Decision Support

Mean-variance model for portfolio optimization problem in the simultaneous presence of random and uncertain returns

Zhongfeng Qin
School of Economics and Management, Beihang University, Beijing 100191, China

ARTICLE INFO
Article history:
Received 28 April 2014
Accepted 11 March 2015
Available online 31 March 2015

Keywords:
Uncertainty modeling
Mean-variance model
Portfolio optimization
Uncertain variable
Uncertain measure

ABSTRACT
The determination of security returns will be associated with the validity of the corresponding portfolio selection models. The complexity of real financial market inevitably leads to diversity of types of security returns. For example, they are considered as random variables when available data are enough, or they are considered as uncertain variables when lack of data. This paper is devoted to solving such a hybrid portfolio selection problem in the simultaneous presence of random and uncertain returns. The variances of portfolio returns are first given and proved based on uncertainty theory. Then the corresponding mean-variance models are introduced and the analytical solutions are obtained in the case with no more than two newly listed securities. In the general case, the proposed models can be effectively solved by Matlab and a numerical experiment is illustrated.

Ef 2015 Elsevier B.V. All rights reserved.

1. Introduction

Portfolio selection problem is to consider how to allocate the fund among the candidate set of risky securities to maximize return and to minimize risk. Since it was originated by Markowitz (1952), there has been a large number of literature to be devoted to developing the problem. Same as the framework of Markowitz's mean-variance model, most of the works are undertaken to measure the investment return and risk by expected value and variance of return of an portfolio, respectively. In spite of the computational difficulty, the model is still successfully applied in reality and stimulate the development of finance.

In portfolio theory, the security returns are considered as random variables and their characteristics such as expected value and variance are calculated based on the sample of available historical data. It remains valid when there are plenty of data such as in the developed financial market. However, there may be lack of enough transaction data in some emerging markets. In the situation, some researchers regarded security returns as fuzzy variables estimated by experienced experts, and developed fuzzy portfolio optimization theory. More specifically, fuzzy portfolio optimization has been studied based on three different methods: fuzzy set theory (Gupta, Mehlawat, & Saxena, 2008), possibility theory (Carlsson, Fullér, & Majlender, 2002; Zhang, Wang, Chen, & Nie, 2007) and credibility theory (Huang, 2006; Qin, Li, & Ji, 2009).

Although fuzzy portfolio optimization provided alternatives to estimate security returns when lack of data, fuzzy theory suffers from criticism since a paradox will appear when fuzzy variable is used to describe the security returns (see Huang & Ying, 2013). In order to better describe the subjective imprecise quantity, Liu (2007) established uncertainty theory as another alternative way to estimate indeterministic quantities subject to experts' estimations. Based on this framework, much works are undertaken to develop the theory (Chen & Dai, 2011; Gao, 2009; Yao, 2012; You, 2009) and related practical applications (Li, Peng, & Zhang, 2013; Liu, 2010b; 2013; Qin & Kar, 2013). In particular, uncertainty theory is also applied to model portfolio selection. Qin, Kar, and Li (2015) first considers mean-variance model in uncertain environment. After that, Huang (2012) established a risk index model for uncertain portfolio selection, and Huang and Ying (2013) further employed the criterion to consider portfolio adjusting problem. Different from risk index model, Liu and Qin (2012) presented semiabsolute deviation of uncertain variable to measure risk and formulated the corresponding return-risk models, and Qin, Kar, and Zheng (In Press) employed the concept to describe and model portfolio adjusting problem. In addition to these single-period optimization models, Huang and Qiao (2012) also modeled the multi-period problem, and Zhu (2010) founded the uncertain optimal control and applied it to solve a continuous-time uncertain portfolio selection problem.

Whether the classical portfolio selection or fuzzy/uncertain one, security returns are considered as the same type of variables. In other words, security returns are assumed to be either random variables or fuzzy/uncertain variables. As stated above, the former makes use of the historical data and the latter makes use of the experiences of
experts. However, the actual situation is that the securities having been listed for a long time have yielded a great deal of transaction data. For these “existing” securities, statistical methods are employed to estimate their returns, which implies that it is reasonable to assume that security returns are random variables. For some newly listed securities, there are lack of data or there are only insufficient data which cannot be used to effectively estimate the returns. Therefore, the returns of these newly listed securities need to be estimated by experts and thus are considered as uncertain variables. If an investor faces such complex situation with simultaneous appearance of random and uncertain returns, how should he/she select a desirable portfolio to achieve some objectives? This paper attempts to consider the hybrid portfolio selection problem and establish mathematical models by means of uncertain random variable which was proposed by Liu (2013a) for modeling complex systems with not only uncertainty but also randomness. Based on this spirit, uncertain random programming is studied by Liu (2013b) and extended by Zhou, Yang, and Wang (2014), also applied to graph and network (Liu, 2014), risk analysis (Liu & Ralescu, 2014) and so on.

The rest of the paper is organized as follows. In Section 2, we review the necessary preliminaries related to uncertain measure, uncertain variable and uncertain random variable. Section 3 in detail describes the problem and notations and then gives two assumptions which are commonly used in portfolio analysis. Further, the formulas of variances of portfolio returns are given and proved. Section 4 considers the formulations of mean-variance models and discusses the solution procedures. A numerical experiment is illustrated to the application of the proposed models in Section 5. Finally, some conclusions are given in Section 6.

2. Preliminaries

In this section, uncertain measure, uncertain variable and uncertain random variable are introduced for the completeness of the paper. More details about existing measures of uncertainties the readers may consult (Liu, 2012). In 2007, Liu (2007) proposed the concept of uncertain measure to indicate the belief degree that a possible event happens. Let \( \mathcal{L} \) be a \( \sigma \)-algebra on a nonempty set \( \Gamma \).

**Definition 1** (Liu, 2007). A set function \( \mathcal{M}: \mathcal{L} \rightarrow [0, 1] \) is called an uncertain measure if it satisfies: (1) \( \mathcal{M} \{\Gamma\} = 1 \) for the universal set \( \Gamma \); (2) \( \mathcal{M} \{A\} + \mathcal{M} \{A^c\} = 1 \) for any event \( A \in \mathcal{L} \); (3) For every countable sequence of events \( A_1, A_2, \ldots \), we have \( \mathcal{M} \left( \bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} \mathcal{M} \{A_i\} \). The triple \( (\Gamma, \mathcal{L}, \mathcal{M}) \) is called an uncertain space.

