GEOPHYSICS, VOL. 75, NO. 4 (JULY-AUGUST 2010); P. S131–S137, 12 FIGS., 5 TABLES.
10.1190/1.3457923

Accelerating migration deconvolution using a nonmonotone gradient method

Yanfei Wang[1] and Changchun Yang[1]

ABSTRACT

New solution methods were considered for migration deconvolution in seismic imaging problems. It is well known that direct migration methods, using the adjoint operator L^*, yield a lower-resolution or blurred image, and that the linearized inversion of seismic data for the reflectivity model usually requires solving a (regularized) least-squares migration problem. We observed that the (regularized) least-squares method is computationally expensive, which becomes a severe obstacle for practical applications. Iterative gradient-descent methods were studied and an efficient method for migration deconvolution was developed. The problem was formulated by incorporating regularizing constraints, and then a nonmonotone gradient-descent method was applied to accelerate the convergence. To test the potential of the application of the developed method, synthetic two-dimensional and three-dimensional seismic-migration-deconvolution simulations were performed. Numerical performance indicates that this method is promising for practical seismic migration imaging.

INTRODUCTION

Standard seismic migration methods give distorted images of the subsurface, even with an accurate velocity model because of limitations in bandwidth, recording time, and aperture of the seismic-reflection experiment. Application of inverse methods can improve the resolution of seismic images by compensating for these distortions (Treitel and Lines, 2001). Migration deconvolution is one possible approach because it uses knowledge of the resolution kernel of the seismic experiment to compensate for the effects mentioned (Schuster, 1997a, b; Hu and Schuster, 1998; Hu and Valasek, 1999; Hu and Schuster, 2000; Hu et al., 2001; Gelius et al., 2002; Sjøberg et al., 2003). However, it is well known that deconvolution approaches to this problem are ill-posed, rendering the results sensitive to noise (Tikhonov and Arsenin, 1977). Least-squares migration (LSM), particularly with the addition of regularization, is another promising approach to improving the resolution of migrated images (Schuster, 1997a; Nemeth et al., 1999; Sacchi et al., 2006; Wang et al., 2009). The key to successful, iterative least-squares migration is the optimization strategy (VanDecar and Snieder, 1994; Treitel and Lines, 2001; Bevc et al., 2007). Because least-squares migration has a cost approximately equal to two migration applications per iteration, it is essential to control the number of iterations required to reach a satisfactory image (Schuster, 1997a, b; Nemeth et al., 1999; Sjøberg et al., 2003; Wang et al., 2009). We have adapted a gradient-descent optimization technique described by Barzilai and Borwein (1988) to least-squares seismic migration. One key feature of this approach, as opposed to the typical steepest-descent or conjugate-gradient (CG) approaches, is that it does not have monotonic convergence. The advantage of relaxing the requirement that each iteration must have a lower residual than the previous step is that the method can converge to a lower overall value of the objective function in fewer overall iterations. This is important for least-squares migration because one wishes to keep the overall number of iterations as small as possible.

ITERATIVE REGULARIZATION SOLVERS FOR MIGRATION DECONVOLUTION AND INVERSION IN SEISMIC IMAGING

Migration deconvolution

In Schuster (1997a, b) and Hu and Schuster (1998), the term *migration deconvolution* describes inverting the blurring operator from the migration image and thus creating a sharper image of the reflectivity. Nemeth et al. (1999) show that by modeling and migration of synthetic data, we can understand better the impact of the noise (e.g., recording footprint). For example, one might specify source-receiver geometry and a reflectivity model m, compute synthetic seismic data d using some modeling operator L, and then migrate these data to obtain the migrated section m_{mig} (Claerbout, 1985;

Manuscript received by the Editor 24 June 2009; revised manuscript received 25 February 2010; published online 18 August 2010.
[1]Chinese Academy of Sciences, Institute of Geology and Geophysics, Key Laboratory of Petroleum Geophysics, Beijing, China. E-mail: yfwang@mail.iggcas.ac.cn, yfwang_ucf@yahoo.com; ccy@mail.iggcas.ac.cn.

Schuster, 1997a, b; Hu and Schuster, 1998; Hu et al., 2001; Yu et al., 2006):

$$m_{\text{mig}} = L^* d = L^* L m, \tag{1}$$

where m_{mig} is the blurred migration image of the reflectivity; L^*L is the integral blurring kernel operator, which is represented by a so-called resolution function (or point-spread function); and the adjoint operator L^* is the integral migration operator. A more general blurring equation is given in Gelius et al. (2002), where they formulate the same problem in different form. The resolution function is important for optimization of survey planning, i.e., defining aperture, sampling, and location. It also can be used to guide the selection of migration parameters (frequency band, frequency sampling, shot/receiver sampling, image sampling) for a well-resolved image (Lecomte and Gelius, 1998). However, the inverse of the operator L^*L should be only approximated (Guitton, 2004; Valenciano et al., 2006).

Gradient-descent least-squares migration

Schuster (1997a), Nemeth et al. (1999), and Sjøberg et al. (2003) use the conjugate-gradient method for least-squares migration. A prototype of the CG algorithm is outlined in Table 1. Nemeth et al. (1999) introduce preconditioning techniques for regularizing this inverse scheme. As pointed out by Schuster (1997a), LSM is expensive because each iteration requires about the same computational cost as two standard migrations. Therefore, a key enabler of practical implementation is to find fast and efficient computational methods for implementing LSM (Nemeth et al., 1999).

