4-1-2007

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

George R. Keller
University of Texas at El Paso, keller@utep.edu

Scott A. Starks

Aaron Velasco
University of Texas at El Paso, aavelasco@utep.edu

Matthew Averill
University of Texas at El Paso

Roberto Araiza
University of Texas at El Paso, raraiza@miners.utep.edu

Comments:
UTEP-CS-06-19b
See next page for additional authors
Original paper published in the *Proceedings of the Second International Conference on Fuzzy Sets and Soft Computing in Economics and Finance FSSCEF'2006*, St. Petersburg, Russia, June 28 - July 1, 2006, pp. 47-54; updated version published in *IEEE Proceedings of the 12th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics*, Duisburg, Germany, September 26-29, 2006.

Authors
George R. Keller, Scott A. Starks, Aaron Velasco, Matthew Averill, Roberto Araiza, Gang Xiang, and Vladik Kreinovich

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

V. Kreinovich[1], S. A. Starks[2], R. Araiza[1], and G. Xiang[1]
Depts. [1]Computer Science and [2]Electrical & Computer Eng.
Univ. of Texas at El Paso
El Paso, TX 79968, USA
contact vladik@utep.edu

A. A. Velasco and M. G. Averill
Dept. Geological Sci.
Univ. of Texas at El Paso
El Paso, TX 79968, USA
mgaverill@utep.edu

G. R. Keller
School of Geology & Geophysics
University of Oklahoma
100 East Boyd
Norman, OK 73019, USA
grkeller@ou.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 and/or 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

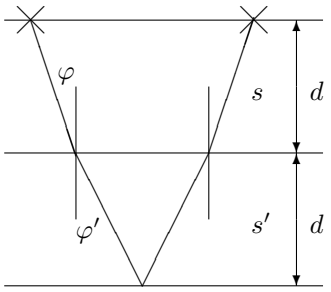
We want to find the values of the velocity $v(\vec{x})$ at different 3-D points $\vec{x}$. Based on the finite number of measurements, we can only reconstruct a finite number of parameters. So, we use a rectangular grid structure to divide the 3-D volume into box-shaped cells. We assume that the value of the velocity v_j is the same within each cell, and we reconstruct the velocities v_j within different cells.

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 $\frac{\sin(\varphi)}{v} = \frac{\sin(\varphi')}{v'}$, 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: $t_i = \sum_j \frac{\ell_{ij}}{v_j}$, where ℓ_{ij} denotes the length of the part of i-th path within cell j.

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

It is worth mentioning, however, that the resulting system of equations is *not* linear in the unknowns s_j. Indeed, the actual geometry of the shortest path between the two given points depends on the actual values of the velocities v_j – i.e., equivalently, on the slownesses s_j. Thus, the lengths ℓ_{ij} of the segments of these shortest paths also depend on the slownesses $s_1, \ldots, s_m$. To be more precise, we should therefore explicitly take this dependence into account and re-write the above system as $t_i = \sum_j \ell_{ij}(s_1, \ldots, s_m) \cdot s_j$ for an appropriate non-linear dependence $\ell_{ij}(s_1, \ldots, s_m)$.



Algorithm for the inverse problem: general description. There are several algorithms for solving this inverse problem; see, e.g., [11, 24, 28]. The most widely used is the following iterative algorithm proposed by John Hole [11].

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 to 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 c^3 of the number of variables c. 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 $\frac{1}{n}\sum_{i=1}^{n}(\Delta t_i)^2$ 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 above auxiliary linear system of equations with unknown Δs_j 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 do 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 our previous papers [2, 3, 13], we described how to do it. Specifically, we proposed a $O(c \log(c))$ time algorithm for taking interval prior knowledge into account.

In this paper, we provide a detailed motivation for that algorithm, and we use this motivation to design a new, faster, linear-time ($O(c)$) for solving this problem.

3. Case of Interval Prior Knowledge: Description and Known Algorithm

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. Often, 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 [2, 3, 13].

