Time domain sparsity promoting LSRTM with source estimation

Mengmeng Yang, Philipp Witte, Zhilong Fang & Felix J. Herrmann
Motivation

Features of RTM:

- **pros**
  - no dip limitation
  - strong lateral velocity variations
- **cons**
  - inaccurate amplitudes & low resolution

Problems of LS-RTM:

- iterations that touch all shots are too expensive
- data can be overfitted
RTM w/ correct wavelet
Sparsity promoting LS-RTM w/ correct wavelet
Sparsity promoting LS-RTM w/ wrong wavelet
\[
\min_{\delta m} \sum_{i=1}^{n_s} \| J_i [m_0, q_i] \delta m - b_i \|^2
\]

- \( m_0 \): background model
- \( J_i \): Born modelling operator for \( i^{th} \) shot
- \( \delta m \): model perturbation
- \( q_i \): source wavelet for \( i^{th} \) shot
- \( b_i \): vectorized reflections for \( i^{th} \) shot
Sparsity promoting inversion

\[
\begin{align*}
\min_x & \quad \|x\|_1 \\
\text{s.t.} & \quad \sum_{i=1}^{ns} \| J_i [m_0, q_i] C^* x - b_i \|_2 \leq \sigma
\end{align*}
\]

\(C^*\): the transpose of Curvelet transform

\(x\): Curvelet coefficients

\(\sigma\): tolerance for noise or modelling error
Randomized subsampling

\[ \hat{J} \]

\[ \mathbf{x} = \mathbf{b} \]

\[ \hat{J}_{r(k)} \]

\[ \mathbf{x} = \mathbf{b}_{r(k)} \]

\[ n'_s \ll n_s \]
Solvers for sparsity promoting inversion

Many solvers for sparse inversion:

- Iterative soft thresholding
  (simple, but slow convergence, cooling of threshold ...)
- Spectral projected gradients w/ L1 constraint – SPGL1
  (expensive, difficult to implement, slow convergence)
- Linearized Bregman (LB)
  (easy to implement, proven convergence w/ subsampling)

Herrmann F J, Tu N, Esser E. Fast “online” migration with Compressive Sensing[J].
Sparsity promoting LS-RTM w/ correct wavelet & SPGL1
Sparsity promoting LS-RTM w/ correct wavelet & LB
Modification

\[
\begin{align*}
\min_x & \quad \lambda \|x\|_1 + \frac{1}{2} \|x\|^2 \\
\text{s.t.} & \quad \|\hat{J}x - b\|_2 \leq \sigma
\end{align*}
\]

- strongly convex objective function because of additional 2-norm term
- for big enough $\lambda$ solves BP problem

Herrmann F J, Tu N, Esser E. Fast "online" migration with Compressive Sensing[J].
Workflow for LB

\[
\begin{align*}
\text{minimize} & \quad \lambda \|x\|_1 + \frac{1}{2} \|x\|^2 \\
\text{s.t.} & \quad \|\hat{J}x - b\|_2 \leq \sigma
\end{align*}
\]

1. Initialize \(x_0 = 0\), \(z_0 = 0\), \(q\), \(\lambda\), batchsize \(n'_s \ll n_s\)
2. for \(k = 0, 1, \ldots\)
3. Randomly choose shot subsets \(I \in [1 \cdots n_s]\), \(|I| = n'_s\)
4. \(\hat{J}_k = \{J_i(m_0, q_i)C^*\}_{i \in I}\)
5. \(b_k = \{b_i\}_{i \in I}\)
6. \(z_{k+1} = z_k - t_k \hat{J}_k^T P_\sigma(\hat{J}_k x_k - b_k)\)
7. \(x_{k+1} = S_\lambda(z_{k+1})\)
8. end

note: \(S_\lambda(z_{k+1}) = \text{sign}(z_{k+1}) \max\{0, |z_{k+1}| - \lambda\}\)
\(P_\sigma(\hat{J}_k x_k - b_k) = \max\{0, 1 - \frac{\sigma}{\|\hat{J}_k x_k - b_k\|}\} \cdot (\hat{J}_k x_k - b_k)\)
Toy example

