数理解析研究所講究録
1043巻 1998年 10-18

Possibility Analysis and Its Applications

大阪府立大学工学部 田中 英夫 (Hideo Tanaka)
郭 沛俊 (Peijun Guo)

1. Introduction

The theory of possibility has been proposed by Zadeh [1], where fuzzy variables are associated with possibility distributions in the same way as random variables are associated with probability distributions [2]. The possibility measure and its dual measure call the necessity measure play important roles in establishing the possibility theory [3]. Condition possibility distributions have been defined in different forms in [4,5]. In [4], the consistency among joint, conditional and marginal possibility distributions is considered. In [5], a conditional possibility is obtained as an interval.

On the other hand, Dempster [6] introduced upper and lower probabilities, which do not satisfy the additivity, and Shafer [7] has interpreted Dempster's work as a theory of evidence. The possibility and necessity measures are special kinds of belief and plausibility measures discussed in [8]. In Dempster-Shafer Theory, the evidence is represented by the basic probability assignment, and the combination rule of evidence is discussed. The theory is applied to obtain a certainty factor in chaining syllogism as a belief interval associated with the composition of chained rules [9].

This paper studied a certain form of evidence theory by exponential possibility distributions. Because possibility distributions are obtained from an expert knowledge, a possibility distribution is regarded as a representation of evidence in this paper. A rule of combination of evidence is given similar to Dempster's rule [6]. Also, the measures of ignorance and fuzziness of evidence are defined by a normality factor and the area of a possibility distribution, respectively. The measures of ignorance is similar to the weight of conflict by Shafer[7], and the measure of fuzziness is the same as one defined by Kaufman and Gupta [10]. Next, marginal and conditional possibilities are defined from a joint possibility distribution, and it is shown that these three definitions are well match to each other. Thus, the posterior possibility distribution is derived from the prior possibility distribution in the same form as Bayes' formula. This fact means that an information-decision theory can be reconstructed from the viewpoint of possibility distributions. Furthermore, linear systems whose variables are defined by possibility distributions are discussed. It shows that the operations of fuzzy vectors defined by multidimensional possibility distributions are well formulated using the extension principle of Zadeh [11].

As applications of possibility analysis, possibilistic regression [12, 13] and possibility portfolio selection [14,15] are formulated in this paper. In regression analysis, it is assumed that regression models are possibilistic linear systems defined by exponential possibility distributions. Thus, the problem of possibilistic regression is to determine an exponential possibility distribution of parametric vector in a model, which reflect the scatting of the given input-output data. In other words, the spread of the given data is transformed into the possibility distribution of the parametric vector in the model.

In portfolio selection problems, it is assumed that security data have associated with possibility grades given by experts. Those data can be described as two types of exponential possibility distributions. In other words, the upper and lower possibility distributions can be identified. Thus, we propose two types of portfolio selection models based on two types of distributions.

In numerical examples of two problems mentioned above are shown in this paper to illustrate our proposed methods.

2. Possibility distribution and its properties

An exponential possibility distribution is regarded as a representation of evidence in this paper. A kind of evidence is represented by an exponential possibility distribution as

$$\Pi_A(\mathbf{x}) = \exp\{-(\mathbf{x}-\mathbf{a})^t \mathbf{D}_A (\mathbf{x}-\mathbf{a})\}, \tag{1}$$

where the evidence is denoted as A, $\mathbf{a}$ is a center vector and $\mathbf{D}_A$ is a symmetrical positive definite matrix. The parametric representation of A is written as follows.

$$A = (\mathbf{a}, \mathbf{D}_A)_e \tag{2}$$

It should be noted that $\Pi_A(\mathbf{x})$ is normal, that is, there is an x such that $\Pi_A(\mathbf{x}) = 1$. Let us assume that A' is not normal. Thus, A' is given as

$$\Pi_{A'}(\mathbf{x}) = c\exp\{-(\mathbf{x}-\mathbf{a})'\mathbf{D}_{A'}(\mathbf{x}-\mathbf{a})\}, \tag{3}$$

where 0<c<1.

Definition 1: Let a measure of ignorance of A' denoted as $I(A')$ be defined by

$$I(A') = -\log c \tag{4}$$

It can be seen from Definition 1 that the possibility distribution given by (1) has no ignorance. The possibility distributions expressed by (1) are dealt with throughout this paper. Thus, given the evidence A' expressed by (3), A' should be normalized to obtain a normal possibility A with $I(A)$, i.e.,

$$\Pi_A(\mathbf{x}) = \Pi_{A'}(\mathbf{x})/c; \;\; I(A) = -\log c. \tag{5}$$

Thus, it should be noted that the given evidence A' is denoted as $\Pi_A(\mathbf{x})$ with $I(A)$.

