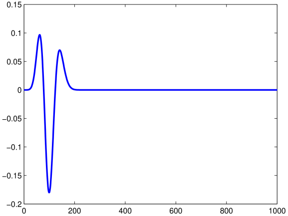
Applications of phase retrieval methods to blind seismic deconvolution

Ernie Esser

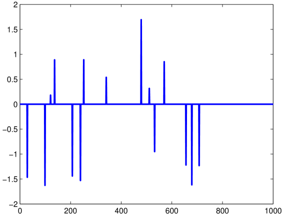
SINBAD Consortium Meeting, December 3, 2013

Data Model

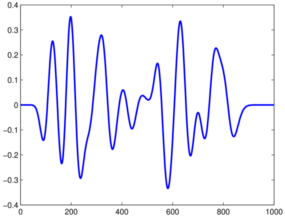
source wavelet



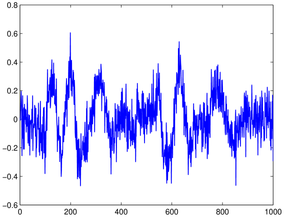
sparse reflectivity series



convolution $w * u_j$



noisy data $f_j = w * u_j + \eta \ , \ j = 1, \ldots, M$



Autocorrelation Estimate

Assuming u_j is white,

$$\frac{1}{M}\sum_{j=1}^{M}|\hat{f}_j|^2 = \frac{1}{M}\sum_{j=1}^{M}|\hat{w}|^2|\hat{u_j}|^2 \approx c|\hat{w}|^2 \qquad \text{for some c}$$

where ˆ denotes the Discrete Fourier Transform.

Assuming $\|w\| = 1$, we can then estimate

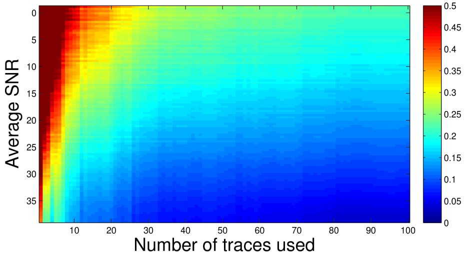
$$|\hat{w}|^2 \approx \frac{\sum_{j=1}^{M}|\hat{f}_j|^2}{\mathrm{mean}(\sum_{j=1}^{M}|\hat{f}_j|^2)}$$

Quality of Autocorrelation Estimate

Generate synthetic data f_j by convolving a fixed wavelet w with randomly generated u_i and adding Gaussian noise. Visualize relative error

$$\frac{\left\| |\hat{w}| - \text{denoise}\left(\sqrt{\frac{\sum_{j=1}^{M}|\hat{f}_j|^2}{\text{mean}(\sum_{j=1}^{M}|\hat{f}_j|^2)}}\right)\right\|}{\sqrt{N}}$$

Relative error vs SNR and M



The high frequency part can be used to estimate noise statistics of the raw estimate, which can then be further improved by applying a denoising method.

A Phase Retrieval Problem

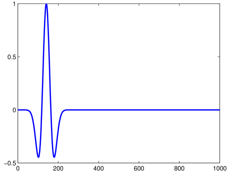
Given an estimate for $|\hat{w}|$, estimate w.

- Problem is not well posed
- w is only determined up to convolution with an all pass filter (magnitude response = 1)
- In particular, w is not determined up to arbitrary shifts

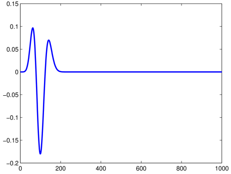
Minimum Phase Approximation

A common assumption is that most of the energy in w is near the beginning. This leads to a strategy of computing a minimum phase filter (minimum group delay) whose magnitude response is the estimate for $|\hat{w}|$.

