日本ファジィ学会誌 Vol.12, No.1, pp.127-132(2000)

論文

感性評価データに対する ファジィ回帰モデル†

中森 義輝* 領家 美奈*

同一の入力に対して複数の異なる出力データが存在する場合におけるファジィ線形回帰分析法を提案する。複数の評価者が商品，環境，あるいは人間を評価するような場合がこれにあたる。本稿では，出力データ集合における評価者間の位置関係をできるだけモデルのパラメータ空間においても保存するような方法の提案をおこなう。田中による可能性線形モデルの同定法を参考にしつつも，線形計画問題に帰着しない，きわめて簡単で見通しのよいアプローチである。

1. はじめに

ファジィ回帰モデルは田中ら[1]により提案されたもので，現在までにいくつかの定式化と同定法が開発されている[2][3]。ここでは，やはり田中ら[4][5]による指数型可能性分布による回帰分析を参考にして，同一の入力に対して複数の異なる出力データが存在する場合におけるファジィ線形回帰分析法を提案する。複数の評価者が，たとえば商品，環境，あるいは人間を評価するような場合がこれにあたる。

本稿では，出力データ集合における評価者間の位置関係をできるだけモデルのパラメータ空間においても保存するような方法を提案する。田中による可能性線形モデルの同定法を参考にしつつも，線形計画問題に帰着しない，きわめて簡単で見通しのよいアプローチである。なお，数量化理論I類への適用も同様におこなえる。本来は説明変数も質的あるいはカテゴリカルな場合の分析が目的であるが，従来法との比較を重視し，本稿では回帰分析を取り上げる。

2. 定義・記号・データ

本研究の目的は，感性評価データの分析手法の確立である。ここでは，目的が明確となるように定義と記号を導入する。

2.1 定義

・商品の評価であれば「見栄えがする」「豪華な」，環境の評価であれば「快適である」「自然とのふれあいを感じる」，また，人間の評価であれば「頼もしい」「愉快な」などという言葉をここでは感性ワードと呼ぶ。厳密にはワードではないが，感性ワードとは主観的評価尺度である。

・商品，地域，人間などの対象に対する，感性ワードによる主観的・感覚的な評価値をここでは感性データと呼ぶ。通常は5段階評価値などの極めてあいまいな数値である。また，同一の対象に対して評価者により異なる評価が与えられるのが特徴である。

2.2 記号

・評価対象(商品サンプル，地域環境など)

$m=1,2,\cdots,M$

・感性ワード(目的変数，外的基準，評価尺度)

$n=1,2,\cdots,N$

・評価者(しばしば被験者と呼ばれる)

$k=1,2,\cdots,K$

・対象の物理的特性(説明変数，デザイン要素)

$i=1,2,\cdots,I$

2.3 データ

・感性データ(目的変数の値；3相データ)

y_{mnk}

本稿では実数連続値として一般的に定式化するが，感性評価データは5段階評価値などの整数値のケースが多く，その場合は，尺度の等間隔性などについて事前の検討が要求される。

・対象の物理的データ(説明変数の値)

x_{mi}

本稿では実数連続値として一般的に定式化するが，感性評価対象は，たとえば，色であれば赤，青，緑のようにカテゴリーデータとして与えられるケースが多い。この場合は数量化理論I類と呼ばれる回帰分析を行うことになるが，データ行列のランク落ちの取り扱い以外は，提案する方法をそのまま適用することができる。

---

† Fuzzy Regression Models for Subjective Evaluation Data
Yoshiteru NAKAMORI and Mina RYOKE

* 北陸先端科学技術大学院大学 知識科学研究科
Graduate School of Knowledge Science, Japan Advanced Institute of Science and Technology

3. 可能性回帰モデル

田中によって提案されている指数型可能性分布による回帰分析を概観する。ただし，本節では目的変数のデータは3相ではなく

$$\{y_{mn}\},\ m=1,2,\cdots,M,\ n=1,2,\cdots,N \tag{1}$$

と与えられているものとする。また，以下ではある一つの感性ワードに注目し，添え字の n を省略する。

3.1 可能性線形システム

ファジィベクトル係数をもつファジィ線形モデル

$$y=a_0+a_1x_1+\cdots+a_Ix_I=\boldsymbol{a}^t\ \boldsymbol{x} \tag{2}$$

を考える。ただし，

$$\boldsymbol{a}=(a_0,a_1,a_2,\cdots,a_I)^t \tag{3}$$

$$\boldsymbol{x}=(x_0,x_1,x_2,\cdots,x_I)^t,\ x_0=1 \tag{4}$$

可能性線形モデルをファジィベクトル A を用いて

$$Y=A\ \boldsymbol{x} \tag{5}$$

と表わす。

ベクトル係数 $\boldsymbol{a}$ は次式で与えられる可能性分布に従っているものとする。

$$\mu_A(\boldsymbol{a})=\exp\{-\|\boldsymbol{a}-\hat{\boldsymbol{a}}\|^2_{D_A^{-1}}\} \tag{6}$$

