Iterative modeling migration and inversion (IMMI): combining full waveform inversion with standard inversion methodology

Wenyong Pan*, Gary F. Margrave, Kristopher A. Innanen, CREWES, University of Calgary

SUMMARY

Full Waveform Inversion (FWI) employs full waveform information to estimate subsurface properties through an iterative process by minimizing the difference between the observed data and synthetic data. This inversion method has been widely studied in recent years but still cannot be practiced effectively in industry. The gradient in FWI is similar to an ungained Reverse Time Migration (RTM) image with the cross-correlation imaging condition. The estimation of the velocity update from the migrated section for FWI is analogous to the common process of impedance inversion. We show that the poorly scaled gradient can be improved by preconditioning with an approximate Hessian matrix (source illumination) forms the gradient or image based on deconvolution imaging condition which can be used to estimate the approximate reflectivity section directly. Thus, the deconvolution based gradient can be employed to estimate the impedance perturbation using the standard inversion methodology, denoted as SM. By examining the key concepts in FWI and SM, a hybrid inversion strategy is proposed by combining FWI with SM, namely, Iterative Modeling Migration and Inversion (IMMI). The IMMI method keeps the step of creating reflectivity image tying to wells control in SM and incorporates the concepts of imaging the data residuals and iteration from FWI. Furthermore, to reduce the computation cost for constructing the gradient and source illumination, the phase-encoding technique is introduced. In this paper, we practice the proposed strategies on a modified Marmousi model and the inversion results show that the IMMI method can reconstruct the velocity model efficiently and stably.

INTRODUCTION

Full waveform Inversion (FWI) is a very important method to reconstruct the velocity model for high resolution seismic imaging through an iterative process by minimizing the difference between the observed data and synthetic data (Tarantola, 1984; Virieux and Operto, 2009; Margrave et al., 2011). However there are various problems existed in FWI such as lack of low frequency information, extensive computational burden, cycle skipping problem etc., which make it difficult to be applied in industry effectively.

It has been widely observed that FWI and reverse time migration (RTM) share the same algorithm structure (Shin et al., 2001). The gradient construction in FWI can be interpreted as a migration process and the gradient is similar to a RTM image with the cross-correlation imaging condition although it lacks gain correction. Margrave et al. (2011) conjecture that any other wave-equation migration technology is also able to calculate an approximate gradient. The FWI gradient image is poorly scaled resulting from lack of correction for the geometrical spreading effects during wave propagation (Gray, 1997).

The classical imaging principle given by Claerbout (1971) states that: the reflector exists where the downgoing wavefields and upgoing wavefields coincide in time and space, which is equivalent to a zero-lag cross-correlation of the two wavefields (Claerbout, 1971; Kaeflin and Guittou, 2006; Chattopadhyay and McMechan, 2008). While this cross-correlation imaging principle actually provides a blurred image for that it employs the adjoint of the forward modeling operator (Lailly, 1983). Claerbout also suggested the deconvolution imaging condition, where the wavefields are divided in the frequency domain, as superior to the correlation condition but less stable. The deconvolution imaging condition automatically corrects for geometrical spreading while use of the correlation condition implies that this correction should be done either before or after imaging. The imaging problem can be posed as an inverse problem based on the minimization of a least-squares function to de-blur the migrated image using the Hessian operator (Tang, 2009). The steepest descent method, which simplifies the Hessian matrix to an identity matrix, has been argued to be a crude strategy in scaling. The Hessian matrix can be viewed as a nonstationary deconvolution operator to compensate the geometrical spreading effects (Pan et al., 2014b). However, direct calculation of the Hessian matrix is considered to be unfeasible for its extensively computational requirements. Shin et al. (2001) proposed to use pseudo-Hessian as a substitution of the full Hessian and the diagonal pseudo-Hessian is actually equivalent to the source-side illumination (one portion of the full Hessian) (Pan et al., 2013, 2014b). And the cross-correlation based gradient preconditioned by the source illumination becomes the RTM image with deconvolution imaging condition, which can provide a direct estimation of the reflectivity. The estimation of the velocity update from the migrated section for FWI is analogous to the common process of standard impedance inversion methodology (SM), in which the reflectivity image is converted into the impedance section.