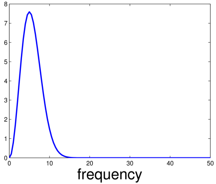
Ricker wavelet



Minimum phase approximation



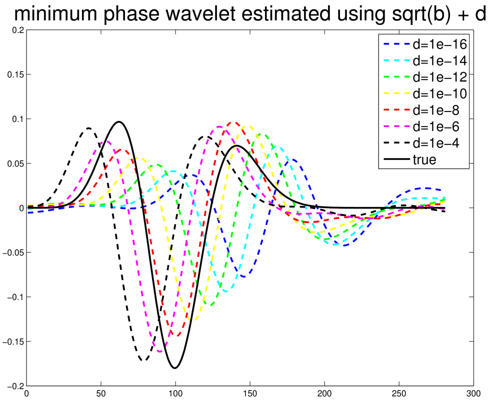
$|\hat{w}|$



Computation

Computing a minimum phase approximation can be tricky:

- Can flip zeros of the filter's transfer function inside the unit circle, but working with roots of a high order polynomials is numerically unstable
- Can apply the Discrete Hilbert Transform to $\log(\sqrt{b})$ where $b \approx |\hat{w}|^2$, but this can be sensitive to errors.



Regularization

One might also consider, given an estimate b for $|\hat{w}|^2$, solving

$$\min_w \sum_n n^2 w_n^2 \qquad \text{such that } |\hat{w}| = \sqrt{b}$$

to directly encourage energy to be concentrated at the beginning. [Lamoureux, Margrave]

This problem can be relaxed to a convex semi-definite program (SDP) and robustly solved. [Lemaréchal, Oustry], [Candes, Strohmer, Voroninski]

SDP Lifting

Let $b = |\hat{w}|^2$. These measurements are linear in ww^T.

Relax ww^T to a symmetric positive semi definite matrix $W \succeq 0$ and let F be the DFT matrix.

We can define a linear operator by $\mathcal{A}(W) = \mathrm{diag}(FWF^*)$ so that $\mathcal{A}(ww^T) = b$.

Convex Model

A possible convex relaxation of the problem of determining w from b is

$$\min_{W \succeq 0, \mathrm{tr}(W)=1} \frac{1}{2}\|\mathcal{A}(W) - b\|^2 + \gamma\, \mathrm{tr}(CW)$$

where $\gamma > 0$, C is a diagonal matrix and $\mathrm{diag}(C) = [1, 2, \ldots, N]^2$. The $\mathrm{tr}(CW)$ penalty encourages the energy in the source wavelet to be near the beginning.

Advantages of convex reformulation:

- Any local minimum is a global minimum
- We can apply robust methods guaranteed to converge to a minimizer

Disadvantages:

- The number of unknowns increased from N to N^2
- We need a rank one minimizer to solve the original problem (but this is empirically observed!)

Gradient Projection

Applying a simple gradient projection method yields the iterations

$$W^{k+1} = \Pi_{\triangle}(W^k - \alpha \mathcal{A}^*(\mathcal{A}(W^k) - b) - \alpha\gamma C)$$

where $\alpha > 0$ is a time step and $\Pi_{\triangle}$ is the orthogonal projection onto symmetric positive semi definite matrices with trace equal to one.

Let $W = V\Lambda V^* = \sum_{i=1}^{N} \lambda_i v_i v_i^*$ denote an eigenvalue decomposition of W.

Then

$$\Pi_{\triangle}(W) = \sum_{i=1}^{N} \max(\lambda_i - \theta, 0) v_i v_i^*$$

for some θ that can be easily computed using a bisection strategy.

Douglas Rachford for Constrained l_2 Model

Alternatively, the Douglas Rachford method can be used to solve a constrained form of the problem: [Demanet, Hand]

$$\min_{W \succeq 0, \mathrm{tr}(W)=1} \mathrm{tr}(CW) \qquad \text{such that} \qquad \|\mathcal{A}(W) - b\| \leq \epsilon$$

which yields the iterations

$$V^{k+1} = \Pi_{\|\mathcal{A}(\cdot)-b\| \leq \epsilon}(2W^k - V^k) - W^k + V^k$$
$$W^{k+1} = \Pi_{\triangle}(V^{k+1} - \alpha C)$$

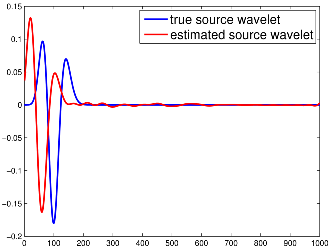
The special structure of the measurement operator $\mathcal{A}$ can be used to efficiently compute the projection onto $\|\mathcal{A}(W) - b\| \leq \epsilon$.

