2021-01-28

Optimal Portfolios of Natural Gas Futures

Li, Xiang

UNIVERSITY OF CALGARY

Optimal Portfolios of Natural Gas Futures

by

Xiang Li

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

JANUARY, 2021

Abstract

In this thesis we investigate portfolios consisting of a collection of positions in natural gas futures. The logarithms of the futures prices follow correlated Ornstein-Uhlenbeck processes which are mean-reverting. Under the assumption that the portfolios are constant and short-selling is allowed, formulae for the portfolio that is optimal in terms of minimizing Capital at Risk (CaR) in a continuous-time context are obtained following an adapted version of the theory in [14]. We show that the results in this paper hold with small modifications in our situation and we apply them to portfolios of positions in AB NIT natural gas futures contracts. Furthermore, we allow the portfolio to be reoptimized periodically, showing dramatically better performance of portfolio with regard to the final wealth. And finally we consider the contract size limitation so that the portfolio is tradeable, and demonstrate the unsurprising underperformance of the portfolio that results.

Acknowledgements

I wish to express sincere gratitude to my supervisor, Dr. Antony Ware, for his generous guidance, for his support and patience. He is always there whenever I need him over the past more than three years. Without his persistent help, this thesis would not have been realized.

Special thanks to my examination committee members: Dr. Anatoliy Swishchuk and Dr. Jinniao Qiu, for their time carefully review my thesis and helpful feedback to make this thesis better.

The Department of Mathematics and Statistics is also truly appreciated not only for their support and funding to my study, but also for providing a great learning and working environment.

Finally, I would like to pay my regards to all the members in Dr. Ware's research group for their brilliant ideas and suggestions. And many thanks to my friends and family for their continuous encouragement throughout years of my study.

Table of Contents

List of Tables

List of Figures and Illustrations

Chapter 1

Introduction

Mean reversion is a concept that has been widely used in commodity markets. It assumes that asset prices and historical returns that stray far from the long term norm will eventually tend to move back to the long-run mean or average level over time. As one of the most powerful and consistent market forces, mean reversion can provide us information on long-run future performance and, therefore, it is regarded as an important indicator of prediction. In this paper we investigate portfolios consisting of commodity futures contracts, taking mean reversion into account. More specifically, the model that we use to capture the stochastic behavior of the futures prices is one-factor Schwartz model introduced in [1], in which the logarithm of the futures price is assumed to follow a mean-reverting Ornstein-Uhlenbeck (OU) process.

In fact, the one-factor Schwartz model belongs to a broader class processes that follow affine distributions which are often used to model the dynamics of commodity spot and futures markets. Schwartz [1] compared three mean reversion models in terms of the ability for commodity futures contracts valuation: a simple one-factor model, a two-factor model that take into account the instantaneous convenience yield and a three-factor model including stochastic interest rates. Gibson and Schwartz [2] empirically tested the two-factor model

using weekly oil futures contract prices. Ribeiro and Hodges [3] developed a reduced form two-factor model for pricing commodity spots and futures by adding two new features to Schwartz's two-factor model. Hikspoors and Jaimungal [4] used an affine two-factor mean-reverting process to model the energy spot prices and further obtained the closed form for the forwards and the spread options valuation. Furthermore, Welemical et al. [5] considered a storable commodity follows Schwartz's one-factor model and proved the futures prices process allows asymptotic exponential arbitrage with geometric decaying failure probability.

The problem of portfolio optimization with mean-reverting assets has been well explored in a number of papers. Wachter [6] produced an exact solution of the optimal portfolio selection problem for the consumption utility under mean-reverting returns by assuming markets are complete. Benth and Karlsen [7] gave a note on Merton's portfolio allocation problem where the risky asset follows the exponential OU process and provided an explicit solution for this optimization problem. d'Aspremont [8] studied the problem of constructing small portfolios with maximum mean reversion from multivariate data sets while constraining the number of assets. Fouque and Hu [9] analyzed the nonlinear optimal portfolio investment problem by modeling the assets return and volatility using a fractional OU process which is fast mean-reverting.

In most of the papers mentioned above, the authors investigated the portfolio selection problem where the risky asset satisfied the price dynamics of Schwartz, and the risk is measured in the case with some utility function. In contrast to these cited papers, we capture the risk by a risk reward approach - Capital at Risk (CaR). CaR, introduced in [10], is defined as the difference between the riskless profit and a low quantile (typically the 5%- or 1%-quantile) of the wealth distribution of a portfolio. And the authors obtained a closed form explicit solution for the optimal portfolio with maximum expected terminal wealth subject to an upper bound for the CaR in the Black–Scholes setting. Emmer and Klüppelberg

[11] also investigated the portfolio allocation problem under CaR, but in the setting that the prices follow exponential Lévy processes, and derived a numerical solution. The results in [10] have been generalized to the case of time dependent parameters and investment strategies, and allow the no-short-selling constraint in [12]. In [13], the authors addressed the portfolio optimization problem in a Black-Scholes continuous-time setting under various quantile based risk measures: Value at Risk(VaR), Capital at Risk and relative Value at Risk. And a counterintuitive behaviour has been shown in VaR constrained optimized portfolios with an increase in time horizon. Dmitrašinović-Vidović and Ware [14] derived expressions for the expected wealth of portfolios consists of instruments whose prices are log normally-distributed and follow correlated one-factor mean reverting diffusion processes, and gave the optimal solutions of portfolio that minimized the global CaR and the portfolio that maximized wealth with constrained CaR. The work Pourbabaee et al.[15] looked at the problem of capital allocation that minimized capital at risk with correlation constraint.

This thesis shows that the results developed in [14] can be adapted to a setting where the portfolio consists of positions in futures contracts, and where changes to the wealth occur through the mark-to-market mechanism. We then apply these results to Alberta natural gas futures contracts. The rest of the thesis is organized as follows.

In Chapter 2, we give a brief introduction to the natural gas market, then describe the most important properties of natural gas prices: mean reversion and seasonality. In addition, we explain the natural gas futures trading basics, especially for AB NIT(Alberta Nova Inventory Transfer) futures.

In Chapter 3, we first review some fundamental knowledge of stochastic process and stochastic calculus that are considered as necessary background to understand our model, such as multidimensional Brownian motion, mean-reverting processes and the Itô-Doeblin formula.

Next, we give the market settings, the definition of the portfolio process and derive the form of the expected wealth process, and then the dynamics of the asset price logarithms.

In Chapter 4, we investigate the log-wealth process of the portfolio and further determine the CaR of the wealth process. Then according to Theorem 4.1 in [14] where the quasiconvexity of CaR is proved, the uniqueness of the optimal solution is guaranteed. Although differing in some of the coefficients in our wealth process, the above theorem still holds. Thus, following Theorem 4.2 in the same paper, we solve the optimal strategy that minimizes the global CaR.

In Chapter 5, we mainly illustrate the empirical results. We first introduce some numerical techniques such as OU process discretization and maximum likelihood method. Then the numerical experiments are presorted, starting with constructing virtual futures prices. Model calibration is followed by futures prices simulation. Finally, we solve the optimization problem using the results in Chapter 4 and get the simulated portfolio wealth. In the next section, we discuss the case that allows the portfolio to be reoptimized periodically, and then the case that takes into account the contract size and makes the portfolio to be tradeable, and compare the performance. Finally, we adapt the model to crude oil futures.

In Chapter 6, we summarize the thesis and discuss the potential research directions. The empirical results of optimal portfolio look successful, and the rebalanced portfolio dramatically enhances the final wealth. However, the underperformance of the tradeable portfolio is not surprising. For the future work, we mainly discuss the possibility of adding assets to the portfolio, especially making hedging portfolios. And we can also consider other risk measures or extend the pricing model.

Chapter 2

Natural Gas Market and Natural Gas Futures

In this chapter, we give a brief introduction of natural gas and natural gas market, then describe the main characteristics of natural gas prices - mean reversion and seasonality. Moreover, we provide details on how AB NIT natural gas futures contracts are traded in real word, such as regulations related to margin, contract size, contract period and delivery time.

2.1 Natural Gas Market

Natural gas is the third most important source of energy after oil and coal. The world's largest producers of natural gas are the United States, Russia, Iran, Qatar, Canada and China. And the three most developed demand centers are Western Europe, North America and North Asia. Thanks to the development of liquefied natural gas (LNG), gas markets around the world no longer operate independently based on local pipeline networks; now all the major production centers are linked to the major demand centers. As a result, the natural gas market is truly evolving into a global gas market.

Natural gas is marketed on a number of trading platforms that are located at 'market hubs' across North America. The prices are mainly driven by market demand and supply. While Henry Hub located in southern Louisiana has served as the benchmark price for the entire North American natural gas market, other locations such as Alberta, Chicago and Ontario have a large number of transactions, and are also important market trading points.

There are two distinct markets for natural gas: the spot market and the futures market. Prices are established via various trading platforms, such as the New York Mercantile Exchange (NYMEX) in the U.S. or the Natural Gas Exchange (NGX) in Alberta. Besides, derivative contracts such as forwards and swaps are also traded in over-the-counter (OTC) markets. The bilateral trading that occurs in OTC markets generally consists of non-standardized contracts that can be tailored to the specific needs of the counter-parties. However, OTC markets are largely unregulated, which make the credit risk, especially the default probability, much higher than that in the exchanges.

2.2 Natural Gas Futures

The futures market of natural gas involves buying and selling natural gas delivered in the future with a predetermined price. In the thesis, we mainly use Alberta's natural gas futures (AB NIT), here we provide more details on how they are traded and the characteristics for prices.

AB NIT Basis Futures is a kind of monthly cash settled exchange futures contract based on the result of subtracting the price of the NYMEX Henry Hub Natural Gas Futures Contract, from NGX AB-NIT Month Ahead Index (7A). AB NIT Futures are traded on NGX and prices are quoted in US dollars and cents per million British thermal unit (MMBtu). The initial margin for trading is returnable and based on a clearing member's open positions.

Additional margin may also be required to provide when the price changes significantly. The contract size is 2500 MMBtus and the minimum upward or downward movement of price, i.e. minimum tick, is 0.0001 dollar per MMBtu. The contracts are monthly listed for the current year and up to 120 consecutive monthly contract periods. The trading for each monthly contract ends one business day prior to the first calendar day of the contract period. Once the final settlement price is determined, delivery is the next step. In contrast to crude oil delivery which is on a certain single day, natural gas is delivered in a specified period of time, usually a month.

Figure 2.1 illustrates the AB NIT Futures curve on 01-Jan-2018. The market data were provided by NRGSTREAM (www.nrgstream.com). From the graph we can clearly see that

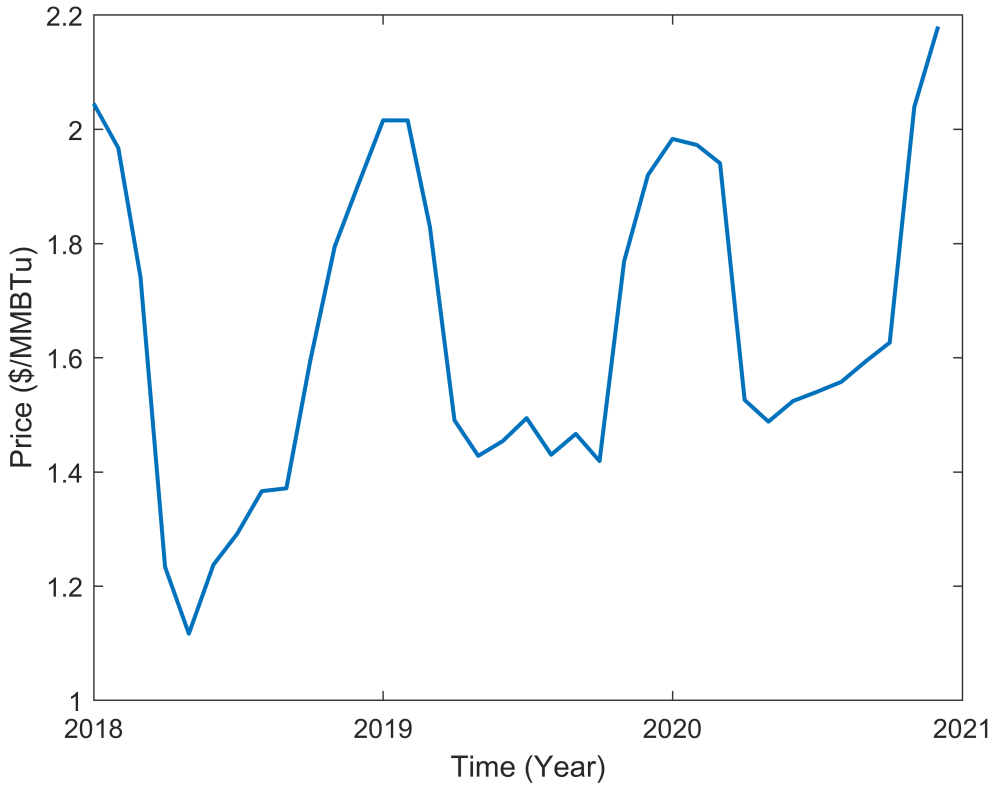

Figure 2.1: AB NIT Futures Curve on January 1st, 2018.

there are two important characteristics for natural gas prices. One is mean reversion: the fluctuations of the prices are always around an average price which is the long term mean. And the meaning-reverting rate depends on how quickly the supply and demand get back to a equilibrium level. When demand for gas is increasing, the prices will rise very soon, then will result in a tendency for the producers to increase their production in order to earn more profit. Thereafter, excessive production will lead to oversupply and push down the prices. The converse is true when the demand is decreasing.

Figure 2.2: AB NIT futures prices of contract which delivery starts on February 1st, 2021.

Another important property of natural gas prices is seasonality. Since natural gas is a widely used fuel for domestic heating and electricity generating, the demand is highly influenced by the season. In winter, heating is crucial in houses and offices, and the day time is much

shorter than summer leads to a higher demand in electricity for lights. However, the supply is not able to perfectly match the sudden change of the demand due to the restriction of transmission and storage capacity, which will result in higher gas prices in winter. That's how the seasonal characteristic of the natural gas prices comes. Compared with the spot prices, seasonality is much easier to be seen in the futures prices. As we can see from the above curve, prices are relatively high in winter and relatively low in summer since the demand is higher in the cold winter and lower in the hot summer.

Figure 2.2 shows the prices of the AB NIT futures for which delivery starts on 01-Feb-2021. It gives us an idea of the dynamics that the futures prices may follow. From the graph we can see that the futures prices always come with some random movement, which makes stochastic process crucial in modelling prices. In the next chapter, we will introduce some background of stochastic process and stochastic calculus that are related to futures prices modeling.

Chapter 3

Preliminaries

This chapter we review some basic knowledge of stochastic process and stochastic calculus, and give the market settings of our model. First, we recall the definition of Brownian motion and multidimensional Brownian Motion, important properties of the Itô's integral and the basic expression of the Itô's formula. Then we introduce the mean-reverting models that we will use: Ornstein-Uhlenbeck process and one-factor Schwartz model. In the next section, we give the model assumptions, the definition of the portfolio process and the dynamics of the wealth process. And last, in order to facilitate further investigation, we introduced the log-price which is normally-distributed.

3.1 Stochastic Process and Basic Stochastic Calculus

3.1.1 Brownian Motion

Brownian motion is one of the most classic and commonly used stochastic process in finance. The following definitions are mainly taken from [17].

Definition 3.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For each $\omega \in \Omega$, suppose there is a continuous function $W(t)$ that satisfies $W(0)$, $t \geq 0$, then $W(t)$ is a Brownian motion if for*

all $t_0 < t_1 < ... < t_m$, $t_0 = 0$, *the increments*

$$W(t_1) - W(t_0), W(t_2) - W(t_1), ..., W(t_m) - W(t_{m-1})$$

are independent and follow normal distribution with

$$\mathbb{E}[W(t_{i+1}) - W(t_i)] = 0,$$

$$\text{Var}[W(t_{i+1}) - W(t_i)] = t_{i+1} - t_i.$$

Definition 3.2. $W(t) = (W_1(t), ..., W_n(t))$ *is an* n*-dimensional Brownian motion if satisfies the following properties.*

(i) Each $W_i(t)$, $i = 1, 2, ..., n$ *is a one- dimensional Brownian motion.*

(ii) The processes $W_i(t)$ *and* $W_j(t)$ *are independent,* $i \neq j$.

Correlated Brownian Motions

Define $\boldsymbol{W}(t) := (W_1(t), ..., W_n(t))$ as an n-dimensional Brownian motion and $\boldsymbol{B}(t) := (B_1(t), ..., B_n(t))$ is an n-dimensional vector of correlated Brownian motions $B_1(t), ..., B_n(t)$ with correlation matrix $\boldsymbol{\rho}$. If $\boldsymbol{\rho}$ has a Cholesky decomposition $\boldsymbol{\rho}_1 \boldsymbol{\rho}_1{}'$(see Appendix B.2 for introduction), then

$$\boldsymbol{B}(t) = \boldsymbol{\rho}_1 \boldsymbol{W}(t).$$

In other words, correlated Brownian motions can be written as linear combinations of independent Brownian motions.

