Comparison between least-squares reverse time migration and full-waveform inversion

Lei Yang, Daniel O. Trad and Wenyong Pan

Summary

The inverse problem in exploration geophysics usually consists of two parts: seismic imaging and velocity model building. In this paper, we compare algorithms for least-squares reverse time migration (LSRTM) and full-waveform inversion (FWI) in the frequency domain and use numerical examples to understand the differences. Usually, LSRTM uses Born approximation as the modelling method because it requires the adjoint of migration (linear inversion), while FWI uses finite-difference modelling because it does not require an adjoint-pair operator (non-linear inversion). In this paper, we perform LSRTM by using an approach similar to FWI for the gradient computation and implement it with a Hessian-free method that uses Hessian vector products during conjugate gradient iterations. The algorithm minimizes a least squares error function in the image domain, at difference of the usual data domain residual normally applied in this technique.

Introduction

Reverse time migration (RTM) was proposed in 1970s, but because of its high computational cost it did not become common until after 2000. At difference of popular one-way wave equation methods, RTM is a two-way migration method. This characteristic makes RTM more accurate than other methods for imaging steep dips, providing high resolution images of complex structures.

On the negative side, RTM suffers, as other migration methods, from operator limitations due to sampling and illumination. These limitations produce amplitude problems, which can be mitigated by applying a least-squares inversion instead of a migration. Also, when the velocity model has strong impedance contrasts, low-frequency artifacts are generated. Post-migration processing, like Laplacian filters, are required to eliminate these artifacts, taking with them useful low frequency information. Least-squares migration (LSM) is an alternative to migration that can, potentially, reduce migration artifacts and improve lateral resolution. It was first proposed by Schuster (1993), and applied for the first time to a real data set by Nemeth et al. (1999). Since then, much research has been done on this topic although its application to real data by industry remains elusive. This problem is often solved by considering migration as a linear problem, which is a common approximation in seismic processing if multiple reflections are not considered. LSM usually uses iterative methods to match the observed data and can improve focusing and amplitude accuracy.

Generally, linear operators are used in forward geophysical modelling calculations. A common task is to find the inverse of these calculations, i.e., given the observed data, find a model that can predict these data. For linear operators, it is possible to use iterative algorithms which require forward and adjoint operators to solve the inverse problem (Claerbout, 1992). Although in general these adjoint operators are different from the inverse operators, often they approximate the inverse quite well and are more stable and faster. This is the case for most migration operators, which are usually very similar to, but not exactly the same as, the respective adjoint operators. In some cases, the exact adjoint operators are not easy to design, and that is the particular case of RTM (Chen and Sacchi, 2017).

The methods mentioned above are all applicable for linear problems. However, recorded seismic data depend nonlinearly on earth parameters. FWI is a nonlinear inversion method which uses all information in the seismogram to get the earth model. Classical FWI involves the minimization of an objective function calculated as the difference between observed and predicted data. The gradient of this cost function provides the direction to move from the initial model to a model that predicts acquired data with minimum error. For forward modelling, FWI uses a finite-difference method, rather than Born modelling as LSM. Also, FWI changes the velocity model in every iteration by using the gradient and Hessian operators. In the following section, we review the theories of LSRTM and FWI in the frequency domain and combine these two methods to implement a Hessian-free algorithm to solve for the reflectivity model in LSRTM.

Theory

In this paper we use the 2D acoustic wave equation with constant density in the frequency domain:

$$\left(\frac{\omega^2}{v^2(x)} + \nabla^2\right) u(x, \omega, v) = f(\omega)\delta(x - x_s), \tag{1}$$

where x is the is the position of the earth, $v(x)$ is the acoustic wave velocity, ω is the angular frequency, $u(x, \omega)$ is the acoustic wave potential, ∇^2 is the Laplacian operator, $f(\omega)$ is the source term and $\delta(x - x_s)$ is the Dirac delta function. We can rewrite this equation in matrix form (Marfurt, 1984):

$$L(x, \omega; v)u(x, x_s, \omega) = f(\omega)\delta(x - x_s), \tag{2}$$

where $L(x, \omega; v) = \left(\frac{\omega^2}{v^2(x)} + \nabla^2\right)$ is the discretized impedance matrix.

