Least-squares RTM with L1 norm regularization

Di Wu and Youcai Tang, Institute of Unconventional Natural Gas, China University of Petroleum(Beijing);
Gang Yao and Yanghua Wang, Department of Earth Science and Engineering, Imperial College London

Summary

Reverse time migration (RTM), for imaging complex Earth models, is a reversal procedure of the forward modeling of seismic wavefields, and hence can be formulated as an inverse problem. The least-squares RTM method attempts to minimize the difference between the observed field data and the synthetic data generated by the migration image. It can reduce the artifacts in the images of a conventional RTM which uses an adjoint operator, instead of an inverse operator, for the migration. However, as the least-squares inversion provides an average solution with minimal variation, the resolution of the reflectivity image is compromised. This paper presents the least-squares RTM method with a model constraint defined by an L1-norm of the reflectivity image. For solving the least-squares RTM with L1 norm regularization, the inversion is reformulated as a 'basis pursuit de-noise (BPDN)' problem, and is solved directly using an algorithm called 'spectral projected gradient for L1 minimization (SPGL1)'. Two numerical examples demonstrate the effectiveness of the method which can mitigate artifacts and produce clean images with significantly higher resolution than the least-squares RTM without such a constraint.

Introduction

Reverse time migration (RTM) solves a two-way wave equation directly, and therefore can handle dipped reflectors up to 90°, severely lateral variation of velocity, multi-arrival events, and even turning waves (Baysal et al, 1983; McMechan, 1983; Zhang et al, 2011). Conventional RTM does not formulate the reversal procedure as an inverse problem. Instead of using a properly solved inverse operator, a conventional migration uses an adjoint operator of the wave propagation operator, for an efficient implementation. This is a crude approximation, which may produce low-resolution seismic images with artifacts and inaccurate amplitudes.

To improve the performance, RTM can be formulated as an inverse problem to seek an optimal model, the reflectivities, so that the synthetic wavefields fit the field seismic data in a least-squares sense. This is called the least-squares RTM (Dai et al, 2012; Yao and Wu, 2015; Zhang et al, 2015; Shi and Wang, 2016; Yao and Jackubowciz, 2016). As expected, least-squares RTM would produce an image with fewer artifacts, more accurate amplitudes and higher resolution than the conventional RTM.

In this paper, seismic wave propagation procedure is presented as a linear relationship between the seismic data and the reflectivity image, following the Born approximation. Regarding to the least-squares RTM, the inverse problem intrinsically is under-determined. First, there are only a limited number of receivers which cannot fully cover the surface. Second, a seismic source can only generate a band-limited signal, propagating through the Earth with attenuation. In summary, the data is incomplete in terms of both the spatial and frequency coverage. To better recover the reflectivity model, Sacchi et al (2006) proposed to incorporate inversion with quadratic and non-quadratic, e.g. Cauchy constraint, regularization on the model. Since least-squares RTM applies iterative solvers, it is computationally intensive. To reduce the computational cost, Herrmann and Li (2012) demonstrated the simultaneous source technique with sparsity promotion and compressive sensing can effectively reduce the computational cost as well as mitigate the artefacts caused by cross-talk. Xue et al (2016) applied a type of shaping regularization on the reflectivity image to achieve the same aim. In this paper, the least-squares RTM is regularized by an L1-norm constraint of the reflectivity model. This regularization will stabilize the inverse problem and promote the sparsity of the migration image (Herrmann and Hennenfent, 2008). This sparse assumption satisfies the actual physical condition since the subsurface media are often consisted of layered strata.

The L1-norm regularized problem can be written as a 'basis pursuit de-noise (BPDN)' problem, and can be solved directly by using an algorithm called 'spectral projected gradient for L1 minimization (SPGL1)'. In this paper, we demonstrate this high resolution RTM method using two synthetic models.

