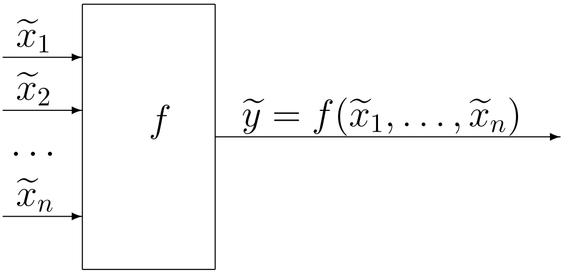
# From Gauging Accuracy of Quantity Estimates to Gauging Accuracy and Resolution of Field Measurements: A Broad Prospective on Fuzzy Transforms

Irina Perfilieva[1] and
Vladik Kreinovich[2]

[1]Inst. for Research and Applications
of Fuzzy Modeling
University of Ostrava, Czech Republic
[2]Department of Computer Science
University of Texas at El Paso
El Paso, Texas 79968, USA
vladik@utep.edu

# 1. General Problem of Data Processing under Uncertainty

- *Indirect measurements:* way to measure $y$ that are are difficult (or even impossible) to measure directly.
- *Idea:* $y = f(x_1, \ldots, x_n)$

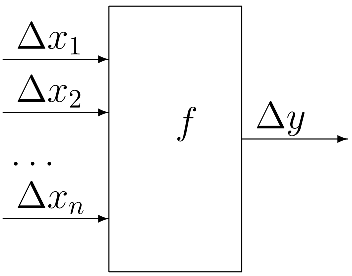



- *Problem:* measurements are never 100% accurate: $\widetilde{x}_i \neq x_i$ ($\Delta x_i \neq 0$) hence

$$\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n) \neq y = f(x_1, \ldots, x_n).$$

What are bounds on $\Delta y \stackrel{\text{def}}{=} \widetilde{y} - y$?

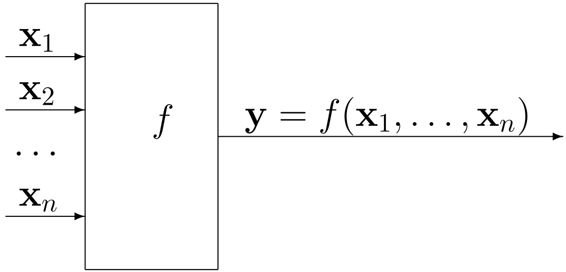
## 2. Probabilistic and Interval Uncertainty




- *Traditional approach:* we know probability distribution for $\Delta x_i$ (usually Gaussian).
- *Where it comes from:* calibration using standard MI.
- *Problem:* calibration is not possible in fundamental science like cosmology.
- *Natural solution:* assume upper bounds $\Delta_i$ on $|\Delta x_i|$ hence

$$x_i \in [\widetilde{x}_i - \Delta_i, \widetilde{x}_i + \Delta_i].$$

# 3. Interval Computations: A Problem




- *Given:* an algorithm $y = f(x_1, \ldots, x_n)$ and $n$ intervals $\mathbf{x}_i = [\underline{x}_i, \overline{x}_i]$.
- *Compute:* the corresponding range of $y$:

$$[\underline{y}, \overline{y}] = \{f(x_1, \ldots, x_n) \mid x_1 \in [\underline{x}_1, \overline{x}_1], \ldots, x_n \in [\underline{x}_n, \overline{x}_n]\}.$$

- *Fact:* NP-hard even for quadratic $f$.
- *Challenge:* when are feasible algorithm possible?
- *Challenge:* when computing $\mathbf{y} = [\underline{y}, \overline{y}]$ is not feasible, find a good approximation $\mathbf{Y} \supseteq \mathbf{y}$.

## 4. In Practice, the Situation is Often More Complex

- *Dynamics:* we measure the values $v(t)$ of a quantity $v$ at a certain moment of time $t$.
- *Spatial dependence:* we measure the value $v(x, t)$ at a certain location $x$.
- *Geophysical example:* we are interested in the values of the density at different locations and at different depth.
- *Traditional:* uncertainty in the measured value, $\widetilde{v} \approx v$.
- *New:* uncertainty in location $x$, $\widetilde{x} \approx x$.
- *Additional uncertainty:* the sensor picks up the "average" value of $v$ at locations close to $\widetilde{x}$.
- *Question:* how to describe and process the new uncertainty (*resolution*)?

## 5. Outline

- *Question* (reminder): how to describe and process uncertainty both
  - in the measured value $\widetilde{v}$ and
  - in the spatial resolution $\widetilde{x}$?
- *Natural idea:* the answer depends on what we know about the spatial resolution.
- *Possible situations:*
  - we know exactly how the measured values $\widetilde{v}_i$ are related to $v(x)$, i.e., $\widetilde{v}_i = \int w_i(x) \cdot v(x)\, dx + \Delta v_i$;
  - we only know the upper bound $\delta$ on the location error $\widetilde{x} - x$ (this is similar to the interval case);
  - we do not even know $\delta$.
- *What we do:* describe how to process all these types of uncertainty.

