华南农业大学学报 13 (2) 1992：112～118
J. South China Agr. Univ.

# 模糊数据分析的可能性线性系统

陈振权
(基础部)

**摘要** H. Tanaka 等[7~10]提出的可能性线性系统是模糊数据分析的重要分支，本文把求解模糊参数的均值和展作先后处理，提出4种可能性线性系统模型，这些模型对模糊数据的拟合，均较 Tanaka 模型为优，本文论证新模型解的存在性、解法以及数据分析的性质，并与 Tanaka 的3种模型作比较。

**关键词** 模糊数据；可能性线性系统；模糊线性回归模型

观测模糊现象便获得模糊数据，如何分析和处理模糊数据，已越来越为科学界所注目，已出现多种模糊数据分析法[1,3,4,6]。依 Zadeh[11]的可能性概念，由模糊数定义的模糊数据，可以视为一种可能性分布。再借助扩张原理[11]和模糊数运算法则[5]，由模糊线性方程建立起来的可能性线性系统，便成为模糊数据分析的合适模型。

一般，模糊数据是可能性的、非统计性的。古典概率回归模型把观测值与估计间的误差，归因于测量误差，Tanaka 等[7,10]则认为数据是可能的，误差在于系统本身参数的波动性，是可能性线性系统的反映。从这个观点出发，Tanaka 等提出了3种可能性线性回归模型，这些模型描述数据的清晰输入——模糊输出。其优点之一是把求解模型参数归结为普通线性规划问题。但是，Tanaka 的方法是一种模糊区间分析法，模型只反映模糊数据的区间拟合，忽视对观测数据可能性最大的均值的拟合，致使模型输出数据的均值与观测数据的均值偏离较大，模型失真大。本文目的在于改进模型输出。

## 1 预备知识

所谓可能性线性系统，是一个其参数由可能性分布定义的线性系统，其可能性分布由模糊数表示。模糊数是实数集的如下模糊集 A：(i) $\forall h\in[0,1]$，$[A]_h=\{a\mid\mu_A(a)\geqslant h\}$ 是闭区间；(ii) 存在数 a 使 $\mu_A(a)=1$；(iii) A 为模糊凸的。

由 Dubois 和 Prade 的 L－R 型模糊数定义[5]，这里记对称 L－R 型模糊数为 $A_i=(\alpha_i, c_i)_L$，其隶属函数定义为

$$\mu_{A_i}(a_i)=L[(a_i-\alpha_i)/c_i],\quad c_i>0$$

其中 $\alpha_i$ 为 $A_i$ 的均值，$c_i$ 为 $A_i$ 的展，$L(a_i)$ 为 $A_i$ 的基准函数，它被定义为满足以下条件的函数：(i) $L(a_i)=L(-a_i)$；(ii) $L(0)=1$；(iii) 当 $a_i\geqslant0$ 时，$L(a_i)$ 是严格减少的；(iv) $\{a_i\mid L(a_i)\geqslant0\}$ 是一闭区间，本文仅使用对称的 L－R 型模糊数。

设有线性函数 $y=a_1x_1+\cdots+a_nx_n=\mathrm{ax}$，其中 $x=(x_1,\cdots,x_n)'$，$\mathrm{a}=(a_1,\cdots,a_n)$，如果

· 1991－03－08收稿

把函数中参数 a 代以模糊数向量 $A=(A_1,\cdots,A_n)$ 则得

$$y=A_1x_1+\cdots+A_nx_n=Ax \tag{1}$$

其中 $A_i=(\alpha_i,c_i)_L$，$x_i$ 为实数。称（1）为可能性线性系统。对给定 x，y 由扩张原理定义如下

定义1：（1）中的 y 定义为如下模糊集

$$\mu_y(y)=\sup_{\{a|y=ax\}}[\mu_{A1}(a_1)\wedge\cdots\wedge\mu_{An}(a_n)]$$

其中若 $\{a|y=ax\}=\varnothing$ 则定义 $\mu_y(y)=0$

命题1：（1）模糊输出 y 为对称 L-R 型模糊数 $y=(\alpha x,\ c|x|)_L$，其中 $\alpha=(\alpha_1,\cdots,\alpha_n)$，$c=(c_1,\cdots c_n)$ 和 $|x|=(|x_1|,\cdots,|x_n|)$·

