Portfolio selection based on fuzzy probabilities and possibility distributions

Hideo Tanaka\textsuperscript{a,}\textsuperscript{*}, Peijun Guo\textsuperscript{a}, I. Burhan Türksen\textsuperscript{b}

\textsuperscript{a}Department of Industrial Engineering, Osaka Prefecture University, 1-1 Gakuuencho, Sakai, Osaka 599-8531, Japan
\textsuperscript{b}Department of Industrial Engineering, University of Toronto, M5S 1A4, Canada

Received June 1997; received in revised form January 1998

Abstract

In this paper, two kinds of portfolio selection models are proposed based on fuzzy probabilities and possibility distributions, respectively, rather than conventional probability distributions in Markowitz’s model. Since fuzzy probabilities and possibility distributions are obtained depending on possibility grades of security data offered by experts, investment experts’ knowledge can be reflected. A numerical example of a portfolio selection problem is given to illustrate our proposed approaches. © 2000 Published by Elsevier Science B.V. All rights reserved.

1. Introduction

Multivariate data analysis is a main tool based on probability theory for analyzing the uncertainty in the real world. Possibility data analysis is an alternative based on possibility distributions. Multivariate data analysis considers the uncertainty as probability phenomena while possibility data analysis considers it as possibility phenomena. Possibility theory based on possibility distributions has been proposed by Zadeh [13] and advanced by Dubois and Prade [1]. Possibility distributions are represented as normal convex fuzzy sets, such as L-R fuzzy numbers, quadratic and exponential functions [3,7,8,10]. The theory of exponential possibility distributions has been proposed by Tanaka et al. and applied to data analysis [4–6,9,11]. In this paper, we propose two kinds of portfolio selection models. One is based on fuzzy probabilities, which can be regarded as a natural extension of Markowitz’s model because of extending probability into fuzzy probability. The other is based on possibility distributions. There are some similarities between the fuzzy probability and possibility models. However, these two kinds of models analyze the security data in very different ways.

Markowitz’s model is based on a probability distribution, while the possibility portfolio selection model is based on a possibility distribution that is used to characterize experts’ knowledge. A possibility distribution is identified using the returns of securities associated with possibility grades offered by portfolio experts. Based on the obtained possibility distribution, we construct a possibility portfolio selection model as a quadratic...
programming problem similar to Markowitz’s model. The possibility portfolio selection problem is to minimize the spread of possibility return of a portfolio subject to the given center return. Because experts’ knowledge is very valuable, it is reasonable that possibility portfolio models are useful in real investment environment. Using a fuzzy probability method and a possibility method, a numerical example of a portfolio selection problem is solved. The simulation results show that the possibility portfolio is more suitable than the fuzzy probability one in real investment problems.

2. Markowitz’s portfolio selection model

Let us give a brief description of Markowitz’s model. Assume that there are \( n \) securities denoted by \( S_j \) \((j = 1, \ldots, n)\), the return of the security \( S_j \) is denoted as \( r_j \) and the proportion of total investment funds devoted to this security is denoted as \( x_j \). Thus,

\[
\sum_{j=1}^{n} x_j = 1. \tag{1}
\]

Since \( r_j \) \((j = 1, \ldots, n)\) vary from time to time, those are assumed to be random variables which can be represented by the pair of the average vector and covariance matrix. For instance, it is assumed that the observation data on returns of securities over \( m \) periods are given. At the discrete time \( i \) \((i = 1, \ldots, m)\), \( n \) kinds of returns are denoted as a vector \( r_i = [r_{i1}, \ldots, r_{in}] \). Thus, the total data over \( m \) periods are denoted as the following matrix:

\[
\begin{pmatrix}
  r_{11} & r_{12} & \cdots & r_{1n} \\
  r_{21} & r_{22} & \cdots & r_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  r_{m1} & r_{m2} & \cdots & r_{mn}
\end{pmatrix}. \tag{2}
\]

The average vector of returns over \( m \) periods is denoted as \( \bar{r} = [r_{1}^0, \ldots, r_{n}^0] \) and is written as

\[
\bar{r} = \left[ \frac{\sum_{i=1}^{m} r_{i1}}{m}, \ldots, \frac{\sum_{i=1}^{m} r_{in}}{m} \right]. \tag{3}
\]

Also the corresponding covariance matrix \( Q = [q_{ij}^2] \) can be written as

\[
q_{ij}^2 = \sum_{k=1}^{m} (r_{ki} - \bar{r}_{i}^0)(r_{kj} - \bar{r}_{j}^0)/m \quad (i = 1, \ldots, n, \ j = 1, \ldots, n). \tag{4}
\]

Therefore, random variables can be represented by the average vector \( \bar{r} \) and the covariance matrix \( Q \), denoted as \((\bar{r}, Q)\).

