Log domain blind decon of echo seismograms

Jon Claerbout, Qiang Fu, Antoine Guitton, and Yang Zhang

SUMMARY

Motivated by a desire for sparse reflectivity models and the need to handle correctly the Ricker wavelet, our earlier work (REF: YANG EAGE PAPER) on bidirectional blind deconvolution showed that seismic polarity can become a more visible aspect of routine seismic data display. The log parameterization introduced here ensures compact wavelets and produces better results. Another innovation here is correctly handling time-variable gain. Since filtering does not commute with TV gain, gain is now done after filtering (not before). An intriguing theoretical aspect shows that log spectral parameterization links penalty functions to crosscorrelation statistics of outputs.

SETTING

In the long-standing problem of blind deconvolution we replace the traditional unknown filter coefficients with lag coefficients u_t in the log spectrum of the deconvolution filter. Given data $D(\omega)$, the deconvolved output is

$$r_t \;=\; \mathrm{FT}^{-1}\, D(\omega)\ \exp\left(\sum_t u_t Z^t\right) \tag{1}$$

where $Z = e^{i\omega}$. The log variables u_t transform the linear least squares (ℓ_2) problem to a non-linear one that requires iteration. Losing the linearity is potentially a big loss, but we lost that at the outset when we first realized we needed to deal with the non-minimum phase Ricker wavelet. We find convergence is rapid, the principal difficulty being the need to attend to a suitable starting estimate.

The source wavelet, inverse to the decon filter above, corresponds to $-u_t$. The positive lag coefficients in u_t correspond to a causal minimum phase wavelet. The the negative lag coefficients correspond to an anticausal filter.

We introduce the complication that seismic data is non-stationary requiring a time variable gain g_t. The deconvolved data is the residual r_t. The gained residual $q_t = g_t r_t$ is "sparsifed" (REFERENCE ELITA GEOPHYSICS PAPER) by minimizing $\sum_t H(q_t)$ where

$$q_t \quad = \quad g_t\ r_t \tag{2}$$

$$H(q_t) \quad = \quad \sqrt{q_t^2+1}-1 \tag{3}$$

$$\frac{dH}{dq} \;=\; H'(q) \quad = \quad \frac{q}{\sqrt{q^2+1}} \;=\; \mathrm{softclip}(q) \tag{4}$$

Our prefered penalty function $H(q_t)$ used for finding u_t, equation (3) is the hyperbolic (or hybrid) penalty function. The output q_t best senses sparsity when its typical numerical value in the penalty function $H(q_t) = \sqrt{q_t^2+1}-1$ is found near the transition level between ℓ_1 and ℓ_2 norms, namely, when $|q_t| \approx 1$.

LOG SPECTRAL PARAMETERIZATION

A minimum phase wavelet can be made from any causal wavelet by taking it to Fourier space, and exponentiating. The proof is straightforward: Let $U(Z) = 1 + u_1 Z + u_2 Z^2 + \cdots$ be the Z transform ($Z = e^{i\omega}$) of any causal function u_t. Consider $e^{U(Z)}$. Although we would always do this calculation in the Fourier domain, the easy proof is in the time domain. The power series for an exponential $e^U = 1 + U + U^2/2! + U^3/3! + \cdots$ has no powers of $1/Z$ (because U has no such powers), and it always converges because of the powerful influence of the denominator factorials. Likewise e^{-U}, the inverse of e^U, always converges and is causal. Thus both the filter and its inverse are causal. This is the essense of minimum phase.

We seek to find two functions, one strictly causal the other strictly anticausal (nothing at $t = 0$).

$$U^+ \quad = \quad u_1 Z + u_2 Z^2 + \cdots \tag{5}$$

$$U^- \quad = \quad u_{-1}/Z + u_{-2}/Z^2 + \cdots \tag{6}$$

Notice U, U^2, etc do not contain Z^0. Thus the coefficient of Z^0 in $e^U = 1 + U + U^2/2! + \cdots$ is unity. Thus $a_0 = b_0 = 1$.

$$e^{U^+} \;=\; A \quad = \quad 1 + a_1 Z + a_2 Z^2 + \cdots \tag{7}$$

$$e^{U^-} \;=\; B \quad = \quad 1 + b_1/Z + b_2/Z^2 + \cdots \tag{8}$$

Define $U = U^- + U^+$. The decon filter is $AB = e^U$ and the source waveform is its inverse e^{-U}.

