Sparse Radon Tansform with Dual Gradient Ascent Method

Yujin Liu[1][2]

Zhimin Peng[2] William W. Symes[2] Wotao Yin[2]

[1]China University of Petroleum (Huadong)

[2]Rice University

TRIP Annual Meeting

Overview

Overview

Introduction

- Radon Transform (RT):
  - Categories: linear RT (slant stack), parabolic RT, hyperbolic RT (stack velocity spectrum)...
  - Implementation: time domain, frequency domain
  - Application: denoising (Random noise and multiples), interpolation, velocity analysis...

Introduction

- Radon Transform (RT):
  - Categories: linear RT (slant stack), parabolic RT, hyperbolic RT (stack velocity spectrum)...
  - Implementation: time domain, frequency domain
  - Application: denoising (Random noise and multiples), interpolation, velocity analysis...
- Problems of RT operator:
  - It's not orthogonal like Fourier transform, wavelet transform ...
  - Loss of resolution and aliasing that arise as a consequence of incomplete information

Introduction

- Solution:
    - Zero-order regularization (Hampson, 1986; Beylkin, 1987)
    - Stochastic inversion (Thorson and Claerbout, 1985)
    - Sparse RT (Sacchi and Ulrych, 1995; Cary, 1998; Yilmaz and Tanner, 1994; Herrmann, 1999; Trad et al. 2003)

Introduction

- Solution:
  - Zero-order regularization (Hampson, 1986; Beylkin, 1987)
  - Stochastic inversion (Thorson and Claerbout, 1985)
  - Sparse RT (Sacchi and Ulrych, 1995; Cary, 1998; Yilmaz and Tanner, 1994; Herrmann, 1999; Trad et al. 2003)
- Our work:
  - Improve resolution with faster sparse-promotion algorithms
  - Combine seismology with compressive sensing

Overview

Hyperbolic Radon Transform

HRT operator:

$$m(\tau, v) = \sum_{x=x_{min}}^{x_{max}} d(t^2 = \tau^2 + \frac{x^2}{v^2}, x)$$

Adjoint of HRT operator:

$$d(t, x) = \sum_{v=v_{min}}^{v_{max}} m(\tau^2 = t^2 - \frac{x^2}{v^2}, v)$$

Matrix form:

$$\mathbf{m} = \mathbf{L}^T \mathbf{d}$$

$$\mathbf{d} = \mathbf{L}\mathbf{m}$$

Sparsity Promotion Methods

Basis Pursuit (BP) problem:

$$\min_{\mathbf{m}}\{\|\mathbf{m}\|_1 : \mathbf{Lm} = \mathbf{d}\}$$

Equivalent form of BP:

$$\min_{\mathbf{m}}\{\|\mathbf{W_m m}\|_2^2 : \mathbf{Lm} = \mathbf{d}\}$$

where $\mathbf{W_m} = diag(m_i^{-\frac{1}{2}})$ is weighting matrix.

Sparsity Promotion Methods

- Iteratively Reweighted Least-Squares (IRLS) method (Claerbout, 1992)
  - Non-linear inverse problem
  - Need to calculate weighting matrix at the outer loop of CG.
- Conjugate Guided Gradient (CGG) method (Ji, 2006)
  - A variant of IRLS
  - Linear inverse problem
  - Only one calculation of $\mathbf{L}$ and $\mathbf{L}^T$ is needed at each iteration.

Implementation of IRLS

Algorithm 1 IRLS method

for $j = 0 \cdots miter$ **do**
    compute $\mathbf{W}_m^j$
    $\mathbf{r}^{j,0} = \mathbf{L}\hat{\mathbf{W}}_m^j \hat{\mathbf{m}}^{j,0} - \mathbf{d}$
    for $k = 0 \cdots niter$ **do**
        $\mathbf{dm}^k = \hat{\mathbf{W}}_m^{T,j} \mathbf{L}^T \mathbf{r}^{j,k}$
        $\mathbf{dr}^k = \mathbf{L}\hat{\mathbf{W}}_m^j \mathbf{dm}^k$
        $(\hat{\mathbf{m}}^{k+1}, \mathbf{r}^{k+1}) \Leftarrow \mathrm{cgstep}(\hat{\mathbf{m}}^{\mathrm{k}}, \mathbf{r}^{\mathrm{k}}, \mathbf{dm}^{\mathrm{k}}, \mathbf{dr}^{\mathrm{k}})$
    end for
    $\hat{\mathbf{m}}^{j+1} = \hat{\mathbf{W}}_m^j \hat{\mathbf{m}}^{j,niter}$
end for

