5-1-2021

How Accurate Are Fuzzy Control Recommendations: Interval-Valued Case

Juan Carlos Figueroa-Garcia
Universidad Distrital Francisco Jose de Caldas Bogota, filthed@gmail.com

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

How Accurate Are Fuzzy Control Recommendations: Interval-Valued Case

Juan Carlos Figueroa Garcia[1] and Vladik Kreinovich[2*]

[1]Universidad Distrital Francisco Jose de Caldas Bogota, Colombia.

[2]Department of Computer Science, University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA.

filthed@gmail.com, vladik@utep.edu

[*]**Corresponding Author:** Vladik Kreinovich.

ABSTRACT

As a result of applying fuzzy rules, we get a fuzzy set describing possible control values. In automatic control systems, we need to defuzzify this fuzzy set, i.e., to transform it to a single control value. One of the most frequently used defuzzification techniques is centroid defuzzification. From the practical viewpoint, an important question is: how accurate is the resulting control recommendation? The more accurately we need to implement the control, the more expensive the resulting controller. The possibility to gauge the accuracy of the fuzzy control recommendation follows from the fact that, from the mathematical viewpoint, centroid defuzzification is equivalent to transforming the fuzzy set into a probability distribution and computing the mean value of the control. In view of this interpretation, a natural measure of the accuracy of a fuzzy control recommendation is the standard deviation of the corresponding random variable.

Computing this standard deviation is straightforward for the traditional [0,1]-based fuzzy logic, in which all experts' degrees of confidence are represented by numbers from the interval [0,1]. In practice, however, an expert usually cannot describe his/her degree of confidence by a single number, a more appropriate way to describe his/her confidence is by allowing to mark an interval of possible degrees. In this paper, we provide an efficient algorithm for estimating the accuracy of fuzzy control recommendations under such interval-valued fuzzy uncertainty.

Keywords: Fuzzy control, Interval-valued fuzzy, Accuracy of fuzzy control recommendation.

INTRODUCTION

THE MAIN OBJECTIVE OF THIS PAPER: IN BRIEF

One of the important measures of uncertainty for a fuzzy set with a membership function $\mu(x)$ is its variance [1]

$$V = \frac{\int x^2 \cdot \mu(x)\, dx}{\int \mu(x)\, dx} - \left(\frac{\int x \cdot \mu(x)\, dx}{\int \mu(x)\, dx} \right)^2 \qquad (1)$$

In many practical situations, we have interval-valued fuzzy sets, in which case, for each value x, we only know the interval $\left[\underline{\mu}(x), \overline{\mu}(x)\right]$ for the corresponding value $\mu(x)$ of the membership function. So, all membership functions $\mu(x)$ for which $\mu(x) \in \left[\underline{\mu}(x), \overline{\mu}(x)\right]$ for all x are possible.

For different possible membership functions $\mu(x)$, in general, the variance is different. The only thing that can compute is the range $\left[\underline{V}, \overline{V}\right]$ of possible values of the variance. The main objective of this paper is to come up with an efficient algorithm for computing this interval.

BACKGROUND AND IMPORTANCE OF THIS RESEARCH

Need for fuzzy control: In many practical situations, we do not have a good algorithm for controlling a system, but we have human experts who can control the corresponding class of systems: drivers drive cars, pilots control planes, engineers control operations of a petrochemical plant, etc.

In each application area, some experts are more skilled, some are less skilled. In the ideal world, every car should be driven by the most skilled driver, every patient should be treated by the most skilled medical doctor, etc. In real life, however, there are not that many extra class drivers – and much more cars, a few top medical doctors – and millions of patients.

To take care of all the remaining cars, all the remaining patients, it is desirable to incorporate the knowledge of the top experts in an automated control system. To implement such a system, we need to be able to indicate, for each possible situation – i.e., for each possible combination of variables describing the current situation – the exact value(s) of the control that we need to apply. It would be nice if we could extract this value from the experts, but this is rarely possible. For example, in the US, most people can drive cars, but if you ask a person a specific numerical question about driving strategy, this person will not be able to answer. Let us give an example. You are driving on a one-lane freeway with a speed of 100 km/h, the car in front of you is 20 m ahead, and it suddenly brakes to 95 km/h. What do you do? The correct answer that everyone gives is: we need to brake a little bit. But computers do not understand "a little bit", an automated system needs to know for how many milliseconds you press the brake and with what exactly force – this no one can explain.

In many situations, all we get from experts are numerous rules formulated in terms of words from the natural language like "a little bit". It is therefore necessary to translate this imprecise ("fuzzy") knowledge into a precise control strategy. This need was the main motivation for Lotfi Zadeh to come up with fuzzy techniques in general and fuzzy control in particular [2-7].

