Greedy Least-Squares and its Application in Radon Transforms

Juefu Wang
Divestco Inc., Calgary, Alberta, Canada
Juefu.wang@divestco.com
and
Mark Ng
Divestco Inc., Calgary, Alberta, Canada

Summary

We propose a greedy inversion method for sparse linear problems. The kernel of the method is based on a conventional iterative algorithm, conjugate gradients (CG), but it is utilized adaptively in amplitude-prioritized local model spaces, thereby giving rise to a greedy algorithm. The adaptive inversion introduces a coherence-oriented mechanism to reduce the crosstalk between seismic events and hence increases both the image resolution and the convergence rate. We adopt the idea in a time-space domain high-resolution Radon transform for multiple attenuation and a local Radon transform for data interpolation. Synthetic examples show that the method can yield high quality solutions at much lower cost than existing standard methods.

Introduction

In recent years, the greedy approach has been adopted by the geophysical community for data decomposition and noise reduction. For example, Li et al. (1998) used a matching pursuit method to decrease the cost of Kirchhoff migration, and Liu and Sacchi (2002) introduced a binary image method to speed up the time domain hyperbolic Radon transform. For the high-resolution Radon transform, it has been shown that a Gauss-Seidel implementation of the greedy approach using prioritized moveout parameters can significantly reduce computational cost (Ng and Perz, 2004). In this research we solve the greedy problem in a more global sense by solving small regions of moveout parameters using the CG (Hestenes, 1952) algorithm. We first examine the possibility of a 2D greedy implementation in a high-resolution Radon transform for multiple attenuation. Then we show another application of the technique for a large-scale problem, namely the local Radon transform (Sacchi et al., 2004) for data interpolation.

Method

According to www.wikipedia.org, there are five key components that make a greedy method:

1. A candidate model set, from which a solution is created.
2. A model selection function to choose the best candidate.
3. A feasibility function to determine if the model candidate can contribute to the solution.
4. An objective (misfit) function to assign a value to a solution, or a partial solution.
5. A stopping criterion to control convergence.

In the context of the Radon transform, we can immediately identify the five key components for a greedy implementation. First we can recognize the amplitude of the Radon panel in a 2D model space (tao and p) as the candidate set. Second the model selection function is a threshold operator which operates on the 2D Radon panel, and which selects a subspace comprising only those model elements whose amplitude exceeds the threshold value. Recalling that the task of the Radon inverse problem is to find an optimal amplitude distribution within the Radon panel such that we can fit the input data, we can use the inverse Radon transform as the feasibility function. The objective function is the L2 norm of the misfit, which is the difference between the observed data and the predicted data using the inverse Radon transform. Finally, the stopping criterion is the ratio of the residual norm to the observation norm. By controlling the misfit to observation ratio, we can control the degree of data fitting. In our implementation, we use conjugate gradients (CG) methods (Hestenes, 1952) to minimize the aftermentioned objective function.

With these ideas in mind, we propose a greedy least-squares method for high-resolution Radon transform as below in pseudo code:

1. Initialize the data residual d_{resi} with the input data and the model m with zeros.
2. Repeat the following loop until the stopping criterion is fulfilled:

   2.1 $\tilde{m}^i = L' d_{resi}$, (1)

   where L' denotes the forward Radon operator and $\tilde{m}^i$ is the adjoint solution (i.e. the Radon panel generated by the Radon operator at the *i*-th iteration).

   2.2 Choose a model subspace S_m^i based on amplitude threshold of the adjoint solution $\tilde{m}^i$.

   2.3 Minimize the following cost function in the chosen subspace using CG algorithm:

   $J(m^i) = \| d_{resi} - Lm^i \|^2$, (2)

   where m^i is solution at the *i*-th iteration within the subspace, L is the inverse Radon operator.

   2.4 Update the solution by $m = m + m^i$.

   2.5 Update the data residual by $d_{resi} = d_{resi} - Lm^i$.

The method is greedy since the least-squares problem posed by equation 2 within each loop is applied in a small subspace (regions of moveout parameters) based on prioritization of the energy level. There are several benefits associated with this subspace strategy. First the small number of model parameters significantly reduces the computational cost for the CG algorithm. Second the convergence rate is faster due to the smaller model size within the subspace. Third, the result will naturally exhibit high resolution in the Radon domain. The resolution is controlled by a careful choice of the threshold value—note that the bigger the threshold value, the higher the resolution. The algorithm is slightly more expensive than Liu and Sacchi's method (2002) since we have two loops. The inner loop solves a least-squares problem (equation 2) within the subspace, and the outer loop (step 2 in general) updates the region of moveout parameters. However the minor extra cost affords significant improvement in resolution. Note that existing sparse inversion methods often tend to suppress weak model elements; by contrast, our method overcomes the problem by iteratively working on the data residual, and it usually recovers weak model elements at later iterations.