## 6. Situations in Which We Have Detailed Knowledge

- *Fact:* all our information about $v(x)$ is contained in the measured values $\widetilde{v}_i$.
- *Linearity assumption:* $\widetilde{v}_i = v_i + \Delta v_i$, where:
  - we have $v_i \stackrel{\text{def}}{=} \int w_i(x) \cdot v(x)\, dx$; and
  - $\Delta v_i$ is the measurement error; e.g., $|\Delta v_i| \le \Delta_i$.
- *Comment:* $v_i$ can be viewed as the value of $v(x)$ at a "fuzzy" point characterized by uncertainty $w_i(x)$.
- *Description of the situation:* we know the weights $w_i(x)$.
- *Find:* range $[\underline{y}, \overline{y}]$ for $y \stackrel{\text{def}}{=} \int w(x) \cdot v(x)\, dx$.
- *LP solution:* minimize (maximize) $\int w(x) \cdot v(x)\, dx$ under the constraints

$$\underline{v}_i \stackrel{\text{def}}{=} \widetilde{v}_i - \Delta_i \le \int w_i(x) \cdot v(x)\, dx \le \overline{v}_i \stackrel{\text{def}}{=} \widetilde{v}_i + \Delta_i.$$

## 7. Situations With Detailed Knowledge (cont-d)

- *Reminder:* find the range of $\int w(x)\cdot v(x)\,dx$ when $\underline{v}_i \le \int w_i(x)\cdot v(x)\,dx \le \overline{v}_i$.
- *General case:* when no bounds on $v(x)$, bounds are infinite – unless $w(x)$ is a linear combination of $w_i(x)$.
- *In practice* (e.g., in geophysics): $v(x)\ge 0$.
- *Similar:* imprecise probabilities (Kuznetsov, Walley).
- *Solution:* dual LP problem provides guaranteed bounds

$$\underline{v} = \sup\left\{\sum y_i\cdot \underline{v}_i : \sum y_i \cdot w_i(x) \le w(x)\right\};$$

$$\overline{v} = \inf\left\{\sum y_i\cdot \overline{v}_i : w(x) \le \sum y_i \cdot w_i(x)\right\}.$$

- *Easier* than in IP: $w_i(x)$ are localized, and we often have $\le 2$ non-zero $w_i(x)$ at each $x$.
- *Piece-wise linear* $w_i(x)$ and $w(x)$ – sufficient to check inequality at endpoints.

## 8. Situations in Which We Only Know Upper Bounds

- *Situation:* we only know;
  - the upper bound $\Delta$ on the measurement inaccuracy $\Delta v \stackrel{\rm def}{=} \widetilde{v} - v$: $|\Delta v| \le \Delta$, and
  - the upper bound $\delta$ on the location error $\Delta x \stackrel{\rm def}{=} \widetilde{x} - x$: $|\Delta v| \le \delta$.
- *Consequence:* the measured value $\widetilde{v}$ is $\Delta$-close to a convex combination of values $v(x)$ for $x$ s.t. $\|x - \widetilde{x}\| \le \Delta x$.
- *Conclusion:* $\underline{v}_\delta(\widetilde{x}) - \Delta \le \widetilde{v} \le \overline{v}_\delta(\widetilde{x}) + \Delta$, where:
  - $\underline{v}_\delta(\widetilde{x}) \stackrel{\rm def}{=} \inf\{v(x) : \|x - \widetilde{x}\| \le \delta\}$, and
  - $\overline{v}_\delta(\widetilde{x}) \stackrel{\rm def}{=} \sup\{v(x) : \|x - \widetilde{x}\| \le \delta\}$.
- *Fact:* measurement errors are random.
- *So:* it makes sense to only consider *essential* ess inf and ess sup (i.e., inf and sup modulo measure 0 sets).

# 9. What If a Model Is Only Known With Interval Uncertainty

- *Reminder:* we can tell when an observation $(\widetilde{v}, \widetilde{x})$ is consistent with a model $v(x)$:

$$\underline{v}_\delta(\widetilde{x}) - \Delta \le \widetilde{v} \le \overline{v}_\delta(\widetilde{x}) + \Delta.$$

- *Fact:* in real life, we rarely have an *exact* model $v(x)$.
- *Usually:* we have *bounds* on $v(x)$, i.e., an interval-valued model $\mathbf{v}(x) = [v^-(x), v^+(x)]$.
- *Question:* when is an observation $(\widetilde{v}, \widetilde{x})$ consistent with an *interval-valued* model?
- *General answer:* when the observation $(\widetilde{v}, \widetilde{x})$ is consistent with *one* of the models $v(x) \in \mathbf{v}(x)$.
- *A checkable answer:* an observation $(\widetilde{v}, \widetilde{x})$ is consistent with an interval-valued model $[v^-(x), v^+(x)]$ when

$$\underline{v}^-_\delta(\widetilde{x}) - \Delta \le \widetilde{v} \le \overline{v}^+_\delta(\widetilde{x}) + \Delta.$$

# 10. Situations in Which We Only Know Upper Bounds (cont-d)

- *Fact:* the actual $v(x)$ is often continuous.
- *Case of continuous $v(x)$:* we can simplify the above criterion.
- *Simplification:* the set $\widetilde{m}$ of all measurement results $(\widetilde{x}, \widetilde{x})$ is consistent with the model $v(x)$ iff

$$\forall(\widetilde{v}, \widetilde{x}) \in \widetilde{m}\, \exists(v(x), x) \in v\,((\widetilde{v}, \widetilde{x}) \text{ is } (\Delta, \delta)\text{-close to } (v(x), x)),$$

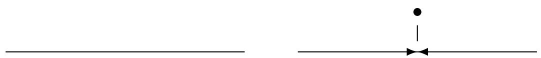
i.e., $|\widetilde{v} - v| \le \Delta$ and $\|x - \widetilde{x}\| \le \delta$.

- *Hausdorff metric:* $d_H(A, B) \le \varepsilon$ means that:

$$\forall a \in A\, \exists b \in B\, (d(a, b) \le \varepsilon) \text{ and } \forall b \in B\, \exists a \in A\, (d(a, b) \le \varepsilon).$$

- *Conclusion:* we have an *asymmetric* version of Hausdorff metric ("'quasi-metric").