3.1.2 Itô's Integral and Itô-Doeblin Formula

Itô's Integral

If a random variable $\mu(X_t, t)$ satisfies the condition $\mathbb{E}[\int_0^T \mu(X_t,t)^2 dt] < \infty$, then the Itô's integral

$$\int_0^T \mu(X_t, t) dW_t$$

can be defined as

$$\lim_{\delta \to 0} \sum_{k=0}^{n-1} \mu\left(X_{\tau_k}, \tau_k\right)\left(W_{t_{k+1}} - W_{t_k}\right),$$

where $0 < t_0 < t_1 < t_2 < ... < t_n$, $t_n = T$, $\tau \in [t_k, t_{k+1}]$ and $\delta = \max_{k=1,2,...,n-1}(t_{k+1} - t_k)$.

The following theorems give the important properties of Itô's integral and the Itô-Doeblin formula for Brownian motion, which can be found in [17].

Theorem 3.1. *Let T be a positive constant and let $\Delta(t)$, $0 \le t \le T$, be an adapted stochastic process that satisfies $\mathbb{E}\int_0^T \Delta^2(t) < \infty$. Then $I(t) = \int_0^t \Delta(u) dW_u$ defined by $\int_0^t \Delta(u) dW_t = \lim_{n\to\infty} \Delta_n(u) dW_u$, $0 \le t \le T$ has the following properties.*

*(i) (**Continuity**) As a function of the upper limit of integration t, the paths of $I(t)$ are continuous.*

*(ii) (**Adaptivity**) For each t, $I(t)$ is $\mathcal{F}$-measurable.*

*(iii) (**Linearity**) If $I(t) = \int_0^t \Delta(u) dW_u$ and $J(t) = \int_0^t \Gamma(u) dW_u$, then $I(t) \pm J(t) = \int_0^t (\Delta(u) \pm \Gamma(u)) dW_u$; furthermore, for every constant c, $cI(t) = \int_0^t c\Delta(u) dW_u$.*

*(iv) (**Martingale**) $I(t)$ is a martingale.*

*(v) (**Itô's Isometry**) $\mathbb{E}\left(I^2(t)\right) = \mathbb{E}\int_0^t \Delta^2(u) du$.*

*(vi) (**Quadratic variation**) $[I, I](t) = \int_0^t \Delta^2(u) du$*

Theorem 3.2. *Let $f(t,x)$ be a function for which the partial derivatives $f_t(t,x)$, $f_x(t,x)$ and $f_{xx}(t,x)$ are defined and continuous, and let W_t be a Brownian motion. Then, f for every*

$T \geq 0$,

$$f(T, W_T) = f(0, W_0) + \int_0^T f_t(t, W_t)dt + \int_0^T f_x(t, W_t)dW_t + \frac{1}{2}\int_0^T f_{xx}(t, W_t)dt.$$

The equation above is the Itô-Doeblin formula for Brownian motion. The Itô-Doeblin formula is also referred as the Itô's lemma, which is one of the most important formulae in Itô's calculus. It can be deduced by the Taylor series expansion and is used to find the derivatives of a function of a stochastic process.

3.1.3 Mean Reverting Process

Although Brownian motion is important in financial modeling, it is not enough when capturing the characteristic of asset prices such as mean reversion and spikes in energy market. We hereby introduce the Ornstein-Uhlenbeck process and the one-factor Schwartz model (see [1]).

Ornstein-Uhlenbeck process

The Ornstein-Uhlenbeck(OU) process X_t as one of the most popular mean-reverting processes, is defined by the following stochastic differential equation(SDE):

$$dX_t = \lambda(\mu - X_t)dt + \sigma dW_t,$$

where W_t is the standard Brownian motion, σ represents the volatility, μ is the long term mean and the mean-reversion rate $\lambda > 0$ measures the speed at which X_t converges to the long term mean μ.

The mean reversion property can be well captured by the Ornstein-Uhlenbeck process. Assume the asset price X_t satisfies the model above. If the asset price is below the long run mean at time $t > 0$ i.e. $X_t < \mu$, the drift term $\lambda(\mu - X_t)$ is positive, and X_t gets pushed upwards

towards the long run mean μ; if X_t is above the long run mean μ at time t, i.e.$X_t > \mu$, then the negative drift term $\lambda(\mu - X_t)$ generates a tendency for X_t to move downwards till $X_t = \mu$. The speed that X_t converge to μ depends on the mean-reversion rate λ. The closer λ is to 0, the weaker the power of mean reversion and the more time for X_t to revert to μ, vice versa.

It is not hard to show that $X_T|X_t$ is normally distributed, with

$$\mathbb{E}[X_T|X_t] = \mu(1 - e^{-\lambda(T-t)}) + X_t e^{-\lambda(T-t)},$$
$$\mathrm{Var}[X_T|X_t] = \frac{\sigma^2}{2\lambda}(1 - e^{-2\lambda(T-t)}).$$

We noticed that as $T \to \infty$, $\mathbb{E}[X_T] \to \mu$, which is the long run mean, and $\mathrm{Var}[X_T] \to \frac{\sigma^2}{2\lambda}$.

One-factor Schwartz model

To introduce the one-factor Schwartz model we first assume that the asset price S_t satisfies the following stochastic process:

$$dS_t = \lambda(\alpha - \ln S_t)S_t dt + \sigma S_t dW_t.$$

Defining $X_t = \ln S_t$ and applying Itô's formula, the log price can be characterized by an OU process:

$$dX_t = \lambda(\mu - X_t)dt + \sigma dW_t,$$

where

$$\mu = \alpha - \frac{\sigma^2}{2\lambda}.$$

The one-factor Schwartz model is also referred as the exponential OU process. In this model $\lambda > 0$ measures the degree of mean reversion to the long run mean log price μ, and the second term in the equation captures the volatility of the process, with dW_t being an increment to a standard Brownian motion.

3.2 Market Setting of the Model

In this section we give the market's setting, define the portfolio and the wealth process, and explore the dynamics of the futures prices. Throughout the thesis we denote vectors and matrices by bold letters, their transposes by $(\cdot)'$, and the Euclidean norm of a matrix or vector by $\|\cdot\|$.

3.2.1 Model Assumptions

Assumptions provide a way to simplify processes and make things easier to study and understand. Checking model assumptions is essential prior to building a model. If assumptions are not met, the model may fail to reflect the data and will likely result in inaccuracy. In this thesis, we assume that the following conditions are satisfied:

Assumption 3.1. *(i) The futures contracts are perfectly divisible.*
(ii) Negative positions in contracts are possible.
(iii) Rebalancing of the holdings does not lead to transaction costs.

In addition, we have the following assumptions.

Assumption 3.2. *(i) All contracts are traded continuously over a finite time horizon* $[0, T]$.
(ii) All futures prices follow the one-factor Schwartz model with correlated Brownian Motions, which can be written in the form:

$$\frac{dF_i(t)}{F_i(t)} = \beta_i \left(L_i - \ln F_i(t)\right) dt + \sum_j \sigma_{ij} dW_j(t), F_i(0) > 0, i = 1, ..., m. \tag{3.1}$$

In this model the mean-reversion rate $\beta_i > 0$ *and* $L_i \in \mathbb{R}$, $i = 1, ..., m$. $\boldsymbol{W}(t) := (W_1(t), ..., W_m(t))'$ *is an m-dimensional Brownian motion, and* $\boldsymbol{\sigma} := (\sigma_{ij})$ *is the volatility matrix.*

(iii) The risk-free instantaneous interest rate r is a positive constant, and that funds in the margin account attract interest at this rate, continuously compounded.
(iv) $\boldsymbol{\sigma}$ is invertible.

If we define

$$a_i = \beta_i L_i,\ b_i = \beta_i,\ \boldsymbol{\sigma_i} = (\sigma_{i1}, ..., \sigma_{im}),\ i = 1, ..., m, \tag{3.2}$$

the system above can be written as

$$\frac{dF_i(t)}{F_i(t)} = (a_i - b_i \ln F_i(t))\, dt + \boldsymbol{\sigma_i} d\boldsymbol{W}(t), i = 1, ..., m. \tag{3.3}$$

Let $N_i(t)$ denote the size of the position at time t in the ith contract in a portfolio, and $X^{\boldsymbol{\pi}}(t)$ denote the wealth at time t in the margin account for this portfolio. The m-dimensional vector-valued function $\boldsymbol{N}(t) = (N_1(t), ..., N_m(t))'$ is called the trading strategy. $[t, t + \Delta t)$ is the mark-to-market period(e.g. trading days), $0 \leq t < T$. We assume that $\boldsymbol{N}(t)$ is set at the start of the period and not changed during it. Then if the change in the ith futures price is ΔF_i during this period, there is a mark-to-market payment of $\sum_i N_i(t)\Delta F_i$. Here we consider the continuous-time analogue of the mark-to-market mechanism, which emerges as we let $\Delta t \to 0$. Then each position in the portfolio is marked-to-market continuously, and we can write for the wealth held in the margin account

$$dX^{\boldsymbol{\pi}}(t) = rX^{\boldsymbol{\pi}}(t)dt + \sum_{i=1}^{m} N_i(t) dF_i(t).$$

Now consider the notional size of each position as a fraction of the wealth. We denote

$$\pi_i(t) = \frac{N_i(t)F_i(t)}{X^{\boldsymbol{\pi}}(t)}, \in \mathbb{R},\ i = 1, ..., m. \tag{3.4}$$

Assumption 3.3. *In order to maintain tractability, we follow the approach of papers such as [10] and [14], and consider constant portfolios, in the sense that $\pi_i(t) \equiv \pi_i$ is kept constant for each i and we call $\boldsymbol{\pi} := (\pi_1, ..., \pi_m)' \in \mathbb{R}^m$ the portfolio.*

Under this assumption, the size of each position, given by $N_i(t)$, is changing at each moment in time, but the positions are changing in such a way as to hold the ratio $\pi_i(t)$ constant. Under the assumption that the trading strategy is self-financing, the dynamics of $X^{\boldsymbol{\pi}}(t)$ can be straightforwardly shown to satisfy

$$\begin{aligned} dX^{\boldsymbol{\pi}}(t) &= X^{\boldsymbol{\pi}}(t)rdt + \sum_{i=1}^{m} N_i(t)F_i(t)[(a_i - b_i \ln F_i(t))dt + \boldsymbol{\sigma_i} d\boldsymbol{W}(t)] \\ &= X^{\boldsymbol{\pi}}(t)\left(rdt + \sum_{i=1}^{m} \pi_i[(a_i - b_i \ln F_i(t))dt + \boldsymbol{\sigma_i} d\boldsymbol{W}(t)]\right), \; X^{\boldsymbol{\pi}}(0) = X_0, \end{aligned} \tag{3.5}$$

where X_0 is the initial wealth.

Remark 3.1. *In equation (2.4) in [14], the SDE that $X^{\boldsymbol{\pi}}(t)$ follows is slightly different: $dX^{\boldsymbol{\pi}}(t) = X^{\boldsymbol{\pi}}(t)\left(rdt + \sum_{i=1}^{m} \pi_i[(a_i - r - b_i \ln F_i(t))dt + \boldsymbol{\sigma_i} d\boldsymbol{W}(t)]\right)$. We does not have the term $-\sum_{i=1}^{m} \pi_i r dt$ in our market settings, but this will not make a significant change in the mathematical deduction for the following steps and still allows us to follow the process in that paper. Any other expressions in the rest of the thesis that are different from the expressions in [14] will be pointed out.*

In particular, the wealth $X^{\boldsymbol{\pi}}(t)$ is log-normal, and we can compute its mean and variance, conditional on $F_i(0), i = 1, ..., m$, and depending also on X_0; more details will be provided in the next chapter.

3.2.2 The Log-price Process

For further investigation, we define $Y_i(t) := \ln F_i(t)$, $i = 1, ..., m$. The dynamics of the processes $Y_i(t)$ are given in the following proposition.

Proposition 3.1. *The log-price process $Y_i(t), i = 1, ..., m$ follows the SDE*

$$dY_i(t) = (\hat{a}_i - b_i Y_i(t))dt + \boldsymbol{\sigma}_i d\boldsymbol{W}(t), i = 1, 2, 3 \tag{3.6}$$

The means and covariances of the normally-distributed random variables $Y_i(t), i = 1, ..., m$ *are*

$$\mathbb{E}[Y_i(t)] = Y_i(0)e^{-b_i t} + \hat{a}_i \boldsymbol{\varepsilon}(t, b_i)$$

$$\mathrm{Cov}[Y_i(t), Y_j(t)] = \boldsymbol{\sigma_i \sigma_j'} \boldsymbol{\varepsilon}(t, b_i + b_j),$$

where

$$\hat{a}_i = a_i - \frac{1}{2} \|\boldsymbol{\sigma_i}\|^2 \tag{3.7}$$

and

$$\boldsymbol{\varepsilon}(t, b) := \int_0^t e^{-sb} ds = \begin{cases} \dfrac{1 - e^{-tb}}{b} & \text{if } b \neq 0 \\ t & \text{if } b = 0. \end{cases} \tag{3.8}$$

Remark 3.2. *Itô's Formula and Itô's Isometry are applied here to complete the proof, details can be found in [14] and it also gives an explicit form for the joint distribution of* $Y_1(t), ..., Y_m(t)$*, which is shown in the following proposition.*

Proposition 3.2. *Let the matrix* $\boldsymbol{E}(t, \boldsymbol{b})$ *have the entries*

$$\boldsymbol{E}_{ij}(t, \boldsymbol{b}) := \mathrm{Cov}[Y_i(t), Y_j(t)] = \boldsymbol{\sigma_i \sigma_j'} \boldsymbol{\varepsilon}(t, b_i + b_j), \quad i, j = 1, ..., m. \tag{3.9}$$

Then, $\boldsymbol{E}(t, \boldsymbol{b})$ *is positive-definite for any* $t > 0$*.*

This proposition is important and will be used when showing the quasiconvexity of CaR (see Theorem 4.1 in the next chapter).

Chapter 4

Wealth and Capital at Risk

In this chapter, we use the results of Proposition 3.1 to determine the Capital at Risk of the wealth process $X^{\boldsymbol{\pi}}(t)$, and get the solution to the global minimum CaR based on the quasiconvexity of CaR. First, we recall that CaR is defined as the difference between the riskless return and the α-quantile of the portfolio wealth distribution, and in order to determine the CaR, we explore the dynamics of log-wealth process and calculate the mean and variance. Then the formula for CaR can be obtained from determine the α-quantile of the wealth process. In [14], the authors proved that in their setting CaR is strongly quasiconvex, which means the local minimum of CaR is its global minimum. Although the wealth process in our setting is missing a coefficient compared with the form in [14], the above conclusion still holds. This is an crucial property for portfolio optimization since it makes the optimal solution unique. At last, following the work of the above paper, we give the optimal solution that minimizes the global CaR.

4.1 Capital at Risk

We recall that the Capital at Risk of $X^{\boldsymbol{\pi}}(t)$ was defined in [10] to be

$$X_0 e^{rt} - q_x \tag{4.1}$$

where q_x is the α-quantile of $X^{\boldsymbol{\pi}}(t)$, X_0 denotes the initial wealth of the portfolio and r denotes the riskless rate of return.

It is clear that the Capital at Risk is defined via the Value at Risk (VaR), i.e. a low quantile (typically the 5%- or 1%-quantile) of the profit-loss distribution of a portfolio. And the crucial point in the application of the CaR model is the determination of the VaR, i.e. the α-quantile of the wealth distribution.

4.2 The Log-wealth Process

In order to determine the CaR of the wealth process $X^{\boldsymbol{\pi}}(t)$ we first investigate the log-wealth process $H(t) := \ln X^{\boldsymbol{\pi}}(t)$. The characteristics of the processes $H(t)$ are given by the following proposition.

Proposition 4.1. *The log-wealth process $H(t) := \ln X^{\boldsymbol{\pi}}(t)$ follows the dynamics*

$$dH(t) = \mu dt + \sum_{i=1}^{m} \pi_i \left[-b_i Y_i(t) dt + \boldsymbol{\sigma_i} d\boldsymbol{W}(t) \right], \quad H(0) = \ln X_0,$$

where

$$\mu = r + \boldsymbol{\pi}' \boldsymbol{a} - \frac{1}{2} \left\| \boldsymbol{\pi}' \boldsymbol{\sigma} \right\|^2. \tag{4.2}$$

The mean and variance of $H(t)$ are given by

$$\mathbb{E}[H(t)] = H(0) + \mu t + \boldsymbol{\pi}' \boldsymbol{A}(t, \boldsymbol{b}),$$

$$\mathrm{Var}[H(t)] = \boldsymbol{\pi}' \boldsymbol{E}(t, \boldsymbol{b}) \boldsymbol{\pi},$$

where

$$\boldsymbol{A}_i = \boldsymbol{\varepsilon}(t, b_i)(\hat{a}_i - b_i Y_i(0)) - \hat{a}_i t, \ i = 1, ..., m. \tag{4.3}$$

Proof. Here is the sketched proof.

