Noise-thresholding sparse-spike inversion with global convergence: calibration and applications

Douglas S. Sassen and Michael Glinsky, Ion Geophysical Corp.*

Summary

We demonstrate an innovative inversion technique for providing a discontinuous representation of Earth reflectivity, consistent with the seismic bandwidth above noise, with proven convergence to the global minima. We also demonstrate a complementary approach for estimating, from well log data, the seismic wavelet and bandwidth above noise required in parameterizing the inversion.

Introduction

It is not uncommon to see rock outcrops with sudden changes in rock character, or to see continuous changes that at the scale of a seismic wavelet (10's to 100's of meters) are effectively a discontinuity in the rock properties. Inspired by these insights, a common objective is to inverse model, or deconvolve, seismic data to recover the discontinuities of the Earth, including by sparse-spike inversion (e.g. Debeye and Van Reil, 1990, and Barrodale and Roberts, 1973). The need for a sparsity constraint within the minimization function of this inverse problem makes it a nonlinear optimization problem, which leads to problems in converging to the global minima that require considerable constraints to converge to an acceptable solution. Daubechies et. al. (2003) has introduced a method for solving this class of problems with proven global convergence. Despite global convergence of this method, accurate parameterization of the model is needed to get solutions consistent with the real Earth.

Theory and Methods

Typical seismic inversion is applied to post-migrated data, which dictates how we model the seismic data and noise. The majority of seismic data is migrated with a weak-scattering and stationary wavelet assumption that does not consider effects such as internal multiples, surface waves or dispersion. Considerable effort is often invested into arresting noise that isn't consistent with this model, without degrading the integrity of the desired signal. Therefore, the seismic model used in our inversion also assumes a weak interaction between reflectors that allows us to simply express seismic data and reflectivity with a convolutional model expressed as,

$$S(t) = W(t) * R(t) + n(t), \qquad (1)$$

where S(t) is the seismic signal, W(t) is the seismic wavelet, R(t) is the reflectivity series and n(t) is the noise. In actual practice, multiples and processing errors are not completely removed, which introduces noise correlated to the model. Special care must be taken to define noise levels and limit model parameters in such a way as to avoid fitting coherent noise, which we address later on.

The inverse problem is to recover R(t) from the seismic and source wavelet by searching for models of R(t) that minimize the norm of the observed seismic and the proposed synthetic seismic. In the frequency domain it is easy to see that equation 1 is actually a linear problem. For an inconsistent set of linear equations, which arise when a system is noisy or the wavelet is not well determined, regularization is usually needed to find a stable solution to the inverse. In the typical L2 norm case, the minimization function becomes,

$$f = |W(t) * R(t) - S(t)|^2 + \lambda |R|^2, \qquad (2)$$

where λ is the weighting term for the regularization. In a Bayesian sense, the L2 regularization term implies a Gaussian model covariance (Tarantola, 2005), and is equivalent to a prior smoothing constraint that essentially removes any sharp discontinuities from the resulting inversion as λ increases to allow convergence to norm. An alternate regularization to the problem that doesn't result in a smoothed representation of the system is to use a L1 norm regularization term. The L1 norm regularization implies a exponential model covariance (Tarantola, 2005), and promotes a sparse and sharp representation of the solution,

$$f = |W(t) * R(t) - S(t)|^2 + \lambda |R|^1. \qquad (3)$$

This functional (equation 3) is a mixed norm, with L2 on the model discrepancy and a L1 on the regularization, that implies a Gaussian error model (L2), while the L1 enforces that the model of reflectivity has a sparse-spike representation.