We examine and compare the key concepts and parallels in FWI and SM and a new hybrid seismic inversion strategy, Iterative Modeling Migration and Inversion (IMMI), is proposed by combing FWI with SM. The IMMI method constructs the deconvolution based gradient by migrating the data residuals, converts the estimated reflectivity gradient into impedance perturbation matching to well data and then updates the impedance section iteratively. Furthermore, in this research, the phase-encoding technology is employed to construct the gradient and diagonal pseudo-Hessian for reducing the computational cost.

This paper is organized as follows: firstly, we review the imaging conditions in seismic imaging and give the analytic solutions for the imaging conditions based on a simple model. Secondly, we review the general principle of FWI and introduce the deconvolution based gradient. And then we state the basic theory for IMMI method and explain how to estimate the impedance perturbation using the deconvolution based gradient. Finally, we applied the proposed strategies on the Marmousi model and the inversion results prove that the IMMI method can rebuild the velocity model stably and effectively.

IMAGING CONDITIONS

The approximation to reflectivity $I$ is always given as the ratio of upgoing wavefields $U(r, \omega)$ to downgoing wavefields $D(r, \omega)$ (Claerbout, 1971), which is also a restatement of deconvolution imaging condition:

$$ I(r) = \sum_{\omega} \frac{U(r, \omega)}{D(r, \omega)}, $$

where $r = (x, y, z)$ is the Cartesian coordinate of a subsurface position, $\omega$ is the temporal frequency, and $\omega_N$ is the Nyquist frequency. Equation (1) becomes unstable when the downgoing wavefields $D(r, \omega)$ in the denominator approaches 0. To avoid this instability, the complex conjugate of downgoing wavefields $D^*(r, \omega)$ ($* \text{ denotes the complex conjugate}$) is always multiplied in denominator and numerator of equation (1) (Guittou et al., 2007; Schleicher et al., 2008) and a small constant is added to the denominator, which give:

$$ I_{dec}(r) = \sum_{\omega} \frac{D^*(r, \omega)U(r, \omega)}{D^*(r, \omega)D(r, \omega) + \lambda A_{max}}, $$

where $\lambda A_{max}$ is the damping term, $\lambda$ is a small constant value between 0 and 1 and $A_{max}$ means the maximum value of $D^*(r, \omega)D(r, \omega)$. Equation (2) is equivalent to equation (1) multiplied by an optimal Wiener
Iterative modeling migration and inversion (IMMI)

Figure 1: Analytic solutions of the cross-correlation and deconvolution imaging conditions. \(k_0\) is the wavenumber, and \(s\) indicates the distance from the receiver to the reflection point.

The model can be written as the summation of the starting model and a model perturbation \(\delta m(r)\):

\[
m_{k+1}(r) = m_k(r) + \mu_k \delta m_k(r),
\]

where \(k\) indicates the number of iteration, \(\mu_k\) is the step length, a scalar constant used to scale the model perturbation \(\delta m_k(r)\) and it can be obtained through a line search method (Gauthier et al., 1986; Pica et al., 1990). The model perturbation can be expressed as:

\[
\delta m = -H^{-1}g,
\]

where \(g = \frac{\partial \phi}{\partial m}\) is the gradient and \(H = \frac{\partial^2 \phi}{\partial m^2}\) is the Hessian matrix. The gradient can be calculated using the adjoint state method by applying a zero-lag cross-correlation between the forward modeling wavefields and backpropagated data residuals (Lailly, 1983; Tarantola, 1984), which avoids the direct computation of the first order partial derivative wavefields:

\[
g = -\sum_{a,r_i} A^2 \omega^2 \mu^T \otimes (B^{-1})^T \Delta a_i^T,
\]

where \(\otimes\) means cross-correlation operation and \(T\) means transpose. Using the simple model shown in Figure 1, the analytic expression of the gradient can be obtained as:

\[
g = -\sum_{a,r_i} A^2 \omega^2 \mu^T \otimes (B^{-1})^T \Delta a_i^T = -\sum_{a,r_i} \Re(\omega^2 \mu^T \otimes (B^{-1})^T \Delta a_i^T),
\]

And furthermore, we can notice that the gradient is indeed a poorly scaled image and it decays as \(F_s^{-2}\) for this analytic example as suggested by Gray (1997).