First we rewrite (3.5) in the following form:

$$dX^{\pi}(t) = X^{\pi}(t)\left(\left(r + \boldsymbol{\pi}'\boldsymbol{a} + \sum_{i=1}^{m} \pi_i(-b_i Y_i(t))\right)dt + \boldsymbol{\pi}'\boldsymbol{\sigma} d\boldsymbol{W}(t)\right).$$

Applying the Itô's Formula to $H(t)$ and using (4.2) we get

$$dH(t) = \mu dt + \sum_{i=1}^{m} \pi_i[-b_i Y_i(t)dt + \boldsymbol{\sigma_i} d\boldsymbol{W}(t)].$$

From (3.6), we have

$$-b_i Y_i(t)dt + \boldsymbol{\sigma_i} d\boldsymbol{W}(t) = dY_i(t) - \hat{a}_i dt,$$

thus

$$\begin{aligned} dH(t) &= \mu dt + \sum_{i=1}^{m} \pi_i[dY_i(t) - \hat{a}_i dt] \\ &= (\mu - \boldsymbol{\pi}'\hat{\boldsymbol{a}})dt + \boldsymbol{\pi}' d\boldsymbol{Y}(t), \end{aligned}$$

so

$$\begin{aligned} H(t) &= H(0) + (\mu - \boldsymbol{\pi}'\hat{\boldsymbol{a}})t + \boldsymbol{\pi}'(\boldsymbol{Y}(t) - \boldsymbol{Y}(0)) \\ &= H(0) + (\mu - \boldsymbol{\pi}'\hat{\boldsymbol{a}})t + \sum_{i=1}^{m} \pi_i(Y_i(t) - Y_i(0)). \end{aligned}$$

Using substitution $Z_i = e^{b_i t} Y_i(t)$ to (3.6), we get

$$dZ_i(t) = e^{b_i t}\hat{a}_i dt + e^{b_i t}\boldsymbol{\sigma_i} d\boldsymbol{W}(t),$$

then

$$Z_i(t) = Z_i(0) + \hat{a}_i \boldsymbol{\varepsilon}(t, -b_i) + \boldsymbol{\sigma_i}\int_0^t e^{b_i u} d\boldsymbol{W}(u),$$

$$Y_i(t) = e^{-b_i t} Y_i(0) + \hat{a}_i \boldsymbol{\varepsilon}(t, b_i) + e^{-b_i t}\boldsymbol{\sigma_i}\int_0^t e^{b_i u} d\boldsymbol{W}(u).$$

So

$$Y_i(t) - Y_i(0) = (\hat{a}_i - b_i Y_i(0))\boldsymbol{\varepsilon}(t, b_i) + e^{-b_i t}\boldsymbol{\sigma_i} \int_0^t e^{b_i u} d\boldsymbol{W}(u),$$

$$H(t) = H(0) + (\mu - \boldsymbol{\pi}'\hat{\boldsymbol{a}})t + \sum_{i=1}^{m} \pi_i(\hat{a}_i - b_i Y_i(0))\boldsymbol{\varepsilon}(t, b_i) + \sum_{i=1}^{m} \pi_i e^{-b_i t}\boldsymbol{\sigma_i} \int_0^t e^{b_i u} d\boldsymbol{W}(u).$$

Thus

$$\mathbb{E}[H(t)] = H(0) + \mu t + \boldsymbol{\pi}'\boldsymbol{A}(t, \boldsymbol{b}),$$

where $\boldsymbol{A}(t, \boldsymbol{b})$ is defined in (4.3). And

$$\begin{aligned}
\mathrm{Var}[H(t)] &= \mathbb{E}[H(t)^2] - \mathbb{E}[H(t)]^2 \\
&= \mathbb{E}\left[\left(\sum_{i=1}^{m} \pi_i e^{-b_i t}\boldsymbol{\sigma_i} \int_0^t e^{b_i u} d\boldsymbol{W}(u)\right)^2\right], \\
&= \boldsymbol{\pi}'\boldsymbol{E}(t, \boldsymbol{b})\boldsymbol{\pi}.
\end{aligned}$$

□

Remark 4.1. *In Proposition 3.1 of [14], the expression for μ is slightly different with an additional term $-\boldsymbol{\pi}' r\mathbf{1}$. Since the log-price $Y_i(t)$ is normally-distributed, it follows that the log-wealth $H(t)$ is also normally-distributed.*

We now turn to investigating the α-quantile of $H(t)$. If a normal random variable $U \sim N(\mu, \sigma^2)$, the α - quantile q_u is

$$q_u = \mu - |z_\alpha|\sigma,$$

where $|z_\alpha|$ denotes the corresponding α-quantile of the standard normal distribution. Thus the α-quantile of $H(t)$ is

$$\begin{aligned}
q_h &= \mathbb{E}[H(t)] - |z_\alpha|\sqrt{\mathrm{Var}[H(t)]} \\
&= H(0) + rt + \boldsymbol{\pi}'[\boldsymbol{a}t + \boldsymbol{A}(t, \boldsymbol{b})] - \frac{t}{2}\left\|\boldsymbol{\pi}'\boldsymbol{\sigma}\right\|^2 - |z_\alpha|\sqrt{\boldsymbol{\pi}'\boldsymbol{E}(t, \boldsymbol{b})\boldsymbol{\pi}}.
\end{aligned}$$

If we define

$$\boldsymbol{g}(t) := \boldsymbol{a}t + \boldsymbol{A}(t, \boldsymbol{b}),$$
$$f(\boldsymbol{\pi}, t) := \boldsymbol{\pi}'\boldsymbol{g}(t) - \frac{t}{2}\left\|\boldsymbol{\pi}'\boldsymbol{\sigma}\right\|^2 - |z_\alpha|\sqrt{\boldsymbol{\pi}'\boldsymbol{E}(t, \boldsymbol{b})\boldsymbol{\pi}},$$

then the α-quantile of the log-wealth process $H(t)$ can be written as

$$q_h = H(0) + rt + f(\boldsymbol{\pi}, t),$$

so that the α-quantile of the wealth process $X^\pi(t)$ is given by

$$q_x = X_0 e^{rt} e^{f(\boldsymbol{\pi}, t)}.$$

According to the definition of CaR, we have the following corollary.

Corollary 4.1. *The capital at risk of the wealth process $X^\pi(t)$ is given by*

$$CaR(\boldsymbol{\pi}, t) = X_0 e^{rt}\left(1 - e^{f(\boldsymbol{\pi}, t)}\right).$$

Remark 4.2. *$f(\boldsymbol{\pi}, t)$ in equation (3.4) of [14] is different through $\boldsymbol{g}(t)$ where $\boldsymbol{g}(t) := (\boldsymbol{a} - r\mathbf{1})t + \boldsymbol{A}(t, \boldsymbol{b})$.*

4.3 Minimal Capital at Risk

In this section we first illustrate that CaR is a strongly quasiconvex function of $\boldsymbol{\pi}$. Then we find the global minimum of CaR, and solve the problem of the maximal portfolio final wealth subject to the minimized CaR.

4.3.1 Quasiconvexity of CaR

We recall that a function $\psi : \mathbb{R}^m \to \mathbb{R}$ is strongly quasiconvex if

$$\varphi(\lambda\boldsymbol{\pi} + (1-\lambda)\boldsymbol{\xi}) < \max\left\{\varphi(\boldsymbol{\pi}), \varphi(\boldsymbol{\xi})\right\},$$

for all $\boldsymbol{\pi}$, $\boldsymbol{\xi} \in \mathbb{R}^m$ for which $\boldsymbol{\pi} \neq \boldsymbol{\xi}$, and for all $\lambda \in (0,1)$.

Theorem 4.1. *The Capital at Risk is a strongly quasiconvex function of the portfolio.*

Proof. Suppose that $X^{\boldsymbol{\pi}}(t)$ and $X^{\boldsymbol{\xi}}(t)$ are two wealth processes defined by portfolios $\boldsymbol{\pi}$, $\boldsymbol{\xi} \in \mathbb{R}^m$, $\boldsymbol{\pi} \neq \boldsymbol{\xi}$, with the same initial wealth $X^{\boldsymbol{\pi}}(0) = X^{\boldsymbol{\xi}}(0) = X_0$, and suppose that

$$\mathrm{CaR}(\boldsymbol{\pi}, T) \geq \mathrm{CaR}(\boldsymbol{\xi}, T). \tag{4.4}$$

This implies that

$$f(\boldsymbol{\pi}, T) \leq f(\boldsymbol{\xi}, T). \tag{4.5}$$

To prove CaR is a strongly quasiconvex function, i.e. for all $\lambda \in (0,1)$,

$$\mathrm{CaR}(\lambda\boldsymbol{\pi} + (1-\lambda)\boldsymbol{\xi}) < \max\left\{\mathrm{CaR}(\boldsymbol{\pi}), \mathrm{CaR}(\boldsymbol{\xi})\right\},$$

under condition (4.4), we only need to prove

$$f(\boldsymbol{\pi} + (1-\lambda)\xi, T) > f(\boldsymbol{\pi}, T). \tag{4.6}$$

Proposition 3.2 shows that $\boldsymbol{E}(t, \boldsymbol{b})$ is positive-definite, it follows that $\boldsymbol{E}(t, \boldsymbol{b})$ has a unique positive-definite square root, and we define

$$\boldsymbol{C}(t, \boldsymbol{b}) := \boldsymbol{E}(t, \boldsymbol{b})^{\frac{1}{2}}. \tag{4.7}$$

Then $f(\boldsymbol{\pi}, T)$ can be writen as

$$\boldsymbol{\pi}'\boldsymbol{g}(T) - \frac{T}{2}\left\|\boldsymbol{\pi}'\boldsymbol{\sigma}\right\|^2 - |z_\alpha|\left\|\boldsymbol{C}(t,\boldsymbol{b})\boldsymbol{\pi}\right\|.$$

Since $\boldsymbol{\sigma}$ is invertible, we have that $f(\boldsymbol{\pi}, T)$ is strictly concave. Under condition (4.5), this implies that

$$f(\boldsymbol{\pi} + (1-\lambda)\xi, T) > \lambda f(\boldsymbol{\pi}, T) + (1-\lambda) f(\boldsymbol{\xi}, T) > f(\boldsymbol{\pi}, T), \forall \lambda \in (0,1).$$

So (4.6) holds, completing the prove. □

This theorem has an immediate and important result: if a function $\psi : U \subset \mathbb{R}^m \to \mathbb{R}$ is strongly quasiconvex, then its local minimum is its unique global minimum(see [29]). Therefore, we have the following corollary.

Corollary 4.2. *If CaR$(\boldsymbol{\pi}, T)$ has a local minimum at $\boldsymbol{\pi}^* \in \boldsymbol{Q}$ then $\boldsymbol{\pi}^*$ is its global minimum.*

Remark 4.3. *Although differing in some of the coefficients, the expressions of our portfolio process still take the same form as those derived in [14], where quasiconvexity of CaR as a function of $\boldsymbol{\pi}$ is established (c.f. Theorem 4.1 and Corollary 4.2 in [14]).*

4.3.2 Global Minimum CaR

We now turn our attention to finding the global minimum of $\mathrm{CaR}(\boldsymbol{\pi}, T)$ at the time horizon T. Note that

$$\begin{aligned} \arg\min_{\boldsymbol{\pi}} \mathrm{CaR}(\boldsymbol{\pi}, T) &= \arg\max_{\boldsymbol{\pi}} f(\boldsymbol{\pi}, T) \\ &= \arg\max_{\boldsymbol{\pi}} \boldsymbol{\pi}'\boldsymbol{g}(T) - \frac{t}{2}\left\|\boldsymbol{\pi}'\boldsymbol{\sigma}\right\|^2 - |z_\alpha|\sqrt{\boldsymbol{\pi}'\boldsymbol{E}\boldsymbol{\pi}}. \end{aligned} \tag{4.8}$$

The solution to the above problem is given by Theorem 4.2 (c.f. Theorem (4.3) in[14]).

Theorem 4.2. *a) If* $\boldsymbol{g}'(T)\boldsymbol{E}^{-1}\boldsymbol{g}(T) > |z_\alpha|^2$*, the optimal solution of (4.8) is equal to*

$$\pi^* = (T\boldsymbol{\sigma}\boldsymbol{\sigma}' + \frac{|z_\alpha|}{\lambda^*}\boldsymbol{E})^{-1}\boldsymbol{g}(T),$$

where λ^* *the unique positive solution of the equation*

$$\left\|\boldsymbol{C}(\lambda T\boldsymbol{\sigma}\boldsymbol{\sigma}' + |z_\alpha|\boldsymbol{E})^{-1}\boldsymbol{g}(T)\right\|^2 = 1,$$

and $\boldsymbol{C}$ *is defined in (4.7).*

b) If $\boldsymbol{g}'(T)\boldsymbol{E}^{-1}\boldsymbol{g}(T) \leq |z_\alpha|^2$*, the optimal solution of (4.8) is* $\boldsymbol{\pi} = 0$*.*

Proof. Here is the sketched proof.

The critical points of $f(\boldsymbol{\pi}, T)$ are $\boldsymbol{\pi} = \mathbf{0}$, and the point at which the derivative is zero, i.e. the solution of the following equation

$$\boldsymbol{g}(T) - T\boldsymbol{\sigma}\boldsymbol{\sigma}'\boldsymbol{\pi} - \frac{|z_\alpha|}{\sqrt{\boldsymbol{\pi}'\boldsymbol{E}\boldsymbol{\pi}}}\boldsymbol{E}\boldsymbol{\pi} = \mathbf{0}.$$

$\Rightarrow$

$$\boldsymbol{\pi} = \left(T\boldsymbol{\sigma}\boldsymbol{\sigma}' + \frac{|z_\alpha|}{\sqrt{\boldsymbol{\pi}'\boldsymbol{E}\boldsymbol{\pi}}}\right)^{-1}\boldsymbol{g}(T).$$

If we define $\lambda := \sqrt{\boldsymbol{\pi}'\boldsymbol{E}\boldsymbol{\pi}}$, we have

$$\begin{aligned}\boldsymbol{\pi}'\boldsymbol{E}\boldsymbol{\pi} = \lambda^2 &= \boldsymbol{g}(T)'\left(T\boldsymbol{\sigma}\boldsymbol{\sigma}' + \frac{|z_\alpha|}{\lambda}\boldsymbol{E}\right)^{-1}\boldsymbol{E}\left(T\boldsymbol{\sigma}\boldsymbol{\sigma}' + \frac{|z_\alpha|}{\lambda}\boldsymbol{E}\right)^{-1}\boldsymbol{g}(T) \\ &= \lambda^2\boldsymbol{g}(T)'(\lambda T\boldsymbol{\sigma}\boldsymbol{\sigma}' + |z_\alpha|\,\boldsymbol{E})^{-1}\boldsymbol{E}(\lambda T\boldsymbol{\sigma}\boldsymbol{\sigma}' + |z_\alpha|\,\boldsymbol{E})^{-1}\boldsymbol{g}(T).\end{aligned}$$

$\Rightarrow$

$$\boldsymbol{g}(T)'(\lambda T\boldsymbol{\sigma}\boldsymbol{\sigma}' + |z_\alpha|\,\boldsymbol{E})^{-1}\boldsymbol{E}(\lambda T\boldsymbol{\sigma}\boldsymbol{\sigma}' + |z_\alpha|\,\boldsymbol{E})^{-1}\boldsymbol{g}(T) = 1. \tag{4.9}$$

Using (4.7), we can write (4.9) in the following way

$$\left\|\boldsymbol{C}(\lambda T\boldsymbol{\sigma}\boldsymbol{\sigma}' + |z_\alpha|\,\boldsymbol{E})^{-1}\boldsymbol{g}(T)\right\|^2 = 1.$$

Define

$$\omega(\lambda) = \left\| \boldsymbol{C}(\lambda T \boldsymbol{\sigma}\boldsymbol{\sigma}' + |z_\alpha|\, \boldsymbol{E})^{-1} \boldsymbol{g}(T) \right\|^2, \tag{4.10}$$

it can be proved that $\omega(\lambda)$ is strictly decreasing in λ, for $\lambda > 0$. And

$$\omega(0) = \frac{\boldsymbol{g}(T)' \boldsymbol{E}^{-1} \boldsymbol{g}(T)}{|z_\alpha|^2}.$$

If $\boldsymbol{g}(T)' \boldsymbol{E}^{-1} \boldsymbol{g}(T) > |z_\alpha|^2$, then $\omega(\lambda) = 1$ has a unique positive solution λ^* and the optimal solution is given by

$$\pi^* = (T \boldsymbol{\sigma}\boldsymbol{\sigma}' + \frac{|z_\alpha|}{\lambda^*} \boldsymbol{E})^{-1} \boldsymbol{g}(T).$$

If $\boldsymbol{g}(T)' \boldsymbol{E}^{-1} \boldsymbol{g}(T) \leq |z_\alpha|^2$, then $\omega(\lambda) = 1$ has no positive solution, so the optimum portfolio is $\boldsymbol{\pi}^* = \mathbf{0}$, another critical point of $f(\boldsymbol{\pi}, T)$. □

In the next chapter we will illustrate the results of the above theorem.