Implementation of CGG

Algorithm 2 CGG method

$\mathbf{r}^0 = \mathbf{L}\hat{\mathbf{m}}^0 - \mathbf{d}$
for $k = 0 \cdots niter$ **do**
    compute $\hat{\mathbf{W}}_m^k$
    $\mathbf{dm}^k = \hat{\mathbf{W}}_m^{T,k} \mathbf{L}^T \mathbf{r}^k$
    $\mathbf{dr}^k = \mathbf{L}\mathbf{dm}^k$
    $(\mathbf{m}^{k+1}, \mathbf{r}^{k+1}) \Leftarrow \mathrm{cgstep}(\mathbf{m}^{\mathrm{k}}, \mathbf{r}^{\mathrm{k}}, \mathbf{dm}^{\mathrm{k}}, \mathbf{dr}^{\mathrm{k}})$
end for

Dual Gradient Ascent (DGA) Method

$\ell_1\ell_2$ problem:

$$\min_{\mathbf{m}}\{\|\mathbf{m}\|_1 + \frac{1}{2\alpha}\|\mathbf{m}\|_2^2 : \mathbf{Lm} = \mathbf{d}\}$$

Dual problem:

$$\min_{\mathbf{y}}\{g(\mathbf{y}) = -\mathbf{d}^T\mathbf{y} + \frac{\alpha}{2}\|\mathbf{L}^T\mathbf{y} - \mathrm{Proj}_{[-1,1]^n}(\mathbf{L}^T\mathbf{y})\|_2^2\}$$

Gradient:

$$\nabla g(\mathbf{y}) = -\mathbf{d} + \alpha\mathbf{L}(\mathbf{L}^T\mathbf{y} - \mathrm{Proj}_{[-1,1]^n}(\mathbf{L}^T\mathbf{y}))$$

where $\mathrm{Proj}_{[-1,1]^n}(\mathbf{x})$ projects $\mathbf{x}$ into $[-1,1]^n$.

Theorem

As long as the smooth parameter α is greater than a certain value, the solutions for BP and $\ell_1\ell_2$ are identical. (Yin, 2010)

Dual Gradient Ascent (DGA) Method

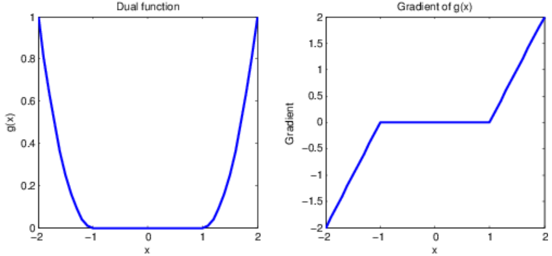

Figure: Dual objective function $g(\mathbf{y})$ (left); and its derivative $\nabla g(\mathbf{y})$ (right)

Dual Gradient Ascent (DGA) Method

Notes

Objective function $g(\mathbf{y})$ is a convex function but its gradient $\nabla g(\mathbf{y})$ is not smooth, hence, we can only apply first order methods to solve the dual problem.

Update scheme:

$$\mathbf{y}^{k+1} = \mathbf{y}^k - \delta \nabla g(\mathbf{y}^k)$$

$$\mathbf{m}^{k+1} = \alpha(\mathbf{L}^T \mathbf{y}^* - \mathrm{Proj}_{[-1,1]^n}(\mathbf{L}^T \mathbf{y}^{k+1}))$$

where $\delta > 0$ is the step size.

Theorem

It has been proved that the objective value of the primal problem given by $\mathbf{m}^*$ *matches the optimal value of the dual objective given by* $\mathbf{y}^*$*. Hence,* $\mathbf{m}^*$ *is also optimal. (Yin, 2010)*

Implementation of DGA with fixed stepsize

Algorithm 3 DGA with fixed stepsize

for $k = 0, 1, \cdots, \textit{niter}$ **do**
    $\mathbf{y}^{k+1} = \mathbf{y}^{k} + \delta(\mathbf{d} - \mathbf{L}\mathbf{m}^{k})$
    $\mathbf{m}^{k+1} = \alpha(\mathbf{L}^{T}\mathbf{y}^{k+1} - \mathrm{Proj}_{[-1,1]^{n}}(\mathbf{L}^{T}\mathbf{y}^{k+1}))$
end for