证：由 L-R 型模糊数的线性运算律[5]和 A 为对称 L-R 型模糊数，即知命题成立。

定义2：$A_1,A_2$ 为两模糊数，$h\in[0,1)$，若 h 水平集 $[A_1]_h\supset[A_2]_h$，则称 $A_1\supseteq_h A_2$

显然，$A_1\supseteq_h A_2$ 等价于（2）成立

$$\alpha_1\leqslant\alpha_2+|L^{-1}(h)|(c_1-c_2);\qquad \alpha_1\geqslant\alpha_2-|L^{-1}(h)|(c_1-c_2), \tag{2}$$

此外，若 $A_1\supseteq_h A_2$ 则 $h'\leqslant h$ 时 $A_1\supseteq_{h'} A_2$

对已给模糊数据 $(y_i,x_i)$，$i=1,\cdots,N$，其中第 i 个样本观测值 $y_i=(y_i,e_i)_L$，$x_i=(x_{i1},\cdots,x_{in})'$。Tanaka[8,10]提出的三种可能性线性回归模型归结为：选取（1）的参数拟合已给数据，三种参数分别记 $\bar{A}_i=(\bar{\alpha}_i,\bar{c}_i)_L$，$\underline{A}_i=(\underline{\alpha}_i,\underline{c}_i)_L$ 和 $\hat{A}_i=(\hat{\alpha}_i,\hat{c}_i)_L$，它们分别满足

$$y_i\subseteq_h\ \bar{y}_i=\bar{A}_1x_{i1}+\cdots+\bar{A}_nx_{in} \tag{3}$$

$$y_i\supseteq_h\ \underline{y}_i=\underline{A}_1x_{i1}+\cdots+\underline{A}_nx_{in} \tag{4}$$

$$[y_i]_h\cap[\hat{y}_i=\hat{A}_1x_{i1}+\cdots+\hat{A}_nx_{in}] \tag{5}$$

使输出的展 $C|x_i|$ 分别达最小、最大和最小，即

(i) Min problem

$$\min_{\bar{a},\bar{c}}\sum_{i=1}^{N}\bar{C}|x_i|=\bar{J}(\bar{c})$$

$$\begin{aligned}\text{s.t.}\quad & y_i+|L^{-1}(h)|e_i\leqslant\bar{\alpha}x_i+|L^{-1}(h)|\bar{c}|x_i|\\ & y_i-|L^{-1}(h)|e_i\geqslant\bar{\alpha}x_i-|L^{-1}(h)|\bar{c}|x_i|\\ & \bar{c}\geqslant0,\ i=1,\cdots,N\end{aligned} \tag{6}$$

(ii) Max problem

$$\max_{\underline{a},\underline{c}}\sum_{i=1}^{N}\underline{c}|x_i|=\underline{J}(\underline{c})$$

$$\begin{aligned}\text{s.t.}\quad & y_i+|L^{-1}(h)|e_i\geqslant\hat{\alpha}x_i-|L^{-1}(h)|\underline{c}|x_i|\\ & y_i-|L^{-1}(h)|e_i\leqslant\underline{\hat{\alpha}}x_i-|L^{-1}(h)|\underline{c}|x_i|\\ & \underline{c}\geqslant0,\ i=1,\cdots,N\end{aligned} \tag{7}$$

(iii) conjunction problem

$$\min_{\hat{a},\hat{c}}\sum_{i=1}^{N}\hat{c}|x_i|=J(\hat{c})$$

$$\begin{aligned}\text{s.t.}\quad & y_i+|L^{-1}(h)e_i\geqslant\hat{\alpha}x_i-|L^{-1}(h)|\hat{c}|x_i|\\ & y_i-|L^{-1}(h)|e_i\leqslant\hat{\alpha}x_i+|L^{-1}(h)\hat{c}|x_i|\\ & \hat{c}\geqslant0,\ i=1\cdots,N\end{aligned} \tag{8}$$

