SIGNAL/NOISE SEPARATION AND SEISMIC INVERSION

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF GEOPHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
William S. Harlan
March, 1986

printed as Stanford Exploration Project Report No. 47
by permission of the author

Signal/noise separation and seismic inversion

William S. Harlan
Stanford University, 1986

ABSTRACT

A non-unique inversion of reflection seismic data can be most easily interpreted if it describes the most important reflections with the fewest physical parameters at unexpected values. Signal should be easily modeled by independent parameters, but noise should lack the spatial coherence (statistical dependence between samples) of signal. A parameter should be perturbed from its simplest value only if this perturbation cannot be easily explained by incoherent noise in the data. Similarly, estimated noise should not be easily explained by the equations that model signal.

Statistical tools are illustrated by an improvement of the "normal-moveout," hyperbolic stack of seismic field gathers. I first find least-squares perturbations of parameters that model independent hyperbolic reflections, the signal. To overestimate the contribution of noise to the inverted model, I invert completely incoherent data and compare amplitude histograms of the two models. Using information theory, I estimate probability density functions for signal and noise in the model perturbations and keep perturbations that contain mostly signal with high probability. Reliable parameters resolve hyperbolic curvatures and velocity information better than the original hyperbolic stack.

I extract from shot gathers those reflections that are easily described as sums of hyperbola segments of all curvatures. Similarly, ground roll and other recorded noise are extracted as those events likely to contain a small amount of residual hyperbolas.

Next, strategies for choosing models are discussed. For example, after migration (extrapolation of wavefields to image sources), diffracted reflections are described by

independent samples. A sum of dipping line segments that are wider than the Fresnel zone models only undiffracted bed reflections. I extract diffractions and estimate seismic velocities from the survey of a growth fault.

Lastly, using the acoustic wave-equation, I model reflections in a recorded vertical seismic profile. To encourage a locally homogeneous impedance function, I treat derivatives of rock impedance with depth as independent parameters. Unreliable non-zero derivatives are replaced by zeros. I also invert for the seismic source and receptivity of geophones and extract tube-wave noise. Reliably perturbed impedances model the data equally well as more complicated functions from conventional gradient optimizations.

ACKNOWLEDGMENTS

Above all, I wish to thank three advisors who made this work possible: Jon Claerbout, Fabio Rocca, and Patrick Lailly. Jon Claerbout posed the original question: if signal and noise behave differently after migration, how can we separate them? Like many questions to follow, this one pointed out previously neglected information in the data. Jon saw that I kept an eye on a large variety of data sets and constantly challenged me to extend my methods to other applications, particularly to the velocity stacks of Chapter 1.

Fabio Rocca followed the early work of Chapter 2 very closely and first called attention to the different statistical properties of diffracted reflections and undiffracted reflections. He also introduced to me an important body of estimation theory and information theory. He quickly penetrated numerical difficulties and found the essential properties that could make a process successful.

I also thank Patrick Lailly and Danielle Macé for patient explanations of their work on the inversion of VSP's. They helped formulate the questions discussed in Chapter 3 and gave innumerable suggestions on both the theory and the programming of the solutions. Both spent considerable time explaining the application of control theory to inversion of the VSP and the construction of a non-dispersive finite-difference algorithm. I thank L'Institut Français du Petrole for bringing us together for six months and for providing the VSP data.

I also thank Robert Burridge, Alphonso Gonzalez-Serrano, Dave Hale, Rick Ottolini, and Jeff Thorson for many crucial early discussions. I thank Francis Muir, Daniel Rothman, and John Toldi for examining and criticizing many ideas as they were later assembled on paper. I particularly thank Clement Kostov and Norm Sleep for very strong suggestions on the presentation of this thesis. I thank Fanny Toldi for very careful editing.

Finally, I thank the sponsors of the Stanford Exploration Project for their continuous support and interest in this research, and I thank Western Geophysical and Trans-Pacific Geothermal for providing data sets.

Table of Contents

List of Figures

Introduction

Many discussions of seismic inversion begin with the choice of a particular earth property and end with the conclusion that seismic data tell something, but not enough, about this property. Much valuable discussion is often presented as a technical aside: how do we attach numbers to earth properties without pretending to have more information than the data have given? How do we avoid misinterpreting irrelevant information? These two questions are discussed here as the business of signal/noise separation.