Chapter 5

Numerical Results

In Chapter 5, we mainly illustrate the empirical results. We start by introducing the numerical techniques that are needed in the numerical experiments: OU process discretization and maximum likelihood method. Then solving the portfolio optimization problem is divided into the following steps. First, we construct three virtual futures prices. Next we calibrate the model parameters, which is followed by simulating the futures prices. Finally, we solve the optimal trading strategy using the results in Chapter 4 and get the simulated portfolio wealth. In the next section, we move on to the case that allow the portfolio to be reoptimized periodically, and compare the results with the case that the portfolio weights are kept constant throughout the time period. Then we investigate what can happen when taking into account the contract size and make the portfolio to be tradeable in the real word, and compare the performance of the portfolio with the idealized case. Finally, we discuss the possibility that our model is also applible to crude oil futures.

5.1 Numerical Techniques

5.1.1 OU Process Discretization

For a given Ornstein-Uhlenbeck process which follows

$$dX_t = \lambda(\mu - X_t)dt + \sigma dW_t,$$

X_t can be formally solved by variation of parameters. Writing

$$f(X_t, t) = X_t e^{\lambda t},$$

we get

$$\begin{aligned} df(X_t, t) &= \lambda X_t e^{\lambda t} dt + e^{\lambda t} dX_t \\ &= \lambda X_t e^{\lambda t} dt + e^{\lambda t} \left[\lambda(\mu - X_t)dt + \sigma dW_t\right] \\ &= \lambda \mu e^{\lambda t} dt + e^{\lambda t} \sigma dW_t \end{aligned}$$

Integrating from t to $t + \Delta t$, we get

$$\begin{aligned} X_{t+\Delta t} e^{\lambda(t+\Delta t)} - X_t e^{\lambda t} &= \int_t^{t+\Delta t} \mu\lambda e^{\lambda s} ds + \int_t^{t+\Delta t} \sigma e^{\lambda s} dW_s \\ &= \mu \left[e^{\lambda(t+\Delta t)} - e^{\lambda t}\right] + \sigma \int_t^{t+\Delta t} e^{\lambda s} dW_s \end{aligned}$$

Then

$$X_{t+\Delta t} = X_t e^{-\lambda \Delta t} + \mu(1 - e^{-\lambda \Delta t}) + \sigma e^{-\lambda(t+\Delta t)} \int_t^{t+\Delta t} e^{\lambda s} dW_s.$$

Let

$$Y_t := \int_t^{t+\Delta t} e^{\lambda s} dW_s,$$

where Y_t is a Gaussian random variable, i.e.,$Y_t \sim N(\mathbb{E}(Y_t), \mathrm{Var}(Y_t))$. And since the mean of a stochastic integral is 0, then $\mathbb{E}(Y_t)$ and according to Itô's Isometry

$$\begin{aligned}\mathrm{Var}(Y_t) &= \mathbb{E}(Y_t^2) - \mathbb{E}^2(Y_t) \\ &= \mathbb{E}(\int_t^{t+\Delta t} e^{2\lambda s} ds) - 0 \\ &= \frac{1}{2\lambda}(e^{2\lambda t + \Delta t} - e^{2\lambda t}),\end{aligned}$$

then

$$\sigma e^{-\lambda(t+\Delta t)} \int_t^{t+\Delta t} e^{\lambda s} dW_s = \sigma \sqrt{\frac{1 - e^{-2\lambda \Delta t}}{2\lambda}} Z_t,\ Z_t \sim N(0,1),$$

so

$$X_{t+\Delta t} = X_t e^{-\lambda \Delta t} + \mu(1 - e^{-\lambda \Delta t}) + \sigma \sqrt{\frac{1 - e^{-2\lambda \Delta t}}{2\lambda}} Z_t,\ Z_t \sim N(0,1).$$

Let

$$a = \mu(1 - e^{-\lambda \Delta t}),\ b = e^{-\lambda \Delta t},\ c = \sigma \sqrt{\frac{1 - e^{-2\lambda \Delta t}}{2\lambda}},$$

If we define

$$\Delta t_n = t_{n+1} - t_n,$$

and

$$\Delta t_n \equiv \Delta t = 1$$

we have the discrete form of the Ornstein-Uhlenbeck process

$$X_{t+1} = a + bX_t + cZ_t,\ Z_t \sim N(0,1).$$

Other discretization methods for OU process can be find in Appendix A.1

5.1.2 Maximum Likelihood Method

Assume $X(t)$ follows the Ornstein-Uhlenbeck process:

$$dX(t) = (\hat{a} - bX(t))dt + \sigma dW_t,$$

and follow the steps in Chapter 5.1.1 we get the discrete form:

$$x_{n+1} = \mu + \lambda x_n + \gamma z_n, \quad n = 0, ..., N, \tag{5.1}$$

where

$$\mu = e^{-b\Delta t},\ \lambda = \frac{\hat{a}}{b}(1 - e^{-b\Delta t}),\ \gamma^2 = \sigma^2(\frac{1 - e^{-2b\Delta t}}{2b}), \tag{5.2}$$

and $x_n = X(t)$, $\Delta t \equiv 1$. We assume z_n is independently and identically distributed over time with a normal distribution with mean 0 and variance 1, i.e, $z_n \sim N(0, 1), n = 0, ..., N$. And we consider x_0 is known.

Our object is to estimate the model parameter $\theta = (\mu, \lambda, \gamma)'$, then the joint density

$$\begin{aligned} f_\theta(x_0, ..., x_N) &= f_\theta(x_N|x_{N-1}, ..., x_0) f_\theta(x_0, ..., x_{N-1}) \\ &= f_\theta(x_N|x_{N-1}, ..., x_0) f_\theta(x_{N-1}|x_{N-2}, ..., x_0) f_\theta(x_0, ..., x_{N-1}) \\ &... \\ &= \prod_{n=0}^{N} f_\theta(x_n|x_{n-1}, ..., x_0) f_\theta(x_0). \end{aligned}$$

Due to the model property that x_n only depends on x_{n-1}, the conditional distribution

$$x_n|x_{n-1}, ..., x_0 = x_n|x_{n-1}.$$

This implies that the conditional density

$$f_\theta(x_n|x_{n-1},...,x_0) = f_\theta(x_0|x_{n-1}).$$

Moreover, due to the fact that z_n is normally distributed, we have that

$$x_n|x_{n-1} \sim N(\mathbb{E}(x_n|x_{n-1}),\ \mathrm{Var}(x_n|x_{n-1})).$$

In particular,

$$\mathbb{E}(x_n|x_{n-1}) = \mu + \lambda x_{n-1},\ \mathrm{Var}(x_n|x_{n-1}) = \gamma.$$

Thus the conditional density of x_n is simply based on the Gaussian density

$$\begin{aligned} f_\theta(x_n|x_{x-1}) &= \frac{1}{\sqrt{2\pi \mathrm{Var}(x_n|x_{n-1})}} \exp\left[-\frac{(x_n - \mathbb{E}(x_n|x_{n-1})^2}{2\mathrm{Var}(x_n|x_{n-1})}\right] \\ &= \frac{1}{\sqrt{2\pi\gamma^2}} \exp\left[-\frac{(x_n - \lambda x_{n-1} - \mu)^2}{2\gamma^2}\right]. \end{aligned}$$

We can then write down the conditional likelihood

$$\begin{aligned} L(\theta) &= f_\theta(x_1,...,x_N|x_0) \\ &= \prod_{n=1}^{N} f_\theta(x_n|x_{n-1}). \end{aligned}$$

Lastly, we can consider the maximum likelihood estimator for θ called the log-likelihood:

$$\begin{aligned} LL(\theta) &= \sum_{n=1}^{N} \ln f_\theta(x_n|x_{n-1}) \\ &= -\sum_{n=1}^{N} \left[\frac{1}{2}\ln(2\pi\gamma^2) + \frac{1}{2\gamma^2}(\hat{x}_n - \lambda\hat{x}_{n-1} - \mu)\right]. \end{aligned}$$

Our optimal problem now is find appropriate θ such that it maximize $LL(\theta)$. It is clear that $LL(\theta)$ gets the optimal when its derivative with respect to μ, λ and γ all equal to 0, i.e.

$$\begin{cases} \dfrac{\partial LL(\theta)}{\partial \mu} = \dfrac{1}{\gamma^2}\displaystyle\sum_{n=1}^{N}(\hat{x}_n - \lambda\hat{x}_{n-1} - \mu) = 0 \\ \dfrac{\partial LL(\theta)}{\partial \lambda} = \dfrac{1}{\gamma^2}\displaystyle\sum_{n=1}^{N}\hat{x}_{n-1}(\hat{x}_n - \lambda\hat{x}_{n-1} - \mu) = 0 \\ \dfrac{\partial LL(\theta)}{\partial \gamma^2} = -\displaystyle\sum_{n=1}^{N}\frac{1}{\gamma^2} + \sum_{n=1}^{N}\frac{1}{\gamma^4}(\hat{x}_n - \lambda\hat{x}_{n-1} - \mu)^2 = 0. \end{cases}$$

From the third equation we get

$$\gamma^2 = \frac{\sum_{n=1}^{N}(\hat{x}_n - \lambda\hat{x}_{n-1} - \mu)^2}{\sum_{n=1}^{N} 1}.$$

In order to simplify the first two equations, let

$$\boldsymbol{A_1} = \begin{pmatrix} \sum_{n=1}^{N} 1 & \sum_{n=1}^{N}\hat{x}_{n-1} \\ \sum_{n=1}^{N}\hat{x}_{n-1} & \sum_{n=1}^{N}\hat{x}_{n-1}^2 \end{pmatrix},$$

$$\boldsymbol{b} = \begin{pmatrix} \sum_{n=1}^{N}\hat{x}_n \\ \sum_{n=1}^{N}\hat{x}_n\hat{x}_{n-1} \end{pmatrix},$$

$$\boldsymbol{x} = \begin{pmatrix} \mu \\ \lambda \end{pmatrix},$$

we get the following linear system

$$\boldsymbol{A_1 x} = \boldsymbol{b}.$$

Remark 5.1. *Instead of directly using the inverse of matrix $\boldsymbol{A_1}$ to solve $\boldsymbol{x}$, we can use QR factorization method to find the solution(see Appendix B.1). And the least square method can also be used to solve OU process(see Appendix C). The least square method and the maximum likelihood method end with the same linear system.*

5.2 Portfolio Optimization

We now tell the story of the numerical experiments. We illustrate the performance of the optimal portfolio using Alberta historical natural gas futures AB NIT. The market data were provided by NRGSTREAM (www.nrgstream.com).

5.2.1 Construct Virtual Futures Prices

In Chapter 2, we introduced how natural gas futures are traded. Once reached the last trading day of one futures contract, then delivery will start soon and this contract will no longer be traded in the market. A new futures contract with delivery period in the far future will come out instead. In order to deal with these roll-over effects and maintain the consistency of the contracts in the portfolio, we generate prices for virtual monthly contracts with fixed time-to-delivery. Interpolation is used here to create a set of futures prices $F_i(t)$, $i = 1, 2, 3$ and we will use these prices instead of real historical prices in the portfolio optimization. This process is done by forming weighted averages of traded futures contracts for months overlapping with the moving windows associated with these virtual contracts. The traded futures contracts are monthly contracts for the period 01-Jan-2015 to 06-Feb-2018, and we created three virtual monthly contracts respectively using these historical contracts, with 30 days, 60 days and 90 days prior to the start of each of the monthly delivery periods.

Figure 5.1 is an example showing the relationship between a virtual contract and real contracts. The red lines represent two historical futures prices on market date of 16-June-2015, the blue line represents one virtual futures price we construct, whose value is the linear combination of the red ones. The price will roll ahead everyday and till the first real contract ends trading; the value of the blue line will then be the combination of the second real contract and the one comes after it. If we want to follow the trading strategy given by the portfolio optimization, it is not hard to find the number of real futures contracts that we

need to hold since the linear relationship between the real and virtual contracts is clear.

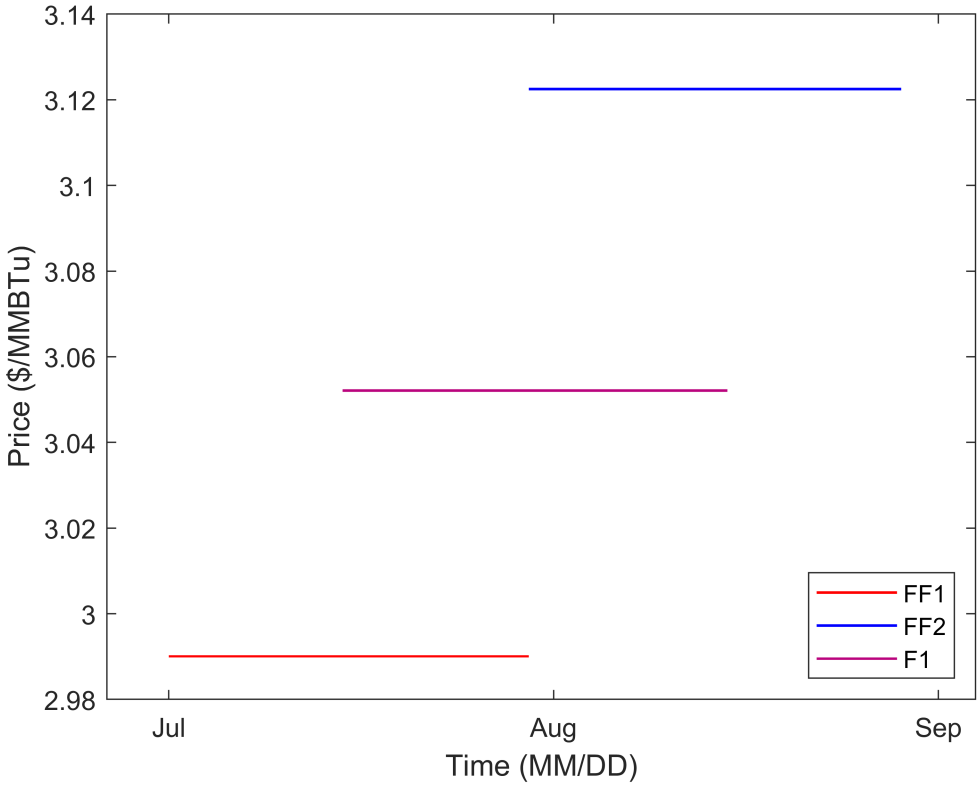

Figure 5.1: Relationship between virtual prices and historical prices. FF denotes the real historical futures prices on the market date of 16-June-2015 and F denotes the virtual futures price we generated.

5.2.2 Model Calibration

Once we get the artificial futures prices above, the one-factor Schwartz model is then fitted to our case. And the log price $Y_i(t) = \ln F_i(t)$ follows the Ornstein-Uhlenbeck process given in (3.6). We will start our model calibration from estimating the parameters of the log price $Y_i(t)$.

First, we rewrite (3.6) into the discrete form (5.1), and then calibrate the model using maximum likelihood estimation. The methodology can be found in Chapter 5.1.2. The

estimated parameters for the discrete form can be found in Table 5.1.

	x_1	x_2	x_3
μ	0.0054	0.0058	0.0020
λ	0.9916	0.9922	0.9966
γ	0.0321	0.0348	0.0326

Table 5.1: Estimated Parameters of x_i

Then using (5.2) we are able to get the resulting values for β, L and the square roots of the diagonal values of $\boldsymbol{\sigma}\boldsymbol{\sigma}'$ (see Table 5.2). L is the long term mean of the log price $Y_i(t)$ and the mean-reversion rate β measures the speed of convergent to the long term mean L.