Sparsity recovery with tall ill-conditioned matrix

A: 20000 X 10000, with Rank 500
x: 10000 X 1, with 20 non-zeros
SPGL1 vs LB
no subsampling
SPGL1 vs LB
50% subsampling
SPGL1 vs LB
80% subsampling
SPGL1 vs LB
90% subsampling
What if source signature is unknown?

Estimate source by solving least-square problem:

$$\min_w \sum_{j}^{N_{tr}} \| w \ast \tilde{b}_j - b_j \|^2 \quad \text{s.t.} \| w \|^2 = 1$$

where $\tilde{b} = J(m_0, q_0) \delta m$

Suppose that $q = w \ast q_0$, and $q$ is the same for all shots $q_0$ is the initial guess of $q$
Initial wavelet setting

Signal in time domain

Approximate duration

Frequency spectrum

Frequency bandwidth wider due to factorization
Combine the image inversion & source estimation

start w/ sufficiently small threshold to allow main reflectors to enter into solution

update $\delta m$

update wavelet

start with initial wavelet
Workflow for sparsity-promoting LS-RTM w/ source estimation

1. Initialize $x_0 = 0$, $z_0 = 0$, $q_0$, $\lambda$, $\lambda_2$, batchsize $n'_s \ll n_s$, weights $r$
2. for $k = 0, 1, \cdots$
3. Randomly choose shot subsets $I \in [1 \cdots n_s]$, $|I| = n'_s$
4. $\hat{J}_k = \{J_i(m_0, q_0)C^*)_{i \in I}$
5. $b_k = \{b_i\}_{i \in I}$
6. $\tilde{b}_k = \hat{J}_kx_k$
7. $w_k = \arg\min_w \sum_{I} \|w \ast \tilde{b}_k - b_k\|^2 + \|r(w \ast q_0)\|^2 + \lambda_2\|w \ast q_0\|^2$
8. $z_{k+1} = z_k - t_k\hat{J}_k^*(w_k \ast P_\sigma(w_k \ast \tilde{b}_k - b_k))$
9. $x_{k+1} = S_\lambda(z_{k+1})$
10. end
Experiments

Data:
- 295 shots with shot interval 15m
- 295 receivers with receiver interval 15m
- 4s record, 15Hz peak frequency designed wavelet
- synthetic linearized data

Experiments:
- one pass through the data with batch sizes 2.5% data
- randomized subset of shots
- normalized true source wavelet & initial guessed wavelet
background model and model perturbation
Sparsity promoting LS-RTM w/ correct wavelet & LB
Sparsity promoting LS-RTM w/ wrong wavelet & LB
Sparsity promoting LS-RTM w/ source estimation w/ LB
Sparsity promoting LS-RTM w/ source estimation, via LB
Sparsity promoting LS-RTM w/ source estimation

- Signal in time domain
- Frequency spectrum

Diagram showing estimated, initial, and true signals and energy spectra.
Residual & model error

![Graphs showing residuals and relative model error over data passes.](image)
Robustness of source estimation starting w/ zero-phase wavelet

signal in time domain

frequency spectrum
Sparsity promoting LS-RTM w/ correct wavelet & LB
Sparsity promoting LS-RTM w/ source estimation & LB
Conclusions

- LB with correct source signature gives image with sharp interfaces w/ correct amplitudes
- Computational complexity is controlled to ~1 RTM w/ randomized source subsampling
- LB improves inversion results compared to other one-norm solvers
- LB can be combined w/ on-the-fly source estimation w/o a large computational overhead
Acknowledgements

This research was carried out as part of the SINBAD project with the support of the member organizations of the SINBAD Consortium.