# 11. Example of Asymmetry

- *Case 1:* 
  - *The actual field:* $v(0) = 1$ and $v(x) = 0$ for $x \neq 0$;
  - *Measurement results:* all zeros, i.e., $\widetilde{v} = 0$ for all $\widetilde{x}$.
  - *Conclusion:* all the measurements are *consistent* with the model.
  - *Reason:* the value $\widetilde{v} = 0$ for $\widetilde{x} = 0$ is consistent with $v(x) = 0$ for $x = \delta$ s.t. $|\widetilde{x} - x| \leq \delta$.
- *Case 2:* 
  - *The actual field:* all zeros, i.e., $v(x) = 0$ for all $x$.
  - *Measurement results:* $\widetilde{v} = 1$ for $\widetilde{x} = 0$, and $\widetilde{v} = 0$ for all $\widetilde{x} \neq 0$.
  - *Conclusion* (for $\Delta < 1$): the measurement $(1, 0)$ is *inconsistent* with the model.
  - *Reason:* for all $x$ which are $\delta$-close to $\widetilde{x} = 0$, we have $v(x) = 0$ hence we should have $|\widetilde{x} - v(x)| = |\widetilde{x}| \leq \Delta$.


Title Page
Page 12 of 14
Go Back
Full Screen
Close

# 12. Situations with No Information about Location Accuracy

- *Example:* when we solve the seismic inverse problem to find the velocity distribution.
- *Natural heuristic idea:*
  - add a perturbation of size $\Delta_0$ to the reconstructed field $\widetilde{v}(x)$;
  - simulate the new measurement results;
  - apply the same algorithm to the simulated results, and reconstruct the new field $\widetilde{v}_{\rm new}(x)$.
- *Case 1:* perturbations are *not visible* in $\widetilde{v}_{\rm new}(x) - \widetilde{v}(x)$.
- *So:* details of size $\Delta_0$ *cannot* be reconstructed: $\delta > \Delta_0$.
- *Case 2:* perturbations are *visible* in $\widetilde{v}_{\rm new}(x) - \widetilde{v}(x)$.
- *So:* details of size $\Delta_0$ *can* be reconstructed: $\delta \le \Delta_0$.

# 13. Towards Optimal Selection of Perturbations

- *Fact:* since perturbations are small, we can safely linearize their effects.
- *Conclusion:*
  - based on the results of perturbations $e_1(x), \ldots, e_k(x)$,
  - we can get the results of any linear combination

$$e(x) = c_1 \cdot e_1(x) + \ldots + c_k \cdot e_k(x).$$

- *Fact:* usually, there is no preferred spatial location.
- *Conclusion:* we can choose different locations as origins $(x = 0)$ of the coordinate system.
- *Natural requirement:* the results of perturbations should not change if we change the origin $x = 0$.

# 14. Towards Optimal Perturbations (cont-d)

- *Reminder:* the class of perturbations should not change when we change the origin $x = 0$.
- *Fact:* in new coordinates, $x_{\text{new}} = x + x_0$.
- *Conclusion:* the set $\{c_1 \cdot e_1(x) + \ldots + c_k \cdot e_k(x)\}$ must be shift-invariant: $e_i(x + x_0) = \sum\limits_{j=1}^{k} c_{ij}(x_0) \cdot e_j(x)$.
- *When* $x_0 \to 0$, we get a system of linear differential equations with constant coefficients.
- *General solution:* linear combination of expressions $\exp\left(\sum a_i \cdot x_i\right)$ with complex $a_i$.
- *Fact:* perturbations must be uniformly small.
- *So:* the only *bounded* perturbations are linear combinations of sinusoids.
- *Conclusion:* use sinusoidal perturbations.

## 15. Acknowledgments

This work was supported in part:

- by NSF Cyber-Share grant HRD-0734825:
  - A Center for Sharing Cyberresources to Advance Science and Education;
- by Grant 1 T36 GM078000-01 from the National Institutes of Health:
  - Enhancement of Qualitative Science;
- by the Japan Advanced Institute of Science and Technology (JAIST) Int'l Joint Research Grant 2006-08.


Title Page
Page 16 of 14
Go Back
Full Screen
Close

# 16. Interval Computations as a Particular Case of Global Optimization

- *Given:* an algorithm $y = f(x_1, \ldots, x_n)$ and $n$ intervals $\mathbf{x}_i = [\underline{x}_i, \overline{x}_i]$.
- *Compute:* the corresponding range of $y$:

$$[\underline{y}, \overline{y}] = \{f(x_1, \ldots, x_n) \mid x_1 \in [\underline{x}_1, \overline{x}_1], \ldots, x_n \in [\underline{x}_n, \overline{x}_n]\}.$$

- *Reduction to optimization:* in the general case, $\underline{y}$ ($\overline{y}$):

$$\text{Minimize} \quad (\text{Maximize}) \quad f(x_1, \ldots, x_n)$$

where $f$ is directly computable, under the constraints

$$\underline{x}_1 \le x_1 \le \overline{x}_1, \quad \ldots, \quad \underline{x}_n \le x_n \le \overline{x}_n.$$

- *Cosmological case:* $f$ is not directly computable:

$$f(x_1, \ldots, x_n) \stackrel{\text{def}}{=} \text{argmin} \quad F(x_1, \ldots, x_n, y_1, \ldots, y_m).$$

## 17. Linearization

- *General case:* NP-hard.
- *Typical situation:* direct measurements are accurate enough, so the approximation errors $\Delta x_i$ are small.
- *Conclusion:* terms quadratic (or of higher order) in $\Delta x_i$ can be safely neglected.
- *Example:* for $\Delta x_i = 1\%$, we have $\Delta x_i^2 = 0.01\% \ll \Delta x_i$.
- *Linearization:*
  - expand $f$ in Taylor series around the point $(\widetilde{x}_1, \ldots, \widetilde{x}_n)$;
  - restrict ourselves only to linear terms:

$$\Delta y = c_1 \cdot \Delta x_1 + \ldots + c_n \cdot \Delta x_n, \text{ where } c_i \stackrel{\text{def}}{=} \frac{\partial f}{\partial x_i}.$$

- *Interval case:* $|\Delta x_i| \le \Delta_i$.
- *Result:* $\Delta \stackrel{\text{def}}{=} \max |\Delta y| = |c_1| \cdot \Delta_1 + \ldots + |c_n| \cdot \Delta_n$.

