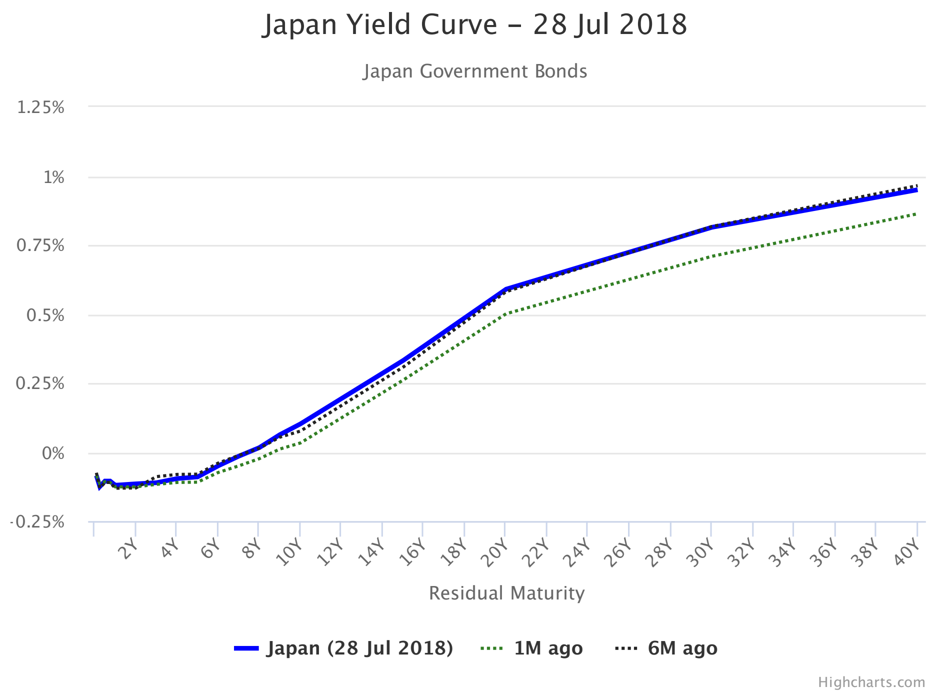


Figure 3.1.5: Japan government bond yield curve as of 2018-7-28. Data source: http://www.worldgovernmentbonds.com/country/japan/

3.2 Common interest rate derivatives

3.2.1 Options on bonds

Definition 3.2.1 (vanilla call option on zero-coupon bond). *A call option is the right to buy a bond at time S which matures at $T > S$. The call option has strike K and expiry $S > t$. The payoff of the call option at expiry S is*

$$\max(P(S,T) - K, 0).$$

Remark 3.2.1. The goal is to determine the price of the call option at $t < S < T$.

3.2.2 Caps and floors

Definition 3.2.2 (cap). *[1, p. 33] A **cap** is a portfolio of n call options on LIBOR. Consider a unit notational amount and a set of dates $T_0 < T_1 < \cdots < T_n$ with $\tau_i = T_i - T_{i-1}, i = 1, 2, ..., n$. At* *time T_i, the holder of an interest rate cap receives $\tau_i \max(L(T_{i-1}, T_i) - K, 0)$, where K is the* ***cap rate****. Each of these n call options are known as* ***caplets****.*

The ith caplet is a European call option with expiry T_i written on the spot LIBOR rate $L(T_{i-1}, T_i)$(a random quantity) with payoff

$$(T_i - T_{i-1}) \max(L(T_{i-1}, T_i) - K, 0).$$

Definition 3.2.3 (floor). *A* ***floor*** *is a portfolio of n put options on LIBOR. Consider a unit notational amount and a set of dates $T_0 < T_1 < \cdots < T_n$ with $\tau_i = T_i - T_{i-1}, i = 1, 2, ..., n$. At time T_i, the holder of an interest rate floor receives $\tau_i \max(K - L(T_{i-1}, T_i), 0)$, where K is the* ***floor rate****. Each of these n put options are known as* ***floorlets****.*

The ith floorlet is a European call option with expiry T_i written on the spot LIBOR rate $L(T_{i-1}, T_i)$ with payoff

$$(T_i - T_{i-1}) \max(K - L(T_{i-1}, T_i), 0).$$

Remark 3.2.2 (interpretation applications of caps and floors). [2, p. 135]

- A caplet is a call on rates. Its payoff on a 3M LIBOR is given by

$$(L(T, T + 3M) - K)^+.$$

  The buyer of a cap is protecting against rates moving higher.
- A floorlet is a put on rates. Its payoff on a 3M LIBOR is given by

$$(K - L(T, T + 3M))^+.$$

  The buyer of a floorlet is protecting against rates moving higher.
- Caps can be used by borrowers to hedge against increasing interest rate; floors can be used by lenders to hedge against decreasing interest rate.
- For example, a borrower who is paying the LIBOR rate on a loan can protect himself against a rise in rates by buying a cap at 2.5%. If the interest rate exceeds 2.5%, the extra money used to repay the loan is compensated by the increasing value of the cap.Therefore, the interest payments are effectively "capped" at 2.5% from the borrowers' point of view.
- Mortgage servicers are exposed to prepayment risk when interest rate is decreasing. Therefore, mortgage servicer can purchase floors on 10Y CMS to hedge the prepayment risk.

3.2.3 Swaption

3.2.3.1 *Basic concepts*

Definition 3.2.4 (swaption). *[1, p. 35] Consider an interest rate swap with payment dates $T_1, T_2, ..., T_n$ and reset dates $T_0, T_1, ..., T_{n-1}$.*

- *A **payer(receiver) swaption on spot-starting swaps** gives the holder the right to enter into an interest rate swap with maturity date $T_s = T_0$. Note that T_0 is the first reset date of the underlying swap. T_0 is also the **expiry date** of the swaption.*
- *A **payer(receiver) swaption on forward-starting swaps** gives the holder the right to enter into an interest rate swap with maturity date $T_s < T_0$. Note that T_0 is the first reset date of the underlying swap. T_0 is also the **expiry date** of the swaption.*

Remark 3.2.3 (summary and interpretation)**.**

- A payer swaption is an option to enter into a payer swap (the swap is paying fixed rate) at a later date.
- A receiver swaption is an option to enter into a receiver swap (the swap is receiving fixed rate) at a later date.
- A payer swaption is like a call option on forward swap rate.
- A receiver swaption is like a put option on forward swap rate.

Remark 3.2.4 (business application of swaption)**.** Swaptions are used to hedge issuance of bonds or to hedge call features in bonds

Remark 3.2.5 (comparison between swaptions and cap/floor)**.** [2, p. 178]

- A buyer of a cap is buying a series of options. For example, the buyer of a 5-year cap on 3-month LIBOR is buying a series of 20 different call options.
- Instead, a buyer of a payer swaption, on the other hand, is buying a single option. For example, a 5 into 10 payer swaption will protect the buyer from an increase in 10-year swap rates, five years from now.
- The choice of which option is more appropriate depends on the needs of the particular investor. If, for example, the investor manages a bank portfolio that has a liability tied to 3-month LIBOR resetting quarterly over the next five years, then it would seem more appropriate for the bank to purchase a 5-year cap on 3-month LIBOR instead of purchasing a swaption. If, however, the investor manages a mortgage portfolio that has exposure to 10-year swap rates increasing, say, five years from now, then it would appear more appropriate for the investor to buy a 5 into 10 payer swaption rather than a cap. (Note that the market does not have call options on 10 year spot rate.)

Definition 3.2.5 (Bermudan interest rate swaption). *[1, p. 59] Consider a unit notional amount and a set of dates $0 < T_0 < T_1 < ... < T_n$. The holder of **a payer (or receiver) Bermudan swaption** with strike K has the right to enter a payer (or receiver) interest rate swap at* *any time T_k for $k = 0, ..., l, l < n$. The underlying swap has reset dates $T_0, ..., T_{n-1}$ and settlement dates $T_1, ..., T_n$ and fair swap rate K.*

3.2.3.2 *Swaption market basics*

3.2.4 Cancellable swap

Definition 3.2.6 (Cancellable swap).

- *An European cancelable payer swap is a payer swap in which the fixed rate payer has the right to terminate the trade at a given date during the swap period prior to its maturity. A Bermudian cancelable payer swap is a payer swap in which the fixed rate payer has the right to terminate the trade at a set prespecified dates prior to its maturity.*
- *More formally, let $T_S = T_0 < T_1, ..., T_n < T_E$ be a set of dates for resetting and coupon payments. A cancellable swap is paying a fixed rate K and receiving a float rate over period T_S and T_E and has cancellable date $T_{C,i}, T_S \leq T_{C_i} \leq T_E, i = 1, 2, ..., K$.*
-

Example 3.2.1. [2, p. 233] Consider a 5-year swap in which the fixed rate payer has the right to cancel the swap after one year. If the fixed rate payer does cancel the swap after one year, then both counterparties will make final coupon payments at the end of the first year, and no further cash flows will be made in the swap. If the fixed rate payer does not cancel the trade at the end of the first year, then the swap will remain outstanding for the remaining four years.

Lemma 3.2.1 (static replication of an European cancellable swaps).

- *Consider an European cancelable swap with start date, end date, cancelable date and fixed rate given by T_S, T_E, T_C, K. Its payoff at cancellation date T_C is*
$$\max(V(T_C, T_C, T_E, K), 0),$$
where $V(T_C, T_C, T_E, K)$ is the time-T_C value of a payer swap, with strike K, starting at T_C and ending at T_E.
- *A long position in cancellable swap can be replicated by a long position payer swap and a long position of receiver European swapation.*
Particularly, at the cancellation date T_C, the payoff of the cancellable swap is
$$\max(V(T_C, T_C, T_E, K), 0),$$

and the replicating portfolio has payoff

$$\underbrace{V(T_C, T_C, T_E, K)}_{swap\ value} + \underbrace{\max(-V(T_C, T_C, T_E, K), 0)}_{receiver\ swaption} = \max(V(T_C, T_C, T_E, K), 0).$$

3.2.5 Eurodollar futures

Definition 3.2.7 (Eurodollar futures). *[4, p. 171][5, p. 141]CMEgroup*

- *Eurodollar futures are cash-settled futures contracts with final futures price based on three-month(or one-month) LIBOR at the expiration date:*

$$P_{futures}(T_0; T_0, T_1) = 100(1 - L(T_0, T_1)),$$

where $L(T_0, T_1)$ is the Libor rate at T_0 for tenor $T_1 - T_0$(usually $T_1 = T_0 + 0.25$ or $T_1 = T_0 + 1/12$, hence the tenor is 3-month or 1-month).See Table 3.2.1.

- *(**quoting convention**) At $t < T_0$, the futures price is quoted as*

$$P_{futures}(t; T_0, T_1) = 100(1 - Fur(t, T_0, T_1))$$

where $Fur(t, T_0, T_1)$ is the futures rate. Therefore, the implied forward interest rate is

$$Fur(t, T_0, T_1) = 1 - \frac{P_{futures}}{100}.$$

- *For example, a price of 95.00 implies an 3M forward libor interest rate of 5%.*
- *A Eurodollar futures contract usually has notational amount 250000 such that the change of one basis point in Fur will lead to change of 25$.*

Table 3.2.1: 3-month Eurodollar futures contract schedule

contract month	**delivery/maturity/expiry date**
Mar 2018	19-Mar-2018
Apr 2018	16-Apr-2018
May 2018	14-May-2018
...	...

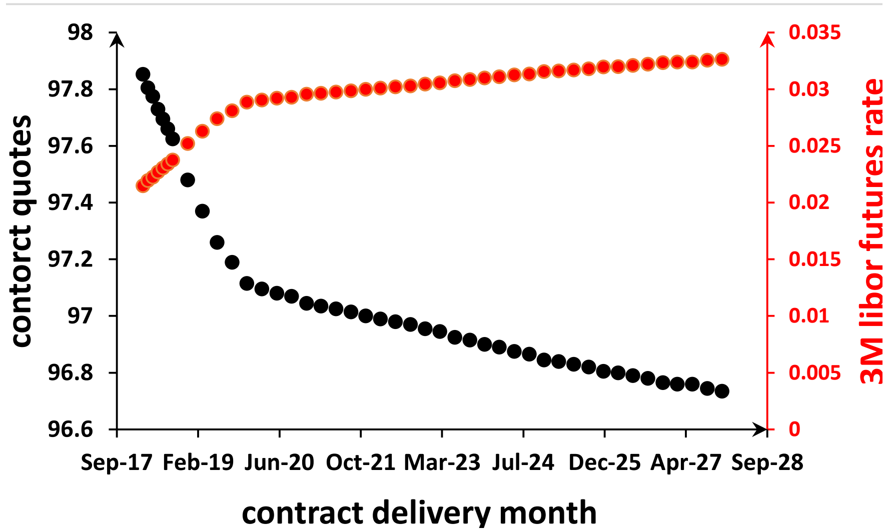


Figure 3.2.1: 3M libor futures rate curve as of Mar,12,2018. Data source http://www.cmegroup.com/trading/interest-rates/stir/eurodollar.html

Example 3.2.2. For example, suppose an investor buys a single three-month Mar2018 contract at 95.00 (implied settlement LIBOR of 5.00%) on Feb, 1, 2018. Then starting from Feb-1-2018 to the delivery/maturity date 19-Mar-2018, there will be cash flow exchange between the investor and the CME exchange. The amount of the exchange is dependent on market 3M libor rate and the contract specified libor rate(here is 5%).

- if at the close of business on Feb-1-2018, the contract price has risen to 95.01 (implying a LIBOR decrease to 4.99%), US$25 will be paid into the investor's margin account;
- if at the close of business on that day, the contract price has fallen to 94.99 (implying a LIBOR increase to 5.01%), US$25 will be deducted from the investor's margin account.
- The process will continue until the end of Mar-19-2018.

3.3 Martingale pricing framework

3.3.1 Change of numeraire

3.3.1.1 *Principles*

Definition 3.3.1 (numeraire). *[1, p. 23] A numeraire is defined as any traded asset that pays no dividends and whose price $A(t)$ is positive at any time $t \geq 0$.*

Proposition 3.3.1 (equivalent martingale measure result, change of numeraire). *[5, p. 661] Consider a market having risk-free asset with short rate r and the market price of risk being σ_g associated with only source of uncertainty. Then under such measure, we have*

$$df = (r + \sigma_g \sigma_f) f dt + \sigma_f f dz$$

and

$$dg = (r + \sigma_g^2) g dt + \sigma_g g dz$$

Then

$$d(\frac{f}{g}) = (\sigma_f - \sigma_g)\frac{f}{g} dz,$$

that is, the quantity

$$\frac{f}{g}$$

is a martingale under the measure associated with market price of risk σ_g.

Proof.

$$d(\ln f) = (r + \sigma_f \sigma_g - \sigma_f^2/2)dt + \sigma_f dz$$

and

$$d(\ln g) = (r + +\sigma_f^2/2)dt + \sigma_f dz$$

Therefore

$$d(\ln f - \ln g) = -\frac{(\sigma_f - \sigma_g)^2}{2} dt + (\sigma_f - \sigma_g) dz$$

and

$$d(\frac{f}{g}) = (\sigma_f - \sigma_g)\frac{f}{g} dz$$

□

Remark 3.3.1. It can be showed that Lemma 1.5.1, under the market price of risk σ_g, we have

$$\frac{r + \sigma_g^2 - r}{\sigma_g} = \frac{r + \sigma_g \sigma_f - r}{\sigma_f} = \sigma_g.$$

Corollary 3.3.1.1. *[5, p. 661]*

$$\frac{f_0}{g_0} = E_g(\frac{f_T}{g_T})$$

or

$$f_0 = g_0 E_g(\frac{f_T}{g_T})$$

where E_g denotes the expected value in a world that is forward risk-neutral with respect to g(i.e., the market price of risk is σ_g).

Remark 3.3.2 (how to do evaluation). Under the market price of σ_g, f_t and g_t are governed by the following SDEs given as

$$df = (r + \sigma_g \sigma_f) f dt + \sigma_f f dz$$

and

$$dg = (r + \sigma_g^2) g dt + \sigma_g g dz$$

or equivalently,

$$d(\frac{f}{g}) = (\sigma_f - \sigma_g)\frac{f}{g} dz$$

We can solve to get the f_T/g_T and g_T.

3.3.1.2 *Money account as numeraire*

Corollary 3.3.1.2 (Money market account as the numeraire).

$$dB = rBdt, B(t) = B(0)\exp(\int_0^t r(s)ds)$$

where r ***can be stochastic****. Then*

$$f_0 = B_0 E_B(\frac{f_T}{B_T}) = B_0 E_B[\frac{f_T}{B_0 \exp(\int_0^T rdt)}] = E_B[\exp(-\int_0^T rdt) f_T]$$

where E_B denotes the expected value in the traditional risk-neutral with respect to B(i.e., the market price of risk is $\sigma_g = 0$). If the short-term interest rate r is constant, then

$$f_0 = e^{-rT} E_B(f_T).$$

Remark 3.3.3 (how to evaluate expectation). Assume interest rate is not stochastic. Under the market price of $\sigma_B = 0$, f_t is governed by the following SDE given as

$$df = rfdt + \sigma_f f dz.$$

We can solve to get the B_T

Remark 3.3.4 (stochastic interest rate issue, theoretical method).

- Propose some short rate model on r and asset dynamics model on f(the dynamics of f should be under risk-neutral measure).
- Evaluate the expectation using **the joint distribution** of $\exp(-\int_0^T r(t)dt)$ and f_T.
- If the short rate dynamics and the asset dynamics do not have the common sources of uncertainty, then

$$E_B[\exp(-\int_0^T r(t)ds)f_T] = E_B[\exp(-\int_0^T r(t)ds)]E_B[f_T] = P(0,T)E_B[f_T].$$

  However, this is impossible since f_t will have drift term including r under risk-neutral measure. This is also the motivation of using zero-coupon bond as the numeraire.

Remark 3.3.5 (stochastic interest rate issue, simulation method).

- Simulate the short term rate r trajectory using some assumed short rate model.
- For each simulated r trajectory:
  - calculate the expected payoff using $E_B[f_T]$
  - discount the expected payoff using $\exp(-\int_0^T rdt)E_B[f_T]$
- Average over many different trajectories of interest rate.

Lemma 3.3.1 (futures contract pricing under stochastic interest rate). *Assume the futures marginal account is cleared in continuous time. Then an asset $S(t)$ has a futures contract price given by*

$$Fu(t) = E_B[S_T|\mathcal{F}_t]$$

where the expectation E_B is taken with respect to the risk-neutral measure.

3.3.1.3 *Zero-coupon bond as numeraire*

Lemma 3.3.2 (pricing using Annuity as numeraire). *Assume there exists a risk-neutral measure such that the price of any traded asset X(without intermediate payments) relative to the money account $B(t)$ is a martingale under Q, i.e.,*

$$\frac{X_t}{B_t} = E_N[\frac{X_T}{B_T}|\mathcal{F}_t], 0 \leq t \leq T.$$

*Let $P(0,T)$ the new numeraire. Then **there exists** a probability measure Q_T, equivalent to Q, such that the price of any attainable claim*

$$\frac{Y_t}{P(t,T)} = E_T[\frac{Y_T}{P(T,T)}|\mathcal{F}_t] = E_T[Y_T|\mathcal{F}_t], 0 \leq t \leq T.$$

Proof. Directly from Theorem 1.6.15. □

Remark 3.3.6 (how to evaluate the expectation and the dynamics of $P(t,T)$)**.** To evaluate the expectation, we usually first propose a dynamics model for $P(t,T)$, then use Theorem 3.3.1 to find out the dynamics of Y_T under this forward measure.

Proposition 3.3.2 (forward price of general assets). *[5, p. 663][1, p. 27] Consider **any** variable S that is not an interesting rate. A forward contract on S with maturity T is defined as a contract that pays off $S(T) - K$ at time T, where $S(T)$ is the value θ at time T. Therefore*

$$F = E_T[S(T)].$$

Or equivalently,

$$S(t) = P(t,T)F \Leftrightarrow F = \frac{S(t)}{P(t,T)} = E_T[\frac{S(T)}{P(T,T)}] = E_T[S(T)]$$

Proof. Define f_0 as the value of this forward contract at $t = 0$. Then

$$f_0 = P(0,T)(E_T[S(T)] - K),$$

where E_T is taking expectation with respect to the forward risk neutral with respect to $P(t,T)$. Note that the forward price F of $S(T)$ is the value of K such that $f_0 = 0$. Therefore

$$F = E_T[\theta_T].$$

Note that $P(t,T)F$ means the value of F discounted to today's value. And under forward measure

$$\frac{S(T)}{P(t,T)}$$

is a martingale(Theorem 1.6.18).

□

Remark 3.3.7 (interpretation)**.**

- The forward price of arbitrary asset is a martingale under forward measure with respect to $P(t,T)$. In other words, the forward price can be regarded as an unbiased estimation of $X(T)$, where the expectation is taken **under the forward measure**.
- Suppose we work under the risk-neutral measure Q, the forward price $F(t,T)$ has to satisfy

$$V(t) = B(t)E_Q[\frac{X(T)-F(t,T)}{B(T)}|\mathcal{F}_t] = 0$$

  where $B(t)$ is the cash account. Because now $B(T)$ is a random variable, and evaluate $E_Q[\frac{X(T)-F(t,T)}{B(T)}|\mathcal{F}_t]$ requires the joint distribution of $X(T)$ and $B(T)$, which can be difficult to calculate. Therefore, under forward measure, we greatly simplify the calculation.

Remark 3.3.8 (constant interest rate case,forward price of stock). Assume the short rate r is constant, then the dynamics of $P(t,T)$ is

$$dP(t,T) = rP(t,T)dt.$$

Under the forward measure with respect to $P(t,T)$, the dynamics of stock is

$$dS_t = rS_t dt + \sigma S_t dW_t,$$

such that $E_T[S_T] = e^{rT}S_0 = F$.

3.3.1.4 *Annuity as numeraire*

Definition 3.3.2 (swap anuity). *[1, p. 12]*

*Given a set of dates $T_0 < T_1 < ... < T_n$ with accrual periods $\tau_i = T_i - T_{i-1}$. The **swap annuity** is defined as*

$$A_{0,n}(t) = \sum_{i=1}^{n} \tau_i P(t, T_i).$$

Note that $A_{0,n}(t)$ is a stochastic process.

Lemma 3.3.3 (pricing using Annuity as numeraire). *Assume there exists a risk-netural measure such that the price of any traded asset X(without intermediate payments) relative to the money account $B(t)$ is a martingale under Q, i.e.,*

$$\frac{X_t}{B_t} = E_N[\frac{X_T}{B_T}|\mathcal{F}_t], 0 \le t \le T.$$

Let $A_{0,n}(t)$ the new numeraire. Then ***there exists*** *a probability measure Q_A, equivalent to Q, such that the price of any attainable claim*

$$\frac{Y_t}{A_{0,n}(t)} = E_U[\frac{Y_T}{A_{0,n}(T)}|\mathcal{F}_t], 0 \leq t \leq T.$$

Proof. Directly from Theorem 1.6.15. □

Remark 3.3.9 (when to use it). The lemma will be useful to price derivatives with a payoff given by $A_{0,n}(T)V(T)$. Then we have

$$V(t) = A_{0,n}(t)E_A[V(T)|\mathcal{F}_t].$$

An example is at Lemma 3.7.8.

3.3.2 Martingale properties under forward measure

3.3.2.1 *Forward price under forward measure*

Proposition 3.3.3 (No-arbitrage forward price for general assets under stochastic interest rate). *[4, p. 112] Consider an asset S_t, its forward price at maturity date T is given by*

- $$F(t,T) = \frac{S_t}{P(t,T)}$$

  for zero dividend assets.

- $$F(t,T) = \frac{S_t \exp(-q(T-t))}{P(t,T)}$$

  for assets with continuous dividend rate q.

- $$F(t,T) = \frac{S_t - \sum_{i=1}^n q_i P(t,T_i)}{P(t,T)}$$

  for assets with discrete dividends $\{q_i\}$ at time $\{T_i\}$; or equivalently,

$$F(t,T) = S_t + \underbrace{P(t,T)((T-t)F(t,t,T)}_{\textit{funding cost}} - \underbrace{\sum_{i=1}^n q_i(1+F(t,T_i,T)(T-T_i)}_{\textit{benefits}},$$

where $F(t, T_i, T)$ is the forward rate with tenor $[T_i, T]$.

Proof. (1) The seller of the contract will take the following strategy:

- enter the contract, which costs zero, as the seller party.
- short $S_t/P(t, T)$ units of $T-$maturity zero-coupon bond.
- long 1 unit of the asset at price S_t.
- At time T, sell the counterparty of the asset at forward price F and repay the short of the zero-coupon bond.

Since the initial cost/position is zero, the payoff for the seller at T is

$$0 = F - S_t/P(t, T)P(T, T) \implies F = \frac{S_t}{P(t, T)}.$$

(2)The seller of the contract will take the following strategy:

- enter the contract, which costs zero, as the seller party.
- short $S_t \exp(-q(T-t))/P(t, T)$ units of $T-$maturity zero-coupon bond.
- long $\exp(-q(T-t))$ unit of the asset at price S_t.
- the dividends from the stock will be reinvested to buy more shares.
- At time T, sell the counterparty of the asset at forward price F and repay the short of the zero-coupon bond.

Let ϕ be the number of shares, then

$$d\phi(t) = (q\phi(t)S_t)/S_t dt = qdt \implies \phi(t) = \exp(qt)\phi(0).$$

Then the total value of the asset at time T is given by

$$\phi(T)S_T = \phi(0)e^{q(T-t)}S_T = S_T.$$

We will sell S_T at price F at time T. Since the initial cost is zero, the payoff for the seller is

$$0 = F - S_t \exp(-q(T-t))/P(t, T) \implies F = S_t \exp(-q(T-t))/P(t, T).$$

(3)(a)The seller of the contract will take the following strategy:

- enter the contract, which costs zero, as the seller party.
- short q_i units of T_i-maturity zero coupon bonds, for $i = 1, 2, ..., n$.
- short $\frac{S_t - \sum_{i=1}^n q_i P(t, T_i)}{P(t, T)}$ units of $T-$maturity zero-coupon bond.
- long 1 unit of the asset at price S_t.
- The received dividends are used to repay the short of zero-coupon bonds at $T_i, \forall i = 1, 2, ..., n$.
- At time T, sell the counterparty of the asset at forward price F and repay the short of the zero-coupon bond.

Since the initial cost is zero, the payoff for the seller is

$$0 = F - \frac{S_t - \sum_{i=1}^n q_i P(t,T_i)}{P(t,T)} \implies F = \frac{S_t - \sum_{i=1}^n q_i P(t,T_i)}{P(t,T)}.$$

(b) The funding cost(borrow money to buy and therefore need to pay interest) is given by

$$P(t,T)((T-t)F(t,t,T).$$

The benefit are coupons paid at T_i and reinvest at forward rate locked at $F(t,T_i,T)$. □

Proposition 3.3.4 (forward price of general assets under forward measure is martingale). *[5, p. 663][1, p. 27] Consider* ***any*** *financial asset (no matter paying dividend or not) S that is not an interesting rate. A forward contract on S with delivery date T is defined as a contract that pays off $S(T)-K$ at time T, where $S(T)$ is the asset value at time T. Let $F(t,T)$ denote the forward price at time t. Then*

$$F(t,T) = E_T[S(T)|\mathcal{F}_t].$$

Proof. (1) From Theorem 3.3.3, we note that no-arbitrage condition requires that

$$F(t,T) = \frac{S_t - \sum_{i=1}^n q_i P(t,T_i)}{P(t,T)}$$

for assets with discrete dividends $\{q_i\}$ at time $\{T_i\}$.

Therefore, $F(t,T)$ is a martingale under T forward measure.

(2) (another method for non-dividend paying asset) Define V_t as the value of this forward contract at t. Then

$$V_t = P(t,T)(E_T[(S(T)-K)|\mathcal{F}_t],$$

where E_T is taking expectation with respect to the forward measure associated with $P(t,T)$. Note that the forward price $F(t,T)$ of $S(T)$ is the value of K such that $V_t = 0$. Therefore

$$F(t,T) = E_T[S(T)|\mathcal{F}_t].$$

□

Remark 3.3.10 (interpretation)**.**

- The forward price of arbitrary asset is a martingale under forward measure with respect to $P(t,T)$. In other words, the forward price can be regarded as an unbiased estimation of $S(T)$, where the expectation is taken **under the forward measure**.

- Suppose we work under the risk-neutral measure Q, the forward price $F(t,T)$ has to satisfy
$$V(t) = B(t)E_Q[\frac{X(T) - F(t,T)}{B(T)}|\mathcal{F}_t] = 0$$
where $B(t)$ is the cash account. Because now $B(T)$ is a random variable, and evaluate $E_Q[\frac{X(T)-F(t,T)}{B(T)}|\mathcal{F}_t]$ requires the joint distribution of $X(T)$ and $B(T)$, which can be difficult to calculate. Therefore, under forward measure, we greatly simplify the calculation.

Example 3.3.1 (constant interest rate case, forward price of stock). Assume the short rate r is constant, then the dynamics of $P(t,T)$ is
$$dP(t,T) = rP(t,T)dt.$$
Under the forward measure with respect to $P(t,T)$, the dynamics of stock is
$$dS_t = rS_t dt + \sigma S_t dW_t,$$
such that $E_T[S_T] = e^{rT}S_0 = F$.

3.3.2.2 *Forward bond and forward rate*

Lemma 3.3.4 (forward bond and forward rate under forward measure). *[1, p. 29] Let $t \leq S \leq T$. We have*

- *The forward bond price FP is a martingale under forward measure with respect to $P(t,S)$; that is,*
$$FP(t;S,T) \triangleq \frac{P(t,T)}{P(t,S)} = E_S[P(S,T)|\mathcal{F}_t].$$
- *The simply compounded forward rate $F(t;S,T)$ is a martingale under forward measure with respect to $P(t,T)$; that is,*
$$F(t;S,T) \triangleq \frac{1}{T-S}(\frac{P(t,S)}{P(t,T)} - 1) = E_T[F(t';S,T)|\mathcal{F}_t], t \leq t' \leq S.$$
- *In particular, if $t = S$, then $F(S;S,T) = L(S,T)$, we have*
$$F(t;S,T) = E_T[L(S,T)|\mathcal{F}_t].$$

Proof. (1) Under forward measure $P(t,S)$, the quantity $FP(t;S,T) \triangleq \frac{P(t,T)}{P(t,S)}$ is a martingale(Theorem 3.3.5). (2) Note that $\frac{P(t,S)}{P(t,T)}$ is martingale under forward measure with respect to

$P(t,T)$. Adding and a constant -1 and multiplying a constant will not change its martingale nature. □

3.3.3 Dynamics under forward measure

3.3.3.1 *Dynamics under forward measure*

Lemma 3.3.5 (changing dynamics from risk-neutral measure to forward measure). *Assume under risk-neutral measure, we have*

$$dS_t/S_t = rdt + \boldsymbol{\sigma}_S^T d\boldsymbol{W}(t)$$
$$dP(t,T)/P(t,T) = rdt + \boldsymbol{\sigma}_P^T d\boldsymbol{W}(t)$$

where $\boldsymbol{\sigma}_S$ and $\boldsymbol{\sigma}_P$ are volatility vectors, and $\boldsymbol{W}(t)$ is the Brownian motion vector.

Then the forward measure Q_T is generated via

$$Z = \frac{dQ_P}{dQ} = \exp(\int_0^t \boldsymbol{\sigma}_P^T d\boldsymbol{W}_s - \int_0^t \frac{1}{2}\boldsymbol{\sigma}_P^T \boldsymbol{\sigma}_P ds)$$

such that under which

- $$d\boldsymbol{W}_t = d\boldsymbol{W}_t^T + \boldsymbol{\sigma}_P dt.$$
- $$d(\frac{S_t}{P_t}) = \frac{S_t}{P_t}(\boldsymbol{\sigma}_S - \boldsymbol{\sigma}_P)^T d\boldsymbol{W}_t$$

Proof. Using Ito rule, we have

$$\begin{aligned} d(\frac{S_t}{P_t}) &= \frac{dS_t}{P_t} - \frac{S_t dP_t}{P_t^2} - \frac{dS_t dP_t}{P_t^2} + \frac{dP_t^2}{P_t^3} \\ &= \frac{S_t}{P_t}(rdt + \boldsymbol{\sigma}_S^T d\boldsymbol{W}_t - rdt - \boldsymbol{\sigma}_P^T d\boldsymbol{W}_t) - \boldsymbol{\sigma}_S^T \boldsymbol{\sigma}_T dt + \boldsymbol{\sigma}_P^T \boldsymbol{\sigma}_T dt) \\ &= \frac{S_t}{P_t}(\boldsymbol{\sigma}_S - \boldsymbol{\sigma}_P)^T (d\boldsymbol{W}_t - \boldsymbol{\sigma}_P dt). \end{aligned}$$

We can define a new measure via

$$Z = \frac{dQ_T}{dQ} = \exp(\int_0^t \boldsymbol{sigma}_T^T d\boldsymbol{W}_s - \int_0^t \frac{1}{2}\boldsymbol{\sigma}_T^T \boldsymbol{\sigma}_T ds)$$

such that

$$d\boldsymbol{W}_t = d\boldsymbol{W}_t^T + \boldsymbol{\sigma}_T dt.$$

Then under the forward measure Q_T,

$$d(\frac{S_t}{P_t}) = \frac{S_t}{P_t}(\boldsymbol{\sigma}_S - \boldsymbol{\sigma}_T)^T d\boldsymbol{W}_t.$$

Also see Theorem 1.6.19. □

Lemma 3.3.6 (changing dynamics from real probability measure to forward measure). *Assume under real probability measure, we have*

$$\begin{aligned} dS_t/S_t &= (r + \lambda^T \boldsymbol{\sigma}_S)dt + \boldsymbol{\sigma}_S^T d\boldsymbol{W}(t) \\ dP(t,T)/P(t,T) &= (r + \lambda^T \boldsymbol{\sigma}_P)dt + \boldsymbol{\sigma}_P^T d\boldsymbol{W}(t) \end{aligned}$$

where $\boldsymbol{\sigma}_S$ and $\boldsymbol{\sigma}_P$ are volatility vectors, and $\boldsymbol{W}(t)$ is the Brownian motion vector, λ is the vector of market price of risks.

Then the forward measure Q_T is generated via

$$Z = \frac{dQ_P}{dQ} = \exp(\int_0^t (\boldsymbol{\sigma}_P - \lambda)^T d\boldsymbol{W}_s - \int_0^t \frac{1}{2}(\boldsymbol{\sigma}_P - \lambda)^T(\boldsymbol{\sigma}_P - \lambda)ds)$$

such that under which

- $$d\boldsymbol{W}_t = d\boldsymbol{W}_t^T + (\boldsymbol{\sigma}_P - \lambda)dt.$$
- $$d(\frac{S_t}{P_t}) = \frac{S_t}{P_t}(\boldsymbol{\sigma}_S - \boldsymbol{\sigma}_P)^T d\boldsymbol{W}_t$$

Proof. Using Ito rule, we have

$$\begin{aligned} d(\frac{S_t}{P_t}) &= \frac{dS_t}{P_t} - \frac{S_t dP_t}{P_t^2} - \frac{dS_t dP_t}{P_t^2} + \frac{dP_t^2}{P_t^3} \\ &= \frac{S_t}{P_t}(rdt + \lambda^T \boldsymbol{\sigma}_S dt + \boldsymbol{\sigma}_S^T d\boldsymbol{W}_t - rdt - \lambda^T \boldsymbol{\sigma}_P dt - \boldsymbol{\sigma}_P^T d\boldsymbol{W}_t) - \boldsymbol{\sigma}_S^T \boldsymbol{\sigma}_T dt + \boldsymbol{\sigma}_P^T \boldsymbol{\sigma}_T dt) \\ &= \frac{S_t}{P_t}(\boldsymbol{\sigma}_S - \boldsymbol{\sigma}_P)^T (d\boldsymbol{W}_t - (\boldsymbol{\sigma}_P - \lambda)dt). \end{aligned}$$

We can define a new measure via

$$Z = \frac{dQ_T}{dQ} = \exp(\int_0^t \boldsymbol{sigma}_T^T d\boldsymbol{W}_s - \int_0^t \frac{1}{2}\boldsymbol{\sigma}_T^T \boldsymbol{\sigma}_T ds)$$

such that

$$d\boldsymbol{W}_t = d\boldsymbol{W}_t^T + \boldsymbol{\sigma}_T dt.$$

Then under the forward measure Q_T,

$$d(\frac{S_t}{P_t}) = \frac{S_t}{P_t}(\boldsymbol{\sigma}_S - \boldsymbol{\sigma}_T)^T d\mathbf{W}_t.$$

Also see Theorem 1.6.19. □

Proposition 3.3.5 (forward bond dynamics under forward measure). *Assume $P(t,T)$ obeys the SDE*

$$dP(t,T) = P(t,T)r(t)dt + P(t,T)\sigma(t,T)dW(t)$$

where $W(t)$ is the standard Brownian motion under the risk-neutral measure Q, and $r(t)$ is the instantaneous short rate. Then,

- *the forward bond price $FP(t;S,T) = \frac{P(t,T)}{P(t,S)}, S \leq T$ has SDE, under measure Q, given by:*

$$dFP/FP = (\sigma(t,T) - \sigma(t,S))dW(t) + \sigma(t,S)(\sigma(t,S) - \sigma(t,T))dt.$$

- *The ratio $\triangleq \frac{P(t,S)}{P(t,T)}, S \leq T$ has SDE, under measure Q, given by:*

$$d\tilde{FP}/\tilde{FP} = (\sigma(t,S) - \sigma(t,T))dW(t) + \sigma(t,T)(\sigma(t,T) - \sigma(t,S))dt.$$

- *Under forward measure with respect to $P(t,S)$.The forward bond price $FP(t;S,T)$ has SDE*

$$dFP/FP = (\sigma(t,T) - \sigma(t,S))dW^S(t).$$

  *Moreover, $FP(t;S,T)$ will follow a geometric Brownian motion (**??**) given by*

$$FP(t';S,T) = FP(t;S,T)\exp(\int_t^{t'} \sigma'(s)dW^S(s) - \frac{1}{2}\int_t^{t'} \sigma'(s)^2 ds),$$

  where $\sigma'(s) = (\sigma(u,T) - \sigma(u,S))$.

- *Use the fact that $FP(S;S,T) = P(S,T)$, we have*

$$P(S,T) = FP(S;S,T) = FP(t;S,T)\exp(\int_t^S \sigma'(s)dW^S(s) - \frac{1}{2}\int_t^S \sigma'(s)^2 ds)$$

- *In the forward measure with respect to $P(t,S)$, the Brownian motion under Q is changed to*

$$dW^S(t) = dW(t) - \sigma(t,S)dt.$$

Proof. (1)(2) Use Ito rule. Particularly, use $X = \log P(t,T), Y = \log P(t,S)$ such that $dX - dY = d(\log(P(t,T)/P(t,S))$ to simplify calculation. (3) Under forward measure of

$P(t,S)$, the drift term $\sigma(t,S)(\sigma(t,S)-\sigma(t,T))$ will add $\sigma(t,S)(\sigma(t,T)-\sigma(t,S))$ based on Theorem 3.3.1Theorem 1.6.18. □

Remark 3.3.11 (caution! forward rate is not martingale in risk-neutral measure)**.** Note that we have shown that for assets other than interest rate derivative, the forward price, under independence assumption, will follow a martingale under risk-neutral measure(see Lemma 1.7.7). For interest rate derivatives, due to the correlation between forward rate the money market numeraire, forward rate is not a martingale under risk-neutral measure.

Remark 3.3.12 (Assumptions on bond dynamics under risk-neutral measure)**.** The validity of this assumption is discussed at Lemma 3.5.1.

Remark 3.3.13 (applications and analog is stocks)**.**

- One of most important result in this theorem is, under forward measure, it gives the distribution of the price at future time S of a zero-coupon bond maturing at T, given by
$$P(S,T) = FP(t;S,T)\exp(\int_t^S \sigma'(s)dW^S(s) - \frac{1}{2}\int_t^S \sigma'(s)^2 ds).$$
- The analog for stock price M prediction under risk-neutral measure Q gives
$$M(S) = M(t)\exp(\int_t^S \sigma_M(s)dW^Q(s) - \frac{1}{2}\int_t^S \sigma_M(s)^2 ds).$$
- The prediction of $P(S,T)$ is useful to price Vanilla bond options.

Remark 3.3.14 (understand $P(S,T)$ and $L(S,T)$)**.** Note that $P(S,T)$ and $L(S,T)$ are random variables; but they **are not** $\mathcal{F}_t, t < S$ **measurable**. On the contrary, $FP(t;S,T)$ and $F(t;S,T)$ are $\mathcal{F}_t$ measurable, and they are the prediction on $P(S,T)$ and $L(S,T)$ based on the information $\mathcal{F}_t$, under the forward measure.

Example 3.3.2 (application examples). Use the fact that $F(t;S,T)$ is a martingale under forward measure with respect to $P(t,T)$.

- An forward rate agreement with rate K has payoff at future time T
$$(T-S)(K-L(S,T)).$$
Note that its current value $V(t)$ satisfies
$$\frac{V(t)}{P(t,T)} = E_T[\frac{(T-S)(K-L(S,T))}{P(T,T)}|\mathcal{F}_t] = (T-S)K - (T-S)F(t;S,T).$$

- An interest rate swap has present value

$$PS(t) = \sum_{i=1}^{n} P(t, T_i) E_{T_i}[\tau_i(L(T_{i-1}, T_i) - K) | \mathcal{F}_t] = \sum_{i=1}^{n} P(t, T_i)\tau_i(F(t; T_{i-1}, T_i) - K).$$

Remark 3.3.15 (drift term of bond SDE). The SDE for zero-coupon bond may have different drift terms instead of $P(t, T)r(t)dt$ in the real probability measure; however, in complete market with no-arbitrage assumption, there exists an unique risk-neutral measure under which the bond SDE will have drift term $P(t, T)r(t)$.

3.3.3.2 *State price interpretation of forward measure*

Lemma 3.3.7 (fundamental solution to term structure equation). *[6, p. 35] Assume the short rate r under risk-neutral measure is governed by*

$$dr = m(r, t)dt + \sigma(r, t)dz.$$

Given the term-structure equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial r^2} + mS\frac{\partial V}{\partial S} - rV = 0,$$

and let $V^\delta(t, r_t, s) > 1$ be the solution for final condition $V_T^\delta(r_T, s) = \delta(r_T - s)$. It follows that

- $$P(t, T, r_t) = \int_{-\infty}^{\infty} V^\delta(t, r_t, s) ds$$

- *Define $p^\delta(t, r_t, s) \triangleq V^\delta(t, r_t, s)/P(t, T, r_t)$, then*

$$\int_{-\infty}^{\infty} p^\delta(t, r_t, s) ds = 1.$$

- *For any payoff function $V_T(r_T)$, we have its current value given by*

$$\begin{aligned} V(t) &= \int_{-\infty}^{\infty} V_T(r_T = s) V^\delta(t, r_t, r_T = s) ds \\ &= P(t, T, r_t) \int_{-\infty}^{\infty} V_T(S_T = s) p^\delta(t, r_t, r_T = s) ds \\ &= P(t, T, r_t) E_T[V_T(S_T) | \mathcal{F}_t] \end{aligned}$$

where E_T denotes taking expectation with respect to p^δ.

Proof. (1)(2) Use the fact that a zero-coupon bond pays 1 at all world states. (3) use linearity of pricing. □

Remark 3.3.16 (interpretation)**.**

- This lemma shows that the nature of forward probability measure, which is the state price(expressed in unit of $P(t,T,r_t)$) resulted from market equilibrium based on current market condition(time t and current short rate r_t).
- The forward probability/state price therefore will depends on time t and short rate r_t.

3.3.4 Applications in pricing

3.3.4.1 *LIBOR payment*

Proposition 3.3.6 (current no-arbitrage value of LIBOR payment, recap). *[1, p. 6] Let $S < T$. Consider the future LIBOR-based payment $(T-S)L(S,T)$ at time T. Its arbitrage-free value at time $t < T$ is $P(t,S) - P(t,T), T > S$. That is*

$$\begin{aligned} V(t) &= E_Q[\frac{M(t)(T-S)L(S,T)}{M(T)}|\mathcal{F}_t] \\ &= P(t,S) - P(t,T) \end{aligned}$$

Proof. (martingale method under forward measure Lemma 3.3.4)

$$\begin{aligned} V(t) &= P(t,T)E_T[\delta L(S,T)|\mathcal{F}_t] \\ &= P(t,T)E_T[\delta F(S;S,T)|\mathcal{F}_t] \\ &= P(t,T)\delta F(t;S,T) \\ &= P(t,T)(\frac{P(t,S)}{P(t,T)} - 1) \\ &= P(t,S) - P(t,T) \end{aligned}$$

(martingale method under risk-neutral measure)

$$\begin{aligned} V(t) &= M(t)E_Q[\frac{(T-S)L(S,T)}{M(T)}|\mathcal{F}_t] \\ &= E_Q[\frac{M(t)}{M(T)}(\frac{M(T)}{M(S)} - 1)|\mathcal{F}_t] \\ &= E_Q[\frac{M(t)}{M(S)} - \frac{M(t)}{M(T)}|\mathcal{F}_t] \\ &= P(t,S) - P(t,T) \end{aligned}$$

where we use no-arbitrage relation between LIBOR and money account (Theorem 3.1.1).

See also the replication method (Theorem 3.1.2). □

Lemma 3.3.8 (current value of LIBOR payment risk-neutral Method). *[1, p. 6] Let $t < S < T$. It follows that*

- *Time t value of $(T-S)L(S,T)$ is given by*

$$V(t) = B(t)E_Q[\frac{(T-S)L(S,T)}{B(T)}|\mathcal{F}_t] = P(t,S) - P(t,T).$$

- $$E_T[L(S,T)|\mathcal{F}_t] = F(t,S,T)$$

Proof. (3) Note that under risk-neutral measure

$$\frac{P(t,T)}{B(t)} = E_Q[\frac{P(T,T)}{B(T)}|\mathcal{F}_t].$$

(4) From risk-neutral pricing and the results from (1)(2)(3), we have

$$\begin{aligned}
V(t) &= B(t)E_Q[\frac{(T-S)L(S,T)}{B(T)}|\mathcal{F}_t] \\
&= E_Q[\frac{B(t)}{B(T)}(\frac{B(T)}{B(S)} - 1)|\mathcal{F}_t] \\
&= E_Q[\frac{B(t)}{B(S)} - \frac{B(t)}{B(T)}|\mathcal{F}_t] \\
&= P(t,S) - P(t,T)
\end{aligned}$$

(4)

$$\begin{aligned}
E_T[L(S,T)|\mathcal{F}_t] &= E_T[\frac{1}{T-S}(\frac{1}{P(S,T)} - 1)|\mathcal{F}_t] \\
&= E_T[\frac{1}{T-S}(\frac{P(S,S)}{P(S,T)} - 1)|\mathcal{F}_t] \\
&= \frac{1}{T-S}(\frac{P(t,S)}{P(t,T)} - 1) \\
&= F(t,S,T)
\end{aligned}$$

where we use the fact that $\frac{P(t,S)}{P(t,T)}$ is a martingale under T forward measure. □

Remark 3.3.17 (interpretation).

- (interpret $L(S,T)$) The quantity $(T-S)L(S,T)$ is stochastic and not $\mathcal{F}_t$ measurable. Note that from definition,
$$P(S,T) = \frac{1}{1+(T-S)L(S,T)}.$$
And $P(t,S) - P(t,T)$ is the market price at time t, which is $\mathcal{F}_t$ measurable(or known at time t).
- (LABOR-in-arrears) If the payment is made at time S, see Lemma 3.9.1.
- **(caution!)** note that $1/P(S,T)$ is not a martingale under T forward measure, but $\frac{P(t,S)}{P(t,T)}$ is .

3.3.4.2 *Interest rate swaps*

Lemma 3.3.9 (the evolution of an interest rate swap value). *Let the current time be t_0. Consider a payer swap starting at T_0 with cash flow at $T_1, T_2, ..., T_n$ given by*
$$\tau_i(L(T_{i-1},T_i) - K), i = 1,2,...,n,$$
such that when $K = S(t_0; T_0, \Delta)$, the value of swap at current time t_0 is zero. It follows that

- *At future time $t_1, t_0 \leq t_1 \leq T_0$, the value of the contract is given by*
$$V(t_1) = A(t_1, T_0, T_n)(S(t_1; T_0, T_n) - S(t_0; T_0, T_n)),$$
where $S(t_1; T_0, \Delta)$ is the fair forward starting rate at t_1, and $A(t, T_0, T_n)$ is the annuity price given by
$$A(t, T_0, T_n) = \sum_{i=1}^{n} \tau_i P(t_i, T_n).$$
- *At future time $t_1, t_1 > T_0$, the value of the contract is given by*
$$V(t_1) = A_{\alpha(t_1),n}(t_1, T_{\alpha(t_1)}, T_n)(S(t_1; T_{\alpha(t_1)}, T_n) - S(t_0; T_0, T_n)),$$
where $\alpha(t_1)$ is the minimum integer such that $T_\alpha \geq t_1$, $S(t_1, T_{\alpha(t_1)}, T_n)$ is the fair forward rate at t_1 for a swap with payment spanning $T_{\alpha(t_1)}, ..., T_n$, $\tau = T - S$, $A_{\alpha(t_1,n)}(t, T_{\alpha(t_1)}, T_n)$ is the annuity price given by
$$A(t, T_{\alpha(t_1)}, T_n) = \sum_{i=\alpha(t_1)}^{n} \tau_i P(t_1, T_n).$$

At future time $t_1, t_0 < t_1 < T_n$, the

Note that** $A(t, T_{\alpha(t_1)}, T_n)$**,** $S(t_1, T_{\alpha(t_1)}, T_n)$ **and** $V(t_1)$ **are all random quantities at time** t_0**.

In particular, $V(t_0) = V(T_n) = 0$.

Proof. Note that at any time $t < T_n$, the fair value swap rate is related to the LIBOR payments via

$$\begin{aligned} V(t) &= \sum_{i=\alpha(t)}^{n} P(t, T_i)\tau_i E_{T_i}[S_{\alpha(t),n}(t) - L(T_{i-1}, T_i)|\mathcal{F}_t] = 0 \\ \implies A_{\alpha(t),n}(t) S_{\alpha(t),n}(t) &= \sum_{i=\alpha(t)}^{n} P(t, T_i)\tau_i E_{T_i}[L(T_{i-1}, T_i)|\mathcal{F}_t] \end{aligned}$$

For a contract entered at time t_0, the fixed leg has t_1-value $A_{\alpha(t_1),n}(t_1) S_{\alpha(t_0),n}(t)$, the floating leg has value

$$\sum_{i=\alpha(t_1)}^{n} P(t_1, T_i)\tau_i E_{T_i}[L(T_{i-1}, T_i)|\mathcal{F}_t] = A_{\alpha(t_1),n}(t) S_{\alpha(t_1),n}(t).$$

Therefore, the value of the forward contract is given by

$$V(t_1) = A(t_1, T_{\alpha(t_1)}, T_n)(S(t_1, T_{\alpha(t_1)}, T_n) - S(t_0, T_0, T_n)).$$

□

Remark 3.3.18 (reduction to forward contract value evolution). Note that forward contract value evolution(Theorem 3.1.4) is given by

$$V(t_1) = P(t_1, T)\tau(F(t_0; S, T) - F(t_1; S, T)),$$

which is the same as one-period interest rate swap with annuity $A(t) = \tau P(t, T)$.

3.3.4.3 *Vanilla bond options*

Lemma 3.3.10 (Black's model for call option on zero-coupon bond). *[1, p. 32] The value at time t of a European call option maturing at S with strike K on a bond maturing at $T > S$ is given by*

$$V(t) = P(t, S)E_S[\max(P(S, T) - K, 0)|\mathcal{F}_t],$$

where the expectation is taken with respect to forward measure of $P(t, S)$. If zero-coupon bonds follows the SDE

$$dP(t, T) = r(t)P(t, T)dt + \sigma(t, T)P(t, T)dW(t)$$

with deterministic volatility $\sigma(t,T)$, then

$$V(t) \triangleq BC(t;S,T,K) = P(t,T)N(d_+) - KP(t,S)N(d_-),$$

where

$$d_+ = \frac{\log \frac{FB(t;S,T)}{K} + \frac{1}{2}v(t,S)^2}{v(t,S)} = \frac{\log \frac{P(t,T)}{KP(t,S)} + \frac{1}{2}v(t,S)^2}{v(t,S)}, d_- = d_+ - v(t,S).$$

and

$$v(t,S)^2 = \int_t^S (\sigma(u,T) - \sigma(u,S))^2 dt$$

Proof. Under the forward measure with respect to $P(t,S)$, we have

$$\begin{aligned}
\frac{V(t)}{P(t,S)} &= E_S[\frac{\max(P(S,T)-K,0)}{P(S,S)}|\mathcal{F}_t] \\
&= E_S[\max(P(S,T)-K,0)] \\
&= E_S[FP(S;S,T)1_{P(S,T)>K}] - E_S[K1_{P(S,T)>K}] \\
&= FP(t;S,T)N(d_+) - KN(d_-)
\end{aligned}$$

Note that $FB(t;S,T)$ is martingale under S forward measure. Then probability distribution of $P(S,T)$ under such forward measure Theorem 3.3.5 will be such that

$$P(S,T) = FB(S;S,T) = FB(t;S,T)\exp(\int_t^S \sigma'(s)dW^S(s) - \frac{1}{2}\int_t^S \sigma'(s)^2 ds),$$

where $\sigma'(s) = (\sigma(u,T) - \sigma(u,S))$. Note that it is easy to see that $\ln \frac{P(S,T)}{FP(t;S,T)}$ has mean $\frac{1}{2}\int_t^S \sigma'(s)^2 ds$ and variance $\int_t^S \sigma'(s)^2 dt$.

Note that $FP(t;S,T)$ is $\mathcal{F}_t$ measurable; therefore, we can calculate the distribution $P(S,T)$.

□

Lemma 3.3.11 (put-call parity and put option pricing). *Let $BC(t;S,T,K)$ and $BP(t;S,T,K)$ denote the time t value of the call and put option (expiration date S and strike K) on a bond matures at time T, then*

$$BC(t;S,T,K) - BP(t;S,T,K) = P(t,T) - KP(t,S).$$

Therefore, the put option price is given by

$$BP(t;S,T,K) = KP(t,S)N(-d_-) - P(t,T)N(-d_+).$$

Proof. We the change of numeraire and forward measure with respect to $P(t,S)$, we haves

$$\begin{aligned}\max(P(S,T)-K,0)-\max(K-P(S,T),0) &= P(S,T)-K\\ E_S[\max(P(S,T)-K,0)|\mathcal{F}_t]-E_S[\max(K-P(S,T),0)|\mathcal{F}_t] &= E_S[P(S,T)|\mathcal{F}_t]-E_S[K|\mathcal{F}_t]\\ (BC(t;S,T,K)-BP(t;S,T,K))/P(t,S) &= P(t,T)/P(t,S)-K\\ BC(t;S,T,K)-BP(t;S,T,K) &= P(t,T)-KP(t,S)\end{aligned}$$

□

3.3.4.4 *Caps and floors*

Lemma 3.3.12 (Black's model for European caps). *[5, p. 680][3, p. 199] Consider a caplet with payoff at t_{k+1} given by*

$$(L(t_k,t_{k+1})-K)^+$$

*where K is the **cap rate**. Assume under the forward measure associated with numeriare $P(t,t_{k+1})$, the forward rate $F(t,t_k,t_{k+1})$ has dynamics*

$$dF = \tilde{\sigma}_F(t)F dW_t.$$

Then the value of a caplet is given by

$$\delta_k P(t,t_{k+1})[F_k N(d_1)-KN(d_2)],$$

where

$$d_1 = \frac{\ln(F_k/K)+\sigma_F^2/2}{\sigma_F}, d_2 = d_1-\sigma_F,$$

$\delta_k = t_{k+1}-t_k$, $F_k = F(t;t_k,t_{k+1}), \sigma_F^2 = \int_t^{t_k}\tilde{\sigma}_F(t)^2 dt.$

The cap consisting of n caplets has price given by

$$\sum_{k=1}^{n}\delta_k P(t,t_{k+1})[F_k N(d_1)-KN(d_2)].$$

Proof. Under the forward measure associated with numeriare $P(t,t_{k+1})$ We have pricing formula

$$V(t) = P(t,t_{k+1})E_{t_{k+1}}[(L(t_k,t_{k+1})-K)^+|\mathcal{F}_t],$$

Note that

$$L(t_k,t_{k+1}) = F(t_k;t_k,t_{k+1}) = F(t;t_k,t_{k+1})\exp(-\frac{1}{2}\sigma_F^2+\sigma_F Z),$$

where $Z\sim N(0,1), \sigma_F^2 = \int_t^{t_k}\tilde{\sigma}_F(t)^2 dt.$ □

Lemma 3.3.13 (Pricing caplets/floorlets using Vanilla options on bonds). *[1, p. 34] Let $BC(t; T_{i-1}, T_i, \frac{1}{1+\tau_i K}), BP(t; T_{i-1}, T_i, \frac{1}{1+\tau_i K})$ denote a call/put option (expiry T_{i-1}) on a zero-coupon bond (maturity T_i) with strike $\frac{1}{1+\tau_i K}$.*

Then

$$Caplet(t) = (1+\tau_i K) BP(t; T_{i-1}, T_i, \frac{1}{1+\tau_i K})$$

$$Flrlet(t) = (1+\tau_i K) BC(t; T_{i-1}, T_i, \frac{1}{1+\tau_i K})$$

Proof. The payoff at T_{i-1} for a caplet with strike K and maturity T_{i-1} is given by

$$\tau_i P(T_{i-1}, T_i)(L(T_{i-1}, T_i) - K)^+$$

From the definitions,

$$P(T_{i-1}, T_i) = \frac{1}{1+\tau L(T_{i-1}, T_i)} \Leftrightarrow L(T_{i-1}, T_i) = \frac{1 - P(T_{i-1}, T_i)}{\tau P(T_{i-1}, T_i)}$$

We have

$$(1+\tau_i K)(\frac{1}{1+\tau_i K} - P(T_{i-1}, T_i))^+.$$

This is equivalent to a vanilla put option with strick $\frac{1}{1+\tau_i K}$. □

Lemma 3.3.14 (put-call parity for caps and floors). *[5, p. 680] Let V_{cap}, V_{floor} and V_{swap} denote the prices of a cap, a floor and a swap. Let the cap and floor have the same strike price, R_K. The swap is an agreement to receive LIBOR and pay a fixed rate of R_K with no exchange of payment on the first reset date.All three instruments have the same life and the same frequency of payments. We have*

$$V_{cap} = V_{floor} + V_{swap}.$$

Proof. Consider we long a cap and short a floor. The cap provides a cash flow of $LIBOR - R_K$ when $LIBOR > R_K$; the short floor provides a cash flow of $-(R_K - LIBOR)$ when $LIBOR < R_K$. Therefore, the position amounts to a swap with cash flow $LIBOR - R_K$. □

3.3.4.5 *Swaption*

Lemma 3.3.15 (value of a swaption). *[1, p. 36] Consider an interest rate swap with payment dates $T_1, T_2, ..., T_n$ and reset dates $T_0, T_1, ..., T_{n-1}$. Let K be the strike and $T_s \leq T_0$ be the maturity of the swaption. It follows that*

- *The payer swaption payoff at maturity date T_s is*

$$PSwpt(T_s, T_0, T_n) = [PS(T_s)]^+ = A(T_s, T_0, T_n)(S(T_s, T_0, T_n) - K)^+,$$

where $A(t, T_0, T_n)$ is the swap annuity given by

$$A(t, T_0, T_n) = \sum_{i=1}^{n} \tau_i P(t, T_i).$$

- *The payer swaption value at time $t < T_s$ is given by*

$$V(t) = A(t, T_0, T_n) E_A[(S_{0,n}(T_s) - K)^+ | \mathcal{F}_t],$$

where E_A is the expectation taken with respect to the martingale measure associated with $A(t, T_0, T_n)$.
- *If K is set to $S(t, T_0, T_n), t < T_s$, then*

$$V(t) = A(t, T_0, T_n) E_A[(S(T_s, T_0, T_n) - S(t, T_0, T_n))^+ | \mathcal{F}_t].$$

Note that the forward swap rate $S(t, T_0, T_n)$ is a stochastic process.

Proof. (1) Under the forward measure of $T_1, T_2, ..., T_n$, we have

$$\begin{aligned}
PS(T_s) &= \sum_{i=1}^{n} P(T_s, T_i) E_{T_i}[\tau_i (L(T_{i-1}, T_i) - K) \mathcal{F}_{T_s}] \\
&= \sum_{i=1}^{n} P(T_0, T_i) \tau_i (F(T_0; T_{i-1}, T_i) - K) \\
&= P(T_s, T_0) - P(T_s, T_n) - K \sum_{i=1}^{n} \tau_i P(T_s, T_i) \\
&= (S_{0,n}(T_s) - K) \sum_{i=1}^{n} \tau_i P(T_s, T_i) \\
&= (S_{0,n}(T_s) - K) A_{0,n}(T_s)
\end{aligned}$$

(2) Under the annuity measure, we have

$$V(t) = A_{0,n}(t) E_A[\frac{A_{0,n}(T_s)(S_{0,n}(T_s) - K)^+}{A_{0,n}(T_s)} | \mathcal{F}_t] = A_{0,n}(t) E_A[(S_{0,n}(T_s) - K)^+ | \mathcal{F}_t],$$

□

Remark 3.3.19 (model dependent valuation). Note that the valuation of $E_A[(S_{0,n}(T_s) - K)^+ | \mathcal{F}_t]$ depends on the model dynamics of $S_{0,n}$. See Lemma 3.7.10 for an example.

Lemma 3.3.16 (put-call parity for swaption). *Consider a payer swaption and a receiver swaption with the same strike K and maturity T_s. It follows that*

- *the payoff at maturity date T_s satisfies*

$$[PS(T_s)]^+ - [-PS(T_0)]^+ = PS(T_s);$$

- *the value at $t_0 < t < T_s$ satisfies*

$$PSwpt_{0,n}(t) - RSwpt_{0,n}(t) = PS(t) = A_{0,n}(t_1)(S_{0,n}(t) - S_{0,n}(t_0)),$$

where $S_{0,n}(t_0), t_0 \leq t$ is the fair value swap rate at t_0 when the swap contract was entered.

-

$$PSwpt_{0,n}(t_0) = RSwpt_{0,n}(t_0).$$

Proof. (1) Straight forward. (2) Using annuity as the measure, we have

$$[PS(T_s)]^+ - [-PS(T_0)]^+ = PS(T_s)$$

$$A(t)E_A[\frac{[PS(T_s)]^+ - [-PS(T_0)]^+}{A(T_s)}|\mathcal{F}_t] = A(t)E_A[\frac{PS(T_s)}{A(T_s)}|\mathcal{F}_t] = PS(t) = A_{0,n}(t)(S_{0,n}(t) - S_{0,n}(t_0))$$

where we use the value of interest rate swap(Lemma 3.3.9). (3) Note that $PS(t_0) = 0$. □

3.3.4.6 *Equity option pricing in stochastic interest rate*

Lemma 3.3.17 (European equity options pricing under stochastic interest rate). *[5, p. 666] Assume the stock price follows geometric Brownian motion and the price of zero-coupon-bond $P(0,T)$ is known.*

Define F_0 and F_T as the forward price of the asset at time 0 and T for a contract maturing at time T.

The European call option price is given by

$$c = P(0,T)E_T[\max(S_T - K, 0)] = P(0,T)E_T[\max(F_T - K, 0)].$$

*where E_T denotes the expectations in a world that is forward risk neutral with respect to $P(t,T)$. Assume that F_t is lognormal distribution **in this forward measure** with variance $\sigma_F^2 T$. Then*

$$c = P(0,T)[F_0 N(d_1) - KN(d_2)],$$

where

$$d_1 = \frac{\ln(F_0/K) + \sigma_F^2 T/2}{\sigma_F\sqrt{T}}, d_2 = d_1 - \sigma_F T.$$

Proof. From the definition of forward price $F_T = S_T$. Under forward measure with respect of $P(t,T)$,

$$F_t \triangleq \frac{S(t)}{P(t,T)}$$

is a martingale(Theorem 3.3.4).

Therefore, $F_T = F_0 \exp(-\frac{1}{2}\sigma_F^2 T + \sigma_F\sqrt{T}Z)$, where $Z \in N(0,1)$

Use the distribution of F_T under forward measure, we can calculate

$$P(0,T)E_T[\max(F_T - K, 0)]$$

.

Note that the $F_0 = S_0/P(0,T)$ and σ_F can be calculated using Theorem 1.6.19. □

Note 3.3.1 (uncertainty increase the value of options). The stochastic nature of the interest rate will make $\sigma_F > \sigma_S$, thus generally increasing the value of the options. See the following example.

Example 3.3.3. link Consider the interest rate model and the stock price model (under risk-neutral measure) as

$$dr_t = (\theta_t - ar_t)dt + \sigma_0 dW_t^1$$
$$dS_t = S_t(r_t dt + \sigma(\rho dW_t^1 + \sqrt{1-\rho^2})dW_t^2)$$

where W_t^1, W_t^2 are two independent Brownian motions.Then,

The zero-coupon bond price $P(t,T)$ is given as

$$dP(t,T) = P(t,T)(r_t dt - \sigma_0 D(t,T)dW_t),$$

where

$$D(t,T) = \int_t^T e^{-a(s-t)} ds$$

- Under the forward risk-neutral of $P(t,T)$, we have

$$dP(t,T) = P(t,T)(r_t - \sigma_0^2 D(t,T)^2 dW_t^1)dt - \sigma_0 D(t,T)dW_t^1),$$

and

$$dS_t = S_t((r_t - \rho\sigma\sigma_0 D(t,T))dt + \sigma(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2)$$

- Let $F(t,T) = S_t/P(t,T)$, then $F(t,T)$ is a martingale under the forward risk-neutral of $P(t,T)$, given as

$$dF(t,T) = F(t,T)((\sigma\rho + \sigma_0 D(t,T))dW_t^1 + \sigma\sqrt{1-\rho^2}dW_t^2),$$

which is a martingale.
Then, we can evaluate the distribution of $F(T,T)$ and the expectation $E_T[\max(F(T,T)-K,0)]$.

3.4 Market yield curve analysis and construction

3.4.1 Concepts and facts

Note 3.4.1 (yield curve concept and market observations). link Let the current time be t and the maturity be T. A yield curve observed at the market today is usually denoted by $y(T)$. It means that if we invest 1$ in the money market for T years, we will get $(1+y)^T$ or $\exp(yT)$ after T years.

As showed in Figure 3.4.1, the market yield curves usually have three types of shapes:

- **(normal yield curve)** A normal or up-sloped yield curve indicates yields on longer-term bonds may continue to rise, responding to periods of economic expansion. When investors expect longer-maturity bond yields to become even higher in the future, many would temporarily park their funds in shorter-term securities in hopes of purchasing longer-term bonds later for higher yields. In a rising interest rate environment, it is risky to have investments tied up in longer-term bonds when their value has yet to decline as a result of higher yields over time. The increasing temporary demand for shorter-term securities pushes their yields even lower, setting in motion a steeper up-sloped normal yield curve.
- **(inverted yield curve)** An inverted or down-sloped yield curve suggests yields on longer-term bonds may continue to fall, corresponding to periods of economic recession. When investors expect longer-maturity bond yields to become even lower in the future, many would purchase longer-maturity bonds to lock in yields before they decrease further. The increasing onset of demand for longer-maturity bonds and the lack of demand for shorter-term securities lead to higher prices but lower

yields on longer-maturity bonds, and lower prices but higher yields on shorter-term securities, further inverting a down-sloped yield curve.

- **(flat(humped) yield curve)** A flat yield curve may arise from normal or inverted yield curve, depending on changing economic conditions. When the economy is transitioning from expansion to slower development and even recession, yields on longer-maturity bonds tend to fall and yields on shorter-term securities likely rise, inverting a normal yield curve into a flat yield curve. When the economy is transitioning from recession to recovery and potential expansion, yields on longer-maturity bonds are set to rise and yields on shorter-maturity securities are sure to fall, tilting an inverted yield curve toward a flat yield curve.

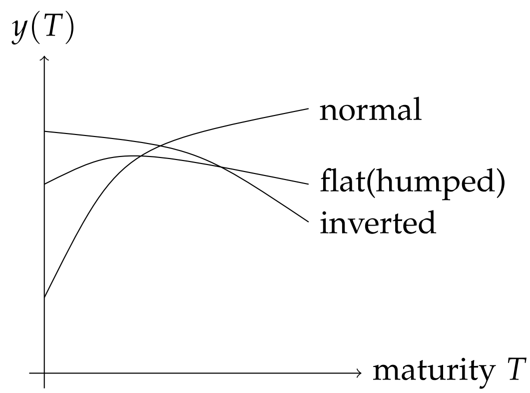


Figure 3.4.1: Different types of market observed yield curve shape

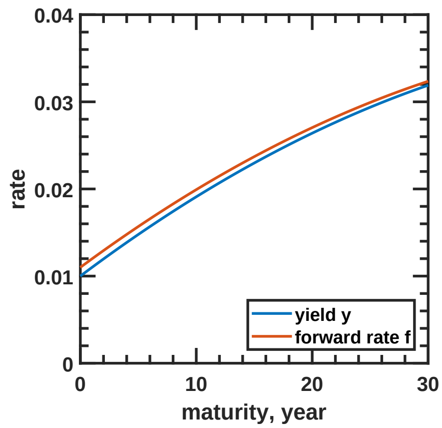
Note 3.4.2 (equivalence of yield curve, zero-coupon curve, and (instantaneous)forward curve, recap of Lemma 3.1.4).

- Given the zero-coupon curve $P(t,T)$, we can calculate the continuously compounding yield curve
$$y(t,T) = R(t,T) \triangleq -\frac{\log P(t,T)}{T-t},$$
and the instantaneous forward curve
$$f(t,T) \triangleq -\frac{\partial}{\partial T}\log P(t,T).$$
- The yield curve and instantaneous forward rate curve are connected by
$$f(t,T) = -\frac{\partial}{\partial T}\log P(t,T) = \frac{\partial}{\partial T}(R(t,T)(T-t)) = R(t,T) + (T-t)\frac{\partial}{\partial T}R(t,T).$$
  - If $f(t,T) = f_0$(i.e., it is flat), then $R(t,T) = f(t,T) = 0$.

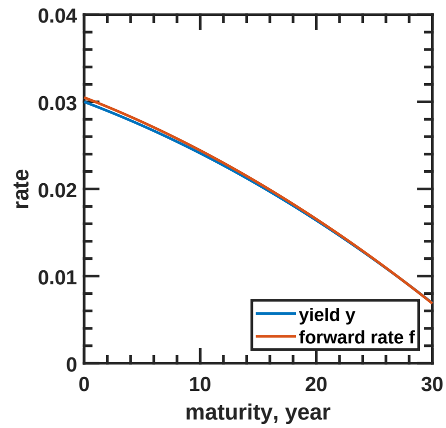
- If $R(t,T)$ is monotonically increasing, then $f(t,T)$ is above $R(t,T)$.
- If $R(t,T)$ is monotonically decreasing, then $f(t,T)$ is below $R(t,T)$.

- The instantaneous forward rate curve can be used to calculate the forward rate of any tenor (T_1, T_2) via
$$F(t,T_1,T_2) = \frac{1}{T_2 - T_1}(\exp(\int_{T_1}^{T_2} f(t,s)ds) - 1).$$
- Similarly, given by the yield curve $R(t,T)$, we have
$$P(t,T) = \exp(-R(t,T)(T-t)).$$

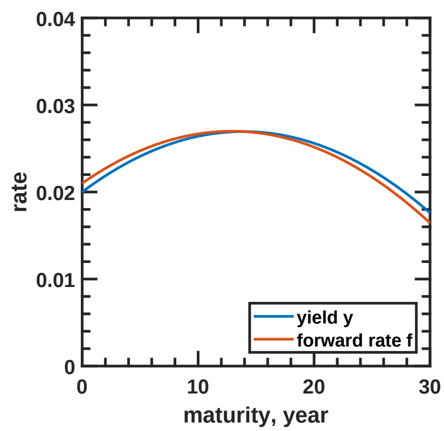
Remark 3.4.1. recall that
$$P(t,T) = \exp(-R(t,T)(T-t)), P(t,T) = \exp(-\int_t^T f(t,s)ds)$$



(a) **Normal** yield curve and forward rate curve.



(b) **Inverted** yield curve and forward rate curve.



(c) **Humped** yield curve and forward rate curve.

Figure 3.4.2: Yield curves and forward rate curves.

3.4.2 Market common instrument data

Table 3.4.1: USD LIBOR spot rate as of March-27-2018

USD LIBOR overnight	1.69750%
USD LIBOR 1week	1.73125%
USD LIBOR 1month	1.87688%
USD LIBOR 2 months	1.99438%
USD LIBOR 3 months	2.30200%
USD LIBOR 6 months	2.45299%
USD LIBOR 12 months	2.67138%

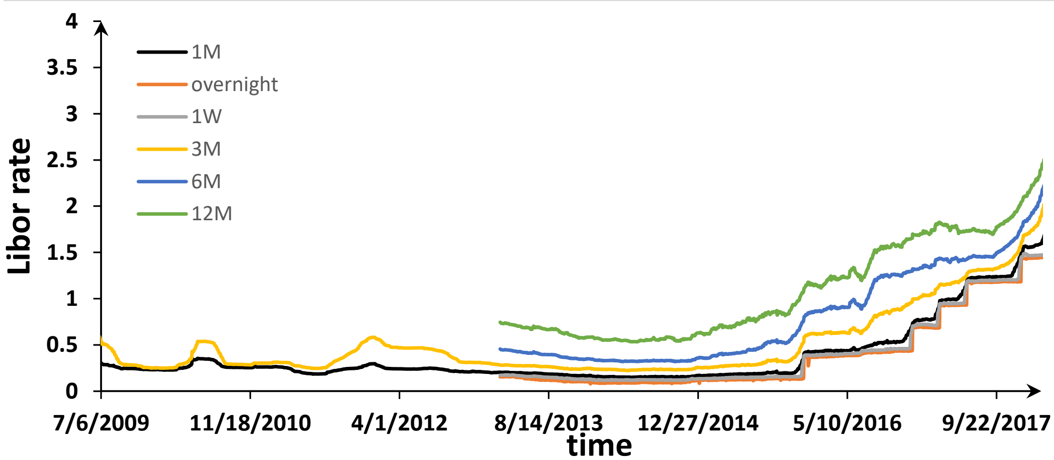


Figure 3.4.3: Historical Libor spot rate with different tenors. Data source https://fred.stlouisfed.org

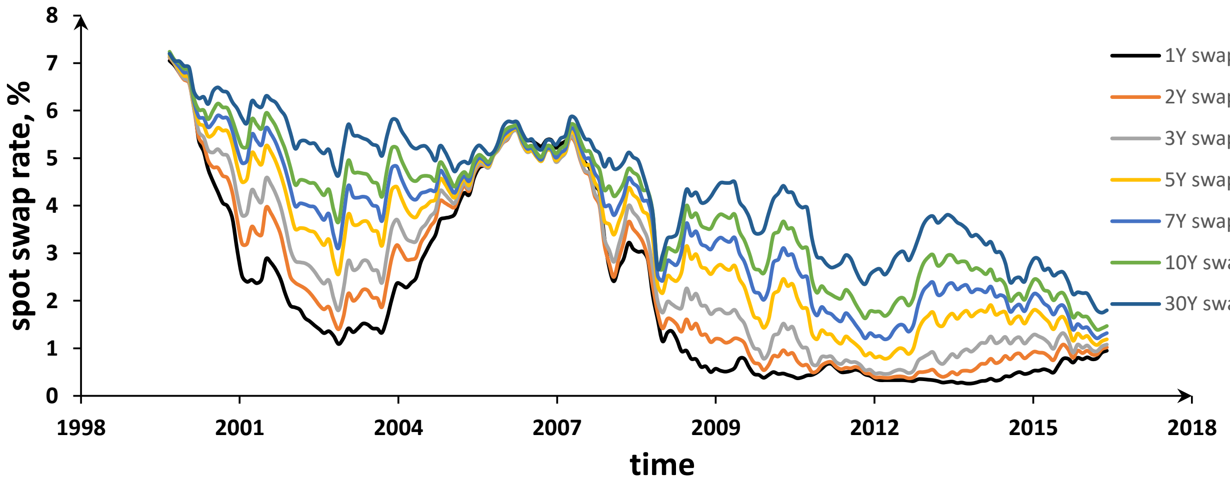


Figure 3.4.4: Historical swap spot rate with different tenors. Data source https://fred.stlouisfed.org

3.4.3 Calibration: instantaneous forward rate

Remark 3.4.2 (how the real world calibrate a zero-coupon curve). [5, p. 164]

- Usually LIBOR rates only available for maturities up to 12 months.
- Eurodollar futures can be used to produce zero-coupon curve up to 5 years.
- Interest rate swap can be then used to extend to 30 years.

Methodology 3.4.1 (forward rate calibration from market). *[1, p. 15]*

Setup:

- *We are given a sequence of co-initial interest rate swaps starting at the spot date t.**They have coincident reset dates but different maturity dates.***
- *Let $T_0, ..., T_{n_i} - 1$ denote reset dates for swap i and let T_{n_i} denote its maturity date.*
- *Let the length of the swaps, i.e $(T_{n_i} - T_0)$ form an increasing sequence $n_1 < n_2 < \cdots$.*
- *Assume the piecewise constant interpolation of the instantaneous forward rate*

$$f(t,T) = \begin{cases} f_1, t \leq T \leq T_{n_1} \\ f_{i+1}, T_{n_i} \leq T \leq T_{n_{i+1}}, i = 1, 2, ... \end{cases}.$$

Procedures to find f_1:

- *For $i = 1$, and $j = 0, ..., n_1$, we have*

$$P(t, T_j) = \exp(-\int_t^{T_j} f(t, u)du) = \exp(-f_1(T_j - t)).$$

From the definition of forward swap rate(Definition 3.1.14), we have

$$\exp(-f_1(T_0 - t)) - \exp(-f_1(T_{n_1} - t)) = S_{0,n_1}(t) \sum_{j=1}^{n_1} \tau_j \exp(-f_1(T_j - t)),$$

where $S_{0,n_1}(t)$ denote the forward swap rate.

- *Therefore, we can solve f_1 using root-find algorithms.*

Procedures to find f_2:

- *For $i = 1$, and $j = n_1, n_1 + 1, ..., n_2$, we have*

$$P(t, T_j) = \exp(-\int_t^{T_j} f(t, u)du) = P(t, T_{n_1}) \exp(-f_2(T_j - T_{n_1})).$$

From the definition of forward swap rate(Definition 3.1.14), we have

$$\begin{aligned} P(t, T_0) - P(t, T_{n_m}) &= S_{0,n_m}(t) \sum_{j=1}^{n_m} \tau_j P(t, T_j) \\ \exp(-f_1(T_0 - t)) - P(t, T_{n_{m-1}}) \exp(-f_m(T_{n_m} - T_{n_{m-1}})) &= S_{0,n_m}(t)(\sum_{j=1}^{n_{m-1}} \tau_j P(t, T_j) \\ &\quad + \sum_{j=n_{m-1}+1}^{n_m} \tau_j P(t, T_j)) \\ P(t, T_{n_{m-1}}) \exp(-f_m(T_{n_m} - T_{n_{m-1}})) &= S_{0,n_m}(t)(\sum_{j=1}^{n_{m-1}} \tau_j P(t, T_j) \\ &\quad + P(t, T_{n_{m-1}}) \sum_{j=n_{m-1}+1}^{n_m} \tau_j \exp(-f_m(T_j - T_{n_{m-}} \end{aligned}$$

where $S_{0,n_2}(t)$ denote the forward swap rate.

- *Therefore, we can solve f_2 using root-find algorithms.*

Procedures to find f_m**:**

- *For* $i = m$*, and* $j = n_m + 1, ..., n_{m+1}$*, we have*

$$P(t, T_j) = \exp(-\int_t^{T_j} f(t,u)du) = P(t, T_{n_{m-1}}) \exp(-f_2(T_j - T_m)).$$

From the definition of forward swap rate(Definition 3.1.14), we have

$$\begin{aligned} P(t,T_0) - P(t,T_{n_2}) &= S_{0,n_2}(t) \sum_{j=1}^{n_2} \tau_j P(t,T_j) \\ \exp(-f_1(T_0 - t)) - P(t,T_{n_1}) \exp(-f_2(T_{n_2} - T_{n_1})) &= S_{0,n_2}(t) (\sum_{j=1}^{n_1} \tau_j P(t,T_j) \\ &\quad + \sum_{j=n_1+1}^{n_2} \tau_j P(t,T_j)) \\ P(t,T_{n_1}) \exp(-f_2(T_{n_2} - T_{n_1})) &= S_{0,n_2}(t) (\sum_{j=1}^{n_1} \tau_j P(t,T_j) + P(t,T_{n_1}) \\ &\quad \sum_{j=n_1+1}^{n_2} \tau_j \exp(-f_2(T_j - T_{n_1}) \end{aligned}$$

where $S_{0,n_2}(t)$ *denote the forward swap rate.*

- *Therefore, we can solve* f_2 *using root-find algorithms.*

3.4.4 Calibration: zero-coupon bond curve

Lemma 3.4.1 (calculate bond price from swap rate bootstrapping). *[1, p. 15] Suppose we have swap rate with maturity dates* $n_1 = 1, n_2 = 2, n_3 = 3, ...$

Then use

$$P(t, T_{n_i}) = \frac{P(t, T_0) - \sum_{j=1}^{n_i - 1} r_i \tau_j P(t, T_j)}{1 + \tau_{n_i} r_i}.$$

to calculate $P(t, n_i)$*.* ***We assume*** $P(t, T_0) = 1$ ***since*** t, T_0 ***are very close***

- ***Procedures to find*** $P(t, 1), n_1 = 1$**:**

$$P(t,1) = \frac{P(t,T_0)}{1 + \tau_{n_1} S_{0,n_1}(t)}.$$

- ***Procedures to find*** $P(t,2), n_2 = 2$**:**

$$P(t,2) = \frac{P(t,T_0) + \sum_{j=1}^{n_2-1} S_{0,n_2}(t)\tau_j P(t,T_j)}{1 + \tau_{n_2} S_{0,n_2}(t)}.$$

- ***Procedures to find*** $P(t,k), n_k = k$**:**

$$P(t,k) = \frac{P(t,T_0) + \sum_{j=1}^{n_k-1} S_{0,n_k}(t)\tau_j P(t,T_j)}{1 + \tau_{n_2} S_{0,n_k}(t)}.$$

Table 3.4.2: OIS curve calibration instruments

Product/quote name	time range
CME Fed Fund Futures	0Y - 2Y
3M LIBOR/Fed fund rate swap	2Y - 50Y
3M LIBOR/SA fixed swap	2Y - 60Y

Table 3.4.3: LIBOR 3M forward curve calibration instruments

Product/quote name	time range
USD money market 3M deposit rate	0 - 3M
CME 3M Eurodollar futures	3M - 3Y
3M LIBOR/SA fixed swap	4Y - 50Y

Table 3.4.4: LIBOR 1M forward curve calibration instruments

Product/quote name	time range
USD money market 1M deposit rate	0 - 1M
USD FRA 1M LIBOR	0 - 3M
3M LIBOR/1M LIBOR basis swap	3M - 30Y

Table 3.4.5: LIBOR 6M forward curve calibration instruments

Product/quote name	time range
USD money market 1M deposit rate	0 - 6M
3M LIBOR/6M LIBOR basis swap	6M - 30Y

3.4.5 Correlation structure from market data

Lemma 3.4.2 (correlation structure in market forward curves). *[4, p. 100] Suppose we have forward curve $f(t, T_i), i = 1, 2, ..., M$ for M maturities(we can obtain these curve from calibration 3.4.1). Denote the sample points in these curves as $f_{n,i} \triangleq f(n\Delta\tau, T_i + n\Delta)$, where $\Delta\tau = 1/12$ represents the interval of one month. Suppose there are N such observations.*

Define the difference quantity

$$\Delta f_{n,i} = f((n+1)\Delta, T_i + (n+1)\Delta) - f(n\Delta, T_i + n\Delta), \forall i = 1, 2, ..., M.$$

We can calculate the following statistical quantities:

- ***(mean drift)**:*

$$\mu(T_i) = \overline{\Delta f_i} = \frac{1}{N}\sum_{n=0}^{N-1} \Delta f_{n,i}.$$

- ***($M \times M$ covariance matrix)**:*

$$C_{ij} = \frac{1}{N}\sum_{n=0}^{N-1} (\Delta f_{n,i} - \overline{\Delta f_i})(\Delta f_{n,j} - \overline{\Delta f_j}).$$

- ***spectral decomposition of covariance matrix***

$$C = \sum_{k=1}^{M} \lambda_k \Delta\tau \boldsymbol{v_k} \boldsymbol{v_k}^T,$$

  where $0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$.
- ***reconstructing forward curve parallel evolution dynamics:***

$$\boldsymbol{f}_{n+1} - \boldsymbol{f}_n = \sum_{k=1}^{M} \sqrt{\lambda_k \Delta\tau} \boldsymbol{v}_k \eta_{k,n},$$

where $\eta_{k,n} \sim N(0,1)$, *and*

$$f_n = \begin{bmatrix} f_{n,1} \\ f_{n,1} \\ \vdots \\ f_{n,M} \end{bmatrix}, \overline{f} = \begin{bmatrix} \overline{f_1} \\ \overline{f_2} \\ \vdots \\ \overline{f_M} \end{bmatrix}.$$

Remark 3.4.3 (observable market data)**.**

- At the beginning, the bond issuer just issue bonds of different maturities. The market prices of these bonds will enable us to construct $f(0,T), \forall T > 0$ via 3.4.1.
- As time goes on, the prices at different t will enable us to trace out $f(t,T), \forall T > 0$.
- The bond issuer will continue to issue new bonds. The market prices of these newly issued bond will enable us to trace out

$$f(\tau, T+\tau), \forall T, \tau > 0.$$

Lemma 3.4.3 (constructing HJM model from market data). *[4, p. 103]*

- ***Parallel shifting dynamics in real probability measure.***

$$df(t, t+T) = \boldsymbol{\sigma}^T d\boldsymbol{W}_t + \mu(T)dt.$$

where $\boldsymbol{\sigma} \in \mathbb{R}^M, \boldsymbol{\sigma}(T) = (v_1(T)\sqrt{\lambda_1}, v_2(T)\sqrt{\lambda_2}, ..., v_M(T)\sqrt{\lambda_M})$

- ***HJM formulation in real probability measure.***

$$df(t,T) = \boldsymbol{\sigma}^T(T-t)d\boldsymbol{W}_t + \mu(T-t)dt.$$

where $\boldsymbol{\sigma} \in \mathbb{R}^M, \boldsymbol{\sigma}(T) = (v_1(T)\sqrt{\lambda_1}, v_2(T)\sqrt{\lambda_2}, ..., v_M(T)\sqrt{\lambda_M})$

- ***HJM formulation in risk-neutral probability measure.***

$$df(t,T) = \boldsymbol{\sigma}^T d\boldsymbol{W}_t + \mu(T)dt.$$

where

$$\mu(t,T) = \boldsymbol{\sigma}^T(T-t)\int_t^T \boldsymbol{\sigma}(s-t)ds$$

Remark 3.4.4 (stationarity assumption)**.** Note that here the correlation structure extraction assumes that the forward curve dynamics is stationary.

3.5 Short rate models: principles and examples

The short-rate model we model the evolution of r_t, as a stochastic process under a risk-neutral measure 0. then the price at time t of a zero-coupon bond maturing at time T with a payoff of 1 is given by where F is the natural filtration for the process. In such framework, the zero-coupon bond and variance interest rate (LIBOR, par swap rate) connected with zero-coupon bond can be derived.

In this section, we will discuss the principles governing a short rate model. Different favors of short rate model including single factor and multiple factors models, will be addressed in the following sections.

In this section, we discuss one class of short rate model, known as equilibrium models. This class of short rate models usually have two to three model parameters; therefore cannot calibrate to the market observed yield curve and admit arbitrage.

We study this class of models is not for practial application but to gain understanding on how short rate models work.

3.5.1 Principles of short-rate model

3.5.1.1 *Understand risk-neutral measure: martingale pricing*

Proposition 3.5.1 (Black model for interest rate derivatives). *[7, p. 513][6, p. 25] Consider the short rate follows a SDE under **real-world** probability measure, given as*

$$dr = u(r,t)dt + \sigma(r,t)dW_t.$$

Then

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial r} + (u - \lambda\sigma)\frac{\partial V}{\partial r} - rV = 0.$$

*There exists an equivalent measure Q,**also known as risk-neutral measure**, under which*

$$dr = (u - \lambda(r,t)\sigma)dt + \sigma dW_t$$

and

$$e^{-\int_0^T r(\tau)d\tau}V(T,r_T)$$

is a martingale under Q.

The pricing formula under risk-neutral measure is given by

$$V(t, r_t) = E_Q[e^{-\int_t^T r(s)ds} V(T, r_T)]$$

Proof. Consider two zero-coupon bonds maturing at T_1 and T_2 with values denoted by $V_1(r, t, T_1)$ and $V_2(r, t, T_2)$. Then

$$\Pi = V_1 - \Delta V_2$$

and

$$d\Pi = \frac{\partial V_1}{\partial t}dt + \frac{\partial V_1}{\partial r}dr + \frac{1}{2}\sigma^2\frac{\partial V_1}{\partial r^2}dr - \Delta(\frac{\partial V_2}{\partial t}dt + \frac{\partial V_2}{\partial r}dr + \frac{1}{2}\sigma^2\frac{\partial^2 V_2}{\partial r^2}dt)$$

then choose

$$\Delta = \frac{\partial V_1}{\partial r} / \frac{\partial V_2}{\partial r}$$

will make $d\Pi$ deterministic.

Then

$$\frac{\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial V_1}{\partial r^2} - rV_1}{\frac{\partial V_1}{\partial r}} = \frac{\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 V_2}{\partial r^2} - rV_2}{\frac{\partial V_2}{\partial r}}$$

Let

$$\frac{\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 + \frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial r^2} - rV}{\frac{\partial V}{\partial r}} = a(r, t).$$

If we write

$$a(r, t) = \sigma(r, t)\lambda(r, t) - u(r, t)$$

for a given $u(r, t)$ and non-zero $\sigma(r, t)$. Then

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial r} + (u - \lambda\sigma)\frac{\partial V}{\partial r} - rV = 0.$$

From the Feyman Kac theorem(Theorem 1.8.1), there exists an equivalent measure Q, also known as risk-neutral measure, under which

$$dr = (u - \lambda(r, t)\sigma)dt + \sigma dW_t$$

and

$$e^{\int_0^T r(\tau)d\tau}P(T)$$

is a martingale under Q. $\square$

Remark 3.5.1. It can be showed that $\lambda(r, t)$ is a proper form; that is

$$\lambda(r, t) = (a(r, t) + u(r, t))/\sigma(r, t)$$

Note 3.5.1 (interpretation).

- In stock option pricing, we can hedge option by traded the underlying stochastic asset. For bond option pricing, the underlying stochastic quantity, i.e., interest rate, is not tradable. Therefore, we have to use bonds of different maturities to hedge the risks.
- **Knowing real-world interest rate dynamics is insufficient for pricing**. We also need to know to λ in order to price. This is because by hedging risks can change the price.
- **market risk has to be fitted from market**. Since λ usually has to be fitted from the market, an alternative method is directly assume the short rate dynamics under risk-neutral measure is given by

$$dr = mdt + \sigma dW_t.$$

  and then fit m from the market. This is the approach in Theorem 3.5.2.
- **Practical difficulties in maringlate pricing formula even we know short rate dynamics under Q**:
  - r_t is not observed, therefore r_T is difficult to compute.
  - **For complex derivatives,** the expectation of $E_Q[e^{-\int_t^T r(s)ds}V(T, r_T)]$ is difficult to evaluate since the joint distribution between $e^{-\int_t^T r(s)ds}$ and $V(T, r_T)$ is unknown.(even though we might be able to evaluate $E_Q[V(T, r_T)]$ and $E_Q[V(T, r_T)]$ separately.)
  - The method is tractable only for simple derivatives, for example, zero coupon bond where $V(T, r_T) = 1$.
- **Monte carlo simulation:** we can certainly use Monte carlo

Remark 3.5.2. Note that similar method has been used in Lemma 2.5.1.

Lemma 3.5.1 (interest rate derivative dynamics under risk-neutral measure). *Consider the short rate follows a SDE under **risk-neutral** probability measure, given by*

$$dr = m(r, t)dt + \sigma(r, t)dW_t.$$

Let V be a derivative on short rate r. Under risk-neutral measure, the dynamics of V is given by

$$dV = rVdt + \frac{\partial V}{\partial r}\sigma dW_t.$$

Proof. From Ito lemma, we have

$$\begin{aligned} dV &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial r}dr + \frac{1}{2}\frac{\partial^2 V}{\partial r^2}(dr)^2 \\ &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial r}dr + \frac{1}{2}\frac{\partial^2 V}{\partial r^2}\sigma^2 dt \end{aligned}$$

From Theorem 3.5.1, we know that $V(r,t)$ has to satisfy

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial r} + (u - \lambda\sigma)\frac{\partial V}{\partial r} - rV = 0.$$

Then the dynamics of V is equivalent to

$$dV = rV + \frac{\partial V}{\partial r}\sigma dW_t.$$

□

3.5.1.2 *Term structure equation*

Proposition 3.5.2 (term structure equation). *[5, p. 707][8, p. 274] Let the short rate r under risk-neutral measure Q follows the general model*

$$dr = m(r,t)dt + s(r,t)dz,$$

where z is a Brownian motion. Then

- *Any derivative with value $f(r(t),t)$ dependent on r and t is governed by*

$$df = (\frac{\partial f}{\partial t} + m\frac{\partial f}{\partial r} + \frac{1}{2}s^2\frac{\partial f}{\partial r^2})dt + s\frac{\partial f}{\partial r}dz.$$

- *Under the risk-neutral measure Q, we have*

$$\frac{\partial f}{\partial t} + m\frac{\partial f}{\partial r} + \frac{1}{2}s^2\frac{\partial f}{\partial r^2} = rf,$$

  with final condition $f(r(T),T) = V(r)$.

Proof. (1) Use Ito's lemma.(**??**).(2) Use Feyman-Kac theorem Theorem 1.8.1. See more details in Theorem 3.5.1. □

Corollary 3.5.2.1 (term-structure equation of a zero-coupon bond). *[8, p. 274][1, p. 43]*
Given the stochastic short rate model under the risk-neutral measure Q

$$dr = m(r,t)dt + s(r,t)dz.$$

It follows that

- *Further assume the price $P(t,T)$ of a zero-coupon bond matures at T is given by $P(t,r(t),T) = f(r(t),t)$. Then $f(r(t),t)$ governed by*

$$\frac{\partial f}{\partial t} + m\frac{\partial f}{\partial r} + \frac{1}{2}s^2\frac{\partial f}{\partial r^2} = rf$$

  with boundary condition $f(r,T) = P(T,r(T),T) = 1, \forall r$.
- *$P(t,T)$ has the following SDE representation:*

$$dP(t,T) = r(t)P(t,T)dt + s(r,t)\frac{\partial P}{\partial r}dW_t.$$

Proof. (1) Direct consequence of Theorem 3.5.2. (2) Using Ito rule,

$$\begin{aligned}
dP(t,T) &= \frac{\partial P}{\partial t}dt + \frac{\partial P}{\partial r}dr + \frac{1}{2}\frac{\partial^2 P(t,T)}{\partial r\partial r}(dr)^2 \\
&= \frac{\partial P}{\partial t}dt + \frac{\partial P}{\partial r}(mdt + sdW_t) + \frac{1}{2}\frac{\partial^2 P(t,T)}{\partial r\partial r}s^2dt \\
&= (\frac{\partial P}{\partial t} + \frac{\partial P}{\partial r}m + \frac{1}{2}\frac{\partial^2 P(t,T)}{\partial r\partial r}s^2)dt + \frac{\partial P}{\partial r}sdW_t \\
&= r(t)P(t,T)dt + \frac{\partial P}{\partial r}sdW_t
\end{aligned}$$

where we use the term structure equation from (1).

□

Remark 3.5.3. The solution will gives the bond price $P(t,r(t),T)$ as a function of short rate at $r(t)$.

Example 3.5.1 (Hull-White interest rate model and zero-coupon bond pricing). [8, p. 274]
In the Hull-White model, the evolution of the interest rate is given by

$$dr(t) = (a(t) - b(t)r(t))dt + \sigma(t)dW(t),$$

where $a(t), b(t)$ and $\sigma(t)$ are deterministic function of time. The governing equation for $P(t,T;r(t)) = f(r(t),t)$ is given by

$$\frac{\partial f}{\partial t} + (a(t) - b(t)r)\frac{\partial f}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial f}{\partial r^2} = rf,$$

with final condition $f(r,T) = 1, \forall r$.

Suppose this PDE has solution

$$P(t,r(t),T) = f(r(t),t) = \exp(-rC(t,T) - A(t,T)),$$

then

$$\begin{aligned}
f_t(t,r) &= (-rC'(t,T) - A'(t,T))f(t,r) \\
f_r(t,r) &= -C(t,T)f(t,r), \\
f_{rr}(t,r) &= C^2(t,T)f(t,r)
\end{aligned}$$

where $C'(t,T) = \frac{\partial}{\partial t}C(t,T)$, $A'(t,T) = \frac{\partial}{\partial t}A(t,T)$.