Analysis of the problem and our main idea. Once we know the current approximations $s_j^{(k)}$ to slownesses, then,

along each path i, we want to find the corrections Δs_{ij} which provide the desired compensation, i.e., for which

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

In Hole's algorithm, we select $\Delta s_{ij} = \dfrac{\Delta t_i}{L_i}$. With the additional knowledge, we may not be able to do this, because we want to make sure that the corrected values of slowness stay within the corresponding intervals

$$\underline{s}_j \le s_j^{(k)} + \Delta s_{ij} \le \overline{s}_j, \tag{2}$$

i.e., equivalently, that

$$\underline{\Delta}_j \le \Delta s_{ij} \le \overline{\Delta}_j, \tag{3}$$

where $\underline{\Delta}_j \stackrel{\text{def}}{=} \underline{s}_j - s_j^{(k)}$ and $\overline{\Delta}_j \stackrel{\text{def}}{=} \overline{s}_j - s_j^{(k)}$. Since $s_j^{(k)} \in [\underline{s}_j, \overline{s}_j]$, we conclude that $\underline{\Delta}_j \le 0$ and $\overline{\Delta}_j \ge 0$ – i.e., all lower endpoints are non-positive and all upper endpoints are non-negative.

How can we achieve this goal?

For each cell j, after an iteration of, say, Hole's algorithm, we have a corrected value of the slowness $s_j^{(k+1)} = s_j^{(k)} + \Delta s_{ij}$ which approximates the actual (unknown) slowness s_j: $s_j \approx s_j^{(k+1)}$. We also know that s_j should be located in the interval $[\underline{s}_j, \overline{s}_j]$. Similar to our previous analysis, it is therefore reasonable to use the Least Squares Method to combine these two piece of information: i.e., we look for the value $s_j \in [\underline{s}_j, \overline{s}_j]$ for which the square $(s_j - s_j^{(k)})^2$ is the smallest possible. In geometric terms, we look for the value within the given interval $[\underline{s}_j, \overline{s}_j]$ which is the closest to $s_j^{(k+1)}$. Thus:

- If the value $s_j^{(k+1)}$ is already within the interval, we keep it intact.
- If the value $s_j^{(k+1)}$ is to the left of the interval, i.e., if $s_j^{(k+1)} < \underline{s}_j$, then the closest point from the interval is its left endpoint $\underline{s}_j$.
- Similarly, if the value $s_j^{(k+1)}$ is to the right of the interval, i.e., if $s_j^{(k+1)} > \overline{s}_j$, then the closest point from the interval is its right endpoint $\overline{s}_j$.

In other words, e.g., for $\Delta t_i > 0$, we first find the universal value Δs and then, for those j for which $\Delta s > \overline{\Delta}_j$, we replace this value with $\overline{\Delta}_j$.

As a result, we arrive at the values Δs_{ij} which are all equal to Δs – except for those values for which $\overline{\Delta}_j < \Delta s$; for these values, $\Delta s_{ij} = \overline{\Delta}_j$.

Complications coming from a straightforward application of this idea. Originally, before we took interval prior knowledge into account, we had a full compensation for Δt_i. Now that we decreased some slownesses Δs_{ij}, the resulting value of $\sum\limits_{j=1}^{c} \ell_{ij} \cdot \Delta s_{ij}$ is, in general, smaller than Δt_i. Thus, there is a remaining discrepancy $\Delta t_i' \stackrel{\text{def}}{=} \Delta t_i - \sum\limits_{j=1}^{c} \ell_{ij} \cdot \Delta s_{ij} > 0$.

To eliminate this discrepancy, we need to repeat the same procedure: divide $\Delta t_i'$ by L_i and again cut down those slownesses that start going outside the corresponding intervals. Because of this cutting down, we may still get some discrepancy remaining, etc.

So, if we apply this idea in a straightforward way, we may need a large number of iterations to fully compensate for the original travel time discrepancy. The need for a large number of iterations leads to a drastic increase in computation time – which, for the seismic inverse problems, is already large.

It is therefore desirable to avoid these iterations and directly come up with a solution which provides the needed compensation of the travel time and at the same time, keeps all the corrected slownesses within the corresponding intervals.

Formulation of the problem in precise terms. For $\Delta t_i > 0$, we would like to find a value $\Delta s > 0$ such that if we take $\Delta s_{ij} = \Delta s$ for all j for which $\Delta s \le \overline{\Delta}_j$ and $\Delta s_{ij} = \overline{\Delta}_j$ for all other j, then we will satisfy the equation (1).

