CHINESE JOURNAL OF GEOPHYSICS Vol.52, No.3, 2009, pp: 704~715

ON ITERATIVE REGULARIZATION METHODS FOR MIGRATION DECONVOLUTION AND INVERSION IN SEISMIC IMAGING

WANG Yan-Fei, YANG Chang-Chun, DUAN Qiu-Liang

Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China

Abstract In this paper, we consider continuous solution methods for migration deconvolution imaging in seismic inverse problems. Direct migration methods, using the adjoint operator L^*, usually yield a lower resolution or blurred image. Linearized migration deconvolution requires solving a least-squares migration (LSM) problem. However, we notice that the direct LSM method is unstable in computation which is a severe obstacle for visual explanation. We study regularized mathematical model. We fist formulate the problem by incorporating regularizing constraints, and then employ iterative gradient methods for migration deconvolution and inversion. A hybrid gradient technique for ill-posed migration inverse problem is proposed. To show the potential for application of the proposed method, we perform synthetic one-, two- and three-dimensional seismogram for seismic migration inversion. Numerical performance indicates that the proposed method is very promising for practical seismic migration imaging.

Key words Migration deconvolution/inversion, Regularization, Optimization

1 INTRODUCTION

The forward model for representing a seismic imaging process is usually written as[1,2]:

$$Lm = d, \tag{1}$$

where L is the forward modeling kernel operator; m is the reflectivity model; d is the seismic data we measured.

A canonical method for the above problem is solving a variational problem

$$\min_m J[m] := \frac{1}{2}\|Lm - d\|^2, \tag{2}$$

such that the energy of $\|m\|$ is minimized to obtain an approximation to the reflectivity model m.

In seismic imaging, we usually solve a migration problem, i.e.,

$$m_{\text{mig}} = L^* d = L^* Lm, \tag{3}$$

which refers to a migration imaging, where L^* is the adjoint operator of L defined by $(Lx, y) = (x, L^* y)$; m_{mig} represents the migrated image; $L^* L$ denotes the blurring integral operator.

It is evident that an inverse operation can be performed on both sides of Eq.(3) if we assume that the inverse of $L^* L$ exists, i.e.,

$$m = (L^* L)^{-1} m_{\text{mig}} = (L^* L)^{-1} L^* d. \tag{4}$$

However we must recall the drawbacks of this operation: (1) $L^* L$ is much more ill-posed than L, thus $(L^* L)^{-1}$ is highly sensitive to the perturbations in data; (2) Huge computational cost. For migration deconvolution and inversion, developed methods include inversion based on resolution function[3], CG-based least squares migration deconvolution[4], nonstationary matching filters method[5] and regularizing seismic inversion imaging[6]. For general geophysical inversion method, many methods were developed, e.g., damped least squares method[7,8], preconditioning conjugate gradient method[9], restarted conjugate gradient method based on discrepancy principle[10], trust region algorithms[11], Bayesian inference and statistical inversion

E-mail: yfwang@mail.iggcas.ac.cn, yfwang_ucf@yahoo.com

techniques[12~15], regularization methods based on a priori information in Sobolev space[16], singular value decomposition method[1,17], least squares and various iterative implementations[18~20], mollification method and Backus-Gilbert inversion[21~23]. Some of the methods were developed in seismic migration imaging. However, the convergence rate of these methods for migration deconvolution and inversion may be slow[24].

We study regularizing methods for migration deconvolution and inversion, discuss the ill-posed nature of the migration deconvolution problem and the limitations of the least squares migration in computation, and propose a regularized hybrid gradient method for computation of migration deconvolution and inversion. Numerical simulations are also performed.

2 FUNDAMENTALS OF MIGRATION DECONVOLUTION AND INVERSION

2.1 Operator Equations of the First Kind

The blurring image processing problems are usually described by the operator equation of the first kind

$$\begin{aligned} &L: \ F \to O \\ &(Lm)(\boldsymbol{r}) := \int_{\Pi} l(\boldsymbol{r}, \boldsymbol{r}_0) m(\boldsymbol{r}_0) \mathrm{d}\boldsymbol{r}_0 = \mathrm{d}(\boldsymbol{r}), \end{aligned} \tag{5}$$

where L is a mapping from the function space F to the observation space O, and F and O are both assumed to be separable Hilbert spaces; $m(\boldsymbol{r})$ represents the input signal or object function; $d(\boldsymbol{r})$ represents the blurred noisy image; $l(\boldsymbol{r}, \boldsymbol{r}_0)$ denotes the so-called blurring kernel function. The integration is on the image plane Π. In engineering and image/signal processing, $l(\boldsymbol{r}, \boldsymbol{r}_0)$ is usually called the point spread function, which represents the approximation to a Dirac delta function[25,26].

