Using Expert Knowledge in Solving the Seismic Inverse Problem

Matthew G. Averill, Kate C. Miller, G. Randy Keller,
Vladik Kreinovich, Roberto Araiza, and Scott A. Starks
Pan-American Center for Earth and Environmental Studies
University of Texas at El Paso
El Paso, TX 79968, USA
averill@geo.utep.edu, miller@geo.utep.edu, keller@utep.edu
vladik@cs.utep.edu, raraiza@cs.utep.edu, sstarks@utep.edu

***Abstract*— In this talk, we analyze how expert knowledge can be used in solving the seismic inverse problem.**

I. Seismic Inverse Problem

To determine the geophysical structure of a region, we measure seismic travel times and reconstruct velocities at different depths from this data. There are several algorithms for solving this inverse problem; see, e.g., [6], [9], [14].

II. Seismic Inverse Problem is an Ill-Posed Problem

The main practical problem with the existing algorithms is that the inverse problem is ill-defined: large changes in the original distribution of velocities can lead to very small changes in the resulting measured values. As a result, based on the same measurement results, we may have many different velocity distributions that are all consistent with the same measurement results.

III. Drawbacks of the Existing Algorithms

Usually, because of this non-uniqueness, the velocity distribution that is returned by the existing algorithm is usually not geophysically meaningful: e.g., it predicts velocities outside of the range of reasonable velocities at this depth. A geophysicist looks at this distribution, and tries to adjust the initial approximation so as to avoid this discrepancy between the actual distribution and the geophysical knowledge.

This adjustment usually requires several iterations. It is a very time-consuming process, because there is no algorithmic way of adjusting the initial data, only heuristic recipes, and as a result, each adjustment requires many time-consuming trial-and-error steps. Moreover, because of the non-algorithmic character of adjustment, it requires special difficult-to-learn skills; as a result, the existing tools for solving the seismic inverse problem are not as widely used as they could be.

IV. It Is Necessary to Take Expert Knowledge Into Consideration

To enhance the use of the seismic data, it is imperative to make the corresponding tools more accessible and their handling more algorithmic. To achieve this goal, we must incorporate the expert knowledge into the algorithm for solving the inverse problem.

V. Explicit Expert Knowledge: Interval Uncertainty

As we have mentioned, one of the reasons that 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. To take this expert knowledge into consideration, we can simply describe these intervals of possible data and modify the inverse algorithms in such a way that the velocities are always within these intervals.

So, first thing we do is modify the inverse algorithm in such a way as to take this interval uncertainty into consideration. How can we do it?

VI. How We Can Use Interval Uncertrainty

Most algorithms for solving the seismic inverse problem start with a geophysically reasonable first approximation to the velocity model. This approximate model usually comes from the prior geophysical knowledge of the area – or, if we do not have much information about this particular area, from the general idea of how velocities increase with depth.

Based on a given approximate model, we then predict the traveltimes between different stations, and use the difference $\Delta x_i \stackrel{\rm def}{=} \widetilde{x}_i - x_i$ between the measured traveltimes $\widetilde{x}_i$ and the predicted traveltimes x_i to adjust the velocity model. For example, in Hole's algorithm [6], for each traveltime, we divide each difference Δx_i by the length L of the corresponding path and add the resulting value $\Delta x_i/L$ to the slownesses $s \stackrel{\rm def}{=} 1/v$ of all the cells along this path.

After this adjustment, we may perform some smoothing or filtering, and then we get a new approximation to the velocity model. We then repeat the same adjustment process starting with this new approximation, etc.

Suppose now that at each pixel j, we know the interval $[\underline{v}_j, \overline{v}_j]$ of geophysically possible velocities – or, equivalently, the interval $[\underline{s}_j, \overline{s}_j]$ of possible slownesses, where $\underline{s}_j = 1/\overline{v}_j$ and $\overline{s}_j = 1/\underline{v}_j$. We can incorporate this information into the existing algorithms as follows: on k-th iteration, once we get the next approximation $s_j^{(k)}$ to the slownesses, we check, for every j, whether the corresponding value is within the desired interval, and if not, we replace it with the nearest value within this interval. In other words:

- if $s_j^{(k)} < \underline{s}_j$, we replace this value with $\underline{s}_j$;
- if $s_j^{(k)} > \overline{s}_j$, we replace this value with $\overline{s}_j$;
- if $\underline{s}_j \le s_j^{(k)} \le \overline{s}_j$, we keep the value $s_j^{(k)}$.

After this additional step, we perform the next iteration, etc.

VII. Explicit Expert Knowledge: Fuzzy Uncertainty

Experts are often not 100% sure about the corresponding intervals. They can usually produce a wider interval 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.

