J. Inv. Ill-Posed Problems **18** (2011), 823 – 838
DOI 10.1515/JIIP.2011.005

Seismic impedance inversion using l_1-norm regularization and gradient descent methods

Yanfei Wang

Abstract. We consider numerical solution methods for seismic impedance inversion problems in this paper. The inversion process is ill-posed. To tackle the ill-posedness of the problem and take the sparsity of the reflectivity function into consideration, an l_1 norm regularization model is established. In computation, a nonmonotone gradient descent method based on Rayleigh quotient for solving the minimization model is developed. Theoretical simulations and field data applications are performed to verify the feasibility of our methods.

Keywords. Impedance inversion, optimization, regularization.

2010 Mathematics Subject Classification. 65J22, 86-08, 86A22.

1 Introduction

The reflection seismic exploration becomes an important method of exploration geophysics. This is because different subsurface layers have different impedance; reflective waves will appear as the seismic waves propagating to the layer interface. The purpose of seismic inversion is to speculate the spatial distribution of underground strata structure and physical parameter by using seismic wave propagation law. A key step for reflectivity inversion is the deconvolution [28]. By deconvolution, we mean that we attempt to recover the reflectivity function from the seismic records. So far, the inversion and deconvolution technique has experienced the development process from direct inversion to model-based inversion, from post-stack inversion to pre-stack inversion, and from linear inversion to nonlinear inversion [1, 6, 9–13, 16]. In recent ten years, nonlinear inversion methods have been utilized effectively. Practice shows that nonlinear inversion is much closer to real situation than linear inversion. Among these methods, nonlinear optimization methods are based on gradient computation, include steepest descent method, Newton's method and conjugate gradient method; statistical methods include neu-

This work is supported by National Natural Science Foundation of China under grant numbers 10871191, 40974075 and Knowledge Innovation Programs of Chinese Academy of Sciences KZCX2-YW-QN107.

ral network method, genetic algorithm (GA), simulated annealing algorithm (SA) and random search algorithm [17, 18, 20, 21, 30]. To improve the solution precision and accelerate convergence speed, many hybrid optimization algorithms were developed [21, 27]. Roughly speaking, there are three classes of commonly used methods [7]. The first class of methods is based on local differential characteristics of the objective function, represented by conjugate gradient method and variable metric method. Since the impedance inversion is quasi-linear, this kind of iterative methods is feasible, but the results strongly depend on the choice of the initial model. The second class of methods is randomized algorithm based on pure random searching, represented by SA. It can obtain global optimal solution, but has difficulty in determining temperature parameter and needs large amount of calculation. The last kind of methods is intelligent algorithms which incorporate randomness and inheritance, represented by GA. One major disadvantage of GA is that it cannot invert too many parameters, so it is usually used to invert some characteristic parameters such as interval velocity and reflection depth.

Though a lot of research works have been done, they still cannot completely satisfy the practical requirements. The results of different methods may lead to differences and may result in incorrect geological explanation. The reasons may come from the low quality of seismic data, inaccurate wavelet extraction, and errors between normal incidence assumption and real situation. In addition, two main influences should be accounted: the first is the band-limited property of seismic data, hence direct inversion can only obtain the mid-frequency component of impedance model and lack the low-frequency and high-frequency component; second, in the Hardmard sense, deconvolution and inversion are ill-posed problems, so the inversion results are highly sensitive to noise. Proper regularization techniques are necessary.

Nowadays, the most effective way to solve band-limited problem is the constrained inversion which can be generalized as the Tikhonov regularization, and the most effective way to solve ill-posed problem is the regularization methods assisted with proper optimization techniques [21]. We study regularization and optimization methods in this paper. Considering the sparsity properties of the reflectivity functions, we develop an equality constrained l_1 norm regularization model. In solving the model, a nonmonotone gradient descent method is developed. Numerical experiments based on synthetic model and filed data are performed.

2 Problem formulation

The seismic impedance usually refers to the characteristic impedance, defined by $I(t) = \rho(t)v(t)$, where $\rho(t)$ refers to the medium density of layers (e.g., determined by analyzing well logging data) and $v(t)$ is the wave velocity. A practical definition of the characteristic impedance in exploration is using the reflectivity function $r(t)$. Since

$$r(t) = \frac{I(t + \Delta t) - I(t)}{I(t + \Delta t) + I(t)},$$

therefore the reflectivity coefficient of a layer can be expressed as

$$r_j = \frac{I_{j+1} - I_j}{I_{j+1} + I_j}.$$

Hence we obtain

$$I_{j+1} = I_j \left(\frac{1 + r_j}{1 - r_j} \right) = \cdots = I_1 \prod_{k=1}^{j} \left(\frac{1 + r_k}{1 - r_k} \right).$$