In geophysical science, $l(\boldsymbol{r}, \boldsymbol{r}_0)$ usually refers to the kernel function which can be calculated by knowing some Green's functions. Under the assumption of shift invariant in horizontal of the imaging system, Eq.(5) reduces to a convolution problem. Hence the key problem is how to recover the input signal by some inversion methods[27,28].

2.2 Ill-Posed Properties

Ill-posedness is the basic property of the first kind operator equations. It is readily seen that the range of L in Eq.(5) is infinite dimensional and L is a compact operator[25,29,30]. Let us denote the singular system of the operator L in Eq.(5) as $\{\sigma_k; u_k, v_k\}$, we have that $Lu_k = \sigma_k v_k, L^* v_k = \sigma_k u_k$, where L^* is the adjoint of L, and

$$\begin{aligned} Lm &= \sum_{k=1}^{\infty} \sigma_k (m, u_k) v_k, \quad m \in F, \\ L^* d &= \sum_{k=1}^{\infty} \sigma_k (d, v_k) u_k, \quad d \in O. \end{aligned}$$

So the least squares solution with minimal norm of the original problem can be written as[25]

$$m^{\mathrm{lse}} = L^+ d = (L^* L)^+ L^* d = \sum_{k=1}^{\infty} \frac{(d, v_k)}{\sigma_k} u_k$$

where L^+ is the Moore-Penrose generalized inverse. A catastrophic result is that, if L^+ is unbounded, then small perturbations in observation will lead to significant deviation in recovering result to the true solution. This may indeed occur since L is a compact operator in infinite spaces[31,26].

2.3 Migration Deconvolution and Inversion Imaging

The aim of migration deconvolution and inversion is to solve a deconvolution problem by some inversion methods to obtain the reflectivity model[32]. The forward model for generating seismic data can be described by the first kind Fredholm integral equation

$$d(x,y,z) = \int_{\Omega} K(x,y,z|x_0,y_0,z_0) m(x_0,y_0,z_0) \mathrm{d}x_0 \mathrm{d}y_0 \mathrm{d}z_0, \tag{6}$$

and $m(x_0, y_0, z_0)$ can be obtained by solving the above equation, where $d(x, y, z)$ is the observed data, $K(x, y, z| x_0, y_0, z_0)$ is the kernel function which incorporates the velocity model, $m(x_0, y_0, z_0)$ is the reflectivity model in the position (x_0, y_0, z_0) and Ω is the integral space. Denote $\boldsymbol{r} = (x, y, z), \boldsymbol{r}_0 = (x_0, y_0, z_0)$, by Born approximation, the above equation can be written as[33,34]

$$d(\boldsymbol{r}_{\rm s}|\boldsymbol{r}_{\rm g},\omega) = \int_{\Omega} w(\omega) G(\boldsymbol{r}_{\rm g}|\boldsymbol{r}_0,\omega) G(\boldsymbol{r}_0|\boldsymbol{r}_{\rm s},\omega) m(\boldsymbol{r}_0) \mathrm{d}V_0, \tag{7}$$

where $d(\boldsymbol{r}_{\rm s}|\boldsymbol{r}_{\rm g},\omega)$ denotes the scattered data for a receiver at $\boldsymbol{r}_{\rm g}$, a source at $\boldsymbol{r}_{\rm s}$ and a reference imaging point at $\boldsymbol{r}_0$; $w(\omega)$ denotes the source wavelet; and $m(\boldsymbol{r}_0)$ denotes the reflectivity distribution function. The Green's function $G(\boldsymbol{r}_{\rm g}|\boldsymbol{r}_{\rm s},\omega)$ satisfies the Helmholtz wave equation and V_0 denotes the three-dimensional integration volume.

Steps for recording and deconvolution for synthetic seismic data are outlined as follows:

(1) Given receiver (geophone) and source, and a point-scatterer model m;

(2) Compute the point-scatterer response $d = Lm$;

(3) Compute the migrated image $m_{\rm mig} = L^* d = L^* L m$;

(4) Inversion computation: $m = (L^* L)^{-1} L^* d$,

where L^* is defined by $[w(\omega)G(\boldsymbol{r}_{\rm g}|\boldsymbol{r}_0,\omega)G(\boldsymbol{r}_0|\boldsymbol{r}_{\rm s},\omega)]^*$. The readers are referred to the recent works regarding migration deconvolution imaging for details[3,5,27,35~37].