Definition 2: Let a measure of fuzziness of A denoted as $H(A)$ be defined by

$$H(A) = \int_{-\infty}^{\infty} \exp\{-(\mathbf{x}-\mathbf{a})'\mathbf{D}_A(\mathbf{x}-\mathbf{a})d\mathbf{x} \tag{6}$$

The characteristic of an evidence A can be represented as

$$\{(\mathbf{a},\mathbf{D}_A)_e, I(A), H(A)\} \tag{7}$$

It follows from the definition (6) that

i) $H(A) = \pi^{n/2} |\mathbf{D}_A^{-1}|^{1/2}$, (8)

ii) If $\mathbf{D}_A \geq \mathbf{D}_B > 0$, $H(A) \leq H(B)$. (9)

Let us define a combination rule of possibility distributions from a similar view to Dempster's rule [7].

Definition 3: Let $A_1 \oplus A_2$ denote the combination of possibility distributions $A_1 = (\mathbf{a}_1, \mathbf{D}_1)_e$ and $A_2 = (\mathbf{a}_2, \mathbf{D}_2)_e$. Then the combination rule is defined as,

$$\Pi_{(A_1 \oplus A_2)} = k\Pi_{A_1} \cdot \Pi_{A_2} \tag{10}$$

where k is a normalizing factor such that

$$\max_{\mathbf{x}} \Pi_{(A_1 \oplus A_2)}(\mathbf{x}) = 1. \tag{11}$$

It is clear from Definition 1 that the measure of ignorance of $A_1 \oplus A_2$ is given by

$$I(A_1 \oplus A_2) = \log k, \tag{12}$$

which is similar to the measure of conflict defined by shafer [7]. $I(A_1 \oplus A_2)$ can be regarded as the weight of conflict between A_1 and A_2.

In order to obtain $\Pi_{(A_1 \oplus A_2)}(\mathbf{x})$, we must solve the optimization problem described in the left-hand side of (11). Thus, we have

$$\mathbf{x}^* = (\mathbf{D}_1 + \mathbf{D}_2)^{-1}(\mathbf{D}_1\mathbf{a}_1 + \mathbf{D}_2\mathbf{a}_2). \tag{13}$$

Substitute $\mathbf{x}^*$ into (10), we have

$$k\exp\{-p\} = 1,$$

where

$$p = -(\mathbf{D}_1\mathbf{a}_1 + \mathbf{D}_2\mathbf{a}_2)'(\mathbf{D}_1 + \mathbf{D}_2)^{-1}(\mathbf{D}_1\mathbf{a}_1 + \mathbf{D}_2\mathbf{a}_2) + \mathbf{a}_1'\mathbf{D}_1\mathbf{a}_1 + \mathbf{a}_2'\mathbf{D}_2\mathbf{a}_2 \tag{14}$$

Thus, we have

$$k = \exp\{p\}. \tag{15}$$

Substituting k into (10) yields

$$\Pi_{A_1 \oplus A_2}(\mathbf{x}) = ((\mathbf{D}_1 + \mathbf{D}_2)^{-1}(\mathbf{D}_1\mathbf{a}_1 + \mathbf{D}_2\mathbf{a}_2), (\mathbf{D}_1 + \mathbf{D}_2))_e. \tag{16}$$

From (15), the measure of ignorance of $A_1 \oplus A_2$ can be written as

$$I(A_1 \oplus A_2) = \mathbf{a}_1'\mathbf{D}_1\mathbf{a}_1 + \mathbf{a}_2'\mathbf{D}_2\mathbf{a}_2 - (\mathbf{D}_1\mathbf{a}_1 + \mathbf{D}_2\mathbf{a}_2)'(\mathbf{D}_1 + \mathbf{D}_2)^{-1}(\mathbf{D}_1\mathbf{a}_1 + \mathbf{D}_2\mathbf{a}_2) \tag{17}$$

In general, the possibility distribution of $A_1 \oplus \cdots \oplus A_n$ can be obtained in the following theorem.

Theorem 1: The combination of n possibility distributions A_i, i=1,...,n can be represented as the following exponential possibility distribution.

$$A_1 \oplus \cdots \oplus A_n = ((\sum_{i=1}^{n} \mathbf{D}_i)^{-1}(\sum_{i=1}^{n} \mathbf{D}_i \mathbf{a}_i), \sum_{i=1}^{n} \mathbf{D}_i)_e \,. \tag{18}$$

The measure of ignorance of $A_1 \oplus \cdots \oplus A_n$ is

$$I(A_1 \oplus \cdots \oplus A_n) = \sum_{i=1}^{n} \mathbf{a}_i' \mathbf{D}_i \mathbf{a}_i - (\sum_{i=1}^{n} \mathbf{D}_i \mathbf{a}_i)' (\sum_{i=1}^{n} \mathbf{D}_i)^{-1} (\sum_{i=1}^{n} \mathbf{D}_i a_i) \,. \tag{19}$$

The measure of fuzziness of $A_1 \oplus \cdots \oplus A_n$ is

$$H(A_1 \oplus \cdots \oplus A_n) = \pi^{n/2} \,|(\sum_{i=1}^{n} \mathbf{D}_i)^{-1}|^{1/2} \,. \tag{20}$$

Marginal and conditional possibility distributions are defined from the given joint possibility distribution. It is shown that these definitions are well matched to each other from the viewpoint of the given joint possibility distribution.

