Wavelet domain linear inversion with application to well logging

Jonathan Kane
Felix J. Herrmann
M. Nafi Töksoz
Earth Resources Laboratory, EAPS, MIT

May 11, 2001

Abstract

Solving linear inversion problems in geophysics is a major challenge when dealing with non-stationary data. Certain non-stationary data sets can be shown to lie in Besov function spaces and are characterized by their smoothness (differentiability) and two other parameters. This information can be input into an inverse problem by posing the problem in the wavelet domain. Contrary to Fourier transforms, wavelets form an unconditional basis for Besov spaces, allowing for a new generation of linear inversion schemes which incorporate smoothness information more precisely. As an example inversion is performed on smoothed and subsampled well log data.

1 Introduction

Multiple length scale variability in earth properties is a predominant characteristic of sedimentary basins. Sedimentary records (well-log data) display variability on length scales ranging from sub-centimeter to many kilometers. Traditional approaches in linear inversion (such as least squares deconvolution or interpolation) are often based on certain homogeneity (stochastic stationarity) assumptions, and may encounter difficulties when evident time/length scale complexity has long range correlations at many scales.

Multifractal analysis shows (Herrmann, 1998) that traditional assumptions about geophysical data, such as that it lies in an L_2 function space, are incorrect. This criterion has been historically used due to its mathematical ease of derivation and manipulation. We will discuss how to move beyond the traditional assumptions of data smoothness and stationarity in order to better characterize the visible complexity and incorporate this information into a linear inversion scheme for estimating unknown geophysical parameters. This involves posing the inverse problem in the wavelet domain.

The set up is as follows. First we will briefly review the discrete wavelet transform and its relation to Besov spaces. We then describe the formulation of a linear inverse problem in the wavelet domain in which we generalize the traditional L_2-norm assumptions. Finally we apply the method to the simple inverse problem of inverting the effects of a well logging tool.

2 The Discrete Wavelet Transform

To implement our methods on a computer we need to deal with finite dimensional data sets. Thus it is necessary to review the discrete wavelet transform (DWT). We follow the exposition in (Choi and Baraniuk, 1999). The DWT represents a 1-D continuous function $f(x)$ as inner products of the function with shifted versions of a lowpass scaling function $\phi(x)$ and shifted and dilated versions of a bandpass wavelet function $\psi(x)$. For appropriate choices of ϕ and ψ the functions $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$, and $\psi(x) = 2^{j/2}\psi(2^j x - k)$, $j, k \in \mathcal{Z}$, form an unconditional orthonormal basis in a Besov space (see below) and we have:

$$f(x) = \sum_k u_{j_0,k}\phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty}\sum_k w_{j,k}\psi_{j,k}(x)$$

where

$$u_{j_0,k} = \int f(x)\phi^*_{j_0,k}(x)dx, \quad w_{j,k} = \int f(x)\psi^*_{j,k}(x)dx.$$

In reality the input to the DWT will be a discrete finite length column vector $\mathbf{f}$ instead of a function $f(x)$. The DWT can still be performed but the analyzing scaling and wavelet functions become finite length vectors. We assume that the vector $\mathbf{f}$ to be analyzed has been prefiltered (Choi and Baraniuk, 1999). We can then write the DWT as a matrix vector product

$$\mathbf{f} = \begin{bmatrix} \Psi\mathbf{w} \\ \Phi\mathbf{u} \end{bmatrix}. \tag{1}$$

Where the columns of Φ are the shifted discrete scaling vectors and the columns of Ψ are the shifted discrete wavelet vectors.

We can combine Φ and Ψ into one matrix $\mathbf{W}^{-1}$ and $\mathbf{u}$ and $\mathbf{w}$ into one vector $\tilde{\mathbf{f}}$. Due to the finite length of the vector $\mathbf{f}$ we must handle edge effects in the DWT just as in the Fourier transform. The method we use is periodization (Strang and Nguyen, 1997). We refer to the matrix $\mathbf{W}^{-1}$ as the *inverse* DWT. Its inverse is $\mathbf{W}$, the *forward* DWT.