	F_1	F_2	F_3
β	3.0890	2.8757	1.2416
L	0.8242	0.8188	0.7487
σ	0.6154	0.6684	0.6243

Table 5.2: Estimated Parameters of F_i

We recall that the futures prices's variance-covariance matrix, which denote by $\boldsymbol{\Gamma}$, is symmetric positive semi-definite, so using Cholesky factorization(see Appendix B.2) it is equal to $\boldsymbol{\sigma}\boldsymbol{\sigma}'$, where $\boldsymbol{\sigma}$ is the volatility matrix that we defined in the market setting section of chapter two. We also recall that $\boldsymbol{\Gamma}$ has its LDL variant(see Appendix B.2.1), which means it can be decomposed as

$$\boldsymbol{\Gamma} = \boldsymbol{\nu}\boldsymbol{\rho}\boldsymbol{\nu} \tag{5.3}$$

where $\boldsymbol{\rho}$ is the futures prices correlation matrix, and $\boldsymbol{\nu}$ is a diagonal matrix with the entries equal to the futures prices' standard deviations. Therefore, from the above equation we get

$$\boldsymbol{\Gamma} = \boldsymbol{\sigma}\boldsymbol{\sigma}' = \boldsymbol{\nu}\boldsymbol{\rho}\boldsymbol{\nu} = \boldsymbol{\nu}\boldsymbol{\rho}_1\boldsymbol{\rho}_1{}'\boldsymbol{\nu} = \boldsymbol{\nu}\boldsymbol{\rho}_1(\boldsymbol{\nu}\boldsymbol{\rho}_1)',$$

where $\boldsymbol{\rho} = \boldsymbol{\rho}_1\boldsymbol{\rho}_1{}'$ is the Cholesky decomposition of $\boldsymbol{\rho}$.

Please note that in formulae (5.2), if we plug in the value of γ_i, we get the entries of $\boldsymbol{\nu}$, i.e. the square roots of the diagonal values of $\boldsymbol{\sigma}\boldsymbol{\sigma}'$ (see σ in Table 5.2), rather than the volatility matrix $\boldsymbol{\sigma}$. When we calibrate the parameters, the correlation between the prices of our three futures contracts is not considered and the calibrations are done separately and independently. So in order to get the correlation matrix $\boldsymbol{\rho}$, we need the following steps.
First in the discrete form (5.1), we write the noise term as the difference of x_{n+1} and $\mu+\lambda x_n$, so for all the three series of log futures prices we get three vectors of noises. Compute the correlation of these vectors will give us the following correlation matrix.

$$\boldsymbol{\rho} = \begin{bmatrix} 1 & 0.9182 & 0.8145 \\ 0.9182 & 1 & 0.9594 \\ 0.8145 & 0.9594 & 1 \end{bmatrix}.$$

According to the discussion in Chapter 3.1.1, once we get the correlation matrix $\boldsymbol{\rho}$, we are able to built the term $\boldsymbol{\sigma} d\boldsymbol{W}(t)$ in (3.3):

$$\boldsymbol{\nu} d\boldsymbol{B}(t) = \boldsymbol{\nu}\boldsymbol{\rho_1} d\boldsymbol{W}(t) = \boldsymbol{\sigma} d\boldsymbol{W}(t),$$

where $\boldsymbol{W}(t) = (W_1(t), W_2(t), W_3(t))'$ is a 3-dimensional vector of three independent Brownian motions, $\boldsymbol{B}(t) = (B_1(t), B_2(t), B_3(t))'$ is a 3-dimensional vector of three correlated Brownian motions with correlation matrix $\boldsymbol{\rho}$, and $\boldsymbol{\rho}_1\boldsymbol{\rho}_1{}'$ is the Cholesky decomposition of $\boldsymbol{\rho}$. Since variance-covariance matrix $\boldsymbol{\Gamma}$ is symmetric positive semi-definite, this leads to a nonunique Cholesky decomposition of $\boldsymbol{\Gamma}$ into the product $\boldsymbol{\sigma}\boldsymbol{\sigma}'$(see Appendix B.2 for details). Despite this fact, the expressions for $\mathrm{CaR}(\boldsymbol{\pi}, t)$ and $\mathbb{E}[X^{\pi}(t)]$ are uniquely determined, since all the terms that involve $\boldsymbol{\sigma}$, i.e., $\|\boldsymbol{\pi}'\boldsymbol{\sigma}\|^2$ and $\boldsymbol{E}(t, \boldsymbol{b})$, can be expressed as

$$\|\boldsymbol{\pi}'\boldsymbol{\sigma}\|^2 = \boldsymbol{\pi}'\boldsymbol{\sigma}\boldsymbol{\sigma}'\boldsymbol{\pi} = \boldsymbol{\pi}'\boldsymbol{\Gamma}\boldsymbol{\pi},$$
$$\boldsymbol{E}(t, \boldsymbol{b}) = \boldsymbol{\sigma}\boldsymbol{\sigma}' \centerdot \boldsymbol{\varepsilon}(t, \boldsymbol{b}) = \boldsymbol{\Gamma} \centerdot \boldsymbol{\varepsilon}(t, \boldsymbol{b}).$$

where the matrix $\boldsymbol{\varepsilon}(t, \boldsymbol{b})$ has the entries equal to $\varepsilon(t, b_i + b_j)$ defined in (3.8), and the symbol '$\boldsymbol{.}$' stands for the Hadamard product of two matrices.

5.2.3 Futures Prices Simulation

Using the calibrated parameters and correlation shown in the above subsection, we simulate prices of the three virtual futures contracts. Figure 5.2 shows one such simulation.

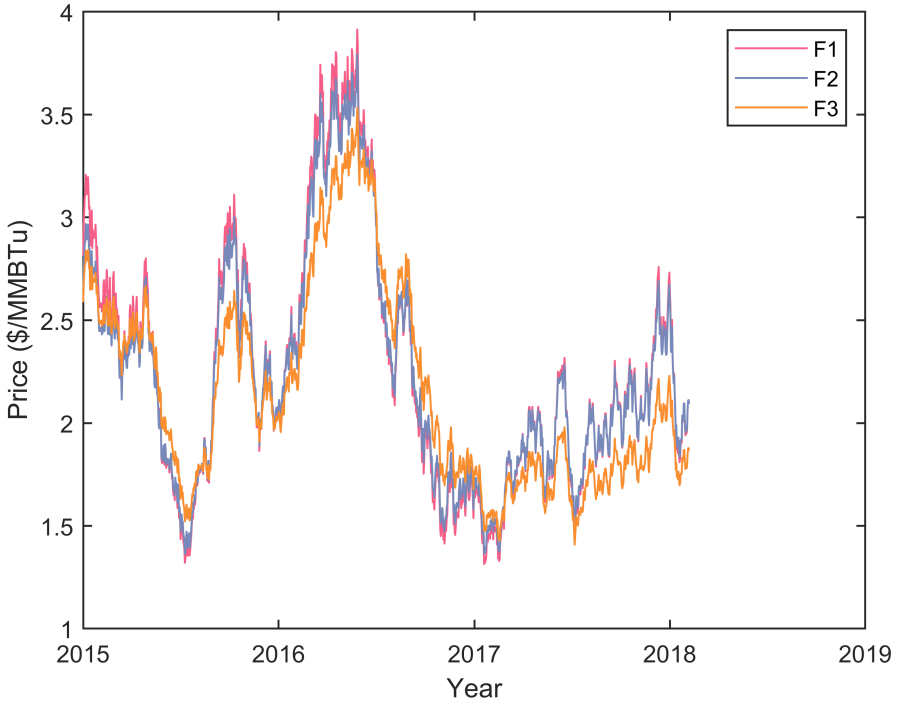

Figure 5.2: Simulated prices of the three virtual futures contracts using the calibrated parameters shown in Table 5.2.

5.2.4 Optimal Trading Strategy

Once the parameters are estimated, we are able to solve the optimization problem with MATLAB. Write a function 'MinCaRPi(t,Covariance,a,b,Y,r,alpha)' with the following inputs: the time window t, the covariance matrix $\boldsymbol{\Gamma}$ defined in (5.3), the vector $\boldsymbol{a}$, $\boldsymbol{b}$ with entries defined in (3.2), the vector of initial futures prices $\boldsymbol{Y}$, the riskless interest rate $r = 2\%$, the

significance level $\alpha = 5\%$. Using these inputs gives $\boldsymbol{E}$, $\boldsymbol{C}$ and $\boldsymbol{g}(T)$ in Theorem 4.2, thus the optimal portfolio $\boldsymbol{\pi}^*$ are able to be calculated(see Table 5.3). And the minimized CaR is -0.0402 for an initial (notional) wealth of \$1.

π_1	π_2	π_3
-0.3745	0.8205	-0.3353

Table 5.3: Optimal Weights of F_i

We now give a brief illustration of the dependence of CaR on the portfolio $\boldsymbol{\pi}$. In Figure 5.3, where we see three cross-sectional contour plots showing CaR as a function of, in turn, (π_1, π_2) (keeping π_3 at its optimal value), (π_2, π_3) (keeping π_1 at its optimal value) and (π_1, π_3) (keeping π_2 at its optimal value). This picture illustrates that the value of $\boldsymbol{\pi}$ that minimises CaR is well-defined.

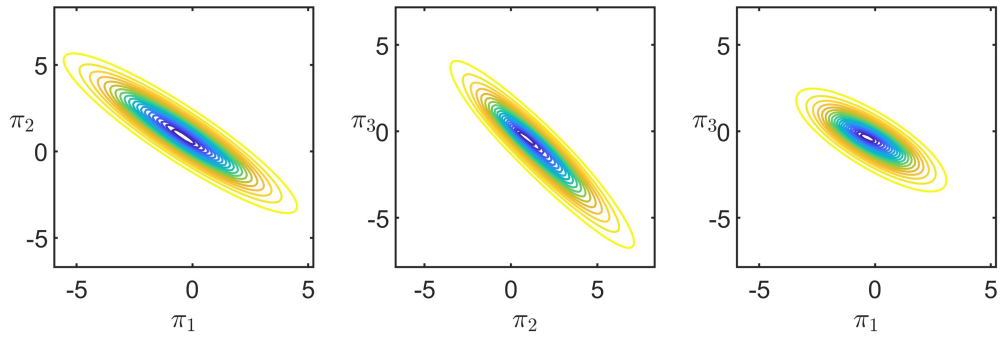

Figure 5.3: CaR as a function of π_1, π_2 and π_3. In each plot, the missing component of π is held constant at its optimal value.

Note that $\boldsymbol{\pi}^*$ then defines the position in each virtual contract, that is a linear combination of positions in actual traded contracts. It is possible to track the result of taking these positions on a daily (or any other) frequency. For example on 20-Sept-2015, the price of the first virtual futures F_1 is 2.7846 and the portfolio wealth $X^{\boldsymbol{\pi}}$ on that day is \$106.5088. Using (3.4) we can get position N_1 is -12.8250. This virtual futures corresponds to two actual futures with delivery start on 01-Oct-2015 and 01-Nov-2015. The proportion of the futures

with delivery start on 01-Oct-2015 in the virtual futures is one third, the other one is two third, so their positions are -4.275 and -8.55.

5.2.5 Portfolio Wealth

Starting with an initial (notional) wealth of \$100, i.e. $X_0 = 100$, Figure 5.4 illustrates a sample simulated path of the wealth process $X^{\pi}(t)$. The performance shown there is compared with an equivalent risk-free investment.

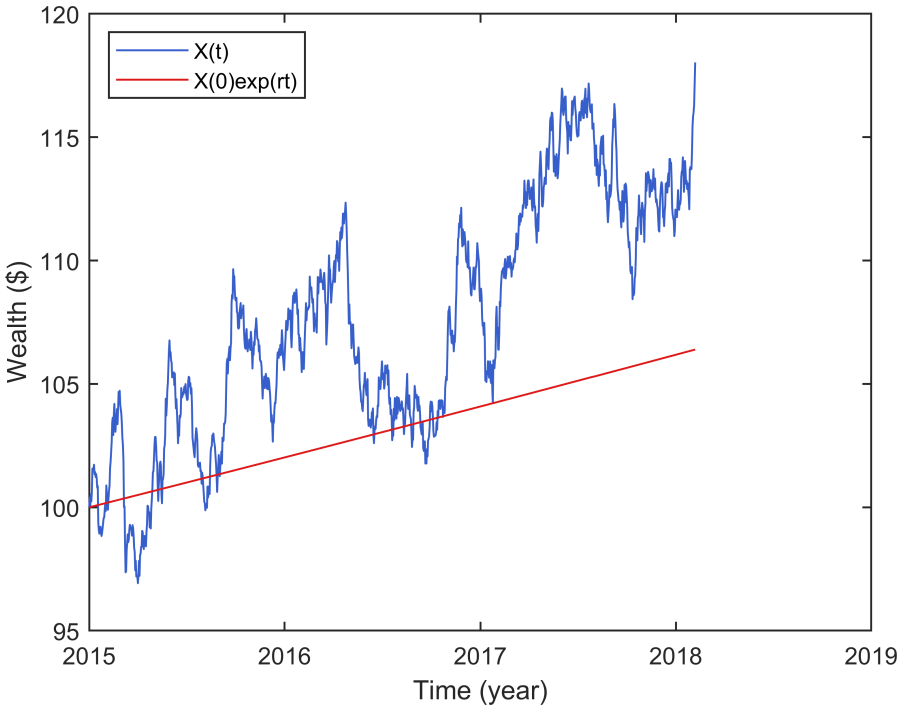

Figure 5.4: Performance of the optimal portfolio with an initial wealth of \$100, compared to a risk-free investment, for a sample simulation.

We can see that for most of the time our portfolio performs better than bank savings. But it can not be guaranteed that the performance of the portfolio is always better than the risk-free investment, especially for the first couple of years. In fact, if the portfolio always did better than the riskless return, it would represent an arbitrage opportunity. So it is expected that it will sometimes underperform.

In Figure 5.5, the ratio between the final wealth of the optimal portfolio and the risk-free investment is shown. This histogram is generated for 1000 simulations. The short red line with the value of 1.0420 denotes the 5% quantile of this distribution, which is corresponding to the quantile in the CaR. It is apparent that the optimized portfolio almost always outperforms the risk-free growth. The mean of the ratio is 1.1168 and the standard deviation is 0.0442.

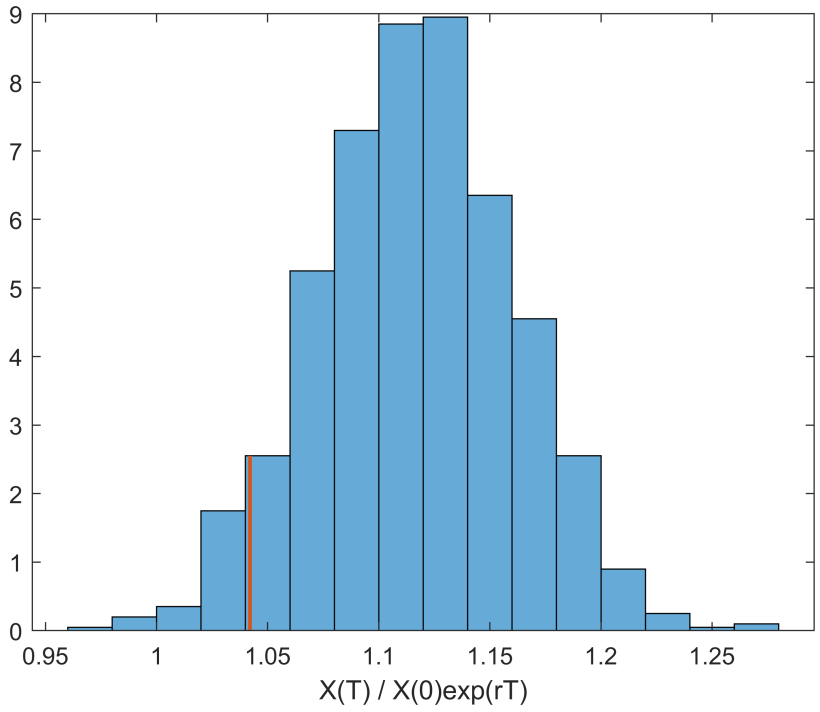

Figure 5.5: A histogram of the ratio of the final wealth of the optimal portfolio to the risk-free final wealth, for 1000 simulations. The short red line denotes the 5% quantile.

The theory developed in [14] and adapted here predicts that the optimal portfolio wealth will be log-normally distributed. This is confirmed in Figure 5.6, where the result of evaluating the portfolio on 1000 sets of simulated prices is illustrated. What is shown is the empirical distribution of the logarithm of the final wealth, with the best fit normal pdf (Probability Density Function) superimposed.

