Released to public domain under Creative Commons license type BY (https://creativecommons.org/licenses/by/4.0).
Copyright (c) 2019 SLIM group @ The University of British Columbia.

Seismic Laboratory of Imaging and Modeling (SLIM)

A numerical solver for least-squares sub-problems in 3D wavefield reconstruction inversion (WRI) and related problem formulations

Bas Peters & Felix J. Herrmann

Introduction

Regular full-waveform inversion (FWI) is based on the minimization of a non-linear least-squares problem. During the past six years, various alternative formulations based on a quadratic penalty formulation have been proposed [2]:

$$\phi_\lambda(\mathbf{m},\mathbf{u}) = \frac{1}{2}\|\mathbf{H}(\mathbf{m})\mathbf{u}-\mathbf{q}\|_2^2 + \frac{\lambda^2}{2}\|\mathbf{P}\mathbf{u}-\mathbf{d}\|_2^2$$

Various researchers claim less dependence on the initial guess (starting model) in certain circumstances. Examples include Wavefield Reconstruction Inversion (WRI, [2]), as well as algorithms based on a data-constrained formulation [1,4]

$$\min_{\mathbf{m},\mathbf{u}} \|\mathbf{H}(\mathbf{m})\mathbf{u}-\mathbf{q}\|_2^2 \quad \text{s.t.} \quad \|\mathbf{P}\mathbf{u}-\mathbf{d}\|_2^2 \leq \gamma$$

(PDE misfit) (data misfit)

The least-squares sub-problem

To avoid storing all fields in memory, we typically do not solve the above problems directly but use alternating minimization or variable projection. For both of the above problems, these approaches require the solution of

$$\mathbf{H}(\mathbf{m})^*\big(\mathbf{H}(\mathbf{m})\mathbf{u}-\mathbf{q}\big) + \lambda^2\mathbf{P}^*\big(\mathbf{P}\mathbf{u}-\mathbf{d}\big) = 0 \quad \Longleftrightarrow \quad \min_{\mathbf{u}} \|\mathbf{H}\mathbf{u}-\mathbf{q}\|_2^2 \quad \text{s.t.} \quad \|\mathbf{P}\mathbf{u}-\mathbf{d}\|_2^2 \leq \gamma$$

which is equivalent to

$$\bar{\mathbf{u}} = \arg\min_{\mathbf{u}} \left\| \begin{pmatrix} \mathbf{H}(\mathbf{m}) \\ \lambda\mathbf{P} \end{pmatrix} \mathbf{u} - \begin{pmatrix} \mathbf{q} \\ \lambda\mathbf{d} \end{pmatrix} \right\|_2^2$$

A solution is possible at the cost of 1 PDE solve per source + 1 PDE solve per receiver.

We derive Algorithm 1 by 1) reformulating the normal equations as a low-rank + identity form; 2) define $\mathbf{W} \equiv \mathbf{H}^{-*}\mathbf{P}^*$ and 3) use the Sherman-Morrison-Woodbury relation. This is similar to an inexact version of the Delves-Barrodale algorithm [5].

Algorithm 1 Least-squares sub-problem

input:
$\mathbf{H} \in \mathbb{C}^{N\times N}$, $\mathbf{q} \in \mathbb{C}^N$, $\mathbf{P} \in \mathbb{R}^{M\times N}$, $\mathbf{d} \in \mathbb{C}^M$
for each receiver (j) (row in $\mathbf{P}$) in parallel **do**
  $\mathbf{H}^*\mathbf{w}_j = \mathbf{p}_j^*$ {solve 1 PDE}
end for
$\mathbf{W} = [\mathbf{w}_1\, \mathbf{w}_2 \dots \mathbf{w}_m]$ {distributed matrix}
$\mathbf{S} = (\mathbf{I}_M + \lambda^2\mathbf{W}^*\mathbf{W})^{-1}$
{adjust λ using Algorithm 2 (optional)}
for source (i) in parallel **do**
  $\mathbf{y}_i = (\mathbf{I}_N - \lambda^2\mathbf{W}\mathbf{S}\mathbf{W}^*)(\mathbf{q}_i + \lambda^2\mathbf{W}\mathbf{d}_i)$
  $\mathbf{H}\mathbf{u}_i = \mathbf{y}_i$ {solve 1 PDE}
end for
output: $\mathbf{u}_i$ {fields for all sources}

