Probabilistic Fuzzy Systems as Additive Fuzzy Systems

Rui Jorge Almeida[1], Nick Verbeek[2], Uzay Kaymak[1], and João Miguel Da Costa Sousa[3]

[1] School of Industrial Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
rjalmeida@tue.nl, u.kaymak@ieee.org
[2] Erasmus University Rotterdam, Erasmus School of Economics, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
291847nv@student.eur.nl
[3] Universidade de Lisboa, Instituto Superior Tecnico, Dept. of Mechanical Engineering, CSI, IDMEC/LAETA
Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
jmsousa@tecnico.ulisboa.pt

Abstract. Probabilistic fuzzy systems combine a linguistic description of the system behaviour with statistical properties of data. It was originally derived based on Zadeh's concept of probability of a fuzzy event. Two possible and equivalent additive reasoning schemes were proposed, that lead to the estimation of the output's conditional probability density. In this work we take a complementary approach and derive a probabilistic fuzzy system from an additive fuzzy system. We show that some fuzzy systems with universal approximation capabilities can compute the same expected output value as probabilistic fuzzy systems and discuss some similarities and differences between them. A practical relevance of this functional equivalence result is that learning algorithms, optimization techniques and design issues can, under certain circumstances, be transferred across different paradigms.

Keywords: Probabilistic Fuzzy Systems, Additive Fuzzy Systems, Conditional Density Approximation.

1 Introduction

Probabilistic fuzzy systems (PFS) can deal explicitly and simultaneously with fuzziness or linguistic uncertainty and probabilistic uncertainty. A probabilistic fuzzy system follows an idea similar to [1–4] where the different concepts [5–8] of fuzzy sets and probabilities are complementary [6].

As a mathematical notion, a fuzzy set F on a finite universe U is unambiguously defined by a membership function $u_F : U \to [0, 1]$. The mathematical object representing the fuzzy set is the membership function $u_F(x)$ indicating the grade of membership of element of $x \in U$ in F. At the mathematical level the domain of the fuzzy sets is $[0, 1]^U$. On the other hand a probabilistic measure Pr

A. Laurent et al. (Eds.): IPMU 2014, Part I, CCIS 442, pp. 567–576, 2014.

of an experiment ϵ yet to be performed, is a mapping $2^U \to [0,1]$ that assigns a number $\Pr(A)$ of event A to each subset of U, satisfying the Kolmogorov axioms. $\Pr(A)$ is the probability that a generic outcome of ϵ, an ill-known single-valued variable x, hits the well-known set A. If the outcome of ϵ is such that $x \in A$, then we say that event A has occurred. In this case there is uncertainty about the occurrence of any particular x and consequently of event A. This uncertainty is described by $\Pr(A)$. At the mathematical level the domain of the mapping Pr is the Boolean algebra 2^U.

Various rule base structures and reasoning mechanisms for fuzzy systems (e.g. [9–11]), emphasize the modelling of the linguistic uncertainty and interpolation capability of fuzzy systems, being typically used for approximating deterministic functions, in which the stochastic uncertainty is ignored. A probabilistic fuzzy system, as it was formally defined in [12], was based on the concept of probability of fuzzy events. This type of system estimates a conditional probability density function for the output variable, given the inputs to the system. Two equivalent additive reasoning mechanism have been proposed for PFS, one based on the concept of fuzzy histograms and another based on the stochastic mapping between fuzzy antecedents and fuzzy consequents.

In this work we follow a different reasoning and derive a probabilistic fuzzy system starting from a additive fuzzy system. This different analysis provides a different insight and understanding of probabilistic fuzzy systems, which can be related to Mamdani fuzzy systems and fuzzy relational models and departs from the concept of probability of fuzzy events. This allows us to formalize the definition of probabilistic fuzzy systems while exposing similarities and differences with different models or concepts. The relation of probabilistic fuzzy system to well known fuzzy systems helps to explain its success for function approximation. A practical relevance of the functional equivalence presented in this work is that learning algorithms, optimization techniques and design issues can be transferred to probabilistic fuzzy systems. Furthermore, it also allows to interpret models transversely across different modeling paradigms.

