Deconvolution Essays

Jon F. Claerbout

DECON — THE REPETITION TEST

Deconvolution is the process of estimating an unknown wavelet shared by a group of seismograms and then removing that wavelet. There are many deconvolution methods but most fail a simple test.

Repetition Test

Take the output of a deconvolution process and use it for the input of a second iteration of the same deconvolution process. The second output should be the same as the first output. If this repetition test fails, then the deconvolution process is not consistent with the basic definition.

Examples

The original definition of spiking deconvolution as unit span prediction error filtering satisfies the repetition test, provided that the filter is infinitely long. We have no quibble with filters that are not infinitely long but are merely as long as computationally feasible. However some decon filters are deliberately chosen to have a finite length to limit the number of statistical degrees of freedom. Such a filter fails to pass the repetition test. The Kolmogorov frequency domain spectral factorization satisfies the repetition test if the spectrum of the data is smoothed over space but not if the spectra are smoothed over time. Certain carefully-specified, gapped filters may pass the repetition test.

Why the repetition test is important

Increasingly we are deconvolving iteratively. Three kinds of iterations are:

1. over various spatial axes
2. with non-linear optimization criteria
3. between radial traces and ordinary traces.

If a deconvolution method does not pass the repetition test, then we will not know whether the goals 1-3 are being met.

PREWHITENING

Linearity

The most linear seismic section plotting device available is the variable brightness CRT display. It offers an advantage that, in principle, it is not necessary to remove very high frequencies before data display. Very high frequencies on a CRT blur together as you step back from the screen. So you choose your preferred high frequency filter cutoff just by changing the distance of your eye from the screen.

Decon

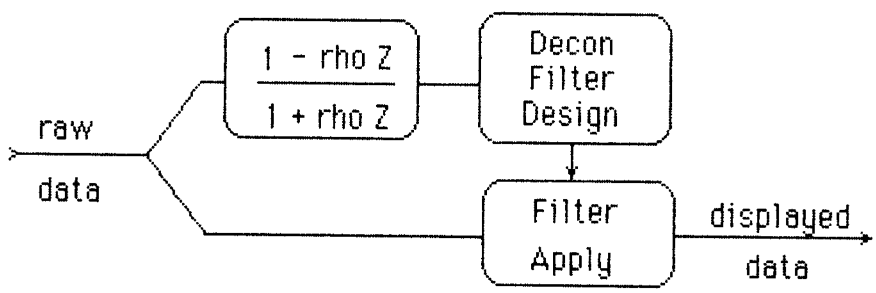
You might guess that an application of this filtering idea would be in post coloration of seismic data after whitening by spiking decon. In principle, the use of a CRT display would mean that post coloration would be unnecessary. In practice there is a tradeoff. As you add more and more high frequencies to the seismic data, the peak time-domain amplitude goes up. As the peak goes up you must scale down the seismograms because the CRT is physically limited to some maximum brightness. This brightness lost by scaling is not recovered when you filter out the high frequencies by moving your head away. The need for a high frequency cutoff filter after deconvolution arises from the limited range of linearity of our plotting devices. With the CRT, which is linear over quite a large range, the issue of design of a high frequency cutoff filter is not a very important one. In practice we have most linearity problems with hard copy plot devices. When traces are closer together the plotting becomes more non-linear and frequency cutoff becomes more important.

Post-filter

Let us design a post-decon coloration filter. Since we will use a CRT, the design need not be elaborate. How about the filter $(1+\rho Z)/(1-\rho Z)$? The free parameter will be ρ. This filter has a damped exponential response. It is monotonic in both time domain and frequency domain. It is minimum phase. Its inverse has the same form with ρ of opposite sign. For a value of ρ equal a half the filter is (1/2, 1/2, 1/4, 1/8, ...) which looks to be roughly suitable for later subsampling the data at alternate points. (In our lab the undesirable practice has arisen of abandoning alternate points with a *Window* program that does not even bother to low-pass filter!)