Let an exponential joint possibility distribution on the $(n+m)$ dimensional space be

$$\Pi_A(\mathbf{x}, \mathbf{y}) = \exp\{-(\mathbf{x} - \mathbf{a}_1, \mathbf{y} - \mathbf{a}_2)' \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{12} & \mathbf{D}_{22} \end{bmatrix} (\mathbf{x} - \mathbf{a}_1, \mathbf{y} - \mathbf{a}_2)\} \,, \tag{21}$$

where $\mathbf{x}$ and $\mathbf{y}$ are n and m dimensional vectors, respectively, $\mathbf{a}_1$ and $\mathbf{a}_2$ are center vectors in n and m dimensional spaces, i.e., X and Y, respectively and the positive definite matrix $\mathbf{D}_A$ is divided into 4 matrices as written in (21).

Definition 4: The marginal possibility distribution on X is defined by

$$\Pi_A(\mathbf{x}) = \max_{\mathbf{y}} \Pi_A(\mathbf{x}, \mathbf{y}) \,. \tag{22}$$

The marginal possibility distribution obtained from the exponential distribution (22) can be represented by

$$\Pi_A(\mathbf{x}) = (\mathbf{a}_1, \mathbf{D}_{11} - \mathbf{D}_{12}\mathbf{D}_{22}^{-1}\mathbf{D}_{12}')_e \,, \tag{23}$$

which is proved as follows. The maximization problem shown in (22) can be reduced to the following minimization problem:

$$\min_{\mathbf{y}} (\mathbf{x} - \mathbf{a}_1)' \mathbf{D}_{11} (\mathbf{x} - \mathbf{a}_1) + 2(\mathbf{y} - \mathbf{a}_2)' \mathbf{D}_{12}' (\mathbf{x} - \mathbf{a}_1) + (\mathbf{y} - \mathbf{a}_2) \mathbf{D}_{22} (\mathbf{y} - \mathbf{a}_2) \,. \tag{24}$$

The optimal solution $\mathbf{y}^*$ of (24) is

$$\mathbf{y}^* = -\mathbf{D}_{22}^{-1}\mathbf{D}_{12}'(\mathbf{x} - \mathbf{a}_1) + \mathbf{a}_2 \tag{25}$$

Substituting (25) into (24) yields (23).

Definition 5: Given the joint possibility distribution (21), the conditional distribution given by $\mathbf{y}$ is defined by

$$\Pi_A(\mathbf{x} \,|\, \mathbf{y}) = k \exp\{-(\mathbf{x} - \mathbf{a}_1, \mathbf{y} - \mathbf{a}_2)' \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{12} & \mathbf{D}_{22} \end{bmatrix} (x - a_1, y - a_2)\} \,, \tag{26}$$

where k is a normalizing factor such that

$$\max_{\mathbf{x}} \Pi_A(\mathbf{x} \,|\, \mathbf{y}) = 1 \,. \tag{27}$$

By solving the optimization problem (27), $\Pi_A(\mathbf{x} \,|\, \mathbf{y})$ can be written as

$$\Pi_A(\mathbf{x} \,|\, \mathbf{y}) = (\mathbf{a}_1 - \mathbf{D}_{11}^{-1}\mathbf{D}_{12}(\mathbf{y} - \mathbf{a}_2), \mathbf{D}_{11})_e \,, \tag{28}$$

which is proved as follows. Consider the problem for obtaining a normalizing factor k. The maximization problem (27) can be reduced to the following minimization problem:

$$\min_{\mathbf{x}} (\mathbf{x} - \mathbf{a}_1)' \mathbf{D}_{11} (\mathbf{x} - \mathbf{a}_1) + 2(\mathbf{y} - \mathbf{a}_2)' \mathbf{D}_{12}' (\mathbf{x} - \mathbf{a}_1) \,. \tag{29}$$

The optimal solution of (29) is

$$\mathbf{x}^* = -\mathbf{D}_{11}^{-1}\mathbf{D}_{12}(\mathbf{y} - \mathbf{a}_2) + \mathbf{a}_1 \,. \tag{30}$$

By substituting $\mathbf{x}^*$ into (26) and setting $\Pi_A(\mathbf{x}^* \,|\, \mathbf{y}) = 1$, we have

$$k = \exp\{(\mathbf{y} - \mathbf{a}_2)'(\mathbf{D}_{22} - \mathbf{D}_{12}'\mathbf{D}_{11}^{-1}\mathbf{D}_{12})(\mathbf{y} - \mathbf{a}_2)\} \,. \tag{31}$$

Substituting (31) into (26) leads to (28).

In what follows, let us show that the marginal and conditional distributions derived from the joint distribution are consistent with each other.

