Analyzing the role of parametrization in elastic full waveform inversion.

Gian Matharu and Mauricio D. Sacchi, Department of Physics, University of Alberta, gian@ualberta.ca

Summary

Elastic full waveform inversion (FWI) is inherently more ill-posed and non-linear than its acoustic counterpart owing to the increased model space and the introduction of more complex wave phenomena. Elastic media can be characterized by as few as three parameters under simplifying assumptions (i.e. isotropy, perfect elasticity). Parameters are unlikely to be uniquely resolvable as perturbations in multiple parameters can generate similar attributes in the data; this is known as parameter trade-off. The data-misfit Hessian can mitigate parameter trade-offs; however, its explicit computation is not feasible for problems of realistic size. Early studies in FWI examined analytical diffraction patterns generated by localized perturbations to evaluate parameter trade-offs. While this approach has been corroborated by empirical results, the analysis is somewhat qualitative. As multi-parameter inversion becomes more prevalent in FWI applications, addressing the resolvability of individual parameters will be essential. In this study we perform synthetic inversions to assess the impact of model parametrization on convergence rates and resolvability.

Introduction

Full waveform inversion is a data fitting technique that seeks to minimize a difference measure between observed and synthetic data (Tarantola, 1986; Mora, 1987; Virieux and Operto, 2009). Recent advances in computational hardware and high performance computing have led to a resurgence in FWI research. FWI is an ill-posed inverse problem owing to the non-uniqueness of inverted solutions. The ill-posedness is exacerbated when multiple independent parameters are inverted for. This study presents preliminary work intended to further understanding on multi-parameter FWI in the absence of the Hessian. We outline FWI theory before presenting results from synthetic inversions.

Theory

FWI can be represented as a P.D.E. constrained optimization problem of the form

$$\begin{aligned} &\underset{\mathbf{m}}{\text{minimize}} && J(\mathbf{m}) = \frac{1}{2}\|\mathbf{u}(\mathbf{m}) - \mathbf{u}^0\|^2 \\ &\text{subject to} && \mathbf{L}(\mathbf{m})\mathbf{u} = \mathbf{f}. \end{aligned}$$

Given an estimated model **m**, we seek to minimize the objective function J(**m**) subject to a P.D.E constraint. $\mathbf{u}^0$ denotes the observed data and **u**(**m**) denotes synthetic data generated using estimated model **m**. **L(m)** represents a linear differential operator that characterizes the acoustic or elastic wave-equation. The optimization problem can be solved via linearized inversion (Tarantola, 1986; Mora, 1987; Pratt, 1999). A second order Taylor expansion of J(**m**) about **m** results in

$$J(\mathbf{m}+\delta\mathbf{m}) = J(\mathbf{m}) + \left(\frac{\partial J(\mathbf{m})}{\partial \mathbf{m}}\right)^T \delta\mathbf{m} + \frac{1}{2}\delta\mathbf{m}^T \frac{\partial^2 J(\mathbf{m})}{\partial \mathbf{m}^2}\delta\mathbf{m} = J(\mathbf{m}+\delta\mathbf{m}) = J(\mathbf{m}) + \mathbf{g}^T\delta\mathbf{m} + \frac{1}{2}\delta\mathbf{m}^T\mathbf{H}\delta\mathbf{m}$$

where we have defined the gradient vector **g** and the Hessian matrix **H**. In the true Earth model, the derivative of the objective function with respect to **m** is zero. In this scenario, the previous expression can be rearranged to solve for a model perturbation.

$$\delta\mathbf{m} = -\mathbf{H}^{-1}\mathbf{g} \approx \mathbf{P}^{-1}\mathbf{g}$$

For problems of realistic size, direct computation of the Hessian matrix (or its inverse) is not computationally feasible. The Hessian is often replaced by a preconditioning matrix **P,** chosen such that it can be readily computed whilst retaining some characteristics of the Hessian. Finally, an iterative update rule is given by

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \nu_k \delta\mathbf{m}_k$$

where *k* denotes the iteration number. $\boldsymbol{\nu}_k$ is a scalar value that signifies the step length.

