3D efficient reflection waveform inversion for initial model building

Pengliang Yang

August 24, 2018

1 Introduction

Reflection waveform inversion (RWI) and its variant - joint full waveform inversion (JFWI), have proved its powerful capability to improve initial velocity model building beyond diving-wave penetration depths, resulting in significantly better ultra-deep images. The key difficulty in RWI and JFWI is the heavy computational cost coming from alternating between the backgroud velocity update and model perturbation (which can be embedded in impedance inversion under $V_p - I_p$ parameterization): To update background velocity at each iteration, we must perform sufficient number of iterations to reconstruct the model perturbation accurately. This computation intensive/inefficient part motives us to present a new scheme in this paper, by combining background velocity update with an efficient reflectivity inversion method using Wiener filtering.

2 RWI Formulation

Assume constant density, we have the modeling and demigration equations based on Born approximation

\[
\begin{align*}
\frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \Delta & \quad p = f \\
\frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \Delta & \quad \delta p = 2r \frac{1}{v^2} \frac{\partial^2}{\partial t^2} p,
\end{align*}
\]

(1)

where $r := \frac{\delta v}{v}$ is the so-called reflectivity corresponding to the least-squares migration image; $p$ is the wavefield simulated in the background velocity model $v$. The initial condition and boundary condition must be applied to the wavefields $p$ and $\delta p$:

\[
\begin{align*}
p|_{t=0} & = \partial_t p|_{t=0} = \partial_t \delta p|_{t=0} = 0; & \quad p|_{x \in \partial \Omega} = \delta p|_{x \in \partial \Omega} = 0.
\end{align*}
\]

(2)

Consider the least-squares of the reflection data misfit

\[
\chi = \frac{1}{2} \| \delta d - R \delta p \|^2
\]

(3)

The corresponding Lagrangian is

\[
L(p, \delta p, \lambda_1, \lambda_2, v) = \chi + \langle \lambda_1, \frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \Delta \rangle p - f \rangle_{T \times \Omega} + \langle \lambda_2, \frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \Delta \rangle \delta p - 2r \frac{1}{v^2} \frac{\partial^2}{\partial t^2} p \rangle_{T \times \Omega}
\]

(4)

leading to the gradient of the misfit w.r.t. the velocity $v$ given the reflectivity $r$

\[
\frac{\partial \chi}{\partial v} = \frac{\partial L}{\partial v} = \langle \lambda_1 - 2r \lambda_2, -2v^{-3} \frac{\partial^2}{\partial t^2} p \rangle_T + \langle \lambda_2, -2v^{-3} \frac{\partial^2}{\partial t^2} \delta p \rangle_T
\]

(5)

with the adjoint equations

\[
\begin{align*}
\frac{\partial L}{\partial p} & = 0 \Rightarrow \frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \Delta \lambda_1 = 2r \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \lambda_2 \\
\frac{\partial L}{\partial \delta p} & = 0 \Rightarrow \frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \Delta \lambda_2 = -\frac{\partial \chi}{\partial \delta p} = R^T (\delta d - R \delta p)
\end{align*}
\]

(6)

which must satisfy the final condition and the boundary condition:

\[
\begin{align*}
\lambda_1|_{t=T} & = \partial_t \lambda_1|_{t=T} = \lambda_2|_{t=T} = \partial_t \lambda_2|_{t=T} = 0; \quad \lambda_1|_{x \in \partial \Omega} = \lambda_2|_{x \in \partial \Omega} = 0.
\end{align*}
\]

(7)

Of course, instead of using least-square misfit one may consider other misfit functions such as cross-correlation misfit function to emphasize the phase information rather than erroneous amplitudes (the use of L2 norm might be inappropriate when the initial model is far away from the ground truth). It only changes the adjoint source $-\partial \chi/\partial \delta p$ to simulate wavefield $\lambda_2$. 

1
3 Replacing iterative reflectivity optimization by Wiener filtering

To obtain the reflectivity $r$, we might need to use the gradient of the misfit with given velocity model $v$

$$\frac{\partial \chi}{\partial r} = \frac{\partial L}{\partial r} = (\lambda_2, -2 \frac{1}{v^2} \partial_t^2 p)_T$$

(8)

to do many iterations in optimization. Without reasonably good estimation of reflectivity $r$, the update of the velocity $v$ is difficult. It implies that we must alternate optimization between $v$ and $r$, while each of them requires many iterations with the other fixed, making RWI very computationally heavy.