Theorem 2: The following relation holds.

$$\Pi_A(\mathbf{x}\,|\,\mathbf{y})\Pi_A(\mathbf{y}) = \Pi_A(\mathbf{x},\mathbf{y}), \tag{32}$$

where $\Pi_A(\mathbf{x}\,|\,\mathbf{y})$ and $\Pi_A(\mathbf{y})$ are derived from $\Pi_A(\mathbf{x},\mathbf{y})$ by **Definition 4** and **5**.

From **Theorem 2**, we know that

$$\Pi_A(\mathbf{x}\,|\,\mathbf{y}) = \Pi_A(\mathbf{y}\,|\,\mathbf{x})\cdot\Pi_A(\mathbf{x})/\Pi_A(\mathbf{y}), \tag{33}$$

which is just the same as Bayes' theorem. Suppose that $\Pi_A(\mathbf{y}\,|\,\mathbf{x})$ is a possibilistic information system, where **x** and **y** are associated with cause and effect, respectively. Assuming the prior possibility distribution $\Pi_A(\mathbf{x})$, we can calculate the posterior possibility distribution $\Pi_A(\mathbf{x}\,|\,\mathbf{y})$ given **y**, using (33). Thus, we can expect to construct an information-decision problem from the viewpoint of possibility [16].

3. Possibilistic linear systems

The possibilistic linear system with exponential distributions are defined by the following extension principle. Let $f: X\times Y \to Z$ be a vector function. Given exponential possibility distributions $\Pi_A(\mathbf{x})$ and $\Pi_B(\mathbf{y})$ on X and Y, respectively, the possibility distribution $\Pi_{f(A,B)}(\mathbf{z})$ on Z is defined as

$$\Pi_{f(A,B)}(\mathbf{z}) = \max_{\{\mathbf{x},\mathbf{y}|\mathbf{z}=f(\mathbf{x},\mathbf{y})\}} \Pi_A(\mathbf{x})\cdot\Pi_B(\mathbf{y}). \tag{34}$$

Suppose that a linear vector function is given as $\mathbf{y}=\mathbf{Tx}$, where $\mathbf{T}$ is an $m\times n$ matrix, $n>m$ and rank[T]=m. The possibility distribution derived from an exponential possibility distribution $\Pi_A(\mathbf{x}) = (\mathbf{a},\mathbf{D_A})_e$ is denoted as

$$\Pi_{\mathbf{T}A}(\mathbf{y}) = \max_{\{\mathbf{x}|\mathbf{y}=\mathbf{Tx}\}} \exp\{-(\mathbf{x}-\mathbf{a})'\mathbf{D_A}(\mathbf{x}-\mathbf{a})\}. \tag{35}$$

In other words, when a fuzzy vector A is used instead of **x**, the fuzzy output vector Y is denoted as

$$\mathrm{Y} = \mathbf{T}A, \tag{36}$$

where the possibility distribution of Y is regarded as $\Pi_{\mathbf{T}A}(y)$. By solving the optimization problem (35), we have

$$\Pi_{\mathbf{T}A}(\mathbf{y}) = \exp\{-(\mathbf{y}-\mathbf{Ta})'(\mathbf{T}\mathbf{D}_\mathbf{A}^{-1}\mathbf{T}^t)^{-1}(\mathbf{y}-\mathbf{Ta})\} = (\mathbf{Ta},(\mathbf{T}\mathbf{D}_\mathbf{A}^{-1}\mathbf{T}^t)^{-1})_e. \tag{37}$$

Suppose that

$$\mathbf{z} = \lambda_1\mathbf{x}+\lambda_2\mathbf{y}, \tag{38}$$

where λ_1 and λ_2 are constants, and **x** and **y** are governed by possibility distribution $\Pi_A(\mathbf{x}) = (\mathbf{a}_1,\mathbf{D}_1)_e$ and $\Pi_A(\mathbf{y}) = (\mathbf{a}_2,\mathbf{D}_2)_e$, respectively. The fuzzy vector of (38) is expressed as

$$Z = \lambda_1 A + \lambda_2 B, \tag{39}$$

whose distribution is denoted as $\Pi_Z(\mathbf{z})$. By solving (34), $\Pi_Z(\mathbf{z})$ can be represented as

$$\Pi_Z = (\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2, \lambda_1^2\mathbf{D}_1+\lambda_2^2\mathbf{D}_2)_e \tag{40}$$

4. Possibility linear regression

Let us define a possibility linear system as

$$\mathrm{Y} = A_1x_1 + A_2x_2 + \ldots\ldots + A_nx_n = A\mathbf{x}, \tag{41}$$

where **x** is a real vector(input vector) and A=$(A_1, A_2,\ldots\ldots,A_n)$ is a fuzzy coefficient vector defined by $(\mathbf{a}_c,\mathbf{D}_A{}^{-1})_e$. The possibility distribution $\Pi_\mathrm{Y}(y)$ of (41) can be obtained as

$$\Pi_\mathrm{Y}(y) = \exp\{-(y-\mathbf{x}_j^t\mathbf{a}_c)^2(\mathbf{x}_j^t\mathbf{D}_A\mathbf{x}_j)^{-1}\} = (\mathbf{x}_j^t\mathbf{a}_c,(\mathbf{x}_j^t\mathbf{D}_A\mathbf{x}_j)^{-1})_e \tag{42}$$