D_Aはメンバシップ関数の拡がりを規定する正定値対称行列であり，

$$\|\boldsymbol{a}-\hat{\boldsymbol{a}}\|^2_{D_A^{-1}}=(\boldsymbol{a}-\hat{\boldsymbol{a}})^tD_A^{-1}(\boldsymbol{a}-\hat{\boldsymbol{a}}) \tag{7}$$

である。

ファジィ出力 Y の可能性分布は，拡張原理を適用すれば

$$\mu_Y(y)=\max_{\{\boldsymbol{a}|y=\boldsymbol{a}^t\boldsymbol{x}\}}\mu_A(\boldsymbol{a})$$
$$=\exp\{-(y-\hat{\boldsymbol{a}}^t\boldsymbol{x})^2(\boldsymbol{x}^tD_A\boldsymbol{x})^{-1}\} \tag{8}$$

によって与えられる[5]。

正定値対称行列 D_Aは，田中らの方法においては以下のように同定している[6]。

1. 与えられた入力データ $\boldsymbol{x}_m$に対して，ファジィ出力は次式で与えられる。

$$Y_m=A\boldsymbol{x}_m \tag{9}$$

2. 各データ y_mが Y_mに属する度合い h_mを与え，つぎの関係を要求する。

$$\mu_{Y_m}(y_m)\ge h_m,\ m=1,2,\cdots,M \tag{10}$$

3. D_Aが正定値行列になるようにつぎの拘束条件を導入する。

$$\boldsymbol{x}^t_m D_A\boldsymbol{x}_{m'}=0,\ m\neq m',\ m,m'\in E \tag{11}$$

ただし，E は独立な入力データベクトルの集合である。

4. ファジィ出力 Y_mの拡がりの和

$$J(\hat{\boldsymbol{a}},D_A)=\sum_{m=1}^{M}\boldsymbol{x}^t_m D_A\boldsymbol{x}_m \tag{12}$$

を最小化するように $\hat{\boldsymbol{a}}$, D_A を求める。

実際の同定においては，通常の回帰分析から得られる回帰係数を中心ベクトル $\hat{\boldsymbol{a}}$ とし，つぎの線形計画問題を解いて正定値行列を求めている。

$$\text{Minimize}\ \sum_{m=1}^{M}\boldsymbol{x}^t_m D_A\boldsymbol{x}_m$$
$$\text{subject to}\ \mu_{Y_m}(y_m)\ge h_m,\ m=1,2,\cdots,M$$
$$\boldsymbol{x}^t_m D_A\boldsymbol{x}_{m'}=0,\ m\neq m',\ m,m'\in E$$

3.2 数値例

4節で提案する方法と比較するため，田中らの方法による数値例を文献[6]から引用する。

文献[6]においては，表1によって与えられる1入力1出力の入出力データを用いて可能性回帰モデル

$$Y=A_0+A_1x \tag{13}$$

を同定している。

表1 入出力データ(田中・石渕[6])

m	1	2	3	4	5	6	7	8	9	10
x_m	2	4	6	9	12	13	14	16	19	20
y_m	4	7	5	8	7	9	12	9	14	10

まず，中心ベクトルは最小2乗法によって

$$\hat{\boldsymbol{a}}=(3.7885,\ 0.4097)^t \tag{14}$$

と与え，すべての m について

$$\mu_{Y_m}(y_m)\ge 0.5 \tag{15}$$

を要求し，直交条件を考慮した線形計画問題を解いて

$$D_A=\begin{pmatrix}5.4446 & -0.3769\\ -0.3769 & 0.0712\end{pmatrix} \tag{16}$$

を得ている。

係数の可能性分布と可能性回帰モデルを図示すれば，それぞれ図1，図2のようになる。

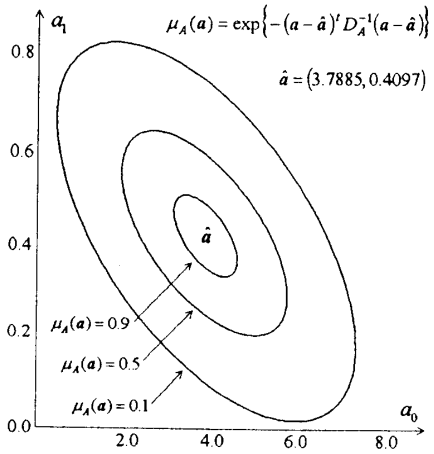


図 1　係数の可能性分布(田中・石渕[6])