The outline of the paper is as follows. In Section 2, we give an overview of the original definition of probabilistic fuzzy systems and present the two equivalent additive reasoning mechanisms of a PFS, as well as the different outputs. In Section 3 we present the new derivation of a probabilistic fuzzy system starting from fuzzy additive systems and discuss in Section 4 several issues related to our findings. Finally we conclude the paper in Section 5.

2 Probabilistic Fuzzy Systems

Probabilistic fuzzy systems combine two different types of uncertainty, namely fuzziness or linguistic vagueness, and probabilistic uncertainty. In this work we consider that the probabilistic uncertainty relate to aleatoric variability, while fuzzy sets are used to represent gradualness, epistemic uncertainty or bipolarity [7, 13].

The PFS consists of a set of rules whose antecedents are fuzzy conditions and whose consequents are probability distributions. Assuming that the input space is a subset of $\mathbb{R}^n$ and that the rule consequents are defined on a finite domain $Y \subseteq \mathbb{R}$, a probabilistic fuzzy system consists of a system of rules R_q, $q = 1, \ldots, Q$, of the type

$$R_q : \text{If } \mathbf{x} \text{ is } A_q \text{ then } f(y) \text{ is } f(y|A_q)\,, \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^n$ is an input vector, $A_q : X \longrightarrow [0,1]$ is a fuzzy set defined on X and $f(y|A_q)$ is the conditional pdf of the stochastic output variable $\underline{y}$ given the fuzzy event A_q. The interpretation is as follows: if fuzzy antecedent A_q is fully valid ($x \in \text{core}(A_q)$), then y is a sample value from the probability distribution with conditional pdf $f(y|A_q)$.

A PFS has been described with two possible and equivalent reasoning mechanisms, namely the fuzzy histogram approach and the probabilistic fuzzy output approach [12]. In both cases, we suppose that J fuzzy classes C_j form a fuzzy partition of the compact output space Y.

2.1 Fuzzy Histogram Model

In the fuzzy histogram approach, we replace in each rule of (1) the true pdf $f(y|A_q)$ by its fuzzy approximation (fuzzy histogram) $\hat{f}(y|A_q)$ yielding the rule set $\hat{R}_q$, $q = 1, \ldots, Q$ defined as

$$\hat{R}_q : \text{If } \mathbf{x} \text{ is } A_q \text{ then } f(y) \text{ is } \hat{f}(y|A_q)\,, \tag{2}$$

where $\hat{f}(y|A_q)$ is a fuzzy histogram conform [14]

$$\hat{f}(y|A_q) = \sum_{j=1}^{J} \frac{\hat{\Pr}(C_j|A_q)u_{C_j}(y)}{\int_{-\infty}^{\infty} u_{C_j}(y)dy}\,. \tag{3}$$

The numerator in (3) describes a superposition of fuzzy events described by their membership functions $u_{C_j}(y)$, weighted by the probability $\hat{\Pr}(C_j|A_q)$ of the fuzzy event. The denominator of (3) is a normalizing factor representing the fuzzified size of class C_j. Because of overlapping membership functions, fuzzy histograms have a high level of statistical efficiency, compared to crisp ones and several classes of fuzzy histograms also have a high level of computational efficiency [15].

The interpretation of this type of reasoning is as follows. Given the occurrence of a (multidimensional) antecedent fuzzy event A_q, which is a conjunction of the fuzzy conditions defined on input variables, an estimate of the conditional probability density function based on a fuzzy histogram $\hat{f}(y|A_q)$ is calculated.