In order to provide the operational law, Liu (2010a) further provided the product axiom as follows. Let \( (\Gamma_k, \mathcal{L}_k, \mathcal{M}_k) \) be uncertainty spaces for \( k = 1, 2, \ldots \). Then the product uncertain measure \( \mathcal{M} \) is an uncertain measure satisfying

\[
\mathcal{M} \left( \bigotimes_{k=1}^{\infty} A_k \right) = \bigwedge_{k=1}^{\infty} \mathcal{M}_k \{A_k\}
\]

where \( A_k \) are arbitrarily chosen events from \( \mathcal{L}_k \) for \( k = 1, 2, \ldots \), respectively. Next, the concept of uncertain variable was defined to describe a quantity with uncertainty.

**Definition 2** (Liu, 2007). An uncertain variable \( \xi \) is a measurable function from an uncertainty space \( (\Gamma, \mathcal{L}, \mathcal{M}) \) to the set of real numbers, i.e., for any Borel set \( B \) of real numbers, the set

\[
\{ \xi \in B \} = \{ \gamma \in \Gamma \mid \xi(\gamma) \in B \} \in \mathcal{L}.
\]

**Definition 3** (Liu, 2007, 2010a). For any \( x \in \mathbb{R} \), the uncertainty distribution of an uncertain variable \( \xi \) is defined by \( \Phi(x) = M \{ \xi \leq x \} \). It is said to be regular if it is a continuous and strictly increasing function with respect to \( x \) at which \( 0 < \Phi(x) < 1 \), and

\[
\lim_{x \to -\infty} \Phi(x) = 0, \quad \lim_{x \to +\infty} \Phi(x) = 1.
\]

The inverse function \( \Phi^{-1}(x) \) is called the inverse uncertainty distribution of \( \xi \) if it exists and is unique for each \( x \in (0, 1) \).

Inverse uncertainty distribution plays a crucial role in the operations of independent uncertain variables. Next we introduce some commonly used uncertainty distributions. An uncertain variable is called linear if it has a linear uncertainty distribution

\[
\Phi(r) = \begin{cases} 
0, & \text{if } r \leq a, \\
(r - a)/(b - a), & \text{if } a \leq r \leq b, \\
1, & \text{if } r \geq b,
\end{cases}
\]

denoted by \( \mathcal{L}(a, b) \) where \( a \) and \( b \) are real numbers with \( a < b \). An uncertain variable is called zigzag if it has a zigzag uncertainty distribution

\[
\Phi(r) = \begin{cases} 
0, & \text{if } r \leq a, \\
(r - a)/2(b - a), & \text{if } a \leq r \leq b, \\
(r + c - 2b)/(c - b), & \text{if } b \leq r \leq c, \\
1, & \text{if } r \geq c,
\end{cases}
\]

denoted by \( \mathcal{Z}(a, b, c) \) where \( a, b, c \) are real numbers with \( a < b < c \). An uncertain variable is called normal if it has a normal uncertainty distribution

\[
\Phi(r) = \left( 1 + \exp \left( \frac{\pi (r - c)}{\sqrt{3} \sigma} \right) \right)^{-1}, \quad r \in \mathbb{R},
\]

denoted by \( \mathcal{N}(e, \sigma) \) where \( e \) and \( \sigma \) are real numbers with \( \sigma > 0 \). These three uncertainty distributions are all regular.

**Lemma 1** (Liu, 2010a). Let \( \xi_1, \xi_2, \ldots, \xi_n \) be independent uncertain variables with regular uncertainty distributions \( \Phi_1, \Phi_2, \ldots, \Phi_n \), respectively. If \( \{x_1, x_2, \ldots, x_n\} \) is strictly increasing with respect to \( x_1, x_2, \ldots, x_n \), then \( \xi = f(\xi_1, \xi_2, \ldots, \xi_n) \) is an uncertain variable with an uncertainty distribution

\[
\Psi(\alpha) = \sup_{f(x_1, x_2, \ldots, x_n) \leq \alpha} \min \Phi_i(x_i),
\]

and with an inverse uncertainty distribution

\[
\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \ldots, \Phi_n^{-1}(\alpha)).
\]

**Definition 4** (Liu, 2007). The expected value of an uncertain variable \( \xi \) is defined by

\[
E[\xi] = \int_{-\infty}^{+\infty} \mathcal{M} \{\xi \geq x\} dx - \int_{-\infty}^{0} \mathcal{M} \{\xi \leq x\} dx
\]

provided that at least one of the two integrals exists.

**Lemma 2** (Liu, 2010a). Let \( \xi \) be an uncertain variable with regular uncertainty distribution \( \Phi \). If the expected value \( E[\xi] \) exists, then \( E[\xi] = \int_{0}^{1} \Phi^{-1}(\alpha) d\alpha \).

Uncertain random variable is employed to describe a complex system with not only uncertainty but also randomness. Let \( (\Gamma, \mathcal{L}, \mathcal{M}) \) be an uncertainty space, and \( (\Omega, \mathcal{P}, \Pr) \) a probability space. The product \( (\Gamma, \mathcal{L}, \mathcal{M}) \times (\Omega, \mathcal{P}, \Pr) \) is called a chance space. We have the following definition.

**Definition 5** (Liu, 2013a). Let \( (\Gamma, \mathcal{L}, \mathcal{M}) \times (\Omega, \mathcal{P}, \Pr) \) be a chance space, and let \( \Theta \in \mathcal{L} \times \mathcal{P} \). Then the chance measure of uncertain random event \( \Theta \) is defined as

\[
\text{Ch}[\Theta] = \int_{0}^{1} \Pr[\omega \in \Omega | \mathcal{M} \{\gamma \in \Gamma | (\gamma, \omega) \in \Theta\} \geq r] \, dr.
\]

**Definition 6** (Liu, 2013a). An uncertain random variable is a function \( \xi \) from the chance space \( (\Gamma, \mathcal{L}, \mathcal{M}) \times (\Omega, \mathcal{P}, \Pr) \) to the set of real numbers, i.e., \( \xi \in B \) is an event in \( \mathcal{L} \times \mathcal{P} \) for any Borel set \( B \). Its chance distribution is defined by \( \Phi(x) = \text{Ch}[\xi \leq x] \) for any \( x \in \mathbb{R} \).
It follows from Definition 6 that random variables and uncertain variables are special cases of uncertain random variables. If $x$ is a random variable and $\tau$ is an uncertain variable, then the sum $\eta + \tau$ and the product $\eta \tau$ are both uncertain random variables.

