Original paper

Soft Computing 8 (2004) 207–216 © Springer-Verlag 2003
DOI 10.1007/s00500-003-0265-3

The use of Kernel set and sample memberships in the identification of nonlinear time series

B. Wu, Y.-Y. Hsu

Abstract The problem of system modeling and identification has attracted considerable attention in the nonlinear time series analysis mostly because of a large number of applications in diverse fields like financial management, biomedical system, transportation, ecology, electric power systems, hydrology, and aeronautics. Many papers have been presented on the study of time series clustering and identification. Nonetheless, we would like to point out that in dealing with clustering time series, we should also take the vague case as they belong to two or more classes simultaneously into account. Because many patterns of grouping in time series really are ambiguous, those phenomena should not be assigned to certain specific classes inflexibly. In this paper, we propose a procedure that can effectively cluster nonlinear time series into several patterns based on kernel set. This algorithm also combines with the concept of a fuzzy set. The membership degree of each datum corresponding to the cluster centers is calculated and is used for performance index grouping. We also suggest a principle for extending the fuzzy set by kernel set and further interpret events in a sensible light through these sets. Finally, the procedure is demonstrated by set off RRI data and its performance is shown to compare favorably with other procedures published in the literature.

Keywords Fuzzy sets, Kernel sets, Clustering, Identification, Nonlinear time series

1 Introduction

Clustering is aimed at organizing and revealing structures within data. Clustering is commonly viewed as an instance of unsupervised learning to cluster a data set into groups of similar individuals. Moreover, the conventional clustering methods restrict that each point of the data set belongs to exactly class and omit the possibility that they belong to two or more classes simultaneously.

Fuzzy sets originated by Zadeh (1965) gave an idea of uncertainty that is described with a membership function. Since hardly ever any disturbance or noise in the data set can be completely eliminated and some inherent data uncertainly cannot be avoided, the use of fuzzy sets therefore provide a solvent for indistinct parts in the data. Therefore, it is quite natural and useful to apply the idea of fuzzy set theory in cluster analysis.

To exhibit the empirical data in an appropriate class in the application of nonlinear system, researchers in this field have been concerned with clustering techniques combining with fuzzy logic. For instance, Cutsem and Gath (1993) proposed a procedure of fuzzy clustering to detect outliers and robustly estimate parameters. Yoshinari, Pedrycz and Hirota (1993) presented fuzzy clustering techniques to construct fuzzy models. Romer et al. (1995) used fuzzy partitions and possibility theory in statistical inference. Cheng et al. (1998) presented a multistage random sampling fuzzy c-means-based clustering algorithm to partition a data set into c classes. Wu and Hsu (1999) utilized the average of cumulative fuzzy entropy to classify and identity Taiwan's unemployment structural changes. Barra and Boire (2000) proposed a possibilistic clustering algorithm using fuzzy theory to test on phantom image of normal and Alzheimer's brains. Besides, Werners (1987), Tseng and Klein (1992), and Yang (1993) conducted more comprehensive research on fuzzy clustering in fuzzy decision analysis.

In the field of time series analysis for system identification, time series are often encountered is important cases where the statistics of the data exhibit nonstationary structure. There is a need for techniques to identify the system of some patterns that could influence decision-making. In the absence of a powerful enough algorithm for this in the past, the usual ploy has been to assume the data under consideration we piecewise stationary and apply identification algorithms which were signed for stationary data but whose estimates converge quickly enough that the assumption of piecewise stationary would be badly abused.

Hence, in order to meet real situation, it has been better to employ the concept of "kernel set" instead of "crisp set" in clustering and identification. In this paper, we propose a procedure that can effectively cluster nonlinear time series into several patterns based on kernel set. This paper is organized as follows. In Sect. 2, we introduce definitions

Published online: 14 July 2003

B. Wu (✉)
Department of Mathematical Sciences,
National Chengchi University, Wenshan,
Taipei 11623, Taiwan, ROC
e-mail: berlin@math.nccu.edu.tw

Y.-Y. Hsu
Department of Applied Mathematics,
National Donghwa University, Hualien 974, Taiwan, ROC

We are grateful to the referees for their careful reading and helpful comments.

and related theorems for clustering and analyzing. In addition, we develop a measuring method for calculating membership degree of each datum corresponding to the cluster centers and further get kernel set from fuzzy set after clustering in Sect. 3. The empirical application and comparative issues for RRI data are covered in Sect. 4. These include computational problems in clustering nonlinear time series, and the issue of whether the procedure combining fuzzy set theory has a better performance than other procedures. Sect. 5 gives the conclusion and suggestions.

2 The nature of kernel sets

2.1 Fuzzy sets and kernel sets

Though there are huge papers discussing about the fuzzy sets and its applications. Few literatures can be found about the crisp definition of fuzzy set. In this paper, we will present an appropriate definition for it. The following definitions are made in order to formalize and simplify nonlinear time series analysis.

Definition 2.1. [Fuzzy set] *Let* $X = \{X_t, t = 1, 2, \ldots, N\}$ *be a universal set and* $L = \{L_k, k = 1, 2, \ldots, r\}$ *be a sequence of linguistic terms with monotone degree of semantic. For each* $k = 1, 2, \ldots, r$, *the fuzzy set* A_k *generated by* L_k *on* X *is a generalized set, which assigns each* X_t *to a value between 0 and 1.*

Definition 2.2. [Kernel set] *Let* $X = \{X_t, t = 1, 2, \ldots, N\}$ *be a universal set and* $L = \{L_k, k = 1, 2, \ldots, r\}$ *be a sequence of linguistic terms with monotone degree of semantic. For any significant level* $\alpha_k \in [0, 1]$, $k = 1, 2, \ldots, r$, *the kernel set of* A_k *generated by* L_k *is denoted by*

$$\mathrm{Ker}(A_k) = \bigvee_{\alpha \in \Lambda(A_k)|_{\alpha \geq \alpha_k}} \alpha \cdot (A_k)_\alpha, \quad k = 1, 2, \ldots, r \ . \qquad (1)$$

where $\Lambda(A_k)$ *is the level set of* A_k, $(A_k)_\alpha$ *is a* α*-cut of* A_k, *and* $\vee$ *denotes the finite fuzzy union.*

Fuzzy set – an extension of kernel set

Let $X = \{X_t, t = 1, 2, \ldots, N\}$ *be a universal set and* $L = \{L_k, k = 1, 2, \ldots, r\}$ *be a sequence of linguistic terms with monotone degree of semantic. The fuzzy set* A_k *generated by* L_k *defined on* X *is an extension of its kernel set and denoted by*

$$A_k = \bigvee_{\alpha \in \Lambda(A_k)} \alpha \cdot (A_k)_\alpha, \quad k = 1, 2, \ldots, r \ . \qquad (2)$$

where $\Lambda(A_k)$ *is the level set of* A_k, $(A_k)_\alpha$ *is a* α*-cut of* A_k, and $\vee$ *denotes the finite fuzzy union.*

Example 2.1. Let $X = \{X_1, X_2, X_3, X_4\} = \{Lavender, Eucalyptus, Thyme, Ylang\}$ be a universal set of essential oil for helping sleep and $L = \{L_1, L_2, L_3, L_4, L_5\} = \{Highly\ \ ineffective, Ineffective, Common, Effective, Highlyeffective\}$ be a sequence of linguistic terms with monotone degree of semantic. After interviewing 10 consumers, we get the membership grades of the foregoing essential oils. The results are showed at Table 1.