## 2 可能性线性回归模型

设已给数据 $(y_i, x_i)$，$i=1, \cdots, N$ 其中 $y_i=(y_i, e_i)_L$ 为模糊数，$x_i=(x_{i1}, \cdots, x_{in})'$ 为 n 维实向量，L 为决策者预先选定的基准函数.为了确定线性系统 (1) 的参数 A 以拟合已给数据，这里分两步解决.首先，求 A 的均值 $\alpha=(\alpha_1, \cdots, \alpha_n)$ 使对数据 $(y_i, x_i)$，$i=1, \cdots, N$ 作最小二乘法拟合，由命题1，即求 α 使下式达最小，为了简单，记 $\sum=\sum_{i=1}^{N}$，

$$\operatorname*{Min}_{\alpha} \sum (y_i-\alpha x_i)^2=m(\alpha) \tag{9}$$

显然，由 (9) 容易解出参数 A 的均值 α。

其次，在均值 α 确定后，A 的展 C 分别取以下4种方法计算，记计算4种参数所得为：$c=(c_1, \cdots, c_n)$；$\bar{c}=(\bar{c}_1, \cdots, \bar{c}_n)$；$\underline{c}$ $(\underline{c}, \cdots, \underline{c}_n)$；$\hat{c}=\hat{c}_1, \cdots, \hat{c}_n)$ 即

（Ⅰ）最小二乘法

$$\operatorname*{Min}_{c} \sum (e_i-c|x_i|)^2=S(c) \tag{10}$$

容易解出 c，且 $c\geqslant 0$，

（Ⅱ）最小值法

$$\operatorname*{Min}_{\bar{c}} \sum \bar{c}|x_i|=\bar{J}(\bar{c})$$

$$\text{s.t.} \quad \begin{aligned} & y_i+|L^{-1}(h)|e_i\leqslant \alpha x_i+|L^{-1}(h)|\bar{c}|x_i| \\ & y_i-|L^{-1}(h)|e_i\geqslant \alpha x_i-|L^{-1}(h)|\bar{c}|x_i| \\ & \bar{c}\geqslant 0, \ i=1, \cdots, N \end{aligned} \tag{11}$$

这可解释为：在 $y_i \subseteq_h \bar{y}_i=\bar{A}x_{i1}+\cdots+\bar{A}_n x_{in}$ 的情况下，使 $\sum \bar{y}_i$ 的展达最小，与 Tanaka 模型不同的是这里的 α 已预先确定了的

（Ⅲ）最大值法

$$\operatorname*{Max}_{\underline{c}} \sum \underline{c}|xi|=\underline{J}(\underline{c})$$

$$\text{s.t.} \quad \begin{aligned} & y_i+|L^{-1}(h)|e_i\geqslant \alpha x_i+|L^{-1}(h)|c|x_i| \\ & y_i-|L^{-1}(h)|e_i\leqslant \alpha x_i-|L^{-1}(h)|c|x_i| \\ & c\geqslant 0, \ i=1, \cdots, N \end{aligned} \tag{12}$$

这可解释为：在 $\underline{y}_i=\underline{A}_1 x_{i1}+\cdots+\underline{A}_n x_{in}\subseteq h \quad y_i$ 的情况下，使$\sum \underline{y}_i$ 的展达最大，其中 α 已确定

（Ⅳ）相交法

$$\operatorname*{Min}_{\hat{c}} \sum \hat{c}|x_i|=\hat{J}(\hat{c})$$

$$\text{s.t.} \quad \begin{aligned} & y_i+|L^{-1}(h)|e_i\geqslant \alpha x_i-|L^{-1}(h)|\hat{c}|x_i| \\ & y_i-|L^{-1}(h)|e_i\leqslant \alpha x_i+|L^{-1}(h)|\hat{c}|x_i| \\ & \hat{c}\geqslant 0, \ i=1, \cdots, N \end{aligned} \tag{13}$$

这可解释为：在 $[\hat{y}_i=\hat{A}_1 x_{i1}+\cdots+\hat{A}_n x_{in}]_h \cap [y_i]_h\neq\varnothing]$ 的条件下，使 $\sum \hat{y}_i$ 的展达最小，α 为已知。

分别称上述方法确定的四种参数，$A_j=(\alpha_j, c_j)_L$，$\bar{A}_j=(\bar{\alpha}_j, \bar{c}_j)_L$，$\underline{A}_j=(\underline{\alpha}_j, \underline{c}_j)_L$，$\hat{A}_j=(\hat{\alpha}_j, \hat{c}_j)_L$ 所决定的可能性线性模型为模型（Ⅰ）、（Ⅱ）、（Ⅲ）、（Ⅳ）.与 Tanaka 的三种模型相比较，上述四种模型的优点是