**Definition 7** (Liu, 2013b). Let $\xi$ be an uncertain random variable. Then its expected value is defined by

$$E[\xi] = \int_0^{+\infty} \text{Ch}(\xi \geq r) dr - \int_{-\infty}^0 \text{Ch}(\xi \leq r) dr$$

provided that at least one of the two integrals is finite.

Let $\xi$ be an uncertain random variable with finite expected value $e$. The variance of $\xi$ is defined by Liu (2013b) as $\text{Var}[\xi] = E[(\xi - e)^2]$. Suppose that $\Phi$ is its chance distribution. Guo and Wang (2014) propose a stipulation to calculate the variance by

$$\text{Var}[\xi] = \int_{\mathbb{R}} (1 - \Phi(e + \sqrt{\xi}) + \Phi(e - \sqrt{\xi})) d\xi.$$ 

Based on the stipulation, we have the following lemma.

**Lemma 3** (Sheng & Yao, 2014). Let $\xi$ be an uncertain random variable with chance distribution $\Phi$ and finite expected value $e$. Then

$$V[\xi] = \int_{-\infty}^{+\infty} (x-e)^2 d\Phi(x).$$

### 3. Problem statement

In this section, we will state the problem of finding out the optimal portfolio in the simultaneous presence of random and uncertain returns and then discuss the variance of uncertain random returns. For sake of convenience, we list the notations used in the rest as follows:

- $m$ the number of existing risky securities which have enough historical data
- $n$ the number of newly listed risky securities which are lack of historical data
- $\xi_i$ the future return of the ith existing risky security, which is a random variable, $i = 1, 2, \ldots, m$
- $\mu_j$ the expected value of random return $\xi_i$ of the ith existing security, $i = 1, 2, \ldots, m$
- $\sigma_{ij}$ the covariance of random returns $\xi_i$ and $\xi_j$, $i, j = 1, 2, \ldots, m$
- $\Psi_j$ the probability distribution of random return $\xi_i$, $i = 1, 2, \ldots, m$
- $\eta_j$ the future return of the jth newly listed risky security, which is an uncertain variable, $j = 1, 2, \ldots, n$
- $\nu_j$ the expected value of uncertain return $\eta_j$ of the jth newly listed security, $j = 1, 2, \ldots, n$
- $\delta^2_j$ the variance of uncertain return $\eta_j$ of the jth newly listed security, $j = 1, 2, \ldots, n$
- $\gamma_j$ the uncertainty distribution of uncertain return $\eta_j$, $j = 1, 2, \ldots, n$
- $x_i$ the holding proportion of the existing security $i$, $i = 1, 2, \ldots, m$
- $y_j$ the holding proportion of the newly listed security $j$, $j = 1, 2, \ldots, n$

Let $'$ be the matrix transpose symbol. Denote by $x = (x_1, x_2, \ldots, x_m)$ the portfolio vector of the existing securities, by $\xi = (\xi_1, \xi_2, \ldots, \xi_m)'$ the random vector of returns of the existing securities, by $y = (y_1, y_2, \ldots, y_n)'$ the portfolio vector of the newly listed securities, by $\eta = (\eta_1, \eta_2, \ldots, \eta_n)'$ the vector of returns of the newly listed securities. Then $(x, y)' = (x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n)'$ represents the portfolio vector of all the candidate securities, which yields the total return

$$r(x, y; \xi, \eta) = x'\xi + y'\eta = x_1\xi_1 + x_2\xi_2 + \cdots + x_m\xi_m + y_1\eta_1 + \cdots + y_n\eta_n$$

where $x \xi$ and $y \eta$ are the total returns yielded by the portfolios of the existing securities and the newly listed securities, respectively. Obviously, $x \xi$ is a random variable and $y \eta$ is an uncertain variable. It follows from Theorem 5 of Liu (2013b) that

$$E[r(x, y; \xi, \eta)] = E[x'\xi + y'\eta] = E[x'\xi] + E[y'\eta].$$

Denote by $\mu = (\mu_1, \mu_2, \ldots, \mu_m)'$ the expected vector of $\xi$, by $\Sigma = (\sigma_{ij})_{m \times m}$ the covariance matrix of $\xi$. We shall make use of the following assumptions.

#### Assumption 1

The random vector $\xi$ has a multivariate normal distribution with the following probability density function

$$\psi_{\xi}(\zeta) = \frac{1}{\sqrt{(2\pi)^m |\Sigma|}} \exp \left(-\frac{1}{2}(\zeta - \mu)'\Sigma^{-1}(\zeta - \mu)\right).$$

#### Assumption 2

Uncertain returns $\eta_1, \eta_2, \ldots, \eta_n$ are independent in the sense of uncertain measure (Liu, 2010a), i.e.,

$$\mathcal{M} \left( \bigcap_{j=1}^n \{ \eta_j \in B_j \} \right) = \prod_{j=1}^n \mathcal{M} \{ \eta_j \in B_j \}$$

for any Borel set $B_1, B_2, \ldots, B_n$ of real numbers, in which $\land$ is the minimum operator.