2.2 Probabilistic Fuzzy Output Model

In the probabilistic fuzzy output approach, sometimes also referred to as Mamdani PFS [16–18], we decompose each rule (1) to provide a stochastic mapping between its fuzzy antecedents and its fuzzy consequents. The rules are written in the following form.

$$\text{Rule } \hat{R}_q\text{: If } \mathbf{x} \text{ is } A_q \text{ then } \begin{array}{l} \underline{y} \text{ is } C_1 \text{ with } \hat{\Pr}(C_1|A_q) \text{ and} \\ \quad\quad\quad\ldots \\ \underline{y} \text{ is } C_J \text{ with } \hat{\Pr}(C_J|A_q). \end{array} \tag{4}$$

These rules specify a probability distribution over a collection of fuzzy sets that partition the output domain. The rules of a PFS also express linguistic information and they can be used to explain the model behaviour by a set of linguistic rules. This way, the system deals both with linguistic uncertainty as well as probabilistic uncertainty.

The interpretation for the probabilistic fuzzy output approach is as follows. Given the occurrence of a (multidimensional) antecedent fuzzy event A_q, which is a conjunction of the fuzzy conditions defined on input variables, each of the consequent fuzzy events C_j is likely to occur. The selection of consequent fuzzy events is done proportionally to the conditional probabilities $\hat{\Pr}(C_j|A_q)$, ($j = 1, 2, \ldots, J$). This applies for all the rules R_q, $q = 1, 2, \ldots, Q$.

The probabilistic fuzzy system in this form resembles a deterministic Mamdani fuzzy system with rule base multiplicative implication and additive aggregation. The difference lies in the fact that in a Mamdani fuzzy system only one of the outputs is considered in each rule, while in a PFS, each fuzzy output C_q can happen with a given conditional probability $\hat{\Pr}(A_q|C_j)$.

2.3 Outputs of Probabilistic Fuzzy Systems

Although the fuzzy rule bases (2) and (4) are different, under certain conditions, the two corresponding probabilistic fuzzy systems implement the same crisp input-output mapping [12]. The output of the fuzzy rules (4) is the same as in the rules (2), if an additive reasoning scheme is used with multiplicative aggregation of the rule antecedents [19].

Given an input vector $\mathbf{x}$, the output of a probabilistic fuzzy system is a conditional density function which can be computed as

$$\hat{f}(y|\mathbf{x}) = \sum_{j=1}^{J}\sum_{q=1}^{Q} \beta_q(\mathbf{x})\hat{\Pr}(C_j|A_q)\frac{u_{C_j}(y)}{\int_{-\infty}^{\infty} u_{C_j}(y)dy}, \tag{5}$$

where

$$\beta_q(\mathbf{x}) = \frac{u_{A_q}(\mathbf{x})}{\sum_{q'=1}^{Q} u_{A_{q'}}(\mathbf{x})} \tag{6}$$

is the normalised degree of fulfillment of rule R_q and u_{A_q} is the degree of fulfillment of rule R_q. When $\mathbf{x}$ is n-dimensional, u_{A_q} is determined as a conjunction of the individual memberships in the antecedents computed by a suitable t-norm, *i.e.*, $u_{A_q}(\mathbf{x}) = u_{A_{q1}}(x_1) \circ \cdots \circ u_{A_{qn}}(x_n)$, where $x_i, i = 1, \ldots, n$ is the i-th component of $\mathbf{x}$ and $\circ$ denotes a t-norm. From the obtained output probability distribution it is possible to calculate a crisp output using the expected value

$$\hat{\mu}_{y|\mathbf{x}} = \hat{E}(y|\mathbf{x}) = \int_{-\infty}^{\infty} y\hat{f}(y|\mathbf{x})dy = \sum_{q=1}^{Q}\sum_{j=1}^{J}\beta_q(\mathbf{x})\hat{\Pr}(C_j|A_q)z_{1,j}\,, \tag{7}$$

where $z_{1,j}$ is the centroid of the jth output fuzzy set defined by

$$z_{1,j} = \frac{\int_{-\infty}^{\infty} y u_{C_j}(y)dy}{\int_{-\infty}^{\infty} u_{C_j}(y)dy}\,. \tag{8}$$