Recovered Source Wavelet (no noise)

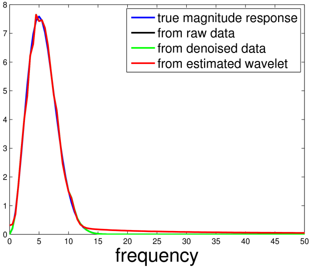
True source: Minimum phase wavelet with same magnitude response as Ricker wavelet

Using Douglas Rachford with $\epsilon = .001\|b\| \approx .2$

True and estimated source wavelet



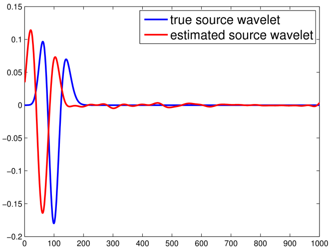
Comparison of magnitude response



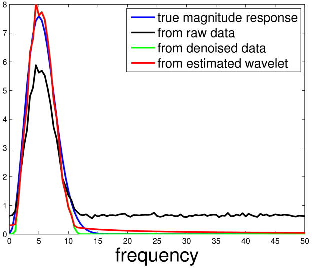
Recovered Source Wavelet (with noise, SNR ≈ 10)

Using Douglas Rachford with $\epsilon = .001\|b\| \approx .2$

True and estimated source wavelet



Comparison of magnitude response

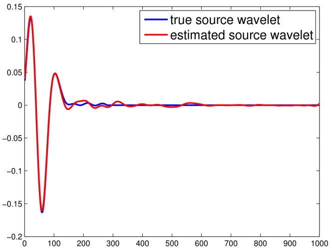


Recovered Source Wavelet (no noise)

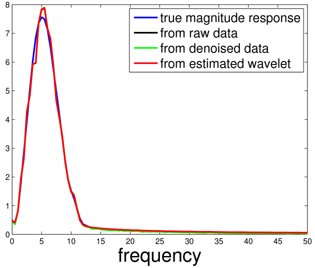
True source: A minimum phase wavelet with similar magnitude response as Ricker wavelet

Using Douglas Rachford with $\epsilon = .001\|b\| \approx .2$

True and estimated source wavelet



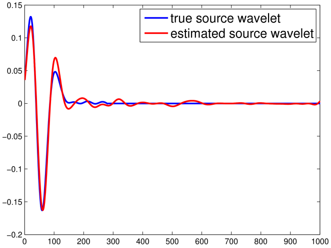
Comparison of magnitude response



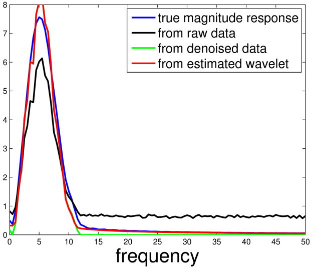
Recovered Source Wavelet (with noise, SNR ≈ 10)

Using Douglas Rachford with $\epsilon = .001\|b\| \approx .2$

True and estimated source wavelet



Comparison of magnitude response

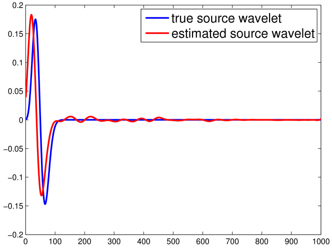


Recovered Source Wavelet (no noise)

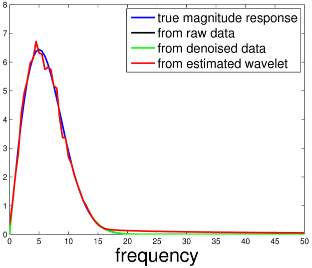
True source: A more impulsive minimum phase wavelet

Using Douglas Rachford with $\epsilon = .001\|b\| \approx .2$

True and estimated source wavelet



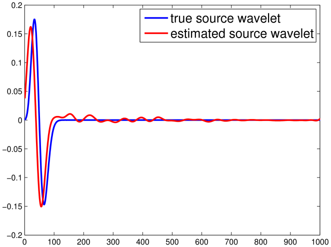
Comparison of magnitude response



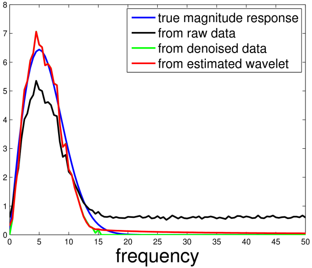
Recovered Source Wavelet (with noise, SNR ≈ 10)

