A NEW STATISTICAL APPROACH TO IDENTIFICATION OF FUZZY MODELS IN RANDOM ENVIRONMENT

Da Q. Qian and M. Mizumoto
Department of Management Engineering
Osaka Electro-Communication University
Neyagawa, Osaka 572, Japan

ABSTRACT: The aim of this paper is at the study of identification of fuzzy models in random environment using statistical principle. A new approach to this topic is proposed. Numerical examples are taken up to show the usefulness of the approach by comparing it with other methods.

1. FUZZY RELATIONAL MODEL

Up to now a number of papers have appeared which are devoted to fuzzy modelling. The fuzzy models are usually described by the use of fuzzy relational models, i.e., the fuzzy relational models have the following forms:

$$X_{n+1}=X_1 \circ X_2 \circ \dots X_n \circ R \tag{1}$$

where X_i is an input fuzzy set of variable X_i defined on the universe of discourse $\mathbf{X}_i$ (i=1, 2, ..., n). X_{n+1} stands for the output fuzzy set of variable X_{n+1} defined on the universe of discourse $\mathbf{X}_{n+1}$. R is a fuzzy relation expressing the relationships existing between them and expressed on the cartesian product of X_1, X_2, ..., X_n and X_{n+1}. They are represented by their membership functions

$$X_i: \mathbf{X}_i \to [0, 1] \ (i=1, 2, \dots, n+1), \quad R: X_1 \times X_2 \times \dots \times X_n \times X_{n+1} \to [0, 1]$$

"$\circ$" stands for the sup-min composition.

In this paper we will propose a new approach to identification of fuzzy models in random environment. The usefulness of the approach is demonstrated by solving numerical examples and comparing it with other methods used in the examples.

2. DESCRIPTION OF THE IDENTIFICATION PROCEDURE

In (1), let the domain $[x_{i1}, x_{im+1}]$ of variable X_i (i=1, 2, ..., n+1) be divided into m subdomains, i.e., the jth subdomain of X_i is $[x_{ij}, x_{ij+1})$, $x_{ij+1}=x_{ij}+dx_i$, $dx_i=(x_{im+1}-x_{i1})/m$, each subdomain $[x_{ij}, x_{ij+1})$ is denoted as d(i, j), i=1, 2, ..., n+1, j=1, 2, ..., m. Therefore, the space of n+1 dimensions can be defined by the n+1 variables, and each dimension corresponds to one of the n+1 variables and can be devided into m subdomains, in total, the space can be divided into m^{n+1} subspaces consisting of $\{d(1, j_1), d(2, j_2), \dots, d(n+1, j_{n+1})\}$, j_1=1, 2, ..., m, j_2=1, 2, ..., m, ..., j_{n+1}=1, 2, ..., m, as the 3-dimensional space shown in Fig. 1.

Then each of the l collected data groups $\{x_1(k), x_2(k), \dots, x_{n+1}(k)\}$, k=1, 2, ..., l, is projected into a corresponding subspace consisting of $\{d(1, j_1), d(2, j_2), \dots, d(n+1, j_{n+1})\}$, where $x_i(k) \in d(i, j_i)$. Assume that $\{x_1(k),$

$x_2(k)$, ..., $x_{n+1}(k)\}$ ($k=k_{1(s)}, k_{2(s)}, \ldots, k_{t(s)}$) falls into the subspace s, we can calculate the mean value vector of the subspace, $\{V_1(s), V_2(s), \ldots, V_{n+1}(s)\}$, $V_i(s)=\sum_{k=k_1(s),\ldots,k_t(s)} x_i(k)/t$, $s=1,2,\ldots,m^{n+1}$. Furthermore, assume that there are t_j data groups falling into the subspace s, $(\{d(1,j_1), d(2,j_2), \ldots, d(n,j_n), d(n+1,j)\}, j=1,2,\ldots,m)$, therefore, the frequence f_s in the subspace, s, $\{x_1 \in d(1,j_1),\ x_2 \in d(2,j_2), \ldots,\ x_n \in d(n,j_n) \to x_{n+1} \in d(n+1,j)\}$, is defined: $t_j/\sum_{j=1,\ldots,m} t_j$.

