APPLICATION OF BASIS PURSUIT IN SPECTRUM ESTIMATION

Scott Shaobing Chen & David L. Donoho IBM T.J. Watson Research Center & Stanford University email: schen@watson.ibm.com & donoho@stat.stanford.edu

ABSTRACT

In this paper, we apply Basis Pursuit, an atomic decomposition technique, for spectrum estimation. Compared with several modern time series methods, our approach can greatly reduce the problem of power leakage; it is able to superresolve; moreover, it works well with noisy and unevenly sampled signals. We present experiments on bizarrely spaced radial velocity data from one of the newly-discoved extrasolar planetary systems.

1. INTRODUCTION

Spectrum analysis has been widely used to detect and characterize periodic behavior present in signals. By examining the peaks in the spectrum, one can locate the frequencies of the periodic components. Therefore, it is essential to obtain good estimates of the spectrum. Over the past decades, researchers have developed various tools for estimating spectra, ranging from the traditional Fourier analysis, to modern time series methods such as such as multi-taper spectral analysis and wavelet denoising of power spectra. In this paper, we address 4 important issues in spectrum estimation:

- Power Leakage. This is a common problem in spectrum estimation. Figure 1 (a) shows a pure sine wave, with its frequency not at the Fourier frequencies; panel (b) shows the idea spectrum: a impulse response at the correct frequency; panel (c) shows the power leakage in the Fourier spectrum: there is energy not only at the correct frequency, but also at the neighboring frequencies; such structure is often called a *sidelobe* in spectrum analysis.
- Superresolution. Sometimes we would like to resolve periodic components with their frequencies very close to each other. However, because of the power leakage problem, most spectrum estimates cannot resolve such fine structures. Indeed, linear methods can resolve up to the so-called *Reyleigh limit*. Figure 2 (a) shows the TwinSine signal, which consists of two sine waves with close frequencies; panel (c) shows there is only one peak in the Fourier spectrum.
- Noisy Observations. We often observe signals contaminated with noise. The power spectrum is well-



Figure 1. Power leakage: analyzing a pure sine wave

known to be noisy; having noise in the observation makes the power spectrum even more noisy. Thus we would like to obtain spectrum estimates robust against noise.

• Uneven Sampling. Normally we work with evenly sampled signals. However, in astronomy and other sciences, one often has to cope with unevely sampled signals. Figure 5 (a) shows a bizarrely spaced radial velocity dataset from a newly-discovered extrasolar planetary system; these data are from the pioneering planetary detection project by Marcy and Butler [2, 10]. By examining the spectrum, one can figure out the orbital periods of the stars. Uneven spacing complicates the estimation and interpretation of the spectrum; the example shown here is particularly challenging because not only is the sampling uneven, but there is large and systematic evolution of the spacing.

Recently there has been great interest in representing signals with *overcomplete* dictionaries. Rather than the traditional orthogonal representations such as the Fourier analysis, atomic decomposition techniques such as Basis Pursuit [1] have been proposed to represent signals as lin-



Figure 2. Superresolution: analyzing the TwinSine signal

ear combinations of *atoms* from the overcomplete dictionary. Since overcomplete dictionaries are often richer than orthogonal bases, these methods can often obtain representations with high resolution.

In this paper, we apply Basis Pursuit with a specially designed dictionary for spectrum estimation; the dictionary includes overcomplete cosine and sine bases, and the Dirac basis. Our approach can greatly reduce the problem of power leakage; it is able to superresolve; moreover, it works well with noisy and unevenly sampled signals.

This paper is organized as follows: section 2 describes the existing methods for spectrum estimation; section 3 describes our approach via Basis Pursuit; section 4 presents experiments with the signals discussed in this section.

2. EXISTING METHODS

In this section, we review several modern time series methods, for both evenly and unevenly sampled signals. Let $\mathbf{t} = t_1, \ldots, t_n$ be the sample times and $\mathbf{x}(t) = x(t_1), \ldots, x(t_n)$ be the signals sampled at \mathbf{t} .

2.1. Fourier Analysis

The traditional method of spectrum analysis is to estimate the power spectrum directly from the Fourier transform of the signal:

$$I(w) = |F(w)|^2.$$
 (1)

where F(w) is the Fourier transform.

However, for unevenly sampled signals, estimation of the power spectrum directly from the Fourier transform of the signal encounters many difficulties: (1) is not well-behaved statistically. One popular method is the Lomb-Scargle algorithm, which modifies the simple direct Fourier sum to make the statistics of the resulting power $|F(w)|^2$ well-behaved [9]. Another scheme is *tapering*: by multiplying the signal by a *window* function before taking the Fourier transform, it reduces the spectral leakage at the expense of a slight broadening of the main lobes [7]. Slepian and Thomson [8] proposed to find the optimum window function by a mathematical optimization. While this optimization problem has a unique solution, it actually leads to a number of window functions nearly as good as the solution. The multitaper spectrum estimate [8] is the average of the Fourier spectra of the signal windowed by these functions.

