A lifted ℓ_1/ℓ_2 constraint for sparse blind deconvolution

Ernie Esser, Tim Lin, Rongrong Wang, and Felix J. Herrmann
University of British Columbia Dept. of Earth and Ocean Sciences

Abstract

We propose a modification to a sparsity constraint based on the ratio of ℓ_1 and ℓ_2 norms for solving blind seismic deconvolution problems in which the data consist of linear convolutions of different sparse reflectivities with the same source wavelet. We also extend the approach to the Estimation of Primaries by Sparse Inversion (EPSI) model, which includes surface related multiples. Minimizing the ratio of ℓ_1 and ℓ_2 norms has been previously shown to promote sparsity in a variety of applications including blind deconvolution. Most existing implementations are heuristic or require smoothing the ℓ_1/ℓ_2 penalty. Lifted versions of ℓ_1/ℓ_2 constraints have also been proposed but are challenging to implement. Inspired by the lifting approach, we propose to split the sparse signals into positive and negative components and apply an ℓ_1/ℓ_2 constraint to the difference, thereby obtaining a constraint that is easy to implement without smoothing the ℓ_1 or ℓ_2 norms. We show that a method of multipliers implementation of the resulting model can recover source wavelets that are not necessarily minimum phase and approximately reconstruct the sparse reflectivities. Numerical experiments demonstrate robustness to the initialization as well as to noise in the data.

Introduction

The problem of recovering a source wavelet w and sparse reflectivity signals x_j from redundant traces f_j can be modeled as a blind deconvolution problem. We assume w is the same for each included trace and consider two convolution models. The simplest model, as described for example in (Ulrych and Sacchi, 2006), is given by

$$f_j = x_j * w \ , \quad j = 1, ..., n \ , \tag{1}$$

which represents seismic data as a linear convolution of a stationary source wavelet with the primary reflectivity. The Estimation of Primaries by Sparse Inversion (EPSI) model used in (Verschuur et al., 1992; van Groenestijn and Verschuur, 2009; Lin and Herrmann, 2014) takes into account surface related multiples, leading to the series $f_j = x_j * w - x_j * x_j * w + x_j * x_j * x_j * w - + \cdots$, which sums to

$$f_j = x_j * w - x_j * f_j \ , \quad j = 1, ..., n \ . \tag{2}$$

Both the standard and EPSI blind deconvolution problems are ill-posed and can't be solved without additional assumptions about w and x_j. A classical strategy is to assume the reflectivity is statistically white, which makes it possible to estimate the autocorrelation of the source wavelet from the data. Then the wavelet can be recovered by assuming it has minimum phase (White and O'Brien, 1974).

If w is known, then sparse x_j can be estimated by solving ℓ_1 minimization problems as in (Claerbout and Muir, 1973; Santosa and Symes, 1986; Donoho, 1992; Dossal and Mallat, 2005). However, we want to recover arbitrarily shaped wavelets, and we also don't want to rely on statistical whiteness assumptions. We will therefore only assume that the x_j are sparse.

Unfortunately, when both w and the x_j are unknown, ℓ_1 regularization on x_j is not sufficient to make the standard blind deconvolution model (1) well posed (Benichoux, 2013). Nonetheless, there are successful blind seismic deconvolution methods based on nonconvex sparsity promoting penalties such as minimum entropy deconvolution (Wiggins, 1978) and variable norm deconvolution (Gray, 1979).

Minimizing the ratio of ℓ_1 and ℓ_2 norms has also been shown to promote sparsity in a variety of applications including blind seismic deconvolution (Repetti et al., 2014). The method in (Repetti et al., 2014) smooths the ℓ_1/ℓ_2 penalty and uses alternating forward backward iterations. An alternative to smoothing an ℓ_1/ℓ_2 penalty is to lift an ℓ_1/ℓ_2 constraint as was done in (D'Aspremont et al., 2007; Long et al., 2014). We propose a related approach that simplifies the constraint by splitting x into positive and negative parts. This leads to a more easily implementable ℓ_1/ℓ_2 constraint, which we use to promote sparsity of the x_j for both the standard and EPSI blind deconvolution models.

The extension to the EPSI convolution model is of particular interest because combining an ℓ_1/ℓ_2 sparsity constraint with the EPSI model eliminates the shift and scale ambiguity that is inherent in (1). With the EPSI model, if the true x_j are sparse, and if the true w is scaled or shifted, then in order to fit the data the x_j usually have to change in a way that increases $\frac{\|x_j\|_1}{\|x_j\|_2}$.

