A pre-stack seismic inversion with L1 constraints and uncertainty estimation using the expectation maximization algorithm
Douglas S. Sassen* and Ben Lasscock, ION Geophysical Corp.

Summary
We present an algorithm that provides improved Earth model reconstruction over Gaussian, or Tikhonov, regulated inversions. By utilizing a dynamic noise estimation, linear error propagation theory and a Bayesian construction this algorithm can: objectively indicate the amount of recoverable independent information, provide a framework on how to weight prior information, and to communicate the uncertainty of the estimates.

Introduction
The distribution of Earth reflectivity has been shown to depart significantly from a Gaussian distribution (Walden and Hosken, 1986, Velis, 2003, Painter et al., 1995). The findings of Painter et al. (1995) are that seismic reflectivity is best represented by a Levy-stable distribution with slowly decaying tails, and Walden and Hoskens (1986) found that mixtures of Laplace distributions are most efficient description of the distribution of reflectivity. These distributions motivate the use of priors other than an L2, or Gaussian, distributions including an L1, or Laplacian, and mixtures of L1 and L2 prior distributions on the Earth model within this paper. Inspired by these insights, a common objective is to inverse model, or deconvolve, seismic data to recover the discontinuities of the Earth, including using sparse-spike inversion (e.g. Debye and Van Reil, 1990, and Barrodale and Roberts, 1973, Taylor, et al., 1979, Perez et al., 2013, Sacchi, 1997) and related basis pursuit methods (Zhang, and Castagna, 2011). 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. We are extending an iterative soft-thresholding algorithm (ISTA), which can be understood as a case of the expectation maximization (EM) algorithm, for solving linear inverse problem with a sparsity constraint (Daubechies et. al. 2003 and Figueiredo and Nowak 2003). This method had first been applied to seismic data by Dossal and Mallat (2005), and an implementation to 2-term prestack seismic inversion was introduced by Perez et al. (2013) using the fast iterative soft-thresholding algorithm (FISTA).

Theory
Our objective then, in pre-stack seismic inversion, is to find the maximum likelihood $L(\theta)$ of properties of the Earth $\theta$ (e.g. Zp, Zs, Vp, Vs, and density) from our observed seismic data $y$. However, the model most often employed for this objective connects the observed seismic $y$ to estimate properties of the Earth $\theta$ with an unobserved reflectivity $z$, depending on $\theta$, and convolved with a presumably known seismic wavelet. $L(\theta; y, z) = p(y, z|\theta)$

We can significantly simplify this problem if we could marginalize out the latent variable of the reflectivity $z$.

$$p(y|\theta) = \int p(y, z|\theta) dz = \int p(y|z, \theta)p(z|\theta)dz$$

In the case of seismic inversion, the presence of noise and inaccurate estimates of the wavelet lead to an inconsistent system of equations for estimating likelihood, where at least some of the properties of interest cannot be found uniquely. To address this problem, prior information can be introduced to help constrain the estimates of our Earth properties to physically defensible bounds. This prior information can be introduced with Bayes theorem, where the marginalized posterior probability from the marginalized likelihood and prior constraint is:

$$p(\theta|y) \propto p(y|\theta)p(\theta).$$

However, to marginalize out the latent variable of the reflectivity $z$ is analytically intractable under certain conditions such as including a sparsity constraint (L1 prior). The Expectation Maximization (EM) algorithm is a method for estimating this integral.

We extend the EM algorithm of Figueiredo and Nowak (2003) for the special problem of prestack seismic inversion. The Figueiredo and Nowak (2003) EM algorithm was originally developed for simultaneous image deblurring and denoising, utilizing the respective strengths of deconvolution for debluring and wavelet thresholding for denoising. This application applies to processing 2-D images; by analogy, we apply the same technique in a 2-D space where the dimensions are time and offset. We will provide a linkage between the two applications and then direct the reader to the work Figueiredo and Nowak (2003) for implementation details.