For $\Delta t_i < 0$, we would like to find a value $\Delta s < 0$ such that if we take $\Delta s_{ij} = \Delta s$ for all j for which $\Delta s \ge \underline{\Delta}_j$ and $\Delta s_{ij} = \underline{\Delta}_j$ for all other j, then we will satisfy the equation (1).

Analysis of the problem. In the desired solution, we have $\Delta s_{ij} = \overline{\Delta}_j$ for the values j for which $\overline{\Delta}_j$ is smaller than a certain threshold.

This desired solution is easier to describe if we first soft all the values $\overline{\Delta}_j$ into a non-decreasing sequence

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

Then, in the desired solution, there is some index p for which $\Delta s_{i(j)} = \overline{\Delta}_{(j)}$ for all $j \le p$. The common value Δs for the indices $j > p$ can be found from the condition (1), i.e., from the condition that $A_p + \mathcal{L}_p \cdot \Delta s = \Delta t_i$, where we denoted $A_p \stackrel{\text{def}}{=} \sum\limits_{i=1}^{p} \ell_{(i)j} \cdot \overline{\Delta}_{(j)}$ and $\mathcal{L}_p \stackrel{\text{def}}{=} \sum\limits_{j=p+1}^{c} \ell_{i(j)}$. Therefore, we will get $\Delta s = \dfrac{\Delta t_i - A_p}{\mathcal{L}_p}$.

For the correctly selected index p, all values $\overline{\Delta}_{(j)}$ for which we "cut off" must be smaller than this Δs, and all

the other values $\overline{\Delta}_{(j)}$ must be larger than (or equal to) this Δs. Since the values $\overline{\Delta}_{(j)}$ are sorted in increasing order, it is sufficient to check that $\overline{\Delta}_{(p)} < \Delta s \le \overline{\Delta}_{(p+1)}$.

If for some p, we get $\Delta s > \overline{\Delta}_{(p+1)}$, this means that need to cut some more – otherwise, for $j = p+1$, we will still have the value outside the desired interval. On the other hand, if we get $\Delta s \le \overline{\Delta}_{(p)}$, then there was no reason to cut off at p-th level – so we need to cut less.

Designing an algorithm. This analysis can be naturally be translated into an algorithm. First, we sort the values $\overline{\Delta}_j$; sorting takes time $O(c \cdot \log(c))$; see, e.g., [7]. Then, for every p from 0 to n, we compute the value $\Delta s = \dfrac{\Delta t_i - A_p}{\mathcal{L}_p}$ and check whether $\overline{\Delta}_{(p)} < \Delta s \le \overline{\Delta}_{(p+1)}$. Once we know A_p, computing A_{p+1} requires just one step – since we need to add one term to the sum. Thus, we to compute all c such values, we need $O(c)$ steps – to the total of $O(c \cdot \log(c)) + O(c) = O(c \cdot \log(c))$. So, we arrive at the following algorithm.

Resulting algorithm. It is sufficient to describe 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: $A_0 = 0$, $\mathcal{L}_0 = L_i$,

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

After that, for each p, we compute $\Delta s = \dfrac{\Delta t_i - A_p}{\mathcal{L}_p}$ and check whether $\overline{\Delta}_{(p)} < \Delta s \le \overline{\Delta}_{(p+1)}$. Once this condition is satisfied, we take $\Delta s_{i(j)} = \overline{\Delta}_{(j)}$ for $j \le p$, and $\Delta s_{i(j)} = \Delta s$ for $j > p$.

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 $\Delta s = \dfrac{\Delta t_i - A_p}{\mathcal{L}_p}$.

Comment. 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.

Example showing efficiency (and feasibility) of the new approach. Let us consider a simple example of two vertical layers of height d (see above picture), with $s > s'$. We assume that the structure below the second layer is so heavy that all the signals simply bounce back from the bottom of the second layer (in real geological situations, this is what happens, e.g., at the Moho surface). For simplicity, we consider only one signal.

