Between Dog and Wolf: A Continuous Transition from Fuzzy to Probabilistic Estimates

Martine Ceberio, Olga Kosheleva,
Luc Longpré, and Vladik Kreinovich
University of Texas at El Paso
El Paso TX 79968, USA
mceberio@utep.edu, olgak@utep.edu
longpre@utep.edu, vladik@utep.edu

1. Computations Based on Expert Estimates: A Typical Situation

- In many practical situations, we have expert estimates $\widetilde{x}_1, \ldots, \widetilde{x}_n$ of several quantities $x_1, \ldots, x_n$.
- Based on $\widetilde{x}_i$, we estimate the values of other quantities y that depend on x_i in a known way: $y = f(x_1, \ldots, x_n)$.
- Namely, as the desired estimate for y, it is natural to take the value

$$\widetilde{y} = f\left(\widetilde{x}_1, \ldots, \widetilde{x}_n\right).$$

- For example, if we estimate the distance x_1 and time x_2, we can estimate the speed as $y = \dfrac{\widetilde{x}_1}{\widetilde{x}_2}$.

2. In Many Situations, Accuracy Estimation Is Important

- In many practical situations, it is important to know the accuracy of the resulting estimate $\widetilde{y}$.
- In economics, we predict the nearest-future change in stock prices.
- Using an inaccurate estimate can lead to huge money losses.
- In geophysics, we estimate the amount of oil $\widetilde{y}$ in a given area.
- If this estimate is reasonably accurate, then it makes sense to invest in this oil field.
- However, if the estimate $\widetilde{y}$ is not very accurate, it is better to perform additional measurements.
- In medicine, we estimate the patient's health.
- By prescribing a wrong treatment, we can make the disease worse or even lose the patient.

3. Resulting Computational Problem

- To estimate the accuracy of $\widetilde{y}$, we need to know how accurate are $\widetilde{x}_1, \ldots, \widetilde{x}_n$.
- Usually, for each of these estimates $\widetilde{x}_i$, we know a number Δ_i that describes its accuracy.
- $\Delta x_i \stackrel{\text{def}}{=} \widetilde{x}_i - x_i$ is approximately of the same order as Δ_i.
- Based on the values Δ_i, we want to estimate the accuracy Δ of $\widetilde{y}$.

4. How This Problem Is Solved Now: General Idea

- There are many techniques for solving the above problem.
- These techniques depend on how exactly the value Δ_i relates to the approximation error.
- This number Δ_i can be the upper bound on the possible values of the approximation error.
- This is the case of *interval uncertainty*.
- The number Δ_i can be the mean squared value of the approximation error, or the most probable value of this error.
- These are the two cases of *probabilistic uncertainty*.
- The number Δ_i can simply be an expert's estimate for the approximation error.
- This is the case of *fuzzy uncertainty*.

5. Remaining Challenges

- At first glance, there are reasonable approaches for estimating accuracy.
- For example:
  - we can use simply probabilistic ideas, or
  - we can use simple fuzzy ideas.
- But here lies the challenge: these two approaches lead to drastically different results.
- Both are intuitively reasonable, so which one should we choose?
- A natural idea is to compare both accuracy estimates with the actual values of uncertainty.
- In several cases that we tried, the probabilistic result is too optimistic and the fuzzy result is too optimistic.
- The actual accuracy estimate is somewhere in between.
- So, we need a new approach to come up with realistic estimates.

6. Estimates Are Usually Reasonably Accurate

- In some cases, the original expert estimates $\widetilde{x}_i$ are really ballpark estimates.
- In such cases, the resulting estimate $\widetilde{y}$ is also not accurate.
- The problem of estimating the accuracy becomes important when the original estimates are accurate.
- Then, the differences Δx_i are reasonably small.
- Then, we can keep only linear terms in the expression

$$\Delta y = f\left(\widetilde{x}_1, \ldots, \widetilde{x}_n\right) - f\left(\widetilde{x}_1 - \Delta x_1, \ldots, \widetilde{x}_n - \Delta x_n\right).$$

- Then, $\Delta y = \sum\limits_{i=1}^{n} \delta x_i$, where $\delta x_i = c_i \cdot \Delta x_i$ and $c_i \stackrel{\text{def}}{=} \dfrac{\partial f}{\partial x_i}$.

7. Estimating the Size of Each Term δx_i

- The approximation error can be positive or negative.
- In most cases, we have no reason to believe that positive values are more probable or less probable.
- So, $-\Delta x_i$ should have the same size Δ_i as Δx_i.
- If we change the measuring unit to a c times smaller one, then all the numerical values multiply by c.
- If Δx_i is of size Δ_i, then $c_i \cdot \Delta x_i$ is of size $|c_i| \cdot \Delta_i$.
- So, we have the sum $\Delta y = \sum\limits_{i=1}^{n} \delta x_i$ of n terms δx_i each of which is of the size δ_i.
- What is the size of the sum?