Theory

RTM utilizes a two-way wave equation for imaging the reflectivity model. Born approximation can be used in order to include the reflectivity parameter into the wave equation. Surface-recorded reflection data can be expressed as

$$d(\mathbf{x}_s, \mathbf{x}_r, \omega) = \int j\omega G_r(\mathbf{x}_r, \mathbf{x}, \omega) I(\mathbf{x}) G_s(\mathbf{x}, \mathbf{x}_s, \omega) s(\mathbf{x}_s, \omega) \mathrm{d}\mathbf{x}, \quad (1)$$

where j is the imaginary symbol, ω is the angular frequency, $s(\mathbf{x}_s)$ is the source signature at $\mathbf{x}_s$, $G_s(\mathbf{x}, \mathbf{x}_s, \omega)$ is the Green's function from the source to the reflector at $\mathbf{x}$, $G_r(\mathbf{x}_r, \mathbf{x}, \omega)$ is the Green's function from the reflector to the receiver at $\mathbf{x}_r$, $I(\mathbf{x})$ is the stacked reflectivity image, $d(\mathbf{x}_s, \mathbf{x}_r, \omega)$ is the reflection seismic data.

Note that the $j\omega$ factor in equation (1) makes a 90-degree phase shift to the modeling data, for migration producing a zero-phase image (Yao and Wu, 2015). The two Green's functions $G_s(\mathbf{x}, \mathbf{x}_s, \omega)$ and $G_r(\mathbf{x}_r, \mathbf{x}, \omega)$ are the solutions of a two-way wave equation while the source is an impulse.

Conventional RTM, which applies an adjoint operator (Claerbout, 1992), can be expressed as

$$I(\mathbf{x}) = \iint [j\omega G_r(\mathbf{x}_r, \mathbf{x}, \omega) G_s(\mathbf{x}, \mathbf{x}_s, \omega) s(\mathbf{x}_s, \omega)]^* \times d(\mathbf{x}_s, \mathbf{x}_r, \omega)\, \mathrm{d}\mathbf{x}_r \mathrm{d}\omega\,, \quad (2)$$

where $*$ denotes the complex conjugate. The RTM migration of equation (2) can be implemented as the following steps: (1) convolving the product of the source and receiver Green's functions with the first-order derivative of the source wavelet ($j\omega s(\mathbf{x}_s, \omega)$), which generates a synthetic wavefield, (2) cross-correlating the synthetic wavefield with the recorded seismic data, (3) summing over all traces, and (4) summing over all frequencies. The last step is called the zero-lag image condition.

This migration can produce correct positions and times but inaccurate amplitudes because the adjoint operator only deals with the kinematic information of the wavefield. For accurate amplitudes, the migration should use a proper inverse operator, instead of that adjoint operator of forward modeling.

Hence, seismic migration should be posed as an inverse problem for finding an optimal reflectivity model through the following objective function

$$\phi_d\big(I(\mathbf{x})\big) = \|d(\mathbf{x}_s, \mathbf{x}_r, \omega) - d_0(\mathbf{x}_s, \mathbf{x}_r, \omega)\|_2^2 \quad (3)$$

where $d_0(\mathbf{x}_s, \mathbf{x}_r, \omega)$ is the recorded data, and $d(\mathbf{x}_s, \mathbf{x}_r, \omega)$ is the predicted data. RTM through this minimization is referred to as least-squares RTM.

The forward modeling equation (1) can be expressed in a vector-matrix form as

$$\mathbf{d} = \mathbf{G}\mathbf{I}, \quad (4)$$

where the matrix $\mathbf{G}$ denotes the kernel of the integral of equation (1), $\mathbf{d}$ and $\mathbf{I}$ are the vectors of data and reflectivity image, respectively. For the details of the matrix implementation, readers are referred to Yao and Jakubowicz (2016). Then, the least-squares solution of the inverse problem (3) can be expressed as

$$\mathbf{I} = [\mathbf{G}^*\mathbf{G}]^{-1}\mathbf{G}^*\mathbf{d}. \quad (5)$$

For field seismic data, the system of equation (4) is essentially underdetermined.