The role of parameter trade-off becomes apparent examining the multi-parameter case involving two generic parameters $\mathbf{m}_1$ and $\mathbf{m}_2$.

$$\begin{bmatrix} \mathbf{H}_{\mathbf{m}_1\mathbf{m}_1} & \mathbf{H}_{\mathbf{m}_1\mathbf{m}_2} \\ \mathbf{H}_{\mathbf{m}_2\mathbf{m}_1} & \mathbf{H}_{\mathbf{m}_2\mathbf{m}_2} \end{bmatrix} \begin{bmatrix} \delta\mathbf{m}_1 \\ \delta\mathbf{m}_2 \end{bmatrix} = -\begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix}$$

The ideal model updates are linear combinations of the individual gradients with coefficients defined by the inverse Hessian. By assuming a steepest descent direction (i.e. Hessian is taken as the identity matrix), the model updates do not capture the aforementioned linear combination. As the Hessian is rarely accessible, an alternative approach to minimizing parameter trade-off, would be to select $\mathbf{m}_1$ and $\mathbf{m}_2$ such that the off-diagonal components of the Hessian are zero or minimal (in comparison to the diagonal components). Reparameterization is investigated in the following section through synthetic inversions.

Synthetic test: SEG/EAGE Overthrust model

We perform inversions on a 2D section of the SEG/EAGE overthrust model v_p model (P-wave velocity). A v_s (S-wave velocity) model is generated by assuming $v_p/v_s = 1.76$ throughout the model; a homogeneous density model is used ($\rho = 1000$ kg m^{-3}). Initial models are obtained by convolving the true model with a 2D Gaussian kernel ($\sigma_x = \sigma_z = 625$ m). Each inversion utilizes 96 explosive sources ($\Delta x_s = 100$ m, $z_s = 50$ m) recorded at 264 receivers along the surface ($\Delta x_r = 75$ m). The source wavelet is a Ricker wavelet with a dominant frequency of 5 Hz. Each inversion implements an L-BFGS algorithm with a backtracking line-search (Nocedal, 2006) and runs for 30 iterations.

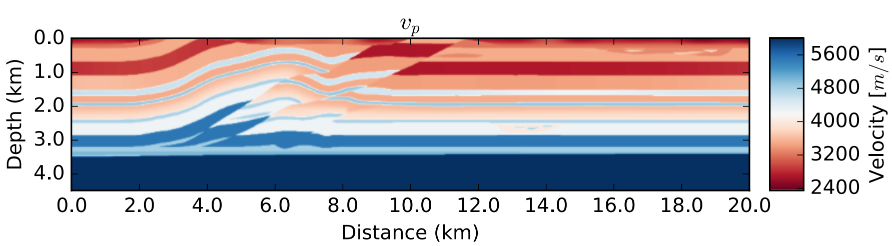

Figure 1: 2D section of the SEG/EAGE overthrust v_p model.

The inversions are performed for four different model parameterizations: seismic velocities (v_p, v_s), Impedances (I_p, I_s), Lamé parameters (λ, μ), and slownesses (p_p, p_s). Where applicable, the p and s subscripts represent the associated P and S-wave variants of each parameter.

We compare the convergence properties for each different parameterization. Figure 2 displays the convergence of data misfit along with model errors, where model error is defined as

$$\mathbf{m}_{err} = \sqrt{\left\| \frac{\mathbf{m}_{est} - \mathbf{m}_{true}}{\mathbf{m}_{true}} \right\|_2}$$

In each case, the parameter sets are converted into seismic velocities prior to comparison.

Figure 2: Convergence plots of data misfit and model errors for various model parameterizations.