The Hessian matrix is the second order partial derivative of the misfit function with respect to the model parameters. Under the assumption of small model perturbation, the nonlinear term in the Hessian matrix can be ignored for computational convenience and the linear term forms the approximate Hessian used in Gauss-Newton method and it can be written as a scalar product between two Jacobian matrices:

\[
H_m = J^T J = \Re(\omega^2 u^T u (B^{-1})^T (B^{-1})),
\]

We can see that the approximate Hessian consists of two forward modeling wavefields and two backward propagation operators which correspond to the source-side and receiver-side geometrical spreading effects respectively.

Shin et al. (2001) proposed to substitute the full Hessian matrix using the pseudo-Hessian, which can be constructed using the forward modeling wavefields, used as the virtual sources in reverse rime migration:

\[
H_pseudo = f_{virtual}^* f_{virtual} = (\frac{\partial B}{\partial m})^* (\frac{\partial B}{\partial m}) = \Re(\omega^2 u^T u),
\]

Compared to the approximation Hessian \(H_m\), we can see that the pseudo-Hessian \(H_pseudo\) actually ignores the back propagation operators. For

---

**GENERAL PRINCIPLE OF FWI**

Full waveform inversion estimates the subsurface parameters through an iterative process by minimizing the difference between the synthetic data \(\mathbf{d}\) and observed data \(\mathbf{d}\) (Lailly, 1983; Tarantola, 1984). The misfit function \(\phi\) is formulated in a least-squares form (Virieux and Operto, 2009):

\[
\phi = \frac{1}{2} \| \mathbf{d} - B \mathbf{u} \|^2,
\]

where \(B\) is the forward modeling operator, \(\mathbf{u}\) mean the wavefields and \(\| \cdot \|^2\) means the \(\ell^2\)-norm. The minimum value of the misfit function is sought in the vicinity of the starting model \(m_0(r)\) and the updated
Iterative modeling migration and inversion (IMMI)

The model perturbation can be obtained by inserting equation (11) and equation (13) into equation (9):

$$\delta m = \langle (J^T \mathbf{R} + \lambda I)^{-1} J^T \Delta d \rangle,$$

where $I$ is an identity matrix and $\lambda$ is a constant value. And $\lambda I$ is the stable term which makes Hessian matrix invertible. When considering high frequency asymptotics, the Hessian matrix is highly diagonal dominant and the inverse of the approximate Hessian can be replaced by its reciprocal. Plugging the gradient and approximate Hessian into the equation (15) gives:

$$\delta m = -\mathcal{R} \left( \frac{1}{(B^{-1})^T} \sum_{\omega, \omega'} \omega^2 u^* \otimes (B^{-1})^T \Delta d^* \right),$$

(16)

The analytic expression of the diagonal pseudo-Hessian for the simple version strategy, denoted as SM. Several interesting differences between these two inversion strategies can be observed (Margrave et al., 2010). Both inversion strategies produce the subsurface properties sections. SM firstly creates the reflectivity section through a sophisticated data processing sequence and then converts the reflectivity image into an impedance section. While FWI updates the impedance section through an iterative process by minimizing the difference between the observed data and synthetic data. It never creates a reflectivity image. In this research, we propose a new seismic inversion strategy, which combines the robust features of SM with most promising concepts from FWI. From SM, we keep the step of creating reflectivity image with matching to well control. In FWI, we incorporate the concepts of imaging the data residuals and iteration. And the new hybrid inversion strategy can be viewed as an iterative (I) process of modeling (M), migrating (M) the data residuals, impedance inversion (I) with well control and velocity model updating in each iteration (Margrave et al., 2011).