Now, the return associated with the portfolio \( x \) is given by

\[
z = x^\top r. \tag{5}
\]

The average and variance of \( z \) is given as

\[
E(z) = E(x^\top r) = x^\top Er = x^\top \bar{r}, \tag{6}
\]

\[
V(z) = V(x^\top r) = x^\top Qx. \tag{7}
\]
Since the variance is regarded as the risk of investment, the best investment is one with the minimum variance (7) subject to a given average return $r_s$. This leads to the following quadratic programming problem:

$$\begin{align*}
\text{Minimize} & \quad x^TQx \\
\text{subject to} & \quad x^Tr^0 = r_s,
\sum_{i=1}^n x_i = 1, \quad x_i \geq 0.
\end{align*}$$

(8)

3. Fuzzy portfolio selection models

The data are given as $(r_i, h_i)$ ($i = 1, \ldots, m$) where $h_i$ is a possibility grade to reflect a similarity degree between the future state of stock markets and the state of the $i$th sample offered by experts. These grades $h_i$ ($i = 1, \ldots, m$) are regarded as weights to determine the fuzzy average vector and covariance matrix for the given data.

3.1. Definition of fuzzy probabilities

Given the data $(r_i, h_i)$ ($i = 1, \ldots, m$), the fuzzy weighted average vector $\mathbf{a} = [a_1, \ldots, a_n]^T$ can be defined as follows [12]:

$$\mathbf{a} = \frac{\sum_{i=1}^m (h_i r_i)}{\sum_{i=1}^m h_i}.$$  

(9)

Similarly, the fuzzy weight covariance matrix $\Sigma = [\sigma_{ij}]$ can be defined by

$$\sigma_{ij} = \frac{\left\{ \sum_{k=1}^m (r_{ki} - a_i)(r_{kj} - a_j)h_k \right\}}{\sum_{k=1}^m h_k} \quad (i = 1, \ldots, n, \ j = 1, \ldots, n).$$

(10)

Thus, the given data $(r_i, h_i)$ ($i = 1, \ldots, m$) can be summarized as parametric representation $(\mathbf{a}, \Sigma)$, which is used to construct the fuzzy portfolio selection model.

3.2. The fuzzy probability portfolio selection model

Given the weight average vector and covariance matrix, $(\mathbf{a}, \Sigma)$, the average and covariance of the return $z$ as in (5) are given as follows:

$$E(z) = x^T \mathbf{a},$$

(11)

$$V(z) = x^T \Sigma x.$$  

(12)

Thus, the fuzzy probability portfolio selection problem can be obtained as:

$$\begin{align*}
\text{Minimize} & \quad x^T \Sigma x \\
\text{subject to} & \quad x^T \mathbf{a} = r_s,
\sum_{i=1}^n x_i = 1, \quad x_i \geq 0.
\end{align*}$$  

(13)

It should be noted that the average vector and covariance matrix in Markowitz’s model are replaced by the weight-average vector and covariance, respectively, in which the expert judgment $h_i$ is contained. The curve obtained from the optimization problem (13) is called a fuzzy efficient frontier.
It is obvious that when the given data have the same important grades, the fuzzy probability portfolio selection model is just equivalent to Markowitz’s model. It means that the fuzzy probability portfolio selection model (13) is an extension of the conventional portfolio selection model (8).

4. Two identification methods of possibility distributions

Let us reconsider the given data \((r_i, h_i) (i = 1, \ldots, m)\) as in Section 3 \((0 \leq h_i < 1)\). Assume that these grades \(h_i (i = 1, \ldots, m)\) can be expressed by a possibility distribution \(A\) defined as

\[
\Pi_A(r) = \exp \{- (r - a)^T D_A^{-1} (r - a)\} = (a, D_A)_e,
\]

where \(a\) is a center vector and \(D_A\) is a symmetric positive-definite matrix.

Given data \((r_i, h_i) (i = 1, \ldots, m)\), let us describe how to identify an exponential distribution as (14), i.e., a center vector \(a\) and a positive-definite matrix \(D_A\). The center vector \(a\) can be determined by (9). For simplification, the following transformation is given

\[
y = r - a.
\]

Then (14) can be rewritten as

\[
\Pi_A(y) = \exp \{- y^T D_A^{-1} y\}.
\]

In order to determine \(D_A\), the following are assumed:

1. \(\Pi_A(y_i) \geq h_i, i = 1, \ldots, m\) (the constraint conditions),
2. minimize \(\Pi_A(y_1) \times \Pi_A(y_2) \times \cdots \times \Pi_A(y_m)\) (the cost function).