Given data $(y_i,\mathbf{x}_i)$, i=1,...,m, let us formulate possibility regression analysis with the following assumptions:

1) $y_i \in [\mathrm{Y}]_{h_i} = \{y\,|\,\Pi_\mathrm{Y}(y_i)\ge h_i\}$,
2) $J = \sum \mathbf{x}_j^t\mathbf{D}_A\mathbf{x}_j \longrightarrow$ minimize (An index of the spread of (42)),
3) $\mathbf{D}_A > 0$.

This formulation is given as follows:

$$\min_{\mathbf{D}_A,\mathbf{a}_c} \quad J=\sum \mathbf{x}_j^t \mathbf{D}_A \mathbf{x}_j$$

$$\text{S.T.} \quad \mathbf{x}_j^t \mathbf{D}_A \mathbf{x}_j \ge (y_j - x_j^t \mathbf{a}_c)^2/(-\ln h_i),\ j=1,\dots,m. \qquad (43)$$
$$\mathbf{D}_A > 0.$$

The center vector $\mathbf{a}_c$ is obtained by a conventional regression method and the condition $\mathbf{D}_A > 0$ is replace by a sufficient condition as follows: Given a real symmetric matrix $[d_{ij}] \in R^{n \times n}$, the sufficient condition for this matrix being positive definite can be described as

$$d_{ii} > \sum_{j=1, j\neq i}^{n} |d_{ij}|,\ (i=1,\dots,n). \qquad (44)$$

In the LP problems, $|d_{ij}|$ will be replaced by $d_{ij}^+ + d_{ij}^-$ because $d_{ij}^+ + d_{ij}^- \ge |d_{ij}^+ - d_{ij}^-| = |d_{ij}|$ where $d_{ij}^+ \ge 0$ and $d_{ij}^- \ge 0$. This is a direct corollary of Gersgorin theorem.

5. Possibility portfolio selection

Given security data (χ_i, h_i), i=1,...,m, where $\mathbf{x}_i = [x_{i1}, \dots\dots x_{in}]^t$ is a vector of returns of securities S_i (i=1,...,n) for the ith period. The h_i is an associated possibility grade given by an expert knowledge. Assume that h_i (i=1,...,m) are expressed by a possibility distribution $(\mathbf{a}, \mathbf{D_A}^{-1})_e$.

Given the data, the upper and the lower possibility distributions denoted as Π_u and Π_l, respectively, are identified to satisfy $\Pi_u(x) \ge \Pi_l(x)$,which is similar to rough sets concept.

The center vector $\mathbf{a}$ can be approximately estimated as $\mathbf{a}=\mathbf{x}_i$ whose possibility degree $h_i = \underset{j=1,\dots,m}{Max}\, h_j$ and the associated possibility degree h_i is revised to be 1. Take $\mathbf{y}_i = \mathbf{x}_i - \mathbf{a}$ and denote $\mathbf{D}_A^{-1}$ as $\mathbf{D}_u$ and $\mathbf{D}_l$ corresponding to the upper and lower distributions, respectively. Let us identify the upper and lower possibility distributions with the following assumptions.

1) $\Pi_l(\mathbf{y}_i) \le h_i$, i=1,...,m
2) $\Pi_u(\mathbf{y}_i) \ge h_i$, i=1,...,m
3) $\Pi_u(\mathbf{y}) \ge \Pi_l(\mathbf{y})$
4) $\Pi_l(\mathbf{y}_1) \times \dots\dots \times \Pi_l(\mathbf{y}_m) \to$ maximize
5) $\Pi_u(\mathbf{y}_1) \times \dots\dots \times \Pi_u(\mathbf{y}_m) \to$ minimize

The upper and lower distributions can be obtained by the following optimization problem.

$$\underset{\mathbf{D}_u, \mathbf{D}_l}{\text{Min}} \quad \sum_{i=1}^{m} \mathbf{y}_i^t \mathbf{D}_l \mathbf{y}_i - \sum_{i=1}^{m} \mathbf{y}_i^t \mathbf{D}_u \mathbf{y}_i$$

$$\text{subject to} \quad \mathbf{y}_i^t \mathbf{D}_u \mathbf{y}_i \le -\ln \mathrm{h}_i, \qquad (45)$$
$$\mathbf{y}_i^t \mathbf{D}_l \mathbf{y}_i \ge -\ln \mathrm{h}_i,\ i=1,\dots\dots m,$$
$$\mathbf{D}_u^{-1} - \mathbf{D}_l^{-1} \ge 0.$$