2.2. Wavelet Shrinkage

Recently, wavelet shrinkage [5] has emerged as a useful tool for recovering signals from noisy observations. By transforming into the wavelet domain, most of the signal becomes concentrated in a few big coefficients whereas the small coefficients are mainly noise. Thus by thresholding the wavelet coefficients and transforming back to the time domain, one can remove noise without sacrificing resolution. Wavelet shrinkage can be realized in O(N) time, since there are fast algorithms for wavelet transforms. Not only does Wavelet shrinkage work well in practice, but also it is proven to be essentially optimal for a variety of signal classes [5].

Gao [6] has studied denoising of power spectrum estimates by wavelet shrinkage. The power spectrum is wellknown to be noisy and have a peculiar statistical distribution; however its logarithm is better behaved. Thus one can apply wavelet shrinkage on the logarithm of an estimate of the power spectrum.

For unevenly sampled noisy observations, Scargle [10] proposed to first compute the the unequally-spaced Haar wavelet transform and apply the shrinkage operator, then compute the ordinary (evenly-sampled) inverse Haar wavelet transform. This yields an evenly spaced reconstruction, on which regular spectrum analysis can be performed.

3. BASIS PURSUIT

Basis Pursuit (BP) is a principle for decomposing a signal into an optimal superposition of dictionary elements, where optimal means having the smallest l^1 norm of coefficients among all such decompositions:

$$\min \|\alpha\|_1 \text{subject to } \Phi \alpha = \mathbf{s} . \tag{2}$$

where Φ is a matrix whose columns are dictionary elements. This optimization principle leads to decompositions that can be very sparse; it can stably super-resolve in ways that other methods usually cannot. Utilizing modern time-frequency dictionaries, Basis Pursuit can find good representations for various signals. The optimization (2) can solved via linear programming [1].

We design a special overcomplete dictionary for spectrum estimation. It consists of 3 components: • Discrete Cosine Dictionary. Let ℓ be a whole number; the ℓ -fold cosine dictionary is the collection of all cosines with $\omega_k = 2\pi k/(\ell n), \ k = 0, \dots, \ell n/2$:

$$\{\cos(\omega_k \mathbf{t}: i=0,\ldots,\ell n/2\}$$

 Discrete Sine Dictionary. Let *l* be a whole number; the *l*-fold cosine dictionary is the collection of all sines with ω_k = 2πk/(*l*n), k = 1,..., *l*n/2 - 1:

$$\{\sin(\omega_k \mathbf{t}: i=1,\ldots,\ell n/2-1\}$$

• The *Dirac* dictionary is simply the collection of waveforms that are zero except in one point:

$$\{1_{\{t=t_i\}}: i=1,\ldots,N\}$$

We choose the discrete cosine dictionary and the discrete sine dictionary because they are able to characterize periodic signals; we include the Dirac dictionary to compensate the noise present in the observation.

We perform Basis Pursuit with this dictionary. Let $\hat{\alpha}$ be the Basis Pursuit solution; the power spectrum can be estimated from the coefficients associated with the cosine components $\hat{\alpha}_{\omega_k}^{cos}$ and the coefficients associated with the sine components $\hat{\alpha}_{\omega_k}^{sin}$:

$$I(\omega_k) = |\hat{\alpha}_{\omega_k}^{cos}|^2 + |\hat{\alpha}_{\omega_k}^{sin}|^2 .$$
(3)

Here are the advantages of our approach:

- Power Leakage. Basis Pursuit tends to find sparse representations, because of the nature of minizing the l¹ norm; i.e. it tries to represent a signal with fewest coefficients as possible. Thus, Basis Pursuit seldom uses all the frequencies in the neighborhood of the true frequency. Figure 1 (d) shows the Basis Pursuit spectrum with l = 1, i.e. without over-sampling the frequency domain in our dictionary. Clearly, there is much less power leakage, compared with the Fourier spectrum in panel (c).
- Superresolution. l^1 methods, along with maximum entropy methods, are well-know to be able to superresolve [4]. Moreover, our dictionary automatically defines a finer resolution in the frequency domain: the bigger the overcompleteness ℓ is, the finer the frequency resolution will be. By increasing the frequency resolution, one can push down the resolution limit of our method. Figure 2 (c) shows the Basis Pursuit spectrum obtained with $\ell = 3$: it indicates that two peaks are present; however, the position of the right peak is slightly off. Panel (d) displays the Basis Pursuit spectrum obtained with $\ell = 10$: not only are the peaks better separated, but also they are positioned closer to the true frequencies.
- Noisy Observations. For noisy observations, Basis Pursuit finds a solution with coefficients in sines and cosines, which represents the periodic components in the signal, and coefficients in Diracs, which represent mainly the noise in the signal.