2.4 Variational Regularization for Numerical Inversion

The complete regularization theory for solving inverse problems is established by Tikhonov and his group[29]. Tikhonov regularization refers to minimizing a regularized functional

$$J^{\alpha}[m] := \frac{1}{2}\|Lm - d\|^2 + \frac{\alpha}{2}\Gamma[m], \tag{8}$$

where $\Gamma[\cdot]$ is the Tikhonov stabilizer; $\alpha \in (0, 1)$ is the regularization parameter for balancing the instability and smoothness.

A simple way for choosing $\Gamma[m]$ is the $\frac{1}{2}\|D^{1/2}m\|^2$, where D is a scale operator, positive semi-definite and bounded. Minimizing $J^{\alpha}[m]$ reduces to the Euler equation

$$L^* L m + \alpha D m = L^* d, \tag{9}$$

and the minimizer can be expressed as

$$m = (L^* L + \alpha D)^{-1} L^* d. \tag{10}$$

For minimizing Eq.(8), a lot of methods can be applied, either direct methods or iterative methods[25]. However for seismic migration deconvolution problem, the computational cost would be huge. Studying fast and efficient methods for solving the inverse problem is meaningful and urgent.

3 LEAST SQUARES MIGRATION DECONVOLUTION

For solving large scale inverse problems, iterative methods are feasible. It is proved that the conjugate gradient method is an efficient method and also a kind of regularization[30]. In the following, we assume that the original problem is discretized into the finite spaces. We denote L as the discretizing from of the operator L, and still use m, d as the discrete arrays. We solve a least squares problem

$$\frac{1}{2}\|Lm - d\|^2 \to \text{minimization} \tag{11}$$

or a normal equation

$$L^{\mathrm{T}}Lm - L^{\mathrm{T}}d = 0. \tag{12}$$

Many methods can be used for solving the positive semi-definite system, e.g., the Cholesky decomposition. However this method is only useful for middle-scale problem with the magnitude of computational cost $O\left(\frac{1}{6}n^3\right)$. For large scale problems, applying stable iterative methods is much better.

The conjugate gradient method was performed in Refs.[4] and [3], where they describe the problem differently. The main computational cost is the matrix-vector multiplication with a lower order of the magnitude of computational cost than Cholesky. This method requires computing:

$$\begin{aligned}
&\text{Iterative stepsize } \ \alpha_k := -g_k^{\mathrm{T}} s_k/((Ls_k)^{\mathrm{T}}(Ls_k)),\\
&\text{New iterative point } \ m_{k+1} := m_k + \alpha_k s_k,\\
&\text{New direction } \ s_{k+1} := -g_{k+1} + \beta_k s_k,\\
&\text{Parameter } \ \beta_k := \frac{\|g_{k+1}\|^2}{\|g_k\|^2},
\end{aligned}$$

where g_k denotes the gradient at the k-th iteration, i.e., $g_k = L^{\mathrm{T}}(Lm_k - d)$; s_k denotes the search direction; the parameter β_k is computed by Fletcher-Reeves formula, we denote it by β_k^{FR}.

In ideal case, the conjugate gradient method should stop its iterations at the norm of residual equal to zero, i.e., $\|L^{\mathrm{T}}(Lm - d)\| = 0$. But for inverse problems with noisy data, the condition is too strict. According to the iterative regularization theory for inverse problems, a saturation state will be reached after a couple of iteration cycles[30]. For conjugate gradient method, this may not occur, but according to the idea of regularization, the iterative step k is closely related with the error/noise level δ in data and operators. Actually, the number k can be written as $O(\delta^{-\frac{1}{2\nu+1}})$, ν is related with the operator source conditions and algorithms[25]. Empirically, we recommend that the maximum iterative number should be less than 50.

4 LIMITATIONS OF STANDARD MIGRATION AND LEAST SQUARES MIGRATION DECONVOLUTION

The original work for migration deconvolution can be found in Ref.[32] and the realization can be found in Ref.[4]. There are some shortcomings of the standard migration and least squares migration:

(1) Incorrect velocity estimation and improper choice of the wavelet function may lead to low quality of image (low resolution);

(2) Lower control of noise propogation of the standard migration;

(3) Large computation cost for solving the least squares problem;

(4) Difficulties in choosing proper preconditioner[10].

To overcome these shortcomings, developing new methods is necessary.