It can be shown [12] that the conditional density output $\hat{f}(y|\mathbf{x})$ of a PFS is a proper probability density function *i.e.,* $\int_{-\infty}^{\infty} \hat{f}(y|\mathbf{x})dy = 1$ and that the crisp outputs, expected value $\hat{\mathrm{E}}(\underline{y}|\mathbf{x})$ and second moment $\hat{\mathrm{E}}(\underline{y}^2|\mathbf{x})$, exist if the output space is well-formed, *i.e.* the output membership values satisfy

$$\sum_{j=1}^{J} u_{C_j}(y) = 1, \qquad \forall y \in Y, y < \infty. \tag{9}$$

3 Probabilistic Fuzzy Systems as Additive Fuzzy Systems

In this work we depart from the previous definition presented in Section 2.2 and instead derive a probabilistic fuzzy system from an additive fuzzy system. This deterministic fuzzy system has rule base multiplicative implication and additive aggregation, where the crisp model output is obtained using the center of gravity defuzzification method. In the following section we present the additive fuzzy system under consideration and how it can be converted in a probabilistic fuzzy system.

3.1 Additive Fuzzy Systems

Let $R = \cup_{q=1}^{Q} R_q$ be a rule base for a additive fuzzy system of the type

$$\begin{array}{ll} \text{Rule } \hat{R}_q\text{: If } \mathbf{x} \text{ is } A_q \text{ then} & \underline{y} \text{ is } C_1 \text{ with } w(A_q, C_1) \text{ and} \\ & \qquad \dots \\ & \underline{y} \text{ is } C_J \text{ with } w(A_q, C_J)\,, \end{array} \tag{10}$$

where $w(A_q, C_j) \in \mathbb{R}_{\geq 0}$ are non-negative weights. The system defined by (10) is similar to the standard additive model [20, 21] but in the former, the consequents are not directly dependent on x.

Although the fuzzy rule base system defined by (10) can be obtained by replacing the conditional probabilities $\hat{\Pr}(C_j|A_q)$ by non-negative weights $w(A_q, C_j) \in \mathbb{R}_{\geq 0}$ in the fuzzy rule system (4), the crisp output of both systems is different, as the following theorem shows.

Theorem 1. *Let $R = \cup_{q=1}^{Q} R_q$ be a fuzzy rule base as defined by (10) such that $u_{A_q}(\mathbf{x}) > 0, \forall q$ and the output space follows (9), and the rule base uses multiplicative implication and additive aggregation. Then the crisp model output y^* obtained using the center of gravity defuzzification method is*

$$y^* = \frac{\sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q, C_j) v_{1,j} z_{1,j}}{\sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q, C_j) v_{1,j}}\,, \tag{11}$$

where $z_{1,j}$ is given by (8) and $v_{1,j}$ is the area of the jth output fuzzy set defined by

$$v_{1,j} = \int_{-\infty}^{\infty} u_{C_j}(y) dy\,. \tag{12}$$

Proof. The center of gravity defuzzification method is given by

$$y^* = \frac{\int_{-\infty}^{\infty} y \chi(x,y) dy}{\int_{-\infty}^{\infty} \chi(x,y) dy}\,, \tag{13}$$

where $\chi(x, y)$ is the output of the fuzzy system under consideration. For the case of the additive fuzzy system (10) using with multiplicative implication and additive aggregation the output is

$$\chi(x,y) = \sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q, C_j) u_{C_j}(y)\,. \tag{14}$$

Substituting (14) into (13) we obtain

$$\begin{aligned} y^* &= \frac{\int_{-\infty}^{\infty} y \sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q,C_j) u_{C_j}(y) dy}{\int_{-\infty}^{\infty} \sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q,C_j) u_{C_j}(y) dy} \\ &= \frac{\sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q,C_j) \int_{-\infty}^{\infty} y u_{C_j}(y) dy}{\sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q,C_j) \int_{-\infty}^{\infty} u_{C_j}(y) dy} \\ &= \frac{\sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q,C_j) \int_{-\infty}^{\infty} u_{C_j}(y) dy \frac{\int_{-\infty}^{\infty} y u_{C_j}(y) dy}{\int_{-\infty}^{\infty} u_{C_j}(y) dy}}{\sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q,C_j) \int_{-\infty}^{\infty} u_{C_j}(y) dy} \\ &= \frac{\sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q,C_j) v_{1,j} z_{1,j}}{\sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q,C_j) v_{1,j}}\,. \end{aligned} \tag{15}$$

3.2 Equivalence to Probabilistic Fuzzy Systems

Starting from an additive fuzzy system defined in (10), it is possible to obtain a probabilistic fuzzy system. Before formalizing this result we introduce the following definition of a probability kernel.