According to Assumption 1, $\mu$ and $\Sigma$ are the mean and covariance matrix of $\xi$, respectively, and $x'\xi$ is a normally distributed random variable with the expected value $\mu'x = x_1\mu_1 + x_2\mu_2 + \cdots + x_m\mu_m$, the variance $\Sigma x = \sum_{i=1}^m x_i\sigma_{ij}$, and the following probability density function

$$\psi(w) = \frac{1}{\sqrt{2\pi |\sigma(x)|}} \exp \left(-\frac{(w - \mu'x)^2}{2\sigma^2(x)}\right)$$

where $\sigma(x) = \sqrt{\Sigma x}$ and $\sigma^2(x) = (\sigma(x))^2$. Denote by $\nu = (\nu_1, \nu_2, \ldots, \nu_n)'$ the expected vector of $\eta$. Further, based on Assumption 2, it follows from the linearity of expected value of uncertain variables that

$$E[r(x, y; \xi, \eta)] = E[x'\xi] + E[y'\eta] = x'\mu + y'\nu.$$  

Next, we consider the variance of the portfolio return $r(x, y; \xi, \eta)$ which is an uncertain random variable. It follows from Lemma 3 that the variance is

$$V[r(x, y; \xi, \eta)] = \int_{-\infty}^{+\infty} [u - (x'\mu + y'\nu)]^2 d\Phi(u).$$

where $\Phi(u)$ is the chance distribution of $r(x, y; \xi, \eta)$ determined by

$$\Phi(u) = \int_{-\infty}^{+\infty} \Upsilon(u - w)d\Psi(w)$$

in which $\Psi(\cdot)$ is the probability distribution of $x'\xi$ and $\Upsilon(\cdot)$ is the uncertainty distribution of $y'\eta$.

Similar to stochastic situation, the uncertain returns $\eta_1, \eta_2, \ldots, \eta_n$ are always assumed to follow the same type of uncertainty distribution, such as linear uncertainty distribution, zigzag uncertainty distribution, normal uncertainty distribution and so forth. In these cases, the analytical expressions of variances $V[r(x, y; \xi, \eta)]$ can be obtained. For simplicity, we write $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)'$, $b = (b_1, b_2, \ldots, b_n)'$, $c = (c_1, c_2, \ldots, c_m)'$, $e = (e_1, e_2, \ldots, e_n)'$ and $\delta = (\delta_1, \delta_2, \ldots, \delta_n)'$. The results are given in the form of theorems stated below.

#### Theorem 1

Assume that $\eta_j \sim C(a_j, b_j)$ is a linear uncertain variable for $j = 1, 2, \ldots, n$. Then

$$V[r(x, y; \xi, \eta)] = \sigma^2(x) + \frac{(b'y - a'y)^2}{12}.$$  

**Proof.** It follows from the operational law of uncertain variables that $y'\eta$ is also a linear uncertain variable with expected value $(a'y + b'y)/2$ and the following uncertainty distribution

$$\Upsilon(z) = \frac{z - a'y}{b'y - a'y} I_{\{a'y < z < b'y\}} + I_{\{z < b'y\}}.$$
where \( l_{ij} \) is the indicator function of the set \{ \}. It follows from Eqs. (2) and (5) that

\[
\Phi(u) = \int_{-\infty}^{\infty} \left( \frac{u - \frac{a'y}{b'y - a'y}}{b'y - a'y} \right) \times l_{[y < s, b'y < b'y]} + l_{[u < w, b'y < b'y]} ) \psi(w) \, dw
\]

\[
= \frac{1}{\sqrt{2\pi}\sigma(y)} \int_{-\infty}^{\infty} \exp\left(-\frac{(u - \mu')^2}{2\sigma^2(y)}\right) \, dw
\]

Taking the derivative with respect to \( u \) on both sides of the above equation, we have

\[
d\Phi(u) = \frac{1}{\sqrt{2\pi}\sigma(y)} \int_{-\infty}^{\infty} \exp\left(-\frac{(u - \mu')^2}{2\sigma^2(y)}\right) \, du
\]

It follows from Eq. (4) that the variance is

\[
\frac{\sigma^2(y)}{\gamma^2} = \frac{5(b'y - a'y)^2 + 5(c'y - b'y)^2 + 6(b'y - a'y)(c'y - b'y)}{48}
\]

The theorem is completed. \( \square \)

**Theorem 2.** Assume that \( \eta_j \sim \mathcal{N}(a_j, b_j, c_j) \) is a zigzag uncertain variable for \( j = 1, 2, \ldots, n \). Then

\[
\frac{\sigma^2(y)}{\gamma^2} = \frac{5(b'y - a'y)^2 + 5(c'y - b'y)^2 + 6(b'y - a'y)(c'y - b'y)}{48}
\]

**Proof.** It follows from the operational law of uncertain variables that \( y' \) is also a zigzag uncertain variable with expected value \( (a'y + 2b'y + c'y)/4 \) and the following uncertainty distribution

\[
\Upsilon(z) = \frac{z - a'y}{2b'y - 2a'y} l_{[y < z, b'y < b'y]} + \frac{z + b'y - 2a'y}{2c'y - 2b'y} l_{[y > z, c'y < c'y]} + l_{[z, c'y]}. \]

It follows from Eqs. (2) and (5) that

\[
\Phi(u) = \int_{-\infty}^{\infty} \left( \frac{u - \frac{a'y}{b'y - a'y}}{b'y - a'y} \right) \times l_{[y < s, b'y < b'y]} + l_{[u < w, b'y < b'y]} ) \psi(w) \, dw
\]

\[
= \frac{1}{\sqrt{2\pi}\sigma(y)} \int_{-\infty}^{\infty} \exp\left(-\frac{(u - \mu')^2}{2\sigma^2(y)}\right) \, du
\]

Taking the derivative with respect to \( u \) on both sides of the above equation, we have

\[
d\Phi(u) = \frac{1}{\sqrt{2\pi}\sigma(y)} \int_{-\infty}^{\infty} \exp\left(-\frac{(u - \mu')^2}{2\sigma^2(y)}\right) \, du
\]

It follows from Eq. (4) that the variance is

\[
\frac{\sigma^2(y)}{\gamma^2} = \frac{5(b'y - a'y)^2 + 5(c'y - b'y)^2 + 6(b'y - a'y)(c'y - b'y)}{48}
\]

The theorem is completed. \( \square \)

**Theorem 3.** Assume that \( \eta_j \sim \mathcal{N}(a_j, b_j) \) is a normal uncertain variable for \( j = 1, 2, \ldots, n \). Then

\[
\frac{\sigma^2(y)}{\gamma^2} = \frac{5(b'y - a'y)^2 + 5(c'y - b'y)^2 + 6(b'y - a'y)(c'y - b'y)}{48}
\]