How fuzzy control works: a very brief description: Based on the original imprecise rules, we estimate, for each possible control value x, to what degree this value is reasonable in a given situation. The corresponding degree is usually described by a number $\mu(x)$ from the interval $[0,1]$, so that 1 means that we are maximally sure that this control value is reasonable, 0 means that we are absolutely sure that it is *not* reasonable, and values between 0 and 1 correspond to intermediate degrees of certainty. Because of the above description, the largest possible value of the degree $\mu(x)$ is always equal to 1.

We know the value $\mu(x)$ for all x, i.e., in other words, we know a *function* that assigns the value $\mu(x)$ to every input x. This function is known as *the membership function*. To implement an automated control, we need to select a single control value X. This selection is known as *defuzzification*. The most widely used defuzzification is the *centroid* defuzzification, in which we take

$$X = \frac{\int x \cdot \mu(x)\, dx}{\int \mu(x)\, dx}.$$ (2)

Need for interval-valued fuzzy techniques: Traditional fuzzy techniques start by experts helping to clarify natural-language words like "a little bit" by assigning a degree of certainty to statement like "0.1 is a little bit", "0.2 is a little bit", etc. The corresponding degrees are then transformed into the final degrees $\mu(x)$.

However, just like an expert cannot describe exactly for how many milliseconds he or she presses the brake, the same expert cannot describe his/her degree of certainty by a single number on the scale from 0 to 1. It is much more realistic to ask the expert to mark a range (= interval) of possible degree values, e.g., from 0.8 to 0.9. Another situation when we get an interval is when we have several top experts and they have somewhat different degrees. In this case, it makes sense to take the interval containing all these values.

If we start with such interval-valued degrees, then for each possible control value x, we also only know the interval $\left[\underline{\mu}(x), \overline{\mu}(x) \right]$ of the possible value of the degree $\mu(x)$. Such a situation is known as *interval-valued membership function*.

This means that, in principle, we can have many different functions $\mu(x)$ as long as for every x, we have $\mu(x) \in \left[\underline{\mu}(x), \overline{\mu}(x) \right]$.

How accurate is the recommended control value: a problem: An important practical question is: how imprecise is it? For example, if the recommended control value is 0.1, it does not necessarily mean that we have to apply this control: if it is 0.1 ± 0.2, then we clearly are not sure whether control is needed in the first place, so it is better not to apply any control – and thus to avoid wasting fuel (and also to avoid unnecessary wiggling).

There are several reasons why control recommendations are not absolutely optimal. In general, the models that we use to design a controller are approximate, and methods that we use to optimize the control are also approximate. For fuzzy control, there is an additional reason why the control recommendations are inaccurate: since we start with the imprecise expert knowledge, of course, the resulting control value X is also imprecise.

One of the proposed measures of this imprecision is the variance (1). It is therefore practically important to be able to estimate this variance.

RESEARCH DONE BY PREVIOUS RESEARCHERS ON SIMILAR PROBLEMS

Why centroid defuzzification and why variance: How can we select a single value based on the membership function? How can we gauge the accuracy of this estimation?

Natural answers to these questions come from the fact – mentioned several times by Zadeh himself – that, from the purely mathematical viewpoint, it is easy to transform each membership function $\mu(x)$ – for which $\max_x \mu(x) = 1$ into a probability density function $f(x)$ for which $\int f(x)\,dx = 1$, and vice versa. Indeed, if we start with a membership function $\mu(x)$, then we can normalize it by diving by its integral and get the probability density function

$$f(x) = \frac{\mu(x)}{\int \mu(y)\,dy}. \tag{3}$$

Vice versa, if we know a probability density function $f(x)$, then we can normalize this function by dividing it by its largest value and get a membership function

$$\mu(x) = \frac{f(x)}{\max_y f(y)}. \tag{4}$$

For a random variable with probability density $f(x)$, there is a natural way to select a single value – namely, we can take the mean

$$X = E[x] \stackrel{def}{=} \int x \cdot f(x)\,dx. \tag{5}$$

Substituting the expression (3) into this formula, we get exactly the centroid value (2).

