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

Irina Perfilieva\textsuperscript{1} and Vladik Kreinovich\textsuperscript{2}

\textsuperscript{1}Inst. for Research and Applications of Fuzzy Modeling
University of Ostrava, Czech Republic
\textsuperscript{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 difficult (or even impossible) to measure directly.

- **Idea**: $y = f(x_1, \ldots, x_n)$

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

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

What are bounds on $\Delta y = \tilde{y} - y$?
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

$$x_i \in [\tilde{x}_i - \Delta_i, \tilde{x}_i + \Delta_i].$$
3. Interval Computations: A Problem

- **Given**: an algorithm $y = f(x_1, \ldots, x_n)$ and $n$ intervals $x_i = [x_i, \bar{x}_i]$.

- **Compute**: the corresponding range of $y$:
  $$[\underline{y}, \bar{y}] = \{ f(x_1, \ldots, x_n) \mid x_1 \in [\underline{x}_1, \bar{x}_1], \ldots, x_n \in [\underline{x}_n, \bar{x}_n] \}.$$  

- **Fact**: NP-hard even for quadratic $f$.

- **Challenge**: when are feasible algorithm possible?

- **Challenge**: when computing $y = [\underline{y}, \bar{y}]$ is not feasible, find a good approximation $Y \supseteq y$. 
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, $\tilde{v} \approx v$.

- **New**: uncertainty in location $x$, $\tilde{x} \approx x$.

- **Additional uncertainty**: the sensor picks up the “average” value of $v$ at locations close to $\tilde{x}$.

- **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 $\tilde{v}$ and
  – in the spatial resolution $\tilde{x}$?

• Natural idea: the answer depends on what we know about the spatial resolution.

• Possible situations:
  – we know exactly how the measured values $\tilde{v}_i$ are related to $v(x)$, i.e., $\tilde{v}_i = \int w_i(x) \cdot v(x) \, dx + \Delta v_i$;
  – we only know the upper bound $\delta$ on the location error $\tilde{x} - x$ (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 $\tilde{v}_i$.

- **Linearity assumption:** $\tilde{v}_i = v_i + \Delta v_i$, where:
  - we have $v_i \overset{\text{def}}{=} \int w_i(x) \cdot v(x) \, dx$; and
  - $\Delta v_i$ is the measurement error; e.g., $|\Delta v_i| \leq \Delta_i$.

- **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 $[y, \bar{y}]$ for $y \overset{\text{def}}{=} \int w(x) \cdot v(x) \, dx$.

- **LP solution:** minimize (maximize) $\int w(x) \cdot v(x) \, dx$ under the constraints

  $$v_i \overset{\text{def}}{=} \tilde{v}_i - \Delta_i \leq \int w_i(x) \cdot v(x) \, dx \leq \bar{v}_i \overset{\text{def}}{=} \tilde{v}_i + \Delta_i.$$
7. Situations With Detailed Knowledge (cont-d)

- **Reminder:** find the range of $\int w(x) \cdot v(x) \, dx$ when $v_i \leq \int w_i(x) \cdot v(x) \, dx \leq \bar{v}_i$.

- **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): $v(x) \geq 0$.

- **Similar:** imprecise probabilities (Kuznetsoov, Walley).

- **Solution:** dual LP problem provides guaranteed bounds

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

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

- **Easier** than in IP: $w_i(x)$ are localized, and we often have $\leq 2$ 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 $\Delta v \overset{\text{def}}{=} \tilde{v} - v$: $|\Delta v| \leq \Delta$, and
  - the upper bound $\delta$ on the location error $\Delta x \overset{\text{def}}{=} \tilde{x} - x$: $|\Delta v| \leq \delta$.

- **Consequence:** the measured value $\tilde{v}$ is $\Delta$-close to a convex combination of values $v(x)$ for $x$ s.t. $\|x - \tilde{x}\| \leq \Delta x$.