2.1 Fractional Splines

Choosing which ϕ and ψ to use is an area of active research. In this work we deal exclusively with *fractional spline* wavelets and scaling functions (Unser and Blu, 2000). Fractional splines are piecewise continuous polynomials of degree α. α is the degree of Hölder continuity, that is the number of bounded fractional derivatives of the function. The α^{th} derivative of a fractional spline yields a function that is piecewise discontinuous with discontinuities at the locations where the piecewise polynomials meet. These points are called *knots* and are usually constrained to lie at the data values of a regularly sampled function. A classical result of integer degree splines is that they are representable in a basis of B-splines, $\beta^\alpha(x)$, which are bell shaped functions of compact support. They can be causal, anti-causal, or non-causal (symmetric about the origin) functions. The B-spline of degree 0 is the boxcar function. All higher, integer degree splines result from multiple convolutions of the boxcar with itself. B-splines are most simply characterized in the Fourier domain where, in the symmetric B-splines case, we have:

$$\hat{\beta}^\alpha(e^{i\omega}) = \left|\frac{sin(\frac{\omega}{2})}{\frac{\omega}{2}}\right|^{\alpha+1}$$

It can be shown that $\beta^\alpha(x) \in L_2$ for $-1/2 < \alpha$ and $\beta^\alpha(x) \in L_1$ for $-1 < \alpha$. More generally, $\beta^\alpha(x) \in H_2^r$ where H_2^r is a Sobolev function space (defined below) and $r < \alpha + 1/2$.

The most simple way to obtain the refinement filter for fractional splines is by calculating the ratio $2\hat{\beta}^\alpha(e^{i2\omega})/\hat{\beta}^\alpha(e^{i\omega})$. This results in:

$$\hat{h}^\alpha(e^{i\omega}) = \sqrt{2}\left|\frac{1+e^{-i\omega}}{2}\right|^{\alpha+1}.$$

In this work we desire our wavelet transform to be orthogonal for reasons outlined below. Thus we need to orthogonalize the above filter. This is done by the following formula:

$$\hat{h}^\alpha_\perp(e^{i\omega}) = \hat{h}^\alpha(e^{i\omega})\sqrt{\frac{\hat{a}^{2\alpha+1}(e^{i\omega})}{\hat{a}^{2\alpha+1}(e^{2i\omega})}}$$

where

$$\hat{a}^{2\alpha+1}(e^{i\omega}) = |\hat{\beta}^\alpha(e^{i\omega})|^2$$

We then compute the wavelet filter by the standard method of "alternating flip" construction Strang and Nguyen (1997):

$$\hat{g}_{\perp}^{\alpha}(e^{i\omega}) = e^{-i\omega}\hat{h}_{\perp}^{\alpha}(-e^{-i\omega})$$

This defines the filter to be applied recursively on the function to be transformed. Filters resulting from wavelets with compact support can be applied more quickly in the state domain by explicit convolution and downsampling. The orthogonal fractional spline filters are IIR and thus must be computed in the Fourier domain. This does not present much of a problem and is still a fast computation thanks to the fast Fourier transform. For an analysis of the computational costs of the fractional spline wavelet transform, see Blu and Unser (2001).

At this point we would like to remind the reader of the connection between wavelet transform defined as the above filter and the wavelet transform defined as a matrix in equation 1. When analyzing a signal of length N, the first $N/2$ rows of $\mathbf{W}$ are the double shifted, state domain filters, $g_{\perp}^{\alpha}(x) = \int \hat{g}_{\perp}^{\alpha}(e^{i\omega})e^{i\omega x}d\omega$. The next $N/4$ rows are $\hat{g}_{\perp}^{\alpha}(x)$ dilated and quadruple shifted. This pattern extends down to the last row which is a single dilated scaling filter.

3 The Covariance Structure of Fractional Splines

If we add further structure to the fractional splines by declaring them to be random vectors in some Hilbert space we can talk about covariances. Fractional splines wavelets can be constructed such that the wavelet transform acts as a *Karhunen-Loeve* transform for fractional splines with a specified α, i.e. $\mathbf{W}$ diagonalizes the covariance matrix of a vector belonging to the Hilbert space.