Substitute into the term structure function we get

$$[(-C'(t,T) + b(t)C(t,T) - 1)r - A'(t,T) - a(t)C(t,T) + \frac{1}{2}\sigma^2(t)C^2(t,T)]f(t,r) = 0.$$

Because this equation holds for all r, then we must have

$$\begin{aligned}
-C'(t,T) + b(t)C(t,T) - 1 = 0 \\
A'(t,T) = -a(t)C(t,T) + \frac{1}{2}\sigma^2(t)C^2(t,T).
\end{aligned}$$

Solve $C(t,T)$ and $A(t,T)$ with boundary condition $C(T,T) = A(T,T) = 0$, we have

$$\begin{aligned}
C(t,T) &= \int_t^T e^{(-\int_t^s b(v)dv)}ds \\
A(t,T) &= \int_t^T (a(s)C(s,T) - \frac{1}{2}\sigma^2(s)C^2(s,T))ds.
\end{aligned}$$

Then, the yield curve is given as

$$Y(t,T) = -\frac{1}{T-t}\log P(t,r(t),T) = \frac{1}{T-t}(rC(t,T) + A(t,T)).$$

That is, $Y(t,T)$ is an **affine function** of r. More deeply, since $r(t)$ is the random variable, $Y(t,T)$ is also a random variable.

Example 3.5.2 (CIR interest rate model and zero-coupon bond pricing). [8, p. 275] In the CIR model, the evolution of the interest rate is given by

$$dr(t) = (a - b(t)r(t))dt + \sigma\sqrt{r(t)}dW(t),$$

where a, b and σ are constants. The governing equation for $P(t,T;r(t)) = f(r(t),t)$ is given by

$$\frac{\partial f}{\partial t} + (a - br)\frac{\partial f}{\partial r} + \frac{1}{2}\sigma^2 r\frac{\partial f}{\partial r^2} = rf,$$

with final condition $f(r,T) = 1, \forall r$.

Suppose this PDE has solution

$$P(t,r(t),T) = f(r(t),t) = \exp(-rC(t,T) - A(t,T)),$$

then

$$\begin{aligned}
f_t(t,r) &= (-rC'(t,T) - A'(t,T))f(t,r)\\
f_r(t,r) &= -C(t,T)f(t,r),\\
f_{rr}(t,r) &= C^2(t,T)f(t,r)
\end{aligned}$$

where $C'(t,T) = \frac{\partial}{\partial t}C(t,T)$, $A'(t,T) = \frac{\partial}{\partial t}A(t,T)$.

Substitute into the term structure function we get

$$[(-C'(t,T) + bC(t,T) - 1 + \frac{1}{2}\sigma^2(t)C^2(t,T))r - A'(t,T) - a(t)C(t,T)]f(t,r) = 0.$$

Because this equation holds for all r, then we must have

$$\begin{aligned}
-C'(t,T) + b(t)C(t,T) - 1 + \frac{1}{2}\sigma^2(t)C^2(t,T) = 0\\
A'(t,T) = -a(t)C(t,T).
\end{aligned}$$

Solve $C(t,T)$ and $A(t,T)$ with boundary condition $C(T,T) = A(T,T) = 0$, we have

$$\begin{aligned}
C(t,T) &= \frac{sinh(\gamma(T-t))}{\gamma cosh(\gamma(T-t)) + \frac{1}{2}bsinh(\gamma(T-t))} \\
A(t,T) &= -\frac{2a}{\sigma^2}\ln(\frac{\gamma\exp(\frac{1}{2}b(T-t))}{\gamma cosh(\gamma(T-t)) + \frac{1}{2}bsinh(\gamma(T-t))}) \\
\gamma &= \frac{1}{2}\sqrt{b^2+2\sigma^2}.
\end{aligned}$$

Remark 3.5.4 (practical difficulties). In the short model method, we can find the yield curve and bond price to be a function of short rate $r(t)$. However, the short rate $r(t)$ cannot be directly observed in the market.

3.5.1.3 *Instruments related to short rates*

Definition 3.5.1 (money market account). *Let the cash account be*

$$M(t) = M(0)\exp(\int_0^t r(s)ds).$$

Lemma 3.5.2 (zero-coupon bond pricing). *Suppose the risk-free rate $r(t)$ is stochastic. Assume no-arbitrage condition holds and the market is complete.*

- *There exists a measure Q, equivalent to P, under which, for each T, the discounted price process $P(t,T)/B(t)$ is a martingale for all $t : 0 < t < T$.*
- $$P(t,T) = B(t)E_Q[\exp(-\int_0^T r(s)ds)P(T,T)|\mathcal{F}_t] = E_Q[\exp(-\int_t^T r(s)ds)|\mathcal{F}_t].$$

Proof. Use the fact that $P(t,T)/B(t)$ is a martingale. Also note that we use the fact that $P(T,T) = 1, B(t)$ is $\mathcal{F}_t$ measurable. □

Remark 3.5.5 (no-arbitrage restrictions for short rate model under risk-neutral measure).

- Unlike dynamic models for tradable assets (e.g., stocks), the dynamic model should satisfy that discounted asset prices are martingales under the equivalent martingale measure. There is almost no restriction on coefficients in a typical short rate model under risk-neutral measure given by

$$dr(t) = \alpha(t)dt + \sigma(t)dW_t.$$

- The discounted zero-coupon bond price is given by

$$\frac{P(t,T)}{M(t)} = E_Q[\exp(-\int_0^T r(s)ds)|\mathcal{F}_t],$$

which is a martingale under risk-neutral measure since $\exp(-\int_0^T r(s)ds)$ is a fixed random variable (using conditional expectation and **??**).

Lemma 3.5.3 (SDE of bond price under risk-neutral measure). *Assume, under real-world measure, the SDE for the price of zero-coupon bond $P(t,T)$ is given by*

$$dP(t,T) = \mu(t,P)P(t,T)dt + \sigma(t,P)P(t,T)dW(t)$$

*where $W(t)$ is the Brownian motion. Let $r(t)$ be the (stochastic)short rate. We further **assume there is no arbitrage opportunities**. Then there exists a risk-neutral measure(i.e. use money account as the numeraire) such that*

$$dP(t,T) = r(t)P(t,T)dt + \sigma(t,P)P(t,T)d\hat{W}(t)$$

where $\hat{W}(t)$ is the Brownian motion under the risk-neutral measure.

Proof. Under no-arbitrage condition(Theorem 1.5.1), we can write the dynamics of bond price and money account as

$$\begin{aligned} dP(t,T) &= (r + \lambda_p\sigma(t,P))P(t,T)dt + \sigma(t,P)P(t,T)dW(t) \\ dB(t) &= rB(t,T)dt \end{aligned}$$

When using money account as numeraire, we should set $\lambda_P = 0$(Theorem 1.6.19). Therefore, under the risk-neutral measure, we have

$$dP(t,T) = r(t)P(t,T)dt + \sigma(t,P)P(t,T)d\hat{W}(t)$$

where $\hat{W}(t)$ is the Brownian motion under the risk-neutral measure. □

3.5.2 Merton model

3.5.2.1 *Basics*

Definition 3.5.2 (Merton model). *[1, p. 44] In Merton model, the SDE for the short rate **under risk-neutral measure** is*

$$dr(t) = \alpha dt + \sigma dW(t),$$

where α and σ are constants, and $W(t)$ is a Brownian motion under the risk-neutral measure Q.

Lemma 3.5.4 (zero-coupon bond price in Merton model). *[1, p. 45] Consider the short rate under risk neutral measure is evolving under the Merton model with drift α and volatility σ. We have*

- *The solution for the short rate is given by*

$$r(s) = r(t) + \alpha(s-t) + \sigma(W(s) - W(t)).$$

  And

$$r(s) \sim N(r(t) + \alpha(s-t), \sigma^2(s-t)), s > t.$$

- $$\int_t^T r(s) = r(t)(T-t) + \frac{1}{2}\alpha(T-t)^2 + \sigma \int_t^T (T-s)dW(s),$$

  which is the Gaussian process with mean

$$m(t) = r(t)(T-t) + \frac{1}{2}\alpha(T-t)^2$$

  and variance

$$s^2 = \frac{1}{3}\sigma^2(T-t)^3$$

  under risk-neutral measure.
- *The zero-coupon bond price in the Merton model is given as*

$$P(t,T) = \exp((-r(t)(T-t) - \frac{1}{2}\alpha(T-t)^2 + \frac{1}{6}\sigma^2(T-t)^3),$$

  and $P(T,T) = 1$.
  Define $A(t,T) = \exp(-\frac{1}{2}\alpha(T-t)^2 + \frac{1}{6}\sigma^2(T-t)^3), B(t,T) = T-t$. Then we can write

$$P(t,T) = A(t,T)\exp(-B(t,T)r(t)).$$

- *The yield curve $y(t,T)$ in the Merton model is given by*

$$y(t,T) \triangleq -\frac{\log P(t,T)}{T-t} = r(t) + \frac{1}{2}\alpha(T-t) - \frac{1}{6}\sigma^2(T-t)^2.$$

- $$dP(t,T) = r(t)P(t,T)dt - (T-t)P(t,T)\sigma dW(t)$$

Proof. (1) See **??**.

(2) Since $r(t)$ is $\mathcal{F}_t$ measurable and X is independent of $\mathcal{F}_t$, We have

$$\begin{aligned}P(t,T) &= E_Q[\exp(-\int_t^T r(s)ds)|\mathcal{F}_t] \\ &= E_Q[\exp(X)] \\ &= M_X(-1) = \exp(-m + \frac{1}{2}s^2)\end{aligned}$$

where $X(t)$ is the Gaussian random variable $\exp(-\int_0^T r(s)ds)$, and M_X is the moment generating function of X (3) straight forward. (4) Use

$$dP = \frac{\partial P}{\partial t} + \frac{\partial P}{\partial r}dr + \frac{\partial^2 P}{\partial r^2}drdr$$

and

$$\begin{aligned}\frac{\partial P}{\partial t} &= (r + \alpha(T-r) - \frac{1}{2}\sigma^2(T-t)^2)P \\ \frac{\partial P}{\partial r} &= -(T-t)P \\ \frac{\partial^2 P}{\partial r^2} &= (T-t)^2 P\end{aligned}$$

□

Note 3.5.2 (equivalence to PDE method). It can be showed that that solution

$$P(t,r,T) = \exp(-r(T-t) - \frac{1}{2}\alpha(T-t)^2 + \frac{1}{6}\sigma^2(T-t)^3)$$

also satisfy the term-structure equation

$$\frac{\partial P}{\partial t} + \alpha\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 P}{\partial r^2} = rP,$$

given by Theorem 3.5.2. Note that

$$\begin{aligned}\frac{\partial P}{\partial t} &= (r + \alpha(T-r) - \frac{1}{2}\sigma^2(T-t)^2)P \\ \frac{\partial P}{\partial r} &= -(T-t)P \\ \frac{\partial^2 P}{\partial r^2} &= (T-t)^2 P\end{aligned}$$

Then

$$\frac{\partial P}{\partial t} + \alpha \frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2 \frac{\partial^2 P}{\partial r^2} = rP.$$

Remark 3.5.6 (risk-neutral process is not a real world process). Note that the model is to model how the short rate will move under risk-neutral probability measure instead of the real-world probability measure. Why we do this is discussed at Theorem 3.5.1.

Remark 3.5.7 (shortcomings of Merton model).

- One problem is that short rate can be negative.
- The model lacks mean-reversion property.
- Not Arbitrage free.

3.5.2.2 *Yield curve in Merton model*

Remark 3.5.8 (construction of the yield curve and model calibration). Note that the yield curve can be constructed by

$$R(t,T) = -\frac{\log P(t,T)}{T-t} = (r(t) + \frac{1}{2}\alpha(T-t) - \frac{1}{6}\sigma^2(T-t)^2).$$

We have the following interpretation

- we can calibrate the yield curve $R(0,T)$ to market yield curve and use this model to predict future yield curve.
- larger volatility in the interest rate will decreases yield; particularly, it will decrease the value of long-term yield more than the short-term.
- **parallel shift property contradicts market observation:** because α, σ is constant, shift t and T at the same time will not change yield curve. However, usually in the market $R(t+\tau, T+\tau)$ will change(that is, the price of newly issue bonds with the same maturities will be different).
- **lack of flexibility:** this model only has two parameters and usually it is difficult to fit the market yield curve. Mathematically,

$$\frac{dP(t+s, T+s)}{ds} = \frac{\partial P}{\partial t}ds + \frac{\partial P}{\partial T}ds = 0.$$

- **possible arbitrage opportunities:** We usually assume the market contains no arbitrages. Since the model usually cannot fit to the current market term structure, it may contains arbitrage opportunities.

Remark 3.5.9 (shape control of yield curve). Let the current time be 0. A yield curve $y(0,T)$ is characterized by three parameters: $r(0), \alpha, \sigma$. As showed in Figure 3.5.1,

- $r(0)$ controls the overall level.
- α controls the up slopping.
- σ controls the bending down. The larger volatility will tend to bend down the yield curve.

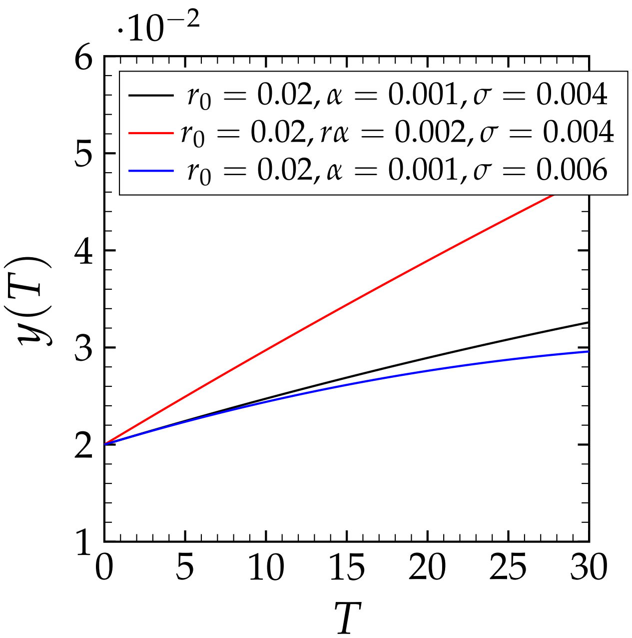


Figure 3.5.1: Current yield curve resulting from Merton's model with different parameters.

Lemma 3.5.5 (evolution of yield curves). *Consider the short rate under risk neutral measure is evolving under the Merton model with drift α and volatility σ.*

- ***(evolution of whole yield curve)*** *Let the current time be 0, then the future yield curve $y(\tau, T+\tau)$ is given by*

$$\begin{aligned} y(\tau, T+\tau) &= r(\tau) + \frac{1}{2}\alpha T + \frac{1}{6}\sigma^2 T^2 \\ &= \alpha\tau + \sigma W(\tau) + y(0, T) \end{aligned}$$

  where $y(0, T)$ is the current yield curve.
- ***evolution of yield associated with a fixed maturity date*** *Let the current time be 0, then the future yield $y(\tau, T)$ is given by*

$$\begin{aligned} y(\tau, T) &= r(\tau) + \frac{1}{2}(T-\tau) + \frac{1}{6}\sigma^2(T-\tau)^2 \\ &= r(0) + \alpha\tau + \sigma W(\tau) + \frac{1}{2}(T-\tau) + \frac{1}{6}\sigma^2(T-\tau)^2 \end{aligned}$$

Particularly, $y(T,T) = r(0) + \alpha T + \sigma W(T)$.

Proof. □

Methodology 3.5.1 (simulate the future term structure in Merton model). *Suppose*

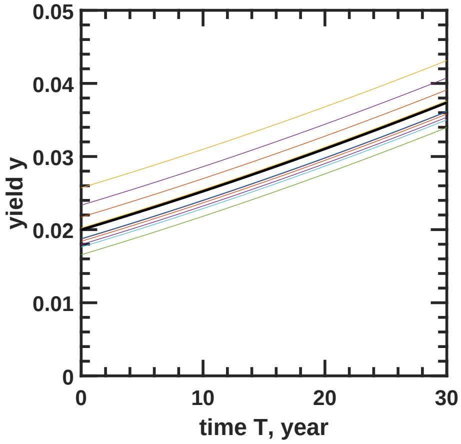
- *current time is 0.*
- *we are given the current term structure* $y(0,T), T \in \mathbb{R}^+$.
- *we are given the Merton model*
$$dr = \alpha dt + \sigma dW_t.$$
- *we are given the initial short rate* $r(0)$.

Then we can generate a ***sample term structure*** *in future time* $t > 0$ *via the following procedure:*

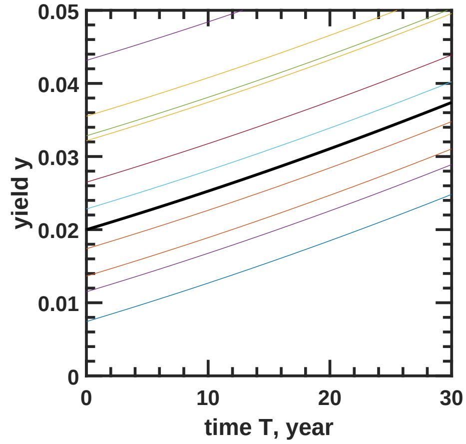
- *simulate* $r(t)$ *by drawing from the normal distribution*
$$r(t) \sim N(r(0) + \alpha t, \sigma^2 t).$$
- *the sample yield curve is given by*
$$y(t,T) = r(t) + \frac{1}{2}(T-t) + \frac{1}{6}\sigma^2(T-t)^2.$$

Remark 3.5.10. As showed in Figure 3.5.2, the initial yield curve is given by
$$y(0,T) = r_0 + \frac{1}{2}\alpha T + \frac{1}{6}\sigma^2 T^2,$$
where $r_0 = 0.02, \alpha = 0.001, \sigma = 0.004$.

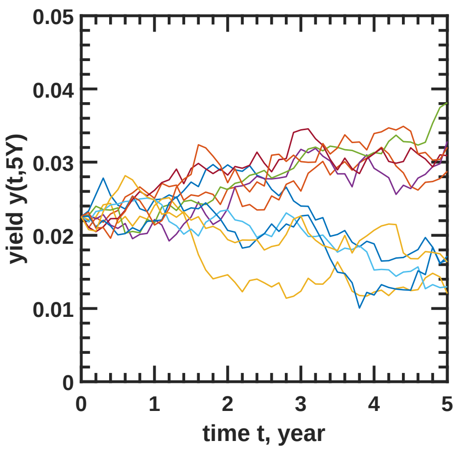
- Figure A shows the different realization of the yield curve $y(1, T+1)$ after 1 year. We can see that only parallel dynamics of yield curve is allowed in the Merton model. There is no mean-reversion in these yield curve realizations..
- Figure B shows the different realization of the yield curve $y(1, T+1)$ after 10 year. Compared to the 1 year's realizations, the yield curve can move far away from the initial yield curve.
- Figure C and D shows the 5Y and 10Y zero rate evolution realizations in the Merton model. These trajectories can also be used to calculate the price trajectories of a zero coupon bond maturing in 5Y and 10Y via
$$P(t,T) = \exp(-y(t,T)(T-t)),$$
as we showed in Figure E and F.



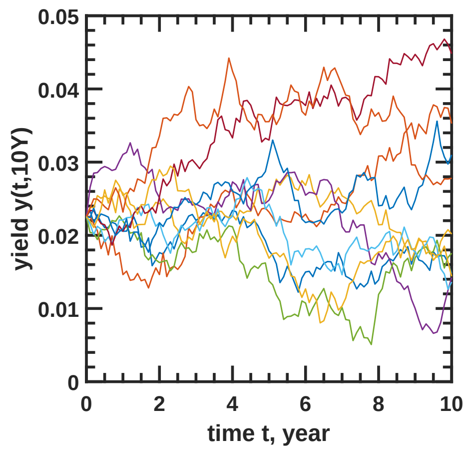
(a) Realizations of yield curve after 1Y (thin lines) from the initial yield curve(thick line).



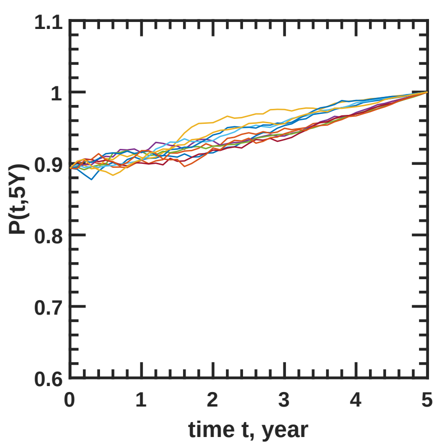
(b) Realizations of yield curve after 5Y (thin lines) from the initial yield curve(thick line).



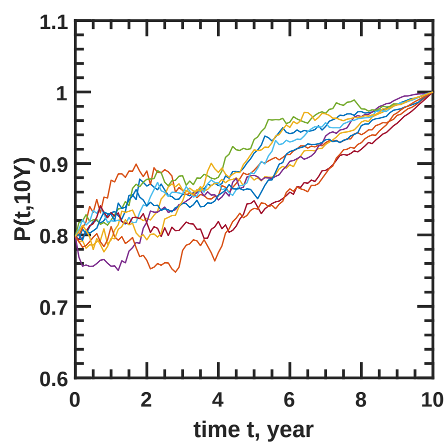
(c) Evolution of yield $y(t, 5Y)$ from the initial value $y(0, 5Y)$.



(d) Evolution of yield $y(t, 10Y)$ from the initial value $y(0, 10Y)$.



(e) 5Y zero-coupon bond price trajectories.



(f) 10Y zero-coupon bond price trajectories.

Figure 3.5.2: Yield curve dynamics in Merton model

3.5.2.3 *Instantaneous forward rate dynamics*

Lemma 3.5.6 (instantaneous forward rate dynamics in Merton model). *Consider the short rate under risk neutral measure is evolving under the Merton model with drift α and volatility σ.*

- *The instantaneous forward rate is given by*

$$f(t,T) = r(t) + \alpha(T-t)^2 - \frac{1}{2}\sigma^2(T-t)^2.$$

- *The dynamics of instantaneous forward rate is given by*

$$df(t,T) = \sigma^2(T-t)dt + \sigma dW_t$$

Proof. (1) Note that in the Merton model, we have zero-coupon bond price given by

$$P(t,T) = \exp((-r(T-t) - \frac{1}{2}\alpha(T-t)^2 + \frac{1}{6}\sigma^2(T-t)^3),$$

and the instantaneous forward rate is given by

$$f(t,T) = -\frac{\ln P(t,T)}{dT} = r(t) + \alpha(T-t) - \frac{1}{2}\sigma^2(T-t)^2.$$

(2)

$$\begin{aligned} df(t,T) &= dr(t) - \alpha dt + \sigma^2(T-t)dt \\ &= \alpha dt + \sigma dW_t - \alpha dt + \sigma^2(T-t)dt \\ &= \sigma^2(T-t)dt + \sigma dW_t \end{aligned}$$

□

3.5.2.4 *Derivative pricing*

Lemma 3.5.7 (zero-coupon bond price and forward price SDE in Merton model). *Given the short rate SDE under risk-neutral measure*

$$dr(t) = \alpha dt + \sigma dW(t),$$

the zero-coupon bond price $P(t,T)$ satisfies SDE

$$dP(t,T) = P(t,T)(r(t) + \alpha(T-t) - \frac{\sigma^2}{2}(T-t)^2)dt - P(t,T)(T-t)\sigma dW(t).$$

Moreover, the forward bond price $FP(t;S,T) \triangleq \frac{P(t,T)}{P(t,S)}$ *satisfies SDE*

$$dFP(t;S,T) = FP(t;S,T)\sigma(T-S)dW^S(t),$$

where W^S *is a Brownian motion under forward measure with respect to* $P(t,S)$.

Proof. (1) From Ito lemma, we have

$$dP(t,r,T) = \frac{\partial P}{\partial t}dt + \frac{\partial P}{\partial r}dr + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}drdr.$$

Note that

$$\begin{aligned}
\frac{\partial P}{\partial t} &= P(t,T)(r(t) + \alpha(T-t) - \frac{\sigma^2}{2}(T-t)^2) \\
\frac{\partial P}{\partial r}dr &= -P(t,T)(T-t)dr \\
\frac{\partial^2 P}{\partial r^2}drdr &= P(t,T)(T-t)^2\sigma^2 dt
\end{aligned}$$

then

$$dP(t,T) = P(t,T)r(t)dt - P(t,T)(T-t)\sigma dW(t).$$

(2) use Theorem 3.3.1, we have

$$dFP(t;S,T) = FP(t;S,T)(\sigma_{P(t,T)} - \sigma_{P(t,S)})dW^S(t).$$

□

Lemma 3.5.8 (Price of European call option on zero-coupon bond). *The value at time t of a European call option maturing at S with strike K on a bond maturing at* $T > S$ *is given by*

$$V(t) = P(t,S)E_S[\max(P(S,T)-K,0)|\mathcal{F}_t],$$

where the expectation is taken with respect to forward measure of $P(t,S)$. *It can be showed that*

$$V(t) = PC(t;S,T,K) = P(t,T)N(d_+) - KP(t,S)N(d_-),$$

where

$$d_+ = \frac{\log\frac{FP(t;S,T)}{K} + \frac{1}{2}v(t,S)^2}{v(t,S)} = \frac{\log\frac{B(t,T)}{KB(t,S)} + \frac{1}{2}v(t,S)^2}{v(t,S)}, d_- = d_+ - v(t,S).$$

and

$$FP(t;S,T) \triangleq P(t,T)/P(t,S)$$

$$v(t,S)^2 = \int_t^S (\sigma(T-S))^2 dt = (\sigma(T-S))^2(S-t)$$

Proof. Note that our bond dynamics is given by

$$dP(t,T) = P(t,T)r(t)dt - P(t,T)(T-t)\sigma dW(t).$$

Then, we can use Lemma 3.3.10 □

Remark 3.5.11 (general discussion). For a general discussion on pricing bond options based on bond dynamics, see Lemma 3.3.10.

3.5.3 Vasicek model

3.5.3.1 *Basics*

Definition 3.5.3 (Vasicek model). *[5, p. 684][1, p. 45] In Vasicek model, the SDE for r under risk-neutral measure Q*

$$dr(t) = (\theta - \alpha r(t))dt + \sigma dW_t$$

where W_t is the Brownian motion under risk-neutral measure, θ, α, σ are constants.

Remark 3.5.12 (general characteristics).

- The model has mean reversion to θ.
- One drawback is that negative risk-free rate can occur.

Lemma 3.5.9 (short rate solution in the Vasicek model). *Assume the short rate, under risk-neutral measure, following Vasicek SDE*

$$dr(t) = (\theta - \alpha r(t))dt + \sigma dW(t).$$

Let $0 \leq t \leq s \leq T$. It follows that

- *The short rate solution in the Vasicek SDE is given by*

$$r(s) = r(t)\exp(-\alpha(s-t)) + \frac{\theta}{\alpha}(1-\exp(-\alpha(s-t))) + \sigma\int_t^s \exp(-\alpha(s-u))dW(u).$$

- ***Conditioning on** $\mathcal{F}_t$, we have*

$$E[r(s)|\mathcal{F}_t] = r(t)\exp(-\alpha(s-t)) + \frac{\theta}{\alpha}(1-\exp(-\alpha(s-t))),$$

$$Var[r(s)|\mathcal{F}_t] = \frac{\sigma^2}{2\alpha}(1 - \exp(-2\alpha(s-t))).,$$
$$r(s)|\mathcal{F}_t \sim N(E[r(s)|\mathcal{F}_t], Var[r(s)|\mathcal{F}_t]).$$

Proof. See the OU process results in **??**. □

Lemma 3.5.10 (zero-coupon bond price in Vasicek model). *[1, p. 48] Assume the short rate, under risk-neutral measure, following SDE*

$$dr(t) = (\theta - \alpha r(t))dt + \sigma dW(t).$$

Denote

$$D(t,T) = \int_t^T e^{-\alpha(s-t)} ds.$$

We have

- *Let $s \geq t$, we have*

$$r(s) = r(t)\exp(-\alpha(s-t)) + \frac{\theta}{\alpha}(1 - \exp(-\alpha(s-t))) + \sigma \int_t^s \exp(-\alpha(s-u))dW(u).$$

- *Let $I(t,T) \triangleq \int_t^T r(s)ds$. Then conditioning on $\mathcal{F}_t$, $I(t,T)$ has Gaussian distribution with mean $M(t,T)$ and variance $V(t,T)$ given by*

$$M(t,T) = r(t)\frac{1 - \exp(-\alpha(T-t))}{\alpha} + \frac{\theta}{\alpha}(T-t) - \frac{\theta}{\alpha}\frac{1 - \exp(-\alpha(T-t))}{\alpha}$$

$$V(t,T) = \frac{\sigma^2}{\alpha^2}(T - t + \frac{2}{\alpha}\exp(-\alpha(T-t)) - \frac{1}{2\alpha}\exp(-2\alpha(T-t)) - \frac{3}{2\alpha})$$

- *The zero-coupon bond price is given by*

$$P(t,T) = E_Q[\exp(-I(t,T))|\mathcal{F}_t] = \exp(-M(t,T) + \frac{1}{2}V(t,T)).$$

Define

$$B(t,T) = \frac{1 - \exp(-\alpha(T-t))}{\alpha},$$

$$A(t,T) = \exp((\frac{\theta}{\alpha} - \frac{\sigma^2}{2\alpha^2})(B(t,T) - (T-t)) - \frac{\sigma^2}{4\alpha}B^2(t,T))),$$

then

$$P(t,T) = A(t,T)\exp(-r(t)B(t,T)).$$

- *The yield curve is given by*

$$y(t,T) = -\frac{\ln P(t,T)}{T-t} = r(t)\frac{B(t,T)}{T-t} - \frac{\ln A(t,T)}{T-t}.$$

Proof. (1) From above lemma. (2)(a)Note that given the observation $r(t)$ at t, we have

$$r(s) = r(t)\exp(-\alpha(s-t)) + \frac{\theta}{\alpha}(1-\exp(-\alpha(s-t))) + \sigma\int_t^s \exp(-\alpha(s-u))dW(u).$$

Therefore,

$$\begin{aligned}
&\int_t^T r(s)ds \\
&= \int_t^T r(t)e^{-\alpha(s-t)}ds + \frac{\theta}{\alpha}\int_t^T (1-e^{-\alpha(s-t)})ds + \int_t^T\int_t^s \sigma e^{-\sigma(s-u)}dW(u)ds \\
&= r(t)\frac{1-e^{-\alpha(T-t)}}{\alpha} + \frac{\theta}{\alpha}(T-t) - \frac{\theta}{\alpha}\frac{1-e^{-\alpha(T-t)}}{\alpha} + \sigma\int_t^T\int_s^T e^{-\alpha(s-u)}dudW(s) \\
&= r(t)\frac{1-e^{-\alpha(T-t)}}{\alpha} + \frac{\theta}{\alpha}(T-t) - \frac{\theta}{\alpha}\frac{1-e^{-\alpha(T-t)}}{\alpha} + \int_t^T \frac{\sigma}{\alpha}(1-e^{-\alpha(T-s)})dW(s)
\end{aligned}$$

where we changed the order of integration. From this, we note that

$$E[\int_t^T x_1(u)du] = r(t)\frac{1-\exp(-\alpha(T-t))}{\alpha} + \frac{\theta}{\alpha}(T-t) - \frac{\theta}{\alpha}\frac{1-\exp(-\alpha(T-t))}{\alpha}.$$

(b) For $Var[\int_t^T r(s)ds]$, we have

$$\begin{aligned}
&Var[\int_t^T r(s)ds \\
&= E[\int_t^T \frac{\sigma}{\alpha}(1-\exp(-\alpha(T-s)))dW(s)\int_t^T \frac{\sigma}{\alpha}(1-\exp(-\alpha(T-s)))dW(s)] \\
&= \frac{\sigma^2}{\alpha^2}(\int_t^T ds + \int_t^T \exp(-2\alpha(T-s))ds - 2\int_t^T \exp(-\alpha(T-s))ds) \\
&= \frac{\sigma^2}{\alpha^2}(T-t+\frac{2}{\alpha}\exp(-\alpha(T-t)) - \frac{1}{2\alpha}\exp(-2\alpha(T-t)) - \frac{3}{2\alpha}).
\end{aligned}$$

(3) If we define $B(t,T) = \frac{1-exp(-\alpha(T-t))}{\alpha}$,then

$$\begin{aligned}
M(t,T) &= r(t)B(t,T) + \frac{\theta}{\alpha}(T-t) - \frac{\theta}{\alpha}B(t,T), \\
V(t,T) &= \frac{\sigma^2}{\alpha^2}(T-t) - \frac{\sigma^2}{\alpha^2}B(t,T) - \frac{\sigma^2}{2k}B^2(t,T).
\end{aligned}$$

□

Remark 3.5.13 (applications to price zero-coupon bond). Use Vasicek model, the stochastic short rate model is given as

$$dr = a(b-r)dt + \sigma dz,$$

the price $P(t,T)$ of a zero-coupon bond matures at T is governed by

$$\frac{\partial P}{\partial t} + a(b-r)\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial P}{\partial r^2} = rP$$

with boundary condition $P(T,T) = 1$.

Remark 3.5.14 (model calibration to market). Suppose we are given prices of N zero-coupon bond with different maturities at $t = 0$. Then, we can use nonlinear least square to solve model parameters θ, α, σ such that the error between the theoretical prediction and market observation is minimized.

Note that because we only have 3 parameters, it is usually difficult to match with the term structure of the market.

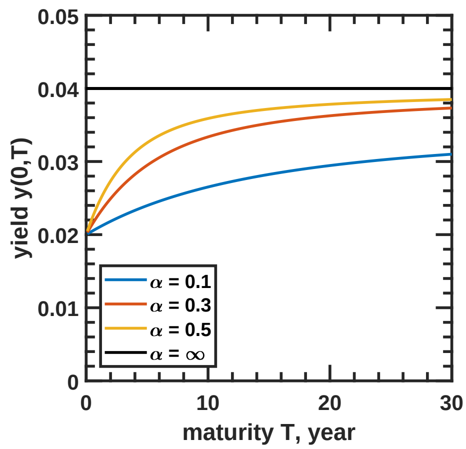
3.5.3.2 *Yield curve dynamics*

Remark 3.5.15 (yield curve shape). From Figure 3.5.3, we can see that

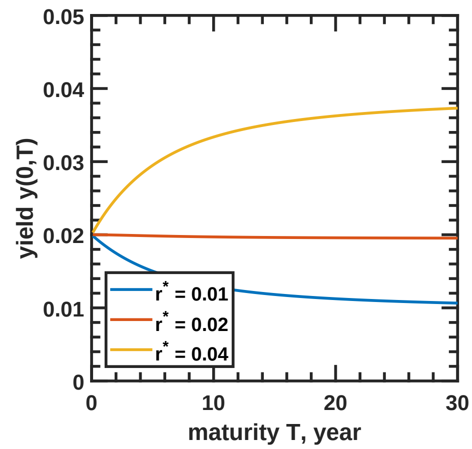
- (reversion speed effect) If $\alpha \to \infty$(but θ/α remains to be a constant), we can see that

$$B(t,T) = 0, A(t,T) = \exp(-\frac{\theta}{\alpha})(T-t)) = \exp(-r^*(T-t)), y(t,T) = r^*.$$

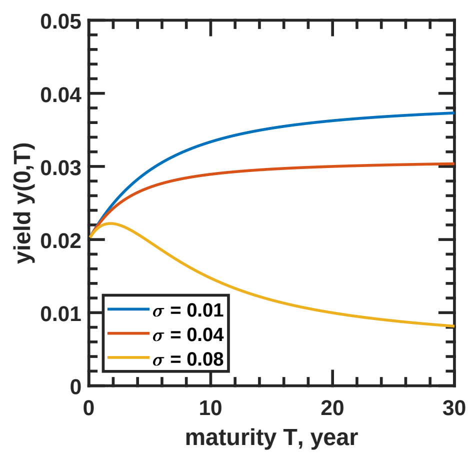
  If reversion speed α increases, then the yield curve will more resemble the equilibrium flat yield curve given by $y(t,T)* = r^*$.
- (mean level effect) The mean level will tend to pull the far-end of the yield curve toward the mean level.
- (volatility level effect) The volatility level tend to pull the far-end of the yield curve down.



(a) Yield curve shape vs. reversion speed parameter.



(b) Yield curve shape vs. mean level $r^* = \frac{\theta}{\alpha}$.



(c) Yield curve shape vs. volatility.

Figure 3.5.3: Control yield curve shape via Vasicek model parameters.

Lemma 3.5.11 (evolution of yield curves). *Let the current time be 0. Consider the short rate under risk neutral measure is evolving under the Vasicek model with drift α and volatility σ such that*

$$dr(t) = (\theta - \alpha r(t))dt + \sigma dW(t).$$

The yield curve is given by

$$y(0,T) = r(0)\frac{B(0,T)}{T} - \frac{\ln A(0,T)}{T}.$$

where

$$B(t,T) = \frac{1 - \exp(-\alpha(T-t))}{\alpha},$$

$$A(t,T) = \exp\left(\left(\frac{\theta}{\alpha} - \frac{\sigma^2}{2\alpha^2}\right)(B(t,T) - (T-t)) - \frac{\sigma^2}{4\alpha}B^2(t,T)\right),$$

- ***(evolution of whole yield curve)*** *The future yield curve $y(\tau, T+\tau)$ is given by*

$$\begin{aligned} y(\tau, T+\tau) &= r(\tau)\frac{B(\tau, T+\tau)}{T} - \frac{\ln A(\tau, T+\tau)}{T} \\ &= r(\tau)\frac{B(0,T)}{T} - \frac{\ln A(0,T)}{T}. \end{aligned}$$

- ***evolution of yield associated with a fixed maturity date*** *The future yield $y(\tau, T)$ is given by*

$$y(\tau, T) = r(\tau)r(\tau)\frac{B(\tau, T+\tau)}{T} - \frac{\ln A(\tau, T+\tau)}{T}$$

where

$$r(\tau) = r(0)\exp(-\alpha\tau) + \frac{\theta}{\alpha}(1 - \exp(-\alpha\tau)) + \sigma\int_0^\tau \exp(-\alpha(\tau - u))dW(u).$$

Methodology 3.5.2 (simulate the future term structure in Vasicek model). *Suppose*

- *current time is 0.*
- *we are given the current term structure $y(0,T), T \in \mathbb{R}^+$.*
- *we are given a Vasicek model*

$$dr(t) = (\theta - \alpha r(t))dt + \sigma dW_t$$

- *we are given the initial short rate $r(0)$*

Then we can generate a ***sample term structure*** *in future time $t > 0$ via the following procedure:*

- *simulate $r(t)$ by drawing from the normal distribution*

$$r(t) \sim N(m(t), v^2(t)),$$

where

$$m(t) = r(0)\exp(-\alpha t) + \frac{\theta}{\alpha}(1 - \exp(-\alpha t)),$$

$$v^2(t) = \frac{\sigma^2}{2\alpha}(1 - \exp(-2\alpha t)).$$

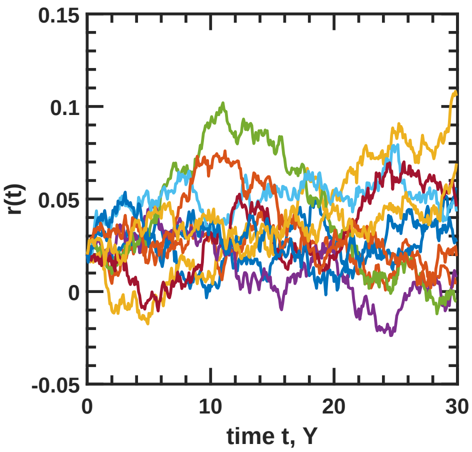
- *the sample yield curve is given by*

$$y(t,T) = r(t)\frac{B(t,T)}{T-t} - \frac{\ln A(t,T)}{T-t}$$

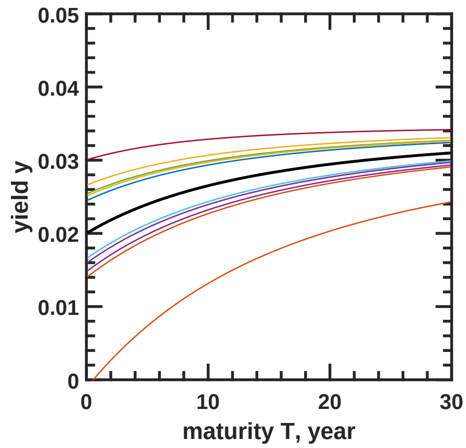
where

$$B(t,T) = \frac{1-\exp(-\alpha(T-t))}{\alpha},$$

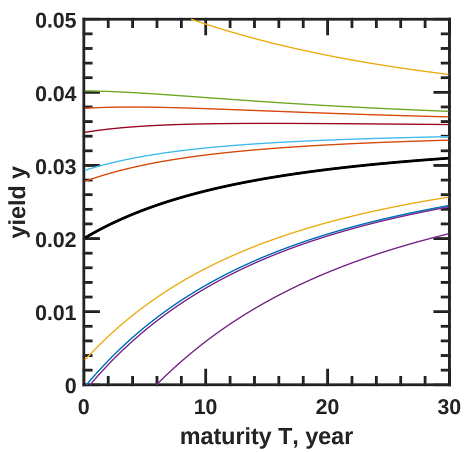
$$A(t,T) = \exp\left(\left(\frac{\theta}{\alpha} - \frac{\sigma^2}{2\alpha^2}\right)(B(t,T)-(T-t)) - \frac{\sigma^2}{4\alpha}B^2(t,T)\right).$$



(a) A. Realizations of short rate trajectories.



(b) A. Realizations of yield curve after 1Y (thin lines) from the initial yield curve(thick line).



(c) B. Realizations of yield curve after 5Y (thin lines) from the initial yield curve(thick line).

Figure 3.5.4: Yield curve dynamics in Vasicek model

3.5.3.3 *Instantaneous forward rate dynamics*

Lemma 3.5.12 (instantaneous forward rate dynamics in Vasicek model). *Consider the short rate under risk neutral measure is evolving under the Vasicek model given by*

$$dr(t) = (\theta - kr(t))dt + \sigma dW_t.$$

It follows that

- *The instantaneous forward rate is given by*

$$f(t,T) = r(t)\exp(-k(T-t)) + \frac{\theta}{k}(1-\exp(-k(T-t))) - \frac{\sigma^2}{2k^2}(1-\exp(-k(T-t)))^2.$$

- *The dynamics of instantaneous forward rate is given by*

$$df(t,T) = \frac{\sigma^2}{k}(1-\exp(-k(T-t)))\exp(-k(T-t))dt + \sigma\exp(-k(T-t))dW_t.$$

Proof. (1) Note that in the Vasicek model, we have zero-coupon bond price given by

$$P(t,T) = \exp((-r(T-t) - \frac{1}{2}\alpha(T-t)^2 + \frac{1}{6}\sigma^2(T-t)^3),$$

and the instantaneous forward rate is given by

$$f(t,T) = -\frac{\ln P(t,T)}{dT} = r(t) + \alpha(T-t) - \frac{1}{2}\sigma^2(T-t)^2.$$

(2)

$$\begin{aligned} df(t,T) &= \exp(-k(T-t))dr(t) + br_t\exp(-k(T-t))dt - \theta\exp(-k(T-t)) \\ &+ \frac{\sigma^2}{k}(1-\exp(-k(T-t)))\exp(-k(T-t))dt \\ &= \exp(-k(T-t))((\theta-\alpha r(t))dt + \sigma dW_t) + br_t\exp(-k(T-t))dt - \theta\exp(-k(T-t)) \\ &+ \frac{\sigma^2}{k}(1-\exp(-k(T-t)))\exp(-k(T-t))dt \\ &= \frac{\sigma^2}{k}(1-\exp(-k(T-t)))\exp(-k(T-t))dt + \sigma\exp(-k(T-t))dW_t \end{aligned}$$

□

3.5.3.4 *Derivative pricing*

Lemma 3.5.13 (zero-coupon bond price and forward price SDE in Vasicek model). *Given the short rate SDE under risk-neutral measure*

$$dr(t) = (\theta - \alpha r(t))dt + \sigma dW(t),$$

we have

- *the zero-coupon bond price* $P(t, r(t), T)$ *satisfies*

$$\frac{\partial P}{\partial t} + (\theta(t) - \alpha r(t))\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 P}{\partial r^2} = rP.$$

- *the zero-coupon bond price* $P(t, T)$ *satisfies SDE*

$$dP(t, T) = P(t, T)r(t)dt - P(t, T)D(t, T)\sigma dW(t).$$

where

$$D(t, T) = \int_t^T e^{-\alpha(s-t)}ds.$$

- *Moreover, the forward bond price* $FP(t; S, T) \triangleq \frac{P(t,T)}{P(t,S)}$ *satisfies SDE*

$$dFP(t; S, T) = FP(t; S, T)\sigma(D(t, S) - D(t, T))dW^S(t),$$

where W^S *is a Brownian motion under forward measure with respect to* $P(t, S)$.

Proof. (1) Use Theorem 3.5.2. (2) From Ito lemma, we have

$$dP(t, r, T) = \frac{\partial P}{\partial t}dt + \frac{\partial P}{\partial r}dr + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}drdr.$$

and (1). (3) use Theorem 3.3.1, we have

$$dFP(t; S, T) = FP(t; S, T)(\sigma_{P(t,T)} - \sigma_{P(t,S)})dW^S(t).$$

□

Lemma 3.5.14 (European call option price). *The value at time t of a European call option maturing at S with strike K on a bond maturing at* $T > S$ *is given by*

$$V(t) = P(t, S)E_S[\max(P(S, T) - K, 0)|\mathcal{F}_t],$$

where the expectation is taken with respect to forward measure of $P(t, S)$. *It can be showed that*

$$V(t) = BC(t; S, T, K) = P(t, T)N(d_+) - KP(t, S)N(d_-),$$

where

$$d_+ = \frac{\log\frac{FP(t;S,T)}{K} + \frac{1}{2}v(t,S)^2}{v(t,S)} = \frac{\log\frac{B(t,T)}{KB(t,S)} + \frac{1}{2}v(t,S)^2}{v(t,S)}, d_- = d_+ - v(t,S).$$

and

$$v(t,S)^2 = \int_t^S (\sigma(D(t,S) - D(t,T)))^2 dt$$

Proof. Under forward measure Q_S,

$$V(t) = P(t,S)E_S[(P(S,T) - K)^+|\mathcal{F}_t] = P(t,S)E_S[(FP_S(S) - K)^+|\mathcal{F}_t],$$

where $FP_S(t) = P(t,T)/P(t,S)$.

Note that the forward bond price dynamics $FP_S(t)$ under Q_S is given by

$$dFP_S(t) = P(t,T)(r(t))dt - FP_S(t)\sigma(D(t,S) - D(t,T))dW^S(t),$$

which is geometric Brownian motion. The rest is routine. □

3.5.4 Cox-Ingersoll-Ross model

3.5.4.1 *The model*

Definition 3.5.4 (Cox-Ingersoll-Ross model). *[5, p. 684][3, p. 64] In Cox-Ingersoll-Ross model, the short rate dynamics under risk-neutral measure is given by*

$$dr(t) = k(\mu - r(t))dt + \sigma\sqrt{r(t)}dW(t), r(0) = r_0$$

where W_t is the Brownian motion, r_0, k, μ, σ are constant model parameters, and $2k\mu > \sigma^2$ is imposed to ensure that the origin is inaccessible.

Lemma 3.5.15 (basic statistical properties of $r(t)$).

- $$r_{t+T} = \frac{Y}{2c},$$

  where Y is a non-central Chi-squared distribution with $4k\mu/\sigma^2$ degrees of freedom and non-centrality parameter $2cr_t\exp(-kT)$, and

  $$c = \frac{2k}{(1 - \exp(-kT))\sigma^2}.$$

- $$E[r_t|r_0] = r_0 \exp(-kt) + \mu(1 - \exp(-kt))$$
- $$Var[r_t|r_0] = r_0\frac{\sigma^2}{k}(\exp(-kt) - \exp(-2kt)) + \frac{\mu\sigma^2}{2k}(1 - \exp(-kt))^2.$$
- *The variance has the following properties:*
  - *Over short time scales, i.e., $t \ll 1/k$, the variance of r grows linearly as*
    $$Var[r_t] \approx r_0\sigma^2 t, \ for\ t \ll 1/k.$$
  - *Over long time scales, i.e., the variances approaches a constant:*
    $$Var[r_t] = \mu\frac{\sigma^2}{2k}, \ for\ t \gg 1/k.$$

Proof. (2) Note that $r(t)$ has solution given by

$$r(t) = e^{-kt}r(0) - \mu(e^{-kt} - 1) + \int_0^t e^{-k(t-s)}\sigma\sqrt{r(s)}dW_s.$$

Therefore,

$$E[r(t)] = e^{-kt}r(0) - \mu(e^{-kt} - 1),$$

and

$$\begin{aligned}
Var[r(t)] &= Var[\int_0^t e^{-k(t-s)}\sigma\sqrt{r(s)}dW_s] \\
&= \sigma^2 e^{-2kt} Var[\int_0^t e^{ks}\sqrt{r(s)}dW_s] \\
&= \sigma^2 e^{-2kt} E[(\int_0^t e^{ks}\sqrt{r(s)}dW_s)^2] \\
&= \sigma^2 e^{-2kt}\int_0^t e^{2ks}E[r(s)]ds \\
&= \sigma^2 e^{-2kt}\int_0^t e^{2ks}(e^{-ks}r(0) - \mu(e^{-ks} - 1))ds \\
&= \sigma^2 e^{-2kt}\int_0^t e^{ks}r(0) - \mu(e^{ks} - e^{2ks}))ds \\
&= r_0\frac{\sigma^2}{k}(\exp(-kt) - \exp(-2kt)) + \frac{\mu\sigma^2}{2k}(1 - \exp(-kt))^2
\end{aligned}$$

(3)(a) Use first order approximation $\exp(-kt) \approx 1 - kt$. (b) Use $\exp(-kt) \approx 0$. □

Lemma 3.5.16 (zero-coupon bond price in CIR model). *[3, p. 64] In the CIR model with parameter k, μ, σ,the price at time t of a zero-coupon bond with maturity T is*

$$P(t,T) = A(t,T)\exp(-B(t,T)r(t)),$$

where

$$\begin{aligned} A(t,T) &= [\frac{2h\exp((k+h)(T-t)/2)}{2h+(k+h)(\exp((T-t)h)-1)}]^{2k\mu/\sigma^2}, \\ B(t,T) &= \frac{2(\exp((T-t)h)-1)}{2h+(k+h)(\exp((T-t)h)-1)} \\ h &= \sqrt{k^2+2\sigma^2} \end{aligned}$$

3.5.4.2 *Yield curve in CIR model*

Lemma 3.5.17 (yield curve in CIR model). *[3, p. 64] In the CIR model with parameter k, μ, σ, the yield curve at time t is*

$$y(t,T) = -\frac{1}{T-t}\ln A(t,T) + \frac{1}{T-t}B(t,T)r(t)),$$

where

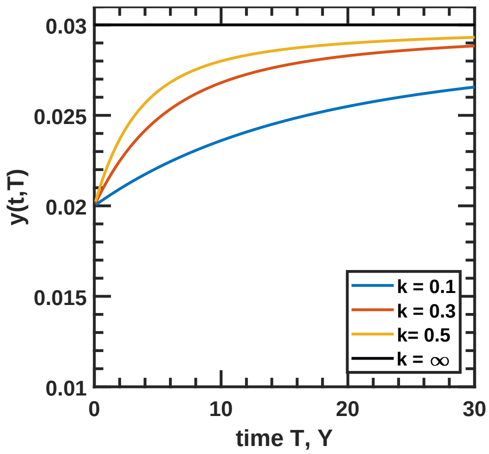
$$\begin{aligned} \ln A(t,T) &= \frac{2k\mu}{\sigma^2}\ln\frac{2h\exp((k+h)(T-t)/2)}{2h+(k+h)(\exp((T-t)h)-1)}, \\ B(t,T) &= \frac{2(\exp((T-t)h)-1)}{2h+(k+h)(\exp((T-t)h)-1)} \\ h &= \sqrt{k^2+2\sigma^2}. \end{aligned}$$

Remark 3.5.16 (yield curve shape). From Figure 3.5.3, we can see that

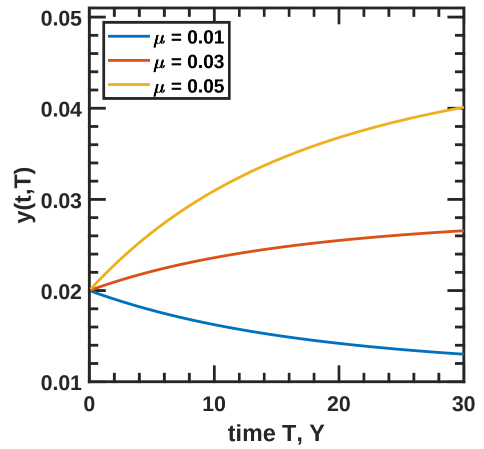
- (reversion speed effect) If $\alpha \to \infty$(but θ/α remains to be a constant), we can see that

  $$B(t,T) = 0, A(t,T) = \exp(-\frac{\theta}{\alpha})(T-t)) = \exp(-r^*(T-t)), y(t,T) = r^*.$$

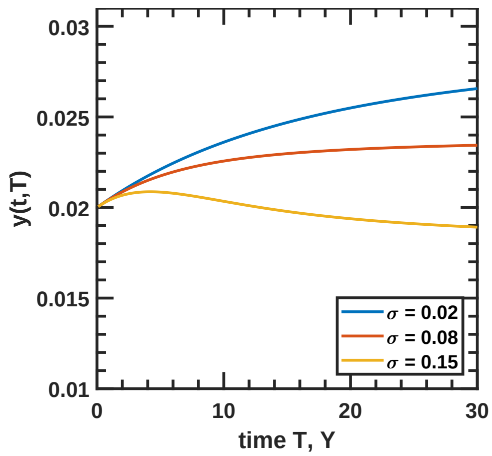
  If reversion speed α increases, then the yield curve will more resemble the equilibrium flat yield curve given by $y(t,T)* = r^*$.
- (mean level effect) The mean level will tend to pull the far-end of the yield curve toward the mean level.
- (volatility level effect) The volatility level tend to pull the far-end of the yield curve down.



(a) Yield curve shape vs. reversion speed.



(b) Yield curve shape vs. mean level μ.



(c) Yield curve shape vs. volatility.

Figure 3.5.5: Control yield curve shape via CIR model parameters.

Lemma 3.5.18 (evolution of yield curves). *Let the current time be 0. Consider the short rate under risk neutral measure is evolving under the CIR model with parameter* (μ, k, σ)*,*

$$dr(t) = k(\mu - r(t))dt + \sigma\sqrt{r(t)}dW(t), r(0) = r_0,$$

where W_t *is the Brownian motion,* r_0, k, μ, σ *are constant model parameters. The yield curve is given by*

$$y(0,T) = r(0)\frac{B(0,T)}{T} - \frac{\ln A(0,T)}{T}.$$

where

$$A(t,T) = [\frac{2h\exp((k+h)(T-t)/2)}{2h+(k+h)(\exp((T-t)h)-1)}]^{2k\mu/\sigma^2},$$
$$B(t,T) = \frac{2(\exp((T-t)h)-1)}{2h+(k+h)(\exp((T-t)h)-1)}$$
$$h = \sqrt{k^2+2\sigma^2}$$

- ***(evolution of whole yield curve)*** *The future yield curve $y(\tau, T+\tau)$ is given by*

$$y(\tau, T+\tau) = r(\tau)\frac{B(\tau, T+\tau)}{T} - \frac{\ln A(\tau, T+\tau)}{T}$$
$$= r(\tau)\frac{B(0,T)}{T} - \frac{\ln A(0,T)}{T}.$$

- ***evolution of yield associated with a fixed maturity date*** *The future yield $y(\tau, T)$ is given by*

$$y(\tau, T) = r(\tau)r(\tau)\frac{B(\tau, T+\tau)}{T} - \frac{\ln A(\tau, T+\tau)}{T}$$

Methodology 3.5.3 (simulate the future term structure in CIR model). *Suppose*

- *current time is 0.*
- *we are given the current term structure $y(0,T), T \in \mathbb{R}^+$.*
- *we are given a CIR model*

$$dr(t) = k(\mu - r(t))dt + \sigma\sqrt{r(t)}dW(t), r(0) = r_0.$$

- *we are given the initial short rate $r(0)$*

Then we can generate a ***sample term structure*** *in future time $t > 0$ via the following procedure:*

- *simulate $r(t)$ by Monte carlo simulation:*

$$r(t+dt) = r(t) + k(\mu - r(t))dt + \sigma\sqrt{r(t)dt}Z, Z \sim N(0,1).$$

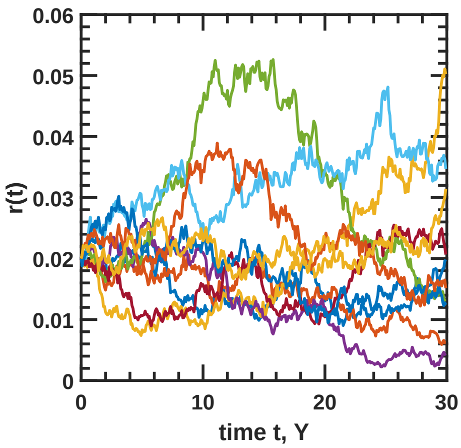
- *the sample yield curve is given by*

$$y(t,T) = r(t)\frac{B(t,T)}{T-t} - \frac{\ln A(t,T)}{T-t}$$

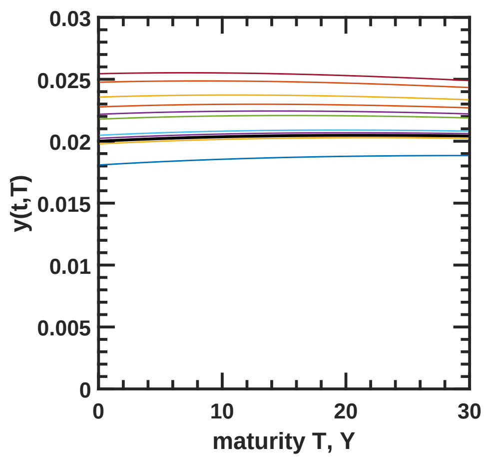
where

$$B(t,T) = \frac{1-\exp(-\alpha(T-t))}{\alpha},$$

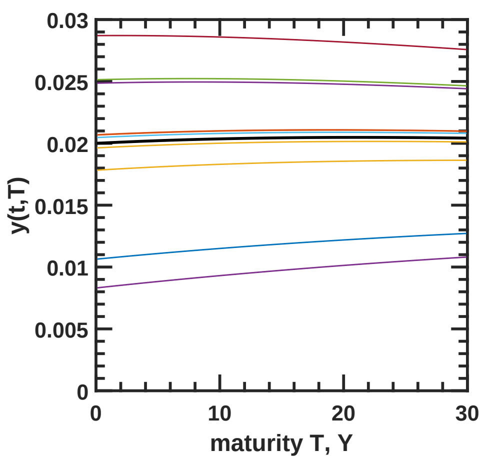
$$A(t,T) = \exp\left(\left(\frac{\theta}{\alpha} - \frac{\sigma^2}{2\alpha^2}\right)(B(t,T)-(T-t)) - \frac{\sigma^2}{4\alpha}B^2(t,T)\right).$$



(a) A. Realizations of short rate trajectories.



(b) A. Realizations of yield curve after 1Y (thin lines) from the initial yield curve(thick line).



(c) B. Realizations of yield curve after 5Y (thin lines) from the initial yield curve(thick line).

Figure 3.5.6: Yield curve dynamics in Merton model

3.5.4.3 *Instantaneous forward rate dynamics*

Lemma 3.5.19 (instantaneous forward rate dynamics in CIR model). *Consider a CIR model under risk-neutral measure given by*

$$dr(t) = k(\mu - r(t))dt + \sigma\sqrt{r(t)}dW(t), r(0) = r_0.$$

Let current time be t.

- *The instantaneous forward rate given by the model is*

$$f(t,T) = \frac{2k\mu(\exp((T-t)h-1))}{2h+(k+h)(\exp((T-t)h)-1)} + r_0\frac{4h^2\exp((T-t)h)}{(2h+(k+h)(\exp((T-t)h)-1))^2},$$

  where $h = \sqrt{k^2+2\sigma^2}$.

Proof. Note that

$$\begin{aligned}
&\frac{\partial}{\partial T}B(t,T) \\
&= \frac{\partial}{\partial T}\frac{2(exp((T-t)h)-1)}{2h+(k+h)(\exp((T-t)h)-1)} \\
&= \frac{(2h+(k+h)(\exp(h(T-t))-1)2h\exp(h(T-t)) - 2(\exp(h(T-t))-1)(k+h)h\exp((T-t)h)}{(2h+(k+h)(\exp((T-t)h)-1))^2} \\
&= \frac{4h^2\exp((T-t)h)}{(2h+(k+h)(\exp((T-t)h)-1))^2}
\end{aligned}$$

Note that

$$\ln A(t,T) = \frac{2k\mu}{\sigma^2}(\ln 2h + \frac{(k+h)(T-t)}{2} - \ln(2h+(k+h)(\exp((T-t)h)-1).$$

Then

$$\begin{aligned}
&\frac{\partial}{\partial T}\ln A(t,T)\\
&= \frac{2k\mu}{\sigma^2}\Big(\frac{k+h}{2} - \frac{(k+h)h\exp((T-t)h)}{(2h+(k+h)(\exp((T-t)h)-1))}\\
&= \frac{2k\mu}{\sigma^2}\Big(\frac{k+h}{2} - \frac{(k+h)h\exp((T-t)h)}{(2h+(k+h)(\exp((T-t)h)-1))}\\
&= \frac{2k\mu}{\sigma^2}\Big(\frac{(k+h)h + \frac{(k+h)^2}{2}(\exp((T-t)h)-1) - (k+h)h\exp((T-t)h)}{(2h+(k+h)(\exp((T-t)h)-1))}\\
&= \frac{2k\mu}{\sigma^2}\frac{(\exp((T-t)h)-1)(k^2-h^2)/2}{(2h+(k+h)(\exp((T-t)h)-1))}\\
&= \frac{2k\mu}{\sigma^2}\frac{(\exp((T-t)h)-1)(-\sigma^2)}{(2h+(k+h)(\exp((T-t)h)-1))}\\
&= \frac{(\exp((T-t)h)-1)(-2k\mu)}{(2h+(k+h)(\exp((T-t)h)-1))}
\end{aligned}$$

□

3.5.4.4 *Applications*

Lemma 3.5.20 (zero-coupon bond price in CIR model). *[3, p. 66] The price at time t of a zero-coupon bond with maturity T is*

$$P(t,T) = E_Q[\exp(-\int_t^T r(s)ds)] = A(t,T)\exp(-B(t,T)r(t),$$

where

$$\begin{aligned}
A(t,T) &= \Big[\frac{2h\exp((k+h)(T-t)/2)}{2h+(k+h)(\exp((T-t)h)-1)}\Big]\\
B(t,T) &= \Big[\frac{2(\exp((T-t)h)-1)}{2h+(k+h)(\exp((T-t)h)-1)}\Big]\\
h &= \sqrt{k^2+2\sigma^2}.
\end{aligned}$$

Under risk-neutral measure Q, the bond dynamics is given by

$$dP(t,T) = r(t)P(t,T)dt - B(t,T)P(t,T)\sigma\sqrt{r(t)}dW(t).$$

3.5.5 Model extension using deterministic shift

3.5.5.1 *Principles*

Proposition 3.5.3. *[3, p. 95] Let current time be 0. Let x_t be a (integrable) stochastic process and let the short rate process $r(t)$ be*

$$r_t = x_t + \phi(t),$$

where $\phi(t)$ is deterministic-shift function

It follows that

- *The zero-coupon bond price is given by*

$$P(t,T) = E_Q[\exp(-\int_t^T r_s ds)|\mathcal{F}_t] = \exp(-\int_t^T \phi(s)ds)F(t,T,x_t).$$

- *Let the current market term structure be represented by $f^*(0,T)$. Then the term structure produced by the model will exactly match with the current market term structure, i.e.,*

$$-\frac{\partial \ln P(0,T)}{\partial T} = f^*(0,T)$$

if

$$\begin{aligned}\phi(T) &= f^*(0,T) + \frac{\partial \ln F(0,t,x_0)}{\partial T}\\ &= f^*(0,T) - f(0,T), f(0,T) \triangleq -\frac{\partial \ln F(0,t,x_0)}{\partial T}.\end{aligned}$$

- *With the choice of ϕ in (2), we have*

$$P(t,T) = \frac{P^*(0,T)F(0,t,x_0)}{P^*(0,t)F(0,T,x_0)}F(t,T,x_t),$$

where $P^*(t,T)$ is the current market term structure.

Proof. (1)

$$\begin{aligned}
P(t,T) &= E_Q[\exp(-\int_t^T r_s ds)|\mathcal{F}_t] \\
&= E_Q[\exp(-\int_t^T \phi(s) + x_s ds)|\mathcal{F}_t] \\
&= \exp(-\int_t^T \phi(s)ds) E_Q[\exp(-\int_t^T x_s ds)|\mathcal{F}_t] \\
&= \exp(-\int_t^T \phi(s)ds) F(t,T,x_t).
\end{aligned}$$

(2)

$$\begin{aligned}
f^*(0,T) &= -\frac{\partial \ln P(0,T)}{\partial T} \\
&= \phi(T) - \frac{\partial \ln F(0,T,x_0)}{\partial T} \\
\implies \phi(T) &= f^*(0,T) + \frac{\partial \ln F(0,T,x_0)}{\partial T}.
\end{aligned}$$

(3)

$$\begin{aligned}
&\exp(-\int_t^T \phi(s)ds) \\
&= \exp(-\int_t^T f^*(0,s) + \frac{\partial \ln F(0,s,x_0)}{\partial s} ds) \\
&= \exp(-\int_t^T f^*(0,s)ds) \exp(-\int_t^T \frac{\partial \ln F(0,s,x_0)}{\partial s} ds) \\
&= \frac{P^*(0,T)}{P^*(0,t)} \exp(-\int_t^T d\ln F(0,s,x_0)) \\
&= \frac{P^*(0,T)}{P^*(0,t)} \exp(-\ln F(0,t,x_0) + \ln F(0,T,x_0)) \\
&= \frac{P^*(0,T)F(0,t,x_0)}{P^*(0,Tt)F(0,T,x_0)}
\end{aligned}$$

where we use the fact that

$$\exp(-\int_t^T f^*(0,s)ds) = \frac{\exp(-\int_0^T f^*(0,s)ds)}{\exp(-\int_0^t f^*(0,s)ds)} = \frac{P^*(0,T)}{P^*(0,t)}.$$

Eventually,

$$P(t,T) = \exp(-\int_t^T \phi(s)ds) F(t,T,x_t) = \frac{P^*(0,T)F(0,t,x_0)}{P^*(0,t)F(0,T,x_0)} F(t,T,x_t).$$

□

Remark 3.5.17 (implication for calibration)**.**

- The deterministic shift extension enables the perfect match to the initial term structure.
- Suppose the stochastic process is given by

$$dx_t = \mu(t)dt + \sigma(t)dW_t.$$

We can choose $\phi(t)$ to fit the initial term structure and choose $\sigma(t), \mu(t)$ to match the volatility term structure.

3.5.5.2 *Extended Vasicek model*

Definition 3.5.5 (extended Vasicek model). *The extended CIR model under risk-neutral model is given by*

$$\begin{aligned} dx(t) &= -kx(t)dt + \sigma dW_t, x(0) = 0. \\ r(t) &= x(t) + \alpha(t). \end{aligned}$$

Lemma 3.5.21. *Consider the short rate $r(t)$ under risk neutral measure is evolving under the following model given by*

$$\begin{aligned} dx(t) &= -kx(t)dt + \sigma dW_t, x(0) = 0. \\ r(t) &= x(t) + \alpha(t) \end{aligned}$$

Let the current time be 0. It follows that

- *If the current market term structure is given by $f^M(0,t)$, then choose*

$$\alpha(t) = f^M(0,t) + \frac{\sigma^2}{2k^2}(1 - \exp(-kt))^2,$$

can match the market term structure.
- *With the choice of ϕ in (1), we have*

$$P(t,T) = \frac{P^*(0,T)A(0,t)\exp(-B(0,t)x_0)}{P^*(0,t)A(0,T)\exp(-B(0,T)x_0)}A(t,T)\exp(-B(t,T)x_t),$$

where $P^*(0,T)$ is the current market term structure, and

$$B(t,T) = \frac{1 - \exp(-k(T-t))}{\alpha},$$

$$A(t,T) = \exp((-\frac{\sigma^2}{2k^2})(B(t,T)-(T-t)) - \frac{\sigma^2}{4k}B^2(t,T))),$$

Proof. (1)Note that the forward rate associated with the dynamics of

$$dx(t) = -kx(t)dt + \sigma dW_t, x(0) = 0$$

is given by(Lemma 3.5.12)

$$f(0,T) = -\frac{\sigma^2}{2k^2}(1-\exp(-k(T-t)))^2.$$

Then we use the principle (Theorem 3.5.3), we have

$$\alpha(t) = f^M(0,t) - f(0,t).$$

(2) Note that the zero-coupon bond price associated with the dynamics of

$$dx(t) = -kx(t)dt + \sigma dW_t, x(0) = 0$$

is given by(Lemma 3.5.10)

$$P(t,T) = A(t,T)\exp(-B(t,T)x(t)),$$

where

$$B(t,T) = \frac{1-\exp(-k(T-t))}{\alpha},$$
$$A(t,T) = \exp((-\frac{\sigma^2}{2k^2})(B(t,T)-(T-t)) - \frac{\sigma^2}{4k}B^2(t,T))).$$

□

3.5.5.3 *Extended CIR model*

Definition 3.5.6 (extended CIR model). *The extended CIR model under risk-neutral model is given by*

$$\begin{aligned} dx(t) &= k(\mu - x(t))dt + \sigma\sqrt{x(t)}dW(t), x(0) = x_0 \\ r_t &= x_t + \phi(t). \end{aligned}$$

Lemma 3.5.22 (extended CIR model properties). *[3, p. 95] Let current time be 0. Let the original model has risk-neutral dynamics be*

$$dx(t) = k(\mu - x(t))dt + \sigma\sqrt{x(t)}dW(t), x(0) = x_0.$$

Denote

$$F(t,T,x_t) = E_Q[\exp(-\int_t^T x_s ds)|\mathcal{F}_t].$$

The deterministic-shift extension consists of defining the short rate by

$$r_t = x_t + \phi(t).$$

It follows that

- *Let the current market term structure be represented by $f^*(0,T)$. Then the term structure produced by the model will exactly match with the current market term structure, i.e.,*

$$-\frac{\partial \ln P(0,T)}{\partial T} = f^*(0,T)$$

if

$$\begin{aligned}\phi(T) &= f^*(0,T) - f(0,T)\\ &= f^*(0,T) - \frac{2k\mu(\exp(th-1))}{2h+(k+h)(\exp(th)-1)} - r_0\frac{4h^2\exp(th)}{(2h+(k+h)(\exp(th)-1))^2},\\ h &= \sqrt{k^2+2\sigma^2}.\end{aligned}$$

- *With the choice of ϕ in (1), we have*

$$P(t,T) = \frac{P^*(0,T)A(0,t)\exp(-B(0,t)x_0)}{P^*(0,t)A(0,T)\exp(-B(0,T)x_0)}A(t,T)\exp(-B(t,T)x_t),$$

where $P^*(0,T)$ is the current market term structure, and

$$\begin{aligned}A(t,T) &= [\frac{2h\exp((k+h)(T-t)/2)}{2h+(k+h)(\exp((T-t)h)-1)}]\\ B(t,T) &= [\frac{2(\exp((T-t)h)-1)}{2h+(k+h)(\exp((T-t)h)-1)}]\\ h &= \sqrt{k^2+2\sigma^2}.\end{aligned}$$

Proof. (1) Note that the forward rate derived from CIR model(Lemma 3.5.19) is given by

$$f(0,T) = \frac{2k\mu(\exp(Th-1))}{2h+(k+h)(\exp(Th)-1)} + r_0\frac{4h^2\exp(Th)}{(2h+(k+h)(\exp(Th)-1))^2}, h = \sqrt{k^2+2\sigma^2}$$

□

3.5.6 Ho-Lee model

Definition 3.5.7 (Ho-Lee model). *[5, p. 689] The Ho-Lee model for short rate, under the risk-neutral measure Q, is given by*

$$dr = \theta(t)dt + \sigma dW_t$$

where W_t is the Brownian motion, and σ is a constant.

Lemma 3.5.23 (zero-coupon bond price in Ho-Lee model). *[6, p. 26] With the Ho-Lee short rate model under the risk-neutral measure Q, given by*

$$dr = \theta(t)dt + \sigma dW_t,$$

the zero-coupon bond price is given by

$$P(t,T) = \exp(-m(t) + \frac{1}{2}s^2)$$

where

$$m(t) = r(t)(T-t) + \int_t^T \theta(u)(T-u)du$$

and variance

$$s^2 = \frac{1}{3}\sigma^2(T-t)^3$$

under risk-neutral measure.

Define $A(t,T) = \exp(-\frac{1}{2}s^2 - \int_t^T \theta(u)(T-u)du), B(t,T) = T-t$. Then we can write

$$P(t,T) = A(t,T)\exp(-B(t,T)r(t)).$$

The zero-coupon bond dynamics is given by

$$dP(t,T) = r(t)P(t,T)dt - \sigma(T-t)P(t,T)dW_t$$

Proof. (1) We have

$$r(s) = r(t) + \int_t^s \theta(u)du + \sigma(W(s) - W(t)),$$

then

$$\int_t^T r(s)ds = r(t)(T-t) + \int_t^T (\int_t^s \theta(u)du)ds + \sigma \int_t^T (T-s)dW(s)$$
$$= r(t)(T-t) + \int_t^T \theta(u)(T-u)du + \sigma \int_t^T (T-s)dW(s)$$

where we exchange the integral. Note that

$$d((T-s)W(s)) = -W(s)ds + (T-s)dW(s),$$

then

$$[(T-s)W(s)]|_t^T = -\int_t^T W(s)ds + \int_t^T (T-s)dW(s).$$

(2) similar to Lemma 3.5.4. (3) Take differential on $P(t,T)$ and use the fact that $P(t,T)$ has to satisfy term structure function such that drift is r. See Lemma 3.5.1. □

Lemma 3.5.24 (fitting θ to initial term structure). *[6, p. 27] Suppose we have market prices of zero-coupon bond $P(t,T), T > 0$, then*

$$\theta(T) = -\frac{\partial^2}{\partial T^2}\log(P(t,T)) + \sigma^2 T$$

will enable Ho-Lee model to fit current term structure exactly.

Proof.

$$\log P = r(t)(T-t) + \int_t^T \theta(u)(T-u)du - \frac{1}{6}\sigma^2(T-t)^3$$
$$\implies \frac{\partial \log P}{\partial T} = -r_0 - \int_t^T \theta(s)ds + \frac{1}{2}\sigma^2 T^2$$
$$= -r_0 - \int_t^T \theta(s)ds + \frac{1}{2}\sigma^2 T^2$$
$$\implies \frac{\partial^2 \log P}{\partial T^2} = -\theta(T) + \sigma^2 T$$

□

Remark 3.5.18 (how to fit σ ?). Once we fit θ to initial term structure, the only variable undetermined is σ, which can be determined from market prices of other vanilla bond options.

Remark 3.5.19 (interpretation).

- The Ho-Lee model is the simplest no-arbitrage short-rate model.
- The value of $\theta(t)$ will be calibrated to make match the Zero-Coupon Bond prices, which depends on σ.

3.5.7 Hull-White model

3.5.7.1 *Fundamentals*

Definition 3.5.8 (Hull-White model). *[5, p. 691] In Hull-White model, the SDE for r under risk-neutral measure is given*

$$dr = (\theta(t) - \alpha r(t))dt + \sigma(t)dW_t$$

where W_t is the Brownian motion under risk-neutral measure, and α is constant, and the coefficients $\theta(t)$ and $\sigma(t)$ are time-dependent variables.

Lemma 3.5.25 (zero-coupon bond pricing in Hull-White model). *[1, p. 50] Assume the short rate r follows the Hull-White model under risk-neutral measure Q. We have* [a]

- $$r(s) = r(t)e^{-\alpha(s-t)} + \int_t^s \theta(u)e^{-\alpha(s-u)}du + \int_t^s \sigma(u)e^{-\alpha(s-u)}dW(u)$$

  In particular, $r(s)$ is a Gaussian process with mean

  $$m(s) = r(t)e^{-\alpha(s-t)} + \int_t^s \theta(u)e^{-\alpha(s-u)}du$$

  and variance

  $$v^2 = \int_t^s \sigma(u)^2 e^{-2\alpha(s-u)}du.$$

- *Let $D(t,T) \triangleq \int_t^T e^{-\alpha(s-t)}ds$, then*

  $$-\int_t^T r(s)ds = -D(t,T)r(t) - \int_t^T \theta(u)D(u,T)du - \int_t^T \sigma(u)D(u,T)dW(u).$$

- *The zero-coupon bond price*

  $$P(t,T) = E_Q[\exp(-\int_t^T r(s)ds)|\mathcal{F}_t] = \exp(m(t) + \frac{1}{2}s^2(t).$$

where

$$m(t) = -D(t,T)r(t) - \int_t^T \theta(u)D(u,T)du$$

and

$$s^2(t) = \int_t^T \sigma(u)^2 D(u,T)^2 du.$$

Define $A(t,T) = \exp(\frac{1}{2}s(t)^2 - \int_t^T \theta(u)D(u,T)du), B(t,T) = D(t,T)$. *Then we can write*

$$P(t,T) = A(t,T)\exp(-B(t,T)r(t)).$$

- *The zero-coupon bond has the dynamics*

$$dP(t,T) = P(t,T)(r(t)dt - \sigma(t)B(t,T)dW_t).$$

[a] Note that if current time is t, then $r(t)$ is observable and $P(t,T)$ is a deterministic value for each T; if current time is 0, then $r(t)$ is unknown and $P(t,T)$ is a random variable quantity for each T.

Proof. (1) We can verify this is the solution. (2)

$$-\int_t^T r(s)ds = -\int_t^T r(t)e^{-\alpha(s-t)}ds - \int_t^T\int_t^s \theta(u)e^{-\alpha(s-u)}duds - \int_t^T\int_t^s \sigma(u)e^{-\alpha(s-u)}dW(u)ds$$

We then use change of variable in integral given by

$$\begin{aligned}\int_t^T\int_t^s \theta(u)e^{-\alpha(s-u)}duds &= \int_t^T\int_u^T \theta(u)e^{-\alpha(s-u)}dsdu \\ &= \int_t^T\int_u^T \theta(u)e^{-\alpha(s-u)}dsdu \\ &= \int_t^T\int_u^T \theta(u)D(u,T)du\end{aligned}$$

Similarly,

$$\int_t^T\int_t^s \sigma(u)e^{-\alpha(s-u)}dW(u)ds = \ = \int_t^T\int_u^T \theta(u)D(u,T)dW(u)$$

(3) Note that $-\int_t^T r(s)ds$ is a Gaussian process with mean

$$m = -D(t,T)r(t) - \int_t^T \theta(u)D(u,T)du$$

and variance

$$s^2 = \int_t^s \sigma(u)^2 D(u,T)^2 du.$$

Use the property of log normal variable [**??**]. □

Lemma 3.5.26 (zero-coupon bond price fitting to the current term structure and its distribution). *[1, p. 54]*

- *The Hull-White short rate model fitting to the current(current time is 0) term structure(characterized by $f(0,t), t \geq 0$) is given by*

$$\theta(t) = \frac{\partial f(0,t)}{\partial t} + \alpha f(0,t) + \int_0^t \sigma^2 e^{-2\alpha(t-s)} ds.$$

- *The future zero-coupon bond price, a random quantity, based on information on time 0 is given by*

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp(-A(t,T) - r(t)D(t,T)),$$

where

$$r(t) = r(0)e^{-\alpha t} + \int_0^t \theta(u)D(0,t)du + \int_0^t \sigma(u)D(0,t)dW(u),$$

$$A(t,T) = -f(0,t)D(t,T) + \frac{1}{2}D^2(t,T)\int_0^t \sigma^2(u)\exp(-2\alpha(t-u))du,$$

$$D(t,T) = \int_t^T \exp(-\alpha(T-s))ds.$$

- *The future yield curve $y(t,T)$, a random quantity, based on information on time 0 is given by*

$$y(t,T) = -\frac{\ln P(t,T)}{T-t} = -\frac{1}{T-t}(\ln\frac{P(0,T)}{P(0,t)} - A(t,T) - r(t)D(t,T)).$$

- *Under risk-neutral measure $P(t,T)$ is lognormal distribution; $\ln P(t,T)$ has variance given by*

$$D^2(t,T)Var[r(t)] = D^2(t,T)\int_0^t \sigma^2(u)\exp(-2\alpha(t-u))du.$$

Proof. (1) See Lemma 3.6.5.(2) See reference. (3)(4) Note that $r(t)$ is the Gaussian random variable(Lemma 3.5.25) and then $Var[\ln P(t,T)] = Var[r(t)D(t,T)] = D(t,T)^2 Var[r(t)]$. □

3.5.7.2 *Yield curve dynamics*

3.5.7.3 *Derivative pricing*

Lemma 3.5.27 (Pricing European call option on zero-coupon bond). *The value at time t of a European call option maturing at S with strike K on a bond maturing at $T > S$ is given by*

$$V(t) = P(t,S)E_S[\max(P(S,T) - K, 0)|\mathcal{F}_t],$$

where the expectation is taken with respect to forward measure of $P(t,S)$. It can be showed that

$$V(t) = BC(t;S,T,K) = P(t,T)N(d_+) - KP(t,S)N(d_-),$$

where

$$d_+ = \frac{\log \frac{FP(t;S,T)}{K} + \frac{1}{2}v(t,S)^2}{v(t,S)} = \frac{\log \frac{B(t,T)}{KB(t,S)} + \frac{1}{2}v(t,S)^2}{v(t,S)}, d_- = d_+ - v(t,S).$$

and

$$v(t,S)^2 = \int_t^S (\sigma(D(t,S) - D(t,T)))^2 dt$$

Proof. Under forward measure Q_S,

$$V(t) = P(t,S)E_S[(P(S,T) - K)^+|\mathcal{F}_t] = P(t,S)E_S[(FP_S(S) - K)^+|\mathcal{F}_t],$$

where $FP_S(t) = P(t,T)/P(t,S)$.

Note that the forward bond price dynamics $FP_S(t)$ under Q_S is given by

$$dFP_S(t) = P(t,T)(r(t))dt - FP_S(t)\sigma(D(t,S) - D(t,T))dW^S(t),$$

which is geometric Brownian motion. The rest is routine. □

Lemma 3.5.28 (Swaption pricing). *[1, p. 59] Consider an call option with strike K and expiry T_0 on a payer swap with unit notional, settlement dates $T_1, ..., T_n$ and reset dates $T_0, ..., T_{n-1}$.Denote $\tau_i = T_i - T_{i-1}$. Let the current time be $t < T_0$.*

- *The swapiton's payoff at time T_0 can be written by*

$$(1 - K\sum_{i=1}^{n} \tau_i P(T_0, T_i) - P(T_0, T_n))^+$$

- *Let $\tilde{r}$ be such that*

$$K\sum_{i=1}^{n} \tau_i F(T_0, T_i; \tilde{r}) + F(T_0, T_n; \tilde{r}) = 1,$$

where[a]

$$F(T_0, T_i; r) = A(T_0, T_i)\exp(-B(T_0, T_i)r).$$

Then

$$(1 - K\sum_{i=1}^{n}\tau_i P(T_0, T_i) - P(T_0, T_n))^+ = K\sum_{i=1}^{n}\tau_i(P_i - P(T_0, T_i))^+ + (P_n - P(T_0, T_n))^+,$$

where we use the notation

$$P_i = F(T_0, T_i; \tilde{r}).$$

- *The current value of the swaption at time t is given by*

$$V(t) = K\sum_{i=1}^{n}\tau_i BP(t, T_0, T_i, P_i) + BP(t, T_0, T_i, P_n)$$

where $BP(t, T_0, T_i, P_i)$ is the time t put option price of the bond $P(T_0, T_i)$ with strike P_i.

a note that $F(T_0, T_i; r)$ is a decreasing function on r, and $K\sum_{i=1}^{n}\tau_i F(T_0, T_i; \tilde{r}) + F(T_0, T_n; \tilde{r})$ is also a decreasing function on r; then we must have a unique root.

Proof. (1) Note that the value of a payer interest rate swap(Lemma 3.1.7) is given by

$$V(T_0, K) = P(T_0, T_0) - P(T_0, T_n) - K\sum_{i=1}^{n}\tau_i P(T_0, T_i).$$

(2)

$$\begin{aligned}
&(1 - K\sum_{i=1}^{n}\tau_i P(T_0, T_i) - P(T_0, T_n))^+ \\
&= (K\sum_{i=1}^{n}\tau_i P_i + P_n - K\sum_{i=1}^{n}\tau_i P(T_0, T_i) - P(T_0, T_n))^+ \\
&= (K\sum_{i=1}^{n}\tau_i(P_i - P(T_0, T_i)) + (P_n - P(T_0, T_n))^+ \\
&= K\sum_{i=1}^{n}\tau_i(P_i - P(T_0, T_i))^+ + (P_n - P(T_0, T_n))^+
\end{aligned}$$

The reason that the last step holds is because

- If $r(T_0) > \tilde{r}$, then

$$(K\sum_{i=1}^{n}\tau_i(P_i - P(T_0, T_i)) + (P_n - P(T_0, T_n)))^+ = 0,$$

since $K\sum_{i=1}^{n}\tau_i P_i + P_n < 1$.

- If $r(T_0) < \tilde{r}$, then

$$\begin{aligned}&(K\sum_{i=1}^{n}\tau_i(P_i - P(T_0,T_i)) + (P_n - P(T_0,T_n)))^+\\&= K\sum_{i=1}^{n}\tau_i(P_i - P(T_0,T_i)) + (P_n - P(T_0,T_n))\\&= 1 - K\sum_{i=1}^{n}\tau_i P(T_0,T_i) - P(T_0,T_n)\end{aligned}$$

(3) Since we already decompose the payoff to a linear combination of put options on bond, then the current value of the swaption is also the linear combination of the current value of these put options on bonds. □

Lemma 3.5.29 (Bermudan swaption pricing). *Consider an Bermudan swaption option with strike K and exercising date $0 < T_0, T_1, ..., T_{n-1}$ on a payer swap with unit notional, settlement dates $T_1, ..., T_n$ and reset dates $T_0, ..., T_{n-1}$. Denote $\tau_i = T_i - T_{i-1}$. Let the current time be $0 < T_0$. Denote the Bermudan swaption value at time T_i by $Berm(T_i)$.*

- *If the Bermudan swaption has not been exercised before T_i, then the (undiscounted)exercise value at T_i is*
$$Ex(T_i) = (PS(T_i, K))^+,$$
where $PS(T_i, K)$ denotes the time T_i value(a random quantity) of the swap with fixed rate K, settlement dates $T_{i+1}, ..., T_n$ and reset dates $T_i, ..., T_{n-1}$.
- *If the Bermudan swaption has not been exercised before T_i, then the (undiscounted)continuation value at T_i is*
$$C(T_i) = P(T_i, T_{i+1})E_{T_{i+1}}(Berm(T_{i+1})|\mathcal{F}_{T_i}).$$
In particular, $C(T_{n-1}) = 0$.
- *The value of the Bermudan swaption is*
$$Berm(T_i) = \max(Ex(T_i), C(T_i)).$$

Proof. This is just the Bellman's principle. Note that the continuation value at time T_i is equivalent to a Bermudan swaption starting at T_{i+1} and then discount it to T_i. □

Methodology 3.5.4 (backward induction to calculate Bermudan swaption value). *Consider an Bermudan swaption option with strike K and exercising date $0 < T_0, T_1, ..., T_{n-1}$ on a payer swap with unit notional, settlement dates $T_1, ..., T_n$ and reset dates $T_0, ..., T_{n-1}$. Denote $\tau_i = T_i - T_{i-1}$. Let the current time be $0 < T_0$. Denote the Bermudan swaption value at time T_i*

by $Berm(T_i)$. $PS(T_i, K)$ denotes the time T_i value(a random quantity) of the swap with fixed rate K, settlement dates $T_{i+1}, ..., T_n$ and reset dates $T_i, ..., T_{n-1}$.

- *We start with T_{n-1}, where*
$$Berm(T_{n-1}) = \max(Ex(T_{n-1}), C(T_{n-1})) = (PS(T_{n-1}, K))^+,$$
*where $PS(T_{n-1}, K)$ denotes the time T_{n-1} value(a random quantity) of **a one-period payer swap** with fixed rate K, settlement dates T_n and reset dates T_{n-1}.*
- *Back to T_{n-2}, we calculate*
$$C(T_{n-2}) = P(T_{n-2}, T_{n-1})E_{T_{n-1}}[Berm(T_{n-1})|\mathcal{F}_{T_{n-2}}].$$
$$Ex(T_{n-2}) = (PS(T_{n-2}, K))^+,$$
$$Berm(T_{n-2}) = \max(Ex(T_{n-2}), C(T_{n-2})).$$
- *Continue the backward induction process to T_0, then we calculate the current time value*
$$Berm(0) = P(0, T_0)E_{T_0}[Berm(T_{n-1})|\mathcal{F}_{T_{n-2}}]$$

Remark 3.5.20 (analytical vs. numerical method)**.**

- Note that it is generally impossible to evaluate the value of a Bermudan swaption analytically; this is because $Berm(T_i), C(T_i), Ex(T_i)$ are all random variables as a function of the stochastic short rate $r(T_i)$ at T_i.
- A simulation pricing approach will be:
  - Simulate N short rate trajectories $r(t), t \in [0, T_{n-1}]$.
  - On each trajectory $r^{(i)}(t)$, calculate $Berm^{(i)}(0)$ using the backward induction method.
  - Aggregate and take average to get the estimate of $Berm(0)$.

3.5.7.4 *Monte carlo simulation*

Methodology 3.5.5 (simulate the future term structure in Hull-White model). *Suppose*

- *current time is 0.*
- *we are given the current term structure $y(0, T), T \in \mathbb{R}^+$(or $P(0, T), T \in \mathbb{R}^+$, or $f(0, t), \mathbb{R}^+$).*
- *we are given a calibrated Hull-White model*
$$dr = (\theta(t) - \alpha r(t))dt + \sigma(t)dW_t.$$
- *we are given the initial short rate $r(0)$(which can be derived from $r(0) = f(0, 0)$).*

*Then we can generate a **sample term structure** in future time $t > 0$ via the following procedure:*

- *simulate $r(t)$ by drawing from the normal distribution*
$$r(t) \sim N(m(t), v^2(t)),$$
where
$$m(t) = r(0)e^{-\alpha t} + \int_0^t \theta(u)e^{-\alpha(t-u)}du,$$
$$v^2 = \int_0^t \sigma(u)^2 e^{-2\alpha(t-u)}du.$$
- *the sample yield curve is given by*
$$y(t,T) = -\frac{\ln P(t,T)}{T-t} = -\frac{1}{T-t}(\ln\frac{P(0,T)}{P(0,t)} - A(t,T) - r(t)D(t,T)),$$
where
$$A(t,T) = -f(0,t)D(t,T) + \frac{1}{2}D^2(t,T)\int_0^t \sigma^2(u)\exp(-2\alpha(t-u))du,$$
$$D(t,T) = \int_t^T \exp(-\alpha(T-s))ds.$$

Proof. See Lemma 3.5.26. □

3.5.7.5 *Trinomial tree model for short rate*

Definition 3.5.9 (trinomial tree model). *[3, p. 78][5, p. 724] Let $r^*(t) = f(0,t) + \frac{\sigma^2}{2a^2}(1 - e^{-at})^2$ denote the instantaneous equilibrium value of the Hull-White model. Let $x(t) \triangleq r(t) - r^*$ be the deviation from the equilibrium value.*

A trinomial tree model for x consists of

- *nodes (i,j): representing date $t = i\Delta t$, and states $x = j\Delta x$, where usually we choose $\Delta x = \sigma\sqrt{3\Delta t}$.*

- *state transition probability:*

$$P((i+1,s)|(i,j)) = \begin{cases} p_u, s = k+1 \\ p_m, s = k \\ p_d, s = k-1 \end{cases}$$

where $k = round(M_{i,j}/\Delta r)$.

Lemma 3.5.30. *In the trinomial tree model, we have*

- $$E[x(t_{i+1})|x(t_i) = x_{i,j}] \approx x_{i,j}\exp(-a\Delta t) \triangleq M_{i,j}$$
$$Var[x(t_{i+1})|x(t_i) = x_{i,j}] \approx \frac{\sigma^2}{2a}(1-e^{-2a\Delta t}) \triangleq V_i^2$$

- *state transition probability:*

$$P((i+1,s)|(i,j)) = \begin{cases} p_u = \frac{1}{6} + \frac{\eta_{j,k}^2}{6V_i^2} + \frac{\eta_{j,k}}{2\sqrt{3V_i}}, s = k+1 \\ p_m = \frac{2}{3} - \frac{\eta_{j,k}^2}{3V_i^2}, s = k \\ p_d = \frac{1}{6} + \frac{\eta_{j,k}^2}{6V_i^2} - \frac{\eta_{j,k}}{2\sqrt{3V_i}}, s = k-1 \end{cases}$$

- *The state transition probability satisfies*

$$p_u + p_m + p_d = 1$$
$$E[x_{i+1,k}|x_{i,j}] = M_{i,j}$$
$$Var[x_{i+1,k}|x_{i,j}] = V_i^2$$

Proof. (1) Note that SDE for x is given by

$$dx = -axdt + \sigma dW_t.$$

(2)(3) Directly verify. □

3.5.7.6 *Hull-White model calibration*

3.5.8 Two factor Gaussian model

3.5.8.1 *The model*

Definition 3.5.10 (two-factor Gaussian additive model). *The two-factor OU process is given by the following SDE*

$$\begin{aligned} r(t) &= x_1(t) + x_2(t) + \psi(t), r(0) = r_0 \\ dx_1(t) &= -a_1 x_1(t)dt + \sigma_1 dW_1(t), x_1(0) = x_{10} \\ dx_2(t) &= -a_2 x_2(t)dt + \sigma_2 dW_1(t), x_2(0) = x_{20} \end{aligned}$$

where $a_1, a_2, \sigma_1, \sigma_2, r_0$ are constant, $\psi(t)$ is a time-dependent function, and W_1, W_2 are correlated Brownian motion such that

$$dW_1(t)dW_2(t) = \rho dt.$$

Lemma 3.5.31 (zero-coupon bond price in two-factor Gaussian model). *Consider a two factor Gaussian additive short rate model in Definition 3.5.10.*

It follows that

- $$\begin{aligned} P(t,T) &\triangleq E[\exp(-\int_t^T r(s)ds)] \\ &= F_1(t,T,x_1(t))F_2(t,T,x_2(t))\exp(U(t,T))\exp(-\int_t^T \psi(u)du). \end{aligned}$$

where

$$\begin{aligned} F_i(t,T,x_1(t)) &\triangleq E[\exp(-\int_t^T x_1(s)ds)] \\ &= \exp(-\frac{1-\exp(-a_1(T-t))}{a_1}x_1(t) + \frac{1}{2}(\frac{\sigma^2}{a_1^2}(T-t+ \\ &\frac{2}{a_1}\exp(-a_1(T-t)) - \frac{1}{2a_1}\exp(-2a_1(T-t)) - \frac{3}{2a_1}))), i = 1,2,. \end{aligned}$$

and

$$U(t,T) = \frac{\rho\sigma_1\sigma_2}{a_1 a_2}(T - t + \frac{\exp(-a_1(T-t)) - 1}{a_1} + \frac{\exp(-a_2(T-t)) - 1}{a_2}$$
$$+ \frac{\exp(-(a_1+a_2)(T-t)) - 1}{a_1 + a_2}).$$

- *If we choose $\psi(T), T \in \mathbb{R}^+$ to match the current term structure given by $f^M(0,T)$, then*

$$\psi(T) = f^M(0,T) + \frac{\sigma_1^2}{2a_1^2}(1 - \exp(-a_1 T))^2 + \frac{\sigma_1^2}{2a_2^2}(1 - \exp(-a_2 T))^2$$
$$+ \rho\frac{\sigma_1\sigma_2}{a_1 a_2}(1 - \exp(-a_1 T))(1 - \exp(-a_2 T))$$

- *The zero-coupon bond price, which agrees with the current term structure, is given by*

$$\exp(-\int_t^T \psi(u)du) = \frac{P^M(0,T)}{P^M(0,t)}\exp(-\frac{1}{2}(V(0,T) - V(0,t)).$$

- *With the choice of ϕ in (2), we have*

$$P(t,T) = \frac{P^*(0,T)F(0,t,x_{10},x_{20})}{P^*(0,t)F(0,T,x_{10},x_{20})}F(t,T,x_1(t),x_2(t)),$$

where $P^*(0,T)$ is the current market term structure and

$$F(t,T,x_1(t),x_2(t)) = F_1(t,T,x_1(t))F_2(t,T,x_2(t))\exp(U(t,T))$$

Proof. Use **??** and Theorem 3.5.3. □

3.5.8.2 *Yield curve dynamics*

3.5.8.3 *Instantaneous forward rate dynamics*

Lemma 3.5.32 (instantaneous forward rate dynamics in two factor Gaussian model). *[9, p. 109] Consider the short rate under risk neutral measure is evolving under the Vasicek model given by*

$$dr(t) = (\theta - \alpha r(t))dt + \sigma dW_t.$$

It follows that

- *The instantaneous forward rate is given by*

$$f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T} = \psi(T) + f_1(t,T) + f_2(t,T) - \rho\frac{\sigma_1\sigma_2}{a_1 a_2}(1-\exp(-a_1(T-t)))(1-\exp(-a_2(T-t))).$$

where

$$f_1(t,T) = x_1(t)\exp(-a_1(T-t)) - \frac{\sigma_1^2}{2a_1^2}(1-\exp(-a_1(T-t))^2)$$

and

$$f_2(t,T) = x_2(t)\exp(-a_2(T-t)) - \frac{\sigma_1^2}{2a_2^2}(1-\exp(-a_2(T-t))^2)$$

- *The dynamics of instantaneous forward rate is given by*

$$d_t f(t,T) = d_t f_1 + d_t f_2 - \rho\sigma_1\sigma_2\exp(-a_1(T-t))\exp(-a_2(T-t))dt$$

where

$$d_t f_1(t,T) = \frac{\sigma_1^2}{a_1}(1-e^{-a_1(T-t)})e^{-a_1(T-t)}dt + \sigma e^{-a_1(T-t)}dW_1$$

and

$$d_t f_2(t,T) = \frac{\sigma_2^2}{a_2}(1-e^{-a_2(T-t))})e^{-a_2(T-t)}dt + \sigma e^{-a_2(T-t)}dW_2$$

- *We can write the forward rate dynamics in a more compact form given by*

$$d_t f(t,T) = \alpha(t,T)dt + \sigma_f(t,T)dW$$

where the forward rate volatility is given by

$$\sigma_f(t,T) = \sqrt{\sigma_1^2 e^{-a_1(T-t)} + \sigma_2^2 e^{-a_2(T-t)} + \sigma_2^2 + 2\rho\sigma_1\sigma_2 e^{-a_1(T-t)}e^{-a_2(T-t)}}$$

and the forward rate drift is given by

$$\alpha(t,T) = \sigma_f(t,T)\int_t^T \sigma_f(t,u)du.$$

Proof. (1) Note that in the two factor Gaussian model, we have zero-coupon bond price given by(Definition 3.5.10)

$$P(t,T) = F_1(t,T,x_1(t))F_2(t,T,x_2(t))\exp(U(t,T))\exp(-\int_t^T \psi(u)du).$$

We can easily verify

$$f_1(t,T) = -\frac{\ln F(t,T,x_1(t))}{dT}, f_2(t,T) = -\frac{\ln F(t,T,x_2(t))}{dT}$$

using the results in Vasicek model(Lemma 3.5.12). and the instantaneous forward rate is given by

$$f(t,T) = -\frac{\ln P(t,T)}{dT} = r(t) + \alpha(T-t) - \frac{1}{2}\sigma^2(T-t)^2.$$

(2)

$$\begin{aligned} df_t(t,T) &= \exp(-k(T-t))dr(t) + br_t \exp(-k(T-t))dt - \theta\exp(-k(T-t)) \\ &+ \frac{\sigma^2}{k}(1-\exp(-k(T-t)))\exp(-k(T-t))dt \\ &= \exp(-k(T-t))((\theta - \alpha r(t))dt + \sigma dW_t) + br_t\exp(-k(T-t))dt - \theta\exp(-k(T-t)) \\ &+ \frac{\sigma^2}{k}(1-\exp(-k(T-t)))\exp(-k(T-t))dt \\ &= \frac{\sigma^2}{k}(1-\exp(-k(T-t)))\exp(-k(T-t))dt + \sigma\exp(-k(T-t))dW_t \end{aligned}$$

□

Remark 3.5.21 (humped forward volatility). [3, p. 149]

- A humped volatility structure similar to what is commonly observed in the market for the caplets volatility, may be only reproduced for negative values of correlation ρ.
- If ρ is positive, then all terms in σ_f is decreasing and the hump is impossible.