## 18. Computations under Linearization: From Numerical Differentiation to Monte-Carlo Approach

- *Linearization:* $\Delta y = \sum\limits_{i=1}^{n} c_i \cdot \Delta x_i$, where $c_i \stackrel{\text{def}}{=} \dfrac{\partial f}{\partial x_i}$.
- *Formulas:* $\sigma^2 = \sum\limits_{i=1}^{n} c_i^2 \cdot \sigma_i^2$, $\Delta = \sum\limits_{i=1}^{n} |c_i| \cdot \Delta_i$.
- *Numerical differentiation:* $n$ iterations, too long.
- *Monte-Carlo approach:* if $\Delta x_i$ are Gaussian w/$\sigma_i$, then $\Delta y = \sum\limits_{i=1}^{n} c_i \cdot \Delta x_i$ is also Gaussian, w/desired $\sigma$.
- *Advantage:* # of iterations does not grow with $n$.
- *Interval estimates:* if $\Delta x_i$ are Cauchy, w/$\rho_i(x) = \dfrac{\Delta_i}{\Delta_i^2 + x^2}$, then $\Delta y = \sum\limits_{i=1}^{n} c_i \cdot \Delta x_i$ is also Cauchy, w/desired $\Delta$.

## 19. Resulting Fast (Linearized) Algorithm for Estimating Interval Uncertainty

- Apply $f$ to $\widetilde{x}_i$: $\widetilde{y} := f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$;
- For $k = 1, 2, \ldots, N$, repeat the following:
  - use RNG to get $r_i^{(k)}$, $i = 1, \ldots, n$ from $U[0,1]$;
  - get st. Cauchy values $c_i^{(k)} := \tan(\pi \cdot (r_i^{(k)} - 0.5))$;
  - compute $K := \max_i |c_i^{(k)}|$ (to stay in linearized area);
  - simulate "actual values" $x_i^{(k)} := \widetilde{x}_i - \delta_i^{(k)}$, where $\delta_i^{(k)} := \Delta_i \cdot c_i^{(k)}/K$;
  - simulate error of the indirect measurement:
  $$\delta^{(k)} := K \cdot \left(\widetilde{y} - f\left(x_1^{(k)}, \ldots, x_n^{(k)}\right)\right);$$
- Solve the ML equation $\sum\limits_{k=1}^{N} \dfrac{1}{1 + \left(\dfrac{\delta^{(k)}}{\Delta}\right)^2} = \dfrac{N}{2}$ by bisection, and get the desired $\Delta$.

## 20. A New (Heuristic) Approach

- *Problem:* guaranteed (interval) bounds are too high.
- *Gaussian case:* we only have bounds guaranteed with confidence, say, 90%.
- *How:* cut top 5% and low 5% off a normal distribution.
- *New idea:* to get similarly estimates for intervals, we "cut off" top 5% and low 5% of Cauchy distribution.
- *How:*
  - find the threshold value $x_0$ for which the probability of exceeding this value is, say, 5%;
  - replace values $x$ for which $x > x_0$ with $x_0$;
  - replace values $x$ for which $x < -x_0$ with $-x_0$;
  - use this "cut-off" Cauchy in error estimation.
- *Example:* for 95% confidence level, we need $x_0 = 12.706$.

# 21. Alternative Approach: Maximum Entropy

- *Situation:* in many practical applications, it is very difficult to come up with the probabilities.
- *Traditional engineering approach:* use probabilistic techniques.
- *Problem:* many different probability distributions are consistent with the same observations.
- *Solution:* select one of these distributions – e.g., the one with the largest entropy.
- *Example – single variable:* if all we know is that $x \in [\underline{x}, \overline{x}]$, then MaxEnt leads to the uniform distribution.
- *Example – multiple variables:* different variables are independently distributed.

## 22. General Limitations of Maximum Entropy Approach

- *Example:* simplest algorithm $y = x_1 + \ldots + x_n$.
- *Measurement errors:* $\Delta x_i \in [-\Delta, \Delta]$.
- *Analysis:* $\Delta y = \Delta x_1 + \ldots + \Delta x_n$.
- *Worst case situation:* $\Delta y = n \cdot \Delta$.
- *Maximum Entropy approach:* due to Central Limit Theorem, $\Delta y$ is $\approx$ normal, with $\sigma = \Delta \cdot \dfrac{\sqrt{n}}{\sqrt{3}}$.
- *Why this may be inadequate:* we get $\Delta \sim \sqrt{n}$, but due to correlation, it is possible that $\Delta = n \cdot \Delta \sim n \gg \sqrt{n}$.
- *Conclusion:* using a single distribution can be very misleading, especially if we want guaranteed results.
- Examples: high-risk application areas such as space exploration or nuclear engineering.