$\mathbf{D}_u^{-1} - \mathbf{D}_l^{-1} \ge 0$ is a nonlinear constraint condition. It is difficult to solve the problem (44). Principle component analysis (**PCA**) is used to rotate the given data $(\mathbf{y}_i, h_i)$ (i=1,...,m) to obtain a positive definite matrix easily. The data can be transformed by the linear transformation **T.** The columns of **T** are eigenvectors of the matrix $\Sigma = [\sigma_{ij}]$ of given data. It should be noted that $\mathbf{T}^t\mathbf{T} = \mathbf{I}$. σ_{ij} is defined as follows:

$$\sigma_{ij} = \{ \sum_{k=1}^{m} (x_{ki} - a_i)(x_{kj} - a_j) h_k \} / \sum_{k=1}^{m} h_k . \qquad (46)$$

It is assumed that $rank[T] = n$. Using the linear transformation, the data $\{\mathbf{y}_i\}$ can be transformed into $\{\mathbf{z}_i = \mathbf{T}^t \mathbf{y}_i\}$. Then we have

$$\Pi_A(\mathbf{z}_i) = \exp\{-\mathbf{z}_i^t \mathbf{T}^t \mathbf{D}_A^{-1} \mathbf{T} \mathbf{z}_i\}. \qquad (47)$$

According to the feature of **PCA**, $\mathbf{T}'\mathbf{D}_A^{-1}\mathbf{T}$ is assumed to be a diagonal matrix as follows:

$$\mathbf{T}'\mathbf{D}_A^{-1}\mathbf{T}=\mathbf{C}_A=\begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix}. \tag{48}$$

Denote $\mathbf{C}_A$ as $\mathbf{C}_u$ and $\mathbf{C}_l$ for the upper and the lower possibility distribution cases, respectively. The corresponding diagonal elements in $\mathbf{C}_u$ and $\mathbf{C}_l$ are denoted as c_{uj} and c_{lj} ($j=1,...,n$), respectively. The integrated model can be rewritten as follows.

$$\begin{aligned} \min_{\mathbf{C}_l,\mathbf{C}_u} \quad & \sum_{i=1}^{m}\mathbf{z}_i'\mathbf{C}_l\mathbf{z}_i - \sum_{i=1}^{m}\mathbf{z}_i'\mathbf{C}_u\mathbf{z}_i \\ \text{subject to} \quad & \mathbf{z}_i'\mathbf{C}_l\mathbf{z}_i \ge -\ln h_i, \\ & \mathbf{z}_i'\mathbf{C}_u\mathbf{z}_i \le -\ln h_i, \\ & c_{uj} \ge \varepsilon \\ & c_{lj} \ge c_{uj}, \\ & i=1,...,m,\ j=1,...,n, \end{aligned} \tag{49}$$

where $c_{lj} \ge c_{uj} \ge \varepsilon > 0$ ($j=1,...,n$) make $\mathbf{D}_u - \mathbf{D}_l$ semi-positive definite and $\mathbf{D}_u$ and $\mathbf{D}_l$ positive. Thus, we have

$$\begin{aligned} \mathbf{D}_u &= \mathbf{T}\mathbf{C}_u^{-1}\mathbf{T}', \\ \mathbf{D}_l &= \mathbf{T}\mathbf{C}_l^{-1}\mathbf{T}'. \end{aligned} \tag{50}$$

Theorem 3: The upper and the lower possibility distribution matrices in (48) always exist.

Theorem 4: Assume the data $(\mathbf{y}_i, h_i)$ $i=1,\ldots,m$, are governed by an exponential possibility distribution $(0,\mathbf{D})_e$. The optimal solutions of $\mathbf{D}_u$ and $\mathbf{D}_l$ in (44) are $\mathbf{D}$.

A possibility return of a portfolio Z can be written as $Z=\sum_{j=1}^{n} r_j x_j$, where r_j denotes the proportion of the total investment funds devoted to the security S_j. Thus, Z is a possibility variable with the following possibility distribution:

$$\Pi_Z = \exp\{-(z-\mathbf{r}'\mathbf{a})^2(\mathbf{r}'\mathbf{D}_A\mathbf{r})^{-1}\}, \tag{51}$$

where $\mathbf{r}'\mathbf{a}$ is a center value and $\mathbf{r}'\mathbf{D}_A\mathbf{r}$ is a spread of Z.

Portfolio selection models based on upper possibility distributions (j=u) and on lower possibility distributions (j=*l*)

$$\begin{aligned} \min_{\mathbf{r}} \quad & \mathbf{r}'\mathbf{D}_j\mathbf{r} \\ \text{subject to} \quad & \mathbf{r}'\mathbf{a} = c, \\ & \sum_{i=1}^{n} r_i = 1, \end{aligned} \tag{52}$$

where c is an expected center value of possibility return rate.

Theorem 5: The spread of the possibility return obtained by the lower model is not larger than the one obtained by the upper model.

The curves from upper and lower possibility models are called possibility frontiers I and II, respectively.