3.5.8.4 *Correlation problems in factor models*

Note 3.5.3 (correlation in one-factor affine model). [9, p. 103] Consider a general one-factor model such that the zero-coupon bond price can be written by

$$P(t,T) = A(t,T)\exp(-B(t,T)r(t)).$$

- The zero-coupon bonds with maturities T_1 and T_2 are related by

$$\frac{P(t,T_2)}{P(t,T_1)} = \frac{A(t,T_2)}{A(t,T_1)}\exp(-r_t(B(t,T_1) - B(t,T_2)).$$

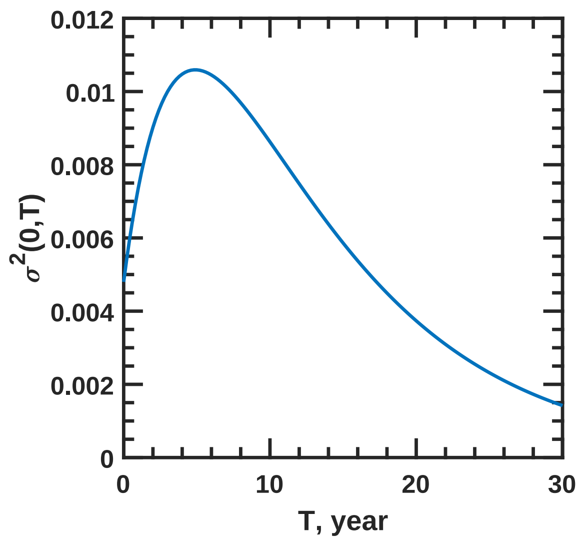


Figure 3.5.7: Humped shaped forward volatility produced with parameters: sig1 = 0.12; sig2 = 0.1; rho = -1; k1 = 0.05; k2 = 0.2;

- Take the log on both sides, we have

$$\begin{aligned}
&\ln P(t,T_2)\\
&= \ln P(t,T_1) + \ln A(t,T_2) - \ln A(t,T_1) - r_t(B(t,T_2) - B(t,T_1))\\
&= \ln P(t,T_1) + \ln A(t,T_2) - \ln A(t,T_1) + \frac{\ln P(t,T_1) - \ln A(t,T_1)}{B(t,T_1)}(B(t,T_2) - B(t,T_1))\\
&= \frac{B(t,T_2)}{B(t,T_1)}\ln P(t,T_1) + \ln A(t,T_2) - \ln A(t,T_1) - \frac{\ln A(t,T_1)}{B(t,T_1)}(B(t,T_1) - B(t,T_2))
\end{aligned}$$

- We can see the $\ln P(t,T_2)$ and $\ln P(t,T_1)$ are linear deterministic function of each other; therefore
$$corr(\ln P(t,T_1), \ln P(t,T_2)) = \pm 1.$$

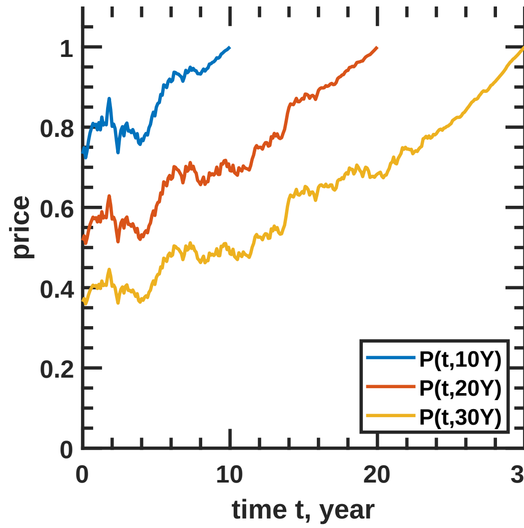


Figure 3.5.8: One realization of zero-coupon bond price evolution until maturity. It is clear that the price movement of bonds(10Y, 20Y, and 30Y) are similar to each other.

3.5.8.5 *Multi-factor Gaussian model*

3.5.8.6 *The model*

Definition 3.5.11 (multi-factor Gaussian additive model). *The two-factor OU process is given by the following SDE*

$$\begin{aligned} r(t) &= x_1(t) + x_2(t) + \psi(t), r(0) = r_0 \\ dx_1(t) &= -a_1 x_1(t)dt + \sigma_1 dW_1(t), x_1(0) = x_{10} \\ dx_2(t) &= -a_2 x_2(t)dt + \sigma_2 dW_1(t), x_2(0) = x_{20} \\ &\cdots\cdots \\ dx_n(t) &= -a_n x_2(t)dt + \sigma_n dW_n(t), x_n(0) = x_{n0} \end{aligned}$$

where $a_1, a_2, \sigma_1, \sigma_2, r_0$ are constant, $\psi(t)$ is a time-dependent function, and W_1, W_2 are correlated Brownian motion such that

$$dW_1(t)dW_2(t) = \rho dt.$$

Lemma 3.5.33 (zero-coupon bond price in multi-factor Gaussian additive model). *Consider a multi-factor Gaussian additive short rate model in Definition 3.5.11.*

It follows that

- $$\begin{aligned} P(t,T) &\triangleq E[\exp(-\int_t^T r(s)ds)] \\ &= \prod_{i=1}^n F_i(t,T,x_i(t))\exp(U(t,T))\exp(-\int_t^T \psi(u)du). \end{aligned}$$

where

$$\begin{aligned} F_i(t,T,x_1(t)) &\triangleq E[\exp(-\int_t^T x_1(s)ds)] \\ &= \exp(-\frac{1-\exp(-a_1(T-t))}{a_1}x_1(t) + \frac{1}{2}(\frac{\sigma^2}{a_1^2}(T-t+ \\ &\frac{2}{a_1}\exp(-a_1(T-t)) - \frac{1}{2a_1}\exp(-2a_1(T-t)) - \frac{3}{2a_1}))), i = 1,2,...,n. \end{aligned}$$

and

$$\begin{aligned} U(t,T) = \sum_{1\leq i<j\leq n} \frac{\rho_{ij}\sigma_i\sigma_j}{a_i a_j}(T-t+\frac{\exp(-a_i(T-t))-1}{a_i} + \frac{\exp(-a_j(T-t))-1}{a_j} \\ + \frac{\exp(-(a_i+a_j)(T-t))-1}{a_i+a_j}). \end{aligned}$$

- *If we choose $\psi(T), T \in \mathbb{R}^+$ to match the current term structure given by $f^M(0,T)$, then*

$$\begin{aligned} \psi(T) = f^M(0,T) + \sum_{i=1}^n \frac{\sigma_i^2}{2a_i^2}(1-\exp(-a_iT))^2 \\ + \sum_{1\leq i<j\leq n} \rho_{ij}\frac{\sigma_i\sigma_j}{a_ia_j}(1-\exp(-a_iT))(1-\exp(-a_jT)) \end{aligned}$$

- *The zero-coupon bond price, which agrees with the current term structure, is given by*

$$\exp(-\int_t^T \psi(u)du) = \frac{P^M(0,T)}{P^M(0,t)}\exp(-\frac{1}{2}(V(0,T)-V(0,t)).$$

- *With the choice of ϕ in (2), we have*

$$P(t,T) = \frac{P^*(0,T)F(0,t,x_{10},x_{20},...,x_{n0})}{P^*(0,t)F(0,T,x_{10},x_{20},x_{n0})}F(t,T,x_1(t),x_2(t),...,x_n(t)),$$

where $P^*(0,T)$ is the current market term structure and

$$F(t,T,x_1(t),x_2(t),...,x_n(t)) = \prod_{i=1}^{n} F_i(t,T,x_i(t)) \exp(U(t,T))$$

Proof. See Lemma 3.5.31. ☐

3.6 Forward rate model: HJM framework

The Heath—Jarrow—Morton (HJM)[10] framework is a general framework to model the evolution of interest rate curve by specifying the dynamics of instantaneous forward rate curve. The key to these techniques is the recognition that the drifts of the no-arbitrage evolution of certain variables can be expressed as functions of their volatilities and the correlations among themselves. In other words, no drift estimation is needed.

When the volatility and drift of the instantaneous forward rate are assumed to be deterministic, this is known as the Gaussian Heath—Jarrow—Morton (HJM) model of forward rates. We can see that short rate model is a dynamical reduction of the HJM model.

3.6.1 Single factor HJM framework

3.6.1.1 *Principles*

Proposition 3.6.1 (HJM framework for one factor forward rate dynamics). *[1, p. 69] Assume that the forward rate $f(t,T)$, **under risk-neutral measure** Q,is given*

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t),$$

where $W(t)$ is a Brownian motion, and $\sigma(t,T)$ is adapted to filtration $\mathcal{F}_t$ for each $T > 0$, and

$$\int_0^T \int_0^T |\alpha(s,u)|\, dsdu < \infty, \int_0^T (\int_0^T |\sigma(s,u)|^2\, ds)^{1/2} du < \infty$$

almost surely. Then,

$$\alpha(t,T) = \sigma(t,T) \int_t^T \sigma(t,u)du.$$

As a consequence,

$$f(t,T) = f(0,T) + \int_0^t \sigma(s,T)(\int_s^T \sigma(s,u)du)dt + \int_0^t \sigma(s,T)dW(s).$$

Furthermore, the zero-coupon bond price $P(t,T)$ satisfies SDE

$$dP(t,T) = r(t)P(t,T)dt + \Sigma(t,T)P(t,T)dW(t)$$

with log-volatility

$$\Sigma(t,T) = -\int_t^T \sigma(t,u)du.$$

Proof. Under risk-neutral measure Q, the bond price is related to forward rate as

$$P(t,T) = \exp(-\int_t^T f(t,u)du).$$

Use

$$\begin{aligned} d(\int_t^T f(t,u)du) &= -f(t,t)dt + \int_t^T (\alpha(t,u)dt + \sigma(t,u)dW(t))du \\ &= -r(t)dt + (\int_t^T \alpha(t,u)du)dt + (\int_t^T \sigma(t,u)du)dW(t) \end{aligned}$$

Then

$$\begin{aligned} dP(t,T) &= d(\exp(-\int_t^T f(t,u)du)) \\ &= P(t,T)d(\int_t^T f(t,u)du) \\ &= (r(t) - \int_t^T \alpha(t,u)du)du)^2)P(t,T)dt - (\int_t^T \sigma(t,u)du)P(t,T)dW(t) \\ &+ \frac{1}{2}(\int_t^T \sigma(t,u)du)dW(t)(\int_t^T \sigma(t,u)du)dW(t) \\ &= (r(t) - \int_t^T \alpha(t,u)du)du)^2)P(t,T)dt - (\int_t^T \sigma(t,u)du)P(t,T)dW(t) \\ &+ \frac{1}{2}(\int_t^T \sigma(t,u)du)(\int_t^T \sigma(t,u)du)dt \end{aligned}$$

Because under risk-neutral measure Q, $P(t,T)$ will have drift $r(t)$(Lemma 3.5.3). Therefore,

$$r(t) - \int_t^T \alpha(t,u)du + \frac{1}{2}(\int_t^T \sigma(t,u)du)^2 = r(t),$$

Then
$$\int_t^T \alpha(t,u)du = \frac{1}{2}(\int_t^T \sigma(t,u)du)^2.$$
Differentiating both sides with respect to T, we have
$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,u)du.$$
□

Remark 3.6.1 (current term structure always matched). From the HJM result, we have the forward curve evolution function as
$$f_{predict}(t,T) = f_{observed}(0,T) + \int_0^t \alpha(s,T)(\int_s^T \sigma(s,u)du)dt + \int_0^t \sigma(s,T)dW(s).$$

If we set $t=0$, then we have
$$f_{predict}(t,T) = f_{observed}(0,T);$$
that is, our prediction on current term structure is consistent with the observation.

Remark 3.6.2 (calculating the short rate dynamics). We can obtain the short rate using the definition
$$r(t) \triangleq f(t,t)$$
to obtain
$$r(t) = f(0,t) + \int_0^t \alpha(s,t)(\int_s^t \sigma(s,u)du)dt + \int_0^t \sigma(s,t)dW(s).$$

Example 3.6.1 (parallel shift arbitrage). Suppose current forward rate curve is given by $f(0,T), T>0$. At a future time $t>0$, the forward curve is predicted(using model) to be **deterministically** given by $f(t,T) = f(0,T)+\epsilon, \epsilon > 0$. Then the parallel shift of the forward rate is known to adopt arbitrage([11]).

From the HJM no-arbitrage condition, it is immediately clear that for a deterministic dynamics(i.e. volatility is zero), the drift has to be 0; that is, the forward curve should remain constant.

Lemma 3.6.1 (correlation structure in one-factor HJM forward model). *Let* $T_1, T_2, ..., T_N$ *be two maturity dates. Then*

- *The* N *dimensional stochastic process* $(f(t,T_1), f(t,T_2), ..., f(t,T_N))$ *has each component being a Gaussian process, but might not be joint Gaussian*

- *The pair correlation is given by*

$$Cov[f(t,T_1), f(t,T_2)] = \int_0^t \sigma(s,T_1)\sigma(s,T_2)ds.$$

Proof. (1)The one-factor HJM forward rate model, for all $T > 0$,

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW_t$$

is a state-independent linear SDE, thus Gaussian process(**??**). Consider $\sigma(t,T_1) = \sigma, \sigma(t,T_2) = \sigma$, then $f(t,T_1) = f(t,T_2)$ such that $(f(t,T_1), f(t,T_2))$ is not a joint normal(their covariance matrix is singular). (2) Using Ito isometry. □

Lemma 3.6.2 (forward rate dynamics under forward measure). *[4, p. 116] The forward rate dynamics under $T-$maturity forward measure Q_T is given by*

$$df(t,T) = \sigma(t,T)dW^T(t),$$

where $W^T(t)$ is a Brownian motion under Q_T.

Proof. Under risk-neutral measure Q, the forward rate dynamics (Theorem 3.6.1)is given by

$$df(t,T) = -\sigma(t,T)\Sigma(t,T)dt + \sigma(t,T)dW(t),$$

where

$$\Sigma(t,T) = -\int_t^T \sigma(t,u)du.$$

Under the forward measure Q_T,

$$dW(t) = dW(t)^T + \Sigma(t,T)dt,$$

where $W^T(t)$ is a Brownian motion under forward measure.

Then

$$df(t,T) = -\sigma(t,T)\Sigma(t,T)dt + \sigma(t,T)(dW^T(t) + \Sigma(t,T)dt) = \sigma(t,T)dW^T(t).$$

□

Note 3.6.1 (single-factor model as a reduced form model). Note that we can also write the single-factor model by

$$df(t,T) + \alpha(t,T)dt + \sigma(t,T)dZ^T(t),$$

where we have a set(infinitely many) of Brownian motions $Z^T(t)$ indexed by the maturity T. By requiring $f(t,T)$ is driven by 1 factors, we are requiring

$$Rank(dZ^T[dZ^T]^T) = 1.$$

3.6.1.2 *Gaussian HJM model*

Lemma 3.6.3 (Reduce to Ho-Lee model). *[1, p. 71] Let $\sigma(t,T) = \sigma$ for all $t \leq T$. Then*

$$\alpha(t,T) = \sigma \int_t^T \sigma du = \sigma^2(T-t),$$

and

$$\begin{aligned} f(t,T) &= f(0,T) + \int_0^t \alpha(s,T)ds + \int_0^t \sigma(s,T)dW(s) \\ &= f(0,T) + \frac{1}{2}\sigma^2 t(2T-t) + \sigma W(t). \end{aligned}$$

Then the short rate is given by

$$r(t) = f(t,t) = f(0,t) + \frac{1}{2}\sigma^2 t^2 + \sigma W(t)$$

and

$$dr(t) = (\frac{\partial f(0,t)}{\partial t} + \sigma^2 t)dt + \sigma dW(t).$$

By Ho-Lee model zero coupon bond price Lemma 3.5.23, we can calculate the zero-coupon bond price as

$$P(t,T) = \exp(-(T-t)(r(t) - f(0,t)) - \int_t^T f(0,s)ds - \frac{1}{2}\sigma^2 t(T-t)^2).$$

Moreover, it can be showed that this prediction is consistent with current term structure, showed by

$$P(0,T) = \exp(-\int_0^T f(0,s)ds).$$

Lemma 3.6.4 (covariance in parallel shift forward rate dynamics). *Consider the Ho-Lee model for forward rate, which is given by*

$$df(t,T) = \sigma^2(T-t)dt + \sigma^2 dW_t,$$

and

$$f(t,T) = f(0,T) + \frac{1}{2}\sigma^2 t(2T-t) + \sigma W(t).$$

Then for given t_0 and T_0,

- $f(t_0, T_0)$ *and* $f(t_0+s, T_0+s), s \geq 0$ *have the covariance given by* $\sigma^2(t_0)$, *even though they might have different means.*
- $f(t_0, T_0)$ *and* $f(t_0+s, T_0+s), s \geq 0$ *have the correlation given by*

$$\rho = \frac{t_0}{\sqrt{t_0(t_0+s)}}$$

- $f(t_0, T_0)$ *and* $f(t_1, T_1)$ *have the covariance given by* $\sigma^2 \min(t_0, t_1)$, *even though they might have different means.*

Proof. (1)(2)We can show that

$$f(t_0, T_0) = f(0, T_0) + \frac{1}{2}\sigma^2 t_0(2T_0 - t_0) + \sigma W(t_0).$$

and

$$f(t_0+s, T_0+s) = f(0, T_0) + \frac{1}{2}\sigma^2(t_0+s)(2T_0 - t_0 + s) + \sigma W(t_0).$$

It is easy to see that they might have different means, but the covariance is given by $\sigma^2 t_0$, independent of s.

(2) similar to (1). The demean portion of $f(t_0, T_0)$ and $f(t_1, T_1)$ are given by $\sigma W(t_0)$ and $\sigma W(t_1)$.

□

Remark 3.6.3 (slow decaying correlation and non-stationary covariance)**.**

- Note that the time series $G(s) = f(t_0+s, T_0+s)$ has a slowly decaying correlation, whereas in Hull white model, the correlation is exponentially decaying.
- The time series $G(s) = f(t_0+s, T_0+s)$ has a non-stationary covariance structure. For example, $Cov(G(s_1), G(0_2))$ is not solely dependent on the difference $s_1 - s_2$.

Lemma 3.6.5 (Reduce to Hull-White model). *[1, p. 71] Take $\sigma(t,T) = \sigma(t)e^{-\alpha(T-t)}$, where $\sigma(t)$ is a deterministic function and α is a constant.*

Then, by Theorem 3.6.1, we have the drift term given by

$$\begin{aligned}\alpha(t,T) &= \sigma(t,T)\int_t^T \sigma(t,u)du \\ &= \sigma^2(t)e^{-\alpha(T-t)}D(t,T),\end{aligned}$$

where $D(t,T)$ is given by $D(t,T) \triangleq \int_t^T e^{-\alpha(s-t)}ds$

It follows that

- *The forward curve is*

$$\begin{aligned}f(t,T) &= f(0,T) + \int_0^t \alpha(s,T)ds + \int_0^t \sigma(s,T)dW(s) \\ &= f(0,T) + \int_0^t \sigma^2(s)e^{-\alpha(T-s)}D(s,T)ds + \int_0^t \sigma(s)e^{-\alpha(T-s)}dW(s).\end{aligned}$$

- *The short rate is*

$$r(t) = f(t,t) = f(0,t) + \int_0^t \sigma^2(s)e^{-\alpha(t-s)}D(s,T)ds + \int_0^t \sigma(s)e^{-\alpha(t-s)}dW(s).$$

- *The short rate model(matching the current term structure)satisfies SDE*

$$dr(t) = (\theta(t) - \alpha r(t))dt + \sigma(t)dW(t),$$

with

$$\theta(t) = \frac{\partial f(0,t)}{\partial t} + \alpha f(0,t) + \int_0^t \sigma(s)^2 e^{-2\alpha(t-s)}ds.$$

Proof. (1)(2) are straight forward. (3) Note that

$$\begin{aligned}dr(t) &= \frac{\partial f(0,t)}{\partial t}dt + \int_0^t \sigma^2(s)\frac{d}{dt}(e^{-\alpha(t-s)D(s,t)})ds + \sigma(t)dW(t) + \int_0^t \sigma\frac{d}{dt}(e^{-\alpha(t,s)})dW(s)) \\ &= \frac{\partial f(0,t)}{\partial t}dt + \sigma(t)dW(t) - \alpha\int_0^t \sigma(s)e^{-\alpha(T-s)}dW(s) \\ &+ \int_0^t \sigma^2(s)(-\alpha e^{-\alpha(t-s)}D(s,t) + e^{-2\alpha(t-s)})ds\end{aligned}$$

where we use

$$\frac{d}{dt}D(s,t) = e^{-\alpha(t-s)}.$$

Further we use the relation

$$\alpha f(0,t) - \alpha r(t) = -\alpha \int_0^t \sigma^2(s) e^{-\alpha(t-s)} D(s,T) ds - \alpha \int_0^t \sigma(s) e^{-\alpha(t-s)} dW(s)$$

to get to the result. □

Lemma 3.6.6 (covariance in parallel shift forward rate dynamics). *Consider the Hull-White model for forward rate, which is given by*

$$f(t,T) = f(0,T) + \int_0^t \sigma^2(s) e^{-\alpha(T-s)} D(s,T) ds + \int_0^t \sigma(s) e^{-\alpha(T-s)} dW(s).$$

Then for given t_0 and T_0,

- $f(t_0, T_0)$ *and* $f(t_0+s, T_0+s), s \geq 0$ *have the covariance given by*

$$\exp(-\alpha s) \int_0^{t_0} \sigma^2(u) \exp(-2\alpha(T_0 - u) du),$$

even though they might have different means.
- $f(t_0, T_0)$ *and* $f(t_0+s, T_0+s), s \geq 0$ *have the correlation given by*

$$\rho = \frac{t_0}{\sqrt{t_0(t_0+s)}}$$

- $f(t_0, T_0)$ *and* $f(t_1, T_1)$ *have the covariance given by* $\sigma^2 \min(t_0, t_1)$*, even though they might have different means.*

Proof. (1)(2)We can show that

$$f(t_0, T_0) = f(0, T_0) + \frac{1}{2}\sigma^2 t_0 (2T_0 - t_0) + \sigma W(t_0).$$

and

$$f(t_0+s, T_0+s) = f(0, T_0) + \frac{1}{2}\sigma^2 (t_0+s)(2T_0 - t_0 + s) + \sigma W(t_0).$$

It is easy to see that they might have different means, but the covariance is given by $\sigma^2 t_0$, independent of s.

$$\begin{aligned}
\rho &= \frac{\exp(-\alpha s)\int_0^{t_0}\sigma^2(u)\exp(-2\alpha(T_0-u)du)}{\sqrt{\int_0^{t_0}\sigma^2(u)\exp(-2\alpha(T_0-u)du)\int_0^{t_0+s}\sigma^2(u)\exp(-2\alpha(T_0+s-u)du)}} \\
&= \frac{\exp(-\alpha s)\int_0^{t_0}\sigma^2(u)\exp(-2\alpha(T_0-u)du)}{\sqrt{\int_0^{t_0}\sigma^2(u)\exp(-2\alpha(T_0-u)du)\exp(-2\alpha s)\int_0^{t_0+s}\sigma^2(u)\exp(-2\alpha(T_0-u)du)}} \\
&= \frac{1}{1+} \\
&= \frac{1}{\sqrt{\int_{t_0}^{t_0+s}\sigma^2(u)\exp(-2\alpha(T_0-u)du)}} \\
&= \frac{1}{\sqrt{\int_0^{s}\sigma^2(u)\exp(-2\alpha(T_0-t_0-u)du)}} \\
&= \frac{1}{\sqrt{\exp(-2\alpha(T_0-t_0))\int_0^{s}\sigma^2(u)\exp(2\alpha u)du)}}
\end{aligned}$$

(3) similar to (1). The demean portion of $f(t_0, T_0)$ and $f(t_1, T_1)$ are given by $\sigma W(t_0)$ and $\sigma W(t_1)$.

□

Remark 3.6.4 (slow decaying correlation and stationary correlation, non-stationary covariance)**.**

- Note that the time series $G(s) = f(t_0 + s, T_0 + s)$ has a slowly decaying correlation, whereas in Hull white model, the correlation is exponentially decaying.
- The time series $G(s) = f(t_0 + s, T_0 + s)$ has a stationary covariance structure. Note that $Cov(G(s_1), G(0_2))$ is not solely dependent on the difference $s_1 - s_2$.

3.6.1.3 *Option pricing in Gaussian HJM models*

Lemma 3.6.7 (Pricing call option on zero coupon bond). *Given the volatility structure $\sigma(t, u)$ in HJM model, the zero coupon bond dynamics under risk-neutral measure Q is given by*

$$dP(t, T) = r(t)P(t, T)dt + \Sigma(t, T)P(t, T)dW_t,$$

where $\Sigma(t, T) = -\int_t^T \sigma(t, u)du$.

The value at time t of a European call option maturing at S with strike K on a bond maturing at $T > S$ is given by

$$V(t) \triangleq BC(t; S, T, K) = P(t,T)N(d_+) - KP(t,S)N(d_-),$$

where

$$d_+ = \frac{\log \frac{FP(t;S,T)}{K} + \frac{1}{2}v(t,S)^2}{v(t,S)} = \frac{\log \frac{P(t,T)}{KP(t,S)} + \frac{1}{2}v(t,S)^2}{v(t,S)}, d_- = d_+ - v(t,S).$$

and

$$v(t,S)^2 = \int_t^S (\Sigma(u,T) - \Sigma(u,S))^2 dt$$

Proof. See Theorem 3.6.1 and Lemma 3.3.10. □

3.6.2 Multi-factor HJM framework

Proposition 3.6.2. *[1, p. 69] Assume that the forward rate $f(t,T)$, under risk-neutral measure Q,is given*

$$df(t,T) = \alpha(t,T)dt + \sum_{i=1}^{n} \sigma_i(t,T)dW_i(t),$$

where $W_1(t), ..., W_n$ are independent Brownian motions, and $\sigma_1(t,T), ..., \sigma_n(t,T)$ are adapted to filtration $\mathcal{F}_t$ for each $T > 0$, and

$$\int_0^T \int_0^T |\alpha(s,u)| \, dsdu < \infty, \int_0^T (\int_0^T |\sigma_i(s,u)|^2 ds)^{1/2} du < \infty, \forall i$$

almost surely. Then,

$$\alpha(t,T) = \sum_{i=1}^{n} (\sigma_i(t,T) \int_t^T \sigma_i(t,u))du.$$

As a consequence,

$$f(t,T) = f(0,T) + \int_0^t \sum_{i=1}^{n} (\alpha_i(s,T)(\int_s^T \sigma_i(s,u)du))dt + \int_0^t \sum_{i=1}^{n} \sigma_i(s,T)dW_i(s).$$

Furthermore, the zero-coupon bond price $P(t,T)$ satisfies SDE

$$dP(t,T) = r(t)P(t,T)dt + \sum_{i=1}^{n} \Sigma_i(t,T)P(t,T)dW_i(t)$$

with log-volatility

$$\Sigma(t,T) = -\int_t^T \sigma_i(t,u)du.$$

Proof. We use the notation $\sigma = (\sigma_1, \sigma_2, ..., \sigma_n)^T, dW_t = (dW_1, dW_2, ..., dW_n)^T$. Under risk-neutral measure Q, the bond price is related to forward rate as

$$P(t,T) = \exp(-\int_t^T f(t,u)du).$$

Use

$$\begin{aligned} d(\int_t^T f(t,u)du) &= -f(t,t)dt + \int_t^T (\alpha(t,u)dt + \sigma^T dW(t))du \\ &= -r(t)dt + (\int_t^T \alpha(t,u)du)dt + (\int_t^T \sigma^T du)dW(t) \end{aligned}$$

Then

$$\begin{aligned} dP(t,T) &= d(\exp(-\int_t^T f(t,u)du)) \\ &= P(t,T)d(\int_t^T f(t,u)du) \\ &= (r(t) - \int_t^T \alpha(t,u)du)du)^2)P(t,T)dt - (\int_t^T \sigma^T(t,u)du)P(t,T)dW(t) \\ &+ \frac{1}{2}(\int_t^T \sigma^T(t,u)du)dW(t)[(\int_t^T \sigma^T(t,u)du)dW(t)]^T \\ &= (r(t) - \int_t^T \alpha(t,u)du)du)^2)P(t,T)dt - (\int_t^T \sigma^T(t,u)du)P(t,T)dW(t) \\ &+ \frac{1}{2}(\int_t^T \sigma^T(t,u)du)(\int_t^T \sigma(t,u)du)dt \end{aligned}$$

Because under risk-neutral measure Q, $P(t,T)$ will have drift $r(t)$(Lemma 3.5.3). Therefore,

$$r(t) - \int_t^T \alpha(t,u)du + \frac{1}{2}\left\|(\int_t^T \sigma(t,u)du)\right\|^2 = r(t),$$

Then

$$\int_t^T \alpha(t,u)du = \frac{1}{2}\left\|\left(\int_t^T \sigma(t,u)du\right)\right\|^2.$$

Differentiating both sides with respect to T, we have

$$\alpha(t,T) = \sigma(t,T)^T \int_t^T \sigma(t,u)du.$$

□

Lemma 3.6.8 (correlation structure in one-factor HJM forward model). *Let $T_1, T_2, ..., T_N$ be two maturity dates. Then*

- *The pair correlation is given by*

$$Cov[f(t,T_1), f(t,T_2)] = \sum_{i=1}^n \int_0^t \sigma_i(s,T_1)\sigma_i(s,T_2)ds.$$

- *The instantaneous correlation is given by*

$$Cov[df(t,T_1), df(t,T_2)] = \sum_{i=1}^n \sigma_i(t,T_1)\sigma_i(t,T_2)dt.$$

Proof. Using Ito isometry. □

Note 3.6.2 (multifactor model as a reduced form model)**.** Note that we can also write the multifactor model by

$$df(t,T) + \alpha(t,T)dt + \sigma(t,T)dZ^T(t),$$

where we have a set(infinitely many) of Brownian motions $Z^T(t)$ indexed by the maturity T. By requiring $f(t,T)$ is driven by n factors, we are requiring

$$Rank(dZ^T[dZ^T]^T) = n.$$

3.6.3 Forward rate model calibration

Lemma 3.6.9 (calibrating a linear interpolating forward curve). *[4, p. 96] Suppose we are given by M yield, denoted by $R(0,T_1), R(0,T_2), ..., R(0,T_M)$, for M maturity dates, where $0 < T_1 < ... < T_M$.*

Further assume the forward curve is taking the following function form:

$$f(0,T) = R(0,T_1), \forall 0 \leq T < T_1$$
$$f(0,T) = f(0,T_{i-1}) + \alpha_i(T - T_{i-1}), \forall T_{i-1} \leq T < T_i, \forall i = 2,...,M$$

The calibrating strategy is:

- *Calculating zero-coupon price $P(t,T_i), \forall i = 1,2,...,M$.*
- *Solve $\alpha_i, i = 2,3,...,M$ from the following equation:*

$$\begin{aligned} P(0,T_i) &= P(0,T_{i-1})\exp(-\int_{T_{i-1}}^{T_i} f(0,s)ds) \\ &= P(0,T_{i-1})\exp(-\int_{T_{i-1}}^{T_i} f(0,T_{i-1}) + \alpha_i(s - T_{i-1})ds) \\ &= P(0,T_{i-1})\exp(-f(0,T_{i-1})(T_i - T_{i-1}) + \frac{1}{2}\alpha_i(T_i - T_{i-1})^2) \end{aligned}$$

Eventually, we have a linear interpolating $f(0,T), 0 \leq T \leq T_M$.

3.6.4 Connection to short-rate model

3.6.4.1 *From short-rate model to forward-rate model*

Proposition 3.6.3. *[4, p. 139] Consider an arbitrage-free short-rate model given by*

$$dr_t = a(t,r_t)dt + b(t,r_t)dW_t.$$

Then it can be converted an HJM model with forward-rate volatility given by

$$\sigma(t,T) = \rho(r_t,t)\frac{\partial^2 g}{\partial x \partial T}(r_t,t,T)$$

where

$$g(x,t,T) = -\ln P(t,T) = -\ln E_Q[\exp(-\int_t^T r_s ds)|r_t = x],$$

Proof. Note that

$$f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T} = \frac{\partial g}{\partial T}(r_t,t,T).$$

Hence, the process for the instantaneous forward rate is

$$\begin{aligned} f(t,T) &= \frac{\partial f}{\partial x} dr_t + (\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2})dt \\ &= \rho \frac{\partial^2 g}{\partial x \partial T} dW_t + \ drift\ term \end{aligned}$$

□

Example 3.6.2 (Hull-White model to HJM model). Consider the Hull-White model given by

$$dr_t = \kappa(\theta_t - r_t)dt + \sigma dW_t.$$

Then

$$g(r_t, t, T) = -\ln P(t,T) = B(t,T)r_t - \ln A(t,T),$$

where

$$B(t,T) = \frac{1 - \exp(-\kappa(T-t))}{\kappa}.$$

Then

$$\begin{aligned} \frac{\partial^2 g}{\partial r_t \partial T}(r_t, t, T) &= \frac{\partial B(t,T)}{\partial T} = \exp(-\kappa(T-t)) \\ \sigma(t,T) &= \sigma \frac{\partial^2 g}{\partial x \partial T}(r_t, t, T) = \sigma \exp(-\kappa(T-t)) \end{aligned}$$

Example 3.6.3 (Ho-Lee model). The Ho-Lee model for short rate, under the risk-neutral measure Q, is given by

$$dr = \theta(t)dt + \sigma dW_t$$

where W_t is the Brownian motion, and σ is a constant.

3.6.4.2 *Criterion for a Markovian model*

Note 3.6.3. [4, p. 138] Note that the short-rate model is connected to the forward-rate model via

$$dr_t = \frac{\partial f(t,T)}{\partial t}|_{T=t} dt + \sigma^T(t,t) dW_t,$$

where $\sigma(t,t) \in \mathbb{R}^n$ and W_t is n-dimensional Brownian motion, and

$$\frac{\partial f(t,T)}{\partial t}|_{T=t}$$

Proposition 3.6.4. *[4, p. 139] Suppose that the forward-rate volatility satisfies*

$$\frac{\partial \sigma_i(t,T)}{\partial T} = -\kappa_i(T)\sigma_i(t,T), i = 1,2,...,n, (*)$$

for some deterministic function $\kappa_i(T)$. Then

$$\begin{aligned} d\psi_i(t) &= (\sigma_i^2(t,t) - 2\kappa_i(t)\psi_i(t))dt \\ d\chi_i(t) &= (\psi_i(t) - \kappa_i(t)\chi_i(t))dt + \sigma_i(t,t)dW_i(t), \end{aligned}$$

for $i = 1,...,n$.

Proof. For $\psi_i(t)$, there is

$$\begin{aligned} d\psi_i(t) &= \sigma_i^2(t,t)dt + (\int_0^t 2\sigma_i(s,t))dt \\ &= \sigma_i^2(t,t)dt - 2\kappa_i(t)(\int_0^t \sigma_i^2(s,t)ds)dt \\ &= (\sigma_i^2(t,t) - 2\kappa_i(t)\psi_i(t))dt, \end{aligned}$$

while for $\chi_i(t)$, we have noticing that $b_i(t,t) = 0$,

$$d\chi_i(t) = (\int_0^t \frac{\partial b_i(s,t)}{\partial t}ds + \frac{\sigma_i(s,t)}{\partial t}dW_i(s)) + \sigma_i(t,t)dW_i(t)\ (**).$$

Because of equation $(*)$ and

$$\begin{aligned} \frac{\partial b_i(s,t)}{\partial t} &= \frac{\sigma_i(s,t)}{\partial t}\int_s^t \sigma_i(s,u)du + \sigma_i^2(s,t) \\ &= -\kappa_i(t)b_i(s,t) + \sigma_i^2(s,t) \end{aligned}$$

we can write Equation $(**)$ as

$$\begin{aligned} d\chi_i(t) &= -\kappa_i(t)(\int_0^t b_i(s,t)ds + \sigma_i(s,t)dW_i(s))dt \\ &\quad + (\int_0^t \sigma_i^2(s,t)ds)dt + \sigma_i(t,t)dW_i(t) \\ &= (\psi_i(t) - \kappa_i(t)\chi_i(t))dt + \sigma_i(t,t)dW_i(t). \end{aligned}$$

□

Remark 3.6.5 (interpretation)**.**

- The connection is established by the price of the zero-coupon bonds resulted from the two types of models.
- An arbitrage-free HJM model is completed by determined by its forward volatility(Theorem 3.6.1).

Proposition 3.6.5. *[4, p. 142] Suppose that the short-rate volatility, $\sigma(t,t)$, is a deterministic function of time. Then a necessary condition for the short rate to be Markovian is*

$$\frac{\sigma_i(t,T)}{\partial T} = -\kappa_i(T)\sigma_i(t,T),$$

for some scalar function, $\kappa_i(T), 1 \leq i \leq n$.

Proof. For the short rate to be Markovian, we require that $r_T - r_t$ depends only on r_t and $\{dW_s, s \in (t,T), dW_s \in \mathbb{R}^n\}$. In fact, we have

$$\begin{aligned} r_T - r_t &= f(0,T) - f(0,t) + \int_0^T \sigma^T(s,T)\Sigma(s,T)ds \\ &\quad - \int_0^t \sigma^T(s,t)\Sigma(s,t)ds + \int_0^T \sigma^T(s,T)dW_s - \int_0^t \sigma^T(s,t)dW_s \\ &= f(0,T) - f(0,t) + \int_0^T \sigma^T(s,T)\Sigma(s,T)ds \\ &\quad - \int_0^t \sigma^T(s,t)\Sigma(s,t)ds + \int_t^T \sigma^T(s,T)dW_s + \int_0^t (\sigma(s,T) - \sigma(s,t))^T dW_s \end{aligned}$$

The last term in the above equation cannot depend on $\{dW_s, s \in (t,T), dW_s \in \mathbb{R}^n\}$, so it can depend only on r_t. Because

$$\int_0^t \sigma^T(s,t)dW_s = r_t + deterministic\ functions,$$

we conclude that

$$\int_0^t \sigma^T(s,t)dW_s$$

is also a deterministic function of r_t. Hence, we have correlation relationship

$$Corr[\int_0^t \sigma^T(s,t)dW_s, \int_0^t \sigma^T(s,t)dW_s] = 1.$$

The last equality can be rewritten into

$$E_Q[\int_0^t \sigma^T(s,T)dW_s \times \int_0^t \sigma^T(s,t)dW_s] = (E_Q[(\int_0^t \sigma^T(s,T)dW_s)^2])^{1/2}(E_Q[(\int_0^t \sigma^T(s,t)dW_s)^2])^{1/2}.$$

By Ito isometry, the above implies that

$$\left|\int_0^t \sigma^T(s,T)\sigma(s,t)ds\right| = (\int_0^t \|\sigma(s,T)\|^2 ds)^{1/2}(\int_0^t \|\sigma(s,t)\|^2 ds)^{1/2},$$

that is, the equality is achieved in the Cauchy equality, and the equality holds if and only if

$$\sigma(s,t) = \alpha(t,T)\sigma(s,T), 0 \leq s \leq t,$$

for some deterministic scalar function α.

Similarly, we also have

$$\sigma(s,t) = \alpha(t,T')\sigma(s,T'), 0 \leq s \leq t,$$

for some other T'. Assume that $\sigma_i(s,t) \neq 0$, we then have

$$\frac{\sigma_i(s,T)}{\sigma_i(s,T')} = \frac{\alpha(t,T')}{\alpha(t,T)} = \frac{\alpha(0,T')}{\alpha(0,T)}, i = 1,2,...,n.$$

Making $T' = s$, we have then proved that $\sigma_i(s,T)$ can be factorized as

$$\sigma_i(s,T) = x_i(s)y_i(T).$$

By differentiating the above equation with respect to T, we obtain

$$\frac{\partial\sigma_i(s,T)}{\partial T} = x_i(s)y_i(T)\frac{\partial \ln y_i(T)}{\partial T}.$$

Denote $\frac{\partial \ln y_i(T)}{\partial T}$ by $-\kappa_i(T)$, we complete the proof. □

Proposition 3.6.6 (zero coupon bond price). *[4, p. 140]*

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp(-\int_t^T (\sum_{i=1}^n \int_0^t b_i(s,u)ds + \sigma_i(s,u)dW_i(s))du)$$

Proof. (1) For the first term in the exponent, we have

□

3.7 LIBOR market models

The LIBOR market model, also known as the BGM Model (Brace Gatarek Musiela Model, in reference to the names of some of the inventors) is a financial model of interest rates.[1] It is used for pricing interest rate derivatives, especially exotic derivatives like Bermudan swaptions, ratchet caps and floors, target redemption notes, autocaps, zero coupon swaptions, constant maturity swaps and spread options, among many others.

3.7.1 LIBOR market model

3.7.1.1 *The model*

Definition 3.7.1 (LIBOR market model). *[1, p. 84] Consider a set of dates* $0 \leq T_0 < T_1 < \cdots < T_n$ *with accrual periods* $\tau_i = T_i - T_{i-1}$*. Denote the forward rate*

$$F_i(t) \triangleq F(t; T_{i-1}, T_i) = \frac{1}{\tau_i}(\frac{P(t, T_{i-1})}{P(t, T_i)} - 1).$$

In the LIBOR model, we assume the forward rate $F_i(t)$ *under the forward measure of* T_j*-maturity zero-coupon bond are given by*

$$dF_i(t) = \mu_i^j(F_i(t), t)dt + C_i(F_i(t), t)dZ_i^j(t),$$

where $C_i(F_i(t), t)$ *is* ***the instantaneous volatilities*** *of the forward rates* $F_i(t)$ *for each* $i = 1, 2, ..., n$*,* $dZ_i dZ_j = \rho_{i,j}dt$ *and* Z_i^j *is a Brownian motion under forward measure* Q_{T_j}*.*

Note 3.7.1 (relation to HJM framework)**.**

- The dynamics of $F(t; T_1, T_2)$ is the **approximate** average of $f(t, T)$ over period of T_1 and T_2.

$$\begin{aligned} F(t; T_1, T_2) &= \frac{1}{T_2 - T_1}(\frac{P(t, T_1)}{P(t, T_2)} - 1) \\ &= \frac{1}{T_2 - T_1}(\frac{\exp(-\int_t^{T_1} f(t, s)ds)}{\exp(-\int_t^{T_2} f(t, s)ds)} - 1) \\ &= \frac{1}{T_2 - T_1}(\exp(\int_{T_1}^{T_2} f(t, s)ds) - 1) \\ &\approx \frac{1}{T_2 - T_1}\int_{T_1}^{T_2} f(t, s)ds \end{aligned}$$

- If $T_2 \to T_1$, then $F(t; T_1, T_2) = f(t, T_1)$.

Remark 3.7.1 (choice of volatility structure)**.** The volatility function $C_j(F_j(t), t)$ can take different forms. For example,

- (normal model)$C_j(F_j(t), t) = \sigma_j(t)$.
- (CEV model)$C_j(F_j(t), t) = \sigma_j(t)$.
- (lognormal model)$C_j(F_j(t), t) = \sigma_j(t)$.

- (shifted normal model)$C_j(F_j(t), t) = \sigma_j(t)$.

where the function $\sigma_j(t)$ is deterministic, $0 \leq \beta_j \leq 1, \delta_j \geq 0$.

3.7.1.2 *Drifts under different measures*

Lemma 3.7.1 (drifts under different forward measure). *[1, p. 89] The drift terms under different Q_{T_i} is given by*

$$\mu_i^j = \begin{cases} -\sum_{k=i+1}^{j} \frac{\tau_k \rho_{k,i} C_i(t) C_k(t) F_i(t) F_k(t)}{1+\tau_k F_k(t)}, & j > i \\ 0, & i=j \\ \sum_{k=j+1}^{i} \frac{\tau_k \rho_{k,i} C_i(t) C_k(t) F_i(t) F_k(t)}{1+\tau_k F_k(t)}, & i > j \end{cases}$$

In particular, if $C_i(F_i(t), t) = \sigma_i(t) F_i(t)$, we have The drift $\mu_i^j(t)$ of the forward rate $F_i(t)$ are given by

$$\mu_i^j = \begin{cases} -\sum_{k=i+1}^{j} \frac{\tau_k \rho_{k,i} \sigma_i(t) \sigma_k(t)}{1+\tau_k F_k(t)}, & j > i \\ 0, & i=j \\ \sum_{k=j+1}^{i} \frac{\tau_k \rho_{k,i} \sigma_i(t) \sigma_k(t)}{1+\tau_k F_k(t)}, & i > j \end{cases}$$

Proof. The drift μ_i under Q_{T_i} is zero since F_i^i is a martingale. μ_i under measure $Q_{T_j}, j > i$ is given by(Theorem 1.6.17)

$$\begin{aligned} \mu_i^j dt &= dF_i d\log(\frac{P(t, T_j)}{P(t, T_i)}) \\ &= -dF_i d\log(\frac{P(t, T_i)}{P(t, T_j)}) \\ &= -dF_i d\log(\prod_{k=i+1}^{j} (1 + \tau_k F_k)) \\ &= -dF_i \sum_{k=i+1}^{j} \frac{dF_k}{1 + \tau_k F_k} \\ &= -\sum_{k=i+1}^{j} \frac{\tau_k \rho_{k,i} C_i(t) C_k(t)}{1 + \tau_k F_k(t)} dt \end{aligned}$$

where we use the relation

$$\frac{P(t,T_i)}{P(t,T_j)} = \prod_{k=i+1}^{j} (1+\delta_k F_k).$$

□

Lemma 3.7.2 (drifts under terminal measure). *[1, p. 90] The measure associated with the zero-coupon bond $P(t,T_n)$ maturing at time T_n is called **terminal measure**. Under terminal measure, the forward rate dynamics are given by*

$$dF_i(t) = \mu_i^n(t)dt + C_i(F_i(t),t)dZ_i^n(t), \forall i,$$

where

$$\mu_i^n = -\sum_{k=i+1}^{n} \frac{\tau_k \rho_{k,i} C_i(t) C_k(t)}{1+\tau_k F_k(t)},$$

and $Z_1^n(t),...,Z_n^n(t)$ are correlated Brownian motions under measure Q_{T_n} such that $dZ_i^n(t)dZ_j^n = \rho_{i,j}(t)$.

In particular, if $C_i(F_i(t),t) = \sigma_i(t)F_i(t)$, we have

$$dF_i(t) = \mu_i^n(t)dt + \sigma_i(t)F_i(t)dZ_i^n(t), \forall i,$$

where

$$\mu_i^n = -\sum_{k=i+1}^{n} \frac{\tau_k \rho_{k,i} \sigma_i(t)\sigma_k(t) F_i(t) F_k(t)}{1+\tau_k F_k(t)}.$$

Proof. Directly from Lemma 3.7.1. □

Definition 3.7.2 (discrete money market account, spot measure). *For any $t \in [0,T_n]$, let $\alpha(t)$ denote the index of the next reset date at time t, that is $\alpha(t) = \min\{j : t \leq T_j, j = 0,...,n\}$. The **discrete money market account** is defined as*

$$L(t) = P(t, T_{\alpha(t)}) \prod_{k=0}^{\alpha(t)} \frac{1}{P(T_{k-1}, T_k)}.$$

*The measure associated with discrete money market account numeraire is called **spot measure**.*

Remark 3.7.2 (intuition before discrete money account).

- At time 0, we begin by investing one dollar to buy an amount $P(0,T_0)^{-1}$ of the T_0 bonds.

- At time T_0, the bonds will be worth $P(0, T_0)^{-1}$, which we reinvest to buy an amount $P(0, T_0)^{-1}P(T_0, T_1)^{-1}$ of T_1 bonds.
- We continue in this way, reinvesting all the proceeds at each date T_{i-1} into zero-coupon bonds maturing at the next date T_i, for $i = 1, 2, ..., n$.

Lemma 3.7.3 (drifts under spot measure). *[1, p. 90] The measure associated with the zero-coupon bond $P(t, T_n)$ maturing at time T_n is called **terminal measure**. Under terminal measure, the forward rate dynamics are given by*

$$dF_i(t) = \mu_i^n(t)F_i(t)dt + C_i(F_i(t), t)dZ_i^n(t), \forall i,$$

where

$$\mu_i^L(t) = \sum_{k=\alpha(t)+1}^{i} \frac{\tau_k \rho_{k,i} C_i(t) C_k(t)}{1 + \tau_k F_k(t)},$$

and $Z_1^L(t), ..., Z_n^L(t)$ are correlated Brownian motions under measure Q_{T_n} such that $dZ_i^L(t)dZ_j^L = \rho_{i,j}(t)$.

In particular, if $C_i(F_i(t), t) = \sigma_i(t)F_i(t)$, we have

$$dF_i(t) = \mu_i^L(t)dt + \sigma_i(t)F_i(t)dZ_i^n(t), \forall i,$$

where

$$\mu_i^L = \sum_{k=\alpha(t)+1}^{i} \frac{\tau_k \rho_{k,i} \sigma_i(t) \sigma_k(t) F_i(t) F_k(t)}{1 + \tau_k F_k(t)}.$$

Proof. The drift μ_i under Q_{T_i} is zero since F_i^i is a martingale. μ_i under measure $Q_{T_j}, j > i$ is given by(Theorem 1.6.17)

$$\begin{aligned}
\mu_i^j dt &= dF_i d\log(\frac{L(t)}{P(t,T_i)}) \\
&= dF_i d\log(\frac{P(t,T_{\alpha(t)})}{P(t,T_i)}) \\
&= dF_i d\log(\prod_{k=\alpha(t)+1}^{i}(1+\tau_k F_k)) \\
&= dF_i d\log(\prod_{k=\alpha(t)+1}^{i}(1+\tau_k F_k)\prod_{k=0}^{\alpha(t)}(1+\tau_k L(T_{k-1},T_k))) \\
&= dF_i d\log(\prod_{k=\alpha(t)+1}^{i}(1+\tau_k F_k)) \\
&= dF_i \sum_{k=\alpha(t)+1}^{i}\frac{dF_k}{1+\tau_k F_k} \\
&= \sum_{k=\alpha(t)+1}^{i}\frac{\tau_k \rho_{k,i} C_i(t) C_k(t)}{1+\tau_k F_k(t)} dt
\end{aligned}$$

where we use the relation

$$\frac{P(t,T_i)}{P(t,T_j)} = \prod_{k=i+1}^{j}(1+\delta_k F_k),$$

and $L(T_{k-1},T_k)$ is known at time t.

□

3.7.1.3 *Pricing applications*

Lemma 3.7.4 (zero-coupon bond price). *[1, p. 125] Let $t < T_0 < T_1 < \cdots < T_n$ be a set of dates. Then*

$$\begin{aligned}
P(T_0,T_n) &= FP(T_0;T_0,T_1)FP(T_0;T_1,T_2)\cdots FP(T_0;T_{n-1},T_n) \\
&= \prod_{i=1}^{n}\frac{1}{1+(T_i-T_{i-1})F(T_0,T_{i-1},T_i)}
\end{aligned}$$

That is, at time t, the random variable $P(T_0,T_n)$ can be expressed as a product of functions involving random variables $F(T_0;T_{i-1},T_i)$.

Proof. See Lemma 3.1.2 □

Remark 3.7.3 (evaluation using Monte Carlo simulation). Suppose we want to generate a random sample of $P(T_0, T_n)$, we can obtain this sample by generating random samples of $F(T_0, T_{i-1}, T_i)$ by simulating forward rate dynamics.

Lemma 3.7.5 (Black's formula for caplet in the LIBOR market model). *Consider a caplet having payoff at T_i given by*

$$\tau_i(L(T_{i-1}, T_i) - K)^+ = \tau_i(F(T_{i-1}; T_{i-1}, T_i) - K)^+,$$

where $\tau_i = T_i - T_{i-1}$.

The price of this caplet at time $t < T_{i-1}$ is given by

$$Cpl_i(t) = \tau_i P(t, T_i)(F(t; T_{i-1}, T_i)N(d_+) - KN(d_-),$$

where

$$d_\pm = \frac{\ln F(t; T_{i-1}, T_i)/K \pm \frac{1}{2}\overline{\sigma}^2}{\overline{\sigma}}, \overline{\sigma}^2 = \int_t^{T_{i-1}} \sigma_i(s)^2 ds$$

Proof. Under the forward measure Q_{T_i},

$$\begin{aligned} Cpl_i(t) &= P(t, T_i)\tau_i E_{T_i}[F_i(T_{i-1}) - K)^+ | \mathcal{F}_t] \\ &= P(t, T_i)E_{T_i}[F_i(T_{i-1})\mathbf{1}_{F_i(T_{i-1})>K} | \mathcal{F}_t] - P(t, T_i)KE_{T_i}[\mathbf{1}_{F_i(T_{i-1})} | \mathcal{F}_t] \end{aligned}$$

In the LIBOR market model, under the forward measure Q_{T_i}, we have

$$dF_i(t) = \sigma_i(t)F_i dW_t,$$

therefore

$$F_i(T_{i-1}) = F_i(t)\exp(\int_t^{T_{i-1}} \sigma_i(s)dW_s - \frac{1}{2}\int_t^{T_{i-1}} \sigma_i^2(s)ds).$$

The rest is just routine. □

Remark 3.7.4 (caplet/cap pricing does not depend on the correlation of different forward rates). [3, p. 221] It can be seen that the prices of caplets/caps do not depend on the correlation matrix of the Brownian motions.Therefore, caplet/caps cannot be used to calibrate correlation.

Lemma 3.7.6 (pricing arbitrary interest rate European derivatives). *Consider a European derivative security with expiry date T_j and payoff of the form $h(F_1(T_j), F_2(T_j), ..., F_n(T_j))$, where*

- $F_i(T_j) = F(T_j; T_{i-1}, T_i), \forall i = j+1, ..., n.$

- $F_i(T_j) = F(T_{i-1}; T_{i-1}, T_i) = L(T_{i-1}, T_i), \forall i = 1, ..., j.$

  The price at time 0 of this option is given by

$$P(0, T_n) E_{P_{T_n}}[\frac{h(F_1(T_j), ..., F_n(T_j))}{P(T_j, T_n)}]$$

Proof. Use change of numeraire technique (Theorem 1.6.15), we have

$$\frac{dP_{T_j}}{dP_{T_n}} = \frac{P(T_j, T_j)}{P(0, T_j)} \frac{P(0, T_n)}{P(T_j, T_n)}.$$

Then,

$$\begin{aligned}
& P(0, T_i) E_{P_{T_i}}[h(F_1(T_j), ..., F_n(T_j))] \\
&= P(0, T_i) E_{P_{T_n}}[h(F_1(T_j), ..., F_n(T_j)) \frac{dP_{T_i}}{dP_{T_n}}] \\
&= P(t, T_i) E_{P_{T_n}}[h(F_1(T_j), ..., F_n(T_j)) \frac{P(T_j, T_j)}{P(0, T_j)} \frac{P(0, T_n)}{P(T_j, T_n)}] \\
&= P(0, T_n) E_{P_{T_n}}[\frac{h(F_1(T_j), ..., F_n(T_j))}{P(T_j, T_n)}]
\end{aligned}$$

□

3.7.2 LIBOR swap market model

3.7.2.1 *The model*

Lemma 3.7.7 (forward swap rate dynamics under forward swap measure). *Assume $P(t, T)$ obeys the SDE*

$$dP(t, T)/P(t, T) = (r(t) + \lambda\sigma(t, T))dt + P(t, T)\sigma(t, T)dW(t)$$

where $W(t)$ is the standard Brownian motion under the real world probability measure, and $r(t)$ is the instantaneous short rate.

*Given a set of dates $T_0 < T_1 < ... < T_n$ with accrual periods $\tau_i = T_i - T_{i-1}$. The **swap annuity** is defined as*

$$A_{0,n}(t) = \sum_{i=1}^{n} \tau_i P(t, T_i).$$

*The **forward swap rate**, denote by $S_{0,n}(t)$, is defined by,*

$$S_{0,n}(t) = \frac{P(t,T_0) - P(t,T_n)}{A_{0,n}(t)}.$$

Then,

- *The swap annuity $A_{0,n}(t)$ has SDE given by*

$$dA_{0,n}(t)/A_{0,n}(t) = (r + \lambda\sigma_A)dt + \sigma_A dW_t.$$

- *The quantity $P(t,T_0) - P(t,T_n)$ has SDE given by*

$$d(P(t,T_0) - P(t,T_n))/(P(t,T_0) - P(t,T_n)) = (r + \lambda\sigma_P)dt + \sigma_P dW_t.$$

- *Under forward measure Q_A with respect to $A_{0,n}(t)$(i.e. set $\lambda = \sigma_A$), the dynamics of $S_{0,n}(t)$ is given by*

$$dS_{0,n}(t) = S_{0,n}(t)(\sigma_P - \sigma_A)dW^S(t).$$

where W^S is a Brownian motion under the measure Q_A.

Proof. (1)(2)Using linearity in no-arbitrage condition(Lemma 1.5.3). (3) use Theorem 1.6.18. □

Lemma 3.7.8 (Black's formula for swaptions). *Assume the forward swap rate $S_{0,n}(t)$ under the forward swap measure(using swap annuity $A_{0,n}(t)$ as the numeraire) has dynamics*

$$dS_{0,n}(t) = \sigma_{0,n}(t)S_{0,n}(t)dW(t),$$

where $\sigma_{0,n}(t)$ is a deterministic time-dependent volatility and $W(t)$ is a Brownian motion under forward swap measure.

The payoff at maturity date T_0 is

$$A_{0,n}(T_0)(S_{0,n}(T_0) - K)^+,$$

where K is the strike.

The current price is given by

$$\begin{aligned} PSwpt_{0,n}(t) &= A_{0,n}(t)E_A[(S_{0,n} - K)^+|\mathcal{F}_t] \\ &= A_{0,n}(t)(S_{0,n}(t)N(d_+) - KN(d_-)) \end{aligned}$$

where

$$d_\pm = \frac{\ln(S_{0,n}(t)/K) \pm \frac{1}{2}\sigma^2}{\sigma}, \sigma^2 = \int_t^{T_0} \sigma_{0,n}(s)^2 ds$$

Proof. Using the change of numeraire method, under forward measure Q_A, we have(Lemma 3.3.3)

$$\frac{V(t)}{A_{0,n}(t)} = E_A[\frac{V(T_0)}{A_{0,n}(t)}|\mathcal{F}_t] = E_A[\frac{A_{0,n}(T_0)(S_{0,n}(T_0)-K)^+}{A_{0,n}(t)}|\mathcal{F}_t] = E_A[(S_{0,n}(T_0)-K)^+|\mathcal{F}_t]$$

Note that under forward measure Q_A, the distribution $S_{0,n}(t)$ is log-normal and the rest is straight-forward. □

Remark 3.7.5 (how to choose numeraire).

- If the final payoff is $(S_{0,t}(T_0)-K)^+$, and we can use forward measure Q_T associated with $P(t, T_0)$. Then
$$V(t) = P(t, T_0)E_T[(S_{0,t}(T_0)-K)^+|\mathcal{F}_t],$$
which can be easily evaluated if we know the distribution of $S_{0,t}(T_0)$ under measure Q_T.(see the example following Theorem 1.6.18).
- If the final payoff is
$$A_{0,n}(T_0)(S_{0,n}(T_0)-K)^+,$$
then it is wise to use forward swap measure.

3.7.2.2 *LMM approximation for swap rate volatility*

Lemma 3.7.9 (swap rate volatility approximation). *[1, p. 110]*

Under the forward measure Q_A with respect to the swap rate, we have

- $S_{0,n}(t)$ *is a martingale, and*
$$dS_{0,n}(t) = \sigma_{0,n}(t)S_{0,n}(t)dW^A(t),$$
- $$w_i(s) = \frac{\tau_i P(s, T_i)}{A_{0,n}(s)}$$
is a martingale.
- $$w_i(s)F_i(s) = \frac{P(s, T_{i-1}) - P(s, T_i)}{A_{0,n}(s)}.$$
is a martingale.

Moreover,

$$\begin{aligned}
\sigma_{0,n}(s)^2 &= \frac{dS_{0,n}(s)}{S_{0,n}(s)}\frac{dS_{0,n}(s)}{S_{0,n}(s)} \\
&\approx \frac{\sum_{i=1}^n \sum_{j=1}^n w_i(s)F_i(s)w_j(s)F_j(s)}{S_{0,n}(s)^2}\sigma_i(s)\sigma_j(s)\rho_{i,j}ds \\
&\approx \frac{\sum_{i=1}^n \sum_{j=1}^n w_i(t)F_i(t)w_j(t)F_j(t)}{S_{0,n}(t)^2}\sigma_i(s)\sigma_j(s)\rho_{i,j}ds, t < s
\end{aligned}$$

Proof. (1)(2)(3) directly from Lemma 3.7.7. (4) From Lemma 3.1.8, we have

$$\begin{aligned}
S_{0,n}(t) &= \sum_{i=1}^n w_i(t)F_i(t) \\
dS_{0,n}(s) &= \sum_{i=1}^n w_i(s)dF_i(s) + \sum_{i=1}^n F_i(s)dw_i(s) + drift \\
&= \sum_{i=1}^n w_i(s)F_i(s)\sigma_i(s)dZ_i^A(s) + \sum_{i=1}^n F_i(s)dw_i(s) \\
&\approx \sum_{i=1}^n w_i(s)F_i(s)\sigma_i(s)dZ_i^A(s)
\end{aligned}$$

where $drift$ represents terms of $O(ds)$ from $dF_i(s)dw_i(s)$, Z_i^A is a Brownian motion at measure Q_A. In the approximation, we neglects $\sum_{i=1}^n F_i(s)dw_i(s)$. $drift$ term disappear because $S_{0,n}(t)$ is a martingale under measure Q_A.

Note that in the last step, we approximate $w_i(s)F_i(s)$ and $S_{0,n}(s)$ by their previous value since they are martingales. □

Lemma 3.7.10 (Black's formula for swaptions with LMM approximate volatility). *Following Lemma 3.7.8, the price for a derivative with payoff*

$$A_{0,n}(T_0)(S_{0,n}(T_0) - K)^+$$

at T_0 is given by

$$PSwpt_{0,n}(t) = A_{0,n}(t)(S_{0,n}(t)N(d_+) - KN(d_-)).$$

where

$$
\begin{aligned}
d_{\pm} &= \frac{\ln(S_{0,n}(t)/K) \pm \frac{1}{2}\sigma^2}{\sigma},\\
\sigma^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{w_i(t)F_i(t)w_j(t)F_j(t)}{S_{0,n}(t)^2} \int_t^{T_0} \sigma_i(s)\sigma_j(s)\rho_{i,j} ds
\end{aligned}
$$

Proof. Use results from Lemma 3.7.9, we have

$$\sigma_{0,n}(s)^2 ds \approx \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i(t)F_i(t)w_j(t)F_j(t)}{S_{0,n}(t)^2}\sigma_i(s)\sigma_j(s)\rho_{i,j}ds, t < s.$$

□

3.7.3 Calibration and implementation

3.7.3.1 *Volatility calibration*

Example 3.7.1. [1, p. 121]

- Assume the instantaneous volatility takes the following time-homogeneous parametric form
$$\sigma_i(t;\alpha) = (a + b(T_{i-1} - t))e^{-c(T_{i-1}-t)} + d,$$
with $\alpha = \{a, b, c, d\}$. Denote
$$\nu_i(\alpha) = \int_t^{T_{i-1}} \sigma_i(s;\alpha) ds.$$
- Denote σ_i^*as the implied caplet volatility such that
$$Cpl_i^{mkt}(t) = Cpl_i(t;\sigma_i^*),$$
where
$$Cpl_i(t,\overline{\sigma}) = \tau_i P(t,T_i)(F(t;T_{i-1},T_i)N(d_+) - KN(d_-),$$
and
$$d_{\pm} = \frac{\ln F(t;T_{i-1},T_i)/K \pm \frac{1}{2}\overline{\sigma}^2}{\overline{\sigma}}.$$
- The parameter α for the volatility is solved from optimization problem
$$\min_{\alpha} \sum_{i=1}^{n} ((\sigma_i^*)^2 - \nu_i(\alpha)^2)^2.$$

- Suppose we already construct the forward curve(3.4.1), then

$$F(t, T_i, T_{i+1}) = \frac{1}{T_{i+1} - T_i}(\exp(\int_{T_i}^{T_{i+1} f(t,s)ds} - 1)).$$

3.7.3.2 *Correlation calibration*

Lemma 3.7.11 (Black's formula for swaptions with LMM approximate volatility). *Following Lemma 3.7.8, the price for a derivative with payoff*

$$A_{0,n}(T_0)(S_{0,n}(T_0) - K)^+$$

at T_0 is given by

$$PSwpt_{0,n}(t) = A_{0,n}(t)(S_{0,n}(t)N(d_+) - KN(d_-)).$$

where

$$d_\pm = \frac{\ln(S_{0,n}(t)/K) \pm \frac{1}{2}\sigma^2}{\sigma},$$
$$\sigma^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{w_i(t)F_i(t)w_j(t)F_j(t)}{S_{0,n}(t)^2} \int_t^{T_0} \sigma_i(s)\sigma_j(s)\rho_{i,j}ds$$

Proof. See Lemma 3.7.10. □

Note 3.7.2 (calibration the correlation structure).

- The correlation of historical forward curve is a good approximation to the correlation in the forward measure since correlation invariance property(**??**).
- The Black's formula for swaptions with LMM approximate volatility also enables the usage of swaption market prices to calibrate correlation.
- Good references on this topic are [12, p. 125][4, p. 212].

Remark 3.7.6 (intuitions behind calibrating correlation using swaption).

- Note that swap rate can be expressed as weighted sum of forward rate(Lemma 3.1.8)

$$S_{0,n}(t) = \sum_{i=1}^{n} w_i(t)F(t; T_{i-1}, T_i)$$

  where

$$w_i(t) = \frac{\tau_i P(t, T_i)}{\sum_{j=1}^{n} P(t, T_j)}.$$

- The valuation of the swaption requires the variance of random variable $\sum_{i=1}^{n} w_i F_{t_{ex};T_{i-1},T_i}$(If we assume each F is log-normal and we lognormal distribution to approximate $\sum_{i=1}^{n} w_i F_{t_{ex};T_{i-1},T_i}$). The calculation of the variance will involves terms in the covariance matrix(See also the moment matching approach(Lemma 2.3.6)).

3.8 Volatility modeling

3.8.1 Implied Black volatility

3.8.1.1 *Cap and floor volatility*

Definition 3.8.1 (implied Black cap and floor volatility).

- *Consider a unit notational amount and a set of dates $T_0 < T_1 < \cdots < T_n$ with $\tau_i = T_i - T_{i-1}, i = 1, 2, ..., n$. At time T_i, the holder of an interest rate cap receives $\tau_i \max(L(T_{i-1}, T_i) - K, 0)$, where K is the* ***cap rate****. Each of these n call options are known as* ***caplets****.*
  The ith caplet is a European call option with expiry T_i written on the spot LIBOR rate $L(T_{i-1}, T_i)$(a random quantity) with payoff
  $$(T_i - T_{i-1}) \max(L(T_{i-1}, T_i) - K, 0).$$
- *Consider the Black formula for caplets*
  $$Cpl_i^B(t;\sigma) = \tau_i P(t, T_i)(F(t; T_{i-1}, T_i)N(d_+) - KN(d_-))$$
  where
  $$d_+ = \frac{\ln \frac{F(t;T_{i-1},T_i)}{K} + \frac{1}{2}\sigma^2(T_{i-1} - t)}{\sigma\sqrt{T_{i-1} - t}}, d_- = d_+ - \sigma\sqrt{T_{i-1} - t}.$$
- *The implied Black spot volatility $\hat{\sigma}_i^{caplet}$ of the ith caplet is defined as the unique solution to the equation*
  $$Cpl_i^m kt(t) = Cpl_i^{Black}(t; \hat{\sigma}_i^{caplet}).$$
- *The implied Black flat volatility $\hat{\sigma}^{cap}$ of the ith caplet is defined as the unique solution to the equation*
  $$Cap^m kt(t) = \sum_{i=1}^{n} Cpl_i^{Black}(t; \hat{\sigma}_n^{cap}).$$

Remark 3.8.1 (the parameters in the implied cap/floor volatility). The implied cap/floor volatility usually can be characterized by three parameter swaption expiry T_s, strike K, and the interest rate accrual period length $\Delta = T_i - T_{i-1}$.

Note 3.8.1 (volatility market observations). [5, p. 681]

- (observation) Usually in the market, volatility term structure like Figure 3.8.1 is usually observed. Particularly, there is a 'hump' at about the 2 to 3year point.
- (explanation) One possible explanation is as follows. Rates at the short end of the zero curve are controlled by central banks. By contrast, 2- and 3-year interest rates are determined to a large extent by the activities of traders. These traders may be overreacting to the changes observed in the short rate and causing the volatility of these rates to be higher than the volatlity of short rates. For maturities beyond 2 to 3 years, the mean reversion of interest rates.

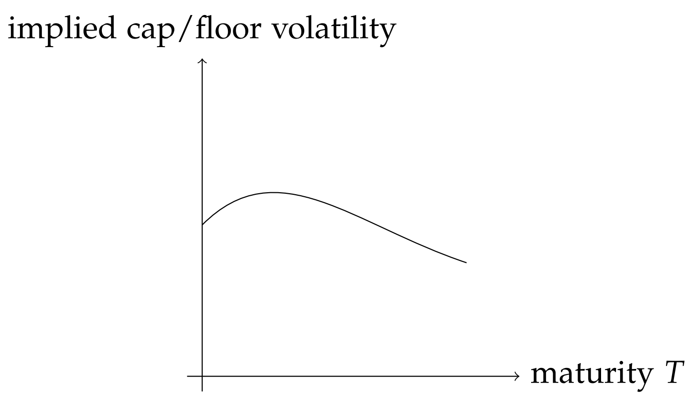


Figure 3.8.1: The cap/floor implied volatility hump

3.8.1.2 *Swaption volatility*

Definition 3.8.2. *Consider an interest rate swap with payment dates $T_1, T_2, ..., T_n$ and reset dates $T_0, T_1, ..., T_{n-1}$. Consider a **payer(receiver) swaption on spot-starting swaps** gives the holder the right to enter into an interest rate swap with maturity date $T_s = T_0$ such that the payoff at T_s is The payoff at maturity date T_0 is*

$$A(T_0, T_0, T_n)(S(T_0, T_0, T_n) - K)^+,$$

where K is the strike, $S_{0,n}(T_0)$ is the fair swap rate at T_0, and

$$A(t, T_0, T_n) = \sum_{i=1}^{n} (T_i - T_{i-1}) P(t, T_i).$$

The Black formula(Lemma 3.7.8) for the value of this swaption at current time $t < T_s$ is given by

$$\begin{aligned} PSwpt^B_{0,n}(t; \sigma) &= A(t, T_0, T_n) E_A[(S_{0,n} - K)^+ | \mathcal{F}_t] \\ &= A(t, T_0, T_n)(S(t, T_0, T_n) N(d_+) - K N(d_-)) \end{aligned}$$

where

$$d_{\pm} = \frac{\ln(S_{0,n}(t)/K) \pm \frac{1}{2}\sigma^2}{\sigma}.$$

*Denote the market price of the swaption is given by $PSwpt^M_{0,n}$. Then **the implied swaption volatility**, denoted by $\sigma^{imp}_{SWap}(K, T_s, \Delta)$, is the volatility that make the following equation hold:*

$$PSwpt^B_{0,n}(t; \sigma^{imp}_{SW}(K, T_s, \Delta)) = PSwpt^M_{0,n},$$

where we use $\Delta = T_n - T_0$ to denote the underlying swap total length.

Note 3.8.2 (the parameters in the implied swaption volatility).

- Usually, the underlying swaps have standard swap dates $T_0, T_1, ..., T_n$; therefore, the swap tensor can be simply represented by its total length $\Delta = T_n - T_0$, which usually takes 1Y, 2Y, 3Y, 5Y, 10Y, and 30Y.
- The option expiry ranges from 0D, 7D, 14D, 1M, 2M, 3M, 6M, 12M, ..., 1Y, 2Y, ..., 20Y, 40Y.
- With the underlying standard swap, the implied swaption volatility usually can be characterized by three parameter swaption expiry T_s, strike K, and the underlying swap length Δ.

3.8.1.3 *Comparisons between swaption and cap/floor implied volatility*

Note 3.8.3 (comparisons between swaption and cap/floor implied volatility). Denote $\sigma^{imp}_{CF}(T_C, K, \Delta_{CF})$ and $\sigma^{imp}_{SW}(T_S, K, \Delta_{SW})$ as the cap/floor implied volatility and the swaption implied volatility respectively.

- Δ_{CF} usually takes values of 1M, 3M, 6M, 1Y; Δ_{SW} usually takes values of 3M, 6M, 1Y, 2Y, 3Y, 5Y, 7Y, 10Y, 15Y, 20Y, 30Y.
- Δ_{CF} represents the volatility of short-tensor interest rate(1M to 1Y) at the expiry T_{CF}; Δ_{CF} represents the volatility of short-tensor interest rate(1M to 1Y) averaged from T_0 to T_n at the expiry T_{SW} since swap rate is 'kind of' average interest rate(Lemma 3.1.8) over a future period.

Expiry \Tenor	0.25Y	1Y	2Y	3Y	4Y	5Y	7Y	10Y	15Y	20Y	25Y	30Y	40Y
0D													
7D													
14D													
1M													
2M													
3M													
6M													
9M													
1Y													
1.5Y													
2Y													
3Y													
...													
10Y													
15Y													
20Y													
25Y													
30Y													
40Y													

Table 3.8.1: An ATM volatility table from market traded swaptions.

Expiry \Tenor	1M	3M	6M	12M
0D				
7D				
14D				
1M				
2M				
3M				
6M				
9M				
1Y				
1.5Y				
2Y				
3Y				
...				
10Y				
15Y				
20Y				
25Y				
30Y				
40Y				

Table 3.8.2: An ATM volatility table from market traded caplets and floorlets.

3.8.2 Constant Elasticity of Variance (CEV) model

Definition 3.8.3 (Constant Elasticity of Variance(CEV) model). *[3, p. 456] Denote a forward LIBOR rate by $F_i(t)$. In the Shifted-Lognormal model model, $F_i(t)$, under forward measure of T_i-maturity zero-coupon bond, is assumed to satisfy*

$$dF_i(t) = F_i(t)^\beta \sigma_i(t) dW_t$$

where $\beta \in [0,1]$ is a real constant, $\sigma(t)$ is a deterministic function of time, and $W(t)$ is a standard Brownian motion.

3.8.3 Shifted-Lognormal model

Definition 3.8.4 (Shifted-Lognormal model). *[3, p. 454] Denote a forward LIBOR rate by $F_i(t)$. In the Shifted-Lognormal model model, $F_i(t)$, under forward measure of T_i-maturity zero-coupon bond, is assumed to satisfy*

$$\begin{aligned} F_i(t) &= X_i(t) - \alpha \\ dX_i(t) &= \beta(t) X_j(t) dW_t \end{aligned}$$

where α is a real constant, $\beta(t)$ is a deterministic function of time, and $W(t)$ is a standard Brownian motion.

Note 3.8.4 (relation between shift-lognormal model and CEV model)**.** Note that the Taylor expansion of $F_i(t)^\beta$ gives

$$\begin{aligned} F_i(t)^\beta &= F_i(0)^\beta + \beta F_i(0)^{\beta-1}(F_i(t) - F_i(0)) + O((F_i(t) - F_i(0)^2) \\ &= F_i(0)^{\beta-1}(F_i(t) + (1-\beta)F_i(0)) + O((F_i(t) - F_i(0)^2) \\ &\approx F_i(0)^{\beta-1}(F_i(t) + (1-\beta)F_i(0)) \end{aligned}$$

Therefore,$dF_i(t) = F_i(t)^\beta \sigma_i(t) dW_t$ can be approximately written by

$$dF_i(t) = \sigma_i(t) F_i(0)^{\beta-1}(F_i(t) + (1-\beta)F_i(0)) dW_t,$$

or equivalently,

$$d(F_i(t) + \alpha) = \sigma_i^D(t)(F_i(t) + \alpha) dW_t,$$

where

$$\alpha = \frac{(1-\beta)F_i(0)}{\beta}, \sigma_i^D(t) = \beta F_i(0)^{\beta-1}\sigma_i(t).$$

Lemma 3.8.1 (solution to Shifted-Lognormal model). *[3, p. 454] In the Shifted-Lognormal model, it follows that*

- *The solution of X_i is given by*

$$X_i(T) = X_i(t)\exp(-\frac{1}{2}\int_t^T \beta^2(u)du + \int_t^T \beta(u)dW(u))$$

- *The solution of F_i is given by*

$$F_i(T) = -\alpha + (F_i(t)+\alpha)\exp(-\frac{1}{2}\int_t^T \beta^2(u)du + \int_t^T \beta(u)dW(u))$$

Proof. Straight forward. □

Lemma 3.8.2 (pricing caplet). *[1, p. 151]*

The current value of a caplet with strike K and maturity T_{i-1} is given by

$$V(t) = P(t,T_i)\tau_i E_{Q_{T_i}}[(F_i(T_{i-1})-K)^+|\mathcal{F}_t] = P(t,T_i)E_{Q_{T_i}}[(X_i(T_{i-1})+\alpha-K)^+|\mathcal{F}_t]$$

Specifically,

$$\begin{aligned} V(t) &= Cpl(K+\alpha, F_i(t)+\alpha, \Sigma(t,T)) \\ &= \tau_i P(t,T_i)((F_i(t)+\alpha)N(d_+) - (K+\alpha)N(d_-)) \end{aligned}$$

where

$$d_\pm = \frac{\ln\frac{F_i(t)+\alpha}{K+\alpha} \pm \frac{1}{2}\Sigma(t,T)}{\sqrt{\Sigma(t,T)}}$$

and

$$\Sigma(t,T)^2 = \int_t^T \beta^2(u)du.$$

Proof. Straight forward. □

Lemma 3.8.3. *The at-the-money(ATM) implied volatility for a caplet with strike K and maturity T_{i-1} is given by*

$$\sigma_{im}(K) = \frac{\sigma F_i(t)}{F_i(t) + \alpha},$$

Remark 3.8.2 (parameterization of the volatility surface). Consider a single maturity date T_{i-1} and a single tenor τ_i, the volatility smile(implied Black volatility as a function of K) is given by

3.8.4 SABR model

Definition 3.8.5 (SABR model). *[1][13] Denote a forward LIBOR rate by $F_i(t)$. In the SABR model, $F_i(t)$, under forward measure of T_i-maturity zero-coupon bond, is assumed to satisfy*

$$\begin{aligned} dF_i(t) &= F_i(t)^\beta \sigma_i(t) dW(t) \\ d\sigma_i(t) &= \nu \sigma_i(t) dZ(t) \\ \sigma_i(0) &= \sigma_i^0 \end{aligned}$$

where $W(t)$ and $Z(t)$ are Brownian motions under the forward measure Q_{T_i} such that $dW(t)dZ(t) = \rho dt$, ν is the constant volatility of $\sigma_i(t)$. The parameter ranges are $\nu \geq 0, 0 \leq \beta < 1, -1 < \rho < 1$.

Lemma 3.8.4 (European call/option pricing).

- *The European call price with strike K and maturity T is given by*

  $$V_c(t) = P(t,T)(F(t,T)N(d_1) - KN(d_2))$$

  where $F(t,T)$ is the forward price the underlying, and

  $$d_{1,2} = \frac{\log(F(t,T)/K) \pm \frac{1}{2}\sigma_B^2(T-t)}{\sigma_B^2\sqrt{T-t}},$$

and the implied volatility $\sigma(F,K)$ is given by

$$\sigma_B(F,K) = \frac{\alpha}{(FK)^{(1-\beta)/2}[1+\frac{(1-\beta)^2}{24}\log^2(F/K)+\frac{(1-\beta)^4}{1920}\log^4(F/K)+\cdots]}$$
$$\cdot \frac{z}{\chi(z)} \cdot \{1+[\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(FK)^{1-\beta}}+\frac{1}{4}\frac{\rho\beta\nu\alpha}{(FK)^{(1-\beta)/2}}+\frac{2-3\rho^2}{24}\nu^2](T-t)+\cdots\}$$
$$z = \frac{\nu}{\alpha}(FK)^{(1-\beta)/2}\log(F/K)$$
$$\chi(z) = \log\frac{\sqrt{1-2\rho z+z^2}+z-\rho}{1-\rho}$$

- *At-the-money volatility*

$$\sigma_B(F,F) = \frac{\alpha}{(FK)^{(1-\beta)/2}}\{1+[\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(FK)^{1-\beta}}+\frac{1}{4}\frac{\rho\beta\nu\alpha}{(FK)^{(1-\beta)/2}}+\frac{2-3\rho^2}{24}\nu^2](T-t)+\cdots\}$$

- *The put price given by*

$$V_p(t) = V_c(t) + P(t,T)(K-F(t,T)).$$

Remark 3.8.3 (choices of parameter for the volatility cube)**.**

- At every maturity T, we use four parameter α_0, β, ν and ρ to characterize the volatility smile.
- For a set of maturity dates $T_1, T_2, ..., T_n$ and different tenors $\tau_1, \tau_2, ..., \tau_m$, we use a set of $4mn$ parameters to characterize the volatility surface. Usually, β is chosen to be constant independent of maturity date and tenors.

Note 3.8.5 (how parameters affects the shape of parameterized volatility surface)**.** [1]

- An increase in the initial volatility σ_i^0 is an approximate upwards parallel shift in the Black implied volatility smile.
- A change in β from 1 to 0 will the cause the slope of smile to steepen.
- The slope of the smile can also be affected by ρ; therefore, we usually fix β and vary ρ.
- The ν parameter introduces curvature of the smile. The larger ν gives larger curvature.

3.9 Exotic interest rate derivatives

3.9.1 LIBOR-in-arrears and in-arrears swap

3.9.1.1 *LIBOR-in-arrears*

Definition 3.9.1 (LIBOR-in-arrears). *A LIBOR-in-arrears contract is a contract that pays amount $\tau L(S,T), S < T$ at time S, where $\tau = T - S$.*

Lemma 3.9.1 (pricing LIBOR-in-arrears). *[1, p. 130] Consider a LIBOR-in-arrears contracts pays $\tau L(S,T)$ at time S.*

Then it time-t value is given by

$$Lia(t) = \tau P(t,S) E_S[L(S,T)|\mathcal{F}_t].$$

where E_S is the expectation taken with respect to forward measure of $P(t,S)$. or equivalently,

$$\begin{aligned} Lia(t) &= \tau P(t,T) E_T[L(S,T)|\mathcal{F}_t] + \tau^2 P(0,S) E_T[L(S,T)^2|\mathcal{F}_t] \\ &= \tau P(t,T) F(t;S,T) + \tau^2 P(0,S) E_T[L(S,T)^2|\mathcal{F}_t] \end{aligned}$$

where E_T is the expectation taken with respect to forward measure of $P(t,T)$.

Proof. Define

$$\frac{dQ_S}{dQ_T} = \frac{P(t,T)P(S,S)}{P(t,S)P(S,T)} = \frac{P(t,T)}{P(t,S)}(1 + \tau L(S,T)).$$

Use change of numeraire technique(Theorem 1.6.15), we have

$$\begin{aligned} P(t,S) E_S[L(S,T)|\mathcal{F}_t] &= P(t,S) E_T[L(S,T) \frac{dQ_S}{dQ_T} |\mathcal{F}_t] \\ &= P(t,S) E_T[L(S,T) \frac{P(t,T)}{P(t,S)} (1 + \tau L(S,T)) |\mathcal{F}_t] \\ &= \tau P(t,T) E_T[L(S,T)|\mathcal{F}_t] + \tau^2 P(0,S) E_T[L(S,T)^2|\mathcal{F}_t] \end{aligned}$$

Then we use the fact that $F(t;S,T) = E_T[L(S,T)|\mathcal{F}_t]$ in Lemma 3.3.4. □

Lemma 3.9.2 (evaluating LIBOR-in-arrears using caplets). *[1, p. 132] It follows that*

- $$L(S,T)^2 = 2\int_0^{L(S,T)} (L(S,T)-K)dK = 2\int_0^{\infty} (L(S,T)-K)^+ dK.$$
- $$\begin{aligned}&\tau^2 P(0,S)E_T[L(S,T)^2|\mathcal{F}_t]\\ =&2\tau\int_0^\infty \tau P(0,T)E_T[(L(S,T)-K)^+|\mathcal{F}_t]dK = 2\tau\int_0^\infty Cpl(t,K,S,T)dK\end{aligned}$$

that is, we can use the caplet prices to price $L(S,T)^2$.

Lemma 3.9.3 (evaluating LIBOR-in-arrears using log-normal model). *[1, p. 132] Assume the forward rate $F(t) \triangleq F(t;S,T)$ has log-normal dynamics under forward measure Q_T, given by*

$$dF(t) = \tilde{\sigma}(t)F(t)dW_t.$$

Then

$$\begin{aligned}E_T[L(S,T)^2|\mathcal{F}_t] &= F(t)^2\exp(-\sigma^2)E_T[\exp(2\sigma Z)|\mathcal{F}_t], Z\sim N(0,1)\\ &= F(t)^2\exp(\sigma^2)\end{aligned}$$

where

$$\sigma^2 = \int_t^S \tilde{\sigma}^2(s)ds.$$

Proof.

$$\begin{aligned} dF(t) &= \tilde{\sigma}(t)F(t)dW_t\\ \implies L(S,T)^2 &= F(S)^2 = [F(t)\exp(-\frac{1}{2}\sigma^2+\sigma Z)]^2, Z\sim N(0,1)\\ &= F(t)^2\exp(-\sigma^2+2\sigma Z)\\ \implies E_T[L(S,T)^2] &= F(t)^2\exp(-\sigma^2)E_T[\exp(2\sigma Z)]\\ &= F(t)^2\exp(-\sigma^2)\exp(2\sigma^2)\\ &= F(t)^2\exp(\sigma^2)\end{aligned}$$

where

$$\sigma^2 = \int_t^S \tilde{\sigma}^2(s)ds.$$

and we use properties of lognormal random variable(**??**). □

Lemma 3.9.4 (convexity adjustment for LIBOR-in-arrears). *Assume the forward rate $F(t) \triangleq F(t; S, T)$ has log-normal dynamics under forward measure Q_T, given by*

$$dF(t) = \tilde{\sigma}(t)F(t)dW_t.$$

Let $t < S < T, \tau = T - S$. Then,

$$\begin{aligned} E_S[L(S,T)|\mathcal{F}_t] &= F(t;S,T) + \frac{\tau F^2}{1+\tau F}(\exp(\sigma^2) - 1) \\ &\approx F(t;S,T) + \frac{\tau F^2}{1+\tau F}\sigma^2 \end{aligned}$$

where

$$\sigma^2 = \int_t^S \tilde{\sigma}^2(s)ds.$$

Proof. Define

$$\frac{dQ_S}{dQ_T} = \frac{P(t,T)P(S,S)}{P(t,S)P(S,T)} = \frac{P(t,T)}{P(t,S)}(1 + \tau L(S,T)).$$

Use change of numeraire technique(Theorem 1.6.15), we have

$$\begin{aligned} E_S[L(S,T)|\mathcal{F}_t] &= E_T[L(S,T)\frac{dQ_S}{dQ_T}|\mathcal{F}_t] \\ &= E_T[L(S,T)\frac{P(t,T)}{P(t,S)}(1+\tau L(S,T))|\mathcal{F}_t] \\ &= \frac{P(t,T)}{P(t,S)}(E_T[L(S,T)|\mathcal{F}_t] + \tau E_T[L(S,T)^2|\mathcal{F}_t]) \\ &= \frac{1}{1+\tau F(t;S,T)}(F(t;S,T) + \tau F(t;S,T)^2 \exp(\sigma^2)) \end{aligned}$$

Then

$$\Delta \triangleq E_S[L(S,T)|\mathcal{F}_t] - F(t;S,T) =$$

Then we use the fact that $F(t;S,T) = E_T[L(S,T)|\mathcal{F}_t]$ in Lemma 3.3.4. □

3.9.1.2 *In-arrears swap*

Definition 3.9.2 (In-arrears swap). *[1, p. 132] The holder of a payer in-arrears swap pays a fixed amount $\tau_i K$ in exchange for a floating payment $\tau_i L(T_{i-1}, T_i)$ at time $T_{i-1}, i = 1, ..., n$ rather than at time T_i.*

Lemma 3.9.5 (pricing In-arrears swap). *[1, p. 132] The current value at time t of an in-arrears swap is*

$$PSia(t) = \sum_{i=1}^{n} \tau_i P(t, T_{i-1}) E_{T_{i-1}}(L(T_{i-1}, T_i) - K) = \sum_{i=1}^{n} Lia_i(t) - K \sum_{i=1}^{n} \tau_i P(0, T_{i-1}).$$

where $Lia_i(t)$ is the current value at time t for a LIBOR-in-arrears payment of $\tau L(T_{i-1}, T_i)$.

Proof. Straight forward. □

3.9.2 Constant-maturity swaps and related products

3.9.2.1 *Constant-maturity swaps*

Lemma 3.9.6 (pricing constant-maturity swap). *The current time t value of a CMS contract is*

$$CMS(t) = \sum_{i=0}^{n-1} \tau_{i+1} P(0, T_{i+1}) E_{P_{T_{i+1}}}[(S_{i,i+m}(T_i) - K)|\mathcal{F}_t].$$

For each term we have,

$$P(0, T_{i+1}) E_{P_{T_{i+1}}}[S_{i,i+m}(T_i)|\mathcal{F}_t] = A_{i,i+m}(T_{i+1}) E_{A_{i,i+m}}[S_{i,i+m}(T_i) \frac{P(T_i, T_{i+m})}{A_{i,i+m}(T_i)}|\mathcal{F}_t].$$

If we assume $P(T_i, T_{i+m}) = (\alpha + \beta S_{i,i+m}(T_i)) A_{i,i+m}(T_i)$, then

$$E_{A_{i,i+m}}[S_{i,i+m}(T_i) \frac{P(T_i, T_{i+m})}{A_{i,i+m}(T_i)}|\mathcal{F}_t] = \alpha S_{i,i+m}(t) + \beta E_{A_{i,i+m}}[S_{i,i+m}(T_i)^2|\mathcal{F}_t].$$

Note that $S_{i,i+m}(t)$ has log-normal dynamics under forward swap measure $Q_{A_{i,i+m}}$(Lemma 3.7.7).

Proof. Note that $S_{i,i+m}(t)$ has nice property under measure $Q_{A_{i,i+m}}$ instead of $Q_{P_{T_{i+1}}}$.

Use change of numeraire technique (Theorem 1.6.15), we have

$$\frac{dP_{T_{i+1}}}{dP_{A_{i,i+m}}} = \frac{P(T_{i+1}, T_{i+1})}{P(t, T_{i+1})} \frac{A_{i,i+m}(t)}{A_{i,i+m}(T_i)}.$$

and

$$E_{P_{T_{i+1}}}[\cdots|\mathcal{F}_t] = E_{A_{i,i+m}}[\cdots \frac{dP_{T_{i+1}}}{dP_{A_{i,i+m}}}|\mathcal{F}_t]$$

□

3.9.2.2 *CMS caps, floors and spread options*

Definition 3.9.3. *[1, p. 136] A CMS cap(or floor) is an agreement where the holder receives cash payment on a set of predefined dates depending on the spot swap rate. A CMS cap consists of a series CMS caplets.*

A CMS caplet with unit notional paying at time T_{i+1} on an m-period swap rate setting at time T_i. Note that the swap rate is associated with an swap usually has final payment much grater than T_{i+1}. The payoff at time T_{i+1} is

$$\tau_{i+1}(S_{i,i+m}(T_i) - K)^+,$$

where $\tau_{i+1} = T_{i+1} - T_i$, K is the strike rate.

Remark 3.9.1 (calibration of swap rate model). Note that CMS caps/floors can be used to calibrate the volatility in the log-normal model of swap rates.

3.9.3 Ratchet floater

Definition 3.9.4 (ratchet floater). *[1, p. 138] Take a unit notional amount and a set of dates, $0 \leq T_0 \leq T_1 \cdots T_n$ with acrual periods $\tau_i = T_i - T_{i-1}$.*

- *At time T_i, the holder receives a payment $\tau_i(L(T_{i-1}, T_i) + X)$, where X is a preassigned constant spread, and pays coupon c_i.*
- *The first coupon $c_1 = \tau_1(L(T_0, T_1) + Y)$; the coupon $c_i, i = 2, ..., n$ satisfies*

$$c_i = c_{i-1} + \min((\tau_i(L(T_{i-1}, T_i) + Y) - c_{i-1})^+, \alpha),$$

where $\alpha > 0$ is preassigned constant.

Lemma 3.9.7. *[1, p. 137] The current time t value of a ratchet floater is given by*

$$V(t) = \sum_{i=1}^{n} P(t, T_i) E_{P_{T_i}}[\tau_i(F(T_{i-1}; T_{i-1}, T_i) + X) - c_i | \mathcal{F}_t],$$

where $E_{P_{T_i}}$ is taking expectation with respect to the forward measure of $P(t, T_i)$.

Or equivalently,

$$V(t) = P(t, T_n) \sum_{i=1}^{n} E_{P_{T_n}}[\frac{\tau_i(F(T_{i-1}; T_{i-1}, T_i) + X) - c_i}{P(T_i, T_n)} | \mathcal{F}_t].$$

Proof. Use change of numeraire technique (Theorem 1.6.15), we have

$$\frac{dP_{T_i}}{dP_{T_n}} = \frac{P(T_i, T_i)}{P(t, T_i)} \frac{P(t, T_n)}{P(T_i, T_n)}.$$

Then,

$$\begin{aligned}
& P(t, T_i) E_{P_{T_i}}[\tau_i(F(T_{i-1}; T_{i-1}, T_i) + X) - c_i | \mathcal{F}_t] \\
&= P(t, T_i) E_{P_{T_n}}[(\tau_i(F(T_{i-1}; T_{i-1}, T_i) + X) - c_i) \frac{dP_{T_i}}{dP_{T_n}} | \mathcal{F}_t] \\
&= P(t, T_i) E_{P_{T_n}}[(\tau_i(F(T_{i-1}; T_{i-1}, T_i) + X) - c_i) \frac{P(T_i, T_i)}{P(t, T_i)} \frac{P(t, T_n)}{P(T_i, T_n)} | \mathcal{F}_t] \\
&= P(t, T_n) \sum_{i=1}^{n} E_{P_{T_n}}[\frac{\tau_i(F(T_{i-1}; T_{i-1}, T_i) + X) - c_i}{P(T_i, T_n)} | \mathcal{F}_t]
\end{aligned}$$

□

Remark 3.9.2 (Monte Carlo evaluation)**.**

- The sample point of $P(T_i, T_n)$ under Q_{T_n} is discussed at Lemma 3.7.4.