Usually, the closer to the surface, the more information we have about the layer. In this example, we assume that we know s exactly, but we only know an approximate value $\widetilde{s}'$ for s' ($\Delta s' \stackrel{\text{def}}{=} \widetilde{s}' - s' \ne 0$). We start with the known values s and $\widetilde{s}'$ and perform iterations following both the original Hole's algorithm and the new interval method.

When the angles φ and φ' are small ($\varphi \ll 1$, $\varphi' \ll 1$), then $\sin(\varphi) \approx \varphi$, $\sin(\varphi') \approx \varphi'$, and we can analytically trace the computations; for details, see [3]. For example, the horizontal distance between the source and the sensor is $2d \cdot (\tan(\varphi) + \tan(\varphi')) \approx 2d \cdot (\varphi + \varphi')$.

In the original Hole's algorithm, the discrepancy in the travel times is uniformly divided between the whole path. As a result, we replace the original approximate slowness $\widetilde{s}' = s' + \Delta s'$ with a more accurate estimate $s' + \dfrac{\Delta s'}{2}$. Hence, the approximation error decreases by a factor of 2. So, e.g., in 7 iterations, we can reduce this error to $< 1\%$ level.

In the new method, we take into account that the value s is already known, i.e., that it is within the given interval $[s, s]$. In this case, the entire discrepancy is corrected by changing only the value s'. Hence, we get the correct value s' in a single iteration.

4. Case of Interval Prior Knowledge: A New Linear Time Algorithm

Motivation: a linear-time algorithm exists for a similar problem of minimizing variance without linear constraints. As we have mentioned, the original Hole's code formulas are related to minimize the variance under a linear constraint (1).

In general, the problem of minimizing variance under interval uncertainty has many other practical applications beyond geophysics. (The only difference is that in most applications, there is no linear constraint similar to (1)). In particular, this general problem has application in geophysics [22, 23].

For this general problem, we have also proposed an $O(c \cdot \log(c))$ algorithm; see, e.g., [15, 16] and references therein.

Recently, we have designed a new algorithm that computes the desired minimum in linear time $O(c)$ [27]. In this paper, we show that a similar linear-time algorithm can be proposed for the case when we want to minimize the variance under an additional linear constraint.

An auxiliary algorithm behind the existing linear-time algorithm. The linear-time algorithm from [27] is based

on the known fact that we can compute the median of a set of n elements in linear time; see, e.g., [7].

The use of median in this algorithm is similar to the one from [6, 10].

A new linear-time algorithm. The proposed algorithm is iterative. At each iteration of this algorithm, we have three sets:

- the set J^- of all the indices j from 1 to c for which we already know that in the desired solution, the corresponding value Δs_{ij} will be cut off (i.e., $\Delta s_{ij} = \overline{\Delta}_j$);
- the set J^+ of all the indices j for which we already know that in the desired solution, the corresponding value Δs_{ij} will *not* be cut off (i.e., $\Delta s_{ij} < \overline{\Delta}_j$);
- the set $J = \{1, \ldots, c\} - J^- - J^+$ of the indices j for which we are still undecided.

In the beginning, $J^- = J^+ = \emptyset$ and $J = \{1, \ldots, c\}$. At each iteration, we also update the values of two auxiliary quantities $A^- \stackrel{\text{def}}{=} \sum\limits_{j \in J^-} \ell_{ij} \cdot \overline{\Delta}_j$ and $\mathcal{L}^+ \stackrel{\text{def}}{=} \sum\limits_{j \in J^+} \ell_{ij}$. In principle, we could compute these values by computing these sums, but to speed up computations, on each iteration, we update these two auxiliary values in a way that is faster than re-computing the corresponding two sums. Initially, since $J^- = J^+ = \emptyset$, we take $A^- = \mathcal{L}^+ = 0$.

