Review

Improved Covariance Matrix Estimation for Portfolio Risk Measurement: A Review

Ruili Sun 1, Tiefeng Ma 2, Shuangzhe Liu 3 and Milind Sathye 4,*

1 Zhongyuan Bank Postdoctoral Programme, Zhongyuan Bank, Zhengzhou 450000, China; sunruili2009@163.com
2 School of Statistics, Southwestern University of Finance and Economics, Chengdu 611130, China; matiefeng@swufe.edu.cn
3 Faculty of Science and Technology, University of Canberra, Canberra 2601, Australia; shuangzhe.liu@canberra.edu.au
4 Faculty of Business, Government and Law, University of Canberra, Canberra 2601, Australia
* Correspondence: milind.sathye@canberra.edu.au

Received: 17 January 2019; Accepted: 15 March 2019; Published: 24 March 2019

Abstract: The literature on portfolio selection and risk measurement has considerably advanced in recent years. The aim of the present paper is to trace the development of the literature and identify areas that require further research. This paper provides a literature review of the characteristics of financial data, commonly used models of portfolio selection, and portfolio risk measurement. In the summary of the characteristics of financial data, we summarize the literature on fat tail and dependence characteristic of financial data. In the portfolio selection model part, we cover three models: mean-variance model, global minimum variance (GMV) model and factor model. In the portfolio risk measurement part, we first classify risk measurement methods into two categories: moment-based risk measurement and moment-based and quantile-based risk measurement. Moment-based risk measurement includes time-varying covariance matrix and shrinkage estimation, while moment-based and quantile-based risk measurement includes semi-variance, VaR and CVaR.

Keywords: portfolio selection; risk measure; fat tail; Copula; shrinkage; semi-variance; CVaR

1. Introduction

This paper is motivated by three stylized facts about the operation of real-world financial markets. First, as real-world financial data are asymmetric and fat-tailed, the return series cannot be approximated by normal distribution. Second, financial time series are marked by volatility clustering and last, dependence structure of multivariate distribution is required to model such data and the model needs to be flexible enough to accommodate different types of financial data. Given these characteristics of financial data, many portfolio selection and risk measurement models have been developed to account for such data. Interestingly, we have not come across literature that reviews these developments in recent years. The present research would help fills this important gap. Accordingly, the aim of this paper is to review of development of the literature in the above areas and identify the directions for future research.

As already stated, financial data are known to exhibit some unique characteristics such as fat tails (leptokurtosis), volatility clustering and possible asymmetry. When the tails of distribution have a higher density than that expected under conditions of normality, it is known as fat tailed data distribution. It is ‘a distribution that has an exponential decay (as in the normal) or a finite endpoint is considered thin tailed, while a power decay of the density function in the tails is considered a fat tailed
distribution’ (LeBaron and Samanta 2004, p. 1). As financial data typically exhibit asymmetry and fat tails, the Gaussian distribution cannot adequately represent it. Consequently, alternative parametric distributions that can account for skewness and fat tails have been suggested in the literature. Over the years, the fat tail phenomenon and the various methods used to capture the characteristics of the fat tail of financial data has generated considerable interest among researchers. Similarly, complex dependency patterns such as asymmetry or dependence in the extremes are found in the financial data. The full characteristics of such data cannot be adequately captured by multivariate Gaussian distributions given that it cannot model extreme events. The multivariate Student’s $t$ and its skewed version could be valid alternatives but also have some disadvantages as outlined by Bauwens and Laurent (2005) and others. Consequently, for such data increasingly the copula approach is being used. It not only can describe the dependence characteristics but also can be combined with other distributions such as Student $t$ distribution to describe fat tails.

We proceed as follows: Section 2 provides an overview of the above characteristics of financial data, Section 3 reviews the literature on portfolio selection models, Section 4 reviews the literature on factor models, Section 5 is devoted to portfolio risk measurement literature review and Section 6 provides directions for future research and conclusions of the study.

2. Fat tail of Financial Data and Data Dependence

In this section, we review the literature on fat tails and data dependence.

2.1. The Concept of Fat Tails

Mandelbrot (1963) first introduced the concept of fat tails in mathematical finance to describe cotton price changes. It was followed by many econometric studies devoted to the quest for suitable classes of models that capture the essential statistical properties of stock and stock index returns, for example, Heyde et al. (2001), McAleer (2005), Liu and Heyde (2008) and Tsay (2010), among many others.

The fat tail distribution may have more than one definition, as there is no universal definition for the term tail in the first place. It generally refers to a probability distribution with a tail that looks fatter than usual or the normal distribution. A good example may be the Student $t$ distribution which is a fat tailed distribution and exhibits tails that are fatter than the normal. It is also a leptokurtic distribution which has excess positive kurtosis as illustrated in Figure 1.

Figure 1. Student $t$ distribution is leptokurtic and has a fatter tail when compared to a standard normal distribution.
Some researchers consider that a fat tail distribution refers to a subclass of heavy tailed distributions that exhibit power law decay behavior as well as infinite variance. One example may be a distribution $X$ defined with a fat right tail by $P(X > x) \sim x^{-\alpha}$ as $x \to \infty$, where $P$ is the probability for the cumulative distribution, $\alpha > 0$ is a (small) constant and referred to as the tail index, and the tilde notation “~” is used to mean that there exists some finite value of $x$ above which the probability distribution follows the right-hand side of the expression, that is, asymptotically the tail of the distribution decays like a power law (for more, see e.g., Kausky and Cooke (2009). It may be noted that a Student $t$ distribution can be considered to be fat tailed by rewriting its cumulative or density function form, and it has finite variance if the degrees of freedom is larger than 2. In addition, a fat-tailed distribution is heavy tailed, although not every heavy tailed distribution has a fat tail. For example, the Weibull distribution is heavy-tailed but not fat-tailed.

Empirically, we present two figures as an example for real-world data. The real data is the daily Tableau Software close price data from 15 March 2018 to 14 March 2019 collected via Yahoo! Finance at https://au.finance.yahoo.com/quote/DATA/history?p=DATA. Figure 2 shows its time series plots, with the upper panel for the simple returns and lower panel for the log returns. Figure 3 compares the empirical Tableau densities with normal densities and shows that the Tableau data reveals obviously a fat tail and high peak.

The initial studies that followed the seminal work of Mandelbrot (1963) used the stable Pareto distribution to simulate the fat tail of such data. Fama (1965), Fama and Roll (1968) also used stable Pareto distribution to study the fat tail characteristics of financial data. Such distributions have many properties exhibited by normal distribution such as closeness under summation. However, Upton and Shannon (1979) as well as Friedman and Vandersteel (1982) claimed that a stable Pareto distribution was inappropriate for simulating the fat tail shape of financial data because the return was more peaked and had fatter tails. Ghose and Kroner (1995) found that the GARCH model and the stability model had something in common, which meant that many of the discoveries of the stable distributions with fat tail in finance were caused by temporary volatility aggregation.
Also, Chen and Yu (2013) proposed a novel nonlinear VaR method to model the risk of option portfolio. A computational technique of functional integrals and Feynman diagrams borrowed from particle physics was used. Most of the empirical applications of the stochastic volatility (SV) model assume that the conditional distribution of returns, given the latent volatility process, is normal. Liesenfeld and Jung (2000) used German stock data to compare stochastic volatility model based on conditional normal distribution and conditional fat tail distribution. These conditional fat tail distributions were mainly Student $t$ distributions and generalized error distributions. Cont (2001) presented a set of stylized empirical facts (including fat tail) emerging from the statistical analysis of price variations in various types of financial markets and analyzed how these stylized empirical facts invalidated many of the common statistical approaches. Chib et al. (2002) discussed a class of generalized stochastic volatility models defined by the horizontal effects of fat tails, fluctuations, observational and evolutionary equations, and the covariate effects of the jumping part of the observational equation and provided two Markov Chain Monte Carlo (MCMC) fitting algorithms for the above models. In addition, simulation-based inference in generalized models of stochastic volatility was considered.

Zhou (2002) used the multivariate normal mixture model to characterize the fat tail characteristics of market risk factors, examined the relationship between risk and return, and established an asset pricing model with fat tail characteristics excluding options. This model provided a new perspective to study asset pricing. Wong et al. (2009) proposed Student $t$ mixture autoregressive model which is also able to capture serial correlations, time-varying means and volatilities, and the shape of the conditional distributions can be time varied from short-tailed to long-tailed, or from unimodal to multimodal. Also, Chen and Yu (2013) proposed a novel nonlinear VaR method to model the risk of option portfolio under fat tailed market risk factors. Multivariate mixture of normal distributions was used to depict the heavy-tailed market risk factors.

Glasserman (2004) used multivariate $t$ distribution to characterize the risk factors of fat tailed market, and indirectly obtained an expression of closed moment generating function. This expression reflects the change of portfolio value when the fat tailed problem was transformed into thin tailed problem. On this basis, the moment generating function was obtained by using the structure of multivariate $t$ distribution. Albanese et al. (2004) used multivariate $t$ distribution to characterize the fat tail of market risk factors. First, the matrix transformation of option portfolio value was derived from Delta-Gamma-Theta model. Thereafter, the density function of option portfolio value was discretized. Finally, the approximate VaR value was calculated by Fourier inverse transformation.