The best way to illustrate this point is to represent a random fractional spline, $\mathbf{f}$, as filtered white noise. The filter in this case is the inverse wavelet transform:

$$\mathbf{f} = \mathbf{W}^{-1}\mathbf{D}\mathbf{n}$$

where $\mathbf{D}$ is a diagonal matrix (defined below) that weights the white noise, $\mathbf{n}$, by scale. Random fractional splines wavelets have indeed been used exactly for this purpose. In Blu and Unser (2001) Blu and Unser use this formulation to approximately reproduce fractional brownian motions. This shows an intimate connection between fractional splines and fractional brownian motions.

We next calulate the expected value of the outer product of the fractional splines to obtain the covariance matrix:

$$E\left[\mathbf{f}\mathbf{f}^{T}\right] = E\left[\mathbf{W}^{-1}\mathbf{D}\mathbf{n}\mathbf{n}^{T}\mathbf{D}\mathbf{W}^{-T}\right] = \mathbf{W}^{-1}\mathbf{D}E\left[\mathbf{n}\mathbf{n}^{T}\right]\mathbf{D}\mathbf{W}^{-T}$$

$$= \mathbf{C_f} = \mathbf{W}^{-1}\mathbf{D}^{2}\mathbf{W}^{-T}$$

or

$$\mathbf{D}^{2} = \mathbf{W}\mathbf{C_f}\mathbf{W}^{T}, \tag{2}$$

This result tells us that there are a class of random processes which are 100% whitened by a wavelet transform. What these processes are is another question. They may be quasi-fractional brownian motions (a sort of band-limited variety) in the case of fractional splines wavelets. Further investigation is needed to fully characterize them.

4 Besov Spaces

A function space is a set (collection) of functions. Such a set is constructed by defining a norm (magnitude) over all possible functions. If this norm is finite for a particular function we say that the function belongs to that function space. The most common norms used are the L_p norms:

$$\|f(x)\|_{L_p} = \left[\int |f(x)|^{p}\right]^{1/p}$$

or l_p norms, if finite vectors are dealt with:

$$\|\mathbf{f}\|_{l_p} = \left[\sum_j |f_j|^p\right]^{1/p}.$$

In this work we characterize a function as belonging to a particular *Besov* space, b_p^β. We choose this space because it contains many realistic functions, such as well data, and it is parameterized by only three numbers: p, q, and β. p describes the distribution of energy in the wavelet coefficients at a particular scale, q describes energy distribution over all scales, and β roughly tells the number of derivatives that also have finite norm. The relationship β and the α parameter of fractional splines is $\beta = \alpha + 1/2$. An equivalent Besov sequence norm is defined in the wavelet domain as Choi and Baraniuk (1999):

$$\|f(x)\|_{b_{p,q}^\beta} \equiv \left[\sum_k |u_{j_0,k}|^p\right]^{1/p} + \left[\sum_{j\geq j_0}^{J}\left[\sum_k |\bar{\sigma_j} 2^{j(\beta+1/2-1/p)} w_{j,k}|^p\right]^{q/p}\right]^{1/q}.$$

$w_{j,k}$ and $u_{j_0,k}$ are the components of the vectors $\mathbf{u}$ and $\mathbf{w}$ defined above. $\bar{\sigma_j}$ is a normalizing constant such that the vector $\mathbf{f}$ has the correct variance. Besov spaces offer perhaps the most general description of functions. It is for this reason that we choose them when performing inversion in the next section.

We can also write equation ** in matrix notation:

$$\|\mathbf{n}\|_{b_{p,q}^\beta(l_p)} = \left[\mathbf{1}^T|\mathbf{u}|^p\right]^{1/p} + \left[\mathbf{1}^T \mid \mathbf{V}|\mathbf{D}\mathbf{w}|^p \mid^{p/q}\right]^{1/q}.$$

The notation $|\cdot|^p$ corresponds to exponentiating each component of a vector to the p^{th} power. The $\mathbf{D}$ matrix consists of the $\bar{\sigma_j} 2^{j(\beta+1/2-1/p)}$ terms along the diagonal and zeros everywhere else. The matrix $\mathbf{V}$ has as each row a vector $\mathbf{v}_j$. These are vectors comprised of ones at the location of the wavelet coefficients in $\tilde{\mathbf{f}}$ corresponding to scale j, and zeros everywhere else. The $\mathbf{v}_j$ vectors have the effect of summing the wavelet coefficients at a particular scale. The $\mathbf{1}^T$ vector performs the action of summing all components of the vector it operates on.