A number of authors have answered the difficulties of incomplete inversion by fundamentally redefining the goals: Lanczos, Backus and Gilbert, Claerbout and Muir, Wiggins, Gray, Godfrey, Rocca, and Thorson. Their inversions encourage models that are simplest to interpret, but that are complicated enough to explain the most important data. Partially inverted parameters prove to be useful so long as they do not mislead.

Clayton (1981) pointed out that early inversion methods could be described as "model fitting." Synthetic seismic data are computed first and compared to recorded data. Model parameters are then selectively perturbed to fit the data better, either with an interpreter's guesses, or by some simplification and partial solution of the modeling equations. Such a strategy places all emphasis on fitting the data, and none on the simplicity of the inverted model. The interpreter is responsible for encouraging simple perturbations, but he can only guess at their effects on the data. Though the number and complexity of perturbations is usually low, so is the accuracy of the synthetic data. Two interpreters can obtain very different results.

Lanczos (1961) noted that linear modeling equations preserve linear components of the model unequally and developed an algorithm to invert the best preserved components first. Some components are destroyed entirely and become part of a "null space"; these components have no effect at all on the data. Some linear components, composed of eigenvectors with small eigenvalues, are suppressed by transformation below the level of noise in the data. Lanczos constructed a decomposition (singular-value) that inverts a matrix first for eigenvectors with large eigenvalues. Unfortunately, eigenvectors rarely represent simple physical structures. The appearance of eigenvectors can change completely for different recording geometries, sampling, and parametrization.

The Lanczos algorithm, once its effect was understood, was reformulated in several equivalent ways. Damped least-squares inversion, one of the simplest, is discussed in any recent survey of linear inverse methods (see Menke, 1984). I shall discuss this method at length in Chapter 1. The inverted model should minimize an objective function equal to the sum of the Cartesian (L^2) norms of the uninverted data and model. The best model inverts only the strongest eigenvectors.

Backus and Gilbert (1968) emphasized the resolution of detail in the model parameters. If possible, information from one parameter should not affect the inverted values for other parameters. They proposed a function called "deltaness" that measures the blurring of isolated details after forward and inverse transformation. The product of matrices for inverse and forward transformation should be close as possible to a diagonal matrix. For overdetermined inversions, one can exchange accuracy in fitting the data for more resolution. Because their solution does not depend on the data, one can distinguish the resolving power of the inversion from statistical dependence between the parameters being inverted.

If the data have resulted from independent, non-Gaussian parameters, then details of the inverted model should always be isolated, sparse, "spiky," and "parsimonious." Claerbout and Muir (1973), Wiggins (1978), and Gray (1979) introduced various functions that measure the simplicity of inverted model parameters directly, chiefly for the deconvolution of seismograms. Whereas previous methods depended only on the modeling equations, these inverses examined the data to improve further the sparseness of the result. Inversions could be modified to encourage more or less simplification of the parameters.

Thorson (1984) showed that some very common geophysical processes could be redefined and improved as inversions—particularly the stacking of surface seismic data and the extrapolation and interpolation of surface and VSP data. He showed that an improved resolution in the model was essential to improving these processes. Thorson used a data-dependent algorithm that, like Backus and Gilbert, could balance the accuracy of fitting the data with the resolution of the inversion.

Most of the preceding algorithms can be derived and justified as statistical estimation problems. Each makes different, though flexible, assumptions about the statistical properties of the inverted model and noise. Godfrey (1979) and Rocca (with Godfrey, 1979, and Harlan, Claerbout, and Rocca, 1983) suggested that constraints on the model be constructed in statistical form, in terms of general probability functions. Statistical information can then be merged from outside sources or from an examination of the data. Such information can determine how much simplicity is statistically reliable for the model.

This thesis presents such an statistical estimation procedure, called signal/noise separation, for choosing simple perturbations of model parameters. I shall use a two-step method, as suggested by Clayton (1981). The first step will concentrate on transforming the data into the domain of the model parameters, so that the data can be reconstructed as well as possible. I shall begin with a simple and stable linear inverse, such as the damped least-squares inverse.