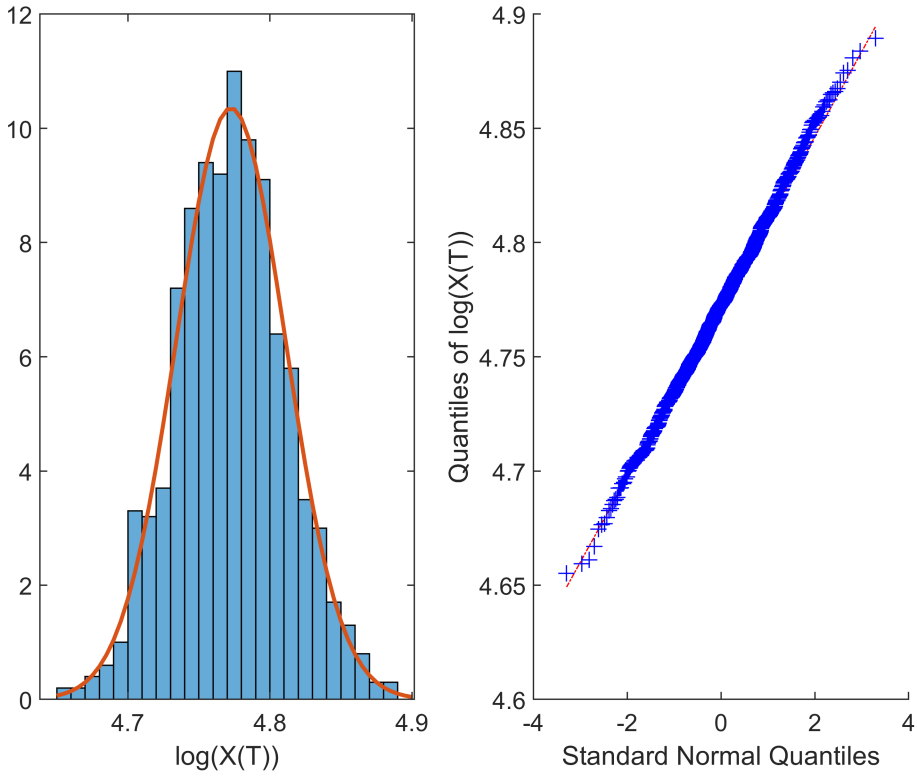

Figure 5.6: A histogram and a QQ plot of the logarithm of the final wealth for 1000 simulations, with a best-fit normal pdf superimposed.

Once we get a histogram like Figure 5.5, we are able to calculated the empirical capital at risk. Since the red line is the 5% quantile q of the ratio $X(T)/(X(0)e^{rT})$, using the definition of CaR given in (4.1), we have the following formula for the empirical CaR:

$$\mathrm{CaR}(\boldsymbol{\pi}, T) = X_0 e^{rT} - q_x = X_0 e^{rT}(1 - q)$$

where q_x is the α-quantile of $X^{\boldsymbol{\pi}}(T)$, X_0 denotes the initial wealth invested in the portfolio and r denotes the riskless rate of return.

In order to check whether the empirical CaR matches the theoretical CaR, we generate

100 histograms like Figure 5.5, calculate the empirical CaR corresponding to each histogram and get the distribution of these empirical CaRs, which is shown in Figure 5.7.

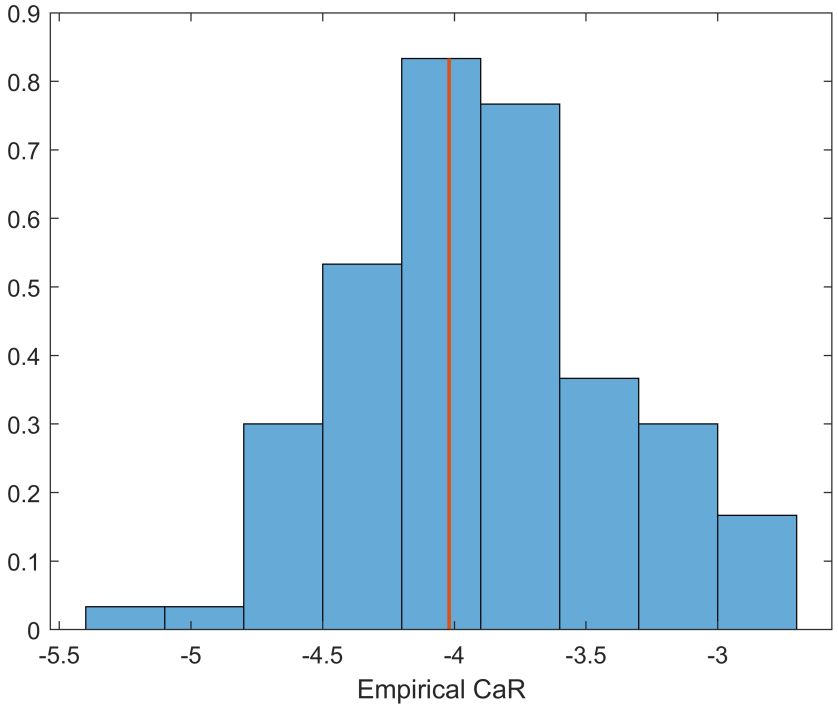

Figure 5.7: A histogram of empirical CaR. This histogram consists of 100 empirical CaRs, and each empirical CaR is given by the 5% quantile of a histogram like Figure 5.5. The short red line denotes the theoretical CaR.

In Chapter 5.2.4, we found the theoretical minimized CaR is -0.0402 for an initial wealth of \$1. So in our optimal portfolio with an initial wealth of \$100, the theoretical CaR should be -4.02, which value is within the distribution of these 100 empirical CaRs. The mean of these 100 empirical CaRs is $\bar{x} = -3.9046$, and the standard deviation is $s = 0.5178$. Assume we have a null hypothesis H_0: the mean of the CaR is $\mu_0 = -4.02$, and the corresponding alternative hypothesis H_1: $\mu_0 \neq -4.02$. Since the sample size is large enough($n = 100$), we can simply assume that CaR is normally distributed, and the test statistics

$$Z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = 2.23.$$

At the 1% significance level, we have the critical value $-2.58 < Z < 2.58$, thus we cannot reject the null hypothesis. This means the empirical CaRs are consistent with the theoretical value at the 99% level of confidence. The numerical implement works well and the results matches the theory.

5.3 Rebalancing the Portfolio

It is notable that the optimal portfolio is computed with reference to the CaR of the wealth at the final time, and the portfolio weights are kept constant throughout the time period. If instead we reoptimize the portfolio periodically through the period (each time with reference to the CaR at the final time), the performance will be affected. We illustrate the results of reoptimizing every 100 days in Figure 5.8 and Figure 5.9.

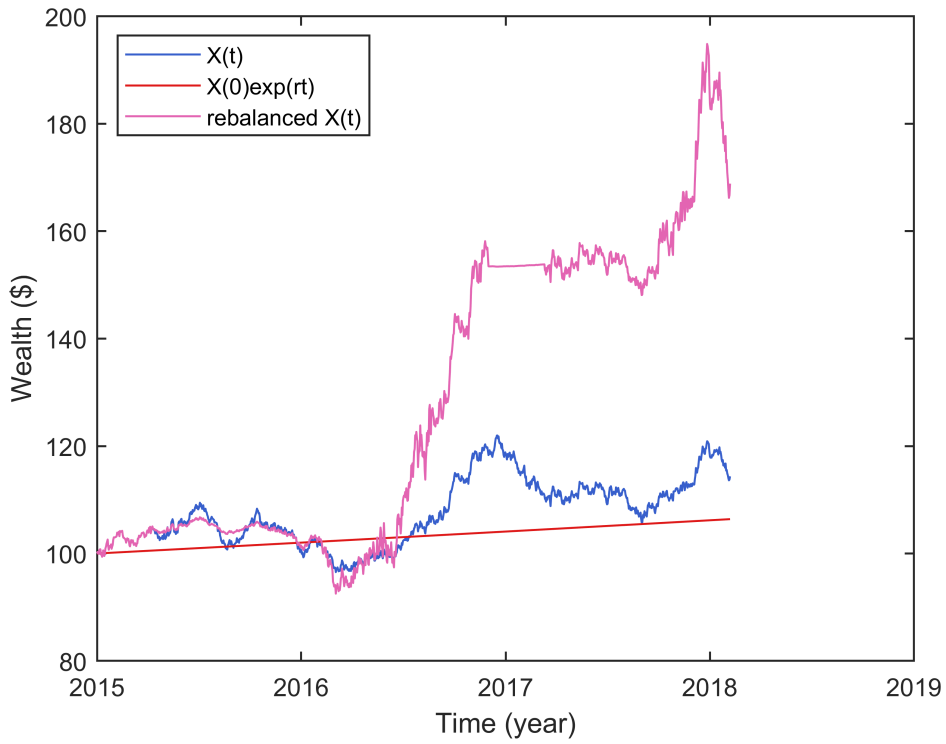

Figure 5.8: Performance of the fixed and periodically rebalanced optimal portfolios relative to an initial wealth of \$100, compared to a risk-free investment, for a sample simulation.

In Figure 5.8, a fixed and a reoptimized portfolio are computed on a single set of sample prices. We can see that the rebalanced portfolio performs better than the portfolio without rebalance mostly but not always. It is because each time we reoptimize the portfolio, it only gives us the the local optimization based on the prices at the rebalancing time. So it is possible that the performance is not better than the fixed optimal portifolio which is the global optimization. We can also see a short straight line within the rebalanced portfolio wealth curve. This illustrates that for this certain reoptimization, the optimal weight is $\boldsymbol{\pi} = 0$. In other words, risk-free investment is the best option to minimize CaR in that period of time.

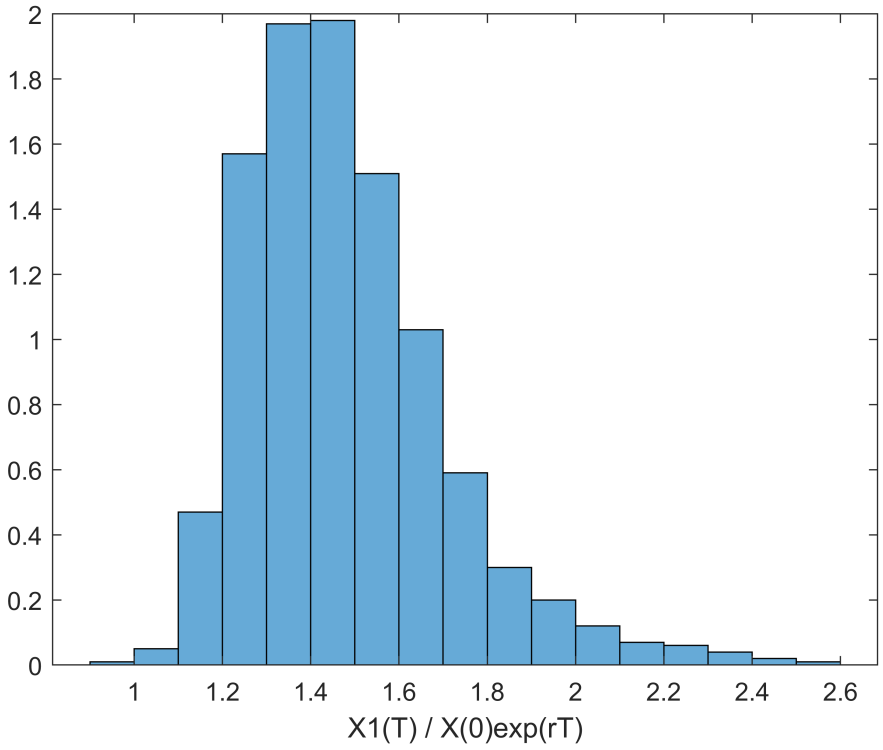

Figure 5.9: A histogram of the ratio of the final wealth of the periodically-rebalanced optimal portfolio to the risk-free final wealth, for 1000 simulations.

In Figure 5.9, a histogram of the ratio between the final wealth of the periodically-rebalanced optimal portfolio and the risk-free final wealth is shown. The mean of this distribution is

1.4813 and the standard deviation is 0.2250. It is apparent that the improved performance relative to the fixed optimal portfolio shown for a single simulation in Figure 5.8 is in fact characteristic-the average final wealth is dramatically higher. But unsurprisingly the variance of the final wealth distribution is also increased.

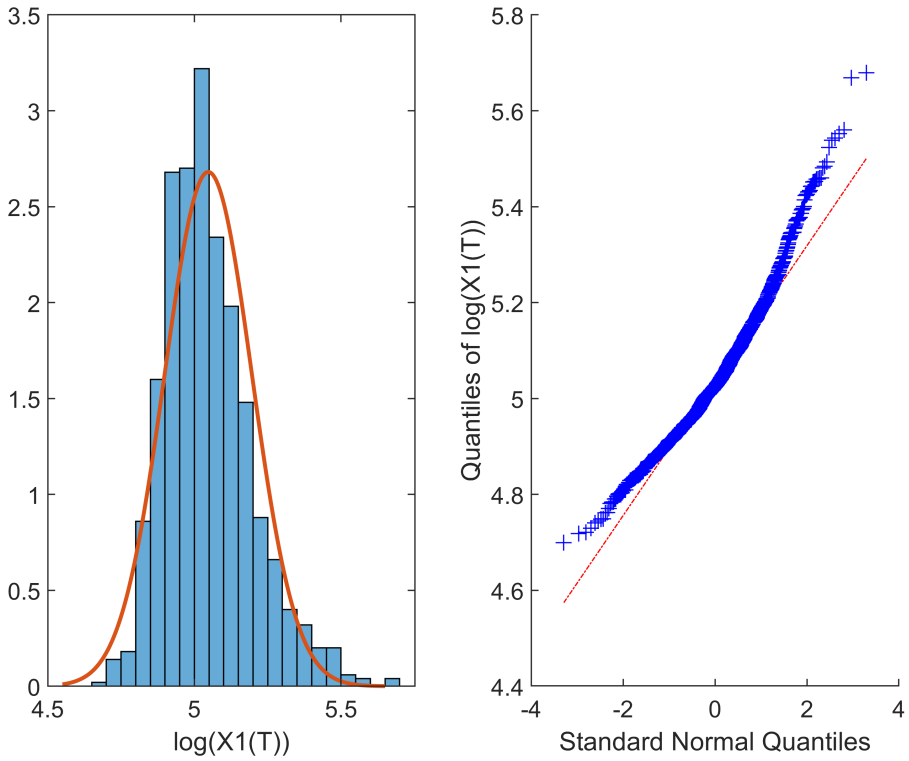

Figure 5.10: A histogram and a QQ plot of the logarithm of the final wealth of the rebalanced optimal portfolio for 1000 simulations, with a best-fit normal pdf superimposed.

In Figure 5.10, a histogram and a QQ plot of the logarithm of the final wealth of the rebalanced optimal portfolio are shown. The red curve represent the best-fit normal pdf. It is clear that the final wealth of the rebalanced optimal portfolio is not log normal anymore. Recall that the optimal portfolio $\boldsymbol{\pi}^*$ is a function of the initial log-prices $Y_i(0)$, $i = 1, 2, 3$. For each rebalancing period, $\boldsymbol{\pi}^*$ is calculated at the beginning of this period of time based on the value of Y_i at that time. These 'initial' values Y_i are, in fact, random and depend

on the values of Y_i in the former period of time. So the wealth of optimal portfolio is not expected to be log-normal in the rebalanced case.

5.4 Tradeable Portfolio

The above discussion about portfolio optimization is idealized, however. There are many limits and additional costs for trading in the real world. In this section, we will make our portfolio optimization a bit more realistic: taking the contract size into account. According to the model assumptions given in the Chapter 3, we assume the futures contracts are perfectly divisible. In fact, the contract size of AB NIT Futures is 2500, which mean the number of contracts we hold for the portfolio not only must be an integer, but also a multiple of 2500, clearly the initial wealth of \$100 is no longer applicable, now we increase the initial value to \$1,000,000. Based on (3.4), we have the following expression for $N_i(t)$:

$$N_i(t) = \frac{\pi_i X^{\boldsymbol{\pi}}(t)}{F_i(t)}, i = 1, 2, 3.$$

Then we are able to get $\bar{N}_i(t)$, which denotes the closest number to $N_i(t)$ that satisfies the contract size restriction. Then the tradeable optimal portfolio weight $\bar{\boldsymbol{\pi}}$ is given by

$$\bar{\pi}_i(t) = \frac{\bar{N}_i(t) F_i(t)}{X^{\boldsymbol{\pi}}(t)}, i = 1, 2, 3.$$

Note that when the futures prices and portfolio wealth change with time, $\bar{\boldsymbol{\pi}}$ changes as well through $\bar{N}_i$ which is not a constant anymore. Replace the π_i with $\bar{\pi}_i(t)$ in (3.5), we get the SDE of the wealth process

$$dX^{\bar{\boldsymbol{\pi}}}(t) = X^{\bar{\boldsymbol{\pi}}}(t) \left(r dt + \sum_{i=1}^{m} \bar{\pi}(t)_i [(a_i - b_i \ln F_i(t)) dt + \boldsymbol{\sigma_i} d\boldsymbol{W}(t)] \right), \; X^{\boldsymbol{\pi}}(0) = X_0,$$

where X_0 is the initial wealth.

