IX. IMPLEMENTATION DETAILS

The two preceding chapters were dedicated to the development of techniques for iteratively maximizing the norm ratio $U(X,\alpha_1,\alpha_2)$. Algorithms for implementing the techniques in both the time and frequency domains were briefly outlined. Further details of implementation and guidelines for choosing input parameters are given in this chapter.

A. Choice of Threshold

In Chapter VI norm ratios such as $U(X,2,\hat{\alpha})$ were proposed where $\hat{\alpha}$ is the estimated distribution of the current reflectivity series. These approaches have a drawback because $\hat{\alpha}$ is often less than 1. This causes problems because the gradient blows up for amplitudes near zero. A partial remedy is to add a small constant to the reflection magnitudes to move them away from zero. The effect of this procedure on the gradient is illustrated in Figure 9-1.

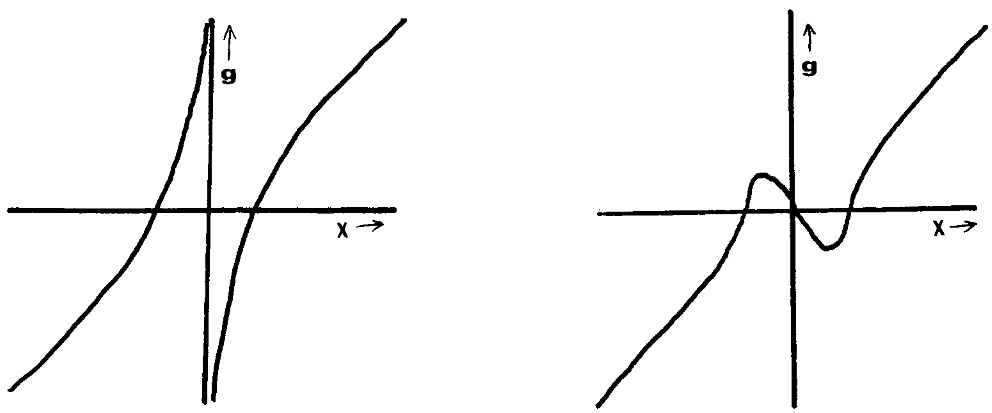

FIGURE 9-1. On the left is the gradient versus reflection amplitude for a norm ratio such as $U(X,2,.55)$. The gradient modified by slightly increasing the amplitudes is shown on the right.

We have tried several methods for implementing this procedure. What worked best was to set the constant to some percentage of the scale estimated from Equation (5-15). The percentage is reduced as the iteration proceeds. This allows large events to be considered first. Smaller events are brought into the gradient in later iterations. When the algorithm converges the constant should be at the noise level. Unfortunately we have found no consistent method for choosing the constant and how to reduce it. What worked for one set of data would fail on another.

B. Choice of Initial Filter

Because the equations being solved are non-linear, they may have several relative maxima. Which of the maxima is found for a solution is determined by the initialization of the inverse filter. The most unbiased choice is a centered unit spike, f = (0.....1.....0). Moving the spike towards the beginning causes the time domain implementation to favor causal solutions, thus reducing precursors. The location of the spike does not affect the frequency domain implementation. We have tried using the estimated minimum phase inverse as an educated initial guess. The results obtained were identical to those using the centered unit spike, and convergence was not accelerated.

When deconvolving common shot gathers the results of processing one gather can be used as an initial estimate for the next gather. This works best when the source waveform is relatively constant along the line. Often convergence can be achieved in two iterations using this procedure.

The time domain implementation allows the length of the inverse filter to vary. We have studied the effects of varying the filter length and often it modifies the solution obtained. Using a longer inverse can increase resolution but requires more data to average out the noise and increases the ability of the filter to combine events into a single spike. Generally the length of the inverse filter should depend on the data window length, number of channels, estimated bandwidth of the source, and noise level present in the dataset.

C. Constraining the Bandwidth

Seismic data is typically oversampled because field recording systems usually suppress frequencies above the half Nyquist. Generally energy present above the field filter cutoff represents noise and is undesirable in any deconvolution procedure. F. Muir (35) has suggested a clever method for constraining the bandwidth of the solution. Only alternate lags of the auto and crosscorrelations are computed. Zeros are inserted for the lags not computed. The solution of the Toeplitz system consequently has half the frequency range of the original data.

