Towards Interval Techniques for Model Validation

Jaime Nava and Vladik Kreinovich

Department of Computer Science
University of Texas at El Paso
El Paso, Texas 79968, USA
jenava@miners.utep.edu
vladik@utep.edu

1. Measurement Under Exact Model: Case Study

- *Case study:* find the unit vector $\vec e_k$ in the direction to a distant astronomical radio-source.
- *Most accurate method:* Very Large Baseline Interferometry (VLBI).
- *How it works:* we measure the time delay $\tau_{i,j,k}$ between the signal observed by antennas i and j:

$$\tau_{i,j,k} = c^{-1} \cdot (\vec b_i - \vec b_j) \cdot \vec e_k + \Delta t_i - \Delta t_j, \text{ where}$$

  - $\vec b_i$ is the location of the i-th antenna, and
  - Δt_i is the bias of the clock on the i-th antenna.
- *Ideal case:* if we knew $\vec b_i$ and Δt_i with high accuracy, we could easily find $\vec e_k$.
- *In practice:* we only know $\vec b_i$ and Δt_i approximately, with much lower accuracy than $\tau_{i,j,k}$.

2. Measurement Under Exact Model: General Case

- *In general:* to find the desired values $x_1, \ldots, x_m$ ($\vec{e}_k$), we use the measured values $z_1, \ldots, z_n$ ($\tau_{i,j,k}$).
- *Fact:* the values z_i depend not only on x_i, they also depend on
  - the values $s_1, \ldots, s_p$ of known auxiliary quantities (e.g., time of the experiment), and
  - the values $y_1, \ldots, y_q$ of the auxiliary quantities which are only approximately known ($\vec{b}_j$ and Δt_i).
- *In VLBI:* we know the exact dependence $z = f(x, s, y)$.
- *Idea:* we can determine both x_i and y_j if we measure each of several (N) different objects x in
  - several (Q) different settings y and
  - several (P) different settings s.

3. Measurement Under Exact Model: General Case (cont-d)

- *Reminder:* we measure n values z_i for each of N objects under Q different settings y and P different settings s.
- *After these measurements:* we get $n \cdot N \cdot P \cdot Q$ measurement results.
- *Thus:* we have $n \cdot N \cdot P \cdot Q$ equations $z = f(x, s, y)$.
- *We want:* to determine $N \cdot m + Q \cdot q$ unknown (to be more exact, approximately known) values x_i and y_j.
- *Fact:* when N and Q are large, the number of equations exceeds the number of unknowns.
- *Conclusion:* we can find the x values for all the objects and and y values for all the settings.

4. Linearizable Case

- Usually, we know the approximate values $\widetilde{x}$ and $\widetilde{y}$ of the quantities x and y with accuracies Δ_x and Δ_y.
- In effect, we know the intervals $[\widetilde{x}-\Delta_x,\widetilde{x}+\Delta_x]$ and $[\widetilde{y}-\Delta_y,\widetilde{y}+\Delta_y]$.
- These intervals contain the actual values x and y.
- In this case,
  - we can linearize the system $z = f(x, s, y)$ and
  - solve the resulting system of linear equations in terms of $\Delta x \stackrel{\text{def}}{=} \widetilde{x} - x$ and $\Delta y \stackrel{\text{def}}{=} \widetilde{y} - y$:

$$\widetilde{z} = f(\widetilde{x}, s, \widetilde{y}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \cdot \Delta x_i + \sum_{j=1}^{q} \frac{\partial f}{\partial y_j} \cdot \Delta y_j.$$

5. Approximate Model: Need for Model Validation

- *In practice:* we often only have an *approximate* model $f(x, s, y)$ for the dependence of z on x, s, and y.
- *In such situations:* it is desirable to *validate* this model.
- *Specifically:* we want to supplement $f(x, s, y)$ with a guaranteed accuracy $\varepsilon > 0$ of this approximate model.
- *Definition:* a model is ε-correct if for every k, $\exists x_i \in [\widetilde{x}_i - \Delta_x, \widetilde{x}_i + \Delta_x]$ and $\exists y_j \in [\widetilde{y}_j - \Delta_y, \widetilde{y}_j + \Delta_y]$ for which

$$\widetilde{z}_k \in [f(x, s, y) - \varepsilon, f(x, s, y) + \varepsilon].$$

- *Objective:* to find the smallest possible $\varepsilon > 0$ for which the given model is ε-correct.

6. How to Solve This Problem: Outline

- *Fact:* in the linearizable case, the above conditions of the type $a \in \mathbf{a}$ become linear inequalities.
- *Hence:* the problem of finding the smallest such ε becomes a linear programming problem.
- *In this talk:* we describe the application of the resulting techniques to a benchmark thermal problem.
- We need to find the accuracy of the approximate model describing the temperature inside the camera.
- This constitutes a model with several approximately known parameters.
- The problem was presented at the 2006 Sandia Validation Challenge Workshop.

