Decon in the log domain with variable gain

Jon Claerbout, Qiang Fu, and Antoine Guitton

SUMMARY

Because predictive decon fails on the Ricker wavelet, earlier we devised an extension to non-minimum phase wavelets (Yang et al, 2011). It showed remarkable clarity of seismogram polarity. Here we improve our method by a log spectral parameterization. Another innovation here is correctly handling time-variable gain. Since filtering does not commute with TV gain, gain is now done after decon (not before). Results at two survey locations confirm the utility of both improvements. An intriguing theoretical aspect shows that log spectral parameterization links penalty functions to crosscorrelation statistics of outputs.

LOG SPACE, SPARSITY, AND GAIN

In the long-standing problem of blind deconvolution we replace the traditional unknown filter coefficients by 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 typically quite rapid.

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 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 "sparsified" (Li et al, 2012) by minimizing $\sum_t H(q_t)$ where

$$q_t \;=\; g_t\, r_t \tag{2}$$

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

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

Our prefered penalty function $H(q)$ used for finding u_t is the hyperbolic (or hybrid) penalty function (equation (3)). The output q_t best senses sparsity when gain is such that the typical penalty $H(q_t)$ value is found near the transition level between ℓ_1 and ℓ_2 norms, namely, when typical $|q_t| \approx 1$.

MINIMUM PHASE EXTENSION

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.

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

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

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 $a_0 = b_0 = 1$.

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

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

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

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 which is 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, say $u_t = 0$, let us see how to update u_t to find a gained output $q_t = g_t r_t$ with better 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).

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

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

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

This follows because Z^τ shifts the data $D(Z)$ by τ units which shifts the residual the same. Output formerly at time t moves to time $t+\tau$. This is not the familiar result that the derivative of an output with respect to a filter coefficient at lag τ is the shifted *input* $d_{t+\tau}$. Here we have the *output* $r_{t+\tau}$. 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.

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

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

Recall $u_0 = 0$ and hence $\Delta u_0 = 0$. To find the update direction at nonzero lags $\Delta \mathbf{u} = (\Delta u_t)$ take the derivative of the hyperbolic penalty function $\sum_t H(q_t)$ by u_τ.

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

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

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

This says to crosscorrelate the physical residual r_t with the statistical residual $g_t H'(q_t)$. Notice in reflection seismology the physical residual r_t generally decreases with time while the gain g_t generally increases to keep the statistical variable q_t roughly constant, so $g_t H'(q_t)$ grows in time(!)

In the frequency domain the crosscorrelation (16) is:

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

Equation (17) is wrong at $t = 0$. It should be brought into the time domain and have Δu_0 set to zero. More simply, the mean can be removed in the Fourier domain.

Causal least squares theory in a stationary world says the signal output r_t is white (Claerbout, GEE); the *auto*correlation of the signal output is a delta function. Noncausal sparseness theory (other penalty functions) in a world of echos (nonstationary gain) says the *cross*correlation of the signal output with its gained softclip is also a delta function (equation (16) upon convergence).

TAKING THE STEP

We adopt the convention that components of a vector $\mathbf{u}$ range over the values of (u_t), likewise for other vectors. Given the gradient direction $\Delta \mathbf{u}$ we need to know the residual change $\Delta \mathbf{r}$ and a distance α to go: $\alpha \Delta \mathbf{r}$ and $\alpha \Delta \mathbf{u}$.

A two-term example demonstrates a required linearization.

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

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

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

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

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

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

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

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

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

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

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

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

It is pleasing that $\Delta \mathbf{r}$ is proportional to $\mathbf{r}$. This might mean we can deal with a wide dynamic range within r_t. The convolution, a physical process, occurs in the physical domain which is only later gained to the statistical domain q_t. Naturally, the convolution may be done as a product in the frequency domain.

To minimize $H'(\mathbf{r} + \alpha \Delta \mathbf{r})$ express it as a Taylor series approximation to quadratic order. Minimizing yields

$$\alpha \quad = \quad - \sum_t \Delta q_t H'_t \;\; / \;\; \sum_t (\Delta q_t)^2 H''_t \tag{29}$$

Update $\mathbf{r} = \mathbf{r} + \alpha \Delta \mathbf{r}$. Update $\mathbf{u} = \mathbf{u} + \alpha \Delta \mathbf{u}$. Optionally iterate to overcome the quadratic truncation (i.e. Newton method).

ALGORITHM

Pseudo code below finds the best single filter for a group of seismograms. Notice $g(t,x)$ could contain mute patterns, etc.