Full waveform inversion

Full waveform inversion is a nonlinear inversion method. It starts from the misfit function:

$$\Phi(m) = \frac{1}{2}\left|\left|Ru - d_{obs}\right|\right|^2, \tag{3}$$

where R is the receiver sampling operator and $d_{syn} = Ru$ is the synthetic data. The goal for FWI is to estimate the subsurface parameters by iteratively minimizing the misfit function in (3). We apply the Taylor-Lagrange expansion of the misfit function

$$\Phi(m + \Delta m) \approx \Phi(m) + g^T\Delta m + \frac{1}{2}\Delta m^T H\Delta m, \tag{4}$$

where g^T is the gradient transpose of the misfit function, Δm is the searching direction and H is the Hessian operator. The searching direction is the solution of the Newton-based optimization method

$$H_k\Delta m_k = -g_k. \tag{5}$$

Least-squares reverse time migration

In LSRTM, the problem is to calculate the model that linearly predicts the data with minimum error as in equation 3, with the following definition of forward modelling:

$$Am = d. \tag{6}$$

A is the born modeling operator, m is the model perturbation (either velocity perturbation or reflectivity) and d is the wavefield perturbation. The least squares solution is

$$A^T Am = A^T d. \tag{7}$$

In this formulation, $A^T d$ can be treated as the result of gradient computation in FWI and $A^T A$ is the Hessian. Using the adjoint state method (Plessix, 2006), the gradient can be calculated by the migration of the residuals. The imaging condition of RTM can be constructed by zero-lag

2

correlation between forward modelled wavefields and backward propagated data residual wavefields:

$$g(x) = -\sum_{x_g}\sum_{x_s}\sum_\omega Re\left(\omega^2 f(\omega)G(x_s,x,\omega)G(x,x_g,\omega)\Delta d^*(x_g,x_s,\omega)\right), \qquad (8)$$

where $g(x)$ is the image of the data residual Δd, also showed as $A^T d$ above, $Re(\cdot)$ means the real part and "$*$" denotes conjugate. In matrix form, this can be rewritten as (Virieux and Operto, 2009)

$$g(x) = -\sum_{x_g}\sum_{x_s}\sum_\omega u^T(x_s,x_g,\omega)\nabla_m L^T(m,\omega)L(m,\omega)^{-1}R^T\Delta d^*(x_s,x_g,\omega), \qquad (9)$$

where $\nabla_m L^T(m,\omega)$ is the impedance matrix derivation with respect to model m. During each iteration in LSM, the search direction is calculated purely from present and past gradients. This provides attractive computation cost but suffers from a slow convergence rate. In this work, we calculate the search direction with both the gradient and Hessian, but without explicitly forming the Hessian (Hessian-free optimization method). We only need Hessian-vector products, which can be calculated by the adjoint state method (Métivier, 2014 and Pan, 2017)

$$H_a v = u^T \nabla_m L^T (L^T)^{-1} R^T R (L^*)^{-1} \nabla_m L^* u^* v, \qquad (10)$$

where v is an arbitrary vector, usually setting as zero vector as the initial guess of the conjugate gradient method and H_a is the approximate Hessian, discarding the second order term of the full Hessian. This truncated-Newton method do not construct the Hessian directly and only need to calculate the vector product in every conjugate gradient iteration and this process can be treated as $A^T A m$ in equation (7).

Therefore, we have an image domain objective function:

$$J(m) = \frac{1}{2}||A^T A m - A^T d||^2, \qquad (11)$$

which is different to the usual data domain objective function:

$$J(m) = \frac{1}{2}||d - Am||^2, \qquad (12)$$

The input to the linear conjugate gradient is the Hessian vector product Hm_n , and the initial guess m_0 is usually set as zero.

