6-2017

How to Gauge the Accuracy of Fuzzy Control Recommendations: A Simple Idea

Patricia Melin
Instituto Tecnologico de Tijuana, pmelin@tectijuana.mx

Oscar Castillo
Instituto Tecnologico de Tijuana, ocastillo@tectijuana.mx

Andrzej Pownuk
University of Texas at El Paso, ampownuk@utep.edu

Olga Kosheleva
University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich
University of Texas at El Paso, vladik@utep.edu

Comments:
Technical Report: UTEP-CS-17-56
To appear in: *Proceedings of the 2017 Annual Conference of the North American Fuzzy Information Processing Society NAFIPS'2017*, Cancun, Mexico, October 16-18, 2017

How to Gauge the Accuracy of Fuzzy Control Recommendations: A Simple Idea

Patricia Melin, Oscar Castillo, Andrzej Pownuk, Olga Kosheleva, and Vladik Kreinovich

Abstract Fuzzy control is based on approximate expert information, so its recommendations are also approximate. However, the traditional fuzzy control algorithms do not tell us how accurate are these recommendations. In contrast, for the probabilistic uncertainty, there is a natural measure of accuracy: namely, the standard deviation. In this paper, we show how to extend this idea from the probabilistic to fuzzy uncertainty and thus, to come up with a reasonable way to gauge the accuracy of fuzzy control recommendations.

1 Formulation of the Problem

Need to gauge accuracy of fuzzy recommendations. Fuzzy logic (see, e.g., [1, 4, 6]) has been successfully applied to many different application areas.

For example, in control – one of the main applications of fuzzy techniques – fuzzy techniques enable us to generate the control value appropriate for a given situation.

A natural question is: with what accuracy do we need to implement this recommendation? In many applications, this is an important question: it is often much easier to implement the control value approximately, by using a simple approximate actuator, but maybe a more accurate actuator is needed? To answer this question, we must have a natural way to gauge the accuracy of the corresponding recommendations.

Patricia Melin and Oscar Castillo
Department of Computer Science, Tijuana Institute of Technology, Tijuana, Baja California, Mexico, e-mail: ocastillo@tectijuana.mx, pmelin@tectijuana.mx

Andrzej Pownuk, Olga Kosheleva, and Vladik Kreinovich
University of Texas at El Paso, 500 W. University,
El Paso, Texas 79968, USA, e-mail: ampownuk@utep.edu, olgak@utep.edu, vladik@utep.edu

Such gauging is possible for probabilistic uncertainty. In a similar case of probabilistic uncertainty, there is such a natural way to gauge the accuracy; see, e.g., [5].

Namely, probabilistic uncertainty means that instead of the exact value x, we only know a probability distribution – which can be described, e.g., by the probability density $\rho(x)$. In this situation, if we need to select a single value x, a natural idea is to select, e.g., the mean value $\overline{x} = \int x \cdot \rho(x)\,dx$.

A natural measure of accuracy of this means is the mean square deviation from the mean, known as the standard deviation:

$$\sigma \stackrel{\text{def}}{=} \sqrt{\int (x - \overline{x})^2\,dx}.$$

What we do in this paper. In this paper, we provide a similar way to gauge the accuracy of fuzzy recommendations, i.e., a recommendations in which, instead of using a probability density function $\rho(x)$, we start with a membership function $\mu(x)$.

2 Main Idea

How we elicit fuzzy degrees: a brief reminder. To explain our idea, let us recall how fuzzy degrees $\mu(x)$ corresponding to different values x are elicited in the first place.

At first glance, the situation may look straightforward: for each possible value x of the corresponding quantity, we ask the expert to mark, on a scale from 0 to 1, his/her degree of confidence that x satisfies the given property. For example, if we are eliciting the membership function describing smallness, we ask the expert to specify the degree to which the value x is small.

In some cases, this is all we need. However, in many other cases, we get a *non-normalized* membership function, for which the largest value $\mu(x)$ is smaller than 1. Most fuzzy techniques assume that the membership function is normalized. So, after the elicitation, we sometimes need to perform an additional step to get an easy-to-process membership function: namely, we *normalize* the original values $\mu(x)$ by dividing them by the largest of the values $\mu(y)$. Thus, we get the function

$$\mu'(x) \stackrel{\text{def}}{=} \frac{\mu(x)}{\max\limits_{y} \mu(y)}.$$

Sometimes, the original fuzzy degrees come from subjective probabilities. Sometimes, the experts have some subjective probabilities assigned to different values x. In this case, when asked to indicate their degree of certainty, they may list the values of the corresponding probability density function $\rho(x)$.