Figure 3. A comparison of simple return and log return of daily Yahoo finance densities (blue) with normal densities (green). The above figure compares the empirical Tableau densities with normal distributions and shows that the Tableau data reveals obviously a fat tail and high peak.
and linear interpolation. This new method does not assume that the characteristic function for the return model is known explicitly. Considering the difference between multivariate normal distribution and multivariate $t$ distribution in the description of market risk factors, Albanese and Campolieti (2006) proposed the probability density function for calculating the change of option portfolio value and the Monte Carlo simulation method for estimating the multivariate VaR at a given confidence level and explored the relationship between a normal distribution and a fat tail distribution. Like Glasserman (2004), Siven et al. (2009) deduced the closed expression of moment generating function in the case of multivariate $t$ distribution and compared Fourier-Inversion method with Monte Carlo simulation method. The results showed that the Fourier-inversion method was much quicker than Monte Carlo simulation method and that Fourier-Inversion was a good way to calculate option VaR. Asai (2008) studied two models for describing fat tail and volatility dependence: autoregressive stochastic volatility model with Student $t$ distribution (ARSV-T) and multifactor stochastic volatility (MFSV) model, and the results showed that ARSV-T model provided a better fit than MFSV model based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

Asai (2009) compared stochastic volatility models defined by normal distribution and other fat tailed distributions, such as Student $t$ distribution and generalized error distribution. Delatola and Griffin (2013) proposed a Bayesian semiparametric stochastic volatility model, and this model allowed the distribution of returns to be fat tail and allowed the correlation between returns and fluctuations. Abanto-Valle et al. (2015) proposed a stochastic volatility model which assumed that the return followed the biased Student $t$ distribution. This model could flexibly control the skewness and the fat tail distribution of the return condition. Meanwhile, an effective MCMC algorithm was given to estimate and predict the parameters. Lafosse and Rodríguez (2018) combined stochastic volatility model with GH Skew Student $t$ distribution to characterize the skewness and fat tail of financial data and showed the evidence of asymmetries and heavy tails of daily stocks returns data. Gunay and Khaki (2018) noted that capturing conditional distributions, fat tails and price spikes was the key to measuring risk and accurately simulating and predicting the volatility of energy futures. These researchers tried to model the volatility of energy futures under different distributions.

2.2. The Dependence of Financial Data

Financial data are usually interdependent. They also exhibit a tendency of volatility clustering, that is, temporal dependence, in which large financial returns are followed by large financial returns. The interdependence of financial data has been extensively researched in various fields of finance, for example, following the US stock market crash of 1987, the contagion spread to the UK and other developed countries (King and Wadhwani 1990). To model the non-linear dependence in data, following Engle (1982), many ARCH-type models have been proposed. However, assumption of iid in such models makes their use inappropriate to model non-linear dependence in a univariate series or simultaneous dependence in two or more timeseries. Copula comes to rescue here. Some important publications are listed in the following table.

Copula proposed by Sklar (1959) can identify the dependency structure, capture the potential nonlinear correlation, and fit the dependency of financial data well, which makes it a good choice to measure correlation (Embrechts 1999). Copula refers to “functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions” (Nelsen 1999, p. 1). The copula decomposes an n-dimensional distribution function in to the marginal distribution functions and the dependence part. It is the latter that the copula describes. In Table 1 below some key research work on copula has been included.
Table 1. Selected work on Copula.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Paper/Book/Thesis Title (Please See References for Details)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sklar</td>
<td>1959</td>
<td>Fonctions dérépartition à n dimensions et leurs marges</td>
</tr>
<tr>
<td>Joe</td>
<td>1997</td>
<td>Multivariate Models and Dependence Concepts</td>
</tr>
<tr>
<td>Embrechts</td>
<td>1999</td>
<td>An Introduction to Copulas</td>
</tr>
<tr>
<td>Mashal and Zeevi</td>
<td>2002</td>
<td>Beyond correlation: Extreme co-movements between financial assets</td>
</tr>
<tr>
<td>Van den Goorbergh et al.</td>
<td>2005</td>
<td>Bivariate option pricing using dynamic copula models</td>
</tr>
<tr>
<td>Kole et al.</td>
<td>2007</td>
<td>Selecting Copulas for risk management</td>
</tr>
<tr>
<td>Hafner and Reznikova</td>
<td>2010</td>
<td>Efficient estimation of a semiparametric dynamic copula model</td>
</tr>
</tbody>
</table>

“Modern risk management calls for an understanding of stochastic dependence going beyond simple linear correlation” (Embrechts et al. 2001). These researchers emphasized the necessity to use Copula to simulate multivariate correlations in financial data given its stochastic dependence and pitfalls. Mashal and Zeevi (2002) showed that Student $t$ Copula had an advantage over other multivariate Copulas in fitting financial data. Kole et al. (2007) used goodness-of-fit test for $t$, Gaussian and Gumbel Copula in risk management of linear assets and found that $t$ Copula had more advantages than Gaussian and Gumbel Copula. Sak et al. (2010) used a flexible and accurate model, such as $t$ Copula dependency structure and generalized hyperbolic distribution, to simulate logarithmic returns. They also calculated the tail probability of the current asset portfolio.

Studies have shown evidence of two types of asymmetries in the joint distribution of stock returns: skewness in the distribution of individual stock returns and an asymmetry in the dependence between stocks. Patton (2004) showed that the rotational Gumbel Copula function was superior than the normal and the Student $t$ Copula, in describing the asymmetric dependency structure of two stock indexes. Trivedi and Zimmer (2006) considered the use of the copula approach for a model with three jointly determined outcomes. The model could handle the discrete case in which outcomes include a mixture of dichotomous choices and discrete count data. They applied this technique to study self-selection and interdependence between health insurance and health care demand among married couples. Hu (2006) proposed a hybrid Copula+ model to capture different types of dependent structures, in which the marginal distribution of each market asset was estimated by nonparametric method and the mixed Copula was estimated by quasi-maximum likelihood method. Patton (2006) extended the Copula theory to allow conditional variables, analyzed two important exchange rates using different forms of Copula-GARCH model, and used them to construct a flexible model of conditional dependency structure of these exchange rates. Liu and Luger (2009) adapted and examined an iterative (fixed-point) algorithm for the maximum-likelihood estimation of copula-based models that circumvents the need to compute second-order derivatives of the full-likelihood function. The algorithm exploits the natural decomposition of a potentially complicated likelihood function: the first part is a working likelihood that only involves the parameters of the marginals and the residual part is used to update estimates from the first part.

Van den Goorbergh et al. (2005) studied the price problem of Binary Options with correlations between assets and used the parameter Copula family with multiple alternative Gaussian dependency structures to fit the correlation. The relationship was assumed to be a function of asset volatility and it changed with time. Since then, time-varying Copula model has been extensively studied. Bartram et al. (2007) used the time-varying Copula model to study the effect of the introduction of the euro on the dependence between 17 European stock markets from 1994 to 2003. The time-varying Copula model used the GJR-GARCH-T model to realize the marginal distribution and Gauss Copula to realize the joint distribution. These could capture the time-varying nonlinear correlations. The correct modeling of non-Gaussian dependences is a key issue in the analysis of multivariate time series, Giacomini et al. (2009) used copula functions with adaptively estimated time-varying parameters for modeling the distribution of returns and applied it to the portfolio VaR. Hafner and Reznikova (2010) proposed a new semiparametric dynamic Copula model in which the marginal of Copula was assigned as a
parameter GARCH-type process while the dependent parameters of Copula could change with time in a nonparametric manner. Negative extreme changes were common in international stock markets, Garcia and Tsafack (2011) pointed out the limitations of some common methods and proposed a regime-switching Copula model, which included a normal system with symmetric dependencies and an asymmetric dependency. The system was applied to allow changes in the market between the international stock and bond markets. Hafner and Manner (2012) proposed a dynamic Copula model in which dependent parameters followed an autoregressive process. Since this kind of model includes Gaussian Copula with stochastic correlation process, it can be regarded as a generalization of multivariate stochastic volatility model. Mendes and Marques (2012) found that the dependency structure between assets was not only linear but also used robust estimation of dual Copula model to fit logarithmic returns. Chen and Tu (2013) used four different types of time-varying Copula to fit the index futures and spot returns by relaxing the traditional normal joint distribution hypothesis and improved the hedging portfolio VaR. Creal and Tsay (2015) constructed a series of Copula families with time-varying dependent parameters by writing Copula as a factor model with random loads.

The Vine Copula (Joe 1997) has great advantages in describing the relationship between multiple financial assets and is widely used in financial risk management. Maugis and Guegan (2010) compared the Vine Copula method with several traditional GARCH models and concluded that the Vine Copula method could give better portfolio VaR prediction. As distinct from the existing Vine Copula structure strategy, Djißmann et al. (2013) proposed automatic Copula selection and estimation technology based on graph theory. It enabled flexible modeling of complex dependencies, that is, even those with larger dimensions. So and Yeung (2014) discussed the construction of Vine Copula structure and studied the relationship between financial markets and Vine Copula theory. Geidosch and Fischer (2016) confirmed the advantages of Vine Copulas over traditional Copula in simulating the dependent structure of credit portfolios. Aiming to measure risk and finding the optimal weights of portfolios containing three financial instruments, Pastpipatkul et al. (2018) used C-D vine Copulas method to establish the dependence relationship of each pair of financial instruments and used Monte Carlo simulation technology to generate simulation data to calculate risk value (VaR) and expected shortfall.

3. Portfolio Selection: A Review of Common Models

A commonly used model for portfolio selection is the mean-variance model, in which, the optimal portfolio weight depends on the mean and covariance matrix of asset returns. Usually, the available portfolio weights are obtained by using sample mean and sample covariance matrix to replace the true mean and covariance matrices of asset returns respectively. However, the estimation of sample means, and covariance usually involves errors and the estimation errors in sample mean are much larger than those in sample covariance. This makes the mean-variance model more sensitive to estimation errors. Therefore, the global minimum variance models whose optimal portfolio weight only depends on the covariance matrix is used. Furthermore, for the measurement of portfolio risk, the factor model is commonly used to estimate covariance matrix.

Accordingly, the literature on portfolio selection has developed in three strands: the traditional mean-variance model and the newer, global minimum variance model and the factor model.

