Separation of Non-Stationary Signals and Noises

Jon F. Claerbout

Despite the many sophisticated papers published on deconvolution, we still have no way to solve single channel problems that humans solve all the time. You hear several voices, and after a while you are able to distinguish several speakers. This paper is an attempt to formulate the problem in a very simplified way.

Let d denote some data which has not been AGCed. Imagine a white noise component ξ in d. The white component has an unknown time-variable variance. The signal will be denoted by a. We have $d = \xi + a$. To begin, we take half of d to be in a and half to be in ξ. Then we construct initial AGC functions with

$$D_\xi^2 \quad = \quad diag\left(\frac{1}{\ll | d-a |^2 \gg} \right) \tag{1}$$

$$D_a^2 \quad = \quad diag\left(\frac{1}{\ll | a |^2 \gg} \right) \tag{2}$$

The signal will be assumed to be adequately sampled, so that its spectrum drops off near the Nyquist. Thus we take the power spectrum of the signal after scaling by D_a, to be proportional to $1/(\omega_0^2 + \omega^2)$. The covariance matrix of the inverse of this spectrum is the familiar tridiagonal matrix encountered in migration programs. Now we set out to minimize

$$E \; = \; (d-a)^* D_\xi^2 (d-a) \; + \; a^* D_a \, T \, D_a \, a \tag{3}$$

Set $\partial E / \partial a^*$ to zero.

$$-D_\xi^2 (d-a) + D_a \, T \, D_a \, a \quad = \quad 0 \tag{4}$$

$$(\, D_\xi^2 + D_a \, T \, D_a \,) \, a \quad = \quad D_\xi^2 \, d \tag{5}$$

Equations (1), (2), and (5) constitute an algorithm which can be iterated. To capture big spikes in the residual, I imagine that we would want to have a short time constant in the smoothing $\ll \gg$ in (1) but a much longer smoothing window in (2).

Sparse Signals

Now suppose that we have been too cavalier in writing the inverse covariance matrix of the signal as $D_a\, T\, D_a$. Perhaps the signal a is produced by a sparse spike train which is smoothed through some wavelet. Maybe the inverse covariance matrix should be represented as $B^*\, D_a^2\, B$ where B is a deconvolution matrix. A tricky point, that I frequently gloss over is how to define D_a^2. I guess it should be the square of the AGC function for $B\, a$. I can also imagine a slow AGC before decon and a fast one afterward.

There are very many unresolved issues of whether and how iterations will work.