Stabilized estimation of primaries by sparse inversion

Tim T.Y. Lin
Felix J. Herrmann
June 16, 2010
Estimation of Primaries by Sparse Inversion (van Groenestijn and Verschuur, 2009)

-based on Amundsen inversion, division of up/down going wavefields

\[
P^- = X_o(Q^+ + R P^-)
\]

- \(P\) total up-going wavefield
- \(Q\) down-going source signature
- \(R\) reflectivity of free surface (assume -1)
- \(X_o\) primary impulse response

(all single-frequency data volume, implicit \(\omega\))
Estimation of Primaries by Sparse Inversion (van Groenestijn and Verschuur, 2009)

-based on Amundsen inversion, division of up/down going wavefields

\[ \mathbf{P}^- = \mathbf{X}_o (\mathbf{Q}^+ + R \mathbf{P}^-) \]

\[ f(\mathbf{X}_o, \mathbf{Q}) = \| \mathbf{P}^- - \mathbf{X}_o (\mathbf{Q}^+ + R \mathbf{P}^-) \|^2 \]
Estimation of Primaries by Sparse Inversion (van Groenestijn and Verschuur, 2009)

-based on Amundsen inversion, division of up/down going wavefields

\[
\begin{align*}
\text{recorded data} & \quad \text{predicted data from primary IR} \\
\mathbf{P}^- &= \mathbf{X}_o (Q^+ + \mathbf{RP}^-) \\
 f(\mathbf{X}_o, Q) &= \|\mathbf{P}^- - \mathbf{X}_o (Q^+ + \mathbf{RP}^-)\|^2_2 \\
 \nabla f \tilde{\mathbf{X}}_o &= \left( \mathbf{P}^- - \tilde{\mathbf{X}}_o (Q^+ + \mathbf{RP}^-) \right) (Q^+ + \mathbf{RP}^-)^H
\end{align*}
\]
EPSI
Primary event estimation step

$\nabla f \tilde{X}_o$

input data
gradient
muted gradient
EPSI

Primary event estimation step

\[ \nabla f \tilde{X}_o \]

input data \hspace{2cm} gradient \hspace{2cm} muted gradient
Primary event estimation step

- muted gradient
- sparsity

4 events picked (per trace)
EPSI
Primary event estimation step

4 events picked (per trace)  windowed picked events
Wavelet matching step

\[ P^- = X_o(Q^+ + RP^-) \]

\[ f(X_o, Q) = \| P^- - X_o(Q^+ + RP^-) \|^2_2 \]

\[ \nabla f_{\tilde{Q}} = X_o^H \left( P^- - X_o(\tilde{Q}^+ + RP^-) \right) \]
**EPSI**

Wavelet matching step

1st wavelet matching gradient

\[ \nabla f \tilde{Q} \]

\[ \tilde{X}_o \]

2nd \(X_o\) Gradient

\[ \nabla f \tilde{X}_o \]

line srch + update

mute

eetc...
**EPSI**

Final results (60 iterations)

Final wavelet

Final Green’s function

Final estimated primary
EPSI

Final results (60 iterations)

Final estimated primary

Data minus estimated primary
Primary event estimation step

$\nabla f \tilde{X}_o$  

muting time

input data  
gradient  
muted gradient
Primary event estimation step

- muted gradient
- sparsity
- # events / trace

4 events picked (per trace)
Primary event estimation step

4 events picked (per trace)   windowed picked events
**EPSI**

Wavelet matching step

\[ \nabla f = Q \]

1st wavelet matching gradient

scaling factor

line srch + update

\[ \nabla f = \tilde{X}_o \]

2nd \( X_o \) Gradient

mute

etc...
Final results (60 iterations)

Final wavelet

Final Green’s function

Final estimated primary

stopping criterion
**EPSI**

Uses sparsity assumption on $X_o$

$$\minimize_{X_o, Q^+ \in Q_\Lambda} \nnz(X_o) \quad \text{s.t.} \quad \|P^- - X_o(Q^+ + RP^-)\|_2^2 \leq \sigma$$

But approximates the solution with $k$ iterations of projected gradient

$$\minimize_{X_o, Q^+ \in Q_\Lambda} \|P^- - X_o(Q^+ + RP^-)\|_2^2 \quad \text{s.t.} \quad \nnz(X_o) \leq \tau$$

This is an NP-hard problem:
- existence of local minima
- no convergence guarantees

$\tau$ number. spike per iteration

$Q_\Lambda$ short time-windowed wavelet
(implies smooth spectrum)
Convex relaxation