In order to further improve the resolution by promoting the sparseness of the model image, we can set the objective function for the migration as an L1-norm model constrained form:

$$\phi_\lambda = \|d(\mathbf{x}_s, \mathbf{x}_r, \omega) - d_0(\mathbf{x}_s, \mathbf{x}_r, \omega)\|_2^2 + \lambda \|I(\mathbf{x})\|_1 \quad (6)$$

where λ is the regularization parameter to balance the data fitting and the model constraint. Here, the L1-norm is a way to measure the sparseness of the reflectivity model (Candes et al, 2006).

Unfortunately, finding an appropriate λ is not an easy task. Alternatively, we can express the sparse inversion as

$$\min \|I(\mathbf{x})\|_1, \quad s.t. \ \|d(\mathbf{x}_s, \mathbf{x}_r, \omega) - d_0(\mathbf{x}_s, \mathbf{x}_r, \omega)\|_2^2 < \sigma \quad (7)$$

where σ is an estimation of the noise energy, which can be easily estimated from the data. This is the basis pursuit de-noise (BPDN) problem (Chen et al, 1998).

The SPGL1 method for solving a BPDN problem

SPGL1 (spectral projected gradient for L1 minimization) is a robust and direct solver to the BPDN problem (7) (van den Berg and Friedlander, 2008).

The SPGL1 method solves a series of the 'least absolute shrinkage and selection operator' (LASSO) problems, given by

$$\min \|d(\mathbf{x}_s, \mathbf{x}_r, \omega) - d_0(\mathbf{x}_s, \mathbf{x}_r, \omega)\|_2^2, \quad s.t. \ \|I(\mathbf{x})\|_1 < \tau \quad (8)$$

where τ is a measurement of the sparseness of the reflectivity model. The parameter τ in this LASSO problem has a tight connection with the σ in the BPDN problem of equation (7). The connection is described by the Pareto curve,

$$\phi_\sigma(\tau) = \sigma. \quad (9)$$

This curve defines the optimal trade-off between the L2-norm of the residual and the L1-norm of the solution (Hennenfent et al, 2008). The BPDN problem and the LASSO problem are two distinct characterizations of the same curve.

The SPGL1 method uses the function ϕ to parameterize the Pareto curve with τ. The solutions of the BPDN problem for any σ can be found by solving the LASSO problem with the corresponding τ. The Pareto curve is continuously differentiable (van den Berg and Friedlander, 2008) and its derivative can be given explicitly. Thus a Newton-based root-finding algorithm can be used to find roots of the Pareto curve.

The LASSO problem is efficiently solved by using a spectral gradient method. First, this method produces the updated model,

$$\tilde{\mathbf{I}}_{n+1}(\mathbf{x}) = \mathbf{I}_n(\mathbf{x}) + \alpha \mathbf{g}_n(\mathbf{x}), \tag{10}$$

where $\tilde{\mathbf{I}}_{n+1}(\mathbf{x})$ is the updated model, $\mathbf{I}_n(\mathbf{x})$ is the model of the n th iteration, $\mathbf{g}_n(\mathbf{x})$ is the gradient of the n-th iteration, and α is the step length (Barzilai and Borwein, 1988). Second, this method projects the updated model $\tilde{\mathbf{I}}_{n+1}(\mathbf{x})$ onto the L1-norm ball with radius τ to generate the final update $\mathbf{I}_{n+1}(\mathbf{x})$ of the iteration of $n+1$ by solving another optimization problem given by

$$\min \ \left\|\tilde{\mathbf{I}}_{n+1}(\mathbf{x}) - \mathbf{I}_{n+1}(\mathbf{x})\right\|_2^2$$
$$s.t. \ \ \|\mathbf{I}_{n+1}(\mathbf{x})\|_1 < \tau \tag{11}$$

Examples

The first model is a nine point model (Figure 1a), in which nine diffractor points are embedded in the constant velocity of 2000 m/s. The model is discretized into grids with the size of $5 \times 5\ m^2$. 201 uniformly-spaced receivers with an interval of 5 m are deployed on the surface of the model. A 30-Hz Ricker wavelet is used as the source signature. Three shot records, one of which is shown in figure 1b, are synthesized by using the acoustic wave equation with different source locations at 200, 500 and 800 m on the surface separately. The recording time duration is 0.65 s with a 0.5 ms sampling interval.