定理1：模型（Ⅰ）、（Ⅱ）、（Ⅲ）、（Ⅳ）输出数据均值与观测数据均值间的误差平方和最小。

定理2：若数据（$y^\circ_i$，$x^\circ_i$），i=1，…，N 满足（1），即

$$y^\circ_i=A^\circ_1x^\circ_{i1}+\cdots+A^\circ_nx^\circ_{in} \tag{14}$$

则 $A^\circ=A=\bar{A}=A=\hat{A}=\bar{A}=\underline{A}=\hat{A}$，$y^\circ=y=\bar{y}=y=\hat{y}=\bar{y}=\underline{y}=\hat{y}$　(15)

证：因（$y^\circ_i$，$x^\circ_i$）满足（14），故 $y^\circ_i=\alpha^\circ x^\circ_i$，$e^\circ_i=c^\circ|x^\circ_i|$，由（9）、（10）知 $\alpha=\alpha^\circ$，$c=c^\circ$，故 $A=A^\circ$，又由 $e^\circ_i=C^\circ|x^\circ_i|$ 代入（11）得

$$y^\circ_i\leqslant\alpha x^\circ_i+|L^{-1}(h)|(\bar{c}-c^\circ)|x_i|,\quad y^\circ_i\geqslant\alpha x^\circ_i+|L^{-1}(h)|(c^\circ-\bar{c})|x^\circ_i| \tag{16}$$

故（$\alpha^\circ$，$c^\circ$）是（Ⅰ）的一个可行解。若另有 $\bar{c}$ 使

$$\sum\bar{C}|x^\circ_i|<\sum c^\circ|x^\circ_i|$$

则必存在某 i 使 $\bar{C}|x^\circ_i|<c^\circ|x^\circ_i|$，这与（16）矛盾，故 $\bar{C}=c^\circ$。从而 $\bar{A}=A^\circ$，其他等式可类似证明。

定理3：若（$y^\circ_i$,$x^\circ_i$），i=1，…，N 满足（14），则 $\hat{J}(\hat{c})=0$

证：因 $\hat{c}=0$ 时，（13）成立，故 $\hat{J}(\hat{c})=0$

定理4：若清晰数据（$y^\circ_i$,$x^\circ_i$），i=1，…N 满足线性方程 $y^\circ_i=a^\circ x^\circ_i$，则

$$y^\circ_i=y_i=\bar{y}_i=\underline{y}_i=\hat{y}_i,\quad a^\circ=\bar{A}=\underline{A}=\underline{A}=\hat{A} \tag{17}$$

证：c=0是（Ⅰ）的解；c=0是（11）、（12）、（13）的可行解，故（17）成立。

## 3　各系统的数据分析

一般，已给数据不一定满足系统（1），因此，有讨论模型解的存在和相互性质的必要

定理5：$\forall_h\in[0,1)$，（Ⅰ），（Ⅳ）必存在最优解。

证：由于 $y_i$、$x_i$ 的个数有限，只要所有 $\bar{C}_i$ 及 $C_i$ 充分大，约束（11）和（13）必成立，即（11）、（13）必有可行解，故（Ⅰ）、（Ⅳ）存在最优解。

定理6：（Ⅱ）有最优解的必要且充分条件是

$$|L^{-1}(h)|e_i\geqslant|y_i-\alpha x_i|,\ i=1,\cdots,N \tag{18}$$

证：首先，可把（12）改写为

$$|L^{-1}(h)|\underline{c}|x_i|\leqslant\min\left[y_i-\alpha x_i+|L^{-1}(h)|e_i,\ -(y_i-\alpha x_i)+|L^{-1}(h)|e_i\right], \tag{19}$$

设（Ⅱ）有最优解 $\underline{c}$，则只对所有 i，$\underline{c}$ 满足（19），因 $\underline{c}\geqslant0$，故（18）成立。反之，若（18）成立，则只要取 $\underline{c}\geqslant0$ 充分地小，（19）必成立，因此，对所有的 i，（12）有可行解，故（Ⅱ）有最优解

定理7：$\bar{J}(\bar{C})\leqslant\hat{J}(\hat{c})$

证：设 $\bar{c}$ 是（11）的可行解，则由（11）有

$$y_i-|L^{-1}(h)|e_i\leqslant y_i+|L^{-1}(h)|e_i\leqslant\alpha x_i+|L^{-1}(h)|\bar{C}|x_i|$$

$$y_i+|L^{-1}(h)|e_i\geqslant y_i-|L^{-1}(h)|e_i\geqslant\alpha x_i-|L^{-1}(h)|\bar{C}|x_i|$$