7. The Thermal Challenge Problem: In Brief

- *Situation:* we need to analyze temperature response $T(x,t)$ of a safety-critical device to a heat flux.
- *Details:* a slab of metal (or other material) of thickness $L = 1.90$ cm is exposed to a heat flux $q = 3500$ W/m^2.
- *We know:*
  - thermal conductivity k,
  - volumetric heat capacity of the material ρC_p,
  - the initial temperature $T_i = 25$ C,
  - an approximate model:

$$T(x,t) = T_i + \frac{q \cdot L}{k} \cdot \left[\frac{(k/\rho C_p) \cdot t}{L^2} + \frac{1}{3} - \frac{x}{L} + \frac{1}{2} \cdot \left(\frac{x}{L}\right)^2 - \frac{2}{\pi^2} \cdot \sum_{n=1}^{6} \frac{1}{n^2} \cdot \exp\left(-n^2 \cdot \pi^2 \cdot \frac{(k/\rho C_p) \cdot t}{L^2}\right) \cdot \cos\left(n \cdot \pi \cdot \frac{x}{L}\right)\right].$$

8. Thermal Challenge Problem: Seemingly Natural Approach

- *Question:* we do not know how accurate is the approximate model.
- *Task:* estimate the accuracy ε of the model.
- *Ideal case:* compare the model with measurement results.
- *Difficulty:* it is very difficult to measure temperatures for the desired flux (corr. to a strong fire).
- *Solution:* we perform lab experiments with smaller flux values.
- *Result:* the accuracy of the model is low: $\approx 25^\circ$.
- *Why this is not perfect:* this makes predictions inaccurate.

9. Thermal Challenge Problem: New Idea

- *Reminder:*
  - we estimate the model's accuracy ε based on the results of the lab experiments;
  - the resulting accuracy is low $\varepsilon \approx 25^\circ$.
- *Observation:* in our computations, we assumed that the given values of k and ρC_p are exact.
- *In practice:*
  - these values k and ρC_p are only approximately known;
  - they may change from sample to sample.
- *Idea:* do not assume any value of these quantities, just assume that for each sample, there are some values.

10. Thermal Problem As Part of General Framework

- *How this problem fits our general framework:*
  - measured quantity z: temperature $z = T$;
  - known auxiliary quantity: time $s_1 = t$;
  - unknown auxiliary quantities: $y_1 = k$; $y_2 = \rho C_p$;
  - we know the $\approx$ dependence $z_1 \approx f(s_1, y_1, y_2)$.
- *Additional complexity:* the model is only approximate:

  $$|z^{(k)} - f(s_1^{(k)}, y_1^{(k)}, y_2^{(k)})| \le \varepsilon$$

  for some (unknown) accuracy ε.
- *Natural idea:* once we know $z^{(k)} = T$ for different moments $t = s^{(k)}$, find y_1, y_2 for which $\varepsilon \to \min$, where:

  $$|z^{(k)} - f(s_1^{(k)}, y_1, y_2)| \le \varepsilon.$$

11. How to Implement the Above Idea: Linearizable Case

- *Reminder:* find y_1 and y_2 with $\varepsilon \to \min$, where
$$|z^{(k)} - f(s_1^{(k)}, y_1, y_2)| \le \varepsilon.$$
- *Linearizable case:*
  - we know the approximate values $y_1^{(0)}$ and $y_2^{(0)}$;
  - the differences $\Delta y_i \stackrel{\text{def}}{=} y_i - y_i^{(0)}$ are small;
  - hence quadratic terms can be ignored.
- *Simplification:* we get a linear programming problem
$$\varepsilon \to \min$$
under the constraints
$$-\varepsilon \le z^{(k)} - f(s^{(k)}, y_1^{(0)}, y_2^{(0)}) - \frac{\partial f}{\partial y_1} \cdot \Delta y_1 - \frac{\partial f}{\partial y_2} \cdot \Delta y_2 \le \varepsilon.$$

12. Linearizable Case (cont-d)

- *Problem (reminder):* find y_1 and y_2 with $\varepsilon \to \min$:
$$|z^{(k)} - f(s_1^{(k)}, y_1, y_2)| \le \varepsilon.$$
- *Solution (reminder):* find the values y_1, y_2, and ε that solve the linearized version of this problem.
- *Difficulty:*
  - we are solving the approximate version of the problem;
  - hence, the resulting value ε is only approximate;
  - the actual accuracy may be larger or smaller than ε;
  - thus, the bound ε is not guaranteed.
- *In practice:* in many critical applications, we want *guaranteed* bounds on ε – e.g., to guarantee fire safety.
- *Solution:* the value $\widetilde{\varepsilon} = \max\limits_k |z^{(k)} - f(s_1^{(k)}, y_1, y_2)|$ is a guaranteed estimate.

13. General Case: Analysis

- *Problem (reminder):* find y_1 and y_2 with $\varepsilon \to \min$:

$$|z^{(k)} - f(s_1^{(k)}, y_1, y_2)| \le \varepsilon.$$

- *Solution (reminder):*
  - find the values y_1, y_2, and ε that solve the linearized version of this problem;
  - compute $\widetilde{\varepsilon} = \max\limits_k |z^{(k)} - f(s_1^{(k)}, y_1, y_2)|$.
- *Fact:* in the linearizable case, this estimate is close to the desired solution: $\widetilde{\varepsilon} \approx \varepsilon_{\rm opt}$.
- *Idea:* in the general case, the value $\widetilde{\varepsilon}$ is also a upper bound: $\varepsilon_{\rm opt} \le \widetilde{\varepsilon}$.
- *Caution:* in the general case, it may not be close to the desired bound – too much excess width: $\varepsilon_{\rm opt} \ll \widetilde{\varepsilon}$.