This function is rarely normalized. After normalizing it, we get the membership function

$$\mu(x) = \frac{\rho(x)}{\max\limits_{y} \rho(y)}. \tag{1}$$

Let us use this idea to gauge the accuracy of fuzzy recommendations. Formula (1) assigns, to each probability density function $\rho(x)$, an appropriate membership function $\mu(x)$. Vice versa, one can easily see if we know that the membership function $\mu(x)$ was obtained by normalizing some probability density function $\rho(x)$, then we can uniquely reconstruct this probability density function $\rho(x)$: namely, since $\mu(x) = c \cdot \rho(x)$ for some normalizing constant c, we thus have $\rho(x) = C \cdot \mu(x)$, for another constant $C = \frac{1}{c}$. So, all we need to find the probability density function is to find the coefficient C.

This coefficient can be easily found from the condition that the overall probability be 1, i.e., that $\int \rho(x)\,dx = 1$. Substituting $\rho(x) = C \cdot \mu(x)$ into this formula, we conclude that $C \cdot \int \mu(x)\,dx = 1$, thus $C = \frac{1}{\int \mu(y)\,dy}$ and therefore,

$$\rho(x) = C \cdot \mu(x) = \frac{\mu(x)}{\int \mu(y)\,dy}. \tag{2}$$

Our idea is then to use the probabilistic formulas corresponding to this artificial distribution.

This makes sense. Does this make sense? The probabilistic measure of accuracy is based on the assumption that we use the mean, but don't we use something else in fuzzy?

Actually, not really. The mean of the distribution (2) is

$$\bar{x} = \int x \cdot \rho(x)\,dx = \frac{\int x \cdot \mu(x)\,dx}{\int \mu(x)\,dx}.$$

This is exactly the centroid defuzzification – one of the main ways to transform the membership function into a single numerical control recommendation.

Since the above idea makes sense, let us use it to gauge the accuracy of the fuzzy control recommendation.

Resulting recommendation. For a given membership function $\mu(x)$, in addition to the result $\bar{x}$ of its centroid defuzzification, we should also generate, as a measure of the accuracy of this recommendation, the value σ which is defined by the following formula

$$\sigma^2 = \int (x-\bar{x})^2 \cdot \rho(x)\,dx = \frac{\int (x-\bar{x})^2 \cdot \mu(x)\,dx}{\int \mu(x)\,dx} =$$

$$\frac{\int x^2 \cdot \mu(x)\,dx}{\int \mu(x)\,dx} - \left(\frac{\int x \cdot \mu(x)\,dx}{\int \mu(x)\,dx}\right)^2. \tag{3}$$

3 But What Should We Do in the Interval-Valued Fuzzy Case?

But what do we do for type-2 fuzzy logic? For the above case of type-1 fuzzy logic, this is just a simple recommendation.

But what do we do if we use a more adequate way to describe uncertainty – namely, type-2 fuzzy logic? In this paper, we consider the simplest case of type-2 fuzzy logic – the interval-valued fuzzy logic (see, e.g., [2, 3]), where for each possible value x of the corresponding quantity, we only know the interval $[\underline{\mu}(x), \overline{\mu}(x)]$ of possible value of degree of confidence $\mu(x)$?

Challenge. In this case, we have a challenge:

- just like to defizzification, we need to find the range of possible values of $\bar{x}$ corresponding to different functions $\mu(x)$ from the given interval [2, 3],
- similarly, we need to find the range of possible values of σ^2 when each value $\mu(x)$ belongs to the corresponding interval.

Analysis of the problem. According to calculus, when the maximum of a function $f(z)$ on the interval $[\underline{z}, \overline{z}]$ is attained at some point $z_0 \in [\underline{z}, \overline{z}]$, then we have one of the three possible cases:

- we can have $z_0 \in (\underline{z}, \overline{z})$, in which case $\dfrac{df}{dz} = 0$ at this point z_0;
- we can have $z_0 = \underline{z}$, in this case, we must have $\dfrac{df}{dz} \le 0$ at this point (otherwise, the function would increase even further when z increases, and so there would no maximum at $\underline{z}$), or
- we can have $z_0 = \overline{z}$, in which case $\dfrac{df}{dz} \ge 0$.

Similarly, when the minimum of a function $f(z)$ on the interval $[\underline{z}, \overline{z}]$ is attained at some point $z_0 \in [\underline{z}, \overline{z}]$, then we have one of the three possible cases:

- we can have $z_0 \in (\underline{z}, \overline{z})$, in which case $\dfrac{df}{dz} = 0$ at this point z_0;
- we can have $z_0 = \underline{z}$, in this case, we must have $\dfrac{df}{dz} \ge 0$ at this point, or
- we can have $z_0 = \overline{z}$, in which case $\dfrac{df}{dz} \le 0$.