A promising application is to multiscale EPSI methods (Lin and Herrmann, 2014) that require solving more difficult yet smaller blind deconvolution problems at coarser scales. The wavelet is effectively wider at coarser scales, so more sophisticated regularization is needed to resolve the sparse signals, and yet the smaller problem sizes mean that more expensive algorithms are still computationally feasible.

Model

The main idea behind simplifying the constraint $\frac{\|x\|_1}{\|x\|_2} \leq \sqrt{\kappa}$ is to rewrite it as $\|x\|_1^2 - \kappa \|x\|_2^2 \leq 0$, which can be lifted to $1^T |xx^T| 1 - \kappa \operatorname{tr}(xx^T) \leq 0$. Here, κ is a sparsity parameter, 1 is a vector of ones and tr is the sum of diagonal entries of a matrix. Although this can be directly implemented using many inequality constraints that are linear in xx^T (Long et al., 2014), the absolute values can be simplified by splitting x into positive and negative parts so that $x = x_p - x_m$ for $x_p \geq 0$ and $x_m \geq 0$. An additional constraint of the form $x_p^T x_m = 0$ ensures that x_p and x_m are never simultaneously nonzero, in which case $|x|$ simplifies to the sum $|x| = x_p + x_m$. In this way we can obtain a lifted sparsity constraint $1^T(x_p+x_m)(x_p+x_m)^T 1 - \kappa \operatorname{tr}((x_p+x_m)(x_p+x_m)^T) \le 0$ that is linear in the matrix $\begin{bmatrix} x_p \\ x_m \end{bmatrix} \begin{bmatrix} x_p^T & x_m^T \end{bmatrix}$.

Lifting here refers to the strategy of relaxing an optimization problem in x to a larger but possibly simpler problem in the positive semidefinite matrix corresponding to xx^T. In (Ahmed et al., 2012) they use lifting to model a blind deconvolution problem as a convex optimization problem. Representing $w = Bh$ and $x = Cm$, the convolution $f = x * w$ can be written as $f = \mathscr{A}_f(hm^T)$ for a linear measurement operator $\mathscr{A}_f$. For good B and C matrices they show that solving $\min_Y \|Y\|_*$ subject to $\mathscr{A}_f(Y) = f$ is likely to have a unique rank one solution at $Y = hm^T$, where the nuclear norm $\|Y\|_*$ is the sum of singular values of Y.

In our setting, B and C are zero padding matrices, and we can't expect nuclear norm minimization to yield a unique rank one solution. We will still write $w = Bh$, $x_p = Cu$ and $x_m = Cv$ and lift (h,u,v) to a matrix Z. Then we can consider an optimization problem that combines a lifted data constraint, the lifted ℓ_1/ℓ_2 sparsity constraint, constraints to ensure u and v have nonoverlapping support and consistent signs, and a low rank penalty on Z.

Solving for the full lifted matrix Z is too computationally expensive. A compromise is to explicitly represent Z as a rank r matrix in factored form. We are already seeing good numerical results in the rank one case, which can be written in a much simpler form than when $r > 1$. For the standard blind deconvolution model, in the $r = 1$ case we solve the nonconvex problem

$$\begin{aligned} &\min_{h,u,v} \frac{1}{2}\|\Gamma h\|_2^2 + \sum_{j=1}^{n} \frac{1}{2}\|u_j + v_j\|_2^2 \qquad \text{subject to} \\ \text{Data constraints:}\quad & \|f_j - \mathscr{A}_f(hu_j^T - hv_j^T)\|_2 \le \varepsilon \\ \text{Sparsity constraints:}\quad & 1^T(u_j+v_j)(u_j+v_j)^T 1 - \kappa(u_j+v_j)^T(u_j+v_j) \le 0 \\ \text{Support and sign constraints:}\quad & u_j^T v_j = 0\ ,\ u_j \ge 0\ , v_j \ge 0 \\ \text{Wavelet normalization:}\quad & h^T h = 1\ . \end{aligned} \tag{3}$$

The matrix Γ in the objective can be zero and the model still works well for n large enough. For small n it helps to choose Γ to penalize energy in the wavelet at later times, thus preferring more impulsive wavelets as in (Lamoureux and Margrave, 2007; Esser and Herrmann, 2014).