Note that $|r(t)| < 1$, applying logarithm to the above expression and using Taylor extension, we have that

$$\ln \left(\frac{1 + x}{1 - x} \right) \approx 2.$$

Therefore

$$\ln \left(\frac{I_{j+1}}{I_1} \right) = \sum_{k=1}^{j} \ln \left(\frac{1 + r_k}{1 - r_k} \right) \approx 2 \sum_{k=1}^{j} r_k.$$

The above formula yields

$$I_{j+1} = I_1 \exp \left(2 \sum_{k=1}^{j} r_k \right). \tag{2.1}$$

Define $\mu s_k = 2r_k$, a more practical formula of the characteristic impedance is given by

$$I_{j+1} = I_1 \exp \left(\mu \sum_{k=1}^{j} s_k \right). \tag{2.2}$$

Therefore, to find the impedance, a key problem is to solve for an accurate reflectivity function r.

A convenient expression for the characteristic impedance inversion is the convolution model

$$Wr = d, \tag{2.3}$$

where $r = [r_0, r_1, \ldots, r_{N-1}]^T$ is the reflectivity coefficient vector, $d = [d_0, d_1, \ldots, d_{N-1}]^T$ is the recorded seismic data and W is the wavelet matrix with length $L+1$ expressed by the wavelet function w as

$$W_{(N+L)\times N} = \begin{bmatrix} w_0 & 0 & 0 & 0 & \cdots & 0 \\ w_1 & w_0 & 0 & 0 & \cdots & 0 \\ 0 & w_1 & w_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ w_L & w_{L-1} & \cdots & w_1 & w_0 & 0 \\ 0 & w_L & w_{L-1} & \cdots & w_1 & w_0 \\ 0 & 0 & w_L & w_{L-1} & \cdots & w_1 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & w_L & w_{L-1} \\ 0 & 0 & 0 & \cdots & 0 & w_L \end{bmatrix}.$$

Practically, the data d may also contains different kinds of noises, hence instead of the exact data d, we have a perturbed version d_δ: $d_\delta = d + \delta \cdot n$, where n represents noise.

A naive approach for finding the reflectivity model r is by

$$r = (W^* W)^{-1} W^* d_\delta. \tag{2.4}$$

It is evident that if W^* is an approximation to the forward operator W, then the reflectivity model can be obtained by $r = W^* d_\delta$. This process in seismic exploration is called the migration [5, 26]. However, for seismic imaging problems, due to limited bandwidth and limited acquisition spaces, the seismic images obtained are blurred, and then direct inversion may cause distortion on the low-frequency component and the high-frequency component as well.

3 Regularization

As noted above, the numerical inversion for finding r is an ill-posed process. Therefore incorporating some kind of regularization is necessary. The general form of the regularization technique is solving a constrained optimization problem

$$\min J(r), \tag{3.1}$$

$$\text{s.t.} \quad Wr = d_\delta, \tag{3.2}$$

$$\Delta_1 \le c(r) \le \Delta_2, \tag{3.3}$$

where $J(r)$ denotes an object function, which is a function of r, $c(r)$ is the constraint to the solution r, Δ_1 and Δ_2 are two constants which specify the bounds of $c(r)$. Usually, $J(r)$ is chosen as the norm of r with different scale. If the parameter r comes from a smooth function, then $J(r)$ can be chosen as a smooth function, otherwise, $J(r)$ can be nonsmooth.

A conventional regularization method is the Tikhonov's regularization [8, 14, 15], which is in the form

$$\min \|r\|_{l_2}^2, \tag{3.4}$$

$$\text{s.t.} \quad Wr = d_\delta, \tag{3.5}$$

or the equivalent form

$$\min \frac{1}{2}\|Wr - d_\delta\|_{l_2}^2 + \frac{\alpha}{2}\|r\|_{l_2}^2, \tag{3.6}$$

where $\alpha > 0$ is the regularization parameter. But this formulation is not suitable for sparse seismic signals.

4 Minimal solution in l_1 space

It is clear that the reflectivity function may possess spiky and may be sparse. In this case the above l_2 minimization model is not proper description of the problem. We consider a special model of (3.1)–(3.3), i.e., an l_1 minimization model

$$\min \|r\|_{l_1}, \tag{4.1}$$

$$\text{s.t.} \quad Wr = d_\delta. \tag{4.2}$$

We remark that the l_1 norm minimization is not a new thing. The use of absolute value error for data fitting had been studied in [4]. However, the l_1 norm has a singular problem when the values of residual vanish. Even if the values of residuals are not zero, the numerical inversion process goes to failure at very small residual. Finding suitable methods for solving the nonlinear optimization problem is desirable.

