Towards Combining Probabilistic, Interval, Fuzzy Uncertainty, and Constraints: On the Example of Inverse Problem in Geophysics

G. Randy Keller, Scott A. Starks, Aaron Velasco, Matthew Averill, Roberto Araiza, Gang Xiang, and Vladik Kreinovich

University of Texas at El Paso, El Paso, TX 79968, USA,
keller@utep.edu sstarks@utep.edu vladik@utep.edu

Abstract. In many real-life situations, we have several types of uncertainty: measurement uncertainty can lead to probabilistic and/or interval uncertainty, expert estimates come with interval and/or fuzzy uncertainty, etc. In many situations, in addition to measurement uncertainty, we have prior knowledge coming from prior data processing, prior knowledge coming from prior interval constraints. In this paper, on the example of the seismic inverse problem, we show how to combine these different types of uncertainty.

1 Seismic Inverse Problem: A Brief Description

In evaluations of natural resources and in the search for natural resources, it is very important to determine Earth structure. Our civilization greatly depends on the things we extract from the Earth, such as fossil fuels (oil, coal, natural gas), minerals, and water. Our need for these commodities is constantly growing, and because of this growth, they are being exhausted. Even under the best conservation policies, there is (and there will be) a constant need to find new sources of minerals, fuels, and water.

The only sure-proof way to guarantee that there are resources such as minerals at a certain location is to actually drill a borehole and analyze the materials extracted. However, exploration for natural resources using indirect means began in earnest during the first half of the 20th century. The result was the discovery of many large relatively easy to locate resources such as the oil in the Middle East.

However, nowadays, most easy-to-access mineral resources have already been discovered. For example, new oil fields are mainly discovered either at large depths, or under water, or in very remote areas – in short, in the areas where drilling is very expensive. It is therefore desirable to predict the presence of resources as accurately as possible before we invest in drilling.

From previous exploration experiences, we usually have a good idea of what type of structures are symptomatic for a particular region. For example, oil and gas tend to concentrate near the top of natural underground domal structures.

So, to be able to distinguish between more promising and less promising locations, it is desirable to determine the structure of the Earth at these locations. To be more precise, we want to know the structure at different depths z at different locations (x, y).

Data that we can use to determine the Earth structure. In general, to determine the Earth structure, we can use different measurement results that can be obtained without actually drilling the boreholes: e.g., gravity and magnetic measurements, analyzing the travel-times and paths of seismic ways as they propagate through the earth, etc.

To get a better understanding of the Earth structure, we must rely on *active* seismic data – in other words, we must make artificial explosions, place sensors around them, and measure how the resulting seismic waves propagate. The most important information about the seismic wave is the *travel-time* t_i, i.e., the time that it takes for the wave to travel from its source to the sensor. to determine the geophysical structure of a region, we measure seismic travel times and reconstruct velocities at different depths from these data. The problem of reconstructing this structure is called the *seismic inverse problem.*

2 Known Algorithms for Solving the Seismic Inverse Problem: Description, Successes, Limitations

We want to find the values of the velocity $v(\boldsymbol{x})$ at different 3-D points $\boldsymbol{x}$. Based on the finite number of measurements, we can only reconstruct a finite number of parameters. So, we take a rectangular grid and to reconstruct the velocities v_j at different grid points.

Algorithm for the forward problem: brief description. Once we know the velocities v_j in each cell j, we can then determine the paths which seismic waves take. Seismic waves travel along the shortest path – shortest in terms of time. It can be easily determined that for such paths, within each cell, the path is a straight line, and on the border between the two cells with velocities v and v', the direction of the path changes in accordance with Snell's law $\dfrac{\sin(\varphi)}{v} = \dfrac{\sin(\varphi')}{v'}$, where φ and φ' are the angles between the paths and the line orthogonal to the border between the cells. (If this formula requires $\sin(\varphi') > 1$, this means that this wave cannot penetrate into the neighboring cell at all; instead, it bounces back into the original cell with the same angle φ.)