Examples

In this section we show some simple numerical examples for LSRTM and FWI using an inaccurate version of the Marmousi velocity model, resampled in a grid with 126*288 cells. The horizontal and vertical sampling are both equal to 8m. We use a constant density in this example. There are 285 shots and 287 receivers on the surface and the frequency is 5-25 Hz for RTM and LSRTM and 3-12 Hz for FWI.
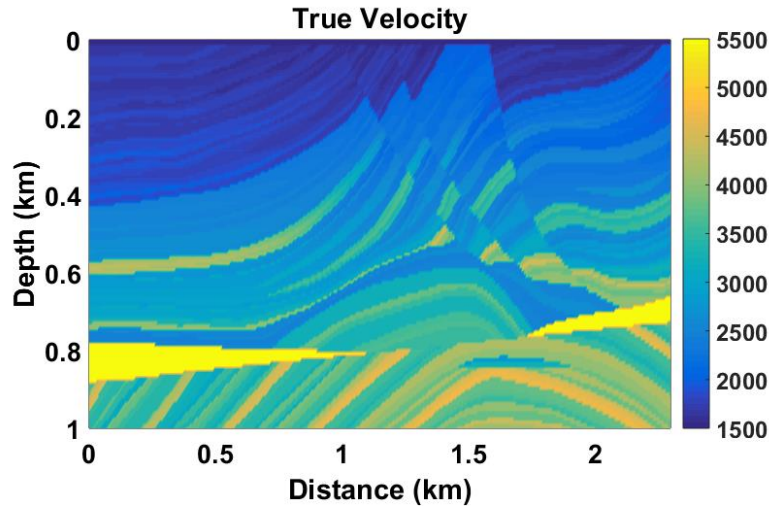
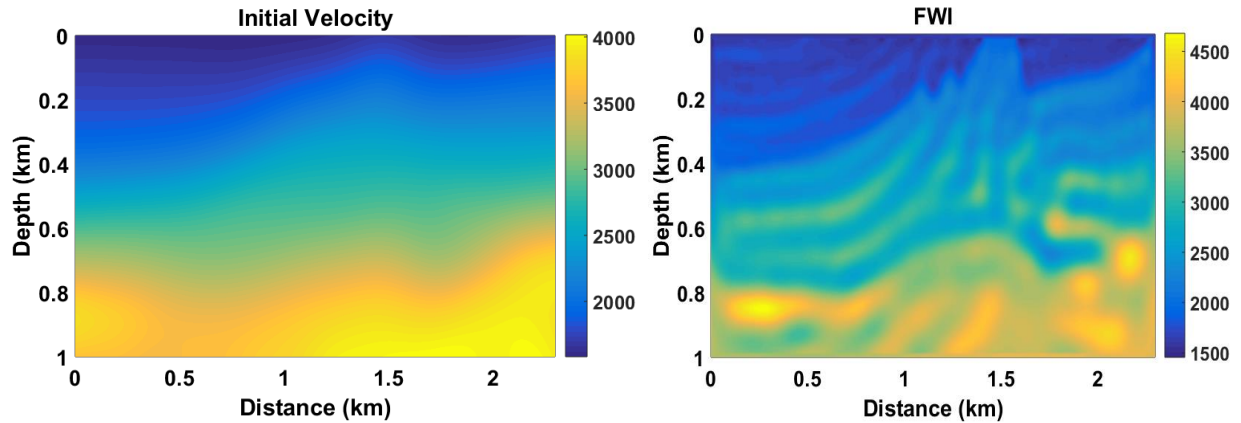Figure 1. The true velocity model of Marmousi