This procedure halves the number of degrees of freedom allowed the inverse filter which yields a statistically more reliable result and allows a longer filter to be computed from a given set of data. The procedure is also computationally attractive because only half as many lags of the auto- and crosscorrelations are required. If much energy is present in the high frequency portion of the spectrum, this method is potentially dangerous because these frequencies are aliased. This aliasing distorts the low frequencies.

D. Truncation and End Effects

Generally, seismic sections are windowed prior to deconvolution. This allows noisy portions to be edited out, which improves stationarity but causes problems because the sharp edges of the editing window can introduce high frequencies not present in the data. The traditional remedy is to apply a triangular or cosine taper to the edges. This removes the high frequencies but introduces low frequencies. These low frequencies may have been suppressed by the field filter and are consequently blown up by the deconvolution. We have found the best approach is to use no taper. The high frequencies can be suppressed by the method explained in the previous section.

Another problem caused by windowing is the inverse filter not being completely meshed with the inputs at the edges. This often creates spurious events at the beginnings and ends of the outputs. These events appear on the first iteration and tend to dominate the succeeding solutions.

This problem can be partially avoided by estimating statistics and a gradient from samples of the reflectivity series computed with the full inverse filter. On the final iteration the ends can be computed using the partial solutions of the Toeplitz system. The Levinson recursion estimates a filter of length n + 1 from one of length n. The first samples of the reflectivity series are computed as

$$\begin{aligned} x_0 &= y_0 f_0^{\,1} \\ x_1 &= y_0 f_1^{\,2} + y_1 f_0^{\,2} \\ x_2 &= y_0 f_2^{\,3} + y_1 f_1^{\,3} + y_2 f_0^{\,3} \end{aligned} \tag{9-1}$$

where the superscript on f represents the n-th partial filter. This procedure is applied to estimate the beginning portion of the reflectivity series. The full length filter is then used to compute the central portion. A simple way to estimate the end portion is to reverse the reflectivity series, RHS, and inputs. The same algorithm used to estimate the beginning portion is then applied. This procedure is computationally expensive because the system of equations must be solved twice and core must be available to hold all inputs and outputs simultaneously.

E. Correcting for Spherical Divergence

The amplitude distortions caused by spherical divergence are a major violation of the stationarity assumption. In the preceding chapters we have tried to find a filter whose outputs maximize spikiness. Here we use the same concepts to estimate a gain constant which minimizes spikiness, thus correcting for spherical divergence.

Let y_{ij} be the i-th sample of the j-th recorded reflection seismogram where there are m channels, each having n samples. The exponentially gained seismograms, X, are defined as

$$x_{ij} = y_{ij}\lambda^{i} \tag{9-2}$$

where λ is a gain constant chosen to optimize some property of X. Assume the gained seismograms are characterized by the generalized Gaussian family of distributions. The norm ratio given by Equation (6-7) is a scale-invariant measure of the probability that X is a member of the distribution characterized by the shape parameter α_2 and not α_1.

Using Equation (9-2), the norm ratio can be written as a nonlinear function of the gain constant λ:

$$W(\lambda,\alpha_1,\alpha_2,Y) = \log \prod_{j=1}^{m} \frac{\left\{\frac{1}{n}\sum_{i=1}^{n} |y_{ij}\lambda^{i}|^{\alpha_1}\right\}^{n/\alpha_1}}{\left\{\frac{1}{n}\sum_{i=1}^{n} |y_{ij}\lambda^{i}|^{\alpha_2}\right\}^{n/\alpha_2}} \quad . \tag{9-3}$$

By minimizing W, the optimum gain constant, λ^*, distorts the seismograms, making the probability very low that they are described by α_2. Thus, the gain constant modifies the inputs so that their distribution is as close as possible to α_1.

1. A General Solution

If there exists some λ^* which minimizes Equation (9-3), its first derivative vanishes,

$$\frac{dW(\lambda^*,\alpha_1,\alpha_2,Y)}{d\lambda} = 0 \ . \tag{9-4}$$

An iterative method is used to solve Equation (9-4) because no analytic solution exists. The classic technique is Newton's method which is often unstable for arbitrary nonlinear functions.