Figure 1c is the least-squares RTM image of the three shots. The least-squares RTM applies an inverse operator to find an optimal image to fit the record. This can mitigate strong artifacts of conventional RTM, caused by the adjoint operator, rather than an inverse operator, applied in migration (Yao and Jakubowicz, 2016). The least-squares RTM is carried out with 30 iterations of the conjugate-gradient method (Wang, 2016).

By comparing it with the true model, it can be seen the least-squares RTM image is still quite different. This is because the seismic record has a limited frequency bandwidth. As demonstrated by Yao and Jakubowicz (2016), the standard least-squares inversion could not recover the missing frequency components. We implement the least-squares RTM with an L1-norm model constraint and can improve the resolution by promoting the sparseness of the image. The image in Figure 1d is produced by solving the BPDN problem with the SPGL1 method with $\sigma = 0.01$. Both the conjugate-gradient and BPDN inversions use 30 iterations. The resultant image is almost identical to the true model in Figure 1a.

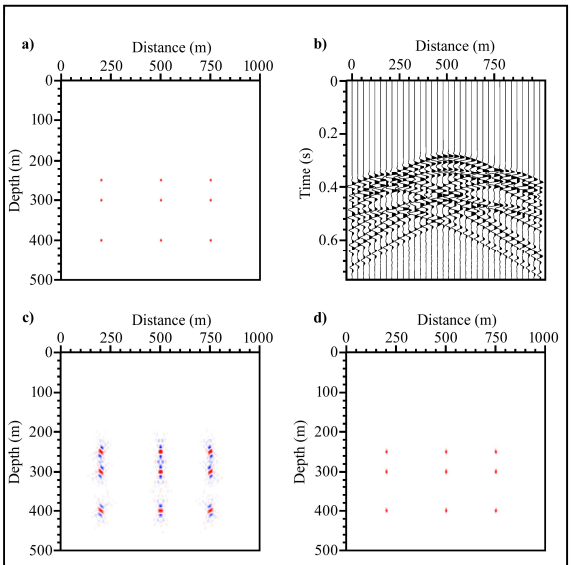

Figure 1: A nine diffractor model. (a) The model containing 9 diffractor points on a constant background of 2000 m/s. (b) One shot gather generated with a 30-Hz Ricker wavelet as the source, located on the surface at the distance of 500 m. The direct arrivals are removed. (c) The reflectivity images of three shots, using the least-squares RTM method. (d) The same image using the least-squares RTM with the L1 norm regularization.

The second example uses the Marmousi velocity model (Figure 2a). We generate 120 shot records with the acoustic wave equation, using an 8-Hz Ricker wavelet as the source wavelet. The first shot is located at the distance of 3 km on the surface. The shot interval is 50 m. Each shot has 96 traces on its left side with the minimal offset of 200 m and the maximal offset of 2575 m. The absorbing boundary is applied, and therefore the record does not include any surface-related multiples but the inter-bed multiples.

A smoothed version of this velocity model is used as the migration velocity for all the migrations. Figure 2b is a stack of 120 images of shot gather RTMs. It indicates that the conventional RTM image is in fact quite good, although the image has low wavenumber artifacts.

By contrast, least-squares RTM image (Figure 2c), obtained after 200 conjugate-gradient iterations, largely eliminates the low wavenumber artifacts and boosts the resolution as the reflectors are slimmer. However, the noise in the image becomes prevalent. This is because the forward modeling formula (Equation 1) for LSRTM is an approximation to the full wave equation. As a result, the data which cannot be predicted by the formula causes the noise in the image.

To further promote the resolution and suppress the noise, we use the least-squares RTM with the L1 norm regularization. The inversion is formulated as the BPDN problem, in which the σ parameter is set as 1% of the data energy. After 200 iterations (Figure 2d), the migration produces a much cleaner image, with higher resolution than the least-squares RTM without such a constraint.

Conclusion

In a least-squares RTM method, because the basic forward modeling is an approximation to a full wave equation, it is prone to produce artifacts in the image, attempting to fully fit the data. In order to reduce artifacts the least-squares RTM image and to promote the resolution, we have added an L1-norm model constraint to the least-squares inverse problem. This is the least-square RTM with the L1-norm regularization.