VIII. How We Can Use Fuzzy Uncertainty

How can we incorporate the fuzzy information into the inverse method? A natural way to do it is as follows: instead of simply getting a solution in which all the slownesses belong to the guaranteed (wide) intervals corresponding to $\alpha = 0$, we try to find the largest possible value α for which all the slownesses belong to the corresponding (narrower) α-cuts.

We can find such α, e.g., by simply trying $\alpha = 0$, $\alpha = 0.1$, $\alpha = 0.2$, 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.

Comment. What we have just described is the basic straightforward way to take fuzzy-valued expert knowledge into consideration. Several researchers have provided other ideas for successfully using fuzzy expert knowledge in geophysical problems; see, e.g., [2], [4], [10] and references therein. We plan to add some of these ideas to our modified algorithms.

IX. Implicit Expert Knowledge: Interval Uncertainty

In other cases, for each 3-D point, the reconstructed velocity is within the corresponding interval, but the geophysical structure is still nor reproduced right. In such cases, it is difficult to explicitly describe, to a computer system, what exactly is wrong, but often, we can describe it implicitly. Namely, the seismic inverse algorithm – like many other algorithms for solving the inverse problem – is based on the assumption that the measured errors are independent and normally distributed. As a result, as a criterion of how well the velocity model fits the measurement results, these algorithms use of mean square error

$$E \stackrel{\rm def}{=} \sum_{i=1}^{N} (x_i - \widetilde{x}_i)^2,$$

where N is the overall number of measured traveltimes, x_i is the i-th traveltime according to the model, and $\widetilde{x}_i$ is the measured traveltime.

For geophysically adequate reconstructions, this mean square error is indeed reasonably small, and the individual differences $x_i - \widetilde{x}_i$ are indeed more or less normally distributed. On the other hand, for geophysically meaningless models, while the mean square error E is also small, several individual differences $|x_i - \widetilde{x}_i|$ are very large in comparison with the others – so that the resulting empirical distribution of these differences is far from normal.

To avoid this problem, it is desirable to require not only that the mean square error be small, but that all individual differences $|x_i - \widetilde{x}_i|$ be small as well. Ideally, we should have an exact upper bound Δ on such a difference, and dismiss a solution as non-physical if at least one of the differences exceeds Δ.

X. How We Can Use Interval Uncertainty

How can we guarantee that we only get solutions which are physical in the above sense?

- Traditionally, once the mean square error is small, we stop iterations.
- Instead, we propose to continue iterations until all the differences are under Δ.

If this does not happen, then we need to do what traditional algorithms do if we do not get a convergence – restart computations with a different starting velocity model.

XI. Implicit Expert Knowledge: Fuzzy Uncertainty

Experts cannot always provide us with exact upper bounds Δ; instead, based on the expert's experience of solving inverse problems, we can have different bounds with different degrees of certainty – i.e., again, in effect, a fuzzy number as an upper bound.

XII. How We Can Use Fuzzy Uncertainty

How can we use this fuzzy information? A natural idea is – like in the use of explicit expert knowledge – to find the largest α for which we can decrease all the differences $x_i - \widetilde{x}_i$ into the corresponding α-cut intervals. Similar to the case of explicit knowledge, we can do it, e.g., by trying $\alpha = 0$, $\alpha = 0.1$, etc.

XIII. Need for Data Fusion

Yet another way to detect velocity models that are not geophysically reasonable is to take into consideration other geophysical and geological data, such as the gravity map, the geological map, etc. – in other words, fuse several different types of data.

XIV. Preliminary Results

Preliminary results of such fusion are indeed very encouraging; see, e.g., [1].

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209 and by NSF grants EAR-0112968, EAR-0225670, and EIA-0321328.

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", *Proceedings of the ESRI International User Conference 2004*, San Diego, California, August 9–13, 2004, Paper 1898.

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

[3] S. D. Cabrera, K. Iyer, G. Xiang, and V. Kreinovich, "On Inverse Halftoning: Computational Complexity and Interval Computations", *Proceedings of the 39th Conference on Information Sciences and Systems CISS'2005*, John Hopkins University, March 16–18, 2005, Paper 164.

[4] R. Demicco and G. Klir (eds.), *Fuzzy Logic in Geology*, Academic Press, 2003.

[5] D. I. Doser, K. D. Crain, M. R. Baker, V. Kreinovich, and M. C. Gerstenberger "Estimating uncertainties for geophysical tomography", *Reliable Computing*, 1998, Vol. 4, No. 3, pp. 241–268.

[6] J. A. Hole, "Nonlinear High-Resolution Three-Dimensional Seismic Travel Time Tomography", *J. Geophysical Research*, 1992, Vol. 97, No. B5, pp. 6553–6562.