Figure 9-2 shows the norm ratio as a function of the gain constant for eight seismic gathers. The function is very smooth and has a global minimum. Newton's method, given a reasonable initial guess, should have no problems minimizing the function.

Let λ^k be the k-th guess for solving Equation (9-4). Newton's method gives the next estimate from

$$\lambda^{k+1} = \lambda^k - \frac{\frac{dW}{d\lambda}(\lambda^k,\alpha_1,\alpha_2,Y)}{\frac{d^2W}{d\lambda^2}(\lambda^k,\alpha_1,\alpha_2,Y)} \quad . \tag{9-5}$$

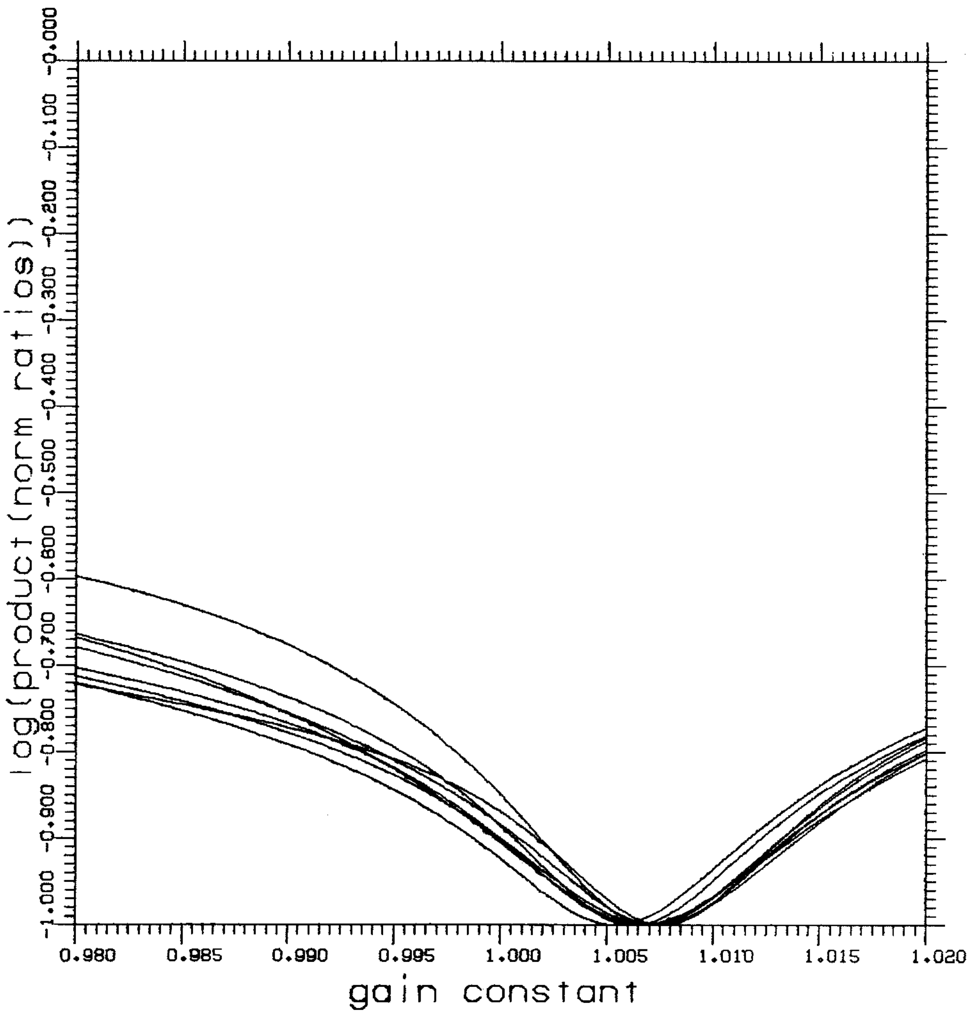

FIGURE 9-2. The norm ratio, $W(\lambda,\alpha_1,\alpha_2,Y)$, plotted versus λ for the seismic gathers of Figure 9-3. The numerical values of W for each gather were scaled between (-1,0).