The basic forward problem for pre-stack seismic inversion is given by:
EM Algorithm for Pre-stack Seismic Inversion

\( y = H x + n \), where \( x = A \theta \).

where \( y \) is the observed data in the time/offset domain (vectorized), \( H \) is the seismic wavelet(s), \( x \) is the true model of Earth reflectivity, \( n \) the noise, \( A \) is our design matrix from the linearization of Zoeppritz equations, and \( \theta \) are the Earth properties. Figueiredo and Nowak (2003) provided an insight to make this solution to this problem more tractable by decomposing the problem into surrogate functions:

\[
\begin{align*}
\{ & z = A \theta + \alpha n_1 \\
& y = Hz + n_2 \}
\end{align*}
\]

where \( z \) is the hidden model of Earth reflectivity, \( n \) is white noise, \( \alpha \) is a normalization by the principle eigenvalue of \( H \), and \( \theta \) are the coefficients of our design matrix. Our objective is to estimate the maximum likelihood of \( \theta \) via the EM algorithm. The EM algorithm is a two-step process that first calculates the expectation and then calculates the maximum likelihood of the first step. An efficient frequency domain representation of the expectation step exists provided the matrix \( H \) is block-circulant (which is true for the wavelet convolution matrix) and an efficient maximization step is possible if \( A \) is orthogonal.

Extensions to Algorithm

Unlike the original derivation, our design matrix \( A \) for seismic reflectivity is not orthogonal. However, we can transform our design matrix into an orthogonal representation utilizing singular value decomposition,

\( A = U \Sigma V^T \).

Daubechies, et al., 2003, provide examples for enforcing sparsity under orthogonal expansions, including a thresholded SVD expansion for the L1 case. The thresholded SVD is given by:

\[
\theta^{(t+1)} = \sum_{k=t}^{\text{max}} \frac{1}{\alpha_k} S_l(\sigma_k, u_k, v_k)
\]

where \( S_l \) is the soft thresholding rule, \( u_k \) are the eigenvectors in the matrix \( U^T \), \( \sigma_k \) are eigenvalues, and \( v_k \) are the eigenvectors of \( V \). This approach has the advantage over traditional linear cut-off in that there may be significant coefficients in our seismic inversion beyond the traditional cutoff that benefit from a soft-thresholding (Daubechies, et al. 2003). It is important to realize that the thresholding of coefficients in the orthogonal domain (\( \theta' \)) is more than just a mathematical convenience. Our prior assumption then is that the drivers of variability in this basis should be sparse. As a consequence, coefficients of \( \theta' \) transformed back into the original reflectivity domain \( x \) will be sparse, and will always give physical results according to linearized Zoeppritz approximations. Opening the way for the application the sparse constraint to pre-stack data AVO and AVAz. An interesting by-product of this approach is that, when viewed in the orthogonal representation, the amount of recoverable independent information beyond noise (i.e. the number and location of the eigencoefficients) can be found. This is beneficial in that it allows one to inspect the rank of a seismic inversion results on each interface.

Figure 1: Synthetic seismic gathers generated using a 2 term linearized Zoeppritz forward model parameterized with intercept (top left) and gradient (top middle) and convolved with 30 Hz Ricker wavelet (top right). The bottom panels add Gaussian random noise at varying signal to noise ratio (20 to 10 – left to right) used in the synthetic testing (after Perez and Sacchi, 2013).

Figure 2: The actual (left) recovered spike series following the 2-term L1 penalized EM inversion for increasing noise levels (left to right) with the intercept coefficients on the top and the gradient coefficients on the bottom. Note the subtle increase in artifacts and decreased accuracy of smaller coefficients with increased noise. Opening the way for the application the sparse constraint to pre-stack data AVO and AVAz. An interesting by-product of this approach is that, when viewed in the orthogonal representation, the amount of recoverable independent information beyond noise (i.e. the number and location of the eigencoefficients) can be found. This is beneficial in that it allows one to inspect the rank of a seismic inversion results on each interface.
EM Algorithm for Pre-stack Seismic Inversion

Error estimation and propagation

Standard techniques for error propagation in the L1 regulated problems is generally considered to be somewhat ill-posed, see Osborne and Presnell (2000), this is a consequence of both the thresholding (to zero) of parameters and the fact that the L1 penalty is not differentiable everywhere. Monte Carlo techniques would be one possible work around, see Park and Casella (2008) for a Gibbs sampling algorithm. However, the computational cost of this technique is currently prohibitive for our applications. Instead we provide a reasonable estimate based on the observed Fisher information computed by the Hessian of the likelihood function evaluated at the (maximum a posteriori) MAP estimate.