For a random variable, a natural way to describe possible deviations from the mean is to take the mean squared difference between the variable x and its means, i.e., the *variance*

$$V \stackrel{def}{=} E[(x - E[x])^2]. \tag{6}$$

Then, possible deviations of x from $E[x]$ can be characterized by the *standard deviation*

$$\sigma \stackrel{def}{=} \sqrt{V}. \tag{7}$$

It is known that the variance can be equivalently described as

$$V = E[x^2] - (E[x])^2 = \int x^2 \cdot f(x)\,dx - \left(\int x \cdot f(x)\,dx \right)^2. \tag{8}$$

Substituting, into this formula, the expression (3) for the probability density function corresponding to the given membership function $\mu(x)$, we get exactly the variance (1). Thus, the variance (1) – and the corresponding standard deviation (7) – describe how accurate is the fuzzy control recommendation.

Known efficient algorithms for computing the range of the centroid: For different functions $\mu(x)$ satisfying this condition, the formula (2) leads, in general, to different values X. All we can do is find the set of possible values X, i.e., the interval

$$\left[\underline{X}, \overline{X}\right] = \left\{\frac{\int x \cdot \mu(x)\, dx}{\int \mu(x)\, dx} : \mu(x) \in \left[\underline{\mu}(x), \overline{\mu}(x)\right]\ for\ all\ \ x\right\}. \tag{9}$$

This formula looks very non-constructive – how can we enumerate all possible functions, there are too many of them. The good news is that there exist reasonably efficient algorithms for computing this interval; see, [8-19,4], for the latest developments.

Specifically, to find the smallest possible value $\underline{X}$ of the centroid X, it is sufficient to find the value x_0 for which the value $X(m_{x_0})$ is the smallest, where the function $m_{x_0}(x)$ is defined as follows:

- for $x < x_0$, we have $m_{x_0}(x) = \overline{\mu}(x)$, and
- for $x > x_0$, we have $m_{x_0}(x) = \underline{\mu}(x)$.

This reduces the original difficult-to-optimize problem of minimizing the centroid value to a much simpler problem of optimizing a function of one variable, a problem for which many efficient algorithms exist.

To find the largest possible value $\overline{X}$ of the centroid X, it is sufficient to find the value x_0 for which the value $X(M_{x_0})$ is the largest, where the function $M_{x_0}(x)$ is defined as follows:

- for $x < x_0$, we have $M_{x_0}(x) = \underline{\mu}(x)$, and
- for $x > x_0$, we have $M_{x_0}(x) = \overline{\mu}(x)$.

This also reduces the original difficult-to-optimize problem of maximizing the centroid value to a much simpler problem of optimizing a function of one variable, a problem for which many efficient algorithms exist.

In contrast, computing bounds on the variance remains an open problem: In the interval-valued case, instead of a single value of the variance V (and of the standard deviation σ), we get an interval $\left[\underline{V}, \overline{V}\right]$ of possible values:

$$\left[\underline{V}, \overline{V}\right] = \left\{V(\mu) : \mu(x) \in \left[\underline{\mu}(x), \overline{\mu}(x)\right]\ for\ all\ \ x\right\}, \tag{10}$$

where $V(\mu)$ is determined by the formula (1).

Here, the situation with computations is not so good: there exist some algorithms, but they are not as simple and efficient as for the formula (9), see, e.g., [1]. In this paper, we design an efficient algorithm for computing the range (10).

METHOD

THIS IS THEORETICAL RESEARCH: NO DATA IS USED

First, it is important to mention that this paper is theoretical: we just develop an efficient algorithm. This development is based on a theoretical analysis, not on any specific data. So, when we talk about methods that we used in this research, we mean theoretical methods, not methods for processing and analyzing data.

FIRST METHOD THAT WE USE IN THIS RESEARCH: CALCULUS TECHNIQUES

Why calculus techniques: We are interested in computing the range of possible values of an expression. This means finding its largest and its smallest values. One of the main techniques for finding the largest and smallest values of a function is calculus – finding the largest and smallest values was one of the main motivations for developing calculus in the first place.

Let us recall how calculus can be used to find these largest and smallest values.

Functions of one variable: analysis: Let us start with the simplest case of functions $F(v)$ of one variable. If we are looking for the maximum or the minimum of a function of one variable, then, according to calculus, both the maximum and the minimum are attained at points v at which the derivative $F'(v)$ is equal to 0.

If we are looking for the minimum or the maximum of a function on an interval, then the situation is somewhat more complicated. If this maximum or minimum is attained inside the interval, then at this point, the derivative is still equal to 0. However, the minimum and maximum can also be attained at the endpoints of an interval $\left[\underline{v}, \overline{v} \right]$ – this is the case, e.g., when the function $F(v)$ is linear.