- **Conclusion:** $v_\delta(\tilde{x}) - \Delta \leq \tilde{v} \leq \overline{v_\delta}(\tilde{x}) + \Delta$, where:
  - $v_\delta(\tilde{x}) \overset{\text{def}}{=} \inf\{v(x) : \|x - \tilde{x}\| \leq \delta\}$, and
  - $\overline{v_\delta}(\tilde{x}) \overset{\text{def}}{=} \sup\{v(x) : \|x - \tilde{x}\| \leq \delta\}$.

- **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 \((\tilde{v}, \tilde{x})\) is consistent with a model \(v(x)\):

\[
v_\delta(\tilde{x}) - \Delta \leq \tilde{v} \leq \overline{v}_\delta(\tilde{x}) + \Delta.
\]

- **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 \(v(x) = [v^-(x), v^+(x)]\).
- **Question:** when is an observation \((\tilde{v}, \tilde{x})\) consistent with an *interval-valued* model?
- **General answer:** when the observation \((\tilde{v}, \tilde{x})\) is consistent with *one* of the models \(v(x) \in v(x)\).
- **A checkable answer:** an observation \((\tilde{v}, \tilde{x})\) is consistent with an interval-valued model \([v^-(x), v^+(x)]\) when

\[
\overline{v}^-_\delta(\tilde{x}) - \Delta \leq \tilde{v} \leq \overline{v}^+_\delta(\tilde{x}) + \Delta.
\]
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 \( \tilde{m} \) of all measurement results \((\tilde{x}, \tilde{x})\) is consistent with the model \( v(x) \) iff

\[
\forall (\tilde{v}, \tilde{x}) \in \tilde{m} \exists (v(x), x) \in v \left((\tilde{v}, \tilde{x}) \text{ is } (\Delta, \delta)-\text{close to } (v(x), x)\right),
\]

i.e., \(|\tilde{v} - v| \leq \Delta \) and \(|x - \tilde{x}| \leq \delta\).

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

\[
\forall a \in A \exists b \in B (d(a, b) \leq \varepsilon) \text{ and } \forall b \in B \exists a \in A (d(a, b) \leq \varepsilon).
\]

- **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., $\tilde{v} = 0$ for all $\tilde{x}$.
  - *Conclusion:* all the measurements are consistent with the model.
  - *Reason:* the value $\tilde{v} = 0$ for $\tilde{x} = 0$ is consistent with $v(x) = 0$ for $x = \delta$ s.t. $|\tilde{x} - x| \leq \delta$.

- **Case 2:**
  - *The actual field:* all zeros, i.e., $v(x) = 0$ for all $x$.
  - *Measurement results:* $\tilde{v} = 1$ for $\tilde{x} = 0$, and $\tilde{v} = 0$ for all $\tilde{x} \neq 0$.
  - *Conclusion (for $\Delta < 1$):* the measurement $(1, 0)$ is inconsistent with the model.
  - *Reason:* for all $x$ which are $\delta$-close to $\tilde{x} = 0$, we have $v(x) = 0$ hence we should have $|\tilde{x} - v(x)| = |\tilde{x}| \leq \Delta$. 
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 $\tilde{v}(x)$;
  - simulate the new measurement results;
  - apply the same algorithm to the simulated results, and reconstruct the new field $\tilde{v}_{\text{new}}(x)$.

- *Case 1:* perturbations are *not visible* in $\tilde{v}_{\text{new}}(x) - \tilde{v}(x)$.
- *So:* details of size $\Delta_0$ *cannot* be reconstructed: $\delta > \Delta_0$.

- *Case 2:* perturbations are *visible* in $\tilde{v}_{\text{new}}(x) - \tilde{v}(x)$.
- *So:* details of size $\Delta_0$ *can* be reconstructed: $\delta \leq \Delta_0$. 
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: $e_i(x + x_0) = \sum_{j=1}^{k} c_{ij}(x_0) \cdot e_j(x)$.

- **When** $x_0 \to 0$, we get a system of linear differential equations with constant coefficients.

- **General solution**: linear combination of expressions $\exp (\sum a_i \cdot x_i)$ 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.
16. Interval Computations as a Particular Case of Global Optimization

- **Given:** an algorithm \( y = f(x_1, \ldots, x_n) \) and \( n \) intervals \( x_i = [x_i, \bar{x}_i] \).