In particular, we can thus determine the paths from the source to each sensor. The travel-time t_i along i-th path can then be determined as the sum of travel-times in different cells j through which this path passes: $t_i = \sum\limits_j \dfrac{\ell_{ij}}{v_j}$, where ℓ_{ij} denotes the length of the part of i-th path within cell j.

This formula becomes linear if we replace the original unknowns – velocities v_j – by their inverses $s_j \stackrel{\text{def}}{=} \dfrac{1}{v_j}$, called *slownesses*. In terms of slownesses, the formula for the travel-time takes the simpler form $t_i = \sum_j \ell_{ij} \cdot s_j$.

Algorithm for the inverse problem: general description. There are several algorithms for solving this inverse problem; see, e.g., [5, 7, 9]. The most widely used is the following iterative algorithm proposed by John Hole [5].

At each stage of this algorithm, we have some approximation to the desired slownesses. We start with some reasonable initial slownesses, and we hope that after several iterations, we will be able to get slownesses which are much closer to the actual values.

At each iteration, we first use the currently known slownesses s_j to find the corresponding paths from the source to each sensor. Based on these paths, we compute the predicted values $t_i = \sum_j \ell_{ij} \cdot s_j$ of travel-times.

Since the currently known slownesses s_j are only approximately correct, the travel-times t_i (which are predicted based on these slownesses) are approximately equal to the measured travel-times $\widetilde{t}_i$; there is, in general, a discrepancy $\Delta t_i \stackrel{\text{def}}{=} \widetilde{t}_i - t_i \neq 0$. It is therefore necessary to use these discrepancies to update the current values of slownesses, i.e., replace the current values s_j with corrected values $s_j + \Delta s_j$. The objective of this correction is eliminate (or at least decrease) the discrepancies $\Delta t_i \neq 0$. In other words, the objective is to make sure that for the corrected values of the slowness, the predicted travel-times are closer to $\widetilde{t}_i$.

Of course, once we have changed the slownesses, the shortest paths will also change; however, if the current values of slownesses are reasonable, the differences in slowness are not large, and thus, paths will not change much. Thus, in the first approximation, we can assume that the paths are the same, i.e., that for each i and j, the length ℓ_{ij} remains the same. In this approximation, the new travel-times are equal to $\sum \ell_{ij} \cdot (s_j + \Delta s_j)$. The desired condition is then $\sum \ell_{ij} \cdot (s_j + \Delta s_j) = \widetilde{t}_i$. Subtracting the formula $t_i = \sum_j \ell_{ij} \cdot s_j$ from this expression, we conclude that the corrections Δs_j must satisfy the following system of (approximate) linear equations: $\sum \ell_{ij} \cdot \Delta s_j \approx \Delta t_i$.

Solving this system of linear equations is not an easy task, because we have many observations and many cell values and thus, many unknowns, and for a system of linear equations, computation time required to solve it grows as a cube n^3 of the number of variables n. So, instead of the standard methods for solving a system of linear equations, researchers use special faster geophysics-motivated techniques (described below) for solving the corresponding systems. These methods are described, in detail, in the next subsection.

Once we solve the corresponding system of linear equations, we compute the updated values Δs_j, compute the new (corrected) slownesses $s_j + \Delta s_j$, and repeat the procedure again. We stop when the discrepancies become small; usually, we stop when the mean square error $\dfrac{1}{n} \sum_{i=1}^{n} (\Delta t_i)^2$ no longer exceeds a given threshold. This threshold is normally set up to be equal to the measurement noise level, so that we stop iterations when the discrepancy between the model and the observations falls below the noise level – i.e., when, for all practical purposes, the model is adequate.

Algorithm for the inverse problem: details. Let us describe, in more detail, how the corresponding linear system of equations is usually solved. In other words, for a given cell j, how do we find the correction Δs_j to the current value of slowness s_j in this cell?

Let us first consider the simplified case when there is only path, and this path is going through the j-th cell. In this case, cells through which this path does not go does not need any correction. To find the corrections Δs_j for all the cells j through which this path goes, we only have one equation $\sum\limits_j \ell_{ij} \cdot \Delta s_j = \Delta t_i$. The resulting system of linear equations is clearly under-determined: we have a single equation to find the values of several variables Δs_j. Since the system is under-determined, we have a infinite number of possible solutions. Our objective is to select the most geophysical reasonable of these solutions.