It is clear that equation (4.1)–(4.2) is equivalent to

$$\min \ e^T r, \tag{4.3}$$

$$\text{s.t.} \quad Wr = d_\delta, \quad c_i^T r + \varepsilon_i \le 0, \tag{4.4}$$

where e is a vector with all components 1, c_i is some vector and $\varepsilon_i \in \mathbb{R}$, $i = 1, 2, \ldots$. Then the optimal solution of problem (4.3)–(4.4) is also a solution of problem (4.1)–(4.2). Solving (4.3)–(4.4) can use the linear programming method [24], however, this method is not applicable for large scale problems in seismology. Instead, we consider an inequality constrained minimization problem

$$\min \|r\|_{l_1}, \tag{4.5}$$

$$\text{s.t.} \quad \|Wr - d_\delta\|_{l_2} \le \delta \tag{4.6}$$

and solve the unconstrained minimization problem

$$\min f(r) := \|Wr - d_\delta\|_{l_2}^2 + \alpha\|r\|_{l_1} \tag{4.7}$$

will yield the solution.

It is evident that the above function f is nondifferentiable at $r = 0$. To make it easy to be calculated by computer, we approximate $\|r\|_{l_1}$ by $\sum_{i=1}^{l} \sqrt{(r_i, r_i) + \epsilon}$ ($\epsilon > 0$) and l is the length of the vector r. For simplification of notations, we let $\gamma(r) = (\frac{r_1}{\sqrt{(r_1)^T r_1 + \epsilon}}, \frac{r_i}{\sqrt{(r_i)^T r_i + \epsilon}}, \ldots, \frac{r_n}{\sqrt{(r_n)^T r_n + \epsilon}})^T$ and

$$\chi_p(r) = \begin{pmatrix} \frac{\epsilon}{((r_1)^T r_1 + \epsilon)^{p/2}} & 0 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \vdots & \vdots \\ 0 & \cdots & \frac{\epsilon}{((r_i)^T r_i + \epsilon)^{p/2}} & \vdots & 0 \\ \vdots & 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \frac{\epsilon}{((r_n)^T r_n + \epsilon)^{p/2}} \end{pmatrix}.$$

Straightforward calculation yields the gradient of f

$$g(r) \approx W^T(Wr - d_\delta) + \alpha\gamma(r)$$

and the Hessian of f

$$H(r) \approx W^T W + \alpha\chi_3(r).$$

With the gradient information, the gradient-based iterative methods can be applied.

4.1 Gradient descent methods

The gradient method is one of the simplest methods for solving the nonlinear minimization problem. Given an iterate point r_k, the gradient method chooses the next iterate point r_{k+1} in the following form:

$$r_{k+1} = r_k - \tau_k g_k, \tag{4.8}$$

where $g_k = g(r_k)$ is the gradient at r_k and $\tau_k > 0$ is a step-length. The gradient method has the advantages of being easy to program and suitable for large scale problems. Different step-lengths τ_k give different gradient algorithms. If $\tau_k = \tau^*$ where τ^* satisfies

$$f(r_k - \tau_k^* g_k) = \min_{\tau>0} f(r_k - \tau g_k), \tag{4.9}$$

the gradient method is the steepest descent method, which is also called the Cauchy's method. However, the steepest descent method, though it uses the "best" direction and the "best" step-length, turns out to be a very bad method as it normally converges very slowly, particularly for ill-conditioned problems.

When we apply the gradient method to large scale problems, the most important issue is which step-length will give a fast convergence rate. Therefore it is vitally important to find what choices of τ_k require less number of iterations to reduce the gradient norm to a given tolerance.

Recently, nonmonotone gradient methods are much popular, see, e.g., [25, 29]. We recall a well-known such kind of method, developed by Barzilai and Borwein [2], which lies in the two choices for the step-length τ_k:

$$\tau_k^{\text{BB1}} = \frac{(s_{k-1}, s_{k-1})}{(s_{k-1}, y_{k-1})}, \quad \tau_k^{\text{BB2}} = \frac{(s_{k-1}, y_{k-1})}{(y_{k-1}, y_{k-1})}, \tag{4.10}$$

where $y_{k-1} = g_k - g_{k-1}$, $s_{k-1} = r_k - r_{k-1}$. This method initially designs for well-posed convex quadratic programming problems. However, it reveals that the method is also applicable for ill-posed problems and non quadratic programming problems provided that the deviation of the non quadratic model is not far away from the quadratic model [22, 25].