3.1. Mean-Variance Model

In the traditional mean-variance model proposed by Markowitz (1952), the return of financial assets is represented by a random variable with Gaussian distribution. The assumption of normal (Gaussian) distribution means that the return of assets depends only on the mean and variance. Markowitz (1959) extended the mean-variance model in a pioneering book on portfolio selection. Merton (1972) studied the application of the mean-variance model allowing short-sale in portfolio selection. Over the years, many studies have used the above models. Some groundbreaking articles are summarized in Table 2 below.
Table 2. Selected works on Portfolio Selection.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Paper/Book/Thesis Title (Please See References for Details)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Markowitz</td>
<td>1952</td>
<td>Portfolio Selection</td>
</tr>
<tr>
<td>Samuelson</td>
<td>1969</td>
<td>Lifetime portfolio selection by dynamic stochastic programming</td>
</tr>
<tr>
<td>Merton</td>
<td>1969</td>
<td>Life time portfolio selection under uncertainty: The continuous-time case</td>
</tr>
<tr>
<td>Pogue</td>
<td>1970</td>
<td>An extension of the Markowitz portfolio selection model to include variable transactions’ costs, short sales, leverage policies and taxes</td>
</tr>
<tr>
<td>Merton</td>
<td>1972</td>
<td>An analytic derivation of the efficient portfolio frontier</td>
</tr>
<tr>
<td>Fernández and Gómez</td>
<td>2007</td>
<td>Portfolio selection using neural networks</td>
</tr>
</tbody>
</table>

Markowitz’s traditional mean-variance model is a static model in which investors can only make investment decisions at the beginning of the investment period and then wait until the end of the investment period. Based on this, the mean-variance model was later extended to the multi-period case. Samuelson (1969) proposed a discrete time multi-period consumption-investment model to maximize the end-of-term expected utility for investors. Grauer and Hakansson (1993) compared the effects of mean-variance asymptotic and quadratic asymptotic in a discrete time dynamic investment model. Yi et al. (2008) used the mean-variance model to consider the discrete time portfolio optimization of asset liability management under uncertain investment level and deduced the analytical optimal strategy by using embedding technology. Wu and Li (2011) studied the discrete time mean-variance portfolio model with regime switching under the assumption of stochastic cash flow.

Merton (1969, 1971) studied the maximized expected return of continuous time model under a given planning period, which is a pioneering work of continuous time research. Karatzas et al. (1987) considered a generalized consumption-investment model with a single member, which aimed to maximize the linear combination of the total expected discount utility and the end-of-term wealth utility from the consumption over a continuous investment period. Li and Ng (2000) firstly used embedding technology to solve the problem of inseparability and constructed a framework with mean-variance model in a discrete case. Xie et al. (2008) used the stochastic optimal linear-quadratic control technique to obtain the optimal dynamic strategy for continuous time mean-variance portfolio selection in incomplete markets. Using dynamic programming and embedding techniques, the closed form optimal strategy and efficient frontier were derived. Xu and Wu (2014) studied the continuous time mean-variance portfolio selection problem with inflation in incomplete markets and obtained the efficient bounds of dynamic optimal strategy and mean-variance model. Wu and Chen (2015) studied the time-consistent multi-period mean-variance portfolio selection problem under the assumption that risk aversion was dynamically dependent on market conditions.

Pogue (1970) first gave a description of the mean-variance portfolio problem in the presence of transaction costs. Davis and Norman (1990) further explored the portfolio selection problem under proportional transaction costs. Dumas and Luciano (1991), Morton and Pliska (1995) studied portfolio selection with proportional transaction costs and fixed transaction costs, respectively. Yoshimoto (1996) first assumed that the transaction cost was a V-shaped function, and then obtained the optimal portfolio strategy. Oksendal and Sulem (2002) studied the optimal consumption and portfolio under fixed and proportional transaction costs, with the objective of maximizing cumulative consumption expected utility within the scope of planning. Xue et al. (2006) constructed a mean-variance portfolio selection model with concave transaction costs to capture real market conditions. The authors provide a branch and bound algorithm as a solution. Dai and Zhong (2008) proposed a numerical penalty method to solve the continuous time portfolio selection problem with proportional transaction costs. Peng et al. (2011) studied portfolio optimization with quadratic transaction costs in the framework of the mean-variance model. Wang and Liu (2013) studied the multi-period mean-variance portfolio selection problem with fixed and proportional transaction costs and defined the indirect utility function to solve the problem by using dynamic programming and Lagrange multiplier. Liagkouras and
Metaxiotis (2018) proposed a new multi-period fuzzy portfolio optimization algorithm for multistage mean-variance fuzzy portfolio optimization with transaction costs. To make the model more practical, different constraints were introduced in the mean-variance model. Fernández and Gómez (2007) generalized the standard mean-variance model including cardinality and boundary constraints, and the constraints guaranteed investment in a given set of different assets and limited the amount of capital invested in each asset. Soleimani et al. (2009) proposed a portfolio selection model based on mean-variance model framework, which included cardinality constraints, minimum trading lot sizes and market (sector) capitalization. Castellano and Cerqueti (2014) studied the mean-variance optimal portfolio selection problem for risky assets with low-frequency trading and low liquidity. To simulate the dynamics of illiquid assets, pure-jump processes were introduced, which enabled the development of portfolio selection models in mixed discrete/continuous time settings.

Simaan (2014) provided a framework that allowed performance comparisons of within and out of sample between mean-variance portfolios and portfolios that maximize expected utility. To develop the best market timing strategy, Gao et al. (2015) considered the mean-variance dynamic portfolio selection problem with management cost time constraints. Lioui and Poncet (2016) proposed a new portfolio decomposition formula to reveal the economics of investor portfolio selection according to the mean-variance criterion and noted that the number of components of the dynamic portfolio strategy could be reduced to two: the first was to hedge the risk of discounted bonds maturing within the investor’s time limit without preference, while the second was to hedge against time variation in pseudo relative risk tolerance.

3.2. Global Minimum Variance Model

The global minimum variance model (GMV) is a specific optimal portfolio with minimum variance on the effective boundary. Haugen and Baker (1991) used the GMV model to verify whether the capitalization weighted (cap weights) portfolio was an efficient investment as claimed by sponsors of such plans. These researchers found that even assuming informationally efficient capital market and that all investors rationally optimized the relationship between risk and expected return, the portfolio of cap weights was not efficient except under extreme restrictive conditions. Chopra and Ziemba (1993) promoted the use of GMV in finance and pointed out that the error in expected returns was 10 times than the error in variance and covariance. Chan et al. (1999) focused on the GMV portfolio and emphasized that the GMV portfolio performed better than the Markowitz mean-variance model. Jagannathan and Ma (2003) pointed out that the weight of the GMV portfolio should be more stable than that of the standard mean-variance model because the estimation error of the covariance was smaller than that of the mean. Kempf and Memmel (2006) noted that the GMV portfolio could provide better out-of-sample results than the tangent portfolio theory and studied the distribution of portfolio weights under the GMV model. Demiguel and Nogales (2009) claimed that GMV model relied only on covariance matrices and were insensitive to estimation errors. These studies have led to the popularity of GMV in portfolios selection.

Traditional GMV only solves the portfolio weights from the perspective of optimization, but many scholars are interested in the distribution and nature of the portfolio weights of GMV. Under the assumption of normal distribution, Okhrin and Schmid (2006) derived the multivariate density function of GMV portfolio. Clarke et al. (2006) noted that the stock weights on the left of the effective boundary under the minimum variance model were independent of the expected safe return. At this point, the portfolio could be obtained only by using the covariance matrix of stocks without involving the equilibrium expectation or the active forecast return. Bodnar and Schmid (2008) discussed portfolio weights in GMV model under the assumption that returns followed a matrix elliptical contoured distribution. Assuming that securities returns were neither normal nor independent, they found that the stochastic nature of the portfolio in GMV model did not depend on the mean vector and the assumption of the distribution of securities returns. Bodnar and Schmid (2009) derived the variance
and expected return of sample GMV portfolio distribution. Frahm (2010) derived a small sample hypothesis test for global and local minimum variance portfolios and calculated the exact distribution of portfolio weight estimation. At the same time, the first two moments of the estimation of portfolio expected return were given. On the assumption that the conditional distribution of logarithmic returns was normal, Bodnar et al. (2017) considered the weight estimation problem of the optimal portfolio from the perspective of Bayes and obtained the posterior distribution of the weight of GMV portfolio by using the standard prior of mean vector and covariance matrix.

Following the research on distribution and nature of the portfolio weights of GMV, parameter uncertainties, Glombek (2014) analyzed the mean, variance, weight and Sharpe ratio estimators of excess returns of GMV portfolio under consistent and asymptotic distributions, discussed the problem of high-dimensional assumptions and demonstrated the applicability of this method. Maillet et al. (2015) proposed a robust approach to mitigate the effects of parameter uncertainties for a decision maker using GMV strategy to optimize portfolio selection. Based on Taylor’s robust M-estimator and Ledoit-Wolf shrinkage estimator, Yang et al. (2015) proposed a hybrid covariance matrix estimator under the GMV model for portfolios, with outliers of financial data, fat tailed distribution of sample data and obtained a consistent estimate of portfolio risk by minimizing the optimum linear shrinkage strength using random matrix theory. Bodnar et al. (2017) analyzed the GMV portfolio model under the Bayesian framework, adding the prior beliefs of investors to the investment decision. Carroll et al. (2017) evaluated the performance of GMV portfolio strategy and equal weight portfolio strategy under time-varying conditions between assets. They found that conditional correlation is more important than conditional variance in portfolio performance. The also found that frequent asset rebalancing does not help improve portfolio performance. Bodnar et al. (2018) estimated the GMV portfolio in high-dimensional case by using the results of random matrix theory, gave a shrinkage estimator in the sense of non-distribution assumption and minimizing the variance of samples, and obtained the asymptotic properties of the estimator under the assumption of the existence of fourth-order moments.