For that, we can use the following idea. Our single observation involves several cells; we cannot distinguish between the effects of slownesses in different cells, we only observe the overall effect. Therefore, there is no reason to assume that the value Δs_j in one of these cells is different from the values in other cells. It is thus reasonable to assume that all these values are close to each other: $\Delta s_j \approx \Delta s_{j'}$. The least squares method enables us to describe this assumption as minimization of the objective function $\sum\limits_{j,j'} (\Delta s_j - \Delta s_{j'})^2$ under the condition that $\sum \ell_{ij} \cdot \Delta s_j = \Delta t_i$. The minimum is attained when all the values Δs_j are equal. Substituting these equal values into the equation $\sum\limits_j \ell_{ij} \cdot \Delta s_j = \Delta t_i$, we conclude that $L_i \cdot \Delta s = \Delta t_i$, where $L_i = \sum\limits_j \ell_{ij}$ is the overall length of i-th path. Thus, in the simplified case in which there is only one path, to the slowness of each cell j along this path, we add the same value $\Delta s_j = \dfrac{\Delta t_i}{L_i}$.

Let us now consider the realistic case in which there are many paths, and moreover, for many cells j, there are many paths i which go through the corresponding cell. For a given cell j, based on each path i passing through this cell, we can estimate the correction Δs_j by the corresponding value $\Delta s_{ij} \stackrel{\text{def}}{=} \dfrac{\Delta t_i}{L_i}$. Since there are usually several paths going through the j-th cell, we have, in general, several different estimates $\Delta s_j \approx \Delta s_{ij}$. Again, the least squares approach leads to $\sum\limits_i (\Delta s_j - \Delta s_{ij})^2 \to \min$, hence to Δs_j as the arithmetic average of the values Δs_{ij}.

Comment. To take into account that paths with larger ℓ_{ij} provide more information, researchers also used weighted average, with weight increasing with ℓ_{ij}.

Successes of the known algorithms. The known algorithms have been actively used to reconstruct the slownesses, and, in many practical situations, they have led to reasonable geophysical models.

Limitations of the known algorithms. Often, the velocity model that is returned by the existing algorithm is not geophysically meaningful: e.g., it predicts velocities outside of the range of reasonable velocities at this depth. To avoid such situations, it is desirable to incorporate the expert knowledge into the algorithm for solving the inverse problem.

In this paper, we describe how to do it.

3 Case of Interval Prior Knowledge

For each cell j, a geophysicist often provides us with his or her estimate of possible values of the corresponding slowness s_j. Sometimes, this estimates comes in the form of an interval $[\underline{s}_j, \overline{s}_j]$ that is guaranteed to contain the (unknown) actual value of slowness.

It is desirable to modify Hole's algorithm in such a way that on all iterations, slownesses s_j stay within the corresponding intervals. Such a modification is described in [1, 2].

Namely, in the original Hole's algorithm, once we know the current approximations $s_j^{(k)}$ to slownesses, then, along each path i, among all corrections Δs_{ij} that provide the desired compensation, i.e.., for which

$$\sum_{j=1}^{c} \ell_{ij} \cdot \Delta s_{ij} = \Delta t_i, \tag{1}$$

we find the assignment that minimizes the objective function $\sum\limits_{j,j'} (\Delta s_{ij} - \Delta s_{ij'})^2$, i.e., equivalently, that minimizes the variance of the values Δs_{ij} along this path:

$$V \stackrel{\text{def}}{=} \frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij}^2 - \left(\frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij} \right)^2. \tag{2}$$

In the presence of the interval prior information, on each iteration of Hole's algorithm, we must still minimize the objective function (2), but this time, we minimize it under two constraint: the same constraint (1) and the new constraints

$$\underline{s}_j \le s_j^{(k)} + \Delta s_{ij} \le \overline{s}_j. \tag{3}$$

We have found the following efficient $O(c \cdot \log(c))$ time algorithm for solving the corresponding constraint optimization problem. We start with the initial slowness values $s_j^{(0)}$ which are within the given intervals $[\underline{s}_j, \overline{s}_j]$.

