Large-Scale Inversion in Exploration Seismology

By Tristan van Leeuwen

Seismic data offer a rich source of information about the subsurface of the earth. By studying its dynamic and kinematic properties, researchers can infer large-scale variations as well as rock properties on a local scale. Seismic measurements for exploration purposes are typically acquired by placing receivers (geophones) on the surface and detonating an explosive source, as seen in Figure 1. This procedure is repeated for various locations, resulting in a large volume of data. This is a typical *multi-experiment* setting, meaning multiple data-sets are collected for a single set of parameters. The rock properties are parametrized in the subsurface by m and the experiment is simulated by solving a linear wave equation $\mathcal{L}[m]u_i = q_i$, where $i = 0,1,\ldots,\ k$ is the experiment index, q_i represents the explosive source, and $\mathcal{L}$ is a differential operator.

The introduction of a linear operator, $\mathcal{P}$ that maps the solution u_i to the measurements formally poses the inverse problem as follows: *For given measurements* d_i, *determine the coefficients m and solutions* u_i *such that* $\mathcal{P}u_i \approx d_i$ *and* $\mathcal{L}(m)u_i = q_i$ *for* $i = 1,2,\ldots k$.

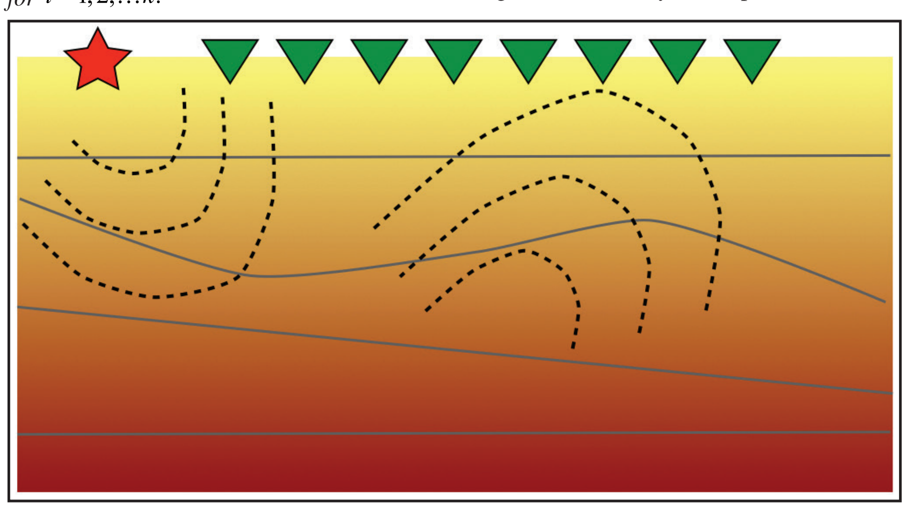

Figure 1. *Schematic depiction of the acquisition process. The seismic source is indicated with a ★ while the receivers are indicated with a ▼.*

Numerically solving the PDEs readily eliminates the u_i and obtains a high-dimensional (m may represent up to 10^9 parameters) nonlinear least-squares problem with k terms:

$$\min_m \sum_{i=1}^{k} \|\mathcal{P}u_i - d_i\|_2^2,$$

where $\mathcal{L}(m)u_i = q_i$ [5]. In principle, any black-box optimization method can be used to solve the resulting optimization problem. Due to the computational cost and severe nonlinearity, however, the seismic problem is not amenable to a black-box approach. The key to developing a better approach is considering the interplay between the formulation, the optimization algorithm, the multi-experiment nature of the data, and the means of (numerically) solving the wave equation. These aspects are traditionally different disciplines' areas of expertise (e.g., statistics, computer-science, machine learning, and numerical analysis), making this a very exciting problem for multidisciplinary research.

The leading computational cost lies in solving the wave equation for all k experiments, where k is potentially very large (easily $k \sim 10^6$). Thus, one can only perform a few iterations to obtain an approximate solution of the optimization problem. Additionally, the severely nonlinear relation between the parameters and the data requires a very good initial parameter estimate. If the initial guess is not 'close' to the true parameters in some sense, the optimization may converge to a local minimum. Failure to find a global minimum is often very hard to detect. The industry therefore spends a considerable amount of time and effort constructing a suitable initial estimate and performing subsequent quality control, both of which involve much specialized manual interference.

An ideal situation would involve running an inversion multiple times from a suite of initial guesses and quantifying the uncertainty of the final result. While mathematical techniques to perform such uncertainty quantification for inverse problems are well established, they often rely on some form of Monte Carlo (MC) sampling. However, the high dimensionality of the problem at hand and the computational cost involved in running even a single simulation prohibit the use of such techniques. My recent research aims at tackling this challenge on multiple fronts:

- Reformulation of the conventional least-squares problem to one that is less nonlinear in the parameters, making the approach less sensitive to the quality of the initial guess [4, 7].
- Reduction of the dimensionality of the data, replacing the full dataset with k terms by a subsampled dataset with $k' \ll k$ terms [6].
- Better quantification of the uncertainty by estimating statistics of the noise and other auxiliary parameters [1].