## 23. Interval Computations: A Brief History

- *Origins:* Archimedes (Ancient Greece)
- *Modern pioneers:* Warmus (Poland), Sunaga (Japan), Moore (USA), 1956–59
- *First boom:* early 1960s.
- *First challenge:* taking interval uncertainty into account when planning spaceflights to the Moon.
- *Current applications* (sample):
  - design of elementary particle colliders: Berz, Kyoko (USA)
  - will a comet hit the Earth: Berz, Moore (USA)
  - robotics: Jaulin (France), Neumaier (Austria)
  - chemical engineering: Stadtherr (USA)

# 24. Interval Arithmetic: Foundations of Interval Techniques

- *Problem:* compute the range

  $[\underline{y}, \overline{y}] = \{f(x_1, \ldots, x_n) \mid x_1 \in [\underline{x}_1, \overline{x}_1], \ldots, x_n \in [\underline{x}_n, \overline{x}_n]\}.$

- *Interval arithmetic:* for arithmetic operations $f(x_1, x_2)$ (and for elementary functions), we have explicit formulas for the range.

- *Examples:* when $x_1 \in \mathbf{x}_1 = [\underline{x}_1, \overline{x}_1]$ and $x_2 \in \mathbf{x}_2 = [\underline{x}_2, \overline{x}_2]$, then:
  - The range $\mathbf{x}_1 + \mathbf{x}_2$ for $x_1 + x_2$ is $[\underline{x}_1 + \underline{x}_2, \overline{x}_1 + \overline{x}_2]$.
  - The range $\mathbf{x}_1 - \mathbf{x}_2$ for $x_1 - x_2$ is $[\underline{x}_1 - \overline{x}_2, \overline{x}_1 - \underline{x}_2]$.
  - The range $\mathbf{x}_1 \cdot \mathbf{x}_2$ for $x_1 \cdot x_2$ is $[\underline{y}, \overline{y}]$, where

$$\underline{y} = \min(\underline{x}_1 \cdot \underline{x}_2, \underline{x}_1 \cdot \overline{x}_2, \overline{x}_1 \cdot \underline{x}_2, \overline{x}_1 \cdot \overline{x}_2);$$

$$\overline{y} = \max(\underline{x}_1 \cdot \underline{x}_2, \underline{x}_1 \cdot \overline{x}_2, \overline{x}_1 \cdot \underline{x}_2, \overline{x}_1 \cdot \overline{x}_2).$$

- The range $1/\mathbf{x}_1$ for $1/x_1$ is $[1/\overline{x}_1, 1/\underline{x}_1]$ (if $0 \notin \mathbf{x}_1$).

## 25. Straightforward Interval Computations: Example

- *Example:* $f(x) = (x-2)\cdot(x+2)$, $x \in [1,2]$.
- How will the computer compute it?
  - $r_1 := x - 2$;
  - $r_2 := x + 2$;
  - $r_3 := r_1 \cdot r_2$.
- *Main idea:* perform the same operations, but with *intervals* instead of *numbers*:
  - $\mathbf{r}_1 := [1,2] - [2,2] = [-1,0]$;
  - $\mathbf{r}_2 := [1,2] + [2,2] = [3,4]$;
  - $\mathbf{r}_3 := [-1,0]\cdot[3,4] = [-4,0]$.
- *Actual range:* $f(\mathbf{x}) = [-3,0]$.
- *Comment:* this is just a toy example, there are more efficient ways of computing an enclosure $\mathbf{Y} \supseteq \mathbf{y}$.

## 26. First Idea: Use of Monotonicity

- *Reminder:* for arithmetic, we had exact ranges.
- *Reason:* $+$, $-$, $\cdot$ are monotonic in each variable.
- *How monotonicity helps:* if $f(x_1, \ldots, x_n)$ is (non-strictly) increasing ($f \uparrow$) in each $x_i$, then

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_n) = [f(\underline{x}_1, \ldots, \underline{x}_n), f(\overline{x}_1, \ldots, \overline{x}_n)].$$

- *Similarly:* if $f \uparrow$ for some $x_i$ and $f \downarrow$ for other $x_j$ $(-)$.
- *Fact:* $f \uparrow$ in $x_i$ if $\dfrac{\partial f}{\partial x_i} \geq 0$.
- *Checking monotonicity:* check that the range $[\underline{r}_i, \overline{r}_i]$ of $\dfrac{\partial f}{\partial x_i}$ on $\mathbf{x}_i$ has $\underline{r}_i \geq 0$.
- *Differentiation:* by Automatic Differentiation (AD) tools.
- *Estimating ranges of* $\dfrac{\partial f}{\partial x_i}$*:* straightforward interval comp.

# 27. Monotonicity: Example

- *Idea:* if the range $[\underline{r}_i, \overline{r}_i]$ of each $\dfrac{\partial f}{\partial x_i}$ on $\mathbf{x}_i$ has $\underline{r}_i \geq 0$, then

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_n) = [f(\underline{x}_1, \ldots, \underline{x}_n), f(\overline{x}_1, \ldots, \overline{x}_n)].$$

- *Example:* $f(x) = (x - 2) \cdot (x + 2)$, $\mathbf{x} = [1, 2]$.
- *Case $n = 1$:* if the range $[\underline{r}, \overline{r}]$ of $\dfrac{df}{dx}$ on $\mathbf{x}$ has $\underline{r} \geq 0$, then

$$f(\mathbf{x}) = [f(\underline{x}), f(\overline{x})].$$

- *AD:* $\dfrac{df}{dx} = 1 \cdot (x + 2) + (x - 2) \cdot 1 = 2x.$
- *Checking:* $[\underline{r}, \overline{r}] = [2, 4]$, with $2 \geq 0$.
- *Result:* $f([1, 2]) = [f(1), f(2)] = [-3, 0]$.
- *Comparison:* this is the exact range.

