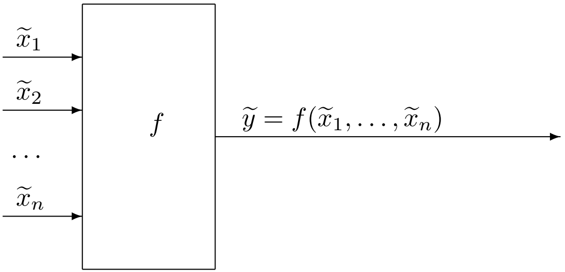
How to Take into Account Dependence Between the Inputs: From Interval Computations to Constraint-Related Set Computations

Vladik Kreinovich
Computer Science, University of Texas
El Paso, TX 79968, USA, vladik@utep.edu

The talk is based on our joint work with Martine Ceberio, Scott Ferson, Gang Xiang, Adrian Murguia, and Jorge Santillan.

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

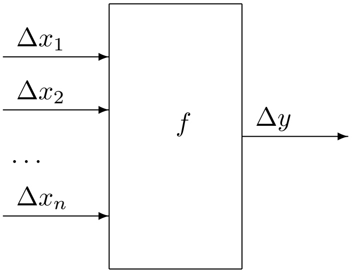


- *Problem:* measurements are never 100% accurate: $\widetilde{x}_i \neq x_i$ ($\Delta x_i \neq 0$) hence

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

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

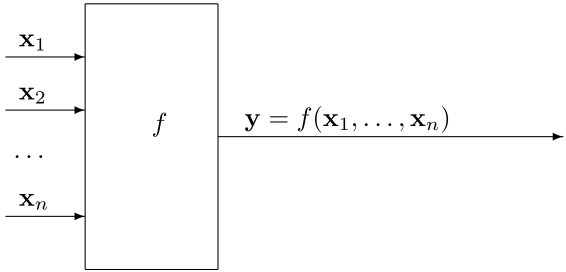
2. Probabilistic and Interval Uncertainty



- *Traditional approach:* we know probability distribution for Δx_i (usually Gaussian).
- *Where it comes from:* calibration using standard MI.
- *Problem:* sometimes we do not know the distribution because no "standard" (more accurate) MI is available. Cases:
  - fundamental science
  - manufacturing
- *Solution:* we know upper bounds Δ_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 $y = f(x_1, \dots, x_n)$ that transforms n real numbers x_i into a number y;
  - n intervals $\mathbf{x}_i = [\underline{x}_i, \overline{x}_i]$.
- *Compute:* the corresponding range of y:

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

- *Fact:* even for quadratic f, the problem of computing the exact range $\mathbf{y}$ is NP-hard.
- *Practical challenges:*
  - find classes of problems for which efficient algorithms are possible; and
  - for problems outside these classes, find efficient techniques for *approximating* uncertainty of y.

4. Why Not 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 a uniform distribution on $[\underline{x}, \overline{x}]$.
- *Example – multiple variables:* different variables are independently distributed.
- *Conclusion:* if $\Delta y = \Delta x_1 + \ldots + \Delta x_n$, with $\Delta x_i \in [-\Delta_i, \Delta_i]$, then due to Central Limit Theorem, Δy is almost normal, with $\sigma = \dfrac{1}{\sqrt{3}} \cdot \sqrt{\sum_{i=1}^{n} \Delta_i^2}$.
- *Why this may be inadequate:* when $\Delta_i = \Delta$, we get $\Delta \sim \sqrt{n}$, but due to correlation, it is possible that $\Delta = n \cdot \Delta_i \sim n \gg \sqrt{n}$.
- *Conclusion:* using a single distribution can be very misleading, especially if we want guaranteed results – e.g., in high-risk application areas such as space exploration or nuclear engineering.

5. General Approach: Interval-Type Step-by-Step Techniques

- *Problem:* it is difficult to compute the range $\mathbf{y}$.
- *Solution:* compute an enclosure $\mathbf{Y}$ such that $\mathbf{y} \subseteq \mathbf{Y}$.
- *Interval arithmetic:* for arithmetic operations $f(x_1, x_2)$, 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 $0 \notin \mathbf{x}_1$).