By Definition 2.1, the fuzzy set A_1 generated by L_1 on X is

$$A_1 = 0.10/X_1 + 0.30/X_2 + 0.20/X_3 \ .$$

and the level set of A_1 is $\Lambda(A_1) = \{0.10, 0.30, 0.20\}$. Moreover, via its membership grades, the α-cuts of A_1 become

$$(A_1)_{0.10} = 1/X_1 + 1/X_2 + 1/X_3 + 0/X_4 \ ,$$

$$(A_1)_{0.30} = 0/X_1 + 1/X_2 + 0/X_3 + 0/X_4 \ ,$$

$$(A_1)_{0.20} = 0/X_1 + 1/X_2 + 1/X_3 + 0/X_4 \ .$$

If we multiply each $(A_1)_\alpha$ by corresponding α value, then we convert each of the α-cuts to a special fuzzy set, i.e.

$$0.10 \cdot (A_1)_{0.10} = 0.10/X_1 + 0.10/X_2 + 0.10/X_3 + 0.00/X_4 \ ,$$

$$0.30 \cdot (A_1)_{0.30} = 0.00/X_1 + 0.30/X_2 + 0.00/X_3 + 0.00/X_4 \ ,$$

$$0.20 \cdot (A_1)_{0.20} = 0.00/X_1 + 0.20/X_2 + 0.20/X_3 + 0.00/X_4 \ .$$

Under the significant level $\alpha_1 = 0.80$ and by Definition 2.2, we can get the kernel set of A_1 is

$$\mathrm{Ker}(A_1) = \bigvee_{\alpha \in \Lambda(A_1)|_{\alpha \geq 0.80}} \alpha \cdot (A_1)_\alpha = \phi \ .$$

and take the finite fuzzy union of above special fuzzy sets, $\alpha \cdot (A_1)_\alpha$, we get

$$\begin{aligned} A_1 &= \bigvee_{\alpha \in \Lambda(A_1)} \alpha \cdot (A_1)_\alpha \\ &= [0.10 \cdot (A_1)_{0.10}] \vee [0.30 \cdot (A_1)_{0.30}] \vee [0.20 \cdot (A_1)_{0.20}] \\ &= (0.10 \vee 0.00 \vee 0.00)/X_1 + (0.10 \vee 0.30 \vee 0.20)/X_2 \\ &\quad + (0.10 \vee 0.00 \vee 0.20)/X_3 + (0.00 \vee 0.00 \vee 0.00)/X_4 \\ &= 0.10/X_1 + 0.30/X_2 + 0.20/X_3 \ . \end{aligned}$$

Similarly, we can get the following Table 2.

Table 1. The membership grades of $\{X_t\}$ for $\{L_k\}$

	L_1	L_2	L_3	L_4	L_5
X_1	0.10	0.35	0.50	0.75	0.65
X_2	0.30	0.50	0.60	0.20	0.10
X_3	0.20	0.45	0.80	0.50	0.30
X_4	0.00	0.10	0.40	0.80	0.92

Table 2. The kernel sets and fuzzy sets for $\{L_k\}$ under the significant levels $\alpha_k = 0.80$, $k = 1, \ldots, 5$

	Kernel set	Fuzzy set
L_1	ϕ	$0.10/X_1 + 0.30/X_2 + 0.20/X_3$
L_2	ϕ	$0.35/X_1 + 0.50/X_2 + 0.45/X_3 + 0.10/X_4$
L_3	$0.80/X_3$	$0.50/X_1 + 0.60/X_2 + 0.80/X_3 + 0.40/X_4$
L_4	$0.80/X_4$	$0.75/X_1 + 0.20/X_2 + 0.50/X_3 + 0.80/X_4$
L_5	$0.92/X_4$	$0.65/X_1 + 0.10/X_2 + 0.30/X_3 + 0.92/X_4$

Theorem 2.1. Let X be a universal set and A be a fuzzy set on X. Then the complement of fuzzy set, A^c, is denoted by

$$A^c = \bigvee_{(1-\alpha)\in\Lambda(A^c)} (1-\alpha)\cdot(A^c)_{(1-\alpha)} \quad . \tag{3}$$

where $\Lambda(A^c)$ is the level set of A^c, $(A^c)_{(1-\alpha)}$ is a $(1-\alpha)$-cut of A^c, and $\vee$ denotes the finite fuzzy union.

Proof. For each $x \in X$, let $a = \mu_{A^c}(x)$. Then,

$$\begin{aligned}
&\mu_{\left(\bigvee_{(1-\alpha)\in\Lambda(A^c)}(1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x)\\
&= \max_{(1-\alpha)\in\Lambda(A^c)} \mu_{\left((1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x)\\
&= \max\left(\max_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)\le a}} \mu_{\left((1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x),\right.\\
&\qquad \left.\max_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)> a}} \mu_{\left((1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x)\right) \quad .
\end{aligned} \tag{4}$$

For each $(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)>a}$, we have

$$\mu_{A^c}(x) = a < 1-\alpha \ ,$$

it implies

$$\mu_{\left((1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x) = 0 \quad . \tag{5}$$

On the other hand, for each $(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)\le a}$, we have

$$\mu_{A^c}(x) = a \ge 1-\alpha \ ,$$

it implies

$$\mu_{\left((1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x) = 1-\alpha \quad . \tag{6}$$

Eq. (4) follows,

$$\begin{aligned}
\mu_{\left(\bigvee_{(1-\alpha)\in\Lambda(A^c)}(1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x) &= \max_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)\le a}} (1-\alpha)\\
&= a = \mu_{A^c}(x) \quad .
\end{aligned}$$

The proof is complete.

Corollary 2.1. Let X be a universal set and A be a fuzzy set on X. For any significant level $(1-\alpha_1)\in[0,1]$, the complement of fuzzy set, A^c, is denoted by

$$A^c = \mathrm{Ker}(A^c) \cup_F \left(\bigvee_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)<(1-\alpha_1)}} (1-\alpha)\cdot(A^c)_{(1-\alpha)}\right) \quad . \tag{7}$$

where $\Lambda(A^c)$ is the level set of A^c, $(A^c)_{(1-\alpha)}$ is a $(1-\alpha)$-cut of A^c and $\cup_F$ denotes the standard fuzzy union.