At each iteration, we do the following:

- first, we compute the median m of the set J (median in terms of sorting by $\overline{\Delta}_j$);
- then, by analyzing the elements of the undecided set J one by one, we divide them into two subsets

$$P^- \stackrel{\text{def}}{=} \{j : \overline{\Delta}_j \le \overline{\Delta}_m\}, \quad P^+ \stackrel{\text{def}}{=} \{j : \overline{\Delta}_j > \overline{\Delta}_m\};$$

- we compute $a^- \stackrel{\text{def}}{=} A^- + \sum\limits_{j \in P^-} \ell_{ij} \cdot \overline{\Delta}_j$ and

$$\ell^+ \stackrel{\text{def}}{=} \mathcal{L}^+ + \sum_{j \in P^+} \ell_{ij};$$

- then, we compute $\Delta s = \dfrac{\Delta_i - a^-}{\ell^+}$; also, among all the values from P^+, we select the smallest value, which we will denote by $\overline{\Delta}_{(p+1)}$;
- if $\Delta s > \overline{\Delta}_{(p+1)}$, then we replace J^- with $J^- \cup P^-$, A^- with a^-, and J with P^+;
- if $\Delta s \le \overline{\Delta}_m$, then we replace J^+ with $J^+ \cup P^+$, $\mathcal{L}^+$ with ℓ^+, and J with P^-;
- finally, if $\overline{\Delta}_m < \Delta s \le \overline{\Delta}_{(p+1)}$, then we replace J^- with $J^- \cup P^-$, J^+ with $J^+ \cup P^+$, and J with $\emptyset$.

At each iteration, the set of undecided indices is divided in half. Iterations continue until all indices are decided, after which we return $\Delta s_{ij} = \overline{\Delta} s_j$ when $\overline{\Delta}_j \le \overline{\Delta}_m$ and $\Delta s_{ij} = \Delta s$ otherwise.

Proof that the new algorithm for computing $\overline{V}$ requires linear time. At each iteration, computing median requires linear time, and all other operations with J require time t linear in the number of elements $|J|$ of J: $t \le C \cdot |J|$ for some constant C. We start with the set J of size c; on the next iteration, we have a set of size $c/2$, then $c/4$, etc. Thus, the overall computation time is $\le C \cdot (c + c/2 + c/4 + \ldots) \le C \cdot 2c$, i.e., linear in c.

5. Case of Fuzzy Prior Knowledge

Main idea. As we have mentioned, one of the reasons why the mathematically valid solution is not geophysically meaningful is that at some points, the velocity is outside the interval of values which are possible at this depth for this particular geological region.

Additional information provided by experts: general case. To take this expert knowledge into consideration, it is reasonable to explicitly solicit, from the experts, the information about possible values of slownesses – and then modify the inverse algorithms in such a way that the velocities are consistent with this knowledge.

Specifically, for each cell j, a geophysicist provides us with his estimate of possible values of the corresponding slowness s_j. As we have mentioned, an expert often describes this information by using words from the natural language, like "most probably, the value of slowness is within 6 and 7, but it is somewhat possible to have values between 5 and 8". To formalize this knowledge, it is natural to use fuzzy set theory, a formalism specifically designed for describing this type of informal ("fuzzy") knowledge; see, e.g., [4, 8, 14, 21]

As a result, for every cell j, we have a fuzzy set $\mu_j(s)$ which describes the expert's prior knowledge about s_j. For every cell j and for each possible value s_j, the number $\mu_j(s_j)$ describes the expert's degree of certainty that s_j is a possible value of the corresponding slowness.

An alternative user-friendly way to represent a fuzzy set is by using its α-cuts $\{s \,|\, \mu_j(s) > \alpha\}$ (or $\{s \,|\, \mu_j(s) \ge \alpha\}$); see, e.g., [5, 14, 19, 20, 21]. For example, the α-cut corresponding to $\alpha = 0$ is the set of all the values which are possible at all, the α-cut corresponding to $\alpha = 0.1$ is the set of all the values which are possible with degree of certainty at least 0.1, etc. In these terms, a fuzzy set can be viewed as

a nested family of intervals $[\underline{s}_j(\alpha), \overline{s}_j(\alpha)]$ corresponding to different level α.

Comment. For some cells – e.g., in some cells which are close to the surface and for which the geophysical structure is well known – we may even know the *exact* values s_j of slowness. Since a crisp number is a particular case of a fuzzy set, this information can also be expressed in fuzzy terms – by saying that for each of these cells, the geophysicist provides us with a crisp set $\{s_j\}$.

In terms of α-cuts, this means that for every degree α, the corresponding intervals are degenerate intervals $[s_j, s_j]$.