6. Interval Approach: Example

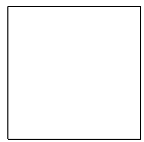
- *Example:* $f(x) = (x-2)\cdot(x+2)$, $x \in [1,2]$.
- How will the computer compute it?
  - $r_1 := x - 2$;
  - $r_2 := x + 2$;
  - $r_3 := r_1 \cdot r_2$.
- *Main idea:* do the same operations, but with *intervals* instead of *numbers*:
  - $\mathbf{r}_1 := [1,2] - [2,2] = [-1,0]$;
  - $\mathbf{r}_2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

7. Interval Computations: Analysis

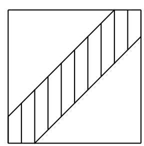
- *Computation time:* ≤ 4 arithmetic operations per original operation, so $O(T)$, where T is the running time of the original algorithm.
- *Result:* often, enclosure $\mathbf{Y} \supseteq \mathbf{y}$ with excess width.
- *Reason:* there is a relation between intermediate results, and we ignore it in straightforward interval computations.
- *Alternative:* we can compute the exact range: e.g., Tarksi algorithm for algebraic f.
- *Computation time:* can be exponential $O(2^T)$.
- *Summarizing:* we have two algorithms:
  - a fast and efficient $O(T)$ algorithm which often has large excess width;
  - a slow and inefficient (often non-feasible) algorithm with no excess width.
- *It is desirable:* to develop a sequence of feasible algorithms with:
  - longer and longer computation time and
  - smaller and smaller excess width.

8. Interval Computations: Limitations

- *Traditional interval computations:*
  - we know the intervals $\mathbf{x}_i$ of possible values of different parameters x_i, and
  - we assume that an arbitrary combination of these values is possible.
- *In geometric terms:* the set of possible combinations $x = (x_1, \ldots, x_n)$ is a box $\mathbf{x} = \mathbf{x}_1 \times \ldots \times \mathbf{x}_n$.



- *In practice:* we also know additional restrictions on the possible combinations of x_i.
- *Example:* in geosciences, in addition to intervals for velocities v_i at different points, we know that $|v_i - v_j| \le \Delta$ for neighboring points:



- *Example:* in nuclear engineering, experts often state that combinations of extreme values are impossible, we have an ellipsoid, not a box.

Title Page
Page 9 of 24
Go Back
Full Screen

9. Similar Situation: Statistics

- Ideally, we should take into account dependence between all the variables.
- In the first approximation, it is often reasonable to consider them independent.
- In the next approximation, we consider pairwise dependencies.
- To get an even better picture, we can consider dependencies between triples, etc.
- As a result, we get a sequence of methods which:
  - require more and more time
  - but at the same time lead to more and more accurate results.

10. Let Us Use a Similar Idea for Interval Uncertainty

- Ideally, we should take the box $\mathbf{x}_1 \times \ldots \times \mathbf{x}_n$ (or appropriate subset of the box), divide it into smaller boxes, estimate the range over each small box, and combine the results.
- This requires C^n subboxes – i.e., exponential time.
- In straightforward interval computations, we consider only intervals of possible values of x_i.
- A natural next approximation is when we consider:
  - sets $\mathbf{x}_i$ of possible values of x_i, and also
  - sets $\mathbf{x}_{ij}$ of possible pairs (x_i, x_j).
- Third approximation: we also consider possible sets of triples, etc.
- As a result, we hope to get a sequence of methods which:
  - require more and more time
  - but at the same time lead to more and more accurate results.

11. How to Represent Sets

- *First idea:* do it in a way cumulative probability distributions (cdf) are represented in RiskCalc package: by discretization.
- In RiskCalc, we:
  - divide the interval $[0, 1]$ of possible values of probability into, say, 10 subintervals of equal width and
  - represent cdf $F(x)$ by 10 values $x_1, \ldots, x_{10}$ at which $F(x_i) = i/10$.
- Similarly, we:
  - divide the box $\mathbf{x}_i \times \mathbf{x}_j$ into, say, 10×10 subboxes and
  - describe the set $\mathbf{x}_{ij}$ by listing all subboxes which contain possible pairs.
- *Comment:*
  - A more efficient idea is to represent this set by a covering paving – in the style of Jaulin et al. – i.e., consider boxes of different sizes starting with larger ones and only decrease the size when necessary.
  - It is also possible (and often efficient) to use ellipsoids.