Examples

We first examine the feasibility of the proposed inversion method using the classic Radon transform problem. Figure 1 shows that the standard least-squares method (similar to Hampson, 1986) is not good enough to separate the multiples with small moveout from the primary. On the other hand, both our method and the iterative least-squares Cauchy norm method (Sacchi and Ulrych, 1995) provide good separation of primary and multiples. Note that the Cauchy norm method is applied in time and space domain, which usually generates better result than the FX domain implementation, but with a higher cost. In this example, the proposed method is about 40 times faster than the time-space domain Cauchy norm method due to the simplicity of the model. For real data, usually the speedup is about 10-20 times.

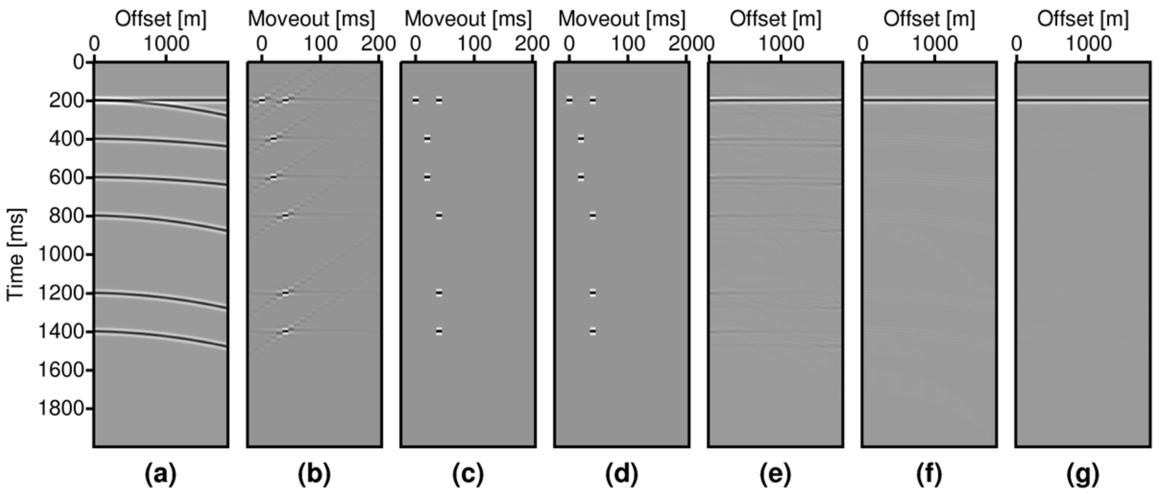
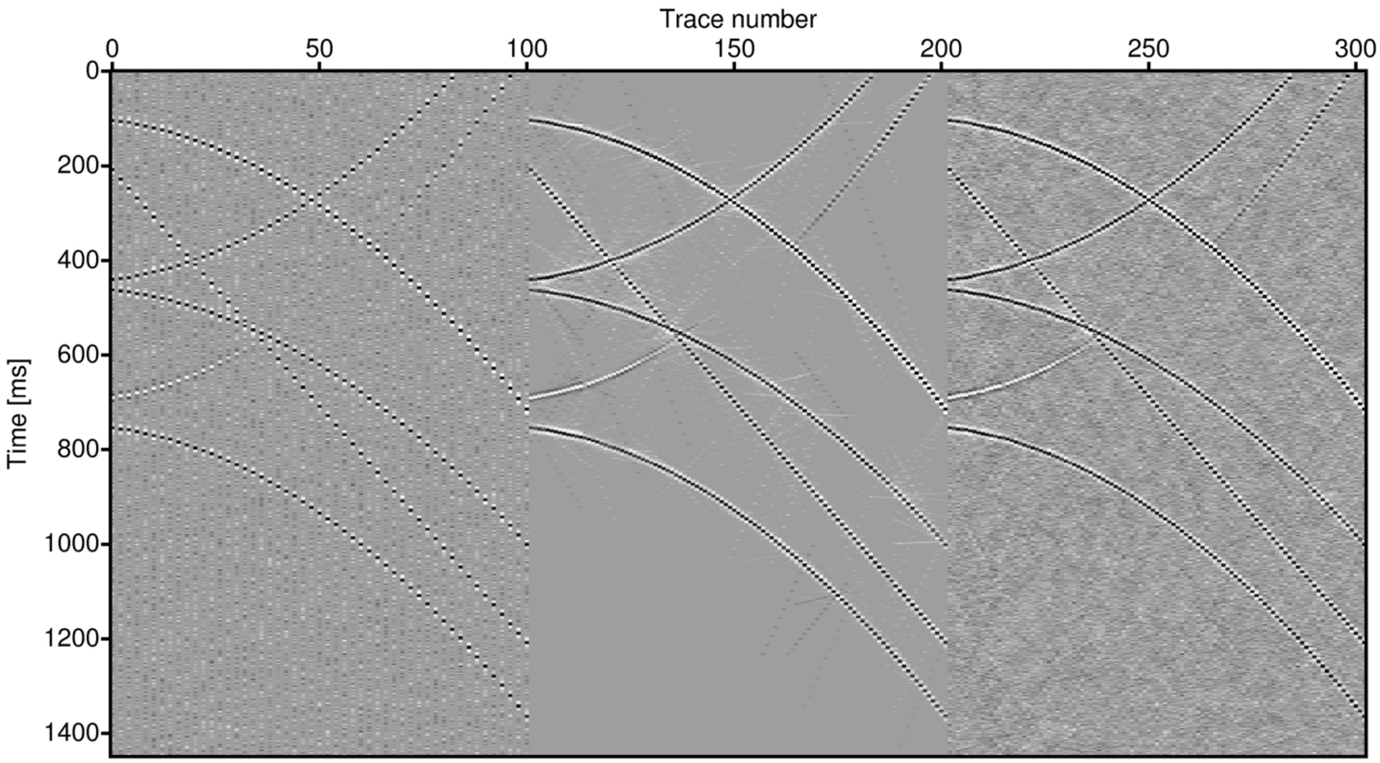


