VARIABLE NORM EXPONENTIAL GAIN ESTIMATION

William Gray

A multichannel norm ratio is presented which measures the distribution of exponentially gained seismograms as a function of the gain constant. Minimization of the ratio yields an optimum gain constant for achieving a uniform output distribution. Newton's method is used to minimize the general norm ratio. A Fibonacci search method is used to minimize the ratio when the inputs are distributed as a double exponential and the outputs are to be as uniform as possible.

Background

In "A Theory for Variable Norm Deconvolution" (see p. 123, this report), a norm ratio is maximized to determine a spiky deconvolution. Here the same ratio is minimized to determine an exponential gain which produces the most uniform seismograms.

Non-adaptive deconvolution methods assume the statistics of reflection seismograms are constant with respect to time. Multichannel methods also assume they are constant not only in time but also across channels. These assumptions are invalid for reflection seismograms; the major violation being the amplitude distortion caused by spherical divergence.

Traditionally, seismograms are edited and an exponential gain is applied to correct approximately for spherical divergence. Often the gain function is determined by trial and error in an attempt to make reflection amplitudes subjectively more uniform over time. This is a costly and time-consuming process which can be automated.

Problem statement

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

$$x_{ij} = y_{ij}\lambda^{i} \tag{1}$$

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 probability distributions. The norm ratio

$$U(X;\alpha_1,\alpha_2) = \log \prod_{j=1}^{m} \frac{\left(\frac{1}{n}\sum_{i=1}^{n}|x_{ij}|^{\alpha_1}\right)^{n/\alpha_1}}{\left(\frac{1}{n}\sum_{i=1}^{n}|x_{ij}|^{\alpha_2}\right)^{n/\alpha_2}} \tag{2}$$

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.

The shape parameter orders the family of distributions. For $\alpha = 1$, the distribution is a double exponential; for $\alpha = 2$ it is a Gaussian; and for $\alpha \to \infty$ it is a uniform. In Equation (2) the shape parameter for the ungained data is α_2, and for the gained data is α_1. The derivation of the norm ratio, a means of estimating α_2, and further properties of the generalized Gaussian are presented in "A Theory for Variable Norm Deconvolution."

Using Equation (1), the norm ratio can be written as a nonlinear function of the gain constant λ:

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

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.