3.9.4 Interest rate futures

3.9.4.1 *Fundamentals*

Lemma 3.9.8 (dynamics of futures under risk neutral measure). *[5, p. 392] Under risk-neutral measure Q,*

- $Fur(t; T_0, T_1)$ *is a martingale such that*

$$Fur(t; T_0, T_1) = E[Fur(s; T_0, T_1)|\mathcal{F}_t], s > t$$

- $$Fur(t; T_0, T_1) = E_Q[L(T_0, T_1)|\mathcal{F}_t]$$

- *In the Brownian motion representation, the futures rate has risk-neutral dynamics given by*

$$dFur(t; S, T) = \sigma(t; S, T)Fur(t; S, T)$$

3.9.4.2 *Convexity adjustment*

Lemma 3.9.9 (convexity adjustment for futures rate). *Let* $Fur(t; S, T) = E_Q[L(S, T)|\mathcal{F}_t]$ *be the futures rate. Let* $F(t; S, T) = E_T[L(S, T)|\mathcal{F}_t]$ *be the forward rate. Denote the convexity adjustment by* $\Delta = Fur(t; S, T) - F(t; S, T)$, *then*

$$\Delta = \frac{1}{\tau}(E_Q[\frac{P(S,S)}{P(S,T)}|\mathcal{F}_t] - \frac{P(t,S)}{P(t,T)}]),$$

or equivalently,

$$\Delta = E_Q[L(S,T)(1 - \frac{B(t)}{B(T)}\frac{1}{P(t,T)})|\mathcal{F}_t].$$

Or equivalently,

$$\Delta = -\frac{1}{P(t,T)}cov(L(S,T)\frac{B(t)}{B(T)}).$$

Proof. (1)

$$\begin{aligned}
\Delta &= E_Q[L(S,T)|\mathcal{F}_t] - E_T[L(S,T)|\mathcal{F}_t] \\
&= E_Q[\frac{1}{\tau}(\frac{P(S,S)}{P(S,T)} - 1)|\mathcal{F}_t] - E_T[\frac{1}{\tau}(\frac{P(S,S)}{P(S,T)} - 1)|\mathcal{F}_t] \\
&= E_Q[\frac{1}{\tau}(\frac{P(S,S)}{P(S,T)} - 1)|\mathcal{F}_t] - \frac{1}{\tau}(\frac{P(t,S)}{P(t,T)} - 1) \\
&= \frac{1}{\tau}(E_Q[\frac{P(S,S)}{P(S,T)|\mathcal{F}_t} - \frac{P(t,S)}{P(t,T)}])
\end{aligned}$$

(2)

$$\begin{aligned}
\Delta &= E_Q[L(S,T)|\mathcal{F}_t] - E_T[L(S,T)|\mathcal{F}_t] \\
&= E_Q[L(S,T)|\mathcal{F}_t] - E_Q[L(S,T)\frac{dQ_T}{dQ}|\mathcal{F}_t] \\
&= E_Q[L(S,T)|\mathcal{F}_t] - E_Q[L(S,T)\frac{P(T,T)B(t)}{P(t,T)B(T)}|\mathcal{F}_t] \\
&= E_Q[L(S,T)(1 - \frac{B(t)}{B(T)}\frac{1}{P(t,T)})|\mathcal{F}_t]
\end{aligned}$$

(3) Note that

$$\begin{aligned}
\Delta &= E_Q[L(S,T)|\mathcal{F}_t] - E_T[L(S,T)|\mathcal{F}_t] \\
&= E_Q[L(S,T)|\mathcal{F}_t] - E_Q[L(S,T)\frac{dQ_T}{dQ}|\mathcal{F}_t] \\
&= E_Q[L(S,T)|\mathcal{F}_t] - E_Q[L(S,T)\frac{P(T,T)B(t)}{P(t,T)B(T)}|\mathcal{F}_t] \\
&= E_Q[L(S,T)|\mathcal{F}_t] - Cov(L(S,T)\frac{P(T,T)B(t)}{P(t,T)B(T)}|\mathcal{F}_t) - E_Q[L(S,T)|\mathcal{F}_t]E_Q[\frac{P(T,T)B(t)}{P(t,T)B(T)}|\mathcal{F}_t] \\
&= E_Q[L(S,T)|\mathcal{F}_t] - Cov(L(S,T)\frac{P(T,T)B(t)}{P(t,T)B(T)}|\mathcal{F}_t) - E_Q[L(S,T)|\mathcal{F}_t] \\
&= -Cov(L(S,T)\frac{B(t)}{P(t,T)B(T)}|\mathcal{F}_t)
\end{aligned}$$

□

Remark 3.9.3 (futures rate usually greater than forward rate). Note that because $L(S,T)$ and $1/B(T)$ are negatively correlated, therefore $\Delta > 0$; that is, futures rate is greater than forward rate.

Lemma 3.9.10 (interest rate futures convexity adjustment in the Gaussian HJM framework). *[4, p. 262] In the HJM framework(Theorem 3.6.1), the forward rate $f(t,T)$,* ***under risk-neutral measure*** *Q,is given*

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t),$$

where $W(t)$ is a Brownian motion,

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,u)du.$$

Furthermore, the zero-coupon bond price $P(t,T)$ under Q satisfies SDE

$$dP(t,T) = r(t)P(t,T)dt + \Sigma(t,T)P(t,T)dW(t)$$

with log-volatility

$$\Sigma(t,T) = -\int_t^T \sigma(t,u)du.$$

It follows that

- $$\frac{P(S,S)}{P(S,T)} = \frac{P(t,S)}{P(t,T)}\exp(\int_t^S (\Sigma(u,S)-\Sigma(u,T)dW_u - \int_t^S \frac{1}{2}(\Sigma(u,S)-\Sigma(u,T))^2 du)$$
$$\exp(\int_t^S \Sigma(u,S)(\Sigma(u,S)-\Sigma(u,T))du)$$

- $$E_Q[\frac{P(S,S)}{P(S,T)}\mathcal{F}_t] = \frac{P(t,S)}{P(t,T)}\exp(\int_t^S \Sigma(u,S)(\Sigma(u,S)-\Sigma(u,T))du)$$

- *The convexity adjustment $\Delta = Fur(t;S,T) - F(t;S,T)$ is given by*

$$\begin{aligned}\Delta &= \frac{1}{\tau}\frac{P(t,S)}{P(t,T)}(\exp(\int_t^S \Sigma(u,S)(\Sigma(u,S)-\Sigma(u,T))du)-1) > 0 \\ &= \frac{\tau F(t;S,T)+1}{\tau}(\exp(\int_t^S \Sigma(u,S)(\Sigma(u,S)-\Sigma(u,T))du)-1)\end{aligned}$$

Proof. (1)

From Theorem 3.3.5, we have The ratio $\triangleq \frac{P(t,S)}{P(t,T)}, S \leq T$ has SDE, under measure Q, given by:

$$d\tilde{FP}/\tilde{FP} = (\Sigma(t,S)-\Sigma(t,T))dW(t) + \Sigma(t,T)(\Sigma(t,T)-\Sigma(t,S))dt.$$

(2) note that

$$\exp(\int_t^S (\Sigma(u,S)-\Sigma(u,T)dW_u - \int_t^S \frac{1}{2}(\Sigma(u,S)-\Sigma(u,T))^2 du)$$

is the martingale with mean value 1 under measure Q.

(3) Use Lemma 3.9.9. □

Remark 3.9.4. For more discussion on convexity adjustment, see [4][3, p. 559].

Lemma 3.9.11 (interest rate futures convexity adjustment in the Ho-Lee model). *[5, p. 144][4, p. 264]*

The Ho-Lee short rate model(Lemma 3.5.23) under the risk-neutral measure Q gives zero-coupon bond price as

$$dP(t,T)/P(t,T) = r(t)dt - \sigma(T-t)dW_t.$$

The convexity adjustment $\Delta = Fur(t;S,T) - F(t;S,T)$ *is given by*

$$\Delta = \frac{1}{\tau}\frac{P(t,S)}{P(t,T)}(\exp(\int_t^S \Sigma(u,S)(\Sigma(u,S)-\Sigma(u,T))du) - 1).$$

Proof. (1)

From Theorem 3.3.5, we have The ratio $\triangleq \frac{P(t,S)}{P(t,T)}, S \leq T$ has SDE, under measure Q, given by:

$$d\tilde{FP}/\tilde{FP} = (\Sigma(t,S)-\Sigma(t,T))dW(t) + \Sigma(t,T)(\Sigma(t,T)-\Sigma(t,S))dt.$$

(2) note that

$$\exp(\int_t^S (\Sigma(u,S)-\Sigma(u,T)dW_u - \int_t^S \frac{1}{2}(\Sigma(u,S)-\Sigma(u,T))^2 du)$$

is the martingale with mean value 1 under measure Q.

(3) Use Lemma 3.9.9. □

3.9.4.3 *Interest rate futures options*

Definition 3.9.5 (European call option on futures). *An European call option on Eurodollar futures obtains the right to enter into a Eurodollar futures at the specified strike price K.*

A call option has maturity T earlier than the delivery date T_1 *of the futures. The total payoff of a call option is* $\max(Fur(T;T_1,T_2)-K,0)$

Example 3.9.1. [5, p. 385] It is February and the futures prices for the June Eurodollar contract is 93.82(corresponding to a 3-month Eurodollar interest rate of 6.18% per annum). The price of a call option on the contract with a strike price of 94.00 is quoted as 10 basis point.

Suppose in June, the Eurodollar futures price increases to 94.78, then the long position of the call will enable the investor to get a futures price of 94, therefore gaining profit.

Lemma 3.9.12 (pricing call option on Eurodollar futures). *The price of call option maturing at T with payoff $(Fur(T)-K)^+$ is given by*

$$\begin{aligned} V(t) &= E_Q[\exp(-\int_t^T r(s)ds)(Fur(T)-K)^+|\mathcal{F}_t] \\ &= P(t,T)E_T[(Fur(T)-K)^+|\mathcal{F}_t] \\ &= P(t,T)E_T[(F(T)+\Delta(T)-K)^+|\mathcal{F}_t] \end{aligned}$$

where $F(t)$ the forward price and $\Delta(T) = \triangleq Fur(T)-F(t)$ is the convexity adjustment.

3.9.4.4 *Interest rate futures as hedging instrument*

3.9.5 Callable bonds

Definition 3.9.6 (callable zero-coupon bond). *A* ***callable zero-coupon bond*** *is a zero-coupon bond that allows the issuer the right, but not the obligation, to buy back from the bond holders at* ***pre-specified prices*** *on the* ***pre-specified call dates****.*

Remark 3.9.5 (understand how callability benefits issuers/borrowers)**.**

- Suppose a company issue a callable zero-coupon bond at price 0.1 with maturity five years later. The embedded call option has call date four year later with call price 0.8.
- Suppose the interest rate drops, and therefore the bond price drops during the next four years. Particularly at the call date, the price of 1-year tenor zero-coupon bond is at 0.9.
- Then the company can issue/sell 1 year zero-coupon bond at 0.9 and then buy back the previous bond at 0.8. In the process, the company gain 0.1.
- In summary, the embedded call option gives the issuer the right to refinance the debt when there is a lower interest rate in the market.

Lemma 3.9.13. *Consider a callable zero-coupon bond with bond maturity date at T. Let $T_c < T$ be the maturity date of the embedded call option and let K be the strike, then the value at time t is given by*

$$V(t) = E_Q[\exp(-\int_t^T r(s)ds)|\mathcal{F}_t] - E_Q[\exp(-\int_t^{T_c} r(s)ds)(P(T_c,T)-K)^+|\mathcal{F}_t].$$

In other words,

$$price\ of\ callable\ bond = price\ of\ normal\ bond - price\ of\ embedded\ option.$$

Proof. We can analyze the discounted cash flow to the bond holders at time T and time T_c:

$$\begin{aligned}
V(t) &= E_Q[\frac{B(t)}{B(T)}\mathbf{1}_{P(T_c,T)<K}|\mathcal{F}_t] + E_Q[\frac{B(t)}{B(T_c)}K\mathbf{1}_{P(T_c,T)>K}|\mathcal{F}_t] \\
&= E_Q[\frac{B(t)}{B(T)}|\mathcal{F}_t] - E_Q[\frac{B(t)}{B(T)}\mathbf{1}_{P(T_c,T)>K}|\mathcal{F}_t] + E_Q[\frac{B(t)}{B(T_c)}K\mathbf{1}_{P(T_c,T)>K}|\mathcal{F}_t] \\
&= E_Q[\frac{B(t)}{B(T)}|\mathcal{F}_t] - E_Q[\frac{B(t)}{B(T_c)}\frac{B(T_c)}{B(T)}\mathbf{1}_{P(T_c,T)>K}|\mathcal{F}_t] + E_Q[\frac{B(t)}{B(T_c)}K\mathbf{1}_{P(T_c,T)>K}|\mathcal{F}_t] \\
&= E_Q[\frac{B(t)}{B(T)}|\mathcal{F}_t] - E_Q[\frac{B(t)}{B(T_c)}E[\frac{B(T_c)}{B(T)}|\mathcal{F}_{T_c}]\mathbf{1}_{P(T_c,T)>K}|\mathcal{F}_t] + E_Q[\frac{B(t)}{B(T_c)}K\mathbf{1}_{P(T_c,T)>K}|\mathcal{F}_t] \\
&= E_Q[\frac{B(t)}{B(T)}|\mathcal{F}_t] - E_Q[\frac{B(t)}{B(T_c)}P(T_c,T)\mathbf{1}_{P(T_c,T)>K}|\mathcal{F}_t] + E_Q[\frac{B(t)}{B(T_c)}K\mathbf{1}_{P(T_c,T)>K}|\mathcal{F}_t] \\
&= E_Q[\frac{B(t)}{B(T)}|\mathcal{F}_t] - E_Q[\frac{B(t)}{B(T_c)}(P(T_c,T)-K)\mathbf{1}_{P(T_c,T)>K}|\mathcal{F}_t] \\
&= E_Q[\exp(-\int_t^T r(s)ds)|\mathcal{F}_t] - E_Q[\exp(-\int_t^{T_c} r(s)ds)(P(T_c,T)-K)^+|\mathcal{F}_t]
\end{aligned}$$

□

3.10 Notes on bibliography

Classical books on interest rate theory includes [3][1][14][15][16][17].

[18]

For comprehensive treatment on the LIBOR market, see [12].

For practical topics, e.g., calibration, hedging, and correlation modeling, see [19].

For convexity correction, see [20][21].

For yield curve construction, see [22].

For multi-curve framework, see[23][24].

Problems and solutions in mathematical finance: interest rates and inflation indexed derivatives

For inflation indexed derivatives, see [25].

BIBLIOGRAPHY

1. McInerney, D. & Zastawniak, T. *Stochastic Interest Rates* ISBN: 9781107002579. https://books.google.com/books?id=8JTwCAAAQBAJ (Cambridge University Press, 2015).
2. Corb, H. *Interest Rate Swaps and Other Derivatives* ISBN: 9780231530361. https://books.google.com/books?id=Gc0W0Y1CsDgC (Columbia University Press, 2012).
3. Brigo, D. & Mercurio, F. *Interest rate models-theory and practice: with smile, inflation and credit* (Springer Science & Business Media, 2007).
4. Wu, L. *Interest rate modeling: Theory and practice* (CRC Press, 2009).
5. Hull, J. *Options, Futures and Other Derivatives 9 edition* (Pearson Education, 2009).
6. Pelsser, A. *Efficient methods for valuing interest rate derivatives* (Springer Science & Business Media, 2000).
7. Wilmott, P. *Paul Wilmott on quantitative finance* (John Wiley & Sons, 2013).
8. Shreve, S. E. *Stochastic calculus for finance II: Continuous-time models* (Springer Science & Business Media, 2004).
9. Privault, N. *An elementary introduction to stochastic interest rate modeling* (World scientific, 2012).
10. Heath, D., Jarrow, R. & Morton, A. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica: Journal of the Econometric Society,* 77–105 (1992).
11. Cairns, A. J. *Interest rate models: an introduction* (Princeton University Press, 2004).
12. Rebonato, R., McKay, K. & White, R. *The SABR/LIBOR Market Model: Pricing, calibration and hedging for complex interest-rate derivatives* (John Wiley & Sons, 2011).
13. Hagan, P. K. & Kumar, D. Managing smile risk. *Wilmott Magazine,* 84–108.
14. Andersen, L. & Piterbarg, V. *Interest Rate Modeling, Volume I Interest Rate Modeling* **v. 1.** ISBN: 9780984422104. https://books.google.com/books?id=oI2fcQAACAAJ (Atlantic Financial Press, 2010).
15. Andersen, L. & Piterbarg, V. *Interest Rate Modeling, Volume II Interest Rate Modeling* **v. 2.** ISBN: 9780984422111. https://books.google.com/books?id=tgmucQAACAAJ (Atlantic Financial Press, 2010).

16. Andersen, L. & Piterbarg, V. *Interest Rate Modeling, Volume III Interest Rate Modeling* **v. 3.** ISBN: 9780984422128. https://books.google.com/books?id=RxHVcQAACAAJ (Atlantic Financial Press, 2010).

17. Jarrow, R. A. *Modeling fixed-income securities and interest rate options* (Stanford University Press, 2002).

18. Filipovic, D. *Term-Structure Models: A Graduate Course* ISBN: 9783540680154. https://books.google.com/books?id=KqcSh6CavaAC (Springer Berlin Heidelberg, 2009).

19. Rebonato, R. *Volatility and correlation: the perfect hedger and the fox* (John Wiley & Sons, 2005).

20. Pelsser, A. Mathematical foundation of convexity correction. *Quantitative Finance* **3,** 59–65 (2003).

21. Hagan, P. S. Convexity conundrums: Pricing CMS swaps, caps, and floors. *The Best of Wilmott,* 305 (2005).

22. Methods for constructing a yield curve. *WILMOTT Magazine, May,* 70–81. http://finmod.co.za/MethodsForConstructingAYieldCurve.pdf (2008).

23. Grbac, Z. & Runggaldier, W. J. *Interest rate modeling: post-crisis challenges and approaches* (Springer, 2015).

24. Henrard, M. *Interest rate modelling in the multi-curve framework: Foundations, evolution and implementation* (Springer, 2014).

25. Kerkhof, J. Inflation derivatives explained (2005).

4

FIXED INCOME ANALYTICS & STRATEGIES

4.1 Bond markets and products

A fixed income security is an investment that provides a return in the form of fixed periodic payments and the eventual return of principal at maturity. Unlike a variable-income security, where payments change based on some underlying measure such as short-term interest rates, the payments of a fixed-income security are known in advance.

The most common type of fixed income instruments are bonds issued by corporates, government, and banks. Fixed income instruments are heavily used by issuers to raise funding for business expansion, government expenditure.

Fixed income instruments analytics have evolve from simple discounting from sophisticated interest rate and credit term structure modeling

We start this chapter with time value of money calculation. We will cover defaultable bond, and mortgage back securities

4.1.1 Vanilla bonds

4.1.1.1 *Fixed rate bonds*

Definition 4.1.1 (bond). *The bond is a debt security, under which the issuer owes the holders a debt and is obliged to pay them fixed amont of interest (termed, the coupon) at specified coupon dates and to repay the principal at the maturity date.*

Definition 4.1.2 (coupon rates: stated coupon rate and effective annual rate).

4.1.1.2 *Treasury bonds*

Definition 4.1.3 (Treasury bonds).

- *On-the-run Treasury bonds are the most recently issued U.S. Treasury bonds or notes of a particular maturity.*
- *Off-the-run Treasury bonds refer to Treasury securities that have been issued before the most recent issue and are still outstanding.*
- *A Treasury transitions from onthe-run to off-the-run once a newer set of Treasuries is released for sale.*
- *Because on-the-run issues are the most liquid, they typically trade at a slight premium and therefore yield a little less than their off-the-run counterparts.*

4.1.2 Floating rate note

Definition 4.1.4. *Floating rate notes (FRNs, floaters) have a variable coupon that is linked to a reference rate of interest, such as Libor or Euribor. For example, the coupon may be defined as three-month USD LIBOR + 0.20%. The coupon rate is resetted periodically, typically every one or three months.*

Remark 4.1.1 (interest rate risk characteristics of floating rate note)**.**

- Floating rate note usually has little exposure to interest rate risk. Consider a sudden rise in market interest rates happened just before one of the floater's reset dates, the holder of the floater would temporarily be receiving a coupon based on a below-market rate. However, at the next reset date, floater's coupon would adjust to prevailing market rates. Such interest rate risk is also called reset risk.

4.1.3 Bonds with options

Definition 4.1.5 (callable bond). *[1, p. 34] A **callable bond** gives the issuer the right to redeem all or part of the bond before the specif ed maturity date.*

The available exercise styles include American call, European call, and Bermudian style call.

Remark 4.1.2 (motivation and intuition of callable bond). [1, p. 34]

- Te primary reason why issuers choose to issue callable bonds rather than non-callable bonds is to protect themselves against a decline in interest rates. If market interest rates fall or credit quality of the issuer improves, the issuer of a callable bond has the right to replace an old, expensive bond issue with a new, cheaper bond issue.
- For example, assume that the market interest rate was 6% at the time of issuance and that a company issued a bond with yield 7%. Now assume that the market interest rate has fallen to 4% and that the company's creditworthiness has not changed; it can still issue at the market interest rate plus 100 bps. If the original bond is callable, the company can redeem it and replace it with a new bond paying 5% annually.
- Callable bonds present investors with a higher level of reinvestment risk than non-callable bonds; that is, if the bonds are called, bondholders have to reinvest funds in a lower interest rate environment. For this reason, callable bonds have to offer at a higher yield and sell at a lower price than otherwise similar non-callable bonds.

Example 4.1.1. [1, p. 36] Consider a 30 Y callable bond(face value 100) issued on Aug 2012 at a price of 98.195. The bond is callable according to the following schedule.

Year	Call Price	Year	Call Price
2022	103.870	2028	101.548
2023	103.485	2029	101.161
2024	103	2030	100.774
2025	102.709	2031	100.387
2026	102.322	$\geq$2032	100
2027	101.955		

We can see that

- The call protection is upto 2022.
- The bond is sold at discount. The call price is always greater than the par in order to protect a decline of interest rate(in such case, the bond price can exceed 100).

Definition 4.1.6 (putable bond). *[1, p. 36] A **putable bond** gives the bondholders the right to sell the bond back to the issuer at a pre-determined price on specified dates.*

Remark 4.1.3 (motivation and intuition of putable bond)**.** [1, p. 36]

- Putable bonds are beneficial for the bondholder by guaranteeing a pre-specified selling price at the redemption dates.
- If interest rates rise after this issue date, thus depressing the bond's price, the bondholders can put the bond back to the issuer and get cash. This cash can be reinvested in bonds that offer higher interest.
- The price of a putable bond will be higher than the price of an otherwise similar bond issued without the put provision. Similarly, the yield on a bond with a put provision will be lower than the yield on an otherwise similar non-putable bond.

Definition 4.1.7 (convertible bond). *[1, p. 37] A **convertible bond** is a hybrid security with both debt and equity features. It gives the bondholder the right to exchange the bond for a specified number of common shares in the issuing company. Thus, a convertible bond can be viewed as the combination of a straight bond (option-free bond) plus an embedded equity call option.*

Remark 4.1.4 (motivation and intuition of convertible bond)**.** [1, p. 37]

- From the investor's perspective, a convertible bond offers several advantages relative to a non-convertible bond.
  - it gives the bondholder the ability to convert into equity in case of share price appreciation. At the same time, the bondholder receives downside protection; if the share price does not appreciate, the convertible bond offers the comfort of regular coupon payments and the promise of principal repayment at maturity.
  - Even if the share price decline, the price of a convertible bond cannot fall below the price of the straight bond because it contains optionality.
- Because the conversion provision is valuable to bondholders, the price of a convertible bond is higher than the price of an otherwise similar bond without the conversion provision. Similarly, the yield on a convertible bond is lower than the yield on an otherwise similar non-convertible bond.
- From the issuer's perspective, convertible bonds offer two main advantages.
  - The first is reduced interest expense. Issuers are usually able to offer below-market coupon rates because of investors' attraction to the conversion feature.
  - The second advantage is the elimination of debt if the conversion option is exercised. But the conversion option is dilutive to existing shareholders

4.2 Cash flow mathematics: growth and discounting

4.2.1 Compounding convention and future values

Definition 4.2.1. *[2, p. 592][3, p. 9] Given a annualized interest rate R, the interest rate calculation depends on the compounding convention.*

- *If the compounding convention is continuous, then unit notional in n years will grow to*
$$\exp(Rn).$$
- *If the compounding convention is once per year, then unit notional in n years will grow to*
$$(1+R)^n.$$
- *If the compounding convention is m times per year, then unit notional in n years will grow to*
$$(1+R/m)^{mn}.$$

Lemma 4.2.1 (conversion). *[2, p. 592] Let R_c be the annual continuous compounding rate. Let R_m be the annual compounding rate with frequency m. Then we have*

$$R_c = m\ln(1+\frac{R_m}{m}),$$

and

$$R_m = m(\exp(R_c/m) - 1).$$

Proof. Consider unit notional, we have

$$\exp(R_c n) = (1+R_m/m)^{mn}.$$

□

Example 4.2.1. [3, p. 9] Consider a bank offers a CD with a two year maturity($N = 2$), a stated annual interest rate of 8 percent compounded quarterly ($m = 4$), and a feature allowing reinvestment of the interest at the same interest rate. Let the notional be 1$, The value of the CD at maturity is given by

$$V = (1+\frac{r}{m})^{mN} = (1.02)^8 = 1.171659.$$

4.2.2 Present value of cash flow

Definition 4.2.2 (present value of a single cash flow in future). *[3, p. 15]*

- *Given a future cash flow FV that is to be received in N periods and an interest rate per period of r, its present value is given by*

$$PV = FV\frac{1}{(1+r)^N}.$$

- *Suppose the compounding frequency is m per year, and the annual interest rate is r. Then a future cash flow FV N years later has present value given by*

$$PV = FV\frac{1}{(1+\frac{r}{m})^{Nm}}.$$

Remark 4.2.1 (present value calculation depends on the convention)**.** Note that different compounding convention can lead to different present value calculation.

Lemma 4.2.2 (present value of annuity and perpetuity). *Consider an annuity with N payments of A occurs on period indexed from $t = 1$ to $t = N$. Let current time be $t = 0$. Then the present value of the annuity is given by*

$$\begin{aligned} PV &= \frac{A}{(1+r)} + \frac{A}{(1+r)^2} + \cdots + \frac{A}{(1+r)^N} \\ &= \frac{A}{r}(1 - \frac{1}{(1+r)^N}) \end{aligned}$$

Particularly for a perpetuity(in which $N \to \infty$, its present value is given by

$$PV = \frac{A}{r}.$$

Proof. Note that from the summation formula for geometric series

$$\begin{aligned} PV &= \frac{A}{(1+r)} + \frac{A}{(1+r)^2} + \cdots + \frac{A}{(1+r)^N} \\ &= \frac{A}{1+r}\left(\frac{1-\frac{1}{(1+r)^N}}{1-\frac{1}{1+r}}\right) \\ &= \frac{A}{1+r}\left(1-\frac{1}{(1+r)^N}\right)\frac{1+r}{r} \\ &= \frac{A}{r}\left(1-\frac{1}{(1+r)^N}\right) \end{aligned}$$

□

4.2.3 Different interest rates

Definition 4.2.3 (repurchase agreement, repo, repo rate). *[4, p. 171][5, p. 79] A repurchase agreement(repo) is an agreement by an owner of securities(usually bonds) to sell the securities to investors(or other counter-parties) at one price and buy them back at another price(usually higher) in a future date.*

Usually in repo, it requires collateral for borrowing. Therefore, the interest rate implied by the repo price is riskless interest rate,called ***repo rate****.*

Definition 4.2.4 (Treasury rate). *[2, p. 178] Treasury rates are the interest rates an investor earns on Treasury bills and Treasury bonds; These are the instruments used by a government to borrow in its own money.*

Remark 4.2.2 (more about Treasury rates). [2, p. 178]

- It is usually assumed that there is no chance that a government will default on a debt denominated its own currency; therefore Treasury rates are usually regarded as risk-free rates.
- In practice, they are regarded as artificially low because:
  - the amount of capital a bank is required to hold to support an investment in Treasury bills and bonds is substantially lower than the capital requirement for a similar investment.
  - In the US, Treasury instruments are given a favorable tax treatment compared with other fixed income investments.

Definition 4.2.5 (overnight indexed swap(OIS), OIS rate). *[5, p. 202] An **overnight indexed swap** is a swap where a fixed rate for a period(e.g., 1 month or 3 months) is exchanged for the geometric average of the overnight unsecured borrowing rate between financial institutions in the US.*

*The fixed rate in an OIS is referred to as the **OIS rate**.*

Remark 4.2.3 (LIBOR, Red Fund rate, repo rate). [5, pp. 79, 201]

- Both LIBOR and fed fund rate are unsecured borrowing rates. On average, overnight LiBOR rate has been about 6 basis point higher than the fed fund rate.
- The repo rate is usually slightly below fed fund rate.
- Interest rate implied from US Treasury bill is considered artificially low. It is the lowest.
- After the 2007 crisis, OIS rate is used for discounting when valuing collateralized derivatives; LIBOR rate is used for discounting when valuing non-collateralized derivatives. OIS rate now is viewed as risk-free rate and it is lower than LIBOR.

4.3 Bond valuation

4.3.1 Valuation via single rate discounting

4.3.1.1 *Basics*

Definition 4.3.1 (market discount rate, required yield, required rate of return). *[1, p. 92] The **market discount rate** is used in the time-value-of-money calculation to obtain the present value. The market discount rate is the rate of return required by investors or market as a whole given the risk of the investment in the bond. It is also called the **required yield, or the required rate of return.***

Definition 4.3.2 (valuation of bond using single rate model). *Let current time be* 0. *Let* $t_1, t_2, ..., t_n$ *be future times when fixed cash flow* $c_1, c_2, ..., c_n$ *of the bond occur . Let* y *be the market required yield for the bond observed in the market. Then*

- *the price for a continuously-compounding bond is*

$$P = \sum_{i=1}^{n} c_i e^{-y t_i}.$$

- *the price y for an annually simple compounding bond is*

$$P = \sum_{i=1}^{n} c_i \frac{1}{(1+y)^{t_i}}.$$

- *the price for a bond with simple compounding m times per year is*

$$P = \sum_{i=1}^{n} c_i \frac{1}{(1+y/m)^{mt_i}}.$$

Remark 4.3.1 (Issues with single discount rate model)**.**

- **(Not taking into account term structure of interest rate)** Note that the single discount rate model is problematic since does not take into account term structure of interest rate; that is, we are implicitly assuming the forward rate is flat (or equivalently, the short rate is constant.). In practice, we would expect different interest rates for different investment horizons.
- **(Pricing depends on compounding convention)**. Note that the compounding convention is not in the bond's contract parameters.

Definition 4.3.3 (implied bond yield-to-maturity). *Let current time be* 0*. Let* $t_1, t_2, ..., t_n$ *be future times when fixed cash flow* $c_1, c_2, ..., c_n$ *of the bond occur .* ***Let*** P ***be the current price of the bond observed on the market****. Then*

- *the* ***yield-to-maturity*** *for a continuously-compounding bond is a number y such that*

$$P = \sum_{i=1}^{n} c_i e^{-yt_i}.$$

- *the* ***yield-to-maturity*** *for an annually simple compounding bond is a number y such that*

$$P = \sum_{i=1}^{n} c_i \frac{1}{(1+y)^{t_i}}.$$

- *the* ***yield-to-maturity*** *for a bond with simple compounding m times per year is a number y such that*

$$P = \sum_{i=1}^{n} c_i \frac{1}{(1+y/m)^{mt_i}}.$$

The bond yield-to-maturity is the implied market discount rate and depends on the bond coupon payment frequency.

Example 4.3.1. [1, p. 97] Consider a bond with a 4Y tenor, a par face 100, a coupon rate 3.5%, and 4 annual coupon payments. Suppose we observe the market price to be 103.75. Then

$$103.75 = \frac{3.5}{1+y} + \frac{3.5}{1+y} + \frac{3.5}{1+y} + \frac{3.5}{1+y} \implies y = 0.02503.$$

Note 4.3.1 (caution! bond yield-to-maturity vs. yield curve). [6, p. 102]

- Only for zero-coupon bond, its yield-to-maturity for different maturities is equivalent to the yield curve.
- For any other coupon bonds, their yield-to-maturity for different maturities is usually larger than the yield curve; the difference is due to the intermediate cash flow.

Remark 4.3.2 (bond yield vs. zero-coupon bond price). For a bond with single payment c occurring at T, we have

$$y = \frac{1}{T}\ln\frac{P(0,T)}{c},$$

where $P(0,T)$ is the zero-coupon bond(maturing at T) price at time 0.

Note 4.3.2 (interpret yield-to-maturity). [1, p. 96] Te yield-to-maturity is the rate of return on the bond to an investor given three critical assumptions:

- The investor holds the bond to maturity.If the bond holder does not hold the bond to maturity, and sell the bond before it matures, then the return might be negative since the bond can depreciates due to credit worsening,interest rate rises, etc.
- The issuer does not default on any of the payments.
- The investor is able to reinvest coupon payments at that same yield. For example(see below table), consider a 2Y coupon bond with coupon payment at 1Y. If we reinvest the coupon at return rate y, then at $T = 2Y$, we will get a cash flow $C(1+y)+N+C$, whose present value is

$$PV = \frac{1}{(1+y)^2}(C(1+y)+N+C) = \frac{C}{1+y} + \frac{N+C}{(1+y)^2} = P.$$

	action	cash flow
T=0	buy the coupon bond at price P	-P
T=1	receive the coupon and reinvest the coupon	+C - C
T=2	receive the final coupon, face value and the proceeds of the reinvestment	C(1+y) + N + C

However, in reality, not all the investors will hold the bond to maturity and reinvest the coupons(even reinvest the coupon at the same rate might not be possible since the rate is changing). Therefore, the yield-to-maturity does not necessarily reflect the rate of return in investing a bond.

Proposition 4.3.1 (No arbitrage condition for bond yield-to-maturity). *Assuming*

- *Bonds are defaultable.*
- *All investors will reinvest them coupons.*
- *The forward rate is flat such that investor can reinvest at the same rate.*
- *The market is free of arbitrage.*

- *For all default-free bonds(for any coupon schedule it has) maturing at T, its yield-to-maturity must equal the market discount rate or required yield corresponding to that maturity T.*
- *If holding and reinvesting until maturity, then investing in any coupon-bearing bond is equivalent to investing in zero-coupon bond with the same maturity. In other words, given the same initial cash, buying zero-coupon bond and the coupon-bearing bond will have the same rate of return.* [a]

[a] Note that coupon-bearing bond will be more expensively than the zero-coupon bond.

Proof. (1) Consider two default-free bonds have the same maturity date T. The first bond has the market required yield, whereas the second bond has the higher yield-to-maturity than the required yield. Then we can short the first bond(or borrow money at the market required yield) and long and hold the second bond to maturity. At maturity, the cash flow we get from the second bond(including reinvesting coupon received) will payoff the borrowed money and gain extra money. See also 4.3.2. (2)

□

Lemma 4.3.1 (bond price and coupon rate relationship). *Consider a coupon bond with coupon rate c,maturity T and face value N. Let the market required yield be $y(T)$. Denote the bond price by P. Then*

- *If $c = y(T)$, then $P = N$. That is, the bond is* ***sold at par.***
- *If $c < y(T)$, then $P < N$. That is, the bond is* ***sold at discount.***
- *If $c > y(T)$, then $P > N$. That is, the bond is* ***sold at premium.***

Proof. □

Example 4.3.2. [1, p. 93]

- Consider a 5 Y bond. The annual coupon rate is 4%, the yield is 6%, and the face value is 100. The bond price is sold at 91.175 because
$$\frac{4}{1.06} + \frac{4}{1.06^2} + \frac{4}{1.06^3} + \frac{4}{1.06^4} + \frac{104}{1.06^5} = 91.575.$$
- Consider a 5 Y bond. The annual coupon rate is 8%, the yield is 8%, and the face value is 100. The bond price is sold at 91.175 because
$$\frac{8}{1.06} + \frac{8}{1.06^2} + \frac{8}{1.06^3} + \frac{8}{1.06^4} + \frac{108}{1.06^5} = 108.425.$$
- Consider a 5 Y bond. The annual coupon rate is 6%, the yield is 6%, and the face value is 100. The bond price is sold at 100 because
$$\frac{6}{1.06} + \frac{6}{1.06^2} + \frac{6}{1.06^3} + \frac{6}{1.06^4} + \frac{106}{1.06^5} = 100.$$

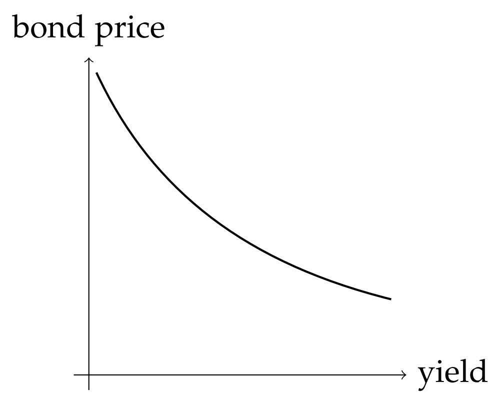


Figure 4.3.1: general price yield relationship. Note its convexity.

4.3.1.2 *Qualitative treatment on risks*

Remark 4.3.3.

- An investor who holds a fixed-rate bond to maturity will earn an annualized rate of return equal to the YTM of the bond when purchased.
- An investor who sells a bond prior to maturity will earn a rate of return equal to the YTM at purchase if the YTM at sale has not changed since purchase.

- If the market YTM for the bond, our assumed reinvestment rate, increases (decreases) after the bond is purchased but before the first coupon date, abuy-and-hold investor's realized return will be higher (lower) than the YTM of the bond when purchased.
- If the market YTM for the bond, our assumed reinvestment rate, increases after the bond is purchased but before the first coupon date, a bond investor will earn a rate of return that is lower than the YTM at bond purchase if the bond is held for a short period.
- If the market YTM for the bond, our assumed reinvestment rate, decreases after the bond is purchased but before the first coupon date, a bond investor will earn a rate of return that is lower than the YTM at bond purchase if the bond is held for a long period

4.3.2 Valuation via zero-coupon bond curve

Definition 4.3.4 (valuation of bond using zero-coupon bond curve). *Let current time be* 0. *Let* $t_1, t_2, ..., t_n$ *be future times when fixed cash flow* $c_1, c_2, ..., c_n$ *of the bond occur. Let* $DF(0,t)$ *be the market observed zero-coupon bond curve. Then the price for a continuously-compounding bond is*

$$P = \sum_{i=1}^{n} c_i DF(0, t_i).$$

4.3.3 Dirty and clean price convention

Definition 4.3.5 (dirty and clean price, accrued interest). *Consider a set of coupon payment dates* $t_0 < t_1 < ... < t_N$. *Let current time be* $t_S, t_0 \leq t_S < t_1$, *and the bond coupon rate be* C.

- ***Accrued interest*** *is defined as*

$$AI(t_S) = \begin{cases} (t_S - t_0) \times C, t_0 \leq t_S < t_1 \\ 0, t_S = t_1 \ (after\ coupon\ payment) \end{cases}$$

- ***Dirty price*** *of a bond is the price of a bond including any interest that has accrued since issue of the most recent coupon payment.*
  - *In the single rate discounting model,*

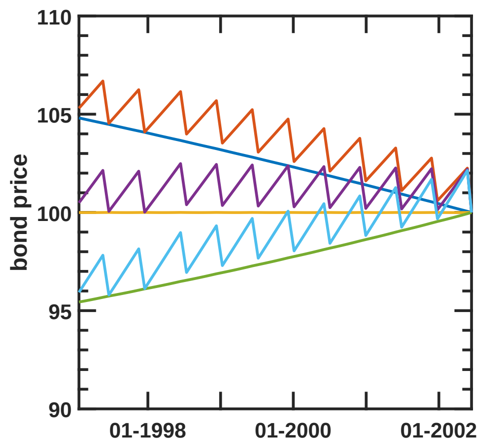
$$PV_{dirty} =$$

  - *In the curve discounting model,*
- ***Clean price*** *of a bond is the price of a bond excluding the accrued interest.* ***Clean prices are usually quoted.***

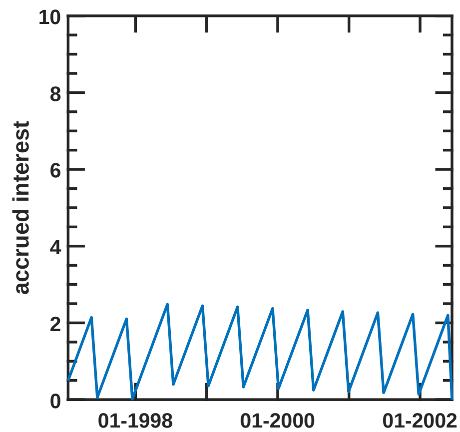
- $$dirty\ price = clean\ price + accrued\ interest$$

Remark 4.3.4 (understand dirty price and clean price).

- Dirty price is the NPV of future cash flows; it is the fair market price.
- Clean price is not the NPV of future cash flows; it is not the fair market price.
- Clean price is solely for quoting purpose. Dirty price is suitable for quoting due to its jigsaw-like pattern (see Figure 4.3.2).
- Dirty price is always greater than clean price.



(a) Bond dirty (with saw-tooth pattern) and clean (smooth) prices trajectories for bonds sold at premium, at par and at discount. The bond coupon rate is fixed and the market required yield is varied.



(b) Bond accured interest as a function of time. Immediately after coupon payment date, it will drop to zero. Accured interst only depends on day count and coupon rate

Figure 4.3.2: Bond clean price, dirty price, and accrued interest.

Example 4.3.3. [1, p. 107] A 6% German corporate bond is priced for settlement on 18 June 2015. The bond makes semiannual coupon payments on 19 March and 19 September of each year and matures on 19 September 2026. The corporate bond uses the 30/360 day-count convention for accrued interest. Suppose the stated annual yield-to-maturity is 5.8%.

Note that there are 23 periodic payments starting from Sep-19-2015 to Sep-19-2026. Also there are 89 days between Mar-19 and Jun-18, or a semi-year fraction of 89/180.

Therefore, the dirty price is

$$PV_{dirty} = (1.029)^{89/180} \times (\frac{3}{1.029} + \frac{3}{1.029^2}) + \cdots + \frac{103}{1.029^{23}} = 103.109;$$

and the accrued interest is

$$AI = 3 \times \frac{89}{180} = 1.4833.$$

4.3.4 Discussion on different approaches

4.3.5 Forward bond price

Proposition 4.3.2 (No-arbitrage forward price for bonds).

- *Consider a zero-coupon coupon has maturity T and spot price $Z(t,T)$. Its forward price with delivery date τ is given by*

$$FB(t,T) = \frac{Z(t,T)}{Z(t,\tau)};$$

if in multi-curve environment, then

$$FB(t,T) = Z(t,T)(1 + F(t,t,\tau)(\tau - t)),$$

where $F(t,t,\tau)$ is the spot LIBOR rate with tenor $[t,\tau]$.
- *Consider a bond with coupon payment c_i at $T_i, i = 1,2,...,N$. The maturity date is $T = T_N$. Let S_t be its* ***spot dirty price*** *its* ***forward dirty price*** *with delivery date $\tau < T$ is given by*

$$F_t = \frac{S_t - \sum_{i=1}^{k} c_i Z(t,T_i)}{Z(t,\tau)},$$

where $T_1 < T_2 < \cdots < T_k < \tau$; If in multi-curve environment,

$$F_t = S_t + \underbrace{S_t(\tau - t)F(t,t,\tau)}_{\text{funding cost}} - \underbrace{\sum_{i=1}^{k} c_i(1 + F(t,T_i,\tau)(\tau - T_i))}_{\textbf{benefits}},$$

where $F(t,T_i,\tau)$ is the forward rate with tenor $[T_i,\tau]$.
- *Let AI denote the accrued interest. Let S_t, F_t denote the clean spot and forward prices. They are given by*

$$F_t = \frac{S_t + AI(t) - \sum_{i=1}^{k} c_i Z(t,T_i)}{Z(t,\tau)} - AI(\tau),$$

and

$$F_t = S_t + AI(t) + S_t(\tau - t)F(t,t,\tau) - \sum_{i=1}^{k} c_i(1 + F(t,T_i,\tau)(\tau - T_i) - AI(\tau).$$

Proof. See Theorem 3.3.3. □

4.4 Bond risks

4.4.1 Time path of bond price

4.4.1.1 *Time path with yield curve unchanged*

Lemma 4.4.1. *Consider a zero-coupon bond with fixed maturity T. Assume constant short rate r. Then its time path is given by*

$$P(t;T) = \exp(-r(T-t)).$$

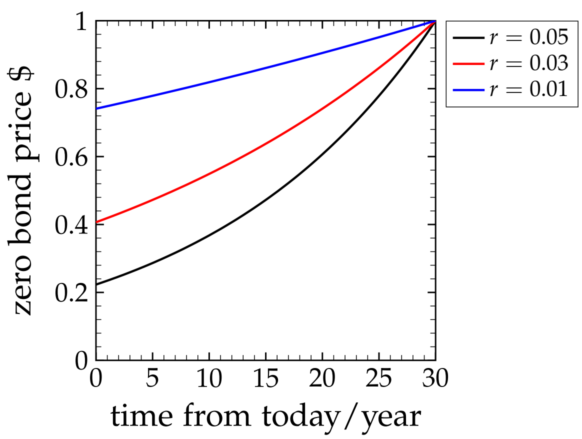


Figure 4.4.1: 30 Y zero-coupon bond time path at different continuous compounding rate r

4.4.1.2 *Time path with changing yield curve*

Methodology 4.4.1 (simulate the future term structure in Hull-White model).

- *Suppose*

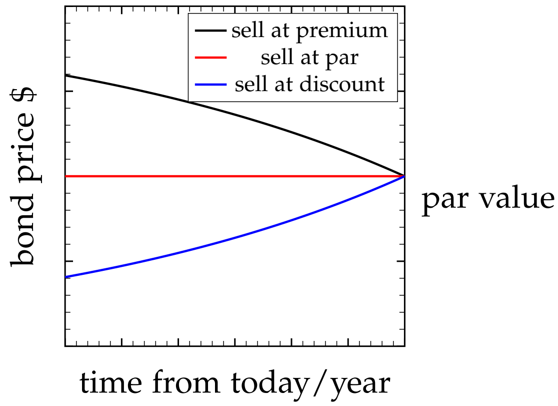


Figure 4.4.2: 30 Y zero-coupon bond time path at different continuous compounding rate r

- *current time is 0.*
- *we are given the current term structure* $y(0,T), T \in \mathbb{R}^+$*(or* $P(0,T), T \in \mathbb{R}^+$*, or* $f(0,t), \mathbb{R}^+$*).*
- *we are given a calibrated Hull-White model*

$$dr = (\theta(t) - \alpha r(t))dt + \sigma(t)dW_t.$$

- *we are given the initial short rate* $r(0)$*(which can be derived from* $r(0) = f(0,0)$*).*

- *Then we can generate a* ***sample term structure*** $y(t,T)$ *and the associated* ***sample bond price*** $V(t)$ *in future time* $t > 0$ *via the following procedure:*
  - *simulate* $r(t)$ *by drawing from the normal distribution*

$$r(t) \sim N(m(t), v^2(t)),$$

where

$$m(t) = r(0)e^{-\alpha t} + \int_0^t \theta(u)e^{-\alpha(t-u)}du,$$

$$v^2 = \int_0^t \sigma(u)^2 e^{-2\alpha(t-u)}du.$$

  - *the sample yield curve is given by*

$$y(t,T) = -\frac{\ln P(t,T)}{T-t} = -\frac{1}{T-t}(\ln \frac{P(0,T)}{P(0,t)} - A(t,T) - r(t)D(t,T)),$$

where

$$A(t,T) = -f(0,t)D(t,T) + \frac{1}{2}D^2(t,T)\int_0^t \sigma^2(u)\exp(-2\alpha(t-u))du,$$

$$D(t,T) = \int_t^T \exp(-\alpha(T-s))ds.$$

- *calculate the sample bond price* $V(t)$ *via*

$$V(t) = \sum_{i=1}^{m} \mathbf{1}(t_i > t)c_i DF(t,t_i),$$

where

$$DF(t,t_i) = \exp(-y(t,t_i)(t_i - t)).$$

- *Let* $V^{(i)}(t), i = 1,2,...,N$ *denote the* N *generated samples. Then we can use these samples to get the distribution of* $V(t)$. *Note that* $V(t)$ *is a random variable.*

Proof. (1) For the future term structure generation method via Hull-White model, see 3.5.5. (2) Note that from risk-neutral pricing theory

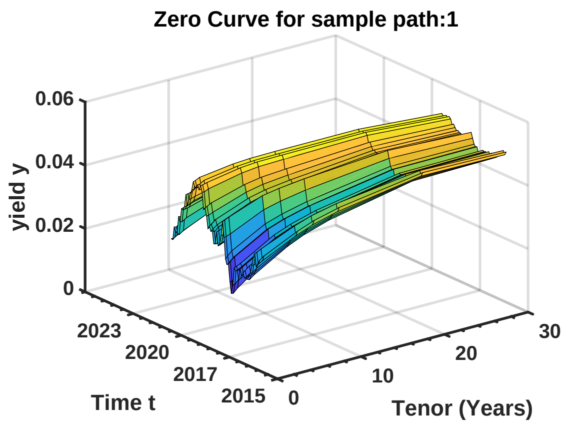
$$\begin{aligned} V(t) &= E_Q[\sum_{i=1}^{m} \mathbf{1}(t_i > t)c_i \exp(-\int_t^{t_i} r(s)ds)|\mathcal{F}_t] \\ &= \sum_{i=1}^{m} \mathbf{1}(t_i > t)c_i E_Q[\exp(-\int_t^{t_i} r(s)ds)|\mathcal{F}_t] \\ &= \sum_{i=1}^{m} \mathbf{1}(t_i > t)c_i DF(t,t_i) \end{aligned}$$

where by definition

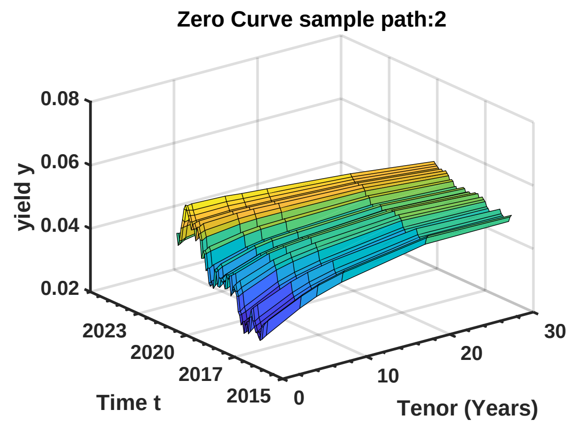
$$DF(t,t_i) = \exp(-y(t,t_i)(t_i - t)) = E_Q[\exp(-\int_t^{t_i} r(s)ds)|\mathcal{F}_t].$$

is a random variable. □

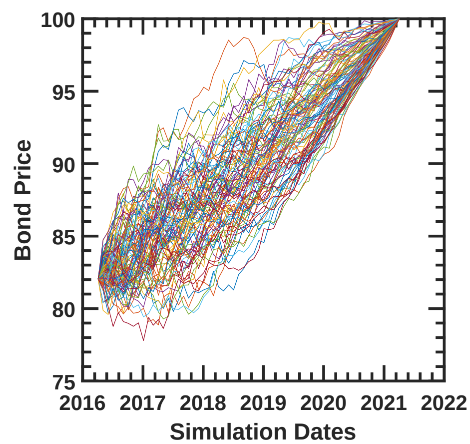
Example 4.4.1. As showed in Figure 4.4.3c, we use HW model to simulate the yield curve evaluation, and then obtain the time path of 5Y bond with evolving yield curves.



(a) The future yield curve in one simulation scenario.



(b) The future yield curve in another simulation scenario.



(c) Time path of a 5Y bond with simulating yield curve.

Figure 4.4.3: Bond dynamics.

4.4.2 Risk analytics in single rate model

4.4.2.1 *Duration*

Definition 4.4.1 (duration, Macaulay's duration). *[5, p. 95][2, p. 182] Given a bond P and the bond yield y such that*

$$P = \sum_{i=1}^{n} c_i e^{-y t_i}$$

where $c_i, i = 1, ..., n$ are cash flows at time $t_i, i = 1, ..., n$.

- *The* ***duration*** *of the bond is defined as*

$$D = -\frac{dP}{dy}\frac{1}{P} = \frac{\sum_{i=1}^{n} c_i t_i e^{-yt_i}}{\sum_{i=1}^{n} c_i e^{-yt_i}}.$$

- *An alternative definition of duration is*

$$D = \sum_{i=1}^{n} t_i(\frac{v_i}{P}),$$

where $v_i = c_i e^{-t_i y}$*;that is, duration can be viewed as a weighted average of the times when payments are made.*

- *The* ***dollar duration****of the bond is defined as*

$$D = -\frac{dP}{dy} = \sum_{i=1}^{n} c_i t_i e^{-yt_i}.$$

Definition 4.4.2 (modified duration). *[2, p. 182] Given a bond P and the bond yield y such that*

$$P = \sum_{i=1}^{n} c_i(1+\frac{y}{m})^{-mt_i}$$

where $c_i, i = 1, ..., n$ are cash flows at time $t_i, i = 1, ..., n$.

- *The* ***duration*** *of the bond is defined as*

$$D = -\frac{dP}{dy}\frac{1}{P} = \frac{\sum_{i=1}^{n} c_i t_i (1+\frac{y}{m})^{-mt_i}(1+\frac{y}{m})^{-1}}{P}.$$

- *An* ***modified duration*** *is defined as*

$$D_m = \sum_{i=1}^{n} t_i(\frac{v_i}{P}),$$

where $v_i = c_i(1+\frac{y}{m})^{-mt_i}$; that is, duration can be viewed as a weighted average of the times when payments are made.

- *The modified duration and duration are connected by $D_m = D(1+y/m)$.*

Remark 4.4.1. When m approaches ∞, duration and modified duration are equivalent.

Lemma 4.4.2. *Intuitively,* ***duration*** *measures how long, on average, a bondholder must wait to receive cash payments.*

- *A zero-coupon bond maturing in T periods has a duration of T;*
- *A coupon bond maturing in T periods has duration T since some payments are received before.*
- *The more coupon received before maturing, the smaller the duration.*

Lemma 4.4.3 (basic properties of duration).

- *Given y and T, duration is higher when C is lower.*
- *Given C and T, duration is lower when y is higher.*
- *Given C and y, duration is generally higher when T is higher (except for some deep discount coupon bonds).*
- *The duration of a portfolio P containing M bonds is*

$$D_P = \sum_{i=1}^{M} w_i D_i,$$

where w_i is the percentage weight of bond i in P.

Example 4.4.2 (weighted summation property). Consider the following three bonds with properties listed in the table.

tenor	**yield,%**	**duration**	**convexity**
2Y	7.71	1.78	0.041
5Y	8.35	3.96	0.195
10Y	8.84	6.54	0.568

Then a portfolio of \$5.4M of 2-year note and \$4.6M of the 10 year note will have duration

$$\frac{5.4 \times 1.78 + 4.6 \times 6.54}{10} = 3.96.$$

Lemma 4.4.4.

- *zero coupon bond*

$$D = T.$$

- *Flat perpetuity*

$$D = \frac{1}{y} + 1.$$

- *Flat annuity*

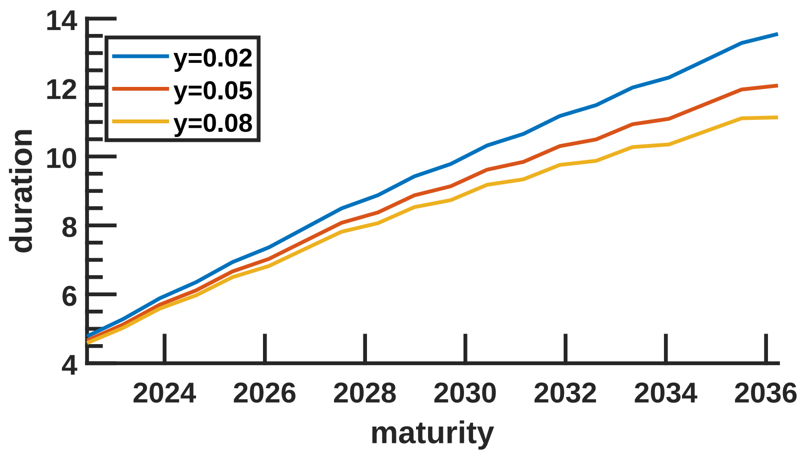
$$D = \frac{1}{y} + 1 - \frac{T}{(1+y)^T - 1}.$$

- *Coupon bond*

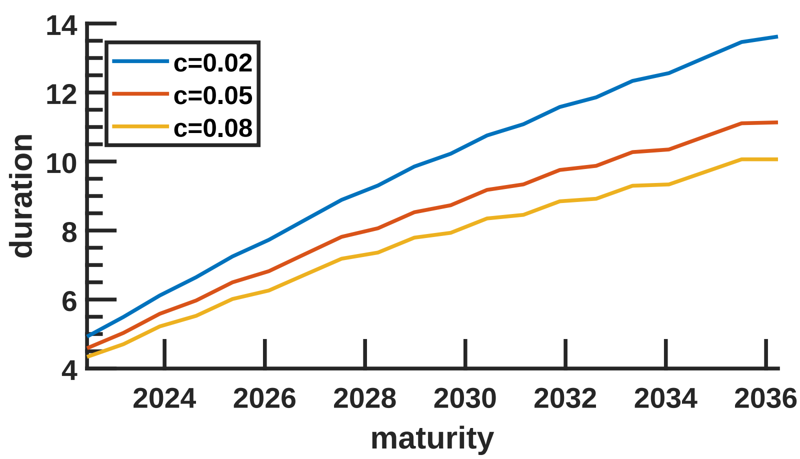
$$D = \frac{1}{y} + 1 - \frac{(1+y) + T(c-y)}{c[(1+y)^T - 1] + y}.$$

- *Coupon bond selling at par*

$$D = \frac{1+y}{y}(1 - \frac{1}{(1+y)^T}).$$



(a) Effects of yield and maturity on duration of a bond



(b) Effects of coupon rate and maturity on duration of a bond

Figure 4.4.4: Effects of coupon rate, yield, and maturity on the duration of a bond

4.4.2.2 *Convexity*

Definition 4.4.3 (convexity). *[5, p. 95] Given a bond P and the bond yield y such that*

$$P = \sum_{i=1}^{n} c_i e^{-y t_i}$$

where $c_i, i = 1, ..., n$ are cash flows at time $t_i, i = 1, ..., n$.

- *The **convexity** of the bond is defined as*

$$C = \frac{d^2 P}{dy^2} \frac{1}{P}.$$

- *The **dollar convexity** of the bond is defined as*

$$C = \frac{d^2 P}{dy^2}.$$

Lemma 4.4.5 (basic properties of convexity).

- *Convexity is always positive.*
- *A zero-coupon bond maturing in T periods has a duration of T;*
- *Given y and T, convexity is higher when C is lower.*
- *Given y and D, convexity is lower when C is lower.*
- *The convexity of a portfolio P containing M bonds is*

$$C_P = \sum_{i=1}^{M} w_i C_i,$$

where w_i is the percentage weight of bond i in P.

Example 4.4.3 (weighted summation property). Consider the following three bonds with properties listed in the table.

tenor	yield,%	duration	convexity
2Y	7.71	1.78	0.041
5Y	8.35	3.96	0.195
10Y	8.84	6.54	0.568

Then a portfolio of \$5.4M of 2-year note and \$4.6M of the 10 year note will have convexity

$$\frac{5.4 \times 0.041 + 4.6 \times 0.568}{10} = 0.283.$$

4.4.2.3 *Duration based hedging*

Lemma 4.4.6 (bond price changes due to yield change). *A change in yield Δy will lead to approximate change of ΔP given by ,*

$$\Delta P = \frac{dP}{dy}\Delta y + \frac{1}{2}\frac{d^2 P}{dy^2}\Delta y^2$$

or equivalently,

$$\frac{\Delta P}{P} = -D\Delta y + \frac{1}{2}C\Delta y^2$$

Proof. Taylor expansion. ☐

Note 4.4.1. Assumptions in duration based hedging:

- the yield curve is flat; that is, $R(t,T) = const$ for all maturities T and a fixed t.
- the yield curve is flat at each point in time; that is, $R(t,T) = const$ for all maturities T and a fixed t.
- the change in yield is small.

(Note that theoretically the price of a interest rate product will depends on the whole yield curve, instead of a single yield value assumed here).

Common hedging instrument includes

- Bonds
- Interest rate swap
- Interest rate futures
- Interest rate options

The strategy is the hold a dynamic portfolio contains the original portfolio and the hedging instruments such that the dynamic portfolio has zero duration.

Lemma 4.4.7 (hedging yield curve: parallel shift vs. non-parallel shift). *Consider a portfolio consisting of two bonds P_1 and P_2 such that $V = n_1 P_1 + n_2 P_2$. Denote the yield of the two bonds by y_1 and y_2, and their durations by D_1 and D_2.*

Define $\Delta V = V(y) - V(\Delta y)$, where y here is the whole yield curve. It follows that

- *Suppose yield curve will only shift in a parallel manner, then to the first order accuracy, if*

$$n_2 = -n_1 \frac{D_1 P_1}{D_2 P_2},$$

then

$$\Delta V = 0.$$

- *If the yield curve shifts in a non-parallel manner such that $\Delta y_1 \neq \Delta y_2$, then choosing*

$$n_2 = -n_1 \frac{D_1 P_1}{D_2 P_2} \frac{Cov(\Delta y_1, \Delta y_2)}{Var[\Delta y_2]}.$$

will minimize $Var[\Delta V]$.

Proof. (1) Note that

$$\Delta V = -n_1 P_1 D_1 \Delta y_1 - n_1 P_1 D_1 \Delta y_1,$$

where Δy_1 denotes the infinitesimal change of the yield curve occurs on y_1.

(2) Note that

$$Var[\Delta V] = (n_1 D_1 P_1)^2 Var[\Delta y_1] + (n_2 D_2 P_2)^2 Var[\Delta y_2] + 2 n_1 n_2 D_1 D_2 P_1 P_2 Cov(\Delta y_1, \Delta y_2).$$

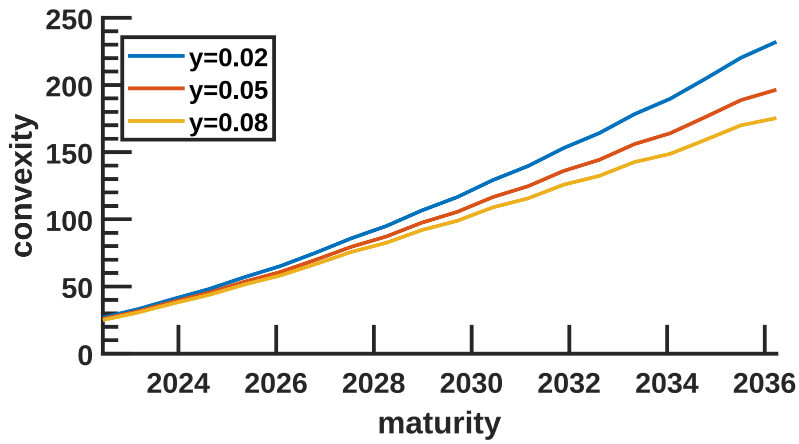
Set the first derivative with respect to n_2 to zero, we have

$$\frac{\partial Var[\Delta V}{\partial n_2} = 2 n_2 (D_2 P_2)^2 Var[\Delta y_2] + 2 n_1 D_1 D_2 P_1 P_2 Cov(\Delta y_1, \Delta y_2) = 0,$$

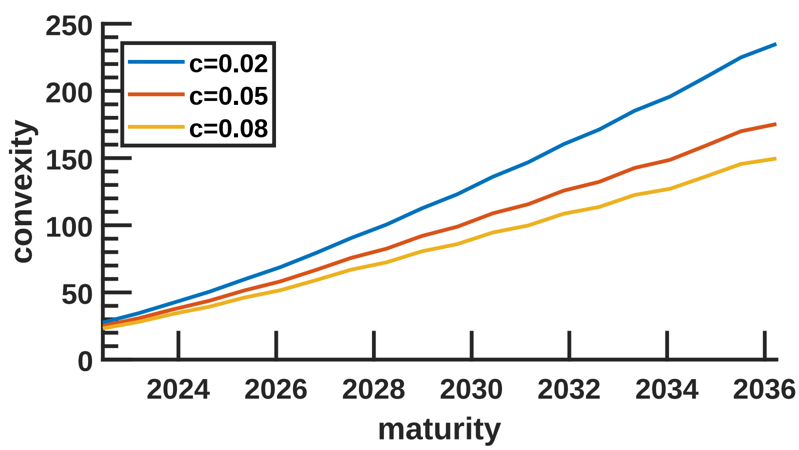
We have

$$n_2 = -n_1 \frac{D_1 P_1}{D_2 P_2} \frac{Cov(\Delta y_1, \Delta y_2)}{Var[\Delta y_2]}.$$

□



(a) Effects of yield and maturity on duration of a bond



(b) Effects of coupon rate and maturity on duration of a bond

Figure 4.4.5: Effects of coupon rate, yield, and maturity on the duration of a bond

4.4.3 Yield curve dynamics modeling

4.4.3.1 *PCA analysis*

Definition 4.4.4 (PCA analysis of curve daily change). *Let swap rate curve(or yield curve) be represented by a vector $y \in \mathbb{R}^N$, where N is the number of different tenors. Suppose we have $M+1$ historical daily observations of curve, denoted by $y_0, y_1, y_2, ..., y_M$. The historical daily change therefore can be constructed via $\Delta y_i = y_i - y_{i-1}$.*

It follows that

- *the covariance matrix Σ of the daily change can be obtained via*

$$\Sigma = \frac{1}{M}\sum_{i=1}^{N}(\Delta y_i - \overline{\Delta y})(\Delta y_i - \overline{\Delta y})^T.$$

- *The PCA factor of the daily change can be obtained via eigendecomposition*

$$\Sigma = V\Lambda V^T.$$

Example 4.4.4. [2, p. 193]

- We analyze **daily spot swap rates** with maturities of 1 year, 2 years, 3 years, 4 years, 5 years, 7 years, 10 years, and 30 years observing between 2000 and 2011. Table 4.4.1 shows the PCA factors and the associated eigenvectors. Figure 4.4.6 plots the first three dominating PCA factors.
- The first eigenvector/factor represents the parallel shift mode; the second eigenvector represents the steepening model; the third eigenvector represents the bending mode.

Note 4.4.2 (interpret eigenvalue as the variance of the factor score)**.** In the PCA framework, we can decompose the daily change of swap rate Δy(we assume there are N tenors, then $\Delta y \in \mathbb{R}^N$)as

$$\Delta y = \sum_{i=1}^{N} v_i s_i,$$

where $v_i \in \mathbb{R}^N$ and $s_i \in \mathbb{R}$ is known as the **factor score**.

Table 4.4.1: Eigenvectors and eigenvalues for swap rate daily change

(a) Eigenvectors for swap rate daily change

	PC1	PC2	PC3	⋯	PC8
1Y	0.216	-0.501	0.627	⋯	-0.034
2Y	0.331	-0.429	0.129	⋯	0.236
3Y	0.372	-0.267	-0.157	⋯	-0.564
4Y	0.392	-0.110	-0.256	⋯	0.512
5Y	0.404	0.019	-0.355	⋯	-0.327
7Y	0.394	0.194	-0.195	⋯	0.422
10Y	0.376	0.371	0.068	⋯	-0.279
30Y	0.305	0.554	0.575	⋯	0.032

(b) Eigenvalues for swap rate daily change

	PC1	**PC2**	**PC3**	⋯	**PC8**
eigenvalue	4.77	2.08	1.29	⋯	-0.034

The variance of the score i can be calculated via

$$\begin{aligned} Var[s_i] &= Var[\Delta y \cdot v_i] \\ &= v_i^T Cov[\Delta y] v_i \\ &= v_i^T (\sum_{i=1}^{N} \lambda_i v_i v_i^T) v_i \\ &= \lambda_i \end{aligned}$$

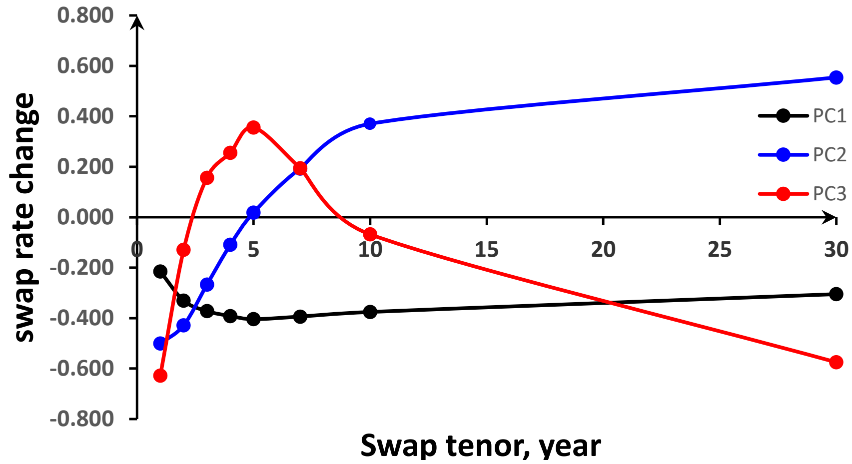


Figure 4.4.6: Demonstration of first three dominating PCA factor in the swap rate curve daily change.

4.5 Mortgage backed security

4.5.1 Mortgage basics

4.5.1.1 *Fixed rate mortgage mechanics*

Lemma 4.5.1 (fixed rate mortgage payment calculation). *[6, p. 584] Denote the original principal amount or loan balance as $B(0)$, f as payment frequency each year, y as the mortgage rate, MP as the periodic mortgage payment.*

- *The eqution to solve X for a T year mortgage is given by*

$$\begin{aligned} MB(0) &= X \cdot \sum_{n=1}^{f\cdot T} \frac{1}{(1+y/f)^{f\cdot n}} \\ &= X\frac{f}{y}(1 - \frac{1}{(1+y/f)^{f\cdot T}}). \end{aligned}$$

- *Let $MB(n)$ denote the principal amount outstanding after n mortgage payments. The interest component in the $n+1$ mortgage payment is*

$$I(n+1) = MB(n) \cdot \frac{y}{f}.$$

The principal component in the $n+1$ *mortgage payment is*

$$P(n+1) = X - MB(n) \cdot \frac{y}{f}.$$

And

$$B(n+1) = B(n) - P(n+1)$$

Example 4.5.1. [6, p. 584]

- A homeowner might borrow \$100,000 from a bank at 4% mortgate rate and agree to make payments of \$477.42 every month for 30 years. The mortgage rate and the monthly payment are related by the following equation:

$$477.42 \cdot \sum_{n=1}^{360} \frac{1}{(1+0.04/12)^n} = 100000.$$

- The interest component in the first payment is

$$100000 \cdot \frac{0.04}{12} = 333.33.$$

- The interest component in the first payment is

$$477.42 - 333.33 = 144.08.$$

- The outstanding loan balance after the first payment is

$$100000 - 144.08 = 99855.92.$$

Lemma 4.5.2 (analytics of mortgage payment).

- $$MP = MB(0)(\frac{z(1+z)^n}{(1+z)^n - 1})$$

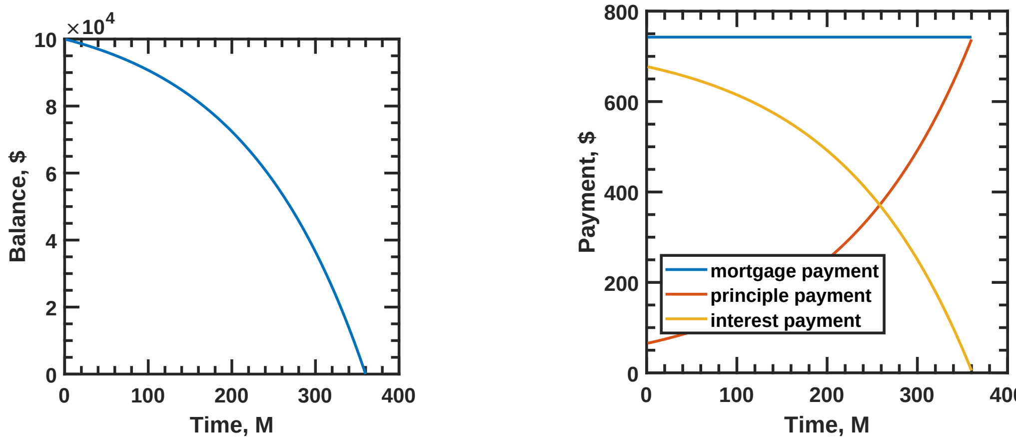
- $$MB(t) = MB(0)(\frac{(1+z)^n - (1+z)^t}{(1+z)^n - 1}).$$

- $$I(t) = MB(t) \cdot z = MB(0)(\frac{z(1+z)^n - (1+z)^{t-1}}{(1+z)^n - 1}), P(t) = MP - I(t) = MB(0)(\frac{z(1+z)^{t-1}}{(1+z)^n - 1}).$$

Proof. Note that

$$\begin{aligned} MB(t) &= \frac{MP}{1+z} + \frac{MP}{(1+z)^2} + \cdots + \frac{MP}{(1+z)^{n-t}} \\ &= MP(\frac{1}{1+z}\frac{(\frac{1}{1+z})^{n-t}-1}{\frac{1}{1+z}-1}) \\ &= MP\frac{1}{z}(\frac{1}{(1+z)^{n-t}}-1) \\ &= MB(0)(\frac{(1+z)^n-(1+z)^t}{(1+z)^n-1}) \end{aligned}$$

□



(a) The future yield curve in one simulation scenario.

(b) The future yield curve in another simulation scenario.

Figure 4.5.1: Bond dynamics.

4.5.1.2 *Prepayment modeling*

Definition 4.5.1.

- *The* ***single monthly mortality rate*** *at month n, denoted by SMM_n, is the percentage of $B(n)$ that is prepaid during month n.*
- *The constant prepayment rate CPR_n is the annualized SMM_n given by*

$$CPR_n = 1 - (1 - SMM_n)^{12}$$

$$CPR_n = 1 - (1 - SMM)$$

Lemma 4.5.3. *The SMM is often annualized to a constant prepayment rate or conditional prepayment rate (CPR).*

Definition 4.5.2 (public securities association (PSA) model). *The model assumes that the prepayment rate starts at 0.2% CPR in the first month and then rises 0.2% CPR per month until month 30; after that, the prepayment rate levels out at 6% CPR. We have*

$$PSA = \frac{CPR(t)}{\min(t, 30M) \times 0.2},$$

and

$$CPR(t) = \begin{cases} 0.2\% \times t, t \leq 30M \\ 6\%, t \geq 31M \end{cases}.$$

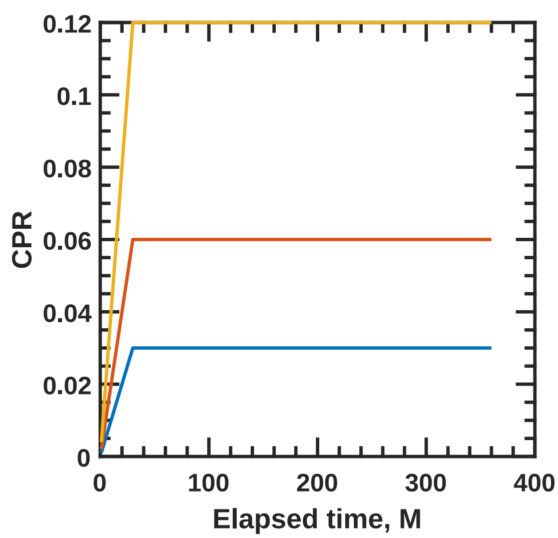


Figure 4.5.2: Conversion between PSA to CRP.

Definition 4.5.3 (Basic relationship among prepayment model). *[7, p. 334] Denote*

- *z monthly gross mortgage rate; $z = c + s$, where c is the monthly coupon rate going to the investors after after the monthly service fee charged.*

- $\overline{MB}(t)$ *as the loan outstanding at month* t *after considering scheduled principle payment and prepayment from first month to month* $t-1$; $MB(t)$ *as the loan outstanding at month* t *after considering scheduled principle payment and excluding any prepayments.*
- $SMM(t)$ *the single monthly mortality rate at month* t;
- $P(t)$ *monthly scheduled principle payment for month* t, ***excluding any prepayments***.
- $\overline{P}(t)$ *projected monthly principle payment for month* t, *including both scheduled principle payment and prepayment;* $\overline{P}(t)$ *projected monthly principle payment for month* t, *including both scheduled principle payment and prepayment.*
- $\overline{SP}(t)$ *projected scheduled monthly principle payment for month* t, *resulting from projected.*
- $\overline{MP}(t)$, *projected monthly mortgage payment for month* t *after considering the scheduled principle payment and prepayment up to month* $t-1$ *and resulting refinancing.*
- $\overline{CF}(t)$ *projected cash flow for month* t.

Then

- $$\overline{MP}(t+1) = \overline{MB}(t)\frac{1}{\frac{1}{1+z} + \frac{1}{(1+z)^2} + \cdots \frac{1}{(1+z)^{N-t}}}.$$
- $$\overline{I}(t+1) = \overline{MB}(t)\cdot z, \overline{SP}(t+1) = \overline{MP}(t+1) - \overline{I}(t+1).$$
- $$\overline{PR}(t+1) \triangleq (\overline{MB}(t) - \overline{SP}(t+1))\cdot SMM(t+1).$$
- $$\overline{P}(t+1) = \overline{SP}(t+1) + \overline{PR}(t+1).$$
- $$\overline{MB}(t+1) = \overline{MB}(t) - \overline{P}(t+1).$$
- $$\overline{CF}(t) = \overline{MP}(t) + \overline{PR}(t) - \overline{S}(t)$$
- *The total undiscounted interest expenses/incomes for the borrower/lender is*
$$\sum_{i=1}^{N}\overline{I}(t).$$

Lemma 4.5.4 (analytics in prepayment model). *[7, p. 334] Denote*

- $\overline{I}(t)$ *projected monthly interest payment after considering the scheduled principle payment and prepayment up to month* $t-1$ *and resulting refinancing;* $I(t)$ *scheduled interest payment for month* t.
- $b(t) = \prod_{n=1}^{t}(1 - SMM(n))$

It follows that

- $$\overline{MP}(t) = \overline{MB}(t-1)\left(\frac{z(1+z)^{N-t+1}}{(1+z)^{N-t+1}-1}\right)$$

- $$\begin{aligned}
\overline{MB}(t) &= MB(t)\cdot b(t), \\
\overline{MP}(t) &= MP\cdot b(t-1), \\
\overline{I}(t) &= I(t)\cdot b(t-1), \\
\overline{SP}(t) &= P(t)\cdot b(t-1).
\end{aligned}$$

Proof. (1) Using definition and summation

$$\frac{\overline{MP}(t+1)}{\overline{MB}(t)} = \frac{MP(t+1)}{MB(t)} = \frac{1}{\frac{1}{1+z}+\frac{1}{(1+z)^2}+\cdots\frac{1}{(1+z)^{N-t}}}.$$

(2)

$$\begin{aligned}
\overline{MB}(t+1) &= \overline{MB}(t) - \overline{P}(t+1) \\
&= \overline{MB}(t) - (\overline{SP}(t+1) + (\overline{MB}(t) - \overline{SP}(t+1))\cdot SMM(t+1)) \\
&= (\overline{MB}(t) + \overline{SP}(t+1))(1 - SMM(t+1))
\end{aligned}$$

For $t=0$, we have

$$\begin{aligned}
\overline{MB}(1) &= (\overline{MB}(0) + \overline{SP}(1))(1 - SMM(1)) \\
&= (MB(0) + P(1))(1 - SMM(1)) \\
&= MB(1)\cdot b(1)
\end{aligned}$$

where we use the definition $\overline{MB}(0) = MB(0), \overline{SP}(1) = P(1)$. Therefore,

$$\overline{MP}(2) = MP\cdot b(1), \overline{I}(2) = I(2)\cdot b(1), \overline{SP}(2) = P(2)\cdot b(1).$$

For $t=1$, we have

$$\begin{aligned}
\overline{MB}(2) &= (\overline{MB}(1) + \overline{SP}(2))(1 - SMM(2)) \\
&= (MB(0) + P(1))(1 - SMM(1))(1 - SMM(2)) \\
&= MB(2)b(2)
\end{aligned}$$

Therefore,

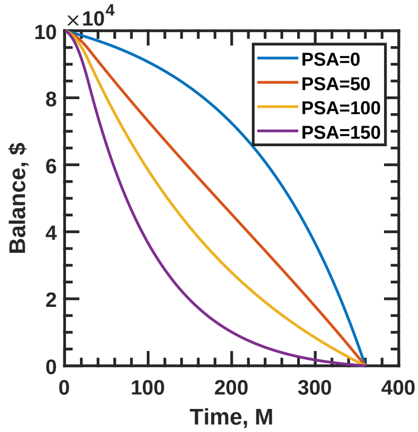
$$\overline{MP}(3) = MP\cdot b(2), \overline{I}(3) = I(3)\cdot b(2), \overline{SP}(3) = P(3)\cdot b(2).$$

Continue the process and we will get the result.

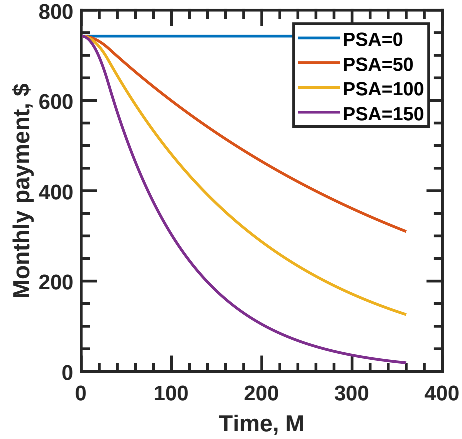
□

Remark 4.5.1 (understand prepayment risk)**.**

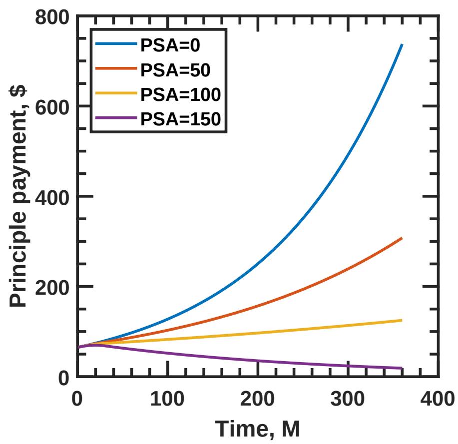
- From Figure Figure 4.5.3, we can see that as prepayment speed increase, the loan balance will decrease faster than the scheduled.
- The higher the prepayment rate, the lower the total interest expenses/incomes.
- The prepayment model predicts that the prepayment amount will first increase for the first 30M and then decrease. The increase is due to the increasing CRP for the first 30M; the decreasing is because the remaining loan balance is decreasing.



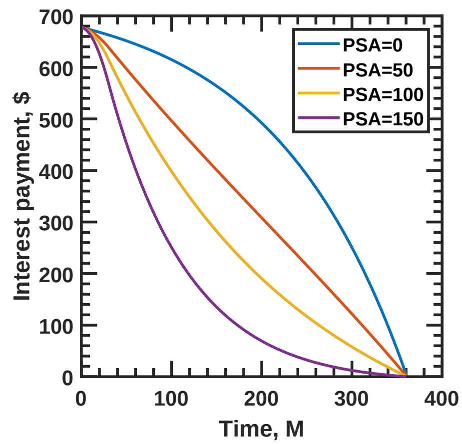
(a) Expected loan balance dynamics after considering prepayment at different speeds.



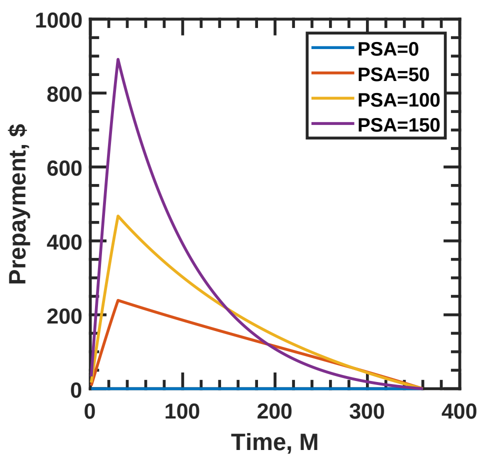
(b) Scheduled monthly mortgage payments (excluding prepayment at this month) at different prepayment speeds.



(c) Expected monthly principle payments after considering prepayment at different speeds.



(d) Expected monthly interest payments after considering prepayment at different speeds.



(e) Expected monthly prepayments after considering prepayment at different speeds.

Figure 4.5.3: Balance and Cash flow with prepayment for a mortgage.

4.6 Notes on bibliography

Major references are[1][8][6]. For mortgage modeling, see [9].

For bond credit spreads, see [10].

BIBLIOGRAPHY

1. Pinto, J., Petitt, B. & Pirie, W. *Fixed Income Analysis 3rd Edition* ISBN: 9781118999509. https://books.google.com/books?id=rOgbBgAAQBAJ (Wiley, 2015).
2. Hull, J. *Risk management and financial institutions,+ Web Site* (John Wiley & Sons, 2012).
3. DeFusco, R. A., McLeavey, D. W., Pinto, J. E., Anson, M. J. & Runkle, D. E. *Quantitative investment analysis* (John Wiley & Sons, 2015).
4. Wu, L. *Interest rate modeling: Theory and practice* (CRC Press, 2009).
5. Hull, J. *Options, Futures and Other Derivatives 9 edition* (Pearson Education, 2009).
6. Tuckman, B. & Serrat, A. *Fixed income securities: tools for today's markets* (John Wiley & Sons, 2011).
7. Fabozzi, F. *Fixed income mathematics, 3rd Ed* (McGraw-hill, 1993).
8. Fabozzi, F. *Bond Markets, Analysis, and Strategies* ISBN: 9780133797138. https://books.google.com/books?id=a7WgBwAAQBAJ (Pearson Education, 2015).
9. Davidson, A. & Levin, A. *Mortgage Valuation Models: Embedded Options, Risk, and Uncertainty* (Oxford University Press, 2014).
10. O'Kane, D. & Sen, S. Credit spreads explained. *Journal of Credit Risk* **1,** 61–78 (2005).

5

EQUITY DERIVATIVES

5.1 Dynamic Models of assets

5.1.1 Geometric Brownian motion model

5.1.1.1 *Basics*

Definition 5.1.1 (geometric Brownian motion (GBM)model). *The GBM model for the stock price S_t is given as*

$$dS_t = \mu(t)S_t dt + \sigma(t)S_t dW_t$$

where W_t is Brownian motion.

Lemma 5.1.1 (solutions to geometric SDE, recap). *The solution to the SDE*

$$dX(t) = \mu(t)X(t)dt + \sigma(t)X(t)dB(t)$$

is given as

$$X(t) = X(0)\exp(\int_0^t [\mu(s) - \frac{1}{2}\sigma(s)^2]ds + \int_0^t \sigma(s)dB(s)).$$

Particularly, if $\mu(t)$ and $\sigma(t)$ is time independent, then

$$X(t) = X(0)\exp((\mu - \frac{1}{2}\sigma^2)t + \sigma B(t))$$

Proof. see **??**. □

Lemma 5.1.2 (probability distribution). *The GBM (with constant coefficients)of the stock price S_t gives that*

$$S(t)/S(0) \sim LN((\mu - \frac{1}{2}\sigma^2)t, \sigma^2)$$

where LN is the lognormal representation.

Proof. See **??**. Note that $(\mu - \frac{1}{2}\sigma^2)t + \sigma B(t) \sim N((\mu - \frac{1}{2}\sigma^2)t, \sigma^2)$. □

Lemma 5.1.3 (martingale properties of the solution).

- *Let $M_t = \exp(\sigma W_t - \frac{1}{2}\sigma^2 t)$, where W_t is the Brownian motion. Then M_t is a martingale with respect the Brownian motion filtration.*
- *The GBM (with constant coefficients)of the stock price S_t is a martingale when $\mu = 0$; If $\mu > 0$, then $E[S_t|\mathcal{F}_u] > S_u, t > u$ (i.e., S_t is a supermartingale); If $\mu < 0$, then $E[S_t|\mathcal{F}_u] < S_u, t > u$ (i.e., S_t is a submartingale);*

Proof. see **??**. □

Remark 5.1.1. In the real-world probability, stock price usually follow supermartingale; that is, holding stock will lead to increase(greater than risk-free rate) of fortune in the expectation sense. This is because

- The market is incomplete: (1) Not all the stocks have derivatives to hedging. (2) Even we have the derivatives, the risks cannot be completely hedged because of transaction cost, stochastic volatility, and incomplete information.
- stock holders are compensated for taking risks.

Definition 5.1.2 (multiple dimension geometric Brownian motion (GBM)model). *[1, p. 183]*

- *The N dimensional GBM model for the stock price $S_i, i = 1, ..., N$ is given as*

$$dS_i(t) = \mu_i(t)S_i(t)dt + \sigma_i(t)S_i(t)dW_i(t)$$

where $W_i, i = 1, ..., n$ is ***correlated Brownian motion*** *such that*

$$E[dW_i] = 0, E[dW_i dW_j] = \rho_{ij}dt.$$

The instantaneous covariance structure is given by

$$Cov[dS_i dS_j] = \sigma_i \sigma_j \rho_{ij} dt.$$

- *(multiple factor representation) A m factor geometric Brownian motion model for the stock price $S_i, i = 1, ..., N$ is given by*

$$dS_i(t)/S_i(t) = \mu_i(t)dt + \sum_{i=1}^{m} \sigma_{ij} dW_j(t)$$

where $W_i, i = 1, ..., m$ is ***uncorrelated Brownian motion*** *such that*

$$E[dW_i] = 0, E[dW_i dW_j] = \delta_{ij}dt.$$

The instantaneous covariance structure is given by

$$Cov[dS_i dS_j] = \sum_{k=1}^{m} \sigma_{im}\sigma_{jm} dt.$$

Remark 5.1.2. The multiple factor model can be used as lower dimensional approximation when actual the instantaneous covariance structure has lower rank.

5.1.1.2 *Model calibration*

Lemma 5.1.4 (parameter estimation for single asset constant coefficient dynamics). *[2, p. 36] In the geometric Brownian motion model of the stock price given by,*

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

the drift μ and volatility σ can be estimated from a sequence of recent prices $\{S_0, S_1, S_2, ..., S_n\}$ with time step Δt, given as

- *Compute returns*
$$x_i = \ln S_i - \ln S_{i-1}, i = 1, 2, ..., n.$$
- *Compute statistics*
$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, S_{xx} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$
- *(MLE)Estimate drift via*
$$\hat{\mu} = \frac{1}{\Delta t}\overline{x} + \frac{S_{xx}}{2\Delta t}.$$
- *(MLE)Estimate variance via*
$$\hat{\sigma}^2 = \frac{S_{xx}}{\Delta t}.$$

Proof. Note that for the SDE

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

we can transform it to the following equivalent form

$$\begin{aligned} X_t &= \ln S_t \\ dX_t &= (\mu - \frac{1}{2}\sigma^2)dt + \sigma dB_t. \end{aligned}$$

X_t will have the following discrete-time form

$$X(t + \Delta t) - X(t) = (\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}Z, Z \in N(0, 1).$$

The rest follows results in maximum likelihood estimation in linear regression. □

Lemma 5.1.5 (parameter estimation for multiple asset constant coefficient dynamics). *[2, p. 36] In the correlated geometric Brownian motion model of the N stock price given by,*

$$dS^i(t) = \mu_i(t)S^i(t)dt + \sigma_i(t)S^i(t)dW_i(t)$$

where $W_i, i = 1, ..., d$ is correlated Brownian motion such that

$$E[dW_i] = 0, E[dW_i dW_j] = \rho_{ij} dt,$$

the drifts $\mu_i, i = 1, ..., N$ and volatilities $\sigma_i, i = 1, ..., N$ can be estimated from a sequence of recent prices $\{S_0^i, S_1^i, S_2^i, ..., S_n^i\}, i = 1, 2, ..., d$ with time step Δt, given as [3]

- *Compute returns*

$$x_i^j = \ln S_i - \ln S_{i-1}, i = 1, 2, ..., n; j = 1, 2, ..., d;$$

- *Compute statistics*

$$\overline{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_i^j,$$
$$[S_{xx}]_{ij} = \frac{1}{n} \sum_{k=1}^{n} (x_k^i - \overline{x}_i)(x_k^j - \overline{x}_j).$$

- *Estimate drift via*

$$\hat{\mu}_i = \frac{1}{\Delta t} \sum_{i=1}^{n} x_i + \frac{[S_{xx}]_{ii}}{2\Delta t}.$$

- *Estimate variance via*

$$\hat{\sigma}_i^2 = \frac{[S_{xx}]_{ii}}{\Delta t}$$

- *Estimate correlation via*

$$\hat{\rho}_{ij} = \frac{[S_{xx}]_{ij}}{\Delta t \hat{\sigma}_i \hat{\sigma}_j}.$$

5.2 Black-Scholes model: application examples

5.2.1 European call and put

5.2.1.1 *The pricing*

Lemma 5.2.1 (European call/put pricing). *[4, p. 219] Let current time t. Denote the spot by S_t and forward by $F(t,T)$. In the Black-Scholes model, an European call option with strike price K and expiry T is given as:*

$$C(S_t,t) = N(d_1)S_t - N(d_2)Ke^{-r(T-t)}$$

where

$$\begin{aligned}
d_1 &= \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r+\sigma^2/2)(T-t)] = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{F(t,T)}{K}) + \sigma^2/2(T-t)] \\
d_2 &= \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r-\sigma^2/2)(T-t)] = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{F(t,T)}{K}) - \sigma^2/2(T-t)] \\
d_2 &= d_1 - \sigma\sqrt{T-t} \\
N(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \\
F(t,T) &= S_t \exp(r(T-t))
\end{aligned}$$

We further have

- *The price of a zero strike is S_t.*
- *The price of the put can be derived based on put-call parity $P_t + S_t = C_t + Ke^{-r(T-t)}$ (**??**) as*

$$P(S_t,t) = N(-d_2)Ke^{-r(T-t)} - N(-d_1)S_t.$$

Proof. See Theorem 2.3.1. □

Remark 5.2.1 (interpretation). The call pricing formula is given as

$$C(S_t,t) = N(d_1)S_t - N(d_2)Ke^{-r(T-t)} = (N(d_1) - N(d_2))S_t + N(d_2)(S_t - Ke^{-r(T-t)}),$$

where $N(d_2)$ is the probability of stock price S_T greater than K(i.e. the probability the call will be exercised.). $N(d_1) - N(d_2)$ is the probability of increasing value from current stock.

Remark 5.2.2 (interpretation).

- The option price will increase as the volatility(as we showed in the Greeks).
- The option price depends only on model parameters σ and r of the underlying assets:risk-free bond and stock. One can use the market price of the option is calibrate σ and r.
- One can improve the model by allowing: (1) σ to follow stochastic volatility model; (2) r to follow stochastic short rate model.

Remark 5.2.3 (interpretation from original publication). [5]

- In general, it seems clear that the higher the price of the stock, the greater the value of the option. When the stock price is much greater than the exercise price, the option is almost sure to be exercised. The current value of the option will thus be approximately equal to the price of the stock minus the price of a pure discount bond that matures on the same date as the option, with a face value equal to the striking price of the option.
- On the other hand, if the price of the stock is much less than the exercise price, the option is almost sure to expire without being exercised, so its value will be near zero.
- If the expiration date of the option is very far in the future, then the price of a bond that pays the exercise price on the maturity date will be very low, and the value of the option will be approximately equal to the price of the stock.
- On the other hand, if the expiration date is very near, the value of the option will be approximately equal to the stock price minus the exercise price, or zero, if the stock price is less than the exercise price.

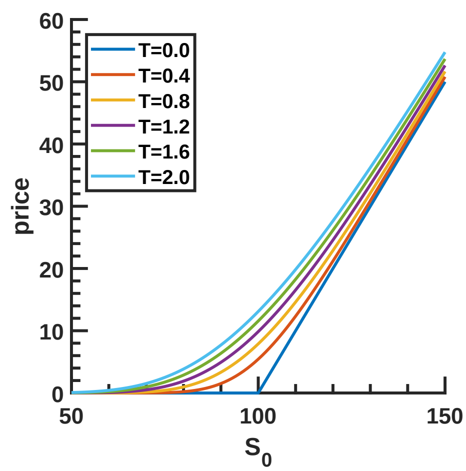


Figure 5.2.1: Price profile of a call option at different spot prices S_0 and different time to maturity T. The strike of the option is $K = 100$.