Numerical Examples

Figure 3: (a), (b), (c) and (d) show the true reflectivity, cross-correlation based gradient, deconvolution based gradient and deconvolution based gradient tying to well data, the impedance update becomes:

$$I_{k+1} = I_k + \Delta I_k = I_k + 2\lambda R_k \mathbf{W} \left( \dot{\mathbf{R}}_k \left( p^j, p_j^H, \omega \right) \right),$$

(24)

where $\mathbf{W}()$ indicates tying to well data (Margrave et al., 2011).
Iterative modeling migration and inversion (IMMI)

Figure 4: Comparison of the cross-correlation based gradient (red line), deconvolution based gradient (blue line), true reflectivity (black line) and reflectivity estimation after well control (green line) at 1km.

Figure 5: (a) shows the true impedance perturbation;(b) shows the estimated impedance perturbation with optimal scaling.

the well log data at 1km to correct the estimated reflectivity. Firstly, we generate the synthetic data with 18 shots and 920 receivers located at the top surface of the model. The source interval and receiver interval are 50m and 5m respectively.

In IMMI, the gradient is considered as the estimation of the reflectivity. So, we can compare the cross-correlation based gradient and deconvolution based gradient with the true reflectivity. Figure 3a shows the normalized true reflectivity section. Figure 3b and c show the cross-correlation based gradient and deconvolution based gradient with normalized amplitudes. And Figure 3e shows the deconvolution based gradient matching to the well data. It is easy to observe that the deconvolution based gradient can recover the amplitudes of the deep reflectors very well. And we extract the vertical lines at 1km from Figure 3a, b, c and d for comparison, as shown in Figure 4. We can see that the amplitude and phase of the deconvolution based gradient with well control (green line) match the true reflectivity (black line) very well.

Then we can convert the estimated reflectivity image into an impedance perturbation and update the impedance section iteratively following equation (22) and equation (23). Figure 5a and b show the true impedance perturbation and the estimated impedance perturbation with proper scaling. Figure 6a, b and c show the inversion result at 1km, 2km and 3km after the first iteration of IMMI with optimal scaling. We can observe that the relative values of the shallow and deep reflectors in the estimated impedance perturbation section match the relative values in the true impedance perturbation section very well. And the updated impedance (the blue lines) has a big improvement comparing with the initial model (the red lines).

Finally, we apply the IMMI method on the modified Marmousi model using the phase-encoding strategy (Pan et al., 2013, 2014b,c). We construct the phase encoded gradient and precondition the phase encoded gradient using the diagonal pseudo-Hessian (Pan et al., 2014a),which is equivalent to the deconvolution imaging condition. The impedance perturbation is estimated using the traditional impedance inversion method. The multiscale approach has also been employed to overcome the cycle skipping difficulty. Figure 7a shows the inversion result using traditional shot by shot and gradient based method after 30 iterations. Figure 7b and c show the inversion results using IMMI method after 30 iterations and 60 iterations respectively. It can be seen that that the IMMI can recover the deep parts of the model very well compared to the traditional FWI with gradient based method. And the IMMI method can reconstruct the main geological structures of the model in a few of iterations stably.

CONCLUSION

In this research, through analyzing the imaging conditions in seismic imaging, we find that the cross-correlation based gradient in FWI lacks the intrinsic gain correction while the deconvolution based gradient can compensate the geometrical spreading effects and estimate the reflectivity directly. Thus, the impedance perturbation can be constructed using traditional impedance inversion method. By examining the parallels between FWI and SM, we propose a new hybrid inversion strategy, namely, Iterative Modelling Migration and Inversion (IMMI) method, in which we keep the step of creating reflectivity image with matching to well control from SM and incorporate the key concepts of imaging the data residuals and iteration from FWI. Furthermore, to reduce the computational cost, the phase-encoding technique is also employed to construct the deconvolution based gradient. We apply this assemblage of strategies on the Marmousi model. Compared to traditional gradient based method for FWI, the IMMI method can reconstruct the velocity model more efficiently.

ACKNOWLEDGMENTS

This research was supported by the Consortium for Research in Elastic Wave Exploration Seismology (CREWES). Thanks also to Dr.Yu Zhang (now ConocoPhillips) for his introduction on the imaging condition.