How to use this expert knowledge in solving the seismic inverse problem: precise formulation of the corresponding optimization problem. In general, the solution $(s_1, s_2, \ldots)$ is satisfactory if s_1 is a possible value of slowness in the first cell, and s_2 is a possible value of slowness in the second cell, etc. The corresponding membership functions $\mu_j(s_j)$ describe to what extent s_j is the possible value of slowness in the j-th cell. So, if we use the simplest possible $\min$ operation to describe "and", we conclude that the degree with which a solution is satisfactory can be described by the value $\min(\mu_1(s_1), \mu_2(s_2), \ldots)$.

When we solve the inverse problem, it is reasonable to look for a solution for which this degree of satisfaction is the largest possible: $\min(\mu_1(s_1), \mu_2(s_2), \ldots) \to \max$.

How can we solve this problem: reduction to the case of interval uncertainty. Maximizing the overall degree of satisfaction means that we want to find the largest value α for which $\mu_j(s_j) \geq \alpha$ for all j, i.e., for which, for every cell j, the slowness s_j belongs to the corresponding interval $[\underline{s}_j(\alpha), \overline{s}_j(\alpha)]$.

For each α, we thus face an auxiliary *interval uncertainty problem:* for each cell, we know the corresponding interval, and we want to find a solution to the seismic inverse problem for which all the slownesses are within the corresponding intervals. It is worth mentioning that this interval problem can be of separate practical interest: it is a particular case of the fuzzy uncertainty problem corresponding to the case when the only information coming from an expert is an interval $[\underline{s}_j, \overline{s}_j]$ of possible value of each slowness s_j.

Once we know how to solve this interval problem, we can easily solve the original fuzzy problem as follows. For each $\alpha = 0$, $\alpha = 0.1$, $\alpha = 0.2$, etc., we solve the interval problem with the corresponding intervals $[\underline{s}_j(\alpha), \overline{s}_j(\alpha)]$. Eventually, we will reach such a value of α that the process no longer converges – so the inverse problem with these too narrow interval restriction does not have a solution. 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.

6. 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., [18]. 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

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

So, to find the corrections Δs_{ij}, we must minimize the objective function (variance)

$$V \stackrel{\text{def}}{=} \frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij}^2 - \left(\frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij} \right)^2. \qquad (5)$$

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

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

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

$$\mu \cdot \frac{1}{n} \cdot \sum_{j=1}^{c} \frac{(s_j^{(k)} + \Delta s_{ij} - \widetilde{s}_j)^2}{\sigma_j^2} \to \min. \qquad (6)$$

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

$$\overline{\Delta s} \stackrel{\text{def}}{=} \frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij}. \quad (7)$$

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 (7), 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.

7. Case of Multiple-Type Prior Knowledge

Practical need for prior knowledge: reminder. In many real-life problems, it is difficult or even impossible to directly measure the desired physical quantities. In such situations, we measure other quantities, which are related to the desired ones by known formulas, and then reconstruct the values of the desired quantities from these measurement results.

The reconstructed values of the desired quantities are sometimes outside the range of what an expert would consider reasonable. In such situations, it is desirable to describe the expert's knowledge (about what is reasonable) as a precisely formulated constraint on the desired values, and to incorporate these constraints into the reconstruction process.

In the previous sections, we have shown that different types of expert knowledge can be naturally formalized in interval, fuzzy, and probabilistic terms. We also showed, on the example of the seismic inverse problem, how each of these types of expert knowledge can be used in the solution process.

Practical need for multiple-type prior knowledge. Previously, we (implicitly) assumed that we have only one type of expert knowledge – e.g., only interval knowledge, or only fuzzy knowledge, etc. In some practical situations, however, we may have multiple-type expert knowledge: e.g., one expert provides interval bounds, another expert provides probabilistic knowledge, etc.

This multiple-type prior knowledge is especially important for *cyberinfrastructure*. The main objective of cyberinfrastructure is to be able to seamlessly move data between different databases (where this data is stored in different formats), to feed the combined data into a remotely located program (which may require yet another data format), and to return the result to the user; see, e.g., [1, 12, 25]. It is also important to gauge the quality and accuracy of this result. We often have different models for describing uncertainty of different databases and programs; it is therefore important to be able to consider multiple-type prior knowledge; see, e.g., [9, 17].