Finally, a set of reference sets r_{ij} is defined, $i=1,2,\ldots,n+1$, $j=1,2,\ldots,q$, where r_{ij} is the jth reference set of the variable X_i, and r_{ij} ($j=1,2,\ldots,q$) can cover the domain $[x_{i1}, x_{im+1}]$ of X_i, the fuzzy relation R is generated via the following formula:

$$R(s)=R(s-1)\cup(f_s\wedge(V_1(s)\times V_2(s)\times\ldots\times V_n(s)\times V_{n+1}(s))),\quad s=2,3,\ldots,m^{n+1} \qquad (2)$$
$$R(1)=f_s\wedge(V_1(1)\times V_2(1)\times\ldots\times V_n(1)\times V_{n+1}(1))$$

or

$$R(s)=R(s-1)\cup(V_1(s)\times V_2(s)\times\ldots\times V_n(s)\times V_{n+1}(s)) \qquad f_s > \text{a threshold} \qquad (3)$$
$$R(1)=(V_1(1)\times V_2(1)\times\ldots\times V_n(1)\times V_{n+1}(1))$$

where $V_i(s)$ is the fuzzy set of variable $V_i(s)$ in subspace s, the $V_i(s)$ is represented by a vector $[p_1(s), p_2(s), \ldots, p_q(s)]$ with respect to the reference sets r_{ij}, each individual $p_j(s)$ is calculated via $r_{ij}(V_i(s))$.

The above identification procedure shows that the reliability of data should first be evaluated by calculating the distribution of the data in space before they are used to build the fuzzy relation R.

3. NUMERICAL EXAMPLES

Example 1: A collection of input-output data is as in Table 1. We will use three approaches including ours to determine R that can match the data.

Table 1

k	input $X_1(k)$	output $X_2(k)$
1	1	1
2	2	2
3	3	1
4	2	4
5	4	1
6	4	3
7	4	4
8	5	5

Analysing the collection of data, we can find that the output is the direct ratio of input if we assume that the pairs of data {3, 1} and {4, 1} are contaminated by noise, therefore, we can guess that there is the rough relationship between input and output described by pairs of data {1, 1}, {2, 3}, {4, 3.5} and {5, 5}. The following is the applications of different approaches to this example and comparison of results of these approaches.

1). At first we determine 4 reference sets r_j of X_1 and X_2 (j=1, 2, 3, 4) as in Fig. 2, then we can generate R according to the recursive formula:

$$R(k)=R(k-1)\cup(X_1(k)\times X_2(k)\times\ldots\times X_n(k)\times X_{n+1}(k)),\quad k=2,3,\ldots,l \tag{4}$$
$$R(1)=X_1(1)\times X_2(1)\times\ldots\times X_n(1)\times X_{n+1}(1)$$

where X_i is the fuzzy set of variable X_i (i=1, 2), the X_i is represented by a vector $[p_1, p_2, p_3, p_4]$ with respect to the reference sets, each individual p_j is calculated via $r_j(X_i)$ (j=1, 2, 3, 4). Using the above formula, the derived R is as follows:

$$\begin{bmatrix} 1.00 & 0.25 & 0.25 & 0.25 \\ 0.50 & 0.75 & 0.75 & 0.25 \\ 0.75 & 0.50 & 0.75 & 0.25 \\ 0.25 & 0.25 & 0.25 & 1.00 \end{bmatrix}$$

By $X_1 \circ R$, we can then infer the output X_2 as follows:

X_2= [1.00 0.25 0.25 0.25] when $X_1=1$
X_2= [0.50 0.75 0.75 0.25] when $X_1=2$
X_2= [0.50 0.50 0.50 0.25] when $X_1=3$
X_2= [0.75 0.50 0.75 0.25] when $X_1=4$
X_2= [0.25 0.25 0.25 1.00] when $X_1=5$

Finally, by using center of gravity method, we can get the following results:

if $X_1=1$ then $X_2=2.14$
if $X_1=2$ then $X_2=2.78$
if $X_1=3$ then $X_2=2.71$
if $X_1=4$ then $X_2=2.63$
if $X_1=5$ then $X_2=3.86$

2). At first using Leszczynski et al's fuzzy clustering method (1985) for above 8 pairs of data in Table 1, we can get the following 6 clusters: (1, 2), (3, 5), (4), (6), (7), (8). According to Hirota et al's method (1983), we can then calculate the mean values of each cluster as in Table 2.