We consider the minimization of a "regularized functional," i.e.,

$$J^\alpha[m] := \tfrac{1}{2}\|Lm - d\|^2 + \tfrac{1}{2}\alpha\|Dm\|^2 \to \min, \tag{2}$$

where ":=" means *defined by*, D is a (semi-) positive-definite-bounded scale operator, and $\alpha \in (0,1)$ is the regularization parameter. A very simple steepest-descent method can be applied to solve equation 2, i.e.,

$$m_{k+1} = m_k + \omega_k s_k, \tag{3}$$

where $s_k = -g_k$, g_k is the gradient of J^α, and ω_k is the step size, which can be obtained by line search. If we set $\alpha = 0$ and restrict step size ω_k to a constant in the interval $(0, \|L\|^{-2})$ in each iteration, we obtain a special gradient method known as the Landweber iteration (Wang, 2007). However, these methods are slow in convergence and difficult to use on practical problems. One might notice that the standard migration is just one step of steepest-descent iteration if we set $m_0 = 0$, $\alpha = 0$, and choose step size ω as unity. It is well known that one step of gradient iteration is generally far from convergence, hence the resolution of standard migration imaging is low.

Table 1. Conjugate-gradient method for migration deconvolution.

Step 1. Given initial point $\mathbf{m}_0 = 0$: compute $\mathbf{g}_0 = \mathbf{g}(\mathbf{m}_0) = -\mathbf{L}^{\mathrm{T}}\mathbf{d}$, choose $\alpha \in (0,1)$, and set $\epsilon > 0$.

Step 2. If $\|\mathbf{g}_0\| \le \epsilon$, output $\mathbf{m}_0$, STOP; otherwise, set $\mathbf{s}_0 := -\mathbf{g}_0$, and $k := 0$.

Step 3. Compute next iteration point:

$$\alpha_k := -\mathbf{g}_k^{\mathrm{T}}\mathbf{s}_k / ((\mathbf{L}\mathbf{s}_k)^{\mathrm{T}}(\mathbf{L}\mathbf{s}_k) + \alpha(\mathbf{D}\mathbf{s}_k)^{\mathrm{T}}(\mathbf{D}\mathbf{s}_k)),$$
$$\mathbf{m}_{k+1} := \mathbf{m}_k + \alpha_k \mathbf{s}_k,$$
$$\mathbf{g}_{k+1} := \mathbf{g}_k + \alpha_k \mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{s}_k + \alpha\alpha_k \mathbf{D}^T\mathbf{D}\mathbf{s}_k$$
$$\beta_k^{\mathrm{FR}} := \|\mathbf{g}_{k+1}\|^2 / \|\mathbf{g}_k\|^2,$$
$$\mathbf{s}_{k+1} := -\mathbf{g}_k + \beta_k^{\mathrm{FR}}\mathbf{s}_k.$$

Step 4. If $\|\mathbf{g}_{k+1}\| \le \epsilon$ or k exceeds the maximum iterative steps, output $\mathbf{m}_k$, STOP; otherwise, set $k := k + 1$, GOTO Step 3.

A fast-iterative migration-deconvolution solver: Nonmonotone gradient method

We apply a nonmonotone gradient method, called the BB method, described by Barzilai and Borwein (1988) to seismic migration-deconvolution imaging by solving equation 2. From now on, we assume that the seismic modeling is formulated in finite space after discretization and the norm $\|.\|$ is in the ℓ_2 sense. We use $\mathbf{L}^{\mathrm{T}}$ to represent transpose of a matrix $\mathbf{L}$ and $(\cdot,\cdot)$ the inner product of two vectors. Let $\mathbf{m}_k$ be the k-th iteration and $\mathbf{g}_k$ the gradient of J^α at $\mathbf{m}_k$, where $\mathbf{g}_k = \mathbf{g}(\mathbf{m}_k) = \mathbf{L}^{\mathrm{T}}(\mathbf{L}\mathbf{m}_k - \mathbf{d}) + \alpha\mathbf{D}^{\mathrm{T}}(\mathbf{D}\mathbf{m}_k)$. The nonmonotone gradient method aims to accelerate the convergence of the steepest-descent method and requires few storage locations and inexpensive computations. The BB method incorporates a quasi-Newton property with the gradient method to obtain second-order information about objective function $J^\alpha[\mathbf{m}]$. Specifically, it approximates the Hessian $\nabla^2 J^\alpha[\mathbf{m}_k]$ by $\nu_k\mathbf{I}$ and solves the two minimization problems $\nu_k = \operatorname{argmin}_\nu\|\mathbf{y}_{k-1} - \nu\mathbf{I}\mathbf{s}_{k-1}\|$ and $\nu_k = \operatorname{argmin}_\nu\|\nu\mathbf{I}\mathbf{y}_{k-1} - \mathbf{s}_{k-1}\|$. This leads to two choices of the step size ν_k

$$\nu_k^{\mathrm{BB1}} = \frac{(\mathbf{s}_{k-1},\mathbf{s}_{k-1})}{(\mathbf{s}_{k-1},\mathbf{y}_{k-1})}, \quad \nu_k^{\mathrm{BB2}} = \frac{(\mathbf{s}_{k-1},\mathbf{y}_{k-1})}{(\mathbf{y}_{k-1},\mathbf{y}_{k-1})}, \tag{4}$$

where $\mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$, $\mathbf{s}_k = \mathbf{m}_{k+1} - \mathbf{m}_k$. Note that from equation 4, the inverse of the scalar ν_k is the Rayleigh quotient of $\mathbf{A}$ or $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ at the vector $\mathbf{g}_{k-1}$, where $\mathbf{A} = \mathbf{L}^{\mathrm{T}}\mathbf{L} + \alpha\mathbf{D}^{\mathrm{T}}\mathbf{D}$. Hence, step sizes inherit the regularized spectrum of the regularization model. Therefore, we believe that this method is very efficient for solving ill-posed convex-quadratic-programming problems (Wang and Ma, 2007).