12. How to Propagate This Uncertainty: A Problem

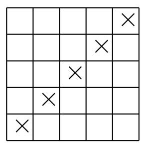
- *In the beginning:* we know the intervals $\mathbf{r}_1, \ldots, \mathbf{r}_n$ corresponding to the input variables $r_i = x_i$, and we know the sets $\mathbf{r}_{ij}$ for i, j from 1 to n.
- *Question:* propagate this information through an intermediate computation step, a step of computing $r_k = r_a * r_b$ for some arithmetic operation $*$ and for previous results r_a and r_b $(a, b < k)$.
- By the time we come to this step, we know the intervals $\mathbf{r}_i$ and the sets $\mathbf{r}_{ij}$ for $i, j < k$.
- We want to find the interval $\mathbf{r}_k$ for x_k, and the sets $\mathbf{r}_{ik}$ for $i < k$.

13. How to Propagate This Uncertainty: An Idea

- First step: computing $\mathbf{r}_k$:
  - In our representation, the set $\mathbf{x}_{ab}$ consists of small 2-D boxes $\mathbf{X}_a \times \mathbf{X}_b$.
  - For each small box $\mathbf{X}_a \times \mathbf{X}_b$, we use interval arithmetic to compute the range $\mathbf{X}_a * \mathbf{X}_b$ of the value $r_a * r_b$ over this box.
  - Then, we take the union (interval hull) of all these ranges.
- Second step: computing $\mathbf{r}_{ik}$:
  - We consider the sets $\mathbf{r}_{ab}$, $\mathbf{r}_{ai}$, and $\mathbf{r}_{bi}$.
  - For each small box $\mathbf{R}_a \times \mathbf{R}_b$ from $\mathbf{r}_{ab}$, we:
    * consider all subintervals $\mathbf{R}_i$ for which $\mathbf{R}_a \times \mathbf{R}_i$ is in $\mathbf{r}_{ai}$ and $\mathbf{R}_b \times \mathbf{R}_i$ is in $\mathbf{r}_{bi}$, and then
    * we add $(\mathbf{R}_a * \mathbf{R}_b) \times \mathbf{R}_i$ to the set $\mathbf{r}_{ki}$.
  - *To be more precise*:
    * since the interval $\mathbf{R}_a * \mathbf{R}_b$ may not have bounds of the type $p/10$,
    * we may need to expand it to get within bounds of the desired type.
- We repeat these computations step by step until we get the desired estimate for the range of the final result of the computations.

14. First Example: Computing the Range of $x - x$

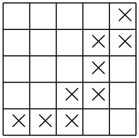
- *Problem:*
  - for $f(x) = x - x$ on $[0, 1]$, the actual range is $[0, 0]$;
  - straightforward interval computations lead to an enclosure $[0, 1] - [0, 1] = [-1, 1]$.
- In straightforward interval computations:
  - we have $r_1 = x$ with interval $\mathbf{r}_1 = [0, 1]$;
  - we have $r_2 = x$ with interval $\mathbf{x}_2 = [0, 1]$;
  - the variables r_1 and r_2 are dependent, but we ignore this dependence.
- In the new approach: we have $\mathbf{r}_1 = \mathbf{r}_2 = [0, 1]$, and we also have $\mathbf{r}_{12}$:



- For each small box, we have $[-0.2, 0.2]$, so the union is $[-0.2, 0.2]$.
- If we divide into more pieces, we get close to 0.

15. Second Example: Computing the Range of $x - x^2$

- In straightforward interval computations:
  - we have $r_1 = x$ with interval $\mathbf{r}_1 = [0, 1]$;
  - we have $r_2 = x^2$ with interval $\mathbf{x}_2 = [0, 1]$;
  - the variables r_1 and r_2 are dependent, but we ignore this dependence and estimate $\mathbf{r}_3$ as $[0, 1] - [0, 1] = [-1, 1]$.
- In the new approach: we have $\mathbf{r}_1 = \mathbf{r}_2 = [0, 1]$, and we also have $\mathbf{r}_{12}$:
  - the union of $\mathbf{R}_1^2$ is $[0, 1]$, so we have $[0, 0.2]$, $[0.2, 0.4]$, etc.;
  - for $\mathbf{R}_1 = [0, 0.2]$, we have $\mathbf{R}_1^2 = [0, 0.04]$, so only $[0, 0.2]$ is affected;
  - for $\mathbf{R}_1 = [0.2, 0.4]$, we have $\mathbf{R}_1^2 = [0.04, 0.16]$, so only $[0, 0.2]$ is affected;
  - for $\mathbf{R}_1 = [0.4, 0.6]$, we have $\mathbf{R}_1^2 = [0.16, 0.25]$, so $[0, 0.2]$ and $[0.2, 0.4]$ are affected, etc.