Incorporating rock physics correlations and trends

Physically we do expect that perturbations in the elastic properties should be significantly correlated. These correlations introduces a known limitation in the existing L1 regulated approaches, see Zou and Hastie (2005), and requires an extension to the model to compensate. Our solution is to extend the algorithm of Figueiredo and Nowak (2003) in a manner analogous to the elastic net regularization of Zou and Hastie (2005). This extension takes rock physics control, which comes in the form of a Gaussian prior distribution, and we incorporate it in the computation as an additional L2 penalty to the existing EM update. This has the effect of regulating the variables $\theta$ (e.g. $\Delta V_p$, $\Delta V_s$, $\Delta \rho$) that are dependent upon smaller eigenvalues and keeping the results within observed physical bounds.

To estimate properties in absolute terms, we need to move from relative changes to absolute changes and provide estimates of the missing low frequency trends. We combine the MAP estimate of perturbations ($\Delta V_p$, $\Delta V_s$, $\Delta \rho$) with a background field, calibrated by well logs, to determine the absolute elastic properties, as shown in Figs. 3 and 4. Since the low frequency background model is uninformative to the high frequency features, we treat this step as a post-process to the optimization discussed in the previous sections. Estimates of uncertainty in the elastic properties are computed using standard linear error propagation.

Examples

We demonstrate the application of this algorithm on synthetic for a two-term inversion using the L1 penalized version (figures 1 and 2), and on field data for three-term P-wave and utilizing the L1-L2 hybrid method (figures 3 and 4). The synthetic data (Figure 1) is calculated using the same linearized forward model for reflectivity and Gaussian random noise of varying SNR is applied to see the effects on the accuracy of the sparse-spike recovery (Figure 2) (after Perez and Sacchi, 2013).

The field data is from an unconventional play in the Anadarko basin, Oklahoma, USA. The field data was PSTM migrated and is pre-conditioned with residual flattening and conservative denoising. The 5-angles stacks within the gathers of the field range from 9 to 41 degrees. The data line selected is intersected by a well log for cross-validation, and exhibits some acquisition and processing issues that exemplify the uncertainty estimation. From the wavelets extraction an initial estimate of the noise used in the inversion are estimated prior to inversion utilizing the open source software waveletExtractor from the Delivery inversion software (Gunning and Glinsky, 2006). The inversion does not utilize a high frequency background model, nor lateral constraints, and is applied on a gather by gather basis. The trade-off term controlling sparsity was calibrated using the Bayesian Information Criterion (Sassen and Glinsky, 2013).

The eigencoefficients estimated from the inversion (not shown here) indicated that, for estimated noise levels and available angles, only two independent coefficients are recoverable. With only two constraints and three unknowns ($V_p$, $V_s$ and density) the problem is underdetermined and will provide poor estimates for properties not well constrained by the seismic and prior information. The effects of the underdetermined system are exemplified in the cross-validation of the inversion against the well data. There the p-wave and s-wave velocity estimates and 95% confidence intervals match the upscaled logs well, while the density estimates are correlated to the velocity estimates rather than the logs and the confidence interval ranges from 2.2 to 3.0 g/cc.

Discussion and Conclusions

We have demonstrated an algorithm that provides improved Earth model reconstruction over Gaussian, or Tikhonov, regulated inversions. However, sparse-spike algorithms, including those based on basis pursuit, have been shown to have a separation resolution, the distance between two reflections above noise, that depends on the support scale, or bandwidth, of the seismic signal above noise (Dossal and Mallat, 2005). By utilizing a dynamic noise estimation, linear error propagation theory and a Bayesian construction this algorithm can: objectively indicate the amount of recoverable independent information, how to weight prior information and communicate the uncertainty of the estimates.
Figure 3. Results of the 3-term L1-L2 penalized form of the EM inversion. The top panel is Vp (ft/s), the middle is Vs (ft/s), and the bottom is density (g/cc).

Figure 4. Cross validation of the inversion results (red) with upscaled well logs (black lines), along with the 95% confidence interval (pink), the low frequency model (green) and the original logs (light gray). Note the mean and 95% confidence interval for the Vp and Vs match the observed upscaled logs well, while the density is poorly identified by the seismic data.
REFERENCES