即 $\bar{C}$ 也是（13）的可行解，即（13）的可行解集包含（11）的可行解集，故 $\bar{J}(\bar{c})\geqslant\hat{J}(\hat{c})$

定理8：（18）成立的充要条件是：$\hat{J}(\hat{c})=0$

证：$\hat{J}(\hat{c})=0$，则 $\hat{c}=0$ 为（Ⅳ）最优解，则（13）成为

$$y_i+|L^{-1}(h)|e_i\geqslant\alpha x_i,\quad y_i-|L^{-1}(h)|e_i\leqslant\alpha x_i \tag{20}$$

从而（18）成立，反之，若（18）成立，则（20）成立，即 $\hat{c}=0$ 是（13）的可行解，故

$\hat{J}(\hat{c})=0$

**定理9**：若观测数据（$y_i, x_i$），$i=1, \cdots, N$ 是清晰的，则 $c=0$，$\bar{c}=\hat{c}$

证：当数据清晰即 $e_i=0$，由（10）知 $C=0$，在此情形下，（11）与（13）完全相同，故 $\bar{c}=\hat{c}$

**定理10**：若（Ⅱ）对 $h'$ 水平有解，则 $0\leqslant h\leqslant h'$ 时，（Ⅱ）对 $h$ 水平也有解

证：设对 $h'$ 水平（Ⅱ）有解 $C'$，因 $h\leqslant h'$ 有 $|L^{-1}(h')|\leqslant|L^{-1}(h)|$，由（12）得

$y_i+|L^{-1}(h)|e_i\geqslant y_i+|L^{-1}(h')|e_i\geqslant \alpha x_i+|L^{-1}(h')|\underline{C}'|x_i|$

$=\alpha x_i+|L^{-1}(h)|\underline{C}'(|L^{-1}(h')|/|L^{-1}(h)|)|x_i|$

$y_i-|L^{-1}(h)|e_i\leqslant y_i-|L^{-1}(h')|e_i\leqslant \alpha x_i-|L^{-1}(h')|\underline{c}'|x_i|$

$=\alpha x_i-|L^{-1}(h)|\underline{c}'(|L^{-1}(h')|/|L^{-1}(h)|)|x_i|$

故 $\underline{C}'(|L^{-1}(h')|/|L^{-1}(h)|$ 是（12）的可行解，即（Ⅱ）对 $h$ 水平有解。

**定理11**：若 $h'\leqslant h$，则

$\bar{J}(h')\leqslant\bar{J}(h)$，$\underline{J}(h')\geqslant\underline{J}(h)$，$\hat{J}(h')\leqslant\hat{J}(h)$

证：这里只证第一式，其余类推，由于 $\bar{y}_i\supseteq y_i$，故 $\bar{c}_h|x_i|-e_i\geqslant0$，由（11）知

$y_i\leqslant \alpha x_i+|L^{-1}(h)|(\bar{C}_h|x_i|-e_i)\leqslant \alpha x_i+|L^{-1}(h')|(\bar{C}_h|x_i|-e_i)$

$y_i\geqslant \alpha x_i-|L^{-1}(h)|(\bar{C}_h|x_i|-e_i)\geqslant \alpha x_i-|L^{-1}(h')|(\bar{C}_h|x_i|-e_i)$

因此，对 $h'$ 水平，$\bar{C}_h$ 也是（11）可行解，即（11）关于 $h'$ 水平的可行解集，包含（11）关于 $h$ 水平的可行解集，故 $\bar{J}(h')\leqslant\bar{J}(h)$

**定理12**：若（Ⅱ）有解，则 $\forall h$，$\bar{J}(\bar{C})\geqslant\underline{J}(\underline{C})$

证：因（Ⅰ）必有解，当（Ⅱ）有解时，对所有的 $i$，分别成立 $\bar{c}|x_i|\geqslant e_i$，$\underline{C}|x_i|\leqslant e_i$，因此

$\bar{J}(\bar{C})=\mathrm{Min}\sum\bar{C}|x_i|\geqslant\mathrm{Max}\sum\underline{C}|x_i|=\underline{J}(\underline{C})$

## 4　举例

这里举两个不同数据例子说明模型性质，并与 Tanaka 模型作比较，所使用的可能性线性系统为

$y=A_0+A_1x_1$

其中参数 $A_i$ 表 $A_i=(\alpha_i, C_i)_L$，取 $L(x)=1-|x|$，因此 $|L^{-1}(h)|=1-h$，仅考虑 $h=0$ 情形