Proof. For each $x\in X$, let $1-\alpha_1 = \mu_{A^c}(x)$ be the significant level. Then,

$$\begin{aligned}
\mu_{A^c}(x) &= \mu_{\left(\bigvee_{(1-\alpha)\in\Lambda(A^c)}(1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x) \ \text{(by Theorem 2.1.)}\\
&= \max_{(1-\alpha)\in\Lambda(A^c)} \mu_{\left((1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x)\\
&= \max\left(\max_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)<(1-\alpha_1)}} \mu_{\left((1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x),\right.\\
&\qquad \left.\max_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)\ge(1-\alpha_1)}} \mu_{\left((1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x)\right)\\
&= \max\left(\mu_{\left(\bigvee_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)<(1-\alpha_1)}}(1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x),\right.\\
&\qquad \left.\mu_{\left(\bigvee_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)\ge(1-\alpha_1)}}(1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x)\right) \quad .
\end{aligned} \tag{8}$$

Since, under the significant level $1-\alpha_1$,

$$\bigvee_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)\ge(1-\alpha_1)}} (1-\alpha)\cdot(A^c)_{(1-\alpha)} = \mathrm{Ker}(A^c) \quad . \tag{9}$$

Eq. (8) follows,

$$\begin{aligned}
\mu_{A^c}(x) &= \max\left(\mu_{\left(\bigvee_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)<(1-\alpha_1)}}(1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x),\right.\\
&\qquad \left.\mu_{\mathrm{Ker}(A^c)}(x)\right)\\
&= \max\left(\mu_{\mathrm{Ker}(A^c)}(x),\right.\\
&\qquad \left.\mu_{\left(\bigvee_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)<(1-\alpha_1)}}(1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)}(x)\right) \quad .\\
&\qquad \text{(by the commutative law)}
\end{aligned} \tag{10}$$

By the standard fuzzy union, Eq. (10) follows,

$$\mu_{A^c}(x) = \mu_{\left(\mathrm{Ker}(A^c)\cup_F\left(\bigvee_{(1-\alpha)\in\Lambda(A^c)|_{(1-\alpha)<(1-\alpha_1)}}(1-\alpha)\cdot(A^c)_{(1-\alpha)}\right)\right)}(x) \ .$$

The proof is complete.

Example 2.2. From Example 2.1, the complement of fuzzy set A_1 generated by L_1 on X can be written as:

$$\begin{aligned}
A_1^c &= 0.90/X_1 + 0.70/X_2 + 0.80/X_3 + 1.00/X_4\\
&= (1-0.10)/X_1 + (1-0.30)/X_2\\
&\quad + (1-0.20)/X_3 + (1-0.00)/X_4 \quad .
\end{aligned}$$

and the level set of A_1^c is $\Lambda(A_1^c) = \{0.90, 0.70, 0.80, 1.00\}$. Thus, the $(1-\alpha)$-cuts of A_1^c according to its membership grades become

$$\left(A_1^c\right)_{(1-0.10)} = 1/X_1 + 0/X_2 + 0/X_3 + 1/X_4 \ ,$$
$$\left(A_1^c\right)_{(1-0.30)} = 1/X_1 + 1/X_2 + 1/X_3 + 1/X_4 \ ,$$
$$\left(A_1^c\right)_{(1-0.20)} = 1/X_1 + 0/X_2 + 1/X_3 + 1/X_4 \ ,$$
$$\left(A_1^c\right)_{(1-0.00)} = 0/X_1 + 0/X_2 + 0/X_3 + 1/X_4 \ .$$

If we multiply each $(A_1^c)_{(1-\alpha)}$ by corresponding $(1-\alpha)$ value, then we convert each of the $(1-\alpha)$-cuts to a special fuzzy set, i.e.

$$\begin{aligned}
&(1-0.10)\cdot\left(A_1^c\right)_{(1-0.10)} \\
&\quad = 0.90/X_1 + 0.00/X_2 + 0.00/X_3 + 0.90/X_4 \ , \\
&(1-0.30)\cdot\left(A_1^c\right)_{(1-0.30)} \\
&\quad = 0.70/X_1 + 0.70/X_2 + 0.70/X_3 + 0.70/X_4 \ , \\
&(1-0.20)\cdot\left(A_1^c\right)_{(1-0.20)} \\
&\quad = 0.80/X_1 + 0.00/X_2 + 0.80/X_3 + 0.80/X_4 \ , \\
&(1-0.00)\cdot\left(A_1^c\right)_{(1-0.00)} \\
&\quad = 0.00/X_1 + 0.00/X_2 + 0.00/X_3 + 1.00/X_4 \ .
\end{aligned}$$

Since, under the significant level $1-\alpha_1 = 0.80$, we can get

$$\begin{aligned}
&\mathrm{Ker}\left(A_1^c\right) \\
&\quad = \bigvee_{(1-\alpha)\in\Lambda(A_1^c)|_{(1-\alpha)\geq 0.80}} (1-\alpha)\cdot\left(A_1^c\right)_{(1-\alpha)} \\
&\quad = \left[(1-0.10)\cdot\left(A_1^c\right)_{(1-0.10)}\right] \vee \left[(1-0.20)\cdot\left(A_1^c\right)_{(1-0.20)}\right] \\
&\qquad \vee \left[(1-0.00)\cdot\left(A_1^c\right)_{(1-0.00)}\right] \\
&\quad = (0.90 \vee 0.80 \vee 0.00)/X_1 + (0.00 \vee 0.00 \vee 0.00)/X_2 \\
&\qquad + (0.00 \vee 0.80 \vee 0.00)/X_3 + (0.90 \vee 0.80 \vee 1.00)/X_4 \\
&\quad = 0.90/X_1 + 0.00/X_2 + 0.80/X_3 + 1.00/X_4 \ .
\end{aligned}$$

and

$$\begin{aligned}
&\bigvee_{(1-\alpha)\in\Lambda\left(A_1^c\right)|_{(1-\alpha)<0.80}} (1-\alpha)\cdot\left(A_1^c\right)_{(1-\alpha)} \\
&\quad = (1-0.30)\cdot\left(A_1^c\right)_{(1-0.30)} \\
&\quad = 0.70/X_1 + 0.70/X_2 + 0.70/X_3 + 0.70/X_4 \ .
\end{aligned}$$