Eqns (1) and (6) are linear systems with respect to reflectivity which are on the right hand side of the second subequation in (1) and the first subequation in (6), generating the secondary (virtual) sources at all locations the velocity perturbation is not zero ($\delta v \neq 0$ or equivalently $v \neq 0$). The iterative update of the reflectivity is a process of the so-called least-squares migration in the sense that we can write the synthetic data extracted at receiver location is linked with the reflectivity through a linear operator

$$L$$

Therefore, we found the reflectivity image $r'$ which gives the minimizer of the misfit function (9) because

$$\chi = \frac{1}{2} \| \delta d - L[v]r \|^2$$

(9)

which is exactly the least-squares migration procedure. The minimizer of the misfit function is the solution of the equation $\delta d = Lr$. Zeroing the gradient of the misfit gives the normal equation

$$\frac{\partial \chi}{\partial r} = L^T (\delta d - Lr) = 0 \iff r = (L^T L)^{-1} L^T \delta d$$

(10)

implying that the linear iterative optimization actually inverts for $(L^T L)^{-1}$ by applying the left pseudo inverse operator $L^\dagger = (L^T L)^{-1} L^T$ to $\delta d$. There exists extensive amount of work to speedup the convergence of this iterative optimization, for example, using approximate direct inverse in extended domain, pseudo Hessian preconditioner. In the following, we consider the filtering approach proposed by Liu and Peter (2018).

The adjoint operator of the Born modeling, namely $L^T$, is the so-called migration operator which transforms observed reflection data $\delta d$ into a reflectivity image $r'$:

$$r' = L^T \delta d$$

(11)

We can obtain the new Born modeling data $\delta d'$ from the first reflectivity image above

$$\delta d' = Lr' = LL^T \delta d \iff \delta d = (LL^T)^{-1} \delta d' = f_m \ast \delta d'$$

(12)

where we assume the application of the composite linear operator $(LL^T)^{-1}$ is realized through the convolution of a matching filter $f_m$, which may be easily found using frequency domain Wiener filter formulation:

$$f_m = F^{-1} \left[ \frac{F[\delta d]^*F[\delta d]}{F[\delta d]^*F[\delta d] + \epsilon} \right]$$

(13)

where $F$ and $F^{-1}$ stand for the forward and the inverse Fourier transform; * indicates complex conjugate and $\epsilon$ is the stabilization factor to avoid division by zero.

Using this filter to filter the observed reflection data and then applying the migration leads to an image

$$m'' = L^T (LL^T)^{-1} \delta d$$

(14)

which gives the minimizer of the misfit function (9) because

$$Lm'' = LL^T(LL^T)^{-1} \delta d = \delta d.$$

(15)

Therefore, we found the reflectivity image $r = m''$ by filtering without any iterations. The cost is only 2 times of RTM or gradient building. Compared with eqn (14), it is easy to see that eqn (14) is equivalent to the application of the right pseudo inverse operator $L^\dagger = L^T(LL^T)^{-1}$ to $\delta d$.

The proposed combination of RWI with filtering by source estimation leads to drastic reduction of the computational cost, making the 3D application of RWI quite manageable. Although still based on Born approximation, this process does not need to rely on direct inverse which is only feasible in extended domain.
4 Extension to JFWI

Following the work of Zhou et al. (2015), slight modification of the least-squares misfit by using both the diving waves and reflections gives also an efficient JFWI scheme:

\[
\chi = \frac{1}{2}(\|W_{d_0}(Rp - d_0)\|^2 + \|W_{\delta d}(R\delta p - \delta d)\|^2) \tag{16}
\]

where we assume the scale separation between the diving wave plus refraction data \(d_0\) and the reflection wave \(\delta d\) in the observed seismic data \(d = d_0 + \delta d\). Through the Lagrangian of the form (4), we have the adjoint equations

\[
\frac{\partial L}{\partial p} = 0 \Rightarrow \left( \frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \Delta \right) \lambda_1 = 2r \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \lambda_2 + R^T W_{d_0}^T W_{d_0}(d_0 - Rp) \\
\frac{\partial L}{\partial \delta p} = 0 \Rightarrow \left( \frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \Delta \right) \lambda_2 = R^T W_{\delta d}^T W_{\delta d}(\delta d - R\delta p) \tag{17}
\]

Assume the weighting operator is identity \((W_{dd} = I)\). What we need to do is simply modifying the adjoint source of the wavefield \(\lambda_1\) in the RWI formulation.

It is important to emphasize a significant difference between the formulation in this report and the initial JFWI in the second term of the Lagrangian: here we put all model perturbation in the right hand side under Born approximation, leading to the linearized inverse problem, while the initial JFWI put the model perturbation in the left hand side in the wave operator, resulting in an iterative nonlinear inverse process which cannot be solved using linear convolutional model by Wiener filtering.

References