Pre-filter

Rather than apply a post filter after decon, it is possible to design the decon filter so that a post filter is not necessary. The way to do this is to use a prewhitening filter, such as the inverse of the postfilter mentioned above. Note below the opposite sign on ρ. Then design the decon filter, then finally apply the filter to the nonprewhitened data. In summary:

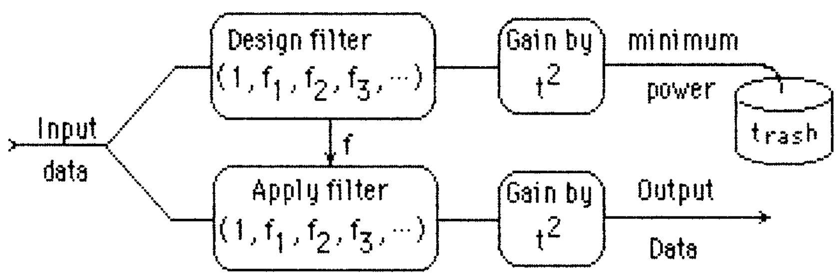


Observe that the overall process will satisfy the repetition test, if the decon filter design (i.e. $\rho = 0$ case) satisfies the repetition test.

DECON AND GAIN

Gain Dilemma

There is a dilemma on when to do the gain control, before or after deconvolution. Simple theory says that data should be transformed to stationarity before Wiener-Levinson filter theory should be applied. This means multiplying the data by time t or t^2 depending on the theory used for spherical divergence correction. (Let us presume, as is well founded by experience, that exponential gain control is unsuitable). If you don't multiply by some such data scaling, you have the problem that the filter design spectrum is strongly biased in favor of the spectrum at early time. Information at late time gets ignored, even though you want your data to be correctly deconvolved at late time. But multiplying your data by these gain control functions destroys the applicability of convolution. Convolution before scaling is not the same as convolution after. There is a solution to this dilemma, but it involves abandoning comfortable old spectral and Toeplitz concepts and tackling a non-trivial optimization problem. Here is what you need to do:



How to do it.

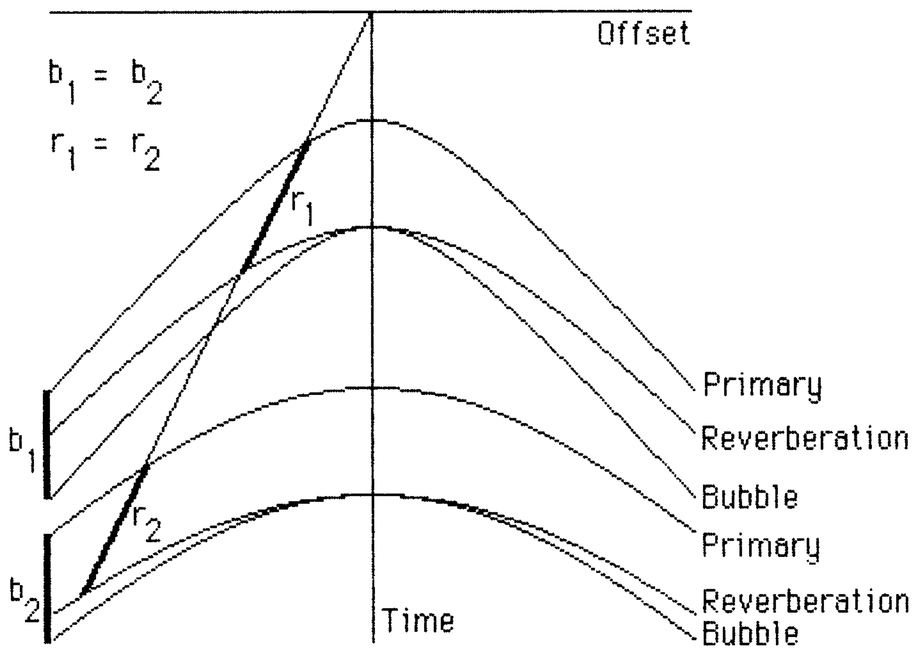
The diagram implies an optimization process to compute the filter coefficients by minimizing the output power after gain control. This optimization problem is linear, but the Levinson technique is apparently not usable. (It is an open question whether a simple "LEVITY type" adaptation of the Burg algorithm will do the job. Since it is not yet proven and published let's assume not.) This may be a fairly slow computation. One way to contain costs is to use a good iterative algorithm such as conjugate gradients or Page-Saunders. In the past I used exact algorithms and contained costs by limiting filter length, but this fails the repetition test. Spectral prewhitening should further hasten the optimization iteration.