The mean-variance models are highly data intensive. Consequently, search was on for models that can capture enough of reality but are simpler. This led to the development of factor models.

4. Factor Model

The literature on factor models has evolved in two stages: single factor models and multi-factor models.

A factor model that is linear in form posits that the return of an asset can be expressed by the following equation:

\[ r = a + b_1 f_1 + \cdots + b_k f_k + e, \]

where \( r \) is the return of an asset, \( a, b_1, \ldots, b_k \) are the parameters, \( e \) is the error term. We call it a single factor model when \( k = 1 \), and we call it a multi-factor model when \( k \geq 2 \).

4.1. Single Factor Models

The factor models have drawn attention of researchers, after Sharpe (1963) used, the factor model to estimate the covariance matrix. Geweke (1977) used dynamic factors to analyze the economic time series data and found that the results supported the methodology.

Geweke and Singleton (1981) proposed the theory of identification, estimation and inference in the dynamic confirmatory factor model of economic time series data and pointed out that the dynamic confirmatory factor model could accommodate the important characteristics of prior constraints in the parameter matrix. Watson and Engle (1983) studied the problem of specification and estimation of the dynamic unobserved component model and provided the method of estimating unknown parameters based on score method and EM algorithm by maximizing the likelihood function. Diebold and Nerlove (1989) identified and estimated the univariate ARCH model, and then used the results of
the univariate ARCH model to propose the multivariate latent variable ARCH model. Engle et al. (1990) suggested the use of Factor-ARCH model as a concise structure of conditional covariance matrix of asset excess returns, which made it possible to study the dynamic relationship between asset risk premium and volatility in multivariate systems. Through a variety of diagnostic tests and compared with the previous empirical results, it was shown that the Factor-ARCH model was better as compared to other models given that it had the advantage of stability over time. Lanne and Saikkonen (2007) proposed a multivariate generalized orthogonal factor GARCH model and gave a program to test the correctness of the number of factors. Also, a mixture of Gaussian distributions was considered, and it was found that some parameters of the conditional covariance matrix that were not identifiable under normality could be identified when the mixture specification was used. Cardinali (2012) used orthogonal factors to model the structure of conditional covariance matrices. The advantage of this approach was that the estimated factors could be simulated using a univariate GARCH process, and the model could be extended to multivariate cases.

4.2. Multi-Factor Models

Litterman and Scheinkman (1991) empirically determined the common factors of treasury bond returns based on past securities. The analysis showed that most of the fluctuations of fixed income securities return could be explained by three factors, and the three-factor model was particularly useful for hedging. Chen and Scott (1993) considered that it was necessary to establish a formal theoretical structure model for bonds and other different types of interest rate options. Based on this, a multi-factor equilibrium model was proposed to estimate the parameters driving the interest rate change process and to determine the number of factors necessary to characterize the interest rate structure model. Duffie and Kan (1996) proposed a consistent and arbitrage-free multifactor model of the term structure of interest rates. The model assumed that the returns on a fixed maturity date followed a parametric multivariate Markov diffusive process with stochastic volatility parameters and provided the necessary and sufficient condition for numerical algorithm and the stochastic affine representation of the model. Fama and French (1993) proposed that the return on excess assets could be explained by three factors: sensitivity to market excess returns, market capitalization and book price ratio. Campbell (1996) noted that adding human capital to common factors could improve the performance of multi-factor asset pricing model in predictability of returns. Chan et al. (1999) found that the market, size and stock market value can capture the common structure of the return covariance matrix and found that the use of three factor model for the minimum variance portfolio model was adequate. Stock and Watson (2002) studied predictions with multiple predictors, observations, and a single time series, and noted that a small number of principal component estimators could be used for prediction when data followed an approximate factor model. Bai (2003) used principal component estimator to establish the inference theory of large-dimensional factor model and derived the finite distribution of convergence rate and factor, factor load and common component. Han (2006) studied the effects of time-varying expected return and volatility on asset allocation in a high-dimensional context, and proposed a dynamic factor multivariate stochastic volatility model, which allowed the first two moments of many assets returns to change with time. Adrian and Franzoni (2009) used conditional CAPM models to allow unobserved changes in risk load factors over time. Based on the assumption that investors could rationally learn the long-term level of the load factor from the observed returns, Kalman filter was used to simulate conditional beta. Fan et al. (2013) used the multifactor model to estimate the covariance matrix in the high-dimensional case. Jungbacker et al. (2014) studied the dynamic factor model and showed how to use cubic spline function to smoothly limit the factor load. Hou et al. (2015) put forward a new four factor model by combining market and scale with investment and profitability. Jungbacker and Koopman (2015) presented a new method of dynamic factor model based on likelihood analysis. These researchers used linear dynamic stochastic processes to simulate latent factors and autoregressive processes with correlations to determine the singular perturbation sequence. The method was found to be effective in estimating the factors and maximum likelihood parameters. Fama and French (2015)
added profitability and investment to the three-factor model and constructed a five-factor model to reveal several abnormal phenomena of average returns. Fama and French (2016) used five-factor model to explain the abnormal phenomenon of average returns. Chiah et al. (2016) confirmed that the five-factor model was superior to the multifactor model in explaining the changes of asset returns in global asset market. Stambaugh and Yuan (2017) proposed a four-factor model that combined the two mispricing factors with market and size factors. They found that the ability of the model to account for many anomalies is much better than the earlier models.

Fama and French (2017) used international data to test the five-factor model. The global three factor and five factor models did not perform well in the test of regional portfolio. Therefore, local variables were used to establish the model that is, the factors and returns to be explained came from the same region. Kubota and Takehara (2018) used five factor model (Fama and French 2015) to test whether the model could well explain the pricing structure of Japanese long-term data stocks. They found that the original version of the five-factor model was not the best benchmark pricing model for Japanese data from 1978 to 2014. Roy and Shijin (2018) proposed a balanced six factor asset pricing model, which explained the change of asset returns by adding human capital to the five-factor model and tested the six-factor asset pricing model with four different portfolios. Tu and Chen (2018) developed a new factor-augmented model for calculating the value at risk (VaR) of bond portfolios based on the Nelson-Siegel structural framework and tested whether the information contained in macroeconomic variables and financial stress shocks could enhance the accuracy of VaR prediction.

5. Portfolio Risk Measure

Several methods are available for the measurement of portfolio risk and we divide them into (a) moment-based risk measurement and (b) moment-based and quantile-based risk measurement. The moment-based methods include time-varying covariance matrix and the shrinkage estimation use the covariance matrix in the risk measurement. The semi-variance method calculates the risk below the target value, and the target can be regarded as a quantile. VaR is based on quantile measures of risk, while CVaR is a measure of risk based on the idea of VaR quantile and mean value. Therefore, semi-variance, VaR and CVaR are risk measures based on moments and quantiles.

5.1. Moment-Based Risk Measurement

5.1.1. Time-Varying Covariance Matrix

Some groundbreaking publications are summarized in Table 3 below.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Paper/Book/Thesis Title (Please See References for Details)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engle</td>
<td>1982</td>
<td>Autoregressive conditional heteroscedasticity and estimates of UK inflation</td>
</tr>
<tr>
<td>Engle et al.</td>
<td>1984</td>
<td>Combining competing forecasts of inflation using a bivariate ARCH model</td>
</tr>
<tr>
<td>Bollerslev</td>
<td>1986</td>
<td>Generalized autoregressive conditional heteroscedasticity</td>
</tr>
<tr>
<td>Bollerslev et al.</td>
<td>1988</td>
<td>A capital asset pricing model with time-varying covariances</td>
</tr>
<tr>
<td>Bollerslev</td>
<td>1990</td>
<td>Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model</td>
</tr>
<tr>
<td>Engle and Kroner</td>
<td>1995</td>
<td>Multivariate simultaneous generalized ARCH</td>
</tr>
<tr>
<td>Tse and Tsui</td>
<td>2002</td>
<td>A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations</td>
</tr>
<tr>
<td>Engle</td>
<td>2002</td>
<td>Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models</td>
</tr>
<tr>
<td>McAleer et al.</td>
<td>2008</td>
<td>Generalized autoregressive conditional correlation</td>
</tr>
</tbody>
</table>
Engle (1982) introduced the autoregressive conditional heteroscedasticity (ARCH) family model and used it to estimate the means and variances of inflation in the U.K. The ARCH effect is found to be significant and the estimated variances increased substantially during the chaotic 1970s. Bollerslev (1986) extended the ARCH family model to the Generalized autoregressive conditional heteroskedasticity (GARCH). To capture the dynamic changes of financial markets, time-varying dynamic covariance matrix has been widely used in portfolio investment. Since the GARCH model can successfully describe one-dimensional time-varying variance, many researchers have tried to extend the time-varying variance to the multivariate case by using the multivariate GARCH model.

Bollerslev et al. (1988) used multivariate GARCH (MGARCH) model to estimate the earnings of bills, bonds and stocks. The expected return of bills, bonds and stocks was proportional to the return of each diversified or market portfolio. The results showed that conditional covariance varied greatly over time and this time-varying factor was an important determinant of time-varying risk premium. Kroner and Claessens (1991) used similar technologies to get a series of optimal dynamic hedge funds. Lien and Luo (1994) assessed the multi-period hedging ratio of currency futures in the framework of MGARCH.

Engle et al. (1984) provided a necessary condition for the conditional covariance matrix in the two-dimensional ARCH model to be a positive definite. However, it was not feasible to extend the necessary condition of positive definite conditional covariance to a more generalized model. Bollerslev (1990) suggested that the constant conditional correlations (CCC) MGARCH model could overcome the difficulty of positive definite. Because of the simplicity of calculation, constant conditional correlations MGARCH model has been widely used in practice, but some researchers find that some assumptions of constant correlations MGARCH model are not supported by financial data in practice. Engle and Kroner (1995) gave the formulas and the theoretical results of estimations for multivariate GARCH models in simultaneous equations, proposed a new parameterization method (BEKK) for multivariate ARCH processes, and discussed the equivalent relations of various ARCH parameterizations. At the same time, the sufficiency constraints for the conditional covariance matrix to guarantee positive definiteness were proposed, and the ‘sufficient and necessary conditions’ for the stability of covariance were given.