For solving this inverse problem, we have formulated it as a BPDN problem, which minimizes the L1-norm of the model, subject to the L2-norm of data misfit smaller than a threshold. The threshold can be found by measuring the noise energy. Then, we have solved this BPDN problem effectively using the SPGL1 algorithm.

We have demonstrated that this migration method with the L1-norm model constraint is robust and can effectively produce clean images with the highest resolution among all RTM methods.

Acknowledgement

The first author is grateful to Natural Science Foundation of China (grant no. 41272099), Foundation of China University of Petroleum (Beijing) 2462015YJRC012 and PRP/indep-3-1508.

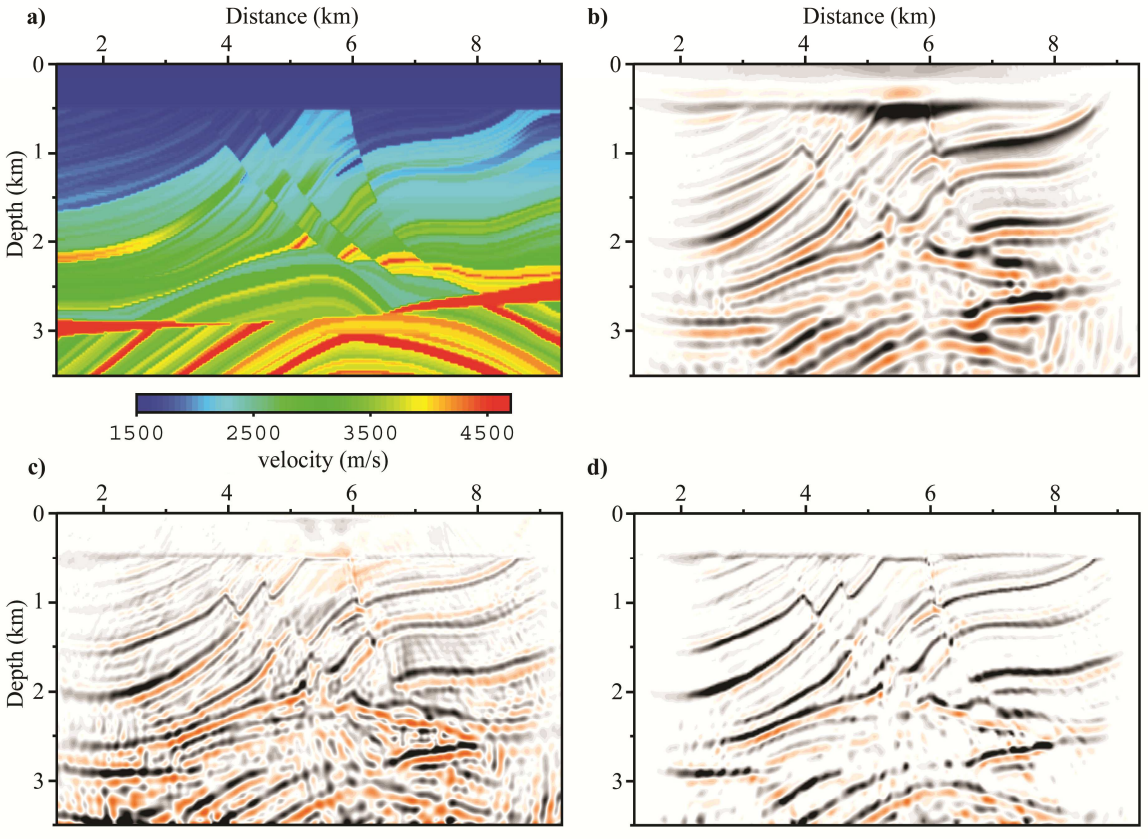

Figure 2: (a) The Marmousi velocity model. Its color scale is displayed below. A 8-Hz Ricker wavelet is used as the source signature, located at a distance of 4.225 km. (b) The reflectivity image of the conventional RTM method. (c) The image of the least-squares RTM method. (d) The image of the least-squares RTM with the L1 norm regularization. The orange red in these images represents the positive amplitude and black for negative.

