SPARSE RADON TRANSFORMS WITH BOUND-CONSTRAINED OPTIMIZATION

ANTOINE GUITTON
Department of Geophysics, Stanford University, Stanford, CA 94305, USA

1 INTRODUCTION

Radon transforms are popular operators for velocity analysis (Taner and Koehler, 1969; Guitton and Symes, 2003), noise attenuation (Foster and Mosher, 1992), and data interpolation (Hindriks and Duijndam, 1998; Trad et al., 2002). One property that is often sought in radon domains is sparseness, where the energy in the model space is focused without transformation artifacts. Sparseness is especially useful for multiple attenuation and interpolation. In practice, depending on the radon transform, sparseness can be achieved either in the Fourier (Herrmann et al., 2000) or time domain (Sacchi and Ulrych, 1995). To estimate sparse radon panels in the time domain, a regularization operator that enforces long-tailed probability density functions for the model parameters is often used. This regularization operator can be the ℓ^1 norm (Nichols, 1994) or the Cauchy norm (Sacchi and Ulrych, 1995).

In this paper, a new time-domain method is presented that yields sparse radon panels. This method estimates a sparse model by adding two models with positive/negative values only obtained with a bound-constrained optimization technique. Therefore, by forcing the model to fall within a certain range of values, the null space and its effects are decreased.

In the section following this introduction, I introduce the problem of finding a bound-constrained model and its resolution by presenting the limited memory L-BFGS-B technique (Zhu et al., 1997). This method aims at finding a solution with simple bounds for linear or non-linear problems. Then, I introduce a method that estimates sparse radon domains. Finally, this technique is tested on a synthetic and field data examples and compared to the Cauchy regularization (Sacchi and Ulrych, 1995). They demonstrate that the proposed strategy yields sparse radon panels comparable to the Cauchy approach. One advantage of this new strategy is that the choice of parameters is much simpler; for instance, no Lagrange multiplier is needed. One drawback is that two inversions need to be carried out as opposed to one for the Cauchy method.

2 FINDING A MODEL WITH SIMPLE BOUNDS

The goal of bound-constrained optimization is to find a vector of model parameters $\mathbf{m}$ such that we minimize

$$\min\ f(\mathbf{m}) \text{ subject to } \mathbf{m} \in \Omega, \tag{1}$$

where

$$\mathbf{m} \in \Omega = \{\mathbf{m} \in \Re^N \mid l_i \leq m_i \leq u_i\}, \tag{2}$$

with l_i and u_i being the lower and upper bounds for the model m_i, respectively. In this case, l_i and u_i are called simple bounds. They can be different for each point of the model space.

The sets of indices i for which the ith constraint are active/inactive are called the active/inactive sets $A(m)/I(m)$. Most of the algorithms used to solve bound constrained problems first identify $A(m)$ and then solve the minimization problem for the free variables of $I(m)$. In the next section, an extension of the quasi-Newton method L-BFGS (Guitton and Symes, 2003) that solves equation (1) is presented.

2.1 THE L-BFGS-B ALGORITHM

The L-BFGS-B algorithm is an extension of the quasi-Newton L-BFGS algorithm (Guitton and Symes, 2003) that yields a model constrained by simple bounds (Zhu et al., 1997). The L-BFGS algorithm is a very efficient algorithm for solving large scale problems. L-BFGS-B borrows ideas from trust region and gradient projection techniques while keeping the L-BFGS update of the Hessian and a line search algorithm.

The L-BFGS-B algorithm is affordable for very large problems. The memory requirement is roughly $(12+2m)N$ where m is the number of BFGS updates kept in memory and N the size of the model space. In practice, $m=5$ is a typical choice. Per iteration, the number of multiplications range from $4mN+N$ when no constraints are applied to m^2N when all variables are bounded. The program offers the freedom to have different bounds for different points of the model space. In addition, some points can be constrained while others are not.