Finally, by Corollary 2.1, we get

$$\begin{aligned}
A_1^c &= \mathrm{Ker}(A_1^c) \cup_F \left(\bigvee_{(1-\alpha)\in\Lambda(A_1^c)|_{(1-\alpha)<0.80}} (1-\alpha)\cdot\left(A_1^c\right)_{(1-\alpha)} \right) \\
&= 0.90/X_1 + 0.70/X_2 + 0.80/X_3 + 1.00/X_4 \ .
\end{aligned}$$

Similarly, we can get the following results.

$$\begin{aligned}
A_2^c &= 0.65/X_1 + 0.50/X_2 + 0.55/X_3 + 0.90/X_4 \ , \\
A_3^c &= 0.50/X_1 + 0.40/X_2 + 0.20/X_3 + 0.60/X_4 \ , \\
A_4^c &= 0.25/X_1 + 0.80/X_2 + 0.50/X_3 + 0.20/X_4 \ , \\
A_5^c &= 0.35/X_1 + 0.90/X_2 + 0.70/X_3 + 0.08/X_4 \ .
\end{aligned}$$

Theorem 2.2. Let X be a universal set, A and B be two fuzzy sets on X. For any significant levels $\alpha_1, \alpha_2 \in [0,1]$, the standard fuzzy union and intersection of A and B are denoted by

(i)

$$\begin{aligned}
A \cup_F B &= \left[\bigvee_{\alpha\in\Lambda(A)} \alpha\cdot(A)_\alpha\right] \cup_F \left[\bigvee_{\beta\in\Lambda(B)} \beta\cdot(B)_\beta\right] \\
&= [\mathrm{Ker}(A) \cup_F \mathrm{Ker}(B)] \\
&\quad \cup_F \left[\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}} \alpha\cdot(A)_\alpha\right)\right. \\
&\quad \left. \cup_F \left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}} \beta\cdot(B)_\beta\right)\right] \ . \qquad (11)
\end{aligned}$$

(ii)

$$\begin{aligned}
A \cap_F B &= \left[\bigvee_{\alpha\in\Lambda(A)} \alpha\cdot(A)_\alpha\right] \cap_F \left[\bigvee_{\beta\in\Lambda(B)} \beta\cdot(B)_\beta\right] \\
&= [\mathrm{Ker}(A) \cap_F \mathrm{Ker}(B)] \\
&\quad \cup_F \left[\mathrm{Ker}(A) \cap_F \left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}} \beta\cdot(B)_\beta\right)\right] \\
&\quad \cup_F \left[\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}} \alpha\cdot(A)_\alpha\right) \cap_F \mathrm{Ker}(B)\right] \\
&\quad \cup_F \left[\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}} \alpha\cdot(A)_\alpha\right)\right. \\
&\quad \left. \cap_F \left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}} \beta\cdot(B)_\beta\right)\right] \ . \qquad (12)
\end{aligned}$$

where $\Lambda(A)$ is the level set of A, $\Lambda(B)$ is the level set of B, $(A)_\alpha$ is a α-cut of A, $(B)_\beta$ is a β-cut of B, $\cup_F$ and $\cap_F$ denote the standard fuzzy union and intersection respectively.

Proof. (i) For each $x \in X$, let $\mu_A(x) = \alpha_1$ and $\mu_B(x) = \alpha_2$ be the significant levels. Then, by the standard fuzzy union, it implies

$$\begin{aligned}
&\mu_{A\cup_F B}(x) \\
&\quad = \max(\mu_A(x), \mu_B(x)) \\
&\quad = \max\left(\mu_{\left(\bigvee_{\alpha\in\Lambda(A)} \alpha\cdot(A)_\alpha\right)}(x), \mu_{\left(\bigvee_{\beta\in\Lambda(B)} \beta\cdot(B)_\beta\right)}(x)\right) \\
&\quad = \max\left(\max_{\alpha\in\Lambda(A)} \mu_{(\alpha\cdot(A)_\alpha)}(x), \max_{\beta\in\Lambda(B)} \mu_{(\beta\cdot(B)_\beta)}(x)\right) \\
&\quad = \max\left[\max\left(\max_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}} \mu_{(\alpha\cdot(A)_\alpha)}(x), \max_{\alpha\in\Lambda(A)|_{\alpha\geq\alpha_1}} \mu_{(\alpha\cdot(A)_\alpha)}(x)\right),\right. \\
&\qquad \left. \max\left(\max_{\beta\in\Lambda(B)|_{\beta<\alpha_2}} \mu_{(\beta\cdot(B)_\beta)}(x), \max_{\beta\in\Lambda(B)|_{\beta\geq\alpha_2}} \mu_{(\beta\cdot(B)_\beta)}(x)\right)\right]
\end{aligned}$$

$$=\max\left[\max\left(\mu_{\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)}(x),\mu_{\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha\geq\alpha_1}}\alpha\cdot(A)_\alpha\right)}(x)\right),\right.$$
$$\left.\max\left(\mu_{\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)}(x),\mu_{\left(\bigvee_{\beta\in\Lambda(B)|_{\beta\geq\alpha_2}}\beta\cdot(B)_\beta\right)}(x)\right)\right]. \tag{13}$$

Since, under the significant level α_1,

$$\bigvee_{\alpha\in\Lambda(A)|_{\alpha\geq\alpha_1}}\alpha\cdot(A)_\alpha=\mathrm{Ker}(A)\ . \tag{14}$$

Similarly, under the significant level α_2,

$$\bigvee_{\beta\in\Lambda(B)|_{\beta\geq\alpha_2}}\beta\cdot(B)_\beta=\mathrm{Ker}(B)\ . \tag{15}$$

Eq. (13) follows,

$$\mu_{A\cup_F B}(x)$$
$$=\max\left[\max\left(\mu_{\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)}(x),\mu_{\mathrm{Ker}(A)}(x)\right),\right.$$
$$\left.\max\left(\mu_{\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)}(x),\mu_{\mathrm{Ker}(B)}(x)\right)\right]$$
$$=\max\left[\max\left(\mu_{\mathrm{Ker}(A)}(x),\mu_{\mathrm{Ker}(B)}(x)\right),\right.$$
$$\left.\max\left(\mu_{\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)}(x),\mu_{\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)}(x)\right)\right].$$
$$\text{(by the commutative and associative laws)} \tag{16}$$

By the standard fuzzy union, Eq. (16) follows,

$$\mu_{A\cup_F B}(x)$$
$$=\max\left(\mu_{(\mathrm{Ker}(A)\cup_F\mathrm{Ker}(B))}(x),\right.$$
$$\left.\mu_{\left(\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)\cup_F\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)\right)}(x)\right)$$
$$=\mu_{\left([\mathrm{Ker}(A)\cup_F\mathrm{Ker}(B)]\cup_F\left[\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)\cup_F\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)\right]\right)}(x)\ .$$

The proof is complete.