Bera et al. (1997) pointed out that the BEKK model proposed by Engle and Kroner (1995) did not perform well in estimating the optimal hedging ratio. Also, Lien et al. (2002) pointed out that BEKK was difficult to converge in estimating the conditional variance structure of spot and futures prices. Tsui and Yu (1999) used the dual GARCH model to study stock returns in two emerging markets, Shanghai and Shenzhen of China, and the information matrix test statistic did not support the hypothesis of the constant conditional correlation of stock returns. Tse (2000) introduced the Lagrange Multiplier test to the multivariate GARCH model with constant correlation assumption and this test verified the limitations of the multivariate GARCH model with constant correlation. The data of stock market returns in China was used to draw a conclusion that the correlation was time-varying.

The fact that the correlation is time-varying has been accepted by many scholars. Tse and Tsui (2002) proposed a MGARCH model with time-varying correlation. The conditional covariance matrix could be decomposed into conditional variance matrix and conditional correlation coefficient matrix. Each conditional variance term was assumed to follow a unitary GARCH model, and the conditional correlation coefficient matrix followed a similar autoregressive moving average. At the same time, Engle (2002) proposed a new family of multivariate GARCH models, a dynamic conditional correlation (DCC) model to estimate time-varying correlations.

MGARCH model settings are usually determined by practical considerations such as easy estimation, which often leads to serious losses in general. The deficiencies and developments on the DCC and BEKK models have been extensively reviewed by McAleer (2019a, 2019b). Alexander (2001) proposed an orthogonal GARCH model in which the time-varying covariance matrix was derived from a small number of uncorrelated factors. Weide (2002) proposed a new MGARCH model: the covariance matrix with many parameters could be parameterized with considerable degrees of
freedom and the estimation of parameters was still feasible. This model could be regarded as a natural
generalization of O-GARCH model and nested in a more general BEKK model. To avoid the difficulty
of convergence, the unconditional information was used to make the number of parameters estimated
by conditional information more than half. Vrontos et al. (2003) proposed a new parameter method
for MGARCH with time-varying covariance: the covariance matrix guaranteed positive definiteness
and the number of parameters of the method was relatively small, which could be easily applied to
high-dimensional time series data model. The parameter estimation of multivariate model was realized
by classical Bayesian technique, and the maximum likelihood estimation was realized by Fisher scoring
method. Ledoit et al. (2003) proposed a new method for estimating the time-varying covariance matrix
in the framework of the MGARCH (1,1) model for the diagonal VECH. This method was numerically
feasible in dealing with large-scale problems and could generate semi-definite conditional covariance
matrix without imposing impractical prior restrictions.

Cappiello et al. (2006) studied the existence of asymmetric conditional second-order moments in
international equity and bond yields and analyzed them by the asymmetric version of Engle (2002)
dynamic conditional correlation (DCC) model. A large amount of evidence showed that the series
of national stock index returns indicated strong asymmetric conditional volatility, while there was
little evidence that bond index returns showed such behavior. Mcaleer et al. (2008) established
the generalized autoregressive conditional correlation (GARCC) model on the assumption that the
normalized residual followed the random coefficient vector autoregressive process. The GARCC
model enabled conditional correlation to change over time. GARCC was also more general than
model and did not impose excessive restrictions on the parameters of DCC models. At the same time,
the structural properties of GARCC model, especially the analytical form of regularity conditions,
were deduced, and the asymptotic theory was established. Hafner and Franses (2009) extended
Engle (2002)’s DCC model to allow asset-specific correlation sensitivity. This model was useful for
investors holding large asset returns. At the same time, they proposed two estimation methods, one
based on complete likelihood maximization and the other based on individual correlation estimation.
Applying the generalized DCC (GDCC) model to the daily data of stock returns of 39 UK firms on
FTSE, they found convincing evidence that the GDCC model was improved on the DCC model and
on the constant conditional correlations MGARCH model of Bollerslev (1990). Haas et al. (2009)
proposed an asymmetric multivariate extension of a new normal mixed GARCH model, discussed
the parameterization and estimation problems, derived the covariance stationarity condition and
the existence of the fourth moment, and gave the expression of the dynamic correlation structure of
the process.

Diamantopoulos and Vrontos (2010) used multiple Student $t$ error distributions to simulate the fat
tail nature of conditional distribution of financial returns data, extended the Vrontos et al. (2003) model,
and then proposed a Student $t$ full factor multivariate GARCH model. Combined with the reduction
parameterization of covariance matrix in the full factor multivariate GARCH model, the model could
be applied to high dimensional problems. Ausin and Lopes (2010) argued that the conditional ellipsoid
joint distribution of MGARCH model required strong symmetry, and financial data did not satisfy
this assumption in many cases. Therefore, they proposed a time-varying correlation Copula GARCH
model to deal with portfolio selection problems. Wei et al. (2010) used a greater number of linear and
nonlinear GARCH class models (see McCulloch (1985); Polasek et al. (2007) to capture the volatility
features of two crude oil markets and found that the nonlinear GARCH-class models exhibited greater
forecasting accuracy than the linear ones. Christoffersen et al. (2012) combined the DCC model
with partial $t$ Copula to study the conditional correlation of emerging stock market indices in 33
developed countries. Santos and Moura (2014) proposed a new method for conditional covariance
matrix estimation based on flexible dynamic multivariate GARCH model.

Klein and Walther (2016) incorporated an Expectation-Maximization algorithm for parameter
estimation of the mixture memory GARCH (MMGARCH) and found MMGARCH was also able to
cover asymmetric and long memory effects. Also, for variance forecasting and Value-at-Risk prediction, they found MMGARCH performed better due to its dynamic approach in varying the volatility level and memory of the process. Conrad and Mammen (2016) established the asymptotic theory of quasi-maximum likelihood estimator for the parametric GARCH-in-Mean model. The asymptotic behavior was based on the study of fluctuations of the parametric process of the model. Although time-varying GARCH-M models are commonly used in econometrics and finance, the recursive nature of conditional variance makes likelihood analysis computationally infeasible. Therefore, Anyfantaki and Demos (2016) suggested using Markov Chain Monte Carlo algorithm, which allowed classical estimators to be computed by simulating EM algorithm or only using simulated Bayes in O(T) operations (T is sample size), and derived the theoretical dynamic properties of time-varying parameter EGARCH (1,1)-M. Dias (2017) proposed an estimation strategy for stochastic time-varying risk premium parameters in time-varying GARCH-in-mean model, and Monte Carlo study showed that the algorithm had good finite sample properties.

Although time-varying covariance matrix performs well in capturing the dynamic change of finance markets, sample covariance matrix is used to replace the truly unknown covariance matrix in practice and there is estimation error in sample covariance matrix. Therefore, shrinkage estimation is introduced to reduce estimation error in sample covariance matrix.

5.1.2. Shrinkage Estimation

Jobson and Korkie (1980) used James-Stein estimator in mean-variance portfolio to demonstrate that the estimator provides more reasonable results than the traditional estimators. Jorion (1985) revealed the disadvantages of replacing expected returns with corresponding sample estimate without considering the inherent uncertainty in these parameter values. On this basis, Stein’s method of estimating initial returns was studied. By shrinking the sample average to a common mean, it was found that the out-of-sample performance of the optimal portfolio increased significantly. Frost and Savarino (1986) studied the portfolio selection problem by maximizing expected returns, based on the forecasting distribution of securities returns under the Bayesian framework, and found that the method could improve the performance of portfolio by reducing the priori information of estimation error. Jorion (1991) compared the active investment strategies of expected return under three alternative models: historical sample mean, shrinkage or Bayesian estimation and CAPM-based estimation and found that exchange risk is not factored in to stock prices despite its significance. Like Jorion (1991), Grauer and Hakansson (1995) compared the investment strategies and returns of the three estimated dynamic investment models. Mori (2004) studied the performance of mean-variance model on the optimal portfolio weights of proportional estimator and Stein estimator when the parameters were unknown. Kan and Zhou (2007) discussed the problem of investing in riskless assets and tangent portfolio funds and proposed a combination of sample tangent portfolio and sample global minimum variance portfolio. Okhrin and Schmid (2007) provided a comparison between the exact and asymptotic distributions of portfolio weight estimates and a sensitivity analysis of asset return moments. At the same time, considering the shrinkage estimation of several types of moments, the portfolio weights and its corresponding estimators were compared based on moment estimation.

The use of linear shrinkage to estimate covariance matrix has also attracted wide attention among researchers. Ledoit and Wolf (2003, 2004) used a linear combination of sample covariance matrices and a target matrix to estimate the covariance matrix, where the target matrix could be the identity matrix, or the covariance matrix estimated by the one factor model and applied this method to the portfolio selection problem. Bai and Shi (2011) summarized the methods commonly used in high-dimensional covariance matrix estimation, including shrinkage, observable and implicit factors, Bayesian method and random matrix theory. Yang et al. (2014) proposed a hybrid covariance matrix estimation method based on robust M estimation and Ledoit and Wolf (2004) shrinkage estimation. Ikeda and Kubokawa (2016) considered a class of general weighted estimators, including the linear...
combination of sample covariance matrices and the model-based estimators and the linear shrinkage estimators without special factors under the factor model.