How to use multiple-type prior knowledge in the seismic inverse problem. We have mentioned that in the traditional approach, we minimize (5) under the constraint (1). Different types of prior knowledge mean adding constraints on Δs_{ij}. Probabilistic prior knowledge is naturally formalized as a constraint (4), and interval prior knowledge is naturally formalized as a constraint (2). Thus, when both probabilistic and interval prior knowledge are present, we must minimize (5) under the constraints (1), (2), and (4).

If we replace the equality in (4) by an inequality (≤ 1 instead of $= 1$), 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., [26].

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 (5), then we restrict the values to the nearest points from the interval $[\underline{s}_j, \overline{s}_j]$ – to satisfy the constraint (2), and after that, find the extra corrections that satisfy the condition (4), after which we repeat the cycle again until the process converges.

8 Conclusion

The paper deals with the difficult *seismic inverse problem*, in which a 3-D field (velocities of the seismic waves) has to be reconstructed. The classical approach is to transform this problem into a huge non-linear system of equation and to use iterative techniques to solve the problem. Often, the classical approach leads to solutions that are not realistic. However, the expert has an idea of what he should not get and he can express this idea as a set of constraints. The main contribution of the paper is to add these additional knowledge, given by the expert, to the classical approach, inside the iterative method.

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.

The authors are thankful to participants of SCAN'06 for valuable discussions and to the anonymous referees for useful suggestions.

References

[1] R. Aldouri, G. R. Keller, A. Gates, J. Rasillo, L. Salayandia, V. Kreinovich, J. Seeley, P. Taylor, and S. Holloway. GEON: Geophysical data add the 3rd dimension in geospatial studies. In: *Proceedings of the ESRI International User Conference 2004*, San Diego, California, August 9–13, 2004, Paper 1898.

[2] M. G. Averill, K. C. Miller, G. R. Keller, V. Kreinovich, R. Araiza, S. A. Starks. Using expert knowledge in solving the seismic inverse problem. In: *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.

[3] M. G. Averill, K. C. Miller, G. R. Keller, V. Kreinovich, R. Araiza, and S. A. Starks. Using expert knowledge in solving the seismic inverse problem. *International Journal of Approximate Reasoning* (to appear).

[4] G. Bardossy and J. Fodor. *Evaluation of Uncertainties and Risks in Geology*. Springer Verlag, Berlin, 2004.

[5] G. Bojadziev and M. Bojadziev. *Fuzzy Sets, Fuzzy Logic, Applications*, World Scientific, Singapore, 1995.

[6] P. van der Broek and J. Noppen. Fuzzy weighted average: alternative approach. In: *Proceedings of the 25th International Conference of the North American Fuzzy Information Processing Society NAFIPS'2006*, Montreal, Quebec, Canada, June 3–6, 2006.

[7] Th. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms*. MIT Press, Cambridge, MA, 2001.

[8] R. Demicco and J. Klir (Eds.). *Fuzzy Logic in Geology*. Academic Press, New York, 2003.

[9] A. Gates, V. Kreinovich, L. Longpré, P. Pinheiro da Silva, and G. R. Keller. Towards secure cyberinfrastructure for sharing border information. In: *Proceedings of the Lineae Terrarum: International Border Conference*, El Paso, Las Cruces, and Cd. Juárez, March 27–30, 2006.

[10] P. Hansen, M. V. P. de Aragao, and C. C. Ribeiro. Hyperbolic 0-1 programming and optimization in information retrieval. *Math. Programming*. 52:255–263, 1991.

[11] J. A. Hole. Nonlinear high-resolution three-dimensional seismic travel time tomography. *J. Geophysical Research* 97:6553–6562, 1992.

[12] G. R. Keller, T. G. Hildenbrand, R. Kucks, M. Webring, A. Briesacher, K. Rujawitz, A. M. Hittleman, D. R. Roman, D. Winester, R. Aldouri, J. Seeley, J. Rasillo, R. Torres, W. J. Hinze, A. Gates, V. Kreinovich, and L. Salayandia. A community effort to construct a gravity database for the United States and an associated Web portal", In: A. K. Sinha (ed.), *Geoinformatics: Data to Knowledge*, Geological Society of America Publ., Boulder, Colorado, 2006, pp. 21–34.