5. Numerical example

5.1 Possibility regression

The input-output data are shown in Table 1. The possibility linear system is

$$\mathrm{Y} = A_1x_1 + A_2x_2 = A\mathbf{x}. \tag{53}$$

The center vector α_c^* is determined by the conventional regression analysis as $\alpha_c^* = (3.7885, 0.4097)'$.

By solving (43), the matrix is obtained as

$$\mathbf{D}_A^{-1} = \begin{bmatrix} 2.93537 & 0.02637 \\ 0.02637 & 0.02637 \end{bmatrix}. \tag{54}$$

Fig. 1 shows the contour line (h=0.5) of the obtained possibility distribution of parameters and Fig. 2 shows the possibility regression model and the given data.

Table 1: input-output data

i	1	2	3	4	5	6	7	8	9	10
x_i	2	4	6	9	12	13	14	16	29	20
y_i	4	7	5	8	7	9	12	9	14	10

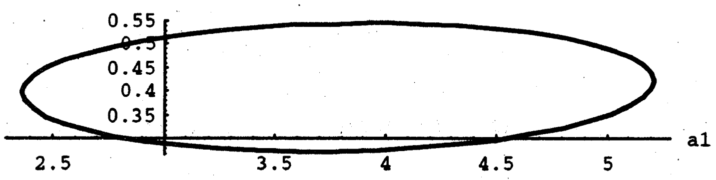

Fig.1 Possibility distribution of with h=0.5

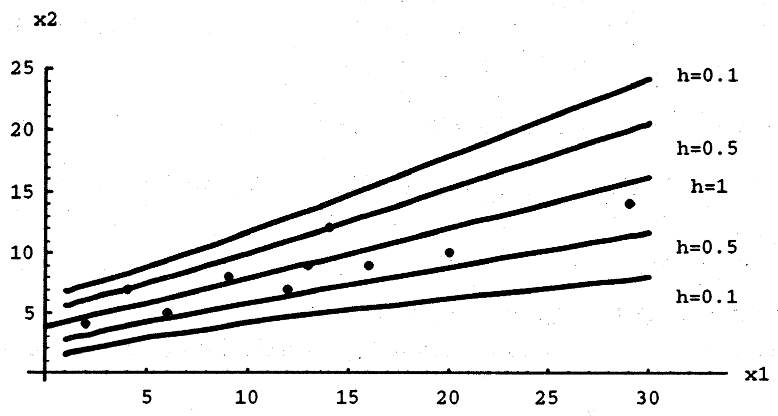

Fig.2 Possibility regression model and the given data

5.2 Possibility portfolio selection

Let us consider the security data shown in Table 2. In order to show the concept of upper and lower possibility distributions graphically let us just consider two securities, namely, Am.T. and At&T in Table 2. The upper and lower possibility are obtained from (49) and (50) as follows:

$\mathbf{a}$ =[0.154,0.176],

$$\mathbf{D}_u^{-1} = \begin{bmatrix} 0.2665 & 0.0972 \\ 0.0972 & 0.1689 \end{bmatrix}, \quad \mathbf{D}_l^{-1} = \begin{bmatrix} 0.0313 & 0.0165 \\ 0.0165 & 0.0148 \end{bmatrix}. \tag{55}$$

Fig. 3 shows the upper and lower possibility distribution of these two securities with h=0.5. From (49) and (50), we can also obtain the upper and lower possibility distributions of nine securities in Table 2. Then, the possibility efficient portfolio frontiers I and II can be obtained from (52) and shown in Fig. 4.

Table 2. Returns on nine securities and possibility grades

hi	year	#1 Am.T	#2 A.T.&T.	#3 U.S.S.	#4 G.M.	#5 A.T.&S.	#6 C.C	#7 Bdn.	#8 Frstn.	#9 S.S.
0.2	1937(1)	-0.305	-0.173	-0.318	-0.477	-0.457	-0.065	-0.319	-0.4	-0.435
0.241	1938(2)	0.513	0.098	0.285	0.714	0.107	0.238	0.076	0.336	0.238
0.282	1939(3)	0.055	0.2	-0.047	0.165	-0.424	-0.078	0.381	-0.093	-0.295
0.324	1940(4)	-0.126	0.03	0.104	-0.043	-0.189	-0.077	-0.051	-0.09	-0.036
0.365	1941(5)	-0.28	-0.183	-0.171	-0.277	0.637	-0.187	0.087	-0.194	-0.24
0.406	1942(6)	-0.003	0.067	-0.039	0.476	0.865	0.156	0.262	1.113	0.126
0.447	1943(7)	0.428	0.3	0.149	0.225	0.313	0.351	0.341	0.58	0.639
0.488	1944(8)	0.192	0.103	0.26	0.29	0.637	0.233	0.227	0.473	0.282
0.529	1945(9)	0.446	0.216	0.419	0.216	0.373	0.349	0.352	0.229	0.578
0.571	1946(10)	-0.088	-0.046	-0.078	-0.272	-0.037	-0.209	0.153	-0.126	0.289
0.612	1947(11)	-0.127	-0.071	0.169	0.144	0.026	0.355	-0.099	0.009	0.184
0.653	1948(12)	-0.015	0.056	-0.035	0.107	0.153	-0.231	0.038	0.000	0.114
0.694	1949(13)	0.305	0.038	0.133	0.321	0.067	0.246	0.273	0.223	-0.222
0.735	1950(14)	-0.096	0.089	0.732	0.305	0.579	-0.248	0.091	0.65	0.327
0.776	1951(15)	0.016	0.09	0.021	0.195	0.04	-0.064	0.054	-0.131	0.333
0.818	1952(16)	0.128	0.083	0.131	0.39	0.434	0.079	0.109	0.175	0.062
0.859	1953(17)	-0.01	0.035	0.006	-0.072	-0.027	0.067	0.21	-0.084	-0.048
0.9	1954(18)	0.154	0.176	0.908	0.715	0.469	0.077	0.112	0.756	0.185

