# Towards Combining Probabilistic, Interval, Fuzzy Uncertainty, and Constraints: An Example Using of Inverse Problem in Geophysics

Vladik Kreinovich, Scott A. Starks,
Matthew Averill, Roberto Araiza, G. Randy Keller,
and Gang Xiang
Pan-American Center for Earth and Environmental Studies
University of Texas at El Paso, El Paso, TX 79968, USA
keller@utep.edu, sstarks@utep.edu, vladik@utep.edu

# 1. Determining Earth Structure Is Very Important

- *Importance:* civilization greatly depends on the things we extract from the Earth: oil, gas, water.
- *Need* is growing, so we must find new resources.
- *Problem:* most easy-to-access mineral resources have been discovered.
- *Example:* new oil fields are at large depths, under water, in remote areas – so drilling is very expensive.
- *Objective:* predict resources before we invest in drilling.
- *How:* we know what structures are promising.
- *Example:* oil and gas concentrate near the top of (natural) underground domal structures.
- *Conclusion:* to find mineral resources, we must determine the structure at different depths $z$ at different locations $(x, y)$.

## 2. Data that We Can Use to Determine the Earth Structure

- *Available measurement results:* those obtained without drilling boreholes.
- *Examples:*
  - gravity and magnetic measurements;
  - travel-times $t_i$ of seismic ways through the earth.
- *Need for active seismic data:*
  - passive data from earthquakes are rare;
  - to get more information, we make explosions, and measure how the resulting seismic waves propagate.
- *Resulting seismic inverse problem:*
  - we know the travel times $t_i$;
  - we want to reconstruct velocities at different depths.

# 3. Algorithm for the Forward Seismic Problem

- *We know:* velocities $v_j$ in each grid cell $j$.
- *We want to compute:* traveltimes $t_i$.
- *First step:* find shortest (in time) paths.
- *Within cell:* path is a straight line.
- *On the border:* between cells with velocities $v$ and $v'$, we have Snell's law $\dfrac{\sin(\varphi)}{v} = \dfrac{\sin(\varphi')}{v'}$.
- *Comment:* if $\sin(\varphi') > 1$, the wave cannot get penetrate into the neighboring cell; it bounces back.
- *Resulting traveltimes:* $t_i = \sum\limits_j \dfrac{\ell_{ij}}{v_j}$, where $\ell_{ij}$ is the length of the part of $i$-th path within cell $j$.
- *Simplification:* use slownesses $s_j \stackrel{\text{def}}{=} \dfrac{1}{v_j}$; $t_i = \sum\limits_j \ell_{ij} \cdot s_j$.

## 4. Algorithm for the Inverse Problem: General Description

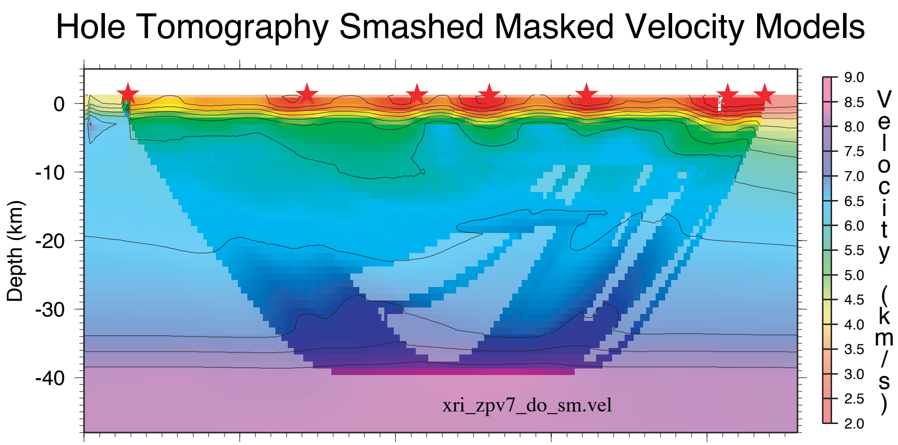
- *The most widely used:* John Hole's iterative algorithm.
- *Starting point:* reasonable initial slownesses.
- *On each iteration:* we use current (approximate) slownesses $s_j$ to compute the travel-times $t_i = \sum_j \ell_{ij} \cdot s_j$.
- *Fact:* measured travel-times $\widetilde{t}_i$ are somewhat different: $\Delta t_i \stackrel{\text{def}}{=} \widetilde{t}_i - t_i \neq 0$.
- *Objective:* find $\Delta s_j$ so that $\sum \ell_{ij} \cdot (s_j + \Delta s_j) = \widetilde{t}_i$.
- *Problem:* we have many observations $n$, and computation time $\sim n^3$ – too long, so we need faster techniques.
- *Stopping criterion:* when average error $\frac{1}{n} \sum_{i=1}^{n} (\Delta t_i)^2$ is below noise.