In other words, the problem can be described as determining \(D_A\) that minimizes the cost function (II) subject to the constraint conditions (I). Fig. 1 illustrates the notion of our identification method in the one-dimensional case.

More detailed, the assumption (I) can be written as follows:

\[
\Pi_A(y_i) \geq h_i \iff y_i^T D_A^{-1} y_i \leq - \ln h_i, \quad i = 1, \ldots, m.
\]

Similarly, the objective function can be rewritten as follows:

\[
\max \sum_{i=1}^{m} y_i^T D_A^{-1} y_i.
\]

Thus, the problem for obtaining \(D_A\) becomes the following optimization problem:

\[
\max_{D_A} \sum_{i=1}^{m} y_i^T D_A^{-1} y_i \quad \text{s.t.} \quad y_i^T D_A^{-1} y_i \leq - \ln h_i, \quad i = 1, \ldots, m, \quad D_A > 0,
\]

where \(D_A > 0\) means that \(D_A\) is a positive-definite matrix. It is obvious that (19) is a nonlinear optimization problem which is difficult to be solved.

In order to transform the nonlinear optimization problem (19) into a linear programming (LP) problem, let us introduce orthogonal conditions. The following auxiliary conditions are added to (19) instead of the
condition $D_A > 0$:

$$y_i^T D_A^{-1} y_i \geq \varepsilon, \quad i \in E, \quad (20)$$

$$y_i^T D_A^{-1} y_j = 0 \quad \text{for all } j \neq i \text{ and } i,j \in E, \quad (21)$$

where $E = \{1, \ldots, n\}$ is the set of subscripts of independent vectors $\{y_1, y_2, \ldots, y_n\}$ among $m$ kinds of given input vectors ($m \gg n$) and $\varepsilon$ is a very small positive value.

Thus, the following LP problem is formed:

$$\begin{align*}
\text{Max} & \quad \sum_{i=1}^{m} y_i^T D_A^{-1} y_i \\
\text{s.t.} & \quad y_i^T D_A^{-1} y_i \leq -\ln h_i, \quad i = 1, \ldots, m, \\
& \quad y_i^T D_A^{-1} y_i \geq \varepsilon, \quad i \in E, \\
& \quad y_i^T D_A^{-1} y_j = 0 \quad \text{for all } i \neq j \text{ and } i,j \in E. \\
\end{align*} \quad (22)$$

This identification procedure is called the orthogonal method. In what follows, let us prove that the matrix $D_A$ obtained from (22) is a positive-definite matrix. Because $\{y_1, y_2, \ldots, y_n\}$ are independent vectors in the $n$-dimensional space, an arbitrary vector $z$ can be represented as

$$z = \lambda_1 y_1 + \lambda_2 y_2 + \cdots + \lambda_n y_n, \quad (23)$$

where $\lambda_i$ is a real number. Thus, using (20) and (21), we have for $z \neq 0$,

$$z^T D_A^{-1} z = (\lambda_1 y_1 + \lambda_2 y_2 + \cdots + \lambda_n y_n)^T \lambda_1 y_1 + \lambda_2 y_2 + \cdots + \lambda_n y_n \\
= \sum_{i=1}^{n} \lambda_i^2 y_i^T D_A^{-1} y_i > 0, \quad (24)$$

which means that $D_A > 0$. 
Here, the orthogonal conditions are added to constraint conditions to confine $D_A$ to a positive-definite matrix. However, since there are many selections among independent vectors as the orthogonal conditions, it is very hard to select appropriate orthogonal conditions.

To cope with this difficulty, we use principle component analysis (PCA) to rotate the given data $(y_i, h_i)$ ($i = 1, \ldots, m$). Fig. 2 illustrates the rotation of orthogonal axes. The data can be transformed by the linear transformation $T$. Columns of $T$ should be eigenvectors of $\Sigma$ defined in (10). Without loss of generality, we assume that $\text{rank}(\Sigma) = n$. It should be noted that $TT^t = I$, where $I$ is an identity matrix.