Only minor changes are needed for the EPSI model. The data constraints change to

$$\|f_j - \mathscr{A}_f(hu_j^T - hv_j^T) + \mathscr{A}_g(f_j u_j^T - f_j v_j^T)\|_2 \le \varepsilon\ , \tag{4}$$

where $\mathscr{A}_g$ is linear and $\mathscr{A}_g(f_j u_j^T - f_j v_j^T)$ corresponds to $x_j * f_j$ in (2). The wavelet normalization constraint is no longer needed and can for instance be relaxed to $h^T h \ge c$ for some small c. Finally, to prevent spikes in x_j at early times, we let C be of the form $C = \begin{bmatrix} 0 & \mathrm{I} & 0 \end{bmatrix}^T$.

Method

We use the Method of Multipliers for problems of the general form

$$\min_x F(x) \qquad \text{subject to} \qquad h_i(x) \in C_i \tag{5}$$

under the assumption that the C_i are convex and the functions F and h_i are differentiable with Lipschitz continuous gradients. The model (3) and its EPSI extension are both of this general form. The method of multipliers finds a saddle point of the augmented Lagrangian

$$L(x,p) = F(x) + \sum_i \frac{1}{2\delta_i}\|D_{C_i}(p_i + \delta_i h_i(x))\|_2^2 - \frac{1}{2\delta_i}\|p_i\|_2^2\ , \tag{6}$$

where $D_{C_i}(p) = p - \Pi_{C_i}(p)$ is the distance from p to C_i and Π_{C_i} denotes the orthogonal projection onto C_i. It requires iterating

$$\begin{aligned} x^{k+1} &= \arg\min_x L(x, p^k) \\ p_i^{k+1} &= D_{C_i}(p_i^k + \delta_i h_i(x^{k+1}))\ . \end{aligned} \tag{7}$$

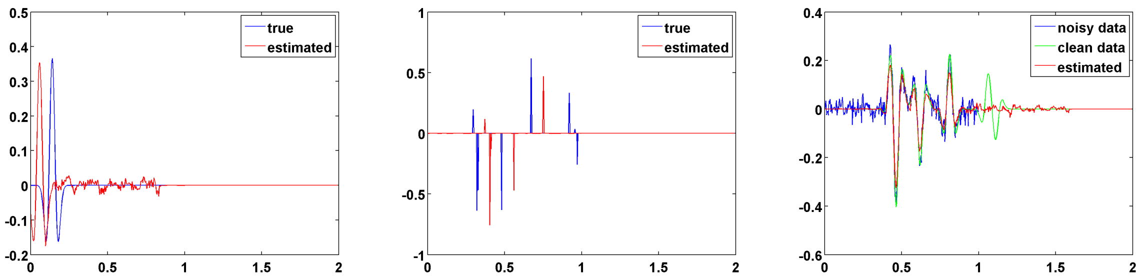
We approximately solve for x^{k+1} using the LBFGS implementation in (Schmidt, 2012), and we increase the parameters δ_i as we iterate according to the progress made towards satisfying the constraints as suggested in (Bertsekas, 1982).

Numerical Experiments

For our numerical experiments, we created synthetic data using randomly generated sparse signals x_j and a Ricker wavelet for w. Gaussian noise was added to the traces f_j, which were then cropped to the desired length in time of the observed data. A good initial guess is not required, and we simply used random initializations for our experiments.

Blind Deconvolution Results

Results for the standard blind deconvolution model using $n = 5$ measurements with moderate noise (SNR $= 13.5$) are shown in Figure 1. There is some shift ambiguity, and for small n we found that it helps to choose Γ to prefer impulsive wavelets with most energy near time zero. With more measurements there is less shift ambiguity and we can typically get good results with $\Gamma = 0$. The last two spikes are not recovered in this example is because the observed data was not long enough to see them.



(a) (b) (c)

Figure 1: Recovered wavelet (a), recovered x_1 (b) and data fit for f_1 (c) using the standard model (1).

EPSI Blind Deconvolution Results

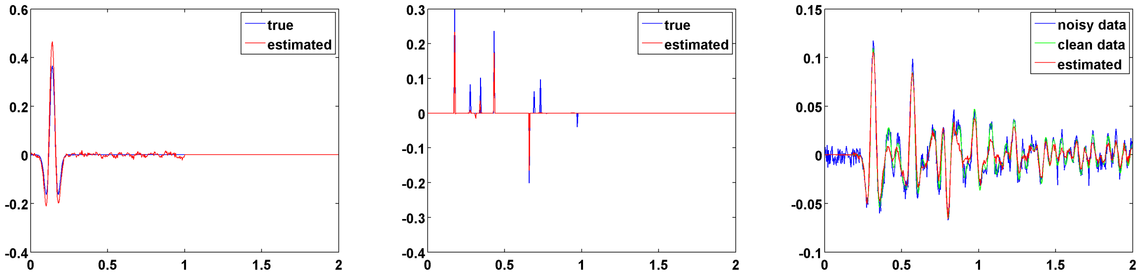
Compared to the standard model, there appears to be no shift or scale ambiguity in the wavelet recovered from the EPSI model with an ℓ_1/ℓ_2 sparsity constraint. Results for the EPSI model with $n = 50$, moderate noise (SNR $= 14.2$) and $\Gamma = 0$ are shown in Figure 2. Despite the noise and the absence of any assumption about the support of the wavelet, both the wavelet and the largest spikes in the x_j are well recovered.