例1：数据如表一中的（$y_i, x_{i1}$），$i=1, \cdots, 5$，数据不满足（18），故（Ⅱ）无解，为了求出其他模型，首先，求参数均值对数据均值的最优拟合，即

$$\mathop{\mathrm{Min}}_{\alpha_0,\alpha_1}\sum(\alpha_0+\alpha_1x_{i1}-y_i)^2$$

容易解得 $\alpha_0=4.25$，$\alpha_1=1.85$，再按第3节方法，在0－水平上求参数的展，分别得各模型参数为（Ⅰ）：$A_0=(4.25,1.84)_L$，$A_1=(1.85,0.16)_L$；（Ⅰ）：$\bar{A}_0=(4.25,5.15)_L$，$\bar{A}_1=(1.85,0)_L$；（Ⅳ）：$\hat{A}_0(4.25,0.75)_L$，$\hat{A}_1=(1.85,0)_L$。又计算 Tanaka 的3个模型，得参数为 (i)：$\bar{A}_0=(3.15,4.55)_L$，$\bar{A}_1=(2.10,0)_L$；(iii)：$\hat{A}_0=(4.35,0.28)_L$，$\hat{A}_1=(1.57,0)_L$。各模型的输出，如表1所示。模型（Ⅱ）与模型 (ii) 无解。在输出数据的均值与观测数据的均值间的误差平方和上，模型（Ⅰ）、（Ⅰ）、（Ⅳ）均比 Tanaka 的模型 (i)、(iii) 小。

表1　各模型输出对照表

| $x_{i1}$ | 1 | 2 | 3 | 4 | 5 | 均值误差平方和 |
|---|---|---|---|---|---|---|
| $Y_i=(y_i,e_i)_L$ | (8,18) | (5,2.2) | (9.5,2.6) | (13.5,2.6) | (13,2.4) | |
| (Ⅰ) | (6.1,1.84) | (7.95,2.16) | (9.8,2.32) | (11.65,2.48) | (13.5,2.64) | 16.075 |
| (Ⅱ) | (6.1,5.15) | (7.95,5.15) | (9.8,5.15) | 11.65,5.15) | (13.5,5.15) | 16.075 |
| (Ⅲ) | 无解 | | | | | |
| (Ⅳ) | (6.1,0.75) | (7.95,0.75) | (9.8,0.75) | (11.65,0.75) | (13.5,0.75) | 16.075 |
| (i) | (5.95,3.85) | (8.05,3.85) | (10.15,3.85) | (12.25,3.85) | (14.35,3.85) | 17.312 |
| (ii) | 无解 | | | | | |
| (iii) | (5.92,0.28) | (7.49,0.28) | (9.06,0.28) | (1[illegible]63,0.28) | (12.2,0.28) | 19.59 |

例2　数据如表2，满足条件（18），故4种模型均有解.首先，按（9）求模型（Ⅰ）、（Ⅱ）、（Ⅲ）、（Ⅳ）参数的均值得 $\alpha_0=4.95$，$\alpha_1=1.71$，再求参数展，分别得参数为（Ⅰ）：$A_0=(4.95,\ 1.84)_L$，$A_1=(1.71,\ 0.16)_L$；（Ⅱ）：$\bar{A}_0=(4.95,\ 4.03)_L$，$\bar{A}_1=(1.71,\ 0.07)_L$；（Ⅲ）：$A_0=(4.95,\ 0)_L$，$A_1=(1.71,\ 0.115)_L$；（Ⅳ）：$\hat{A}_0=(4.95,\ 0)_L$，$\hat{A}_1=(1.71,\ 0)_L$。Tanaka 的3模型参数为 (i) $\bar{A}_0=(3.85,\ 3.85)_L$，$\bar{A}_1=(2.10,\ 0)_L$；(ii)：$\underline{A}_0=(4.63,\ 0)_L$，$\underline{A}_1=(1.78,\ 0.21)_L$；(iii) $\hat{A}_0=(4.63,\ 0)L$，$\hat{A}_1=(1.57,\ 0)_L$。各模型的输出如表2，均值间误差平方和，前4种模型均较后3种 Tanaka 模型小，拟合较好。