Table 2

k	input X_1	output X_2
1	1.5	1.5
2	3.5	1.0
3	2	4
4	4	3
5	4	4
6	5	5

Using (4) for data in Table 2, R can be derived as follows:

$$\begin{bmatrix} 0.625 & 0.375 & 0.250 & 0.250 \\ 0.375 & 0.375 & 0.750 & 0.250 \\ 0.875 & 0.500 & 0.750 & 0.250 \\ 0.000 & 0.250 & 0.250 & 1.000 \end{bmatrix}$$

Then, by $X_1 \circ R$, we can infer the output X_2 as follows:

X_2= [0.625 0.375 0.25 0.25] when X_1=1
X_2= [0.375 0.375 0.75 0.25] when X_1=2
X_2= [0.50 0.50 0.50 0.25] when X_1=3
X_2= [0.75 0.50 0.75 0.25] when X_1=4
X_2= [0.00 0.25 0.25 1.00] when X_1=5

Finally, by using center of gravity method, we can get the following results:

if X_1=1 then X_2=2.44
if X_1=2 then X_2=3.0
if X_1=3 then X_2=2.71
if X_1=4 then X_2=2.63
if X_1=5 then X_2=4.33

3). Here we use our approach proposed above. At first we divide the domains of X_1 and X_2 into 3 subdomains: [1, 2], (2, 3], (3, 4]. By projecting the data in Table 1 into the subdomains, we can get the results in Table 3.

Table 3

subdomain of X_1	subdomain of X_2	mean of X_1	mean of X_2	frequency
[1, 2]	[1, 2]	1.5	1.5	2/3
[1, 2]	(3, 4]	2	4	1/3
(2, 3]	[1, 2]	3	1	1
(3, 4]	[1, 2]	4	1	1/4
(3, 4]	(2, 3]	4	3	1/4
(3, 4]	(3, 4]	4.5	4.5	1/2

Using (2) where the factor, frequency, is considered, the R is derived as follows:

$$\begin{bmatrix} 0.625 & 0.375 & 0.250 & 0.250 \\ 0.500 & 0.375 & 0.333 & 0.250 \\ 0.250 & 0.250 & 0.375 & 0.375 \\ 0.250 & 0.250 & 0.375 & 0.500 \end{bmatrix}$$

Then, by $X_1 \circ R$, we can infer the output X_2 as follows:

X_2= [0.625 0.375 0.250 0.250] when X_1=1
X_2= [0.500 0.375 0.333 0.250] when X_1=2
X_2= [0.500 0.375 0.375 0.325] when X_1=3
X_2= [0.500 0.250 0.375 0.375] when X_1=4
X_2= [0.250 0.250 0.375 0.500] when X_1=5

Finally, by using center of gravity method, we can get the following results:

if $X_1=1$ then $X_2=2.44$
if $X_1=2$ then $X_2=2.64$
if $X_1=3$ then $X_2=2.85$
if $X_1=4$ then $X_2=2.90$
if $X_1=5$ then $X_2=3.42$

Analysing the results of above three approaches, it can been found that $X_2=2.85$ and $X_2=2.90$ derived by our approach are larger than the values of X_2 derived by the first and second approaches when $X_1=2$ and $X_1=3$. It means that the influence of data pairs {3, 1} and {4, 1}, the possible noise, is decreased.

Example 2: A collection of input-output data generated by $X_2(k)=2*(X_1(k)+r(k))$ where r(k) is non-measureable noise shown in Table 4.

Table 4

k	input $X_1(k)$	output $X_2(k)$
1	0.5	1.0
2	1.0	2.0
3	1.5	3.0
4	2.5	3.4
5	3.0	6.0
6	3.5	7.0
7	4.5	9.0
8	5.0	10.0
9	5.5	12.6
10	6.5	13.0
11	7.0	14.0
12	7.5	15.0
13	8.5	15.8
14	9.0	18.0
15	9.5	19.0

As in the first example, we use the above aproaches to determine R as follows.