There are three different stopping criteria for the L-BFGS-B algorithm. First the program stops when the maximum number of iterations is reached. Or, the program stops when the decrease of the objective function becomes small enough. Or, the program stops when the norm of the projected gradient (in a ℓ^∞ sense) is small enough.

Tests indicate that the L-BFGS-B algorithm ran in single precision with no constraints is not quite twice as slow as a conjugate gradient solver per iteration. This result is quite remarkable when considering that L-BFGS-B works for any type of non-linear (or linear) problem with line searches. In addition, the number of iterations needed to convergence is almost identical for both L-BFGS-B and the conjugate gradient solver. In the next section, I present a method to estimate sparse radon transforms.

3 PROPOSED METHOD TO ESTIMATE SPARSE RADON DOMAINS

Given a CMP gather $\mathbf{d}$ and a radon transform operator $\mathbf{L}$, we want to minimize the objective function

$$f(\mathbf{m}) = \|\mathbf{Lm}-\mathbf{d}\|^2, \tag{3}$$

where $\mathbf{m}$ is the unknown radon domain. The main idea of this paper is to decompose $\mathbf{m}$ into its positive and negative parts by imposing simple bounds on $\mathbf{m}$ with the L-BFGS-B algorithm. Therefore, the two problems

$$\min f(\mathbf{m}_{(-)}) \text{ subject to } \mathbf{m}_{(-)} \in \ ]-\infty,0[\ , \tag{4}$$

and

$$\min f(\mathbf{m}_{(+)}) \text{ subject to } \mathbf{m}_{(+)} \in \ [0,+\infty[\ , \tag{5}$$

need to be solved. Note that we could decompose $\mathbf{m}$ into more subdomains as well. Here, the main idea is to decrease the null space and its effects by constraining the model, similar to what is accomplished with the Cauchy regularization. Once the two models $\mathbf{m}_{(+)}$ and $\mathbf{m}_{(-)}$ are estimated with the L-BFGS-B algorithm, the sparse model is obtained by computing $\mathbf{m}_{\mathbf{sparse}} = \mathbf{m}_{(-)}+\mathbf{m}_{(+)}$. In the following section, I illustrate this technique with a synthetic and real data example using the hyperbolic radon transform.

4 EXAMPLES

Figure 1a shows a synthetic CMP gather with five hyperbolas. First, the model is constrained to have positive values in Figure 1b. Note that this domain is artifacts free and extremely focused. Second, the model is constrained to have negative values in Figure 1c. Again, the model is very sparse. Finally, the sparse model obtained by adding Figures 1b and 1c is shown in Figure 1d. As expected, this model is very sparse compared to the radon panel obtained without sparseness constraints in Figure 1e.

Now, this method is tested on one CMP gather from a marine dataset in the Gulf of Mexico. Here, the proposed method is also compared with the sparse result with the Cauchy regularization (Sacchi and Ulrych, 1995). Figure 2a shows the input data. The sparse models obtained by adding the bounded models and by using the Cauchy regularization are shown in Figures 2d and 2e, respectively. Both results are almost identical, with the new technique yielding a better panel. The residual, i.e., the difference between the input data and the remodeled data, is also very similar in both cases.

5 CONCLUSION

A new method to estimate sparse radon panels has been presented. This method is based on (1) the decomposition of the model space into positive and negative values with a bound-constrained optimization technique, and (2) the summation of the two estimated models. This decomposition has the property of reducing the null space and its effects. As illustrated with synthetic and field data examples, this method yields sparse radon panels and compares favorably with the Cauchy regularization technique. Compared to the Cauchy regularization, the proposed method is simpler to parameterize where, for instance, no Lagrange multiplier is estimated. However, more iterations are needed for the bound-constrained approach where two models are computed.

6 ACKNOWLEDGMENTS

I would like to thank the sponsors of the Stanford Exploration Project for their financial support and Westerngeco for providing the field data example.