The wavelet AB is as compact as any wavelet of its spectrum can possibly be. This follows from Robinson's minimum delay theorem (FGDP) that: *Comparing all causal wavelets of the same spectrum, being minimum-phase gets energy out the quickest.*

Consider $U(\omega) = \ln AB$ the log spectrum of the filter. We will be adjusting the various u_t, all of them but not u_0. What does u_0 represent? It dictates the average of the log spectrum. The other u_t cannot change the average; they merely cause the log spectrum to oscillate.

The gradient

Having data d_t, having chosen gain g_t, and having a starting log filter $u_t = 0$, let us see how to update u_t to find the gained output $q_t = g_t r_t$ with the best hyperbolicity. Our forward modeling operation with model parameters u_t acting upon data d_t (in the Fourier domain $D(Z)$ where $Z = e^{i\omega}$) produces deconvolved data r_t (the residual).

$$r_t \quad = \quad \mathrm{FT}^{-1}\ D(Z)\ e^{\cdots + u_2 Z^2 + u_3 Z^3 + u_4 Z^4 + \cdots} \tag{9}$$

$$\frac{dr_t}{du_\tau} \quad = \quad \mathrm{FT}^{-1}\ D(Z)\ Z^\tau e^{\cdots + u_2 Z^2 + u_3 Z^3 + u_4 Z^4 + \cdots} \tag{10}$$

$$\frac{dr_t}{du_\tau} \quad = \quad r_{t+\tau} \tag{11}$$

The last step follows because Z^τ simply shifts the data $D(Z)$ by τ units which shifts the residual the same. An output formerly at time t gets moved to time $t+\tau$. This result may look familiar, but it is not. The familiar result is that the derivative of a filter output with respect to the filter coefficient at lag τ is the shifted input $d_{t+\tau}$ not the shifted output $r_{t+\tau}$ we see above. This difference leads to remarkable consequences below.

It is the gained residual $q_t = g_t r_t$ that we are trying to sparsify. So we need its derivative by the model parameters u_τ.

$$q_t \quad = \quad g_t\ r_t \ = \ r_t\ g_t \tag{12}$$

$$\frac{dq_t}{du_\tau} \quad = \quad \frac{dr_t}{du_\tau}\, g_t \ = \ r_{t+\tau}\ g_t \tag{13}$$

We may select any penalty function $H(q)$ (which to avoid difficulty is most certain to be convex $d^2H/dq^2 > 0$). Recall $u_0 = 0$ and hence $\Delta u_0 = 0$. To find the update direction at nonzero lags $\Delta\mathbf{u} = (\blacksquare\mathbf{u_t})$ take the derivative of the hyperbolic penalty function $\sum_t H(q_t)$ by u_τ.

$$\Delta\mathbf{u} \quad = \quad \sum_t \frac{dH(q_t)}{du_\tau} \qquad \tau \neq 0 \tag{14}$$

$$= \quad \sum_t \frac{dq_t}{du_\tau}\ \frac{dH(q_t)}{dq_t} \tag{15}$$

$$\Delta\mathbf{u} \quad = \quad \sum_t\ (r_{t+\tau})\ \ \big(g_t H'(q_t)\big) \qquad \tau \neq 0 \tag{16}$$

This says to crosscorrelate the physical residual r_t with a statistical residual, namely, correlate it with the statistical function $H'(q_t)$ (called the softclip) gained by g_t. Notice in reflection seismology the physical residual r_t generally decreases with time while the gain g_t generally increases the statistical variable with time.

The components of a crosscorrelation may be computed in the time domain as shown above. Alternately we bring the two factors into the Fourier domain and do the job more quickly by multiplication.

$$\Delta U \quad = \quad \overline{\mathrm{FT}(r_t)}\ \mathrm{FT}(g_t\, \mathrm{softclip}(q_t)) \tag{17}$$

Actually, equation (17) as it stands is wrong. Conceptually it should be brought into the time domain and have Δu_0 set to zero. More simply, the mean could be removed in the Fourier domain.

Least squares theory in a stationary world (GEE) says the output r_t is white; its autocorrelation is a delta function. Equation (16) is the generalization to other penalty functions with echo data (gained signals). The crosscorrelation of gained functions of the output is a delta function.