1). At first we determine 6 reference sets r_{ij} of X_1 and X_2 ($i=1,2; j=1,\ldots,6$) as in Fig. 3 in which the 6 reference sets of X_1 cover the range of 0-10 and the 6 reference sets of X_2 cover the range of 0-20, then we can generate R as follows according to (4).

$$\begin{bmatrix} 0.75 & 0.50 & 0 & 0 & 0 & 0 \\ 0.50 & 0.75 & 0.50 & 0 & 0 & 0 \\ 0.15 & 0.50 & 0.75 & 0.50 & 0.15 & 0 \\ 0 & 0 & 0.50 & 0.75 & 0.50 & 0 \\ 0 & 0 & 0 & 0.50 & 0.75 & 0.50 \\ 0 & 0 & 0 & 0.05 & 0.50 & 0.75 \end{bmatrix}$$

By $X_1 \bullet R$ and center of gravity method, we can then infer the output X_2 as follows:

Table 5

X_1	0.5	1.0	1.375	1.5	2.5	3.0	3.25	3.5	3.665	4.5	4.75	5.0	5.5
X_2	2.67	4.00	4.00	4.00	5.77	6.70	2.18	7.63	7.93	8.37	8.84	9.30	10.23

X_1	6.5	7.0	7.5	8.5	9.0	9.25	9.5
X_2	13.0	14.0	15.0	16.0	16.0	16.67	17.33

The error, $\sum_{k=1,2,\ldots,20} |2*X_1(k)-X_2(k)|$ is equal to 21.07.

2). At first using Leszczynski et al's fuzzy clustering method (1985) for above 15 pairs of data in Table 4, we can get the following 6 clusters: (1, 2, 3, 4), (5, 6, 7), (8), (9), (10, 11, 12), (13, 14, 15). According to Hirota et al's method (1983), we can then calculate the mean values of each cluster as in Table 6.

Table 6

k	input X_1	output X_2
1	1.375	2.35
2	3.665	7.33
3	5.0	10.0
4	5.5	12.6
5	7.0	14.0
6	9.0	17.6

Using (4) for data in Table 6, R can be derived as follows:

$$\begin{bmatrix} 0.3125 & 0.3125 & 0.0000 & 0.000 & 0.0000 & 0.0000 \\ 0.4125 & 0.5875 & 0.1675 & 0.000 & 0.0000 & 0.0000 \\ 0.0000 & 0.1675 & 0.8325 & 0.500 & 0.1500 & 0.0000 \\ 0.0000 & 0.0000 & 0.5000 & 0.750 & 0.5000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.500 & 0.5000 & 0.4000 \\ 0.0000 & 0.0000 & 0.0000 & 0.000 & 0.5000 & 0.4000 \end{bmatrix}$$

Then, by $X_1 \bullet R$ and center of gravity method, we can infer the output X_2 as follows:

Table 7

X_1	0.5	1.0	1.375	1.5	2.5	3.0	3.25	3.5	3.665	4.5	4.75	5.0	5.5
X_2	3.27	3.09	3.16	3.16	5.91	6.98	3.00	8.11	8.66	10.0	10.6	11.2	11.3

X_1	6.5	7.0	7.5	8.5	9.0	9.25	9.5
X_2	13.0	13.68	14.55	15.71	15.71	16.08	16.52

The error, $\sum_{k=1,2,\ldots,20} |2*X_1(k)-X_2(k)|$ is equal to 24.6.

3). Here we use our approach proposed above. At first we divide the domains of X_1 and X_2 into 5 subdomains: [0,2],(2,4],(4,6],(6,8] and (8,10] for X_1 and [0,4],(4,8],(8,12],(12,16] and (16,20] for X_2. By projecting the data in Table 6 into the subdomains, we can get the results in Table 8.

Table 8

subdomain of X_1	subdomain of X_2	mean of X_1	mean of X_2	frequency
[0,2]	[0,4]	1	2	1
(2,4]	[0,4]	2.5	3.4	1/3
(2,4]	(4,8]	3.25	6.5	2/3
(4,6]	(8,12]	4.75	9.5	2/3
(4,6]	(12,16]	5.5	12.6	1/3
(6,8]	(12,16]	7.0	14.0	1
(8,10]	(12,16]	8.5	15.8	1/3
(8,10]	(16,20]	9.25	18.5	2/3

Using (2) where the factor, frequency, is considered, the R is derived as follows:

$$\begin{pmatrix} 0.500 & 0.500 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.500 & 0.500 & 0.375 & 0.000 & 0.000 & 0.000 \\ 0.150 & 0.375 & 0.625 & 0.375 & 0.150 & 0.000 \\ 0.000 & 0.000 & 0.375 & 0.500 & 0.500 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.500 & 0.500 & 0.375 \\ 0.000 & 0.000 & 0.000 & 0.050 & 0.375 & 0.625 \end{pmatrix}$$