7 REFERENCES

Foster, D. J., and Mosher, C. C., 1992, Suppression of multiple reflections using the Radon transform: Geophysics, **57**, no. 03, 386–395.

Guitton, A., and Symes, W., 2003, Robust inversion of seismic data using the Huber norm: Geophysics, **68**, no. 4, 1310–1319.

Herrmann, P., Mojesky, T., Magesan, M., and Hugonnet, P., 2000, De-aliased, high-resolution Radon transforms: Soc. of Expl. Geophys., 70th Ann. Internat. Mtg, 1953–1956.

Hindriks, C. O. H., and Duijndam, A. J. W., 1998, Radon domain reconstruction of 3-D irregularly sampled VSP data: Soc. of Expl. Geophys., 68th Ann. Internat. Mtg, 2003–2006.

Nichols, D., 1994, Velocity-stack inversion using $\mathbf{L_p}$ norms: SEP–**82**, 1–16.

Sacchi, M. D., and Ulrych, T. J., 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, no. 04, 1169–1177.

Taner, M. T., and Koehler, F., 1969, Velocity spectra - Digital computer derivation and applications of velocity functions: Geophysics, **34**, no. 06, 859–881.

Trad, D., Ulrych, T. J., and Sacchi, M. D., 2002, Accurate interpolation with high-resolution time-variant Radon transforms: Geophysics, **67**, no. 2, 644–656.

Zhu, H., Byrd, R. H., and Nocedal, J., 1997, Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization: ACM Transactions on Mathematical Software, **23**, 550–560.