Let us consider the quasi-Newton equation of the minimization problem (4.7). It is easy to deduce that the quasi-Newton equation satisfies

$$H_{k+1} s_k = y_k, \tag{4.11}$$

where $H_k = H(r_k) = W^T W + \alpha \chi_3(r_k)$. Noting that $s_k = -\tau_k g_k$, we have that

$$\tau_k^{\text{BB1}'} = \frac{(g_{k-1}, g_{k-1})}{(g_{k-1}, H_k g_{k-1})}, \quad \tau_k^{\text{BB2}'} = \frac{(g_{k-1}, H_k g_{k-1})}{(g_{k-1}, H_k^T H_k g_{k-1})}. \tag{4.12}$$

This indicates that the two step-lengths inherit different information from the seismic wavelet. In literature, people usually favor τ_k^{BB1} or $\tau_k^{\text{BB1}'}$. It is readily to see that $\tau_k^{\text{BB1}'} > \tau_k^{\text{BB2}'}$, i.e., the BB1 step-length is usually larger than BB2 step-length. However, there is no reason to disregard τ_k^{BB2} or $\tau_k^{\text{BB2}'}$ since the method using BB2 step-length would be efficient if we want to obtain a very accurate solution of a very large-scale and ill-conditioned problem [29].

Further, we observe that the inverse of the scalar τ_k is the Rayleigh quotient of H_k or $H_k^T H_k$ at the vector g_{k-1}. This indicates that more choices can be obtained by combining Rayleigh quotients. To make a trade-off, we develop a new choice of the step-length by

$$\tau_k^{\text{Rayleigh}} = \beta_1 \frac{(g_{k-1}, g_{k-1})}{(g_{k-1}, H_k g_{k-1})} + \beta_2 \frac{(g_{k-1}, H_k g_{k-1})}{(g_{k-1}, H_k^T H_k g_{k-1})}, \tag{4.13}$$

where β_1 and β_2 are two positive parameters assigned by users.

4.2 Chaotic nature

In recent communications, van den Doel and Ascher notice the chaotic nature of the nonmonotone gradient methods when they study the Poisson problem

$$-\Delta u(\xi, t) = q(s, t), \quad \xi, t \in (0, 1)$$

solved by the nonmonotone gradient method [19]. In which, $q(\xi, t)$ is known and subject to homogeneous Dirichlet boundary conditions. The chaotic nature lies in the sensitivity of the total number of iterations required to achieve a fixed accuracy to small changes in the initial vector ξ_0. We give an example of 1d case:

$$Ax(s) = \int_\Omega a(s-t)x(t)dt = z(s),$$

where $a(s) = 1/(\sqrt{2\pi}\sigma)\exp(-1/2(s/\sigma)^2)$, $\Omega = [-\pi/2, \pi/2]$ and the true solution is $2\exp(-6(t-0.8)^2) + \exp(-2(t+0.5)^2)$.

The initial guess value is chosen as $x_0 = a_1 e_1 + a_2 e_2$, where $a_i \in [-10^{-5}, 10^{-5}]$ $(i = 1, 2)$, e_1 and e_2 are two random orthogonal vectors. The chaotic nature can be vividly seen from Figure 1.

The chaotic nature indicates that the pure nonmonotone iteration may be not sufficient to guarantee fast convergence, some safeguard techniques maybe useful.

4.3 Safeguard

From equations (4.1)–(4.2) we learn that their solution is also the solution of the problem

$$\min \|Wr - d_\delta\|_{l_2}, \tag{4.14}$$

$$\text{s.t.} \quad \|r\|_{l_1} \le \epsilon, \ \epsilon > 0. \tag{4.15}$$

Therefore, r^* solves (4.14) within the feasible set $S = \{r : \|r\|_{l_1} \le \epsilon\}$. To maintain this property, we apply a projection technique. Note that the set S is bounded

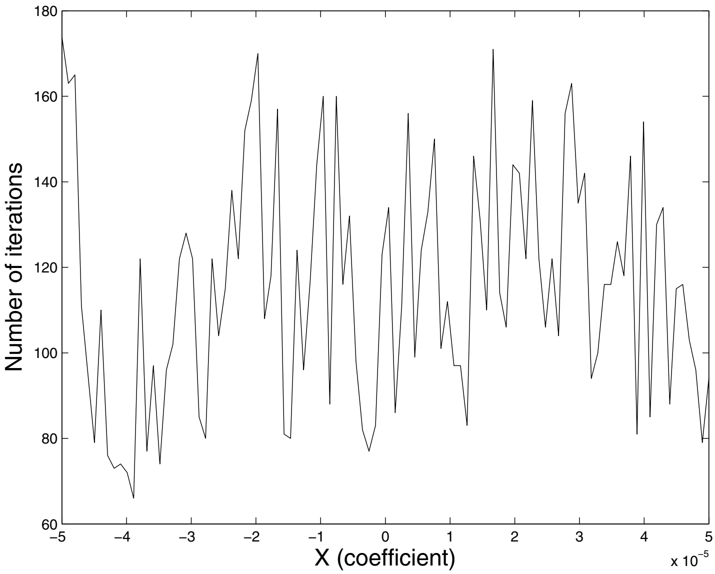

Figure 1. Chaotic nature of the nonmonotone gradient method.