5 ITERATIVE REGULARIZATION FOR MIGRATION DECONVOLUTION AND INVERSION

5.1 Iterative Regularization

Regularization refers to finding a group of regularizing operators $R(\alpha, d)$ such that as the error/noise level in data d approaches zero, $R(\alpha, d)$ can converge to the true reflectivity model m. This requires solving a nonlinear programming problem

$$J^{\alpha}[m] \to \text{minimization}, \tag{13}$$

about m, where α can be adjusted and $\Gamma[m]$ is chosen by users. If we choose $\Gamma[m]$ as $\frac{1}{2}\|D^{1/2}m\|^2$, where D represents the discretizing for of the operator D, then $R(\alpha, \cdot)$ can be expressed as

$$R(\alpha) = (L^{\mathrm{T}}L + \alpha D)^{-1}L^{\mathrm{T}}\cdot, \alpha > 0. \tag{14}$$

So the reflectivity model can be obtained by

$$m = R(\alpha)d, \ \alpha > 0. \tag{15}$$

The commonly called iterative regularization method refers to the iterative Tikhonov regularization[25,26], which is in the form

$$(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L} + \alpha\boldsymbol{D})m_{k+1} = \boldsymbol{L}^{\mathrm{T}}d + \alpha m_k, \quad k = 1, 2, \cdots, \alpha > 0. \tag{16}$$

The advantage of the method is that the optimal convergence rate can be obtained under source conditions of operators. The disadvantage of the method is that it is not suitable for large scale computation. In addition, the steepest descent method and the Landweber-Fridman iterative method are also commonly used iterative regularization methods[25,26,30]. These two methods possess the same iterative formula

$$m_{k+1} = m_k + f(\tau)\boldsymbol{L}^{\mathrm{T}}(d - \boldsymbol{L}m_k), \ \ k = 1, 2, \cdots, \alpha > 0. \tag{17}$$

If $f(\tau)$ is chosen as a constant in $(0, 1/\|\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}\|)$, it corresponds to the Landweber-Fridman iterative method; if $f(\tau) = \mathrm{argmin}_{\tau} J^{\alpha}(m_k - \tau r_k)$, where $r_k = \boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}m_k - \boldsymbol{L}^{\mathrm{T}}d$, it corresponds to the steepest descent method. It is easily seen that if we set the initial guess value $m_0 = 0$, and set $k = 1, f(\tau) \equiv 1$, the one-step steepest descent method and the one-step Landweber-Fridman iterative method lead to the Kirchhoff time migration. It is well-known that the one-step steepest descent method or the one-step Landweber-Fridman iterative method are far from the accuracy of convergence, so it is no doubt that the standard migration leads to low resolution imaging.

5.2 Hybrid Conjugate Gradient Iterative Regularization

The conjugate gradient method is convergent in finite steps for quadratic programming problems[38]. The Fletcher-Reeves conjugate gradient method is mostly used, e.g., the literature [4]. But this method cannot guarantee the descent property of the algorithm. Therefore we re-consider this kind of method. It is known that the Polak-Ribiere-Polyak method can yield the parameter β_k approaching zero for small iterative stepsize, hence restart the negative gradient and accelerate convergence. The good property of Fletcher-Reeves method is its good convergence. Therefore, it is expected that combining these two methods yields better results. For well-posed nonlinear programming problems, this method is proved to be successful[39]. We propose applying it to ill-posed problems. The Polak-Ribiere-Polyak method requires that the parameter β_k satisfies

$$\beta_k = \frac{g_{k+1}(g_{k+1} - g_k)}{\|g_k\|^2}, \tag{18}$$

and we denote it as β_k^{PRP}. According to the suggestion from Ref.[39], we select the parameter as

$$\beta_k = \max\{0, \min\{\beta_k^{\mathrm{FR}}, \beta_k^{\mathrm{PRP}}\}\}. \tag{19}$$

Since $J^{\alpha}[m]$ is a quadratic function in high dimension, therefore, the iterative stepsize can be expressed explicitly as

$$\alpha_k := -g_k^{\mathrm{T}} s_k / ((\boldsymbol{L}s_k)^{\mathrm{T}}(\boldsymbol{L}s_k) + \alpha s_k^{\mathrm{T}}(\boldsymbol{D}s_k)). \tag{20}$$

The regularizing term in the denominator can help to avoid leap-type jump and ensure convergence.