The second step will examine the perturbed model and keep those details (signal) that seem most consistent with the modeling equations. I must introduce some additional statistical tools, but their goals can be stated simply.

- Keep parameter perturbations that can be interpreted independently, that do not obscure each other.
- Keep perturbations that are unlikely to have resulted from the transformation of useless information—noise or misinterpreted signal.

The first goal can be satisfied by a correct choice of modeling equations, so that each model parameter represents a fundamental, independent physical detail. For instance, seismic midpoint gathers can be described simply if each point of the model corresponds to a single hyperbolic reflection. Such a model must treat all non-hyperbolic features as noise. The most efficient and most interpretable inversion should perturb the fewest independent parameters.

Let "signal" be those features of the data that are easily described by independent model details, and "noise" as components that are not. The ability of the model to describe an event depends on the event's statistical predictability—its "coherence." Coherence is the shape of an event in the data, the visible dependence between amplitudes in different samples. The characteristic coherence of signal can be easily created by the modeling equations; that of the noise cannot.

A least-squares inversion affects signal and noise differently: the coherence of signal simplifies, noise becomes more coherent. By examining distributions of amplitudes before and after transformation, one can estimate, with a known reliability, the contribution of noise to every least-squares perturbation. Perturbations should be kept as a reliable estimate of signal if they are unlikely to have resulted from the transformation of noise.

The three chapters of this thesis examine three applications of increasing complexity. Chapter 1 begins with the familiar hyperbolic model of the normal moveout (NMO) stack. The NMO stack is familiar as perhaps the most common single process applied to reflection seismic data; yet, its shortcomings are widely lamented. Reflections are grossly oversimplified: much valuable information is destroyed, and much non-hyperbolic noise is interpreted as if it were hyperbolic.

The first section of Chapter 1 reviews the posing and solution of least-squares inversion for a process that is not often thought of as an inversion. As did Thorson (1984), I first present a linear inverse as an alternative to the NMO stack. I use the example to present damped least-squares inverses in a form that can be applied directly to later, more complicated examples. This first step improves considerably the quality of the data that can be reconstructed from the stack.

Next, statistical methods are introduced to eliminate from the hyperbolic model those details that could have resulted from the random alignment of noise or misinterpreted signal. The same methods are applied later to a more complicated hyperbolic model that allows coherence to change laterally. Finally, the same statistical tools are used to extract ground roll and other noise—features that are not easily explained as a superposition of hyperbolic events in the data.

Chapter 2 emphasizes the crucial choice of models for the coherence of signal and noise. The best model for signal should maximize the independence of model parameters. I find that extrapolation of wavefields back in time tends to simplify the spatial coherence of signal, while dispersing and adding coherence to noise. This simplification of signal increases the independence of samples as well as their non-Gaussianity. To recognize the best model of signal, one can measure the non-Gaussianity of the inverted parameters from amplitude histograms.

Zero-offset (or stacked) data prove to have two varieties of coherent signal: diffracted reflections that contain velocity information, and non-diffracted reflections that do not. Non-diffracted reflections are easily described by the superposition of dipping lines that are wider that the Fresnel zone of reflected diffractions. Diffractions are easily described by the least-squares inverse of "migration," which extrapolates the image sources of reflections. Edge diffractions and seismic velocities are successfully extracted from reflections of a growth fault.

Chapter 3 generalizes the methods of the previous chapters for a non-linear model, to invert acoustic impedances from vertical seismic profiles. Signal extraction gives the impedance function the fewest structural complications necessary to explain the data. Derivatives of impedance with depth are used as independent model parameters. Unless the data give a strong probability to the contrary, the inverted impedance function is kept as homogeneous as possible. Noise extractions remove useless, obscuring events such as tube waves.

These chapters discuss a single method of inversion in three different contexts. The first chapter explains the algorithm with a simple and familiar seismic model. The second explores tools for the definition of new and flexible models of recorded wavefields. The third examines a difficult, non-linear modeling equation and shows how its inversion can use the simple tools developed in the earlier chapters.