The $\gamma - \lambda$ relation

The penalty sub-problem is equivalent to the constrained sub-problem for a certain $\gamma - \lambda$ pair [3]

This connection pair is generally not known in advance. Various numerical methods exist [6] to compute λ, given γ, but they are iterative with each iteration typically as costly as solving the entire least-squares problem without direct factorizations. The core of the method is:
1) form Lagrangian for the constrained problem above:

$$\mathcal{L}(\mathbf{u}) = \|\mathbf{H}\mathbf{u}-\mathbf{q}\|_2^2 + \mu(\|\mathbf{P}\mathbf{u}-\mathbf{d}\|_2^2 - \gamma^2)$$

2) then note that the first-order optimality condition

$$(\mathbf{H}^*\mathbf{H} + \mu\mathbf{P}^*\mathbf{P})\mathbf{u}_\mu = \mathbf{H}^*\mathbf{q} + \mu\mathbf{P}^*\mathbf{d},$$

3) is satisfied using a penalty parameter that is found as a root of

$$f(\mu) = \|\mathbf{P}\mathbf{u}_\mu - \mathbf{d}\|_2 - \gamma$$

where $\mathbf{u}_\mu$ satisfies the optimality condition above.

Algorithm 2 (secant) to obtain the penalty parameter λ corresponding to data fit constraint parameter γ using precomputed quantities.

Input: $\mathbf{W}$, $\mathbf{X} = \mathbf{W}^*\mathbf{W}$, γ, $\mathbf{v} \equiv \mathbf{W}^*\mathbf{q} - \mathbf{d}$
$f(\lambda) \equiv \|\mathbf{S}(\lambda)\mathbf{v}\|_2 - \gamma$, $\mathbf{S}(\lambda) = (\mathbf{I}_M + \lambda_k^2\mathbf{X})^{-1}$
$\lambda_0 = 0$, $\lambda_1 \in (0, \lambda_*)$, $k = 1$
$\mathbf{S}_0(\lambda) = (\mathbf{I}_M + \lambda_0^2\mathbf{X})^{-1}$,
while $\|\mathbf{S}_k(\lambda)\mathbf{v}\|_2 > \gamma$ **do**
  $\lambda_{k+1} = \lambda_k - \frac{f(\lambda_k)(\lambda_k - \lambda_{k-1})}{f(\lambda_k) - f(\lambda_{k-1})}$
  $k \leftarrow k+1$
end while
Output: λ

Our primary contribution is constructing Algorithm 2 and merging it into Algorithm 1. Given a constraint on the data fit, Algorithm 2 finds the corresponding penalty parameter using root-finding with pre-computed quantities and one very small matrix inversion. Some manipulations lead to the following important identity that enables Algorithm 2 at an insignificant computational cost:

$$\mathbf{P}\mathbf{u}_\mu - \mathbf{d} = \mathbf{S}(\mathbf{W}^*\mathbf{q} - \mathbf{d})$$

Conclusions

We propose an algorithm that solves least-squares sub-problems for WRI and data-constrained FWI at the cost of 1 PDE solve per source + 1 PDE solve per receiver in parallel for each source and receiver. Our algorithm automatically finds the penalty-constraint correspondence at virtually no additional cost.

References

1.Wang, R., and F. Herrmann, 2017, A denoising formulation of full-waveform inversion.
2.van Leeuwen, T., and F.J. Herrmann, 2013, Mitigating local minima in full-waveform inversion by expanding the search space.
3. Gander, W., 1980, Least squares with a quadratic constraint
4.Fu, L., and W.W. Symes, 2017, A discrepancy-based penalty method for extended waveform inversion,
5.LM Delves and I Barrodale, 1979, A fast direct method for the least squares solution of slightly overdetermined sets of linear equations.
6.Bjorck, A.A., 1996, Numerical methods for least squares problems

THE UNIVERSITY OF BRITISH COLUMBIA
Georgia Tech