# 28. Non-Monotonic Example

- *Example:* $f(x) = x \cdot (1 - x)$, $x \in [0, 1]$.
- How will the computer compute it?
  - $r_1 := 1 - x$;
  - $r_2 := x \cdot r_1$.
- *Straightforward interval computations:*
  - $\mathbf{r}_1 := [1, 1] - [0, 1] = [0, 1]$;
  - $\mathbf{r}_2 := [0, 1] \cdot [0, 1] = [0, 1]$.
- *Actual range:* min, max of $f$ at $\underline{x}$, $\overline{x}$, or when $\dfrac{df}{dx} = 0$.
- Here, $\dfrac{df}{dx} = 1 - 2x = 0$ for $x = 0.5$, so
  - compute $f(0) = 0$, $f(0.5) = 0.25$, and $f(1) = 0$.
  - $\underline{y} = \min(0, 0.25, 0) = 0$, $\overline{y} = \max(0, 0.25, 0) = 0.25$.
- *Resulting range:* $f(\mathbf{x}) = [0, 0.25]$.

# 29. Second Idea: Centered Form

- *Main idea:* Intermediate Value Theorem

$$f(x_1, \ldots, x_n) = f(\widetilde{x}_1, \ldots, \widetilde{x}_n) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\chi) \cdot (x_i - \widetilde{x}_i)$$

  for some $\chi_i \in \mathbf{x}_i$.

- *Corollary:* $f(x_1, \ldots, x_n) \in \mathbf{Y}$, where

$$\mathbf{Y} = \widetilde{y} + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_1, \ldots, \mathbf{x}_n) \cdot [-\Delta_i, \Delta_i].$$

- *Differentiation:* by Automatic Differentiation (AD) tools.
- *Estimating the ranges of derivatives:*
  - if appropriate, by monotonicity, or
  - by straightforward interval computations, or
  - by centered form (more time but more accurate).

## 30. Centered Form: Example

- *General formula:*

$$\mathbf{Y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_1, \ldots, \mathbf{x}_n) \cdot [-\Delta_i, \Delta_i].$$

- *Example:* $f(x) = x \cdot (1 - x)$, $\mathbf{x} = [0, 1]$.
- Here, $\mathbf{x} = [\widetilde{x} - \Delta, \widetilde{x} + \Delta]$, with $\widetilde{x} = 0.5$ and $\Delta = 0.5$.
- *Case $n = 1$:* $\mathbf{Y} = f(\widetilde{x}) + \dfrac{df}{dx}(\mathbf{x}) \cdot [-\Delta, \Delta]$.
- *AD:* $\dfrac{df}{dx} = 1 \cdot (1 - x) + x \cdot (-1) = 1 - 2x$.
- *Estimation:* we have $\dfrac{df}{dx}(\mathbf{x}) = 1 - 2 \cdot [0, 1] = [-1, 1]$.
- *Result:* $\mathbf{Y} = 0.5 \cdot (1 - 0.5) + [-1, 1] \cdot [-0.5, 0.5] = 0.25 + [-0.5, 0.5] = [-0.25, 0.75]$.
- *Comparison:* actual range $[0, 0.25]$, straightforward $[0, 1]$.

## 31. Third Idea: Bisection

- *Known:* accuracy $O(\Delta_i^2)$ of first order formula

$$f(x_1, \ldots, x_n) = f(\widetilde{x}_1, \ldots, \widetilde{x}_n) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\chi) \cdot (x_i - \widetilde{x}_i).$$

- *Idea:* if the intervals are too wide, we:
  - split one of them in half ($\Delta_i^2 \to \Delta_i^2/4$); and
  - take the union of the resulting ranges.
- *Example:* $f(x) = x \cdot (1 - x)$, where $x \in \mathbf{x} = [0, 1]$.
- *Split:* take $\mathbf{x}' = [0, 0.5]$ and $\mathbf{x}'' = [0.5, 1]$.
- *1st range:* $1 - 2 \cdot \mathbf{x} = 1 - 2 \cdot [0, 0.5] = [0, 1]$, so $f \uparrow$ and $f(\mathbf{x}') = [f(0), f(0.5)] = [0, 0.25]$.
- *2nd range:* $1 - 2 \cdot \mathbf{x} = 1 - 2 \cdot [0.5, 1] = [-1, 0]$, so $f \downarrow$ and $f(\mathbf{x}'') = [f(1), f(0.5)] = [0, 0.25]$.
- *Result:* $f(\mathbf{x}') \cup f(\mathbf{x}'') = [0, 0.25]$ – exact.

## 32. Alternative Approach: Affine Arithmetic

- *So far:* we compute the range of $x \cdot (1-x)$ by multiplying ranges of $x$ and $1-x$.
- *We ignore:* that both factors depend on $x$ and are, thus, dependent.
- *Idea:* for each intermediate result $a$, keep an explicit dependence on $\Delta x_i = \widetilde{x}_i - x_i$ (at least its linear terms).
- *Implementation:*

$$a = a_0 + \sum_{i=1}^{n} a_i \cdot \Delta x_i + [\underline{a}, \overline{a}].$$

- *We start:* with $x_i = \widetilde{x}_i - \Delta x_i$, i.e.,

$$\widetilde{x}_i + 0 \cdot \Delta x_1 + \ldots + 0 \cdot \Delta x_{i-1} + (-1) \cdot \Delta x_i + 0 \cdot \Delta x_{i+1} + \ldots + 0 \cdot \Delta x_n + [0,0].$$

- *Description:* $a_0 = \widetilde{x}_i$, $a_i = -1$, $a_j = 0$ for $j \neq i$, and $[\underline{a}, \overline{a}] = [0,0]$.