- **Compute:** the corresponding range of \( y \):
  \[ [\underline{y}, \overline{y}] = \{ f(x_1, \ldots, x_n) \mid x_1 \in [x_1, \bar{x}_1], \ldots, x_n \in [x_n, \bar{x}_n] \}. \]

- **Reduction to optimization:** in the general case, \( y (\overline{y}) \):
  
  Minimize (Maximize) \( f(x_1, \ldots, x_n) \)

  where \( f \) is directly computable, under the constraints

  \[ x_1 \leq x_1 \leq \bar{x}_1, \ldots, x_n \leq x_n \leq \bar{x}_n. \]

- **Cosmological case:** \( f \) is not directly computable:

  \[ f(x_1, \ldots, x_n) \overset{\text{def}}{=} \text{argmin } F(x_1, \ldots, x_n, y_1, \ldots, y_m). \]
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 $(\tilde{x}_1, \ldots, \tilde{x}_n)$;
  - restrict ourselves only to linear terms:
    \[
    \Delta y = c_1 \cdot \Delta x_1 + \ldots + c_n \cdot \Delta x_n, \quad \text{where } c_i \overset{\text{def}}{=} \frac{\partial f}{\partial x_i}.
    \]
- **Interval case:** $|\Delta x_i| \leq \Delta_i$.
- **Result:** $\Delta \overset{\text{def}}{=} \max |\Delta y| = |c_1| \cdot \Delta_1 + \ldots + |c_n| \cdot \Delta_n$. 
18. Computations under Linearization: From Numerical Differentiation to Monte-Carlo Approach

- **Linearization**: \( \Delta y = \sum_{i=1}^{n} c_i \cdot \Delta x_i \), where \( c_i = \frac{\partial f}{\partial x_i} \).

- **Formulas**: \( \sigma^2 = \sum_{i=1}^{n} c_i^2 \cdot \sigma_i^2 \), \( \Delta = \sum_{i=1}^{n} |c_i| \cdot \Delta_i \).

- **Numerical differentiation**: \( n \) iterations, too long.

- **Monte-Carlo approach**: if \( \Delta x_i \) are Gaussian w/\( \sigma_i \), then \( \Delta y = \sum_{i=1}^{n} c_i \cdot \Delta x_i \) is also Gaussian, w/desired \( \sigma \).

- **Advantage**: # of iterations does not grow with \( n \).

- **Interval estimates**: if \( \Delta x_i \) are Cauchy, w/\( \rho_i(x) = \frac{\Delta_i}{\Delta_i^2 + x^2} \),
  then \( \Delta y = \sum_{i=1}^{n} c_i \cdot \Delta x_i \) is also Cauchy, w/desired \( \Delta \).

- Apply \( f \) to \( \tilde{x}_i \): \( \tilde{y} := f(\tilde{x}_1, \ldots, \tilde{x}_n) \);
- 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 \( K := \max |c_i^{(k)}| \) (to stay in linearized area);
  - simulate “actual values” \( \tilde{x}_i^{(k)} := \tilde{x}_i - \delta_i^{(k)} \), where \( \delta_i^{(k)} := \Delta_i \cdot c_i^{(k)}/K \);
  - simulate error of the indirect measurement:
    \[
    \delta^{(k)} := K \cdot \left( \tilde{y} - f \left( \tilde{x}_1^{(k)}, \ldots, \tilde{x}_n^{(k)} \right) \right);
    \]
- Solve the ML equation
  \[
  \sum_{k=1}^{N} \frac{1}{\left( 1 + \left( \frac{\delta^{(k)}}{\Delta} \right) \right)^2} = \frac{N}{2}
  \]
  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 $x_0 = 12.706$. 