(ii) For each $x\in X$, let $\mu_A(x)=\alpha_1$ and $\mu_B(x)=\alpha_2$ be the significant levels. Then, by the standard fuzzy intersection, it implies

$$\mu_{A\cap_F B}(x)$$
$$=\min(\mu_A(x),\mu_B(x))$$
$$=\min\left(\mu_{\left(\bigvee_{\alpha\in\Lambda(A)}\alpha\cdot(A)_\alpha\right)}(x),\mu_{\left(\bigvee_{\beta\in\Lambda(B)}\beta\cdot(B)_\beta\right)}(x)\right)$$
$$=\min\left(\max_{\alpha\in\Lambda(A)}\mu_{(\alpha\cdot(A)_\alpha)}(x),\ \max_{\beta\in\Lambda(B)}\mu_{(\beta\cdot(B)_\beta)}(x)\right)$$
$$=\min\left[\max\left(\max_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\mu_{(\alpha\cdot(A)_\alpha)}(x),\ \max_{\alpha\in\Lambda(A)|_{\alpha\geq\alpha_1}}\mu_{(\alpha\cdot(A)_\alpha)}(x)\right),\right.$$
$$\left.\max\left(\max_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\mu_{(\beta\cdot(B)_\beta)}(x),\ \max_{\beta\in\Lambda(B)|_{\beta\geq\alpha_2}}\mu_{(\beta\cdot(B)_\beta)}(x)\right)\right]$$
$$=\min\left[\max\left(\mu_{\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)}(x),\mu_{\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha\geq\alpha_1}}\alpha\cdot(A)_\alpha\right)}(x)\right),\right.$$
$$\left.\max\left(\mu_{\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)}(x),\mu_{\left(\bigvee_{\beta\in\Lambda(B)|_{\beta\geq\alpha_2}}\beta\cdot(B)_\beta\right)}(x)\right)\right]. \tag{17}$$

Since, under the significant level α_1,

$$\bigvee_{\alpha\in\Lambda(A)|_{\alpha\geq\alpha_1}}\alpha\cdot(A)_\alpha=\mathrm{Ker}(A)\ . \tag{18}$$

Similarly, under the significant level α_2,

$$\bigvee_{\beta\in\Lambda(B)|_{\beta\geq\alpha_2}}\beta\cdot(B)_\beta=\mathrm{Ker}(B)\ . \tag{19}$$

Eq. (17) follows,

$$\mu_{A\cap_F B}(x)$$
$$=\min\left[\max\left(\mu_{\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)}(x),\mu_{\mathrm{Ker}(A)}(x)\right),\right.$$
$$\left.\max\left(\mu_{\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)}(x),\mu_{\mathrm{Ker}(B)}(x)\right)\right]$$
$$=\max\left[\min\left(\mu_{\mathrm{Ker}(A)}(x),\mu_{\mathrm{Ker}(B)}(x)\right),\right.$$
$$\min\left(\mu_{\mathrm{Ker}(A)}(x),\mu_{\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)}(x)\right),$$

$$\min\left(\mu_{\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)}(x),\mu_{\mathrm{Ker}(B)}(x)\right),$$

$$\min\left(\mu_{\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)}(x),\right.$$

$$\left.\left.\mu_{\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)}(x)\right)\right]\ .$$

(by the distributive law) (20)

By the standard fuzzy intersection, Eq. (20) follows,

$$\mu_{A\cap_F B}(x)$$
$$=\max\left(\mu_{(\mathrm{Ker}(A)\cap_F\mathrm{Ker}(B))}(x),\right.$$
$$\mu_{\left(\mathrm{Ker}(A)\cap_F\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)\right)}(x),$$
$$\mu_{\left(\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)\cap_F\mathrm{Ker}(B)\right)}(x),$$
$$\left.\mu_{\left(\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)\cap_F\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)\right)}(x)\right)\quad(21)$$

By the standard fuzzy union, Eq. (21) follows,

$$\mu_{A\cap_F B}(x)=\mu_{\left([\mathrm{Ker}(A)\cap_F\mathrm{Ker}(B)]\cup_F\left[\mathrm{Ker}(A)\cap_F\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)\right]\right.}$$
$$\cup_F\left[\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)\cap_F\mathrm{Ker}(B)\right]$$
$$\cup_F\left[\left(\bigvee_{\alpha\in\Lambda(A)|_{\alpha<\alpha_1}}\alpha\cdot(A)_\alpha\right)\right.$$
$$\left.\left.\cap_F\left(\bigvee_{\beta\in\Lambda(B)|_{\beta<\alpha_2}}\beta\cdot(B)_\beta\right)\right]\right)(x)\ .$$

The proof is complete.

Example 2.3. From Example 2.1, the standard fuzzy union and intersection of A_3, A_5 can be written, respectively, as:

$$A_3\cup_F A_5=0.65/X_1+0.60/X_2+0.80/X_3+0.92/X_4\ .$$
$$A_3\cap_F A_5=0.50/X_1+0.10/X_2+0.30/X_3+0.40/X_4\ .$$

and the level set of A_3 is $\Lambda(A_3)=\{0.50,0.60,0.80,0.40\}$, the level set of A_5 is $\Lambda(A_5)=\{0.65,0.10,0.30,0.92\}$ Thus, the α-cuts of A_3 and A_5 according to their membership grades, respectively, become

$$(A_3)_{0.50}=1/X_1+1/X_2+1/X_3+0/X_4\ ,$$
$$(A_3)_{0.60}=0/X_1+1/X_2+1/X_3+0/X_4\ ,$$
$$(A_3)_{0.80}=0/X_1+0/X_2+1/X_3+0/X_4\ ,$$
$$(A_3)_{0.40}=1/X_1+1/X_2+1/X_3+1/X_4\ ;$$
$$(A_5)_{0.65}=1/X_1+0/X_2+0/X_3+1/X_4\ ,$$
$$(A_5)_{0.10}=1/X_1+1/X_2+1/X_3+1/X_4\ ,$$
$$(A_5)_{0.30}=1/X_1+0/X_2+1/X_3+1/X_4\ ,$$
$$(A_5)_{0.92}=0/X_1+0/X_2+0/X_3+1/X_4\ .$$

If we multiply each $(A_k)_\alpha$ by corresponding α value, $k=3,5$, then we convert each of the α-cuts to a special fuzzy set, i.e.