Implementation of DGAN

Several ways to speedup the fixed stepsize gradient ascent method:

- Line search
- Quasi-Newton methods, like LBFGS
- Nesterov's acceleration scheme

Main idea of DGAN

Instead of only using information from previous iteration, Nesterov's method use the usual projection-like step, evaluated at an auxiliary point which is constructed by a special linear combination of the previous two points. (Nesterov, 2007)

Implementation of DGAN

Algorithm 4 DGA with Nesterov's acceleration

$\theta_0 = 1, h > 0$
for $k = 0, 1, \cdots, niter$ **do**
  $\beta_k = \frac{(1-\theta_k)(\sqrt{\theta_k^2+4}-\theta_k)}{2}$
  $\mathbf{z}^{k+1} = \mathbf{y}^k + \delta(\mathbf{d} - \mathbf{L}\mathbf{m}^k)$
  $\mathbf{y}^{k+1} = \mathbf{z}^{k+1} + \beta_k(\mathbf{z}^{k+1} - \mathbf{z}^k)$
  $\mathbf{m}^{k+1} = \alpha(\mathbf{L}^T\mathbf{y}^{k+1} - \mathrm{Proj}_{[-1,1]^n}(\mathbf{L}^T\mathbf{y}^{k+1}))$
  $\theta_{k+1} = \theta_k \frac{\sqrt{\theta_k^2+4}-\theta_k}{2}$
end for