In this case, for minimum, we have the following conclusions:
- if the minimum is attained at the lower endpoint $\underline{v}$, this means that at this point, the function cannot be decreasing – otherwise, the values for nearby points $\underline{v} + \varepsilon$ will be even smaller; thus, we must have $F'(\underline{v}) \geq 0$;
- similarly, if the minimum is attained at the upper endpoint $\overline{v}$, this means that at this point, the function cannot be increasing – otherwise, the values for the nearby points $\overline{v} - \varepsilon$ will be even smaller; thus, we must have $F'(\overline{v}) \leq 0$.

For maximum, we have similar conclusions:
- if the maximum is attained at the lower endpoint $\underline{v}$, this means that at this point, the function cannot be increasing – otherwise, the values for nearby points $\underline{v} + \varepsilon$ will be even larger; thus, we must have $F'(\underline{v}) \leq 0$;
- similarly, if the maximum is attained at the upper endpoint $\overline{v}$, this means that at this point, the function cannot be decreasing – otherwise, the values for the nearby points $\overline{v} - \varepsilon$ will be even larger; thus, we must have $F'(\overline{v}) \geq 0$.

Functions of one variable: conclusions: For a function of one variable, at the point m where a function $F(v)$ attains its minimum on a given interval $\left[\underline{v}, \overline{v} \right]$:

- either this minimum is attained inside the interval, and $F'(m) = 0$;
- or this minimum is attained at the left endpoint $m = \underline{v}$ of this interval, and $F'(m) \geq 0$;
- or this minimum is attained at the right endpoint $m = \overline{v}$ of this interval, and $F'(m) \leq 0$.

Similarly, at the point M where a function $F(v)$ attains its maximum on a given interval $\left[\underline{v}, \overline{v}\right]$:

- either this maximum is attained inside the interval, and $F'(M) = 0$;
- or this maximum is attained at the left endpoint $M = \underline{v}$ of this interval, and $F'(M) \leq 0$;
- or this maximum is attained at the right endpoint $M = \overline{v}$ of this interval, and $F'(M) \geq 0$.

General case: functions of several variables: When a function of several variables $F(v_1, \ldots, v_n)$ attains its minimum on a given box

$$\left[\underline{v}_1, \overline{v}_1\right] \times \ldots \times \left[\underline{v}_n, \overline{v}_n\right],$$

at some point $m = (m_1, \ldots, m_n)$, then clearly for each i, this means that the value $F(m_1, \ldots, m_n)$ is the smallest possible value of the given function. This implies that the function $F_i(v_i) \overset{def}{=} F(m_1, \ldots, m_{i-1}, v_i, m_{i+1}, \ldots, m_n)$ of one variable attains its minimum for $v_i = m_i$. For this auxiliary function of one variable, the derivative is nothing else but the i-th partial derivative $\frac{\partial F}{\partial v_i}$ of the original function $F(v_1, \ldots, v_n)$. Thus, we can apply the previously described calculus-based conclusions and imply that for each i:

- either the value m_i is attained inside the corresponding interval $\left[\underline{v}_i, \overline{v}_i\right]$, and $\frac{\partial F}{\partial v_i} = 0$;
- or the value m_i is located at the left endpoint $m_i = \underline{v}_i$ of this interval, and $\frac{\partial F}{\partial v_i} \geq 0$;
- or the value m_i is located at the right endpoint $m_i = \overline{v}_i$ of this interval, and $\frac{\partial F}{\partial v_i} \leq 0$.

Similarly, when a function of several variables $F(v_1, \ldots, v_n)$ attains its minimum on a given box at some point $M = (M_1, \ldots, M_n)$, then for each i:

- either the value M_i is attained inside the corresponding interval $\left[\underline{v}_i, \overline{v}_i\right]$, and $\frac{\partial F}{\partial v_i} = 0$;
- or the value M_i is located at the left endpoint $M_i = \underline{v}_i$ of this interval, and $\frac{\partial F}{\partial v_i} \leq 0$;
- or the value M_i is located at the right endpoint $M_i = \overline{v}_i$ of this interval, and $\frac{\partial F}{\partial v_i} \geq 0$.

SECOND METHOD THAT WE USE IN THIS RESEARCH: DISCRETIZATION

The centroid and variance – as described by the equations (1) and (2) – are, in effect, functions of infinitely many variables $\mu(x)$ – the values of the membership function at different values x. However, from the mathematical viewpoint, each integral

is the limit of the corresponding integral sums, and the limit means that if we consider a sufficiently dense grid, we get the value of this integral with any given accuracy. This is how integrals are computed in a computer in the first place.