Innovative Sparse-Spike Inversion

We are utilizing a soft iterative thresholding technique for solving linear inverse problem with a sparsity constraint, first proposed by Daubechies et. al. (2003), and first applied to seismic by Mallat (personal communication, 2005) and Dossal and Mallat (2005). While solving a set of linear equations that are naturally diagonal in an orthonormal basis with L1 regularization is relatively straight forward (e.g. thresholded SVD), the more general inversion problem with a L1 regularization (equation 3), such as sparse-spike deconvolution, is actually a set of coupled non-linear equations (Daubechies et. al. 2003), for which convergence to norm is not guaranteed. The ground-

DOI http://dx.doi.org/10.1190/segam2013-0617.1

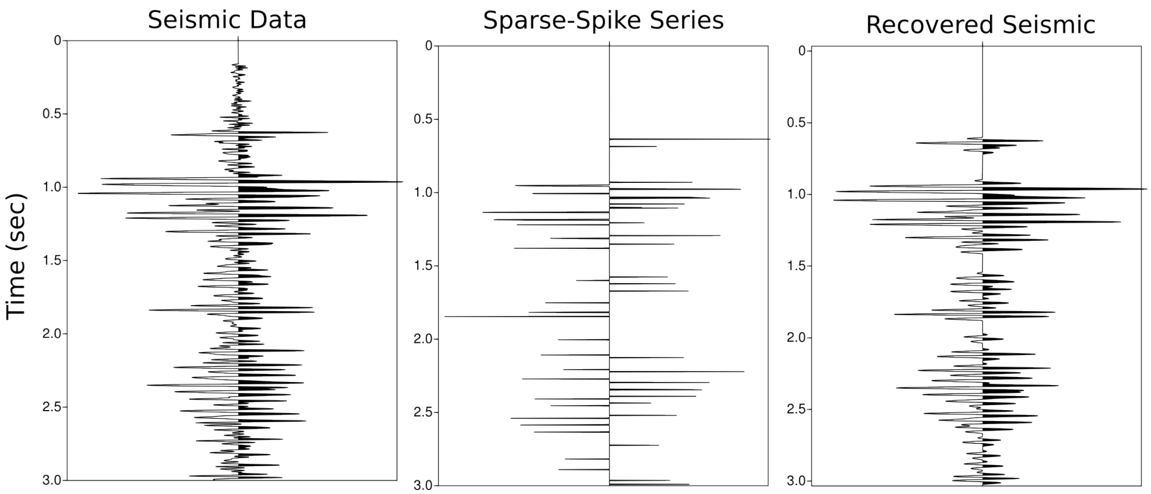

Figure 1. An example of application of the noise-thresholding sparse-spike inversion, (left) with the original migrated seismic data, (center) the inverted sparse-spike series, and (right) the synthetic seismic recovered from the wavelet and the sparse-spike series. Note that the recovered seismic (right) closely matches the seismic data for more significant events while removing the background noise.

breaking insight by Daubechies et. al. (2003) was to bypass this problem, by using surrogate functionals that orthogonalizes the problem. These surrogates are relatively simple to minimize via an iterative convex optimization method with proven convergence to norm. Therefore, converging to local minima does not plague this technique, nor does it rely on local well log constraint to find an acceptable solution.

In more simple terms, this technique uses a soft thresholding rule to eliminate wavelet coefficients (i.e. spikes) that fall below noise levels while iterating towards a sparse set of coefficients that fit observed seismic data. The end result is a model that has preserved reflections resolvable by the bandwidth of the seismic above noise - without smoothing (Figure 1). In order for this technique to fulfill its potential, accurate estimates of both the source wavelet and the proper noise thresholding level are required. These requirements can be met with a wavelet extraction technique that estimates noise levels from downhole log data co-located with the seismic data.

Wavelet Extraction and Noise Estimation

We utilize the open-source software WaveletExtractor (Gunning and Glinsky, 2006) to make estimates of the wavelet, noise, and time-to-depth registration, with uncertainty. This Bayesian inverse method constrains possible wavelet and noise levels using well log data, seismic check-shots, and the effective bandwidth of the signal. Only parsimonious wavelets, of minimal length and fitting coefficients, are selected, which avoids excessive side-lobes or details that may fit the signal from short-path internal multiples or processing errors within the coda of the reflected signal. The noise is estimated simultaneously during the inversion process and depends on the mismatch between the synthetic seismic generated from the log derived reflectivity and the extracted wavelet against the observed seismic.