EDITED REFERENCES

Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2016 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Barzilai, J., and J. M. Borwein, 1988, Two-point step size gradient methods: IMA Journal of Numerical Analysis, **8**, 141–148, http://dx.doi.org/10.1093/imanum/8.1.141.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524, http://dx.doi.org/10.1190/1.1441434.

Candes, E. J., J. Romberg, and T. Tao, 2006, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information: IEEE Transactions on Information Theory, **52**, 489–509, http://dx.doi.org/10.1109/TIT.2005.862083.

Chen, S. S., D. L. Donoho, and M. A. Saunders, 1998, Atomic decomposition by basis pursuit: SIAM Journal on Scientific Computing, **20**, 33–61, http://dx.doi.org/10.1137/S1064827596304010.

Claerbout, J. F., 1992, Earth Soundings Analysis: Processing versus Inversion: Blackwell Science.

Dai, W., P. Fowler, and G. T. Schuster, 2012, Multi-source least-squares reverse time migration: Geophysical Prospecting, **60**, 681–695, http://dx.doi.org/10.1111/j.1365-2478.2012.01092.x.

Hennenfent, G., E. van den Berg, M. P. Friedlander, and F. J. Herrmann, 2008, New insights into one-norm solvers from the Pareto curve: Geophysics, **73**, no. 4, A23–A26, http://dx.doi.org/10.1190/1.2944169.

Herrmann, F., and G. Hennenfent, 2008, Non-parametric seismic data recovery with curvelet frames: Geophysical Journal International, **173**, 233–248, http://dx.doi.org/10.1111/j.1365-246X.2007.03698.x.

Herrmann, F., and X. Li, 2012, Efficient least-squares imaging with sparsity promotion and compressive sensing: Geophysical Prospecting, **60**, 696–712, http://dx.doi.org/10.1111/j.1365-2478.2011.01041.x.

McMechan, G. A., 1983, Migration by exploration of time-depedent boundary values: Geophysical Prospecting, **31**, 413–420, http://dx.doi.org/10.1111/j.1365-2478.1983.tb01060.x.

Sacchi, M., J. Wang, and H. Kuehl, 2006, Regularized migration/inversion: New generation of imaging algorithms: CSEG Recorder, **31**, 54–59.

Shi, Y., and Y. H. Wang, 2016, Reverse time migration of 3D vertical seismic profile data: Geophysics, **81**, no. 1, S31–S38, http://dx.doi.org/10.1190/geo2015-0277.1.

van den Berg, E., and M. P. Friedlander, 2009, Probing the Pareto Frontier for basis pursuit solutions: SIAM Journal on Scientific Computing, **31**, 890–912, http://dx.doi.org/10.1137/080714488.

Wang, Y. H., 2016, Seismic Inversion: Theory and Applications: Wiley-Blackwell.

Xue, Z., Y. Chen, S. Fomel, and J. Sun, 2016, Imaging incomplete data and simultaneous-source data using least-squares reverse-time migration with shaping regularization: Geophysics, **81**, no. 1, S11–S20, http://dx.doi.org/10.1190/geo2014-0524.1.

Yao, G., and D. Wu, 2015, Least-squares reverse-time migration for reflectivity imaging: Science China. Earth Sciences, **58**, 1982–1992, http://dx.doi.org/10.1007/s11430-015-5143-1.

Yao, G., and H. Jakubowicz, 2016, Least-squares reverse-time migration in a matrix-based formulation: Geophysical Prospecting, **64**, 611–621, http://dx.doi.org/10.1111/1365-2478.12305.

Zhang, Y., H. Zhang, and G. Zhang, 2011, A stable TTI reverse time migration and its implementation: Geophysics, **76**, no. 3, WA3–WA11, http://dx.doi.org/10.1190/1.3554411.

Zhang, Y., L. Duan, and Y. Xie, 2015, A stable and practical implementation of least-squares reverse time migration: Geophysics, **80**, no. 1, V23–V31, http://dx.doi.org/10.1190/geo2013-0461.1.