below and convex, therefore, there exists a projection operator $P_S : \mathbb{R}^N \to S$ onto S such that

$$P_S(x) = \text{argmin}_z \|x - z\|, \quad z \in S.$$

The projection is easy to be calculated [3]. Assume that the current iterate r_k is feasible, then the next iteration point can be obtained by

$$r_{k+1} = P_S(r_k - \tau_k g_k).$$

4.4 Choosing the regularization parameter

There are many ways to choosing the regularization parameter. Roughly speaking, the techniques can be classified as either *a priori* way or a posteriori way. An *a priori* choice of the regularization parameter α requires that $\alpha > 0$ and is fixed. An a posteriori choice of the regularization parameter usually requires solving nonlinear equations about α which involves the computations of derivatives of the solution r^α. For our gradient descent method, we apply a simple posteriori technique, i.e., a geometric choice manner,

$$\alpha_k = \alpha_0 \cdot \xi^{k-1}, \ \xi \in (0, 1), \tag{4.16}$$

where $\alpha_0 > 0$ is a preassigned value of regularization parameter.

4.5 Remarks

The minimization model based on combination of l_2 and l_1 norm possesses several advantages. A simple fact is that the l_1 norm is robust to eliminate outliers and large amplitude anomalies when the necessity of the signal to have minimum energy is unnecessary. A more general form is the l_p-l_q model, which can be in the form

$$\min f(r) := \|Wr - d_\delta\|_{l_p}^p + \alpha\|L(r - r^0)\|_{l_q}^q, \tag{4.17}$$

where $p,\ q > 0$ which are specified by users, $\alpha > 0$ is the regularization parameter, L is the scale operator and r^0 is an *a priori* estimated solution of the original model. This model relaxes the convexity requirements on the usual models l_2-l_2 and l_2-l_1. Details about implementation of the minimization model are given in [23].

5 Numerical experiments

5.1 Synthetic simulations

The seismogram is generated by a velocity model with 6 layers, see Figure 2. The thickness of each layer varies. The velocity parameter in each layer is shown in Table 5.1. To generate the seismogram, a theoretical Ricker wavelet

$$w(t) = (1 - 2\pi^2 f_m^2 t^2)\exp(-(\pi f_m t)^2)$$

is used to perform a convolution, where f_m represents the central frequency. The seismic records with additive Gaussian random noise is shown in Figure 3. Using our algorithm, the retrieved reflectivity function is shown in Figures 4 and 5. It is evident that our algorithm can reduce the noise and retrieve the reflectivity stably. With the reflectivity function we could rebuild the velocity. The recovered velocity is shown in Figure 6. Comparison of the true velocity model (Figure 2) with the recovered velocity (Figure 6) reveals that our algorithm is robust and could be used for seismic impedance inversion.

layers	1	2	3	4	5	6
V ($\mathrm{km \cdot s^{-1}}$)	2.5	3.0	2.7	3.2	3.8	4.0

Table 1. Parameters of the velocity model.

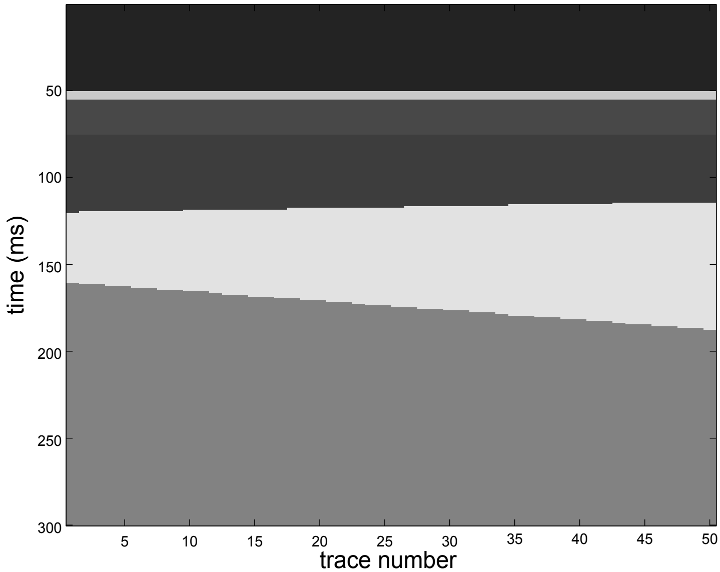

Figure 2. Velocity model.

