How Accurate Are Fuzzy Control Recommendations: Interval-Valued Case

Juan Carlos Figueroa Garcia[1] and Vladik Kreinovich[2]
[1]Universidad Distrital, Departamento de Ingenieria Industrial
Bogota, Colombia, filthed@gmail.com
[2]Department of Computer Science, University of Texas at El Paso
500 W. University, El Paso, TX 79968, USA, vladik@utep.edu

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 control. In view of this interpretation, a natural measure of 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' degree 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.

1 Formulation of the Problem

Need for fuzzy control. In many practical situations, we have human experts who can control a certain 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 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; see, e.g., [1, 4, 7, 8, 9, 11].

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

$$X = \frac{\int x \cdot \mu(x)\, dx}{\int \mu(x)\, dx}. \tag{1}$$

How accurate is the recommended control value: a problem. Since we started with the imprecise expert knowledge, of course, the resulting control value X is imprecise. An important question is: how imprecise is it? For example, if the recommended control value is 0.1, it does not necessarily means that we have to apply this control: if it is 0.1 ± 0.2, then we clearly are not

sure whether a 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).

How accurate is the recommended control value: how to estimate this accuracy. The possibility to estimate the accuracy comes 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 $\max\limits_{x} \mu(x) = 1$ 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

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

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

$$\mu(x) = \frac{f(x)}{\max\limits_{y} f(y)}. \tag{3}$$

In these terms, as one can easily check, the centroid value X for a given membership function $\mu(x)$ is nothing else but the mean value $E[x]$ of the random variable described by the corresponding probability density function $f(x)$:

$$X = E[x] \stackrel{\text{def}}{=} \int x \cdot f(x)\,dx. \tag{4}$$

In probability theory, 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*

$$V \stackrel{\text{def}}{=} E\left[(x - E[x])^2\right]. \tag{5}$$

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

$$\sigma \stackrel{\text{def}}{=} \sqrt{V}. \tag{6}$$

It is known that the variance can be equivalently described as

$$V = E\left[x^2\right] - (E[x])^2 = \int x^2 \cdot f(x)\,dx - \left(\int x \cdot f(x)\,dx\right)^2. \tag{7}$$

Substituting, into this formula, the expression (2) for the probability density function corresponding to the given membership function $\mu(x)$, we conclude that

$$V = \frac{\int x^2 \cdot \mu(x)\,dx}{\int \mu(x)\,dx} - \left(\frac{\int x \cdot \mu(x)\,dx}{\int \mu(x)\,dx}\right)^2. \tag{8}$$

This value (8) – and the corresponding standard deviation (6) – describe how accurate is the fuzzy control recommendation.

Interval-valued case. So far so good, but there is a subtlety that we skipped in the above analysis. Namely, 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 then 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 $\big[\underline{\mu}(x), \overline{\mu}(x)\big]$ of possible value of the degree $\mu(x)$. Such 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 $\mu(x) \in \big[\underline{\mu}(x), \overline{\mu}(x)\big]$. For different functions $\mu(x)$ satisfying this condition, the formula (1) leads, in general, to different values X. All we can do is find the set of possible values X, i.e., the interval

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

This formula looks very non-constructive – how can we enumerate all possible functions, there are too many of them. Good news is that there exist reasonably efficient algorithms for computing this interval; see, e.g., [6, 7].

Remaining problem. Similarly, instead of a single value of V (and of σ), in the interval-valued case, we get an interval $\big[\underline{V}, \overline{V}\,\big]$ of possible values:

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

where $V(\mu)$ is determined by the formula (8).

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).

What we do in this paper. The main objective of this paper is to come up with the algorithm for computing the range (10) which is as straightforward – and (almost) as efficient as the known algorithms for computing the range (9).

2 Analysis of the Problem

Plan. To come up with an efficient algorithm for computing the range (10) of the variance, let us recall the main ideas behind the current efficient algorithms

for computing the range (9) of the centroid. Then, we will see how these ideas can be used to find the range (10).

Main calculus-based ideas used in estimating the range of the centroid. Computing the range of possible values of an expression 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 largest and smallest values was one of the main motivations for developing calculus in the first place.

Let us recall the use of calculus of optimization on the example 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 $[\underline{v}, \overline{v}]$ – 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 $F'(\underline{v}) \leq 0$.

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 $F'(\underline{v}) \geq 0$.

Calculus ideas: conclusion. For a function of one variable, at the point m where a function $F(v)$ attains its minimum on a given interval $[\underline{v}, \overline{v}]$:

- 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 $F'(v) \geq 0$;

- or this minimum is attained at the right endpoint $m = \overline{v}$ of this interval, and $F'(v) \le 0$.

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

- 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 $F'(M) \le 0$;
- or this maximum is attained at the right endpoint $M = \overline{v}$ of this interval, and $F'(M) \ge 0$.

Functions of several variables. When a function of several variables $F(v_1, \ldots, v_n)$ attains its minimum on a given box

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

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 $F_i(v_i) \stackrel{\text{def}}{=} F(m_1, \ldots, m_{i-1}, v_i, m_{i+1}, \ldots, m_n)$ 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 $\dfrac{\partial F}{\partial v_i}$ of the original function $F(v_1, \ldots, v_n)$. Thus, we can apply the previously described calculus-based conclusion and imply that for each i:

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

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 $[\underline{v}_i, \overline{v}_i]$, and $\dfrac{\partial F}{\partial v_i} = 0$;
- or the value M_i is located at the left endpoint $M_i = \underline{v}_i$ of this interval, and $\dfrac{\partial F}{\partial v_i} \le 0$;

- or the value M_i is located at the right endpoint $M_i = \overline{v}_i$ of this interval, and $\dfrac{\partial F}{\partial v_i} \geq 0$.