Definition 1. *A kernel is a mapping $\mathcal{K} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_{\geq 0}$ from the measurable space $(\mathbf{X}, \mathcal{X})$ to the measurable space $(\mathbf{Y}, \mathcal{Y})$. The kernel $\mathcal{K}$ is a probability kernel if it is defined as a probability measure on $(\mathbf{Y}, \mathcal{Y})$.*

Given this definition we can now prove that a probabilistic fuzzy system can be obtained starting from the fuzzy system defined in (10).

Theorem 2. *If the mapping $w(A_q, C_j)$ is defined as a probability kernel and each output consequent C_j are functions defined on a random variable space then the output of the PFS is a conditional probability density for y given x. Under this definition, the fuzzy rule base in (10) has a functional equivalent to the PFS in (4) and the crisp output (11) has a functional equivalent to the conditional output of the PFS in (5).*

Proof. The defined non-negative weights $w(A_q, C_j) : (\mathbf{X} \times Y) \to R_{\geq 0}$ form a kernel on the measurable space $(\mathbb{R}^n \times \mathbb{R})$. If $w(A_q, C_j)$ is also defined as a probability measure on $(Y, \mathcal{Y})$, such that $\sum_{j=1}^{J} w(A_q, C_j) = 1, \forall q = 1, \ldots, Q$ then according to Definition 1, $w(A_q, C_j)$ is a probability kernel. We recall that using (6) we obtain $\sum_{q=1}^{Q} \beta_q(\mathbf{x}) = 1$. Furthermore, since the output fuzzy sets C_j are admissible functions for defining random variables then they are limited to those for which a probability distribution exists. A simple form to ensure this is to normalize them

$$u_{C_j'} = \frac{u_{C_j}(y)}{\int_{-\infty}^{\infty} u_{C_j}(y)dy} . \tag{16}$$

The output of the fuzzy system $\chi(x, y)$ in (13) is then a conditional density function for Y given X such that:

$$\begin{aligned} \int_{-\infty}^{\infty} \chi(x, y)dy &= \int_{-\infty}^{\infty} \sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q, C_j) u_{C_j'} dy \\ &= \sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q, C_j) \frac{\int_{-\infty}^{\infty} u_{C_j}(y)}{\int_{-\infty}^{\infty} u_{C_j}(y)dy} = 1 . \end{aligned} \tag{17}$$

In the case that $w(A_q, C_j)$ is defined as a probability kernel, the additive fuzzy system defined by the rule base (10) is a probabilistic fuzzy system as presented in (4). Furthermore, the center of gravity output (11) of the additive fuzzy system has a functional equivalent to the expectation of the conditional output of the PFS (5)

$$\begin{aligned} y^* &= \frac{\sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q, C_j) v_{1,j} z_{1,j}}{\sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q, C_j) v_{1,j}} \\ &= \sum_{q=1}^{Q} \sum_{j=1}^{J} \beta_q(\mathbf{x}) w(A_q, C_j) z_{1,j} . \end{aligned} \tag{18}$$

Since $w(A_q, C_j)$ is a probability kernel, (18) is equivalent to (7).

A practical relevance of the functional equivalence result is that learning algorithms, optimization techniques and design issues can be transferred across different paradigms. Furthermore, this result helps to explain the success of fuzzy systems for function approximation in the presence of probabilistic uncertainty.