8. Simple Probabilistic Approach

- Errors of different measurement are, in general, independent.
- The distribution of a sum of a large number of small independent random variables is close to Gaussian.
- This result is known as the *Central Limit Theorem.*
- In the independent case, the variances add, so Δ^2 is the sum of variances $c_i^2 \cdot \Delta_i^2$ of the terms $\delta x_i = c_i \cdot \Delta x_i$:

$$\Delta = \sqrt{\sum_{i=1}^{n} \delta_i^2} = \sqrt{\sum_{i=1}^{n} c_i^2 \cdot \Delta_i^2}.$$

9. Simple Fuzzy Approach

- In the fuzzy case, uncertainty is characterized by a membership function $\mu_i(\Delta x_i)$.
- We assume that the information about Δx_i is the same as about $-\Delta x_i$, so $\mu_i(\Delta x_i) = \mu_i(|\Delta x_i|)$.
- The larger the deviation, the less possible it is, so $\mu_i(z)$ is decreasing for $z \geq 0$.
- We assume that uncertainty is characterized by one parameter Δ_i.
- Let $\mu_0(\Delta x_0)$ be a membership function corresponding to the value 1 of this parameter.
- Then, by re-scaling, we get $\mu_i(\Delta x_i) = \mu_0\left(\frac{\Delta x_i}{\Delta_i}\right)$.
- So, for $\delta x_i = c_i \cdot \Delta x_i$, we get $\mu_i'(\delta x_i) = \mu_0\left(\frac{\delta x_i}{\delta_i}\right)$.
- By using Zadeh's extension principle, for y, we get

$$\Delta = \sum_{i=1}^{n} \delta_i = \sum_{i=1}^{n} |c_i| \cdot \Delta_i.$$

10. Resulting Challenge

- The above two formulas are different.
- E.g., if all the value δ_i are the same $\delta_1 = \ldots = \delta_n$, then:
  - in the probabilistic case, we get $\Delta = \sqrt{n} \cdot \delta_i$, while
  - in the fuzzy case, we get $\Delta = n \cdot \delta_i$.
- The difference is a factor of $\sqrt{n}$.
- When n is large – and we can have $n \approx 100$ – the difference is order of magnitude.
- So which of the two approaches should we choose?

11. We Compared the Two Approaches on Several Examples

- Our conclusion was that both methods are imperfect:
  - the probabilistic formula usually underestimated the uncertainty, while
  - the fuzzy formula usually overestimated the uncertainty.
- Lotfi Zadeh always emphasized:
  - that fuzzy logic is not a substitute for probabilities (or for any other uncertainty formalism),
  - that an ideal way to deal with uncertainty is to combine different techniques.
- So, instead of selecting one or another, let us try to combine the two approaches.

12. How to Combine the Uncertainty $\delta_i > 0$ of δx_i into an Uncertainty $\delta_1 * \delta_2$ of $\delta x_1 + \delta x_2$.

- The sum cannot be more accurate than each of the values: $\delta_1 * \delta_2 \geq \delta_i$.
- Small changes in δ_1 or in δ_2 should not lead to drastic changes in the result; so, the operation should be *continuous*.
- The sum does not depend on the order in which we add the quantities, so:

$$\delta_1 * \delta_2 = \delta_2 * \delta_1. \text{ and } (\delta_1 * \delta_2) * \delta_3 = \delta_1 * (\delta_2 * \delta_3).$$

- The result should not change if we change the measuring unit:

$$c \cdot (\delta_1 * \delta_2) = (c \cdot \delta_1) * (c \cdot \delta_2).$$

- It turns out that every operation with these properties is

$$\delta_1 * \delta_2 = \max(\delta_1, \delta_2) \text{ or } \delta_1 * \delta_2 = (\delta_1^p + \delta_2^p)^{1/p}.$$

- $p = 2$ is probabilistic case, $p = 1$ is fuzzy case, min is $p \to \infty$.
- For each domain, we need to empirically select p; e.g., for seismic data, $p \approx 1.1$.

13. Conclusions

- In many practical situations:
  - we know that the quantity y depends on the quantities $x_1, \ldots, x_n$,
  - we know the exact dependence $y = f(x_1, \ldots, x_n)$;
  - we know the approximate values $\widetilde{x}_1, \ldots, \widetilde{x}_n$ of the quantities x_i, and
  - we know the accuracies $\Delta_1, \ldots, \Delta_n$ of these estimates.
- We can then compute the estimate $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ for y.
- What is the accuracy Δ of this estimate?
- In this paper, we justify the following formula: $\Delta = \left(\sum\limits_{i=1}^{n} |c_i|^p \cdot \Delta_i^p\right)^{1/p}$.
- Here $c_i \stackrel{\text{def}}{=} \dfrac{\partial f}{\partial x_i}$ are the partial derivatives of the function $f(x_1, \ldots, x_n)$ computed for $x_i = \widetilde{x}_i$.
- p can be determined as the value for which the above formula is the closest to the actual accuracy of y.
- For example, for the analysis of seismic data, the optimal value p is $p \approx 1.1$.