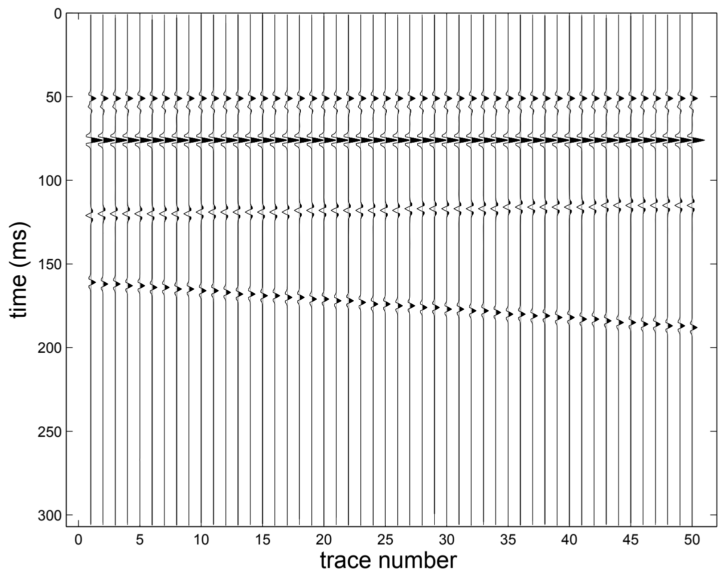

Figure 3. Seismic data.

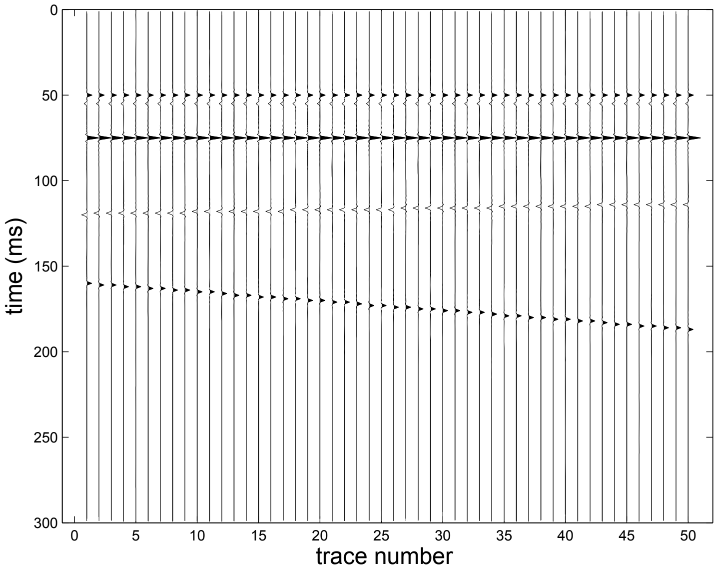

Figure 4. Reflectivity function.

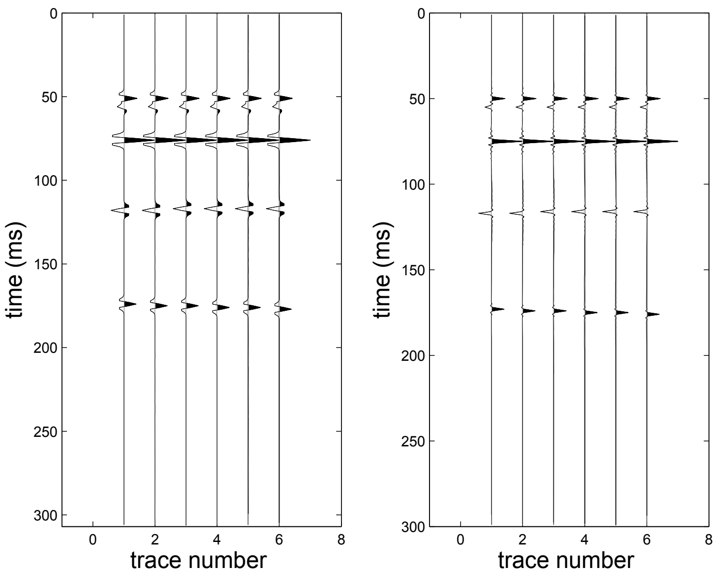

Figure 5. Comparison of the reflectivity sections with the seismogram sections.

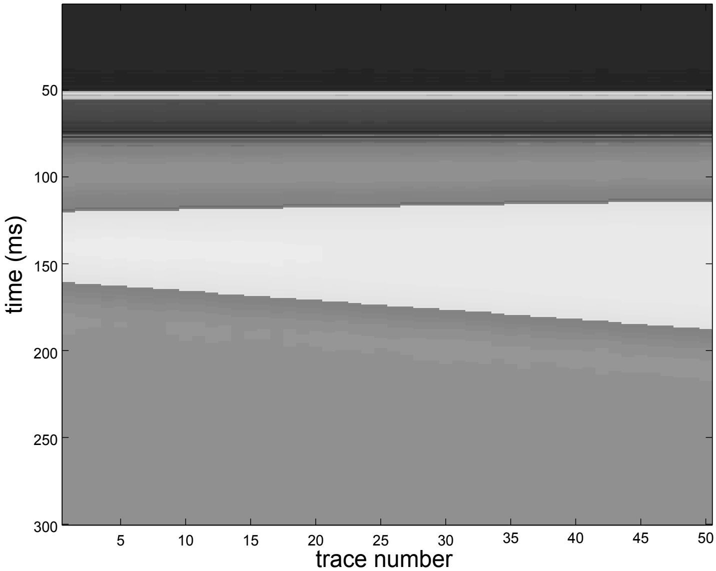

Figure 6. Recovered velocity.