We consider an extremely ill-conditioned linear algebraic N-dimensional problem

$$\mathbf{L}\mathbf{m}_{\text{input}} = \mathbf{d}_{\text{observe}} = \mathbf{d}_{\text{true}} + \mathbf{noise},$$

where $\mathbf{L} = [\ell_{ij}]_{N\times N}$, $\ell_{ij} = 1/(i+j-1)$; $m_{\text{input}}(i) = \frac{1}{2}\exp(-2*t(i))$, $t(i) = a + h*i, i = 0,1,\cdots,N-1$, $h_i = (b-a)/(N-1)$, and $a = -3.14$, $b = 3.14$; $\mathbf{noise} = [n_0,n_1,\cdots,n_{N-1}]^T$;. $\mathbf{d}_{\text{observe}}(i) = \mathbf{d}_{\text{true}}(i) + \mathbf{noise}(i) = \Sigma_{j=0}^{N-1}\ell_{ij}\mathbf{m}_{\text{input}}(j) + n_i$. Choosing $\mathbf{d}$ as the identity, gradient $\mathbf{g}(\mathbf{m})$ of the functional $J^\alpha[\mathbf{m}]$ can be computed easily as $\mathbf{g}(\mathbf{m}) = (\mathbf{L}^{\mathrm{T}}\mathbf{L} + \alpha\mathbf{I})\mathbf{m} - \mathbf{L}^{\mathrm{T}}\mathbf{d}_{\text{observe}}$. In this simulation, the value of N (the number of unknowns) is set to 4000, leading to an extremely ill-conditioned system of equations. The condition number for $\mathbf{L}$ is up to 1.2517×10^{22}. Using the BB nonmonotone iteration, the nonmonotonicity of the object functional $J^\alpha[\mathbf{m}]$ can be seen vividly in Figure 1. The monotonicity of the object functional $J^\alpha[\mathbf{m}]$ for steepest-descent and CG methods are shown in Figures 2 and 3, respectively. The efficiency comparison between steepest-descent, BB nonmonotone, and CG methods is listed in Table 2. This example reveals that the BB nonmonotone iterative method is superior to the steepest-descent and CG meth-

ods. This observation motivates us to apply BB nonmonotone iteration to the seismic-migration-deconvolution problem.

In our algorithm, the initial search direction is the negative gradient direction. In addition, we use the following stopping condition in our numerical tests:

$$\|\mathbf{g}_k\| \le \varrho \|\mathbf{g}_1\|, \tag{5}$$

where $\varrho > 0$ is a preassigned tolerance. Now we state our algorithm conveniently for solving realistic seismic inverse problems in Table 3.

The regularization property of the algorithm is controlled by regularization parameter α and dominant parameter ϱ. The role of ϱ is to control further iteration at preassigned precision. Choice of ϱ depends on the degree of ill-posedness of problems. Our algorithm with lower values of ϱ could yield better approximation to the solution, but it might take more iterative steps and more CPU time in the implementation. The algorithm with larger values of ϱ can be implemented with fewer iterative steps and less CPU time, but could lead to insufficient approximation. From our simulations, we recommend choosing ϱ between 1.0×10^{-5} and 1.0×10^{-2}.

Vital advantages of the nonmonotone BB iterative method are its simplicity of implementation, fast nonmonotonic convergence, and extremely low memory requirements. The BB method was designed originally for well-posed quadratic programming problems which have received more attention in recent years (Yuan, 2008). However, the BB method is applicable to ill-posed image-restoration problems, and Wang and Ma (2007) demonstrate it to be a regularization method. These observations make us believe the algorithm will be very useful for solving large-scale, ill-posed migration-deconvolution problems.

NUMERICAL SIMULATIONS

Our simulation consists of two steps. First, a simulated seismic signal (input signal) is generated by forward modeling of acoustic data. Then, the input seismic signal is processed through our algorithm and the retrieved distribution signal is compared with the input one. Note that **d** is the observation, thus different kinds of noise also could be recorded besides the true signal. Here, we consider a simple case only, i.e., we assume that the noise is mainly additive Gaussian noise, i.e., $\mathbf{d} = \mathbf{d}_{\text{true}} + \delta \cdot \text{rand}(\text{size}(\mathbf{d}_{\text{true}}))$, where δ is the noise level in (0,1), $\text{rand}(\text{size}(\mathbf{d}_{\text{true}}))$ is the Gaussian random noise with the same size as $\mathbf{d}_{\text{true}}$.