On the other hand, the optimal portfolio is directly dependent on the inverse of the covariance matrix. Accordingly, the direct shrinkage of the inverse covariance matrix is also a good strategy. Stevens (1998) dealt with portfolio optimization problems by primitively constructing several direct characteristics of the inverse covariance matrix. Kourtis et al. (2012) used the linear combination of the inverse covariance matrix and the target matrix to estimate the inverse covariance matrix, where the target matrix could be the identity matrix, the inverse of the covariance matrix estimated by the one factor model and the linear combination of the former two and applied this method to the portfolio selection problem. Bodnar et al. (2016) gave an explicit stochastic representation of the weights of mean-variance portfolios by using the linear transformation distribution of the inverse covariance matrix.

Bickel and Levina (2008a) discussed the regularization of covariance matrices with \( n \) observation samples and \( p \) variables by hard threshold method. Under the conditions that: the true covariance matrix was sparse in a proper sense, the variables were Gaussian or sub-Gaussian distribution, \((\log p)/n\) approached zero and the explicit rate could be obtained, then the threshold estimation was consistent. Bickel and Levina (2008b) studied the estimation of banded and tapered sample covariance matrix and the banded inverse covariance matrix. Rothman et al. (2009) proposed a new generalized threshold algorithm combining shrinkage and threshold and studied the generalized threshold of sample covariance matrix in high-dimensional case. The generalized threshold of the covariance matrix has good theoretical properties and almost no computational burden. At the same time, an explicit convergence rate can be obtained in the operator norm, showing the tradeoff between sparsity, dimensionality and sample size of the real model. It was found that the generalized threshold is consistent in a large class of models if the dimension \( p \) and sample size \( n \) satisfy \( \log (p/n) \) approaching zero. Konno (2009) considered the estimation of large-dimensional covariance matrices for multivariate real normal and complex normal distributions when the dimensions of variables were larger than the number of samples. For real and complex cases, the Stein-Haff equations and eigen structures of singular Wishart matrices were respectively, provided. By using these techniques, unbiased risk estimates for some classes of global covariance matrices under real and complex invariant quadratic loss functions were obtained. Chen et al. (2010) considered the shrinkage method in a high-dimensional case and, proposed a covariance estimation method based on minimizing mean square error in Gaussian samples. Firstly, under the condition of sufficiency of statistics, Rao-Blackwell theory was used to propose a new method, RBLW estimator, which was superior to Ledoit-Wolf method under mean square error. Secondly, the iterative method of the clairvoyant shrinkage estimator was proposed to reduce estimation error. At the same time, the convergence of the iterative method was established, the closed-form expression of the limit was determined, and this method was an Oracle approximate contraction (OAS) estimator. Fisher and Sun (2011) used the convex combination of the sample covariance matrix and the well-conditioned target matrix to estimate the covariance matrix and introduced a new set optimal convex combination estimates of three commonly used target matrix. Cai and Liu (2011) considered the estimation of sparse covariance matrix using the threshold step of single element change. The estimator is completely data-driven and has good data and theoretical results. Moreover, the estimator adaptively achieves the optimal convergence rate on a large class of sparse covariance matrices under spectral norm. Ledoit and Wolf (2012) further studied the linear shrinkage of Ledoit and Wolf (2004) by nonlinear transformation of sample eigenvalues and extended the nonlinear shrinkage method to the precision matrix. Fan et al. (2013) proposed the principal orthogonal complement thresholding method (POET) to discuss the estimation of high-dimensional covariance with conditional sparse structure and fast divergent eigenvalues. By assuming the sparse error covariance matrix in the approximate factor model, some cross-sectional correlations were allowed even after the common but unobservable factors were excluded. The POET estimator included sample covariance matrix, factor covariance matrix, threshold estimator and adaptive threshold
estimator. At the same time, the convergence rates of sparse residual covariance matrix and conditional sparse covariance matrix under different norms were studied. With the increase of dimension, the sample covariance matrix becomes ill-conditioned and even singular. A commonly used method to estimate covariance matrix is Stein-type type compression estimation when the dimension is high. Touloumis (2015) proposed a new family of nonparametric Stein-type shrinkage covariance estimators, which were convex linear combinations of sample covariance matrices and predefined reversible target matrices. Under the Frobenius norm, the optimal shrinkage strength for defining the optimal convex linear combination depended on the unobserved covariance matrix and must be estimated from the data. At the same time, a simple and effective estimation process was proposed, which could obtain the nonparametric uniform estimator of the optimal contraction intensity for three commonly used target matrices. Zhang and Zhang (2018) combined the advantages of shrinkage estimation, vine copula structure and Black-Litterman model that could satisfy three investment objectives: estimation sensitivity, asymmetric risks appreciation, and portfolio stability.

5.2. Moment-Based and Quantile-Based Risk Measurement

5.2.1. VaR and CVaR

Some important articles are summarized in Table 4 below.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Paper/Book/Thesis Title (Please See References for Details)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baumol</td>
<td>1952</td>
<td>The transactions demand for cash: An inventory theoretic approach</td>
</tr>
<tr>
<td>Goldfarb and Iyengar</td>
<td>2003</td>
<td>Robust portfolio selection problems</td>
</tr>
<tr>
<td>McKay and Keefer</td>
<td>1996</td>
<td>VaR Is a Dangerous Technique</td>
</tr>
<tr>
<td>Artzner et al.</td>
<td>1997</td>
<td>Thinking coherently</td>
</tr>
<tr>
<td>Pflug</td>
<td>2000</td>
<td>Some remarks on the value-at-risk and the conditional value-at-risk</td>
</tr>
<tr>
<td>Uryasev</td>
<td>2000</td>
<td>Optimization of Conditional Value-at-Risk</td>
</tr>
<tr>
<td>Chen and Yang</td>
<td>2017</td>
<td>Multiperiod portfolio investment using stochastic programming with conditional value at risk</td>
</tr>
<tr>
<td>Zhang and Gao</td>
<td>2017</td>
<td>Portfolio selection based on a benchmark process with dynamic value-at-risk constraints</td>
</tr>
</tbody>
</table>

With the in-depth study of portfolio optimization theory and financial data, VaR attracted increasing attention in risk measurement. Baumol (1952) put forward the idea of VaR and used it to study the choice of securities. After Morgan (1996)’s risk measurement system, VaR has gained more attention (see Beder 1995; Jorion 1996). The basic principle recommended by the Basel Committee on Banking Supervision in 2001 states that VaR is a key indicator of risk (Szegö 2002). Goldfarb and Iyengar (2003) considered the robust VaR portfolio selection problem under the assumption of normal distribution and the objective of these robust formulations was to systematically combat the sensitivity of the optimal portfolio to statistical and modeling errors in the estimates of the relevant market parameters. Ghaoui et al. (2003) studied the portfolio selection problem with the worst case VaR when some of the distributed information was known. VaR is also used in insurance contracts. Wang et al. (2005) designed an optimal insurance contract by maximizing the expected final wealth of the insured under VaR constraints. Wang et al. (2005), Huang (2006) established an insurance contract under the risk constraint of VaR assuming that the insured was risk-averse. Giot (2005) used GARCH model and Riskmetrics model with residuals following normal distribution and Student t distribution to study VaR of three stocks traded on NYSE in 15 and 30 min. Chin (2008) compared the power of VaR under quantile and nonlinear time-varying volatility, proposed a simple Pareto distribution to explain the fat tail property in the empirical distribution of return, and implemented the measure of non-parametric quantile estimation of VaR using interpolation method. Batten et al. (2014) used the
modified version of the multifractal model of asset returns (MMAR) with a series of asset returns data characterized by second. Considering the fat tail of financial data, long-term dependence and inconsistency with the MMAR scale, the out-of-sample VaR prediction was derived, and the difference between this method and GARCH (1,1) position scale VaR model was compared. Zhao and Xiao (2016) proposed an optimal portfolio selection model with VaR constraints and the asset price process was modeled by a non-generalized statistical mechanics rather than a classical Wiener process. This model could describe the characteristics of fat tails of returns. Also, a Hamilton Jacobi Bell equation was obtained by using the dynamic programming principle and the closed form solution of logarithmic utility was obtained by Lagrange multiplier method. Jang and Park (2016) incorporated VaR constraints into the wealth and fuzziness of fund management and provided an optimal portfolio selection model for fund managers who divided assets into risk and risk-free. Chang et al. (2016) used the Granularity Adjustment (GA) method to calculate VaR in portfolio credit risk model and used Monte Carlo simulation to study the impact of concentrated risk on risk value. Naimy (2016) used CDS portfolio data from March 2013 to November 2015 in the United States, Europe and Asia to study the accuracy of VaR by measuring risk under the Delta normal and historical methods. Based on the VaR calculation method of portfolio composed of options and bonds, Wang et al. (2017) proposed a Monte Carlo simulation method to allow jump diffusion in underlying assets and provided a layout suitable for various models, including non-parametric and semi-parametric structures.

Several nonlinear VaR models for the calculation of option portfolio VaR have also been extensively studied. These models concentrate on relaxing the assumption that the change of option portfolio value is linear with the change of market risk factors while maintaining the computational feasibility. They improve the correlation between market risk factors and option portfolio value, including quadratic and linear terms, and are called Delta-Gamma-Theta-Normal VaR models. Morgan (1996) calculated the VaR value of option portfolio using the method of Johnson distribution transformation. Hardle et al. (2002) evaluated the accuracy and speed of computational methods for nonlinear VaR, including Johnson transformation, Cornish-Fisher, Monte Carlo and Fourier-Inversion methods. Data experiments showed that Johnson transformation and Cornish-Fisher method were faster but inexact, and Monte Carlo method was accurate but inefficient. Accordingly, the Fourier-Inversion method was found to be the best choice. Castellacci and Siclari (2003) used the Cornish-Fisher method to calculate the first-order moments of the distribution of portfolio value changes. Cui et al. (2013) studied Delta-Normal and Delta-Gamma-Theta-Normal VaR as well as parametric VaR asymptotic methods for nonlinear portfolios and discussed their computational effectiveness.