Figure 3. Analyzing the noisy TwinSine signal, SNR = 2

• Uneven Sampling. Our dictionary handles uneven sampling naturally: the dictionary elements are generated as cosines and sines on the sampled time locations.

We comment that for evenly sampled observations, fast algorithms are available for the discrete cosine and discrete sine transforms; thus the analysis operator Φ^T s and the synthesis operator Φ^T a can be computed in $O(N \log(N))$ time. This is important when N is large, since the Basis Pursuit implementation requires a series of such operations. However, for unevenly sampled observations, it is not clear that fast algorithms exist for the discrete cosine and discrete sine transforms. Fortunately, in the case of the Star dataset, the sample size is small (55), and the Basis Pursuit solution can still be obtained in reasonable time.

4. EXPERIMENTS

In this section, we present experiments on noisy and unevenly sampled signals.

4.1. TwinSine

Figure 3 (b) shows the noisy TwinSine signal at signal to noise ratio 2. Panel (c) shows the Fourier spectrum; there is only one peak with wide decaying tails; the position of the peak is away from the two true frequencies, which are indicated by the two dotted lines. Panel (d) shows the Basis Pursuit spectrum with $\ell = 10$. Clearly it indicates two peaks; the positions of the peaks are very close to the true frequencies.

4.2. Synthetic Star Data

Figure 4 (a) displays a synthetic dataset for the Kepler orbit found in [3]; the computed radial velocity was sampled at the same times as the actual data in Figure 5 (a) (this



Figure 4. Analyzing the synthetic radial velocity data

dataset was constructed by Scargle [10]). The exact spectrum is indicated by the frequency impulses in Panel (b). Scargle [10] estimated the spectrum by first multiplying the signal by a taper, then computing the Lomb-Scargle periodogram; the resulting spectrum is shown in panel (c) (same as Figure 2 (a) in [10]); the dotted lines indicate the positions of the true harmonics. Clearly there are peaks associated with the true harmonics; however, the spectrum appears noisy, with many other spurious peaks which are not present in the signal. [10] considered these already reasonably good, specially in view of the peculiarity of the sampling. Panel (d) shows the Basis Pursuit spectrum with $\ell = 10$; the first 3 harmonics are clearly indicated at the right locations, with nearly no power leakage.

4.3. Star Data

Figure 5 (a) displays the real radial velocity data. Panel (b) displays the Lomb-Scargle periodogram, obtained the same way as Figure 5 (c) (same as Figure 3 in [10]). Scargle [10] used the wavelet method to denoise and interpolate the data, then computed the regular Fourier spectrum. Panel (c) displays the resulting spectrum (same as Figure 5 in [10]). (c) appears less noisy than (b); both reveal a dominant harmonics and possibly a few others. Panel (d) displays the Basis Pursuit spectrum with $\ell = 10$. The first 3 harmonics are exact multiples of $\omega = .0014 : \{i \times \omega : i = 1, 3, 5\}$. The other peaks have rather different locations than the peaks in (b) and (c).

5. ACKNOWLEDGEMENTS

We give special thanks to Jeff Scargle, who kindly provided us with the star radial velocity datasets, and his software for reproducing the figures in his work [10].



Figure 5. Analyzing the real radial velocity data

REFERENCES

 S. Chen, D. Donoho and M. Saunders, "Atomic decomposition by basis pursuit", to appear in SIAM J. Scientific Computing.

(http://www-stat.stanford.edu/~schen/).

- [2] R. Butler and G. Marcy, "A planet orbiting 47 urae majoris", Ap. J. Letters, vol. 464, pp. 153-156, 1996.
- [3] W. Cochran et. al., "The discovery of a planetary companion to 16 Cygni B", ap. J., in press, 1997.
- [4] D.L. Donoho, I.M. Johnstone, J.C. Hoch and A.M. Stern, "Maximum Entropy and the Nearly Black Object", J. Roy. Stat. Soc. Ser B, vol. 54, pp. 41-81, 1992.
- [5] D.L. Donoho, "De-Noising by soft thresholding", IEEE Trans. Information Theory, vol. 41, pp. 613– 627, 1995.
- [6] H. Gao, "Wavelet estimation of spectral densities in times series analysis", Ph.D. thesis, Univ. of California, Berkeley, 1993.
- [7] F. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform", Proc. IEEE, vol. 66, no. 1, 1978.
- [8] D. Percival and A. Walden, Spectral analysis for physical applications: multitaper and conventional univariate techniques, Cambridge Univ. Press, Cambridge, 1993.
- [9] J. Scargle, "Studies in astronomical time series analysis. II. statistical aspects of spectral analysis of unevenly spaced data", Astrophysical J., vol. 263, pp. 835-853, 1982.
- [10] J. Scargle, "Astronomical time series analysis: new methods for studying periodic and aperiodic systems", Manuscript, 1996. (http://ccf.arc.nasa.gov/~scargle).