$$0.50\cdot(A_3)_{0.50}=0.50/X_1+0.50/X_2+0.50/X_3+0.00/X_4\ ,$$
$$0.60\cdot(A_3)_{0.60}=0.00/X_1+0.60/X_2+0.60/X_3+0.00/X_4\ ,$$
$$0.80\cdot(A_3)_{0.80}=0.00/X_1+0.00/X_2+0.80/X_3+0.00/X_4\ ,$$
$$0.40\cdot(A_3)_{0.40}=0.40/X_1+0.40/X_2+0.40/X_3+0.40/X_4\ ;$$
$$0.65\cdot(A_5)_{0.65}=0.65/X_1+0.00/X_2+0.00/X_3+0.65/X_4\ ,$$
$$0.10\cdot(A_5)_{0.10}=0.10/X_1+0.10/X_2+0.10/X_3+0.10/X_4\ ,$$
$$0.30\cdot(A_5)_{0.30}=0.30/X_1+0.00/X_2+0.30/X_3+0.30/X_4\ ,$$
$$0.92\cdot(A_5)_{0.92}=0.00/X_1+0.00/X_2+0.00/X_3+0.92/X_4\ .$$

Since, under the significant levels $\alpha_k=0.80$, $k=3,5$, we can get

$$\mathrm{Ker}(A_3)=\bigvee_{\alpha\in\Lambda(A_3)|_{\alpha\geq0.80}}\alpha\cdot(A_3)_\alpha=0.80\cdot(A_3)_{0.80}=0.00/X_1+0.00/X_2+0.80/X_3+0.00/X_4\ .$$

$$\mathrm{Ker}(A_5)=\bigvee_{\alpha\in\Lambda(A_5)|_{\alpha\geq0.80}}\alpha\cdot(A_5)_\alpha=0.92\cdot(A_5)_{0.92}=0.00/X_1+0.00/X_2+0.00/X_3+0.92/X_4\ .$$

and

$$\bigvee_{\alpha\in\Lambda(A_3)|_{\alpha<0.80}}\alpha\cdot(A_3)_\alpha$$
$$=\left[0.50\cdot(A_3)_{0.50}\right]\vee\left[0.60\cdot(A_3)_{0.60}\right]\vee\left[0.40\cdot(A_3)_{0.40}\right]$$
$$=(0.50\vee0.00\vee0.40)/X_1+(0.50\vee0.60\vee0.40)/X_2$$
$$+(0.50\vee0.60\vee0.40)/X_3+(0.00\vee0.00\vee0.40)/X_4$$
$$=0.50/X_1+0.60/X_2+0.60/X_3+0.40/X_4\ .$$

$$\bigvee_{\alpha\in\Lambda(A_5)|_{\alpha<0.80}} \alpha\cdot(A_5)_\alpha$$
$$= \left[0.65\cdot(A_5)_{0.65}\right]\vee\left[0.10\cdot(A_5)_{0.10}\right]\vee\left[0.30\cdot(A_5)_{0.30}\right]$$
$$= (0.65\vee 0.10\vee 0.30)/X_1 + (0.00\vee 0.10\vee 0.00)/X_2 + (0.00\vee 0.10\vee 0.30)/X_3 + (0.65\vee 0.10\vee 0.30)/X_4$$
$$= 0.65/X_1 + 0.10/X_2 + 0.30/X_3 + 0.65/X_4 \ .$$

Finally, by Theorem 2.2 (i), we get

$$A_3 \cup_F A_5 = [\mathrm{Ker}(A_3)\cup_F \mathrm{Ker}(A_5)] \cup_F \left[\left(\bigvee_{\alpha\in\Lambda(A_3)|_{\alpha<0.80}} \alpha\cdot(A_3)_\alpha\right) \cup_F \left(\bigvee_{\alpha\in\Lambda(A_5)|_{\alpha<0.80}} \alpha\cdot(A_5)_\alpha\right)\right]$$
$$= 0.65/X_1 + 0.60/X_2 + 0.80/X_3 + 0.92/X_4 \ .$$

Similarly, by Theorem 2.2 (ii), we get

$$A_3 \cap_F A_5 = [\mathrm{Ker}(A_3)\cap_F \mathrm{Ker}(A_5)] \cup_F \left[\mathrm{Ker}(A_3)\cap_F\left(\bigvee_{\alpha\in\Lambda(A_5)|_{\alpha<0.80}} \alpha\cdot(A_5)_\alpha\right)\right] \cup_F \left[\left(\bigvee_{\alpha\in\Lambda(A_3)|_{\alpha<0.80}} \alpha\cdot(A_3)_\alpha\right)\cap_F \mathrm{Ker}(A_5)\right] \cup_F \left[\left(\bigvee_{\alpha\in\Lambda(A_3)|_{\alpha<0.80}} \alpha\cdot(A_3)_\alpha\right) \cap_F \left(\bigvee_{\alpha\in\Lambda(A_5)|_{\alpha<0.80}} \alpha\cdot(A_5)_\alpha\right)\right]$$
$$= 0.50/X_1 + 0.10/X_2 + 0.30/X_3 + 0.40/X_4 \ .$$

Definition 2.3. [The Crisp set] *Let $X = \{X_t, t = 1, 2, \ldots, N\}$ be a universal set and $L = \{L_k, k = 1, 2, \ldots, r\}$ be a sequence of linguistic terms with monotone degree of semantic. For each $k = 1, 2, \ldots, r$, the crisp set H_k generated by L_k on X is a set, which contains all elements X_t map to the value 1.*

Note. For any fuzzy set A on X and the corresponding membership function μ_A. If the significant level $\alpha = 1$ and $\mu_A(x) = 1$ for each $x \in X$, then A and $\mathrm{Ker}(A)$ are equal. Moreover, A is called a crisp set on X.

3 Sample memberships estimation with respect to the kernel set

3.1 How to measure the membership for a sample with respect to its cluster center

It is interesting to see how far for a sample (sets) away from the kernel set in the sampling survey or time series analysis. Especially in the pattern identification processing, when we have several typical patterns (here means kernel set of the object), we are eager to know how far are the data different from the proposed type.

3.2 How to get the kernel set after clustering

In order to solve this problem, we will use the sample memberships with respect to the kernel set to estimate their distance. Firstly, we have to construct the kernel set under the features of the samples we have gathered. Which means we will learn the pattern from the experiments or the experience as the neural network did. An integrated procedure for deriving a kernel set with samples pattern is presented as follows:

A procedure for kernel set construction

Step 1. For time series $\{X_{it}\}_{i=1}^n$, do $i = 1, \ldots, n$.
Step 2. Input the time series $\{X_{it}\}$. Find the cluster center C_i for $\{X_{it}\}$.
Step 3. Let μ_{ti} be the degree of membership of each observation of $\{X_{it}\}$ to the cluster center C_i. Compute the membership μ_{ti} by

$$\mu_{ti} = \frac{1}{\|X_{it} - C_i\|}, \quad t = 1, \ldots, N \ .$$

where $\|\cdot\|$ denote Euclidean distance and if $\mu_{ti} > 1$, then let $\mu_{ti} = 1$.
Step 4. Constructing the fuzzy set A_i by its memberships.
Step 5. Choose a proper significant level α for A_i, and decide the kernel set $\mathrm{Ker}(A_i)$ of A_i.
Step 6. The kernel set learned from these samples will be $\mathrm{Ker}(A) = \bigcup_{i=1}^{n}{}_F\mathrm{Ker}(A_i)$.