The simplest stopping condition is when

$$|\lambda^{k+1} - \lambda^{k}| < \varepsilon \tag{9-6}$$

where ε is the desired accuracy.

Taking the derivatives of Equation (9-3) and simplifying gives

$$\frac{dW}{d\lambda} = \sum_{j=1}^{m} \left(\frac{C_j}{A_j} - \frac{D_j}{B_j} \right) \tag{9-7}$$

and

$$\frac{d^2W}{d\lambda^2} = \sum_{j=1}^{m} \left[\frac{E_j}{A_j} - \frac{F_j}{B_j} - \alpha_1 \left(\frac{C_j}{A_j} \right)^2 + \alpha_2 \left(\frac{D_j}{B_j} \right)^2 \right] \tag{9-8}$$

where

$$A_j = \sum_{i=1}^{n} |y_{ij}|^{\alpha_1} \lambda^{\alpha_1 i}$$

$$B_j = \sum_{i=1}^{n} |y_{ij}|^{\alpha_2} \lambda^{\alpha_2 i}$$

$$C_j = \sum_{i=1}^{n} |y_{ij}|^{\alpha_1} \lambda^{\alpha_1 i-1} i$$

$$D_j = \sum_{i=1}^{n} |y_{ij}|^{\alpha_2} \lambda^{\alpha_2 i-1} i$$

$$E_j = \sum_{i=1}^{n} |y_{ij}|^{\alpha_1} \lambda^{\alpha_1 i-2} (\alpha_1 i^2-1)$$

$$F_j = \sum_{i=1}^{n} |y_{ij}|^{\alpha_2} \lambda^{\alpha_2 i-2} (\alpha_2 i^2-1) \quad .$$

A program using Newton's method was implemented and run on several sets of real data. Figure 9-3 displays seismograms before and after exponential gain. Eight gathers having six traces each were used. The initial guess was $\lambda = 1$. α_2 was estimated from the data as .6. α_1 was arbitrarily given the value 2. The iteration converged with an accuracy greater than 1×10^{-6} after five iterations. The average distribution of the output was $\alpha = 1.47$, about halfway between a double exponential and a Gaussian.

Newton's method is a general minimization technique. For specific values of α_1 and α_2, more efficient search procedures can be implemented.

2. A Specific Solution

The generalized Gaussian with infinite shape parameter is a uniform distribution. Several experiments have shown that seismic data generally has a shape parameter near 1. It is believed that the optimum function to minimize is $W(\lambda,\infty,1,Y)$.

If all channels are of the same length, an equivalent function to $W(\lambda,\infty,1,Y)$ is

$$V(\lambda,Y) = \sum_{j=1}^{m} \log \frac{\max\limits_{i} (|y_{ij}\lambda^{i}|)}{\sum\limits_{i=1}^{n} (|y_{ij}\lambda^{i}|)} \tag{9-9}$$