Lemma 5.2.2 (European call and put with dividends). *[4, p. 236][6, p. 373] Let current time t. Denote the spot by S_t and forward by $F(t, T)$. In the Black-Scholes model, an European call option with strike price K for an asset paying continuous dividends a is given as:*

$$C(S_t, t) = N(d_1)S_t e^{-a(T-t)} - N(d_2)Ke^{-r(T-t)}$$

where

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r - a + \sigma^2/2)(T-t)] = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{F(t,T)}{K}) + \sigma^2/2(T-t)]$$
$$d_2 = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r - a - \sigma^2/2)(T-t)] = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{F(t,T)}{K}) - \sigma^2/2(T-t)]$$
$$d_2 = d_1 - \sigma\sqrt{T-t}$$
$$N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$
$$F(t,T) = S_t \exp((r-a)(T-t))$$

The price of the put can be derived based on put-call parity

$$P_t + S_t e^{-a(T-t)} = C_t + Ke^{-r(T-t)}$$

*(**??**) as*

$$P(S_t,t) = N(-d_2)Ke^{-r(T-t)} - N(-d_1)S_t e^{-a(T-t)}.$$

Proof. See Theorem 2.3.2. □

Example 5.2.1. [6, p. 374] Consider an European call option on S&P500 that is two months from maturity. The current value of the index is 930, the exercise price is 900, the risk-free rate is 8% per annum, and the volatility of the index is 20% per annum. Dividend yields is 3% per annum. Then we have $S_0 = 930, K = 900, r = 0.08, \sigma = 0.2$, and $T = 2/12$. And

$$d_1 = \frac{1}{0.2\sqrt{2/12}}[\ln(\frac{930}{900}) + (0.08 - 0.03 + 0.02^2/2)(2/12) = 0.5444;$$

$$d_2 = \frac{1}{0.2\sqrt{2/12}}[\ln(\frac{930}{900}) + (0.08 - 0.03 - 0.02^2/2)(2/12) = 0.4628.$$

$$N(d_1) = 0.7069, N(d_2) = 0.6782;$$

and the call price is

$$c = 930 \times N(d_1)\exp(-q \times 2/12) - 900 \times N(d_2)\exp(-r \times 2/12) = 51.83.$$

Lemma 5.2.3 (implied dividend yield). *[6, p. 376] If we want to get the market perspective on the dividend yield, we can use*

$$q = -\frac{1}{T}\ln\frac{C - P + K\exp(-rT)}{S_0},$$

where S_0 is the spot price, and C/P are the call/put prices with the same strike K and maturity T.

Proof. Use the put-call parity

$$C - P = S_0 \exp(-qT) - K\exp(-rT).$$

□

5.2.1.2 *Other relations*

Lemma 5.2.4.
- *Let X_T be a martingale and $Pr(X_T < 0) = 0$. Then for all positive K, T such that*

$$(X_0 - K)^+ < E[(X_T - K)^+] < X_0.$$

- *When the expectation is taken under the risk-neutral measure, we have*

$$(F(0,T) - K)^+ < E_Q[(F(T,T) - K)^+] < F(0,T).$$

$$(S_0 \exp((r-a)T) - K)^+ < E_Q[(S_T - K)^+] < S_0 \exp((r-a)T).$$

- *The call option price bound*

$$(S_0 - K))^+ < (S_0 - K\exp(-(r-a)T))^+ < \exp(-(r-a)T)E_Q[(S_T - K)^+] = C < S_0.$$

Proof. (1) The function $f(x) = (x-K)^+$ is convex. Jansen inequality gives that

$$E[(X_T - K)^+] > [(E[X_T] - K)^+] = (X_0 - K)^+.$$

And

$$E[(X_T - K)^+] < E[(X_T - K = K)^+] = E[(X_T)^+] = E[X_T] = X_0.$$

(2) (a) Use the fact that $F(t,T)$ is a martingale under risk-neutral measure. (b) use $F(t,T) = S_t \exp((r-a)(T-t))$. (b) The call option price is

$$C = \exp(-(r-a)T)E_Q[(S_T - K)^+],$$

therefore is smaller than spot S_0 and greater than intrinsic $(S_0 - K)^+$. □

Lemma 5.2.5. *Consider the black formula for call and put pricing given by*

$$C(S_t, K, \sigma, t) = N(d_1)S_t e^{-q\tau} - N(d_2)Ke^{-r\tau}, C(F(t,T), K, \sigma, \tau) = N(d_1)F(t,T)e^{-r\tau} - N(d_2)Ke^{-r\tau},$$

$$P(S_t, K, \sigma, t) = -N(-d_1)S_t e^{-q\tau} + N(-d_2)Ke^{-r\tau}, P(F(t,T), K, \sigma, \tau) = -N(-d_1)F(t,T)e^{-r\tau} + N(-d_2)Ke^{-}$$

where

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r - a + \sigma^2/2)(T-t)] = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{F(t,T)}{K}) + \sigma^2/2(T-t)]$$

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

$$N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

$$F(t,T) = S_t \exp((r-a)(T-t))$$

It follows that

- $$P(S_t, K, \sigma, \tau) = C(-S_t, -K, -\sigma, \tau)$$
- $$C(F(t,T), K, \sigma, \tau) = P(K, F(t,T), \sigma, T)$$

Proof. (1) Note that when we replace S_t, K, σ by $-S_t, -K, -\sigma$, d_1 will change sign. (2) Note that when we replace the order of $F(t,T)$ and K, d_1 will change to $-d_2$. □

5.2.1.3 *Option value after initialization*

Lemma 5.2.6. *Suppose the underlying S_t has real-world dynamics given by*

$$dS_t/S_t = \mu dt + \sigma dW_t.$$

Consider a call option with strike K and expiry T.

- *the value $C(t)$ of a call option in the real world will be*

$$C(t) = N(d_1)S_t - N(d_2)Ke^{-r(T-t)}.$$

- *the value $C(t)$ of a call option will have real-world dynamics given by*

$$dC_t = (rC_t + \frac{\partial C_t}{\partial S_t}\lambda\sigma S_t)dt + \frac{\partial C_t}{\partial S_t}\sigma dW_t$$

  where λ is the market price of the stock risk given by

$$\lambda = \frac{\mu - r}{\sigma}.$$

Proof. (1) Straight forward from Lemma 5.2.1.(2) Because $C(t) = C(t, S(t))$, Ito lemma (**??**)gives

$$\begin{aligned}
dC &= \frac{\partial C}{\partial t}dt + \frac{\partial C}{\partial S_t}dS_t + \frac{1}{2}\frac{\partial^2 C}{\partial S_t^2}dS_t dS_t \\
&= \frac{\partial C}{\partial t}dt + \frac{\partial C}{\partial S_t}(\mu S_t dt + \sigma S_t dW_t) + \frac{1}{2}\frac{\partial^2 C}{\partial S_t^2}\sigma^2 S_t^2 dt \\
&= (\frac{\partial C}{\partial t} + \frac{\partial C}{\partial S_t}rS_t + \frac{1}{2}\frac{\partial^2 C}{\partial S_t^2}\sigma^2 S_t^2)dt + \frac{\partial C}{\partial S_t}(\mu - r)S_t dt + \frac{\partial C}{\partial S_t}\sigma S_t dW_t \\
&= (rC_t + \frac{\partial C}{\partial S_t}(\mu - r)S_t)dt + \frac{\partial C}{\partial S_t}\sigma S_t dW_t \\
&= (rC_t + \frac{\partial C_t}{\partial S_t}\lambda\sigma S_t)dt + \frac{\partial C_t}{\partial S_t}\sigma dW_t
\end{aligned}$$

where we used the the Black-Scholes equation

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 C}{\partial S_t^2} + rS_t\frac{\partial C}{\partial S_t} - rC = 0.$$

□

Note 5.2.1 (simulating European option value). To generate the trajectories of the forward contract value evolution, we can simulate the underlying, S_t, and then use the relation between S_t and $V(t)$ in Theorem 1.7.3.

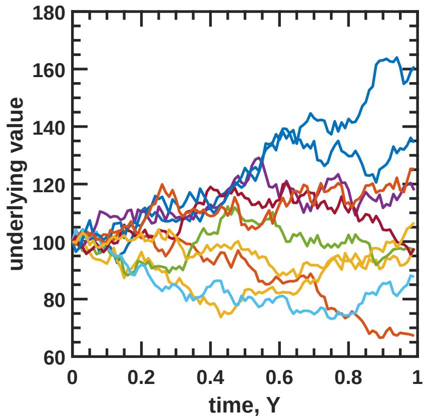
- As showed in Figure 5.2.2, we simulate the underlying in the real world via model

  $$dS_t/S_t = \mu dt + \sigma dW_t,$$

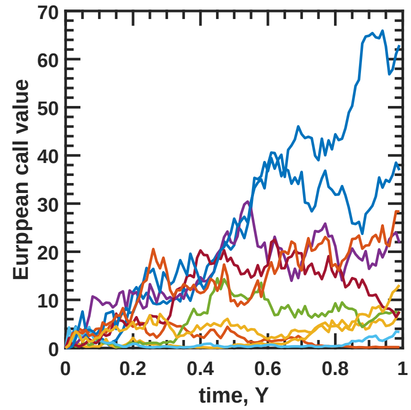
  and then evaluate the European call contract via

  $$V(t) = N(d_1)S_t e^{-q(T-t)} - N(d_2)Ke^{-r(T-t)}.$$

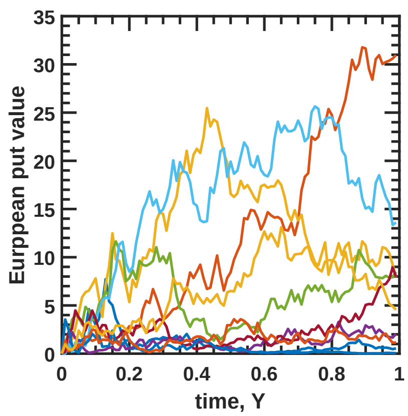
- Note that **we should not simulate the underlying in the risk-neutral measure.** This is because in the risk neutral measure $V(t)$ will have expected growth rate of r. If an option contract has risk-free rate expected return with risks, then no one wants to enter due to risk aversion.
- On the other hand, if we simulate the underlying in the real world measure where S_t will have a different expected growth rate from r. Then the market participants will enter either short or long positions based on their estimation of the S_t dynamics.
- Also see [7, p. 145] for a discussion.



(a) Underlying price trajectories.



(b) European call value trajectories.



(c) European put value trajectories.

Figure 5.2.2: Demonstration of European option value evolution. Underlying simulation parameters: $\mu = 0.06, r = 0.02, \sigma = 0.2, q = 0.0, t_0 = 0, S(t_0) = 100, K = 100, T = 1Y.$

5.2.1.4 *Profit and loss bounds*

Table 5.2.1: Maxium gain and loss for call and put options

(a) Call option profit and loss bounds

	maximum loss	maximum gain
buyer(long)	premium	unlimited
seller(short)	unlimited	premium

(b) Put option profit and loss bounds

	maximum loss	maximum gain
buyer(long)	premium	strike-premium
seller(short)	strike-premium	premium

5.2.2 Greeks for European call and put

5.2.2.1 *Basics*

Definition 5.2.1 (Greeks).

- ***Delta***:
$$\Delta = \frac{\partial V}{\partial S}$$
- ***Gamma***:
$$\Gamma = \frac{\partial^2 V}{\partial S^2}$$
- ***Theta***:
$$\Theta = \frac{\partial V}{\partial t}$$
- ***Vega***:
$$\nu = \frac{\partial V}{\partial \sigma}$$
- ***Rho***:
$$\rho = \frac{\partial V}{\partial r}$$

- ***Volga***:
$$Volga = \frac{\partial^2 V}{\partial \sigma^2}$$

Lemma 5.2.7 (preparation results). *[8] Define*

$$d_1 = \frac{1}{\sigma\sqrt{\tau}}[\ln(\frac{S_t}{K}) + (r+\sigma^2/2)\tau], d_2 = d_1 - \sigma\sqrt{\tau}, \tau = T - t.$$

We have

- $$S_t \exp(-q\tau) n(d_1) = K \exp(-r\tau) n(d_2),$$
$$F(t,T) n(d_1) = K n(d_2).$$
where forward rate $F(t,T) = S(t,T)\exp(-(r-q)(T-t)$.
- $$\frac{\partial d_2}{\partial S_t} = \frac{\partial d_1}{\partial S_t}$$
$$\frac{\partial d_2}{\partial K} = \frac{\partial d_1}{\partial K}$$
$$\frac{\partial d_2}{\partial \tau} = \frac{\partial d_1}{\partial \tau} - \frac{1}{2}\frac{\sigma}{\sqrt{\tau}}$$
$$\frac{\partial d_2}{\partial \sigma} = \frac{\partial d_1}{\partial \sigma} - \sqrt{\tau}$$
$$\frac{\partial d_2}{\partial r} = \frac{\partial d_1}{\partial r}$$

Proof. Note that

$$\begin{aligned} &\frac{\partial}{\partial S_t} N(d_1) \\ &= N'(d_1)\frac{\partial}{\partial S_t} d_1 \\ &= \frac{1}{\sqrt{2\pi}}\exp(-d_1^2/2)\frac{1}{S_t\sigma\sqrt{\tau}} \\ &= \frac{1}{\sqrt{2\pi\tau}S_t\sigma}\exp(-d_1^2/2) \end{aligned}$$

and

$$
\begin{aligned}
&\frac{\partial}{\partial S_t} N(d_2) \\
&= N'(d_2)\frac{\partial}{\partial S_t} d_2 \\
&= \frac{1}{\sqrt{2\pi}} \exp(-d_2^2/2) \frac{1}{S_t \sigma \sqrt{\tau}} \\
&= \frac{1}{\sqrt{2\pi\tau} S_t \sigma} \exp(-(d_1 - \sigma\sqrt{\tau})^2/2) \\
&= \frac{1}{\sqrt{2\pi\tau} S_t \sigma} \exp(-d_1^2/2) \exp(-\sigma^2\tau/2) \exp(\sigma\sqrt{\tau} d_1) \\
&= \frac{1}{\sqrt{2\pi\tau} S_t \sigma} \exp(-d_1^2/2) \frac{S_t}{K} \exp(r\tau)
\end{aligned}
$$

□

5.2.2.2 *Delta*

Lemma 5.2.8 (Delta). *Let current time be t. Consider a call/put option with strike K and expiry T. Let the short rate be r.*

- *The Delta for a call option in the Black-Scholes pricing is given as*

$$\Delta_C = \frac{\partial C_t}{\partial S_t} = N(d_1), d_1 = \frac{1}{\sigma\sqrt{\tau}}[\ln(\frac{S_t}{K}) + (r + \sigma^2/2)\tau], \tau = T - t.$$

  And the Delta for a put option is given as

$$\Delta_P = \frac{\partial P_t}{\partial S_t} = \frac{\partial C_t}{\partial S_t} - 1 = N(d_1) - 1 = -N(-d_1).$$

- *If the underlying asset is paying continuous dividends with rate q, then*

$$\Delta_C = \exp(-q(T-t))N(d_1), \Delta_P = \exp(-q(T-t))(1 - N(d_1)) = -\exp(-q(T-t))N(-d_1),$$

  where

$$d_1 = \frac{1}{\sigma\sqrt{\tau}}[\ln(\frac{S_t}{K}) + (r - q + \sigma^2/2)\tau], \tau = T - t.$$

- *The Delta with respect to forward rate $F(t, T) = S(t, T)\exp(-(r-q)(T-t)$ is given by*

$$\Delta_{C,F} = \frac{\partial C_t}{\partial F(t, T)} = \exp(-r\tau)N(d_1), \Delta_{P,F} = \frac{\partial P_t}{\partial F(t, T)} = -\exp(-r\tau)N(-d_1).$$

- *The Delta with respect to strike K is given by*

$$\Delta_{C,K} = \frac{\partial C_t}{\partial K} = \exp(-r\tau)N(d_1), \Delta_{P,K} = \frac{\partial P_t}{\partial F(t,T)} = -\exp(-r\tau)N(-d_1).$$

Proof. (1)(a) direct valuation. Note that

$$\frac{\partial C_t}{\partial S_t} = N(d_1) + S_t\frac{\partial}{\partial S_t}N(d_1) - K\exp(-r\tau)\frac{\partial}{\partial S_t}N(d_2).$$

Note that

$$\begin{aligned}
&\frac{\partial}{\partial S_t}N(d_1) \\
&= N'(d_1)\frac{\partial}{\partial S_t}d_1 \\
&= \frac{1}{\sqrt{2\pi}}\exp(-d_1^2/2)\frac{1}{S_t\sigma\sqrt{\tau}} \\
&= \frac{1}{\sqrt{2\pi\tau}S_t\sigma}\exp(-d_1^2/2)
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\partial}{\partial S_t}N(d_2) \\
&= N'(d_2)\frac{\partial}{\partial S_t}d_2 \\
&= \frac{1}{\sqrt{2\pi}}\exp(-d_2^2/2)\frac{1}{S_t\sigma\sqrt{\tau}} \\
&= \frac{1}{\sqrt{2\pi\tau}S_t\sigma}\exp(-(d_1-\sigma\sqrt{\tau})^2/2) \\
&= \frac{1}{\sqrt{2\pi\tau}S_t\sigma}\exp(-d_1^2/2)\exp(-\sigma^2\tau/2)\exp(\sigma\sqrt{\tau}d_1) \\
&= \frac{1}{\sqrt{2\pi\tau}S_t\sigma}\exp(-d_1^2/2)\frac{S_t}{K}\exp(r\tau)
\end{aligned}$$

Therefore,

$$S_t\frac{\partial}{\partial S_t}N(d_1) - K\exp(-r\tau)\frac{\partial}{\partial S_t}N(d_2) = 0, \frac{\partial C_t}{\partial S_t} = N(d_1).$$

(1)(b) For the put: From put-call parity(**??**), we have $P_t + S_t = C_t + Ke^{-r(T-t)}$. Take first differential on both sides, we get

$$\frac{\partial P_t}{\partial S_t} + 1 = \frac{\partial C_t}{\partial S_t}.$$

(2) Similar to (1). (3) Use the chain rule that

$$\frac{\partial C_t}{\partial F(t,T)} = \frac{\partial C_t}{\partial F(t,T)} \frac{\partial F(t,T)}{\partial S_t}.$$

(4)(a)

$$\frac{\partial C}{\partial K} = S_t \exp(-q\tau) n(d_1) \frac{\partial d_1}{\partial K} - \exp(-r\tau) N(d_2) - K \exp(-r\tau) n(d_2) \frac{\partial d_2}{\partial K} = -\exp(-r\tau) N(d_2.$$

where we use

$$S_t \exp(-q\tau) n(d_1) \frac{\partial d_1}{\partial K} = K \exp(-r\tau) n(d_2) \frac{\partial d_2}{\partial K}$$

in Lemma 5.2.7. (b) Use put-call parity such that

$$\frac{\partial C}{\partial K} - \frac{\partial P}{\partial K} = -\exp(-r\tau).$$

□

Remark 5.2.4 (different ATM convention)**.** There are different interpretations when we refer to ATM strike.

- ATM-spot: $K = S_0$.
- ATM-forward: $K = F_0$.
- ATM-value-neutral: K such that call value = put value.
- ATM-Delta-neutral: K such that call delta = -put delta. In this case,

$$\Delta_C = -\Delta_P = 0.5 \implies d_1 = 0, K = S_0 \exp((r + \frac{\sigma^2}{2})T).$$

Remark 5.2.5 (interpretation)**.**

- The Delta for a call is always positive, which mean if S_t increases, C_t will increase.
- The Delta for a put is always negative, which mean if S_t increases, P_t will decrease.

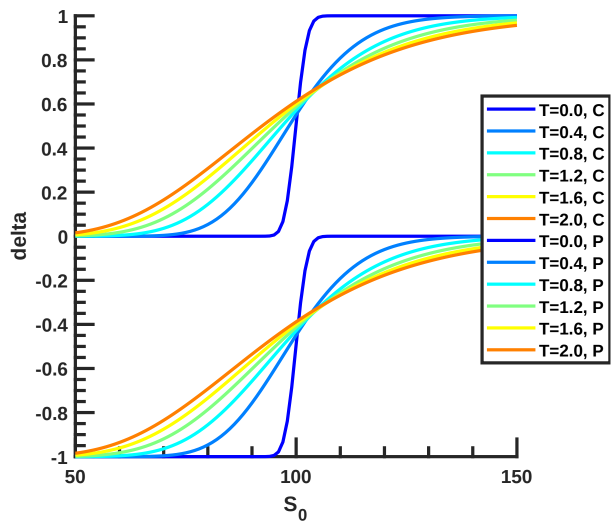


Figure 5.2.3: Delta profile of a call option(upper) and a put option(lower) at different spot prices S_0 and different time to maturity T. The strike of the option is $K = 100$.

5.2.2.3 *Gamma*

Lemma 5.2.9 (Gamma and convexity). *Let current time be t. Consider a call/put option with strike K and expiry T. Let the short rate be r and the dividend rate be q. In the Black-Scholes framework, the Gamma of the call/put option is given as*

$$\Gamma = \frac{\partial^2 C}{\partial S_t^2} = \frac{\partial^2 C}{\partial S_t^2} = \exp(-q(T-t))\frac{N'(d_1)}{S_t\sigma\sqrt{T-t}} = K\exp(-r(T-t))\frac{N'(d_2)}{S_t^2\sigma\sqrt{T-t}},$$

where

$$d_1 = \frac{1}{\sigma\sqrt{\tau}}[\ln(\frac{S_t}{K}) + (r - q + \sigma^2/2)\tau], \tau = T - t.$$

Moreover, we have

- *C is a convex function of the spot price S_t; that is, $\frac{\partial^2 C}{\partial S^2} \geq 0$.*
- *The Γ for a put option is same as that of the call option due to put-call parity..*

Proof. (1)Since

$$\frac{\partial C}{\partial S} = N(d_1), d_1 = \frac{1}{\sigma\sqrt{T-t}}[\ln(\frac{S_t}{K}) + (r + \sigma^2/2)(T-t)],$$

we have

$$\frac{\partial^2 C}{\partial S^2} = \frac{N'(d_1)}{S\sigma\sqrt{T-t}}.$$

It is easy to see that $\Gamma \geq 0$, therefore, C_t is convex. c (2)For the put: From put-call parity(**??**), we have $P_t + S_t = C_t + Ke^{-r(T-t)}$. Take twice differential on both sides, we get

$$\frac{\partial^2 P_t}{\partial S_t^2} = \frac{\partial^2 C_t}{\partial S_t^2}.$$

□

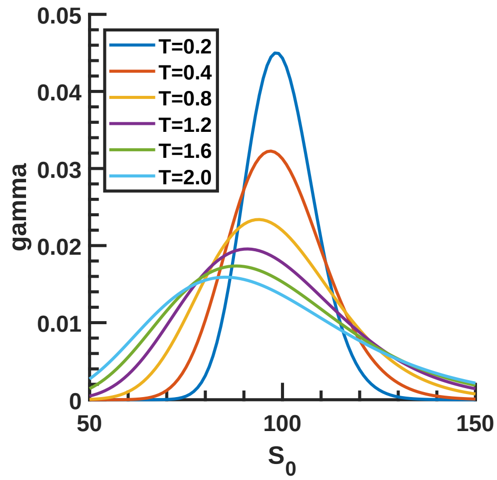


Figure 5.2.4: Gamma profile of a call option(or a put option) at different spot prices S_0 and different time to maturity T. The strike of the option is $K = 100$.

5.2.2.4 *Theta*

Lemma 5.2.10 (Theta). *[6, p. 409]Let current time be t. Consider a call/put option with strike K and expiry T. Let the short rate be r and the dividend rate be q. In the Black-Scholes framework, the Theta of the call/put is given by,*

$$\Theta_C = -\frac{S_t N'(d_1)\sigma}{2\sqrt{T-t}} - rKe^{-r(T-t)}N(d_2) + qS_t \exp(-q(T-t))N(d_1)$$
$$\Theta_P = -\frac{S_t N'(d_1)\sigma}{2\sqrt{T-t}} + rKe^{-r(T-t)}N(d_2) - qS_t \exp(-q(T-t))N(-d_1)$$

where

$$d_1 = \frac{1}{\sigma\sqrt{\tau}}[\ln(\frac{S_t}{K}) + (r - q + \sigma^2/2)\tau], \tau = T - t$$

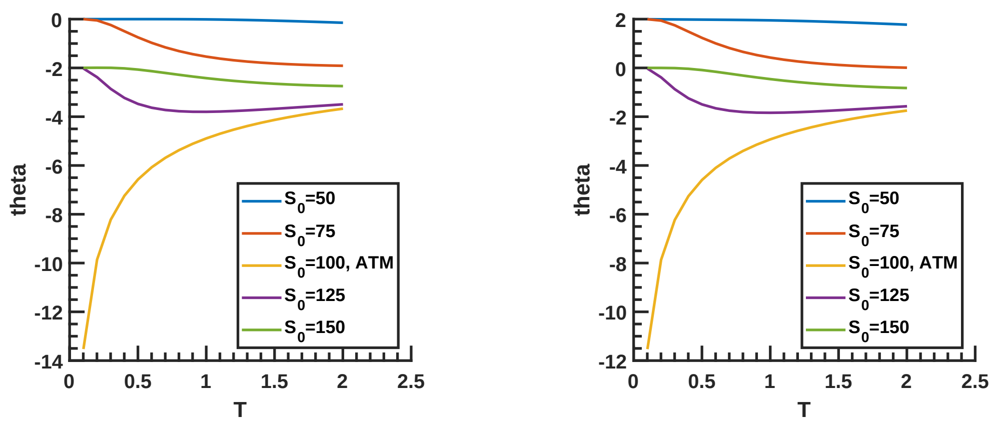
Note that

$$\Theta_C - \Theta_P = -\frac{d}{dt}(Ke^{-r(T-t)}) = -qS_t \exp(-q(T-t)) - rKe^{-r(T-t)}$$

due to put-call parity.

Remark 5.2.6 (example theta and interpretation). As showed in Figure 5.2.5, we examine the Theta for call and put options on a non-dividend-paying underlying. We have the following observation.

- ATM call and put have large theta values, which indicate the dramatic decreasing the of the time value for ATM options approaching expiry.
- Keeping other factors remain fixed, call options always decrease value when approaching expiry.
- For ATM and out-of-money put options, their value will decrease when approaching expiry; For deep-in-the-money put options, put options might increase value when approaching expiry.



(a) Underlying price trajectories. (b) European call value trajectories.

Figure 5.2.5: Theta profile of a call option(a) and a put option(b) at different spot prices S_0 and different time to maturity T. The strike of the option is $K = 100$. The dividend rate $q = 0$.

Remark 5.2.7 (interpretation). Given a Geometric Brownian motion with constant drift μ and volatility σ. Its mean and variance is given by(**??**)

- $E[S(t)] = S(0)e^{rt}$
- $Var[S(t)] = S(0)^2 e^{2rt}(e^{\sigma^2 t} - 1)$

- If $r \geq 0$, then an European call will always decrease in value; Intuitively, geometric Brownian motion has an upward drift , the larger volatility, the more likely it is going to arrive at the higher asset value.
- If $r \geq 0$, then an European put might not lose value. The result depends on the magnitude of r. Intuitively, geometric Brownian motion will spread the underlying and increase the value of put(due to $Gamma > 0$); a large positive drift will increase the likelihood of out-of- money situation.

Corollary 5.2.0.1 (never early exercise an American call in the Black-Scholes world). *Let $C(T;K)$ denote the price (the price calculated in Black-Scholes world)of a European call with the same strike K but a continuum of maturity. Since $\frac{dC}{dT} = -\Theta_C > 0$, calls with more distant exercising date will have more value; that is, never early exercising an American call.*

Note 5.2.2 (what the real world will be). In the real world, the asset dynamics need to have a state-dependent volatility $\sigma(S_t.t)$ in order to match the market prices (see Theorem 2.4.1).

$$\frac{\partial C}{\partial T} = \frac{\partial C_{BS}}{\partial \sigma^*}\frac{\partial \sigma^*}{\partial T}.$$

Since the term structure of implied volatility can be anything, the time-value of call option in real world is uncertain.

5.2.2.5 *Vega and Volga*

Lemma 5.2.11 (Vega). *Let current time be t. Consider a call/put option with strike K and expiry T. Let the short rate be r and the dividend rate be q. In the Black-Scholes framework, the Vega of the call/put is given by,*

$$\nu = \frac{\partial C_t}{\partial \sigma} = \frac{\partial P_t}{\partial \sigma} = S_t e^{-q(T-t)}\sqrt{T-t}N'(d_1) = Ke^{-r(T-t)}\sqrt{T-t}N'(d_2),$$

where

$$d_1 = \frac{1}{\sigma\sqrt{\tau}}[\ln(\frac{S_t}{K}) + (r-q+\sigma^2/2)\tau], \tau = T-t.$$

Remark 5.2.8 (interpretation).

- The Vega is always positive, indicating that the option prices will increase as volatility increases.
- The Vega of all options decrease as approaching expiration since a long-term option is more sensitive to change in the volatility.

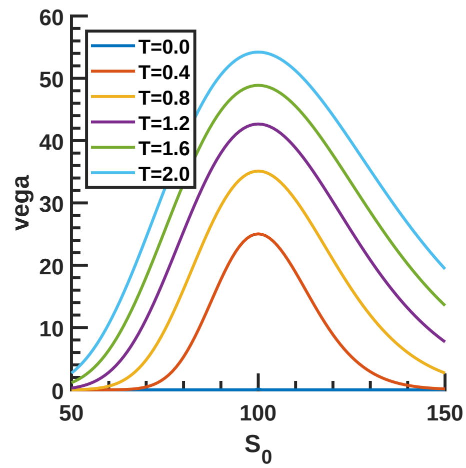


Figure 5.2.6: Vega profile of a call option(or a put option) at different spot prices S_0 and different time to maturity T. The strike of the option is $K = 100$.

Lemma 5.2.12 (Volga). *For European call and put options,*

$$Volga = \frac{\partial^2 C_t}{\partial \sigma^2} = \frac{\partial^2 P_t}{\partial \sigma^2} = \nu \frac{d_1 d_2}{\sigma}$$

Proof. direct calculation. □

Remark 5.2.9 (convexity on volatility)**.**

- For most out-of-money call options, both d_1 and d_2 are negative; for most in-the-money call options, both d_1 and d_2 are positive. So $d_1 d_2 > 0$ in most cases and C_t is a convex function on σ.
- When close at-the-money, we might have $d_1 d_2 < 0$.

5.2.2.6 *Rho*

Lemma 5.2.13 (Rho). *[6, p. 417] Let current time be t. Consider a call/put option with strike K and expiry T. Let the short rate be r and the dividend rate be q. In the Black-Scholes framework, the Rho of the call/put is given by,*

$$Rho_C = \frac{\partial V}{\partial r} = K(T-t)e^{-r(T-t)}N(d_2) > 0, Rho_P = \frac{\partial V}{\partial r} = -K(T-t)e^{-r(T-t)}N(-d_2) < 0.$$

where

$$d_1 = \frac{1}{\sigma\sqrt{\tau}}[\ln(\frac{S_t}{K}) + (r - q + \sigma^2/2)\tau], \tau = T - t.$$

Note that

$$Rho_C - Rho_P = K(T-t)\exp(-r(T-t))N(d_2).$$

5.2.2.7 *Additional relationship*

Lemma 5.2.14 (relationship between Delta, Theta, and Gamma). *[6, p. 414] Consider a portfolio $\pi = C - \Delta S$, where $\Delta = \frac{\partial C}{\partial S}$. We have*

- $$\begin{aligned} d\pi &= \frac{\partial \pi}{\partial t}dt + (\frac{\partial C}{\partial S} - \Delta)dS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} \\ &= \frac{\partial \pi}{\partial t}dt + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} \\ &= r\pi dt; \end{aligned}$$

  that is, the value of the portfolio will grow at a rate r.
- $\frac{\partial \pi}{\partial S} = 0$; *that is, the portfolio π is the delta-neutral.*
- $\frac{\partial^2 \pi}{\partial S^2} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}$; *that is, the portfolio π is not Gamma-neutral.*
- *For a delta-neutral portfolio π,*

  $$\Theta + \frac{1}{2}\sigma^2 S^2 \Gamma = r\pi,$$

  where

  $$\Theta = \frac{\partial \pi}{\partial t}, \Gamma = \frac{\partial^2 \pi}{\partial S^2}$$

Proof.

$$\begin{aligned} d\pi \triangleq d(C - \frac{\partial C}{\partial S_t}S_t) &= (\frac{\partial C}{\partial t}dt + \frac{\partial V}{\partial S_t}dS_t + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2}dt) + \frac{\partial C}{\partial S_t}dS_t \\ &= (\frac{\partial C}{\partial t}dt + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2}dt) \\ &= rV - rS_t\frac{\partial C}{\partial S_t}S_t = r\pi \end{aligned}$$

where in the last step we use Black-Scholes equation

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0.$$

□

5.2.2.8 *Greeks with implied volatility*

Lemma 5.2.15 (Delta and Gamma with implied volatility). *Let $V(S_t, K, \sigma(K, S_t))$ be the value at time t of an European option of strike K and maturity T on the spot S_t, where $\sigma(K, F_t)$ denotes the implied volatility that make $V(S_t, K, \sigma(K, S_t))$ equal the market price. The Delta and Gamma calculation are given by*

- $$\Delta = \frac{\partial}{\partial S_t} V(S_t, K, \sigma(K, S_t)) = \frac{\partial V}{\partial S_t} + \frac{\partial V}{\partial \sigma} \frac{\sigma}{\partial S_t}.$$
- $$\Gamma = \frac{\partial^2}{\partial S_t^2} V(S_t, K, \sigma(K, S_t)) = \frac{\partial^2 V}{\partial S_t^2} + \frac{\partial^2 V}{\partial S_t \partial \sigma} \frac{\partial \sigma}{\partial S_t} + \frac{\partial^2 V}{\partial \sigma^2} \frac{\partial^2 \sigma}{\partial S_t^2} + \frac{\partial V}{\partial \sigma} \frac{\partial^2 \sigma}{\partial S_t^2}.$$

Proof. Use calculus and chain rule. □

5.2.3 Log contract

Definition 5.2.2 (log contract). *[9, p. 70] A log contract L is a derivative that, at expiration T, pays the value $L_T = \ln(S_T/K)$, where K is the strike.*

Lemma 5.2.16 (value of a log contract). *Assume Black-Scholes world assumptions hold. Then current value of a log contract with strike K is given by*

$$V_0 = \ln(\frac{S_0}{K}) + rT - \frac{1}{2}\sigma^2 T.$$

Proof. In the Black-Scholes world, S_T under risk-neutral measure is given by

$$\ln S_T = \ln S_0 + rT - \frac{1}{2}\sigma^2 T + \sigma W_T.$$

Take expectation, we have

$$E_Q[\ln S_T] = \ln S_0 + rT - \frac{1}{2}\sigma^2 T.$$

Finally,

$$V_0 = E_Q[\frac{\ln S_T}{K}] = \ln S_0/K + rT - \frac{1}{2}\sigma^2 T$$

□

Lemma 5.2.17 (replicating log contract using vanilla options).

- *The final payoff of a log contract, $L_T = \ln(S_T/K)$, can be decomposed as*

$$\ln(S_T/K) = \frac{S_T - K}{K} - \int_K^\infty \frac{1}{v^2}(S_T - v)^+ dv - \int_0^K \frac{1}{v^2}(v - S_T)^+ dv.$$

- *The replicating strategy is:*
  - *long $1/K$ unit of forward with strike price K.*
  - *short $1/v$ unit of put options at strike v where v ranges from 0 to K.*
  - *short $1/v$ unit of call options at strike v where v ranges from K to ∞.*

Proof. Use the result in Theorem 2.1.2, we have for any twice-continuously differentiable function $f(x)$, we have

$$f(x) = f(\kappa) + f'(\kappa)(x - \kappa) + \int_0^\kappa f''(K)(K - x)^+ dK + \int_\kappa^\infty f''(K)(x - K)^+ dK.$$

Take $x = S_T, \kappa = K$, we will get the result. □

5.2.4 Options on futures and forwards

Lemma 5.2.18 (forward price dynamics at constant interest rate). *[10, p. 101] Assume constant interest rate. The forward price for an asset at T determined at t is given as*

$$F(t) = S_t \exp(r(T - t)).$$

Assuming S_t, under real-world probability measure, follows a geometric Brownian motion with drift μ and volatility σ, then we have

$$dF_t = (\mu - r)F_t dt + \sigma F_t dW_t.$$

Particularly, under the risk-neutral measure Q such that S_t has a drift $rS_t dt$, we have

$$dF_t = \sigma F_t d\hat{W}_t,$$

where $\hat{W}$ is the Brownian motion under the measure Q.

Proof. Note that

$$dF_t = -rSe^{r(T-t)}dt + e^{r(T-t)}dS,$$

then we have

$$dF_t = (\mu - r)S\exp(r(T - t))dt + \sigma S\exp(r(T - t))dW_t.$$

Under the measure Q, $dW_t = d\hat{W}_t - (\mu - r)dt$. □

Lemma 5.2.19 (Black-Scholes equation for options on forwards/futures). *[10, p. 101] Let $V(F(t), t)$ be the value of the derivative as a function of the forward/futures price $F(t)$ and time t. Assume $F(t)$ is governed by*

$$dF_t = (\mu - r)F_t dt + \sigma F_t dW_t$$

where W_t is the Brownian motion. Then V is governed by

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 F^2 \frac{\partial^2 V}{\partial F^2} - rV = 0$$

with final condition $V(F(T), T) = V_T(F(T))$ and boundary condition $V(S,t) = V_a(t)$ on $S = a$ and $V(S,t) = V_b(t)$ on $S = b$.

Proof. (1) (PDE method) Consider the portfolio $\Pi = V - \Delta F$, (we are hedging by entering forward contracts at different time)then

$$dV = (\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 F^2)dt + \frac{\partial V}{\partial F}dF.$$

Let $\Delta = \frac{\partial V}{\partial F}$, then

$$d\Pi = dV - \Delta dF = (\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 F^2)dt = rVdt.$$

Note that the right-hand side is $rVdt$ instead of $r\Pi dt$, since it takes zero money to enter the contract. (2) (martingale method) Note that $D(t)V(F(t), t)$is a martingale under the risk-neutral measure Q. Using martingale pricing theorem(Theorem 1.6.13), we have

$$D(t)V(F(t), t) = E_Q[D(T)V(F(T), T)|\mathcal{F}_t].$$

Because $F(t)$ is following SDE under Q given by

$$dF_t = \sigma F_t d\hat{W}_t,$$

we can derive the associated PDE using Feyman-Kac theorem(Theorem 1.8.1). □

Lemma 5.2.20 (Solutions to Black-Scholes equation for futures and forwards). *[10, p. 101] The call option with strike K and matures at T will have price:*

$$C(F,t) = e^{r(T-t)}(FN(d_1) - EN(d_2)).$$

Proof. Since the Black-Scholes equation is similar to the case of asset paying dividends, we can use results from Theorem 1.6.13. □

5.2.5 Options of multiple assets

Lemma 5.2.21 (options on multiple assets). *[1, p. 183] The option value $V(S_1, ..., S_d, t)$ on d assets with Geometric Brownian motion is given as*

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sigma_i\sigma_j\rho_{ij}S_iS_j\frac{\partial^2 V}{\partial S_i \partial S_j} + \sum_{i=1}^{d}(r - D_i)S_i\frac{\partial V}{\partial S_i} - rV = 0$$

under the risk-neutral measure Q, the asset dynamics become

$$dS_i(t) = (r(t) - D_i)S_i(t)dt + \sigma(t)S_i(t)dW_i(t), i = 1, ..., d.$$

We have

$$V(t) = E_Q[V(S_1(T), ..., S_d(T)|\mathcal{F}_t].$$

Proof. Let

$$\Pi = V(S_1, ..., S_d, t) - \sum_{i=1}^{d}\Delta_i S_i,$$

then

$$d\Pi = \frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sigma_i\sigma_j\rho_{ij}S_iS_j\frac{\partial^2 V}{\partial S_i \partial S_j}dt + \sum_{i=1}^{d}(\frac{\partial V}{\partial S_i} - \Delta_i)dS_i.$$

If we choose $\Delta_i = \frac{\partial V}{\partial S_i}, \forall i$, then the portfolio is risk-free. Then $d\Pi = r\Pi dt$. We have

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sigma_i\sigma_j\rho_{ij}S_iS_j\frac{\partial^2 V}{\partial S_i \partial S_j} + \sum_{i=1}^{d}(r - D_i)S_i\frac{\partial V}{\partial S_i} - rV = 0.$$

Then we can use Theorem 1.8.6 to show the rest. □

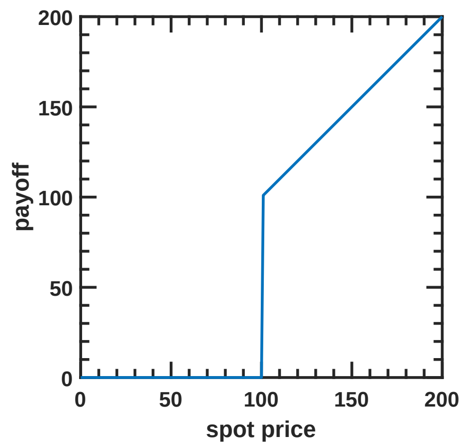
5.3 Exotic Option pricing

5.3.1 Digital option

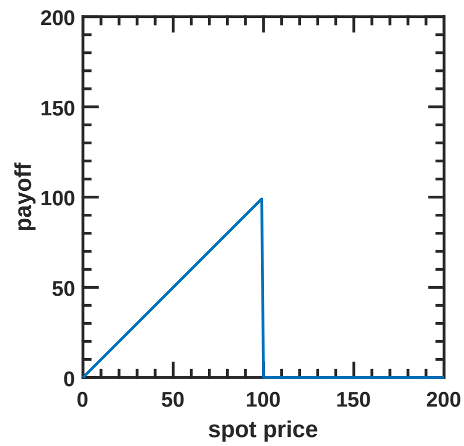
5.3.1.1 *Valuation*

Definition 5.3.1 (cash or nothing digital call option). *A cash-or-nothing digital call or put option on the underlying asset S_t at maturity T with strike K has payoff $1_{S_T > K}$ or $1(S_T - K)$.*

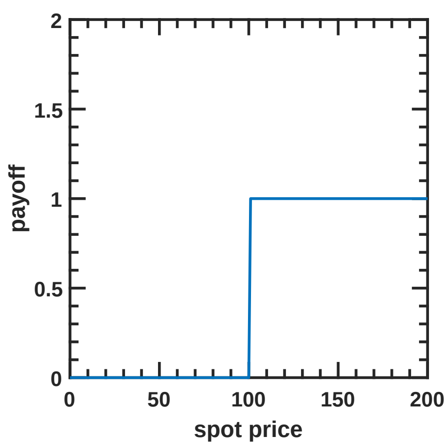
Definition 5.3.2 (asset or nothing digital call option). *An asset-or-nothing digital call or put option on the underlying asset S_t at maturity T with strike K has payoff $S_T 1_{S_T>K}$ or $S_T 1_{S_T-K}$.*



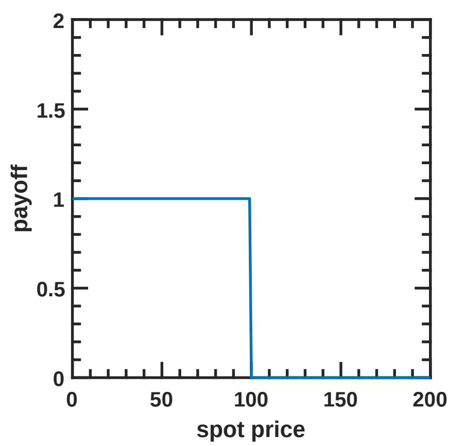
(a) Asset-or-nothing digital call option payoff at maturity. The strike $K = 100$.



(b) Asset-or-nothing digital put option payoff at maturity. The strike $K = 100$.



(c) Cash-or-nothing digital call option payoff at maturity. The strike $K = 100$.



(d) Cash-or-nothing digital put option payoff at maturity. The strike $K = 100$.

Figure 5.3.1: Payoff function for digital options.

Lemma 5.3.1 (put-call parity for digital options). *Assume current time is t and the risk-free rate is constant r. It follows that*

- *The cash-or-nothing digital call and put with the same maturity T and strike K satisfy*

$$C_t + P_t = \exp(-r(T-t)),$$

where C_t, P_t denote the current values for the call and the put.

- *The asset-or-nothing digital call and put with the same maturity T and strike K satisfy*

$$C_t + P_t = \exp(-D(T-t))S_t,$$

where C_t, P_t, S_t denote the current values for the call, the put and the asset, D is the dividend rate.

Proof. (1) Note that $C_T + P_T$ has payoff of 1 at all possible situations. Therefore

$$C_t + P_t = \exp(-r(T-t))E_Q[C_T + P_T|\mathcal{F}_t] = \exp(-r(T-t)).$$

(2)Note that $C_T + P_T$ has payoff of S_T at all possible situations. Therefore

$$C_t + P_t = \exp(-r(T-t))E_Q[C_T + P_T|\mathcal{F}_t] = \exp(-r(T-t))E_Q[S_T|\mathcal{F}_t] = \exp(-D(T-t))S_t.$$

□

Lemma 5.3.2 (cash or nothing digital call/put pricing). *Assume that the underlying asset S_t is following*

$$dS_t/S_t = \mu dt + \sigma dW_t.$$

The risk-neutral price of a digital call option is

$$V(t) = e^{-r(T-t)}P_Q(S_T > K) = e^{-r(T-t)}N(d),$$

where

$$d = (\log(S_t/K) + (r - \frac{\sigma^2}{2})(T-t))/\sigma\sqrt{T-t}.$$

or equivalently,

$$d = (\log(F_t/K) - \frac{\sigma^2}{2}(T-t))/\sigma\sqrt{T-t}$$

where F_t is the forward price.

Similarly, the risk-neutral price of a digital put option is

$$V(t) = e^{-r(T-t)}P_Q(S_T < K) = e^{-r(T-t)}N(-d).$$

Proof. (1) Under the risk-neutral measure, the S_T is given by

$$S_T = S_t\exp((r - \frac{\sigma^2}{2})(T-t) + \sigma(W_T - W_t)).$$

Then

$$\begin{aligned}
P(S_T > K) &= P(S_t \exp((r - \frac{\sigma^2}{2})(T-t) + \sigma(W_T - W_t)) > K) \\
&= P(S_t \exp((r - \frac{\sigma^2}{2})(T-t) + \sigma\sqrt{T-t}Z) > K) \\
&= P(Z > -(\log(S/K) + (r - \frac{\sigma^2}{2})(T-t))/\sigma\sqrt{T-t}) \\
&= P(Z > -d) = N(d)
\end{aligned}$$

where

$$d = (\log(S_t/K) + (r - \frac{\sigma^2}{2})(T-t))/\sigma\sqrt{T-t}.$$

(2)

$$\begin{aligned}
P(S_T < K) &= P(S_t \exp((r - \frac{\sigma^2}{2})(T-t) + \sigma(W_T - W_t)) < K) \\
&= P(S_t \exp((r - \frac{\sigma^2}{2})(T-t) + \sigma\sqrt{T-t}Z) < K) \\
&= P(Z < -(\log(S/K) + (r - \frac{\sigma^2}{2})(T-t))/\sigma\sqrt{T-t}) \\
&= P(Z < -d) = N(-d)
\end{aligned}$$

□

Lemma 5.3.3 (asset or nothing digital call/put pricing). *Assume that the underlying asset S_t under risk neutral measure is following*

$$dS_t/S_t = (r-a)dt + \sigma dW_t,$$

where r is the short rate and a is the dividend rate. The risk-neutral price of a digital call option is

$$V_C(t) = e^{-r(T-t)}E_Q[S_T 1_{(S_T>K)}] = e^{-a(T-t)}S(t)N(d),$$

where

$$d = (\log(S_t/K) + (r - a - \frac{\sigma^2}{2})(T-t))/\sigma\sqrt{T-t}.$$

or equivalently,

$$d = (\log(F_t/K) + \frac{\sigma^2}{2}(T-t))/\sigma\sqrt{T-t}$$

where F_t is the forward price.

Similarly, the risk-neutral price of a digital put option(using put-call parity Lemma 5.3.1) is

$$V_P(t) = (\exp(-a(T-t))S_t - V_C(t)) = e^{-a(T-t)}S(t)N(-d).$$

Proof. (1) Under the risk-neutral measure, the S_T is given by

$$S_T = S_t \exp((r - a - \frac{\sigma^2}{2})(T-t) + \sigma(W_T - W_t)).$$

Let $Y = (W(T) - W(t))/\sqrt{T-t}$, then $Y \sim N(0,1)$, and we can write

$$S(T) = S(t)\exp((r - a - \frac{1}{2}\sigma(s)^2)\tau + \sigma\tau Y)),$$

where $\tau = T - t$.

Let

$$d_2 = \frac{\ln\frac{S(t)}{K} + (r - a - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}$$

such that

$$S(t)\exp((r - a - \frac{1}{2}\sigma^2)\tau + \sigma\sqrt{\tau}(-d_2)) = K.$$

Then the for the standard normal Y, We have

$$Pr(S(t)\exp((r - a - \frac{1}{2}\sigma(s)^2)(T-t) + \sigma\tau Y)) > K) = Pr(Y > -d_2) = Pr(Y < d_2).$$

Then

$$\begin{aligned}
V(t) &= \frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty} e^{-r\tau}(S(t)\exp(+\sigma\tau y + (r - a - \frac{1}{2}\sigma^2)\tau))e^{-\frac{1}{2}y^2}dy \\
&= \frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty} S(t)\exp(-\tau a)\exp(-\sigma\tau y - \frac{1}{2}\sigma^2\tau)e^{-\frac{1}{2}y^2}dy \\
&= \frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty} S(t)\exp(-\tau a)\exp(-\frac{1}{2}(y - \sigma\sqrt{\tau})^2)dy \\
&= \frac{1}{\sqrt{2\pi}}\int_{-d_2-\sigma\sqrt{\tau}}^{\infty} S(t)\exp(-\tau a)\exp(-\frac{1}{2}z^2)dz \\
&= S(t)\exp(-\tau a)N(d)
\end{aligned}$$

where we used the fact that $-d = -d_2 - \sigma\sqrt{\tau}$. (2)

$$V_P(t) = (\exp(-a(T-t))S_t - V_C(t)) = e^{-a(T-t)}S(t)(1 - N(d)) = e^{-a(T-t)}S(t)N(-d).$$

□

We use martingale method.

$$C(S_T, T) = E_Q[e^{-r(T-t)} \max(S_T - K, 0) | \mathcal{F}_t].$$

where

$$S(T) = S(t) \exp((r - a - \frac{1}{2}\sigma(s)^2)(T - t) + \sigma(W(T) - W(t))).$$

Let $Y = (W(T) - W(t))/\sqrt{T-t}$, then $Y \sim N(0,1)$, and we can write

$$S(T) = S(t) \exp((r - a - \frac{1}{2}\sigma(s)^2)(T - t) + \sigma\tau Y)).$$

Let

$$d_2 = \frac{\ln \frac{S(t)}{K} + (r - a - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}$$

such that

$$S(t) \exp((r - a - \frac{1}{2}\sigma^2)\tau + \sigma\sqrt{\tau}(-d_2)) = K.$$

Then the for the standard normal Y, We have

$$Pr(S(t) \exp((r - a - \frac{1}{2}\sigma(s)^2)(T - t) + \sigma\tau Y)) > K) = Pr(Y > -d_2) = Pr(Y < d_2).$$

$$\begin{aligned}
C(S_T, T) &= \frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} e^{-r\tau}(S(t) \exp(+\sigma\tau y + (r - a - \frac{1}{2}\sigma^2)\tau) - K)e^{-\frac{1}{2}y^2} dy \\
&= [\frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} S(t) \exp(-\tau a) \exp(-\sigma\tau y - \frac{1}{2}\sigma^2\tau)e^{-\frac{1}{2}y^2} dy] - [\frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} e^{-r\tau}e^{-\frac{1}{2}y^2} dy] \\
&= [\frac{1}{\sqrt{2\pi}} \int_{-d_2}^{\infty} S(t) \exp(-\tau a) \exp(-\frac{1}{2}(y - \sigma\sqrt{\tau})^2) dy] - e^{-r\tau}KN(d_2) \\
&= [\frac{1}{\sqrt{2\pi}} \int_{-d_2-\sigma\sqrt{\tau}}^{\infty} S(t) \exp(-\tau a) \exp(-\frac{1}{2}z^2) dz] - e^{-r\tau}KN(d_2) \\
&= S(t) \exp(-\tau a)N(d_1) - e^{-r\tau}KN(d_2)
\end{aligned}$$

where we used the fact that $-d_1 = -d_2 - \sigma\sqrt{\tau}$.

Lemma 5.3.4 (hedging digital call option, connection to European call). *[11, p. 2] Let $D(K, T)$ be the price of a digital call option. Let $C(K, T)$ be the price of an European call option. We have*

- $$D = -\frac{\partial C}{\partial K}.$$
- *To (approximately) hedge a short position of D, we can long a call at strike K and short a call at strike $K + \epsilon$ in quantity $1/\epsilon$.*

*Note that **these results are model independent**.*

Proof. We can use the result at Lemma 2.1.1. If we use a Black-Scholes model(which satisfies the no-arbitrage condition), then

$$C(S_t, t, T, K) = N(d_1)S_t - N(d_2)Ke^{-r(T-t)},$$

and we can show

$$-\frac{PaC}{\partial K} = N(d_2)e^{-r(T-t)} = D(K, T).$$

□

Remark 5.3.1 (theoretical importance of digital option). Digital options have important theoretical consequence in pricing exotic option and volatility model. See **??**.

5.3.1.2 *Risk analysis*

Lemma 5.3.5 (greeks for cash-or-nothing digital options). *Let the current time be t. Consider a cash-or-nothing digital call/put option with strike K and expiry T.*

- *(Delta)*
$$\Delta = \frac{\partial C_t}{\partial S_t} = \frac{\exp(-r(T-t))}{\sigma S_t\sqrt{T-t}}N'(d_2).$$
- *(Gamma)*
$$\Delta = \frac{\partial^2 C_t}{\partial S_t^2} = -\frac{\exp(-r(T-t))d_1}{\sigma S_t(T-t)}N'(d_2).$$
- *(Theta)*
$$\Delta = \frac{\partial C_t}{\partial \tau} = r\exp(-r\tau)N(d_2) + \exp(-r\tau)N'(d_2)(\frac{d_1}{2\tau} - \frac{r}{\sigma\sqrt{\tau}}).$$
- *(Vega)*
$$\Delta = \frac{\partial C_t}{\partial \sigma} = -\frac{\exp(-r(T-t))d_1}{\sigma}N'(d_2).$$

where

$$d_1 = \frac{\ln(S_t/K) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, d_2 = d_1 + \sigma\sqrt{T-t}.$$

Proof. (1) Note that $C_t = \exp(-r(T-t))N(d_2)$. Then

$$\frac{\partial C_t}{\partial S_t} = \exp(-r(T-t))N'(d_2)\frac{\partial d_2}{\partial S_t} = \exp(-r(T-t))N'(d_2)\frac{1}{S_t\sigma\sqrt{T-t}}.$$

□

5.3.2 Asian option

5.3.2.1 *Canonical approach*

Definition 5.3.3 (Asian option). *An Asian call option with strike K and maturity T has payoff given by*

$$V(T) = \max\left(\frac{1}{T}\int_0^T S(u)du - K, 0\right),$$

where $S(t)$ is the stochastic process for the underlying asset.

Remark 5.3.2 (business needs). An Asian option can be used when managing the risk exposure(e.g., stock price) for a period of time. For example, a company needs to buy a stock in the next 12 months, with each month a fixed quantity. Then the company has the exposure to the stock price in the next 12 months.

Lemma 5.3.6 (price equation). *[4, pp. 279, 322] The Asian call price $v(t,x,y)$ satisfies*

$$v_t(t,x,y) + rxv_x(t,x,y) + xv_y(t,x,y) + \frac{1}{2}\sigma^2 x^2 v_{xx}(t,x,y) = rv(t,x,y), 0 \leq t \leq T, x \geq 0, y \in \mathbb{R},$$

with boundary conditions

$$\begin{aligned} v(t,0,y) &= e^{-r(T-t)}\left(\max\left(\frac{y}{T} - K, 0\right)\right) \\ \lim_{y\to-\infty} v(t,x,y) &= 0, 0 \leq t \leq T, x \geq 0 \\ v(T,x,y) &= \max\left(\frac{y}{T} - K, 0\right), x \geq 0, y \in \mathbb{R}. \end{aligned}$$

Proof. Define $Y(t) = \int_0^T S(u)du$ such that $dY(t) = Y(t)dt$. Under risk-neutral measure Q,

$$e^{-rt}V(t) = E_Q[e^{-rT}\max\left(\frac{1}{T}Y(T) - K, 0\right)|\mathcal{F}_t]$$

is a martingale.

$$\begin{aligned} d(e^{-rt}v(t,x,y)) &= e^{-rt}(-rvdt + v_t dt + v_x dX + v_y dY + \frac{1}{2}v_{xx}dXdX + v_{xy}dXdY + \frac{1}{2}v_{yy}dYdY) \\ &= e^{-rt}(-rvdt + v_t dt + v_x rxdt + v_y xdt + \frac{1}{2}v_{xx}\sigma^2 x^2 dt) + e^{-rt}v_x \sigma x dW(t) \end{aligned}$$

set the drift to zero, we get the PDE. $\square$

Remark 5.3.3 (interpretation). Under risk-neutral measure Q, the dynamics of $(Y(t), S(t))$ is given by

$$dS(t) = rS(t)dt + \sigma S(t)dW(t)$$
$$dY(t) = S(t)dt.$$

In terms of quadratic variation properties, we have

$$dYdY = 0, dYdS = 0$$

Remark 5.3.4 (how to hedge).

5.3.2.2 *put-call parity*

Lemma 5.3.7 (put-call parity for Asian options). *Consider an Asian call with expiration date t_N and payoff given by*

$$\max(A(t_N) - K, 0),$$

where

$$A(t_m) = \frac{1}{m+1}\sum_{i=0}^{m} S(t_i).$$

Similarly, consider an Asian put with payoff given by

$$\max(K - A(t_N), 0).$$

It follows that

- *at maturity t_N, we have*
$$C(t_N) - P(t_N) = A(t_N) - K.$$
- *at any earlier date $t < t_0$, we have*

$$C(t) - P(t) = \frac{1}{N+1}\sum_{i=0}^{N} \exp(-r(t_N - t_i))S(t) - K\exp(-r(t_N - t)).$$

Proof. (1) straight forward. (2)

$$C(t_N) - P(t_N) = \frac{1}{N+1}\sum_{i=0}^{N} S(t_i) - K$$

$$E[\exp(-r(t_N-t))C(t_N)|\mathcal{F}_t] - E[\exp(-r(t_N-t))P(t_N)|\mathcal{F}_t]] = \frac{1}{N+1}\sum_{i=0}^{N} E[\exp(-r(t_N-t))S(t_i)|\mathcal{F}_t] - E[$$

$$C(t) - P(t) = \frac{1}{N+1}\sum_{i=0}^{N} E[\exp(-r(t_i - t - t_i + t_N))S(t_i)|$$

$$C(t) - P(t) = \frac{1}{N+1}\sum_{i=0}^{N} \exp(-r(t_N - t_i))S(t) - K\exp(-$$

□

5.3.2.3 *Approximation via geometric average*

Lemma 5.3.8. *Suppose*

$$S_t = S_0\exp(vt + \sigma W_t), v = r - \frac{1}{2}\sigma^2.$$

Then the geometric average

$$G_m = (S_0 \times S_1 \times \cdots \times S_m)^{1/(m+1)},$$

where $S_i = S(ih), h = T/m$, has log-normal distribution given by

$$G_m = S_0\exp(v\frac{T}{2} + \sigma\sqrt{\frac{2m+1}{6(m+1)}}W_T).$$

If $m \to \infty$, we have

$$G_\infty = S_0\exp(v\frac{T}{2} + \sigma\sqrt{\frac{1}{3}}W_T).$$

Proof. (1) The coefficient for the drift is given by

$$\frac{1}{m+1}(h + 2h + \cdots + mh) = \frac{mh}{2} = \frac{T}{2}.$$

(2) Note that

$$
\begin{aligned}
&(W_0 + W_1 + W_2 + \cdots + W_m)\\
=&m(W_1 - W_0) + (m-1)(W_2 - W_1) + (m-2)(W_3 - W_2) + \cdots + (W_m - W_{m-1})\\
=&\sqrt{m^2 + (m-1)^2 + (m-2)^2 + \cdots + 1^2)}Z, Z \sim (0, T/m)\\
=&\sqrt{\frac{m(m+1)(2m+1)}{6}}Z\\
=&\sqrt{\frac{m(m+1)(2m+1)}{6}}\frac{\sqrt{W_T}}{\sqrt{m}}
\end{aligned}
$$

□

Remark 5.3.5 (alternative derivation). To derive the coefficient for Brownian motion term, we can consider the fact(**??**) of

$$\frac{1}{T}\int_0^T W_u du = \frac{1}{\sqrt{3}} W_T.$$

5.3.2.4 *Moment matching method*

5.3.3 Basket option

5.3.3.1 *Basics*

Definition 5.3.4 (basket option). *A basket call option with strike K, maturity T, and n underlying $S_t^{(1)}, S_t^{(2)}, ..., , S_t^{(n)}$ has payoff*

$$V_T = (\sum_{i=1}^n w_i S_T^{(i)} - K)^+,$$

where $w_i > 0, w_i \in \mathbb{R}$ are the weights.

Note 5.3.1 (business needs: the price of basket option vs. a portfolio of vanilla options). Consider a basket option with payoff at maturity T given by

$$V_T^B = (\sum_{i=1}^n \frac{1}{n} S_T^{(i)} - K)^+;$$

and a portfolio of call options with payoff at maturity T given by

$$V_T^P = \sum_{i=1}^{n} \frac{1}{n}(S_T^{(i)} - K)^+.$$

Also we assume $S_0^{(i)}, i = 1, 2, ..., n$ are close to K.

Then in general, the basket's option value will be much lower than the value of vanilla option portfolio due to the following arguments:

- If each asset $S^{(i)}$ are independent from each other, then we can view $S^{(1)}, ..., S^{(n)}$ as iid random sample of S. Then the term $\sum_{i=1}^{n} \frac{1}{n} S_T^{(i)}$ will have a $1/\sqrt{n}$ variance of S.
- If each asset $S^{(i)}$ are perfectly correlated with each other, then we can view $S^{(1)}, ..., S^{(n)}$ as exact copy of S. Then the term $\sum_{i=1}^{n} \frac{1}{n} S_T^{(i)}$ is the same as S. In this case, the basket option will have similar price to the portfolio of options.

Therefore, a company usually prefers basket option when managing the risk of a set of assets due to its lower cost.

5.3.3.2 *Moment matching method for pricing*

Lemma 5.3.9 (black formula with extended lognormal distribution family). *[12]*

- $$c = \exp(-rT)[M_1(T)N(d_1) - KN(d_2)],$$

  where

  $$d_{1,2} = \frac{\log M_1(T) - \log K}{\sigma} \pm \frac{1}{2}\sigma$$

  and

  $$\sigma = \sqrt{\log(\frac{M_2(T)}{M_1(T)^2})}.$$

- $$c = \exp(-rT)[(M_1(T) - \tau)N(d_1) - (K - \tau)N(d_2)],$$

  where

  $$d_{1,2} = \frac{\log(M_1(T) - \tau) - \log(K - \tau)}{\sigma} \pm \frac{1}{2}\sigma$$

  and

  $$\sigma = \sqrt{\log(\frac{M_2(T)}{M_1(T)^2})}.$$

- $$c = \exp(-rT)[M_1(T)N(d_1) - KN(d_2)],$$

  where

  $$d_{1,2} = \frac{\log -M_1(T) - \log -K}{\sigma} \pm \frac{1}{2}\sigma$$

  and

  $$\sigma = \sqrt{\log(\frac{M_2(T)}{M_1(T)^2})}.$$

- $$c = \exp(-rT)[M_1(T)N(d_1) - KN(d_2)],$$

  where

  $$d_{1,2} = \frac{\log M_1(T) - \log K}{\sigma} \pm \frac{1}{2}\sigma$$

  and

  $$\sigma = \sqrt{\log(\frac{M_2(T)}{M_1(T)^2})}.$$

Lemma 5.3.10 (log-normal approximate pricing for vanilla call and put). *Assume under risk-neutral measure Q that the underlying dynamics are given by*

$$dS_t^{(i)}/S_t^{(i)} = (r-q)dt + \sigma_i dW_i(t), i = 1, 2, ..., n,$$

and $dW_i dW_j = \rho_{ij}dt$. Denote $B(T) = \sum_{i=1}^n w_i S_T^{(i)}$. It follows that

- *The random variable $B(T)$ can be approximated by Y by matching first two moments defined by*

  $$Y = M_1 \exp(-\frac{1}{2}\sigma^2 + \sigma Z), Z \in N(0,1),$$

  where

  $$M_1 = \sum_{i=1}^n w_i F_i, M_2 = \sum_{i=1}^n \sum_{j=1}^n F_i F_j w_i w_j \exp(\rho_{ij}\sigma_i\sigma_j T), F_i = S^{(i)}(0)\exp((r-q)T).$$

- $$C = \exp(-(r-q)T)(M_1(T)N(d_1) - KN(d_2)).$$

- $$P = \exp(-(r-q)T)(-M_1(T)N(-d_1) + KN(-d_2)).$$

where

$$d_{1,2} = \frac{\log M_1(T) - \log K}{\sigma} \pm \frac{1}{2}\sigma$$

and

$$\sigma = \sqrt{\log(\frac{M_2(T)}{M_1(T)^2})}.$$

Proof. (1) The moments of $B(T)$ is discussed in Lemma 2.3.3. The log-normal approximation is discussed in **??**. □

5.3.3.3 *Copula method for pricing*

Lemma 5.3.11 (copula method for pricing basket option). *Consider a basket option with payoff at maturity T given by*

$$V_T^B = (\sum_{i=1}^n \frac{1}{n} S_T^{(i)} - K)^+;$$

and a portfolio of call options with payoff at maturity T given by

$$V_T^P = \sum_{i=1}^n \frac{1}{n} (S_T^{(i)} - K)^+.$$

We make the following assumptions:

- *We can construct the implied cdf $F_i$for each $S_T^{(i)}, i = 1, 2, ..., n$.(Lemma 2.1.5).*
- *The joint distribution of $S_T^{(1)}, S_T^{(2)}, ..., S_T^{(n)}$ has Gaussian copula with correlation matrix Σ.*[a]

Then we can use the following simulation method to generate one pricing sample

- *First generate $(Y_1, Y_2, ..., Y_n) \sim MN(0, \Sigma)$.*
- *Return $S_T^{(1)} = F_1^{-1}(\phi(Y_1)), S_T^{(2)} = F_2^{-1}(\phi(Y_2)), ..., S_T^{(n)} = F_n^{-1}(\phi(Y_n))$,, where ϕ is standard normal cdf.*

[a] the correlation matrix can be estimated from historical data.

5.3.4 Exchange(Margrabe) option

Definition 5.3.5 (exchange option). *[6, p. 611] Let U_t and V_t be the price processes of two assets. An exchange option with maturity T has payoff at T given by*

$$\max(V_T - U_T, 0).$$

Lemma 5.3.12 (price of exchange option). *[6, p. 612] Consider an exchange option on two assets S_1 and S_2 with maturity T. Assume S_1 and S_2 are following Geometric Brownian motion in the real-world given by*

$$dS_1(t) = (r - q_1 + \lambda_1\sigma_1)S_1(t)dt + \sigma_1 S_1(t)dW_1(t)$$
$$dS_2(t) = (r - q_2 + \rho\lambda_1\sigma_2 + \lambda_2\sqrt{1-\rho^2}\sigma_2)S_2(t)dt + \rho\sigma_2 S_2(t)dW_1(t) + \sqrt{1-\rho^2}\sigma_2 S_2(t)dW_2(t)$$

where W_1 and W_2 are independent Brownian motions.

Then the value of the exchange option at time 0 is given by

$$V_0 = S_1(0)\exp(-q_1T)N(d_1) - S_2(0)\exp(-q_2T)N(d_2)$$

where

$$d_1 = \frac{\ln(S_1(0)/S_2(0) + (q_2 - q_1 + \hat{\sigma}^2/2)}{\hat{\sigma}\sqrt{T}}, d_2 = d_1 - \hat{\sigma}\sqrt{T},$$

and

$$\hat{\sigma} = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$$

Proof. Two stocks,$S_1(t)$ and $S_2(t)$, follows the SDE in real-world measure.

$$dS_1(t) = (r + \lambda_1\sigma_1)S_1(t)dt + \sigma_1 S_1(t)dW_1(t)$$
$$dS_2(t) = (r + \rho\lambda_1\sigma_2 + \lambda_2\sqrt{1-\rho^2}\sigma_2)S_2(t)dt + \rho\sigma_2 S_2(t)dW_1(t) + \sqrt{1-\rho^2}\sigma_2 S_2(t)dW_2(t)$$

where W_1 and W_2 are independent Brownian motions.

A derivative D_t pays $\max(S_2(T) - S_1(T), 0)$ at time T. Develop a formula for its current price at time t.

Solution:

Use the stock S_1 as the numeraire, then under this new measure Q_S,

$$\frac{D(T)}{S_1(T)}$$

is a martingale. Therefore,

$$\frac{D(t)}{S_1(t)} = E_{S_1}[\frac{D(T)}{S(T)}|\mathcal{F}_t] = E_{S_1}[\max(\frac{S_2(T)}{S_1(T)} - 1)|\mathcal{F}_t].$$

Note that under measure Q_S (we take $\lambda_1 = \sigma_1, \lambda_2 = 0$ following Theorem 1.6.19), the dynamics of $S_1(t)$ and $S_2(t)$ follows

$$\begin{aligned}
dS_1(t) &= (r + \sigma_1^2)S_1(t)dt + \sigma_1 S_1(t)dW_1(t) \\
dS_2(t) &= (r + \rho\sigma_1\sigma_2)S_2(t)dt + \ \rho\sigma_2 S_2(t)dW_1(t) + \sqrt{1-\rho^2}\sigma_2 S_2(t)dW_2(t) \\
d\frac{S_2}{S_1} &= \frac{S_2}{S_1}((\rho\sigma_2 - \sigma_1)dW_1 + \sqrt{1-\rho^2}\sigma_2\sigma dW_2)
\end{aligned}$$

Denote $Y = \frac{S_2}{S_1}$, then Y is a geometric Brownian motion with volatility $\sigma_Y = \sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}$.

Then we have $D(t) = S_1(t)E_{S_1}[\max(Y(T) - 1, 0)]$, which can be evaluated. □

Remark 5.3.6 (effects of correlation). Note that negative correlation will increase the value of an exchange option. Consider the following example in Figure 5.3.2, where $S_1(0) = 120, \sigma_1 = 0.3, S_2(0) = 100, \sigma_2 = 0.4, r = 1\%$ and correlation ranges from -1 to 1.

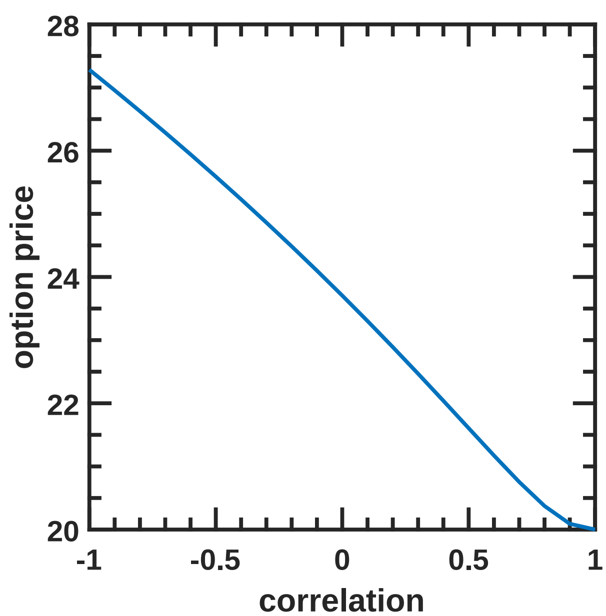


Figure 5.3.2: Correlation effect on exchange option price.

5.3.5 Chooser option

Lemma 5.3.13 (price). *[6, p. 604] Let current time T_0 be zero. A chooser option has right at T_1 to choose a pair of call and put options with strike K and maturity $T_2, T_2 > T_1$. Let r denote* *the risk-free interest rate and q the continuous dividend rate. It follows that the choose option has payoff at T_1 given by*

$$\max(c, p),$$

where c and p denote the price at T_1 of a call and a put with strike K and maturity T_2. The payoff can be decomposed as

- *A call option with strike price K and maturity T_2;*
- $\exp(-q(T_2 - T_1))$ *unit of put option with strike price* $K\exp(-(r-q)(T_2 - T_1))$ *and maturity T_1.*

Proof. Let c and p denote the price at T_1 of a call and a put with strike K and maturity T_2. The payoff at T_1 is given by

$$\begin{aligned}\max(c, p) &= \max(c, c + K\exp(-r(T_2 - T_1)) - S_1\exp(-q(T_2 - T_1))) \\ &= c + \exp(-q(T_2 - T_1))\max(0, K\exp(-(r-q)(T_2 - T_1)) - S_1)\end{aligned}$$

where use the put-call parity Lemma 5.2.2 given by

$$c - p = S_1\exp(-q(T_2 - T_1)) - K\exp(-r(T_2 - T_1)).$$

□

Lemma 5.3.14 (price of general chooser option). *[6, p. 612] For general chooser options, we can have the following decompositions*

$$\begin{aligned}\min(U_T, V_T) &= V_T - \max(V_T - U_T, 0) \\ \max(U_T, V_T) &= U_T + \max(V_T - U_T, 0).\end{aligned}$$

That is, we can decompose it into an asset and an exchange option(Lemma 5.3.12).

Proof. (1)

$$\begin{aligned}\min(U_T, V_T) &= \min(U_T - V_T, V_T - V_T) + V_T \\ &= \min(U_T - V_T, V_T - V_T) + V_T \\ &= V_T - \max(V_T - U_T, 0)\end{aligned}$$

(2)

$$\max(U_T, V_T) = \max(U_T - U_T, V_T - U_T) + U_T = U_T + \max(V_T - U_T, 0)$$

□

5.3.6 Compound option

Definition 5.3.6 (compound option). *Suppose the compound option has strike K and maturity date T while the underlying option has strike K^* and maturity date $T^* > T$.*

- *a call-on-call option gives holder the right buy a call option with strike K^* and maturity T^* at time T; it has payoff at T given by*

$$V_T(S_T) = \max\{0, C_{BS}(S_T, K^*, T^*) - K\}.$$

- *a put on call option has payoff at T*

$$V_T(S_T) = \max\{0, K - C_{BS}(S_T, K^*, T^*)\}.$$

- *a call on put option has payoff at T*

$$V_T(S_T) = \max\{0, P_{BS}(S_T, K^*, T^*) - K\}.$$

- *a put on put option has payoff at T*

$$V_T(S_T) = \max\{0, K - P_{BS}(S_T, K^*, T^*)\}.$$

where

$$\begin{aligned} C_{BS}(S_T, K^*, T) &= N(d_1)S_T - N(d_2)Ke^{-r(T^*-T)}, \\ P_{BS}(S_T, K^*, T) &= -N(-d_1)S_T + N(-d_2)Ke^{-r(T^*-T)}, \end{aligned}$$

where

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T^* - T}}[\ln(\frac{S_T}{K}) + (r + \sigma^2/2)(T^* - T)] \\ d_2 &= d_1 - \sigma\sqrt{T^* - T} \\ N(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy. \end{aligned}$$

Remark 5.3.7 (business application). Consider a company bids to complete a large project in one year. If they win the bid, they will need financing for \$100 million for 3 years. They could buy a three-year interest rate cap beginning the date of the contract but this could be very expensive if they do not win the contract.

In this situation,the company could buy a compoun call option on a three-year interest cap, giving their right to buy the three-year cap at the time of bidding. If they win the contract, they then exercise the option for the interest rate cap at the predetermined premium. And if they do not win the contract, they can let the option expire. The advantage is a lower initial outlay and reduced risk.

Lemma 5.3.15 (compound option put-call parity).

- $$C_{CallonCall}(S_0, T_1, K_1; T_2, K_2) - P_{PutonCall}(S_0, T_1, K_1; T_2, K_2) = C(S_0; T_2, K_2) - K_1 \exp(-rT_1)$$
- $$C_{CallonPut}(S_0, T_1, K_1; T_2, K_2) - P_{PutonPut}(S_0, T_1, K_1; T_2, K_2) = P(X; T_2, K_2) - K_1 \exp(-rT_1)$$

Proof. Straightforward. □

Lemma 5.3.16. *Let $f(x)$ be the PDF of $N(\mu, \sigma^2)$. Then*

$$\int_a^\infty f(x) N(bx+c) dx = \int_a^\infty f(x) \int_{-\infty}^{bx+c} f(y) dy dx = N_2(\frac{\mu - a}{\sigma}, \frac{b\mu + c}{\sqrt{1 + b^2\sigma^2}}; \rho),$$

where $N_2(u, v; \rho)$ is the joint cmf for bivariate standard normal variables (U, V) with correlation

$$\rho = b\sigma / \sqrt{1 + b^2\sigma^2}.$$

Proof. In the original formulation $X = simN(\mu, \sigma^2), Y \sim N(0, 1)$, X and Y are independent. Introduce

$$U = -X, V = -bX + Y,$$

the above integral becomes

$$Pr(X \geq a, Y \leq bX + c) = Pr(U \leq -a, V \leq c),$$

Further note that we can express U, V via

$$U = \sigma Z_1 - \mu, V = b\sigma Z_2 - b\mu, (Z_1, Z_2) \sim MN(0, I; \rho), \rho = \frac{b\sigma}{\sqrt{b^2\sigma^2 + 1}}.$$

Then

$$Pr(U \leq -a, V \leq c) = Pr(Z_1 \leq \frac{\mu - a}{\sigma}, Z_2 \leq \frac{b\mu + c}{\sqrt{1 + b^2\sigma^2}}).$$

□

Lemma 5.3.17. *Consider a call on call compound option has strike K and maturity date T while the underlying option has strike K^* and maturity date $T^* > T$. Then its price is given by*

$$V(t) = e^{-r(T-t)} E_Q[V_T(S_T)|\mathcal{F}_t] = e^{-r(T-t)} \int_{S^*}^{\infty} (C_{BS}(x, K, T) - K) f_{S_T}(x) dx,$$

where S^* is the break-even asset price such that $C_{BS}(S^*, K^*, T^*) = K$.

The above integral can be further reduce to

$$e^{-r(T-t)} \int_{S^*}^{\infty} (C_{BS}(x, K, T) - K) f_{S_T}(x) dx = e^{-r(T-t)} \int_{\ln(S^*/S_t)}^{\infty} (C_{BS}(S_t e^x, K, T) - K) f(x) dx.$$

where $f(x)$ is the pdf of $N(\mu, \sigma_S^2)$.

The evaluation is given by

$$V(t) = S_t N_2(D_1, D_1^*; \rho) - K^* e^{-r(T^*-t)} N_2(D_2, D_2^*; \rho) - K e^{-r(T-t)} N(D_2^*).$$

where

$$D_1^* = D_2^* = \frac{\mu - a}{\sigma}, D_1 = D_2 = \frac{b\mu + c}{\sqrt{1 + b^2\sigma^2}},$$

and

$$\mu = (r + \sigma_S^2/2)(T - t), \sigma = \sigma_S\sqrt{T - t}, \rho = \sqrt{(T - t)/(T^* - t)},$$
$$a = \ln(S^*/S_t), b = 1/(\sigma_S\sqrt{T^* - T})$$

5.3.7 Variance swap

5.3.7.1 *Basics*

Definition 5.3.7 (variance swap). *[11, p. 61] The payoff of a long position on variance swap on an underlying S_t with strike K_{Var} and maturity T is given by*

$$V_T = N \times (\sigma_R^2 - K_{Var}^2),$$

where N is the notional amount, σ_R^2 is the annualized volatility of daily log-return between current $t = 0$ and $t = T$ under zero-mean assumption:

$$\sigma_R^2 = \frac{252}{T-2} \sum_{i=0}^{T-1} \ln^2(S_{i+1}/S_i)^2.$$

Remark 5.3.8 (other variance products). [13, p. 119]

- **volatility swap**, where two parties exchange future realized volatility with a fixed value.
- **CBOE VIX futures**, which enables investors to bet the short-term implied volatility at a future date.
- **Options on realized volatility**, options on VIX index.

Lemma 5.3.18 (the fair value of a variance swap via risk-neutral pricing). *Let $\sigma^2(t)$ be a stochastic process modeling the realized variance via*

$$dS_t/S_t = rdt + \sigma(t)dW_t,$$

under risk-neutral measure Q. Further assume constant interest rate r. It follows that

- *The current value at $t = 0$ is given by*

$$V_0 = E_Q[\exp(-rT)(\sigma_R^2 - K_{Var})],$$

where

$$\sigma_R^2 = \frac{1}{T}\int_0^T \sigma^2(t)dt.$$

- *The fair strike that makes $V_0 = 0$ is given by*

$$K_{Var} = E_Q[\sigma_R^2] = \frac{1}{T}E_Q[\int_0^T \sigma^2(t)dt].$$

Proof. (1) Note that $\sigma_R^2 = \frac{1}{T}\int_0^T \sigma^2(t)dt$ is the continuous-time version of realized variance starting from $t = 0$ to T. (2)

$$V_0 = E_Q[\exp(-rT)(\sigma_R^2 - K_{Var})] = \exp(-rT)E_Q[(\sigma_R^2 - K_{Var})] = 0 \implies K_{Var} = \sigma_R^2.$$

□

5.3.7.2 *Pricing via replication*

Lemma 5.3.19 (continuous replicating realized variance). *Let $\sigma^2(t)$ be a stochastic process modeling the realized variance via*

$$dS_t/S_t = rdt + \sigma(t)dW_t,$$

under risk-neutral measure Q. Further assume constant interest rate r. It follows that

- *the fair strike that makes $V_0 = 0$ is given by*

$$K_{Var} = \frac{2}{T}(rT - E_Q[\ln\frac{S_T}{S_0}]).$$

- *the fair strike equals the value of the following portfolios:*
  - *cash value $2r$.*
  - *a short position of $\frac{2\exp(rT)}{T}\frac{1}{S_0}$ unit of forward with strike price S_0.*
  - *long $\frac{2\exp(rT)}{T}\frac{1}{v}$ unit of put options at strike v where v ranges from 0 to S_0.*
  - *long $\frac{2\exp(rT)}{T}\frac{1}{v}$ unit of call options at strike v where v ranges from S_0 to ∞.*

Proof. (1)Under the dynamic model for S_t, we have

$$\begin{aligned} d(\ln S_t) &= (\mu - \frac{1}{2}\sigma^2)dt + \sigma dZ_t \\ \implies \sigma^2 dt &= \frac{2dS_t}{S_t} - d(\ln S_t) \\ K_{Var} = \frac{1}{T}\int_0^T \sigma^2 dt &= \frac{2}{T}[\int_0^T \frac{dS_t}{S_t} - d(\ln S_t)] \\ &= \frac{2}{T}[\int_0^T \frac{dS_t}{S_t} - (\ln S_T - \ln S_0)]. \end{aligned}$$

Now note

$$E[\frac{dS_t}{S_t}] = E[rdt + \sigma(t)dW_t] = rdt \implies \int_0^T \frac{dS_t}{S_t} = rT.$$

Therefore,

$$K_{Var} = \frac{2}{T}(rT - E_Q[\ln\frac{S_T}{S_0}])$$

(2) the details on replicating the payoff $\ln(S_T/S_0)$ is at Lemma 5.2.17.

□

5.4 Notes on bibliography

For practical investment strategy using options, see [14].

For advanced treatment of options, see [15].

For volatility and correlation modeling, see [11][9][16].

Practical option books:[17][18][19][20].

For dynamic hedging, see [21][22].

For equity option quotes data, see NYSE.

BIBLIOGRAPHY

1. Wilmott, P. *Paul Wilmott on quantitative finance* (John Wiley & Sons, 2013).
2. Remillard, B. *Statistical methods for financial engineering* (CRC Press, 2013).
3. Shonkwiler, R. W. *Finance with Monte Carlo* (Springer, 2013).
4. Shreve, S. E. *Stochastic calculus for finance II: Continuous-time models* (Springer Science & Business Media, 2004).
5. Black, F. & Scholes, M. The pricing of options and corporate liabilities. *Journal of political economy* **81,** 637–654 (1973).
6. Hull, J. *Options, Futures and Other Derivatives 9 edition* (Pearson Education, 2009).
7. Hull, J. *Risk management and financial institutions,+ Web Site* (John Wiley & Sons, 2012).
8. Yu, X. & Xie, X. On Derivations of Black-Scholes Greek Letters. *Research Journal of Finance and Accounting* **4,** 80–85 (2013).
9. Derman, E., Miller, M. & Park, D. *The Volatility Smile* ISBN: 9781118959169. https://books.google.com/books?id=WKfBDAAAQBAJ (Wiley, 2016).
10. Lalley, S. *The mathematics of financial derivatives: a student introduction* (Chicago University Statistics 390/ Mathematical Finance 345, 2001).
11. Bossu, S. *Advanced Equity Derivatives: Volatility and Correlation* (John Wiley & Sons, 2014).
12. Borovkova, S., Permana, F. J. & Weide, H. V. A closed form approach to the valuation and hedging of basket and spread options. *Journal of Derivatives* **14,** 8 (2007).
13. Bossu, S. & Henrotte, P. *An Introduction to Equity Derivatives: Theory and Practice* ISBN: 9781119961857. https://books.google.com/books?id=SgG8cg0uAEgC (Wiley, 2012).
14. McMillan, L. G. *Options as a strategic investment* (Penguin, 2002).
15. Guyon, J. & Henry-Labordère, P. *Nonlinear option pricing* (CRC Press, 2013).
16. Rebonato, R. *Volatility and correlation: the perfect hedger and the fox* (John Wiley & Sons, 2005).
17. Shover, L. *Trading Options in Turbulent Markets: Master Uncertainty Through Active Volatility Management* (John Wiley & Sons, 2012).
18. Haug, E. G. *The complete guide to option pricing formulas* (McGraw-Hill Companies, 2007).
19. Haug, E. G. *Derivatives models on models* (John Wiley & Sons, 2013).
20. Bennett, C. *Trading Volatility: Trading Volatility, Correlation, Term Structure and Skew* (CreateSpace Independent Publishing Platform, 2014).

21. Passarelli, D. *Trading options Greeks: How time, volatility, and other pricing factors drive profits* (John Wiley & Sons, 2012).

22. Taleb, N. *Dynamic hedging: managing vanilla and exotic options* (John Wiley & Sons, 1997).

6

FOREIGN EXCHANGE MODELING

6.1 Overview

The foreign exchange (FX) market and its currency option market are probably the world's largest market. The major market participants include investment institutions, banks, international corporations, and private investors.

FX market shares many similarities with equity market since foreign currency can be viewed as risky assets with additional FX market risk factors, e.g., FX spot rate, FX spot rate volatility etc. Additionally, FX market can be perceived from both perspective of domestic investor and foreign investors; therefore, some symmetric properties can be employed or have to be satisfied.

This chapter will start with basics of FX market (spot rate, forward rate, interest rate parity, and option delta conventions); then we move to the martingale pricing framework addressing the general pricing theorem in the FX market; We Finally, we extending single-currency short-rate and forward rate modeling framework to cross-country setting as the preparation for credit valuation adjustment in

6.2 The FX basics and market

6.2.1 FX market

6.2.1.1 *Spot market*

Definition 6.2.1 (FX spot rate).

- *The first currency of a currency pair is called the base currency, and the second currency is called the quote currency. The currency pair shows how much of the quote currency is needed to purchase one unit of the base currency.*
- *The* ***FX spot rate*** S_t *or* $S(t)$*, represents the number of units of domestic currency needed to buy one nit of foreign current at time* t*. Essentially,* ***spot rate*** *is the price of foreign currency in the unit of domestic currency.*
- *The FX spot rate is usually denoted by symbol* ***FOR-DOM*** *or* ***FOR/DOM****. For example, EUR-USD = 1.39 means that one EUR is worth 1.39 USD. The EUR-USD = 1.39 quote is equivalent to USD-EUR = 0.7194.*

Remark 6.2.1 (commonly traded currency pairs). In the FX market, the four most popular traded currency pairs are:

- EUR/USD (euro/dollar)
- USD/JPY (U.S. dollar/Japanese yen)
- GBP/USD (British pound/dollar)
- USD/CHF (U.S. dollar/Swiss franc)

The three less popular commodity pairs are:

- AUD/USD (Australian dollar/U.S. dollar)
- USD/CAD (U.S. dollar/Canadian dollar)
- NZD/USD (New Zealand dollar/U.S. dollar)

These currency pairs, along with their various combinations (such as EUR/JPY, GBP/JPY and EUR/GBP) account for the majority of all speculative trading in FX.

Definition 6.2.2 (FX forward rate).

- *The **FX forward rate** $F(t, T)$, is the exchange rte between the domestic currency and the foreign currency at some future point of time T as observed at the present time $t(t < T)$.*
- *The FX forward rate is usually derived from forward/futures contract traded on the market.*

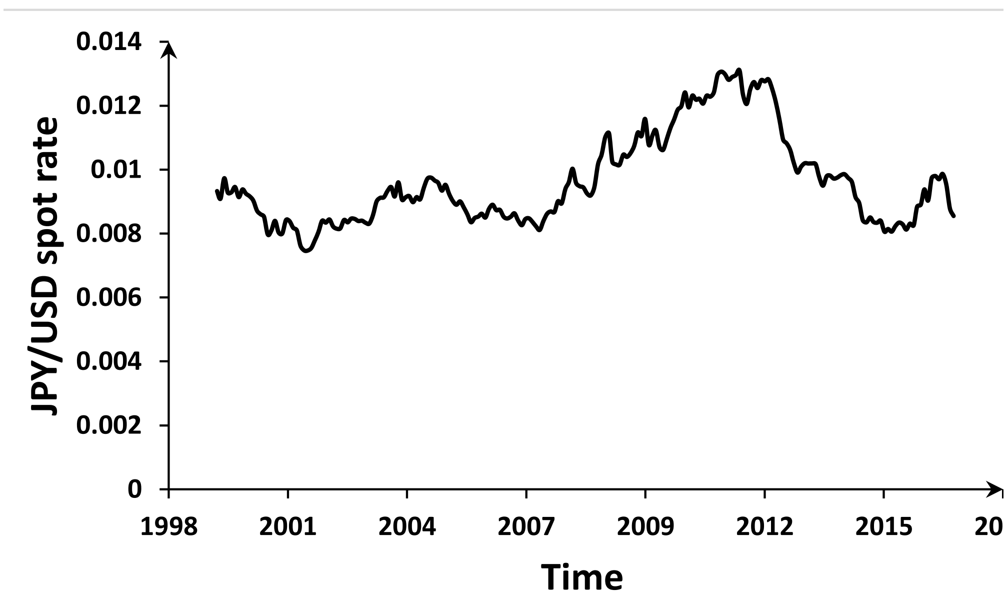


Figure 6.2.1: FX JPY/USD historical quote time series.

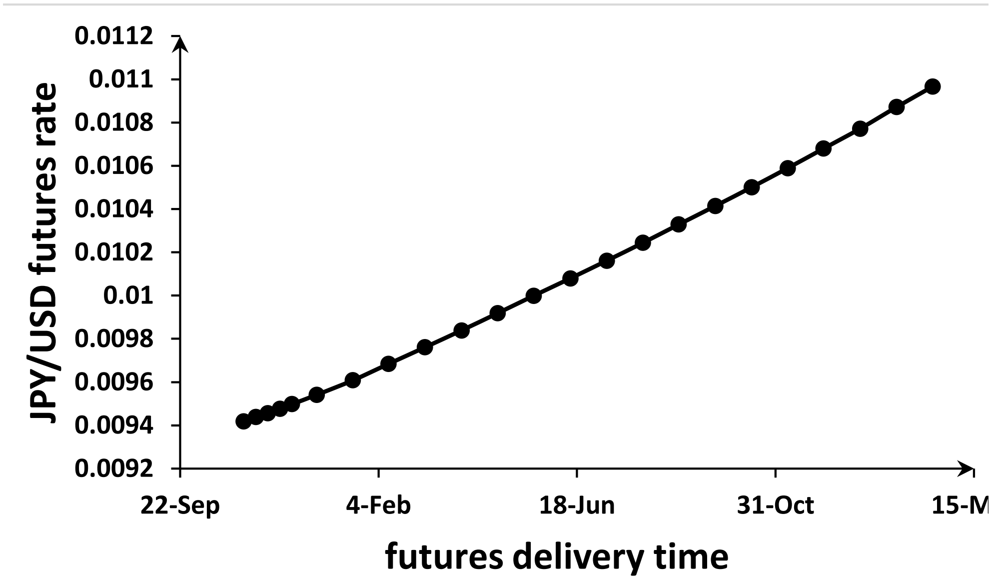


Figure 6.2.2: FX JPY/USD forward rate curve observed on Mar,11,2018

Table 6.2.1: FX futures and European options on FX futures. data source: CME group http://www.cmegroup.com/trading/fx/g10/japanese-yen.html

contracts	**futures settlement date**	**futures delivery date**	**option expiry date**
MAR18	19-Mar-2018	22-Mar-2018	09-Mar-2018
APR18	16-Apr-2018	18-Apr-2018	06-Apr-2018
MAY18	14-May-2018	16-May-2018	04-May-2018
JUN18	18-Jun-2018	20-Jun-2018	08-Jun-2018
...	...	...	...

6.2.1.2 *Forward market*

6.2.1.3 *FX trading PnL accounting*

Remark 6.2.2.

- In the FX trading, domestic investors usually will calculate the profit&loss(PnL) of their position in domestic currency. Economically, domestic currency is the currency investors will ultimately hold and use.

- However, a domestic investors can also calculate PnL using foreign currency if he wants to ultimately emigrate to foreign country and use foreign currency.
- Consider the holding of domestic money market account. Investors will not have FX exposure based on domestic currency accounting; However, they have have FX exposure based on foreign currency accounting.

6.2.2 The integrated market

6.2.2.1 *The setup of the integrated market*

6.2.2.2 *Fundamental relationships in the integrated market*

Proposition 6.2.1 (no-arbitrage relation for asset spot prices denominated in different currencies). *Suppose there is a market where an asset is quoted in both domestic currency and foreign currency. Let $V^f(t)$ denote its foreign currency denominated price at time t. Let $V^d(t)$ denote its domestic currency denominated price at time t. Let $S(t)$ denote the spot rate. Then*

$$S(t)V^f(t) = V^d(t).$$

Let current time be zero. For time $t > 0$, the equality means for all sample path $\omega \in \Omega$, we have

$$S(t,\omega)V^f(t,\omega) = V^d(t,\omega).$$

Proof. If such relation does not hold (suppose $S(t)V^f(t) > V^d(t)$), we can short 1 units of asset using foreign quote and get $S(t)V^f(t)$ domestic currency; buy 1 unit of the asset using domestic quote and cover the short position. There is a riskless profit of $V^d(t) - S(t)V^f(t) - V^d(t)$. □

Proposition 6.2.2 (no-arbitrage relation for asset spot prices denominated in different currencies). *Suppose at current time t there is a market where an asset is quoted in both domestic currency and foreign currency. Let $F^f(t,T)$ denote its foreign currency denominated forward price at time t. Let $V^d(t)$ denote its domestic currency denominated forward price at time t. Let $F^X(t,T)$ denote the forward FX rate. Then for all $t \leq T$, we have*

$$F^X(t,T)F^f(t,T) = F^d(t,T).$$

Let current time be zero. For time $t > 0$, the equality means for all sample path $\omega \in \Omega$, we have

$$F^X(t,T,\omega)F^f(t,T,\omega) = F^d(t,T,\omega).$$

Proof. If such relation does not hold (suppose $F^X(t,T)F^f(t,T) > F^d(t,T)$), we can short 1 unit of forward asset contract using foreign quote, enter the FX forward contract to buy $F^X(t,T)$ units of domestic currency at time T, and enter 1 unit of forward asset contract using domestic quote.

At time T, there is a riskless profit of

$$\underbrace{F^X(t,T)F^f(t,T)}_{\text{gain from short forward asset contract}} - \underbrace{F^d(t,T)}_{payforthelongforwardassetcontract}.$$

□

6.2.3 Volatility skew

6.2.3.1 *Delta types*

Definition 6.2.3 (spot delta). *[1]*

- *In the Black model, the sensitivity of the European option price (denominated in domestic currency) with respect to the spot rate S_t is given by(Lemma 6.5.7)*

$$\Delta_S(K,\sigma,\phi) \triangleq \frac{\partial v}{\partial S_t} = \phi \exp(-r_f(T-t))N(\phi d_+),$$

where $\phi = 1$ for call, $\phi = -1$ for put,

$$v(K,\sigma,S_t,\phi) = \phi(e^{-r^f\tau}S_tN(d_+) - e^{-r\tau}KN(d_-)),$$

and

$$d_\pm = \frac{1}{\sigma\sqrt{\tau}}[\log(\frac{S_t}{K}) + (r^d - r^f \pm \frac{1}{2}\sigma^2)\tau], \tau = T - t.$$

- *And we have put-call delta parity*

$$\Delta_S(K,\sigma,+1) - \Delta_S(K,\sigma,-1) = \exp(-r_f(T-t)).$$

Remark 6.2.3 (comparison with hedging equity option).

- In equity markets, one would buy Δ_S units of the stock, using $\Delta_S \cdot S(t)$ units of currency, to hedge a short vanilla option position.

- In FX markets, this is equivalent to buying Δ_S times the foreign notional N, in the mean time, selling of $\Delta_S N S_t$ units of domestic currency.

Definition 6.2.4 (forward delta). *[1]*

- *In the black model, the sensitivity of the European option price (denominated in domestic currency) with respect to the forward rate $F(t,T)$ is given by(Lemma 6.5.7)*

$$\Delta_F(K,\sigma,\phi) \triangleq \frac{\partial v}{\partial F(t,T)} = \frac{\partial v}{\partial S_t}\frac{\partial S_t}{\partial F(t,T)} = \phi N(\phi d_+),$$

where $\phi = 1$ for call, $\phi = -1$ for put,

$$v(K,\sigma,S_t,\phi) = \phi(e^{-r\tau}(F(t,T)N(d_+) - KN(d_-)),$$

and

$$d_\pm = \frac{1}{\sigma\sqrt{\tau}}[\log(\frac{F(t,T)}{K}) \pm \frac{1}{2}\sigma^2)\tau], \tau = T - t.$$

- *And we have put-call delta parity*

$$\Delta_F(K,\sigma,+1) - \Delta_F(K,\sigma,-1) = 1.$$

Remark 6.2.4. In FX markets, to hedge a short option, we need to long futures with foreign currency notion $\Delta_F \times N$, where N is the foreign currency notional of the option.