From this viewpoint, it makes sense to consider the corresponding integral sums in the formulas (1) and (2):

$$X = \frac{\sum_{i=1}^{n} x_i \cdot \mu_i \cdot \Delta x}{\sum_{i=1}^{n} \mu_i \cdot \Delta x} \qquad (11)$$

and

$$V = \frac{\sum_{i=1}^{n} x_i^2 \cdot \mu_i \cdot \Delta x}{\sum_{i=1}^{n} \mu_i \cdot \Delta x} - \left(\frac{\sum_{i=1}^{n} x_i \cdot \mu_i \cdot \Delta x}{\sum_{i=1}^{n} \mu_i \cdot \Delta x} \right)^2. \qquad (12)$$

where we consider grid points $x_i = x_1 + (i - 1) \cdot \Delta x$, and μ_i denotes $\mu(x_i)$.

LET US APPLY THESE METHODS TO OUR PROBLEM

Let us start with the problem of computing the range of the centroid: In accordance with the discretization method, the value of the centroid can be described by the formula (11). If we divide both numerator and denominator of this formula by Δx, we will get a simplified expression

$$X = \frac{\sum_{i=1}^{n} x_i \cdot \mu_i}{\sum_{i=1}^{n} \mu_i}. \qquad (13)$$

We want to find the minimum and maximum of this function when for each i, we have $\mu_i \in \left[\underline{\mu}_i, \overline{\mu}_i \right]$, where we similarly denoted $\underline{\mu}_i \overset{\text{def}}{=} \underline{\mu}(x_i)$ and $\overline{\mu}_i \overset{\text{def}}{=} \overline{\mu}(x_i)$. As we have mentioned earlier, to find the minimum and maximum, it is beneficial to know the partial derivatives of the function (13) with respect to its variables μ_i. Here,

$$\frac{\partial}{\partial \mu_i} \left(\frac{\sum_{j=1}^{n} x_j \cdot \mu_j}{\sum_{j=1}^{n} \mu_j} \right) = \frac{x_i \cdot \left(\sum_{j=1}^{n} \mu_j \right) - \sum_{j=1}^{n} x_j \cdot \mu_j}{\left(\sum_{j=1}^{n} \mu_j \right)^2}. \qquad (14)$$

This expression can be described as

$$\frac{\partial X}{\partial \mu_i} = a \cdot x_i - b, \qquad (15)$$

where we denoted

$$a \overset{\text{def}}{=} \frac{\sum_{j=1}^{n} \mu_j}{\left(\sum_{j=1}^{n} \mu_j \right)^2} = \frac{1}{\sum_{j=1}^{n} \mu_j} \qquad (16)$$

and

$$b \overset{\text{def}}{=} \frac{\sum_{j=1}^{n} x_j \cdot \mu_j}{\left(\sum_{j=1}^{n} \mu_j\right)^2}. \tag{17}$$

The values μ_j are non-negative and some of them are positive, so $a > 0$.

Thus, for the values $m_1, \ldots, m_n$ for which the expression (13) attains its minimum, for each i:

• either the value m_i is attained inside the corresponding interval $\left[\underline{\mu}_i, \overline{\mu}_i\right]$, and $a \cdot x_i - b = 0$;

• or the value m_i is located at the left endpoint $m_i = \underline{\mu}_i$ of this interval, and $a \cdot x_i - b \geq 0$;

• or the value m_i is located at the right endpoint $m_i = \overline{\mu}_i$ of this interval, and $a \cdot x_i - b \leq 0$.

The equality $a \cdot x_i - b = 0$ can only be satisfied for one value $x_i = b/a$. For all larger values x_i, we will have $a \cdot x_i - b > 0$. Thus, in this case, the minimum cannot be attained inside the interval, and it cannot be attained at the right endpoint – so it must be attained at the left endpoint $m_i = \underline{\mu}_i$.

Similarly, for values x_i which are smaller than b/a, we have $a \cdot x_i - b < 0$. Thus, in this case, the minimum cannot be attained inside the interval, and it cannot be attained at the left endpoint – so it must be attained at the right endpoint $m_i = \overline{\mu}_i$.

We do not know the threshold value $x_0 \overset{\text{def}}{=} b/a$, but we can conclude that for the function $\mu(x)$ for which the centroid attains its minimum, we have $\mu(x) = \overline{\mu}(x)$ for all $x < x_0$ and $\mu(x) = \underline{\mu}(x)$ for all $x > x_0$. Thus, we arrive at the known algorithm – that we described earlier.

For the values $M_1, \ldots, M_n$ for which the expression (13) attains its maximum, for each i:

• either the value M_i is attained inside the corresponding interval $\left[\underline{\mu}_i, \overline{\mu}_i\right]$, and $a \cdot x_i - b = 0$;

• or the value M_i is located at the left endpoint $M_i = \underline{\mu}_i$ of this interval, and $a \cdot x_i - b \leq 0$;

• or the value M_i is located at the right endpoint $M_i = \overline{\mu}_i$ of this interval, and $a \cdot x_i - b \geq 0$.