Use L1-norm relaxation for the sparsity objective

\[
\begin{align*}
\text{minimize} & \quad \|X_o\|_1 \\
\text{s.t.} & \quad \|P^- - X_o(Q^+ + RP^-)\|_2^2 \leq \sigma
\end{align*}
\]

Bi-convex problem, but turns into two convex problems we know how to solve via alternating optimization

- Standard approach in blind image deconvolution
- No need for windowing primary events at each iteration
Use L1-norm relaxation for the sparsity objective

\[
\minimize_{X_o} \|X_o\|_1 \quad \text{s.t.} \quad \|P^- - X_o(Q_k^+ + R P^-)\|_2^2 \leq \sigma
\]

Fix source signature, turns into \( \ell_1 \)-minimization
Operator form

\[ \mathbf{P}^- = X_o (S^+ + R \mathbf{P}^-) \]

Define linear operator \( A \) that maps Green’s func to up-going wavefield

\[ A x_o := \mathcal{F}_t^* \text{BlockDiag}_\omega \left[(Q^+ - \mathbf{P}^-)^* \otimes \mathbf{I}\right] \mathcal{F}_t x_o = \mathbf{p}^- \]

\[ \mathbf{p}^- := \text{vec}(\mathbf{P}^-) \]
\[ x_o := \text{vec}(X_o) \]
L1 minimization

\[ \min \|x_o\|_1 \quad \text{s.t.} \quad \|p - Ax_o\|_2^2 \leq \sigma \]

Use SPGL1 (van den Berg, Friedlander, 2008)
- a projected gradient based method (seismic data-volumes are huge)
- uses root-finding to find the final one-norm

\[
\min \|Ax_o - p\|_2^2 \quad \text{s.t.} \quad \|x_o\|_1 \leq \tau_k
\]
L1 minimization

\[
\min \|x_o\|_1 \quad \text{s.t.} \quad \|p^- - Ax_o\|_2^2 \leq \sigma
\]

Use SPGL1 (van den Berg, Friedlander, 2008)
- a projected gradient based method (seismic data-volumes are huge)
- uses root-finding to find the final one-norm
Alternating optimization

Wavelet matching at Pareto curve

\[
\text{minimize} \quad \|X_{Ok}\|_1 \quad \text{s.t.} \quad \|P^- - X_{Ok}(Q^+ + RP^-)\|_2^2 < \sigma_k
\]

Fix primary impulse response, get least-squares matching for \(Q^+\) past \(\ell_2\) mismatch tolerance \(\sigma_k\)
Wavelet ambiguity

same $\ell_1$ norm
Wavelet ambiguity

reset Green’s Func

wavelet matching

pick only max event

\[ \mathbf{X}_{k=1} = \text{zero vector} \]

\[ = \mathbf{Q}_{k=1}^+ \]
IR estimation 1 - start
IR estimation 1 - Pareto
Wavelet matching 2
IR estimation 2 - start
IR estimation 2 - Pareto
Wavelet matching 3
IR estimation 3 - Pareto
Wavelet matching 4
IR estimation 4 - Pareto
Wavelet matching 5
IR estimation 5 - Pareto

total ~60 gradient updates
Sparsity vs L1

Data minus estimated primary (one shot)

Sparse EPSI

$l_1$ EPSI
Sparsity vs L1

Data minus estimated primary (zero-offset)

Sparse EPSI

$l_1$ EPSI
Sparsity vs L1

Gulf of Suez

Sunday, October 31, 2010
Sparsity vs L1

Gulf of Suez (one shot)

Sparse EPSI

$l_1$ EPSI
Sparsity vs L1

Gulf of Suez (zero-offset)

Sparse EPSI

$l_1$ EPSI
Sparsity vs L1

Gulf of Suez (zero-offset zoomed)
Sparsity vs L1

Data minus estimated primary (zero-offset)

Sparse EPSI

$\ell_1$ EPSI
EPSI requires

- tight muting window
- physical locations of the primaries (window at each gradient update)
- number of reflection events inside this window
- source wavelet length
L1 reformulation

- tight muting window
- physical locations of the primaries (window at each gradient update)
- number of reflection events inside this window
- noise level in input data (use GCV in the future)
- source wavelet length
conclusions

- **less** parameters to tweak
- **improved convergence** properties of EPSI by convex relaxation
- **improved quality** of first arrivals
- cast EPSI into blind deconvolution **framework** using alternating optimization
- **removed gradient scaling issues** btw wavelet matching and IR estimation
- future work on intelligently setting $\sigma$ (GCV)
acknowledgements

Special thanks to G.J. van Groenestijn, Eric Verschuur, and the rest of the members of DELPHI

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE (334810-05) of Felix J. Herrmann. This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BP, Chevron, Petrobras, and Schlumberger.