- **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 $x \in [\underline{x}, \bar{x}]$, 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 $\sigma = \Delta \cdot \frac{\sqrt{n}}{\sqrt{3}}$.
- **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
  \[ [y, \bar{y}] = \{ f(x_1, \ldots, x_n) \mid x_1 \in [x_1, \bar{x}_1], \ldots, x_n \in [x_n, \bar{x}_n] \}. \]

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

- **Examples**: when \( x_1 \in \mathbf{x}_1 = [\underline{x}_1, \bar{x}_1] \) and \( x_2 \in \mathbf{x}_2 = [\underline{x}_2, \bar{x}_2] \), then:
  - The range \( \mathbf{x}_1 + \mathbf{x}_2 \) for \( x_1 + x_2 \) is \([\underline{x}_1 + \underline{x}_2, \bar{x}_1 + \bar{x}_2] \).
  - The range \( \mathbf{x}_1 - \mathbf{x}_2 \) for \( x_1 - x_2 \) is \([\underline{x}_1 - \underline{x}_2, \bar{x}_1 - \bar{x}_2] \).
  - The range \( \mathbf{x}_1 \cdot \mathbf{x}_2 \) for \( x_1 \cdot x_2 \) is \([\underline{y}, \bar{y}]\), where
    \[ \underline{y} = \min(\underline{x}_1 \cdot \underline{x}_2, \underline{x}_1 \cdot \bar{x}_2, \bar{x}_1 \cdot \underline{x}_2, \bar{x}_1 \cdot \bar{x}_2); \]
    \[ \bar{y} = \max(\underline{x}_1 \cdot \underline{x}_2, \underline{x}_1 \cdot \bar{x}_2, \bar{x}_1 \cdot \underline{x}_2, \bar{x}_1 \cdot \bar{x}_2). \]

- The range \( 1/\mathbf{x}_1 \) for \( 1/x_1 \) is \([1/\bar{x}_1, 1/\underline{x}_1] \) (if \( 0 \not\in \mathbf{x}_1 \)).
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*:
  - \( r_1 := [1, 2] - [2, 2] = [-1, 0]; \)
  - \( r_2 := [1, 2] + [2, 2] = [3, 4]; \)
  - \( r_3 := [-1, 0] \cdot [3, 4] = [-4, 0]. \)

- **Actual range:** \( f(x) = [-3, 0]. \)

- **Comment:** this is just a toy example, there are more efficient ways of computing an enclosure \( Y \supseteq y. \)
26. First Idea: Use of Monotonicity

- **Reminder:** for arithmetic, we had exact ranges.
- **Reason:** +, −, · 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
  \[
  f(x_1, \ldots, x_n) = [f(x_1, \ldots, x_n), f(x_1, \ldots, x_n)].
  \]
- **Similarly:** if \( f \uparrow \) for some \( x_i \) and \( f \downarrow \) for other \( x_j \) (−).
- **Fact:** \( f \uparrow \) in \( x_i \) if \( \frac{\partial f}{\partial x_i} \geq 0 \).
- **Checking monotonicity:** check that the range \([r_i, \bar{r}_i] \) of \( \frac{\partial f}{\partial x_i} \) on \( x_i \) has \( r_i \geq 0 \).
- **Differentiation:** by Automatic Differentiation (AD) tools.
- **Estimating ranges of** \( \frac{\partial f}{\partial x_i} \): straightforward interval comp.
27. Monotonicity: Example

- Idea: if the range \([r_i, \bar{r}_i]\) of each \(\frac{\partial f}{\partial x_i}\) on \(x_i\) has \(\bar{r}_i \geq 0\), then

\[
f(x_1, \ldots, x_n) = [f(x_1, \ldots, x_n), f(\bar{x}_1, \ldots, \bar{x}_n)].
\]

- Example: \(f(x) = (x - 2) \cdot (x + 2)\), \(x = [1, 2]\).

- Case \(n = 1\): if the range \([r, \bar{r}]\) of \(\frac{df}{dx}\) on \(x\) has \(r \geq 0\), then

\[
f(x) = [f(x), f(\bar{x})].
\]

- \(AD\): \(\frac{df}{dx} = 1 \cdot (x + 2) + (x - 2) \cdot 1 = 2x\).