[13] G. R. Keller, S. A. Starks, A. Velasco, M. Averill, R. Araiza, G. Xiang, and V. Kreinovich. Towards combining probabilistic, interval, fuzzy uncertainty, and constraints: on the example of inverse problem in geophysics. In: *Proceedings of the Second International Conference on Fuzzy Sets and Soft Computing in Economics and Finance FSSCEF'2006*, St. Petersburg, Russia, June 28–July 1, 2006, pp. 47–54.

[14] G. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic*. Prentice Hall, New Jersey, 1995.

[15] V. Kreinovich, L. Longpré, S. A. Starks, G. Xiang, J. Beck, R. Kandathi, A. Nayak, S. Ferson, and J. Hajagos. Interval versions of statistical techniques, with applications to environmental analysis, bioinformatics, and privacy in statistical databases. *Journal of Computational and Applied Mathematics*, 199(2):418–423, 2007.

[16] V. Kreinovich, G. Xiang, S. A. Starks, L. Longpré, M. Ceberio, R. Araiza, J. Beck, R. Kandathi, A. Nayak, R. Torres, and J. Hajagos. Towards combining probabilistic and interval uncertainty in engineering calculations: algorithms for computing statistics under interval uncertainty, and their computational complexity. *Reliable Computing*, 12(6):471–501, 2006.

[17] L. Longpré and V. Kreinovich. How to efficiently process uncertainty within a cyberinfrastructure without sacrificing privacy and confidentiality. In: N. Nedjah, A. Abraham, and L. de Macedo Mourelle (Eds.), *Computational Intelligence in Information Assurance and Security*, Springer-Verlag, Berlin-Heidelberg (to appear).

[18] M. Maceira, S. R. Taylor, C. J. Ammon, X. Yang, A. Velasco. High-resolution Rayleigh wave slowness tomography of Central Asia. *Journal of Geophysical Research*, Vol. 110, paper B06304, 2005.

[19] R. E. Moore and W. A. Lodwick. Interval analysis and fuzzy set theory. *Fuzzy Sets and Systems*, 135(1):5–9, 2003.

[20] H. T. Nguyen and V. Kreinovich. Nested intervals and sets: concepts, relations to fuzzy sets, and applications. In: R. B. Kearfott and V. Kreinovich (eds.). *Applications of Interval Computations*, Kluwer, Dordrecht, 1996, pp. 245–290.

[21] H. T. Nguyen and E. A. Walker. *A First Course in Fuzzy Logic*. CRC Press, Boca Raton, Florida, 2005.

[22] P. Nivlet, F. Fournier, and J. Royer. A new methodology to account for uncertainties in 4-D seismic interpretation. In: *Proc. 71st Annual Int'l Meeting of Soc. of Exploratory Geophysics SEG'2001*, San Antonio, TX, September 9–14, 2001, 1644–1647.

[23] P. Nivlet, F. Fournier, and J. Royer. Propagating interval uncertainties in supervised pattern recognition for reservoir characterization. In: *Proc. 2001 Society of Petroleum Engineers Annual Conf. SPE'2001*, New Orleans, LA, September 30–October 3, 2001, paper SPE-71327.

[24] R. L. Parker. *Geophysical Inverse Theory*. Princeton University Press, Princeton, New Jersey, 1994.

[25] A. K. Sinha (ed.), *Geoinformatics: Data to Knowledge*, Geological Society of America Publ., Boulder, Colorado, 2006.

[26] S. A. Vavasis. *Nonlinear Optimization: Complexity Issues*. Oxford University Press, New York, 1991.

[27] G. Xiang, M. Ceberio, and V. Kreinovich. *Computing Population Variance and Entropy under Interval Uncertainty: Linear-Time Algorithms*. University of Texas at El Paso, Department of Computter Science, Technical Report UTEP-CS-06-28b, November 2006.

[28] C. A. Zelt and P. J. Barton. Three-dimensional seismic refraction tomography: A comparison of two methods applied to data from the Faeroe Basin. *J. Geophysical Research* 103:7187–7210, 1998.