How a gradient move changes the residual

Now we adopt the convention that components of a vector $\mathbf{u}$ range over the values of (u_t), and a likewise shorthand for other vectors. We have found the gradient direction $\Delta\mathbf{u}$. Now we need to choose a good distance α to go in that direction. Given a model change $\alpha\Delta\mathbf{u}$, what residual change $\alpha\Delta\mathbf{q}$ would we see? First we will find $\Delta\mathbf{q}$, after that, α. The expression e^U is in the Fourier domain. A simple two-term example introduces a required linearization (neglect of α^2).

$$e^{\alpha\Delta U} \quad = \quad e^{\alpha(\Delta u_1 Z + \Delta u_2 Z^2)} \tag{18}$$

$$e^{\alpha\Delta U} \quad = \quad 1 + \alpha(\Delta u_1 Z + \Delta u_2 Z^2) + \alpha^2(\cdots) \tag{19}$$

$$\mathrm{FT}^{-1}\ e^{\alpha\Delta U} \quad = \quad (1, \alpha\Delta u_1, \alpha\Delta u_2) + \alpha^2(\cdots) \tag{20}$$

$$\mathrm{FT}^{-1}\ e^{\alpha\Delta U} \quad = \quad (1, \alpha\Delta\mathbf{u}) + \alpha^{\mathbf{2}}(\cdots) \tag{21}$$

With that background, ignoring α^2, and knowing the gradient $\Delta\mathbf{u}$, let us work out the forward operator to find $\Delta\mathbf{q}$. Let "$*$" denote convolution.

$$\mathbf{r} + \alpha\Delta\mathbf{r} \quad = \quad \mathrm{FT}^{-1}(De^{U+\alpha\Delta U}) \tag{22}$$

$$= \quad \mathrm{FT}^{-1}(De^{U}\, e^{\alpha\Delta U}) \tag{23}$$

$$= \quad \mathrm{FT}^{-1}(De^{U}) * \mathrm{FT}^{-1}(e^{\alpha\Delta U}) \tag{24}$$

$$= \quad \mathbf{r} * (1, \alpha\Delta\mathbf{u}) \tag{25}$$

$$= \quad \mathbf{r} + \alpha\, \mathbf{r} * \Delta\mathbf{u} \tag{26}$$

$$\Delta\mathbf{r} \quad = \quad \mathbf{r}\ *\ \blacksquare\mathbf{u} \tag{27}$$

$$\Delta q_t \quad = \quad g_t\ \Delta r_t \tag{28}$$

It is pleasing to notice that $\Delta\mathbf{r}$ is proportional to $\mathbf{r}$. This consequence of the logarithmic variables seems to imply we can deal with a wide dynamic range within r_t. Because seismograms span a large range of amplitudes in both time and frequency we may require that wide dynamic range. The convolution, a physical process, occurs in the physical domain which is only later gained to the statistical variable q_t. Naturally, the convolution may be done as a product in the frequency domain.

ALGORITHM

A shot waveform may be estimated from a single seismogram or a group of them. We look forward to gaining more experience on this issue. Meanwhile, the pseudo code below handles groups.

Lower case letters are used for variables in time and space like $\mathtt{d} = d(t,x)$, $\mathtt{g} = g(t,x)$, $\mathtt{q} = q(t,x)$, $\mathtt{dq} = \Delta q(t,x)$. while upper case for functions of frequency $\mathtt{D} = D(\omega,x)$, $\mathtt{R} = R(\omega,x)$, $\mathtt{dR} = \Delta R(\omega,x)$, $\mathtt{U} = U(\omega)$, $\mathtt{dU} = \Delta U(\omega)$. Asterisk $*$ means multiply within an implied loop on t or ω.

```
D = FT(d)
u=0
iteration {
        U = FT(u)
        dU = 0
        for all x
                r = IFT( D * exp(U))
                q = g * r
                dU = dU + conjg( FT(r)) * FT( g * softclip(q) )
        remove mean from dU(omega)
        for all x
                dR = FT(r) * dU
                dq = g * IFT(dR)
        alpha = Newton line search (  H( q+alpha*dq ))
        u = u + alpha * du
        }
```

CONCLUSIONS (THEORETICAL)

As you see, the basic algorithm is compact and straightforward. Naturally traditional regularization, preconditioning, and computational enhancements are possible. What is needed now is more experience to better identify the oportunities and limitations of this new method. Will we be able to integrate the reflectivity to get the impedance? Hopefully, but it is too early to say.