- For each possible pair of small boxes $\mathbf{R}_1 \times \mathbf{R}_2$, we have $\mathbf{R}_1 - \mathbf{R}_2 = [-0.2, 0.2]$, $[0, 0.4]$ and $[0.2, 0.6]$, so the union of $\mathbf{R}_1 - \mathbf{R}_2$ is $\mathbf{r}_3 = [-0.2, 0.6]$.
- If we divide into more pieces, we get closer to $[0, 0.25]$.

16. How to Compute $\mathbf{r}_{ik}$

- Since $\mathbf{r}_3 = [-0.2, 0.6]$, we divide this range into 5 subintervals $[-0.2, -0.04]$, $[-0.04, 0.12]$, $[0.12, 0.28]$, $[0.28, 0.44]$, $[0.44, 0.6]$.
- For $\mathbf{R}_1 = [0, 0.2]$, the only possible $\mathbf{R}_2$ is $[0, 0.2]$, so $\mathbf{R}_1 - \mathbf{R}_2 = [-0.2, 0.2]$. This covers $[-0.2, -0.04]$ and $[-0.04, 0.12]$.
- For $\mathbf{R}_1 = [0.2, 0.4]$, the only possible $\mathbf{R}_2$ is $[0, 0.2]$, so $\mathbf{R}_1 - \mathbf{R}_2 = [0, 0.4]$. This covers $[-0.04, 0.12]$, $[0.12, 0.28]$, and $[0.28, 0.44]$.
- For $\mathbf{R}_1 = [0.4, 0.6]$, we have two possible $\mathbf{R}_2$:
  - for $\mathbf{R}_2 = [0, 0.2]$, we have $\mathbf{R}_1 - \mathbf{R}_2 = [0.2, 0.6]$; this covers $[0.12, 0.28]$, $[0.28, 0.44]$, and $[0.44, 0.6]$;
  - for $\mathbf{R}_2 = [0.2, 0.4]$, we have $\mathbf{R}_1 - \mathbf{R}_2 = [0, 0.4]$; this covers $[-0.04, 0.12]$, $[0.12, 0.28]$, and $[0.28, 0.44]$.
- For $\mathbf{R}_1 = [0.6, 0.8]$, we have $\mathbf{R}_1^2 = [0.36, 0.64]$, so three possible $\mathbf{R}_2$: $[0.2, 0.4]$, $[0.4, 0.6]$, and $[0.6, 0.8]$, to the total of $[0.2, 0.8]$. Here, $[0.6, 0.8] - [0.2, 0.8] = [-0.2, 0.6]$, so all 5 subintervals are affected.
- For $\mathbf{R}_1 = [0.8, 1.0]$, we have $\mathbf{R}_1^2 = [0.64, 1.0]$, so two possible $\mathbf{R}_2$: $[0.6, 0.8]$ and $[0.8, 1.0]$, to the total of $[0.6, 1.0]$. Here, $[0.8, 1.0] - [0.6, 1.0] = [-0.2, 0.4]$, so the first 4 subintervals are affected.

		×	×	
	×	×	×	×
	×	×	×	×
×	×	×	×	×
×		×	×	×

Title Page
◀◀ ▶▶
◀ ▶

Go Back
Full Screen

17. Distributivity: $a \cdot (b + c)$ vs. $a \cdot b + a \cdot c$

- *Problem:* compute the range of $x_1 \cdot (x_2 + x_3) = x_1 \cdot x_2 + x_1 \cdot x_3$ when $x_1 \in \mathbf{x}_1 = [0, 1]$, $\mathbf{x}_2 = [1, 1]$, and $\mathbf{x}_3 = [-1, -1]$.
- *Actual range:* we have $x_1 \cdot (x_2 + x_3) = 0$ for all possible x_i hence the actual range is $[0, 0]$.
- *Straightforward interval computations:*
  - for $\mathbf{x}_1 \cdot (\mathbf{x}_2 + \mathbf{x}_3)$, we get $[0, 1] \cdot [0, 0] = [0, 0]$;
  - for $\mathbf{x}_1 \cdot \mathbf{x}_2 + \mathbf{x}_1 \cdot \mathbf{x}_3$, we get $[0, 1] \cdot 1 + [0, 1] \cdot (-1) = [0, 1] + [-1, 0] = [-1, 1]$, i.e., excess width.
- *Reason*: we have $r_4 = x_1 \cdot x_2$, $r_5 = x_1 \cdot x_3$, but we ignore the dependence between r_4 and r_5.