5.2 Field applications

We apply our method to the field data. The data is taken from the Tarimu Oilfield. The time sampling is 2 milliseconds and the spacing of traces is 20 meters. Figure 7 plots the original seismic section. The wavelet function was extracted using the autocorrelation method. Then we apply our algorithm to the seismic data. The retrieved reflectivity function is shown in Figure 8. It is illustrated from the inversion results that the recovered section well represents reflection of layers and much noise is reduced using our algorithms.

6 Conclusion

We develop an l_1-norm constrained regularization model for seismic impedance inversion problems. Nonmonotone gradient descent methods are used to solve the regularization problem. Numerical results indicate the feasibility of our algorithm to practical applications.

We argue that to tackle the ill-posedness of the inversion problem, choosing a proper regularization parameter α is needed. In this paper, we use a geometric technique. Clearly this kind of choice is not optimal. It is desirable to find a better regularization parameter α by a posteriori techniques.

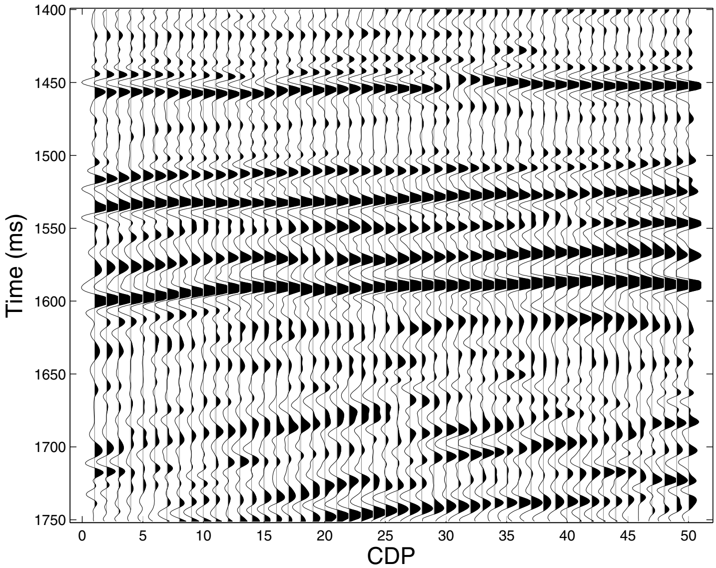

Figure 7. Seismic section.

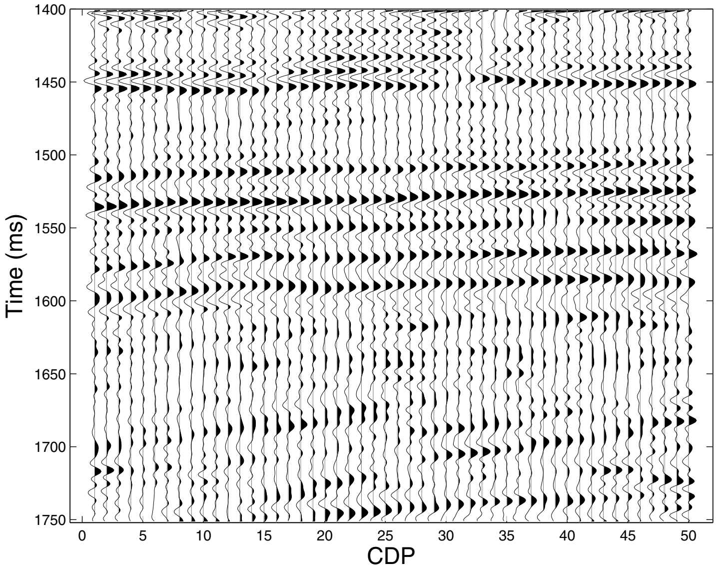

Figure 8. Recovered reflectivity function.

Bibliography

[1] K. Aki and P. G. Richards, *Quantitative Seismology: Theory and Methods*, W. H. Freeman and Company, San Francisco, 1980.

[2] J. Barzilai and J. Borwein, Two-point step size gradient methods, *IMA Journal of Numerical Analysis*, **8** (1988), 141–148.

[3] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*, Institute of Physics Publishing, Bristol and Philadelphia, 1998.