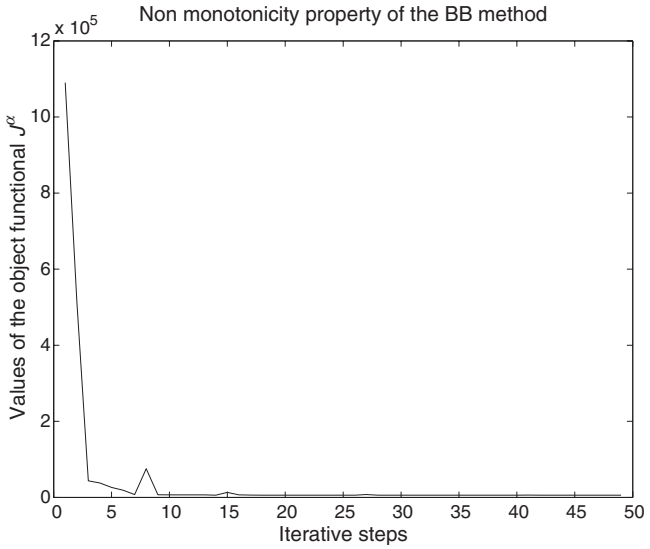


Figure 1. Fast nonmonotonic decreasing of the object function.

Two-dimensional simulations

The first example is a 2D migration-deconvolution problem. We assume receivers are uniformly distributed on a line with maximum line length of 1200 m. A sampling interval of 30 m in the x-coordinate is assumed. We use a Ricker wavelet with central frequency of 27 Hz that generates 40 traces of seismic data; the background velocity is homogeneous with $c = 3000$ m/s, and the time sampling interval is $dt = 4$ ms. In this simulation, grid dimensions of the model are 1200×1200 points with x and z gridpoint spacing of $dx = dz = 30$ m. Nine point scatterers are buried at depths from 270 to 870 m. The seismogram from the nine-point-scatterer model with noise level equal to 0.005 is illustrated in Figure 4. It is seen clearly that noises are very strong and parts of the data are contaminated by noise.

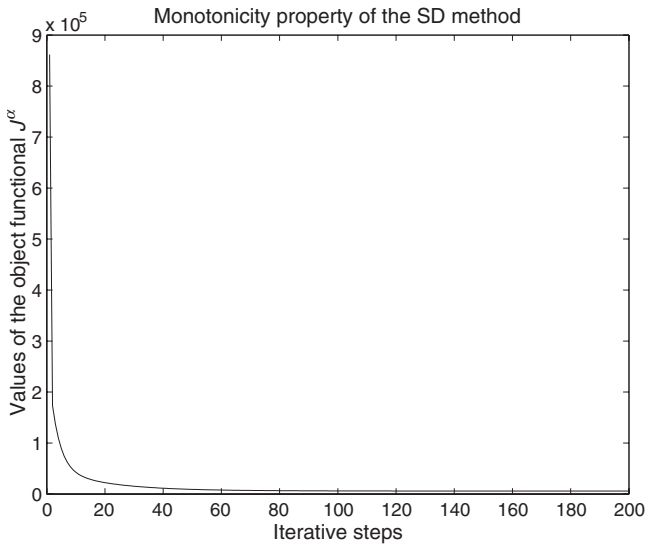


Figure 2. Monotonic decreasing of the object function.

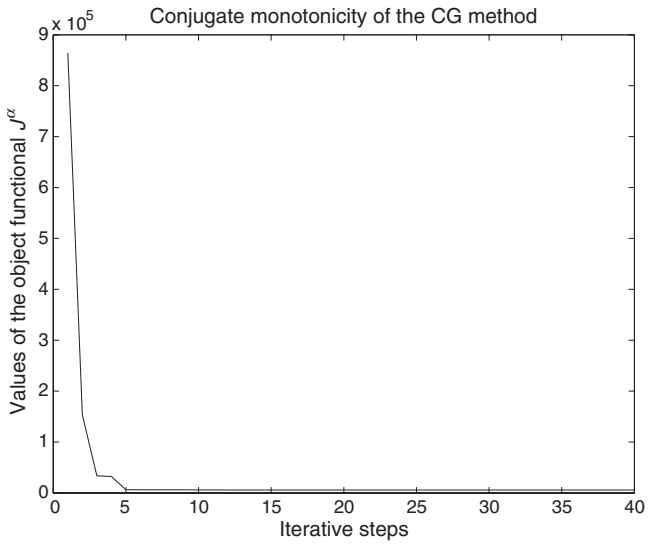


Figure 3. Conjugately monotonic decreasing of the object function.

Numerical simulations are made for the CG-based LSM algorithm (Table 1) and the fast gradient algorithm (Table 3). The noise level δ for both algorithms is assumed to be 0.005. For the CG-based LSM algorithm, the iteration process is stopped at 43 iteration cycles. For the fast gradient algorithm, the iteration process is terminated after 34 iteration cycles. The CPU time costs for LSM and for the fast gradient algorithm are 12.9688 s and 5.7969 s, respectively. The fast algorithm converges for all noise levels in fewer than 40 iterative steps. The CG-based LSM algorithm converges before reaching the maximum iterative number for small noise levels. For large noise levels, 40 iteration cycles are not enough for ensuring convergence of LSM. However, more iteration cycles induce more CPU time.

Table 2. Comparison of CPU time (s), iterative steps, and rms errors of the fast nonmonotone-gradient method with conjugate-gradient and steepest-descent method.

	Noise level	0.005	0.01	0.05
CPU time (s)	Nonmonotone gradient	3.6094	3.6719	3.5938
	Steepest descent	27.0938	27.0000	27.2344
	Conjugate gradient	25.1094	25.2812	25.3594
Steps	Nonmonotone gradient	50	50	49
	Steepest descent	200	200	200
	Conjugate gradient	40	40	40
RMS errors	Nonmonotone gradient	0.0016	0.0016	0.0016
	Steepest descent	0.0036	0.0036	0.0036
	Conjugate gradient	0.0013	0.0013	0.0013

Table 3. Nonmonotone-gradient method for migration deconvolution.