Conclusions and Future Work

We showed that a Method of Multipliers implementation of a lifted ℓ_1/ℓ_2 constrained blind deconvolution model can be used to solve sparse blind deconvolution problems that have redundant measurements. Numerical results generally improve as the number of measurements increase. With more redundancy, more noise can be handled. The wavelet estimates are also substantially better when using the EPSI model, which doesn't suffer from the same kind of shift and scaling ambiguities as the classical blind deconvolution model. Future work will compare to alternative approaches and other ℓ_1/ℓ_2 implementations as well as incorporate the method into a multilevel EPSI algorithm.

Acknowledgements

This work was partly supported by the NSERC of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as



(a) (b) (c)

Figure 2: Recovered wavelet (a), recovered x_1 (b) and data fit for f_1 (c) using the EPSI model (2).

part of the SINBAD II project with support from BG Group, BGP, CGG, Chevron, ConocoPhillips, Ion, Petrobras, PGS, Statoil, Sub Salt Solutions, Total SA, Woodside and WesternGeco.

References

Ahmed, A., Recht, B. and Romberg, J. [2012] Blind deconvolution using convex programming. *arXiv preprint arXiv:1211.5608*, 40.

Benichoux, A. [2013] A fundamental pitfall in blind deconvolution with sparse and shift-invariant priors. *... on Acoustics, Speech, ...*.

Bertsekas, D.P. [1982] *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, Inc.

Claerbout, J. and Muir, F. [1973] Robust modeling with erratic data. *Geophysics*, **38**.

D'Aspremont, A., Ghaoui, L.E., Jordan, M. and Lanckriet, G. [2007] A direct formulation for sparse PCA using semidefinite programming. *SIAM review*, 1–15.

Donoho, D. [1992] superresoution via sparsity constraints. *SIAM J. Math. Anal*, **23**(5), 1309–1331.

Dossal, C. and Mallat, S. [2005] Sparse spike deconvolution with minimum scale. *In Proceedings of Signal Processing with Adaptive Sparse Structured Representations*, (4), 123 – 126.

Esser, E. and Herrmann, F. [2014] Application of a Convex Phase Retrieval Method to Blind Seismic Deconvolution.

Gray, W. [1979] Variable norm deconvolution.

Lamoureux, M. and Margrave, G. [2007] An analytic approach to minimum phase signals. *CREWES Research Report*, **19**(1), 1–12.

Lin, T. and Herrmann, F. [2014] Multilevel Acceleration Strategy for the Robust Estimation of Primaries by Sparse Inversion.

Long, X., Solna, K. and Xin, J. [2014] Two l1 based nonconvex methods for constructing sparse mean reverting portfolios. *UCLA CAM Reports [14-55]*, 1–29.

Repetti, A., Pham, M. and Duval, L. [2014] Euclid in a Taxicab: Sparse Blind Deconvolution with Smoothed l1/l2 Regularization, 1–11.

Santosa, F. and Symes, W.W. [1986] Linear Inversion of Band-Limited Reflection Seismograms. *SIAM Journal on Scientific and Statistical Computing*, **7**(4), 1307–1330.

Schmidt, M. [2012] minFunc. Http://www.cs.ubc.ca/ schmidtm/Software/minFunc.html.

Ulrych, T. and Sacchi, M. [2006] *Information-Based Inversion and Processing with Applications, Volume 36 (Handbook of Geophysical Exploration: Seismic Exploration)*. Elsevier Science.

van Groenestijn, G. and Verschuur, D. [2009] Estimating primaries by sparse inversion and application to near-offset data reconstruction. *Geophysics*, **74**(3).

Verschuur, D.J., Berkhout, a.J. and Wapenaar, C.P.a. [1992] Adaptive surface related multiple elimination. *Geophysics*, **57**(9), 1166–1177.

White, R. and O'Brien, P. [1974] Estimation of the primary seismic pulse. *Geophysical Prospecting*.

Wiggins, R. [1978] Minimum entropy deconvolution. *Geoexploration*, **16**, 21–35.