**Proof.** It follows from the operational law of uncertain variables that \( y' \) is also a normal uncertain variable with expected value \( e'y \) and the following uncertainty distribution

\[
\Upsilon(z) = \left( 1 + \exp\left( \frac{(e'y - z)}{\sqrt{3\delta y}} \right) \right)^{-1}
\]

It follows from Eqs. (2) and (5) that

\[
\Phi(u) = \frac{1}{\sqrt{2\pi}\sigma(y)} \int_{-\infty}^{\infty} \exp\left(-\frac{(u - \mu')^2}{2\sigma^2(y)}\right) \, du
\]
Taking the derivative with respect to $u$ on both sides of the above equation, we have

$$\frac{d\Phi(u)}{du} = \frac{\sqrt{\pi}}{\sqrt{6d(\pi)\delta y}} \int_{-\infty}^{+\infty} \left(1 + \exp\left(\frac{\pi(e'y - u + w)}{\sqrt{3d} y}\right)\right)^2 \times \exp\left(\frac{\pi(e'y - u + w)}{\sqrt{3d} y}\right) \exp\left(-\frac{(w - \mu'x)^2}{2\sigma^2(x)}\right) \, dw.$$ 

It follows from Eq. (4) that the variance is

$$V[r(x, y; \xi, \eta)] = \int_{-\infty}^{+\infty} \left[u - (\mu'x + e'y)\right]^2 d\Phi(u)$$

$$= \frac{\sqrt{\pi}}{\sqrt{6d(\pi)\delta y}} \int_{-\infty}^{+\infty} \left[1 + \exp\left(\frac{\pi(e'y - u + w)}{\sqrt{3d} y}\right)\right]^2 \times \exp\left(\frac{\pi(e'y - u + w)}{\sqrt{3d} y}\right) \exp\left(-\frac{(w - \mu'x)^2}{2\sigma^2(x)}\right) \, dw du$$

$$= \frac{\sqrt{\pi}}{\sqrt{6d(\pi)\delta y}} \int_{-\infty}^{+\infty} \left[1 + \exp\left(\frac{\pi(e'y - u + w)}{\sqrt{3d} y}\right)\right]^2 \times \exp\left(\frac{\pi(e'y - u + w)}{\sqrt{3d} y}\right) \exp\left(-\frac{(w - \mu'x)^2}{2\sigma^2(x)}\right) \, dw du$$

By making the change of variables $u = e'y + w - \sqrt{3d}y/\pi$, we obtain

$$V[r(x, y; \xi, \eta)] = \frac{1}{\sqrt{2\pi(\pi x)}(\pi)} \int_{-\infty}^{+\infty} \exp\left(-\frac{(w - \mu'x)^2}{2\sigma^2(x)}\right) \left(\frac{\sqrt{3d}y}{\pi}x + \mu'x - w\right)^2 \frac{e^x}{(1 + e^x)^2} \, dx dw$$

$$= \frac{1}{\sqrt{2\pi(\pi x)}(\pi)} \int_{-\infty}^{+\infty} \exp\left(-\frac{(w - \mu'x)^2}{2\sigma^2(x)}\right) \left(\frac{\sqrt{3d}y}{\pi}x + \mu'x - w\right)^2 \frac{e^x}{(1 + e^x)^2} \, dx dw$$

$$= \sigma^2(x) + (\delta y)^2.$$ 

The theorem is completed. $\square$

4. Model formulations

Mean-variance model aims at finding out the most desirable portfolio only by the first two moments. Expected value of the total return is regarded as the investment return, and the variance is used to measure the investment risk. A typical mean-variance model is to compromise between risk and return. For example, we may choose a portfolio with minimum investment risk on the condition that the acceptable return level is given. Suppose that $\beta_m$ and $\beta_n$ are respectively the $m$-dimensional column vector and $n$-dimensional column vector whose all elements are 1. Following the spirit, we can formulate the following mean-variance model for hybrid portfolio optimization problem,

$$\begin{align*}
&\min \left\{ V[r(x, y; \xi, \eta)] = V[x'\xi + y'\eta] \\
&\text{s.t. } E[r(x, y; \xi, \eta)] = E[x'\xi + y'\eta] = \mu_0 \\
&\quad 1_m x + 1_n y = 1 \right\}
\end{align*}$$

(9)

where $\mu_0$ is the acceptable return level and the second constraint ensures that the net expected return is exactly equal to the return level $\mu_0$.

We may also choose a portfolio with maximum investment return on the condition that the tolerable risk level is given. The detailed formulation is as follows,

$$\begin{align*}
&\max \left\{ E[x'\xi + y'\eta] \\
&\text{s.t. } V[x'\xi + y'\eta] = \lambda_0 \\
&\quad 1_m x + 1_n y = 1 \right\}
\end{align*}$$

(10)

where $\lambda_0$ is the tolerable risk level.

In practice, the investor does not always know how to set $\mu_0$ and $\lambda_0$. An alternative is to provide a factor representing the risk aversion by trading off return and risk. Thus, we can establish the following compromise model,

$$\begin{align*}
&\min \left\{ \phi V[x'\xi + y'\eta] - E[x'\xi + y'\eta] \\
&\text{s.t. } 1_m x + 1_n y = 1 \right\}
\end{align*}$$

(11)

where $\phi$ is a factor of risk aversion.

Next, we consider the solution procedures of the proposed models. We first give the third assumption.

**Assumption 3.** The covariance matrix $\Sigma$ is positive definite and the expectation $\mu$ is not a multiple of $1_m$.

**Assumption 3** guarantees a nondegenerate situation and implies that all the existing securities are indeed risky. The detailed remark can be found in Steinbach (2001). Based on **Assumption 3**, we write $\rho = \mu'\Sigma^{-1} \mu, \kappa = 1_m \Sigma^{-1} 1_m, \varsigma = 1_m \Sigma^{-1} \mu$ that will be used throughout this section.