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 {
     dU = 0
     for all x
          r = IFT( D * exp(U))
          q = g * r
          dU = dU + conjg(FT(r)) * FT(g*softclip(q))
     remove the 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
     U = FT(u)
     }
```

UNIQUENESS

As the figures show, our results are excellent, amazing even, but we've had a continuing problem with uniqueness. We find our present solution can spike any of the three lobes of the Ricker wavelet defining the sea floor. This is particularly annoying as it amounts to apparent time shifts and polarity changes. For about a year we have ascribed this difficulty to descent in a nonlinear problem. Now it looks like something else is responsible.

First we tried as a starting guess various approximations to the inverse of the Ricker wavelet, an inverse that would spike at the middle lobe of the Ricker wavelet. These didn't work. Then we tried various preconditionings, such as changing variables so the dependent variable would be the industry standard predictive decon. That made things worse. We tried various regularizations $\sum_t w_t u_t^2$, but they didn't seem to help. We tried different ways of truncating u(t) at negative times, but that didn't work. Then we returned to addressing again the initial conditions. Although U=0 often worked, it was not wholly reliable. Then we realized the initialization $U = -\log(|D|)$ amounts to starting with a standard but symmetrical deconvolution $R = D/\sqrt{D^*D}$. That begins us with a symmetrical spiking wavelet much nearer our goal than $U = 0$, namely $R = D$. We were dismayed to find although iteration seemed to converge rapidly, iteration being allowed to continue, the spiking might switch to first or third lobes of the Ricker wavelet with their accompanying polarity change. To make matters worse, only slight changes in the gain function g_t would determine the selection of which final lobe. Arrg!

These studies led us to conclude that we are not facing a problem with the non-linear descent. Instead, our hyperbolic penalty function has not the power to choose the lobe. Data spiked on other lobes is also well spiked. Thus we need some kind of regularization to make that choice. In the days preceding this abstract deadline, we came up with ideas for that, but had insufficient time to test them adequately. Come to our talk to find out if we have overcome this perplexing aspect.