Using Douglas Rachford with $\epsilon = .001\|b\| \approx .2$

True and estimated source wavelet



Comparison of magnitude response



Sparse Blind Deconvolution

To recover w that is not minimum phase and/or if we don't have a good estimate of $|\hat{w}|$, we need to use more information and assumptions such as sparsity of u_j.

A classical approach is to solve a problem of the general form

$$\min_{w,u} \sum_{j=1}^{M} \frac{1}{2} \|w * u_j - f_j\|^2 + R(u_j) \ ,$$

by alternating minimization or variable projection, where R is a penalty designed to promote sparsity of u.

Is lifting another viable strategy here?

Lifting for Blind Deconvolution

The previous model can also be lifted to a convex model by noting that the bilinear convolution measurements $w * u_j$ are linear measurements of the matrix

$$X = \begin{bmatrix} w \\ u \end{bmatrix} \begin{bmatrix} w^T & u^T \end{bmatrix} .$$

If w and u belong to certain known lower dimensional subspaces, then minimizing the nuclear norm of X subject to data constraints can recover w and u up to a scalar multiple. [Ahmed, Recht, Romberg]

From $b \approx |\hat{w}|^2$, we can estimate the support of w in the frequency domain.

Difficulty: The supports of the sparse u_j are unknown.

Goals for Lifted Model

With ww^T and $w(\text{flip}(u_j))^T$ lifted to matrices W and V_j, we want the following:

- W should be positive semi-definite
- $\text{tr}(W) = 1$
- $\begin{bmatrix} W & V_1 & \cdots & V_M \end{bmatrix}$ should have rank 1
- $\mathcal{A}(W) \approx b$ (agrees with autocorrelation estimate)
- $\mathcal{A}(V_j) \approx \hat{f}_j$ (data fidelity)
- V_j should be sparse

It is difficult to find V that is simultaneously sparse and low rank.

Convex relaxation for simultaneously sparse and low rank matrices has limits [Oymak, Jalali, Fazel, Eldar, Hassibi]
but might work with the right formulation as with SDP for sparse PCA [D'Aspremont, Ghaoui, Jordan, Lanckriet]

Conclusions

- Lifting to convexify blind deconvolution can work when the unknowns are restricted to certain known lower dimensional subspaces (as in [Ahmed, Recht, Romberg]), but so far it does not appear to be a good strategy when combining with an l_1 penalty to promote sparsity.
- SDP relaxation does, however, yield an interesting method for approximating the source wavelet from an estimate of its magnitude response and the assumption that most of its energy is concentrated near the beginning.
    - Lifted problem with 1 million unknowns can be solved in minutes
    - Robust to noise
    - Potentially a good strategy if the true source wavelet has its energy concentrated at the start

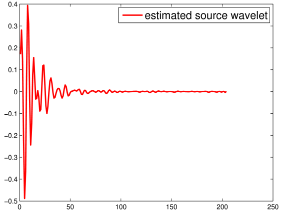
Future Work:

- Show minimizer is rank one and corresponds to a minimum phase filter
- Continue to explore models for lifted blind deconvolution

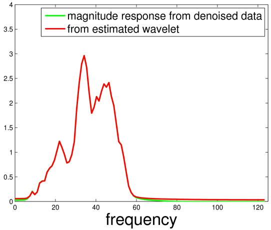
Application to Gulf of Suez Data

Select 100 random traces from the Gulf of Suez data set as input noisy data f_n and apply Douglas Rachford to the l_2 constrained lifted convex model with $\epsilon = .001\|b\| \approx .04$.

True and estimated source wavelet



Comparison of magnitude response



Deconvolution Using Estimated Source Wavelet

Estimate sparse u_j by solving

$$\min_u \mu \|u\|_1 + \frac{1}{2}\|w * u - f_j\|^2 ,$$

then debias by solving a least squares problem on the recovered support.