Remark 5.2. *In fact, our tradeable portfolio is simi-realistic. Recall that $N_i(t)$ is the position in the artificial futures contract which is the weighted average of actual traded contracts. For the tradeable case that we discuss here, the positions in the virtual futures contracts are the ones that would be rounded, not the positions in traded contracts.*

The following steps are the same as before. Figure 5.11 is a fixed and a tradeable portfolio wealth computed on a set of sample prices. The performance is compared with an equivalent risk-free investment. It is not surprising that we lost some profit in the tradeable portfolio for most of the time. However, it is still reasonable that for some time the tradeable portfolio even performs slightly better than the fixed portfolio, because our criteria for optimizing the portfolio is minimize capital at risk not maximize the profit.

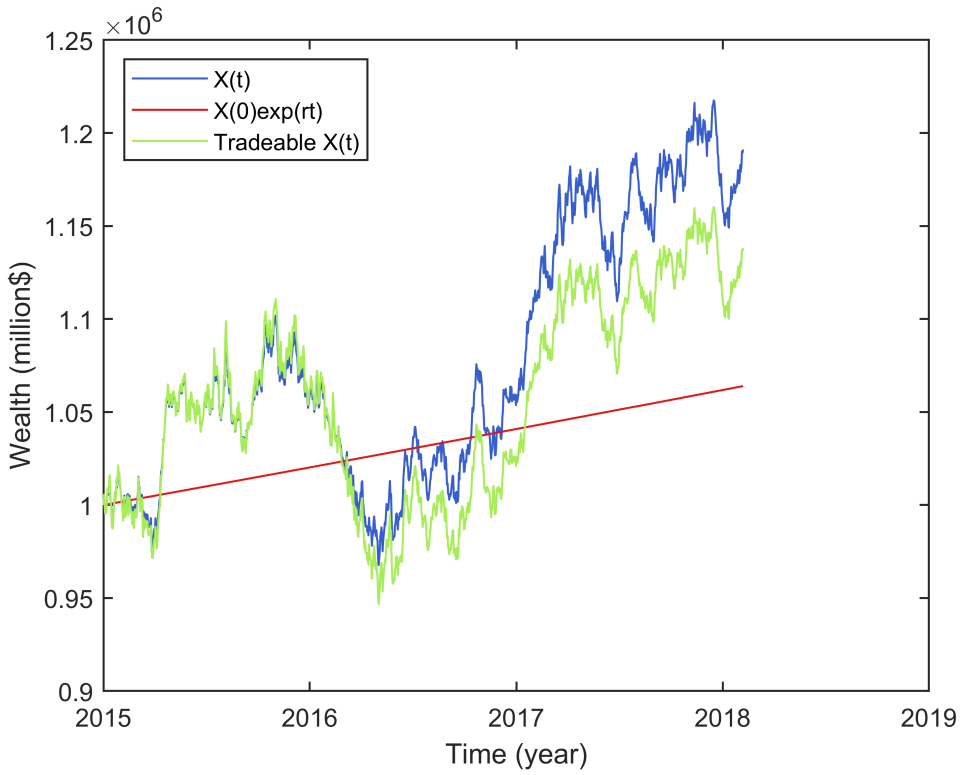

Figure 5.11: Performance of the fixed and tradeable optimal portfolios relative to an initial wealth of one million dollars, compared to a risk-free investment, for a sample simulation.

Figure 5.12 shows a histogram of the ratio between the final wealth of the tradeable optimal

portfolio and the risk-free final wealth for 1000 simulations. As a comparison, the ratio between the final wealth of the fixed optimal portfolio and the risk-free final wealth using the same simulated futures prices, is also plotted in the same figure. The mean of the distribution is 1.077 and the standard deviation is 0.0408. It is apparent that the underperformance of the tradeable optimal portfolio relative to the fixed optimal portfolio shown for a single simulation in Figure 5.11 is in fact characteristic. The average final wealth is significantly lower. But the difference in the standard deviation is not so obvious. Besides, the number of times this ratio is less than 1 is 39, which means there is a 4% probability that the tradeable portfolio will perform worse than an risk-free investment.

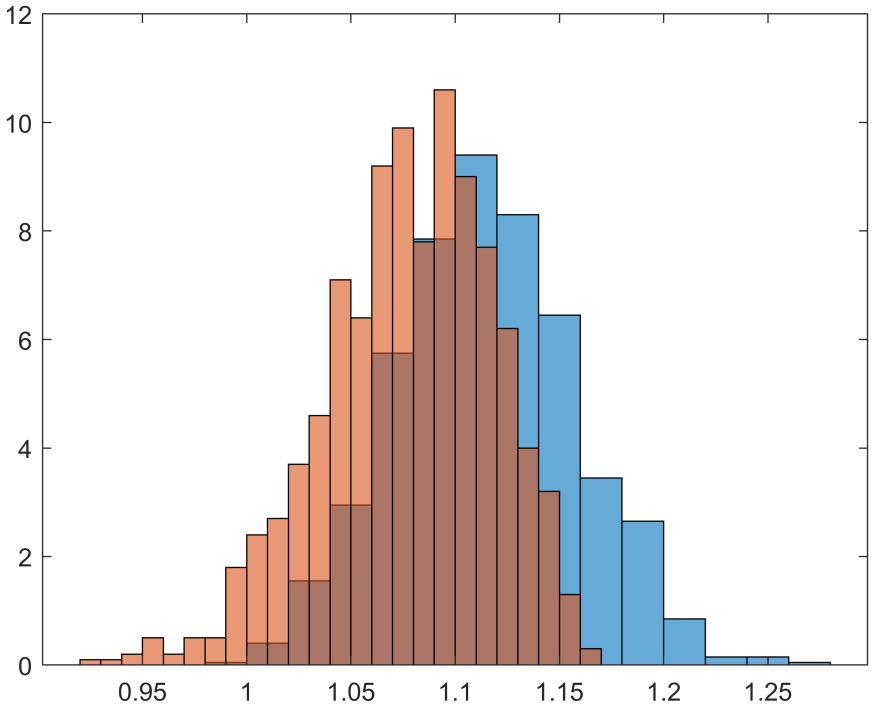

Figure 5.12: A histogram of the ratio of the final wealth of the tradeable optimal portfolio to the risk-free final wealth, and the ratio of the final wealth of the fixed optimal portfolio to the risk-free final wealth as a comparison, for 1000 simulations.

Figure 5.13 illustrates the distribution of the empirical CaR of the tradeable portfolio. This histogram consists of 100 empirical CaRs, and each empirical CaR is given by the 5% quantile

of a histogram for the final wealth of the tradeable portfolio, for 1000 simulation. It is obvious that the overall value range of the empirical CaR has shifted to the right compared to that of the optimal portfolio(see Figure 5.7). The risk is no longer minimized while the profits have been lost to some degree as well.

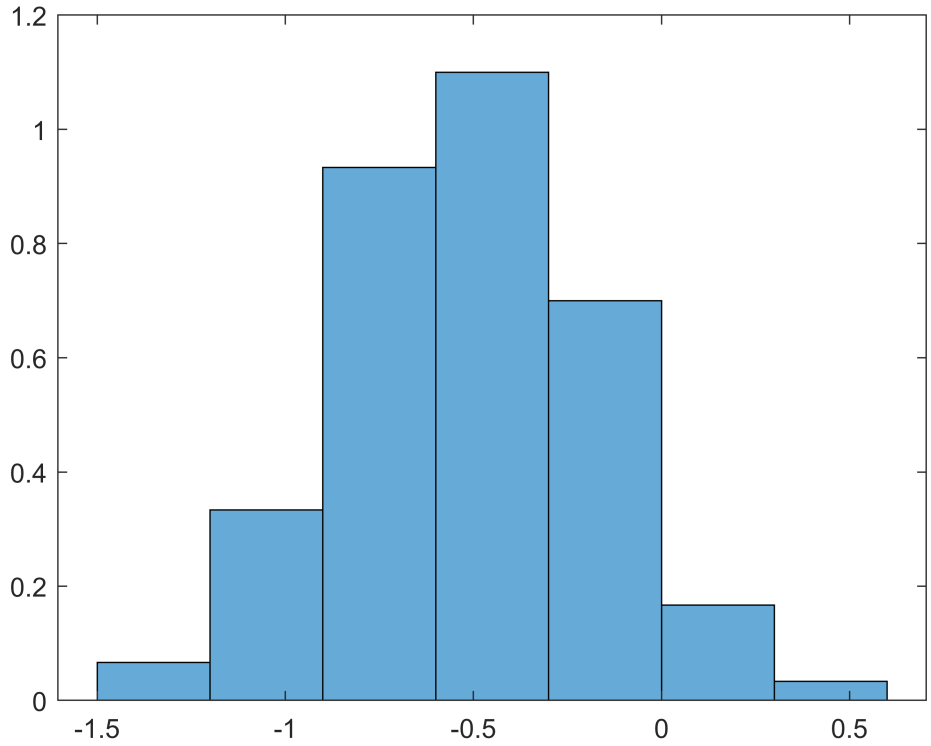

Figure 5.13: A histogram of empirical CaR for tradeable portfolio. This histogram consists of 100 empirical CaRs, and each empirical CaR is given by the 5% quantile of a histogram for the final wealth of the tradeable portfolio for 1000 simulation.

In Figure 5.14, a histogram and a QQ plot of the logarithm of the final wealth of the tradeable portfolio are shown. The red curve represent the best-fit normal pdf. It is not surprising that the final wealth of the rebalanced optimal portfolio is not log normal anymore. A negative skew can be seen in the figure. The fat tail on the left side of the distribution means our tradeable portfolio has a greater chance of losing money than the original portifolio, and the loss is likely to be greater.

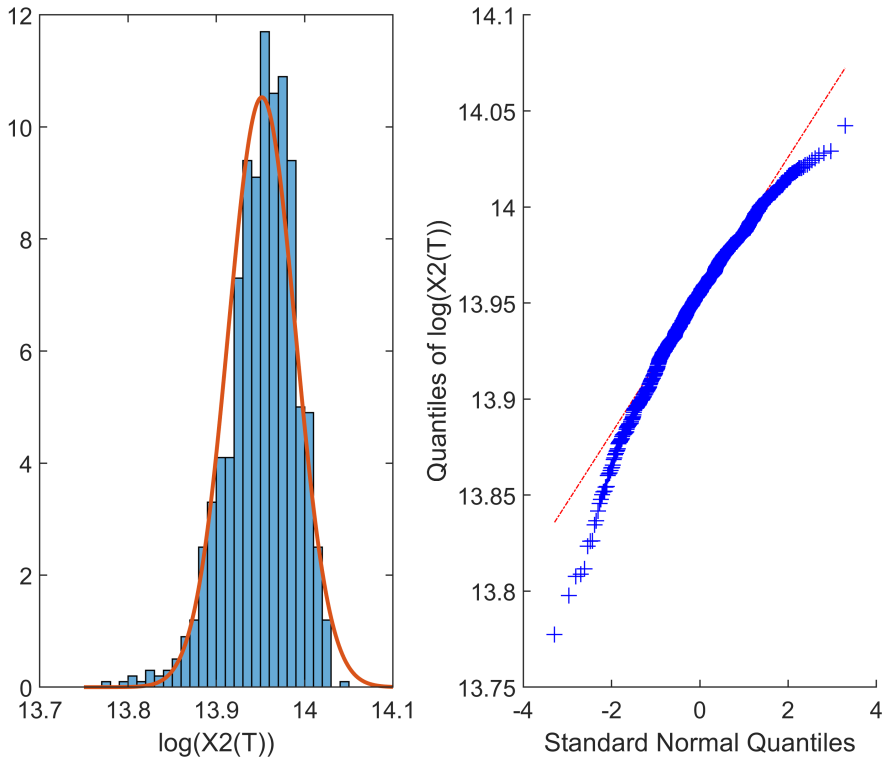

Figure 5.14: A histogram and a QQ plot of the logarithm of the final wealth of the tradeable portfolio for 1000 simulations, with a best-fit normal pdf superimposed.

5.5 Optimal Portfolio of Crude Oil Futures

In addition to being applicable to natural gas futures, our portfolio optimization model is also applicable to crude oil futures, although the delivery procedure is different - while natural gas futures deliver over a period of time, crude oil futures deliver on a single day.

Take NYMEX WTI crude oil futures as an example. Trading terminates 3 business days prior to the 25th calendar day of the month ahead the contract month. Once a futures contract reaches its last trading day, the price will be fixed and it will no longer be traded in the market. Another new futures contract with delivery time in the far future will come out. So constructing virtual futures prices will still be needed in order to deal with the roll-over effects. Interpolation is used here to create a set of futures prices with fixed time-to-delivery.

The virtual contract price is the linear combination of two traded futures contracts' prices. The weight of each traded contract rolls ahead every day: as the former traded contract approaches the last trading day, its proportion in the virtual contract gradually decreases to zero, while the proportion of the latter traded contract gradually increases to one.

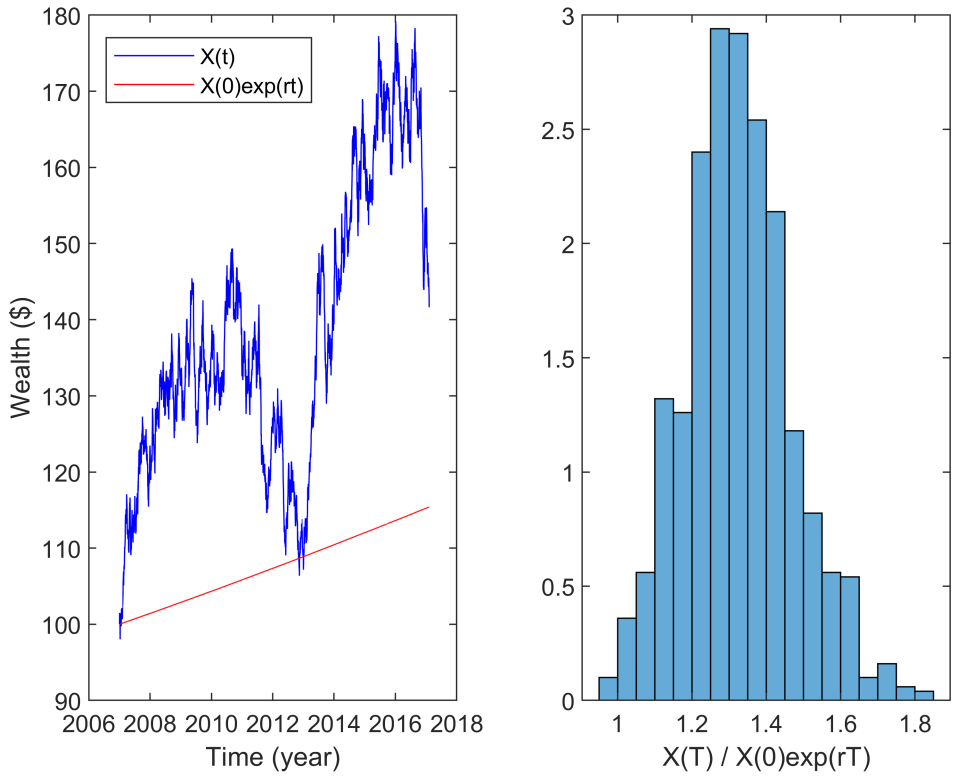

Figure 5.15: The performance of the optimal portfolio of WTI Crude Oil futures compared to a risk-free investment and a histogram of the ratio of the final wealth of the optimal portfolio to the risk-free final wealth.

The remaining process is the same as before. Figure 5.15 illustrates a sample simulation that shows the performance of the optimal portfolio of WTI crude oil futures with an initial wealth of $100 comparing to a risk-free investment; and a histogram of the ratio of the final wealth of the optimal portfolio to the risk-free final wealth, for 1000 simulation. The market data were provided by NRGSTREAM. The performance of the optimal portfolio looks reasonable and this confirms that our model can also adapt to crude oil futures.

Chapter 6

Conclusion and Future Work

In this thesis we have investigated the portfolio allocation problem when the portfolio consists of positions in futures contracts. We first introduced the natural gas market and natural gas futures, then gave some background of basic stochastic process and stochastic integral, and the market setting of our model. The futures prices processes follow correlated exponential OU processes and the prices are log normally-distributed. Assuming that short-selling is allowed and trading is continuous over a finite time horizon, we derived analytic expressions for the wealth process and the capital at risk of a portfolio where the positions are held in constant proportion. Then the solution for the portfolio optimization problem that minimize the Capital at Risk is shown, which is a straight forward reformulation of Theorem 4.1 and Theorem 4.2 in [14].

