A log spectral approach to bidirectional deconvolution

Jon Claerbout, Qiang Fu, and Yi Shen

ABSTRACT

The blind-deconvolution problem for non-minimum-phase-source was established in the time domain. This is a Fourier domain formulation. Changing variables from $A(Z)B(1/Z)$ to $U(Z) = \ln(A(Z)B(1/Z))$ leads to a different kind of whiteness — the output being orthogonal to its shifted soft clip.

INTRODUCTION

We three have been impressed by the excellent field data results of Zhang and Claerbout (2010), making us feel that proper blind deconvolution will give much more reliable presentation of seismogram polarity, and hopefully impedance. Additionally Fu et al. (2011) have estimated a field-data shot waveform of remarkable plausibility. Exploring with programs similar to Zhang and Claerbout (2010), however, we discovered a variety of unexpected irregularities that have sent us back to many investigations with synthetic data. Indeed, we now have three quite independent approaches. We believe the heart of the problem lies in the fundamental non-linearity of the basis of the formulation. We are not aware of any realistic linear formulation of it. To allieviate the difficulties we have (1) come up with a much improved starting solution in the form of Ricker wavelet inverses, (2) come up with a smooth "simultaneous" descent method, and (3) come up with a log spectral formulation. Unfortunately, we are not out of the woods yet. Hopefully, adding preconditioning to our codes will guide us away from spurious solutions enabling us to make more firm conclusions about the theory which would then allow us to plunge confidently into many field data investigations.

THE LOG SPECTRAL APPROACH

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. Then $e^{U(Z)}$ will be minimum phase. 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 $1/Z$, 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. Q.E.D.

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

$$U^+ = u_1 Z + u_2 Z^2 + \cdots \tag{1}$$
$$U^- = u_{-1}/Z + u_{-2}/Z^2 + \cdots \tag{2}$$

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 $a_0 = b_0 = 1$.

$$e^{U^+} = A = 1 + a_1 Z + a_2 Z^2 + \cdots \tag{3}$$
$$e^{U^-} = B = 1 + b_1/Z + b_2/Z^2 + \cdots \tag{4}$$

Define $U = U^- + U^+$. The decon filter is $AB = e^U$ and the source waveform is its inverse e^{-U}. With the Fourier transform of the data $D(\omega)$, the decon output is:

$$\mathbf{r} = (r_t) = \mathrm{FT}^{-1}\ D(\omega)\ e^{U(Z(\omega))} \tag{5}$$

where U is found with a penalty function, our choice being the hyperbolic penalty function.

$$\mathrm{argmin(U)} = \mathrm{hyp}(\mathbf{r}) = \sum_t H(r_t) \tag{6}$$

where $H(r) = \sqrt{r^2 + R^2} - R$, and R is the ℓ_1/ℓ_2 threshold parameter.

Take the gradient of the penalty function assuming there is only one variable, u_3 giving a single regression equation:

$$0 \approx \sum_t \frac{\partial H}{\partial r}\frac{\partial r}{\partial u_3} = \frac{\partial r}{\partial u_3}\frac{\partial H}{\partial r} \tag{7}$$
$$0 \approx \sum_t (\mathrm{FT}^{-1}\ D(\omega)\ \frac{\partial}{\partial u_3}\ e^{U(Z)})_t\ H'(r_t) \tag{8}$$
$$0 \approx \sum_t (\mathrm{FT}^{-1}\ D(\omega)\ Z^3\ e^{U(Z)})_t\ H'(r_t) \tag{9}$$

so the deconvolution output selected at time $t+3$ multiplies $H'(r_t)$ (also known as the "soft-clip" function).

Equation (9) requires us to do an inverse Fourier transform to find the gradient for only u_3. For u_4 there is an analogous expression, but it is time shifted by Z^4 instead of Z^3. Clearly we need only do one Fourier transform and then shift it to get the time function required for other filter lags. Thus the gradient for all nonzero lags is:

$$0 \approx \Delta\mathbf{u} = (\Delta u_\tau) = (\sum_t\ r_{t+\tau}\ H'(r_t)) \tag{10}$$
$$\Delta U = \overline{\mathrm{FT}(\mathbf{r})}\ \mathrm{FT}(\mathrm{softclip}(\mathbf{r})) \tag{11}$$

where τ measures some filter lag. Actually, equation (11) is wrong as it stands. Conceptually 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.

Equation (10) says if we were doing least squares, the gradient would be simply the autocorrelation of the residual. When the gradient at nonzero lags drops to zero, the residual is white. Hooray! We long understood that limit. What is currently new is that we now have a two-sided filter. Likewise the ℓ_1 limit must be where the output is uncorrelated with the clipped output at all lags but zero lag.

(I'm finding it fascinating to look back on what we did all these years with the causal filter $A(Z)$ and comparing it to the non-causal exponential filter $e^{U(Z)}$. In an ℓ_2 norm world for filter $A(Z)$ we easily saw the shifted output was orthogonal to the fitting function *input*. For filter $e^{U(Z)}$ we easily see now the shifted output is orthogonal to the *output*. The whiteness of the output comes easily with $e^{U(Z)}$ but with $A(Z)$ the Claerbout (2011) contains a lengthy and tricky proof of whiteness.)