GOALS

A long range goal is to successfully integrate the reflectivity to get the log impedance. This requires good low frequency handling. Recording equipment often suppresses low frequencies for various practical reasons whose validity is likely location dependent. Our decon is pulling back some of these low frequencies but should be stopping before pulling up noise. Figure 2 demonstrates doing gain after non-minimum phase decon makes a valuable first step. To find impedance may require the additional statistical assumption of sparseness, but by solving the physical problem correctly, we have reduced the need for that.

ACKNOWLEDGEMENT

We thank Western Geophysical for the Gulf of Mexico data set. Jon Claerbout and Qiang Fu thank the sponsors of the Stanford Exploration Project. Antoine Guitton thanks Repsol Sinopec Brasil SA and Geo Imaging Solucoes Tecnologicas em Geociencias Ltda. We'd like to thank Yang Zhang for continued interest.

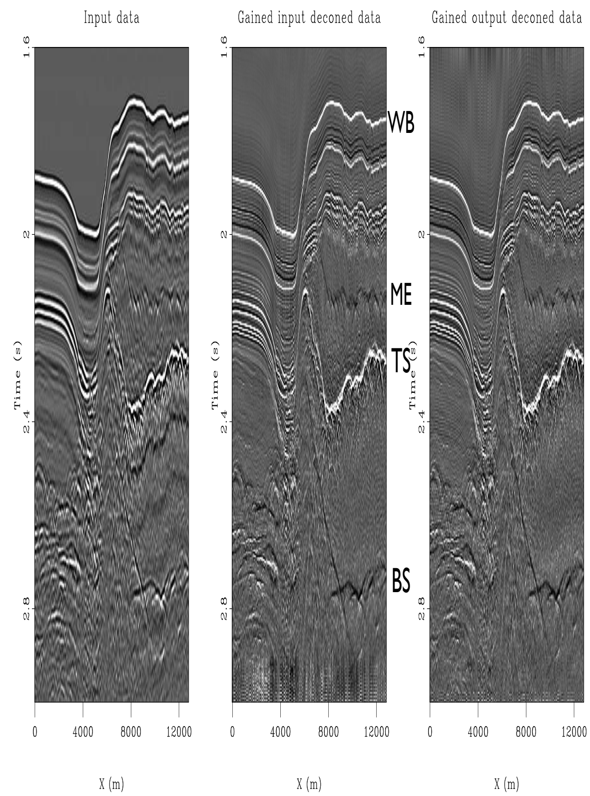

Figure 1: Gulf of Mexico. Decon produces plain white reflections from hard boundaries, and plain black boundaries from soft ones. WB= Water Bottom (white), TS= Top Salt (white), BS= Bottom Salt (black), ME= Mystery Event (black), soft reflector could be rugose salt solution of a former salt layer.

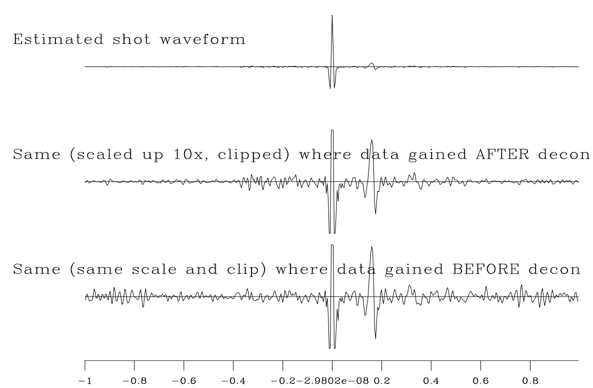

Figure 2: Wavelet from Cabo data. Although causality is not imposed, the estimated shot waveform is near causal (discounting the leading lobe of the Ricker wavelet). The importance of gain (here t^2) after deconvolution instead of before is shown by the lower two traces. There is much less noise when we gain AFTER decon, not BEFORE. Notice also that gain before decon estimates a slightly larger bubble (which is wrong).

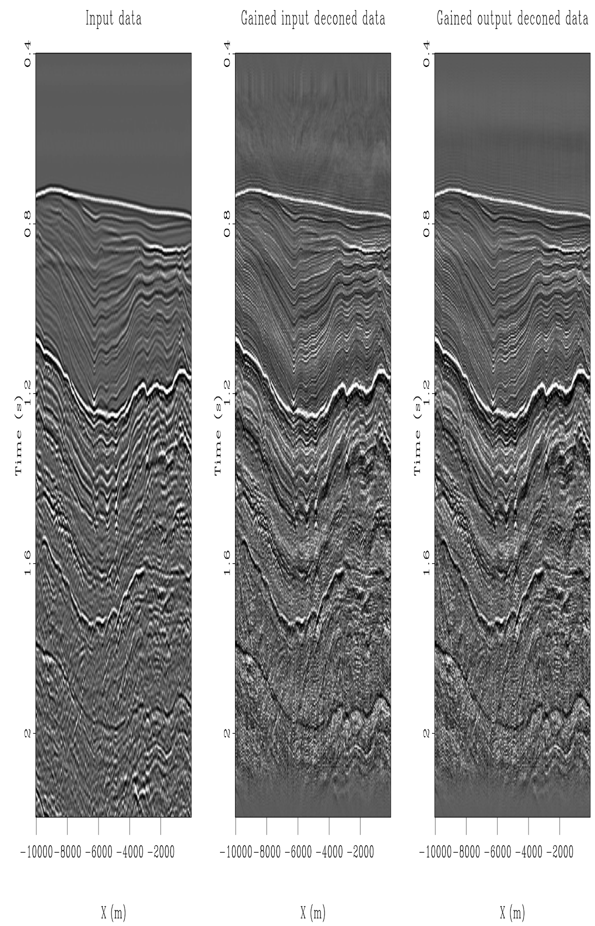

Figure 3: Cabo data. One filter on all traces. Bubble (at about 0.9s) removed. Enhanced high frequency at 1.1s. Gained-input method gave low frequency event precursors especially clear above the event at 1.2s but also visible above the water bottom. The problem is overcome by the gained-output method. (Guitton)

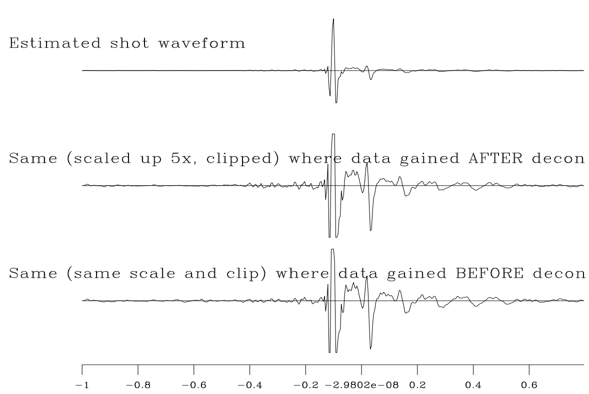

Figure 4: The Gulf of Mexico data produces a very different bubble but the same conclusions as Figure 2. The lack of symmetry in the Ricker wavelet may be related to the unresolved uniqueness issue. (Awaits better regularization.)