When all the necessary formulae are ready, we provided some empirical results. In order to deal with roll-over effects of futures contracts and maintain the consistency of the contracts in the portfolio, we generated futures prices for virtual monthly contracts with fixed time-to-delivery. Then we discretized the OU process and used maximum likelihood method to calibrate the model parameters. Mean-revertion can be seen in these estimated parameters that were subsequantly used for futures price simulations. Next, we solved the optimal

portfolio and got the simulated portfolio wealth. The empirical distribution of the logarithm of the final wealth showed it fits perfectly with the normal distribution, and the theoretical value of CaR also falls within the interval of experimental distribution of CaR. This means the numerical implementation works well and the results match the theory. In the next section, we looked at the case that allows the portfolio to be reoptimized periodically, and compare the results with the constant portfolio. The average final wealth of the rebalanced portfolio is dramatically higher; however the variance of the final wealth distribution has also increased. Then we considered the contract size and made the portfolio to be tradeable in real word. It is apparent that there is underperformance of tradeable optimal portfolio relative to the fixed optimal portfolio. At last, we discussed the possibility of adapting our model to crude oil futures. Although the delivery procedure of crude oil futures is different from natural gas futures, our portfolio optimization model can also work with crude oil futures.

There are several possibilities for future work. One interesting task will be add more assets to the portfolio and test the performance. Furthermore, incorporating existing risk exposures and making hedging portfolios. For instance, consider the risk exposure X_T of futures and risk exposure Y_T of the corresponding underlying assets, we might be interested in minimizing some risk measure calculated on $X_T + Y_T$. According to our discussion, X_T is log normal-distributed, and we can set Y_T to be log normal-distributed as well. Maybe $Y_T = \exp(y_T)$ and $X_T = \exp(x_T)$, and the SDEs satisfied by x_T and y_T have correlated noise term. Then the mean and variance of $X_T + Y_T$ can be expressed in terms of $\boldsymbol{\pi}$, and hence risk measures based on these information can be optimized.

One further possible future research direction is to explore the portfolio allocation problem where the constraint is some other risk measure. Take conditional capital at risk as an example. It can eliminate the drawback of CaR that does not describe the tail loss and

does not penalize the worst conditions. So comparing the results with respect to conditional CaR and the results in this paper would be interesting. Another important task could be to extend the modeling assumption to allow time-dependent portfolios, and even consider more general pricing models such as Lévy processes.

Bibliography

[1] E.S. Schwartz. *The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging*. Journal of Finance, Vol 52, No 3, pp 923-973, 1997.

[2] R. Gibson, E. S. Schwartz. *Stochastic Convenience Yield and the Pricing of Oil Contingent Claims*. Journal of Finance, Vol 45, No 3, pp 959-976, 1990.

[3] D. R. Ribeiro, S. D. Hodges. *A Two-Factor Model for Commodity Prices and Futures Valuation*. Warwick Business School, University of Warwick, Coventry, United Kingdom, 2004.

[4] S. Hikspoors, S.Jaimungal. *Energy Spot Price Models and Spread Options Pricing* . International Journal of Theoretical and Applied Finance, Vol 10, issue 07, pp 1111-1135, 2007.

[5] T.T. Welemical, J. A. Aduda, M. L. D. M. Bidima. *Asymptotic Exponential Arbitrage in the Schwartz Commodity Futures Model*. International Journal of Mathematics and Mathematical Sciences., 2019.

[6] J. A. Wachter. *Portfolio and Consumption Decisions under Mean-Reverting Returns: An Exact Solution for complete Markets*. The Journal of Financial and Quantitative Analysis, Vol 37, No 1, 2002.

[7] F. E. Benth and K H. Karlsen. *A Note on Merton's Portfolio Selection Problem for the Schwartz Mean Reversion Model*. Stochastic Analysis and Applications, Vol 23, No 34, 2004.

[8] A. d'Aspremont. *Identifying Small Mean Reverting Portfolios*. Economics, Computer Science, Mathematics, 2008.

[9] J. Fouque, R. Hu. *Optimal Portfolios with Bounded Capital at Risk*. AMathematical Finance, Vol 11, pp 365-384, 2001.

[10] S. Emmer, C. Klüppelberg and R. Korn. *Optimal Portfolios with Bounded Capital at Risk*. Mathematical Finance, Vol 11, pp 365-384,2001.

[11] S. Emmer and C. Kluppelberg. *Optimal Portfolios when Stock Prices Follow an Exponential L'evy Process*. Finance and Stochastics, No 8, 2004.

[12] G. Dmitrašinović-Vidović, A. Lari-Lavassani, X. Li and A. Ware. *Dynamic Portfolio Selection Under Capital at Risk*. University of Calgary Yellow Series. Report 833. 2003.

[13] G. Dmitrašinović-Vidović and A. Ware. *Asymptotic Behaviour of Mean-Quantile Efficient Portfolios*. Finance and Stochastics, Vol. 10, No. 4, 2006.

[14] G. Dmitrašinović-Vidović and A. Ware. *Optimal portfolios of mean-reverting instruments*. SIAM Journal of Financial Mathematics, Vol 2, No 1, pp 748-767, 2011.

[15] JF. Pourbabaee, M. Kwak, T. A. Pirvu. *Risk minimization and portfolio diversification*. Quantitative Finance, Vol 16, No 9, pp 1325–1332, 2016.

[16] J. C. Hull. *Options, Futures, & Other Derivatives*. Prentice Hall, Upper Saddle River, 1999.

[17] J. S. E. Shreve. *Stochastic calculus for finance II: Continuous-time models*. Springer Science & Business Media, 2004.

[18] A. Hirsa. *Computational methods in finance*. CRC Press, 2012.

[19] A. Cartea, T. Williams. *UK Gas Markets: The market price of risk and applications to multiple interruptible supply contracts*. Energy Economics, Vol 30, pp 829-846, 2008.

[20] E.S. Schwartz, J. E. Smith. *Short-Term Variations and Long-Term Dynamics in Commodity Prices*. Management Science, Vol 46, No 7, pp 893-912, 2000.

[21] A. Lari-Lavassani, A. A. Sadeghi and A. Ware. *Modeling and Implementing Mean Reverting Price Processes in Energy Markets*. Electronic Publications of the International Energy Credit Association, 2001.

[22] D. Pilipovic. *Energy Risk*. McGraw-Hill Professional, 2007.

[23] R. Korn, and E. Korn. *Option Pricing and Portfolio Optimization*. Graduate Studies in Mathematics. V 31. AMS. Providence, 2000.

[24] G. Dmitrašinović-Vidović, A. Lari-Lavassani, X. Li, and A. Ware. *Continuous Time Portfolio Selection under Conditional Capital at Risk*. Journal of probability and statistics, special issue, Actuarial and Financial Risks: Models, Statistical Inference, and Case Studies, Vol 2010, pp Article ID 976371, 2010.

[25] G. Dmitrašinović-Vidović. *Portfolio Selection Under Downside Risk Measures*. PhD Thesis. University of Calgary, 2004.

[26] R. Korn. *Optimal Portfolios*. World Scientific, Singapore, 1997.

[27] I. Karatzas and S. E. Shreve. *Methods of Mathematical Finance*. Springer-Verlag, New York, 1999.

[28] R. T. Rockafellar. *Convex Analysis*. Princeton University Press. Princeton, 1970.

[29] M. S. Bazaraa and C. M. Shetty. *Nonlinear Programming*. John Wiley & Sons, New York, 1979.

Appendix A

Other Discretization Methods for OU Process

A.1 Euler-Maruyama Method

The Euler-Maruyama method is a commonly used numerical method to approximate the solution of stochastic differential equations. It's a simple generalization of the Euler method for ordinary differential equation to stochastic differential equation. For an Ornstein-Uhlenbeck process

$$dX_t = \lambda(\mu - X_t)dt + \sigma dW_t,$$

its Euler-Maruyama discrete form is

$$X_{t_{n+1}} = X_{t_n} + \lambda(\mu - X_t)\Delta t + \sigma \Delta W_t$$

where $\Delta t = t_{n+1} - t_n$, W_t is a standard Brownian motion and its increments $\Delta W_t = W_{t_{n+1}} - W_{t_n}$ are independent and each of these increments is normally distributed with

$$\mathbb{E}[W(t_{n+1}) - W(t)_n] = 0,$$

$$\mathrm{Var}[W(t_{n+1}) - W(t_n)] = t_{n+1} - t_n = \Delta t.$$

Then

$$\begin{aligned} X_{t_{n+1}} &= X_{t_n} + \lambda(\mu - X_t)\Delta t + \sigma \Delta W_t \\ &= X_{t_n} + \lambda(\mu - X_t)\Delta t + \sigma\sqrt{\Delta t} z, \end{aligned}$$

where z follows a standard normal distribution with mean zero and variance one, i.e.,$z \sim N(0, 1)$.

Appendix B

QR Decomposition and Cholesky Decomposition

B.1 QR Decomposition

Theorem B.1. *Let $\boldsymbol{A}$ be a real non-singular matrix, there exists a unique matrix pair $(\boldsymbol{Q}, \boldsymbol{R})$ such that*

$$\boldsymbol{A} = \boldsymbol{QR}$$

where $\boldsymbol{Q}$ is an orthogonal matrix and $\boldsymbol{R}$ is an upper triangular matrix whose diagonal entries are positive.

QR decomposition is often referred to as QR factorization. Compared to the direct matrix inverse, QR factorization has its own advantage when solving linear systems such as numerical stability. It can be shown that QR factorization reduces the condition numbers which gives a bound on how inaccurate the solution $\boldsymbol{x}$ will be after approximation. Because $\boldsymbol{Q}$ is an orthogonal matrix, it satisfies that $\boldsymbol{Q^{-1}} = \boldsymbol{Q'}$. If we have a linear system $\boldsymbol{Ax} = \boldsymbol{b}$, using QR factorization we get $\boldsymbol{QRx} = \boldsymbol{b}$, let $\boldsymbol{Rx} = \boldsymbol{y}$, then solving the linear system is equivalent to solving $\boldsymbol{Qy} = \boldsymbol{b}$ and $\boldsymbol{Rx} = \boldsymbol{y}$. Therefore $\boldsymbol{y} = \boldsymbol{Q'b}$ and since $\boldsymbol{R}$ is an upper triangular matrix, we can use back substitution to find the solution without explicitly in-

verting $\boldsymbol{R}$, this can make the entire process quickly and accurately.

Moreover, QR decomposition can be extended to rectangular matrix. If $\boldsymbol{A}_{n\times m}$ is a rectangular matrix, with $n \neq m$, then there exists $\boldsymbol{Q}_{n\times m}$ and $\boldsymbol{R}_{n\times m}$ such that $\boldsymbol{A} = \boldsymbol{QR}$, where $\boldsymbol{Q}$ is an orthogonal matrix, $\boldsymbol{R}$ is an upper triangular matrix with positive diagonal entries.

B.2 Cholesky Decomposition

Theorem B.2. *Let $\boldsymbol{A}$ be a real symmetric positive definite matrix, it has a unique Cholesky decomposition*

$$\boldsymbol{A} = \boldsymbol{BB}'$$

where matrix $\boldsymbol{B}$ is a real lower triangular matrix with positive diagonal entries.

The Cholesky decomposition, which is also referred to as Cholesky factorization, is mainly used for numerically solving the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ where A is symmetric and positive definite. Since $\boldsymbol{A} = \boldsymbol{BB}'$, let $\boldsymbol{B}'\boldsymbol{x} = \boldsymbol{y}$, then solving the linear system is equivalent to solving $\boldsymbol{By} = \boldsymbol{b}$ for $\boldsymbol{y}$ by forward substitution and $\boldsymbol{B}'\boldsymbol{x} = \boldsymbol{y}$ for $\boldsymbol{x}$ by back substitution.

When solving linear systems in a symmetric form, the Cholesky decomposition is recommended because it is more numerically stable and efficient. If the matrix $\boldsymbol{A}$ is symmetric positive semi-definite, then it still has a decomposition of the form $\boldsymbol{A} = \boldsymbol{BB}'$ if the diagonal entries of $\boldsymbol{B}$ are allowed to be zero. However, the decomposition need not be unique.

B.2.1 LDL decomposition

There is a closely related variant of the classical Cholesky decomposition, it is called LDL decomposition. The LDL decomposition of a real symmetric positive definite matrix $\boldsymbol{A}$ is a decomposition of the form

$$\boldsymbol{A} = \boldsymbol{LDL}'$$

where L is a lower triangular matrix having real and positive diagonal entries, and D is a diagonal matrix.

Given the classical Cholesky decomposition $\boldsymbol{B}$, we can find $\boldsymbol{LDL'}$ form. Using the property that the diagonal of $\boldsymbol{L}$ must be 1 and that both the Cholesky and the $\boldsymbol{LDL'}$ form are lower triangles, and let $\boldsymbol{C}$ be a diagonal matrix that contains the main diagonal of $\boldsymbol{B}$, then

$$\boldsymbol{D} = \boldsymbol{C}^2,$$

$$\boldsymbol{L} = \boldsymbol{B}\boldsymbol{C}^{-1}.$$

Or if we have the LDL decomposition, the corresponding Cholesky decomposition can be found as follows:

$$\boldsymbol{A} = \boldsymbol{LDL'} = \boldsymbol{L}\boldsymbol{D}^{\frac{1}{2}}(\boldsymbol{D}^{\frac{1}{2}})'\boldsymbol{L'} = \boldsymbol{L}\boldsymbol{D}^{\frac{1}{2}}(\boldsymbol{L}\boldsymbol{D}^{\frac{1}{2}})' = \boldsymbol{B}\boldsymbol{B'}$$

If efficiently implemented, the LDL decomposition avoids extracting square roots compared with Cholesky decomposition. And for some indefinite matrices which do not have Cholesky decomposition, there exists an LDL decomposition with negative entries in $\boldsymbol{D}$. For these reasons, the LDL variant may be preferred.

Appendix C

Least Square Method

Assume x_n satisfied the equation below

$$x_{n+1} = a + bx_n + cz_n,\ z_n \sim N(0,1),\ n = 0, ..., N$$

Let

$$\boldsymbol{A} = \begin{pmatrix} 1 & x_0 \\ 1 & x_1 \\ ... & ... \\ 1 & x_n \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_{n+1} \end{pmatrix}, \boldsymbol{x} = \begin{pmatrix} a \\ b \end{pmatrix},$$

and $\boldsymbol{r}$ represents the residual. Then the least square problem can be written in the matrix form as

$$\boldsymbol{Ax} \approx \boldsymbol{b}$$

If we assume that the rank of matrix $\boldsymbol{A}$ is n, the best-fit solution is obtained by minimize the 2-norm of the residual $\boldsymbol{r}$:

$$\|\boldsymbol{r}\|_2^2 = \boldsymbol{r}'\boldsymbol{r},$$

where

$$\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{Ax}.$$

To minimize it, we first rewrite the 2-norm of residual

$$\|\boldsymbol{r}\|_2^2 = \boldsymbol{r}'\boldsymbol{r} = (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})'(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}) = \boldsymbol{b}'\boldsymbol{b} - 2\boldsymbol{x}'\boldsymbol{A}'\boldsymbol{b} + \boldsymbol{x}'\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}.$$

We take the derivative with respect to $\boldsymbol{x}$ and set it to zero:

$$2\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{A}'\boldsymbol{b} = 0$$

which can be simplified to

$$\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}'\boldsymbol{b}$$

$\Rightarrow$

$$\begin{pmatrix} 1 & 1 & ... & 1 \\ x_0 & x_1 & ... & x_n \end{pmatrix} \begin{pmatrix} 1 & x_0 \\ 1 & x_1 \\ ... & ... \\ 1 & x_n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & 1 & ... & 1 \\ x_0 & x_1 & ... & x_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_{n+1} \end{pmatrix}$$

$\Rightarrow$

$$\begin{pmatrix} \sum_{n=1}^{N} 1 & \sum_{n=1}^{N} x_{n-1} \\ \sum_{n=1}^{N} x_{n-1} & \sum_{n=1}^{N} x_{n-1}^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{n=1}^{N} x_n \\ \sum_{n=1}^{N} x_n x_{n-1} \end{pmatrix}.$$

Now we notice that the linear system above is exactly the same as the one we get using maximum likelihood estimation. The equation above is known as the the system of normal equation. Apparently, the normal equations system is obtained by multiplying $\boldsymbol{A}'$ to the original equations system. $\boldsymbol{A}'\boldsymbol{A}$ is an $n \times n$ matrix with rank n, so it is a symmetric non-singular matrix, while $\boldsymbol{A}'\boldsymbol{b} \in \mathbb{R}^n$, so there is a unique solution to the normal equation. Moreover $\boldsymbol{A}'\boldsymbol{A}$ is SPD (symmetric positive definite), so one can compute the Cholesky factorization

$$\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}'$$

where $\boldsymbol{L}$ is an $n \times n$ lower triangular matrix. So the original problem is solved as

$$\boldsymbol{L}\boldsymbol{y} = \boldsymbol{A}'\boldsymbol{b},\ \boldsymbol{L}'\boldsymbol{x} = \boldsymbol{y}$$

Then we are able to calculate the residual

$$\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}.$$

Since $z_n \sim N(0,1),\ n = 0, ..., N$, so c equals to the standard deviation of $\boldsymbol{r}$.