According to Assumptions 1 and 2, it follows from Theorems 1–3 that Model (11) is equivalent to the following form,

$$\begin{align*}
&\min \left\{ \phi \left[x'\Sigma x + y'p'y\right] - (x'\mu + y'v) \\
&\text{s.t. } 1_m x + 1_n y = 1 \right\}
\end{align*}$$

(12)

in which $p = (b - a)/2\sqrt{3}$ in Theorem 1, $pp' = [5(b - a)(b - a)' + 5(c - b)(c - b)' + 6(b - a)(c - b)]/48$ in Theorem 2 and $p = \delta$ in Theorem 3, respectively. On this basis, it follows from Assumption 3 that Model (12) is a convex quadratic programming. We may use the method of Lagrange multipliers to find out its solution.

The Lagrange function of Model (12) is

$$L(x, y, \lambda) = \phi(x'\Sigma x + y'p'y) - (x'\mu + y'v) - \lambda(1_m x + 1_n y - 1).$$

Further, the first-order necessary conditions are

$$\frac{\partial}{\partial x} L(x, y, \lambda) = 2\phi\Sigma x - \mu - \lambda 1_m = 0,$$

$$\frac{\partial}{\partial y} L(x, y, \lambda) = 2\phi p y - v - \lambda 1_n = 0,$$

$$\frac{\partial}{\partial \lambda} L(x, y, \lambda) = (1_m x + 1_n y) - 1 = 0,$$

in which Eq. (15) is essentially the original constraint of Model (12). If $(x^*, y^*, \lambda^*)$ is the solution of Eqs. (13)–(15), then it is called a stationary point for the Lagrange function $L(x, y, \lambda)$. It is known that a stationary point of a convex function is always a global minimum point. Thus, the problem of finding the optimal solution of Model (12) is reduced to find stationary points of Eqs. (13)–(15).

**Theorem 4.** Let $n = 1$ and $p = (p) > 0$. Then the solution of Model (11) is

$$x^* = \frac{1}{2\phi} \Sigma^{-1} \mu + \frac{\lambda^*}{2\phi} \Sigma^{-1} 1_m,$$

$$y^* = \frac{1}{2\phi p^2} (v + \lambda^*),$$

(16)

where

$$\lambda^* = \left(2\phi - \varsigma\right)p^2 - v \quad \frac{p^2}{\kappa + 1}.$$ 

**Proof.** First, Eqs. (13) and (14) yield

$$x = \frac{1}{2\phi} \Sigma^{-1} \mu + \frac{\lambda}{2\phi} \Sigma^{-1} 1_m,$$

$$y = \frac{1}{2\phi p^2} (v + \lambda),$$

(17)
Substituting into Eq. (15) yields

\[
\frac{1}{2\phi} \Sigma^{-1} \mu + \frac{\lambda}{2\phi} \Sigma^{-1} \mathbf{1}_m + \frac{1}{2\phi p^2} (v + \lambda) = 1
\]

which implies that \( \lambda \) must be equal to be \((2\phi p^2 - \zeta p^2 - v)/(p^2 \kappa + 1)\).

**Theorem 5.** Let \( n \) be an integer. In particular, they have the following relationship,

\[
\left\{
\begin{array}{ll}
x^* > y^*, & \text{if } \mu + \phi p^2 > v + \phi \sigma^2 \\
x^* = y^*, & \text{if } \mu + \phi p^2 = v + \phi \sigma^2 \\
x^* < y^*, & \text{if } \mu + \phi p^2 < v + \phi \sigma^2.
\end{array}
\right.
\]

**Example 1.** We consider a simplified portfolio selection problem with one existing security and one newly listed one. In this case, \( m = 1 \) and \( n = 1 \). Write \( \Sigma = \sigma_{11} = \sigma^2, \mu = \mu \). Then we have

\[
\rho = \frac{\mu^2}{\sigma^2}, \kappa = \frac{1}{\sigma^2}, \zeta = \frac{\mu}{\sigma^2}, \lambda^* = \frac{(2\phi \sigma^2 - \mu)p^2 - \nu \sigma^2}{p^2 + \sigma^2}.
\]

The optimal proportions invested these two securities are

\[
x^* = \frac{\mu - v + 2\phi \sigma^2}{2\phi (p^2 + \sigma^2)} \quad y^* = \frac{v - \mu + 2\phi \sigma^2}{2\phi (p^2 + \sigma^2)}
\]

respectively. In particular, they have the following relationship,

\[
\left\{
\begin{array}{ll}
x^* > y^*, & \text{if } \mu + \phi p^2 > v + \phi \sigma^2 \\
x^* = y^*, & \text{if } \mu + \phi p^2 = v + \phi \sigma^2 \\
x^* < y^*, & \text{if } \mu + \phi p^2 < v + \phi \sigma^2.
\end{array}
\right.
\]

**Remark 1.** Assume that \( p_1 = p_2 \), i.e., the variances are same for two newly listed securities. Then investors generally prefer the security with larger expected return. Therefore, the security with smaller expected return is naturally not considered and the problem is reduced to the one in Example 1.

**Example 2.** Let \( m = 1 \). That is, we consider a simplified portfolio selection problem with one existing security and two newly listed ones with different risks. According to Theorem 5, we have

\[
x^* = \frac{p_1 (v_2 - \mu) + p_2 (\mu - v_1)}{2\phi \sigma^2 (p_2 - p_1)},
\]

\[
y_1^* = \frac{(v_1 - v_2) \sigma^2 - p_1 (v_1 - v_2) p_2 (p_1 - p_2)}{2\phi \sigma^2 (p_2 - p_1)^2} + \frac{(2\phi - \mu) p_2}{2\phi \sigma^2 (p_2 - p_1)},
\]

\[
y_2^* = \frac{(v_1 - v_2) \sigma^2 + p_1 (v_1 - v_2) p_2 (p_1 - p_2)}{2\phi \sigma^2 (p_2 - p_1)^2} - \frac{(2\phi - \mu) p_1}{2\phi \sigma^2 (p_2 - p_1)}.
\]

Let \( v_1 = 2, p_1 = 1, v_2 = 3, p_2 = 2, \mu = 3 \) and \( \sigma^2 = 4 \). Fig. 1 visually shows the investment proportions of different securities with the changing of the values of the risk aversion factor \( \phi \).