Using the linear transformation, the data $\{y_i\}$ can be transformed into $\{z = Ty\}$. Then we have

$$\Pi_A(z) = \exp\{-z^t T^t D_A^{-1} Tz\}. \tag{25}$$

According to the feature of PCA, we can assume that $T^t D_A^{-1} T$ is a diagonal matrix as follows:

$$T^t D_A^{-1} T = C_A = \begin{pmatrix} c_1 & 0 \\ \vdots & \ddots \\ 0 & \cdots & c_n \end{pmatrix}. \tag{26}$$

Using $C_A$, the constraint conditions can be written as

$$\Pi_A(z_i) \geq h_i \iff c_i C_A z_i \leq -\ln h_i, \quad i = 1, \ldots, m. \tag{27}$$

Similarly, we have the following performance function to be maximized:

$$\sum_{i=1}^{m} c_i C_A z_i. \tag{28}$$

Thus, the problem for obtaining $C_A$ leads to the following LP problem:

$$\begin{align*}
\text{Max} & \quad \sum_{i=1}^{m} c_i C_A z_i \\
\text{s.t.} & \quad c_i C_A z_i \leq -\ln h_i, \quad i = 1, \ldots, m, \\
& \quad c_j \geq c, \quad j = 1, \ldots, n,
\end{align*} \tag{29}$$

where $c$ is a small positive number and $c_j \geq c$ makes $C_A$ positive definite.

It is very easy to obtain $C_A$ by solving the LP problem (29). After obtaining $C_A$, we can have

$$D_A = (TC_A T^t)^{-1} = TC_A^{-1} T^t, \tag{30}$$

which is the obtained possibility distribution matrix of (14). This identification procedure is called the PCA method.

It is true that an optimal solution of (29) always exists. In order to show it, let us take $C = qI$ in (29), where $I$ is an $n \times n$ unity matrix. Thus, the constraint conditions of (29) become

$$q c_i z_i \leq -\ln h_i, \quad i = 1, \ldots, m, \quad q \geq c, \tag{31}$$

which means that if we take $c \leq q \leq \min_{i=1,\ldots,m} (-\ln h_i/c_i z_i)$ ($z_i \neq 0$), (31) can hold. Therefore, there is an admissible set in the constraint conditions of (29).

PCA method is simpler than the orthogonal method so that it is used in portfolio selection problems.
5. Possibility portfolio selection model

Assume that the returns $r_i$ ($i = 1, \ldots, m$) are governed by a possibility distribution. The possibility distribution denoted as $\Pi_A(r) = (a_i, D_A)$ is obtained by PCA method. The possibility return of a portfolio $x = [x_1, \ldots, x_n]^T$ can be written as

$$z = x^T r.$$  

(32)

The possibility distribution of $Z$, denoted as $\Pi_Z(z)$, can be defined by the extension principle as follows:

$$\Pi_Z(z) = \text{Max}_{\{r|z=x^T r\}} \Pi_A(r).$$  

(33)

Solving the simple optimization problem (33), we have

$$\Pi_Z(z) = \exp\{-(z - x^T a)^2 \cdot (x^T D_A x)^{-1}\},$$  

(34)

where $x^T a$ is the center value and $x^T D_A x$ is the spread of the possibility return $Z$. Following Markowitz’s model, the following possibility portfolio selection model is given:

$$\begin{align*}
\text{Min}_x & \quad x^T D_A x \\
\text{s.t.} & \quad x^T a = r_s, \quad \sum_{i=1}^{n} x_i = 1, \quad x_i \geq 0,
\end{align*}$$  

(35)

which is also a quadratic programming problem minimizing the spread of possibility return subject to a given center return $r_s$. The curve obtained from the quadratic programming problem (35) is called possibility efficient frontier. The points in this frontier are nondominated solutions with considering two objective functions, i.e., the spread and the center of a possibility portfolio return.

6. Numerical example

In order to illustrate the proposed methods, let us consider an example shown in Table 1 introduced by Markowitz [2]. Since we consider that the recent sample is more similar to the future state, it is assumed that the possibility grade $h_i$ is defined by

$$h_i = 0.1 + 0.3(t - 1)/17 \quad (i = 1, \ldots, 18).$$  

(36)

In Table 1, Nos. 1–9 are American Tobacco, AT&T, United States Steel, General Motors, Atchison&Topeka &Santa Fe, Coca-Cola, Borden, Firestone and Sharon Steel, respectively. The return of the security $S_j$ during a year $t$ is defined to be

$$r_{t,j} = (p_{t+1,j} + d_{t,j} - p_{t,j})/p_{t,j},$$  

(37)

where $p_{t,j}$ is the closing price of the security $S_j$ in the year $t$ and $d_{t,j}$ is its dividends in year $t$. 