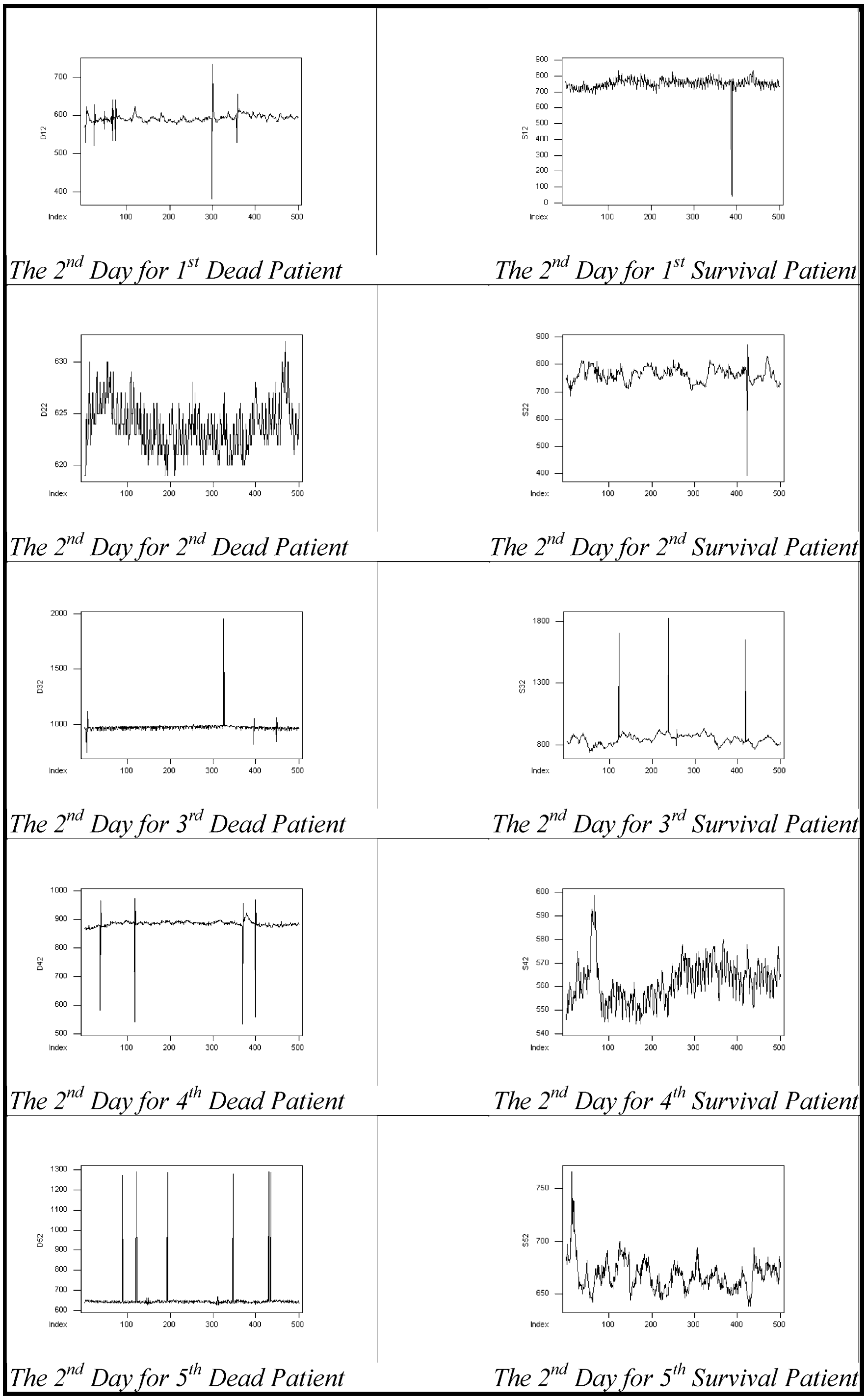
After we decide the kernel set form the samples we have known, we will compute the sample memberships. And under the significant level α, we will see how many data exceed the threshold value that will give us a useful suggestion for the final decision-making.

4 Applications with the RRI data

The data analyzed here comes from the ICU of Taipei Veterans General Hospital, 2001. The data records RRI of the dead patients and survival patients for the first four days of ICU. The RRI data for each patient is measured with 30 minutes. By discarding the first 100 observation, we analysis the 101 to 600 observations from each patient which contains about 1800–3000 observations of RRI. The purpose of this study is to extract features and identify nonlinear time series for the ICU patients. Figure 1a and b plot respectively the dead and survival patients' RRI. For the 500 observations, we can find the cluster centers for each data set. Now, under the significant level $\alpha = 0.9$, we can construct our kernel sets by the proposed procedures in the Sect. 3. In following, the dead patients' and the survival patients' cluster centers, radiuses of kernel set and ratios are showed in Table 3.

The kernel set learned from the dead patients is $\mathrm{Ker}(D) = \bigcup_{i=1}^{4}{}_F\mathrm{Ker}(D_{i2})$. Then, we can give the following testing-hypothesis procedure:

H_0: the data belongs to the dead patterns.
H_1: the data doesn't belong to the dead patterns.



(a) (b)

Fig. 1. Plots of RRI for dead and survival patients

Decision rule: for the new sample $\mathrm{Ker}(D_{new})$, under the significant level α_D, if there exist i, such that $\mathrm{Ker}(D_{new})/\mathrm{Ker}(D_{i2}) > 1 - a_D$ then we accept H_0. Otherwise, we reject H_0.

Similarly, the kernel set learned from the survival patients is $\mathrm{Ker}(S) = \bigcup_{i=1}^{4} {}_F\mathrm{Ker}(S_{i2})$, and the testing-hypothesis procedures:

H_0: the data belongs to the survival patterns.
H_1: the data doesn't belong to the survival patterns.

Decision rule: for the new sample $\mathrm{Ker}(S_{new})$, under the significant level α_D, if there exist i, such that $\mathrm{Ker}(S_{new})/\mathrm{Ker}(S_{i2}) > 1 - \alpha_D$ then we accept H_0. Otherwise, we reject H_0.

Now, we examine two new RRI samples of patients from ICU. First, we can find the cluster centers for each data set and construct their kernel sets under the significant level $\alpha = 0.9$. Then we can find radiuses of kernel sets for these samples. Finally, we can get their patterns according to their features by above the testing-hypothesis procedures.