Suppose that \( \mu = v_1 = v_2 \). Then

\[
x^* = 0, \quad y_1^* = \frac{p_2}{p_2 - p_1}, \quad y_2^* = -\frac{p_1}{p_2 - p_1}
\]

which implies that the existing security is not selected if it has the same expected return with the newly listed securities regardless of its variance.

Next we consider the case that the number \( n \) of newly listed securities is more than 2. First note that the rank of matrix \( pp^T \) is no more than 1 which implies that the solution of simultaneous Eqs. (13)–(15) is either no solution or infinitely many solutions. In the latter case, we may find the stationary point to further obtain the optimal solution because of the convexity of the proposed models. We also may directly solve the original problem since the proposed models are convex quadratic programmings, which can be solved by available softwares such as Matlab.
nonnegativity. Model (9) is reformulated as follows, in this case, we need to added the constraint of 

Table 3 including their expected values and return rates are denoted by zigzag uncertain variables. The simu-

April 1, 2014 and 51 monthly return rates are obtained for each

trate the application of the proposed models. The existing secu-

Table 2

The sample covariance matrix of the 20 stocks.

Table 3

Uncertain returns, expected values, variances of simulated 5 newly listed stocks.

5. Numerical experiment

In this section, a numerical experiment is presented to illus-

Where $\mu$, $\Sigma$ and $a$, $b$, $c$, $v$ can be found in Tables 1, 2 and 3, respec-

\begin{equation}
\begin{aligned}
\min_{x} & \quad \left( b^\top y - a^\top y \right)^2 + 5 \left( c^\top y - b^\top y \right)^2 + 6 \left( b^\top y - a^\top y \right) (c^\top y - b^\top y) \\
\text{s.t.} & \quad x^\top \Sigma x + \frac{5}{48} (b^\top y - a^\top y)^2 + 5 (c^\top y - b^\top y)^2 + 6 (b^\top y - a^\top y) (c^\top y - b^\top y) \\
& \quad 1^\top x + 1^\top y = 1, \\
& \quad x, y \geq 0.
\end{aligned}
\end{equation}

We also employ Model (12) to construct the desirable portfolio 
by changing the factor of risk aversion. In fact, we can obtain the

\[\begin{array}{c}
\eta_1 188.5 \\
\eta_2 19.2 \\
\eta_3 15.6 \\
\eta_4 6.093 \\
\eta_5 2.581
\end{array}\]

\[\begin{array}{c}
\eta_1 1.566 \\
\eta_2 1.040 \\
\eta_3 2.390 \\
\eta_4 2.850 \\
\eta_5 0.854
\end{array}\]

\[\begin{array}{c}
\eta_1 1.953 \\
\eta_2 1.040 \\
\eta_3 2.390 \\
\eta_4 2.850 \\
\eta_5 0.854
\end{array}\]

\[\begin{array}{c}
\eta_1 6.093 \\
\eta_2 2.581 \\
\eta_3 0.854 \\
\eta_4 1.566 \\
\eta_5 1.040
\end{array}\]

\[\begin{array}{c}
\eta_1 1.953 \\
\eta_2 1.040 \\
\eta_3 2.390 \\
\eta_4 2.850 \\
\eta_5 0.854
\end{array}\]

\[\begin{array}{c}
\eta_1 6.093 \\
\eta_2 2.581 \\
\eta_3 0.854 \\
\eta_4 1.566 \\
\eta_5 1.040
\end{array}\]
Table 4

The investment proportions of the selected stocks in the optimal portfolio of Model (20).