A General solution

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

$$\frac{dW(\lambda^*;\alpha_1,\alpha_2,Y)}{d\lambda} = 0 \qquad (4)$$

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

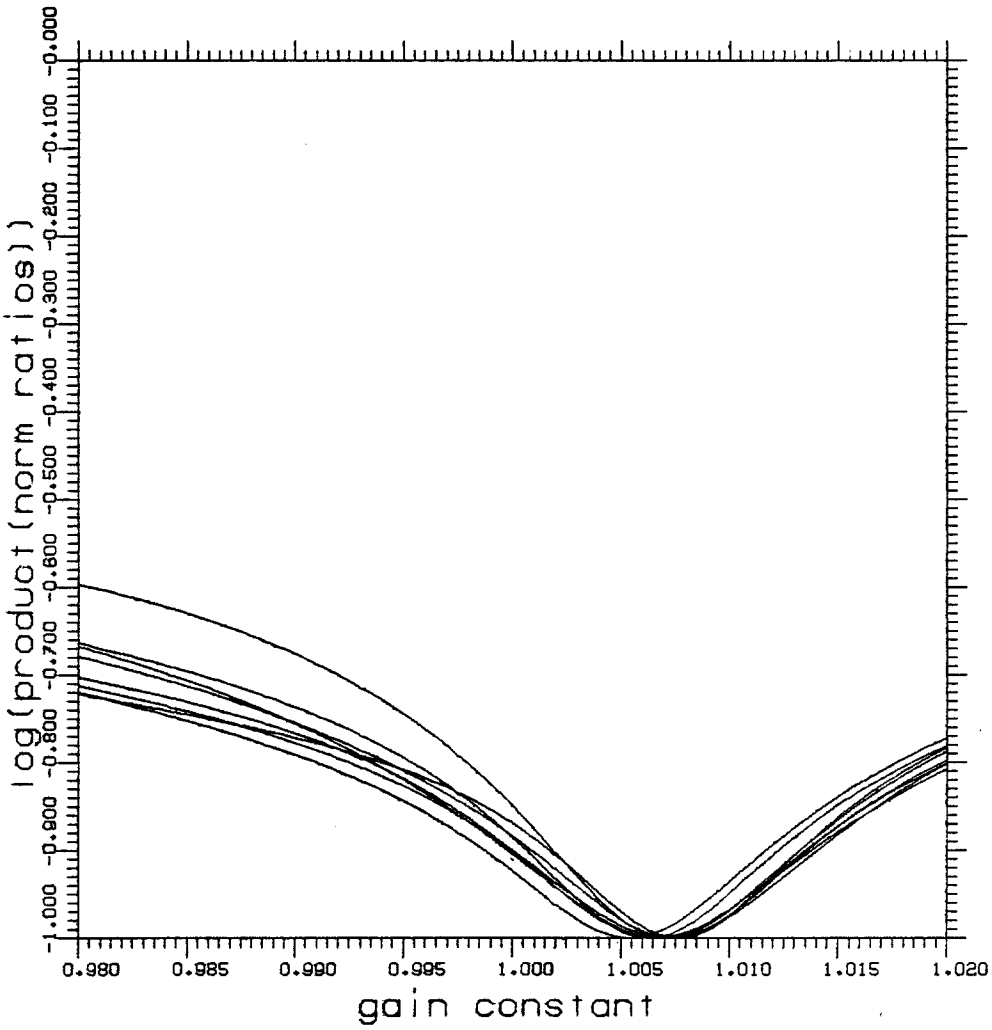

FIGURE 1.--The norm ratio, $W(\lambda;\alpha_1,\alpha_2,Y)$, plotted versus λ for the seismic gathers of Figure 2. The numerical values of W for each gather were scaled between (-1,0).

Figure 1 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 (4). Newton's method gives the next estimate from

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

The simplest stopping condition is when

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

where ε is the desired accuracy.

Taking the derivatives of Equation (3) and simplifying gives

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

and

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

where

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

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

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

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

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

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

A program using Newton's method was implemented and run on several sets of real data. Figure 2 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.

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

$$V(\lambda;Y) = \sum_{j=1}^{m} \log \frac{\max_i (|y_{ij}\lambda^i|)}{\sum_{i=1}^{n} (|y_{ij}\lambda^i|)} \qquad (9)$$