The widespread use of VaR has led many researchers to expand it to multidimensional case. However, there are many definitions of multidimensional quantile that make it difficult to generalize one-dimensional VaR to multivariate and still maintain many good properties of one-dimensional VaR (see Serfling 2002; Hallin et al. 2010; Fraiman and Pateiro-López 2012). In addition, McKay and Keefer (1996) believed that VaR may be ineffective in portfolio selection. Artzner et al. (1997) pointed out that VaR did not have subadditivity and convexity. Basak and Shapiro (2001) noted VaR-based optimal decisions had greater losses than expected expectation-based optimal decisions. Miller and Liu (2006) argued that under the current general assumption of joint normal distribution, there were many deviations in the VaR model of portfolio due to model set up errors.

Fortunately, Artzner et al. (1997) and Embrechts et al. (1999) explained that CVaR could be a reasonable alternative to VaR. Pflug (2000) proved that CVaR was a coherent risk measure and emphasized several properties of CVaR, such as convexity and monotonicity. Uryasev (2000) provided a comprehensive description of CVaR. Assuming the distribution was ellipsoid and the VaR was computable, Embrechts et al. (2001) showed the result of CVaR under the restricted condition was consistent with that of VaR. Rockafellar and Uryasev (2002) argued that CVaR was superior to VaR as a risk measure and gave the properties of CVaR under the distribution of financial losses involving prudent behavior. Gaivoronski and Pflug (2005) gave a method to calculate the minimum VaR portfolio under specific returns. By removing local outliers and smoothing VaR, the VaR’s efficient
boundaries were calculated. At the same time, the differences of VaR, CVaR and standard deviation as risk measures were compared. Kibzun and Kuznetsov (2006) compared the standards of VaR and CVaR and identified the links between them. Topaloglou et al. (2008) used CVaR as a risk measure to solve the international portfolio selection problem under the stochastic programming model. Huang et al. (2010) considered the relatively robust CVaR model when the potential distribution of asset returns belonged to a particular set. Under the worst case of uncertain distribution, the possible optimal decision was given according to the realization of each distribution. Mainik and Schaanning (2012) compared two possible concepts of CVaR available in the current literature, studied their general dependency consistency, and presented their performance in several stochastic models. Nguyen and Samorodnitsky (2013) proposed a multivariate tail estimator involving CVaR sequential statistical tests. Bernardino et al. (2014) constructed two multivariate CVaRs at the level of multivariate distribution functions and provided new risk measures based on Copula structure and random ordering of marginal distributions. Wang and Huang (2016) endogenously formulated the best form of insurance contract to maximize the expected utility of insurance under VaR and CVaR constraints. Date and Bustreo (2016) studied how to approximate VaR and CVAR using new heuristic methods when the net return of portfolio investment may be a nonlinear function of non-Gaussian risk factors. Chen and Yang (2017) used CVaR as a risk measure to propose portfolio stochastic programming and stage wise portfolio stochastic programming based on the stock investment data. Zhang and Gao (2017) used the dynamic CVaR risk-constrained benchmarking process to deal with the dynamic portfolio problem. Using the dynamic programming technique, they derived the corresponding Hamilton Jacoby Bellman equation and obtained the optimal portfolio strategy by Lagrange multiplier method. Li et al. (2018) proposed a hybrid intelligent algorithm using genetic algorithm design and adaptive penalty function, Simulated Annealing Back Propagation neural network and fuzzy simulation technic to solve the fuzzy mean CVaR efficient portfolio model. At the same time, in order to further improve the computing speed, MPI technology was used to parallelize the hybrid intelligent algorithm.

For both the VaR and CVaR, a probability level of cumulative loss needs to be specified (Benninga and Wiener 1998). Furthermore, (i) the optimal portfolios with the VaR constraints are sensitive to the confidence level selected (Campbell et al. 2001), (ii) the CVaR model requires either an assumption about the asset returns distribution or a substantial amount of return observations below the target return (Boasson et al. 2011). Comparison with VaR and CVaR, semi-variance is a good substitution for measuring the one-sided risk.

5.2.2. Semi-Variance

Some groundbreaking articles are summarized in Table 5 below.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Paper/Book/Thesis Title (Please See References for Details)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roy</td>
<td>1952</td>
<td>Safety First and the Holding of Assets</td>
</tr>
<tr>
<td>Markowitz</td>
<td>1959</td>
<td>Portfolio selection: Efficient diversification of investments</td>
</tr>
<tr>
<td>Bawa and Lindenberg</td>
<td>1977</td>
<td>Capital market equilibrium in a mean-lower partial moment framework</td>
</tr>
<tr>
<td>Bawa</td>
<td>1978</td>
<td>Safety-first, stochastic dominance, and optimal portfolio choice</td>
</tr>
<tr>
<td>Chen et al.</td>
<td>1991</td>
<td>A Method for Approximating Semivariance in Project Portfolio Analysis</td>
</tr>
<tr>
<td>Hamza and Janssen</td>
<td>1998</td>
<td>The mean-semivariances approach to realistic portfolio optimization subject to transaction costs</td>
</tr>
<tr>
<td>Ballestero</td>
<td>2005</td>
<td>Mean-semivariance efficient frontier: A downside risk model for portfolio selection</td>
</tr>
<tr>
<td>Huang</td>
<td>2008</td>
<td>Mean-semi variance models for fuzzy portfolio selection</td>
</tr>
<tr>
<td>Sayilgan and Mut</td>
<td>2010</td>
<td>Uses of Variance and Lower Partial Moment Measures for Portfolio Optimization</td>
</tr>
</tbody>
</table>

Roy (1952) proposed the concept of downside risk and defined it as a risk below the target value. Markowitz (1959) proposed a well-known mean-semi-variance model to estimate the weights...
of portfolio. Hogan and Warren (1972) pointed out the advantage of using the mean-semi-variance criterion in portfolio selection over the mean-variance model. Stone (1973) gave two interrelated three-parameter risk measures, in which the semi-variance was a special case. Hogan and Warren (1974) compared the difference between mean-variance model and mean-semi-variance model. Porter (1974) analyzed the relationship between stochastic dominance and mean-semi-variance model. Jahankhani (1976) empirically verified the relationship between return and risk in the mean-variance model and mean-semi-variance asset pricing model. Bawa and Lindenberg (1977) extended the semi-variance to the generalized lower partial moment framework, developed a Capital Asset Pricing Model (CAPM) using a mean-lower partial moment framework and derive explicitly formulæ for the equilibrium values of risky assets that hold for arbitrary probability distributions. Fishburn (1977) applied the downside risk to the utility function model. Bawa (1978) extended the downside risk to higher order and showed its usability. Nantell and Price (1979) calculated variance and semi-variance by means of the distribution of prior portfolio returns and found that asset market portfolio prices with semi-variance were higher than variance at a certain risk level. Choobineh and Branting (1986) provided a simple form of semi-variance approximation by using mean, variance, critical value and cumulative probability below the critical value. Lee and Rao (1988) proposed a new asset pricing model in the framework of mean lower partial moment, which used semi-variance and semi-deviation to measure risk. Lewis (1990) used semi-variance as a measure of risk, applied it to the capital market and utility theory, and explained its advantages and disadvantages. Chen et al. (1991) proposed a set of linear regression models to approximate the semi-variance of the total returns of items with independent distribution. Chow et al. (1992) pointed out that in the absence of prior knowledge about the parametric structure of asset return distribution and the form of investor preference function, variance may no longer be an appropriate risk measure. They used various risk-return measures independent of distribution to test the efficiency and decentralization effect of international portfolio investment and found that semi-variance could effectively and conveniently identify risks. Tse et al. (1993) put forward an optimal strategy for personal investment using downside risk and proposed a model for accurate calculation of failure probability under the assumption of Brown’s motion process. Markowitz (1993) transformed the mean-semi-variance portfolio optimization problem into the mean-variance optimization problem and used the critical line algorithm to obtain the optimal solution. Josephy and Aczel (1993) proposed an unbiased, consistent and effective estimators for the semi-variance.