## 5. Algorithm for the Inverse Problem: Details

- *Objective (reminder):* find $\Delta s_j$ s.t. $\sum \ell_{ij} \cdot \Delta s_j = \Delta t_i$.
- *Simplest case:* one path.
- *Specifics:* under-determined system: 1 equation, many unknowns $\Delta s_j$.
- *Idea:* no reason for $\Delta s_j$ to be different: $\Delta s_j \approx \Delta s_{j'}$.
- *Formalization:* minimize $\sum\limits_{j,j'} (\Delta s_j - \Delta s_{j'})^2$ under the constraint $\sum \ell_{ij} \cdot \Delta s_j = \Delta t_i$.
- *Solution:* $\Delta s_j = \dfrac{\Delta t_i}{L_i}$ for all $j$, where $L_i = \sum\limits_j \ell_{ij}$.
- *Realistic case:* several paths; we have $\Delta s_{ij}$ for different paths $i$.
- *Idea:* least squares $\sum\limits_i (\Delta s_j - \Delta s_{ij})^2 \to \min$.
- *Solution:* $\Delta s_j$ is the average of $\Delta s_{ij}$.

## 6. Successes, Limitations, Need for Prior Knowledge

- *Successes:* the algorithm usually leads to reasonable geophysical models.
- *Limitations:* often, the resulting velocity model is not geophysically meaningful.
- *Example:* resulting velocities outside of the range of reasonable velocities at this depth.
- *It is desirable:* incorporate the expert knowledge into the algorithm for solving the inverse problem.


Hole Tomography Smashed Masked Velocity Models
Depth (km)
0
-10
-20
-30
-40
xri_zpv7_do_sm.vel
Velocity (km/s)
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0

# 7. Case of Interval Prior Knowledge

- *Idea:* for each cell $j$, a geophysicist provides an interval $[\underline{s}_j, \overline{s}_j]$ of possible values of $s_j$.
- *Hole's code:* along each path $i$, we find corrections $\Delta s_{ij}$ that minimize
$$\sum_{j,j'} (\Delta s_{ij} - \Delta s_{ij'})^2$$
under the constraint
$$\sum_{j=1}^{c} \ell_{ij} \cdot \Delta s_{ij} = \Delta t_i.$$
- *Modification:* we must minimize under the additional constraints
$$\underline{s}_j \le s_j^{(k)} + \Delta s_{ij} \le \overline{s}_j.$$
- *What we designed:* an $O(c \cdot \log(c))$ algorithm for solving this new problem.

## 8. New Algorithm: For Each Path on Each Iteration

- *Case:* $\Delta t_i > 0$; for $\Delta t_i < 0$, we have similar formulas.
- Compute, for each cell $j$,

$$\underline{\Delta}_j = \underline{s}_j - s_j^{(k-1)} \text{ and } \overline{\Delta}_j = \overline{s}_j - s_j^{(k-1)}.$$

- Sort values $\overline{\Delta}_j$ into

$$\overline{\Delta}_{(1)} \le \overline{\Delta}_{(2)} \le \ldots \le \overline{\Delta}_{(c)}.$$

- For every $p$ from 0 to $c$, compute:

$$A_0 = 0,\ \mathcal{L}_0 = L_i,\ A_p = A_{p-1} + \ell_{i(p)} \cdot \overline{\Delta}_{(p)},\ \mathcal{L}_p = \mathcal{L}_{p-1} - \ell_{i(p)}.$$

- Compute $S_p = A_p + \mathcal{L}_p \cdot \Delta_{(p+1)}$, and find $p$ s.t.

$$S_{p-1} \le \Delta t_i < S_p.$$

- Take $\Delta s_{i(j)} = \overline{\Delta}_j$ for $j \le p$; $\Delta s_{i(j)} = \dfrac{\Delta t_i - A_p}{\mathcal{L}_p}$ else.
- Then, average $\Delta s_{i(j)}$ over paths $i$.