## 33. Affine Arithmetic: Operations

- *Representation:* $a = a_0 + \sum\limits_{i=1}^{n} a_i \cdot \Delta x_i + [\underline{a}, \overline{a}]$.
- *Input:* $a = a_0 + \sum\limits_{i=1}^{n} a_i \cdot \Delta x_i + \mathbf{a}$ and $b = b_0 + \sum\limits_{i=1}^{n} b_i \cdot \Delta x_i + \mathbf{b}$.
- *Operations:* $c = a \otimes b$.
- *Addition:* $c_0 = a_0 + b_0$, $c_i = a_i + b_i$, $\mathbf{c} = \mathbf{a} + \mathbf{b}$.
- *Subtraction:* $c_0 = a_0 - b_0$, $c_i = a_i - b_i$, $\mathbf{c} = \mathbf{a} - \mathbf{b}$.
- *Multiplication:* $c_0 = a_0 \cdot b_0$, $c_i = a_0 \cdot b_i + b_0 \cdot a_i$,

$$\mathbf{c} = a_0 \cdot \mathbf{b} + b_0 \cdot \mathbf{a} + \sum_{i \neq j} a_i \cdot b_j \cdot [-\Delta_i, \Delta_i] \cdot [-\Delta_j, \Delta_j] +$$

$$\sum_i a_i \cdot b_i \cdot [-\Delta_i, \Delta_i]^2 +$$

$$\left( \sum_i a_i \cdot [-\Delta_i, \Delta_i] \right) \cdot \mathbf{b} + \left( \sum_i b_i \cdot [-\Delta_i, \Delta_i] \right) \cdot \mathbf{a} + \mathbf{a} \cdot \mathbf{b}.$$

## 34. Affine Arithmetic: Example

- *Example:* $f(x) = x \cdot (1 - x)$, $x \in [0, 1]$.
- Here, $n = 1$, $\widetilde{x} = 0.5$, and $\Delta = 0.5$.
- How will the computer compute it?
  - $r_1 := 1 - x$;
  - $r_2 := x \cdot r_1$.
- *Affine arithmetic:* we start with $x = 0.5 - \Delta x + [0, 0]$;
  - $\mathbf{r}_1 := 1 - (0.5 - \Delta) = 0.5 + \Delta x$;
  - $\mathbf{r}_2 := (0.5 - \Delta x) \cdot (0.5 + \Delta x)$, i.e.,

$$\mathbf{r}_2 = 0.25 + 0 \cdot \Delta x - [-\Delta, \Delta]^2 = 0.25 + [-\Delta^2, 0].$$

- *Resulting range:* $\mathbf{y} = 0.25 + [-0.25, 0] = [0, 0.25]$.
- *Comparison:* this is the exact range.

## 35. Affine Arithmetic: Towards More Accurate Estimates

- *In our simple example:* we got the exact range.
- *In general:* range estimation is NP-hard.
- *Meaning:* a feasible (polynomial-time) algorithm will sometimes lead to excess width: $\mathbf{Y} \supset \mathbf{y}$.
- *Conclusion:* affine arithmetic may lead to excess width.
- *Question:* how to get more accurate estimates?
- *First idea:* bisection.
- *Second idea* (Taylor arithmetic):
  - *affine arithmetic:* $a = a_0 + \sum a_i \cdot \Delta x_i + \mathbf{a}$;
  - *meaning:* we keep linear terms in $\Delta x_i$;
  - *idea:* keep, e.g., quadratic terms

$$a = a_0 + \sum a_i \cdot \Delta x_i + \sum a_{ij} \cdot \Delta x_i \cdot \Delta x_j + \mathbf{a}.$$

## 36. Interval Computations vs. Affine Arithmetic: Comparative Analysis

- *Objective:* we want a method that computes a reasonable estimate for the range in reasonable time.
- *Conclusion – how to compare different methods:*
  - how accurate are the estimates, and
  - how fast we can compute them.
- *Accuracy:* affine arithmetic leads to more accurate ranges.
- *Computation time:*
  - *Interval arithmetic:* for each intermediate result $a$, we compute two values: endpoints $\underline{a}$ and $\overline{a}$ of $[\underline{a}, \overline{a}]$.
  - *Affine arithmetic:* for each $a$, we compute $n + 3$ values:

$$a_0 \quad a_1, \ldots, a_n \quad \underline{a}, \overline{a}.$$

- *Conclusion:* affine arithmetic is $\sim n$ times slower.

# 37. Solving Systems of Equations: Extending Known Algorithms to Situations with Interval Uncertainty

- *We have:* a system of equations $g_i(y_1, \ldots, y_n) = a_i$ with unknowns $y_i$;
- *We know:* $a_i$ with interval uncertainty: $a_i \in [\underline{a}_i, \overline{a}_i]$;
- *We want:* to find the corresponding ranges of $y_j$.
- *First case:* for exactly known $a_i$, we have an algorithm $y_j = f_j(a_1, \ldots, a_n)$ for solving the system.
- *Example:* system of linear equations.
- *Solution:* apply interval computations techniques to find the range $f_j([\underline{a}_1, \overline{a}_1], \ldots, [\underline{a}_n, \overline{a}_n])$.
- *Better solution:* for specific equations, we often already know which ideas work best.
- *Example:* linear equations $Ay = b$; $y$ is monotonic in $b$.