- Checking: \([r, \bar{r}] = [2, 4]\), 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:**
  - \( r_1 := [1, 1] - [0, 1] = [0, 1]; \)
  - \( r_2 := [0, 1] \cdot [0, 1] = [0, 1]. \)

- **Actual range:** min, max of \( f \) at \( x, \overline{x} \), or when \( \frac{df}{dx} = 0. \)

- Here, \( \frac{df}{dx} = 1 - 2x = 0 \) for \( x = 0.5 \), so
  - compute \( f(0) = 0, \ f(0.5) = 0.25, \) and \( f(1) = 0. \)
  - \( y = \min(0, 0.25, 0) = 0, \ \overline{y} = \max(0, 0.25, 0) = 0.25. \)

- **Resulting range:** \( f(x) = [0, 0.25]. \)
29. **Second Idea: Centered Form**

- **Main idea:** Intermediate Value Theorem
  \[
  f(x_1, \ldots, x_n) = f(\tilde{x}_1, \ldots, \tilde{x}_n) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\chi) \cdot (x_i - \tilde{x}_i)
  \]
  for some $\chi_i \in x_i$.

- **Corollary:** $f(x_1, \ldots, x_n) \in Y$, where
  \[
  Y = \tilde{y} + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n) \cdot [-\Delta_i, \Delta_i].
  \]

- **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:**
  \[ Y = f(\tilde{x}_1, \ldots, \tilde{x}_n) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n) \cdot [-\Delta_i, \Delta_i]. \]

- **Example:** \( f(x) = x \cdot (1 - x), \ x = [0, 1]. \)

- **Here,** \( x = [\tilde{x} - \Delta, \tilde{x} + \Delta], \) with \( \tilde{x} = 0.5 \) and \( \Delta = 0.5. \)

- **Case \( n = 1: \)** \( Y = f(\tilde{x}) + \frac{df}{dx}(x) \cdot [-\Delta, \Delta]. \)

- **AD:** \( \frac{df}{dx} = 1 \cdot (1 - x) + x \cdot (-1) = 1 - 2x. \)

- **Estimation:** we have \( \frac{df}{dx}(x) = 1 - 2 \cdot [0, 1] = [-1, 1]. \)

- **Result:** \( Y = 0.5 \cdot (1 - 0.5) + [-1, 1] \cdot [-0.5, 0.5] = 0.25 + [-0.5, 0.5] = [-0.25, 0.75]. \)

- **Comparison:** actual range \([0, 0.25]\), straightforward \([0, 1]\).
31. Third Idea: Bisection

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

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

• **Idea:** if the intervals are too wide, we:
  – split one of them in half ($\Delta_i^2 \rightarrow \Delta_i^2/4$); and
  – take the union of the resulting ranges.

• **Example:** $f(x) = x \cdot (1 - x)$, where $x \in x = [0, 1]$.

• **Split:** take $x' = [0, 0.5]$ and $x'' = [0.5, 1]$.

• **1st range:** $1 - 2 \cdot x = 1 - 2 \cdot [0, 0.5] = [0, 1]$, so $f \uparrow$ and $f(x') = [f(0), f(0.5)] = [0, 0.25]$.

• **2nd range:** $1 - 2 \cdot x = 1 - 2 \cdot [0.5, 1] = [-1, 0]$, so $f \downarrow$ and $f(x'') = [f(1), f(0.5)] = [0, 0.25]$.

• **Result:** $f(x') \cup f(x'') = [0, 0.25] -$ 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 $\Delta x_i = \tilde{x}_i - x_i$ (at least its linear terms).
- **Implementation:**

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

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

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

- **Description:** $a_0 = \tilde{x}_i$, $a_i = -1$, $a_j = 0$ for $j \neq i$, and $[a, \bar{a}] = [0, 0]$. 
33. **Affine Arithmetic: Operations**

- **Representation:** \( a = a_0 + \sum_{i=1}^{n} a_i \cdot \Delta x_i + [a, \bar{a}] \).

- **Input:** \( a = a_0 + \sum_{i=1}^{n} a_i \cdot \Delta x_i + a \) and \( b = b_0 + \sum_{i=1}^{n} b_i \cdot \Delta x_i + b \).

- **Operations:** \( c = a \otimes b \).

- **Addition:** \( c_0 = a_0 + b_0, c_i = a_i + b_i, c = a + b \).

- **Subtraction:** \( c_0 = a_0 - b_0, c_i = a_i - b_i, c = a - b \).