The equality $a \cdot x_i - b = 0$ can only be satisfied for one value $x_i = b/a$. For all larger values x_i, we will have $a \cdot x_i - b > 0$. Thus, in this case, the minimum cannot be attained inside the interval, and it cannot be attained at the left endpoint – so it must be attained at the right endpoint $M_i = \overline{\mu}_i$.

Similarly, for values x_i which are smaller than b/a, we have $a \cdot x_i - b < 0$. Thus, in this case, the minimum cannot be attained inside the interval, and it cannot be attained at the right endpoint – so it must be attained at the left endpoint $M_i = \underline{\mu}_i$.

We do not know the threshold value $x_0 \overset{\text{def}}{=} b/a$, but we can conclude that for the function $\mu(x)$ for which the centroid attains its minimum, we have $\mu(x) = \underline{\mu}(x)$ for all $x < x_0$ and $\mu(x) = \overline{\mu}(x)$ for all $x > x_0$. Thus, for computing $\overline{X}$, we also arrive at the above algorithm.

Comment. Instead of approximating the integrals by integral sums, we could use *variational calculus* – an extension of calculus specifically intended for the situations when we want to find a function that minimizes or maximizes a given objective function; [20-23].

Let us apply these methods to computing the variance: Let us now apply these methods to computing the variance (1), which, in terms of integral sums, takes the form (12). If we divide both the numerator and denominator of both fractions in this formula by Δx, we get a simplified expression

$$V = \frac{\sum_{i=1}^{n} x_i^2 \cdot \mu_i}{\sum_{i=1}^{n} \mu_i} - \left(\frac{\sum_{i=1}^{n} x_i \cdot \mu_i}{\sum_{i=1}^{n} \mu_i} \right)^2. \tag{18}$$

Here,

$$\frac{\partial V}{\partial \mu_i} = \frac{x_i^2 \cdot \left(\sum_{j=1}^{n} \mu_j \right) - \sum_{j=1}^{n} x_j^2 \cdot \mu_j}{\left(\sum_{j=1}^{n} \mu_j \right)^2} - 2 \cdot \left(\frac{\sum_{j=1}^{n} x_j \cdot \mu_j}{\sum_{j=1}^{n} \mu_j} \right) \cdot \frac{x_i \cdot \left(\sum_{j=1}^{n} \mu_j \right) - \sum_{j=1}^{n} x_j \cdot \mu_j}{\left(\sum_{j=1}^{n} \mu_i \right)^2}. \tag{19}$$

Thus, the dependence of this derivative on x_i has the form

$$\frac{\partial V}{\partial \mu_i} = a \cdot x_i^2 + b \cdot x_i + c, \tag{20}$$

where

$$a \overset{\text{def}}{=} \frac{\sum_{j=1}^{n} \mu_j}{\left(\sum_{j=1}^{n} \mu_j \right)^2} = \frac{1}{\sum_{j=1}^{n} \mu_j}. \tag{21}$$

Similar to the case of the centroid, this value a is positive. So, the quadratic expression (20) is either always positive, or has two roots $r_1 \le r_2$, so that this expression is positive for $x < r_1$ and $x > r_2$ and negative for $r_1 < x < r_2$. The first case can be described as a particular case of the second case if we take $r_1 = r_2$.

Thus, for the values $m_1, \dots, m_n$ for which the expression (18) attains its minimum, for each i:
- either the value m_i is attained inside the corresponding interval $\left[\underline{\mu}_i, \overline{\mu}_i \right]$, and $a \cdot x_i^2 + b \cdot x_i + c = 0$;
- or the value m_i is located at the left endpoint $m_i = \underline{\mu}_i$ of this interval, and $a \cdot x_i^2 + b \cdot x_i + c \ge 0$;
- or the value m_i is located at the right endpoint $m_i = \overline{\mu}_i$ of this interval, and $a \cdot x_i^2 + b \cdot x_i + c \le 0$.

The equality $a \cdot x_i^2 + b \cdot x_i + c = 0$ can only be satisfied for at most two values r_1 and r_2. For values $x_i < r_1$ and $x_i > r_2$, we will have $a \cdot x_i^2 + b \cdot x_i + c > 0$. Thus, in this case, the minimum cannot be attained inside the interval, and it cannot be attained at the right endpoint – so it must be attained at the left endpoint $m_i = \underline{\mu}_i$.