5 Linear inversion

We let $\mathbf{f}$ represent the model parameters we wish to estimate. Let $\mathbf{d}$ represent observed values of this data. We have a linear forward modelling operator in the form of a matrix, $\mathbf{P}$, that operates on $\mathbf{f}$. There is also possibly noise, $\mathbf{n}$, in the experiment. We have the following relationship:

$$\mathbf{d} = \mathbf{P}\mathbf{f} + \mathbf{n}. \tag{3}$$

The solution of an inverse problem involves minimizing

$$\|\mathbf{n}\|_{b_{p,q}^\beta} = \|(\mathbf{d} - \mathbf{P}\mathbf{f})\|_{b_{p,q}^\beta} \tag{4}$$

over all possible values of $\mathbf{f}$.

In many inverse problems, there is insufficient information in the data to allow a unique solution, i.e., there are an infinite number of solution vectors $\mathbf{f}$ that satisfy equation 4. In order for there to be a unique solution we must provide some other constraint to the problem. This usually comes in the form of a smoothness constraint in which we minimize the norm of the model vector in some function space:

$$\|\mathbf{f}\|_{b_{p',q'}^{\beta'}}. \tag{5}$$

In order to obtain a solution to an inverse problem we must find the model vector, $\hat{\mathbf{f}}$ that simultaneously minimizes a norm on the noise vector and the unknown model:

$$\hat{\mathbf{f}} = min\left[\|(\mathbf{d} - \mathbf{P}\mathbf{f})\|_{b_{p,q}^\beta} + \|\mathbf{f}\|_{b_{p',q'}^{\beta'}}\right].$$

In transforming the inverse problem to the wavelet domain we change equation 3 to

$$\mathbf{d} = \mathbf{P}\mathbf{W}^{-1}\mathbf{W}\mathbf{f} + \mathbf{n},$$

or,

$$\mathbf{d} = \tilde{\mathbf{P}}\tilde{\mathbf{f}} + \mathbf{n}$$

where $\tilde{\mathbf{P}} = \mathbf{P}\mathbf{W}^{-1}$ is the wavelet transform of the each row of the matrix $\mathbf{P}$ and $\tilde{\mathbf{f}} = \mathbf{W}\mathbf{f}$. This expresses the forward problem in terms of the wavelet coefficients. We then need to redefine the minimization in eqn. 5 in terms of the wavelet coefficients. This would normally present a problem because the Besov norm in the wavelet domain is *equivalent* to the norm of the function in the state domain, not *equal* to it. Thus $A\|\mathbf{f}\|_{b^{\beta}_{p,q}}\| \leq \|\tilde{\mathbf{f}}\|_{b^{\beta'}_{p',q'}}\| \leq B\|\mathbf{f}\|_{b^{\beta}_{p,q}}$ for some constants A and B. Only in the case of an orthogonal transform is the energy in both domains equal. This is the reason we chose to use orthogonal fractional spline wavelets. We then have

$$\hat{\mathbf{f}} = min\left[\|(\mathbf{d} - \mathbf{P}\mathbf{f})\|_{b^{\beta}_{p,q}} + \|\tilde{\mathbf{f}}\|_{b^{\beta'}_{p',q'}}\right]. \tag{6}$$

The simplest and easiest inverse methods assume that $p = q = 2$. In this special case the Besov spaces reduce to simpler Sobolev spaces. It is much easier to perform inversion in Sobolev spaces because we can use standard least squares methods. For this reason we use these values of p and q to illustrate the inversion. We also assume that the measurement noise is white and thus set $\beta = -0.5$ in equation 6. Thus the measurement noise, $\mathbf{n}$, lies in $b^{-0.5}_{2,2}$.

We make the same assumption for the vector $\mathbf{f}$ itself. We assign $p' = q' = 2$ also. The value used for β' is described below.

5.1 A Statistical Viewpoint

We have defined our fractional splines in the wavelet domain by norms on the wavelet coefficients. We have also declared out fractional splines to be discrete random vectors. We now have to connect these two concepts. That means defining the probablistic structure of the vectors.