Using (10) we obtained the fuzzy covariance matrix as follows:

\[
\Sigma = \begin{bmatrix}
0.04042 & 0.01593 & 0.01933 & 0.03277 & 0.01046 & 0.02665 & 0.0183 & 0.02816 & 0.02077 \\
0.01593 & 0.01119 & 0.01701 & 0.01875 & 0.00771 & 0.00705 & 0.00912 & 0.02106 & 0.01518 \\
0.01933 & 0.01701 & 0.09852 & 0.06287 & 0.04545 & 0.00694 & 0.0035 & 0.07752 & 0.03114 \\
0.03277 & 0.01875 & 0.06287 & 0.07954 & 0.04465 & 0.02072 & 0.00904 & 0.07894 & 0.01964 \\
0.01046 & 0.00771 & 0.04545 & 0.04465 & 0.09713 & 0.00585 & 0.01096 & 0.086 & 0.0285 \\
0.02665 & 0.00705 & 0.00694 & 0.02072 & 0.00585 & 0.0405 & 0.00953 & 0.01952 & 0.01095 \\
0.0183 & 0.00912 & 0.0035 & 0.00904 & 0.01096 & 0.00953 & 0.01986 & 0.01784 & 0.0079 \\
0.02816 & 0.02106 & 0.07752 & 0.07894 & 0.086 & 0.01952 & 0.01784 & 0.13343 & 0.03455 \\
0.02077 & 0.01518 & 0.03114 & 0.01964 & 0.0285 & 0.01095 & 0.0079 & 0.03455 & 0.06211 
\end{bmatrix}
\]

Using (29) and (30) the possibility distribution \( D_A \) is obtained as follows:

\[
D_A = \begin{bmatrix}
618.74 & -136.82 & -40.2 & -154.4 & 237.87 & -227.49 & -267.42 & -78.87 & -73.76 \\
-136.82 & 51.75 & 5.15 & 83.18 & 20.01 & 35.88 & 125.82 & -75.11 & 13.55 \\
-40.2 & 5.15 & 435.82 & 226.47 & -273.73 & -166.28 & -140.08 & -137.99 & -226.64 \\
-154.4 & 83.18 & 226.47 & 267.87 & -20.89 & -71.91 & 143.3 & -275.17 & -108.93 \\
237.87 & 20.01 & -273.73 & -20.89 & 478.23 & -32.67 & 205.95 & -253.99 & 93.76 \\
-227.49 & 51.75 & -166.28 & -71.91 & -32.67 & 172.09 & 115.12 & 160.31 & 127.91 \\
-267.42 & 125.82 & -140.08 & 143.3 & 205.95 & 115.12 & 372 & -200.3 & 101.22 \\
-78.87 & -75.11 & -137.99 & -275.17 & -253.99 & 160.31 & -200.3 & 471.78 & 108.06 \\
-73.76 & 13.55 & -226.64 & -108.93 & 93.76 & 127.91 & 101.22 & 108.06 & 134.35 
\end{bmatrix}
\]
Using the fuzzy probability model (13) and the possibility model (35), we obtained two portfolio efficient frontiers. Results are shown in Fig. 3 where the variance and the spread of a portfolio return are used to represent the risk of investment for fuzzy probability and possibility portfolio selection problems, respectively. It is seen from Fig. 3 that the spread of the return of a possibility portfolio is larger than the variance of the return of a fuzzy probability portfolio. In some sense it can be said that possibility models consider more risk than fuzzy probability ones.

Fig. 4 shows the securities selected by fuzzy probability and possibility portfolio models in the case of \( r_s = 0.17 \) and shows that the fuzzy portfolio has five securities while the possibility portfolio has six securities. It means that the possibility portfolio tends to take more distributive investment than the fuzzy probability one. From the simulation results, we know that possibility portfolio selection models are more reasonable than fuzzy probability ones in real investment problems.
7. Conclusions

We have proposed two kinds of portfolio selection approaches based on fuzzy probabilities and possibility distributions. The remarkable distinction between our models and Markowitz’s model can be said as follows: Markowitz’s model deals with the securities data according to the statistic viewpoints, while we try to model investment experts’ knowledge. In detail, both of fuzzy probability and possibility portfolio selection models are attempts to deal with the possibility grades associated with security data. The possibility grade $h_i$ offered by experts reflects experts’ knowledge representing the similarity degree between the future state of stock markets and the $i$th sample. The average vector and covariance matrix in Markowitz’s model are replaced with the weighted average vector and covariance matrix by the experts’ judgement $h_i$. For possibility portfolio models the security data associated with possibility grades are used to construct a possibility distribution. Based on a fuzzy probability and a possibility distribution, portfolios are selected to minimize the variance of the return of a portfolio in a fuzzy probability model and the spread of the return of a portfolio in a possibility model, respectively. Because the experts’ judgment in data analysis is very important, possibility analysis methods will be suitable for such an attempt shown in this paper.
References