The numerator of Equation (9-9) simplifies because the maximum term of the sum,

$$\sum_{i=1}^{n} (|y_{ij}|\lambda^{i})^{\frac{n}{\alpha_1}} \tag{9-10}$$

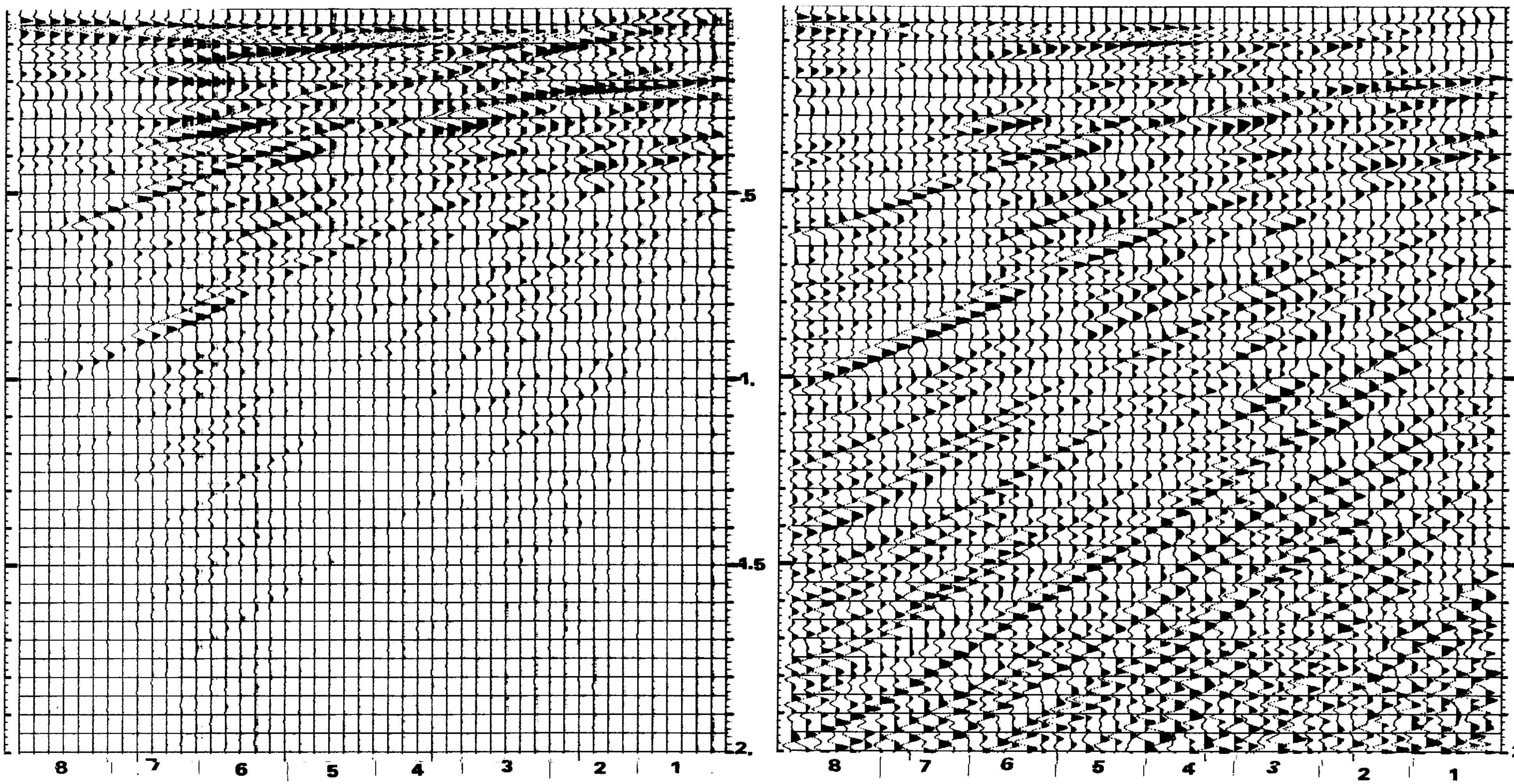

FIGURE 9-3. On the left is an edited seismic gather. Eight gain constants were computed for successive groups of six traces. On the right is the exponentially gained gather.

dominates as α goes to infinity. Hogg (24) gives empirical results and Uthoff (36) theoretical results which imply that Equation (9-9) is optimal. For typical seismic data, a more robust function might use the lower bound of the upper percentile of the magnitudes of the gained seismogram instead of the maximum value.

The Fibonacci search method is a more efficient way of minimizing V than Newton's method. It requires a search interval within which V is unimodal. Given an initial search interval (1.,1.01) for the data of Figure 9-3, the maximum error after 11 function evaluations is 1×10^{-4}. After 16 evaluations, the error is less than 1×10^{-6}. The gain constants differed by less than 1×10^{-4} from those found by Newton's method. The outputs were slightly more uniform because the average shape parameter was 1.62.

The algorithm was tested on uniform random numbers. The estimated gain constant differed from 1 by less than 1×10^{-5}. This implies the minimization will do nothing to data which are already uniformly distributed.

Care should be taken when using this method before deconvolution as the statistics of the noise are not affected by spherical divergence. Boosting the gain of later portions of seismograms can raise the noise to unacceptable levels and can degrade the deconvolution.