An investigation of internal parameters for sparse separation algorithms on VLF data

Matthew Staib
Mentored by David Strauss and Bob Marshall
EE REU 2013

Motivation

Stanford’s VLF Group relies on electromagnetic data that is composed of two significant parts: signals that are sparse in frequency (VLF transmitters; referred to as “Fourier components”) and signals that are sparse in time (lightning discharges; referred to as “wavelet components”). Sparse separation algorithms can be used to separate this data into these two types of components in order to:

- Reduce noise
- Compress data
- Better understand the more “interesting” time-sparse components [1].

To understand this visually, consider the spectrogram of a sample of VLF data after sparse separation (with two different sets of parameters):

![Waveform Comparisons](Image)

The current implementations of these algorithms work well in practice, but there is room for improvement. In this project we explored several internal parameters in these algorithms with the goal of improving the quality of the separation.

Testing Different Wavelets

First, we expected that by trying different wavelets (to comprise the wavelet components of \( \Phi \)), we might find that some wavelets fit the data better than others (yielding better separation). To compare wavelets, we did the following:

1. Randomly select \( n = 300 \) segments of data (of length \( 2^{15} \)) from a pool of several hours of VLF data from different sites.
2. For each wavelet, produce an average tradeoff curve:
   - (a) Run FISTA on each segment of data, for a range of values of \( \eta \).
   - (b) For each value of \( \eta \), average \( \| y - \Phi \hat{x} \|_2^2 \) and \( \text{card}(x) \) over all data segments.
   - (c) Plot this pair of means on the tradeoff curve for this wavelet.
3. Rank wavelets by the area underneath the tradeoff curve.

![Sparsity vs. \( \eta \) error for different wavelet families](Image)

Testing Different Problem Sizes

Second, we expected that running the algorithms on samples of data of different sizes would produce different tradeoff curves, and require different values of \( \eta \) to operate in the knee of the curve. In order to create a heuristic for choosing a reasonable value of \( \eta \), we did the following:

1. For each of a range of problem sizes, compute the average tradeoff curve over \( n = 300 \) segments of that size.
2. Automatically choose a point in the “knee” of the tradeoff curve:
   - (a) Choose tangent point of curve with line \( y = mx + b \) (visually chose \( m = -1/5000 \) from the curves comparing wavelets).
   - (b) Choose point on curve with least norm.
   - (c) Choose point on curve with least error satisfying \( \| x \|_0 \leq 0.25 \).

![Sparsity vs. \( \eta \) error for different problem sizes](Image)

Background

Given VLF data \( y \) and a “dictionary” of Fourier and Wavelet components \( \Phi \), we ideally want to find a weight vector \( x \) so that \( \| y - \Phi x \|_2^2 \) and \( \text{card}(x) \) are both small (where \( \text{card}(x) \) is the number of nonzero elements of \( x \)). However, problems of this form are intractable, as \( \text{card}(x) \) is not convex. Instead, we solve the “relaxed” problem

\[
\text{minimize } \| y - \Phi x \|_2^2 + \lambda \| x \|_1,
\]

where \( \lambda \) is chosen to decide a tradeoff between error and the \( \ell_1 \) norm of \( x \) (and hopefully also the sparsity of \( x \)).

There are several methods available to solve this optimization problem. We used an algorithm called FISTA (Fast Iterative Shrinkage-Thresholding Algorithm) [2]. One clarification: in creating tradeoff curves with FISTA, we did not directly alter the parameter \( \lambda \) as above, but instead a parameter we will call \( \eta \) that functions in a very similar way.

Conclusions

The Haar wavelet faced noticeably worse than the other wavelets tested, which performed very similarly to each other. The Meyer wavelet (dme) produced the tradeoff curve with the least area, but only by a slight margin (\( \approx 1\% \) less than the next best wavelet).

In testing different problem sizes, we have created useful heuristics for choosing a value of \( \eta \) when running FISTA on VLF data. Namely, linear fit (for methods (a) and (b) above) yields

\[
\eta = -0.063 \log_2 N + 2.393,
\]

where \( N \) is the length of the data \( y \), while method (c) suggests always choosing \( \eta = 1.3 \). Together these results suggest \( \eta \leq 1.3 \) for large problem sizes.

Overall, we have confirmed our previous choice in wavelets and made it possible to easily choose \( \eta \) so that we avoid unfavorable sections of the tradeoff curve.

Selected References


Acknowledgements

Thank you to David Strauss and Bob Marshall for getting me started and advising me throughout the summer, and thank you to the Stanford EE REU program for making this project possible.