<table>
<thead>
<tr>
<th>μ₀</th>
<th>600008</th>
<th>600016</th>
<th>600027</th>
<th>600085</th>
<th>600271</th>
<th>600887</th>
<th>601299</th>
<th>η₁</th>
<th>η₂</th>
<th>η₃</th>
<th>η₄</th>
<th>η₅</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.20</td>
<td>1.97</td>
<td>7.86</td>
<td>6.67</td>
<td>3.34</td>
<td>7.76</td>
<td>7.46</td>
<td>6.47</td>
<td>18.86</td>
<td>12.04</td>
<td>8.45</td>
<td>12.09</td>
<td>7.22</td>
<td>26.04</td>
</tr>
<tr>
<td>0.30</td>
<td>2.04</td>
<td>8.84</td>
<td>4.99</td>
<td>3.69</td>
<td>7.88</td>
<td>8.76</td>
<td>6.39</td>
<td>18.37</td>
<td>12.13</td>
<td>6.89</td>
<td>12.89</td>
<td>6.57</td>
<td>27.14</td>
</tr>
<tr>
<td>0.40</td>
<td>2.11</td>
<td>10.02</td>
<td>3.30</td>
<td>4.03</td>
<td>8.00</td>
<td>10.05</td>
<td>6.31</td>
<td>17.88</td>
<td>12.21</td>
<td>6.66</td>
<td>13.70</td>
<td>5.91</td>
<td>28.55</td>
</tr>
<tr>
<td>0.50</td>
<td>2.19</td>
<td>11.20</td>
<td>1.62</td>
<td>4.38</td>
<td>8.12</td>
<td>11.35</td>
<td>6.23</td>
<td>17.39</td>
<td>12.39</td>
<td>6.46</td>
<td>13.98</td>
<td>5.26</td>
<td>30.27</td>
</tr>
<tr>
<td>0.60</td>
<td>2.25</td>
<td>12.35</td>
<td>0</td>
<td>4.70</td>
<td>8.25</td>
<td>12.66</td>
<td>6.16</td>
<td>16.89</td>
<td>12.47</td>
<td>6.32</td>
<td>14.16</td>
<td>4.59</td>
<td>32.28</td>
</tr>
<tr>
<td>0.70</td>
<td>1.97</td>
<td>12.89</td>
<td>0</td>
<td>4.56</td>
<td>8.57</td>
<td>14.45</td>
<td>6.25</td>
<td>16.02</td>
<td>12.33</td>
<td>6.06</td>
<td>16.23</td>
<td>3.66</td>
<td>34.65</td>
</tr>
<tr>
<td>0.80</td>
<td>1.69</td>
<td>13.43</td>
<td>0</td>
<td>4.42</td>
<td>8.90</td>
<td>16.23</td>
<td>6.34</td>
<td>15.16</td>
<td>12.30</td>
<td>5.76</td>
<td>17.15</td>
<td>2.72</td>
<td>37.40</td>
</tr>
<tr>
<td>0.90</td>
<td>1.41</td>
<td>13.97</td>
<td>0</td>
<td>4.28</td>
<td>9.23</td>
<td>18.02</td>
<td>6.43</td>
<td>14.30</td>
<td>12.26</td>
<td>5.46</td>
<td>18.06</td>
<td>1.78</td>
<td>40.54</td>
</tr>
<tr>
<td>1.00</td>
<td>0.99</td>
<td>14.50</td>
<td>0</td>
<td>4.01</td>
<td>9.53</td>
<td>20.23</td>
<td>6.44</td>
<td>13.23</td>
<td>12.00</td>
<td>5.16</td>
<td>19.04</td>
<td>0.42</td>
<td>44.11</td>
</tr>
<tr>
<td>1.10</td>
<td>0.37</td>
<td>14.98</td>
<td>0</td>
<td>3.54</td>
<td>9.79</td>
<td>23.03</td>
<td>6.34</td>
<td>10.48</td>
<td>11.40</td>
<td>4.86</td>
<td>20.08</td>
<td>0.08</td>
<td>48.23</td>
</tr>
<tr>
<td>1.20</td>
<td>0</td>
<td>15.38</td>
<td>0</td>
<td>2.96</td>
<td>10.00</td>
<td>25.96</td>
<td>6.14</td>
<td>7.76</td>
<td>10.65</td>
<td>4.56</td>
<td>21.15</td>
<td>0.00</td>
<td>53.02</td>
</tr>
<tr>
<td>1.30</td>
<td>0</td>
<td>15.71</td>
<td>0</td>
<td>2.34</td>
<td>10.18</td>
<td>28.81</td>
<td>5.88</td>
<td>4.97</td>
<td>9.88</td>
<td>4.26</td>
<td>22.22</td>
<td>0.00</td>
<td>58.51</td>
</tr>
<tr>
<td>1.40</td>
<td>0</td>
<td>16.04</td>
<td>0</td>
<td>1.72</td>
<td>10.36</td>
<td>31.67</td>
<td>5.62</td>
<td>2.19</td>
<td>9.11</td>
<td>3.96</td>
<td>23.29</td>
<td>0.00</td>
<td>64.70</td>
</tr>
<tr>
<td>1.50</td>
<td>0</td>
<td>16.25</td>
<td>0</td>
<td>0.93</td>
<td>10.44</td>
<td>34.81</td>
<td>5.22</td>
<td>0</td>
<td>8.02</td>
<td>3.66</td>
<td>24.31</td>
<td>0.00</td>
<td>71.59</td>
</tr>
<tr>
<td>1.60</td>
<td>0</td>
<td>15.91</td>
<td>0</td>
<td>0</td>
<td>10.16</td>
<td>38.97</td>
<td>4.19</td>
<td>0</td>
<td>5.60</td>
<td>3.36</td>
<td>25.17</td>
<td>0.00</td>
<td>79.47</td>
</tr>
<tr>
<td>1.70</td>
<td>0</td>
<td>15.45</td>
<td>0</td>
<td>0</td>
<td>9.82</td>
<td>42.97</td>
<td>2.94</td>
<td>0</td>
<td>2.85</td>
<td>3.06</td>
<td>25.97</td>
<td>0.00</td>
<td>88.57</td>
</tr>
<tr>
<td>1.80</td>
<td>0</td>
<td>14.98</td>
<td>0</td>
<td>0</td>
<td>9.48</td>
<td>46.97</td>
<td>1.69</td>
<td>0</td>
<td>0.11</td>
<td>2.76</td>
<td>26.77</td>
<td>0.00</td>
<td>98.91</td>
</tr>
<tr>
<td>1.90</td>
<td>0</td>
<td>12.84</td>
<td>0</td>
<td>0</td>
<td>7.79</td>
<td>52.60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.67</td>
<td>26.77</td>
<td>0.00</td>
<td>110.81</td>
</tr>
<tr>
<td>2.00</td>
<td>0</td>
<td>9.70</td>
<td>0</td>
<td>0</td>
<td>4.93</td>
<td>58.98</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.60</td>
<td>26.77</td>
<td>0.00</td>
<td>124.90</td>
</tr>
<tr>
<td>2.10</td>
<td>0</td>
<td>6.56</td>
<td>0</td>
<td>0</td>
<td>2.07</td>
<td>65.36</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.54</td>
<td>26.01</td>
<td>0.00</td>
<td>141.28</td>
</tr>
<tr>
<td>2.20</td>
<td>0</td>
<td>2.83</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>71.82</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.48</td>
<td>25.35</td>
<td>0.00</td>
<td>159.97</td>
</tr>
<tr>
<td>2.30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>78.89</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.42</td>
<td>24.70</td>
<td>0.00</td>
<td>181.72</td>
</tr>
<tr>
<td>2.40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>86.40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.36</td>
<td>24.00</td>
<td>0.00</td>
<td>209.00</td>
</tr>
<tr>
<td>2.50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>93.91</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.30</td>
<td>23.60</td>
<td>0.00</td>
<td>224.26</td>
</tr>
<tr>
<td>2.58</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.24</td>
<td>23.20</td>
<td>0.00</td>
<td>273.60</td>
</tr>
</tbody>
</table>

6. Conclusions

Different from separate indeterministic environment such as stochastic or fuzzy/uncertain situation, this paper considered the complex situation in which some securities possess a great deal of transaction data and the others are newly listed with insufficient data. The corresponding returns were assumed to be random variables and uncertain variables, respectively. This paper first deduced the formulas of variances of hybrid portfolio returns based on two assumptions, and then formulated mean-variance models. The analytical solutions were given in the case of no more than two newly listed securities. Other cases could be solved by using Matlab such as quadprog function. A numerical experiment was illustrated to the application of the proposed method. The computational results indicate that the proposed models are meaningful and able to be applied in reality.
Acknowledgments

This work was supported by National Natural Science Foundation of China (No. 71371019) and by the Program for New Century Excellent Talents in University (No. NCET-12-0026).

References