## 38. Solving Systems of Equations When No Algorithm Is Known

- *Idea:*
  - parse each equation into elementary constraints, and
  - use interval computations to improve original ranges until we get a narrow range (= solution).
- *First example:* $x - x^2 = 0.5$, $x \in [0, 1]$ (no solution).
- *Parsing:* $r_1 = x^2$, $0.5\ (= r_2)\ = x - r_1$.
- *Rules:* from $r_1 = x^2$, we extract two rules:

$$(1)\ x \to r_1 = x^2; \quad (2)\ r_1 \to x = \sqrt{r_1};$$

from $0.5 = x - r_1$, we extract two more rules:

$$(3)\ x \to r_1 = x - 0.5; \quad (4)\ r_1 \to x = r_1 + 0.5.$$

## 39. Solving Systems of Equations When No Algorithm Is Known: Example

- (1) $r = x^2$; (2) $x = \sqrt{r}$; (3) $r = x - 0.5$; (4) $x = r + 0.5$.
- *We start with:* $\mathbf{x} = [0, 1]$, $\mathbf{r} = (-\infty, \infty)$.

(1) $\mathbf{r} = [0, 1]^2 = [0, 1]$, so $\mathbf{r}_{\text{new}} = (-\infty, \infty) \cap [0, 1] = [0, 1]$.

(2) $\mathbf{x}_{\text{new}} = \sqrt{[0, 1]} \cap [0, 1] = [0, 1]$ – no change.

(3) $\mathbf{r}_{\text{new}} = ([0, 1] - 0.5) \cap [0, 1] = [-0.5, 0.5] \cap [0, 1] = [0, 0.5]$.

(4) $\mathbf{x}_{\text{new}} = ([0, 0.5] + 0.5) \cap [0, 1] = [0.5, 1] \cap [0, 1] = [0.5, 1]$.

(1) $\mathbf{r}_{\text{new}} = [0.5, 1]^2 \cap [0, 0.5] = [0.25, 0.5]$.

(2) $\mathbf{x}_{\text{new}} = \sqrt{[0.25, 0.5]} \cap [0.5, 1] = [0.5, 0.71]$;
round $\underline{a}$ down $\downarrow$ and $\overline{a}$ up $\uparrow$, to guarantee enclosure.

(3) $\mathbf{r}_{\text{new}} = ([0.5, 0.71] - 0.5) \cap [0.25, 5] = [0.0.21] \cap [0.25, 0.5]$,
i.e., $\mathbf{r}_{\text{new}} = \emptyset$.

- *Conclusion:* the original equation has no solutions.

# 40. Solving Systems of Equations: Second Example

- *Example:* $x - x^2 = 0$, $x \in [0, 1]$.
- *Parsing:* $r_1 = x^2$, $0\ (= r_2)\ = x - r_1$.
- *Rules:* (1) $r = x^2$; (2) $x = \sqrt{r}$; (3) $r = x$; (4) $x = r$.
- *We start with:* $\mathbf{x} = [0, 1]$, $\mathbf{r} = (-\infty, \infty)$.
- *Problem:* after Rule 1, we're stuck with $\mathbf{x} = \mathbf{r} = [0, 1]$.
- *Solution:* bisect $\mathbf{x} = [0, 1]$ into $[0, 0.5]$ and $[0.5, 1]$.
- *For 1st subinterval:*
  - Rule 1 leads to $\mathbf{r}_{\text{new}} = [0, 0.5]^2 \cap [0, 0.5] = [0, 0.25]$;
  - Rule 4 leads to $\mathbf{x}_{\text{new}} = [0, 0.25]$;
  - Rule 1 leads to $\mathbf{r}_{\text{new}} = [0, 0.25]^2 = [0, 0.0625]$;
  - Rule 4 leads to $\mathbf{x}_{\text{new}} = [0, 0.0625]$; etc.
  - we converge to $x = 0$.
- *For 2nd subinterval:* we converge to $x = 1$.

# 41. Optimization: Extending Known Algorithms to Situations with Interval Uncertainty

- *Problem:* find $y_1, \ldots, y_m$ for which

$$g(y_1, \ldots, y_m, a_1, \ldots, a_m) \to \max .$$

- *We know:* $a_i$ with interval uncertainty: $a_i \in [\underline{a}_i, \overline{a}_i]$;
- *We want:* to find the corresponding ranges of $y_j$.
- *First case:* for exactly known $a_i$, we have an algorithm $y_j = f_j(a_1, \ldots, a_n)$ for solving the optimization problem.
- *Example:* quadratic objective function $g$.
- *Solution:* apply interval computations techniques to find the range $f_j([\underline{a}_1, \overline{a}_1], \ldots, [\underline{a}_n, \overline{a}_n])$.
- *Better solution:* for specific $f$, we often already know which ideas work best.

# 42. Optimization When No Algorithm Is Known

- *Idea:* divide the original box $\mathbf{x}$ into subboxes $\mathbf{b}$.
- If $\max\limits_{x \in \mathbf{b}} g(x) < g(x')$ for a known $x'$, dismiss $\mathbf{b}$.
- *Example:* $g(x) = x \cdot (1 - x)$, $\mathbf{x} = [0, 1]$.
- Divide into 10 (?) subboxes $\mathbf{b} = [0, 0.1], [0.1, 0.2], \ldots$
- Find $g(\widetilde{b})$ for each $\mathbf{b}$; the largest is $0.45 \cdot 0.55 = 0.2475$.
- Compute $G(\mathbf{b}) = g(\widetilde{b}) + (1 - 2 \cdot \mathbf{b}) \cdot [-\Delta, \Delta]$.
- Dismiss subboxes for which $\overline{Y} < 0.2475$.
- *Example:* for $[0.2, 0.3]$, we have
  $$0.25 \cdot (1 - 0.25) + (1 - 2 \cdot [0.2, 0.3]) \cdot [-0.05, 0.05].$$
- Here $\overline{Y} = 0.2175 < 0.2475$, so we dismiss $[0.2, 0.3]$.
- *Result:* keep only boxes $\subseteq [0.3, 0.7]$.
- *Further subdivision:* get us closer and closer to $x = 0.5$.