Let us figure out how a scaled gradient $\alpha\Delta\mathbf{u}$ leads to a residual modification $\alpha\Delta\mathbf{r}$. The expression e^U is in the Fourier domain. We first view a simple two term example.

$$e^{\alpha\Delta U} = e^{\alpha(\Delta u_1 Z + \Delta u_2 Z^2)} \tag{12}$$

$$e^{\alpha\Delta U} = 1 + \alpha(\Delta u_1 Z + \Delta u_2 Z^2) + \alpha^2(\cdots) \tag{13}$$

$$\mathrm{FT}^{-1}\ e^{\alpha\Delta U} = (1, \alpha\Delta u_1, \alpha\Delta u_2) + \alpha^2(\cdots) \tag{14}$$

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

$$\mathbf{r} + \alpha\Delta\mathbf{r} = \mathrm{FT}^{-1}(De^{U+\alpha\Delta U}) \tag{15}$$

$$= \mathrm{FT}^{-1}(De^{U}e^{\alpha\Delta U}) \tag{16}$$

$$= \mathrm{FT}^{-1}(De^{U}) * \mathrm{FT}^{-1}(e^{\alpha\Delta U}) \tag{17}$$

$$= \mathbf{r} * (1, \alpha\Delta\mathbf{u}) \tag{18}$$

$$= \mathbf{r} + \alpha\mathbf{r} * \Delta\mathbf{u} \tag{19}$$

In familiar ℓ_2 problems we would find α as $\alpha = -(\mathbf{r}\cdot d\mathbf{r})/(d\mathbf{r}\cdot d\mathbf{r})$, now we must find it by

$$\mathrm{argmin}(\alpha) = H(\mathbf{r} + \alpha\mathbf{r} * \Delta\mathbf{u}) \tag{20}$$

This we do by the Newton method which iteratively fits the hyperbola to a parabola.

```
iterate {
  alpha = - (Sum_i H'(r_i) dr_i ) / ( Sum H''(r_i) dr_i^2 )
  r = r + alpha dr
  }
```

In the pseudocode above, in the ℓ_2 limit $H'(r) = r$ and $H''(r) = 1$, so so the first iteration gets the correct α, and changes the residual accordingly so all subsequent α values are zero.

We are not finished because we need to assure the constraint $u_0 = 0$. In a linear problem it would be sufficient to set $\Delta u_0 = 0$, but here we soon do a linearization which breaks the constraint. In the frequency domain the constraint is $\sum_\omega U = 0$. We meet this constraint by inserting a constant β in equation (15) and chosing β to get a zero sum over frequency of $U + \alpha\Delta U + \beta$. Let $\sum$ denote a normalized summation over frequency. By normalized, I mean $\sum \beta = \beta$. We must choose β so that $0 = \sum U + \alpha \sum \Delta U + \beta$. Clearly, $\beta = -\sum U - \alpha \sum \Delta U$. Pick up again from equation (15) including β.

$$\begin{aligned}
\text{argmin}(\alpha) &= H(\ \text{FT}^{-1}(De^U\ e^{\alpha\Delta U}\ e^{\beta})\ ) && (21)\\
&= H(\ \text{FT}^{-1}(De^U\ e^{\alpha\Delta U}\ e^{-\sum U}\ e^{-\alpha\sum\Delta U})\ ) && (22)\\
&= H(\ \text{FT}^{-1}(De^{U-\sum U}\ \ e^{\alpha(\Delta U-\sum\Delta U)})\ ) && (23)\\
&= H(\ \text{FT}^{-1}(De^{U-\sum U})\ *\ \text{FT}^{-1}(e^{\alpha(\Delta U-\sum\Delta U)})\ ) && (24)
\end{aligned}$$

Proceed now along the lines of equation (17) through (20), but with means removed in Fourier space leading to slightly different vectors $\tilde{\mathbf{r}}$ and $\Delta\tilde{\mathbf{u}}$

$$\text{argmin}(\alpha) = H(\ \tilde{\mathbf{r}} + \alpha\tilde{\mathbf{r}} * \Delta\tilde{\mathbf{u}}\ ) \qquad (25)$$

which is solveable by the method of the same pseudocode above.

ALGORITHM

Here we fill in more details of the algorithm. After we are certain of its behavior we would naturally switch over to conjugate directions.

```
D(omega,x) = FT d(t,x)
u=0;
iteration {
        U = FT(u)
        remove mean from U(omega)
        exp(U(Z))
        dU = 0
        for all x
                r(t,x) = IFT( D exp(U) )
                softclip( r )
                dU += conjg(FT(r)) * FT(softclip)        #  "*" means multiply
        remove mean from dU(omega)
        for all x
                dR = FT(r) * dU                          #  "*" means multiply
                dr = IFT(dR)
        argmin(alpha) = H(r+alpha*dr)
        u = u + alpha du
        }
```

CONCLUSIONS

It is too early to draw reliable conclusions about this theory. It arose a few weeks before the progress report deadline, and got coded a few days before. The coding (steepest descent) illuminated a bug in the initial theory regarding the constraint. A few test cases were successfully run, but we do not know how this method compares to Shen et al. (2011) and Fu et al. (2011).

The problem formulation itself is nonlinear and thereby susceptible to many minima, mostly bad ones. We don't know whether preconditioning applies in the usual way to this method. Perhaps some kind of tapering of Δu_t is a means of directing the method towards a more appropriate optimum, but for now this is no more than speculation.

REFERENCES

Claerbout, J. F., 2011, Image Estimation by Example: `http://sepwww.stanford.edu/ sep/prof/`.

Fu, Q., Y. Shen, and J. Claerbout, 2011, An approximation of the inverse ricker wavelet as an initial guess for bidirectional deconvolution: SEP-Report, **143**, 283–296.

Shen, Y., Q. Fu, and J. Claerbout, 2011, A new algorithm for bidirectional deconvolution: SEP-Report, **143**, 271–282.

Zhang, Y. and J. Claerbout, 2010, A new bidirectional deconvolution method that overcomes the minimum phase assumption: SEP-Report, **142**, 93–104.