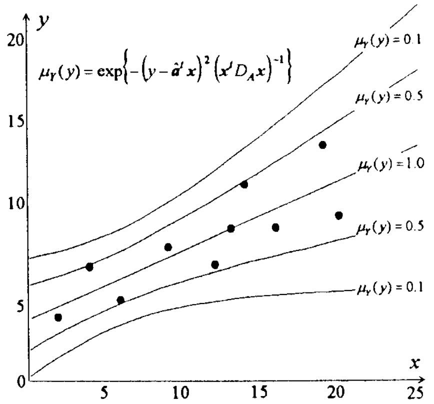


図 2　可能性回帰モデル(田中・石渕[6])

4．感性評価データに対する回帰分析

本節においては，目的変数のデータは

$$\{y_{mk}\},\quad m=1,2,\cdots,M,\ k=1,2,\cdots,K \tag{17}$$

であるとする。

4.1　平均データによる回帰分析

目的変数データの評価者についての平均データ

$$y_m = \frac{1}{K}\sum_{k=1}^{K} y_{mk},\quad m=1,2,\cdots,M \tag{18}$$

を用いた回帰モデルを

$$y_m = a_0 + \sum_{i=1}^{I} a_i x_{mi} + e_m,\quad m=1,2,\cdots,M \tag{19}$$

とする。ここで以下の記号を導入する。

$$\boldsymbol{y} = (y_1, y_2, \cdots, y_M)^t \tag{20}$$

$$X = \left\{\begin{array}{ccccc} 1 & x_{11} & x_{12} & \cdots & x_{1I} \\ 1 & x_{21} & x_{22} & \cdots & x_{2I} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{M1} & x_{M2} & \cdots & x_{MI} \end{array}\right\} \tag{21}$$

$$\boldsymbol{a} = (a_0, a_1, \cdots, a_I)^t \tag{22}$$

$$\boldsymbol{e} = (e_1, e_2, \cdots, e_M)^t \tag{23}$$

X^tX が正則であれば，回帰係数は最小 2 乗法

$$\|\boldsymbol{e}\|^2 \rightarrow \min. \tag{24}$$

によりつぎのように求められる。

$$\hat{\boldsymbol{a}} = (X^tX)^{-1}X^t\boldsymbol{y} \tag{25}$$

$$\hat{\boldsymbol{e}} = \boldsymbol{y} - X\hat{\boldsymbol{a}} \tag{26}$$

4.2　評価者データの写像

ここで，各評価者による目的変数の評価値

$$\boldsymbol{y}_k = (y_{1k}, y_{2k}, \cdots, y_{Mk})^t \tag{27}$$

を，次式によりパラメータ空間に写像する。

$$\hat{\boldsymbol{a}}_k = (X^tX)^{-1}X^t(\boldsymbol{y}_k - \hat{\boldsymbol{e}}) \tag{28}$$

このとき，以下の等式が成り立つことが確認できる。

$$\frac{1}{K}\sum_{k=1}^{K} \hat{\boldsymbol{a}}_k = \hat{\boldsymbol{a}} \tag{29}$$

$$\|\boldsymbol{y}_k - \boldsymbol{y}\|^2_{X(X^tX)^{-1}X^t} = \|\hat{\boldsymbol{a}}_k - \hat{\boldsymbol{a}}\|^2_{X^tX} \tag{30}$$

すなわち，

主張 1：評価者全員の回帰パラメータの平均は，平均データによる回帰パラメータに一致する。

主張 2：距離を上式のノルムにより定義したとき，評価データの評価者間の分散共分散構造はパラメータ空間においても保存される。

ここで，(30)式は以下のように導出される。(28)式に(26)式を代入して，

$$\hat{\boldsymbol{a}}_k = (X^tX)^{-1}X^t(\boldsymbol{y}_k - \hat{\boldsymbol{y}}) + \hat{\boldsymbol{a}} \tag{31}$$