18. Distributivity: New Approach

- *Reminder:* $r_4 = r_1 \cdot r_2$, $r_5 = r_1 \cdot r_3$, $r_6 = r_4 + r_5$, $\mathbf{r}_1 = [0, 1]$, $r_2 = 1$, $r_3 = -1$.
- When we get $r_4 = r_1 \cdot r_2$, we compute the ranges $\mathbf{r}_{14}$, $\mathbf{r}_{24}$, and $\mathbf{r}_{34}$; the only non-trivial range is $\mathbf{r}_{14}$:

				×
			×	
		×		
	×			
×				

- For $r_5 = r_1 \cdot r_3$, we get $\mathbf{r}_5 = [-1, 0]$.
- To compute the range $\mathbf{r}_{45}$, for each possible box $\mathbf{R}_1 \times \mathbf{R}_3$, we:
  - consider all boxes $\mathbf{R}_4$ for which $\mathbf{R}_4 \times \mathbf{R}_1$ is possible and $\mathbf{R}_4 \times \mathbf{R}_3$ is possible;
  - add $\mathbf{R}_4 \times (\mathbf{R}_1 \cdot \mathbf{R}_3)$ to the set $\mathbf{r}_{45}$.
- *Result:*

×				
	×			
		×		
			×	
				×

- Hence, for $r_6 = r_4 + r_5$, we get $[-0.2, 0.2]$.
- If we divide into more pieces, we get the enclosure closer to 0.

19. Toy Example with Prior Dependence

- *Case study:* find the range of $r_1 - r_2$ when $\mathbf{r}_1 = [0,1]$, $\mathbf{r}_2 = [0,1]$, and $|r_1 - r_2| \le 0.1$.
- *Actual range:* $[-0.1, 0.1]$.
- *Straightforward interval computations:* $[0,1] - [0,1] = [-1,1]$.
- *New approach:*
  - First, we describe the constraint in terms of subboxes:

			×	×
		×	×	×
	×	×	×	
×	×	×		
×	×			

  - Next, we compute $\mathbf{R}_1 - \mathbf{R}_2$ for all possible pairs and take the union.
- *Result:* $[-0.6, 0.6]$.
- If we divide into more pieces, we get the enclosure closer to $[-0.1, 0.1]$.

Title Page
◀◀ ▶▶
◀ ▶
Page 20 of 24
Go Back
Full Screen

20. Computation Time

- *Straightforward interval computations:*
  - we need to compute T intervals $\mathbf{r}_i$, $i = 1, \ldots, T$;
  - so, it requires $O(T)$ steps.
- *New idea:*
  - we need to compute T^2 sets $\mathbf{r}_{ij}$, $i, j = 1, \ldots, T$;
  - so, it requires $O(T^2)$ steps.
- *Conclusion:*
  - the new method is longer than for straightforward interval computations, but
  - it is still feasible.

21. What Next?

- *Known fact:* the range estimation problem is, in general, NP-hard (even without any dependency between the inputs).
- *Corollary:* our quadratic time method cannot completely avoid excess width.
- To get better estimates, in addition to sets of pairs, we can also consider sets of *triples* $\mathbf{r}_{ijk}$.
- This will be a T^3 time version of our approach.
- We can also go to *quadruples* etc.
- Similar ideas can be applied to the case when we also have partial information about probabilities.

22. Probabilistic Case: In Brief

- *Traditionally:* expert systems use technique similar to straightforward interval computations.
- We parse F and replace each computation step with corresponding probability operation.
- *Problem:* at each step, we ignore the dependence between the intermediate results F_j.
- *Result:* intervals are too wide (and numerical estimates off).
- *Example:* the estimate for $P(A \vee \neg A)$ is not 1.
- *Solution:* similarly to the above algorithm, besides $P(F_j)$, we also compute $P(F_j \,\&\, F_i)$ (or $P(F_{j_1} \,\&\, \ldots \,\&\, F_{j_k})$).
- On each step, use all combinations of l such probabilities to get new estimates.
- *Result:* e.g., $P(A \vee \neg A)$ is estimated as 1.

23. Acknowledgments

This work was supported in part:

- by NASA under cooperative agreement NCC5-209,
- by NSF grant EAR-0225670,
- by NIH grant 3T34GM008048-20S1,
- by Army Research Lab grant DATM-05-02-C-0046,
- by Star Award from the University of Texas System,
- by Texas Department of Transportation grant No. 0-5453, and
- by the workshop organizers.