Reformulations

Many reformulations of the seismic inverse problem have been proposed over the years. One class of reformulations uses a different distance metric to measure the difference between the observed and simulated data. This is very useful when certain features of the data are of primary interest, or when the data contain large outliers.

Another line of research focuses on relaxing the physics and putting more emphasis on fitting the data. The conventional approach insists on obeying the physics for a given set of parameters by solving the PDE $\mathcal{L}(m)u_i = q_i$ and finding the parameters such that u_i fits the data. Instead, we can relax the constraints and construct solutions u_i that fit the data but fail to obey the physics, i.e., $\mathcal{L}(m)u_i \approx q_i$. The goal is now to find parameters m so the solution u_i obeys the physics. We state the problem here as follows: *For given measurements* d_i, *determine the coefficients m such that* $\mathcal{P}u_i \approx d_i$ *and* $\mathcal{L}(m)u_i \approx q_i$ *for* $i = 1,2,\ldots k$.

Such approaches, which place the data and physics on equal footing, are well-known in data assimilation but new in inverse problems. They can be used to solve the original inverse problem while being less sensitive to the initial guess.

Dimensionality reduction

We can formulate a *multi-experiment* inverse problem generically as

$$\min_m \frac{1}{k}\sum_{i=1}^{k} f_i(m),$$

where f_i measures the data fit for given parameters m. Evaluation of a single f_i requires the solution of a PDE which constitutes the dominant computational cost when solving this optimization problem. The idea is to replace the objective by an unbiased approximation

$$\frac{1}{k}\sum_{i=1}^{k} f_i(m) \approx \frac{1}{|I|}\sum_{i\in I} f_i(m),$$

where $I \subset \{1,2,\ldots,k\}$ is a randomly-chosen subset of size $|I| \ll k$. Using a relatively small number of terms can obtain very good results and lead to an order of magnitude speedup. To guarantee convergence to a solution of the full problem, special optimization techniques have to be developed. Of special interest are techniques to adaptively choose the number of samples based on the required accuracy [2].

Estimating nuisance parameters

Many formulations of the inverse problem involve additional *nuisance* parameters that may not be of primary interest but are crucial for finding a meaningful reconstruction. Such parameters include calibration weights or characteristics of noise, such as variance. Solving for these additional parameters alongside the primary ones leads to a *bi-level* optimization problem

$$\min_{m,w} f(m,w).$$

Rather than solve this as a generic nonlinear optimization problem, a more attractive approach involves introducing an optimal value function $\overline{f}(m) = \min_w f(m,w)$ and solving a reduced problem for m alone. In many instances optimization in w is easy, and it turns out that the derivatives of $\overline{f}$ with respect to m do not involve derivatives of w with respect to m. In particular, $\partial_m \overline{f}(m) = \partial_m f(m,\overline{w})$ where $\overline{w}$ is the optimal w, implicitly defined through $\partial_w f(m,w) = 0$. Employing the chain rule easily verifies the latter statement, but similar statements can be made when f is not smooth in w (under suitable assumptions). This results in an extremely powerful framework for handling additional parameters in the context of large-scale optimization.

The aforementioned challenges of the seismic inverse problem call for unconventional reformulations of the inverse problem and new computational techniques to handle the large amounts of data. While some of the challenges are unique to exploration seismology, other issues are generally encountered in inverse problems with wave equations. Being able to quantify the uncertainty in the solution is important in many applications. Together with faster methods to solve the inverse problem, this may ultimately lead to computationally-feasible approaches for uncertainty mitigation.

References

[1] Aravkin, A., & van Leeuwen, T. (2012). Estimating nuisance parameters in inverse problems. *Inverse Problems, 28*(11), 115016.

[2] Aravkin, A., Friedlander, M.P., Herrmann, F.J., & van Leeuwen, T. (2012). Robust inversion, dimensionality reduction, and randomized sampling. *Mathematical Programming, 134*(1), 101–125.

[3] Martin, J., Wilcox, L.C., Burstedde, C., & Ghattas, O. (2012). A Stochastic Newton MCMC Method for Large-Scale Statistical Inverse Problems with Application to Seismic Inversion. *SIAM Journal on Scientific Computing, 34*(3), A1460–A1487.

[4] Symes, W.W. (2014). Seismic inverse problems: recent developments in theory and practice. In A.K. Louis, S. Arridge, & B. Rundell (Eds.) *Proceedings of the Inverse Problems from Theory to Applications Conference* (pp. 2-6). Bristol, UK: IOP Publishing.

[5] Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics, 49*(8), 1259–1266.

[6] Van Leeuwen, T., & Herrmann, F.J. (2014). 3D Frequency-Domain Seismic Inversion with Controlled Sloppiness. *SIAM Journal on Scientific Computing, 36*(5), S192–S217.

[7] Van Leeuwen, T., & Herrmann, F.J. (2016 January). A penalty method for PDE-constrained optimization in inverse problems. *Inverse Problems, 32*(1), 015007.

Tristan van Leeuwen is an assistant professor in the Department of Mathematics at Utrecht University, where he researches inverse problems in imaging. In July 2015 Tristan received the SIAM Geosciences Junior Geoscientist prize for his research on computational methods for seismic inversion.