6 NUMERICAL SIMULATIONS

Numerical simulation is a necessary step to prove the efficiency of our proposed method before applying it to solving practical application problems. Since d is the observation, it inevitably contains various noises. We assume that the noise is additive and mainly Gaussian, i.e.,

$$d = d_{\mathrm{true}} + \delta \cdot \mathrm{rand}(\mathrm{size}(d_{\mathrm{true}})),$$

where δ is the noise level in (0,1), rand(size(d_{true})) denotes the Gaussian noise with the same dimension as d_{true}, d_{true} is computed by computer through forward model. In these numerical simulations, we choose the noise level between 0.001 and 0.03.

In constructing the regularization operator $R(\alpha, d)$, we need to choose the regularization parameter α, which is vital for successful inversion[25]. If α is significantly greater than 1, the approximation will be poor to the original problem; on the contrary if α is far below 1, the spectrum of the operator cannot be improved and the noise/error cannot be well controlled. In this numerical simulation, we choose α as a priori value in (0,1) and $\alpha > \delta^2$. This choice can guarantee the convergence of the algorithm[30].

6.1 One-Dimensional Seismic Migration Deconvolution/Inversion Imaging

The one-dimensional synthetic seismograms are very important in seismic interpretation. We consider a simple synthetic seismogram generated by a Ricker wavelet convolved with an input signal with 3 peaks and 3 valleys. The central frequency of the Ricker wavelet is 50 Hz, the sampling interval is 0.001 second. The initialization of our algorithm is as follows: the initial model is chosen as zero, the regularization parameter is set as 0.005. The noise-free seismogram, the noisy seismogram with noise level equal to 0.01 and 0.03 are plotted in Figs. 1, 2a and 2b, respectively. The reconstructions corresponding to noise level 0.01 and 0.03 are plotted in Figs. 3a and 3b, respectively. We also generate standard migration images for noise level equal to 0.01 and 0.03 shown in Figs. 4a and 4b. It is apparent that the reconstructions by regularizing migration deconvolution and inversion show correct amplitude and better resolution.

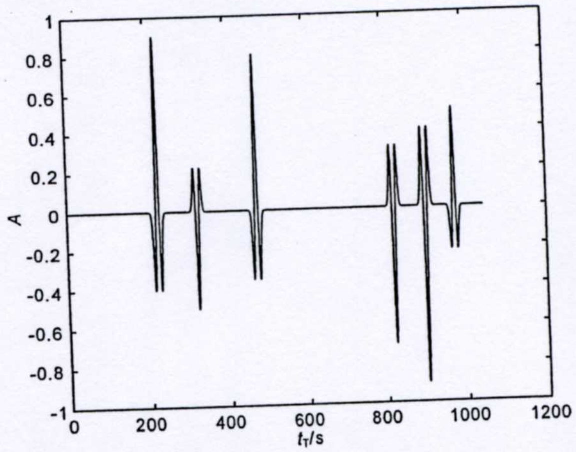

Fig. 1 The noise-free 1D seismogram

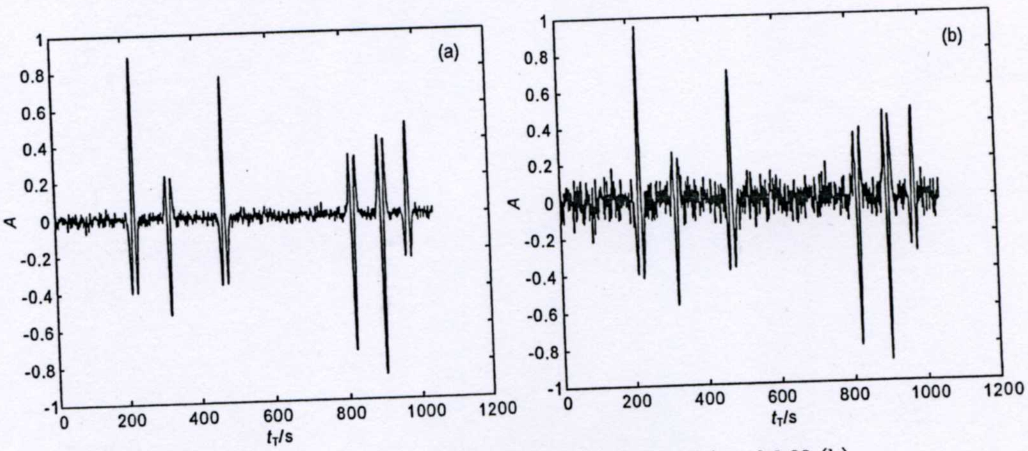

Fig. 2 The seismogram as noise level equal to 0.01 (a) and 0.03 (b)