Then, by $X_1 \bullet R$ and center of gravity method, we can infer the output X_2 as follows:

Table 9

X_1	0.5	1.0	1.375	1.5	2.5	3.0	3.25	3.5	3.665	4.5	4.75	5.0	5.5
X_2	3.2	3.64	3.64	3.64	5.86	6.37	7.05	7.55	7.92	8.45	8.95	9.63	10.14

X_1	6.5	7.0	7.5	8.5	9.0	9.25	9.5
X_2	13.54	14	14.46	15.64	16.0	16.73	17.2

The error, $\sum_{k=1,2,\ldots,20} |2*X_1(k)-X_2(k)|$ is equal to 19.83.

Or we can also use the formula (3) where the threshold is 1/3, the R is as follows:

$$\begin{pmatrix} 0.500 & 0.500 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.500 & 0.500 & 0.375 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.375 & 0.625 & 0.375 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.375 & 0.500 & 0.500 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.500 & 0.500 & 0.375 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.375 & 0.625 \end{pmatrix}$$

Then, by $X_1 \circ R$ and center of gravity method, we can infer the output X_2 as follows:

Table 10

X_1	0.5	1.0	1.375	1.5	2.5	3.0	3.25	3.5	3.665	4.5	4.75	5.0	5.5
X_2	3.2	3.64	3.64	3.64	4.92	5.60	6.29	6.77	7.13	9.23	9.71	10.4	11.08
X_1	6.5	7.0	7.5	8.5	9.0	9.25	9.5						
X_2	13.54	14	14.46	15.64	16.0	16.73	17.2						

The error, $\Sigma_{k=1,2,\ldots,20}\ |2*X_1(k)-X_2(k)|$ is equal to 16.68.

By comparison of the computational errors as above, the X_2 derived by our approach has the minimal errors, 19.83 and 16.68.

4. CONCLUDING REMARKS

So far identification methods can be divided into two types. One is that the fuzzy relation R of a fuzzy model is derived by a recursive formula as (4). Another is to determine of the fuzzy relation R by solving fuzzy relational equations (Pedrycz, 1983). Both of them fail to eliminate noise contaminating data used for system identification, so that the fuzzy models identified by them are not identical to their real models. Filter methods can be used to remove some noise from data, the same idea can also be used in (4) (Xu and Lu 1987), i.e., replace (4) with

$$R(k)=a*R(k-1)\mathbf{U}(1-a)(X_1(k)\times X_2(k)\times\ldots\times X_n(k)\times X_{n+1}(k)), \qquad k=1,2,\ldots 1$$

but they are all based on time series analysis compared with our satistical approach. The clustering method (Hirota and Pedrycz (1983)), as used in above examples, can be used in determination of R in (4) when there are unknown disturbances imposing (4) not to hold. The difference between our approach and the clustering method is that our approach tries to eliminate noise but the clustering method tries to determine the R that can best match with the input-output data including noise; moreover, the clustering method can only be used to analyse the limited number of data since the memory space needed by the clustering program is the direct ratio of the number of data mutiplied by the number of clusters, and the result of clustering is dependent on the number of clusters chosen by users (Leszczynski 1985). Czogala and Pedrycz (1983) proposed the concept of fuzzy probabilistic controllers, but they did not deal with how to build the probabilistic controller by making the use of statistics.

This paper aims at the study of an approach to identification of fuzzy systems in noise environment. The application of the proposed approach in above numerical examples and the comparison of our approach with other methods prove its usefulness.

REFERENCES

Czogala, E. and Pedrycz, W., On the Concept of Fuzzy Probabilistic Controllers.

Fuzzy Sets and Systems 10(1983)109-121.

Hirota, K. and Pedrycz, W., Analysis and Synthesis of Fuzzy Systems by the Use of Probabilistic Sets. Fuzzy Sets and Systems 10(1983)1-13.

Leszczynski, K., Penczek, P. and Grochulski, W., Sugeno's Fuzzy Measure and Fuzzy Clustering. Fuzzy Sets and Systems 15(1985)147-158.

Pedrycz, W., Numerical and Applicational Aspects of Fuzzy Relational Equations. Fuzzy Sets and Systems 11(1983)1-18.

Pedrycz, W., An Identification Algorithm in Fuzzy Relational Systems. Fuzzy Sets and Systems 13(1984)153-167.