On each iteration of the new procedure, we start with the slowness values $s_j^{(k-1)}$ which are within given intervals $[\underline{s}_j, \overline{s}_j]$. Based on these slownesses, we find the paths from the sources to the sensors, compute the predicted travel-times t_i along each path, and the discrepancies $\Delta t_i = \widetilde{t}_i - t_i$.

We then compute, for each cell j, the values $\underline{\Delta}_j = \underline{s}_j - s_j^{(k-1)}$ and $\overline{\Delta}_j = \overline{s}_j - s_j^{(k-1)}$. We will consider the case when $\Delta t_i > 0$; the case when $\Delta t_i < 0$ is treated similarly. In this case, we first sort all c values $\overline{\Delta}_j$ along the i-th path into a non-decreasing sequence

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

Then, for every p from 0 to c, we compute the values A_p and $\mathcal{L}_p$ as follows:

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

After that, for each p, we compute $S_p = A_p + \mathcal{L}_p \cdot \Delta_{(p+1)}$, and we find p for which $S_{p-1} \le \Delta t_i < S_p$. Once this p is found, we take $\Delta s_{i(j)} = \overline{\Delta}_j$ for $j \le p$, and for $j > p$, we take $\Delta s_{(j)} = \dfrac{\Delta t_i - A_p}{\mathcal{L}_p}$.

When $\Delta t_i < 0$, we similarly sort the values $\underline{\Delta}_j$ into a decreasing sequence, and find p so that the first p corrections are "maxed out" to $\underline{\Delta}_j$, and the rest $c - p$ corrections are determined from the condition $\Delta s_{i(j)} = \dfrac{\Delta t_i - A_p}{\mathcal{L}_p}$.

Once we have computed these corrections for all the paths, then for each cell j, we take the average (or weighted average) of all the corrections coming from all the paths which pass through this cell.

4 Case of Fuzzy Prior Knowledge

In general, experts are often not 100% sure about the corresponding intervals. They can usually produce a wider interval $[\underline{s}_j, \overline{s}_j]$ of which they are practically 100% certain, but in addition to that, they 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 – i.e., in effect, a fuzzy interval (of which different intervals are α-cuts).

So, instead of simply saying that a given solution to the seismic inverse problem is satisfying or not, we provide a *degree* to which the given solution is satisfying – as the largest α for which the velocity at every point is within the corresponding α-cut intervals.

To solve the seismic inverse problem under such fuzzy uncertainty, we apply the interval algorithm for α-cuts corresponding to $\alpha = 0$, $\alpha = 0.1$, $\alpha = 0.2$, etc., until we reach such a value of α that the process no longer converges. Then, the solution corresponding to the previous value α – i.e., to the largest value α for which the process converged – is returned as the desired solution to the seismic inverse problem.

5 Case of Probabilistic Prior Knowledge

Often, prior information comes from processing previous observations of the region of interest. In this case, before our experiments, for each cell j, we know a prior (approximate) slowness value $\widetilde{s}_j$, and we know the accuracy (standard deviation) σ_j of this approximate value $\widetilde{s}_j$. It is known that this prior information can lead to much more accurate velocity models; see, e.g., [6]. How can we modify Hole's algorithm so that it takes this prior information into account?

Due to the prior knowledge, for each cell j, the ratio $\dfrac{(s_j^{(k)} + \Delta s_{ij}) - \widetilde{s}_j}{\sigma_j}$ is normally distributed with 0 mean and variance 1. Since each path i consists of a reasonable number of cells, we can thus conclude that the sample variance of this ratio should be close to σ_j, i.e., that

$$\frac{1}{n} \cdot \sum_{j=1}^{c} \frac{((s_j^{(k)} + \Delta s_{ij}) - \widetilde{s}_j)^2}{\sigma_j^2} = 1. \tag{4}$$

So, to find the corrections Δs_{ij}, we must minimize the objective function (2) under the constraints (1) and (4).