Step 1. Initialization: Given initial point $\mathbf{m}_0$, tolerance $\varrho > 0$, set $k := 1$; Compute $\mathbf{g}_k$.

Step 2. Check whether the stopping condition holds:
If equation 5 is satisfied, STOP; otherwise, set $\mathbf{s}_k = -\mathbf{g}_k$

Step 3. Compute a BB step: ν_k by equation ν_k^{BB1} or ν_k^{BB2}.

Step 4. Update the current iteration point by setting $\mathbf{m}_{k+1} = \mathbf{m}_k + \mathrm{v}_k\mathbf{s}_k$ with projection.

Step 5. Compute a new search direction $\mathbf{s}_{k+1} = -\mathbf{g}_{k+1}$ and go to Step 6.

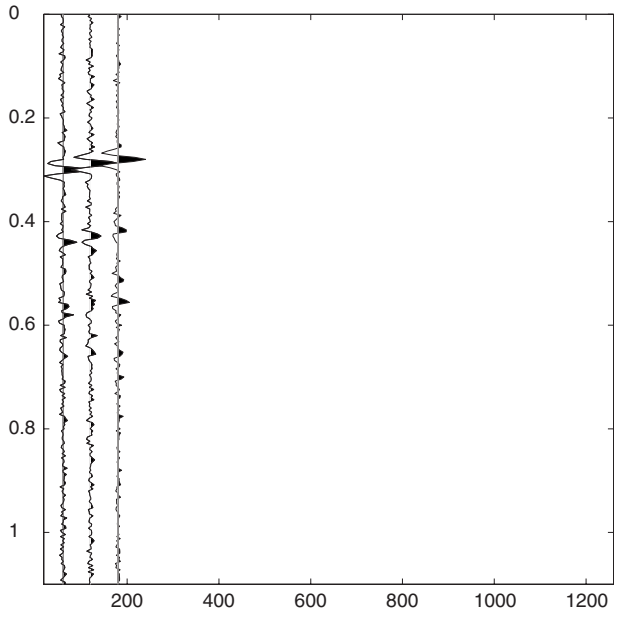
Step 6. Loop: $k := k + 1$, and go to Step 2.

Precision of the approximation is characterized by the root mean-square (rms) error

$$\mathrm{RMSE} = \sqrt{1/(MN)\Sigma_{i,j}(\mathbf{d}_{\mathrm{comp}}(x_i,t_j) - \mathbf{d}_{\mathrm{meas}}(x_i,t_j))^2/(\mathbf{d}_{\mathrm{comp}}(x_i,t_j))^2}.$$

This describes the average relative deviation of retrieved seismic signals from true seismic signals. In the above expression of rms error, $\mathbf{d}_{\mathrm{comp}}$ refers to retrieved (computed) signals, $\mathbf{d}_{\mathrm{meas}}$ refers to synthetic (measured) signals, M is the trace number, and N is traveltime length. Simulation results are shown in Table 4. Both algorithms yield the same values of rms error for different noise levels. This indicates that both algorithms are stable for solving LSM deconvolution problems. Comparison of computational cost for recovering the point-scatterer model also is recorded in Table 4. It is clear from the table that our method needs less CPU time than LSM for a successful inversion. Therefore, the fast gradient method supplies us with another choice when solving the LSM model.

Simulation results in Figures 5–7 represent, respectively, the standard migration image, the least-squares migration-deconvolution image, and the fast migration-deconvolution image, for a noise level equal to 0.005. It is clear that the fast gradient-descent migration-deconvolution method yields the best resolution result because of its fast convergence property.



Three-dimensional simulations

We show the applicability of our algorithm for 3D migration-deconvolution imaging problem. We perform experiments on a ten-point scatterer model. These scatterers are buried at a depth of 2000 m. An impulsive source wavelet with central frequency of 30 Hz is used to generate the data. Then the data are a superposition of ten hyperbolic events. We assume receivers are distributed uniformly on a 21×21 orthogonal grid with a sampling interval of 40 m in the x-direction and 20 m in the y-direction. Background velocity is homogeneous with $c = 5000$ m/s, and the time sampling interval is 1.7 ms. Grid dimensions of the model are $21 \times 21 \times 6$, with x, y, and z gridpoint spacing of 40 m, 20 m, and 200 m, respectively.

Results for the standard migration, LSM deconvolution, and fast migration deconvolution when noise level equals 0.01 are illustrated in Figures 8–10, respectively. To show convergence characteristics of our algorithm, we plot the relative errors in each step for a noise level equal to 0.01 in Figure 11. For comparison, we also plot the rel-

Table 4. Comparison of CPU time (s), iterative steps, and rms errors of our algorithm with LSM for 2D migration-deconvolution imaging.

	CPU time (s)		Steps		rms errors	
Noise level	LSM	Fast method	LSM	Fast method	LSM	Fast method
0	9.8906	5.5938	34	34	0.0135	0.0135
0.001	12.5938	4.3225	40	27	0.0135	0.0135
0.003	12.7813	4.7344	42	28	0.0135	0.0135
0.005	12.9688	5.7969	43	34	0.0135	0.0135
0.01	13.7812	6.2656	44	39	0.0135	0.0135

ative errors of the LSM in each step for a noise level equaling 0.01 in Figure 12. To illustrate the computational efficiency, we also apply our method to a fine-gridded synthetic problem with dimensions $41 \times 41 \times 6$, which yields 10,086 unknowns. The comparison of computational cost for the ten-point scatterer model is given in Table 5. Because CPU time is similar for the noise case and the noise-free case, we show only the values of the noise-free case. Table 5 indicates that our method needs approximately one-half the CPU time of the LSM to get similar migration-deconvolution results. Hence, our method would be promising for practical applications.