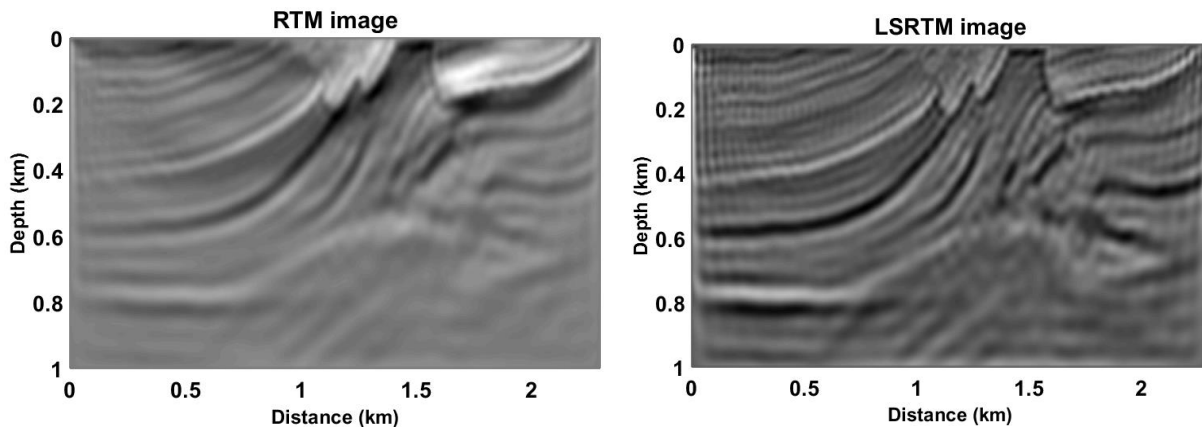Figure 2. The initial model and the FWI result from the initial model

Figure 3. The RTM image and the LSRTM image from the model in Figure 2a.

Figure 1 shows the true velocity model and figure 2 is the initial model, which was obtained by applying a Gaussian smoother on the true velocity model and the FWI result from the initial model. In Figure 3 we see the results for RTM and LSRTM when using the inaccurate smooth velocity model in Figure 2a. The RTM image is distorted by the low-frequency components severely. The LSRTM image is better resolved.

Conclusions

We have discussed differences between LSRTM and FWI. Usually, LSRTM uses a linearized wave equation based on the Born approximation, which allows one to use a linear inversion method with parameterized step size calculation. In this work, we have instead formulated LSRTM with the gradient and Hessian forms from FWI and solved the image domain least squares problem. The method is Hessian-free because only Hessian vector products are required. We don't have yet a clear conclusion on how this approach compares to traditional LSRTM in terms of cost, convergence and resolution, but we expect that it could behave better than the traditional method when the velocity model is inaccurate.

Acknowledgements

References

Chen, K. and Sacchi, M.D., 2017, Elastic least-squares reverse time migration via linearized elastic full waveform inversion with pseudo-Hessian preconditioning: Geophysics, 82, No.5, 1-89

Claerbout, J.F., 1992. Earth soundings analysis: Processing versus inversion (Vol. 6). London: Blackwell Scientific Publications.

Métivier, L., Bretaudeau, F., Brossier, R., Operto, S. and Virieux, J., 2014. Full waveform inversion and the truncated Newton method: Quantitative imaging of complex subsurface structures. *Geophysical Prospecting*, *62*(6), 1353-1375.

Nemeth, T., Wu, C. and Schuster, G.T., 1999. Least-squares migration of incomplete reflection data. Geophysics, 64(1), pp.208-221.

Pan, W., K. A. Innanen, and W. Liao, 2017, Accelerating Hessian-free Gauss-Newton full-waveform inversion via l-BFGS preconditioned conjugate-gradient algorithm: Geophysics, 32, no. 2, R49–R64.

Plessix, R.E. and Mulder, W.A., 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophysical Journal International, 157, No.3, 975-987

Schuster, G.T., 1993. Least-squares cross-well migration. In SEG Technical Program Expanded Abstracts 1993 (pp. 110-113). Society of Exploration Geophysicists.

Virieux, J. and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, 74, No. 6, WCC1-WCC26

Geoconvention 2018