Xu, C. W. and Lu, Y. Z., Fuzzy Model Identification and Self-Learning for Dynamic Systems. TEEE Trans. Systems, Man and Cybernetics, 17(1987) 683-689.

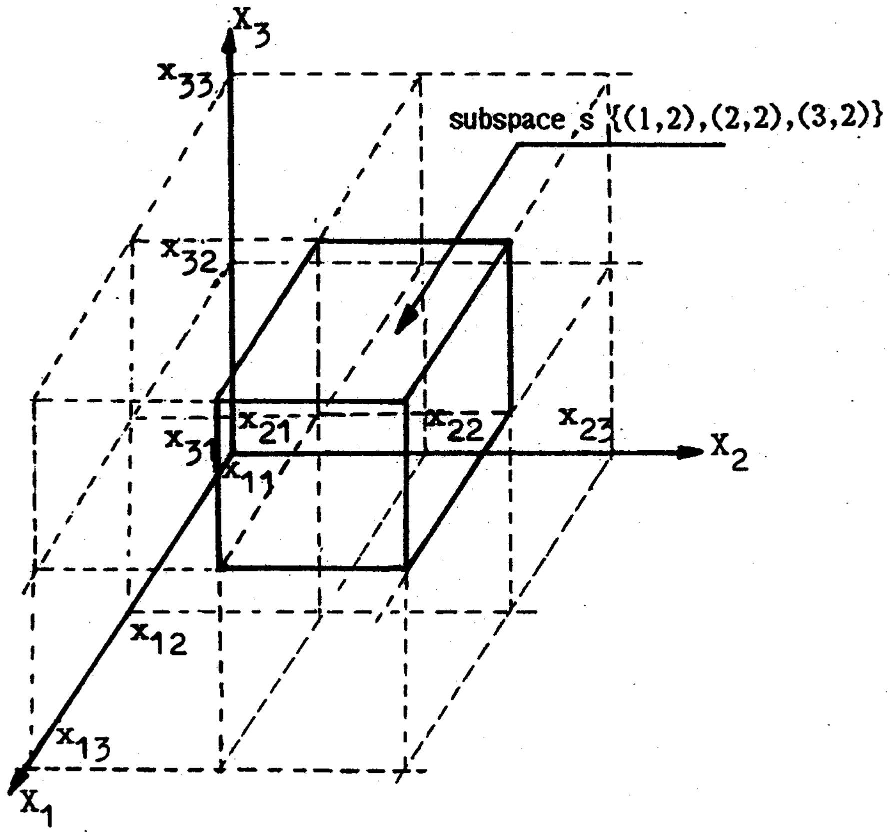

Fig. 1 A 3 dimensional space

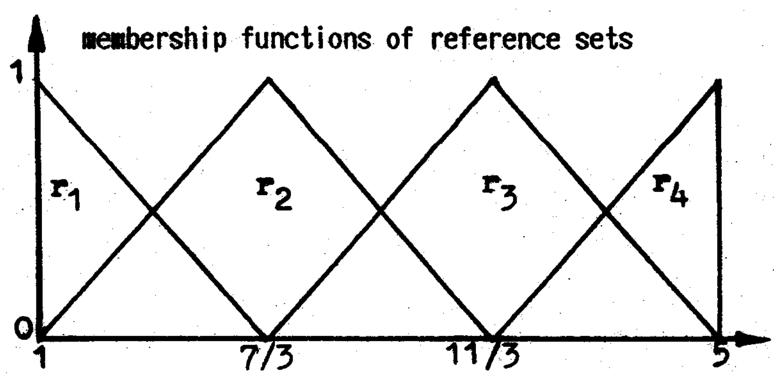

Fig. 2 Reference sets of example 1

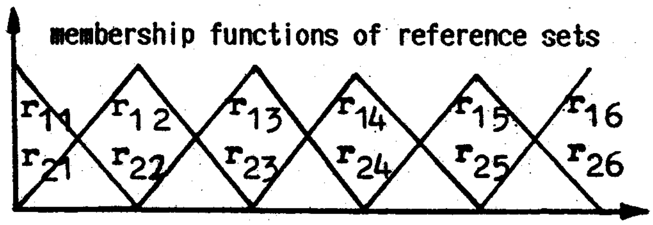

Fig. 3 Reference sets of example 2