Definition 6.2.5 (premium-adjusted spot delta). *[1]*

- *In the Black model, the sensitivity of the European option price (denominated in domestic currency) with respect to the forward rate $F(t,T)$ is given by(Lemma 6.5.7)*

$$\Delta_{S,pa}(K,\sigma,\phi) \triangleq \frac{\partial v/S_t}{\partial S_t} = \Delta_S - \frac{v}{S_t}(FOM\ to\ buy) = \phi\exp(-r_f\tau)\frac{K}{F(t,T)}N(\phi d_-),$$

where $\phi = 1$ for call, $\phi = -1$ for put,

$$v(K,\sigma,S_t,\phi) = \phi(e^{-r\tau}(F(t,T)N(d_+) - KN(d_-)),$$

and

$$d_\pm = \frac{1}{\sigma\sqrt{\tau}}[\log(\frac{F(t,T)}{K}) \pm \frac{1}{2}\sigma^2)\tau], \tau = T - t.$$

- *And we have put-call delta parity*

$$\Delta_{S,pa}(K,\sigma,+1)-\Delta_{S,pa}(K,\sigma,-1)=\exp(-r_f\tau)\frac{K}{f}.$$

Remark 6.2.5. Suppose the value of an option with a notional of 1,000 EUR was calculated as 74 EUR. Assuming a short position with a delta of 60% means, that buying 600 EUR is necessary to hedge. However the final hedge quantity will be 526 EUR which is the delta quantity reduced by the received premium in EUR. Consequently, the premium-adjusted delta would be 52.63%. T

Remark 6.2.6.

- In the un-adjusted spot delta, we assume the option is quoted in domestic currency and its payoff is in foreign currency.
- The definition of the premium-adjusted spot delta takes care of the correction induced by payment of the premium in foreign currency, which is the amount by which the delta hedge in foreign currency has to be corrected.

Definition 6.2.6 (premium-adjusted forward delta). *[1]*

- *In the black model, the sensitivity of the European option with respect to the forward rate $F(t,T)$ is given by(Lemma 6.5.7)*

$$\Delta_F(K,\sigma,\phi)\triangleq\frac{\partial v}{\partial F(t,T)}=\frac{\partial v}{\partial S_t}\frac{\partial S_t}{\partial F(t,T)}=\phi N(\phi d_+),$$

where $\phi=1$ for call, $\phi=-1$ for put,

$$v(K,\sigma,S_t,\phi)=\phi(e^{-r\tau}(F(t,T)N(d_+)-KN(d_-)),$$

and

$$d_\pm=\frac{1}{\sigma\sqrt{\tau}}[\log(\frac{F(t,T)}{K})\pm\frac{1}{2}\sigma^2)\tau],\tau=T-t.$$

- *And we have put-call delta parity*

$$\Delta_F(K,\sigma,+1)-\Delta_F(K,\sigma,-1)=1.$$

6.2.3.2 *conversion Delta to strike*

Methodology 6.2.1. *[1]*

- *From*
$$\Delta_F(K,\phi) = \phi N(\phi d_+) = \phi N(\phi \frac{\ln(F(t,T)/K) + \frac{1}{2}\sigma^2\tau}{\sigma\sqrt{\tau}}),$$
we can solve for the strike
$$K = F(t,T)\exp(-\phi\sigma\sqrt{\tau}N^{-1}(\phi\Delta_f) + \frac{1}{2}\sigma^2\tau).$$
- *From*
$$\Delta_{F,pa}(K,\phi) = \phi\frac{K}{F(t,T)}N(\phi d_-) = \phi\frac{K}{f}N(\phi \frac{\ln(F(t,T)/K) - \frac{1}{2}\sigma^2\tau}{\sigma\sqrt{\tau}}),$$
we need to numerically solve for the strike.

6.2.3.3 *Market convention*

Definition 6.2.7 (risk-reversal).

- ***ATM quote*** *is usually volatility of the call option with strike such that call spot/forward delta is equal to negative put spot/forward delta.*
- ***Risk reversal quote*** *is usually given by.*
$$R_{25} = \sigma_{C,25} - \sigma_{P,25}$$
that is for a given maturity, the 25 risk reversal is the vol of the 25 delta call less the vol of the 25 delta put. The 25 delta put is the put whose strike has been chosen such that the delta is -25%.
- ***Strangle quote*** *is usually given by*
$$\sigma_{25-S-Q} = \frac{\sigma_{25C} + \sigma_{25P}}{2} - \sigma_{ATM}.$$
- ***Market strangle quote*** *is usually given by.*
$$\sigma_{25-S-M} = \sigma_{ATM} + \sigma_{25-S-Q}.$$

Note 6.2.1 (market convention of quoting delta)**.** [2, p. 848]

- In the FX market, instead of quoting an option price by specifying its expiry, strike, and volatility, denoted by (T,K,σ), it is also common to quote its expiry, delta and volatility, denoted by (T,Δ,σ).
- For example, consider a six-month 25 delta call has a volatility of 13%.

- We can solve for the strike from delta formula

$$\exp(-D(T-t))N(d_1) = 0.25,$$

where

$$d_1 = \frac{\ln(S_t/K) + (r - D + \frac{1}{2}0.13^2)(T-t)}{0.13\sqrt{T-t}}.$$

Then we can solve for the strike K.

Maturity	$\cdots$	$35\Delta_P$	$40\Delta_P$	$45\Delta_P$	ATM	$45\Delta_C$	$40\Delta_C$	$35\Delta_C$	$\cdots$
1W	$\cdots$	15.5	15.3	15.1	15	15.2	15.4	15.6	$\cdots$
1M	$\cdots$	20.8	20.4	20.1	20	20.2	20.5	20.9	$\cdots$

Table 6.2.2: FX implied volatility for call or put options with different Δ and different time-to-maturities(1 week and 1 month). 45Δ

6.2.4 Dynamical model for spot rate

Unlike the dynamic models for other asset classes, e.g., equity, interest rate, commodity etc, the choice of dynamical model for FX spot rate has to possess some symmetric properties inherent to the FX market. For example, any dynamics model for spot rate S_t should guarantee that the dynamics of $1/S_t$ belong to the same type of model. Further, let S_1 stand for USD/JPY, S_2 stand for EUR/USD, S_3 stand for GBP/USD, then the dynamics of $S_4 \triangleq S_1 \cdot S_2$, representing EUR/JPY, and $S_5 = S_2/S_3$, representing EUR/GBP, should belong to the same type of model.

Lognormal types of model, including constan-volatility lognormal model and local volatility model, naturally satisfy these symmetric requirements. However, dynamical models, like normal model, shifted lognormal model, stochastic volatility model, do not satisfy these requirements and cannot be used to model FX spot rates.

6.3 Martingale pricing framework

6.3.1 Numeraires and martingales

Definition 6.3.1 (different numeraire). *Let M be the domestic money account, M^f be the foreign money account and $S(t)$ the exchange rate. Let $P^d(t,T), P^f(t,T)$ be zero-coupon bond prices in domestic currency and in foreign currencies such that*

$$P^d(t,T) = M(t)E_{Q^d}[\frac{1}{M(T)}|\mathcal{F}_t], P^f(t,T) = M^f(t)E_{Q^f}[\frac{1}{M^f(T)}|\mathcal{F}_t].$$

Then

- ***domestic (risk-neutral) measure*** *is the measure associated with doemstic money account numeraire $M(t)$. We denote expectation with respect to this measure as $E_{Q^d}[\cdot]$.*
- ***foreign (risk-neutral) measure*** *is the measure associated with foreign money account numeraire $S(t)M^f(t)$. We denote expectation with respect to this measure as $E_{Q^f}[\cdot]$.*
- ***domestic T-forward measure*** *is the measure associated with doemstic T maturity zero-coupon bond numeraire $P(t,T)$. We denote expectation with respect to this measure as $E_{T^d}[\cdot]$.*
- ***foreign T-forward measure*** *is the measure associated with foreign T maturity zero-coupon bond numeraire $S(t)P^f(t)$. We denote expectation with respect to this measure as $E_{T^f}[\cdot]$.*
- ***foreign-denomenated foreign risk-neutral measure*** *is the measure associated with foreign money account numeraire $M^f(t)$. We denote expectation with respect to this measure as $E_{Q^{f/f}}[\cdot]$.*
- ***foreign-denomenated foreign T-forward measure*** *is the measure associated with foreign T maturity zero-coupon bond numeraire $P^f(t)$. We denote expectation with respect to this measure as $E_{T^{f/f}}[\cdot]$.*

Proposition 6.3.1 (martingale properties associated with different measures). *Let X be an asset price denomiated in domestic currency. Let Y be an asset price denomiated in foreign currency. It follows that*

- *$\frac{X(t)}{M(t)}$ and $\frac{S(t)Y(t)}{M(t)}$ are martingales under domestic (risk-neutral) measure.*
- *$\frac{X(t)}{M^f(t)S(t)}$ and $\frac{S(t)Y(t)}{M^f(t)S(t)}$ are martingales under foreign (risk-neutral) measure.*
- *$\frac{X(t)}{P^d(t,T)}$ and $\frac{S(t)Y(t)}{P^d(t,T)}$ are martingales under domestic T forward measure.*
- *$\frac{X(t)}{P^f(t,T)S(t)}$ and $\frac{S(t)Y(t)}{S(t)P^f(t,T)}$ are martingales under foreign T forward measure.*
- *$\frac{X(t)/S(t)}{M^f(t)}$ and $\frac{Y(t)}{M^f(t)}$ are martingales under foreign-denomenated foreign (risk-neutral) measure.*
- *$\frac{X(t)/S(t)}{P^f(t,T)}$ and $\frac{Y(t)}{P^f(t,T)}$ are martingales under foreign-denomenated foreign T forward measure.*
- ***(numeriare invariance)*** *Let Z and W be two assets denominated in the same currency, then denominating in an arbitrary currency, $\frac{Z}{W}$ is a martingale under W numeriaire denominated in the same currency.*

Proof. See Theorem 1.6.15. ☐

Note 6.3.1 (equivalence of foreign (risk-neutral) measure and foreign-denominated foreign (risk-neutral) measure). From above, we can see that foreign (risk-neutral) measure and foreign-denominated foreign (risk-neutral) measure are equivalent; Under either of them, $\frac{X(t)}{M^f(t)S(t)}$ and $\frac{Y(t)}{M^f(t)}$ are martingales.

6.3.2 Change of numeraire

Lemma 6.3.1 (cross-country change of numeraire). *[3] Let M be the domestic money account, M^f be the foreign money account and S the exchange rate.*

- *The Radon-Nikodym derivative to change measure from domestic money account M to the foreign money account M^f is given by*

$$Z(T) = \frac{dQ^f}{dQ^d} = \frac{S(T)M^f(T)}{S(t)M^f(t)}\frac{M(t)}{M(T)}$$

- $$E_{Q^d}[Z_T|\mathcal{F}_t] = 1$$

- $$E_{Q^f}[X] = E_{Q^d}[X\frac{dQ^f}{dQ^d}]$$

Proof. (3) Note that $S(T)M^f(T)/M(T)$ is martingale under Q^d, we have

$$E_{Q^d}[Z_T|\mathcal{F}_t] = E_{Q^d}[\frac{S(T)M^f(T)}{S(t)M^f(t)}\frac{M(t)}{M(T)}|\mathcal{F}_t] = \frac{S(T)M^f(T)}{S(t)M^f(t)}\frac{M(t)}{M(T)} = 1.$$

☐

Proposition 6.3.2 (change of measure for dynamic models). *Let $X_1, X_2, ..., X_n$ be domestic assets having domestic risk-neutral dynamics*

$$dX_i/X_i = r^d dt + \sigma_{X,i} dW_i^d.$$

Let $Y_1, Y_2, ..., Y_n$ be foreign assets having foreign risk-neutral dynamics

$$dY_i/Y_i = r^f dt + \sigma_{Y,i} dZ_i^f.$$

It follows that

- *Under domestic measure* Q^d,

$$dS/S = (r^d - r^f)dt + \sigma_S dB^d,$$

where B^d *is the Brownian motion under* Q^d; *under foreign measure* Q^f,

$$d(1/S)/(1/S) = (r^f - r^d)dt - \sigma_S dB^f,$$

where B^d *is the Brownian motion under* Q^d;

- *Under domestic measure, the foreign asset dynamics is given by*

$$dY_i/Y_i = (r^f - \rho_{B,Z_i}\sigma_S\sigma_{Y,i})dt + \sigma_{Y,i}dZ_i^d.$$

- *Under foreign measure, the domestic asset dynamics is given by*

$$dX_i/X_i = (r^d + \rho_{B,Zi}\sigma_S\sigma_{Y,i})dt + \sigma_{Y,i}dZ_i^f,$$

where $dB^d dZ_i^f = \rho_{B,Z_i}dt$

- *Under domestic measure,*

$$d(SY_i)/(SY_i) = r^d dt + \sigma_{Y,i}dZ_i^d + \sigma_B dB^f.$$

- *Under foreign measure,*

$$d(X_i/S)/(X_i/S) = r^f dt + \sigma_{Y,i}dZ_i^d - \sigma_B dB^f.$$

Proof. Similar to the proof of Theorem 1.6.17. Let $N(t) = M^f(t)S(t)$ denote the foreign money account numeraire. Then

$$\begin{aligned}
E_{Q^f}[\frac{dX_t}{X_t}] &= E_{Q^d}[\frac{dX_t}{X_t}\frac{dQ^f}{dQ^d}] \\
&= E_{Q^d}[\frac{dX_t}{X_t}\frac{S(T)M^f(T)}{S(t)M^f(t)}\frac{M(t)}{M(T)}] \\
&= E_{Q^d}[\frac{dX_t}{X_t}(1 + d\ln(N(t)/M(t))] \\
&= E_{Q^d}[\frac{dX_t}{X_t}(1 + d\ln(N(t)/M(t))] \\
&= rdt + \rho_{XS}\sigma_X\sigma_S
\end{aligned}$$

We can similarly prove other facts.

(4) can be directly obtained from (1) and (2) using Ito process product rule. (5) can be directly obtained from (3) and (4) using Ito process product rule. □

Lemma 6.3.2 (cross-country change of numeraire for zero-coupon bonds). *[4, p. 700] Let $P^d(t,T), P^f(t,T)$ be zero-coupon bond prices in domestic currency and in foreign currencies such that*

$$P^d(t,T) = M(t)E_{Q^d}[\frac{1}{M(T)}|\mathcal{F}_t], P^f(t,T) = M^f(t)E_{Q^f}[\frac{1}{M^f(T)}|\mathcal{F}_t].$$

Let M be the domestic money account, M^f be the foreign money account and S the exchange rate.

- *The $S(T)P^f(t,T)$ is the foreign T-zero-coupon-bond numeraire. The Radon-Nikodym derivative to change measure from domestic T-zero-coupon-bond numeraire $P^d(t,T)$ to the foreign T-zero-coupon-bond numeraire $P^f(t,T)$ is given by*

$$Z(T) = \frac{dS_{T^d}}{dS_{T^f}} = \frac{S(T)P^f(T,T)}{S(t)P^f(t,T)}\frac{P^d(t,T)}{P^d(T,T)} = \frac{S(T)}{S(t)P^f(t,T)}P^d(t,T) = \frac{S(T)}{F(t,T)}$$

- $$1/Z(T) = \frac{dS_{T^f}}{dS_{T^d}} = \frac{F(t,T)}{S(T)}$$

- *(Todo!)*

$$E_{T^f}[Z_T|\mathcal{F}_t] = 1$$

- $$E_{T^f}[X] = E_{T^d}[X\frac{dS_{T^f}}{dS_{T^d}}]$$

- *Support $S(T)$ is a martingale under measure T^d and $\frac{dQ^f}{dQ^d}$ has log-normal process with volatility Σ_Z and correlation ρ with $S(t)$. Then*

$$E_{T^f}[S(T)|\mathcal{F}_t] = S(t)\exp(\rho\sigma_Z\sigma_S(T-t))$$

Proof. (1)(2) Use Theorem 1.6.15 . □

Lemma 6.3.3 (cross-country change of numeraire for zero-coupon bonds and money market account). *[4, p. 700] Let $P^d(t,T), P^f(t,T)$ be zero-coupon bond prices in domestic currency and in foreign currencies such that*

$$P^d(t,T) = M(t)E_{Q^d}[\frac{1}{M(T)}|\mathcal{F}_t], P^f(t,T) = M^f(t)E_{Q^f}[\frac{1}{M^f(T)}|\mathcal{F}_t].$$

Let M be the domestic money account, M^f be the foreign money account and S the exchange rate. The $S(T)P^f(t,T)$ is the foreign T-zero-coupon-bond numeraire.

It follows that

- $$M^f(t)E_{Q^f}[\frac{V(T)}{M^f(T)}|\mathcal{F}_t] = P^f(t,T)E_{T^f}[V(T)|\mathcal{F}_t].$$

- $$M(t)E_{Q^d}[\frac{V(T)}{M(T)}|\mathcal{F}_t] = P^d(t,T)E_{T^d}[V(T)|\mathcal{F}_t].$$

- *The Radon-Nikodym derivative to change measure from domestic T-zero-coupon-bond numeraire $P^d(t,T)$ to the foreign money market account numeraire $S(t)M^f(t)$is given by*

$$Z(T) = \frac{dS_{T_d}}{dQ^f} = \frac{P^d(T,T)}{P^d(t,T)}\frac{S(t)M^f(t)}{S(T)M^f(T)}.$$

- *The Radon-Nikodym derivative to change measure from foreign T-zero-coupon-bond numeraire $P^f(t,T)S(t)$ to the domestic money market account numeraire $M(t)$ is given by*

$$Z(T) = \frac{dS_{T_f}}{dQ^d} = \frac{S(T)P^f(T,T)}{S(t)P^f(t,T)}\frac{M(t)}{M(T)}.$$

Proof. The key theorem we are using is Theorem 1.6.15 (1) Note that

$$\frac{dS_{T^f}}{dQ^f} = \frac{S(T)P^f(T,T)}{S(t)P^f(t,T)}\frac{S(t)M^f(t)}{S(T)M^f(T)} = \frac{P^f(T,T)}{P^f(t,T)}\frac{M^f(t)}{M^f(T)},$$

then

$$\begin{aligned} M^f(t)E_{Q^f}[\frac{V(T)}{M^f(T)}|\mathcal{F}_t] &= M^f(t)E_{T^f}[\frac{V(T)}{M^f(T)}\frac{dQ^f}{dS_{T^f}}|\mathcal{F}_t] \\ &= M^f(t)E_{T^f}[\frac{V(T)}{M^f(T)}\frac{dQ^f}{dS_{T^f}}|\mathcal{F}_t] \\ &= M^f(t)E_{T^f}[\frac{V(T)}{M^f(T)}\frac{P^f(t,T)}{P^f(T,T)}\frac{M^f(T)}{M^f(t)}|\mathcal{F}_t] \\ &= P^f(t,T)E_{T^f}[\frac{V(T)}{P^f(T,T)}|\mathcal{F}_t] \\ &= P^f(t,T)E_{T^f}[V(T)|\mathcal{F}_t] \end{aligned}$$

(2) Similar to (1), note that

$$\frac{dS_{T^d}}{dQ^d} = \frac{P^d(T,T)}{P^d(t,T)}\frac{M^d(t)}{M^d(T)}.$$

(3)(4) straight forward. □

6.3.3 FX rate dynamics under different measure

6.3.3.1 *Two countries*

Proposition 6.3.3 (martingale property of spot rates under domestic measure). *Let current time be $t, t \leq T$.*

- $$S(t)\frac{M^f(t)}{M(t)} = E_{Q^d}[\frac{S(T)M^f(T)}{M(T)}|\mathcal{F}_t].$$
- *If the interest rates are constants given by r^d, r^f, we have*
$$E_{Q^d}[S(T)|\mathcal{F}_t] = \exp((r^d - r^f)(T-t))S(t).$$
- *If we assume $S(t)$ is governed by geometric SDE, then under domestic risk-neutral measure, we have representation*
$$dS(t)/S(t) = (r^d - r^f)dt + \sigma(t)dW^Q(t).$$

Proof. Under the domestic risk-neutral measure, the domestic value of the foreign money market account $M^f(T)S(T)$ will have martingale property

$$S(t)\frac{M^f(t)}{M(t)} = E_{Q^d}[\frac{S(T)M^f(T)}{M(T)}|\mathcal{F}_t].$$

□

Proposition 6.3.4 (martingale property of spot rates under domestic measure). *Let current time be $t, t \leq T$.*

- $$\frac{M(t)}{S(t)M^f(t)} = E_{Q^f}[\frac{M(T)}{M^f(T)S(T)}|\mathcal{F}_t].$$
- *If the interest rates are constants given by r^d, r^f, we have*
$$E_{Q^f}[\frac{1}{S(T)}|\mathcal{F}_t] = \exp((r^f - r^d)(T-t))\frac{1}{S(t)}.$$
- *If we assume $S(t)$ is governed by geometric SDE has representation under domestic risk-neutral measure given by*
$$dS(t)/S(t) = (r^d - r^f)dt + \sigma(t)dW^d(t).$$

If we assume $1/S(t)$ is governed by geometric SDE, then under foreign risk-neutral measure, we have representation

$$d(1/S(t))/(1/S(t)) = (r^f - r^d)dt - \sigma(t)dW^f(t).$$

Proof. (1) Under the domestic risk-neutral measure, the domestic value of the foreign money market account $M^f(T)S(T)$ will have martingale property

$$S(t)\frac{M^f(t)}{M(t)} = E_{Q^d}[\frac{S(T)M^f(T)}{M(T)}|\mathcal{F}_t].$$

(3) Using Ito quotient rule, we have

$$d(1/S(t))/(1/S(t)) = (r^f - r^d + \sigma^2(t))dt - \sigma(t)dW^d(t).$$

to do □

6.3.3.2 *Multiple countries*

Notation:

$$S^{A/B} \triangleq \textit{number of units of currency A required to purchase 1 unit of currency B.}$$

Lemma 6.3.4 (triangle relationship).

- $$X_t^{A/B} = X_t^{A/C} \times X_t^{C/B}.$$
- *If*

$$\begin{aligned} dX_t^{A/C}/X_t^{A/C} &= \mu^{A/C}dt + \sigma^{A/C}dW_t^{A/C} \\ dX_t^{C/B}/X_t^{C/B} &= \mu^{C/B}dt + \sigma^{C/B}dW_t^{C/B}, \end{aligned}$$

then

$$dX_t^{A/B}/X_t^{A/B} = (\mu^{A/C} + \mu^{C/B} + \rho\sigma^{A/C}\sigma^{C/B})dt + \sigma^{A/C}dW_t^{A/C} + \sigma^{C/B}dW_t^{C/B}$$

Proof. (1) Use no-arbitrage argument. (2) Use **??**. □

Remark 6.3.1 (interpretation)**.**

- In practice, exchange rates do not necessarily follow geometric Brownian motion; so above triangle relationship might not be useful.
- In theory, if geometric Brownian motion actually describes exchange rate dynamics, then two known exchange rate dynamics should fully determine the third.

6.3.4 Asset dynamics under different measure

6.3.4.1 *Foreign money money account under domestic measure*

Definition 6.3.2 (basic concepts).

- ***Domestic money account*** *denominated in domestic currency:*

$$M(t) = \exp(\int_0^t r(u)du).$$

- ***Foreign money account*** *denominated in foreign currency:*

$$M^f(t) = \exp(\int_0^t r^f(u)du).$$

- *The* ***exchange rate*** *$S(t)$ gives units on domestic currency per unit of foreign currency.*

$$dS(t) = \gamma(t)S(t)dt + \sigma_2(t)S(t)(\rho(t)dW_1(t) + \sqrt{1-\rho^2}dW_2),$$

where W_1 and W_2 are independent Brownian motions.
- *$M^f(t)S(t)$ gives the value of the foreign money account denominated in domestic currency.*

Lemma 6.3.5 (governing SDE for foreign money market account). *Assume exchange rate $S(t)$, under real-world measure, is given by*

$$dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t).$$

And let $r^f(t)$ be the deterministic foreign interest rate. It follows that

- *Then the dynamics for the foreign money market account under real-world measure is*

$$d(M^f(t)S(t))/(M^f(t)S(t)) = (r^f(t) + \mu(t))dt + \sigma(t)dW(t).$$

- *Then the dynamics for the foreign money market account under domestic risk-neutral measure is*

$$d(M^f(t)S(t))/(M^f(t)S(t)) = r^d(t)dt + \sigma(t)dW^Q(t).$$

Therefore, investing in foreign money account becomes risky due to the risk from exchange rate.

Proof. (1) Ito rule.

$$\begin{aligned} d(M^f(t)S(t)) &= SdM^f + M^f dS + dM^f dS \\ &= SM^f r dt + M^f(\gamma(t)S(t)dt + \sigma_2(t)S(t)dW(t)) + 0 \end{aligned}$$

(2) In the domestic risk-neutral measure, the geometric SDE for $S(t)$ is given by (Theorem 6.3.3)

$$dS(t)/S(t) = (r^d - r^f)dt + \sigma(t)dW^Q(t).$$

□

Remark 6.3.2 (understand exchange rate).

- If we view the foreign currency as an asset, then its current price is S, and its price process **under domestic risk-neutral measure** is given by

$$dS(t) = S(t)[(r(t) - r^f(t)dt + \sigma_2\rho(t)d\hat{W}_1(t) + \sigma_2(t)\sqrt{1-\rho^2(t)}d\hat{W}_2].$$

- Moreover, we view it as a dividend-paying asset. Because holding one unit of foreign currency can generate continuous dividends by investing in the foreign money market.

6.3.4.2 *Asset dynamics for two countries*

Proposition 6.3.5 (no-arbitrage governing SDE for asset dynamics under domestic measure). *Assume domestic assets $X_1^d, X_2^d, ..., X_M^d$, exchange rate S, domestic money account M, foreign money account M^f, and foreign asset $Y_1, Y_2, ..., Y_M$, in the domestic measure, follow the below dynamics*

$$\begin{aligned} dS_t/S_t &= (r^d - r^f)dt + \sigma_S dB^d \\ dX_i^d/X_i^d &= r^d dt + \sigma_{X,i} dW_1^d, i = 1,2,...,M \\ dY_i^f/Y_i^f &= r^f - \rho_{S,Y_i}\sigma_S\sigma_{Y,i}dt + \sigma_{Y,i}dZ_i^d, i = 1,2,....,N \\ dS_tY_i^f/S_tY_i^f &= r^d dt + \sigma_{Y,i}dZ_i^d + \sigma_S dB^d, i = 1,2,....,N \\ dM/M &= r^d dt \\ dM^f/M^f &= r^f dt \end{aligned}$$

where

- $dW_i^d dW_j^d = \sigma_{ij}^X dt$
- $dZ_i^d dZ_j^d = \sigma_{ij}^Y dt$
- W_i^d*s and* Z_i^d*s are independent*
- $dW_i^d dB^d = \rho_{X,S} dt$, $dZ_i^d dB^d = \rho_{Y,S} dt$.

It follows that

- $X_i^d(t)/M(t), i = 1,2,...,M, S(t)Y_i^f(t)/M(t), j = 1,...,N,$ $S(t)M^f(t)$ *are martingales under domestic risk-neutral measure.*
- $X_i^d(t)/S(t)M^f(t), i = 1,2,...,M, Y_i^f(t)/M^f(t), j = 1,...,N,$ $M(t)/S(t)$ *are martingales under foreign risk-neutral measure. And they adopt dynamics*

$$
\begin{aligned}
d(1/S_t)/(1/S_t) &= (r^f - r^d)dt - \sigma_S dB^f \\
dX_i^d/X_i^d &= (r^d + \rho_{S,X_i}\sigma_S\sigma_{X,i})dt + \sigma_{X,i}dW_1^f, i = 1,2,...,M \\
dY_i^f/Y_i^f &= r^f dt + \sigma_{Y,i}dZ_i^f, i = 1,2,....,N \\
dX_i^d/S_t/X_i^d/S_t &= r^f dt + \sigma_{X,i}dZ_i^f - \sigma_S dB^f, i = 1,2,....,N \\
dM/M &= r^d dt \\
dM^f/M^f &= r^f dt
\end{aligned}
$$

where
  - $dW_i^f dW_j^f = \sigma_{ij}^X dt$
  - $dZ_i^f dZ_j^f = \sigma_{ij}^Y dt$
  - W_i^f*s and* Z_i^f*s are independent*
  - $dW_i^f dB^f = \rho_{X,S}dt$, $dZ_i^f dB^f = \rho_{Y,S}dt$.
- *Therefore, for both domestic and foreign investors, there do not exist self-financing strategies that are arbitrages.*

Proof. □

6.3.4.3 *On the existence of risk-neutral measure*

Lemma 6.3.6 (existence of domestic and foreign measures in the geometric SDE framework). *Assume domestic asset* S_1*, exchange rate* S*, domestic money account* M*, foreign money account* M^f*(measured in foreign currency), and foreign asset* S_2*(measured in foreign currency), in the real-world measure, follow the below dynamics*

$$
\begin{aligned}
dS_1^d/S_1^d &= \mu_1 dt + \sigma_1 dW_1 \\
dS_2^f/S_2^f &= \mu_2 dt + \sigma_2 dW_2 \\
dX/X &= \gamma dt + \sigma_X dW_3 \\
dM/M &= rdt \\
dM^f/M^f &= r^f dt
\end{aligned}
$$

where $dW_i dW_j = \rho_{ij} dt$.

It follows that

- *The foreign assets* S_2^f, M^f *have value dynamics in domestic currency given by*

$$d(XS_2^f)/(XS_2^f) = (\mu_2 + \gamma + \rho_{23}\sigma_2\sigma_X)dt + \sigma_2 dW_2 + \sigma_X dW_3$$
$$d(XM^f)/(XM^f) = (r^f + \gamma)dt + \sigma_X dW_3$$

- *If there exists* $\theta_1, \theta_2, \theta_2$ *such that*

$$\sigma_1\theta_1 = \mu_1 - r$$
$$\sigma_2\theta_2 + \sigma_X\theta_3 = \mu_2 + \gamma - r + \rho_{23}\sigma_2\sigma_X$$
$$\sigma_X\theta_3 = r^f - r + \gamma$$

then there exist a domestic risk-neutral measure Q^d *such that under which*

$$\sigma_2 dW_2 = \sigma_2 dW_2^d - (r^f - r + \gamma + \rho_{23}\sigma_2\sigma_X)dt$$
$$\sigma_X dW_3 = \sigma_X dW_3^d - (\mu_2 - r^f)dt$$
$$\sigma_2 dW_2 + \sigma_X dW_3 = \sigma_2 dW_2^d + \sigma_X dW_3^d - (\mu_2 + \gamma + \rho_{23}\sigma_2\sigma_3 - r)dt$$

and

$$dS_1/S_1 = rdt + \sigma_1 dW_1^d$$
$$d(S_2)/(S_2) = (r^f - \rho_{23}\sigma_2\sigma_X)dt + \sigma_2 dW_2^d + \sigma_X dW_3^d$$
$$d(XS_2)/(XS_2) = rdt + \sigma_2 dW_2^d + \sigma_X dW_3^d$$
$$d(XM^f)/(XM^f) = rdt + \sigma_X dW_3^d$$
$$dS/S = (r - r^f)dt + \sigma_X dW_3^d$$
$$d(1/X)/(1/X) = (r^f - r + \sigma_X^2)dt - \sigma_X dW_3^d$$

- *In the foreign measure (the existence is guaranteed by Theorem 1.6.15), we have*

$$dW_1^f = dW_1^d - \rho_{13}\sigma_X dt$$
$$dW_2^f = dW_2^d - \rho_{23}\sigma_X dt$$
$$dW_3^f = dW_3^d - \sigma_X dt$$

and

$$\begin{aligned}
d(1/X)/(1/X) &= (r^f - r)dt - \sigma_X dW_3^f \\
dS_1/S_1 &= (r + \rho_{13}\sigma_1\sigma_X)dt + \sigma_1 dW_1^f \\
d(S_1/X)/(S_1/X) &= (r^f)dt + \sigma_1 dW_1^f \\
d(S_2)/(S_2) &= (r^f)dt + \sigma_2 dW_2^f \\
d(M^f)/(M^f) &= r^f dt \\
d(M/X)/(M/X) &= r^f dt - \sigma_X dW_3^f \\
dX/X &= (r - r^f + \sigma_X^2)dt + \sigma_X dW_3^f
\end{aligned}$$

- *Under domestic measure, assets denominated in domestic currency S_1, SS_2, SM^f, all have drift r; under foreign measure, assets denominated in foreign currency $S_1/S, S_2, M/S$, all have drift r^f.*

Proof. Use the Brownian motion under different measure(Theorem 1.6.17), we have

$$\begin{bmatrix} dW_1^f \\ dW_2^f \\ dW_3^f \end{bmatrix} = \begin{bmatrix} dW_1^d \\ dW_2^d \\ dW_3^d \end{bmatrix} + \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \sigma_S \end{bmatrix}.$$

□

Remark 6.3.3 (asset interpretation)**.**

- M/X is the value of domestic money account measured in foreign currency.
- S_1/X is the value of domestic asset measured in foreign currency.
- $M^f S$ is the value of foreign money account measured in domestic currency.
- $S_2 X$ is the value of foreign asset measured in domestic currency.

Remark 6.3.4 (inconsistence and arbitrage)**.** Note that if there are no solutions of θ_1, θ_2 due to inconsistence, then there are arbitrage opportunities(1.5.4).

Lemma 6.3.7 (existence of domestic and foreign measures in the geometric SDE framework, multiple assets). *Assume domestic asset $X_1, X_2, ..., X_M$, exchange rate S, domestic money* *account M, foreign money account M^f, and foreign asset $Y_1, Y_2, ..., Y_M$, in the real-world measure, follow the below dynamics*

$$\begin{aligned} dX_i/X_i &= \mu_{X,i}dt + \sigma_{X,i}dW_1, i = 1, 2, ..., M \\ dY_i/Y_i &= \mu_{Y,i}dt + \sigma_{Y,i}dZ_i, i = 1, 2, ..., N \\ dS/S &= \gamma dt + \sigma_S dB \\ dM/M &= rdt \\ dM^f/M^f &= r^f dt \end{aligned}$$

where $dW_i dW_j = \Sigma^X_{ij}dt$, $dZ_i dZ_j = \Sigma^Y_{ij}dt$.

It follows that

- $$\begin{aligned} d(SY_i)/(SY_i) &= (\mu_{Y,i} + \gamma + \rho_{SY_i}\sigma_{Y_i}\sigma_S)dt + \sigma_{Y_i}dZ_i + \sigma_S dB \\ d(SM^f)/(SM^f) &= (r^f + \gamma)dt + \sigma_S dB \end{aligned}$$

- *If there exists* $\theta_{X,i}, \theta_{Y,i}, \theta_S$ *such that*

  $$\begin{aligned} \sigma_1\theta_{X_i} &= \mu_{X_i} - r \\ \sigma_2\theta_{Y_i} + \sigma_S\theta_S &= \mu_{Y_i} + \gamma - r \\ \sigma_S\theta_S &= r^f - r + \gamma \end{aligned}$$

  then there exist a domestic risk-neutral measure Q^d *such that under which*

  $$\begin{aligned} \sigma_{Y,i}dZ_i &= \sigma_{Y,i}dZ_i^d - (r^f - r + \gamma)dt \\ \sigma_S dB &= \sigma_S dB^d - (\mu_2 - r^f)dt \\ \sigma_2 dW_2 + \sigma_S dW_3 &= \sigma_2 dW_2^d + \sigma_S dW_3^d - (\mu_2 + \gamma + \rho_{23}\sigma_2\sigma_3 - r)dt \end{aligned}$$

  and

  $$\begin{aligned} d(SS_2)/(SS_2) &= rdt + \sigma_2 dW_2^d + \sigma_S dW_3^d \\ d(SM^f)/(SM^f) &= rdt + \sigma_S dW_3^d \\ dS/S &= (r - r^f)dt + \sigma_S dW_3^d \end{aligned}$$

Remark 6.3.5 (induced drift when change of measure for correlated Brownian motion). **??**

6.3.5 Martingale pricing method

6.3.5.1 *General principle*

Proposition 6.3.6 (two equivalent methods to evaluate domestic value of foreign cash flow). *Consider a random foreign cash flow $V^f(T)$ at T. Let $V^f(t)$ denote its foreign value at time $t < T$. Let $V^d(t)$ denote its domestic value at time $t < T$. Then its t-value in domestic currency can be evaluated in the following two equivalent ways:*

- *(convert to domestic quote and then discount)*

$$V^d(t) = E_{Q^d}[\frac{M(t)S(T)V^f(T)}{M(T)}|\mathcal{F}_t].$$

- *(discount and then covert to domestic quote)*

$$\begin{aligned} V^d(t) &= S(t)V^f(t) \\ &= S(t)E_{Q^f}[\frac{M^f(t)V^f(T)}{M^f(T)}|\mathcal{F}_t] \end{aligned}$$

- *That is, we can either view $S(T)V^f(T)$ as a domestic asset or value $V^f(T)$ in foreign currency and then convert it to domestic value.*

Proof. Use Lemma 6.3.1 we can show that

$$\begin{aligned} E_{Q^f}[\frac{S(t)V^f(T)M^f(t)}{M^f(T)}|\mathcal{F}_t] &= E_{Q^d}[\frac{S(t)V(T)M^f(t)}{M^f(T)}\frac{dQ^f}{dQ^d}|\mathcal{F}_t] \\ &= E_{Q^d}[\frac{S(t)V(T)M^f(t)}{M^f(T)}\frac{S(T)M^f(T)}{S(t)M^f(t)}\frac{M(t)}{M(T)}|\mathcal{F}_t] \\ &= E_{Q^d}[\frac{M(t)S(T)V(T)}{M(T)}|\mathcal{F}_t] \end{aligned}$$

□

Proposition 6.3.7 (martingale pricing in foreign exchange models). *Consider an asset with payoff $V^d(T)$ at maturity T, denominated by domestic currency. Assume the existence of risk-neutral measure Q^d and Q^f. Let $V^d(t)$ denote its price denominated in domestic value at time $t < T$. It follows that*

- $$\frac{V^d(t)}{M(t)} = E_{Q^d}[\frac{V^d(T)}{M(T)}|\mathcal{F}_t].$$

- $$\frac{V^d(t)}{M^f(t)S(t)} = E_{Q^f}[\frac{V(T)}{M^f(T)Q^d(T)}|\mathcal{F}_t].$$
- *Consider a random foreign cash flow $V^f(T)$ at T. Let $V^f(t)$ denote its price denominated in foreign currency at time $t < T$. We have*

$$V^f(t) = E_{Q^f}[\frac{M^f(t)V^f(T)}{M^f(T)}|\mathcal{F}_t].$$

and

$$V^d(t) = V^f(t)S(t).$$

Proof. Directly from Theorem 1.6.15, Theorem 6.3.2, and Theorem 6.3.6. Note that (3) is directly implied by (2) via

$$\begin{aligned}\frac{V^f(t)S(t)}{M^f(t)S(t)} &= E_{Q^f}[\frac{V^d(T)S(T)}{M^f(T)S(T)}|\mathcal{F}_t]\\ \frac{V^f(t)}{M^f(t)} &= E_{Q^f}[\frac{V^d(T)}{M^f(T)}|\mathcal{F}_t]\\ V^f(t) &= M^f(t)E_{Q^f}[\frac{V^f(T)}{M^f(T)}|\mathcal{F}_t]\end{aligned}$$

□

Proposition 6.3.8 (martingale pricing in foreign exchange models with stochastic interest rate). *Consider an asset with payoff $V(T)$ at maturity T, denominated by domestic currency. Assume the existence of risk-neutral measure Q^d and Q^f. It follows that*

- $$V(t) = P^d(t,T)E_{T^d}[V(T)|\mathcal{F}_t].$$
- $$V(t) = S(t)P^d(t,T)E_{T^f}[\frac{V(T)}{S(T)}|\mathcal{F}_t].$$
- *Consider a random foreign cash flow $V^f(T)$ at T. Let $V^f(t)$ denote its foreign value at time $t < T$. Let $V^d(t)$ denote its domestic value at time $t < T$. We have*

$$\begin{aligned}V^d(t) = S(t)P^f(t,T)E_{T^f}[V^f(T)|\mathcal{F}_t] &= S(t)V^f(t)\\ V^f(t) &= P^f(t,T)E_{T^f}[V^f(T)|\mathcal{F}_t]\end{aligned}$$

Proof. (1) To change the domestic risk-netural measure to domestic forward measure, we have
(1) To change the foreign risk-netural measure to foreign forward measure, we have

□

6.4 Cross-country interest rate modeling

6.4.1 Basic concepts

Definition 6.4.1 (zero coupon bond).

- $P^d(t,T)$ *value at* t*, denominated in domestic currency, of a zero-coupon bond paying 1 unit domestic currency at maturity* T.
- $P^f(t,T)$ *value at* t*, denominated in foreign currency, of a zero-coupon bond paying 1 unit foreign currency at maturity* T.

Lemma 6.4.1 (zero-coupon bond dynamics under domestic and foreign measure).

- $$P^d(t,T) = M(t)E_{Q^d}[\frac{1}{M(T)}|\mathcal{F}_t].$$
  $$P^f(t,T) = M^f(t)E_{Q^f}[\frac{1}{M^f(T)}|\mathcal{F}_t].$$
  $$P^f(t,T) = S(t)M^f(t)E_{Q^d}[\frac{1}{M^f(T)S(T)}|\mathcal{F}_t].$$
- *Under domestic measure* Q^d,
  $$dP^d(t,T)/P^d(t,T) = r^d dt + \sigma_d dW_t^d$$
  $$dP^f(t,T)/P^f(t,T) = (r^f - \rho_{Sd}\sigma_S\sigma_d)dt + \sigma_d dY_t^d$$
- *Under foreign measure* Q^f,
  $$dP^d(t,T)/P^d(t,T) = rdt + \sigma_d dW_t^d$$
  $$dP^f(t,T)/P^f(t,T) = r^f dt + \sigma_d dY_t^f$$

Proof. See Theorem 6.3.2. □

6.4.2 Cross-country short rate dynamics

6.4.2.1 *Principles*

Proposition 6.4.1 (foreign short rate dynamics under domestic measure). [5] *Let the foreign short rate model, under foreign measure, have dynamics given by*

$$dr_t^f = mdt + \sigma_r dW_r^f,$$

such that the foreign zero coupon bond price dynamics is given by (Theorem 3.5.2)

$$dP^f(t,T)/P^f(t,T) = r^f dt + \sigma_r \frac{\partial P^f}{dr^f} dW_r^f.$$

It follows that

- *Under the domestic risk-neutral measure,*

$$dP^f(t,T)/P^f(t,T) = (r^f - \rho\sigma_r \frac{\partial P^f}{dr^f}\sigma_S)dt + \sigma_r \frac{\partial P^f}{dr^f} dW_t^d,$$

  where W_t^d is a Brownian motion under risk-neutral measure, and under domestic risk-neutral measure,

$$dW_r^f = dW_r^d - \rho\sigma_S.$$

- *Under the domestic risk-neutral measure,*

$$dr_t^f = (m - \rho\sigma_r\sigma_S)dt + \sigma_r dW_r^d,$$

Proof. (1) Under the domestic risk-neutral measure, the FX rate has representation

$$dS(t)/S(t) = (r^d - r^f)dt + \sigma_S dW_S^d.$$

Then

$$d(P^f(t,T)S(t)/M^d(t))/(P^f(t,T)S(t)/M^d(t)) = \sigma_r \frac{\partial P^f}{dr^f} dW_r^f + \sigma_S dW_S^d + \rho\sigma_r \frac{\partial P^f}{dr^f}\sigma_S dt,$$

In order for $P^f(t,T)S(t)/M^d(t)$ to be a martingale under domestic risk-neutral measure, we need to have

$$dW_r^f = dW_r^d - \rho\sigma_S dt.$$

(we can also derive this using Theorem 1.6.17.)

(2) directly from (1). □

6.4.2.2 *Hull-White model*

Lemma 6.4.2 (cross-currency Hull-White model under domestic measure). *Let there be $N+1$ countries. Assume the domestic short rate dynamics, under the domestic risk-neutral measure Q^d, is given by Hull-White model:*

$$dr_0 = (\theta_0(t) - k_0(t)r_0(t))dt + \sigma_0(t)dW_0(t),$$

where the subscript 0 denote domestic and

$$\begin{aligned}
\theta_0(t) &= \frac{\partial f_0(0,t)}{\partial t} + \kappa_0(t)f_0(0,t) + \phi_0(t) \\
\phi_0(t) &= \int_0^t \sigma_0^2(u)\exp(-2\beta_0(u,t)dt)du \\
\beta_0(t,T) &= \int_t^T k_0(u)du
\end{aligned}$$

Similarly, let the short rate model in these N countries, under their associated foreign measure, have dynamics given by

$$dr_t^{f,i} = (\theta_i - k_i(t)r_i(t))dt + \sigma_i dW_i(t).$$

Assume the exchange rate $S_i(t)$ of the ith foreign currency with respect to domestic currency, under the domestic risk-neutral measure Q^d, is given by

$$dS_i/S_i = (r_0(t) - r_i(t))dt + \xi_i(t)dZ_i, i = 1, 2, ..., N.$$

*Then the short rate dynamics in those foreign countries under **domestic measure** are given by*

$$dr_t^{f,i} = (\theta_i - k_i(t)r_i(t) - \rho_{Si,ri}\sigma_i(t)\xi_i(t))dt + \sigma_i dW_i(t), i = 1, 2, ..., N,$$

where $E[dW_i(t)dZ_i] = \rho_{Si,ri}dt$.

Proof. See Theorem 6.4.1. □

6.4.3 Cross-country HJM framework

Lemma 6.4.3 (foreign forward rate dynamics under domestic measure). *Let the foreign forward rate model, under foreign measure, have dynamics given by*

$$df^f(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t),$$

where $W(t)$ is a Brownian motion and

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,u)du.$$

The foreign zero coupon bond price dynamics is given by

$$dP^f(t,T)/P^f(t,T) = r^f(t)dt + \Sigma(t,T)dW(t)$$

with

$$\Sigma(t,T) = -\int_t^T \sigma(t,u)du.$$

It follows that

- *Under the domestic measure,*

$$dP^f(t,T)/P^f(t,T) = (r^f - \rho\sigma_S\Sigma(t,T))dt + \Sigma(t,T)dW_t^d.$$

- *Under the domestic measure,*

$$df_t^f = (\alpha(t,T) - \rho\sigma(t,T)\sigma_S)dt + \sigma(t,T)dW_t^d,$$

Proof. See the change of measure for dynamic models(Theorem 6.3.2). ☐

6.5 Pricing examples

6.5.1 FX forward contract and FX forward rate

Definition 6.5.1 (FX forward contract and FX forward rate).

- ***FX forward contracts*** *are transactions in which two party agree to exchange a specified amount of different currencies at some future date, with the exchange rate K being set at the time the contract is entered into.*

- *The payoff denominated in domestic currency at the delivery date T for a domestic investor purchasing foreign currency via a forward contract is given by*

$$(\underbrace{N \cdot S(T)}_{\text{domestic value of N foreign currency}} - \underbrace{K \cdot N}_{\text{demestic currency paid}}).$$

- *The fixed rate K such that forward contract has zero value is called forward FX rate.*
- *Forward FX rate can be viewed as the forward price of 1 unit foreign currency.*

Remark 6.5.1 (symmetric relationship on forwards).

- To long CCY1/CCY2 forward contract (buy CCY1 by selling CCY2) is equivalent to short CCY2/CCY1 (sell CCY2 and receive CCY1) forward contract.
- To long CCY2/CCY1 forward contract is equivalent to short CCY1/CCY2 forward contract.

Proposition 6.5.1 (no arbitrage constraint on forward FX rate). *Let current time be t and let the forward delivery date be T. Let $S(t)$ be the spot rate and $F(t,T)$ be the FX forward rate. Denote zero-coupon bond prices in domestic currency and in foreign currency are given by*

$$P^d(t,T) = M(t)E_{Q^d}[\frac{1}{M(T)}|\mathcal{F}_t], P^f(t,T) = M^f(t)E_{Q^f}[\frac{1}{M^f(T)}|\mathcal{F}_t].$$

- *The no-arbitrage condition requires that*

$$F(t,T) = S(t)\frac{P^f(t,T)}{P^d(t,T)}.$$

- *(martingale property) The FX forward rate $F(t,T)$ is a martingale under the martingale measure associated with domestic zero-coupon bond $P^d(t,T)$. And it satisfy*

$$F(t',T) = E_{T^d}[S(T)|\mathcal{F}_{t'}] = E_{T^d}[F(T,T)|\mathcal{F}_{t'}],$$

and

$$E[F(t',T)|\mathcal{F}_t] = E[S(T)|\mathcal{F}_t], t \leq t' < T.$$

- *Assume constant interest rate.*[a] *Under domestic risk-neutral measure,*

$$F(t,T) = E_{Q^d}[S(T)|\mathcal{F}_t] = e^{(r+r^f)(T-t)}S(t),$$

and $F(t,T)$ is a martingale.

[a] the case of deterministic interest rate can be similarly derived.

Proof. (1)(a) (no-arbitrage replication method) We can replicate the forward contract payoff in the following way:

- At time t, borrow $X(t)P^f(t,T)$ amount of domestic currency and buy $1/P^f(t,T)$ unit of foreign currrency; the $1/P^f(t,T)$ unit of foreign currency will buy 1 unit of foreign zero-coupon bond.
- At time T, get 1 unit of foreign currency from foreign money market; at the same time pay for the principle and the interest cost in total $X(t)P^f(t,T)/P^d(t,T)$.

Therefore, the forward FX rate (i.e., forward price of 1 unit foreign currency) is given by

$$F(t,T) = X(t)P^f(t,T)/P^d(t,T)$$

. (b) (martingale method)

$$\begin{aligned} & M(t)E_{Q^d}[(S(T)-F(t,T))/M(T)|\mathcal{F}_t] = 0 \\ \implies & M(t)E_{Q^d}[S(T)/M(T)|\mathcal{F}_t] = M(t)F(t,T)E_{Q^d}[1/M(T)] = F(t,T)P^d(t,T) \end{aligned}$$

Use change of measure(Theorem 1.6.15), we have

$$\begin{aligned} E_{Q^d}[\frac{M(t)S(T)}{M(T)}|\mathcal{F}_t] &= E_{Q^f}[\frac{M(t)S(T)}{M(T)}\frac{dQ^d}{dQ^f}|\mathcal{F}_t] \\ &= E_{Q^f}[\frac{M(t)S(T)}{M(T)}\frac{M(T)}{M(t)}\frac{M^f(t)S(t)}{M^f(T)S(T)}|\mathcal{F}_t] \\ &= E_{Q^f}[\frac{M^f(t)S(t)}{M^f(T)}|\mathcal{F}_t] \\ &= S(t)E_{Q^f}[\frac{M^f(t)}{M^f(T)}|\mathcal{F}_t] \\ &= S(t)P^f(t,T) \end{aligned}$$

(2)

$$\begin{aligned} E_{Q^d}[M(t)(S(T)-F(t,T))/M(T)|\mathcal{F}_t] &= 0 \\ P^d(t,T)E_{T^d}[S(T)-F(t,T)|\mathcal{F}_t] &= 0 \\ \implies F(t,T) &= E_{T^d}[S(T)|\mathcal{F}_t]. \end{aligned}$$

(3) Use the result from Theorem 6.3.3. (1)

$$E_{Q^d}[e^{-rT}(S(T)-F)|\mathcal{F}_t] = 0 \implies E_{Q^d}[S(T)|\mathcal{F}_t] = F(t,T)$$

Then we use Theorem 6.3.3 to get

$$F(t,T) = e^{(r+r^f)(T-t)}S(t)$$

□

Remark 6.5.2 (from the perspective of foreign investors).

- Under foreign risk-neutral measure, the forward price F at time t (measured in foreign currency) of one unit of domestic currency, to be delivered at time T, is determined by

  $$E_{Q^f}[e^{r^f T}(\frac{1}{S(T)} - F^f(t,T)|\mathcal{F}_t] = 0,$$

  or equivalently,

  $$F^f(t,T) = E_{Q^f}[\frac{1}{S(T)}|\mathcal{F}_t] = e^{(r^f-r)(T-t)}\frac{1}{S(t)},$$

  where $F^f(t,T)$ is a martingale under foreign risk-neutral measure.
- forward rate is martingale under forward measure

Note 6.5.1 (triangle relationship between forward exchange, domestic zero-curve, and foreign zero-curve). Note that $F(t,T), P^d(t,T), P^f(t,T), S(t)$ are not independent. Their relation is given by

$$F(t,T) = S(t)\frac{P^f(t,T)}{P^d(t,T)}.$$

Lemma 6.5.1 (value of a forward contract). *Let current time be t. Consider a unit notional forward contract starting at $t_0 \leq t$, with forward exchange rate $F(t_0,T)$ and delivery date T. Its value(long position) denominated in domestic currency at time t is given by*

$$V(t) = P^d(t,T)(F(t,T) - F(t_0,T)) = S(t)P^d(t,T) - P^d(t,T)F(t_0,T).$$

Proof.

$$\begin{aligned} V(t) &= P^d(t,T)E_{T^d}[(S(T) - F(t_0,T))|\mathcal{F}_t] \\ &= P^d(t,T)E_{T^d}[(F(T,T) - F(t_0,T))|\mathcal{F}_t] \\ &= P^d(t,T)(F(t,T) - F(t_0,T)) \end{aligned}$$

where we used the fact that forward rate is a martingale under forward measureTheorem 6.5.1.

□

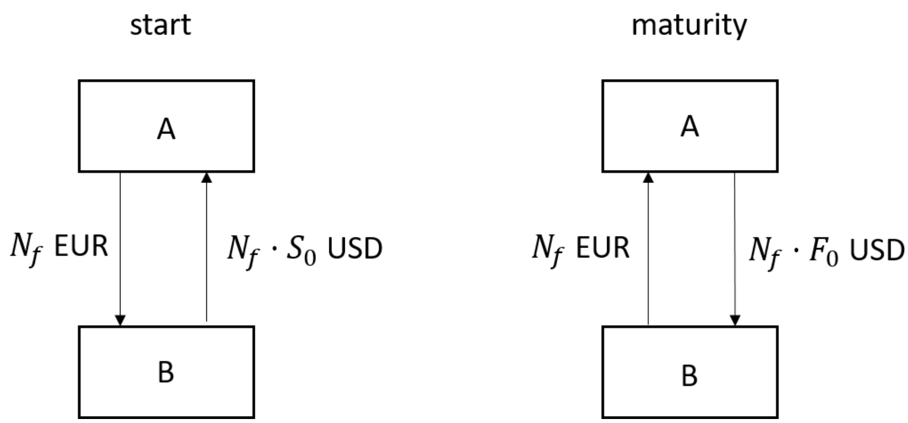


Figure 6.5.1: An illustration of FX swap. At the begining, two parties, A and B, will exchange equivalent value of currencies. At the maturity of the FX swap, the two parties will exchange the currencies with amount set by the FX forward rate.

6.5.2 Value of cash flows

Lemma 6.5.2 (domestic discounted value of fixed foreign cash flow). *Consider fixed foreign cash flow at future times $T_1, T_2, ..., T_n$ given by $V_1, V_2, ..., V_n$. Then at $t < T_1$ the domestic value is given by*

$$V(t) = S(t)\sum_{i=1}^{n} P^f(t, T_i)V_i = V_i F(t, T_i)P^d(t, T_i),$$

where $F(t, T_i)$ is the forward exchange rate with maturity date T_i.

Proof. Note that we use the triangle relationship(Theorem 6.5.1)

$$F(t, T) = S(t)\frac{P^f(t, T)}{P^d(t, T)}.$$

□

Lemma 6.5.3 (domestic value random foreign cash flow). *Consider random foreign cash flow at future times $T_1, T_2, ..., T_n$ given by $V^f(T_1), V^f(T_2), ..., V^f(T_n)$. Then at $t < T_1$ the domestic value is given by*

$$V^d(t) = S(t)\sum_{i=1}^{n} V_i^f(t),$$

where

$$V_i^f(t) = M^f(t)E_{Q^f}[\frac{V(T_i)}{M^f(T)}|\mathcal{F}_t].$$

Other equivalent formulation includes:

- $$V^d(t) = \sum_{i=1}^{n} P^d(t, T_i) \sum_{i=1}^{n} E_{T_i^d}[S(T_i)V(T_i)|\mathcal{F}_t].$$
- $$V^d(t) = M(t) \sum_{i=1}^{n} E_{Q^d}[S(T_i)V(T_i)/M(T_i)|\mathcal{F}_t].$$

Proof. (1) Note that $V_i^f(t) = M^f(t)E_{Q^f}[\frac{V(T_i)}{M^f(T)}|\mathcal{F}_t]$ is the present value dominated in foreign currency for foreign cash flow $V(T_i)$ based on the martingale pricing method(Theorem 6.3.7). (2)(3) Directly from the equivalent approach for evaluating domestic values of foreign cash flow. □

6.5.3 Cross currency swaps

Definition 6.5.2 (fixed-fixed cross-currency swap). *[6, p. 210]Given a set of dates $T_0, T_1, ..., T_m$ and denote $\tau_i = T_i - T_{i-1}$. A typical fixed-fixed cross currency swap between two counterparties domestic A and foreign B involves three phases:*

- *Initial exchange:*
  - *At T_0, B pays N^f in foreign currency to A.*
  - *At T_0, A pays N in domestic currency to B.*
- *Running period:*
  - *At $T_i, i = 1, 2, ..., M$, B pays fixed $N\tau_i K_i$ in domestic currency to A.*
  - *At $T_i, i = 1, 2, ..., M$, A pays fixed $N^f \tau_i K_i^f$ in foreign currency to B.*
- *Final exchange:*
  - *At T_M, B pays fixed N in domestic currency to A.*
  - *At T_M, A pays fixed N^f in foreign currency to B.*

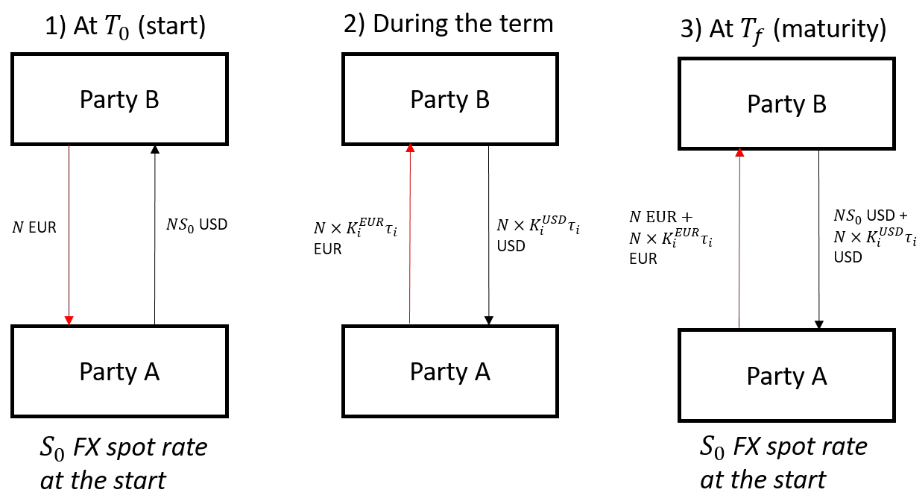


Figure 6.5.2: Cross currency fixed-fixed swap.

Remark 6.5.3 (application of cross currency swap). link

- A cross currency swap allows two parties to borrow foreign currency by issuing domestic bonds (borrowing foreign currency sometimes can be more difficult for companies than borrowing domestic money.)
- Suppose the British Petroleum Company needs $150 million US dollars to finance its new refining facility in the U.S. Also, suppose that the Piper Shoe Company, a US company, plans to issue 100 million pounds to set up its distribution center in London. To meet each other's needs, suppose that both companies go to a swap bank that sets up the following agreements:
- **Initial stage:** The British Petroleum Company will issue 5-year 100 million pound bonds with 7.5% interest rate. It will then deliver the 100 million pound to the swap bank who will pass it on to the U.S. Piper Company to finance the construction of its British distribution center. The Piper Company will issue 5-year 150 million dollar bonds paying 10% interest. The Piper Company will then pass the 150 million dollars to swap bank that will pass it on to the British Petroleum Company who will use the funds to finance the construction of its U.S. refinery.
- **Running periods:** The British company, with its U.S. asset (refinery), will pay the 10% interest on $150 million to the swap bank who will pass it on to the American company so it can pay its U.S. bondholders. The American company, with its British asset , will pay the 7.5% interest on 100 million pounds, to the swap bank who will pass it on to the British company so it can pay its British bondholders.
- **Final exchange:** At maturity, the British company will pay $150 million to the swap bank who will pass it on to the American company so it can pay its U.S. bondholders. At

maturity, the American company will pay 100 million pound to the swap bank who will pass it on to the British company so it can pay its British bondholders.

Lemma 6.5.4 (valuing fixed-fixed cross-currency swap). *Consider a fixed-fixed cross-currency swap where two parties domestic A and foreign B exchange cash flow at a set of dates $T_0, T_1, ..., T_m$(denote $\tau_i = T_i - T_{i-1}$). Then from the perspective of A, we have*

- *the initial exchange has value denominated in domestic currency*

$$V_{init}(t) = \underbrace{-NP^d(t, T_0)}_{pay\ to\ B} + \underbrace{N^f P^f(t, T_0)S(t)}_{receive\ from\ B}.$$

- *the running period has value denominated in domestic currency*

$$V_{run}(t) = \underbrace{-\sum_{i=1}^{M} N^f \tau_i K_i P^f(t, T_i)S(t)}_{pay\ to\ B} + \underbrace{\sum_{i=1}^{M} N\tau_i K_i P^d(t, T_i)}_{receive\ from\ B}.$$

- *the final exchange has value denominated in domestic currency*

$$V_{final}(t) = \underbrace{-N^f P^f(t, T_M)S(t)}_{pay\ to\ B} + \underbrace{NP^d(t, T_M)}_{receive\ from\ B}.$$

- *the current value of the swap at time t is*

$$\begin{aligned}
V(t) &= V_{init}(t) + V_{run}(t) + V_{final}(t) \\
&= N(-P^d(t, T_0) + \sum_{i=1}^{M} N\tau_i K_i P^d(t, T_i) + P^d(t, T_M)) \\
&+ N^f S(t)(-P^f(t, T_0) + \sum_{i=1}^{M} N\tau_i K_i P^f(t, T_i) + P^f(t, T_M))
\end{aligned}$$

Proof. Straight forward application of cash discounting(Lemma 6.5.2). □

Remark 6.5.4 (other equivalent evaluation method). Note that under no-arbitrage condition, there are always two equivalent ways to evaluate the domestic value of foreign cash flow(Theorem 6.3.6). For example,

- For the initial exchange, we can also write

$$V_{init}(t) = -NP^d(t, T_0) + N^f P^f(t, T_0)S(t) = -NP(t, T_0) + N^f F(t, T_0)P^d(t, T_0).$$

where $F(t,T)$ is the forward exchange rate and we have used the triangle relation

$$F(t,T) = S(t)\frac{P^f(t,T)}{P^d(t,T)}.$$

- Similarly, for the running period, we can also write

$$\begin{aligned} V_{run}(t) &= -\sum_{i=1}^{M} N^f \tau_i K_i P^f(t,T_i)S(t) + \sum_{i=1}^{M} N\tau_i K_i P(t,T_0) \\ &= -\sum_{i=1}^{M} N^f \tau_i K_i P^d(t,T_i)F(t,T_i) + \sum_{i=1}^{M} N\tau_i K_i P(t,T_0) \end{aligned}$$

6.5.4 Cross-currency basis swaps

Definition 6.5.3 (cross-currency basis swap). *A cross-currency basis swap is a floating/floating swap where two parties borrow from and simultaneously lend to each other an equivalent amount of money denominated in two different currencies for a predefined period of time. Given a set of dates* $T_0, T_1, ..., T_m$ *and denote* $\tau_i = T_i - T_{i-1}$. *A typical float-float cross currency basis swap between two counterparties A and B involves three phases:*

- *Initial exchange:*
  - *At* T_0, *B pays* $N^f = N/S_0$ *in foreign currency to A, where* S_0 *is FX spot rate at* T_0.
  - *At* T_0, *A pays N in domestic currency to B.*
- *Running period:*
  - *At* $T_i, i = 1,2,...,M$, *B pays fixed* $N\tau_i L^d(T_{i-1},T_i)$ *in domestic currency to A.*
  - *At* $T_i, i = 1,2,...,M$, *A pays fixed* $N^f\tau_i L^f(T_{i-1},T_i)$ *in foreign currency to B.*
- *Final exchange:*
  - *At* T_M, *B pays fixed N in domestic currency to B.*
  - *At* T_M, *A pays fixed* N^f *in foreign currency to B.*

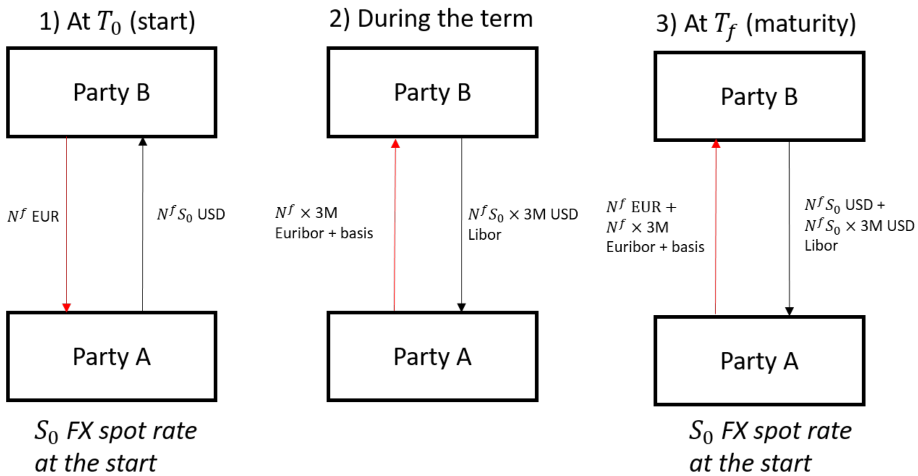


Figure 6.5.3: Cross currency basis swap.

Remark 6.5.5 (understand the cash flow in cross-currency basis swap). Let A be the domestic party. The cash flows from A's point of view are

- For undiscounted domestic currency leg(cash flow in domestic currency)
  - At T_0, $C_d(T_0) = -S(T_0)N_f$
  - At $T_i, i = 1, 2, ..., M$, $C_d(T_i) = L^d(T_{i-1}, T_i)\tau_i S(T_0)N^f$.
  - At T_M, $C_d(T_M) = S(T_0)N^f$.
- For undiscounted foreign currency leg
  - At T_0, $C_f(T_0) = N_f$
  - At $T_i, i = 1, 2, ..., M$, $C_f(T_i) = -L^f(T_{i-1}, T_i)\tau_i N^f$
  - At T_M, $C_f(T_M) = -N_f$

Lemma 6.5.5 (valuing float-float cross-currency swap). *Consider a fixed-fixed cross-currency swap where two parties domestic A and foreign B exchange cash flow at a set of dates $T_0, T_1, ..., T_M$(denote $\tau_i = T_i - T_{i-1}$). Then from the perspective of A, we have*

- *the initial exchange has value denominated in domestic currency*

$$V_{init}(t) = \underbrace{-NP^d(t, T_0)}_{pay\ to\ B} + \underbrace{N^f P^f(t, T_0)S(t)}_{receive\ from\ B}.$$

- *the running period has value denominated in domestic currency*

$$V_{run}(t) = \underbrace{-\sum_{i=1}^{M} N^f S(t)P^f(t,T_i)\tau_i E_{T^f}[L^f(T_{i-1},T_i)|\mathcal{F}_t]}_{pay\ to\ B} + \underbrace{\sum_{i=1}^{M} N\tau_i P^d(t,T)E_{T^d}[L^d(T_{i-1},T_i)]}_{receive\ from\ B}$$

$$= -\sum_{i=1}^{M} N^f S(t)P^f(t,T_i)\tau_i F^f(t,T_{i-1},T_i) + \sum_{i=1}^{M} N\tau_i P^d(t,T)F^d(,tT_{i-1},T_i)$$

- *the final exchange has value denominated in domestic currency*

$$V_{final}(t) = \underbrace{-N^f P^f(t,T_M)S(t)}_{pay\ to\ B} + \underbrace{NP^d(t,T_M)}_{receive\ from\ B}.$$

- *the current value of the swap at time t is*

$$\begin{aligned} V(t) &= V_{init}(t) + V_{run}(t) + V_{final}(t) \\ &= \underbrace{N(-P^d(t,T_0) + \sum_{i=1}^{M} N\tau_i F^d(t,T_{i-1},T_i)P^d(t,T_i) + P^d(t,T_M))}_{domestic\ leg} \\ &+ \underbrace{N^f S(t)(-P^f(t,T_0) + \sum_{i=1}^{M} N\tau_i F^f(t,T_{i-1},T_i)P^f(t,T_i) + P^f(t,T_M))}_{foreign\ leg} \end{aligned}$$

Proof. Straight forward application of cash discounting(Lemma 6.5.2). □

Definition 6.5.4 (mark-to-market(MtM) cross-currency basis swap). *A mark-to-market(MtM) cross-currency basis swap is a floating/floating swap where two parties borrow from and simultaneously lend to each other an equivalent amount of money denominated in two different currencies for a predefined period of time.*

The major currency has a notional resettable leg, while the minor currency

Given a set of dates $T_0, T_1, ..., T_m$ and denote $\tau_i = T_i - T_{i-1}$. Assume the domenstic currency has a notional resettable leg, a typical MtM cross currency swap between two counterparties foreign A and domestic B involves three phases:

- *Initial exchange:*
  - *At T_0, B pays N^f in foreign currency to A,*
  - *At T_0, A pays $N^f \times S_0$ in domestic currency to B, where S_0 is FX spot rate at T_0.*
- *Running period:*

  - *At $T_i, i = 1, 2, ..., M$, B pays $N^f S_{i-1}\tau_i L_{dom}(T_i, T_{i+1})$ in domestic currency to A. B pays $N^f \times S_{i-1}$ domesitc currency to A, and B receives $N^f \times S_i$ domesitc currency from B.*
  - *At $T_i, i = 1, 2, ..., M$, A pays fixed $N\tau_i L_{for}(T_i, T_{i+1})$ in foreign currency to B.*
- *Final exchange:*
  - *At T_M, B pays $N^f \times S_{M-1}$ in domestic currency to A.*
  - *At T_M, A pays fixed N^f in foreign currency to B.*

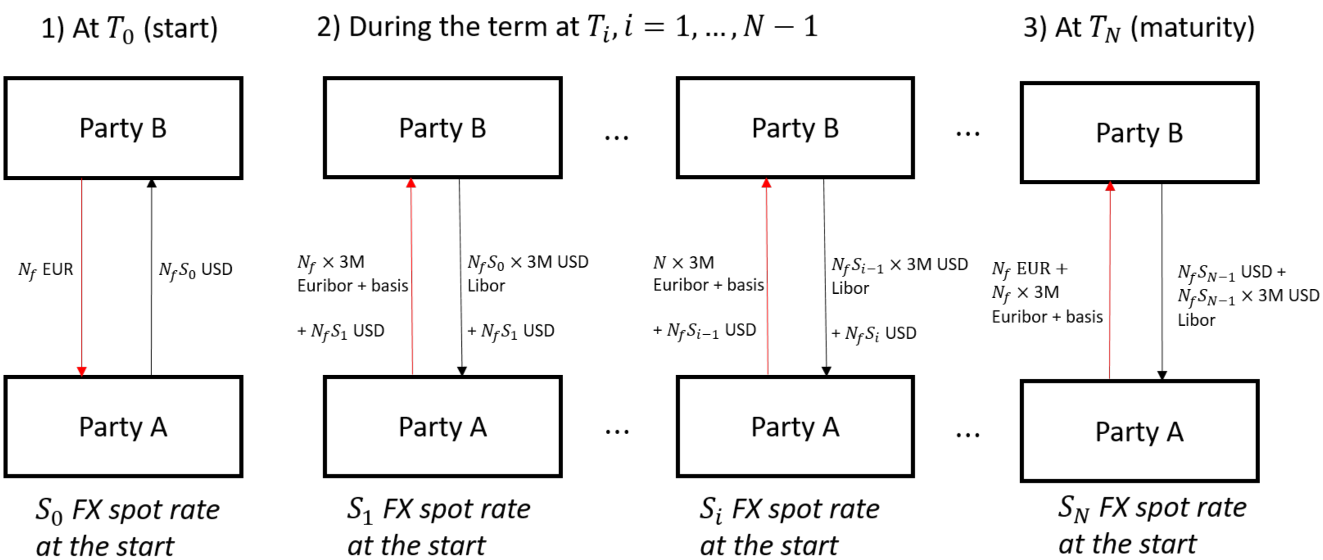


Figure 6.5.4: MtM Cross currency basis swap.

Remark 6.5.6 (understand the cash flow in MtM cross-currency basis swap). Let A be the domestic party. The cash flows from A's point of view are

- For undiscounted domestic currency resettable leg(cash flow in domestic currency)
  - At T_0, $C_d(T_0) = -S(T_0)N_f$
  - At $T_i, i = 1, 2, ..., M$, $C_d(T_i) = L^d(T_{i-1}, T_i)\tau_i S(T_{i-1})N^f + (S(T_{i-1}) - S(T_i))N^f$
  - At T_M, $C_d(T_M) = S(T_M)N^f$.
- For undiscounted foreign currency leg
  - At T_0, $C_f(T_0) = N_f$.
  - At $T_i, i = 1, 2, ..., M$, $C_f(T_i) = -L^f(T_{i-1}, T_i)\tau_i N^f$.
  - At T_M, $C_f(T_M) = -N_f$.

Note that for the resettable leg cash flow, we can decompose into N single period swaps over T_{i-1}, T_i, each one with discounted cash flow given by

$$\underbrace{-S(T_i)N_f}_{occurs\ at\ T_{i-1}} + \underbrace{L^d(T_{i-1},T_i)\delta_i S(T_{i-1})N_f + S(T_{i-1})N_f}_{occurs\ at\ T_i}.$$

Lemma 6.5.6 (valuing MtM float-float cross-currency swap). *Consider a cross-currency swap with $N^f = 1$.*

- *For the domestic party, the present value of a one-period leg going between T_{i-1} and T_i(see Remark 6.5.6), $i = 0,1,2,...,M$, is given by*

$$V_i^d(t) = \underbrace{P^d(t,T_{i-1})E_{T^d}[-N^f S(T_{i-1})|\mathcal{F}_t]}_{pay\ out\ new\ notional\ at\ T_{i-1}} +$$
$$\underbrace{P^d(t,T_i)E_{T^d}[S(T_{i-1})|\mathcal{F}_t]}_{receive\ notional\ at\ T_i} + \underbrace{P^d(t,T_i)E_{T^d}[S(T_{i-1})L^d(T_{i-1},T_i)\tau_i|\mathcal{F}_t]}_{receive\ interest\ at\ T_i}$$

- *If **assuming the independence between foreign exchange rate S and interest rate**, we have*

$$V_i^d(t) = -P^d(t,T_{i-1})N^f F_X(t,T_{i-1}) + P^d(t,T_i)F_X(t,T_{i-1}) + P^d(t,T_i)F_X(t,T_{i-1})F^d(t,T_{i-1},T_i)\tau_i$$

  where $F_X(t,T)$ is the forword exchange rate, $F^d(t,T_i,T_{i+1})$ is the domestic forward Libor rate.
- *The present value of the domestic leg is*

$$\begin{aligned} V^d(t) &= \sum_{i=0}^{N} V_n \\ &= -N^f(P^d(t,T_0)F_X(t,T_0) - P(t,T_N)F_X(t,T_{N-1}) \\ &- \sum_{i=1}^{N-1} P^d(t,T_i)(F_X(t,T_i) - F_X(t,T_{i-1}) \\ &+ \sum_{i=1}^{N} P^d(t,T_i)F_X(t,T_{i-1})F^d(t,T_{i-1},T_i)\tau_i \end{aligned}$$

- *The present value of the foreign leg is*

$$V^f(t) = N^f S(t)(-P^f(t,T_0) + \sum_{i=1}^{M} N\tau_i F^f(t,T_{i-1},T_i)P^f(t,T_i) + P^f(t,T_M)).$$

- *The present value of the swap is*

$$V(t) = V^d(t) + V^f(t).$$

Proof. (1) direct cash flow discounting Remark 6.5.6. (2) With independence assumption, we have

$$\begin{aligned}
&P^d(t,T_i)E_{T^d}[S(T_{i-1})|\mathcal{F}_t] \\
=&E_{Q^d}[S(T_{i-1})\frac{M(t)}{M(T_i)}|\mathcal{F}_t] \\
=&E_{Q^d}[S(T_{i-1})|\mathcal{F}_t]E_{Q^d}[\frac{M(t)}{M(T_i)}|\mathcal{F}_t] \\
=&F_X(t,T_{i-1})P^d(t,T_i),
\end{aligned}$$

where use the martingale property of foreign exchange rate in Theorem 6.3.3. Similarly,

$$\begin{aligned}
&P^d(t,T_i)E_{T^d}[S(T_{i-1})L^d(T_{i-1},T_i)\tau_i|\mathcal{F}_t] \\
&= P^d(t,T_i)E_{T^d}[S(T_{i-1})|\mathcal{F}_t]P^d(t,T_i)E_{T^d}L^d(T_{i-1},T_i)\tau_i|\mathcal{F}_t] \\
&= P^d(t,T_i)F_X(t,T_{i-1})F^d(t,T_{i-1},T_i).
\end{aligned}$$

□

6.5.5 Options on spot and forward exchange rate

6.5.5.1 *Basic concepts*

Definition 6.5.5 (European call/put option).

- *A* ***call option on the exchange rate*** *on $S(t)$ with N notional is a financial contract with payoff $N(S(T)-K)^+$ in domestic currency, where K is the strike parameter, at the maturity date T.*
- *Or equivalently, a call option on exchange rate gives the buyer the right to use K units of domestic currency to buy 1 unit of foreign currency whose value is $S(T)$. Therefore, when the $S(T) > K$, the option buyer will buy foreign currency at price K, which is lower than its market price $S(T)$.*
- *Similarly, a* ***put option on the exchange rate*** *on $S(t)$ with N notional gives the holder to right to sell N units of foreign currency at K domestic value. It has payoff given by $N(K-S(T))^+$.*

Remark 6.5.7 (premium and payment conventions)**.**

- The premium of option can be paid in either domestic currency or foreign currency.
- From the perspective of domestic investors, the notional of the option contract is foreign currency.
- The payoff of the option can be in either domestic currency or foreign currency.

Remark 6.5.8 (symmetric relationship on options).

- A CCY1/CCY2 call option (right to buy CCY1 by selling CCY2) is equivalent to a CCY2/CCY1 put option (right to sell CCY2 and receive CCY1).
- The relation between CCY1/CCY2 call and CCY1/CCY2 is connected by put-call parity.

6.5.5.2 *Pricing*

Lemma 6.5.7 (European option on spot exchange rate in Black model with payoff and premium paid in domestic currency). *[7, p. 390][4, p. 377] Let current time be t. Assume constant domestic interest rate r, constant foreign interest rate r^f, and constant volatility σ_S. Consider a call or put option on the FX spot rate S_t with strike K and expiry T whose payoff in domestic currency is $(S(T)-K)^+$. At time t, the value(in domestic currency) is*

$$C(t,S(t)) = e^{-r(T-t)}E_{Q^d}[(S(T)-K)^+|\mathcal{F}_t],$$

where the expectation is taken with respect to domestic risk-neutral measure Q^d. In particular, given by the fact that(Theorem 6.3.3), under Q^d, the dynamics of S is

$$dS(t) = (r-r^f)S(t)dt + \sigma_S S(T)d\hat{W},$$

then current value is given by

$$C(t,S(t)) = e^{-r^f\tau}S(t)N(d_+) - e^{-r\tau}KN(d_-),$$

$$P(t,S(t)) = e^{-r^f\tau}S(t)N(d_+) - e^{-r\tau}KN(d_-),$$

where

$$d_\pm = \frac{1}{\sigma_S\sqrt{\tau}}[\log(\frac{S(t)}{K}) + (r-r^f \pm \frac{1}{2}\sigma_S^2)\tau], \tau = T-t.$$

Proof. Use martingale pricing at Theorem 6.3.7 such that

$$C(t,S(t)) = \exp(-r_d(T-t))E_{Q^d}[(S(T)-K)^+|\mathcal{F}_t].$$

In the Black model, we assume the dynamics of $S(t)$ under domestic risk-neutral measure will follow

$$dS(t)/S(t) = (r_d - r_f)dt + \sigma_S dW_t.$$

Then we use the result from equity option(Lemma 5.2.2). □

Remark 6.5.9 (price in terms of forward exchange rate)**.** [4, p. 378] Using the forward exchange rate expression given by $F(t) = S(t)\exp((r - r_f)(T - t)$, we have

$$C = \exp(-r\tau)(F(t)N(d_1) - KN(d_2))$$
$$P = \exp(-r\tau)(-F(t)N(-d_1) + KN(-d_2))$$

where

$$d_{1,2} = \frac{1}{\sigma_S\sqrt{\tau}}[\log(\frac{F(t)}{K}) + (\pm\frac{1}{2}\sigma_S^2)\tau], \tau = T - t.$$

Remark 6.5.10 (value of the option if premium is in foreign currency)**.** The value of the option at current time t will be worth $\frac{V(t)}{S_t}$ in the unit of foreign currency.

Remark 6.5.11 (business needs)**.** [8, p. 113]

Consider the company Microsoft.

- In some countries (e.g., Europe, Japan, and Australia), it bills in the local currency and converts its net revenue to U.S. dollars monthly. For these currencies, Microsoft has a exposure to exchange rate movements.
- In other countries (e.g., those in Latin America, Eastern Europe, and Southeast Asia), it bills in U.S. dollars. Suppose the U.S. dollar appreciates against the currency of a country in which it is billing in U.S. dollars. People in the country will find Microsoft's products more expensive. As a result, Microsoft is likely to reduce its (U.S. dollar) price in the country or face a decline in sales. Microsoft therefore also has a foreign exchange exposure-both when it bills in U.S. dollars and in foreign currency.
- Microsoft therefore will use options to manage the foreign exchange rate risk.

6.5.5.3 *Put-call duality*

Lemma 6.5.8 (put-call duality). *[7, p. 391] Consider a call on FX rate $S(T)$ with strike K and expiry T. Its price denominated in domestic currency equals its price of a K units of put on $1/S(T)$ with strike $1/K$ dominated in foreign currency. That is,*

$$\frac{1}{S_t}v_{call}^d(S_t, K, T, r_d, r_t) = Kv_{put}^d(\frac{1}{S}, \frac{1}{K}, T).$$

Proof. Change of measure approach:

$$\begin{aligned}
&C(S(T),K) \textit{ in domestic currency} \\
=&E_{Q^d}[\frac{M(t)}{M(T)}(S(T)-K)^+|\mathcal{F}_t] \\
=&E_{Q^f}[\frac{M(t)}{M(T)}(S(T)-K)^+\frac{dQ^d}{dQ^f}|\mathcal{F}_t] \\
=&E_{Q^f}[\frac{M(t)}{M(T)}(S(T)-K)^+\frac{S(t)M(T)M^f(t)}{M^f(T)S(T)M(t)}|\mathcal{F}_t] \\
=&E_{Q^f}[\frac{M^f(t)S(t)}{M^f(T)S(T)}(S(T)-K)^+|\mathcal{F}_t] \\
=&E_{Q^f}[\frac{M^f(t)S(t)}{M^f(T)}K(\frac{1}{K}-\frac{1}{S(T)})^+|\mathcal{F}_t] \\
=&S(t)KE_{Q^f}[\frac{M^f(t)}{M^f(T)}K(\frac{1}{K}-\frac{1}{S(T)})^+|\mathcal{F}_t] \\
=&K\times P(\frac{1}{S(T)},\frac{1}{K}) \textit{ dominated in foreign currency}
\end{aligned}$$

□

Remark 6.5.12 (interpretation)**.**

- An option to buy one unit of EUR and sell K units of USD (which is a call) is equivalent to an option to sell K units of USD and buy one unit of EUR (which is a put).
- For a domestic (USD) investor, the payoff of the call in USD is

$$(S_T-K)^+;$$

for a foreign (EUR) investor, the payoff of the put in EUR is

$$(1-KS'_T)^+, S^T = 1/S_T.$$

- Let
$$V^d = e^{-r_dT}E_{Q^d}[(S_T-K)^+], V^f = e^{-r_fT}E_{Q^f}[(1/K-1/S_T)^+],$$
then $V^d = S_0K\cdot V^f$ based on Theorem 6.3.6.

6.5.5.4 *Hedging*

Remark 6.5.13 (Delta hedging for exchange rate option)**.**

- Consider hedging a long position of call with payoff $C_T = (S_T - K)^+$. Let $\Delta = \partial C_t / \partial S_t$. We can short Δ units of foreign currency and receive payment in $\Delta \cdot S_t$ of domestic currency.
- If S_t increases, our call option will increase value in the unit of domestic currency; however, to close our short position in the foreign currency, we need more domestic currency than before to buy(therefore we are losing money).

6.5.6 Forwards on foreign assets

6.5.6.1 *Forwards with domestic forward price*

Definition 6.5.6 (Forwards with domestic forward price).

- *Let S_t^f denote the foreign price of a foreign asset. A forward contract is an agreement to buy a foreign asset with domestic price K^d at future time T.*
- *The payoff in domestic currency at time T is given by*

$$V^d(T) = (S_T^f X_T - K^d),$$

where X_T is the FX rate at time T.

Lemma 6.5.9 (Forward price for forwards with domestic forward price). *Let S_t^f denote the foreign price of a foreign asset. Let X_t denote FX spot rate. Consider a forward with maturity T and domestic price K^d. Let current time be t.*

- *The current value of the forward contract is*

$$\begin{aligned} V^d(t) &= E_{Q^d}[\frac{M(t)(X_T S_T^f - K^d)}{M(T)} | \mathcal{F}_t] \\ &= P^d(t,T)(E_{T^d}[X_T S_T^f] - K^d) \end{aligned}$$

- *The forward price make the forward contract have zero value is given by*

$$F^d(t,T) = E_{T^d}[X_T S_T^f].$$

- *Assume deterministic interest rate, we have*

$$F^d(t,T) = E_{Q^d}[X_T S_T^f] = X_t S_t^f \exp(r^d(T-t)).$$

6.5.6.2 *Forwards with foreign forward price*

Definition 6.5.7 (Forwards with foreign forward price).

- *Let S_t^f denote the foreign price of a foreign asset. A forward contract is an agreement to buy a foreign asset with foreign price K^f at future time T.*
- *The payoff in domestic currency at time T is given by*

$$V^d(T) = (S_T^f X_T - K^d),$$

where X_T is the FX rate at time T.

Methodology 6.5.1 (Forward price for forwards with domestic forward price). *Let S_t^f denote the foreign price of a foreign asset. Let X_t denote FX spot rate. Consider a forward with maturity T and domestic price K^d. Let current time be t.*

- *The current value of the forward contract is*

$$\begin{aligned} V^d(t) &= X_t E_{Q^f}[\frac{M^f(t)(S_T^f - K^f)}{M^f(T)}|\mathcal{F}_t] \\ &= X_t P^f(t,T)(E_{T^f}[X_T S_T^f] - K^f) \end{aligned}$$

- *The forward price make the forward contract have zero value is given by*

$$F^f(t,T) = E_{T^f}[S_T^f].$$

- *Assume deterministic interest rate, we have*

$$F^f(t,T) = E_{Q^f}[S_T^f] = S_t^f \exp(r^f(T-t)).$$

6.5.7 Options on foreign assets

6.5.7.1 *Basic principles*

Lemma 6.5.10 (option on foreign asset paying domestic currency). *Let the foreign asset S_t^f, denominated in foreign currency, have dynamics given by*

$$dS_t^f / S_t^f = \mu dt + \sigma_S dW_1.$$

Let the exchange rate X_t have dynamics given by

$$dX_t/X_t = \gamma dt + \sigma_X dW_2, dW_1 dW_2 = \rho dt.$$

It follows that

- *Under domestic measure Q^d, S_t and X_t have dynamics given by*

$$d(S_t)/(S_t) = (r^f - \rho\sigma_S\sigma_S)dt + \sigma_S dW_1^d$$
$$d(X_t)/(X_t) = (r^d - r^f)dt + \sigma_X dW_2^d$$

- *Consider an option will payoff $V^d(S_T^f, X_T^f)$ in domestic currency, then its current value is*

$$V^d(t) = M(t)E_{Q^d}[\frac{V^d(S_T)}{M(T)}|\mathcal{F}_t].$$

- *Consider an option will payoff $V^f(S_T^f, X_T^f)$ in foreign currency, then its current domestic value is*

$$V^d(t) = M(t)E_{Q^d}[\frac{V^f(S_T, X_T)X_T^f}{M(T)}|\mathcal{F}_t],$$

or

$$V^d(t) = S(t)V^f(t), V^f = M^f(t)E_{Q^f}[\frac{V^f(S_T, X_T)}{M^f(T)}|\mathcal{F}_t].$$

Proof. From Lemma 6.3.6. □

Remark 6.5.14. Note that S_tX_t is a martingale under domestic measure (Lemma 6.3.6), but S_t is not.

6.5.7.2 *European options with domestic strike*

Definition 6.5.8 (European options with domestic strike).

- *Let S_t^f denote the foreign price of a foreign asset. An European call option with domestic strike K^d gives the holder the right to buy the foreign asset at future time T with domestic price K^d.*
- *The payoff in domestic currency at time T is given by*

$$V^d(T) = (S_T^f X_T - K^d)^+,$$

where X_T is the FX rate at time T.

Lemma 6.5.11 (European options with domestic strike pricing). *Let S_t^f denote the foreign price of a foreign asset. Let X_t denote FX spot rate. Consider an European call with domestic strike K^d and maturity T Assume X_tS_t has domestic risk-neutral dynamics (Theorem 6.3.5) given by*

$$dX_tS_t^f/X_tS_t^f = (r^d - q)dt + \overline{\sigma}_S dW_t,$$

where $\overline{\sigma}_S^2 = \sigma_X^2 + 2\rho\sigma_S\sigma_X + \sigma_S^2$, σ_S is the volatility of asset S_t under foreign risk-neutral measure. then

$$V_t = BS(S_t^f X_t, r^d, q, K^d, T, \overline{\sigma}_S),$$

where $BS(S_t^f X_t, r^d, q, K^d, T, \overline{\sigma}_S)$ is the Black-Scholes formula with spot $S_t^f X_t$, drift r^d, and dividend rate q.

Proof. We can view $X_tS_t^f$ as a domestic asset with drift $r^d - q$, where $q = r^d - r^f - \rho\sigma_S\sigma_X$ is the dividend rate. □

6.5.7.3 *European options with foreign strike*

Definition 6.5.9 (European options with foreign strike).

- *Let S_t^f denote the foreign price of a foreign asset. An European call option with foreign strike K^f gives the holder the right to buy the foreign asset at future time T with foreign price K^f.*
- *The payoff in domestic currency at time T is given by*

$$V^d(T) = X_T(S_T^f - K^f)^+,$$

where X_T is the FX rate at time T.

Lemma 6.5.12 (European options with foreign strike pricing). *Let S_t^f denote the foreign price of a foreign asset. Let X_t denote FX spot rate. Consider an European call with foreign strike K^f and maturity T Assume S_t has foreign risk-neutral dynamics (Theorem 6.3.5) given by*

$$dS_t^f/S_t^f = (r^f - q)dt + \sigma_S dW_t,$$

where σ_S is the volatility of asset S_t under foreign risk-neutral measure. then

$$V_t^d = X_t BS(S_t, r^f, q, K^f, T, \sigma_S),$$

where $BS(S_t^f, r^f, q, K^f, T, \sigma_S)$ is the Black-Scholes formula with spot S_t^f, drift r^f, and dividend rate q.

Proof. From Theorem 6.3.6, we have

$$\begin{aligned}
V^d(t) &= X(t)V^f(t) \\
&= X(t)E_{Q^f}[\frac{M^f(t)V^f(T)}{M^f(T)}|\mathcal{F}_t] \\
&= X(t)E_{Q^f}[\frac{M^f(t)(S_T^f - K^f)^+}{M^f(T)}|\mathcal{F}_t] \\
&= X_t BS(S_t^f, r^f, q, K^f, T, \sigma_S)
\end{aligned}$$

□

6.5.8 Quanto securities

6.5.8.1 *Quanto forwards*

Definition 6.5.10 (quanto forward). *Let S_t^f denote the foreign price of a foreign asset. From the perspective of domestic investor, a Quanto forward is agreement to purchase the foreign asset at a fixed domestic price at a fixed exchange rate $\overline{X}$.*

The fixed domestic price chosen such that the price of the contract is zero at the time of entry called quanto-forward price.

Definition 6.5.11 (quanto-forward curve). *A quanto-forward curve observed at current time t on foreign asset S_t^f is defined by*

$$F^d(t,T) = E_{Q^d}[S(T)|\mathcal{F}_t].$$

Lemma 6.5.13 (quanto-forward price and quanto-forward curve). *[9, p. 118] Consider a quanto-forward contract on foreign asset S_t^f with delivery date T and fixed FX rate $\overline{X}$. Let current time be t. Then the quanto-forward price is given by*

$$F^d(t,T;\overline{X}) = \overline{X}E_{Q^d}[S_T|\mathcal{F}_t].$$

Assume S_T^f has domestic risk-neutral dynamics (Theorem 6.3.5) given by

$$dS_t^f/S_t^f = (r^f - \rho\sigma_S\sigma_X)dt + \sigma_S dW_t,$$

then

$$F^d(t,T) = S_t^f \exp((r_f - \rho\sigma_X\sigma_S)(T-t)) = F^f(t,T)\exp(-\rho\sigma_X\sigma_S(T-t)),$$

where $F^f(t,T)$ is the forward curve on foreign asset denominated in foreign currency given by $F^f(t,T) = S_t^f \exp(r_f(T-t))$

Proof. In the domestic measure, we have

$$0 = E_{Q^d}[e^{-r^d(T-t)}(\overline{X}S_T - F^d(0,T))];$$

therefore

$$F^d(t,T) = E_{Q^d}[\overline{X}S_T] = \overline{X}E_{Q^d}[S_T|\mathcal{F}_t].$$

□

6.5.8.2 *Quanto options*

Definition 6.5.12 (European quanto option with domestic strike).

- *Let S_t^f denote the foreign price of a foreign asset. An European call quanto-option with domestic strike K^d gives the holder the right to buy the foreign asset at future time T with domestic price K and fixed FX rate $\overline{X}$.*
- *The payoff in domestic currency at time T is given by*

$$V^d(T) = (S_T^f\overline{X} - K^d)^+.$$

Lemma 6.5.14 (quanto option pricing). *Let S_t^f denote the foreign price of a foreign asset. Consider an European call quanto-option with domestic strike K^d, maturity T, and fixed FX rate $\overline{X}$. Assume S_T has domestic risk-neutral dynamics (Theorem 6.3.5) given by*

$$dS_t/S_t = (r^f - \rho\sigma_S\sigma_X)dt + \sigma_S dW_t,$$

where σ_S is the volatility of asset S_t under foreign risk-neutral measure. then

$$V_t = BS(S_t\overline{X}, r^d, q, K^d, T, \sigma_S),$$

where $BS(S_t\overline{X}, r^d, q, K^d, T, \sigma_S)$ is the Black-Scholes formula with spot $S_t\overline{X}$, drift r^d and dividend rate $q = r^d - r^f - \rho\sigma_S\sigma_X$.

Proof. We can view S_t as a domestic asset with drift $r^d - q$, where $q = r^d - r^f - \rho\sigma_S\sigma_X$ is the dividend rate. □

6.6 Notes on bibliography

Major references are [7][2][10][11, ch 14]. [12][1][13].

BIBLIOGRAPHY

1. Reiswich, D. & Uwe, W. FX volatility smile construction. *Wilmott* **2012,** 58–69 (2012).
2. Wilmott, P. *Paul Wilmott on quantitative finance* (John Wiley & Sons, 2013).
3. Pelsser, A. Mathematical foundation of convexity correction. *Quantitative Finance* **3,** 59–65 (2003).
4. Hull, J. *Options, Futures and Other Derivatives 9 edition* (Pearson Education, 2009).
5. Grzelak, L. A. & Oosterlee, C. W. On cross-currency models with stochastic volatility and correlated interest rates. *Applied Mathematical Finance* **19,** 1–35 (2012).
6. Brigo, D., Morini, M. & Pallavicini, A. *Counterparty Credit Risk, Collateral and Funding: With Pricing Cases For All Asset Classes* ISBN: 9780470661789. https://books.google.com/books?id=Y-X9zqVcOFkC (Wiley, 2013).
7. Shreve, S. E. *Stochastic calculus for finance II: Continuous-time models* (Springer Science & Business Media, 2004).
8. Hull, J. *Risk management and financial institutions,+ Web Site* (John Wiley & Sons, 2012).
9. Back, K. *A course in derivative securities: Introduction to theory and computation* (Springer Science & Business Media, 2006).
10. Clark, I. J. *Foreign exchange option pricing: a practitioner's guide* (John Wiley & Sons, 2011).
11. Musiela, M. & Rutkowski, M. *Martingale methods in financial modelling* (Springer Science & Business Media, 2006).
12. Reiswich, D. & Wystup, U. A guide to FX options quoting conventions. *Journal of Derivatives* **18,** 58 (2010).
13. Castagna, A. *Fx options and smile risk* (John Wiley & Sons, 2010).