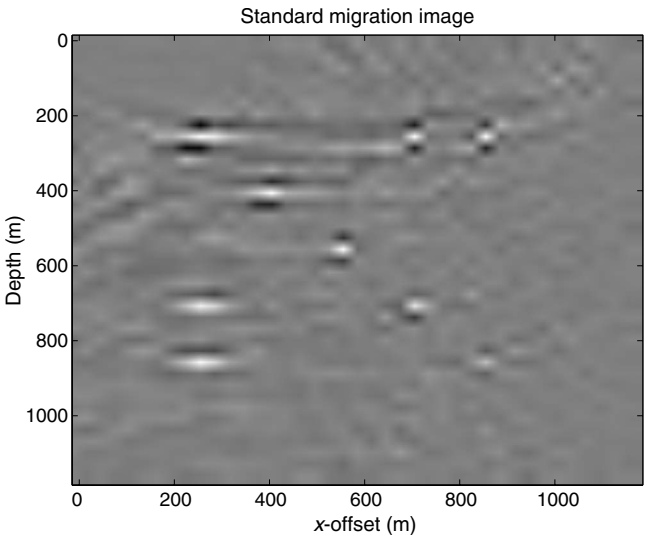


Figure 5. The standard migration image for the nine-point scatterer model with noise level equal to 0.005.

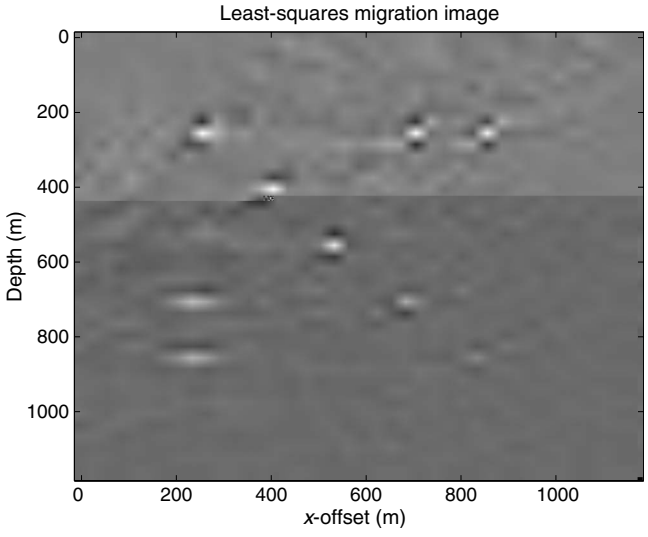


Figure 6. The least-squares migration-deconvolution image for the nine-point scatterer model with noise level equal to 0.005.

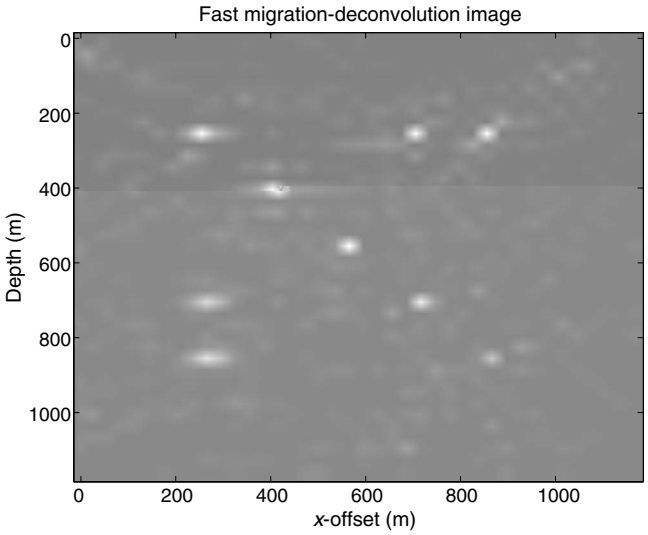


Figure 7. The fast migration-deconvolution image for the nine-point scatterer model with noise level equal to 0.005.