Let us recall how this idea can help compute the range of the centroid. The centroid – as described by the equation (1) – is, in effect a function 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 formula (1):

$$X = \frac{\sum\limits_{i=1}^{n} x_i \cdot \mu_i \cdot \Delta x}{\sum\limits_{i=1}^{n} \mu_i \cdot \Delta x}, \tag{11}$$

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

If we divide both numerator and denominator by Δx, we get a simplified expression

$$X = \frac{\sum\limits_{i=1}^{n} x_i \cdot \mu_i}{\sum\limits_{i=1}^{n} \mu_i}. \tag{12}$$

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 $\underline{\mu}_i \stackrel{\text{def}}{=} \underline{\mu}(x_i)$ and $\overline{\mu}_i \stackrel{\text{def}}{=} \overline{\mu}(x_i)$. As we have mentioned earlier, to find the minimum and maximum, it is beneficial to know the partial derivatives of the function (12) with respect to its variables μ_i. Here,

$$\frac{\partial}{\partial \mu_i}\left(\frac{\sum\limits_{i=1}^{n} x_i \cdot \mu_i}{\sum\limits_{i=1}^{n} \mu_i}\right) = \frac{x_i \cdot \left(\sum\limits_{j=1}^{n} \mu_j\right) - \sum\limits_{j=1}^{n} x_j \cdot \mu_j}{\left(\sum\limits_{j=1}^{n} \mu_j\right)^2}. \tag{13}$$

This expression can be described as

$$\frac{\partial X}{\partial \mu_i} = a \cdot x_i - b, \tag{14}$$

where we denoted

$$a \stackrel{\text{def}}{=} \frac{\sum\limits_{j=1}^{n} \mu_j}{\left(\sum\limits_{j=1}^{n} \mu_j\right)^2} = \frac{1}{\sum\limits_{j=1}^{n} \mu_j} \tag{15}$$

and

$$b \stackrel{\text{def}}{=} \frac{\sum\limits_{j=1}^{n} x_j \cdot \mu_j}{\left(\sum\limits_{j=1}^{n} \mu_j\right)^2}. \tag{16}$$

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 (12) 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 $x_0 \stackrel{\text{def}}{=} b/a$, 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 following algorithm.

Algorithm for computing $\underline{X}$. To find the smallest possible value 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.

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; see, e.g., [2, 3, 5, 10].

How to compute $\overline{X}$: idea. Similarly, for the values $M_1, \ldots, M_n$ for which the expression (12) 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 $a \cdot x_i - b \le 0$;
- or the value M_i is located at the right endpoint $M_i = \overline{\mu}_i$ of this interval, and $a \cdot x_i - b \ge 0$.

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 $x_0 \stackrel{\text{def}}{=} b/a$, 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, we arrive at the following algorithm.

Algorithm for computing $\overline{X}$. To find the smallest possible value 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 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.

Let us apply this idea to the variance. Let us now apply this idea to computing the variance (8), which, in terms of integral sums, takes the form

$$V = \frac{\sum\limits_{i=1}^{n} x_i^2 \cdot \mu_i \cdot \Delta x}{\sum\limits_{i=1}^{n} \mu_i \cdot \Delta x} - \left(\frac{\sum\limits_{i=1}^{n} x_i \cdot \mu_i \cdot \Delta x}{\sum\limits_{i=1}^{n} \mu_i \cdot \Delta x} \right)^2. \tag{17}$$

If we divide both numerator and denominator of both fractions by Δx, we get a simplified expression

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

Here,

$$\frac{\partial V}{\partial \mu_i} = \frac{x_i^2 \cdot \left(\sum\limits_{j=1}^{n} \mu_j\right) - \sum\limits_{j=1}^{n} x_j^2 \cdot \mu_j}{\left(\sum\limits_{j=1}^{n} \mu_j\right)^2} -$$

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

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

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

Similarly 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 $m_1, \ldots, m_n$ 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 $a \cdot x_i^2 + b \cdot m_i + c = 0$;
- or the value M_i is located at the left endpoint $m_i = \underline{\mu}_i$ of this interval, and $a \cdot x_i^2 + b \cdot_i + c \le 0$;
- or the value M_i is located at the right endpoint $m_i = \overline{\mu}_i$ of this interval, and $a \cdot x_i^2 + b \cdot x_i + c \ge 0$.

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:

3 Resulting Algorithms for Computing the Range $\left[\underline{V}, \overline{V}\right]$ of the Desired Accuracy Measure

Algorithm for computing $\underline{V}$. To find the smallest possible value of the variance V, it is sufficient to find the value $r_1 \le r_2$ 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 $r_1 \le r_2$ 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.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), and HRD-1834620 and HRD-2034030 (CAHSI Includes).

It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478.

References

[1] R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.

[2] R. Feynman, R. Leighton, and M. Sands, *Feynman Lectures on Physics*, Basic Books, New York, 2005.

[3] I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Dover, New York, 2000.

[4] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[5] O. Kosheleva and V. Kreinovich, "Finding the best function: a way to explain calculus of variations to engineering and science students", *Applied Mathematical Sciences*, 2013, Vol. 7, No. 144, pp. 7187–7192.

[6] R. Lea, V. Kreinovich, and R. Trejo, "Optimal interval enclosures for fractionally-linear functions, and their application to intelligent control", *Reliable Computing*, 1996, Vol. 2, No. 3, pp. 265–286.

[7] J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.

[8] H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

[9] V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.

[10] I. M. Ross, *A Primer on Pontryagin Principle in Optimal Control*, Collegiate Publishers, San Francisco, California, 2009.

[11] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.