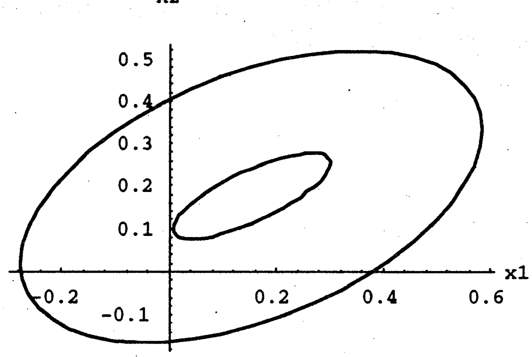

Fig. 3 Upper and lower possibility distributions with h=0.5

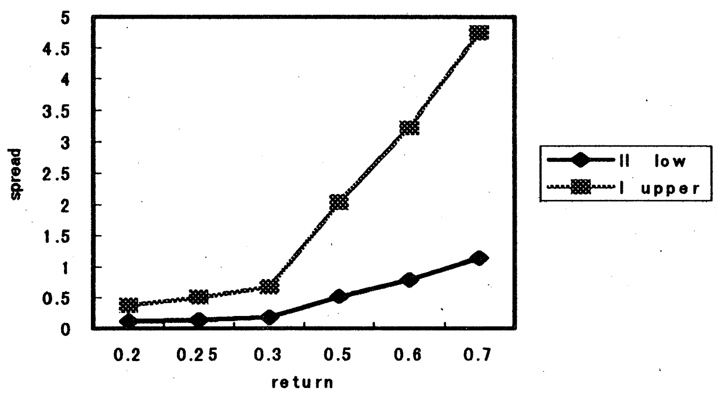

Fig. 4 Possibility efficient portfolio frontiers I and II

References:

[1] Zadeh, L. A., Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems* 1 (1978) 3-28.
[2] Nahmias, S., Fuzzy variables, *Fuzzy Sets and Systems* 1 (1978) 97-110.
[3] Dubois, D., and Prade, H., *Possibility Theory* (Plenum Press, New York, 1988).
[4] Nguyen, H. T., On conditional possibility distributions, Fuzzy Sets and Systems 1 (1978) 283-297.
[5] Hisdal, E., conditional possibility, independence and noninteraction, Fuzzy Sets and Systems, 1 (1978) 283-297.
[6] Dempster, A. P. Upper and lower probabilities induced by a multivalued mapping, Anal. Math. Statist. 38 (1967) 447-511.
[7] Shafer, G., *A mathematical Theory of Evidence*, (Princeton University Press, 1976) 447-511.
[8] Tanaka, H., Fuzzy Modelling and Its Applications (Asakura, 1990) 106-112. (In Japanese)
[9] Hau, H. Y. and Kashyap, R. L. Belief combination and propagation in a lattice-structures inference networks, *IEEE Trans. On SMC* 20 (1990) 45-58
[10] Kaufman, A., and Gupta, M. M., *Fuzzy Mathematical models in engineering and management Science*, (North-Holland,1988) 47-48.
[11] Zadeh, L. A., The concept of a linguistic variable and its application to approximate reasoning-I. *Information Science*, 8 (1975) 199-249.
[12] Tanaka, H., Hayashi, I. and Watada, J. Possibilistic linear regression analysis for fuzzy data, *European Journal of Operational Research* 40 (1989) 389-396.
[13] Tanaka, H., Ishibuchi, H. and Yoshikawa, S. Exponential possibility regression analysis, *Fuzzy Sets and Systems* 69 (1995) 305-318.
[14] Tanaka, H., Guo, P. and Turksen, I. B. Portfolio selection based on fuzzy probability distributions and possibility distributions, Fuzzy Sets and Systems (to appear)
[15] Tanaka, H. and Guo, P., Portfolio selection based on upper and lower exponential possibility distributions, *European Journal of Operational Research* (to appear)
[16] Tanaka, H., and Ishibuchi, H., Evidence theory of exponential possibility distributions, *International Journal of Approximate Reasoning* 8 (1993) 123-140.