DEBUBBLE VERSUS DEREVERBERATION

Deconvolution is the seismic process to compensate for (1) shot waveform and (2) near-surface reverberation. Separately the processes are called debubble and dereverberation. Generally, the industry does not distinguish between the two processes, perhaps because they both involve convolution or because under certain highly-idealized, one-dimensional assumptions they are hardly distinguishable. Consideration of moveout correction and of spherical divergence correction shows massive differences between the processes. These massive differences suggest that we distinguish the two processes, see which is the more important, and think about which should be done first or whether they should be done simultaneously.

NMO

To see the massive differences that arise from non-zero offset, consider a two layer model with the same velocity in each layer. The top layer will be said to be shallow, to illustrate a shallow multiple reflection, i.e. a shallow reverberation. Let the bubble delay equal the vertical incidence two-way travel time to the shallow layer. So at zero offset, the bubble arrives at the same time as the multiple reflection. A bubble curve is always delayed by a fixed amount from its primary. At infinite offset, a multiple reflection has the same travel time as its primary. A quantitative analysis yields the following travel time curves:



We observe that the timing differences are indeed massive and marvel that radial traces are so widely neglected. Large amplitude differences also distinguish debubble from dereverb. Debubbling fits the convolutional model. Convolution with the debubbling filter should be done before spherical divergence correction. (In Decon and Gain it is explained that statistics for finding the debubbling filter should be based on filter outputs that have been divergence corrected).

Reverberation models are basically one dimensional models, which means that real 3-D seismograms must be transformed to one dimension before the reverberation model applies. So dereverberation, unlike debubble, should be done after spherical divergence correction. Dereverberation is a form of inversion that is simplified to the convolutional model. To derive the convolutional model you must assume first that the only multiple reflections worth worrying about are those near the source or the receiver. Deep multiples demand a more complicated non-convolutional theory such as that in FGDP chapter 8. There are two means of transforming field data to one dimension. First is slant stack, a method that is theoretically ideal for depth dependent velocity. Second is radial traces, less troubled by data imperfections, but theoretically limited to constant velocity. Considering that the industry commonly ignores the difference between field traces and radial traces, we can hope that radial traces are as suitable for dereverberation as slant stacks.

How to get started

A way to go after these two different phenomena is to deconvolve twice, say once before spherical divergence, and once after transformation to radial traces and spherical divergence correction. I don't know how this would work out in practice, but we can readily anticipate some theoretical difficulties. Suppose the data does not contain wide offsets. Then the models for debubble and dereverb are about the same, so whichever process is done first will capture the combined wavelets. Since we don't need the two wavelets, we really only need the deconvolved data, the quality of the result may not depend strongly on the order of the processes. But what if the two phenomena have comparable significance? Then it seems to be necessary to iterate between debubble and dereverb. It seems to me that a worthy goal for us at SEP is to define an optimization procedure that will simultaneously estimate and remove both wavelets. Then we will see which is the more significant and we will be able to evaluate the approximations inherent to many "practical" approaches. In practice decon is rarely done on radial traces, and it usually is done after gain control to the inputs. This doesn't seem like a realistic physical model. Does it matter? If not, why not? What does deconvolution mean when it is applied to a CDP stack?

Worries

Perhaps the the assumption of a constant velocity model is misleading us. On land the near surface is very slow and even in the marine case, the water is generally somewhat slower than the sediments. The effect may be to improve the conventional constant offset model compared to the radial model. We should think this through, both with regard to the moveout correction and the divergence correction.

REFERENCES

Paige, C.C., and M.A. Saunders, 1982, ALGORITHM 583 LSQR: Sparse linear equations and least squares problems ACM Transactions on Mathematical Software, Vol. 8, No. 2, June 1982, pp 195-209.

Paige, C.C., and M.A. Saunders, 1982, LSQR: An algorithm for sparse linear equations and sparse least squares, ACM Transactions on Mathematical software, vol 8, No 1, March 1982, pp 43-71.

Claerbout, J.F., 1976, LEVITY: Levinson recursion reprogrammed, SEP-10, p 90-98.

Stanford

Exploration

Project