The reflectivity series, for deconvolving the source from the seismic, is modeled using linearized approximations to Zoeppritz equations. The seismic properties for parameterizing these equations are calculated from blocked and upscaled log properties (p-wave velocity, s-wave velocity, and density). The blocking of the logs is done by detection of change-points (e.g. Hawkins, 2001) within the data, and has no explicit sense of the support scale of the seismic data. The blocking scheme used in the WaveletExtraction software errors on the side of over-parameterizing the reflectivities to get the most-likely wavelets without the excessive cost of an exhaustive model selection procedure. This over-parameterization of the reflectivities provides the inversion with the potential to fit noise of similar signal to the desired data (e.g. long-path multiples or processing artifacts) and therefore gives a biased estimate of the actual noise limits. Since we are interested in finding accurate estimates of the bandwidth of the seismic signal above noise, we have designed an external blocking and model selection algorithm that interfaces with WaveletExtractor's text input and output files.

Blocking and Model Selection

In a model selection procedure, we desire a model that is of sufficient complexity to adequately describe the observed seismic data, but not so complex that it has the potential to

DOI http://dx.doi.org/10.1190/segam2013-0617.1

fit noise. Noise, in this case is any signal not consistent with the assumptions of our model – namely multiples or processing artifacts.

A formal model selection technique that considers the entire range of the number and location of reflectors is computationally prohibitive. Therefore heuristic methods such as a statistical change-point model or simply blocking at a predetermined scale (e.g. ¼ wavelength) are more commonly used. These approaches either do not have an explicit consideration of the seismic support scale, or they are not they physically consistent with seismic. A blocking method called scale-space segmentation has a direct consideration of support scale and is reconcilable with a physical model of the seismic response to reflectivity. Scale-space segmentation (Witkin 1984), or diffusion filtering (e.g. Mallat , 1999), is an edge detection method useful for recovering sharp edges in the presence of significant noise.

Scale-Space Blocking

Scale-space segmentation scales, or low pass filters, a time-series via convolution with a Gaussian wavelet, given by:

$$W(t,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}}, \text{ and } \sigma = \frac{1}{2\pi f},$$

where σ is the support scale of the wavelet, t is time, f is frequency, and μ is the center time of the wavelet. This provides a scale dependent representation of the time series to be blocked. The zero-crossing points of the second derivative with respect to time correspond to the extrema, or discontinuities, of the scaled data. The areas in-between the discontinuities correspond to smoothly varying parts of the series. As one goes up in scale to coarser and coarser representations of the data, the zero crossing points will decrease in number and drift in position. To compensate for the drift, the zero crossing points are "localized" to the originating point by following contours of the zero-crossing points from the scale of interest back to Nyquist.

The scale-space process has a seismic modeling analogy that instructs how to condition the log data for segmentation. If one takes half the natural log of seismic impedance that has been interpolated into time by a velocity function, the first derivative of the scale-space filtered impedance gives linearized reflectivity,

$$R(t) = \frac{1}{2}\frac{d}{dt}\ln(Imp),$$

and when convolved with the Gaussian wavelet produces a scaled synthetic seismic response,

$$S(t,\sigma) = W(t,\sigma) * R(t).$$

The second derivative then simply corresponds to the maxima of the salient reflections at that scale, and the localization procedure places those reflection points at their unfiltered positions, which defines block edges. The segmented blocks are then filled with a Backus average of the seismic properties extracted from logs for parameterizing the reflection series used by WaveletExtractor. Using this blocking procedure reduces the expensive task of selecting the location and number of block edges to a simple task of determining the proper scale that represents the seismic data.