By applying the Lagrange multiplier method to this problem, we can reduce this problem to the unconstrained minimization problem

$$\frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij}^2 - \left(\frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij}\right)^2 + \lambda \cdot \left(\sum_{j=1}^{c} \ell_{ij} \cdot \Delta s_{ij} - \Delta t_i\right) +$$

$$\mu \cdot \frac{1}{n} \cdot \sum_{j=1}^{c} \frac{(s_j^{(k)} + \Delta s_{ij} - \widetilde{s}_j)^2}{\sigma_j^2} \to \min. \tag{5}$$

Differentiating this equation by Δs_{ij} and equating the derivative to 0, we conclude that

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

where

$$\overline{\Delta s} \stackrel{\text{def}}{=} \frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij}. \tag{6}$$

Once we fix λ, μ, and $\overline{\Delta s}$, we get an explicit expression for the values Δs_{ij}. Substituting these expressions into the equations (1), (4), and (6), we get an easy-to-solve system of 3 non-linear equations with 3 unknowns, which we can solve, e.g., by using Newton's method.

Now, instead of explicit formulas for a transition from $s_j^{(k)}$ to $s_j^{(k+1)}$, we need a separate iteration process – so the computation time is somewhat larger, but we get a more geophysically meaningful velocity map – that takes prior knowledge into account.

6 Combination of Different Types of Prior Knowledge

In many real-life situations, we have both the prior measurement results – which lead to the probabilistic prior knowledge, and expert estimates – which lead to interval and fuzzy prior knowledge. In the presence of probabilistic and interval prior knowledge, we must minimize (2) under the constraints (1), (3), and (4).

If we replace the equality in (4) by an inequality ≤ 1, then we get a problem of minimizing a convex function under convex constraints, a problem for which there are known efficient algorithms; see, e.g., [8].

For example, we can use a method of alternating projections, in which we first add a correction that satisfy the first constraint, then the additional correction that satisfies the second constraint, etc. In our case, we first add equal values of Δs_{ij} to satisfy the constraint (2), then we restrict the values to the nearest points from the interval $[\underline{s}_j, \overline{s}_j]$ – to satisfy the constraint (3), and after that, find the extra corrections that satisfy the condition (4), after which we repeat the cycle again until the process converges.

Acknowledgments. This work was supported in part by NASA under cooperative agreement NCC5-209, NSF grants EAR-0225670 and DMS-0532645, Star Award from the University of Texas System, and Texas Department of Transportation grant No. 0-5453.

References

1. Averill, M. G., Miller, K. C., Keller, G. R., Kreinovich, V., Araiza, R., Starks, S. A.: Using Expert Knowledge in Solving the Seismic Inverse Problem. Proceedings of the 24nd International Conference of the North American Fuzzy Information Processing Society NAFIPS'2005, Ann Arbor, Michigan, June 22–25, 2005, pp. 310–314.
2. Averill, M. G., Miller, K. C., Keller, G. R., Kreinovich, V., Araiza, R., Starks, S. A.: Using Expert Knowledge in Solving the Seismic Inverse Problem. International Journal of Approximate Reasoning (to appear).
3. Bardossy, G., Fodor, J.: *Evaluation of Uncertainties and Risks in Geology*, Springer Verlag, Berlin, 2004.
4. Demicco, R., Klir, J., Eds.: *Fuzzy Logic in Geology*, Academic Press, 2003.
5. Hole J. A.: Nonlinear High-Resolution Three-Dimensional Seismic Travel Time Tomography. J. Geophysical Research **97** (1992) 6553–6562.
6. Maceira, M., Taylor, S. R., Ammon, C. J., Yang, X., Velasco, A. A.: High-resolution Rayleigh wave slowness tomography of Central Asia. Journal of Geophysical Research **110** (2005) paper B06304.
7. Parker R. L.: *Geophysical Inverse Theory*, Princeton University Press, Princeton, New Jersey, 1994.
8. Vavasis, S. A.: *Nonlinear Optimization: Complexity Issues*, Oxford University Press, New York, 1991.
9. Zelt, C. A., Barton, P. J.: Three-dimensional seismic refraction tomography: A comparison of two methods applied to data from the Faeroe Basin. J. Geophysical Research **103** (1998) 7187–7210.