4 Discussion

The previous sections have shown that a probabilistic fuzzy system defined by (4) can be obtained starting from a additive fuzzy system (10). An important aspect is that since $w(A_q, C_j)$ is defined as a probability kernel then it has a functional equivalent to $\Pr(C_j|A_q)$ in (4), implying that $\sum_{j=1}^{J} \hat{\Pr}(C_j|A_q) = 1$ and $\hat{\Pr}(C_j|A_q) \geq 0$. In this paper we do not assume any particular algebraic structure for the conditional probability of fuzzy events. There are several examples of definitions of conditional probabilities of fuzzy events that satisfy the classical axioms of conditional probabilities as given by [22] and [23] that can be found in [24–26]. This is an important issue that needs to be studied closely in the future.

It is also interesting to note that the system defined by (10) can be transformed in a fuzzy relational model [11] when $w(A_q, C_j)$ is replaced by the fuzzy relation $u(A_q, C_j)$. Similarly to a fuzzy relational model, a probabilistic fuzzy system can also be fine tuned by modifying the probability parameters $\hat{\Pr}(C_j|A_q)$, while maintaining the fuzzy input and fuzzy output space constant. We stress that a fuzzy relational model and a probabilistic fuzzy system have different interpretations, based on the nature of the uncertainty of the relation and output being modelled, as described in Section 1. In a fuzzy relational model the elements of the relation represent the strength of association between the fuzzy sets, while in the case of a fuzzy probabilistic model they are a stochastic mapping between fuzzy sets. Furthermore, the output fuzzy sets of a probabilistic fuzzy system are defined in the space of a stochastic variable $\underline{y}$. These differences leads to different nature of outputs, albeit under certain circumstances, there is a functional equivalence between both models crisp output. In the general case that $w(A_q, C_j)$ are non-negative weights, or in the case of a fuzzy relational model $u(A_q, C_j)$ are fuzzy relations, the output of such a system is not a proper probability density function.

As a result of theorem 1 and theorem 2, a Mamdani fuzzy model can be regarded as a special case of the fuzzy system defined in (10), or equivalently the system defined by (4). A Mamdani fuzzy model is recovered when the system is purely deterministic by setting setting for all $q = 1, \ldots, Q$, $\exists \kappa \in \{1 \ldots, J\}$ such that $\hat{\Pr}(C_\kappa|A_q) = 1$ and $\hat{\Pr}(C_j|A_q) = 0$, $j \neq \kappa$ *i.e.,* only one of the possible consequents is certain for each rule Q.

5 Conclusions

This paper presents a new form to derive a probabilistic fuzzy system starting from an additive fuzzy system. This new reasoning departs from the original derivation of a PFS which was based on Zadehs' concept of probability of a fuzzy event. We show that in certain cases an additive fuzzy system can compute the same expected output value as a PFS. We discuss some similarities between Mamdani and fuzzy relation models with probabilistic fuzzy systems. A practical relevance of the functional equivalence result is that learning algorithms, optimization techniques and design issues can be transferred across different paradigms. Furthermore, our results provide insight why additive deterministic fuzzy systems have proven to be so successful for function approximation purposes.

Acknowledgements. The authors would like to thank Didier Coquin for the discussions and help provided in this work. This work was supported by French National Agency for Research with the reference ANR-10-CORD-005 (ReVeS project).