7

CREDIT RISK MODELING

7.1 Default term structure modeling

7.1.1 Concepts of default

Credit risk can be defined as the possibility that a contractual counterparty does not meet its obligations stated in the contract, therefore causing the creditor a financial loss. Credit default swap is Credit derivatives are financial instruments that essentially provide insurance against the credit deterioration and default of the counterparty.

We start this chapter with a basic probabilistic model on default. We then incorporate the credit default probability description into the risk-neutral pricing framework.

We also introduce the Merton's structure model and credit migration model.

Definition 7.1.1 (default concepts).

- *The* ***default indicator process*** *among a time interval* $[0, T]$ *is defined by*

$$N(u) = \mathbf{1}_{(\tau < u)},$$

where the stopping time τ *is the default event.*
- *The* ***hazard rate function*** $\lambda(t)$ *associated with jump process* $N(u)$ *describes the rate of a jump occurs at the next instant given that the life has survived to time* t*.*
- *The* ***hazard function*** *is defined as* $\Lambda(t) = \int_0^t \lambda(s)ds$ *is the accumulated default intensity.*
- $e^{-\Lambda} = P(N(t) = 0)$ *is the probability that there is no default in the interval* $(0, t)$*.*

- ***Survival probability*** *$S(t,u)$ is defined as*

$$S(t,u) = P(\tau > u|\tau > t) = \exp(-\int_t^u \lambda(s)ds).$$

Note that $\lambda(u) = -\frac{\partial \ln S(t,u)}{\partial u}$.

- ***Default probability*** *$D(t,u)$ is defined as*

$$D(t,u) = P(\tau < u|\tau > t) = 1 - S(t,u)$$

- ***Survival indicator function*** *$I(t)$ is defined as*

$$I(t) = \mathbf{1}_{\tau > t}.$$

Note that $S(t,u) = E[I(u)|I(t) = 0]$.

Definition 7.1.2 (default loss). *[1, p. 6] Consider a set of obligors, indexed by $i = 1, ..., I$. Define*

- *$N_i(t)$: Default indicator process. $N_i(t) = 1$ if obligor i has defaulted by time t, and 0 otherwise.*
- *$E_i(t)$: Exposure process. $E_i(t)$ is the amount we would lose if obligor i defaulted at time t with zero recovery. Also called exposure at default.*
- *$L_i(t)$: Loss given default of obligor i at time t. $L_i(t) \leq 1$ since a fraction of $E_i(t)$ may be recovered in bankruptcy proceedings. The **recovery rate** is given by $R_i(t) = 1 - L_i(t)$.*

*Then the **default loss** of obligor i*

$$D_i(T) \triangleq N_i(T) \times E_i(\tau_i) \times L_i(\tau_i),$$

where T is some time horizon, say 1 year, and τ_i is the time of default.

7.1.2 Probability characterization of default

Definition 7.1.3. *Let τ be a random variable be the default time.*

- *Let F_τ be the **default cumulative density function** such that*

$$F_\tau(t) \triangleq Pr(\tau < t), F(0) = 0$$

- *Let $S : \mathbb{R} \to [0,1]$ be the **survival probability** such that*

$$S(t) = Pr(\tau > t) = 1 - F(t), S(0) = 1.$$

- *Let f_τ be the* ***default probability density function*** *such that*

$$f_\tau(t) \triangleq \lim_{\Delta\to 0} \frac{Pr(t < \tau < t+\Delta)}{\Delta} = \frac{dF_\tau(t)}{dt} = -\frac{dS_\tau(t)}{dt}.$$

- *Let $h(t)$ be the* ***conditional instantaneous default probability*** *or* ***hazard rate*** *such that*

$$h(t) = \lim_{\Delta\to 0} \frac{Pr(t < \tau < t+\Delta t | \tau > t)}{\Delta t}.$$

Lemma 7.1.1 (basic properties of default probabilities in the hazard curve model). *Let the current time be 0. Assume the default time τ has cdf given by*

$$F_\tau(t) = 1 - \exp(-\int_0^t h(s)ds),$$

where $h(t)$ is the hazard rate function. It follows that

- *(recover hazard rate function)*

$$h(t) = \frac{1}{S(t)}\frac{dF_\tau(t)}{dt} = \frac{f_\tau(t)}{S(t)} = -\frac{S'(t)}{S(t)}.$$

- *(calculation of survival probability)*

$$Pr(\tau > t) = S(t) = \exp(-\int_0^t h(s)ds).$$

  If $h(s) = h_0$ is flat, then

$$Pr(\tau > t) = S(t) = \exp(-h_0 t).$$

- *(calculation of default probability density)*

$$f_\tau(t) = S(t)h(t) = h(t)\exp(-\int_0^t h(s)ds).$$

- ***Prior default probability*** *at future time interval* $[T_1, T_2]$ *is given by*

$$\begin{aligned} Pr(T_1 < \tau < T_2) &= \int_{T_1}^{T_2} f(s)ds \\ &= S(T_1) - S(T_2) \\ &= S(T_1)(1 - \exp(-\int_{T_1}^{T_2} h(s)ds)) \end{aligned}$$

- ***Conditional default probability*** *at future time interval* $[T_1, T_2](T_0 < T_1)$ *is given by*

$$Pr(T_1 < \tau < T_2 | \tau > T_0) = \frac{S(T_1) - S(T_2)}{S(T_0)}.$$

- *Conditional survival probability at future time* T_1 *given survival till* T_0 *is given by*

$$Pr(T_1 < \tau | \tau > T_0) = \frac{S(T_1)}{S(T_0)}.$$

Proof. (1) From definition, we have

$$h(t) = \lim_{\Delta \to 0} \frac{Pr(t < \tau < t + \Delta t | \tau > t)}{\Delta t} = \lim_{\Delta \to 0} \frac{1}{\Delta t} \frac{Pr(t < \tau < t + \Delta t)}{Pr(t > \tau)} = \frac{f(t)}{S(t)}.$$

(2) From (1), since $S' = -Sh, S(0) = 1$, we can solve

$$S(t) = S(0) \exp(-\int_0^t h(s)ds) = \exp(-\int_0^t h(s)ds).$$

(3) From (1), we have $f(t) = S(t)h(t)$. (3) Note that

$$\int_{T_1}^{T_2} f_\tau(s)ds = \int_{T_1}^{T_2} h(s)S(s)ds = -\int_{T_1}^{T_2} S'(s)ds = S(T_1) - S(T_2).$$

(4) Note that

$$P(T_1 < \tau < T_2 | \tau > T_0) = \frac{P(T_1 < \tau < T_2)}{P(\tau > T_0)} = \frac{S(T_1) - S(T_2)}{S(T_0)}.$$

□

7.2 Risk-neutral pricing framework

7.2.1 Risk-neutral pricing fundamentals

7.2.1.1 *Principle*

Lemma 7.2.1. *Consider an asset with payoff $V(T)$ at maturity T. Let τ be the default time. Then the current value of the asset is*

$$V(t) = E_Q[\exp(-\int_t^T r(s)ds)V(T)\mathbf{1}_{\tau>T} + \exp(-\int_t^\tau r(s)ds)V(\tau)R(\tau)\mathbf{1}_{t<\tau<T}],$$

where R is the recovery rate.4

Proof. Note that $\exp(-\int_t^\tau r(s)ds)V(\tau)R(\tau)\mathbf{1}_{t<\tau<T}$ denotes the recovered value at the default time. □

7.2.1.2 *Defaultable zero-coupon bond*

Definition 7.2.1 (zero coupon bond). *[2]*

- *The **default-free zero-coupon bonds** maturing at T have price at $t < T$ given by*
$$P(t,T).$$
- *The **defaultable zero-recovery zero-coupon bonds** maturing at T have price at $t < T$ given by*
$$P_0^d(t,T).$$
- *The **defaultable zero-coupon bonds** maturing at T have price at $t < T$ given by*
$$P^d(t,T).$$
- *The **credit spread** is defined by*
$$s(t,T) = \frac{1}{T-t}\ln(\frac{1}{P^d(t,T)}) - \frac{1}{T-t}\ln(\frac{1}{P(t,T)}).$$

Lemma 7.2.2 (defaultable zero-coupon bond pricing). *[3, p. 62][2, p. 54]*

- *Under risk-neutral measure Q, the default-free zero-coupon bonds have prices given by*
$$P(t,T) = E_Q[\exp(-\int_t^T r(s)ds)|\mathcal{F}_t].$$

- *Let τ denote the random default time. Under risk-neutral measure Q, the defaultable zero-recovery zero-coupon bonds have prices given by*

$$P_0^d(t,T) = E_Q[\exp(-\int_t^T r(s)ds)\mathbf{1}_{\tau>T}|\mathcal{F}_t].$$

- *Let τ denote the random default time. Under risk-neutral measure Q, the defaultable zero-recovery bonds have prices given by*

$$P_0^d(t,T) = E_Q[\exp(-\int_t^T r(s)ds)(\mathbf{1}_{\tau>T} + R\mathbf{1}_{\tau<T})|\mathcal{F}_t],$$

where R is the recovery rate at maturity.

Proof. $\tau > T$ indicates no default occurs before maturity. □

Lemma 7.2.3 (defaultable zero-coupon bond pricing with independence and deterministic assumption). *[2, p. 54] Assume the default process is characterized by the deterministic intensity parameter $\lambda(t)$ independent of stochastic process $r(t)$. Let τ denote the random default time. It follows that*

- *The defaultable zero-recovery zero-coupon bonds have prices given by*

$$\begin{aligned} P_0^d(t,T) &= E_Q[\exp(-\int_t^T r(s)ds)\mathbf{1}_{\tau>T}|\mathcal{F}_t] \\ &= E_Q[\exp(-\int_t^T r(s)+\lambda(s)ds)|\mathcal{F}_t] \\ &= P(t,T)\exp(-\int_t^T \lambda(s)ds) \\ &= P(t,T)S(t,T). \end{aligned}$$

- *Then the defaultable zero-coupon bond has price given by*

$$P^d(t,T) = P(t,T)E_Q[\mathbf{1}_{\tau>T} + R\mathbf{1}_{\tau<T}|\mathcal{F}_t] = P(t,T)(S(t,T) + R(1-S(t,T))).$$

where R is the external given recovery rate paid at T.
- *If $\lambda(t)$ and $r(t)$ are constant, then*

$$P_0^d(t,T) = P^d(t,T)\exp(-\lambda(T-t)).$$

Proof. (1)

$$
\begin{aligned}
P_0^d(t,T) &= E_Q[\exp(-\int_t^T r(s)ds)\mathbf{1}_{\tau>T}|\mathcal{F}_t] \\
&= P(t,T)E_Q[\mathbf{1}_{\tau>T}|\mathcal{F}_t] \\
&= P(t,T)S(t,T) = P(t,T)\exp(-\int_t^T \lambda(s)ds),
\end{aligned}
$$

where $S(t,T) = Q(\tau > T|\tau > t) = \exp(-\int_t^T \lambda(s)ds)$.

□

Definition 7.2.2 (implied survival probability, implied default term structure). *[2, p. 54]* *Denote the default time by τ.*

- *The **implied survival probability** between from t to T as seen from t(i.e. based on information accumulated up to t)*

$$S^*(t,T) = \frac{P_0^d(t,T)}{P(t,T)}.$$

- *The **implied default probability** over $[t,T]$ is given by*

$$D^*(t,T) = 1 - S^*(t,T).$$

- *The **implied density of default time** is*

$$Q(\tau \in [T, T+dt]|\mathcal{F}_t) = -\frac{\partial S^*(t,T)}{\partial T}dt.$$

- *The **implied hazard rate of default** at time T is*

$$h^*(T) = -\frac{\partial \ln S(t,T)}{\partial T}$$

Lemma 7.2.4 (decomposition of defaultable zero-coupon bond). *[2, p. 63]* *Let $t = T_0 < T_1 < T_2 < \cdots < T_n$ be a set of dates. Let $\delta_i = T_{i+1} - T_i$.*

- $$P(t,T_k) = \prod_{i=1}^{k} \frac{1}{1+\delta_{i-1}F(t,T_{i-1},T_i)}$$

- $$P_0^d(t,T_k) = P(t,T_k)S(t,T_k) = P(t,T_k)\prod_{i=1}^{k}\frac{1}{1+\delta_{i-1}H(t,T_{i-1},T_i)}$$

Proof. (1) see the zero-coupon bond decomposition (Lemma 3.1.2). (2) Use the that

$$S_c(t;T_1,T_2) = \frac{S(t,T_2)}{S(t,T_1)},$$

and

$$S_c(t;T_1,T_2) = \frac{1}{1+(T_2-T_1)H(t;T_1,T_2)}.$$

□

Lemma 7.2.5 (current value of recovery at default). *[2, pp. 60, 118][4, pp. 48, 52] Let $t = T_0 < T_1 < T_2 < \cdots < T_n$ be a set of dates. Let $\delta_i = T_{i+1} - T_i$. Then*

- *the current value of $1 payoff at T_{k+1} if and only if a default occurs in $(T_k, T_{k+1}]$.*

$$e(t,T_k,T_{k+1}) = \delta_k H(t,T_k,T_{k+1})P_0^d(t,T_{k+1}).$$

- *Assume the default process is characterized by an deterministic intensity parameter $\lambda(t)$ independent of stochastic process $r(t)$. We have*

$$\begin{aligned} e(t,T_k,T_{k+1}) &= E_Q[\exp(-\int_t^{T_{k+1}} r(s)ds)\mathbf{1}_{T_k<\tau<T_{k+1}}|\mathcal{F}_t] \\ &= P(0,T_{k+1})(\exp(-\int_t^{T_k}\lambda(s)ds) - \exp(-\int_t^{T_{k+1}}\lambda(s)ds)) \end{aligned}$$

- ***(alternative, payment at default)*** *the current value of $1 payoff at τ if and only if a default occurs in $(T_k, T_{k+1}]$.*

$$e_1(t,T_k,T_{k+1}) = E_Q[\exp(-\int_t^{\tau} r(s)ds)\mathbf{1}_{T_k<\tau<T_{k+1}})|\mathcal{F}_t].$$

In particular, if assume the default is characterized by the hazard rate $h(t,s)$, then

$$\begin{aligned} e_1(t,T_k,T_{k+1}) &= E_Q[\int_{T_k}^{T_{k+1}}\exp(-\int_t^{\tau} r(s)ds)S(t,\tau)h(t,\tau)d\tau|\mathcal{F}_t] \\ &= E_Q[\int_{T_k}^{T_{k+1}}\exp(-\int_t^{\tau} r(s)+h(t,s)ds)h(t,\tau)d\tau|\mathcal{F}_t] \end{aligned}$$

- ***(alternative, random payment at default)****the current value of a random payoff $\Phi(\tau)$ at τ if and only if a default occurs in $(T_k, T_{k+1}]$.*

$$e_2(t, T_k, T_{k+1}) = E_Q[\exp(-\int_t^\tau r(s)ds\Phi(\tau)\mathbf{1}_{T_k<\tau<T_{k+1}})|\mathcal{F}_t],$$

Proof. (1)Note that the probability of default at $(T_k, T_{k+1}]$ is given by

$$Pr(T_k < \tau < T_{k+1}|\mathcal{F}_t) = S(t, T_k) - S(t, T_{k+1}) = S(t, T_k)(1 - S_c(t, T_k, T_{k+1})).$$

Then the current value is

$$\begin{aligned} e(t, T_k, T_{k+1}) &= P(t, T_{k+1})S(t, T_k)(1 - S_c(t, T_k, T_{k+1})) \\ &= P(t, T_{k+1})\delta_k H(t, t_k, t_{k+1})S(t, T_{k+1}) \\ &= \delta_k H(t, T_k, T_{k+1})P_0^d(t, T_{k+1}) \end{aligned}$$

(2)Note that the probability of default at $(T_k, T_{k+1}]$ is given by

$$Pr(T_k < \tau < T_{k+1}|\mathcal{F}_t) = S(t, T_k) - S(t, T_{k+1}) = (\exp(-\int_t^{T_k} \lambda(s)ds) - \exp(-\int_t^{T_{k+1}} \lambda(s)ds)).$$

(3) Note that the conditional instantaneous default probability at τ is given by

$$S(t, \tau)h(t, \tau)d\tau.$$

□

Remark 7.2.1. The specific choice of e depends on the contract and the specific scenario.

7.2.2 Risk neutral vs. real default probability

Note 7.2.1. [5, p. 228]

- If one is interested in estimating the economic capital and risk charges, one should use real-world probability; If the objective is to price and hedge credit-related securities, we need to use the risk-neutral probability.
- The real-world probability is estimated from historical information; the risk-neutral probability can be implied from existing market price of credit product.

Example 7.2.1.

- Consider a year bond with face value 100 and 0.07 coupon at maturity. The one-year risk-free rate is 0.05. Assume the real default probability PD = 0.01, and recovery at default is 0.5.
- Discounting the expected payoff using the real default probability is given by

$$\frac{107 \times 0.99 + 0.5 \times 100 \times 0.01}{1 + 0.05} = 101.36.$$

- This price is higher for the buyers since it does not account for the risk aversion for the buyers. Taking expectation with respect to real probability is mistake since it assumes buyer are not risk averse.
- Suppose the market price is 100. Denote PD^* as the implied default probability or risk-neutral probability. We have

$$\frac{107 \times (1 - PD^*) + 0.5 \times 100 \times PD^*}{1 + 0.05} = 100,$$

gives $PD^* = 0.0351$. Such implied probability then can be used to price other credit product related to the same firm by taking expectation.

7.2.3 Survival curve construction

7.2.3.1 *Survival curve construction from CDS spread*

Note 7.2.2 (bootstrap method to construct hazard curve from CDS spread). Note that

$$s^{CDS} = (1 - \pi)\frac{\sum_{i=1}^{n} \delta_{i-1} H(t, T_{i-1}, T_i) P_0^d(t, T_i)}{\sum_{i=1}^{n} \delta_{i-1} P_0^d(t, T_i)},$$

where

$$H(t, T, T + \Delta t) = \frac{1}{\Delta t}(\exp(\int_T^{T+\Delta t} h(t, s) ds) - 1).$$

By assuming piece-wise linear form of $h(t, s), s \geq t$ and using market data of s^{CDS} for various maturities and tenors, we can solve the $h(t, s)$.

Lemma 7.2.6. *[4, p. 117] Let $z(t,T)$ be* ***flat default rate*** *such that*

$$S(t,T) = \exp(-z(t,T)(T-t)).$$

We have

- $$S(t,T) = \exp(-\int_t^T h(t,s)ds)$$
- $$h(t,T) = \frac{\partial}{\partial T}(z(t,T)(T-t))$$
- $$h(t,T) = -\frac{\partial}{\partial T}\ln S(t,T))$$

Proof. Note that

$$-z(t,T)(T-t) = -\ln S(t,T) = \int_t^T h(t,s)ds.$$

□

Remark 7.2.2 (analogs in interest rate modeling)**.**

- $z(t,T)$ here resembles the yield, $h(t,T)$ here resembles the forward rate.
- The construction of $h(t,T)$ from $z(t,T)$ is the same in 3.4.2

7.2.3.2 *Survival curve construction from bonds*

Lemma 7.2.7 (survival curve construction from defaultable zero-coupon bond). *Suppose we are given defaultable zero coupon bond price $P^d(t,T_1), P^d(t,T_2), ..., P^d(t,T_N)$ on a set of maturities $T_1, T_2, ..., T_N$. Further assume the recovery rate are given by $R_1, R_2, ..., R_N$ and default behavior is independent of exposure. Then the survival probability on $T_1, T_2, ..., T_N$ are given by*

$$S(t,T_i) = (\frac{P^d(t,T_i)}{P(t,T_i)} - R_i)/(1-R_i),$$

where $P(t,T_i)$ is the non-defaultable zero coupon bond price.

Proof. Note that from defaultable zero-coupon bond pricing(Lemma 7.2.3), we have

$$P^d(t,T_i) = P(t,T_i)(S(t,T_i) + R_i(1-S(t,T_i))).$$

Rearrange and we get

$$S(t, T_i) = (\frac{P^d(t, T_i)}{P(t, T_i)} - R_i)/(1 - R_i).$$

□

7.3 Risk neutral pricing examples

7.3.1 Defaultable fixed-coupon bond

Definition 7.3.1 (defaultable fixed-coupon bond). *[2, p. 65] Let $t < T_0 < T_1 < T_2 < \cdots < T_n$ be a set of dates. Let $\delta_i = T_{i+1} - T_i$. A defaultable fixed-coupon bond has coupon payments c_i at $T_i, i = 1, ..., n$. At maturity T_n, there is also a principal \$1 payment. If a default occurs at $(T_i, T_{i+1}]$, then the payment at T_{i+1} is π.*

Lemma 7.3.1 (price of a defaultable fixed-coupon bond). *[2, p. 65] The current price of a defaultable fixed-coupon bond is*

$$\begin{aligned} C(t) = & \sum_{i=1}^{n} \delta_{i-1} F(t, T_{i-1}, T_i) P_0^d(0, T_i) \ (coupons) \\ & + P_0^d(t, T_n) \ (principal) \\ & + \pi \sum_{i=1}^{n} e(t, T_{i-1}, T_i) \ (recovery) \end{aligned}$$

Proof. Straight forward. □

7.3.2 Defaultable floater

Definition 7.3.2 (defaultable floater). *[2, p. 65] Let $t < T_0 < T_1 < T_2 < \cdots < T_n$ be a set of dates. Let $\delta_i = T_{i+1} - T_i$. A defaultable fixed-coupon bond has coupon payments*

$$c_i = \delta_{i-1}(L(T_{i-1}, T_i) + s^{par})$$

at $T_i, i = 1, ..., n$. At maturity T_n, there is also a principal \$1 payment. If a default occurs at $(T_i, T_{i+1}]$, then the payment at T_{i+1} is π.

Lemma 7.3.2 (price of a defaultable floater bond). *[2, p. 66] The current price of a defaultable floater is*

$$
\begin{aligned}
C(t) = & \sum_{i=1}^{n} c_i P_0^d(0, T_i) \ (defaultable\ LIBOR\ payment) \\
& + s^{par} \sum_{i=1}^{n} \delta_{i-1} P_0^d(0, T_i) \ (coupon\ spread) \\
& + P_0^d(t, T_n) \ (principal) \\
& + \pi \sum_{i=1}^{n} e(t, T_{i-1}, T_i) \ (recovery)
\end{aligned}
$$

Proof. Note that the current value of LIBOR payment $\tau L(S,T), \tau = T - S$ with default can be expressed as

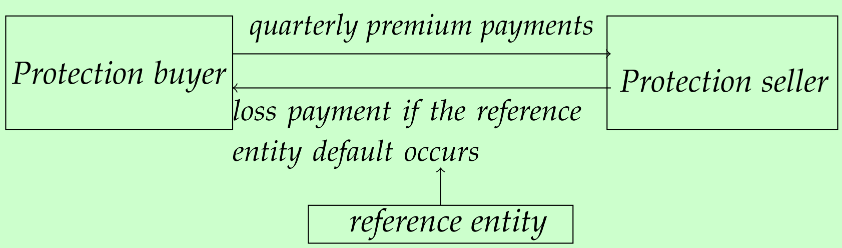
$$
\begin{aligned}
V(t) &= E_Q[\exp(-\int_t^T r(s)ds)\tau L(S,T)\mathbf{1}_{\tau > T}|\mathcal{F}_t] \\
&= E_Q[\exp(-\int_t^T r(s)ds)\tau L(S,T)|\mathcal{F}_t]E_Q[\mathbf{1}_{\tau > T}|\mathcal{F}_t] \\
&= P(t,T)\delta F(t,S,T)S(t,T) \\
&= \delta F(t,S,T)P_0^d(t,T).
\end{aligned}
$$

where we use result in Theorem 3.1.2. □

7.3.3 Credit default swaps

7.3.3.1 *Business of CDS*

Definition 7.3.3 (CDS contract).



CDS convention

- When long CDS, we mean buying protection; we pay premium in exchange for loss payment.
- When short CDS, we mean selling protection; we receive premium in exchange for receipt of loss payment.

Remark 7.3.1 (How to use CDS). link CDS can be used by investor for speculation, hedging, and arbitrage.

- **speculation** CDS allow investors to speculate on changes in CDS spreads of single names or of market indices such as the North American CDX index or the European iTraxx index. For a long position, the widening spread will increase the CDS value.
- **hedging**. CDS are often used to manage the risk of default that arises from holding debt. A bank may hedge its risk that a borrower may default on a loan by buying protection. By offloading a particular credit risk, a bank is not required to hold as much capital in reserve against the risk of default.
- **arbitrage**. A company's stock price and its CDS spread should exhibit negative correlation; i.e., if the outlook for a company improves then its share price should go up and its CDS spread should tighten. For example, if a company has announced some bad news and its share price has dropped by 25%, but its CDS spread has remained unchanged, then an investor can long CDS and wait for the spread to widen and increase the value of the CDS.

7.3.3.2 *standard CDS swap*

Definition 7.3.4 (premium leg coupon payment convention). *Consider a set of coupon payment dates $t_0 < t_1 < \cdots < t_N$ and the protection starting date t_S. t_S satisfies $t_0 \leq t_S < t_1$.*

- *(initial coupon payment convention) Suppose the reference entity defaults after t_1. Even the protection date t_s starts between coupon payment dates t_0, t_1, the contract still requires the buyer to pay the full coupon $(t_1 - t_0) \times C$ instead of a partial coupon of $(t_1 - t_S) \times C$.*
- *Note that only when default $\tau \geq t_i$, the protection buyer will pay the i coupon(protection fee) for protection period i. If default time $\tau < t_{i-1}$, the buyer does not pay the i coupon for protection period i. If default time $t_{i-1} \leq \tau < t_i$, the buyer only pay the partial i coupon for protection period from t_{i-1} to τ.*
- *Details of coupon payment relation and default time are given by*

Default time τ	*PV of full premium payment*	*PV of accrued premium payment*
$t_0 \leq \tau < t_1$	0	$C \times \Delta(t_0, \tau) \times DF(\tau)$
$t_1 \leq \tau < t_2$	$C \times \Delta_1 \times DF(t_1)$	$C \times \Delta(t_0, \tau) \times DF(\tau)$
$t_2 \leq \tau < t_3$	$C \times \Delta_2 \times DF(t_2)$	$C \times \Delta(t_2, \tau) \times DF(\tau)$
$\cdots$	$\cdots$	$\cdots$
$t_{n-1} \leq \tau < t_n$	$C \times \Delta_{n-1} \times DF(t_{n-1})$	$C \times \Delta(t_{n-1}, \tau) \times DF(\tau)$
$t_n \leq \tau$	$C \times \Delta_n \times DF(t_n)$	0

where $\Delta_i = t_i - t_{i-1}$, and $\Delta(t_i, \tau)$ is the year fraction for period $[t_i, \tau]$, and $DF(t)$ is the discount factor from t_S to t.

Lemma 7.3.3 (present value of premium leg). *[6]Consider a set of coupon payment dates $t_0 < t_1 < \cdots < t_N = T$ and the protection starting date and the valuation date be $t_S, t_0 < t_S < t_1$.*

- *The **present value of the premium leg excluding the partial coupon** is given by*

$$\begin{aligned} PV_{fullCoupon}(t_S) &= C \times E_Q[\sum_{i=1}^{N} \Delta_i df(t_S, t_i)\mathbf{1}(\tau > t_i)] \\ &= C \times \sum_{i=1}^{N} \Delta_i DF(t_S, t_i) Q(t_i) \end{aligned}$$

where

$$df(t_S, t_i) = \exp(-\int_{t_0}^{t_i} r(s)ds), DF(t_S, t_i) = E_Q[df(t_S, t_i)|\mathcal{F}_S], Q(t_i) = E[\mathbf{1}(\tau > t_i)] = Pr(\tau > t_i).$$

- *The **present value of the partial coupon**, which pays at default time, is given by*

$$\begin{aligned} PV_{partialCoupon}(t_S) &= C \times E_Q[\sum_{i=1}^{N} (\tau - t_{i-1}) df(t_S, \tau)\mathbf{1}(t_{i-1} \leq \tau \leq t_i)] \\ &= -C \times \sum_{i=1}^{N} (\int_{t_{i-1}}^{t_i} DF(t_S, s)(s - t_{i-1}) dQ(s) \end{aligned}$$

- *The full present **clean value** of the premium leg is given by*

$$\begin{aligned} PV_{premium,clean} &= PV_{fullCoupon} + PV_{partialCoupon} \\ &= C \times \sum_{i=1}^{N} (\Delta_i DF(t_S, t_i) Q(t_i) - \int_{t_{i-1}}^{t_i} DF(t_0, s)(s - t_{i-1}) dQ(s)). \end{aligned}$$

*In particular, we define **RPV01** by*

$$RPV01 = \sum_{i=1}^{N} (\Delta_i DF(t_S, t_i) Q(t_i) - \int_{t_{i-1}}^{t_i} DF(t_0, s)(s - t_{i-1}) dQ(s)).$$

- *If the hazard curve $h(t)$ is piece-wise constant with kinks at $t_1, ..., t_N$, i.e,*

$$h(t) = \begin{cases} \lambda_1, t_S < t < t_1 \\ \lambda_2, t_1 < t < t_2 \\ \cdots \\ \lambda_n, t_{n-1} < t < t_n \end{cases}$$

And further assume the partial coupon is paid at the next immediate coupon date (instead of default time) then

$$PV_{premium,clean} = C \sum_{i=1}^{N} DF(t_S, t_i) [\frac{Q(t_{i-1}) - Q(t_i)}{\lambda_i}]$$

and

$$RPV01 = \sum_{i=1}^{N} DF(t_S, t_i) [\frac{Q(t_{i-1}) - Q(t_i)}{\lambda_i}].$$

Note that we call this clean value because it does not follow the contract spefication.

Proof. (1) straight forward from the cash flow in premium leg (see Remark 7.3.2). Denote $DF_i = DF(t_S, t_i), df_i = df(t_S, t_i), \Delta_i = t_i - t_{i-1}, \Delta_1 = t_1 - t_S$.. Let F_τ, f_τ be the cmf and pdf of default time τ. Let $Q_i = Pr(\tau > t_i) = \int_{t_i}^{\infty} f_\tau(t)dt$. We have

$$\begin{aligned}
PV_{prem.clean} &= \sum_{i=1}^{N} E_Q[C \cdot \Delta_i \cdot \mathbf{1}_{\tau > t_i} \cdot df_i] + E_Q[C \cdot (\tau - t_{i-1}) \cdot \mathbf{1}_{t_{i-1} \leq \tau \leq t_i} \cdot df_i] \\
&= C \sum_{i=1}^{N} df_i [\int_{t_i}^{\infty} \Delta_i f_\tau(t)dt + \int_{t_{i-1}}^{t_i} (t - t_{i-1}) f_\tau(t)dt] \\
&= C \sum_{i=1}^{N} DF_i [\Delta_i PND_i + \int_{t_{i-1}}^{t_i} (t - t_{i-1}) dF_\tau(t)] \\
&= C \sum_{i=1}^{N} DF_i [\Delta_i PND_i + (t_i - t_{i-1}) F(t_i) - \int_{t_{i-1}}^{t_i} F_\tau(t)dt] \\
&= C \sum_{i=1}^{N} DF_i [\Delta_i - \int_{t_{i-1}}^{t_i} F_\tau(t)dt] \\
&= C \sum_{i=1}^{N} DF_i [\int_{t_{i-1}}^{t_i} 1 dt - \int_{t_{i-1}}^{t_i} F_\tau(t)dt] \\
&= C \sum_{i=1}^{N} DF_i [\int_{t_{i-1}}^{t_i} (1 - F_\tau(t))dt] \\
&= C \sum_{i=1}^{N} DF_i [\int_{t_{i-1}}^{t_i} \exp(-\int_{t_0}^{t} \lambda(s)ds)dt] \\
&= C \sum_{i=1}^{N} DF_i [\frac{Q_{i-1} - Q_i}{\lambda_i}].
\end{aligned}$$

□

Remark 7.3.2 (understand the coupon cash flow in the premium leg). Let $t_1 < t_2 < \cdots < t_n$ denote the coupon payment dates. Let τ denote the default time.

- Let t_0 be the current time. The discounted cash flow is

$$\begin{aligned}
CL = &\mathbf{1}(t_1 \leq \tau < t_2) C \Delta_1 DF(t_0, t_1) \\
&+ \mathbf{1}(t_2 \leq \tau < t_3) C(\Delta_1 DF(t_0, t_1) + \Delta_2 DF(t_0, t_2)) \\
&+ \mathbf{1}(t_3 \leq \tau < t_4) C(\Delta_1 DF(t_0, t_1) + \Delta_2 DF(t_0, t_2) + \Delta_3 DF(t_0, t_3)) \\
&+ \cdots \\
&+ \mathbf{1}(t_n \leq \tau) C(\sum_{i=1}^{n} \Delta_i DF(t_0, t_i))
\end{aligned}$$

- The cash flow can be rewritten by

$$
\begin{aligned}
CL = &\, C\Delta_1 DF(t_0,t_1)(\mathbf{1}(t_1 \le \tau < t_2) + \mathbf{1}(t_1 \le \tau < t_2) + \cdots + \mathbf{1}(t_n \le \tau)) \\
&+ C\Delta_2 DF(t_0,t_2)(\mathbf{1}(t_2 \le \tau < t_3) + \mathbf{1}(t_3 \le \tau < t_4) + \cdots + \mathbf{1}(t_n \le \tau)) \\
&+ \cdots \\
&+ C\Delta_n DF(t_0,t_n)\mathbf{1}(t_n \le \tau) \\
&= C\Delta_1 DF(t_0,t_1)\mathbf{1}(t_1 \le \tau) \\
&+ C\Delta_2 DF(t_0,t_2)\mathbf{1}(t_2 \le \tau) \\
&+ \cdots \\
&+ C\Delta_n DF(t_0,t_n)\mathbf{1}(t_n \le \tau)
\end{aligned}
$$

  where we used the equality

$$(\mathbf{1}(t_1 \le \tau < t_2) + \mathbf{1}(t_1 \le \tau < t_2) + \cdots + \mathbf{1}(t_n \le \tau)) = \mathbf{1}(t_1 \le \tau).$$

Definition 7.3.5 (present value of a CDS contract). *Consider the perspective of a protection buyer.*

- *Let t be the current date. Let t_S be the protection starting date and $t \le t_S$. Let t_1 be the next coupon date and t_0 be the previous coupon payment date. Then* ***accrued premium***[a]*, denoted by AP, is given by*

$$AP = C \times (t_S - t_0) \times DF(t,t_S)Q(t,t_S),$$

  where In the case where the current date is the protection starting date, i.e., $t = t_s$, we have

$$AP = C \times (t_S - t_0),$$

  where C is the premium coupon rate and $(t_S - t_0)$ is the year fraction of the no-protection period.
- *The* ***present clean value*** *of a CDS contract is*

$$V_{clean}(t) = PV_{default}(t) - PV_{premium,clean}(t).$$

- *The* ***present dirty value*** *of a CDS contract is*

$$
\begin{aligned}
V(t) &= PV_{default}(t) - PV_{premium,dirty} \\
&= PV_{default}(t) - PV_{premium,clean} - AP.
\end{aligned}
$$

a AP is the positive amount the protection seller should compensate the protection buyer for the no-protection between t_0 and t_S

Remark 7.3.3 (understand dirty price and clean price).
- Dirty value is calculated based on the cash flow specificed by the contract; **Dirty value the fair market value of the cash flows**.

- Clean value calculation is not completely following the contract because at the first coupon date (t_1), only fraction of coupon $(t_1 - t_S) \times C$ is paid instead of the full coupon $(t_1 - t_0) \times C$, which is specified in the contract. Therefore, clean value is not the fair value.
- From the perspective of the buyer, dirty value is smaller than clean value.

7.3.3.3 *Continuous-time CDS pricing*

Lemma 7.3.4 (price of a continuous-time CDS). *Consider a CDS with protection period from 0 to time T. Denote the CDS spread by S^{CDS}, which is the premium paid per unit time by the protection buyer. Assuming deterministic hazard rate function $h(t)$, recovery rate R, and short rate function $r(t)$. We have*

- *The value of the fixed/premium leg at current time $t = 0$ is*

$$V_{fixed}(0) = s^{CDS} \int_0^T \exp(-\int_0^u (h(s) + r(s))ds)du.$$

- *The value of the floating/default leg (paying (1-R) if default is*

$$V_{float}(0) = (1 - R) \int_0^T h(u) \exp(-\int_0^u (h(s) + r(s))ds)du.$$

- $$s^{CDS} = (1 - R) \frac{\int_0^T h(u) \exp(-\int_0^u (h(s) + r(s))ds)du}{\int_0^T \exp(-\int_0^u (h(s) + r(s))ds)du}$$

Proof. (1) We use $0 = T_0 < T_1 < \cdots < T_n = T$ to partition the interval $[0, T]$, we have

$$\begin{aligned}
V_{fixed}(0; n) &= S^{CDS} \sum_{i=1}^{n} (T_i - T_{i-1}) \exp(-\int_0^{T_i} r(s)ds) Pr(\tau > T_i) \\
&= S^{CDS} \sum_{i=1}^{n} (T_i - T_{i-1}) \exp(-\int_0^{T_i} r(s)ds) \exp(-\int_0^{T_i} h(s)ds \\
&= S^{CDS} \sum_{i=1}^{n} (T_i - T_{i-1}) \exp(-\int_0^{T_i} (r(s) + h(s))ds)
\end{aligned}$$

where we the default probability characterization in Lemma 7.1.1 such that

$$Pr(\tau > T_i) = \exp(-\int_0^{T_i} h(s)ds).$$

As we refine the partition $n \to \infty$, we have

$$V_{fixed}(0; \infty) = V_{fixed}(0) = s^{CDS} \int_0^T \exp(-\int_0^u (h(s) + r(s))ds)du.$$

(2)We use $0 = T_0 < T_1 < \cdots < T_n = T$ to partition the interval $[0, T]$, we have

$$\begin{aligned}
V_{float}(0;n) &= (1-R)\sum_{i=1}^{n}\exp(-\int_0^{T_i} r(s)ds)Pr(T_{i-1} < \tau < T_i) \\
&= (1-R)\sum_{i=1}^{n}\exp(-\int_0^{T_i} r(s)ds)S(T_{i-1})(1-\exp(-\int_{T_{i-1}}^{T_i} h(s)ds)) \\
&= (1-R)\sum_{i=1}^{n}\exp(-\int_0^{T_i} r(s)ds)\exp(-\int_0^{T_{i-1}} h(s)ds)(1-\exp(-\int_{T_{i-1}}^{T_i} h(s)ds) \\
&= (1-R)\sum_{i=1}^{n}\exp(-\int_0^{T_i} (r(s)+h(s))ds)(1-\exp(-\int_{T_{i-1}}^{T_i} h(s)ds) \\
&\approx (1-R)\sum_{i=1}^{n}\exp(-\int_0^{T_i} (r(s)+h(s))ds)h(T_i)(T_i - T_{i-1})
\end{aligned}$$

where we the default probability characterization in Lemma 7.1.1 such that

$$Pr(T_1 < \tau < T_2) = S(T_1)(1-\exp(-\int_{T_1}^{T_2} h(s)ds)).$$

As we refine the partition $n \to \infty$, we have

$$V_{float}(0;\infty) = V_{float}(0) = (1-R)\int_0^T h(u)\exp(-\int_0^u (h(s)+r(s))ds)du.$$

□

Lemma 7.3.5 (credit spread triangle). *Consider a CDS with protection period from 0 to time T. Denote the CDS spread by S^{CDS}, which is the premium paid per unit time by the protection buyer. Assuming deterministic hazard rate function $h(t)$, recovery rate R, and short rate function $r(t)$.* ***If $h(t), r(t)$ are constant functions****, we have*

$$h = \frac{S^{CDS}}{(1-R)}.$$

Proof. (1) Note that for constant hazard rate and short rate, we have

$$\begin{aligned}
&\int_0^T h(u)\exp(-\int_0^u (h(s)+r(s))ds)du \\
&= \int_0^T h\exp(-(h+r)u)du \\
&= \frac{h}{h+r}(1-\exp(-(h+r)T))
\end{aligned}$$

and

$$\begin{aligned}
&\int_0^T \exp(-\int_0^u (h(s)+r(s))ds)du \\
&= \int_0^T \exp(-(h+r)u)du \\
&= \frac{1}{h+r}(1-\exp(-(h+r)T))
\end{aligned}$$

therefore,

$$s^{CDS} = (1-R)\frac{\int_0^T h(u)\exp(-\int_0^u (h(s)+r(s))ds)du}{\int_0^T \exp(-\int_0^u (h(s)+r(s))ds)du} = (1-R)h.$$

□

Remark 7.3.4 (Another interpretation for nonconstant function in infinitesimal time intervals)**.**

- The protection buyer pays(on expectation sense) $S^{CDS}\delta t(1-h\delta t)$ during the interval $[t, t+\delta t]$, where $(1-h\delta t)$ is the probability of not defaulting conditioned on no default occurs before t. The protection seller pays(on expectation sense) $h\delta t(1-R)$ during the interval $[t, t+\delta t]$,where $(h\delta t)$ is the probability of defaulting conditioned on no default occurs before t.
  Equaling the two payment and we will get the result.
- Note that we get the equation under the assumption that the discounted cash flow in an infinitesimal time interval should equal; Usually, we require the discounted cash flow in the whole protection period to be equal.

Methodology 7.3.1 (value evolution for a continuous-time CDS). *Consider a CDS with protection period from 0 to time T. Let current time be $t, t \in [0, T]$. Let C be the contract spread.*

- *The current value of CDS contract is given by*

$$\begin{aligned}
V(t) &= V_{float}(t) - V_{fixed}(t) \\
&= (1-R)\int_t^T h(u)\exp(-\int_t^u (h(s)+r(s))ds)du - C\int_t^T \exp(-\int_t^u (h(s)+r(s))ds)du.
\end{aligned}$$

- *Assume the hazard rate and short rate are flat, i.e.,* $h(t) = h_0, r(t) = r_0$. *Then the value of the floating/default leg (paying (1-R) if default is*

$$\begin{aligned} V(t) &= ((1-R)h_0 - C)\int_t^T \exp(-\int_t^u (h_0 + r_0)ds)du \\ &= ((1-R)h_0 - C)\frac{1 - \exp(-(r_0 + h_0)(T-t))}{r_0 + h_0} \end{aligned}$$

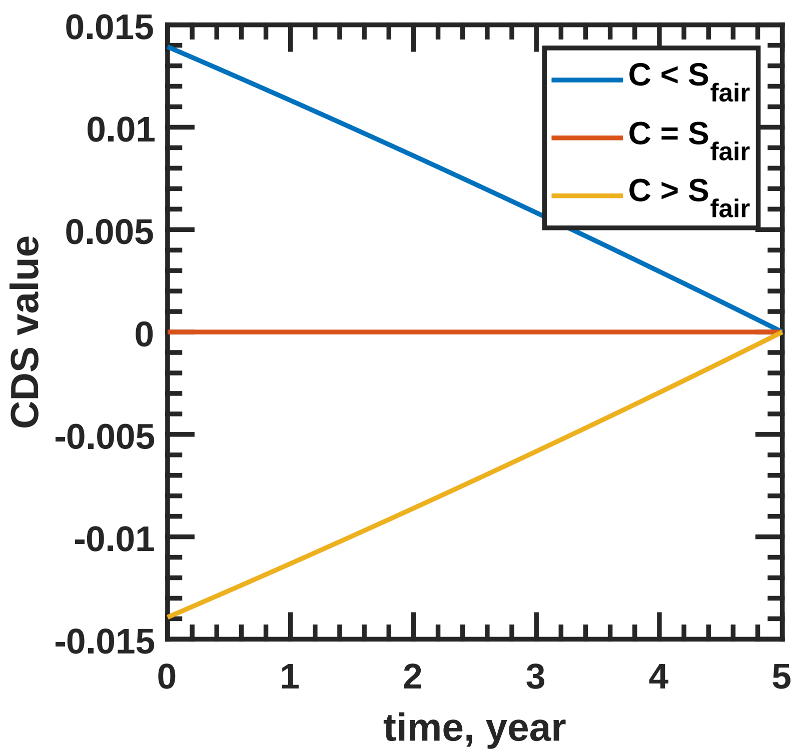


Figure 7.3.1: Value evolution for a continuous-time CDS

7.3.3.4 *Monte Carlo method for CDS pricing*

Methodology 7.3.2. *Suppose we are given by the cdf of default time* τ, *denoted by* F_τ.

We can use the following procedure to generate

- *generate a sample* $\tau^{(i)}$ *using* F_τ.
- *evaluate the premium leg value* $V^{(i)}_{premium}$ *and default leg value* $V^{(i)}_{protection}$
- *generate a sample value* $V^{(i)} = V^{(i)}_{protection} - V^{(i)}_{premium}$.

- $$\hat{V}_{clean} = \frac{1}{N}\sum_{i=1}^{N} V^{(i)}.$$

7.3.3.5 *Risk management of CDS*

Lemma 7.3.6 (risk due to interest rate, qualitative result). *Consider a long position in CDS such that is value is default payment minus premium payment(see Figure 7.3.2).*

- *If the reference entity has relatively low probability to default, then the value of CDS will increase when interest rate increases.*
- *If the reference entity has relatively high probability to default, then the value of CDS will decrease when interest rate increases.*

Proof. (informal) Note that when interest rate increases, the default payment and the premium payment will both decrease since discount factors are decreasing. We can view default payment and premium as certain form of bonds and we compare their dollar duration. (1) for entity with low probability to default, the dollar duration of default payment is small since it is very likely not to receive any default payment; on the other hand, the dollar duration for the premium payment is large, since it is very likely the CDS buyer has to finish all the payments. (2) for entity with high probability to default, the dollar duration of default payment is big since it is very likely to receive the default payment; on the other hand, the dollar duration for the premium payment is small, since it is very likely the CDS buyer will stop premium payment when default occurs half-way. □

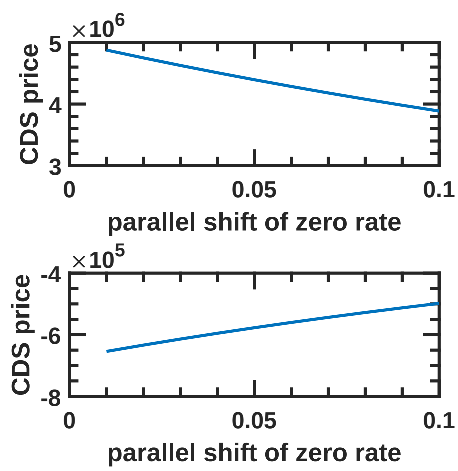


Figure 7.3.2: Effect of interest rate parallel shifting up on CDS price. (top) high probability of default case; (below) low probability of default case.

Lemma 7.3.7 (CDS time risk).

- *In general, for increasing hazard rate curve, the clean price of a CDS long position will increase when approaching maturity; for decreasing hazard rate curve, the clean price of a CDS long position will decrease when approaching maturity; For perfectly flat hazard curve, the clean price*
- *Assume zero interest rate and flat hazard curve, the time path for a continuously CDS is given by*

Proof. (1)(informal) the long position of a CDS will receive default payment and pay out premium. When approaching maturity, the premium payment is decreasing, □

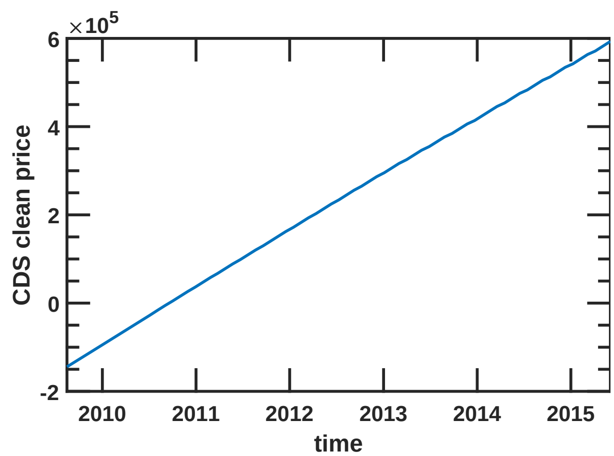


Figure 7.3.3: CDS clean price(buy protection) time path for a CDS contract maturing on 20-Sep-2015.

7.4 Default correlation modeling

7.4.1 Sampling correlated default time

Let τ be the random default time with cdf F_τ. Then we can sample τ in the following two ways:

- Draw a random number U with uniform distribution $U([0,1])$, then

$$\tau = F_\tau^{-1}(u).$$

Or equivalently, $\tau < t$ if and only if $U < F_\tau(t)$,i.e.,

$$Pr(\tau < t) = Pr(u < F_\tau(t)) = F_\tau(t).$$

- Draw a random number $Z \sim N(0,1)$, then

$$\tau = F_\tau^{-1}(\phi(u)),$$

where ϕ is the cdf for a standard normal distribution. Or equivalently, $\tau < t$ if and only if $Z < \phi^{-1}(F_\tau(t))$,i.e.,

$$Pr(\tau < t) = Pr(Z < \phi^{-1}(F_\tau(t)) = F_\tau(t).$$

Methodology 7.4.1 (generation of dependent default time). *Let $T_1, T_2, ..., T_n$ denote the random default time for n parties. Assume the hazard curve for each party is given by $h_i(t), t \geq 0$ such that the marginal cdf of default time is given by*

$$F_i(t) = Pr(T_i \leq t) = 1 - \exp(-\int_0^t h(s)ds).$$

Further assume the copula associated with the joint cdf is Gaussian copula with correlation matrix R.

Consider the following random number generating process

- *Simulate $Y_1, Y_2, ..., Y_n$ from $MN(0, R)$.*
- *Obtain $T_1, T_2, ..., T_n$ using $T_i = F_i^{-1}(\phi(Y_i))$, where ϕ is the cdf of a standard normal.*

$T_1, T_2, ..., T_n$ will follow the joint cdf F.

Note 7.4.1 (first default time random sample from simulation). In each simulation we generate one sample of default times $t_1, t_2, ..., t_n$. The first-to-default time is simply $t = \min(t_1, t_2, ..., t_n)$.

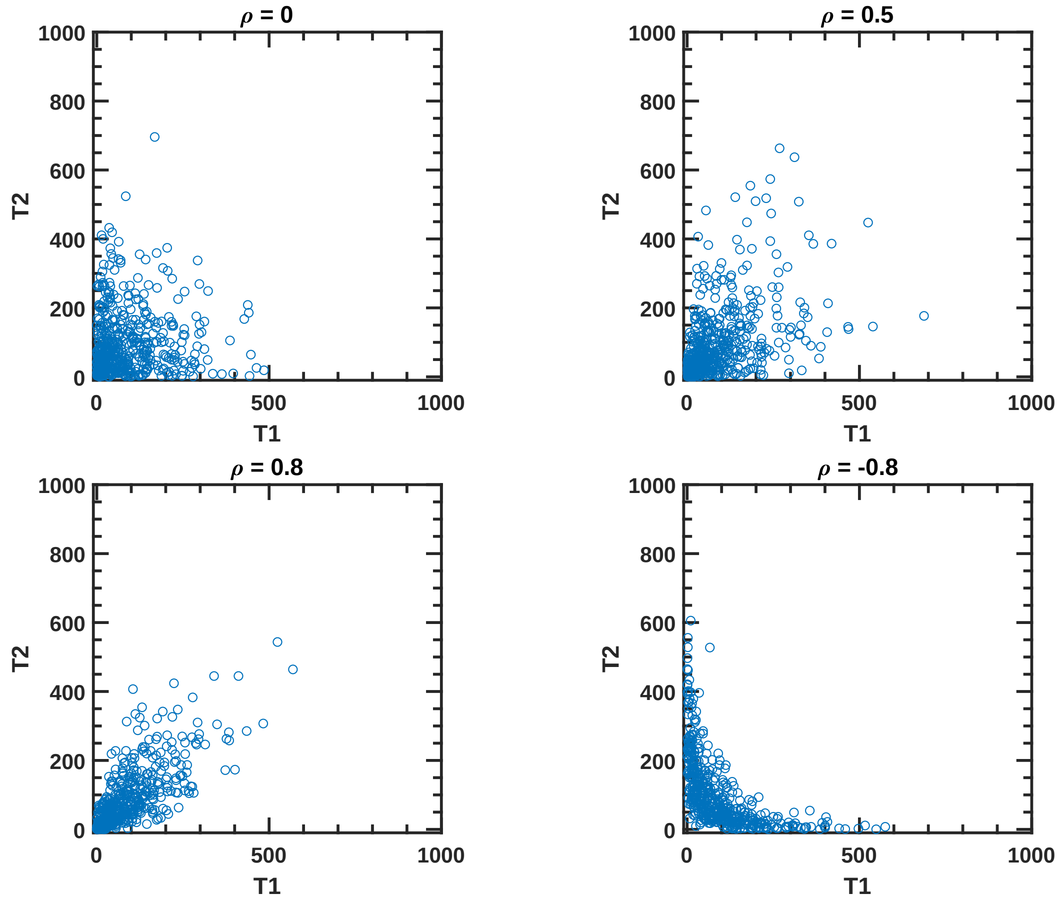


Figure 7.4.1: Generated correlated default time via Gaussian copula with different correlations. The hazard rate for both parties is $h(t) = 0.01$.

7.4.2 First-to-default modeling and valuation

7.4.2.1 *General principles*

Lemma 7.4.1 (first-to-default probability and contract valuation under independence assumption). *Let $T_1, T_2, ..., T_n$ denote the **independent** random default time for n parties. Assume the hazard curve for each party is given by $h_i(t), t \geq 0$ such that the marginal cdf of default time is given by*

$$F_i(t) = Pr(T_i \leq t) = 1 - \exp(-\int_0^t h(s)ds).$$

It follows that

- *The first-to-default time* $T = \min\{T_1, T_2, ..., T_n\}$ *has cdf and density given by*

$$F_T(t) = 1 - \exp(-\int_0^t \sum_{i=1}^n h_i(s)ds), f_T(t) = (\sum_{i=1}^n h_i(s))\exp(-\int_0^t \sum_{i=1}^n h_i(s)ds).$$

- *Consider a contract maturing at* T_c *which pays 1 if the first default among n parties occur between* 0 *and* T_c*. The value of the contract at time 0 is*

$$V(0) = \int_0^{T_c} P(0,u)f_T(t)dt.$$

where $P(0,u)$ *is the zero coupon bond price at 0 with maturity* u*.*

Proof. (1) We know that

$$F_T(t) = 1 - \prod_{i=1}^n (1 - F_i(t)) = 1 - \prod_{i=1}^n \exp(-\int_0^t h_i(s)ds) = 1 - \exp(-\int_0^t \sum_{i=1}^n h_i(s)ds).$$

(2) Straight forward. □

7.4.2.2 *Multivariate Gaussian copula approximation*

Lemma 7.4.2 (Multivariate Gaussian copula approximation for joint default modeling). *Let* $T_1, T_2, ..., T_n$ *denote the random default time for n parties. Assume marginal cdf of default time is given by* $F_i(t)$*. Further* ***assume the copula associated with the joint cdf*** *is Gaussian copula with correlation matrix R.*

- *Then* $T_1, T_2, ..., T_n$ *will follow the joint cdf F given by*

$$F_{T_1,T_2,...,T_n}(t_1, t_2, ..., t_n) = \Phi(\phi^{-1}(F(T_1)), \phi^{-1}(F(T_2)), ..., \phi^{-1}(F(T_n))).$$

- *Let* $T = \min(T_1, T_2, ..., T_n)$ *be the first default time. Then*

$$F_T(t) = 1 - F_{T_1,T_2,...,T_n}(t, t, ..., t)$$

Proof. (1) Directly from multivariate Gaussian copula properties. (2)

$$\begin{aligned} 1 - F_T(t) &= Pr(T > t) \\ &= Pr(T_1 > t, T_2 > t, ..., T_n > t) \\ &= \end{aligned}$$

□

7.4.2.3 *Gaussian one-factor model*

Lemma 7.4.3 (Multivariate Gaussian copula approximation for Bernoulli default modeling). *Consider n parties that will default in the next period with unconditional probability p_i. Define a new proxy random variable X_i by*

$$X_i = a_i F + \sqrt{1-a_i^2} Z_i, i = 1, 2, ..., n,$$

where F is a common factor affecting defaults for all firms and Z_i is a factor affecting only firm i. F and Z_i are independent standard normal variables. It follows that

- $$Pr(X_i \text{ will default}) \triangleq p_i = Pr(X_i < \phi^{-1}(p_i)),$$

  where ϕ is the cdf of the standard normal variable.
- *The **conditional default probability** of firm i before t conditioning on the observation of F is given by*

  $$Pr(X_i \text{ will default}|F = f) \triangleq Pr(T_i < t|F = f) = \phi(\frac{\phi^{-1}(Q_i(t)) - a_i f}{\sqrt{1-a_i^2}}),$$

- *The **unconditional probability of no default** is given by*

  $$Pr(\text{no default}) \triangleq \int_{-\infty}^{\infty} \prod_{i=1}^{n} (1 - Pr(X_i \text{ will default}|F = f)) g(f) df$$

  where $g(f) = \frac{1}{\sqrt{2\pi}} \exp(-f^2/2)$.

Proof. (1) Note that $X_i \sim N(0,1)$. Therefore,

$$Pr(X_i < \phi^{-1}(p_i)) = \phi(\phi^{-1}(p_i)) = p_i.$$

(2)

$$\begin{aligned} Pr(X_i < \phi^{-1}(p_i)) &= Pr(a_i F + \sqrt{1-a_i^2} Z_i < \phi^{-1}(p_i)) \\ &= Pr(Z_i < \frac{\phi^{-1}(p_i) - a_i F}{\sqrt{1-a_i^2}}) \end{aligned}$$

(3) straight forward. □

Lemma 7.4.4 (Multivariate Gaussian copula approximation for joint default modeling with single factor). *[7, p. 563] Let $T_1, T_2, ..., T_n$ denote the random default time for n parties. Assume marginal cdf of default time is given by $Q_i(t)$. Define a new proxy random variable $X_i = \phi^{-1}(Q_i(T_i)), i = 1, 2, ..., n$, and **assume** X_i can be modeled by*

$$X_i = a_i F + \sqrt{1 - a_i^2} Z_i, i = 1, 2, ..., n,$$

where F is a common factor affecting defaults for all firms and Z_i is a factor affecting only firm i. F and Z_i are independent standard normal variables. It follows that

- *The **conditional default probability** of firm i before t conditioning on the observation of F is given by*

$$Q_i(t|F = f) \triangleq Pr(T_i < t|F = f) = \phi(\frac{\phi^{-1}(Q_i(t)) - a_i f}{\sqrt{1 - a_i^2}}),$$

  where ϕ is the standard normal cdf.
- *The **unconditional default probability** is given by*

$$Q_i(t) = \int_{-\infty}^{\infty} \phi(\frac{\phi^{-1}(Q_i(T)) - a_i f}{\sqrt{1 - a_i^2}}) g(f) df,$$

  where $g(f) = \frac{1}{\sqrt{2\pi}} \exp(-f^2/2)$.
- *Conditioning on the observation of F, the conditional probability of having m defaults among N firms is given by*

$$Pr(X = m|F = f) = \binom{N}{m} (p(y))^m (1 - p(y))^{N-m}.$$

  and the unconditional probability is given by

$$Pr(X = m) = \int_{-\infty}^{\infty} \binom{N}{m} (p(y))^m (1 - p(y))^{N-m} g(f) df.$$

- *Let* $T = \min(T_1, T_2, ..., T_n)$. *Then the* ***conditional and unconditional first-to-default probabilities*** *are given by*

$$\begin{aligned} Q_T(t|F=f) &\triangleq Pr(T<t|F=f) \\ &= 1-(1-Q_i(t|F=f))^n \\ Q_T(t) &\triangleq Pr(T<t) \\ &= 1-\int_{-\infty}^{\infty}(1-Q_i(t|F=f))^n g(f)df. \end{aligned}$$

Proof. (1)

$$\begin{aligned} Pr(T_i<t) &= Pr(T_i<t) \\ &= Pr(Q_i(T_i)<Q_i(t)) \\ &= Pr(\phi^{-1}(Q_i(T_i))<\phi^{-1}(Q_i(t))) \\ &= Pr(X_i<\phi^{-1}(Q_i(t))) \\ &= Pr(a_iF+\sqrt{1-a_i^2}Z_i<\phi^{-1}(Q_i(t))) \\ &= Pr(Z_i<\frac{\phi^{-1}(Q_i(t))-a_iF}{\sqrt{1-a_i^2}}) \\ \implies Q_i(T|F=f)\triangleq Pr(T_i<t|F) &= Pr(Z_i<\frac{\phi^{-1}(Q_i(T))-a_iF}{\sqrt{1-a_i^2}}|F=f) \\ &= \phi(\frac{\phi^{-1}(Q_i(t))-a_if}{\sqrt{1-a_i^2}}) \end{aligned}$$

(2) Use the fact that

$$Pr(T_i<t,F=f)=Pr(T_i<t|F=f)f_F(f),$$

and then marginalize out F. (3) Note that when conditioning on the F, the default behavior among firms are independent. Therefore the number of defaulting firms can be characterized by binomial distribution. (4) Note that when conditioning on the F, the default time T_is are independent from each other. Then we the first-to-default probability calculation under independence assumption result(see Lemma 7.4.1). □

Remark 7.4.1 (interpret proxy random variable and its representation)**.** Note that we define a proxy random variable X_i by $X_i = \phi^{-1}(Q_i(T_i))$. We can see that

- X_i is a standard normal variable because

$$Pr(X_i<x)=Pr(\phi^{-1}(Q_i(T_i))<x)=Pr(Q_i(T_i)<\phi(x))=\phi(x),$$

where we used the fact that $Q_i(T_i)$ is a uniform random variable.

- We represent $X_i = a_i F + \sqrt{1-a_i^2} Z_i$ is consistent. Because $a_i F + \sqrt{1-a_i^2} Z_i \sim N(0,1)$.

Remark 7.4.2 (the correlation structure for the Gaussian copula implied by the factor model)**.** Consider a one factor model given by

$$V_i = a_i Y + \sqrt{1-a_i^2} Z_i, i = 1,2,...,n,$$

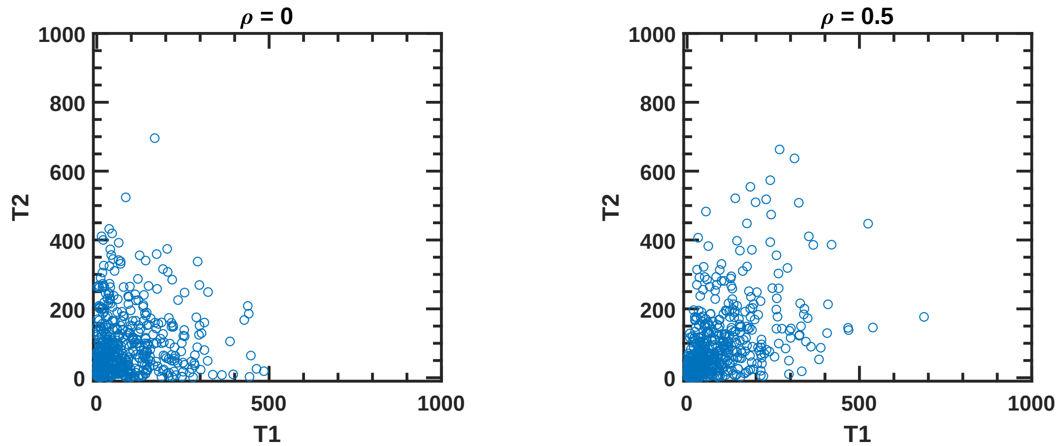
where $(F, Z_1, Z_2, ..., Z_n)$ are independent standard normal variables.

Then the correlation structure implied by the factor model is given by

$$\begin{aligned}
Cov &= \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{bmatrix} [a_1, a_2, ..., a_n] + \begin{bmatrix} 1-a_1^2 & & & \\ & 1-a_2^2 & & \\ & & \ddots & \\ & & & 1-a_n^2 \end{bmatrix} \\
&= \begin{bmatrix} a_1^2 & a_1 a_2 & \cdots & a_1 a_n \\ a_2 a_1 & a_2 a_2 & \cdots & a_2 a_n \\ \vdots & \cdots & \ddots & \vdots \\ a_n a_1 & a_n a_2 & \cdots & a_n a_n \end{bmatrix} \begin{bmatrix} 1-a_1^2 & & & \\ & 1-a_2^2 & & \\ & & \ddots & \\ & & & 1-a_n^2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & a_1 a_2 & \cdots & a_1 a_n \\ a_2 a_1 & 1 & \cdots & a_2 a_n \\ \vdots & \cdots & \ddots & \vdots \\ a_n a_1 & a_n a_2 & \cdots & 1 \end{bmatrix}
\end{aligned}$$

(a) Calculation of first default probability(thick solid black) from individual default probability of 10 reference names using Gaussian one factor model with $\rho = 0.5$.

(b) First default probability as a funciton of correlation of the underlying names.



Remark 7.4.3 (the correlation structure for the Gaussian copula implied by the factor model). Consider a one factor model given by

$$V_i = \sqrt{\rho} Y + \sqrt{1-\rho} Z_i, i = 1, 2, ..., n,$$

where $(F, Z_1, Z_2, ..., Z_n)$ are independent standard normal variables.

Then the correlation structure implied by the factor model is given by

$$Cov = \begin{bmatrix} \sqrt{\rho} \\ \sqrt{\rho} \\ \cdots \\ \sqrt{\rho} \end{bmatrix} [\sqrt{\rho}, \sqrt{\rho}, ..., \sqrt{\rho}] + \begin{bmatrix} 1-\rho & & & \\ & 1-\rho & & \\ & & \ddots & \\ & & & 1-\rho \end{bmatrix} = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \cdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}$$

7.4.2.4 *Gaussian multi-factor model*

Lemma 7.4.5 (Multivariate Gaussian copula approximation for joint default modeling with multiple factors). *Let $T_1, T_2, ..., T_n$ denote the random default time for n parties. Assume* *marginal cdf of default time is given by* $Q_i(t)$. *Define a new proxy random variable* $X_i = \phi^{-1}(Q_i(T_i)), i = 1,2,...,n$, *and* ***assume*** X_i *can be modeled by*

$$X_i = \sum_{j=1}^{m} a_{ij}F_j + \sqrt{1 - \sum_{j=1}^{m} a_{ij}^2} Z_i, i = 1,2,...,n,$$

where $F_1, F_2, ..., F_m$ *are common factors affecting defaults for all firms and* Z_i *is a factor affecting only firm i.* $F_1, F_2, ..., F_m$ *and* Z_i *are mutually independent standard normal variables. It follows that*

- *The* ***conditional default probability*** *of firm i before t conditioning on the observation of* $F_1, F_2, ..., F_m$ *is given by*

$$Q_i(t|F_1 = f_1, ..., F_m = f_m) \triangleq Pr(T_i < t|F_1 = f_1, ..., F_m = f_m) = \phi(\frac{\phi^{-1}(Q_i(t)) - \sum_{j=1}^m a_{ij}f_j}{\sqrt{1 - \sum_{j=1}^m a_{ij}^2}}),$$

  where ϕ *is the standard normal cdf.*
- *The* ***unconditional default probability*** *is given by*

$$Q_i(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \phi(\frac{\phi^{-1}(Q_i(t)) - \sum_{j=1}^m a_{ij}f_j}{\sqrt{1 - \sum_{j=1}^m a_{ij}^2}})g(f_1)...g(f_m)df_1...df_m,$$

  where $g(f) = \frac{1}{\sqrt{2\pi}}\exp(-f^2/2)$.
- *Let* $T = \min(T_1, T_2, ..., T_n)$. *Then the* ***conditional and unconditional first-to-default probabilities*** *are given by*

$$\begin{aligned}
&Q_T(t|F_1 = f_1, ..., F_m = f_m) \\
&\triangleq Pr(T < t|F_1 = f_1, ..., F_m = f_m) \\
&= 1 - (1 - Q_i(t|F_1 = f_1, ..., F_m = f_m))^n \\
&Q_T(t) \\
&\triangleq Pr(T < t) \\
&= 1 - \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (1 - Q_i(t|F_1 = f_1, ..., F_m = f_m))^n g(f_1)...g(f_m)df_1...df_m.
\end{aligned}$$

Proof. See the single factor model(Lemma 7.4.4). □

7.4.3 Structural model approximation

7.4.3.1 *Factor model*

Lemma 7.4.6 (one-factor default model). *[2, p. 308] Assume the asset values of the firms are driven by one common driving factor Y such that the firm's asset is modeled by*

$$V_i = \sqrt{\rho}Y + \sqrt{1-\rho}Z_i, i = 1, 2, ..., N$$

where $Y \sim N(0, \sigma_Y^2)$, $\epsilon_i \sim N(0, w_i^2)$, and $(Y, Z_1, Z_2, ..., Z_N)$ are independent.

- *If a firm will default when its asset value is below K, then the default probability conditioning on the observation of factor value Y is given by*

$$p(y) = \Phi(\frac{K - \sqrt{\rho}y}{\sqrt{1-\rho}w_i}),$$

  where Φ is the cdf for standard normal.
- *The probability of having m defaults among N firms is given by*

$$Pr(X = m|F = f) = \binom{N}{m}(p(y))^m(1-p(y))^{N-m}.$$

Proof. (1)

$$\begin{aligned}
p(y) &\triangleq Pr(V_i < K|Y = y) \\
&= Pr(\sqrt{\rho}Y + \sqrt{1-\rho}Z_i < K|Y = y) \\
&= Pr(Z_i < \frac{K - \sqrt{\rho}y}{\sqrt{1-\rho}}|Y = y) \\
&= \Phi(\frac{K - \sqrt{\rho}y}{\sqrt{1-\rho}w_i})
\end{aligned}$$

(2) Note that when conditioning on the Y, the default behavior among firms are independent. Therefore the number of defaulting firms can be characterized by binomial distribution. □

Remark 7.4.4 (interpretation). [2, p. 307] The systematic risk factor Y can be viewed as an indicator of the state of the business cycle, such as GDP, unemployment rate etc; and the idiosyncratic factor ϵ_n as a firm-specific effects factor such as the management strategy of the firm or the innovation of firm.

Example 7.4.1. Consider a one factor model given by

$$V_i = \sqrt{\rho}Y + \sqrt{1-\rho}Z_i, i = 1, 2, ..., n,$$

where $(F, Z_1, Z_2, ..., Z_n)$ are independent standard normal variables.

Then the covariance structure implied by the factor model is given by

$$\begin{aligned}
Cov &= \begin{bmatrix} \sqrt{\rho} \\ \sqrt{\rho} \\ \cdots \\ \sqrt{\rho} \end{bmatrix} [\sqrt{\rho}, \sqrt{\rho}, ..., \sqrt{\rho}] + \begin{bmatrix} 1-\rho & & & \\ & 1-\rho & & \\ & & \ddots & \\ & & & 1-\rho \end{bmatrix} \\
&= \begin{bmatrix} \rho & \rho & \cdots & \rho \\ \rho & \rho & \cdots & \rho \\ \vdots & \cdots & \ddots & \vdots \\ \rho & \rho & \cdots & \rho \end{bmatrix} \begin{bmatrix} 1-\rho & & & \\ & 1-\rho & & \\ & & \ddots & \\ & & & 1-\rho \end{bmatrix} \\
&= \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \cdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}
\end{aligned}$$

which is also a valid correlation matrix.

Lemma 7.4.7 (multiple factor default model). *[2, p. 312] Assume the asset values of the firms are driven by M driving factors $Y_j, j \leq M$ such that the firm's asset is modeled by*

$$V_i = \sum_{j=1}^{M} \beta_j Y_j + \epsilon_i, i = 1, 2, ..., N$$

where $Y = (Y_1, Y_2, ..., Y_N) \sim MN(0, \Omega_Y)$, $\epsilon_i \sim N(0, w_i^2)$, and $(Y, Z_1, Z_2, ..., Z_N)$ are independent. It follows that

- *If a firm will default when its asset value is below K, then the default probability conditioning on the observation of factor value $Y_1, Y_2, ..., Y_n$ is given by*

$$p(y) = \Phi(\frac{K - \sum_{i=1}^{M} \beta_i y_i}{w_i}),$$

where Φ is the cdf for standard normal.

- *The probability of having m defaults among N firms is given by*

$$Pr(X = m | Y = y) = \binom{N}{m} (p(y))^m (1 - p(y))^{N-m}.$$

Proof. (1)

$$\begin{aligned} p(y) &\triangleq Pr(V_i < K | Y = y) \\ &= Pr(\sum_{i=1}^{M} \beta_i y_i + Z_i < K | Y = y) \\ &= Pr(Z_i < K - \sum_{i=1}^{M} \beta_i y_i | Y = y) \\ &= \Phi(\frac{K - \sum_{i=1}^{M} \beta_i y_i}{w_i} \end{aligned}$$

(2) Note that when conditioning on the Y, the default behavior among firms are independent. Therefore the number of defaulting firms can be characterized by binomial distribution. □

7.4.3.2 *Asset classes model*

Definition 7.4.1. *[2, p. 313]Assume that the portfolio consists of two classes of obligors C_1 and C_2. There are N_1 obligors of class C_1 and N_2 obligors of class C_2. We assume the asset value follows*

$$V_{n_1} = \frac{1}{\sqrt{\beta_{11}^2 + \beta_{12}^2 + 1}} (Y_1 \beta_{11} + Y_2 \beta_{12} + \epsilon_{n_1}), \forall n_1 \in C_1$$
$$V_{n_2} = \frac{1}{\sqrt{\beta_{21}^2 + \beta_{22}^2 + 1}} (Y_1 \beta_{21} + Y_2 \beta_{22} + \epsilon_{n_2}), \forall n_2 \in C_2$$

where two factors Y_1 and Y_2 and the noises ϵ_{n_1} and ϵ_{n_2} are independently standard normal variables.

Lemma 7.4.8. *[2, p. 313] Conisder a two-asset-classes model. It follows that*

- *The obligors' assets within one class are correlated with a correlation coefficient of ρ_1 and ρ_2 respectively, where*

$$\rho_1 = \frac{\beta_{11}^2 + \beta_{12}^2}{1 + \beta_{11}^2 + \beta_{12}^2}$$
$$\rho_2 = \frac{\beta_{21}^2 + \beta_{22}^2}{1 + \beta_{21}^2 + \beta_{22}^2}$$

- *The correlation of two obligors of different classes is*

$$\rho_{12} = \frac{\beta_{11}\beta_{21} + \beta_{12}\beta_{22}}{\sqrt{1 + \beta_{11}^2 + \beta_{12}^2}\sqrt{1 + \beta_{21}^2 + \beta_{22}^2}}$$

Proof. (1) Note that $Var[V_{n_1}] = Var[V_{n_2}] = 1$. Therefore

$$\begin{aligned}
\rho_1 &= E[\frac{1}{\sqrt{\beta_{11}^2 + \beta_{12}^2 + 1}}(Y_1\beta_{11} + Y_2\beta_{12} + \epsilon_{n_1})\frac{1}{\sqrt{\beta_{11}^2 + \beta_{12}^2 + 1}}(Y_1\beta_{11} + Y_2\beta_{12} + \epsilon_{n_1'})] \\
&= \frac{1}{1 + \beta_{11}^2 + \beta_{12}^2}(\beta_{11}^2 Var[Y_1] + \beta_{12}^2 Var[Y_2]) \\
&= \frac{\beta_{11}^2 + \beta_{12}^2}{1 + \beta_{11}^2 + \beta_{12}^2}.
\end{aligned}$$

Similarly,we can derive ρ_2. (2) Note that

$$\begin{aligned}
\rho_{12} &= E[\frac{1}{\sqrt{\beta_{11}^2 + \beta_{12}^2 + 1}}(Y_1\beta_{11} + Y_2\beta_{12} + \epsilon_{n_1})\frac{1}{\sqrt{\beta_{21}^2 + \beta_{22}^2 + 1}}(Y_1\beta_{21} + Y_2\beta_{22} + \epsilon_{n_2})] \\
&= \frac{1}{\sqrt{\beta_{11}^2 + \beta_{12}^2 + 1}}\frac{1}{\sqrt{\beta_{21}^2 + \beta_{22}^2 + 1}}(\beta_{11}\beta_{21} Var[Y_1] + \beta_{12}\beta_{22} Var[Y_2]) \\
&= \frac{\beta_{11}\beta_{21} + \beta_{12}\beta_{22}}{\sqrt{1 + \beta_{11}^2 + \beta_{12}^2}\sqrt{1 + \beta_{21}^2 + \beta_{22}^2}}.
\end{aligned}$$

□

7.5 Notes on bibliography

Major references are [4][2][8].

For default correlation modeling, see [9].

For CDS valuation, see [6][10].

For estimation of default term structure, see [11][12][13][14]

CDS spread data, Markit

BIBLIOGRAPHY

1. Pereira, J. P. *Credit Risk* (Universidade Nova de Lisboa, 2016).
2. Schönbucher, P. J. *Credit derivatives pricing models: models, pricing and implementation* (John Wiley & Sons, 2003).
3. Ammann, M. *Credit risk valuation: methods, models, and applications* (Springer Science & Business Media, 2013).
4. O'Kane, D. *Modelling single-name and multi-name credit derivatives* (John Wiley & Sons, 2011).
5. Bluhm, C., Overbeck, L. & Wagner, C. *Introduction to credit risk modeling* (Crc Press, 2016).
6. White, R. The pricing and risk management of credit default swaps, with a focus on the isda model. *OpenGamma Quantitative Research* **16** (2013).
7. Hull, J. *Options, Futures and Other Derivatives 9 edition* (Pearson Education, 2009).
8. Wilmott, P. *Paul Wilmott on quantitative finance* (John Wiley & Sons, 2013).
9. Li, D. X. On default correlation: A copula function approach (1999).
10. O'Kane, D. & Turnbull, S. Valuation of credit default swaps. *Lehman Brothers quantitative credit research quarterly* **2003** (2003).
11. Berd, A, Mashal, R. & Wang, P. Estimating implied default probabilities from credit bond prices. *Fixed income-quantitative credit research, Lehman Brothers* **7** (2003).
12. Berd, A. M., Mashal, R. & Wang, P. Defining, estimating and using credit term structures. Part 1: Consistent Valuation Measures (2009).
13. Berd, A. M., Mashal, R. & Wang, P. Defining, estimating and using credit term structures. Part 2: Consistent Risk Measures (2004).
14. Berd, A. M., Mashal, R. & Wang, P. Defining, estimating and using credit term structures. Part 3: Consistent CDS-Bond Basis (2004).

8

COMMODITY MODELING

8.1 Commodity market basics

8.1.1 Overview

Common commodities include:

- Agricultural products - corn, soybean, wheat
- Energy products - WTI crude oil, Brent crude oil, natural gas
- Precious metals - gold, silver, platinum
- Industrial metals - copper, aluminum, tin
- Soft commodities - coffee, cocoa, sugar

Spot market generally does not exist since commodities need to be transported and stored. Major markets are futures, including electricity, crude oil, heating oil, natural gas. Contracts may be for physical delivery or financially settlement.

8.1.2 Forward curve basics

8.1.2.1 *Forward curve*

Definition 8.1.1 (futures contract and forward curve for commodities). *[1, p. 71]*

- *Futures contracts on commodities have delivery dates or maturiteis ranging from 1W to 3Y.*
- *The future contract with the nearest delivery date is called prompt forward contract.*
- *Let current time be t. Denote the forward/futures price associated with maturity T contract as $F(t,T)$. The forward curve prevailing at date t for a given commodity is a graphical representation of the set $\{F(t,T), T \geq t\}$ of forward prices for different maturities T.*

Definition 8.1.2 (contango, backwardation). *[1, p. 73]*

- *Most commodities forward curve have volatile short-ends, and quasi-stable long-ends. One important explanation is that the time period prior to delivery allows for the production process to adjust for shocks in supply and demand.*
- *Long-end is determined by marginal cost of production.*
- *Short-end governed by short termed supply and demand:*
  - *When there are excesses supply of commodity, curve is upward sloping(**contango**). From the spot-forward relation(Lemma 8.1.1) $F(t,T) = S(t)\exp((r-y)(T-t))$, when $r - y \geq 0$, we can observe contango.*

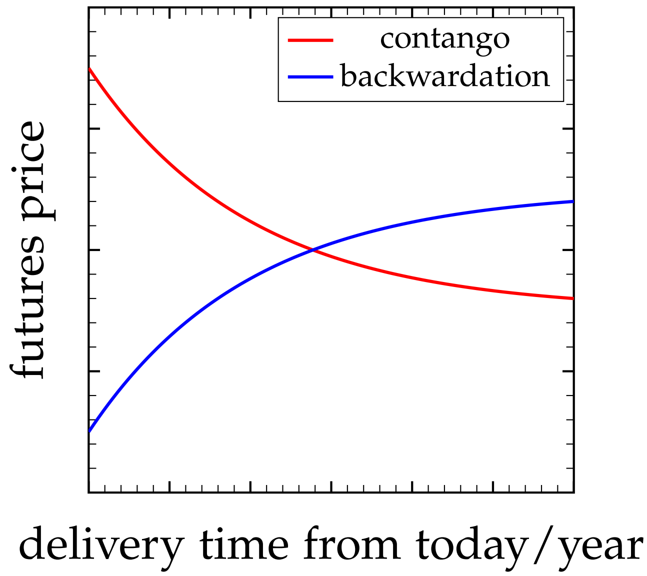


Figure 8.1.1: contango and backwardation forward curves

– *When there are shortages of commodity, curve is downward sloping(**backwardation**).From the spot-forward relation(Lemma 8.1.1) $F(t,T) = S(t)\exp((r-y)(T-t))$, when $r-y<0$, that is we have large convenience yield due to the shortage of supply, we can observe backwardation.*

See Figure 8.1.1 for contango and backwardation.

Definition 8.1.3 (commodity swap).

- *Swaps entail exchanging a fixed payment stream for a floating payment stream.*
- *Floating stream is typically linked to a commodity spot price or price index*
- *The floating leg can be viewed as a series of forward contract.*
- *The fixed leg can be viewed as a series of coupon payments*
- *Fixed leg payments determined such that both legs have equal value.*

Definition 8.1.4 (commodity options).

- *European Call options on commodity forward contracts*
- *Asian Call options on commodity forward contracts*
- *Floating strike call options are call options with the strike level set at a future date based on some price index. Both the commodity forward price and the index price are uncertain.*

- *A calendar spread option is an option to exchange a T_2 maturity forward contract for a T_1 maturity forward contract at a cost K at time T, where $T_1, T_2 > T$. The payoff at time T is*

$$V(T) = (F(T, T_1) - F(T, T_2) - K)^+.$$

- *Intercommodity spread options are options based on the difference in two commodities. The payoff at maturity T is given by*

Definition 8.1.5 (convenience yield). *[1, p. 24]*

- ***Convenience yield** is a benefit accrues to the owner of the physical commodity but not to the holder of a forward contract. Analogously, the dividend yield is paid to the owner of a stock but not the holder of a derivative contract written on the stock.*
- *The convenience yield is expressed as rate y such that the benefit of holding the commodity worth spot $S(t)$ will equal to $S(t)ydt$ over the interval $(t, t + dt)$.*
- *(economical interpretation) The benefit of holding the physical commodity will have a productive value since they allow us to meet the unexpected demand.*
- *The convenience yield is defined as the positive gain attached to the physical commodity minus the cost of storage.*

8.1.2.2 *Relationship between forward and spot*

Lemma 8.1.1 (spot-forward relationship for a storable commodity). *[1, pp. 37, 44]*

- *Under the assumption of constant interest rate, the spot-forward relation is given by*

$$F(t, T) = S(t) \exp((r - y)(T - t))$$

  where $y = y_1 - c$, y_1 is the benefit from the holding the physical commodity and c is the storage cost.
- *Under the assumption of stochastic interest rate, the spot-forward relation is given by*

$$F(t, T) = S(t) \exp((-y)(T - t))/P(t, T),$$

  where $y = y_1 - c$, y_1 is the benefit from the holding the physical commodity and c is the storage cost, and $P(t, T)$ is the T zero-coupon bond price at time t.

Note 8.1.1 (interpretation). [1, p. 37] Based on the spot-forward relation, we can approximately write

$$F(t, T) \approx S(t)(1 + r(T - t) + c(T - t) - y_1(T - t)).$$

We can interpret the difference between spot and the forward as:

- $S(t)r(T-t)$ is the funding cost of financing the purchase of S.
- $S(t)c(T-t)$ is the storage cost of holding the physical commodity.
- $S(t)y_1(T-t)$ is the benefit of holding the physical commodity.

Note 8.1.2 (how storage fee assets prices). [2, p. 878] Assume zero interest rate. Assume further that today's spot price of gas is $S_0 = 30$ and the storage cost is 1 per year. Then the no-arbitrage price for the natural gas will be $S_1 = 31$ one year later.

The following strategy ensures that the price is free of arbitrage.

$F_T < 31$			$F_T > 31$		
T=0	short S_0; enter futures to buy F_T	+30	T=0	long S_0 and store it, enter futures to sell at F_T	-30
T=1	settle the futures, receive payment for storage fee, and return S_T	+1-F_T	T=1	settle the futures, and pay the storage fee	+F_T - 1
payoff		$31 - F_T$	payoff		$F_T - 31$

8.1.3 Oil market

The hallmarks of oil market are the following[3, p. 157]

- The most commonly encountered oils by far are West Texas Intermediate (WTI) and Brent. West Texas Intermediate trades on both the New York Metals Exchange (NYMEX) and the Intercontinental Exchange (ICE), whereas Brent trades predominantly on the Intercontinental Exchange (ICE).
- WTI is a light sweet North American crude oil; Brent crude is a European crude oil.
- Futures on WTI/Brent for a sequence of calendar monthsTable 8.1.1.
- WTI futures contracts on NYMEX are for physical delivery; WTI contracts also trade on the ICE, but these are cash settled.
- Options on WTI futures are also traded on NYMEX, as showed in Table 8.1.2. These options expire three business days before the expiry for the underlying futures contract.

[3, p. 162]

Table 8.1.1: Cal 12 WTI-NYMEX strip of futures contract

contract	expiry date	cash date	first notice	delivery date
JAN12	20-Dec-11	21-Dec-11	22-Dec-11	Jan-12
FEB12	20-Jan-12	23-Jan-12	24-Jan-12	Feb-12
MAR12	21-Feb-12	22-Feb-12	23-Feb-12	Mar-12
APR12	20-Mar-12	21-Mar-12	22-Mar-12	Apr-12
...	...	...	...	...

Table 8.1.2: Cal 12 WTI-NYMEX strip of options on futures contract

contracts	option expiry date	futures expiry date	cash date
JAN12	15-Dec-11	20-Dec-11	22-Dec-11
FEB12	17-Jan-12	20-Jan-12	24-Jan-12
MAR12	15-Feb-12	21-Feb-12	23-Feb-12
APR12	15-Mar-12	20-Mar-12	22-Mar-12
...	...	...	...

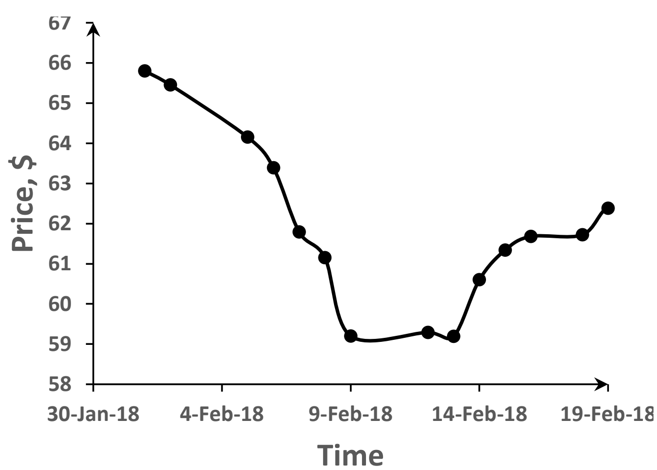


Figure 8.1.2: WTI APR18 contract historical prices starting from Feb,1,2018 to Mar,2,2018

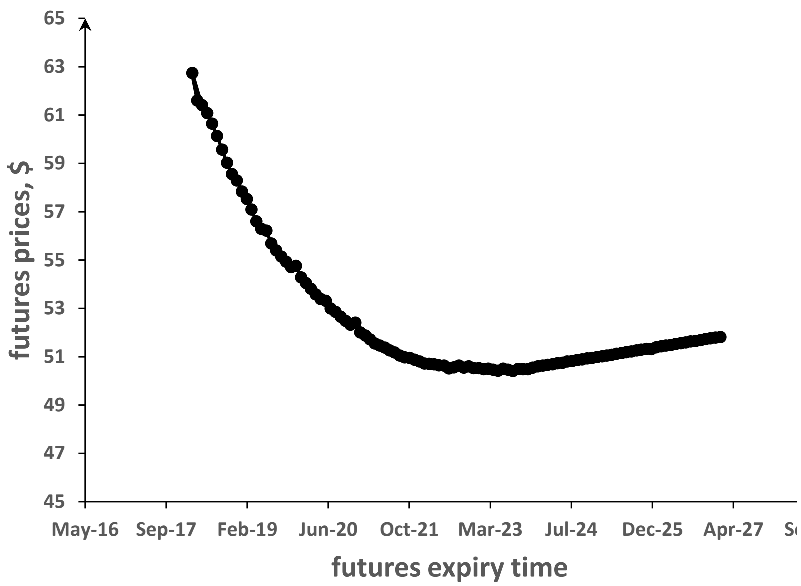


Figure 8.1.3: WTI forward curve observed on Mar,2,2018

8.1.4 Natural gas market

Remark 8.1.1 (general remarks).

- Natural gas contracts are traded in various geographical regions. The most liquid contract is the Henry Hub contract on the NYMEX for delivery of natural gas to the pipeline interconnector at **Henry Hub**.
- Henry Hub futures contracts traded on NYMEX(see Table 8.1.3) are for physical delivery of natural gas containing 10,000 MMBtu of extractable energy, into the Henry Hub pipeline complex operated by Sabine Hub Services at Erath, Louisiana.
- Natural gas futures contract also trades on the ICE, but is cash settled and is for a smaller contract size of 2500 MMBtu.
- Options on Henry Hub futures contracts trade on NYMEX, the expiry date being the business day before the futures contract expiry date.

Table 8.1.3: Cal 12 NG-NYMEX strip of futures contract

contract	expiry date	cash date	delivery date
JAN12	28-Dec-11	29-Dec-11	Jan-12
FEB12	27-Jan-12	30-Jan-12	Feb-12
MAR12	27-Feb-12	28-Feb-12	Mar-12
APR12	28-Mar-12	29-Mar-12	Apr-12
...	...	...	...

Definition 8.1.6 (seasonality).

- *Seasonality occurs in many commodities(crude oil is the main exception)*
- *If the storage capacity exceeds the wavelength, then no humps*
- *As showed in Figure 8.1.4, natural gas has large hump in winter, small hump in summer for the forward curve; however, the prompt futures price does not display significant seasonality(Figure 8.1.5).*
- *Gasoline has large hump in summer*
- *Electricity has humps in winter and summer, negative hump on weekends, and intro-day structure.*

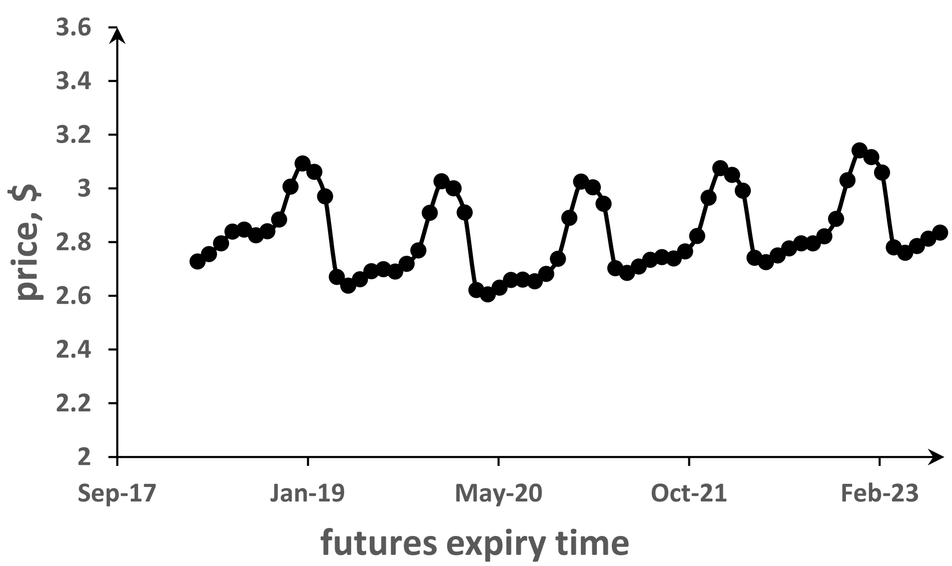


Figure 8.1.4: NG forward curve observed on Mar,2,2018

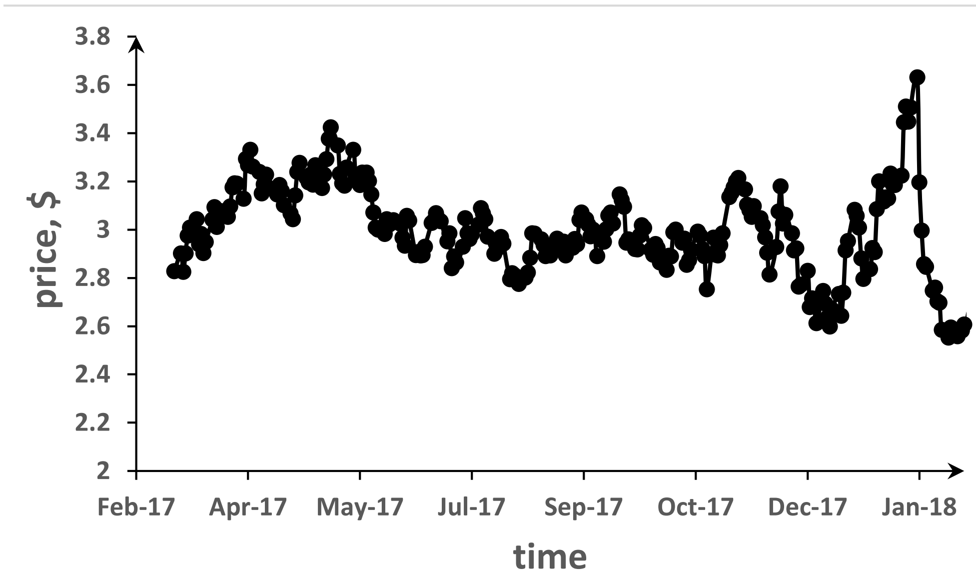


Figure 8.1.5: NG APR18 contract historical prices starting from Mar,2,2017 to Mar,2,2018

8.1.5 Base metal market

Remark 8.1.2 (general remarks). Futures contracts for base metals are primarily traded on the London Metal Exchange

maturity dates for futures contracts on the LME.

Futures contracts on COMEX are subject to cash clearing and daily margining

ot every futures contract has an option traded against it. The futures contracts which do have options available to trade against them are known as the monthly value dates, these being those futures with maturity dates on the third Wednesday of the month.

Table 8.1.4: Base metal market in LME and COMEX

metal	maximum tenor	
	futures	options
London Metal Exchange (LME)		
copper	123M	63M
aluminium high grade	123M	63M
aluminium alloy	27M	27M
NASAAC	27M	27M
zinc special high grade	63M	63M
nickel	63M	63M
lead	63M	63M
tin	15M	15M
New York Mercantile Exchange (COMEX)		
copper high grade	60M	24M

Table 8.1.5: LME base metal futures delivery date specifications

	copper aluminium	lead nickel zinc	NASAAC aluminium alloy	tin	
cash/spot	T+2	T+2	T+2	T+2	
0-3M	daily	daily	daily	daily	
3M-6M	weekly	weekly	weekly	weekly	every Wednesday
7-15M	monthly	monthly	monthly	monthly	3rd Wednesday
16-27M	monthly	monthly	monthly	-	3rd Wednesday
28-63M	monthly	monthly	-	-	3rd Wednesday
64-123M	monthly	-	-	-	3rd Wednesday

Table 8.1.6: Option expiry on futures in LME

option expiry	underlying future maturity	underlying future
Wed 05-Jun-2013	Wed 19-Jun-2013	JUN13
Wed 03-Jul-2013	Wed 17-Jul-2013	JUL13
Wed 07-Aug-2013	Wed 21-Aug-2013	AUG13
Wed 04-Sep-2013	Wed 18-Sep-2013	SEP13
...	...	....

8.1.6 Basis risk

Definition 8.1.7 (basis risk). *link*

Basis is the difference between the price of an energy commodity in one market and the price of an energy commodity in different market.

- *The different 'market' can be a different location, also known as* ***locational basis****, a different product or quality which we can be referred to as* ***product or quality basis*** *or a different tenor or time frame, which we refer to as* ***calendar basis****.*

Definition 8.1.8 (location risk). *link*

Locational basis risk *is the risk that you encounter when you hedge with a contract that doesn't have the same or similar delivery point as the risk you are seeking to hedge.*

- *As an example, if a US Gulf Coast oil producer decides to hedge their crude oil price risk with NYMEX WTI futures (which are deliverable in Cushing, Oklahoma), the producer is exposed to the locational basis risk between Cushing and their local market price (i.e. LLS – light Louisiana sweet). To quantify this example, the August NYMEX WTI crude oil swap closed yesterday at $46.61 while the August LLS crude oil basis swap (LLS-WTI) closed yesterday at $1.54 which means that the forward market for August LLS swaps is trading at a $1.54 premium to August WTI swaps.*

Definition 8.1.9 (product or quality basis risk). *link*

Product or quality basis risk *is the risk that you encounter when you hedge with a contract that isn't the same product or quality as the product that you are seeking to hedge.*

- *As an example, jet fuel is often hedged with crude oil, gasoil or ultra-low sulfur diesel fuel. While jet fuel, gasoil and ULSD are similar and highly correlated they are not one in the same. As a result, if one chooses to hedge jet fuel with crude oil, gasoil, ULSD or another product, they are often exposing themselves to significant product or quality basis risk, in addition to locational basis risk.*

Definition 8.1.10 (calendar basis risk). *link*

Calendar basis risk, also known as calendar spread risk, is the risk that arises from hedging with a contract that doesn't expire, settle or mature on the same date as the underlying exposure.