Similarly, for values x_i which are between r_1 and r_2, we have $a \cdot x_i^2 + b \cdot x_i + c < 0$. Thus, in this case, the minimum cannot be attained inside the interval, and it cannot be attained at the left endpoint – so it must be attained at the right endpoint $m_i = \overline{\mu}_i$.

We do not know the threshold values r_1 and r_2, but we can conclude that for the function $\mu(x)$ for which the variance attains its minimum, we have $\mu(x) = \underline{\mu}(x)$ for all $x < r_1$ and $x > r_2$, and we have $\mu(x) = \overline{\mu}(x)$ for all x between r_1 and r_2.

Comment. This makes sense: to minimize the variance, we must assign the smallest possible weight to values far away from the centroid, and the largest possible weight to the values close to the centroid.

Similarly, for the values $M_1, \ldots, M_n$ for which the expression (18) attains its maximum, for each i:
- either the value M_i is attained inside the corresponding interval $\left[\underline{\mu}_i, \overline{\mu}_i \right]$, and $a \cdot x_i^2 + b \cdot x_i + c = 0$;
- or the value M_i is located at the left endpoint $m_i = \underline{\mu}_i$ of this interval, and $a \cdot x_i^2 + b \cdot x_i + c \leq 0$;
- or the value M_i is located at the right endpoint $m_i = \overline{\mu}_i$ of this interval, and $a \cdot x_i^2 + b \cdot x_i + c \geq 0$.

The equality $a \cdot x_i^2 + b \cdot x_i + c = 0$ can only be satisfied for at most two values r_1 and r_2. For values $x_i < r_1$ and $x_i > r_2$, we will have $a \cdot x_i^2 + b \cdot x_i + c > 0$. Thus, in this case, the maximum cannot be attained inside the interval, and it cannot be attained at the left endpoint – so it must be attained at the right endpoint $M_i = \overline{\mu}_i$.

Similarly, for values x_i which are between r_1 and r_2, we have $a \cdot x_i^2 + b \cdot x_i + c < 0$. Thus, in this case, the maximum cannot be attained inside the interval, and it cannot be attained at the right endpoint – so it must be attained at the left endpoint $m_i = \underline{\mu}_i$.

We do not know the threshold values r_1 and r_2, but we can conclude that for the function $\mu(x)$ for which the variance attains its minimum, we have $\mu(x) = \overline{\mu}(x)$ for all $x < r_1$ and $x > r_2$, and we have $\mu(x) = \underline{\mu}(x)$ for all x between r_1 and r_2.

Comment. This also makes sense: to maximize the variance, we must assign the largest possible weight to values far away from the centroid, and the smallest possible weight to the values close to the centroid.

Thus, we arrive at the following algorithms.

RESULTS

Algorithm for computing $\underline{V}$. : To find the smallest possible value of the variance V, it is sufficient to find the value $r_1 \leq r_2$ for which the value $V(m_{r_1,r_2})$ is the smallest, where the function $m_{r_1,r_2}(x)$ is defined as follows:

- for $x < r_1$ and for $x > r_2$, we have $m_{r_1,r_2}(x) = \underline{\mu}(x)$, and
- for $x \in (r_1, r_2)$, we have $m_{r_1,r_2}(x) = \overline{\mu}(x)$.

This reduces the original difficult-to-optimize problem of minimizing the centroid value to a much simpler problem of optimizing a function of two variables, a problem for which many efficient algorithms exist.

Algorithm for computing $\overline{V}$. :To find the largest possible value of the variance V, it is sufficient to find the value $r_1 \leq r_2$ for which the value $V(M_{r_1,r_2})$ is the largest, where the function $M_{r_1,r_2}(x)$ is defined as follows:

- for $x < r_1$ and for $x > r_2$, we have $M_{r_1,r_2}(x) = \overline{\mu}(x)$, and
- for $x \in (r_1, r_2)$, we have $M_{r_1,r_2}(x) = \underline{\mu}(x)$.

This reduces the original difficult-to-optimize problem of maximizing the centroid value to a much simpler problem of optimizing a function of two variables, a problem for which many efficient algorithms exist.

DISCUSSION

So, we indeed designed efficient algorithms for computing the range $\left[\underline{V}, \overline{V}\right]$ of the variance for interval-valued fuzzy sets.

Comment. The range of the possible values of a random variable is often described by an interval $[X - k \cdot \sigma, X + k \cdot \sigma]$, for some value k. One can see that in the fuzzy case, the derivatives of both expressions $X - k \cdot \sigma$ and $X + k \cdot \sigma$ with respect to μ_I are also quadratic in x_i, and so, similar efficient algorithms can be used to compute the ranges of possible values of each of these expressions.