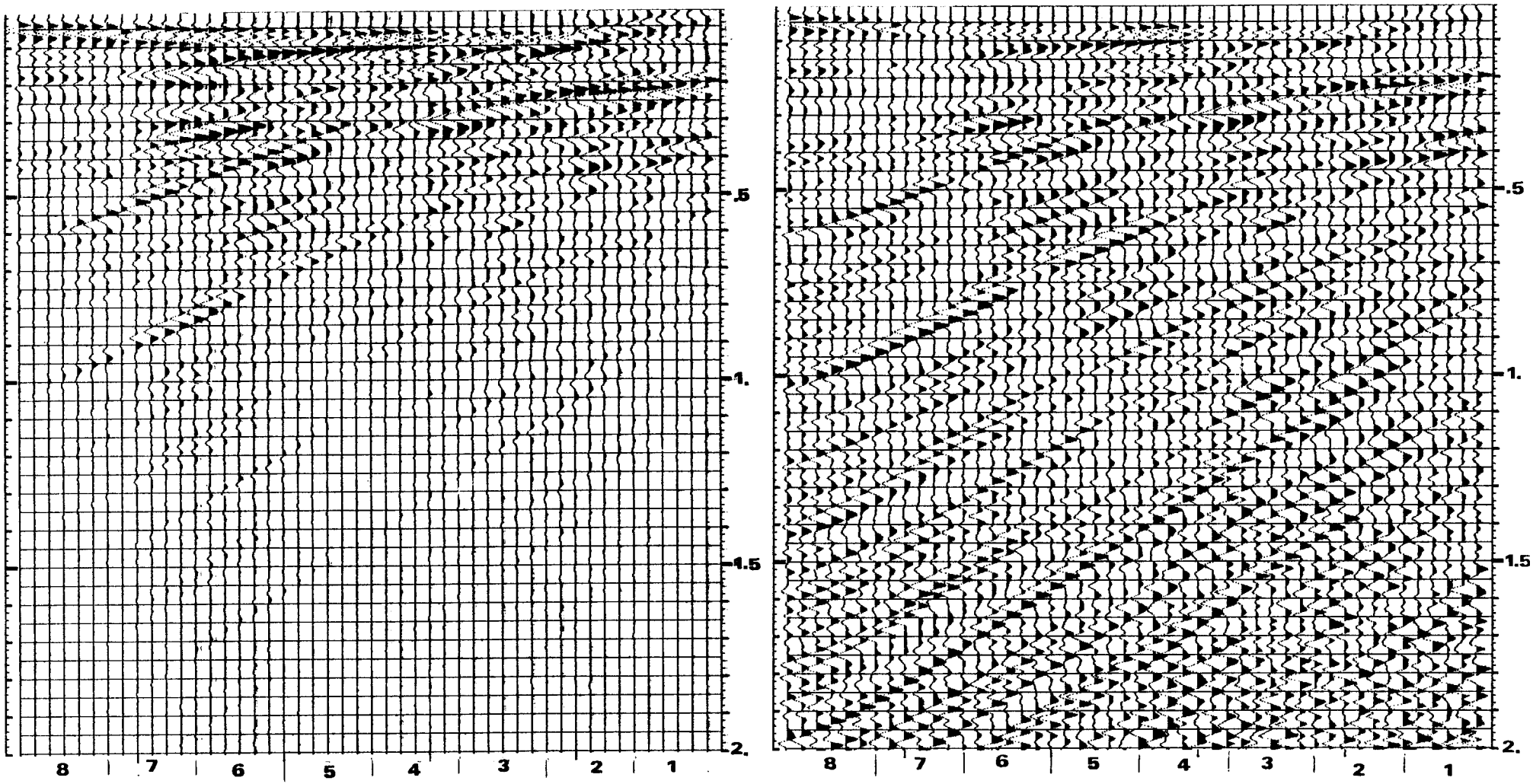

FIGURE 2.--On the left is an edited seismic gather. Eight gain constants were computed for successive groups of six traces. On the right is the exponentially gained gather.

The numerator of Equation (3) simplifies because the maximum term of the sum,

$$\sum_{i=1}^{n} \left(|y_{ij}|\ \lambda^{i}\right)^{\frac{n}{\alpha}1}$$

dominates as α goes to infinity. Hogg (1972, pp. 422-424) gives empirical results and Uthoff (1970, pp. 1597-1600) theoretical results which imply that Equation (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. An excellent description of the method is given by Avriel (1976, pp. 225-240).

Given an initial search interval (1.,1.01) for the data of Figure 2, 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.

Conclusions

A scale invariant, multichannel method for determining an exponential gain has been presented. It has worked well for several sets of seismic data. Care should be taken when using it 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.

REFERENCES

AVRIEL, M. (1976), *Nonlinear Programming: Analysis and Methods* (Englewood Cliffs, New Jersey: Prentice-Hall).

HOGG, R. V. (1972), "More light on the Kurtosis and related statistics," *Journ. Amer. Statist. Assoc.*, 67.

UTHOFF, V. A. (1970), "An optimum test property of two well-known statistics," *Journ. Amer. Statist. Assoc.*, 65.