[7] R. B. Kearfott and V. Kreinovich, "Beyond Convex? Global Optimization Is Feasible Only for Convex Objective Functions: A Theorem", *Journal of Global Optimization* (to appear).

[8] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational complexity and feasibility of data processing and interval computations*, Kluwer, Dordrecht, 1997.

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

[10] M. Nikravesh, "Soft computing-based computational intelligence for reservoir characterization", *Expert Syst. Appl.*, 2004, Vol. 26, No. 1, pp. 19–38.

[11] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, Dover Publications, Inc., Mineola, New York, 1998.

[12] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, W. H. Whinston & Sons, Washington, D.C., 1977.

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

[14] 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*, 1998, Vol. 103, No. B4, pp. 7187–7210.

Appendix: Computational Complexity of the Seismic Inverse Problem and Why Traditional Methods of Solving Inverse Problems Do Not Work Well in the Seismic Case

A.I. Most Inverse Problems in Science and Engineering are Ill-Posed

The above ill-posedness of the seismic problem is a common feature in applications: most inverse problems in science and engineering are ill-posed; see, e.g., [12].

A.II. Smoothness: Traditional Approach to Solving Ill-Posed Inverse Problems

A typical way to solve an inverse problem is to find a natural physically meaningful property of actual solution, and use this *a priori* information to select a single most physically meaningful solution among many mathematically possible ones. This process is called *regularization.*

Typically, in inverse problems, this natural property is smoothness. Smoothness can be naturally described in precise mathematical terms. For example, when we reconstruct a 1-D signal $x(t)$, then the degree of smoothness can be defined as follows. At a given moment of time t, the larger the absolute value $|x'(t)|$ of the derivative $x'(t)$, the less smooth the signal is. Thus, at a given time t, the value $|x'(t)|$ is a natural degree of the signal's non-smoothness. Overall, a natural degree of non-smoothness can be defined as a mean square of these degrees corresponding to different moments t, i.e., as $J \stackrel{\text{def}}{=} \int (x'(t))^2\,dt$.

Most regularization techniques try to find, among many signals that are consistent with given observations, the smoothest signal, i.e., the signal with the smallest possible value of the degree of non-smoothness J.

A.III. Smoothness: Discrete Case

In real life, we only have the values

$$x(t_1), x(t_2), \ldots,$$

of the signal $x(t)$ at discrete moment of time

$$t_1, t_2 = t_1 + \Delta t, \ldots, t_{i+1} = t_i + \Delta t, \ldots$$

Based on this discrete data, we can approximate the derivative $x'(t)$ as a difference

$$\frac{x(t_{i+1}) - x(t_i)}{\Delta t},$$

so minimizing the integral J is equivalent to minimizing the corresponding integral sum

$$J_{\text{discr}} \stackrel{\text{def}}{=} \sum_i (x(t_{i+1}) - x(t_i))^2.$$

A.IV. Smoothness: 2-D Case

For a 2-D velocity distribution $f(n_1, n_2)$, similarly, a natural assumption is that this distribution is smooth. Similarly to the 1-D case, a natural way to describe the degree of smoothness of a given distribution is to use the integral sum

$$J \stackrel{\text{def}}{=} \sum_{n_1, n_2} s(n_1, n_2),$$

where

$$s(n_1, n_2) \stackrel{\text{def}}{=}$$

$$(f(n_1+1, n_2) - f(n_1, n_2))^2 + (f(n_1, n_2+1) - f(n_1, n_2))^2.$$

Alternatively, we can describe this criterion as the sum of the squares of the differences in intensity between all possible pairs (p, p') of neighboring pixels $p = (n_1, n_2)$ and $p' = (n'_1, n'_2)$:

$$J = \sum_{p,p' \text{ are neighbors}} (f(p) - f(p'))^2.$$

A.V. Smoothness Makes Problems Computationally Solvable

A practically useful property of the above degrees of non-smoothness J is that the expression J is a convex function of the signal $x(t_i)$ or $f(n_1, n_2)$. Thus, if the conditions describing the fact that the unknown velocity distributions is consistent with the observations is also described by linear or, more generally, smooth inequalities, then the problem of finding the regularized solution can be reformulated as a problem of minimizing a convex function J on the convex set.

Similarly, if we fix the degree of non-smoothness and look, among all the solutions with a given degree of non-smoothness, for the one that is the closest to the original approximate solution, we also have a problem of minimizing a convex function (distance) on the convex set (of all functions that are consistent with the observations and have the desired degree of smoothness).

It is known that, in general, the problems of minimizing convex functions over convex domains are algorithmically solvable (see, e.g., [13]), and smoothness-based regularization has indeed been efficiently implemented; see, e.g., [12].