[4] J. Claerbout and F. Muir, Robust modeling with erratic data, *Geophysics*, **38** (1973), 826–844.

[5] Y. Cui , Y. F. Wang and C. C. Yang, Regularizing method with *a priori* knowledge for seismic impedance inversion, *Chinese Journal of Geophysics*, **52** (2009), 2135–2141.

[6] V. Dimri, *Deconvolution and Inverse Theory-Applications to Geophysical Problems*, Elsevier, Amsterdam, 1992.

[7] D. E. Goldberg, *Genetic Algorithms in Search. Optimization & Machine Learning*, Addison-Wesley Publishing Company, INC, 1989.

[8] S. I. Kabanikhin, Definitions and examples of inverse and ill-posed problems, *J. Inv. Ill-posed Problems*, **16** (2008), 317–357.

[9] K. Levenberg, A method for the solution of certain nonlinear problems in least squares, *Qart. Appl. Math.*, **2** (1944), 164–166.

[10] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear inequalities, *SIAM J. Appl. Math.*, **11** (1963), 431–441.

[11] G. Nolet, *Seismic Tomography*, Reidel Publishing, Boston, 1987.

[12] G. Nolet and R. Snieder, Solving large linear inverse problems by projection, *Geophys. J. Int.*, **103** (1990), 565–568.

[13] A. Tarantola, *Inverse Problems Theory: Methods for Data Fitting and Model Parameter Estimation*, Elsevier, Amsterdam, 1987.

[14] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-posed Problems*, John Wiley and Sons, New York, 1977.

[15] A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, Kluwer, Dordrecht, 1995.

[16] J. Trampert and J. J. Leveque, Simultaneous iterative reconstruction technique: physical interpretation based on the generalized least squares solution, *Journal of Geophysical Research*, **95** (1990), 12553–12559.

[17] T. J. Ulrych, M. D. Sacchi and M. Grau, Signal and noise separation: art and science, *Geophysics*, **64** (1999), 1648–1656.

[18] T. J. Ulrych, M. D. Sacchi and A. Woodbury, A Bayesian tour to inversion, *Geophysics*, **66** (2000), 55–69.

AUTHOR’S COPY | AUTORENEXEMPLAR

[19] K. van den Doel and and U. Ascher, The chaotic nature of faster gradient descent methods, *Preprint*, 2010.

[20] J. Y. Wang, *Inversion Theory for Geophysics*, Higher Education Press, Beijing, 2002, (in Chinese).

[21] Y. F. Wang, *Computational methods for inverse problems and their applications*, Higher Education Press, Beijing, 2007, (in Chinese).

[22] Y. F. Wang, An efficient gradient method for maximum entropy regularizing retrieval of atmospheric aerosol particle size distribution function, *Journal of Aerosol Science*, **39** (2008), 305–322.

[23] Y. F. Wang, J. J. Cao, Y. X. Yuan, C. C. Yang and N. H. Xiu, Regularizing active set method for nonnegatively constrained ill-posed multichannel image restoration problem, *Applied Optics*, **48** (2009), 1389–1401.

[24] Y. F. Wang, S. F. Fan and X. Feng, Retrieval of the aerosol particle size distribution function by incorporating *a priori* information, *Journal of Aerosol Science*, **38** (2007), 885–901.

[25] Y. F. Wang and S. Q. Ma, Projected Barzilai–Borwein methods for large scale nonnegative image restorations, *Inverse Problems in Science and Engineering*, **15** (2007), 559–583.

[26] Y. F. Wang and C. C. Yang, Accelerating migration deconvolution using a nonmonotone gradient method, *Geophysics*, **75** (2010), S. 131–S. 137.

[27] W. C. Yang, On inversion methods for nonlinear geophysics, *Progress in Geophysics*, **17** (2002), 255–261, (in Chinese).

[28] O. Yilmaz, *Seismic Data Processing, Investigations in Geophysics No. 2, Society of Exploration Geophysicists*, Tulsa, Okla, 1987.

[29] Y. Yuan, Gradient methods for large scale convex quadratic functions, in Y. F. Wang, A. G. Yagola and C. C. Yang editors, *Optimization and Regularization for Computational Inverse Problems and Applications*, Springer, Berlin, 2010, pp. 141–155.

[30] H. B. Zhang, Z. P. Shang and C. C. Yang, A non-linear regularized constrained impedance inversion, *Geophysical Prospecting*, **55** (2007), 819–833.

Received August 20, 2010; revised December 2, 2010.

Author information

Yanfei Wang, Key Laboratory of Petroleum Resources Research,
Institute of Geology and Geophysics, Chinese Academy of Sciences,
P. O. Box 9825, Beijing, 100029, P. R. China.
E-mail: yfwang@mail.iggcas.ac.cn