In setting $p = p' = q = q' = 2$ we have declared our functions to be in the Sobolev-2 space, H^{β}_{2}. We have also declared them to be Gaussian. The probability distribution on the wavelet coefficients is then:

$$p(\tilde{\mathbf{f}}) = \frac{1}{(2\pi)^{\frac{N}{2}}|\mathbf{D^2}|^{\frac{1}{2}}} exp[-\frac{1}{2}(\tilde{\mathbf{f}})^T\mathbf{D}^{-2}(\tilde{\mathbf{f}})] \tag{7}$$

The exponent of this expression is in fact the Sobolev-2 norm.

The distribution of the measurment noise was also defined to lie in a Sobolev-2 splace and has the following Gaussian distribution:

$$p(\mathbf{d}|\tilde{\mathbf{f}}) = \frac{1}{(2\pi)^{\frac{N}{2}}} exp[-\frac{1}{2}(\mathbf{d} - \tilde{\mathbf{P}}\tilde{\mathbf{f}})^T\mathbf{C}_{\mathbf{n}}^{-1}(\mathbf{d} - \tilde{\mathbf{P}}\tilde{\mathbf{f}})]. \tag{8}$$

where $\mathbf{C}_{\mathbf{n}}^{-1} = 1/\sigma_{\mathbf{n}}^2\mathbf{I}$ is the covariance function of the measurement noise.

Equation 8 is often called the *likelihood* function. Equation 7 is called the *prior*. Standard Bayesian inverse theory then defines the *posterior* probability distribution as

$$p(\tilde{\mathbf{f}}|\mathbf{d}) = k\ p(\mathbf{d}|\tilde{\mathbf{f}})p(\tilde{\mathbf{f}}) \tag{9}$$

Maximizing the probability in equation 9 corresponds to minimizing its exponent, which is the Besov norm in equation 6.

6 Application to well logging

We shall now show an application in which we attempt to undo the bandlimiting effect that a logging tool has on a parameter field it is sampling. The parameter field is the wave velocity of the rock formation. In this case the operator $\mathbf{P}$ above is a cascade of two other operators $\mathbf{P}_1$ and $\mathbf{P}_2$. $\mathbf{P}_1$ is a subset of the rows of the identity matrix that picks rows out the vector it operates on. $\mathbf{P}_2$ is a convolution matrix - a constant diagonal Toeplitz matrix which low-pass filters the vector it operates on. The matrix $\mathbf{P} = \mathbf{P}_1\mathbf{P}_2$ is then a rectangular matrix comprised of a subset of the rows of the convolution matrix. The cascade of these two operator mimics the effect of a logging tool that continuously measures velocity in a bore hole with an array of sensors and then samples at regular time intervals.

In figure 1 we first show a typical velocity well log sampled at $0.5ft.$ intervals in the earth over roughly $500ft.$. We will assume this to be the "true" velocity vector and then then apply our logging tool to smooth and sample it. Our measured response of the velocity is shown in figure 2. The variance of the noise vector in this case is $\sigma_{\mathbf{n}}^2 = 100ft.^2/s^2$. This is quite small and means that we assume our measuring tool to be accurate and noise free.

In order for inversion to proceed we need prior knowledge of the β' parameter in equation 9 that determines the velocity field's smoothness. From previous analysis of similar wells we know that the exponent lies around $\beta' = 0.1$. Using this in the inversion we obtain the results in figure 3. As can be seen, a great deal of the original detail can be reconstructed.

In essence, what has been done here is a joint deconvolution and interpolation. Deconvolution in that the bandlimiting effects of a measurement have been undone. Interpolation, because the data was subsampled and we were required to interpolate at unknown locations. These results foreshadow the application of this method to least-squares deconvolution of seismic traces and interpolation of well logs at interwell locations.

7 Conclusions

Formulating a linear inverse problem in the wavelet domain via fractional splines provides a flexible and easy way to incorporate smoothness information into an inverse problem. By using a correct fractional spline wavelet, the wavelet transform whitens a random vector in a particular Besov space $b^{\beta}_{p,q}$. The results of the inversion on a simple problem are promising and future work will concentrate on least-squares seismic deconvolution and interpolation.

8 Acknowledgements

We would like to thank Christie Callender, Texaco, and Geco Prakla for the well data used in this work. We also thank Michael Unser and Thierry Blu for use of their fractional spline freeware which can be found on the web at http://bigwww.epfl.ch/.