Figure 1: Comparison of three Radon transforms for multiple attenuation. (a) Input data. (b) Radon panel obtained with standard least-squares method. (c) Radon panel obtained with Cauchy norm sparse Radon transform (Sacchi and Ulrych, 1995). (d) Radon panel obtained with the proposed method. (e) Primary from the standard least-squares method. (f) Primary from the Cauchy norm sparse Radon transform. (g) Primary from the proposed method.

We also test the algorithm on another sparse inverse problem, namely the local Radon transform (or generalized convolution, Sacchi et al., 2004). The idea is to simulate seismic data using a spatially and temporarily variant convolution operator. Instead of using a single Radon panel to fit the data, the method assigns weights to local Radon panels for data within local spatial windows. Sacchi et al. (2004) proposed a global problem in which the seismogram is generated by summing up the contributions from all local Radon panels. Typically if the algorithm is implemented in FX domain, one needs to think about overlapping time windows to address structural variation with time. To avoid this problem, we propose a TX domain implementation so that time-variant information is naturally incorporated. The problem is huge since the dimension of the model is higher than that of the conventional Radon transform. We find that the greedy least-squares method is suitable for this kind of sparse problem. Figure 2 shows the result of the local Radon transform for data upsampling in space. The input data is seriously aliased and there are strong amplitude variations in two events. In addition, some random noise is added to complicate the problem. Figure 2b shows the result after two iterations of the greedy method. All main features are recovered, but the noise is not. Figure 2c is the result after six iterations of the algorithm. We can see that the noise begins to appear. This is reasonable since random noise usually generates weak amplitude in the Radon panel, and consequently the inversion tends to assign a low priority to its reconstruction.

Conclusions

We have proposed a method for sparse linear problems and have validated the idea using two synthetic examples for multiple attenuation and for the local Radon interpolation. The method is efficient and not sensitive to noise. It also works for field data (not shown in the paper due to limited space) at an affordable computational cost. One shortcoming of the algorithm is that one needs to play with the threshold value to acquire optimal results. Different datasets may demand different threshold values. Except for this caveat, the method opens a door for solving many sparse problems in an efficient way. An ongoing research is focused on comparing with the Gauss-Seidel method (Ng and Perz, 2004) to better understand the benefits of the proposed method.



(a) (b) (c)

Figure 2. Greedy least-squares method for local Radon interpolation. (a) Input TX domain data. Note that every second trace is removed. (b) Interpolation result after 2 iterations of the greedy method. Most of the significant information is recovered. (c) Interpolation result after 6 iterations of the greedy method with minimum misfit.

References

Hampson, D., 1986, Inverse velocity stacking for multiple estimation: 66th Ann. Internat. Mtg., Soc. Expl. Geophys SEG Expanded Abstracts, 422-424.

Hestenes, M. and E. Steifel, 1952, Methods of conjugated gradients for solving linear systems: Nat. Bur. Standards J. Res., **49**, 403-436.

Ng, M. and M. Perz, 2004, High resolution Radon transform in the t-x domain using "intelligent" prioritization of the Gauss-Seidel estimation sequence: 74th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 2160-2163.

Li, X., B. Wang, K. Pann, J. Anderson, and L. Deng, 1998, Fast migration using a matching pursuit algorithms: 68th Ann. Internat. Mtg., Soc. Expl. Geophys SEG Expanded Abstracts, 1732-1735.

Liu, Y. and M. D. Sacchi, 2002, De-multiple via a fast least-squares hyperbolic Radon transform: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys SEG Expanded Abstracts, 2182-2185.

Sacchi, M. D. and T. J. Ulrych, 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, 1169-1177.

Sacchi, M. D., D. J. Verschnuur and P. M. Zwartjes, 2004, Data reconstruction by generalized deconvolution: 74th Ann. Internat. Mtg., Soc. Expl. Geophys SEG Expanded Abstracts. 1989-1992.