したがって，

$$X(\hat{\boldsymbol{a}}_k - \hat{\boldsymbol{a}}) = X(X^tX)^{-1}X^t(\boldsymbol{y}_k - \hat{\boldsymbol{y}}) \tag{32}$$

これから，

$$\begin{aligned}&(\hat{\boldsymbol{a}}_k - \hat{\boldsymbol{a}})^t X^t X(\hat{\boldsymbol{a}}_k - \hat{\boldsymbol{a}}) \\ &= (\boldsymbol{y}_k - \hat{\boldsymbol{y}})^t X(X^tX)^{-1}X^t(\boldsymbol{y}_k - \hat{\boldsymbol{y}})\end{aligned} \tag{33}$$

ところで，平均データによるモデル式

$$\boldsymbol{y} = X\hat{\boldsymbol{a}} + \hat{\boldsymbol{e}} \tag{34}$$

の右辺に $\hat{\boldsymbol{a}}_k$ を代入したものを

$$\hat{\boldsymbol{y}}_k = X\hat{\boldsymbol{a}}_k + \hat{\boldsymbol{e}} \tag{35}$$

とおくとき，

$$\|\hat{\boldsymbol{y}}_k - \boldsymbol{y}_k\|^2 \tag{36}$$

を最小化するものが(28)式で与えられる

$$\hat{\boldsymbol{a}}_k = (X^tX)^{-1}X^t(\boldsymbol{y}_k - \hat{\boldsymbol{e}}) \tag{37}$$

である。すなわち，$\hat{\boldsymbol{a}}_k$ は(36)式を最小化する意味で，データベクトル $\boldsymbol{y}_k$ の近似値を提供するものである。

4.3 ファジィ線形モデル

このようにして，回帰係数はパラメータ空間においてファジィベクトル A として認識される。A のメンバシップ関数を(6)式と同様に

$$\mu_A(\boldsymbol{a})=\exp\{-\|\boldsymbol{a}-\hat{\boldsymbol{a}}\|^2_{D_A^{-1}}\} \tag{38}$$

と定義すれば，ファジィ線形モデル

$$Y=A\boldsymbol{x} \tag{39}$$

の出力メンバシップ関数は(8)式と同様に

$$\begin{aligned}\mu_Y(y)&=\max_{\{\boldsymbol{a}\,\|\,y=\boldsymbol{a}^t\boldsymbol{x}\}}\mu_A(\boldsymbol{a})\\&=\exp\{-(y-\hat{\boldsymbol{a}}^t\boldsymbol{x})^2(\boldsymbol{x}^tD_A\boldsymbol{x})^{-1}\}\end{aligned} \tag{40}$$

によって与えられる。

正定値対称行列 D_Aは，田中らの方法[6]においては，各データ y_mに対する拘束を導入し，線形計画問題を解くことにより求められる。本稿の提案によれば，すでにファジィベクトル A としてデータのばらつきがパラメータ空間に写像されているので，ファジィベクトル A のメンバシップ関数を直接同定すればよい。たとえば，

1. $(I+1)$次元空間におけるパラメータ値の分散共分散行列(必要あれば正則化)に基づき，
2. 最終的には，Y の拡がりに対するモデラーの判断により決定する。

以上により，田中の可能性線形モデルとは異なるファジィ線形モデルを得る。

ところで，提案手法によれば分散共分散行列がランク落ちするケースがある。その場合には，小さな正数を対角成分に加えるなどの正則化が必要となる。いずれにしても，D_Aの同定においてはモデラーの主観が入らざるを得ない。

高次元のパラメータ空間における D_Aの決定については検討が必要であるが，以下のような方法が考えられる。パラメータベクトルの分散共分散行列を S_Aとし，

$$D_A=cS_A \tag{41}$$

とおく。すべての k に対して

$$\mu_A(\hat{\boldsymbol{a}}_k)\geq h \tag{42}$$

を満たすように，パラメータ c を選択する。この場合，h の与え方に主観が入ることになる。

4.4 数値例

表2は同じ説明変数に対して5人が目的変数の値を回答したケースである。5人の回答の平均 $y_1, y_2, \cdots, y_{10}$は表1の出力データと一致させている。