Let us apply this general idea to the dependence of the expression (3) on each value $\mu(a)$.

Here, taking into account that for $\int \mu(x)\,dx \approx \sum \mu(x_i) \cdot \Delta x_i$, we get

$$\frac{\partial(\int \mu(x)\,dx)}{\partial(\mu(a))} = \Delta x, \quad \frac{\partial(\int x \cdot \mu(x)\,dx)}{\partial(\mu(a))} = a \cdot \Delta x \text{ and}$$

$$\frac{\partial(\int x^2 \cdot \mu(x)\,dx)}{\partial(\mu(a))} = a^2 \cdot \Delta x.$$

Now, by using the usual rules for differentiating the ratio, for the composition, and for the square, we conclude that:

$$\frac{\partial(\sigma^2)}{\partial(\mu(a))} = \Delta x \cdot S(a),$$

where we denoted

$$S(a) \stackrel{\text{def}}{=} \frac{a^2}{\int \mu(x)\,dx} - \frac{\int x^2 \cdot \mu(x)\,dx}{\left(\int \mu(x)\,dx\right)^2} - 2 \cdot \overline{a} \cdot \left(\frac{x}{\int \mu(x)\,dx} - \frac{\int x \cdot \mu(x)\,dx}{\left(\int \mu(x)\,dx\right)^2} \right). \quad (4)$$

We are only interested in the sign of the derivative, so we can as well consider the sign of the expression $S(a)$ instead of the sign of the desired derivative $\dfrac{\partial(\sigma^2)}{\partial(\mu(a))}$.

Similar, the sign of the expression $S(a)$ is the same as the sign of the expression $s(a) \stackrel{\text{def}}{=} S(a) \cdot \int \mu(y)\,dy$ which has a simpler form

$$s(a) = a^2 - ((\overline{x})^2 + \sigma^2) - 2 \cdot \overline{x} \cdot (a - \overline{x}).$$

If we know the roots $\underline{x} < \overline{x}$ of this quadratic expression, we can conclude that this quadratic expression $s(a)$ is:

- positive when $a < \underline{x}$ and
- negative when $a > \overline{x}$.

Here, the value $a = \overline{x}$ is between $\underline{x}$ and $\overline{x}$, since for this value a, we have

$$s(\overline{x}) = -\sigma^2 < 0.$$

Thus, in accordance with the above fact from calculus:

- when $a < \underline{x}$ or $a > \overline{x}$, then to find the upper bound for σ^2, we must take $\mu(a) = \overline{\mu}(a)$ and to find the lower bound, we must take $\mu(a) = \underline{\mu}(a)$;
- when $\underline{x} < a < \overline{a}$, then, vice versa, we need to take $\mu(a) = \underline{\mu}(a)$ to find the upper bound for σ^2 and we must take $\mu(a) = \overline{\mu}(a)$ to find the lower bound.

This mathematical conclusion makes perfect sense: to get the largest standard deviation, we must concentrate the distribution as much as possible on values outside the mean, and to get the smallest possible standard deviation, we concentrate it as much as possible on values close to the mean.

Thus, we arrive at the following algorithm.

Resulting algorithm. For all possible values $\underline{x} < \overline{x}$, we use the formula (3) to compute the values $\sigma^2(\mu^-)$ and $\sigma^2(\mu^+)$ for the following two functions $\mu^-(x)$ and $\mu^+(x)$:

- $\mu^+(x) = \overline{\mu}(x)$ when $x < \underline{x}$ or $x > \overline{x}$, and $\mu^+(x) = \underline{\mu}(x)$ when $\underline{x} < x < \overline{x}$;
- $\mu^-(x) = \underline{\mu}(x)$ when $x < \underline{x}$ or $x > \overline{x}$, and $\mu^-(x) = \overline{\mu}(x)$ when $\underline{x} < x < \overline{x}$.

Then:

- as the upper bound for σ^2, we take the maximum of the values $\sigma^2(\mu^+)$ corresponding to different pairs $\underline{x} < \overline{x}$, and
- as the lower bound for σ^2, we take the minimum of the values $\sigma^2(\mu^-)$ corresponding to different pairs $\underline{x} < \overline{x}$.

Acknowledgments

This work was supported in part by NSF grant HRD-1242122.

References

1. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
2. J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice-Hall, Upper Saddle River, 2001.
3. J. M. Mendel and D. Wu, *Perceptual Computing: Aiding People in Making Subjective Judgments*, IEEE Press and Wiley, New York, 2010.
4. H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.
5. D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.
6. L. A. Zadeh, “Fuzzy sets”, *Information and Control*, 1965, Vol. 8, pp. 338–353.