Overview

Synthetic data

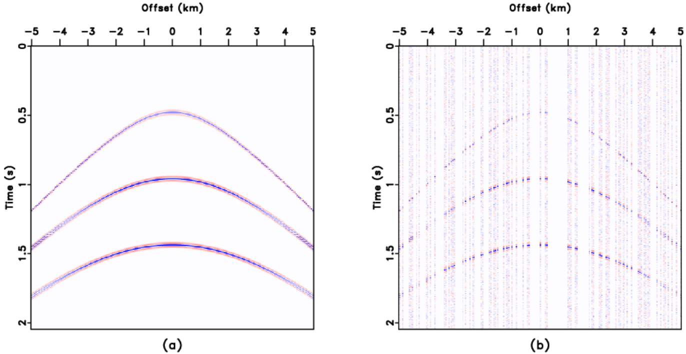

Figure: (a) Original data; (b) Input data

Inversion results

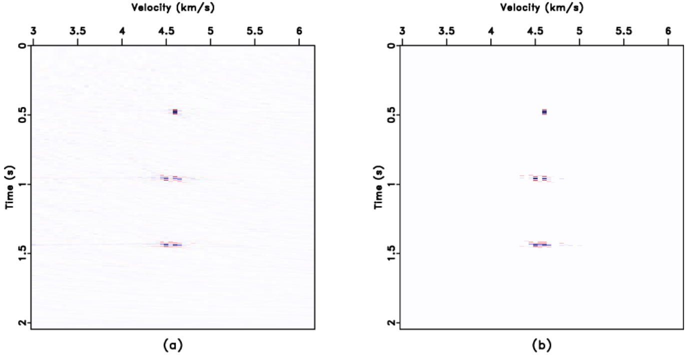

Figure: Inversion result with (a) CGG method; (b) DGAN method

Reconstructed results

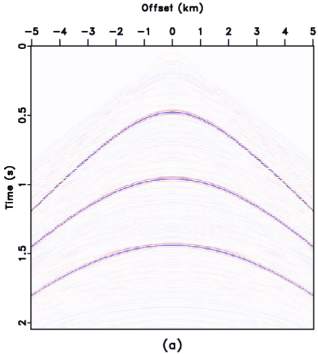

Figure: Reconstructed data with (a) CGG method; (b) DGAN method

Residual

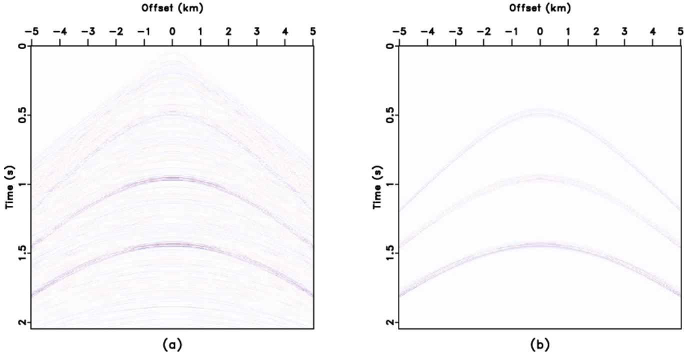

Figure: Residual with (a) CGG method; (b) DGAN method

Residual model error

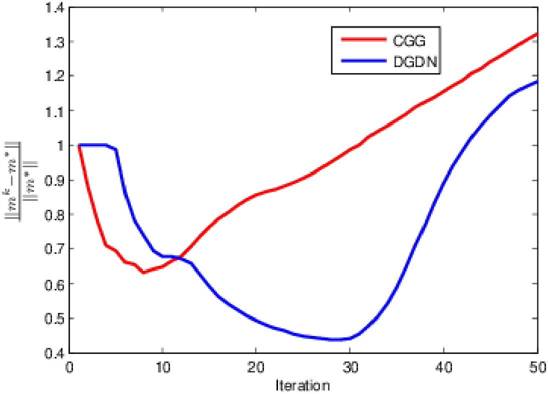

Figure: Relative model error curve of CGG method and DGAN method

Real data

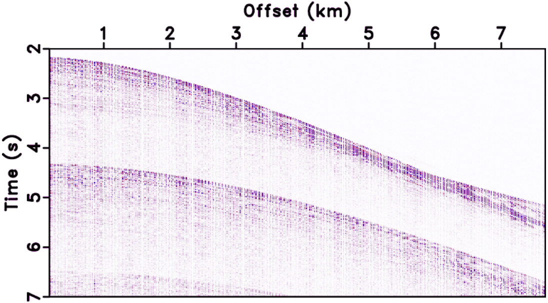

Figure: CMP gather after data binning

Real data

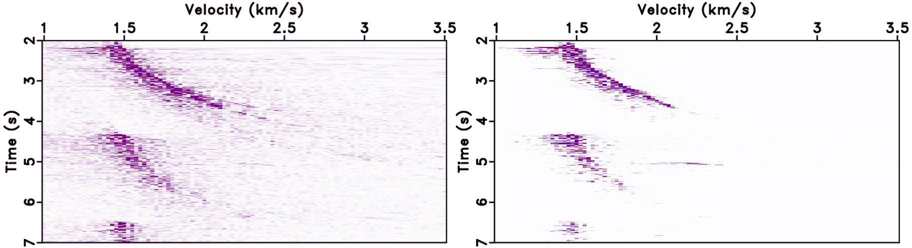

Figure: Inversion result with (a) CGG method; (b) DGAN method

Real data

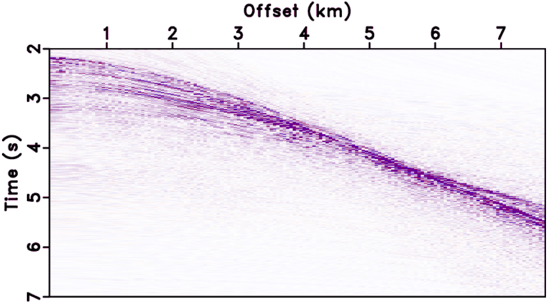

Figure: Denoising and interpolation result with CGG method

Real data

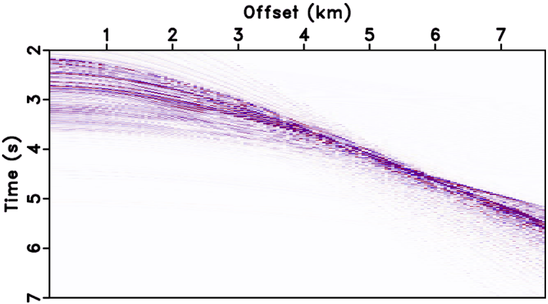

Figure: Denoising and interpolation result with DGAN method

Overview

Conclusion and Discussion

- With the help of different transformation, most signals, including seismic wave, can be expressed in a sparse form, which implies that sparsity is a very important prior information in many applications.
- A recently developed sparsity promotion method in compressive sensing is introduced into geophysics. Compared to CGG method, DGAN outperforms it in the following aspects:
    - The sparsity level of the solution is higher;
    - Reconstruction results have much less coherent noise;
    - More accurate solution can be obtained with a few iterations.
- Challenges $\Rightarrow$ Sparse representation of seismic wave.

Acknowledgments

- Thank Min Yan and Hui Zhang for helpful discussion!
- Thank CSC for supporting my visit to TRIP!
- Thank TRIP for hosting me!
- Thank the sponsors of TRIP for their support!

Thank you!