- *As an example, a large consumer (i.e. a vehicle fleet) of gasoline might decide to hedge their exposure to gasoline price by purchasing NYMEX RBOB gasoline futures. In this example, the consumer is exposed to calendar basis risk as NYMEX gasoline futures expire on the last day of the month prior to the delivery month i.e. the August RBOB gasoline futures contract expired on July 29, the last trading day of the month.*

8.2 Forward price normal model

8.2.1 The model

Definition 8.2.1 (Black model). *[1, p. 73]*

- *(single factor model)The Black forward price single factor model assumes that under risk-neutral measure*
$$\frac{dF(t,T)}{dt} = \sigma(t,T)dW_t$$
where W_t is a Brownian motion under risk-neutral measure Q.
- *(multi-factor model) The Black forward price multi-factor model assumes that under risk-neutral measure*
$$\frac{dF(t,T)}{dt} = \sum_{k=1}^{K} \sigma_k(t,T)dW_k(t),$$
*where $W_1, W_2, ..., W_K$ are **independent** Brownian motion under risk-neutral measure Q such that $dW_i dW_j = \rho_{ij} dt$.*

Remark 8.2.1 (choices of parameters)**.**

- The covariance structure[1, p. 73] can be estimated from principle component analysis of historical forward price curves(Lemma 3.4.2).

- Volatility functions often assumed to be deterministic
- The **incremental stationarity** of the $\ln F(t,T)$ with fixed time-to-maturity $T-t$ can be realized by assuming
$$\sigma_k(t,T) = \sigma_k(T-t).$$

Lemma 8.2.1 (basic statistical properties of forward prices). *Consider forward price model*

$$\frac{dF(t,T)}{dt} = \sum_{k=1}^{K} \sigma_k(t,T) dW_k(t),$$

where $W_1, W_2, ..., W_K$ are independent Brownian motion under risk-neutral measure Q such that $dW_i dW_j = \rho_{ij} dt$. It follows that

- *For any $s \geq t$ and fixed T*
$$E[F(s,T)|\mathcal{F}_t] = F(t,T).$$
That is, $F(t,T)$ is a martingale with fixed T.
- *For any $s \geq t \geq t_0$, the convariance structure is given by*
$$E[F(t,T_1)F(s,T_2)] = F(t_0,T)F(t_0,T)\exp(\Sigma_{ij}(t)),$$
where
$$\Sigma_{ij}(t) = \int_{t_0}^{t} \sum_{k=1}^{K} \sigma_k(u,T_i)\sigma_k(u,T_j) du$$

Proof. Directly from the multi-dimensional geometric SDE(**??**) □

8.2.2 Interprete the covariance structure

Definition 8.2.2 (rolling futures price). *[4, p. 15] Define*

$$\begin{aligned} \ln f(t;\tau_j) &\triangleq \ln F(t, t+\tau_j) \\ &\approx \frac{(t+\tau_j - T_j)\ln F(t,T_{j+1}) + (T_{j+1} - t - \tau_j)\ln F(t,T_j))}{T_{j+1} - T_j}, T_j < t+\tau_j < T_{j+1}. \end{aligned}$$

We say $f(t;\tau_j)$ is ***the rolling futures price*** *at time t for contracts maturing at $\tau_j + t$.*

Remark 8.2.2 (the dynamics of full forward curve via interpolation).

- Note that given the dynamics of finite number of points $F(t,T_1), F(t,T_2), ..., F(t,T_n)$ on the full forward curve $F(t,T)$. We can approximate $F(t,\tau)$ by weighted sum of nearby points.

- If $\sigma_i(t, T_i)$ is state independent, then $F(t, T_i)$ will be a Gaussian process; the interpolated dynamics $F(t, \tau)$ will also be a Gaussian process.

Lemma 8.2.2 (rolling futures prices dynamics and covariance structure). *[4, p. 15]*

- *The rolling futures price dynamics for N futures expiring at $\tau_1, \tau_2, ..., \tau_N$ are given by*

$$d \ln f(t; \tau_j) = v(t; \tau_j)dt + \sum_{k=1}^{M} \sigma_k(\tau_j) dW_k(t), j = 1, 2, ..., N$$

$$v_j(t) = [\mu(t; \tau_j) - \frac{1}{2}\sum_{k=1}^{M} \sigma_k^2(\tau_j)] + \frac{\partial \ln F(t, t + \tau_j)}{\partial T_j}$$

where $W_1, W_2, ..., W_M$ are M independent Brownian motions.

- *The vector form is given by*

$$d \ln f = vdt + \Sigma dW, \Sigma \in \mathbb{R}^{N \times M}$$

where

$$f(t) = \begin{bmatrix} f(t; \tau_1) \\ f(t; \tau_2) \\ \vdots \\ f(t; \tau_N) \end{bmatrix}, v(t) = \begin{bmatrix} v(t; \tau_1) \\ v(t; \tau_2) \\ \vdots \\ v(t; \tau_N) \end{bmatrix}, \Sigma = \begin{bmatrix} \sigma_1(\tau_1) & \sigma_2(\tau_1)) & \cdots & \sigma_M(\tau_1)) \\ \sigma_1(\tau_2) & \sigma_2(\tau_2) & \cdots & \sigma_M(\tau_2) \\ \vdots & \ddots & & \vdots \\ \sigma_1(\tau_N)) & \sigma_2(\tau_N)) & \cdots & \sigma_M(\tau_N)) \end{bmatrix}, dW = \begin{bmatrix} dW_1 \\ dW_2 \\ \vdots \\ dW_N \end{bmatrix}$$

- *The covariance matrix for vector $d \ln f - vdt$ is given by*

$$Cov(d \ln f - vdt, d \ln f - vdt) = \Sigma \Sigma^T dt.$$

Specifically,

$$Cov(d \ln f(t; \tau_i) - v_i dt, d \ln f(t; \tau_j) - v_j dt) = \sum_{k=1}^{M} \sigma_k(\tau_i)\sigma_k(\tau_j)dt = C(\tau_i, \tau_j)dt$$

is a constant.

- *An alternative vector form is given by*

$$d \ln f = vdt + DdZ,$$

or equivalently

$$d \ln f(t; \tau_i) = vdt + D_i dZ_i, i = 1, 2, ..., N$$

where $D = diag(D_1, D_2, ..., D_N), D_i = \sqrt{(\Sigma\Sigma^T)_{ii}}$, *and* $dZdZ^T = \rho dt$ *is correlated Brownian motion,* $\rho = D^{-1/2}\Sigma\Sigma^T D^{-1/2}$*is the correlation matrix associated with covariance matrix* $\Sigma\Sigma^T$.

- *An PCA representation is given by*

$$d\ln f = vdt + \sqrt{\Lambda}UdW,$$

where $\Sigma\Sigma^T = U\Lambda U^T, U \in \mathbb{R}^{N\times M}, \Lambda \in \mathbb{R}^{M\times M}$.

Proof. To show the equivalence between the two forms, we have

$$DdZdZ^T D = D^{1/2}\rho D^{1/2}dt = \Sigma\Sigma^T dt = \Sigma dWdW^T\Sigma^T.$$

□

Remark 8.2.3. Note that $\frac{\partial \ln F(t,t+\tau_j)}{\partial T_j}$ is deterministic quantity since $F(t, t+\tau_j)$ is known at time t.

8.3 Pricing examples

8.3.1 European options on futures

Lemma 8.3.1 (Options on futures). *[5, p. 101] Assume the futures price* $F(t, T_1)$ *has the risk-neutral dynamics of*

$$dF(t, T_1)/F(t, T_1) = \sigma(t)dW_t.$$

Let current time be t and the constant short rate be r The call/put option with strike K and matures at $T, T < T_1$ *will have price*

$$C(t) = e^{r(T-t)}(F(t, T_1)N(d_1) - KN(d_2).$$

$$P(t) = -e^{r(T-t)}(F(t, T_1)N(-d_1) + KN(-d_2).$$

where

$$d_1 = \frac{1}{\Sigma\sqrt{T-t}}[\ln(\frac{F(t, T_1)}{K}) + \Sigma^2/2(T-t)]$$
$$d_2 = d_1 - \Sigma\sqrt{T-t}$$

$$\Sigma^2 = \frac{1}{T-t}\int_t^T \sigma^2(u)du.$$

Proof. Note that $F(T, T_1)$ is the log-normal random variable $LN(F(t, T_0), \Sigma^2(T-t)$. The pricing derivation is same as Lemma 5.2.1. □

8.3.2 Asian options

Lemma 8.3.2 (distribution of arithmetic average of single futures). *Let $t_0 \leq t_1 < t_2 < \cdots < t_m \leq T$ be a set of dates. Let*

$$X = \sum_{i=1}^{m} w_i F(t_i, T).$$

Then

- $E_Q[X|\mathcal{F}_0] = \sum_{i=1}^{m} w_i E_Q[F(t_i, T)|\mathcal{F}_0] = \sum_{i=1}^{m} w_i F(t_0, T).$
- $$\begin{aligned} E[X^2] &= \sum_{i=1}^{m}\sum_{j=1}^{m} w_i w_j E_Q[F(t_i, T)F(t_j, T)|t_0] \\ &= \sum_{i=1}^{m}\sum_{j=i+1}^{m} 2 w_i w_j F^2(t_0, T)\exp(\int_{t_0}^{t_i} g(s)g(s)D(T-s)D(T-s)ds) \end{aligned}$$

Proof. (1) Use linearity of expectation and the fact that $F(t, T)$ is a martingale in measure Q. (2)

$$\begin{aligned} F(t_i, T) &= F(t_0, T)\exp(-\frac{1}{2}\int_{t_0}^{t_i} g^2(s)D^2(T-s)ds + \int_{t_0}^{t_i} g(s)D(T-s)dW(s)) \\ \implies F(t_i, T)F(t_j, T) &= F(t_0, T)^2\exp(-\frac{1}{2}\int_{t_0}^{t_i} g^2(s)D^2(T-s)ds + \int_{t_0}^{t_1} g(s)D(T-s)dW(s)) \\ &\exp(-\frac{1}{2}\int_{t_0}^{t_j} g^2(s)D^2(T-s)ds + \int_{t_0}^{t_j} g(s)D(T-s)dW(s)) \\ &= F(t_0, T)^2\exp(-\int_{t_0}^{t_i} g^2(s)D^2(T-s)ds + 2\int_{t_0}^{t_1} g(s)D(T-s)dW(s)) \\ &\exp(-\frac{1}{2}\int_{t_i}^{t_j - t_i} g^2(s)D^2(T-s)ds + \int_{t_i}^{t_j - t_i} g(s)D(T-s)dW(s)) \\ &= F(t_0, T)^2\exp(-\int_{t_0}^{t_i} g^2(s)D^2(T-s)ds + 2\int_{t_0}^{t_1} g(s)D(T-s)dW(s)) \\ E[F(t_i, T)F(t_j, T)] &= F(t_0, T)^2\exp(\int_{t_0}^{t_i} g^2(s)D^2(T-s)ds) \end{aligned}$$

□

Lemma 8.3.3 (distribution of arithmetic average of multiple futures). *Let $t_0 \leq t_1 < T_1 < t_2 < T_2 < \cdots < t_m < T_m$ be a set of dates. Let*

$$X = \sum_{i=1}^{m} w_i F(t_i, T_i).$$

Then

- $E_Q[X|\mathcal{F}_0] = \sum_{i=1}^{m} w_i E_Q[F(t_i, T)|\mathcal{F}_0] = \sum_{i=1}^{m} w_i F(t_0, T_i)$.
-

$$\begin{aligned} E[X^2] &= \sum_{i=1}^{m}\sum_{j=1}^{m} w_i w_j E_Q[F(t_i, T)F(t_j, T)|t_0] \\ &= \sum_{i=1}^{m}\sum_{j=i+1}^{m} 2w_i w_j F(t_0, T_i)F(t_0, T_j)\exp(\int_{t_0}^{t_i} g(s)g(s)D(T_i - s)D(T_j - s)\rho_{ij} ds) \end{aligned}$$

Proof. (1) Use linearity of expectation and the fact that $F(t, T)$ is a martingale in measure Q. (2)

$$\begin{aligned} F(t_i, T_i) &= F(t_0, T_i)\exp(-\frac{1}{2}\int_{t_0}^{t_i} g^2(s)D^2(T_i - s)ds + \int_{t_0}^{t_i} g(s)D(T_i - s)dZ_i(s)) \\ \implies F(t_i, T)F(t_j, T) &= F(t_0, T)^2 \exp(-\frac{1}{2}\int_{t_0}^{t_i} g^2(s)D^2(T_i - s)ds + \int_{t_0}^{t_1} g(s)D(T_i - s)dZ_i(s)) \\ &\exp(-\frac{1}{2}\int_{t_0}^{t_j} g^2(s)D^2(T_j - s)ds + \int_{t_0}^{t_j} g(s)D(T_j - s)dZ_j(s)) \\ E[F(t_i, T_i)F(t_j, T_j)] &= F(t_0, T_i)F(t_0, T_j)\exp(\int_{t_0}^{t_i} g(s)^2 D(T_i - s)D(T_j - s)\rho_{ij} ds) \end{aligned}$$

To get $E[F(t_i, T_i)F(t_j, T_j)]$, we use the fact(**??**) that If $X_1 \in N(\mu_1, \sigma_1^2), X_2 \in N(\mu_2, \sigma_2^2)$, then

$$E[\exp(X_1 + X_2)] = \exp(E[X_1] + E[X_2] + \frac{1}{2}Var[X] + \frac{1}{2}Var[X_2] + Cov(X_1, X_2)).$$

□

8.3.3 Commodity swaps

A single futures contract enables the holder to get exposure to the futures price change on a single delivery date. Considering the highly volatile nature of commodity prices, exposure to futures prices averaging on multiple dates might be more desired. Commodity swap meets such needs, and can be viewed as a collection of multiple futures on the same commodity.

Definition 8.3.1 (commondity swap). *[3, p. 42]*

- *A **commodity swap** involves exchanging floating payments indexed to commodity prices against a fixed known price K, with the cash settlement either at the end of the swap or on a regular interval basis.*
- *The floating payment is usually indexed to spot price or prompt futures price.*

Lemma 8.3.4 (commodity swap value and fair swap rate). *[3, p. 42] Let $t_0 < t_1 < t_2 < \cdots < t_n$ be a set of dates. Let $\mathcal{T}(t_i)$ be a function map t_i to the nearest futures contract maturity date. Let t_0 be current time.*

- *Consider a swap exchanging a floating leg with a fixed leg. The floating leg is paying $\frac{1}{n}\sum_{i=1}^n F(t_i, \mathcal{T}(t_i))$ at time T_{stl} with weight $T_{stl} < t_1$; the fixed leg is paying amount K at time T_{stl}. We assume $t_0 < T_{stl}$.*
  - *The value of the swap at time t_0 is*

$$\begin{aligned} V_0 &= \exp(-r(T_{stl} - t_0))E_Q[\frac{1}{n}\sum_{i=1}^n F(t_i, \mathcal{T}(t_i))|\mathcal{F}_0] \\ &- \exp(-r(T_{stl} - t_0))K = \exp(-r(T_{stl} - t_0))[\frac{1}{n}\sum_{i=1}^n F(t_0, \mathcal{T}(t_i)) - K]. \end{aligned}$$

  - *The fair swap rate is*

$$V_0 = 0 \implies K_{swap} = \frac{1}{n}\sum_{i=1}^n F(t_0, \mathcal{T}(t_i))$$

-
  - *Consider a swap exchanging floating legs $F(t_i, \mathcal{T}(t_i))$ with a fixed leg amount K at time $T_{stl,i}, i = 1, 2, ..., n$. Then the value of swap at time t_0 is given by*

$$V_0 = E_Q[\sum_{i=1}^n \exp(-r(T_{stl,i} - t_0))V_i] = \sum_{i=1}^n \exp(-r(T_{stl,i} - t_0))[F(t_0, \mathcal{T}(t_i)) - nK],$$

  where

$$V_i = F(t_i, \mathcal{T}(t_i)) - K.$$

  - *The fair swap rate is*

$$K_{swap} = \frac{\sum_{i=1}^n \exp(-r(T_{stl,i} - t_0))F(t_0, \mathcal{T}(t_i))}{\sum_{i=1}^n \exp(-r(T_{stl,i} - t_0))}.$$

Proof. (1) Note that under log-normal model, the futures price is a martingale; that is,

$$F(t_0, \mathcal{T}(t_i)) = E[F(t_i, \mathcal{T}(t_i))|\mathcal{F}_0].$$

To get the fair swap rate, we have

$$V_0 = 0 \implies K_{swap} = \frac{1}{n}\sum_{i=1}^{n} F(t_0, \mathcal{T}(t_i))$$

(2) Similar to (1). □

Lemma 8.3.5 (Greeks associated with a commodity swap). *Let $t_0 < t_1 < t_2 < \cdots < t_n$ be a set of dates. Let $\mathcal{T}(t_i)$ be a function map t_i to the nearest futures contract maturity date. Let t_0 be current time. Consider a swap exchanging floating legs $F(t_i, \mathcal{T}(t_i))$ with a fixed leg amount K at time $T_{stl,i}, i = 1, 2, ..., n$.*

- *Then the value of swap at time t_0 is given by*

$$V_0 = E_Q[\sum_{i=1}^{n} \exp(-r(T_{stl,i} - t_0))V_i] = \sum_{i=1}^{n} \exp(-r(T_{stl,i} - t_0))[F(t_0, \mathcal{T}(t_i)) - nK],$$

  where

$$V_i = F(t_i, \mathcal{T}(t_i)) - K.$$

- *The swap has exposure the forward price curve; that is,*

$$\frac{\partial V_0}{\partial F(t_0, \mathcal{T}(t_i))} = \exp(-r(T_{stl,i} - t_0)).$$

- *The swap has zero Gamma and vega risk; that is,*

$$\frac{\partial^2 V_0}{\partial F(t_0, \mathcal{T}(t_i))^2} = 0, \frac{\partial V_0}{\partial \sigma} = 0.$$

Proof. Straight forward from Lemma 8.3.4. □

8.3.4 Spread options

8.3.4.1 *Basics*

Spread options are very popular in all commodity markets. Common spreads include crackspread, which is the difference between the prices of refined products and the price of crude oil input, sparkspread, which is the difference between the price of electricity (output) and the price of the corresponding quantity of primary fuel (input).

Spread options can be used by oil refineries, power plant, etc to hedge the risks associated with the input and ouput commondiites prices, hence locking the profit.

Calendar spread options are options on the prices spread of futures with two different maturities on the same commodity. They can be used by companies providing stroage services, such as gas storage, water reservoirs, for valuation and hedging purpose.

In summary, we can define spread options as follows:

Definition 8.3.2 (spread option on futures prices).

- *A spread is the difference between two commodity prices (either spot price or futures price).*
- *A spread on two futures prices $F(T, T_1)$ and $F(T, T_2), T_1 \neq T_2$ is called calendar spread.*
- *A spread option on two futures prices $F_1(t, T_1), F_2(t, T_2)$ with strike K and maturity $T < T_1, T < T_2$ has payoff*

$$V(T) = \max(F(T, T_2) - F(T, T_1) - K, 0).$$

8.3.4.2 *Pricing*

Lemma 8.3.6 (pricing when strike is zero). *Consider a spread option on two futures prices $F(t, T_1), F(t, T_2)$ with strike 0, maturity $T < T_1, T < T_2$, and payoff*

$$V(T) = \max(F(T, T_2) - F(T, T_1), 0).$$

Assume the following risk-neutral model for $F_1(t) \triangleq F(t, T_1)$, $F_2(t) \triangleq F(t, T_2)$,

$$\begin{aligned} dF_1(t) &= \sigma_1 F_1(t) dW_1(t) \\ dF_2(t) &= \rho\sigma_2 F_2(t) dW_1(t) + \sqrt{1-\rho^2}\sigma_2 F_2(t) dW_2(t) \end{aligned}$$

*where W_1 and W_2 are **independent** Brownian motions.*

Then the value of the exchange option at time 0 is given by

$$V_0 = F_1(0)N(d_1) - F_2(0)N(d_2)$$

where

$$d_1 = \frac{\ln(F_1(0)/F_2(0) + (\hat{\sigma}^2/2)}{\hat{\sigma}\sqrt{T}}, d_2 = d_1 - \hat{\sigma}\sqrt{T},$$

and

$$\hat{\sigma} = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}.$$

Proof. Use the F_1 as the numeraire, then under this new measure Q_F,

$$\frac{V(T)}{S_1(T)}$$

is a martingale. Therefore,

$$\frac{V(0)}{F_1(t)} = E_F[\frac{V(T)}{F_1(T)}|\mathcal{F}_t] = E_F[\max(\frac{F_2(T)}{F_1(T)} - 1)|\mathcal{F}_t].$$

Note that under measure Q_F Theorem 1.6.19), the dynamics of $F_1(t)$ and $F_2(t)$ follows

$$\begin{aligned} dF_1(t) &= (\sigma_1^2)F_1(t)dt + \sigma_1 F_1(t)dW_1(t) \\ dF_2(t) &= (r + \rho\sigma_1\sigma_2)F_2(t)dt + \rho\sigma_2 F_2(t)dW_1(t) + \sqrt{1-\rho^2}\sigma_2 S_2(t)dW_2(t) \\ d\frac{F_2}{F_1} &= \frac{F_2}{F_1}((\rho\sigma_2 - \sigma_1)dW_1 + \sqrt{1-\rho^2}\sigma_2\sigma dW_2) \end{aligned}$$

Denote $Y = \frac{F_2}{F_1}$, then Y is a geometric Brownian motion with volatility $\sigma_Y = \sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}$.

Then we have $V(t) = F_1(t)E_F[\max(Y(T) - 1, 0)]$, which can be evaluated. □

Lemma 8.3.7 (pricing with Kirk approximation). *Consider a spread option on two futures prices $F(t, T_1), F(t, T_2)$ with strike K, maturity $T < T_1, T < T_2$, and payoff*

$$V(T) = \max(F(T, T_2) - F(T, T_1), 0).$$

Assume the following risk-neutral model for $F_1(t) \triangleq F(t, T_1)$, $F_2(t) \triangleq F(t, T_2)$,

$$\begin{aligned} dF_1(t) &= \sigma_1 F_1(t)dW_1(t) \\ dF_2(t) &= \rho\sigma_2 F_2(t)dW_1(t) + \sqrt{1-\rho^2}\sigma_2 F_2(t)dW_2(t) \end{aligned}$$

*where W_1 and W_2 are **independent** Brownian motions.*

If $K \ll F_1(T)$, then the approximate value of the exchange option at time 0 is given by

$$V_0 = (F_1(0) + K)N(d_1) - F_2(0)N(d_2)$$

where

$$d_1 = \frac{\ln(F_1(0) + K/F_2(0) + (\hat{\sigma}^2/2)}{\hat{\sigma}\sqrt{T}}, d_2 = d_1 - \hat{\sigma}\sqrt{T},$$

and

$$\hat{\sigma} = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}.$$

Proof. In the Kirk approximation, we assume

$$d(F_1 + K)/(F_1 + K) = \sigma_1 dW_1.$$

View $F_1 + K$ as a new random variable, then we can use the previous result in Lemma 8.3.6. □

8.4 Notes on bibliography

Major references are .

[6][3]

[7][8][1]

For multi-factor model, see [9].

Problems and solutions in mathematical finance: commodity and foreign exchange

Market data sources: investing.com

BIBLIOGRAPHY

1. Geman, H. *Commodities and commodity derivatives: modeling and pricing for agriculturals, metals and energy* (John Wiley & Sons, 2009).
2. Roncoroni, A., Fusai, G. & Cummins, M. *Handbook of multi-commodity markets and products: structuring, trading and risk management* (John Wiley & Sons, 2015).
3. Clark, I. J. *Commodity Option Pricing: A Practitioner's Guide* (John Wiley & Sons, 2014).
4. Ellefsen, P. E. & Sclavounos, P. D. Multi-Factor Model of Correlated Commodity-Forward Curves for Crude Oil and Shipping Markets (2009).
5. Lalley, S. *The mathematics of financial derivatives: a student introduction* (Chicago University Statistics 390/ Mathematical Finance 345, 2001).
6. Burger, M., Graeber, B. & Schindlmayr, G. *Managing energy risk: An integrated view on power and other energy markets* (John Wiley & Sons, 2008).
7. Jewson, S. & Brix, A. *Weather derivative valuation: the meteorological, statistical, financial and mathematical foundations* (Cambridge University Press, 2005).
8. Parodi, P. *Pricing in general insurance* (CRC Press, 2014).
9. Bukkapatanam, V. & Doraiswamy, S. T. Multi factor models for the commodities futures curve: Forecasting and pricing. *Futures* **15,** 20.

9

VALUE AT RISK

9.1 Foundations

Definition 9.1.1 (value at risk). *[1, p. 65] The -value-at-risk(VaR) $VaR_\alpha(w;h)$ or $VaR_\alpha(L)$ is defined as the potential loss which the portfolio w can suffer for a given confidence level α and a fixed holding period h. Three parameters are needed for the computation VaR:*

- *the **holding period** h, which indicates the time period to calculate the loss;*
- *the **confidence level** α, which gives the probability that the loss is lower than the VaR(or equivalently, there is $1-\alpha$ probability that the loss is greater than the VaR).*
- *the **portfolio** w, which gives the allocation in terms of risky assets and is related to risk factors.*

Remark 9.1.1. Typical values of α are 90%, 95%, 99%.

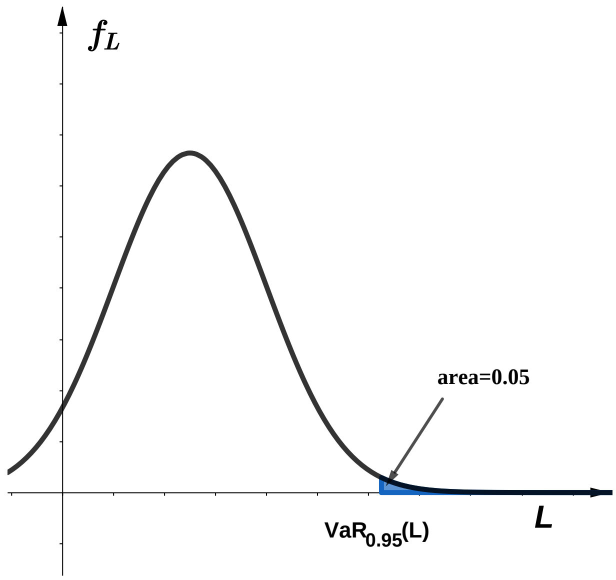


Figure 9.1.1: VaR demo. Given the loss L distribution, we can determine VaR.

Definition 9.1.2 (VaR calculation). *[1, p. 65] Let $P_t(w)$ be the mark-to-market value of the portfolio w at time t. We define the loss of the portfolio(which is a random variable) at future time $t+h$ as*

$$L(w) = P_t(w) - P_{t+h}(w).$$

Then $VaR_\alpha(w;h)$ is given by

$$VaR_\alpha(w;h) = F_L^{-1}(\alpha),$$

where F_L is the CDF of the random loss L. Often, we write $VaR_\alpha(w;h)$ as $VaR_\alpha(L)$.

Remark 9.1.2.

- VaR is a 'tail' risk measure
- VaR_p is increasing in p

Definition 9.1.3 (expected shortfall, conditional value at risk). *The conditional value at risk $CVaR_p(L)$ of random variable L at the confidence level $\alpha \in (0,1)$ is defined as*

$$\begin{aligned} ES = CVaR_\alpha(L) &= E[L|L \geq VaR_\alpha(L)] \\ &= \frac{\int_{VaR_\alpha(L)}^{\infty} x f_L(x)dx}{P(L \geq VaR_\alpha(L))} \\ &= \frac{1}{1-\alpha}\int_{VaR_\alpha(L)}^{\infty} x f_L(x)dx \\ &= \frac{1}{1-\alpha}\int_{\alpha}^{1} VaR_u(L)du \end{aligned}$$

Remark 9.1.3 (the transformation details). We use the following procedure to do the transformation

$$\begin{aligned} ES &= \frac{1}{1-\alpha}\int_{VaR_\alpha(L)}^{\infty} x f_L(x)dx \\ &= \frac{1}{1-\alpha}\int_{VaR_\alpha(L)}^{\infty} x dF_L(x) \\ &= \frac{1}{1-\alpha}\int_{\alpha}^{1} F_L^{-1}(u)du \\ &= \frac{1}{1-\alpha}\int_{\alpha}^{1} VaR_u(L)du \end{aligned}$$

where we used the variable transformation for integrals.

Lemma 9.1.1 (VaR basic properties). *VaR satisfies* ***homogeneity, monotonicity, and translational invariance*** *but fails sub-additivity; therefore,* ***VaR is not a coherent risk measure****. Specifically,*

- *For any $\lambda \geq 0$, $VaR_\alpha(\lambda L) = \lambda VaR_\alpha(\lambda L)$.*
- *For random variables L_1 and L_2 such that $L_1(\omega) < L_2(\omega) \forall \omega \in \Omega$, we have $VaR_\alpha(L_1) < VaR_\alpha(L_2)$.*
- *For any $C \in \mathbb{R}$, $VaR_\alpha(L + C) = VaR_\alpha(L) + C$.*

Proof. (1)(homogeneity) For any $\lambda \geq 0$, we have

$$VaR_\alpha(\lambda L) = F_{\lambda L}^{-1}(\alpha) = \lambda F_L^{-1}(\alpha) = VaR_\alpha(L),$$

where we use the scaling property of cdf. (2)(monotonicity) Consider L_1 and L_2 for the losses of two portfolios and assume $L_1(\omega) < L_2(\omega) \forall \omega \in \Omega$. Then given a fixed α, we have

$$VaR_\alpha(L_1) < VaR_\alpha(L_2).$$

(3)(translational invariance)Consider a constant loss $C \in \mathbb{R}$. We have

$$VaR_\alpha(L+C) = F^{-1}_{L+C}(\alpha) = C + F^{-1}_L(\alpha) = C + VaR_\alpha,$$

where we use the translational property of cdf. □

Lemma 9.1.2 (ES properties). *Expected shortfall satisfies* ***sub-additivity, homogeneity, monotonicity, and translational invariance****; Expect shortfall is a coherent risk measure. Specifically,*

- *For random variables L_1 and L_2, $ES(L_1+L_2) \leq ES(L_1) + ES(L_2)$.*
- *For any $\lambda \geq 0$, $ES(\lambda L) = \lambda ES(\lambda L)$.*
- *For random variables L_1 and L_2 such that $L_1(\omega) < L_2(\omega) \forall \omega \in \Omega$, we have $ES(L_1) < ES(L_2)$.*
- *For any $C \in \mathbb{R}$, $ES(L+C) = ES(L) + C$.*

Proof. Note that $ES = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(L)du$. Using the VaR properties of homogeneity, monotonicity, and translational invariance, we have (1) (homogeneity) For any $\lambda \geq 0$, we have

$$ES(\lambda L) = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(\lambda L)du = \frac{1}{1-\alpha}\int_\alpha^1 \lambda VaR_u(L)du = \lambda ES(L),$$

(2) (monotonicity) Consider L_1 and L_2 for the losses of two portfolios and assume $L_1(\omega) < L_2(\omega)\forall\omega \in \Omega$. Then given a fixed α, we have

$$ES(L+C) = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(L_1)du < \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(L_2)du = ES(L_2).$$

(3) (translational invariance)Consider a constant loss $C \in \mathbb{R}$. We have

$$ES(L+C) = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(L+C)du = \frac{1}{1-\alpha}\int_\alpha^1 (VaR_u(L)+C)du = ES(L) + C.$$

(4) (subadditivity)

□

9.2 Analytical VaR and ES

9.2.1 Loss with normal distribution

Lemma 9.2.1 (VaR with normal distribution of loss). *[1, p. 73] Assume $L(w) \sim N(\mu, \sigma^2)$(usually $\mu < 0$). Then*

$$VaR_\alpha(w; h) = \mu + \Phi^{-1}(\alpha)\sigma,$$

where $\Phi(x)$ is the cdf for the standard normal distribution.

Proof. Denote $F_L(x) : \mathbb{R} \to [0, 1]$ as the cdf of random variable $L(w)$. Then $Var_\alpha(w; h) = F_L^{-1}(\alpha)$. We have

$$\begin{aligned}
Pr(L \leq F_L^{-1}(\alpha)) &= \alpha \\
Pr(\mu + \sigma Z \leq F_L^{-1}(\alpha)) &= \alpha, Z \sim N(0, 1) \\
Pr(Z \leq \frac{F_L^{-1}(\alpha) - \mu}{\sigma}) &= \alpha \\
\Phi(\frac{F_L^{-1}(\alpha) - \mu}{\sigma}) &= \alpha \\
\Phi^{-1} \circ \Phi(\frac{F_L^{-1}(\alpha) - \mu}{\sigma}) &= \Phi^{-1}(\alpha) \\
\frac{F_L^{-1}(\alpha) - \mu}{\sigma} &= \Phi^{-1}(\alpha) \\
F_L^{-1}(\alpha) &= \mu + \Phi^{-1}(\alpha)\sigma
\end{aligned}$$

□

Corollary 9.2.0.1 (extensions for VaR). *[1, p. 76] Let $R_{t+h} \in \mathbb{R}^n$ be the vector of asset returns in a horizon h. We note that the loss for a portfolio characterized by weight vector $W_t \in \mathbb{R}^n$ is given by*

$$L(W) = -W_t^T R_{t+h}.$$

- *Assume $R_{t+h} \sim MN(\mu, \Sigma)$. Then VaR with confidence level α is given by*

$$VaR_\alpha(W; h) = -W_t^T \mu + \Phi^{-1}(\alpha)\sqrt{W_t^T \Sigma W_t}.$$

- *Assume $R_{t+h} = B\mathcal{F}_t + \epsilon_t, B \in \mathbb{R}^{n\times m}, \mathcal{F}_t \sim MN(\mu, \Omega), \epsilon_t \sim MN(0, D)$, $\mathcal{F}_t, \epsilon_t$ are m-dimension random vectors independent to each other. Then VaR with confidence level α is given by*

$$VaR_\alpha(W; h) = -W_t^T B\mu + \Phi^{-1}(\alpha)\sqrt{W_t^T (B\Omega B^T + D)W_t}.$$

Proof. (1)Note that $L(w) \sim N(W^T\mu, W^T\Sigma W)$. (2) Note that

$$E[L(w)] = E[-W_t^T R_{t+h}] = W^T B\mu, Cov[L(w)] = W_t^T (B\Omega B^T + D)W_t.$$

□

Lemma 9.2.2 (ES with normal distribution of loss). *[2, p. 264] Assume $L(w) \sim N(\mu, \sigma^2)$(usually $\mu < 0$). Then*

$$ES_\alpha(w; h) = \mu + \sigma \frac{\exp(-y^2/2)}{\sqrt{2\pi}(1-\alpha)}, y = \Phi^{-1}(\alpha),$$

where $\Phi(x)$ is the cdf for the standard normal distribution.

Proof. Note that from definition of ES, we have

$$\begin{aligned}
ES &= \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(L)du \\
&= \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(\mu + \sigma Z)du, Z \sim N(0,1) \\
&= \frac{1}{1-\alpha}\int_\alpha^1 \sigma VaR_u(Z) + \mu du \\
&= \mu + \sigma\frac{1}{1-\alpha}\int_\alpha^1 VaR_u(Z)du \\
&= \mu + \sigma\frac{1}{1-\alpha}\int_{VaR_\alpha(Z)}^\infty x f_Z(x)dx \\
&= \mu + \sigma\frac{1}{1-\alpha}\int_y^\infty \frac{x}{\sqrt{2\pi}}\exp(-\frac{x^2}{2})dx, y = VaR_\alpha(Z) \\
&= \mu + \sigma\frac{1}{(1-\alpha)\sqrt{2\pi}}\exp(-\frac{y^2}{2})dx
\end{aligned}$$

where we use the homogeneity and translational invariance of VaR(Lemma 9.1.1). □

Corollary 9.2.0.2 (extensions for ES). *[1, p. 76] Let $R_{t+h} \in \mathbb{R}^n$ be the vector of asset returns in a horizon h. We note that the loss for a portfolio characterized by weight vector $W_t \in \mathbb{R}^n$ is given by*

$$L(W) = -W_t^T R_{t+h}.$$

- *Assume $R_{t+h} \sim MN(\mu, \Sigma)$. Then ES with confidence level α is given by*

$$ES = -W_t^T\mu + \sqrt{W_t^T \Sigma W_t}\frac{\exp(-y^2/2)}{\sqrt{2\pi}(1-\alpha)}, y = \Phi^{-1}(\alpha),$$

where $\Phi(x)$ is the cdf for the standard normal distribution.

- *Assume* $R_{t+h} = B\mathcal{F}_t + \epsilon_t, B \in \mathbb{R}^{n\times m}, \mathcal{F}_t \sim MN(\mu, \Omega), \epsilon_t \sim MN(0, D)$*,* $\mathcal{F}_t, \epsilon_t$ *are m-dimension random vectors independent to each other. Then ES with confidence level* α *is given by*

$$ES = -W_t^T B\mu + \sqrt{W_t^T(B\Omega B^T + D)W_t}\frac{\exp(-y^2/2)}{\sqrt{2\pi}(1-\alpha)}, y = \Phi^{-1}(\alpha),$$

where $\Phi(x)$ *is the cdf for the standard normal distribution.*

Proof. (1)Note that $L(w) \sim N(W^T\mu, W^T\Sigma W)$. (2) Note that

$$E[L(w)] = E[-W_t^T R_{t+h}] = W^T B\mu, Cov[L(w)] = W_t^T(B\Omega B^T + D)W_t.$$

□

9.2.2 Loss with t distribution

Lemma 9.2.3 (VaR with t distribution of loss). *[1, p. 73] Assume* $L(w) \sim t_v(\mu, \sigma^2), v > 2$*(usually* $\mu < 0$*). Let* $L_1, L_2, ..., L_n$ *denote the random sample of* L*. Then*

$$VaR_\alpha(w; h) = \mu + T^{-1}(\alpha)\sigma,$$

where $T(x)$ *is the cdf for the standard normal distribution.*

If σ *is estimated using mothod of moments, we have*

$$VaR_\alpha(w; h) = \mu + T^{-1}(\alpha)\hat{\sigma}$$

where

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (L_i - \bar{L})^2 \frac{v-2}{v}$$

.

Proof. (1) Note that $L = \mu + \sigma Z, Z \sim t_v(0,1)$. Then $Var_\alpha(w;h) = F_L^{-1}(\alpha)$. We have

$$\begin{aligned}
Pr(L \leq F_L^{-1}(\alpha)) &= \alpha \\
Pr(\mu + \sigma Z \leq F_L^{-1}(\alpha)) &= \alpha, Z \sim t_v(0,1) \\
Pr(Z \leq \frac{F_L^{-1}(\alpha) - \mu}{\sigma}) &= \alpha \\
T(\frac{F_L^{-1}(\alpha) - \mu}{\sigma}) &= \alpha \\
T^{-1} \circ T(\frac{F_L^{-1}(\alpha) - \mu}{\sigma}) &= T^{-1}(\alpha) \\
\frac{F_L^{-1}(\alpha) - \mu}{\sigma} &= T^{-1}(\alpha) \\
F_L^{-1}(\alpha) &= \mu + T^{-1}(\alpha)\sigma
\end{aligned}$$

(2) Use method of moments estimator for t distribution. □

9.2.3 Loss with approximate normal distribution

Lemma 9.2.4 (Application of Cornish-Fisher expansion for approximate normal distribution of loss). *[3, p. 170] Suppose that the random loss L for a given horizon h has mean μ and variance σ^2. Then*

$$VaR_\alpha(L) \approx \mu + \sigma z_\alpha^{cf}$$

where

$$z_\alpha^{cf} = q_\alpha + \frac{(q_\alpha^2 - 1)S(L)}{6} + \frac{(q_\alpha^3 - 3q_\alpha)K(L)}{24} - \frac{(2q_\alpha^3 - 5q_\alpha)S^2(L)}{36},$$

where $S(L)$ is skewness, $K(L)$ is kurtosis, z_α^{cf} is the Cornish-Fisher approximate quantile value for the confidence level α, and q_α is the quantile value for the standard normal distribution with confidence level α.

Proof. Directly from **??**. □

9.2.4 Loss with polynomial tails

Lemma 9.2.5. *[4, p. 566] Assume the loss L of a portfolio have a polynomial right tail such that*

$$f_L(y) \sim Ay^{-(\beta+1)}, y \to \infty,$$

where $A, \beta > 0$ are constants. β is known as ***tail index****. It follows that*

- $$Pr(L > y) = \int_y^\infty f_L(u)du = \frac{A}{a}y^{-\beta}, y \to \infty.$$
- $$VaR(\alpha) \triangleq F_L^{-1}(\alpha) = [\frac{(1-\alpha)\beta}{A}]^{-1/\beta}.$$
- *Consider two confidence levels α_0, α_1, we have*

$$\frac{VaR(\alpha_1)}{Var(\alpha_0)} = (\frac{(1-\alpha_1)}{1-\alpha_0})^{1/\beta}.$$

That is, if we know the tail index α, and a VaR at α_0, then we can get $VaR(\alpha_1)$.

Proof. (1)

$$\begin{aligned} Pr(L > y) &= \int_y^\infty f_L(u)du \\ &= \int_y^\infty Au^{-(\beta+1)}du \\ &= -\frac{A}{\beta}u^{-\beta}|_y^\infty \\ &= \frac{A}{\beta}y^{-\beta} \end{aligned}$$

(2)

$$\begin{aligned} F_L(L < y) = 1 - \frac{A}{\beta}y^{-\beta} &= \alpha \\ \implies \frac{A}{\beta}y^{-\beta} &= 1-\alpha \\ y^{-\beta} &= \frac{(1-\alpha)\beta}{A} \\ VaR(\alpha) = y &= [\frac{(1-\alpha)\beta}{A}]^{-1/\beta} \end{aligned}$$

(3)

$$\frac{VaR(\alpha_0)}{VaR(\alpha_1)} = (\frac{(1-\alpha_0)}{1-\alpha_1})^{-1/\beta} = (\frac{(1-\alpha_1)}{1-\alpha_0})^{1/\beta}.$$

□

Lemma 9.2.6 (regression estimating of the tail index). *Assume the loss L of a portfolio have a polynomial right tail such that*

$$f_L(y) \sim Ay^{-(\beta+1)}, y \to \infty.$$

We know that

$$\log(Pr(L \geq y) = \log(A/\beta) - \beta \log y.$$

If $L_{(1)}, L_{(2)}, ..., L_{(n)}$ are the order statistics of losses sorted in descending order, then we can construct a set of paired data for regression $(k/n, L_k), k > m$, and use the following the regression model

$$\log(k/n) = \log(A/\beta) - \beta \log(L_{(k)}),$$

to get the slope β.

9.3 Historical simulation approach

Remark 9.3.1 (general approach).

- Historical simulation them involves the day-to-day changes in the values of market variables that occurred in the past in a direct way to estimate the probability distribution of the change in the value of the current portfolio between today to tomorrow.
- Historical simulation approach is the most popular non-parametric approach.

Definition 9.3.1 (scenario, market variable simulation). *[2, pp. 278, 282]*

- *Define v_i as the value of a market variable on Day i and suppose that today is day $n(i < n)$. The i scenario in the historical simulation approach assumes that the value of the market variable tomorrow will be*

$$\text{market value under } i \text{ scenario} = v_n \frac{v_i}{v_{i-1}}.$$

-

Remark 9.3.2 (weighting scenario). [2, p. 285] Suppose we have observations/scenarios for n day-to-day changes ordered in $i = 1, 2, ..., n$.

- in ordinary historical simulation, each scenario is given by $1/n$.
- in weighting historical simulation, each scenario is given by

$$\frac{\lambda^{n-i}(1-\lambda)}{1-\lambda^n}, 0 < \lambda < 1.$$

- in volatility weighting historical simulation, each scenario is given by

$$\textit{market value under i scenario} = v_n \frac{v_{i-1} + (v_i - v_{i-1})\sigma_{n+1}/\sigma_i}{v_{i-1}},$$

where σ_{n+1} denotes the volatility of market variable between today and tomorrow.

Lemma 9.3.1 (standard deviation of percentile estimation). *[2, p. 282] Suppose that the q-percentile of the distribution is estimated as x. The standard error of the estimate is*

$$\frac{1}{f(x)}\sqrt{\frac{(1-q)q}{n}},$$

where n is the number of observations and $f(x)$ is the estimate of the probability density function of the loss evaluated at x.

9.3.1 Delta-gamma approximation

Lemma 9.3.2 (delta-gamma approximation for VaR). *[2, p. 289] Suppose an instrument depends on several market variables, $S_i, i \leq i \leq k$.*

- *the price change of the instrument is given by*

$$\Delta P = \sum_{i=1}^{k} \delta_i \Delta S_i + \sum_{i=1}^{k}\sum_{j=1}^{k} \frac{1}{2}\gamma_{ij}\Delta S_i \Delta S_j,$$

  where δ_i and γ_{ij} are defined by

$$\delta_i = \frac{\partial P}{\partial S_i}, \gamma_{ij} = \frac{\partial^2 P}{\partial S_i \partial S_j},$$

  and ΔS_i is the market variable change.
- *Suppose for each market variable S_i, we have n historical scenarios indexed by $S_i^{(j)}, j = 1, 2, ..., n$. Then the price change under scenario m is given by*

$$\Delta P^{(m)} = \sum_{i=1}^{k} \delta_i \Delta S_i^{(m)} + \sum_{i=1}^{k}\sum_{j=1}^{k} \frac{1}{2}\gamma_{ij} S_i^{(m)} \Delta S_i^{(m)}.$$

- *The VaR with α percentile is the α percentile among the n losses given by $-\Delta P_1, -\Delta P_2, ..., -\Delta P_n$.*

Proof. Directly from the Taylor expansion and definition of VaR. □

Example 9.3.1 (the price change of a call option). [1, p. 104] Consider a call option price, denoted by C_t. Its price change is due to the market risk can be approximated by

$$C_{t+h} - C_t \approx Delta(S_{t+h} - S_t) + \frac{1}{2}Gamma(S_{t+h} - S_t)^2 + \frac{\partial C_t}{\partial t}h + Vega(\sigma_{t+h}^{imp} - \sigma_t^{imp}).$$

Lemma 9.3.3 (ladder gamma). *Suppose a function V depends on N factors denoted by $F_1, F_2, ..., F_N$. Define*

$$D_{u,i} = \frac{V(F_{j\neq i} + \delta, F_i + 2\delta) - V(F_{j\neq i} + \delta, F_i)}{(2\delta)},$$

and

$$D_{d,i} = \frac{V(F_{j\neq i} - \delta, F_i - 2\delta) - V(F_{j\neq i} - \delta, F_i)}{(-2\delta)}.$$

It follows that

- $$\frac{D_{u,i} - D_{d,i}}{2\delta} = \sum_{j\neq i} \frac{\partial^2 V}{\partial F_j \partial F_i} + \frac{\partial^2 V}{\partial F_i^2}.$$
- $$\sum_{i=1}^{N} \frac{D_{u,i} - D_{d,i}}{2\delta} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2 V}{\partial F_j \partial F_i}.$$

Proof. Use Taylor expansion, we have

$$\begin{aligned} V(F_{j\neq i} + \delta, F_i + 2\delta) = V(F) &+ \sum_{j\neq i} \frac{\partial V}{\partial F_j}\delta + \frac{\partial V}{\partial F_i}2\delta + \frac{1}{2}\sum_{j\neq i}\frac{\partial^2 V}{\partial F_j^2}\delta^2 \\ &+ \sum_{j\neq i}\sum_{k\neq i,j} \frac{\partial^2 V}{\partial F_j \partial F_k}\delta^2 + \sum_{j\neq i} \frac{\partial^2 V}{\partial F_j \partial F_k}\delta \cdot 2\delta + \frac{1}{2}\frac{\partial^2 V}{\partial F_i^2}(2\delta)^2 \end{aligned}$$

$$\begin{aligned} V(F_{j\neq i} - \delta, F_i - 2\delta) = V(F) &- \sum_{j\neq i} \frac{\partial V}{\partial F_j}\delta - \frac{\partial V}{\partial F_i}2\delta + \frac{1}{2}\sum_{j\neq i}\frac{\partial^2 V}{\partial F_j^2}\delta^2 \\ &+ \sum_{j\neq i}\sum_{k\neq i,j} \frac{\partial^2 V}{\partial F_j \partial F_k}\delta^2 + \sum_{j\neq i} \frac{\partial^2 V}{\partial F_j \partial F_k}\delta \cdot 2\delta + \frac{1}{2}\frac{\partial^2 V}{\partial F_i^2}(2\delta)^2 \end{aligned}$$

$$V(F_{j\neq i}+\delta, F_i) = V(F) + \sum_{j\neq i}\frac{\partial V}{\partial F_j}\delta + \frac{1}{2}\sum_{j\neq i}\frac{\partial^2 V}{\partial F_j^2}\delta^2 + \sum_{j\neq i}\sum_{k\neq i,j}\frac{\partial^2 V}{\partial F_j \partial F_k}\delta^2$$
$$V(F_{j\neq i}-\delta, F_i) = V(F) - \sum_{j\neq i}\frac{\partial V}{\partial F_j}\delta + \frac{1}{2}\sum_{j\neq i}\frac{\partial^2 V}{\partial F_j^2}\delta^2 + \sum_{j\neq i}\sum_{k\neq i,j}\frac{\partial^2 V}{\partial F_j \partial F_k}\delta^2$$

Then

$$\begin{aligned}
&\frac{D_{u,i}-D_{d,i}}{2\delta}\\
&= \frac{V(F_{j\neq i}+\delta, F_i+2\delta) - V(F_{j\neq i}+\delta, F_i)}{(2\delta)^2} + \frac{V(F_{j\neq i}-\delta, F_i-2\delta) - V(F_{j\neq i}-\delta, F_i)}{(2\delta)^2}\\
&= \frac{2\sum_{j\neq i}\frac{\partial^2 V}{\partial F_j\partial F_k}\delta\cdot 2\delta + \frac{\partial^2 V}{\partial F_i^2}(2\delta)^2}{(2\delta)^2}\\
&= \sum_{j\neq i}\frac{\partial^2 V}{\partial F_j \partial F_i} + \frac{\partial^2 V}{\partial F_i^2}
\end{aligned}$$

□

9.4 VaR calculation via extreme value theory

Definition 9.4.1 (conditional tail cdf). *[2, p. 290] Let L be the random loss with cdf F_L. The conditional tail cdf of L with loss threshold parameter u is defined by*

$$F_L(y;u) \triangleq Pr(L < u+y | L > u) = \frac{Pr(u < L < u+y)}{Pr(L > u)} = \frac{F(u+y)-F(u)}{1-F(u)}.$$

Remark 9.4.1. Note that, for a fixed parameter u, $F_L(y;u)$ is a qualified cdf since it is nondecreasing with respect to y and $F_L(\infty;u) = 1$.

Proposition 9.4.1 (approximate parametric tail distribution from Extreme value theory). *[2, p. 290]*

- *Let $u \in \mathbb{R}$ denote the loss threshold parameter. When u is large[a],*

$$Pr(L < u+y | L > u) \approx G(y;\xi,\beta) = 1 - [1+\xi\frac{y}{\beta}]^{-1/\xi},$$

where ξ is the shape parameter and β is the scale parameter.

- *The associated density function for the excess loss $Y = L - u$ is given by*

$$g(y;\xi,\beta) = \frac{1}{\beta}(1+\frac{\xi y}{\beta})^{-1/\xi-1}$$

- *The resulting parametric distribution for loss with threshold parameter u is given by*

$$\begin{aligned} Pr(L<x) &= 1-(1-G(x-u;\xi,\beta))Pr(L>u) \\ &\approx 1-(1-G(x-u;\xi,\beta))\frac{N_u}{N} \\ &= 1-\frac{N_u}{N}[1+\xi\frac{x-u}{\beta}]^{-1/\xi} \end{aligned}$$

  where N is the total sample number and N_u is the number of samples with loss greater than u.

[a] by large, we usually mean at least 0.95 percentile of L

Proof. (1) Todo; (2) take derivative of G with respect to y. (3) Note that

$$\begin{aligned} Pr(L<x) &= 1-Pr(L>x) \\ &= 1-Pr(L>x|L>u)Pr(L>u) \\ &= 1-Pr(L-u+u>x|L>u)Pr(L>u) \\ &= 1-Pr(L-u>x-u|L>u)Pr(L>u) \\ &= 1-(1-G(x-u;\xi,\beta))Pr(L>u) \\ &\approx 1-(1-G(x-u;\xi,\beta))\frac{N_u}{N} \end{aligned}$$

□

Note 9.4.1 (interpretation)**.**

- Note that $G(y;\xi,\beta)$ is a qualified cdf since it is nondecreasing with respect to y and $G(\infty;\xi,\beta) = 1$.
- **(Connection to power law)** If we set $u = \beta/\xi$, then

$$Pr(L>x) = \frac{N_u}{N}[\frac{\xi}{\beta}]^{-1/\xi},$$

  which is an example of power law distribution.

Lemma 9.4.1 (maximum likelihood parameter estimation for parametric tail distribution). *[2, p. 290] For a given loss threshold parameter u, the likelihood function for parameter ξ, β with N_u observed losses above u, denoted by $l_1, l_2, ..., l_{N_u}$, is given by*

$$L(\xi,\beta)\prod_{i=1}^{N_u}\frac{1}{\beta}(1+\frac{\xi(l_i-u)}{\beta})^{-1/\xi-1}.$$

Proof. From extreme value theory(Theorem 9.4.1), the excess loss denoted by $l_1 - u, l_2 - u, ..., l_{N_u} - u$ will have density function given by

$$g(y;\xi,\beta)=\frac{1}{\beta}(1+\frac{\xi y}{\beta})^{-1/\xi-1}.$$

□

Proposition 9.4.2 (calculation of VaR and ES). *[2, p. 290]*

Suppose we have parameter u, β, ξ characterizing the tail distribution of the loss L. Let N be the total sample number and N_u be the number of samples with loss greater than u. Then

- *The VaR with confidence level q is given by*

$$VaR_L(q)=F_L^{-1}(q)=u+\frac{\beta}{\xi}(\frac{N}{N_u}(1-q)^{-\xi}-1).$$

- *The ES with confidence level q is given by*

$$ES=\frac{VaR+\beta-\xi u}{1-\xi}.$$

Proof. (1)From tail distribution (Theorem 9.4.1), we know that

$$Pr(L<x)=1-\frac{N_u}{N}[1+\xi\frac{x-u}{\beta}]^{-1/\xi}.$$

With confidence q, we solve

$$Pr(L<VaR)=1-\frac{N_u}{N}[1+\xi\frac{VaR-u}{\beta}]^{-1/\xi}=q,$$

and get

$$VaR_L(q)=F_L^{-1}(q)=u+\frac{\beta}{\xi}(\frac{N}{N_u}(1-q)^{-\xi}-1).$$

(2) From the definition of ES(Definition 9.1.3), we have

$$\begin{aligned}
ES &= \frac{1}{1-q}\int_q^1 VaR(s)ds \\
&= u - \frac{\beta}{\xi} + \frac{\beta}{\xi}(\frac{N}{N_u}(1-q)^{-\xi})/(1-\xi) \\
&= VaR_L(q)/(1-\xi) + u(1-1/(1-\xi)) - \frac{\beta}{\xi}(1-1/(1-\xi)) \\
&= \frac{VaR+\beta-\xi u}{1-\xi}
\end{aligned}$$

□

9.4.1 Credit portfolios

Lemma 9.4.2 (large uniform portfolio approximation for default modeling). *[5, p. 564] Consider a bank with a very large portfolio of similar loans. As an approximation, lLet $T_1, T_2, ..., T_n$ denote the random default time for n parties. Assume the marginal cdfs of default time are all given by $Q_i(t) = Q(t), i = 1, 2, ..., n$. Let $x_i = \phi^{-1}(Q(t)), i = 1, 2, ..., n$, and assume x_i follows*

$$x_i = \sqrt{\rho}F + \sqrt{1-\rho}Z_i, i = 1, 2, ..., n,$$

where F is a common factor affecting defaults for all firms and Z_i is a factor affecting only firm i. F and Z_i are independent standard normal variables. It follows that

- *Conditioning on the value $F = f$, the probability of default for each loan before T years is*

$$Pr(T_i < t | F = f) = \phi(\frac{\phi^{-1}(Q(t)) - \sqrt{\rho}f}{\sqrt{1-\rho}}).$$

- ***the percentage of defaults** among n loans with confidence level α before t years on this large portfolio will be*

$$V(\alpha, t) = \phi(\frac{\phi^{-1}(Q(t)) - \sqrt{\rho}\phi^{-1}(1-\alpha)}{\sqrt{1-\rho}}),$$

  where ϕ is the standard normal cdf.
- *The VaR for the loan size L before T years is given by*

$$VaR(\alpha) = L(1-R)V(\alpha, t).$$

Proof. (1) See Lemma 7.4.3. (2) Note that conditioning on F, each loan's default is independent from others. In the large number approximiation, the default percentage conditioning on F is also $Pr(T_i < t|F = f)$. Because $Pr(T_i < t|F = f)$ is the decreasing function of f, therefore we take f to be $1 - \alpha$ percentile. (3) Note that $L(1 - R)$ is the loss at default for each loan. □

Example 9.4.1. [5, p. 565] Suppose that a bank has a large pool of outstanding loans. The 1-year marginal probability of default is average 0.02%. The average correlation among the loans is estimated to be 0.1. Then the percentage of defaults among all loans with confidence level 0.999 before 1 years is given by

$$V(0.999, t) = \phi(\frac{\phi^{-1}(0.02) - \sqrt{0.1}\phi^{-1}(1 - 0.999)}{\sqrt{1 - 0.1}}) = 0.128.$$

9.5 Notes on bibliography

The major reference are [1][6] [3].

BIBLIOGRAPHY

1. Roncalli, T. Lecture Notes on Risk Management & Financial Regulation (2016).
2. Hull, J. *Risk management and financial institutions,+ Web Site* (John Wiley & Sons, 2012).
3. Alexander, C. *Market Risk Analysis, Value at Risk Models* ISBN: 9780470745076. https://books.google.com/books?id=j5l82vMfcbQC (Wiley, 2009).
4. Ruppert, D. *Statistics and data analysis for financial engineering, 2ed* (Springer, 2015).
5. Hull, J. *Options, Futures and Other Derivatives 9 edition* (Pearson Education, 2009).
6. Jorion, P. *Value at Risk, 3rd Ed.: The New Benchmark for Managing Financial Risk* ISBN: 9780071736923. https://books.google.com/books?id=nnblKhI7KP8C (McGraw-Hill Education, 2006).

A

SUPPLEMENTAL MATHEMATICAL FACTS

A.1 Basic logic for proof

[1, p. 60]The negation of

for any $\epsilon > 0$, there exist $N > 0$, such that for all $n > N$, we have $|a_n - a| < \epsilon$

is

there exist an $\epsilon > 0$, such that for every $N > 0$, such that for all $n > N$, we have $|a_n - a| > \epsilon$.

[1, p. 60]The negation of

for any $\epsilon > 0$, there exist $\delta > 0$, such that for all $|x - x_0| < \delta$, we have $|f(x) - f(x_0)| < \epsilon$

is

there exist an $\epsilon > 0$, such that for every $\delta > 0$, such that for all $|x - x_0| < \delta$, we have $|f(x) - f(x_0)| > \epsilon$.

A.2 Some common limits

Lemma A.2.1 (Stirling approximation).
- *For positive integer n,*

$$\ln n! = n \ln n - n + O(\ln n).$$

-
$$\sqrt{2\pi} n^{n+1/2} e^{-n} \leq n! \leq e n^{n+1/2} e^{-n}, \forall n > 0.$$

Lemma A.2.2 (common limits summary).

-
$$\lim_{n\to\infty} \frac{\ln n}{n} = 0.$$
-
$$\lim_{n\to\infty} \frac{x^n}{n!} = 0, \forall x \in \mathbb{R}.$$
-
$$\lim_{n\to\infty} \frac{n!}{n^n} = 0.$$
-
$$\lim_{n\to\infty} M^{1/n} = 1$$
  for any $M > 0$.
-
$$\lim_{n\to\infty} \frac{\ln n!}{n} = \infty, \lim_{n\to\infty} (n!)^{1/n} = \infty.$$

Proof. (2)see **??** and **??**.(3)**??**.(4) **??**. (5) (a)Use Stirling approximation $\ln n! = n \ln n - n + O(\ln n)$ and $\ln n!/n = n - 1 + O(\ln n/n) \to \infty$.(b) Note that $(n!)^{1/n} = \exp(\ln(n!)^{1/n}) = \exp(\frac{\ln n!}{n})$. □

Note A.2.1. A helpful and general summary, as $n \to \infty$

$$\ln n \ll n^r (r > 0) \ll a^n (a > 1) \ll n! \ll n^n.$$

Lemma A.2.3 (property of e). *Define*

$$\lim_{n\to\infty}(1+1/n)^n = e$$

and then

$$\lim_{n\to\infty}(1+x/n)^n = e^x$$

for any real x.

Proof.

$$\lim_{n\to\infty}((1+x/n)^{n/x})^x = e^x$$

use the fact the $f(y) = y^x$ is continuous, such that function evaluation and limit can be exchanged. □

A.3 Common series summation

Lemma A.3.1. *[2, p. 1]*

- $$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$
- $$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$
- $$\sum_{k=1}^{n} k^3 = [\frac{n(n+1)}{2}]^2$$

Lemma A.3.2. *[2, p. 1]* *Assume $q \neq 1$.*

- $$\sum_{k=1}^{n} aq^{k-1} = a\frac{q^n-1}{q-1}$$
- $$\sum_{k=0}^{n-1} kq^k = \frac{(n-1)q^n}{q-1} + \frac{(q-q^n)}{(q-1)^2}$$

- $$\sum_{k=0}^{n-1}(n-1-k)q^k = -(n-1)\frac{1}{q-1} - \frac{(q-q^n)}{(q-1)^2}$$

Proof. (3) use (1)(2), we have

$$\sum_{k=0}^{n-1}(n-1-k)q^k = (n-1)\frac{q^n-1}{q-1} - \frac{(n-1)q^n}{q-1} - \frac{(q-q^n)}{(q-1)^2}$$

□

A.4 Some common spaces

The metric space $(\mathbb{R}^n, d_2)$ is the set $\mathbb{R}^n$ with metric $d_2(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$.

[3, p. 122]The metric space l^2 is the set of all infinite sequence of real or complex numbers $\{x_1, x_2, ...\}$ such that $\sum_{i=1}\infty x_i^2 < \infty$,i.e., $\sum_{i=1}\infty x_i^2$ converges. The metric is usually defined as

$$d_2(\{x_n\},\{y_n\}) = \sqrt{\sum_{k=1}^{\infty}(x_i - y_i)^2}$$

The metric space $l^p, 1 \leq p < \infty$,is the set of all infinite sequence of real or complex numbers $\{x_1, x_2, ...\}$ such that $\sum_{i=1}\infty |x_i|^p < \infty$,i.e., $\sum_{i=1}\infty |x_i|^p$ converges. The metric is usually defined as

$$d_p(\{x_n\},\{y_n\}) = \sqrt[p]{\sum_{k=1}^{\infty}(x_i - y_i)^p}$$

The metric space l^∞, is the set all infinite sequence of real or complex numbers $\{x_1, x_2, ...\}$ such that every x_i is bounded. The metric is defined as

$$d_\infty(\{x_n\},\{y_n\}) = \sup_n |x_n - y_n|$$

[4, p. 75].The metric space $C[a,b] = (C[a,b], d_\infty)$ denote the set of real-valued(or complex valued) functions defined on the interval $[a,b]$. The metric d_∞ is given as

$$d_\infty(x,y) = \sup_t |x(t) - y(t)|$$

Remark A.4.1. **Caution!** Sometimes $C[a,b]$ refers to only continuous functions.[5, p. 23]

The metric space $(C[a,b], d_p)$ denote the set of real-valued(or complex valued) functions defined on the interval $[a,b]$. The metric d_p is given as

$$d_p(x,y) = [\int_a^b |x(t) - y(t)|^p \, dt]^{1/p}$$

where $1 \leq p < \infty$.[4, p. 75].

The vector space $\mathcal{L}(V,W)$ usually denotes the set of all linear operators from V into W.

A.4.1 Notations on continuously differentiable functions

- C^0 refers to continuous function
- C^1 refers to functions having continuous first derivatives, also called continuously differentiable functions.
- C^2 refers to functions having continuous second derivatives
- C^∞ refers to smooth functions

A.5 Different modes of continuity

Chain of inclusions for functions over a closed and bounded subset of the real line

$$continuouslyDifferntiable \subseteq LipschitzContinouous \subseteq UniformlyContinuous$$

Remark A.5.1.

- Continuously differentiable on a closed interval indicates the derivative is bounded $f' \leq M$, then we have
$$|f(x) - f(y)| = f'(s)|x-y| \leq M|x-y|$$
hence Lipschitz continuous.
- $f(x) = |x|$ is Lipschitz continuous but is not differentiable everywhere except at $x = 0$, therefore it is not continosly differentiable.
- Lipschitz continuous $\rightarrow$ continous:
$$|f(x) - f(y)| \leq L|x-y| \rightarrow 0$$
as $|x-y| \rightarrow 0$

Lemma A.5.1 (differentiable implies continuous). *If f is differentiable on $[a,b]$, then it is continuous on $[a,b]$.*

Proof:

$$\lim_{y \to x} f(y) - f(x) = \lim_{y \to x} (y-x)(f(y) - f(x))/(y-x) = \\ \lim_{y \to x}(y-x) \lim_{y \to x}(f(y)-f(x))/(y-x) = 0$$

where we have use the property that if two limits exist then they can multiply.[3, p. 42].

Remark A.5.2. This lemma indicates that a function differentiable everywhere will be continous everywhere.

Lemma A.5.2 (differentiable everywhere NOT implies continuously differentiable). *A function is differentiable everywhere NOT implies it is continuously differentiable function.*

The standard example is

$$f(x) = \begin{cases} x^2 \sin(\frac{1}{x}), x \neq 0 \\ 0, x = 0 \end{cases}$$

This function can be differentiated every where and $f'(0) = 0$, but $\lim_{x \to 0} f'(x)$ does not exist. See link.

A.5.1 continuity vs. uniform continuity

Definition A.5.1. *A function $f : X \to Y$ is* uniformly continuous *if for every $\epsilon > 0$ there exist a $\delta > 0$ such that for every $x, x_0 \in X$,*

$$\rho(x, x_0) \leq \delta \Rightarrow \rho(f(x), f(x_0)) < \epsilon$$

Proposition A.5.1. *[3, p. 154] If f is a continuous function from a compact metric space M_1 into a metric space M_2, then f is uniformly continuous on M_1.*

Corollary A.5.1.1. *[3, p. 154] If f is a continuous real-value function on a closed and bounded subset X of $\mathbb{R}^n$, then f is uniformly continuous on X.*

Example A.5.1. The function $f(x) = x^2$ is continuous but not uniformly continuous on the interval $(0, \infty)$.

Lemma A.5.3 (sufficient condition). *Let $S = \mathbb{R}$. if f is global Lipschitz continuous, i.e.*

$$|f(x_1) - f(x_2)| < M|x_1 - x_2|$$

$\forall x_1, x_2 \in S$, then f is uniformly continuous.

Proof: $|f(x_1) - f(x_2)| < M|x_1 - x_2| \to 0$

A.6 Exchanges of limits

A.6.1 Overall remark

Remark A.6.1.

- Usually, the necessary conditions for exchanging limits is difficult to find, therefore only sufficient conditions are given.
- Many operations are in nature taking limits, for example, summing infinite terms is taking limits on partial sums; integrals is taking limits on both summation and partitions; derivative is taking limits on quotient expressions.

A.6.2 exchange limits with infinite summations

Let $\lim_{m\to\infty} \sum_{n=1}^{\infty} f(m,n) = \sum_{n=1}^{\infty} \lim_{m\to\infty} f(m,n)$ Based on dominated convergence, if there is a $g(n)$ such that $f(m,n) < g(n), \forall m$ and $\sum_{n=1}^{\infty} g(n)$ exists, then we can exchange.

To use the dominated convergence theorem in Lebesgue integral, we can define a simple function $s_n$on $[0,\infty]$ take $f(m,n)$ on the interval $[m-1,m)$. Then the integral of s_n with respect to Legesbue measure on real lime will give the $\int_{[0,\infty)} s_n d\mu = \sum_{m=1}^{\infty} f(m,n)$

Proposition A.6.1. *[3, pp. 94, 373]Let $a_{m,n}$ be non-negative and $\sum_m^{\infty} \sum_n^{\infty} a_{m,n}$ exists, then*

$$\sum_m^{\infty} \sum_n^{\infty} a_{m,n} = \sum_n^{\infty} \sum_m^{\infty} a_{m,n}$$

Corollary A.6.1.1. *Let $a_{m,n}$ be increasing on both m, n and $\lim_{m\to\infty} \lim_{n\to\infty} a_{m,n}$ exists, then*

$$\lim_{m\to\infty} \lim_{n\to\infty} a_{m,n} = \lim_{n\to\infty} \lim_{m\to\infty} a_{m,n}$$

Proof: by constructing partial sums.

A.6.3 Exchange limits with integration and differentiation

Proposition A.6.2. *[3, p. 249] Let α be a function of bounded variation on $[a, b]$ and let f_n be a sequence of functions in $\mathcal{R}_\alpha[a, b]$ which converges uniformly to a function f. Then $f \in \mathcal{R}_\alpha[a, b]$ and*

$$\lim_{n\to\infty} \int_a^b f_n d\alpha = \int_a^b \lim_{n\to\infty} f d\alpha$$

Proposition A.6.3. *[3, p. 249] Let $\{f_n\}$ be a sequence of differentiable functions on (a, b). Suppose that*

- *f_n' is continuous on (a, b)*
- *$\{f_n\}$ converges pointwise to f*
- *$\{f_n'\}$ converges uniformly*

then f is differentiable on (a, b) and f_n' converges uniformly to f'.

A.6.4 Exchange differentiation with integration

Proposition A.6.4. *Let $f(x, y)$ be continuous on $[a, b] \times [c, d]$. Then*

$$\phi(y) = \int_a^b f(x, y) dx$$

defined above is continuous function on $[c, d]$

Proof: for any $\epsilon > 0$, there exist δ, such that

$$|\phi(y) - \phi(y')| \leq \int_a^b |f(x, y) - f(x, y')| \, dx \leq \epsilon(b - a) \forall |y - y'| < \delta$$

where we have the fact of $f(x, y) - f(x, y)$ is bounded (since continuous function on a compace set is uniformly continuous and will have maximum and minimum) which shows $\phi(y)$is uniformly continuous.

Proposition A.6.5. *Let f and f_y be continuous on $[a,b] \times [c,d]$. Then ϕ is differentiable and*

$$\phi_y = \int_a^b f_y(x,y)dx$$

Proof:

$$\frac{\phi(y+h)-\phi(y)}{h} = \frac{1}{h}\int_a^b f(x,y+h) - f(x,y)dx = \int_a^b f_y(x,z)$$

due to Taylor theorem, where $z \in [y, y+h]$. Then

$$\left|\frac{\phi(y+h)-\phi(y)}{h} - \int_a^b f_y(x,y)dx\right| \leq \int_a^b \left|f_y(x,z) - f_y(x,y)\right| dx$$

Because f_y is continuous on compact set, then it is uniformly continuous. Therefore given $\epsilon > 0$, there exists δ such that

$$\left|f_y(x,y') - f_y(x,y)\right| < \epsilon/(b-a), \forall \left|y-y'\right| < \delta$$

Taking $h < \delta$, we have

$$\left|\frac{\phi(y+h)-\phi(y)}{h} - \int_a^b f_y(x,y)dx\right| < \epsilon.$$

Take the limit on h and we get the result.

A.6.5 Exchange limit and function evaluations

Lemma A.6.1. *Let $\{x_n\}$ be a sequence with limit x, let f be a continuous function*

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(x)$$

Proof: from the definition of continuous function.

A.7 Useful inequalities

Lemma A.7.1 (arithmetic-geometric mean inequality). *For $x_1, ..., X_n \geq 0$, we have*

$$(x_1 x_2 ... x_n)^{1/n} \leq \sum_{i=1}^{n} x_i / n.$$

Specifically,

$$\frac{x_1 + x_2}{x} \geq \sqrt{x_1 x_2}$$

Proof. use $y = \ln(x)$ and concavity of $\ln(x)$ □

A.7.1 Gronwall's inequality

see [6]

A.7.2 Inequality for norms

Lemma A.7.2. *[7]For L^p normed space, we have*

$$\|x\|_2 \leq \|x\|_1$$

where

$$\|x\|_2 = (\int_{-\infty}^{\infty} |f(x)|^2 \, d\mu(x))^{0.5}$$

and

$$\|x\|_1 = \int_{-\infty}^{\infty} |f(x)| \, d\mu(x)$$

Proof: for finite dimensional normed space cases: we need to prove

$$\sqrt{x_1^2 + x_2^2 + ... + x_n^2} \leq |x_1| + |x_2| + ... + |x_n|$$

By squaring both sides, we can get the result. For continuous case, TODO

Proposition A.7.1. *[7]For L^p normed space, we have*

$$\|x\|_q \leq \|x\|_q$$

whenever $p \leq q$ where

$$\|x\|_q = (\int_{-\infty}^{\infty} |f(x)|^q \, d\mu(x))^{1/q}$$

Proof: todo

Remark A.7.1. For complete description on L^p norms, see [7]

A.7.3 Young's inequality for product

Lemma A.7.3. *If $a, b \geq 0$, and $p, q > 1, 1/p + 1/q = 1$, then*

$$ab \leq a^p/p + b^q/q$$

Proof:

$$\log(a^p/p + b^q/q) \geq \log(a^p)/p + \log(a^q)/q = \log(a) + \log(b) = \log(ab)$$

where we use the fact of log is concave.

A.8 Useful properties of matrix

A.8.1 Matrix derivatives

Lemma A.8.1 (common matrix derivative in quadratic forms). *[8] For $A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n$, we have:*

$$\frac{\partial a^T x}{\partial x} = \frac{\partial x^T a}{\partial x} = a$$

$$\frac{\partial Ax}{\partial x} = A$$

$$\frac{\partial BAx}{\partial x} = BA$$

$$\frac{\partial x^T Ax}{\partial x} = (A + A^T)x$$

$$\frac{\partial x^T Ax}{\partial x} = 2A$$

Lemma A.8.2. *If $f(x) = g(Ax), A \in \mathbb{R}^{n\times n}, x \in \mathbb{R}^n$ for some differentiable function $g(y)$, then*

$$\nabla f = A^T \nabla g$$

In particularly, $a \in \mathbb{R}^n$, then

$$\nabla a^T Ax = A^T x$$

A.8.2 Matrix inversion lemma

Lemma A.8.3 (matrix inversion lemma). *[9, p. 120]*

- $(E - FH^{-1}G)^{-1} = E^{-1} + E^{-1}F(H - GE^{-1}F)^{-1}GE^{-1}$
- $(E - FH^{-1}G)^{-1}FH^{-1} = E^{-1}F(H - GE^{-1}F)^{-1}$

Proof. (1) can be verified (2) expand the parenthesis using (1). □

Corollary A.8.0.1 (matrix inversion of rank one update). *Let $H = I$, $F = \pm u \in \mathbb{R}^n$, and $G = \pm v \in \mathbb{R}^n$, we have*

- $$(E - uv)^{-1} = E^{-1} - \frac{E^{-1}uv^TE^{-1}}{1 + v^TE^{-1}u}$$
- $$(E - uv)^{-1} = E^{-1} + \frac{E^{-1}uv^TE^{-1}}{1 - v^TE^{-1}u}$$

A.8.3 Block matrix

Lemma A.8.4. *Given an $(m \times p)$ matrix A with q row partions and s colun partitions and a $(p \times n)$ matrix B with s row partions and r colun partitions,*

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{q1} & A_{q2} & \cdots & A_{qs} \end{pmatrix}, B = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1r} \\ B_{21} & B_{22} & \cdots & B_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{sr} \end{pmatrix},$$

then the matrix product

$$C = AB$$

can be formed blockwise, giving C *as an* $(m \times n)$ *matrix with* q *row partitions and* r *column partitions. In particular,*

$$C_{\alpha\beta} = \sum_{\gamma=1}^{s} A_{\alpha\gamma} B_{\gamma\beta}.$$

Lemma A.8.5 (sum of vector product to matrix product). *Consider column vectors* $x_1, x_2, ..., x_N \in \mathbb{R}^d$ *and column vectors* $x_1, x_2, ..., x_N \in \mathbb{R}^d$*. It follows that*

- $$\sum_{i=1}^{N} x_i^T y_i = X_C^T Y_C,$$

  where $X_C \in \mathbb{R}^{Nd}$ *is a vector stacking all the* $x_1, ..., x_N$ *(similarly* Y_C*).*
- $$\sum_{i=1}^{N} x_i y_i^T = X_R^T Y_R,$$

  where $X_R \in \mathbb{R}^{N \times d}$ *is a matrix stacking all the* $x_1^T, ..., x_N^T$ *(similarly* Y_R*).*

Lemma A.8.6 (block matrix inversion formula).

$$\begin{aligned} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} &= \begin{pmatrix} A_{11}^{-1} + B_{12} B_{22}^{-1} B_{21} & -B_{12} B_{22}^{-1} \\ -B_{22}^{-1} B_{21} & B_{22}^{-1} \end{pmatrix} \\ &= \begin{pmatrix} C_{11}^{-1} & -C_{11}^{-1} C_{12} \\ -C_{21} C_{11}^{-1} & A_{22}^{-1} + C_{21} C_{11}^{-1} C_{12} \end{pmatrix} \end{aligned}$$

where

$$B_{22} = A_{22} - A_{21} A_{11}^{-1} A_{12}, B_{12} = A_{11}^{-1} A_{12}, B_{21} = A_{21} A_{11}^{-1}$$

and

$$C_{11} = A_{11} - A_{12} A_{22}^{-1} A_{21}, C_{12} = A_{12} A_{22}^{-1}, C_{21} = A_{22}^{-1} A_{21}$$

A.8.4 Matrix trace

Lemma A.8.7. • $\|A\|_F^2 = Tr(AA^T)$

Lemma A.8.8 (matrix trace).

- *(linearity)* $Tr(aA + bB) = aTr(A) + bTr(B)$
- *(commutative)* $Tr(AB) = Tr(BA)$
- *(invariance under transposition)* $Tr(A) = Tr(A^T)$
- *(cyclic rule)* $Tr(ABCD) = Tr(DABC) = Tr(CDAB) = Tr(BCDA)$ *or* $Tr(ABC) = Tr(CAB) = Tr(BCA)$

Proof. (1)(2)(3) can be proved directly from definition. (4) We can group three elements together and commute with the fourth. For example, we can group (ABC) together and commute with D to prove the first equality. □

Corollary A.8.0.2.

- *(invariance under similar transformation)* $Tr(PAP^{-1}) = Tr(A)$
- $Tr(X^TY) = Tr(XY^T) = Tr(Y^TX) = Tr(XY^T)$

Proof. (1) Use cyclic rule. (2) Use invariance under transposition and commutative rule. □

Lemma A.8.9 (common matrix derivative involving matrix trace). *[8] Let* $A, X, B \in \mathbb{R}^{m\times m}$*. We have*

$$\frac{\partial Tr(X)}{\partial X} = I$$

$$\frac{\partial Tr(XA)}{\partial X} = \frac{\partial Tr(AX)}{X} = A^T$$

$$\frac{\partial Tr(X^TA)}{\partial X} = \frac{\partial Tr(AX^T)}{X} = A$$

$$\frac{\partial Tr(AXB)}{\partial X} = \frac{\partial Tr(BAX)}{X} = A^TB^T$$

$$\frac{\partial Tr(AX^TB)}{\partial X} = \frac{\partial Tr(BAX^T)}{X} = BA$$

$$\frac{\partial Tr(XX^T)}{\partial X} = 2X$$

$$\frac{\partial Tr(XX)}{\partial X} = 2X^T$$

Additional, we have chain rule given by

$$\frac{\partial Tr(X^T A^T AX)}{\partial X} = \frac{\partial Tr(XX^T A^T A)}{\partial X} = \frac{\partial Tr(XX^T A^T A)}{\partial XX^T}\frac{\partial XX^T}{\partial X} = 2A^T AX.$$

A.8.5 Matrix elementary operator

Lemma A.8.10 (elementary operator matrix). *Left multiply a matrix A by an elementary matrix is equivalent to performing row operation. The elementary operator matrix is given by*

- *(Interchange row i and j) For example, exchange row 2 and row 3:*

$$R_1 = \begin{bmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 1 \end{bmatrix}$$

- *(Multiply row i by s) For example*

$$R_2 = \begin{bmatrix} 1 & & & \\ & s & & \\ & & 1 & \\ & & & 1 \end{bmatrix}$$

- *(Add s times row i to row j) For example, add s times row 2 to row 3*

$$R_3 = \begin{bmatrix} 1 & & & \\ & 1 & 0 & \\ & s & 1 & \\ & & & 1 \end{bmatrix}.$$

Note that $R_3 = R_1 R_2 \neq R_2 R_1$.

Lemma A.8.11 (elementary operator matrix). *Right multiply a matrix A by an elementary matrix is equivalent to performing row operation. The elementary operator matrix is given by*

- *(Interchange column i and j) For example, exchange row 2 and row 3:*

$$C_1 = \begin{bmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 1 \end{bmatrix}$$

- *(Multiply column i by s) For example*

$$C_2 = \begin{bmatrix} 1 & & & \\ & s & & \\ & & 1 & \\ & & & 1 \end{bmatrix}$$

- *(Add s times column i to column j) For example, add s times column 2 to column 3*

$$C_3 = \begin{bmatrix} 1 & & & \\ & 1 & s & \\ & & 1 & \\ & & & 1 \end{bmatrix}.$$

Note that $C_3 = C_1C_2 \neq C_2C_1$.

A.8.6 Matrix determinant

Lemma A.8.12 (properties of determinant).

- *All elementary operator matrix has determinant 1.*
- *For matrix $A \in \mathbb{R}^{n\times n}$,*
$$det(kA) = k^n det(A).$$
- $det(AB) = det(A)det(B)$.

- *All elementary operation on a matrix will not change its determinant.*

A.9 Numerical integration

Definition A.9.1 (Newton-Cotes Formula). *Suppose we want to evaluate $\int_a^b f(x)dx$. We can evaluate $f(x)$ at $n+1$ equally spacing points $x_i = a + i(b-a)/n$, and then we approximate $f(x)$ by n degree of Lagrange polynomial and do the integral. Specifically, we have*

$$\int_a^b f(x)dx \approx \int_a^b L(x)dx = \int_a^b (\sum_{i=0}^n f(x_i)l_i(x)) = \sum_{i=1}^n f(x_i) \int_a^b l_i(x)dx = \sum_{i=1}^n f(x_i)w_i$$

where L is the Lagrange polynomial of degree n, and $l_i(x), i = 0, ..., n$ is the (n+1) Lagrange polynomial basis, given as **??**.

Example A.9.1. Consider we use degree 1 Lagrange polynomial to approximate $f(x)$, then

$$L(x) = f(a)\frac{x-a}{b-a} + f(b)\frac{x-b}{a-b}$$

where $l_0(x) = \frac{x-a}{b-a}$ and $l_1(x) = \frac{x-b}{a-b}$. Then

$$w_0 = \int_a^b l_0(x)dx = \frac{1}{2}, w_1 = \int_a^b l_1(x)dx = \frac{1}{2}.$$

Table A.9.1: Closed Newton-Cotes Formula

Notation: $\int_a^b f(x)dx$, $f_i = f(x_i), x_i = a + i(b-a)/n$			
Degree	Name	Formula	Error term
1	Trapezoid rule	$\frac{b-a}{2}(f_0 + f_1)$	$-\frac{(b-a)^3}{12} f^{(2)}(\eta)$
2	Simpson's rule	$\frac{b-a}{6}(f_0 + 4f_1 + f_2)$	$-\frac{(b-a)^5}{2880} f^{(4)}(\eta)$
3	Simpson's 3/8 rule	$\frac{b-a}{8}(f_0 + 3f_1 + 3f_2 + f_3)$	$-\frac{(b-a)^3}{6480} f^{(5)}(\eta)$

Remark A.9.1 (Error analysis). For detailed error analysis, see [10, p. 252].

Remark A.9.2 (how to use). Usually, given the integral $\int_a^b f(x)dx$, we will first divide into smaller intervals and do the numerical integral on each interval and add them up. For example

$$\int_a^b f(x)dx = \int_a^{a+h} f(x)dx + \int_{a+h}^{a+2h} f(x)dx ... + \int_{b-h}^{b} f(x)dx.$$

Lemma A.9.1 (Trapezoid rule and the error bound). *Given the integral $\int_a^b f(x)dx$, we have*

$$\int_a^b f(x)dx \approx \frac{b-a}{n}(\frac{f(a)}{2} + \sum_{k=1}^{n-1}(f(a+k\frac{b-a}{n})) + \frac{f(b)}{2})$$

where we divide $b-a$ into n subintervals, and we do the Trapezoid rule on each subinterval. And the error bound is

$$\frac{(b-a)^3}{12n^2} \max_{x\in[a,b]} f^{(2)}(x)$$

Proof. Note that on each subinterval, the error is $-\frac{(b-a/n)^3}{12} f^{(2)}(\eta)$. Sum up n terms, and we have upper bound

$$\frac{(b-a)^3}{12n^3} n \max_{x\in[a,b]} f^{(2)}(x)$$

□

Lemma A.9.2 (Midpoint rule and the error bound). *Given the integral $\int_a^b f(x)dx$, we have*

$$\int_a^b f(x)dx \approx \frac{b-a}{n}(\sum_{k=1}^{n}(f(a+(k-0.5)\frac{b-a}{n})))$$

where we divide $b-a$ into n subintervals, and we do the Trapezoid rule on each subinterval. And the error bound is

$$\frac{(b-a)^3}{n^2} K$$

A.9.1 Gaussian quadrature

$$\int_a^b w(x)f(x)dx \approx \sum_{i=1}^{n} w_i f(x_i)$$

which is exact when f is a polynomial.

Remark A.9.3. In Newton-Cotes formulas, we fix nodes and try to find suitable weights; in Gaussian quadrature, we use a weighted sum of function values at specified points within the domain of integration.

A.10 Vector calculus

Lemma A.10.1 (divergence theorem).

$$\iiint_V (\nabla \cdot \mathbf{F}) dV = \oiint_{S(V)} \mathbf{F} \cdot \hat{\mathbf{n}} dS$$

$$\iiint_V (\nabla \times \mathbf{F}) dV = \oiint_{S(V)} \hat{\mathbf{n}} \times \mathbf{F} dS$$

$$\iiint_V (\nabla f) dV = \oiint_{S(V)} \hat{\mathbf{n}} f dS$$

Lemma A.10.2 (Lapacian product rule). *Given functions* $u : \mathbb{R}^n \to \mathbb{R}$ *and* $v : \mathbb{R}^n \to \mathbb{R}$, *we have*

$$\nabla^2(uv) = u\nabla v + 2\nabla u \cdot \nabla v + v\nabla^2 u.$$

Proof. Directly use product rule. □

A.11 Numerical linear algebra computation complexity

Note A.11.1. [11, p. 606]

- For a $m \times n$ matrix multiplying a n dimensional vector, mn.
- For a $n \times n$ matrix multiplying a $n \times n$ matrix, n^3(without optimization).
- For a $n \times n$ matrix, LU decomposition $2n^3/3$ (for symmetric matrix $n^3/3$).
- For a $m \times n$ matrix, Cholesky decomposition $4m^2n/3$ (for square matrix $4n^3/3$).
- For a $m \times n$ matrix, QR decomposition $4m^2n/3$ (for square matrix $4n^3/3$).

Note A.11.2 (solving triangular linear system). Let L be a $n \times n$ lower triangle matrix, the forward substitution algorithm for solving

$$Ly = d,$$

is given by

```
y(1) = d(1) / L(1,1);
for i=2:n
y(i) = (d(i) - L(i,1:i-1)* y(1:i-1))/L(i,i)
end
```

This algorithm has complexity of $O(n^2)$.

Let U be a $n \times n$ upper triangle matrix, the backward substitution algorithm for solving

$$Ux = d,$$

is given by

```
x(n) = d(n)/U(n,n);
for i = n - 1: -1 :1
x(i) = (d(i) - U(i,i + 1:n)*x(i + 1:n) )/U(i,i)
end
```

This algorithm has complexity of $O(n^2)$.

A.12 Distributions

Lemma A.12.1. *[12, p. 579] Let K be an externally given parameter. We have*

- $$\int_{-\infty}^{\infty} \delta(x)dx = 1, x\delta(x) = 0, \int_{-\infty}^{\infty} f(x)\delta(x-K)dx = f(K).$$
- $\delta(g(x)) = \sum_i \frac{\delta(x-x_i)}{|g'(x_i)|}$, *where* $x_i, i = 1, 2, ...$ *are the zeros of the function* $g(x)$.
- $\delta(\lambda x) = \frac{\delta(x)}{|\lambda|}, \delta(x-K) = \delta(K-x)$.
- *(step function definition)*

$$H(x) \triangleq \frac{d}{dx}\max\{x,0\}, H(x-A) \triangleq \frac{d}{dx}\max\{x-A,0\}$$

- $H(x-K)+H(K-x)=1$.
- $$\frac{dH(x-K)}{dx}=\delta(x-K), \frac{dH(K-x)}{dx}=-\delta(x-K).$$

Proof. Use $H(x-K)+H(K-x)=1$ to prove $\frac{dH(K-x)}{dx}$. □

A.13 Common integrals

Lemma A.13.1.

- $$\int_0^\infty e^{-ax^2}dx=\frac{1}{2}\sqrt{\frac{\pi}{a}}, \int_{-\infty}^\infty e^{-ax^2}dx=\sqrt{\frac{\pi}{a}}$$
- $$\int_0^\infty xe^{-ax^2}dx=\frac{1}{2a}, \int_{-\infty}^\infty xe^{-ax^2}dx=0$$
- $$\int_0^\infty x^2e^{-ax^2}dx=\frac{1}{4a}\sqrt{\frac{\pi}{a}}$$
- $$\int_0^\infty x^n e^{-ax}dx=\frac{n!}{a^{n+1}}$$
- $$\int_0^\infty x^m e^{-ax^2}=\frac{\Gamma((m+1)/2)}{2a^{(m+1)/2}}$$

A.14 Nonlinear root finding

A.14.1 Bisection method

Methodology A.14.1.

- ***(Goal)****: We want to find a $x^*\in\mathbb{R}$ such that $f(x^*)=0$.*
- ***(Initial input)****: Initial guess of l_0 and r_0 such that*

$$f(l_0)<0, f(r_0)>0;\ or\ f(l_0)>0, f(r_0)<0.$$

- ***Repeat (i) is the iteration index****:*

- *Let* $m = \frac{r_i + l_i}{2}$.
- *If* $f(l_i)f(m) < 0$, *then* $l_{i+1} = l_i, r_{i+1} = m$.
- *If* $f(l_i)f(m) > 0$*(then we must have* $f(r_i)f(m) < 0$*), then* $l_{i+1} = m, r_{i+1} = r_i$.
- *If* $f(l_i)f(m) = 0$, *then* m *is the root.*

A.14.2 Newton method

Methodology A.14.2.

- ***(Goal)****: We want to find a* $x^* \in \mathbb{R}$ *such that* $f(x^*) = 0$. f *is assumed to be differentiable.*
- ***(Initial input)****: Initial guess of* x_0.
- ***Repeat****(i is the iteration index):*
  - *Let* $x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$.
  - *If* $f(x_{i+1}) = 0$, *then* x_{i+1} *is the root.*

A.14.3 Secant method

Methodology A.14.3 (Secant method).

- ***(Goal)****: We want to find a* $x^* \in \mathbb{R}$ *such that* $f(x^*) = 0$. f *is assumed to be differentiable.*
- ***(Initial input)****: Initial guess of* x_0, x_1.
- ***Repeat****(i is the iteration index):*
  - *Let*
  $$x_{i+1} = x_i - f(x_i)\frac{x_i - x_{i-1}}{f(x_i) - f(x_{i-1})}$$
  .
  - *If* $f(x_{i+1}) = 0$, *then* x_{i+1} *is the root.*

Remark A.14.1 (derivation). Starting with initial guesses x_0, x_1, we construct a first order approximation of $f(x)$ via

$$y = \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_1) + f(x_1).$$

And we solve the root for the first-order approximation problem via

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_1) + f(x_1) = 0 \implies x_2 = x_1 - f(x_1)\frac{x_1 - x_0}{f(x_1) - f(x_0)}.$$

Then we continue the process with x_1, x_2.

Remark A.14.2 (convergence property).

- There is no guarantee on the global convergence to the root of f.
- Only when the initial values x_0 and x_1 are sufficiently close to the root, the iterates x_n will converge to the root.

A.15 Interpolation

A.15.1 cubic interpolation

Definition A.15.1 (the cubic spine line functional form). *[13]*

- *Suppose $x_1, x_2, ..., x_n$ and $y_1, y_2, ..., y_n$ are known.*
- *A cubic spine line is given by*

$$y(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3, x_i \leq x \leq x_{i+1}, i = 1, 2, ..., n - 1.$$

- *There are $4n - 4$ unknowns.*
- *Note that*

$$\begin{aligned} y'(x) &= b_i + 2c_i\,(x - x_i) + 3d_i\,(x - x_i)^2\,, x_i < x < x_{i+1} \\ y''(x) &= 2c_i + 6d_i\,(x - x_i)\,, x_i < x < x_{i+1} \\ y'''(x) &= 6d_{i,}x_i < x < x_{i+1} \end{aligned}$$

Definition A.15.2 (natural cubic spline condition). *[13]* *Let $h_i = x_{i+1} - x_i$*

- ***(spline line passing data points)**: for $i = 1, 2, ..., n - 1$, $a_i = y_i$; $a_{n-1} + b_{n-1}h_{n-1} + c_{n-1}h_{n-1}^2 + d_{n-1}h_{n-1}^3 = y_n$.*
- ***(interpolating function is continuous)**; that is,*

$$\lim_{x \to x_i -} y(x) = \lim_{x \to x_i +} y(x) \implies a_i + b_i h_i + c_i h_i^2 + d_i h_i^3 = a_{i+1}, \forall i = 1, 2, ..., n - 2.$$

- ***(interpolating function is differentiable)**; note that the interpolating function is differentiable on interval, therefore we require that,*

$$\lim_{x \to x_i -} y'(x) = \lim_{x \to x_i +} y'(x) \implies b_i + 2c_i h_i + 3d_i h_i^2 = b_{i+1}, \forall i = 1, 2, ..., n - 2$$

:

- ***(interpolating function is twice differentiable and the second derivative at each endpoint is 0)****; that is,*

  $$\lim_{x \to x_i -} y''(x) = \lim_{x \to x_i +} y''(x) \implies c_i + 3d_i h_i = c_{i+1}, \forall i = 1, 2, ..., n-2,$$

  and $y''(x_1+) = y''(x_n-) = 0$.
- *these* $4n-4$ *equations will solve the* $4n-4$ *unknowns.*

A.16 Statistical learning

A.16.1 Algorithm computational time complexity

Algorithm	Training	Testing	Comments
Decision Tree	$O(N^2 d)$	O(d)	For regression and classification
Random Forest	$O(N^2 d_R n_T)$	$O(dn_T)$	
Extremly Random Trees	Ndn_T	$O(dn_T)$	
Gradient Boosting Trees			
Linear Regression	$O(d^2 N + d^3)$	$O(d)$	$d^2 N$ for $X^T X$ and d^3 for $(X^T X)^{-1}$
SVM (Kernel)	$O(N^2 d + N^3)$	$(N^2 d + N^3)$	$N^2 d$ for XX^T and N^3 for $(XX^T)^{-1}$
k-Nearest Neighbours	0	$O(Nd)$	
Naive Bayes	$O(np)$	$O(p)$	

Table A.16.1: Computational time complexity for different machine learning algorithms.

BIBLIOGRAPHY

1. Hammack, R. H. *Book of Proof* (Richard Hammack, 2013).
2. Gradshteyn, I. S. & Ryzhik, I. M. *Table of integrals, series, and products* (Academic press, 2014).
3. Johnsonbaugh, R. & Pfaffenberger, W. Foundations of mathematical analysis (2010).
4. Moon, T. K. S. & Wynn, C. *Mathematical methods and algorithms for signal processing* **621.39: 51 MON** (2000).
5. Luenberger, D. G. *Optimization by vector space methods* (John Wiley & Sons, 1969).
6. Wikipedia. *Gronwall's inequality — Wikipedia, The Free Encyclopedia* [Online; accessed 10-July-2016]. 2016. https://en.wikipedia.org/w/index.php?title=Gr%C3%B6nwall%27s_inequality&oldid=717747531.
7. Wikipedia. *Lp space — Wikipedia, The Free Encyclopedia* [Online; accessed 15-August-2016]. 2016. https://en.wikipedia.org/w/index.php?title=Lp_space&oldid=733853566.
8. Petersen, K. & Pedersen, M. The matrix cookbook, nov 2012. *URL http://www2. imm. dtu. dk/pubdb/p. php* **3274.**
9. Murphy, K. P. *Machine learning: a probabilistic perspective* (MIT press, 2012).
10. Atkinson, K. E. *An introduction to numerical analysis* (John Wiley & Sons, 1989).
11. Nocedal, J. & Wright, S. *Numerical optimization* (Springer Science & Business Media, 2006).
12. Prosperetti, A. Advanced mathematics for applications. *SIAM Review Vol. 55, Issue 2 (June 2013)* **55,** 403 (2013).
13. Hagan, P. S. & West, G. Interpolation methods for curve construction. *Applied Mathematical Finance* **13,** 89–129 (2006).

ALPHABETICAL INDEX