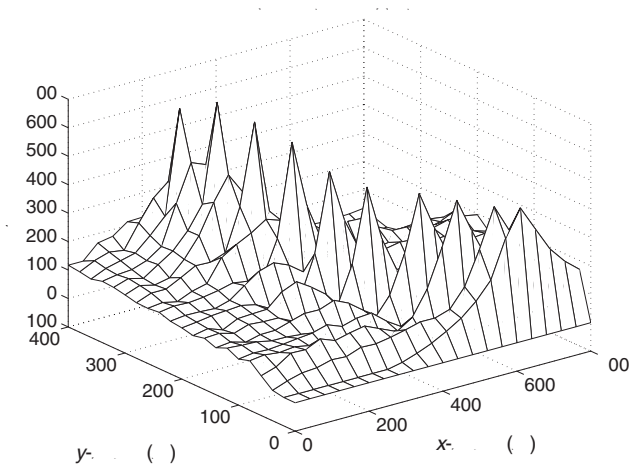
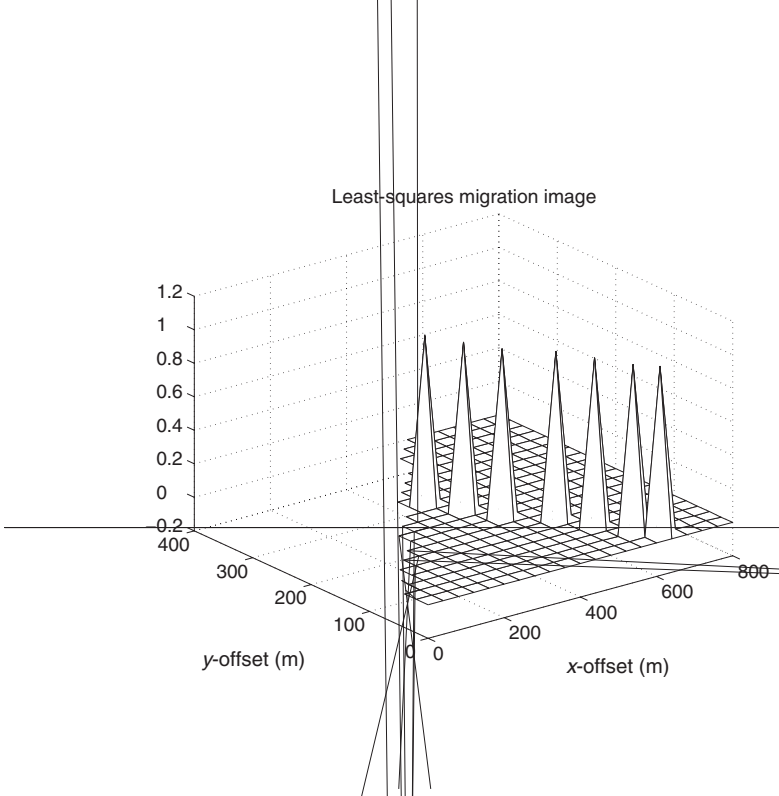


Figure 8. The standard migration image for the ten-point scatterer model with noise level equal to 0.01.



DISCUSSION

Our simulations reveals that CG-based LSM deconvolution and the fast nonmonotone gradient descent method yield better resolution images than standard migration. For n-dimensional quadratic programming problems, the CG method can converge in a finite n steps. However, CG iterations should not be driven that far for ill-posed deconvolution problems. At the same time, the CG iterations should not be terminated at a very early stage because earlier terminations could filter out noise components of the solution corresponding to the small singular values of matrix operator $\mathbf{L}$.

In our comparison, we do not consider preconditioning of the regularized conjugate-gradient iteration. We give a theoretical explanation as follows. We assume that $\mathbf{D}$ and $\mathbf{L}$ are $p \times n$ and $m \times n$ matrices, respectively. By singular value decomposition of the matrix pair $(\mathbf{L},\mathbf{D})$ (Golub and Van Loan, 1989; Hanke and Hansen, 1993), we have

$$\mathbf{L} = \mathbf{U}\begin{pmatrix} \mathbf{\Sigma}_a & 0 \\ 0 & \mathbf{I}_{n-p} \end{pmatrix}\mathbf{W}^{-1}$$

and

$$\mathbf{D} = \mathbf{V}(\mathbf{\Sigma}_b \quad 0)\mathbf{W}^{-1},$$

where $\mathbf{U} = (\mathbf{u}_1,\ldots,\mathbf{u}_m)$, $\mathbf{W} = (\mathbf{w}_1,\ldots,\mathbf{w}_n)$, $\mathbf{V} = (\mathbf{v}_1,\ldots,\mathbf{v}_p)$, $\mathbf{I}$ is the identity, $\mathbf{\Sigma}_a = \mathrm{diag}(\sigma_1^a,\ldots,\sigma_p^a)$, $\mathbf{\Sigma}_b = \mathrm{diag}(\sigma_1^b,\ldots,\sigma_p^b)$, and $i = 1,2,\ldots,p$. Denoting the Tikhonov filter function as

Table 5. Comparison of CPU time (s) of our algorithm with LSM for ten-point scatterers model in 3D case.

LSM	Fast method	Problem scale
125.7500	68.0000	$nx = ny = 21$; $nz = 6$; $dx = 40$; $dy = 20$
1629.1250	722.6406	$nx = ny = 41$; $nz = 6$; $dx = 40$; $dy = 20$

$$f^{\alpha}(\lambda) = (\alpha + \lambda)^{-1}, \quad \alpha > 0$$

and defining

$$\sigma_i = \frac{\sigma_i^a}{\sigma_i^b},$$

the approximate solution $\mathbf{m}^{\alpha}$ can be written as

$$\mathbf{m}^{\alpha} = \sum_{j=1}^{p} \sigma_j f^{\alpha}(\sigma_j^2)(\mathbf{d},\mathbf{u}_j)\mathbf{w}_j + \sum_{j=p+1}^{n} (\mathbf{d},\mathbf{u}_j)\mathbf{w}_j.$$

It is seen clearly from the above expression that the regularized solution actually consists of two parts. The second part $\Sigma_{i=p+1}^{n}(\mathbf{d},\mathbf{u}_j)\mathbf{w}_j$ is not affected by regularization. Therefore, to maintain the filtering properties of the CG iteration for regularized minimization problem (equation 2), it is required that components of the data corresponding to the very small singular values of $\mathbf{L}$ should remain in a subspace exclusively spanned by the left singular vectors. This signifies that the preconditioner should consist of two parts: one part should precondition only the well-posed component of $\mathbf{L}$, and the other should behave like an identity matrix applying to the ill-posed component of $\mathbf{L}$. Such a requirement is not being satisfied for conventional preconditioners that have been suggested so far in literature for well-posed problems. Therefore, we believe choosing a proper preconditioner for seismic ill-posed inverse problems in imaging remains an interesting topic.