For these two samples, we get the cluster centers 592.624 and 761.658 respectively. We can calculate the membership for each observation via the distance between observation and its cluster center. Under the significant level $\alpha = 0.9$, if the membership of observation is lager than 0.9, then this observation is a member of the kernel set. Therefore, the results of two new samples are showed in Table 4.

From Table 4 we can find that:

(1) For two new samples, the radiuses of kernel sets are 0.624 and 0.658 respectively.
(2) For new sample one, the ratio of observations which belongs to its kernel set and total observations is $0.094(= 47/500)$, which is larger than 0.05. This indicates that the patient has some features of dead patients.
(3) From Table 3, we obtain the significant level $\alpha_D = 0.05$. Under this condition, there exists D_{32} and D_{42} such that the ratio of observations which belongs to its kernel set and $\text{Ker}(D_{i2})$ is lager than 0.95, $i = 3, 4$. By way of decision rule, the patient will not be surviving.
(4) For new sample two, the ratio of observations which belongs to its kernel set and total observations is $0.034(= 17/500)$, which is smaller than 0.05. This indicates that the patient has some features of survival patients.
(5) Under the significant level $\alpha_D = 0.05$, there exists S_{22} such that the ratio of observations which belongs to its kernel set and $\text{Ker}(S_{22})$ is lager than 0.95. By way of decision rule, the patient will be alive.

5 Conclusion

In the medical science analysis discussed above the time series data have the uncertain property. If we use the conventional clustering methods to analyze these data, it will not solve the orientation problem. The contribution of

Table 3. The dead and the survival patients' cluster centers, radiuses and ratios

Patient	Cluster center	Radius of kernel set	Its ratio $\lvert\text{Ker}(D_{i2})\rvert/\lvert D_{i2}\rvert$	Patient	Cluster center	Radius of kernel set	Its ratio $\lvert\text{Ker}(S_{i2})\rvert/\lvert S_{i2}\rvert$
D_{12}	623.734	0.734	0.348	S_{12}	750.546	0.546	0.078
D_{22}	976.018	1.018	0.118	S_{22}	850.132	0.868	0.006
D_{32}	883.592	0.592	0.088	S_{32}	561.882	0.882	0.066
D_{42}	651.442	0.558	0.046	S_{42}	667.570	0.570	0.060
Average of $\lvert Ker(D_{i2})\rvert/\lvert D_{i2}\rvert \doteq 0.15$				Average of $\lvert Ker(S_{i2})\rvert/\lvert S_{i2}\rvert \doteq 0.05$			

Table 4. The sample memberships and kernel set for RRI of new samples

	The sample memberships and kernel set for RRI of new sample one (Cluster center: 592.624, $\alpha = 0.9$)			The sample memberships and kernel set for RRI of new sample two (Cluster center: 761.658, $\alpha = 0.9$)		
	Data	Memberships	Is a member of the kernel set?	Data	Memberships	Is a member of the kernel set?
1	569	0.042	no	751	0.094	no
2	573	0.051	no	740	0.046	no
3	571	0.046	no	760	0.603	no
4	529	0.016	no	734	0.036	no
5	622	0.034	no	730	0.032	no
6	598	0.186	no	729	0.031	no
7	609	0.061	no	718	0.023	no
8	614	0.047	no	713	0.021	no
9	608	0.065	no	708	0.019	no
10	605	0.081	no	741	0.048	no
⋮	⋮	⋮	⋮	⋮	⋮	⋮
491	591	0.616	no	737	0.041	no
492	591	0.616	no	725	0.027	no
493	591	0.616	no	722	0.025	no
494	595	0.421	no	717	0.022	no
495	598	0.186	no	721	0.025	no
496	598	0.186	no	721	0.025	no
497	595	0.421	no	717	0.022	no
498	592	1.000	yes	725	0.027	no
499	598	0.186	no	736	0.039	no
500	596	0.296	no	728	0.030	no
Total			$47(0.094 > 0.05)$			$17(0.034 \leq 0.05)$

this paper is that it provides a new method to cluster and identify time series. In this paper, we presented a procedure that can effectively cluster nonlinear time series into several patterns based on kernel set. The proposed algorithm also combines with the concept of a fuzzy set. We have demonstrated how to find a kernel set to help to cluster nonlinear time series into several patterns.

Our algorithm is highly recommended practically for clustering nonlinear time series and is supported by the empirical results. A major advantage of this framework is that our procedure does not require any initial knowledge about the structure in the data and can take full advantage of much more detailed information for some ambiguity.

However, certain challenging problems still remain open, such as:

(1) Since hardly ever any disturbance or noise in the data set can be completely eliminated, therefore, for the case of interacting noise, the complexity of multivariate filtering problems still remains to be solved.
(2) The convergence of the algorithm for clustering and the proposed statistics have not been well proved, although the algorithms and the proposed statistics are known as fuzzy criteria. This needs further investigation.

Although there remain many problems to be overcome, we think fuzzy statistical methods will be a worthwhile approach and will stimulate more future empirical work in time series analysis.

References

Barra V, Boire JY (2000) Tissue segmentation on MR images of the brain by possibilistic clustering on a 3D wavelet representation, J Magnetic Resonance Imaging **11**: 267–278

Cheng TW, Goldgof DB, Hall LO (1998) Fast fuzzy clustering, Fuzzy Sets and Systems **93**: 49–56

Cutsem BV, Gath I (1993) Detection of outliers and robust estimation using fuzzy clustering, Comput Statistics Data Anal **15**: 47–61

Klir GJ, Yuan B (1995) Fuzzy Sets and Fuzzy Logic. Prentice Hall, Upper Saddle River, NJ

Romer C, Kandel A, Backer E (1995) Fuzzy partitions of the sample space and fuzzy parameter hypotheses, IEEE Trans Systems, Man, and Cybernetics **125**: 1314–1321

Tseng T, Klein C (1992) A new algorithm for fuzzy multicriteria decision making, Int J Approximate Reasoning **6**: 45–66

Werners B (1987) An interactive fuzzy programming system, Fuzzy Sets and Systems **23**: 131–147

Wu B, Hsu YY (1999) Fuzzy statistical cluster analysis and its application to Taiwan's unemployment, J Chinese Statistical Association **37**: 37–52

Yang MS (1993) A survey of fuzzy clustering, Mathematical and Computer Modelling **18**: 1–16

Yoshinari Y, Pedrycz W, Hirota K (1993) Construction of fuzzy models through clustering techniques, Fuzzy Sets and Systems **54**: 157–165

Zadeh LA (1965) Fuzzy Sets. Information and Control **8**: 338–353

Zimmermann HJ (1991) Fuzzy Set Theory and Its Applications. Kluwer Academic, Boston