With the deepening of risk research, the downside risk has attracted more and more attention (Rom and Ferguson 1994). Kaplan and Alldredge (1997) used a specific risk-based index, which could maintain a certain level of risk in different periods of time, to make a series of trade-offs between risk and return and studied its properties and performance in the case of semi-variance. Hamza and Janssen (1998) took transaction cost into consideration and applied the mean-semi-variance model to the portfolio selection problem, introduced a series of binary variables and separable constraints, and finally solved the portfolio optimization problem using separable techniques. Grootveld and Hallerbach (1999) analyzed the similarities and differences of using variance and downside risk as risk measures from empirical data and theory. Costa and Nabholz (2002) considered different computational forms of mean and semi-variance with errors and formulated robust mean-semi-variance portfolio selection problems based on linear matrix inequality optimization problems. Estrada (2004) noted that semi-variance was supported by theoretical facts and practical considerations and was a feasible measure of risk, and that the mean-semi-variance behavior criterion was perfectly consistent with the expected utility and the average compound return utility. Ballestero (2005) defined semi-variance as a weighted sum of squares deviating from the objective value of return on assets and applied it to portfolio selection. Jin et al. (2006) proved that no matter the market conditions and the distribution of stock returns, the effective strategy of mean-semi-variance in a single period could always be realized. They also established the realizability of the mean-semi-variance model under the condition of no arbitrage and extended it to the general downside risk measurement
problem. Sira (2006) described the significant differences in portfolio outcomes using variance and semi-variance to measure risk and emphasized that using semi-variance as a risk measure could lead to more robust and effective boundaries. Chabaane et al. (2006) used a group of hedge funds with significant deviations from normal to consider the portfolio problem by maximizing expected return under the constraints of standard deviation, semi-variance, VaR and CVaR. However, if the asset return data do not follow the normal distribution, the mean-semi-variance model may produce inefficient portfolios. Consequently, Eldomiaty (2007) proposed the mean-semi-deviation model to measure the average loss rate. Huang (2008) proposed two fuzzy mean-semi-variance models and proved the properties of semi-variance in the case of fuzzy variables. Sayilgan and Mut (2010) regarded the portfolio problem as a multi-objective optimization, used the semi-variance and the lower partial moment as the risk measurement, and took genetic algorithm to solve the multi-objective optimization to achieve Pareto efficient portfolio. Cumova and Nawrocki (2011) transformed the exogenous asymmetric matrix into a symmetric matrix and proved that there was indeed a closed form of solution. On this basis, the critical line algorithm could be used to solve the mean semi-variance problem. Assuming investment capital and net cash flow as fuzzy variables, Zhang et al. (2011) proposed the reliability return index and the reliability risk index by using the expected value of credibility and the lower semi-variance of the fuzzy variables and gave the comprehensive risk return index for selecting the optimal investment strategy. Zhang et al. (2012) proposed a probabilistic mean-semi-variance entropy model to deal with multi-period portfolio selection under fuzzy returns. Metaxiotis and Liagkouras (2012) used a multi-objective evolutionary algorithm to solve the constrained mean-semi-variance portfolio optimization problem. Alimi et al. (2012) used fuzzy programming technology to solve multi-objective fuzzy mean semi-variance portfolio optimization model. Brito et al. (2016) proposed a flexible approach to portfolio selection using skewness-semi-variance bio-objective optimization framework, which allowed investors to analyze the effective balance between biases and semi-variables. Salah et al. (2016) noted that estimating portfolio risk by conditional variance or conditional semi-variance could obtain information about the future development of different asset returns and help investors to obtain more effective portfolio. Chen et al. (2017) considered that stock returns limited by expert estimates were described as uncertain variables and then verified three properties of semi-variances of uncertain variables. Based on the concept of semi-variances of uncertain variables, the mean-semi-variance models of two types uncertain portfolio selection were proposed.

The semi-variance is also used in the multi-period case. Bi et al. (2013) discussed the continuous time mean-semi-variance portfolio selection problem with probability distorted by nonlinear transformation, provided ‘necessary and sufficient’ conditions for the existence of feasibility and optimal strategy, and gave the general form of the solution when the optimal solution existed. Zhang (2015) considered the multi-period portfolio selection problem in a fuzzy investment environment, in which the return and risk of assets were characterized by probability mean and semi-variance, respectively. At the same time, based on the possibility theory, a new multi-period possible portfolio selection model was proposed, which includes risk control, transaction cost, borrowing constraints, threshold constraints and cardinality constraints. Liu and Zhang (2015) considered the multi-period fuzzy portfolio optimization problem with the shortest trading lot. Based on the possibility theory, a mean-semi-variance portfolio selection model was proposed to maximize the final wealth and minimize the cumulative risk within the entire investment level. Najafi and Mushakhian (2015) proposed a multi-stage stochastic mean-semi-variance CVaR model to deal with portfolio optimization problems. The parameters of semi-variance and CVaR were controlled at a certain confidence level. Huang et al. (2016) took the correlation between items and time sequence into account to propose a new mean-variance and mean-semi-variance model. Chen et al. (2018) took securities returns as uncertain variables to establish a multi-period mean-semi-variance portfolio optimization model with realistic constraints: transaction costs, cardinality and boundary constraints. Furthermore, if the security return was zigzag uncertain variable, they gave the equivalent
deterministic form of mean-semi-variance model and proposed a modified imperialist competitive algorithm to solve the corresponding optimization problems.

6. Conclusions

In finance literature, the issues of portfolio selection, and risk measurement have always attracted attention of researchers globally. Accordingly, the present paper set out to review the development of related literature in the above areas and to identify the directions for future research. The study focused on three themes: (a) a review of literature on stylized facts that is, fat tails, volatility clustering and dependence structure of returns data, thereafter (b) a review of literature on portfolio selection and finally on (c) portfolio risk measurement. The objective was not only to trace the historical development but also identify possible research issues for future research.

The two important models for portfolio selection are the mean-variance model and global minimum variance model. The portfolio risk is measured by the covariance matrix in these models. From the literature review of these two models, we stressed that the covariance matrix estimation is important because the optimal portfolio weights rely on the covariance matrix. Accordingly, one of our focuses is on the estimation of covariance matrix. However, the estimation error in the covariance matrix estimation of asset returns is so large that the portfolio weights are likely inefficient. Therefore, the shrinkage methods are adopted to cope the estimation error in the estimation of covariance matrix. The shrinkage methods include Stein-type shrinkage methods and linear shrinkage methods. In linear shrinkage methods, we find that the factor model can be used to estimate the covariance matrix and the estimation is used as the target matrix. Consequently, the factor model is also included in the present paper. Also, to reflect the rapid changes of financial markets, we consider that the time-varying structure of covariance matrix is effective, and we take it as one useful improvement of the estimation of covariance matrix.

In addition to the covariance matrix to measure risk of portfolio, VaR and CVaR is another approach from the quantile perspective. Furthermore, many researchers think the risk is not symmetric and the risk should be the downside risk which measures the risk of falling below a target value. If the investor cares more about the loss, the downside risk measure could be a good solution.

The fat tail feature of financial data has received considerable attention in the relevant literature and many studies are based on the multivariate \( t \) distribution. In the presence of fat tails, the risk measure becomes more difficult to examine and the dependence of financial data is more important because co-movements exacerbate negative portfolio returns. Consequently, the Copula method has become a popular tool to describe the dependence structure of financial data appropriately.

However, there are many interesting questions that remain unsolved. For instance, in the stein-type shrinkage estimation of covariance matrix for portfolio selection problem, could we give an explicit shrinkage parameters selection method with maximizing investors’ utility? How to measure the asymmetric relationship of asset returns and apply it to portfolio selection problems?

One of the co-authors of this paper, Sun et al. (2018) have derived the Stein-type shrinkage strategy for optimal portfolio selection using the Cholesky decomposition of the covariance matrix under the mean-variance framework. The Stein-type shrinkage strategy is applied to simulation experiments and an empirical study to test its feasibility. Their proposed method works well in the simulation study and in the empirical analysis; however, there still exist interesting questions. For future work, the assumption of \( n > p \) may be replaced by \( p > n \) for high dimensional cases, where \( n \) is the sample size and \( p \) is the number of variables. A reasonable statistical loss function with a different objective function may be studied to take advantage of the proposed approach. In addition, the assumption of the normal distribution can be extended to elliptically symmetric or skewed distributions and take robustness into consideration as well.

Please note that in the minimum variance model, the covariance matrix plays an important role because it measures the risk and relationship of asset returns simultaneously under the normality assumption. However, as discussed earlier, the distribution of asset returns is non-normal and has an
obvious fat tail nature. In addition, the risk is one-sided. Hence it should be beneficial to study further and use a better tool to replace the covariance matrix, by involving the semi variance and distance correlation as discussed by e.g., Huang et al. (2016) and Sun et al. (2019).

Similarly, studies are required to examine the extent to which investment managers in the real-world incorporate the findings from the academic literature in practice.

**Funding:** The first two authors received support from the National Natural Science Foundation of China (11471264).

**Acknowledgments:** The authors would like to thank the editor and referees for the opportunity and their constructive comments which led to an improved version of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


Anyfantaki, Sofia, and Antonis Demos. 2016. Estimation and properties of a time-varying EGARCH (1,1) in mean model. *Econometric Reviews* 35: 293–310. [CrossRef]


Delatola, Eleni Ioanna, and Jim E. Griffin. 2013. A bayesian semiparametric model for volatility with a leverage effect. *Computational Statistics and Data Analysis* 60: 97–110. [CrossRef]


Haas, Markus, Stefan Mittnik, and Marc S. Paolella. 2009. Asymmetric multivariate normal mixture GARCH. *Computational Statistics and Data Analysis* 53: 2129–54. [CrossRef]


J. Risk Financial Manag. 2019, 12, 48


Karatzas, Ioannis, John P. Lehoczky, and Steven E. Shreve. 1987. Optimal portfolio and consumption decisions for a “small investor” on a finite horizon. SIAM Journal on Control and Optimization 25: 1557–86. [CrossRef]


Klein, Tony, and Thomas Walther. 2016. Oil price volatility forecast with mixture memory GARCH. Energy Economics 58: 46–58. [CrossRef]


Nguyen, Tilo, and Gennady Samorodnitsky. 2013. Multivariate tail estimation with application to analysis of CoVaR. *Astin Bulletin* 43: 245–70. [CrossRef]


Oksendal, Bernt, and Agnès Sulem. 2002. Optimal consumption and portfolio with both fixed and proportional transaction costs. *SIAM Journal on Control and Optimization* 40: 1765–90. [CrossRef]


Pogue, Gerald A. 1970. An extension of the Markowitz portfolio selection model to include variable transactions’ costs, short sales, leverage policies and taxes. *The journal of Finance* 25: 1005–27. [CrossRef]


So, Mike Ka Pui, and Cherry Y. T. Yeung. 2014. Vine-copula GARCH model with dynamic conditional dependence. *Computational Statistics and Data Analysis* 76: 655–71. [CrossRef]


Sun, Ruili, Tiefeng Ma, and Shuangzhe Liu. 2018. A Stein-type shrinkage estimator of the covariance matrix for portfolio selections. *Metrika*, 931–52. [CrossRef]

Sun, Ruili, Tiefeng Ma, and Shuangzhe Liu. 2019. Portfolio selection based on semivariance and distance correlation under minimum variance framework. *Statistica Neerlandica*. [CrossRef]


Touloumis, Anestis. 2015. Nonparametric Stein-type shrinkage covariance matrix estimators in high-dimensional setting. *Computational Statistics and Data Analysis* 83: 251–61. [CrossRef]


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).