# 9. Explicit Expert Knowledge: Fuzzy Uncertainty

- Experts can usually produce an wider interval of which they are practically 100% certain.
- In addition, experts can also produce narrower intervals about which their degree of certainty is smaller.
- As a result, instead of a *single* interval, we have a *nested* family of intervals corresponding to different levels of uncertainty.
- In effect, we get a *fuzzy* interval (of which different intervals are $\alpha$-cuts).
- *Previously:* a solution is satisfying or not.
- *New idea:* a satisfaction *degree* $d$.
- *Specifics:* $d$ is the largest $\alpha$ for which all $s_i$ are within the corresponding $\alpha$-cut intervals.

## 10. How We Can Use Fuzzy Uncertainty

- *Objective:* find the largest possible value $\alpha$ for which the slownesses belong to the $\alpha$-cut intervals.
- *Possible approach:*
  - try $\alpha = 0$, $\alpha = 0.1$, $\alpha = 0.2$, etc., until the process stops converging;
  - the solution corresponding to the previous value $\alpha$ is the answer.
- *Comment:*
  - this is the basic straightforward way to take fuzzy-valued expert knowledge into consideration;
  - several researchers successfully used fuzzy expert knowledge in geophysics (Nikravesh, Klir, et al.);
  - we plan to add their ideas to our algorithms.

## 11. Case of Probabilistic Prior Knowledge

- *Description:* from prior observations, we know $\widetilde{s}_j \approx s_j$, and we know the st. dev. $\sigma_j$ of this value.

- *Minimize:* $\sum\limits_{j,j'}(\Delta s_{ij} - \Delta s_{ij'})^2$ s.t. $\sum\limits_{j=1}^{c} \ell_{ij} \cdot \Delta s_{ij} = \Delta t_i$ and

$$\frac{1}{n} \cdot \sum_{j=1}^{c} \frac{((s_j^{(k)} + \Delta s_{ij}) - \widetilde{s}_j)^2}{\sigma_j^2} = 1.$$

- *Solution* (Lagrange multipliers): $\overline{\Delta s} \stackrel{\text{def}}{=} \frac{1}{n} \cdot \sum\limits_{j=1}^{c} \Delta s_{ij}$,

$$\frac{2}{n} \cdot \Delta s_{ij} - \frac{2}{n} \cdot \overline{\Delta s} + \lambda \cdot \ell_{ij} + \frac{2\mu}{n \cdot \sigma_j^2} \cdot (s_j^{(k)} + \Delta s_{ij} - \widetilde{s}_j) = 0.$$

- *Fact:* $\Delta s_{ij}$ is an explicit function of $\lambda$, $\mu$, $\overline{\Delta s}$.

- *Algorithm:* solve 3 non-linear equations (above one + 2 constraints) with unknowns $\lambda$, $\mu$, $\overline{\Delta s}$.

# 12. Combination of Different Types of Prior Knowledge

- *Need:* we often have both:
  - prior measurement results – i.e., *probabilistic* knowledge, and
  - expert estimates – i.e., *interval* and *fuzzy* knowledge.
- *Minimize:* $\sum\limits_{j,j'}(\Delta s_{ij} - \Delta s_{ij'})^2$ s.t. $\sum\limits_{j=1}^{c} \ell_{ij} \cdot \Delta s_{ij} = \Delta t_i$,

$$\frac{1}{n} \cdot \sum_{j=1}^{c} \frac{((s_j^{(k)} + \Delta s_{ij}) - \widetilde{s}_j)^2}{\sigma_j^2} \le 1,$$

and $\underline{s}_j \le s_j^{(k)} + \Delta s_{ij} \le \overline{s}_j$.

- *Idea:* we minimize a convex function under convex constraints; efficient algorithms are known.

# 13. Combination of Different Types of Prior Knowledge: Algorithm

- *Idea* – method of alternating projections:
  - first, add a correction that satisfy the first constraint,
  - then, the additional correction that satisfies the second constraint,
  - etc.
- *Specifics:*
  - first, add equal values $\Delta s_{ij}$ to minimize $\Delta t_i$;
  - restrict the values to the nearest points from $[\underline{s}_j, \overline{s}_j]$,
  - find the extra corrections that satisfy the probabilistic constraint,
  - repeat until converges.

## 14. Acknowledgments

This work was supported in part:

- by NASA under cooperative agreement NCC5-209,
- by NSF grants EAR-0112968, EAR-0225670, and EIA-0321328,
- by Star Award from the University of Texas System, and
- by Texas Department of Transportation grant No. 0-5453.