表 2 入力データと複数の出力データ

m	1	2	3	4	5
x_m	2	4	6	9	12
y_m	4	7	5	8	7
y_{m1}	2.0	8.0	4.0	9.0	6.5
y_{m2}	5.5	4.5	6.0	10.0	3.5
y_{m3}	5.0	6.5	7.0	6.0	8.5
y_{m4}	4.5	8.5	3.5	6.5	9.0
y_{m5}	3.0	7.5	4.5	8.5	7.5
m	6	7	8	9	10
x_m	13	14	16	19	20
y_m	9	12	9	14	10
y_{m1}	10.0	11.0	12.0	14.5	12.5
y_{m2}	11.5	14.5	11.0	15.5	10.5
y_{m3}	7.5	12.5	7.5	15.0	9.0
y_{m4}	9.5	10.5	8.0	13.0	8.5
y_{m5}	6.5	11.5	6.5	12.0	9.5

平均データによる回帰係数は(14)式と同じく

$$\hat{\boldsymbol{a}}=(3.7885,\ 0.4097)^t \tag{43}$$

であり，(28)式により5人のデータをパラメータ空間に写像した回帰係数ベクトルは以下のように得られる。

$$\boldsymbol{a}_1=(2.4908,\ 0.5617)^t \tag{44}$$
$$\boldsymbol{a}_2=(3.6351,\ 0.4883)^t \tag{45}$$
$$\boldsymbol{a}_3=(4.4056,\ 0.3517)^t \tag{46}$$
$$\boldsymbol{a}_4=(4.4771,\ 0.3194)^t \tag{47}$$
$$\boldsymbol{a}_5=(3.9342,\ 0.3275)^t \tag{48}$$

図 4 5人の回帰係数と回帰係数のメンバシップ関数

図 5 5人のデータと目的変数のメンバシップ関数

5人のデータとこれらの回帰直線を図3に示す。直線 $\hat{y}$ が平均データによる回帰直線である。なお，表2のデータに関しては，直線 $y_1, \cdots, y_5$は，独立に最小 2 乗法で求めた回帰直線とそれぞれほぼ同一であった。

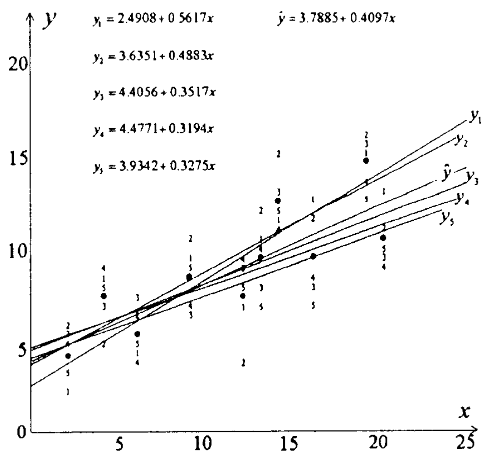


図3 5人のデータと提案手法による回帰直線群

5人の回帰係数を a_0a_1 軸にプロットしたときの分散共分散の5倍を D_A としたとき，図4に示されるように5人の回帰係数は0.5以上のメンバシップ値を持つ。なお，このとき

$$D_A=\begin{pmatrix} 2.5839 & -0.3192 \\ -0.3192 & 0.0476 \end{pmatrix} \quad (49)$$

であった。また，出力のメンバシップ関数を5人のデータとともに示したものが図5である。

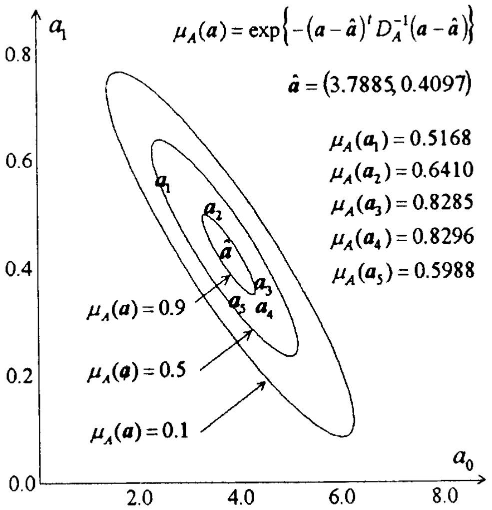


図4　5人の回帰係数と回帰係数のメンバシップ関数

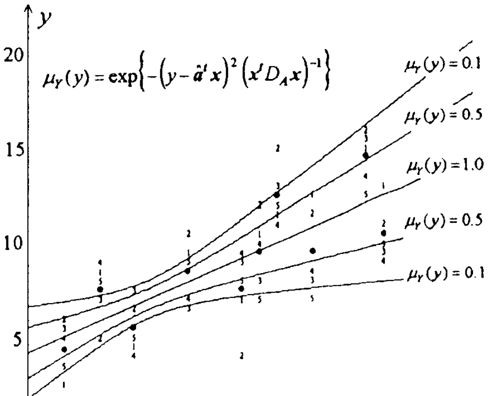


図5　5人のデータと目的変数のメンバシップ関数

5．おわりに

感性評価における個人差や言葉のあいまい性を多変量解析モデルのパラメータ空間で表現する手法を提案した。本稿では回帰分析のみを取り上げたが，他の手法に対しても同様な適用が可能である。なお，田中らの可能性回帰分析手法によれば，線形計画問題を解くことにより，データのばらつきは直接係数のばらつきに変換されるが，本稿の手法によれば，個人データのばらつきを保存するという意味を付け加えることができる。