表2　各模型输出对照

| $x_{i1}$ | 1 | 2 | 3 | 4 | 5 | 均值误差平方和 |
|---|---|---|---|---|---|---|
| $Y_i=(y_i,e_i)_L$ | (8,1.8) | (6.4,2.2) | (9.5,2.6) | (13.5,2.6) | (13,2.4) | |
| (Ⅰ) | (6.66,2) | (8.37,2.16) | (10.08,2.32) | (11.79,2.48) | (13.5,2.64) | 9.187 |
| (Ⅱ) | (6.66,4.1) | (8.37,4.17) | (10.08,4.24) | (11.79,4.31) | (13.5,5.38) | 9.187 |
| (Ⅲ) | (6.66,0.115) | (8.37,0.23) | (10.08,0.345) | (11.79,0.46) | (13.5,0.575) | 9.187 |
| (Ⅳ) | (66.6,0) | (8.37,0) | (10.08,0) | (11.79,0) | (13.5,0) | 9.187 |
| (i) | (5.95,3.85) | (8.05,3.85) | (10.15,3.85) | (12.25,3.85) | (14.35,3.85) | 10.7325 |
| (ii) | (6.41,0.21) | (8.19,0.42) | (9.97,0.63) | (11.75,0.84) | (13.53,1.05) | 9.2965 |
| (iii) | (6.2,0) | (7.77,0) | (9.34,0) | (10.91,0) | (12.48,0) | 12.121 |

## 5　结论

本文提出四种用于模糊数据分析的可能性线性回归模型，研究它们解的存在、解法和数据分析的性质.这些模型不仅保留模糊区间分析的原有特点，而且，较 Tanaka 模型使输出的模糊数据与观测数据，在最大可能性的均值间，有较优的拟合，更能反映实际。

与传统建立在概率模式的回归分析不同，可能性线性系统认为，模糊数据中的每个数据的发生都是可能的.这种模型特别适用于由专家意见给出的模糊数据，只有几次观测数据也行。

## 参考文献


1 陈振权.模糊最小二乘法模型.华南农业大学学报，1990，11（1）：50～60

2 Bandemer H. Evaluating explicit functional relationships from fuzzy observations, Fuzzy Sets and Systems, 1985, 16: 41～52

3 Celmins A. Multidimensional least squares fitting of fuzzy models, Math Modelling, 1987, 9: 669～690

4 Diamond P. Fuzzy least squares, Information Sciences, 1988, 46: 145～157

5 Dubois D, Prade H. Fuzzy Sets and Systems. New York, Academic Press, 1980. 36～66

6 Heshmaty B, Kandel A. Fuzzy linear regression and its applicatoins to forecasting in uncertain environment, Fuzzy Sets and Systems, 1985, 15: 159～191

7 Tanaka H, Uejima S, Asai, K. Linear regression analysis with fuzzy model, IEEE Trans Systems Man Cybernet. 1982, 12 (6): 903～907

8 Tanaka H. Fuzzy data analysis by possibilistic linear model, Fuzzy Sets and Systems, 1987, 24 (3): 363～375

9 Tanaka H, Watada J, Hayashi I. On three formulations of fuzzy linear regression analysis, Journal of the Society of Instrument and Control Engineering, 1986, 22 (10): 389～396

10 Tanaka H, Hayashi I, Watada J. Possibilistic linear regression analysis for fuzzy data, European Journal of operational Research, 1989, 40: 389～396

11 Zadeh L A. The concept of a linguistic variable and its application to approximate reasoning I, Information Sciences, 1975, 8: 199～249


## POSSIBILISTIC LINEAR SYSTEMS FOR FUZZY DATA ANALYSIS


Chen Zhenquan

(Department of Basic Courses)



**Abstract** Possibilistic linear regression analysis for fuzzy data originated from a fuzzy phenomenon was introduced by Tanaka et al. [7~10]. In this paper, four types of possibilistic linear regression model are proposed to develop Tanaka' s models which are called the Min problem, the Max problem and the Conjunction problem. The method for obtaining the parameters of the models consists of two steps: (1) determine the mean values of the parameters by using a linear least squares approach so that the merit of the models is to be able to fit fuzzy observed values more realistically; (2) compute the spreads of the parameters by a linear least squares approach or by reducing the problem to a conventional linear programming as treated by Tanaka' s techniques while mean values are determined. The advantage of these models is that they can not only fit the main ideas of the data but also retain the advantages of fuzzy interval analysis. The exixtence of the models' s solution and mutual relations in these models are discussed. To illustrate the approach for dealing with fuzzy data, two numerical examples are shown finally.



**Key words** Fuzzy data; Possibilistic linear systems; Fuzzy linear regression model.