CONCLUSIONS

The benefits of treating the seismic migration process as an inverse problem are improved resolution of images, compensation for irregularities in acquisition geometries, and treatment of variations in illumination due to velocity-model complexity. However, iterative least-squares migration is expensive, requiring at least the cost of two migrations per iteration. Therefore, any method that can reduce the number of iterations required to converge to an improved image and also help deal with the ill-conditioned nature of the LSM inverse problem is worth examining. The nonmonotone BB method detailed here has the useful properties: it converges quickly and allows straightforward implementation of regularization schemes. This method could be the basis for even more advanced nonmonotone methods applied to the least-squares migration/inversion problem. The examples we show, although only synthetic, show the power of this particular gradient-descent method and large-scale implementation of this technique is straightforward.

ACKNOWLEDGMENTS

We thank reviewers and editors for their concrete and helpful comments. We give special thanks to an anonymous associate editor for his suggestions and instructions. We thank I. Lecomte for discussions and instructions, J. X. Hu, Q. Y. Sun, and J. B. Chen for personal communications. This work is supported by the National Basic Research Program of China under grant 2005CB422104, National Natural Science Foundation of China under grants 10871191 and 40974075, and the Knowledge Innovation Program of the Chinese Academy of Sciences, grant KZCX2-YW-QN107.

REFERENCES

Barzilai, J., and J. Borwein, 1988, Two-point step size gradient methods: IMA Journal of Numerical Analysis, **8**, 141–148.
Bevc, D., F. Ortigosa, A. Guitton, and B. Kaelin, 2007, Next generation seismic imaging: High fidelity algorithms and high-end computing: Presented at the 2007 Joint Assembly, AGU.
Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell.
Gelius, L.-J., I. Lecomte, and H. Tabti, 2002, Analysis of the resolution function in seismic prestack depth imaging: Geophysical Prospecting, **50**, 505–515.
Golub, G. H., and C. F. Van Loan, 1989, Matrix computations: Johns Hopkins University Press.
Guitton, A., 2004, Amplitude and kinematic corrections of migrated images for nonunitary imaging operators: Geophysics, **69**, 1017–1024.
Hanke, M., and P. C. Hansen, 1993, Regularization methods for large-scale problems: Surveys on Mathematics for Industry, **3**, 253–315.
Hu, J. X., and G. T. Schuster, 1998, Migration deconvolution: Mathematical Methods in Geophysical Imaging Conference, SPIE, Proceedings, 118–124.
———, 2000, Prestack migration deconvolution: 70th Annual International Meeting, SEG, Expanded Abstracts, 984–987.
Hu, J. X., G. T. Schuster, and P. A. Valasek, 2001, Poststack migration deconvolution: Geophysics, **66**, 939–952.
Hu, J. X., and P. A. Valasek, 1999, Migration deconvolution of 3D seismic data: 69th Annual International Meeting, SEG, Expanded Abstracts, 1118–1121.
Lecomte, I., and L.-J. Gelius, 1998, Have a look at the resolution of prestack depth migration for any model, survey and wavefields: 68th Annual International Meeting, SEG, 1112–1115.
Nemeth, T., C. J. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.
Sacchi, M. D., J. Wang, and H. Kuehl, 2006, Regularized migration/inversion: New generation of imaging algoritihms: CSEG Recorder, **31**, 54–59.
Schuster, G. T., 1997a, Acquisition footprint removal by least-squares migration: Utah Tomography and Modeling/Migration (UTAM) Research Report, University of Utah.
———, 1997b, Green's functions for migration: 67th Annual International Meeting, SEG, Expanded Abstracts, 1754–1758.
Sjøberg, T., L.-J. Gelius, and I. Lecomte, 2003, 2D deconvolution of seismic image blur: 73rd Annual International Meeting, SEG, Expanded Abstracts, 1055–1058.
Tikhonov, A. N., and V. Y. Arsenin, 1977, Solutions of ill-posed problems: John Wiley and Sons.
Treitel, S., and L. R. Lines, 2001, Past, present, and future of geophysical inversion — A new millennium analysis: Geophysics, **66**, 21–24.
Valenciano, A., B. Biondi, and A. Guitton, 2006, Target oriented wave-equation inversion: Geophysics, **71**, no. 4, A35–A38.
VanDecar, J. C., and R. Snieder, 1994, Obtaining smooth solutions to large linear inverse problems: Geophysics, **59**, 818–829.
Wang, Y. F., 2007, Computational methods for inverse problems and their applications: Higher Education Press.
Wang, Y. F., and S. Q. Ma, 2007, Projected Barzilai-Borwein methods for large scale nonnegative image restorations: Inverse Problems in Science and Engineering, **15**, 559–583.
Wang, Y. F., C. C. Yang, and Q. L. Duan, 2009, On iterative regularization methods for migration deconvolution and inversion in seismic imaging: Chinese Journal of Geophysics, **52**, 704–715.
Yu, J. H., J. X. Hu, G. T. Schuster, and R. Estill, 2006, Prestack migration deconvolution: Geophysics, **71**, no. 2, S53–S62.
YuanY. X., 2008, Step-sizes for the gradient method, *in* K.-S. Lau, Z.-P. Xin, and S.-T. Yau, eds., Proceedings of the third international congress of Chinese mathematicians: American Mathematical Society and International Press, 785–796.