References

Thierry Blu and Michael Unser. The fractional spline wavelet transform: Definition and implementation. Technical report, Swiss Federal Institute of Technology EPFL, 2001.

H. Choi and R. Baraniuk. Interpolation and denoising of nonuniformly sampled data using wavelet-domain processing. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP'99*, 1999.

F.J. Herrmann. Multiscale analysis of well and seismic data. In Siamak Hassanzadeh, editor, *Mathematical Methods in Geophysical Imaging V*, volume 3453, pages 180–208. SPIE, 1998.

Gilbert Strang and Truong Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, 1997.

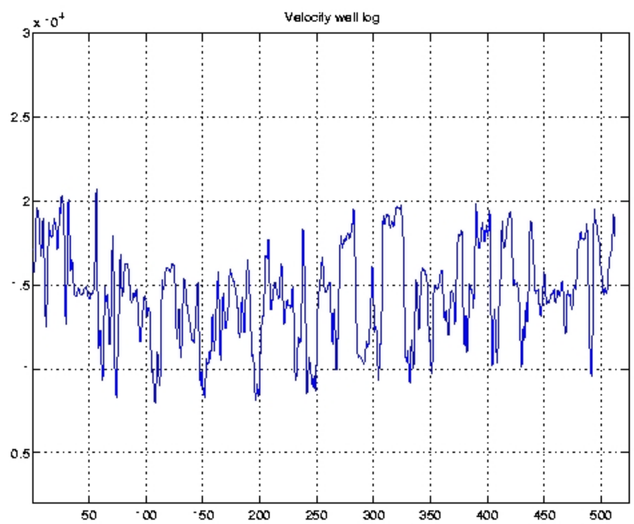

Figure 1: Velocity well log

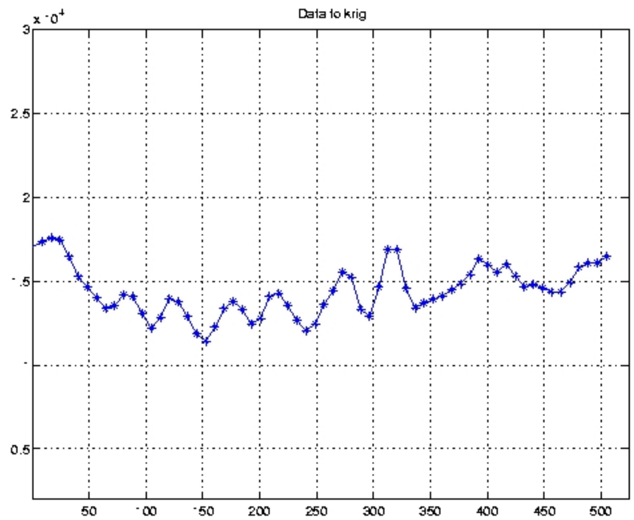

Figure 2: Smoothed and sampled well data

Michael Unser and Thierry Blu. Fractional splines and wavelets. *SIAM Review*, 42(1):43–67, 2000.

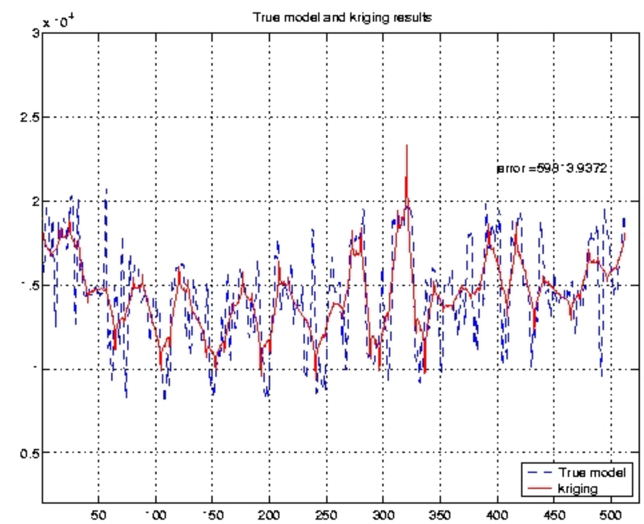

Figure 3: Reconstructed well log