14. General Case: Idea

- *Newton's method* for solving an equation $F(x) = 0$:
  - pick an initial approximation $x^{(0)}$;
  - for $p = 0, 1, \ldots$, once we have $x^{(p)}$, solve the linearized problem, with $x = x^{(p)} + \Delta x$:

$$F(x^{(p)}) + \frac{\partial F}{\partial x}(x^{(p)}) \cdot \Delta x = 0;$$

  - use the solution $x^{(p)} + \Delta x$ as the next approx. $x^{(p+1)}$;
  - iterate until $|\Delta x|$ is small enough.
- *Similar idea:* start with $y_1^{(0)}$, $y_2^{(0)}$; for $p = 0, 1, \ldots$:
  - find Δy_1 and Δy_2 with $\varepsilon \to \min$, where:

$$-\varepsilon \le z^{(k)} - f(s^{(k)}, y_1^{(p)}, y_2^{(p)}) - \frac{\partial f}{\partial y_1}(y_1^{(p)}, y_2^{(p)}) \cdot \Delta y_1 - \frac{\partial f}{\partial y_2}(\cdot) \cdot \Delta y_2 \le \varepsilon;$$

  - take $y_1^{(p+1)} = y_1^{(p)} + \Delta y_1$ and $y_2^{(p+1)} = y_2^{(p)} + \Delta y_2$;
  - iterate until $|\Delta y_1|$ and $|\Delta y_1|$ are small enough.

time (in sec)	measured temperature	prediction: original model	prediction: new model
0	25.0	25.0	25.0
100	105.5	97.3	105.5
200	139.3	127.4	138.8
300	165.5	150.9	165.2
400	188.7	172.1	188.7
500	210.6	192.2	211.1
600	231.9	211.9	233.1
700	253.0	231.4	254.9
800	273.9	250.8	276.6
900	294.9	270.3	298.3
1000	315.8	289.7	319.9

15. Results

- *Original model:*
  - assumes that the parameters k and ρC_p are known;
  - leads to predictions with accuracy 25°.
- *New model:*
  - takes into account that we only know the values k and ρC_p with uncertainty;
  - leads to predictions with a much higher accuracy 5°.

16. Additional Idea: How to Simplify Computations

- *Reminder:* our main formula has the form

$$T(x,t) = T_i + \frac{q \cdot L}{k} \cdot \left[\frac{(k/\rho C_p) \cdot t}{L^2} + \frac{1}{3} - \frac{x}{L} + \frac{1}{2} \cdot \left(\frac{x}{L} \right)^2 - \right.$$

$$\left. \frac{2}{\pi^2} \cdot \sum_{n=1}^{6} \frac{1}{n^2} \cdot \exp\left(-n^2 \cdot \pi^2 \cdot \frac{(k/\rho C_p) \cdot t}{L^2} \right) \cdot \cos\left(n \cdot \pi \cdot \frac{x}{L} \right) \right].$$

- *Observation:* in this formula, the parameter ρC_p always appears in a ratio $\dfrac{k/\rho C_p}{L^2}$.
- *Resulting idea:*
  - instead of $y_1 = k$ and $y_2 = \rho C_p$,
  - we should use $Y_1 = \dfrac{q \cdot L}{k}$ and $Y_2 = \dfrac{k/\rho C_p}{L^2}$.

17. How to Simplify Computations (cont-d)

- *Idea (reminder):*
  - instead of $y_1 = k$ and $y_2 = \rho C_p$,
  - we use $Y_1 = \dfrac{q \cdot L}{k}$ and $Y_2 = \dfrac{k/\rho C_p}{L^2}$.
- *Resulting simplified formula:*

$$T(x,t) = T_i + Y_1 \cdot \left[Y_2 \cdot t + \frac{1}{3} - x_0 + \frac{1}{2} \cdot x_0^2 - \right.$$

$$\left. \frac{2}{\pi^2} \cdot \sum_{n=1}^{6} \frac{1}{n^2} \cdot \exp(-n^2 \cdot \pi^2 \cdot Y_2 \cdot t) \cdot \cos\left(n \cdot \pi \cdot x_0\right) \right],$$

where $x_0 \stackrel{\text{def}}{=} \dfrac{x}{L}$.

18. Acknowledgments

This work was supported in part

- by the National Science Foundation grants HRD-0734825 and DUE-0926721,
- by Grant 1 T36 GM078000-01 from the National Institutes of Health,
- by Grant MSM 6198898701 from MŠMT of Czech Republic, and
- by Grant 5015 "Application of fuzzy logic with operators in the knowledge based systems" from the Science and Technology Centre in Ukraine (STCU), funded by European Union.