References

1. Meghdadi, A.H., Akbarzadeh-T., M.R.: Probabilistic fuzzy logic and probabilistic fuzzy systems. In: Proceedings of the Tenth IEEE International Conference on Fuzzy Systems, Melbourne, Australia, pp. 1127–1130 (2001)
2. Liu, Z., Li, H.X.: A probabilistic fuzzy logic system for modeling and control. IEEE Transactions on Fuzzy Systems 13, 848–859 (2005)
3. Zhang, G., Li, H.X.: A probabilistic fuzzy learning system for pattern classification. In: 2010 IEEE International Conference on Systems Man and Cybernetics (SMC), pp. 2336–2341 (2010)
4. Zhang, G., Li, H.X.: An efficient configuration for probabilistic fuzzy logic system. IEEE Transactions on Fuzzy Systems 20, 898–909 (2012)
5. Zadeh, L.A.: Probability measures of fuzzy events. Journal of Mathematical Analysis and Applications 23, 421–427 (1968)
6. Zadeh, L.A.: Discussion: Probability theory and fuzzy logic are complementary rather than competitive. Technometrics 37, 271–276 (1995)
7. Dubois, D., Prade, H.: The three semantics of fuzzy sets. Fuzzy Sets and Systems 90, 141–150 (1997)
8. Bertoluzza, C., Gil, M.A., Ralescu, D.A. (eds.): Statistical Modeling, Analysis and Management of Fuzzy Data. STUDFUZZ. Physica Verlag, Heidelberg (2002)
9. Mamdani, E.H., Gaines, B.R. (eds.): Fuzzy Reasoning and its Applications. Academic Press, London (1981)
10. Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modelling and control. IEEE Transactions on Systems, Man and Cybernetics 15, 116–132 (1985)
11. Pedrycz, W.: Applications of fuzzy relational equations for methods of reasoning in presence of fuzzy data. Fuzzy Sets and Systems 16, 163–175 (1985)
12. van den Berg, J., Kaymak, U., Almeida, R.J.: Conditional density estimation using probabilistic fuzzy systems. Accepted for Publication in IEEE Transactions on Fuzzy Systems (2013)
13. Dubois, D., Prade, H.: Gradualness, uncertainty and bipolarity: Making sense of fuzzy sets. Fuzzy Sets and Systems 192, 3–24 (2012)
14. van den Berg, J., van den Bergh, W.M., Kaymak, U.: Probabilistic and statistical fuzzy set foundations of competitive exception learning. In: Proceedings of the Tenth IEEE International Conference on Fuzzy Systems, Melbourne, Australia, vol. 2, pp. 1035–1038 (2001)
15. Waltman, L., Kaymak, U., van den Berg, J.: Fuzzy histograms: A statistical analysis. In: EUSFLAT-LFA 2005 Joint 4th EUSFLAT 11th LFA Conference, pp. 605–610 (2005)
16. Kaymak, U., van den Bergh, W.M., van den Berg, J.: A fuzzy additive reasoning scheme for probabilistic Mamdani fuzzy systems. In: Proceedings of the 2003 IEEE International Conference on Fuzzy Systems, St. Louis, USA, vol. 1, pp. 331–336 (2003)

17. Xu, D., Kaymak, U.: Value-at-risk estimation by using probabilistic fuzzy systems. In: Proceedings of the 2008 World Congress on Computational Intelligence, Hong-Kong, pp. 2109–2116 (2008)
18. Almeida, R., Kaymak, U.: Probabilistic fuzzy systems in Value-at-Risk estimation. Intelligent Systems in Accounting, Finance & Management 16, 49–70 (2009)
19. van den Berg, J., Kaymak, U., van den Bergh, W.M.: Financial markets analysis by using a probabilistic fuzzy modelling approach. International Journal of Approximate Reasoning 35, 291–305 (2004)
20. Kosko, B.: Fuzzy Engineering. Prentice Hall, Upper Saddle River (1997)
21. Kosko, B.: Global stability of generalized additive fuzzy systems. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 28, 441–452 (1998)
22. de Finetti, B.: Sull'impostazione assiomatica del calcolo delle probabilità. Annali Univ. Trieste 19, 3–55 (1949); (English translation: Probability, Induction, Statistics. Wiley, London, 1972 (Chapter 5)
23. Popper, K.: The Logic of Scientific Discovery. Routledge, London (1959)
24. Coletti, G., Scozzafava, R.: Conditional probability and fuzzy information. Computational Statistics & Data Analysis 51, 115–132 (2006)
25. Baldwin, J., Lawry, J., Martin, T.: A mass assignment theory of the probability of fuzzy events. Fuzzy Sets and Systems 83, 353–367 (1996)
26. Singpurwalla, N.D., Booker, J.M.: Membership functions and probability measures of fuzzy sets. Journal of the American Statistical Association 99, 867–877 (2004)