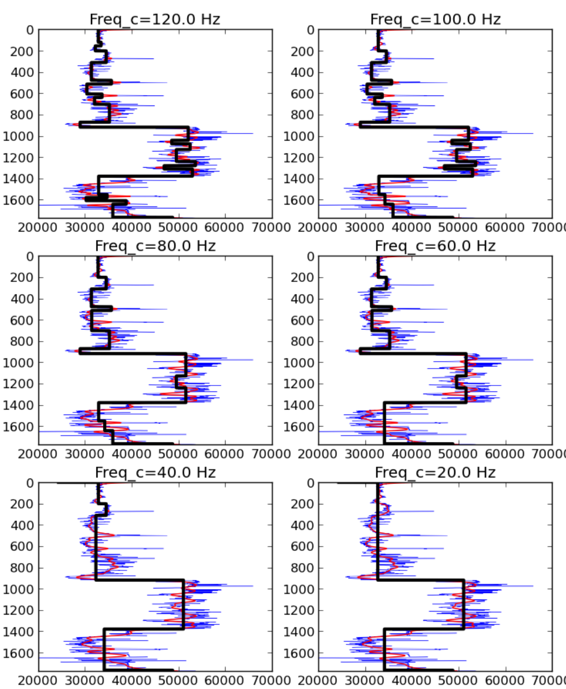

Figure 2. Scale-space blocked impedance profiles for various scales. The blue is the acoustic impedance calculated from wireline logs. The red is the scaled, or low pass filtered, version of the original data, and the black lines represent the blocks with Backus averaged acoustic impedance.

Finding the Support Scale

Starting at the highest frequencies (smallest scale of the scale-space), we run the WaveletExtractor to estimate the wavelet and noise for the corresponding upscaled and blocked data. The procedure is iterated towards the lowest frequencies while estimating the goodness of fit. The tradeoff between model complexity and fit is quantified with the Akaike information criterion (AIC) that considers both the χ^2 metric and the number of parameters

DOI http://dx.doi.org/10.1190/segam2013-0617.1

Optimal Sparse-Spike Inversion

(reflections). A plot of the AIC as a function of the support-scale cutoff frequency is given below (figure 3).

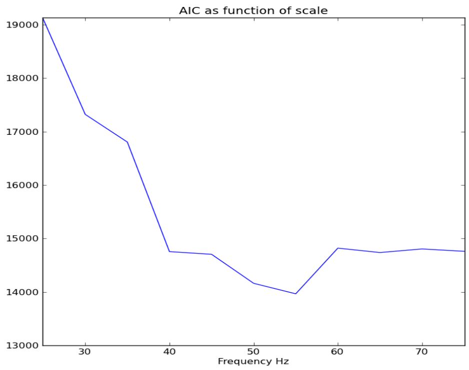

Figure 3. Akaike information criterion as function of cutoff frequency. The best trade-off was interpreted as 50Hz.

One could choose an overly conservative model that filters out all coherent noise, or conversely choose an overly parameterized model that reveals all detectable reflection with a significant number of coherent noise events. The knee point (~50Hz) effectively gives us the optimal trade-off between resolving real reflections and coherent noise.

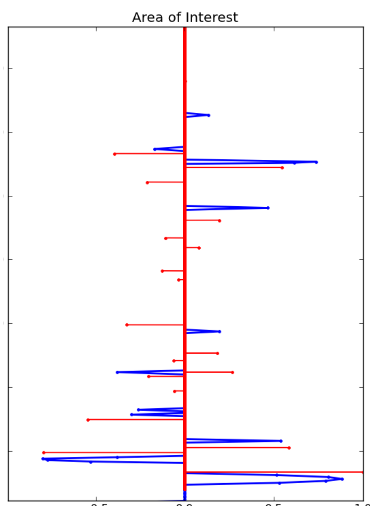

Figure 4. Plot of both the sparse-spike reflectivity (blue) and the scale-space blocked reflectivity (red).

This scale-space block reflectivity provides the model to which the sparse-spike is calibrated. The mean noise level and sparsity estimated from the WaveletExtractor, that used optimal selected model (50Hz), is used as prior information to calibrate the thresholding of the sparse-spike inversion. Comparisons between the sparse-spike and scale-space reflections show excellent correlation (Figure 4), confirming that the inversion is identifying discontinuities within the limits of the support scale of the seismic data. This supports the argument that not only is the inversion converging to global minima, but it also exhibits accuracy when the wavelet and noise have been properly estimated.