CONCLUSIONS

To gauge the accuracy of the fuzzy control recommendations, it is useful to know the variance V of the corresponding fuzzy set. In the practically important interval-valued fuzzy case, instead of a single value of the variance, we have a range $\left[\underline{V}, \overline{V}\right]$ of its possible values. In this paper, we provide an efficient algorithm for computing this range.

ACKNOWLEDGEMENT

REFERENCES

1. Wu D, Mendel J. Uncertainty Measures For Interval Type-2 Fuzzy Sets. Information Sciences. 2007;177:5378–5393.

2. Belohlavek R, Dauben JW, Klir GJ. Fuzzy Logic and Mathematics: A Historical Perspective, Oxford University Press, New York. 2017.

3. Klir G and Yuan B. Fuzzy Sets and Fuzzy Logic, Prentice Hall, Upper Saddle River, New Jersey. 1995.

4. Mendel JM. Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions, Springer, Cham, Switzerland. 2017.

5. Nguyen HT, Walker CL, Walker EA. A First Course in Fuzzy Logic, Chapman and Hall/CRC, Boca Raton, Florida. 2019.

6. Novák V, Perfilieva I, Močkoř J. Mathematical Principles of Fuzzy Logic, Kluwer, Boston, Dordrecht, 1999.

7. Zadeh LA. Fuzzy sets, Information and Control. 1965;8: 338–353.

8. Kreinovich V, Lakeyev A, Rohn J, Kahl P. Computational Complexity and Feasibility of Data Processing and Interval Computations, Kluwer, Dordrecht. 1998.

9. Lea R, Kreinovich V, Trejo R. Optimal interval enclosures for fractionally-linear functions, and their application to intelligent control. Reliable Computing. 1996;2:265–286.

10. Bernal E, Lagunes MI, Castillo O, Soria J, Valdez F. Optimization of Type-2 Fuzzy Logic Controller Design Using the CSO And FA Algorithms. International Journal of Fuzzy Systems. 2020; 23:52–57.

11. Carvajal O, Melin P, Moramontes I, Prado-Arechiga G, Optimal design of a general type-2 fuzzy classifier for the pulse level and its hardware implementation, Engineering Applications of Artificial Intelligence. 2021;97:104069.

12. Castillo O, Melin P, Ontiveros E, Peraza C, Ochoa P, et. Al. A High-Speed Interval Type 2 Fuzzy System Approach for Dynamic Parameter Adaptation In Metaheuristics. Engineering Applications of Artificial Intelligence. 2019;85: 666–680.

13. Castillo O, Valdez F, Peraza C, Yoon JH, Geem ZW. High-speed interval type-2 fuzzy systems for dynamic parameter adaptation in harmony search for optimal design of fuzzy controllers. Mathematics. 2021;9:758.

14. Chen C, John R, Twycross J, Garibaldi J. A direct approach for determining the switch points in the Karnik-Mendel algorithm. IEEE Transactions on Fuzzy Systems. 2018;26:1079-1085.

15. Liu C, Lin Y. New Efficient Algorithms for The Centroid of An Interval Type-2 Fuzzy Set. Information Sciences. 2021;570:468–486.

16. Moreno JE, Sanchez MA, Mendoza O, Rodriguez-Diaz A, Castillo O. Design of An Interval Type-2 Fuzzy Model With Justifiable Uncertainty. Information Sciences. 2020;513:206–221.

17. Ontiveros E, Melin P, Castillo O. High Order A-Planes Integration: A New Approach to Computational Cost Reduction of General Type-2 Fuzzy Systems, Engineering Applications of Artificial Intelligence. 2018;74:186-197.

18. Ontiveros E, Melin P, Castillo O. Comparative Study of Interval Type-2 and General Type-2 Fuzzy Systems in medical diagnosis. Information Sciences. 2020; 525: 37-53.

19. Valdez F. A Review of Optimization Swarm Intelligence-Inspired Algorithms With Type-2 Fuzzy Logic Parameter Adaptation. Soft Computing. 2020; 24:215-226.

20. Feynman R, Leighton R, Sands M. Feynman Lectures on Physics, Basic Books, New York. 2005.

21. Gelfand M, Fomin SV. Calculus of Variations, Dover, New York. 2000.

22.   Kosheleva O and Kreinovich V. Finding the best function: a way to explain calculus of variations to engineering and science students, Applied Mathematical Sciences. 2013;7:7187-7192.

23.   Ross IM. A Primer on Pontryagin Principle in Optimal Control, Collegiate Publishers, San Francisco, California. 2009.