ところで，数量化理論Ⅰ類の場合は目的変数値の予測よりも，カテゴリー・スコアの推定により重要な意味がある。したがって，回帰係数空間で多次元のメンバシップ関数を求めるだけでなく，各回帰係数のあいまい性を表現する必要がある。このようなケースについては別の機会に議論したい。

参 考 文 献

[1] H. Tanaka, S. Uejima and K. Asai：Linear regression analysis with fuzzy model. *IEEE Trans. on Systems, Man and Cybernetics*, Vol.SMC-12, No.6, pp.903-907, 1982.

[2] 田中英夫・和多田淳三・林勲：ファジィ線形回帰分析の三つの定式化．計測自動制御学会論文集，Vol.22, No.10, pp.1051-1057, 1986.

[3] H. Tanaka, I. Hayashi and J. Watada：Possibilistic linear regression analysis for fuzzy data. *European Journal of Operational Research*, Vol.40, No.3, pp.389-396, 1989.

[4] 田中英夫・石渕久生：2次形式メンバシップ関数による可能性線形システムの同定．計測自動制御学会論文集，Vol.26, No.1, pp.93-100, 1990.

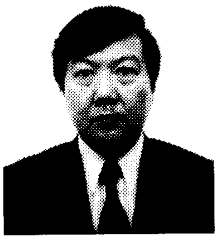
[5] 田中英夫：可能性回帰分析．日本ファジィ学会誌，Vol.5, No.5, pp.1260-1272, 1993.

[6] 田中英夫・石渕久生：ソフトデータ解析．朝倉書店，1995.

(1999年4月26日　受　付)
(1999年8月23日　再受付)

[問い合わせ先]
〒923-1292
石川県能美郡辰口町旭台 1-1
北陸先端科学技術大学院大学
知識科学研究科
中森 義輝
TEL　：0761-51-1755
FAX　：0761-51-1149
E-mail：nakamori@jaist.ac.jp

著者紹介



中森 義輝（なかもり よしてる）

北陸先端科学技術大学院大学知識科学研究科

1979年 京都大学大学院工学研究科数理工学専攻博士課程修了。工学博士。甲南大学理学部応用数学科勤務を経て，1998年4月より北陸先端科学技術大学院大学教授。1984年10月より1985年11月まで，国際応用システム解析研究所研究員。1986年4月より環境庁国立環境研究所客員研究員。1992年9月より大連理工大学客員教授兼務。日本ファジィ学会，計測自動制御学会，環境科学会，IEEEなどの会員。

領家 美奈（りょうけ みな）

北陸先端科学技術大学院大学知識科学研究科

1998年 大阪大学大学院基礎工学研究科物理系専攻制御工学分野博士後期課程修了。博士(工学)。1998年4月より北陸先端科学技術大学院大学助手。1993年度日本ファジィ学会論文賞，1999年度システム制御情報学会学会賞奨励賞受賞。日本ファジィ学会，システム制御情報学会，環境科学会会員。

Fuzzy Regression Models for Subjective Evaluation Data

by

Yoshiteru NAKAMORI and Mina RYOKE

Abstract：

A fuzzy linear regression method is proposed for the data in which plural different output data exist for the same input. Examples of such a situation include some evaluators make their judgment of articles of trade, environment, or students. This paper proposes a technique to map the relations between locations of evaluators in the data set, preserving them in the model parameter space as much as possible. This approach is based on Tanaka's identification method and provides a quite easy way without using the linear programming, and gives a good perspective of the data.

Contact Address：**Yoshiteru NAKAMORI**
Graduate School of Knowledge Science,
Japan Advanced Institute of Science and Technology
1-1 *Asahidai, Tatsunokuchi, Ishikawa* 923-1292, *Japan*
TEL ：0761-51-1755
FAX ：0761-51-1149
E-mail：nakamori@jaist.ac.jp