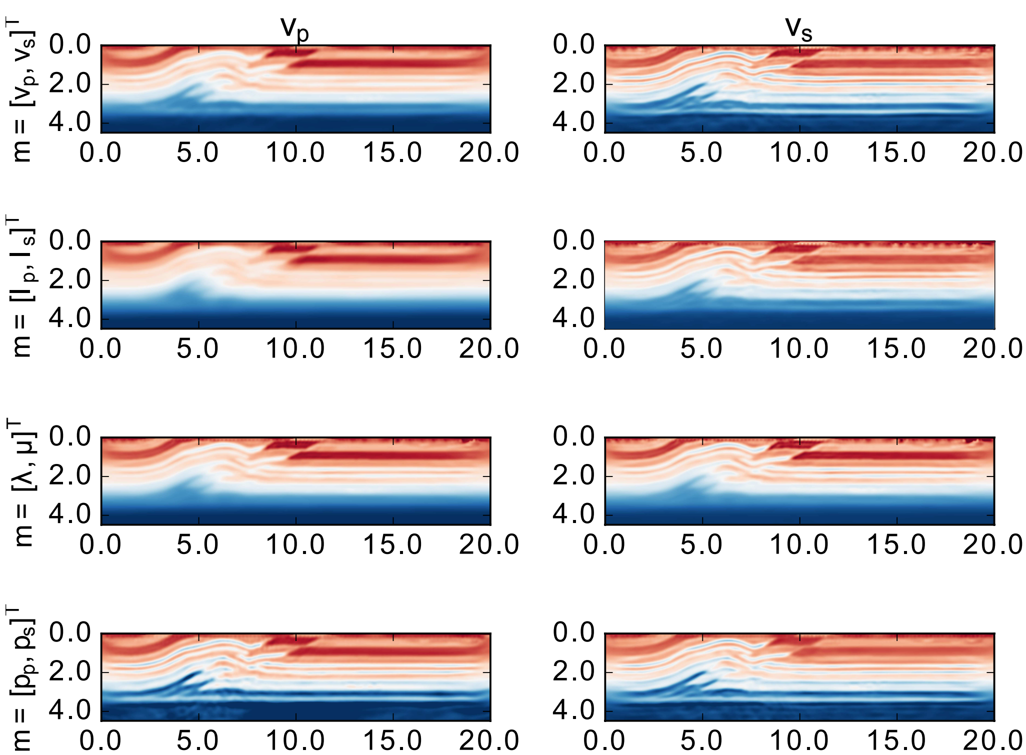

Figure 3: Inverted v_p and v_s models after 15 iterations for various model parameterizations.

After 30 iterations, velocity and slowness parametrizations have reduced the misfit to ~10% of the initial value whereas for impedance and Lamé parameters, a reduction to ~20% of the initial value is achieved. The difference arises from the inability of the impedance and Lamé parametrizations to invert deeper parts of the model (figure 3). Slowness shows a greater reduction in the data misfit over the velocity parametrization, owing to the more accurate v_p model inverted. The impedance parametrization displays the slowest reduction rate in model errors. The effect of parametrization is more pronounced in the inverted v_p model errors.The slowness parametrization dramatically improves the rate at which the v_p model error is reduced when compared to all other parametrizations. All parametrizations display various rates at which v_p model errors are reduced relative to v_s model errors. For example, v_p model errors reduce at a noticeably slower rate than v_s model errors. Velocity and Lamé parametrization show more comparable model error reduction update rates for v_p and v_s.

Conclusions

We have performed a series of synthetic inversions to evaluate the role of model parametrization in multi-parameter FWI. Changing the model parametrization resulted in varying convergence rates for the data misfit. Slowness and velocity parametrizations outperformed impedance and Lamé parametrizations. By converting the parameter sets to seismic velocities we observed variable behaviour in the convergence of v_p and v_s model errors. The slowness parametrization rapidly reduced v_p model errors and excelled at imaging deeper parts of the model. In future work, it will be important to compare traditional diffraction pattern analyses with more direct approaches presented here to form a more complete understanding of parameter trade-offs. By combining such approaches, it may eventually be possible to identify optimal parametrizations. The acquisition geometry influences the distribution of sampled by the data and will also need to be considered in future work.

Acknowledgements

The authors recognize support from NSERC (Alexander Graham Bell Canada Graduate Scholarship), Alberta Innovates and the industrial sponsors of the Seismic Analysis and Imaging Group.

References

Tarantola, A. (1986), A strategy for nonlinear elastic inversion of seismic reflection data,Geophysics, 51 (10), 1893-1903.

Mora, P. (1987), ”Nonlinear two-dimensional elastic inversion of multioffset seismic data.”, Geophysics, 52(9), 1211-1228.

Pratt, R. G. (1999), Seismic waveform inversion in the frequency domain, part 1: Theory and verication in a physical scale model, Geophysics, 64, 888-901.

Virieux, J., and S. Operto (2009), An overview of full-waveform inversion in exploration geophysics: Geophysics, 74, WCC1-WCC26.

Nocedal, J., and S. J. Wright (2006), Numerical optimization, 2nd ed.: Springer. Springer Series in Operations Research and Financial Engineering

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent (1995), Multiscale seismic waveform inversion, Geophysics, 60, 1457-1473.

Fichtner, A. and Trampert, J. (2011), Resolution analysis in full waveform inversion. Geophysical Journal International, 187: 1604–1624.