Discussion

The scale-space segmentation process and search for the appropriate support scale during wavelet extraction provides insights into the noise-thresholding sparse-spike inversion results. The sparse spike inversion method here is closely related to another denoising method commonly called wavelet shrinkage (Chambolle et al, 1998). A demonstrated equivalence between scale-space filtering and wavelet shrinkage, for the case of Haar wavelets, has been provided by Mrázek et. al (2003). This provides insight into the correlation of reflectivities derived from the scale-space blocking and the noise-thresholding sparse-spike inversion, where only a single noise-thresholding term is adjusted. Additionally, the sparse-spike algorithm from Daubachies work has been shown to have a separation resolution, the distance between two reflections above noise, that depends only on the support scale of the seismic signal above noise (Dossal and Mallat, 2005). So, the minimum separation resolution of the sparse-spike should be approximately the same as the selected scale-space blocked reflectivity series.

The blocking and model selection procedure discussed herein provides a method to evaluate how to register PP and PS seismic data together. Synthetic studies of PP and PS reflectivity show that, while not equivalent in AVO pattern, there exists a strong spatial correlation in expected reflection events. This would suggest that most events in PP and PS data could be registered if one has a good handle on the Vp/Vs ratio. However, PP and PS seismic data typically have differing effective support scales because one component (typically PS) has higher noise levels and lower bandwidth. By independently applying the blocking and model selection procedure to both PP and PS data sets, one can evaluate the reflections that are common to both scales. This provides a map of which horizons one can confidently register between the two. The sparse-spike deconvolution of the respective data sets can then be used to register these common horizons away from the well.

Conclusions

We have reviewed the application of a technically superior sparse-spike inversion technique with proven global convergence. We have demonstrated how to objectively calibrate this inversion based on a model selection procedure that uses scale-space blocking of log data.

DOI http://dx.doi.org/10.1190/segam2013-0617.1

http://dx.doi.org/10.1190/segam2013-0617.1

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2013 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Barrodale, I., and F. D. Roberts, 1973, An improved algorithm for discrete *l*1 linear approximation: SIAM Journal on Numerical Analysis, **10**, no. 5, 839–848, http://dx.doi.org/10.1137/0710069.

Chambolle, A., R. A. De Vore, N. Y. Lee, and B. J. Lucier, 1998, Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage: IEEE Transactions on Image Processing, **7**, no. 3, 319–335.

Daubechies, I., M. Defrise, and C. De Mol, 2004, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint: Communications on Pure and Applied Mathematics, **57**, no. 11, 1413–1457, http://dx.doi.org/10.1002/cpa.20042.

Debeye, H. W. J., and P. Riel, 1990, L_p-norm deconvolution: Geophysical Prospecting, **38**, no. 4, 381–403, http://dx.doi.org/10.1111/j.1365-2478.1990.tb01852.x.

Dossal, C., and S. Mallat, 2005, Sparse spike deconvolution with minimum scale: Signal Processing with Adaptive Sparse Structured Representations Workshop, 123–126.

Gunning, J., and M. E. Glinsky, 2006, Wavelet extractor: A Bayesian well-tie and wavelet extraction program: Computers & Geosciences, **32**, no. 5, 681–695, http://dx.doi.org/10.1016/j.cageo.2005.10.001.

Hawkins, D. M., 2001, Fitting multiple change-point models to data: Computational Statistics & Data Analysis, **37**, no. 3, 323–341, http://dx.doi.org/10.1016/S0167-9473(00)00068-2.

Mallat, S., 1999, A wavelet tour of signal processing: Academic Press.

Mrázek, P., J. Weickert, and G. Steidl, 2003, Correspondences between wavelet shrinkage and nonlinear diffusion, *in* L. D. Griffin and M. Lillholm, eds., Scale-space methods in computer vision: Springer-Verlag, 101–116.

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: SIAM.

Witkin, A., 1984, Scale-space filtering: A new approach to multi-scale description: International Conference on Acoustics, Speech, and Signal Processing, IEEE, vol. 9, 150–153.

DOI http://dx.doi.org/10.1190/segam2013-0617.1