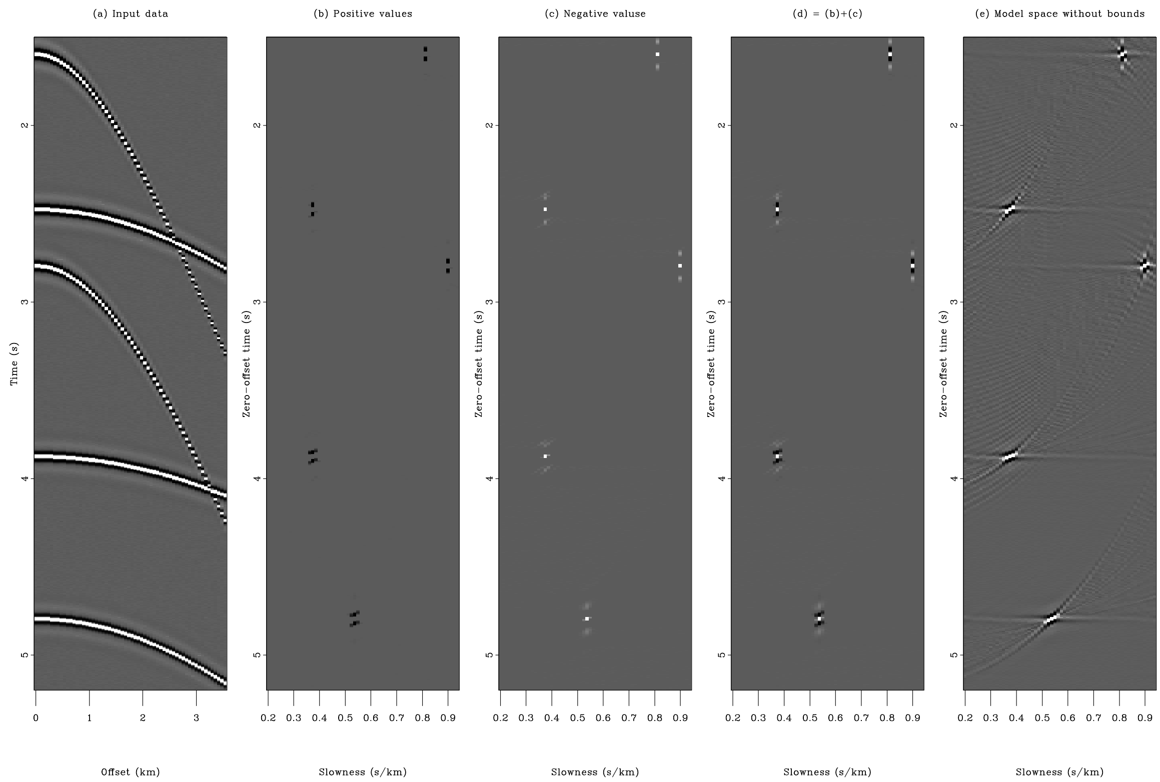

Figure 1: (a) Input data. Estimated model for (b) positive values and (c) negative values only. (d) Estimated sparse domain (b)+(c). (e) Estimated model without sparseness constraints.

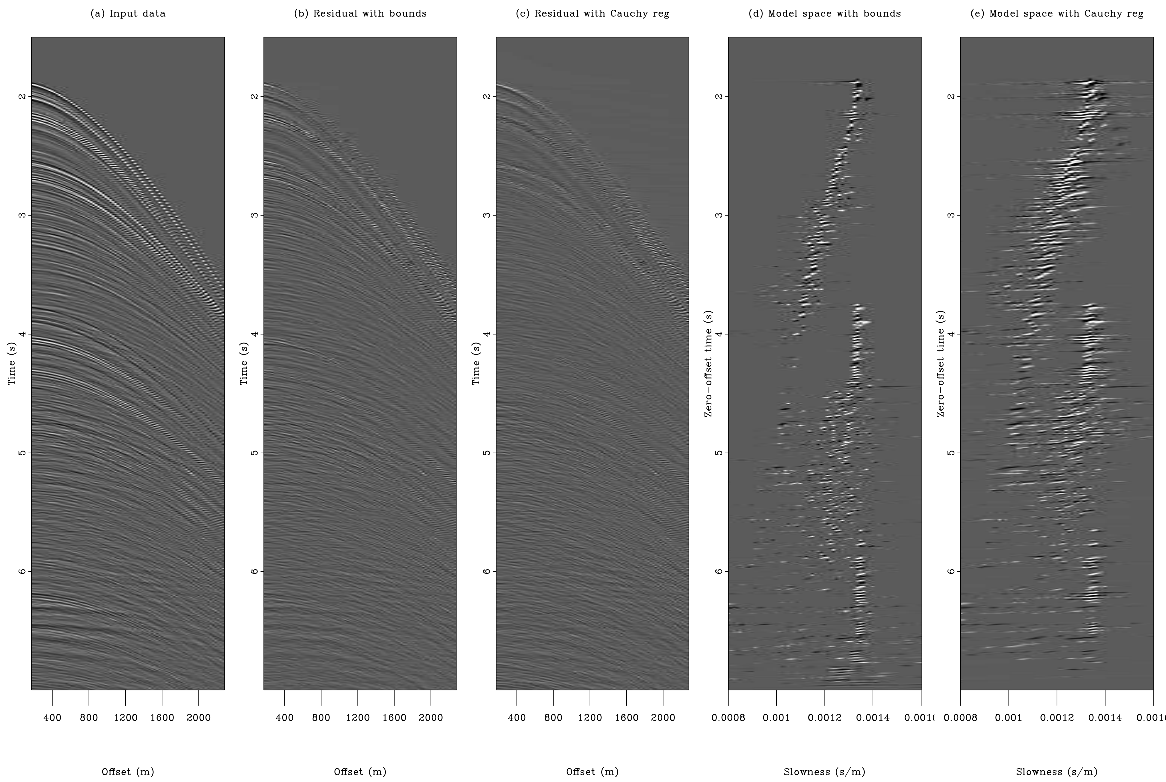

Figure 2: (a) Input CMP gather from a marine data experiment. Residual panels for the sparse model estimated with (b) the bounded models and (c) the Cauchy regularization. Sparse models estimated with (d) the bounded models and (e) the Cauchy regularization.