A.VI. For the Seismic Inverse Problem, We Only Have Piecewise Smoothness

In geophysics, we have clear layers of different types of rocks, with sharp difference between different layers, so we face an inverse problem with only piecewise smoothness; see, e.g., [9].

A.VII. Traditional Smoothness Measures are Not Adequate for Piecewise Smoothness

In the piecewise smooth case, the above measure of non-smoothness is not applicable, because it would include neighboring pixels on the different sides of the border between the two layers.

A.VIII. Appropriate Smoothness Measures for Piecewise Smoothness Case

To avoid the above problem, we need to only take into account the pairs of neighboring pixels that belong to the same zone (layer), i.e., consider the sum

$$J(Z) = \sum_{p,p' \text{ are neighbors in the same zone}} (f(p) - f(p'))^2,$$

where Z denotes the information about the zones. This measure makes computational sense only if we know beforehand where the zones are – i.e., where is the border between the two zones.

However, in real life, finding the border is a part of the problem. In this case, we can use the same smoothness criterion not only to reconstruct the original velocity distribution, but also to find the border location. Specifically, we want to look for the zone distribution *and* for the zone location for which the above criterion J takes the smallest possible value.

In other words, we fix the number of zones, and we characterize the non-smoothness of an velocity distribution by a criterion

$$J^* = \min_{\text{all possible divisions } Z \text{ into zones}} J(Z).$$

A.IX. The Resulting Problem is No Longer Convex

The resulting functional is no longer convex, because the division into zones is a discrete problem. It is known that non-convex problems are, in general, more computationally difficult than the corresponding convex ones (see, e.g., [7]), and adding discrete variables makes the problems even more computationally difficult; see, e.g., [11].

A.IX. Complexity of Piecewise Smooth Inverse Problems

In the following sections, we show that in general, the inverse problem for piecewise smooth case is computationally intractable (NP-hard) even when the relation expressing the consistency between the measured results and the desired velocity distribution is linear.

This proof will follow the proof of NP-hardness of different signal processing problems described in our previous publications [3], [8].

Let us prove that in general, the inverse problem for piecewise smooth case is computationally intractable (NP-hard).

A.X. Main Idea of the Proof: Reduction to a Subset Problem

To prove NP-hardness of our problem, we will reduce a known NP-hard problem to the problem whose NP-hardness we try to prove: namely, to the inverse problem for piecewise smooth velocity distributions.

Specifically, we will reduce, to our problem, the following *subset sum* problem [8], [11] that is known to be NP-hard:

- Given:
  - m positive integers $s_1, \ldots, s_m$ and
  - an integer $s > 0$,
- check whether it is possible to find a subset of this set of integers whose sum is equal to exactly s.

For each i, we can take $x_i = 0$ if we do not include the i-th integer in the subset, and $x_i = 1$ if we do. Then the subset problem takes the following form: check whether there exist values $x_i \in \{0, 1\}$ for which

$$\sum s_i \cdot x_i = s.$$

We will reduce each instance of this problem to the corresponding piecewise smooth inverse problem.

A.XI. Reduction to a Subset Problem: Details

Let us consider the following problem. We want to reconstruct an $m \times m$ velocity distribution $f(n_1, n_2)$. Let $d = \lfloor m/2 \rfloor$. We want a piecewise smooth velocity distribution $f(n_1, n_2)$ that consists of two zones.

The following linear constraints describe the consistency between the observations and the desired velocity distribution:

- $f(n_1, n_2) = 1$ for $n_2 > d$;
- $\sum\limits_{i=1}^{m} s_i \cdot f(i, d) = s$; and
- $f(n_1, n_2) = 0$ for $n_2 < d$.

The problem that we consider is to find the solution with the smallest possible value of smoothness J^* among all the velocity distributions that satisfy these linear constraints.

Let us show that the minimum of J^* is 0 if and only if the original instance of the subset problem has a solution.

Indeed, if J^* is 0, this means that all the values within each zone must be the same. Since we have values 1 for $n_2 > d$ and values 0 for $n_2 < d$, we must therefore have every value to be equal either to 0 or to 1. Thus, if we have such a solution, the corresponding values $f(i, d) \in \{0, 1\}$ provide the solution to the original subset problem $\sum s_i \cdot x_i = s$.

Vice versa, if the selected instance of the original subset problem has a solution x_i, then we can take $f(i, d) = x_i$ and get the solution of the inverse problem for which the degree of non-smoothness is exactly 0.

So, if we can solve the inverse problem for piecewise smooth velocity distributions, we will thus be able to solve the subset sum problem.

This reduction proves that the inverse problem for piecewise smooth velocity distributions is indeed NP-hard.