- **Multiplication:** \( c_0 = a_0 \cdot b_0, c_i = a_0 \cdot b_i + b_0 \cdot a_i, c = a_0 \cdot b + b_0 \cdot a + \sum_{i \neq j} a_i \cdot b_j \cdot [-\Delta_i, \Delta_i] \cdot [-\Delta_j, \Delta_j] + \sum_{i} a_i \cdot b_i \cdot [-\Delta_i, \Delta_i]^2 + \left( \sum_{i} a_i \cdot [-\Delta_i, \Delta_i] \right) \cdot b + \left( \sum_{i} b_i \cdot [-\Delta_i, \Delta_i] \right) \cdot a + a \cdot b \).
34. **Affine Arithmetic: Example**

- *Example:* \( f(x) = x \cdot (1 - x), \ x \in [0, 1] \).
- Here, \( n = 1, \tilde{x} = 0.5, \) 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 \( x = 0.5 - \Delta x + [0, 0]; \)
  - \( r_1 := 1 - (0.5 - \Delta) = 0.5 + \Delta x; \)
  - \( r_2 := (0.5 - \Delta x) \cdot (0.5 + \Delta x), \) i.e.,
    \[
r_2 = 0.25 + 0 \cdot \Delta x - [-\Delta, \Delta]^2 = 0.25 + [-\Delta^2, 0].
    \]
- *Resulting range:* \( y = 0.25 + [-0.25, 0] = [0, 0.25]. \)
- *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 \( a \) and \( \overline{a} \) of \([a, \overline{a}]\).
  - *Affine arithmetic:* for each \( a \), we compute \( n + 3 \) values:
    \[
    a_0, a_1, \ldots, a_n, a, \overline{a}.
    \]

- **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: \( a_i \in [a_i, \bar{a}_i] \);
- **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 \( f_j([a_1, \bar{a}_1], \ldots, [a_n, \bar{a}_n]) \).
- **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:* \( x = [0, 1], \ r = (-\infty, \infty) \).

1. \( r = [0, 1]^2 = [0, 1] \), so \( r_{\text{new}} = (-\infty, \infty) \cap [0, 1] = [0, 1] \).
2. \( x_{\text{new}} = \sqrt{[0, 1]} \cap [0, 1] = [0, 1] \) – no change.
3. \( r_{\text{new}} = ([0, 1] - 0.5) \cap [0, 1] = [-0.5, 0.5] \cap [0, 1] = [0, 0.5] \).
4. \( x_{\text{new}} = ([0, 0.5] + 0.5) \cap [0, 1] = [0.5, 1] \cap [0, 1] = [0.5, 1] \).

- 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:** \( x = [0, 1], r = (-\infty, \infty) \).
- **Problem:** after Rule 1, we’re stuck with \( x = r = [0, 1] \).
- **Solution:** bisect \( x = [0, 1] \) into \([0, 0.5]\) and \([0.5, 1]\).
- **For 1st subinterval:**
  - Rule 1 leads to \( r_{\text{new}} = [0, 0.5]^2 \cap [0, 0.5] = [0, 0.25] \);
  - Rule 4 leads to \( x_{\text{new}} = [0, 0.25] \);
  - Rule 1 leads to \( r_{\text{new}} = [0, 0.25]^2 = [0, 0.0625] \);
  - Rule 4 leads to \( x_{\text{new}} = [0, 0.0625] \); 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

  $$ g(y_1, \ldots, y_m, a_1, \ldots, a_m) \rightarrow \text{max}.$$  

- **We know:** $a_i$ with interval uncertainty: $a_i \in [a_i, \bar{a}_i]$;

- **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 $f_j([a_1, \bar{a}_1], \ldots, [a_n, \bar{a}_n])$.

- **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 $\max_{x \in \mathbf{b}} g(x) < g(x')$ 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 $g(\tilde{\mathbf{b}})$ for each $\mathbf{b}$; the largest is $0.45 \cdot 0.55 = 0.2475$.
- Compute $G(\mathbf{b}) = g(\tilde{\mathbf{b}}) + (1 - 2 \cdot \mathbf{b}) \cdot [-\Delta, \Delta]$.
- 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 $x = 0.5$. 