Lecture Notes on

Nonlinear Inversion and Tomography:

I. Borehole Seismic Tomography

Developed from a Series of Lectures by

James G. Berryman
University of California
Lawrence Livermore National Laboratory
Livermore, CA 94550

Originally Presented at

Earth Resources Laboratory
Massachusetts Institute of Technology

July 9–30, 1990

Revised and Expanded

October, 1991

Contents

Acknowledgments

This book developed from a series of lectures presented at MIT/Earth Resources Laboratory in July, 1990. It is my pleasure to thank Professor M. Nafi Toksöz for inviting me to present these lectures and for his warm hospitality during my stay. My sincere thanks also go to Bill Rodi for his efforts to help produce the first set of notes based on the lectures.

An expanded set of lectures was subsequently presented at UC Berkeley from March through May, 1993. My host at Berkeley was Professor Jamie Rector. These lectures also included a more extensive discussion of electrical impedance tomography (the notes for which remain in an unfinished state at this writing).

I also want to thank my many colleagues both at Lawrence Livermore National Laboratory and elsewhere who have shared their insights into inverse problems and tomography with me. It is often difficult to remember who taught me what, but it is not difficult to remember that I have learned much of what I know from conversations, discussions, correspondence, and arguments with the following people: J. A. Beatty, W. B. Beydoun, S. N. Blakeslee, N. Bleistein, N. R. Burkhard, R. Burridge, M. Cheney, W. D. Daily, A. J. DeGroot, A. J. Devaney, K. A. Dines, D. M. Goodman, F. A. Grünbaum, P. E. Harben, J. Harris, D. Isaacson, J. S. Kallman, P. W. Kasameyer, R. V. Kohn, L. R. Lines, S.-Y. Lu, R. J. Lytle, G. J. F. MacDonald, J. R. McLaughlin, Ch. Y. Pichot, A. L. Ramirez, W. L. Rodi, F. Santosa, G. Schuster, W. W. Symes, A. Tarantola, J. E. Vidale, M. Vogelius, T. J. Yorkey, G. Zandt, and J. J. Zucca.

Most of this work was performed under the auspices of the U. S. Department of Energy by the Lawrence Livermore National Laboratory under contract No. W-7405-ENG-48 and supported specifically by the DOE Office of Basic Energy Sciences, Division of Engineering and Geosciences. LLNL is managed by the Regents of the University of California. Some of the work was performed at Stanford University while the author was a Consulting Professor of Geophysics with partial support from the sponsors of the Stanford Exploration Project. Some of the finishing work was completed while on sabbatical leave from LLNL at the Institut de Physique du Globe de Paris, with partial support from the French government. The author gratefully acknowledges all sources of support for this work.

Preface

> In general we look for a new law by the following process: First we guess it. Then we compute the consequences of the guess to see what would be implied if this law that we guessed is right. Then we compare the result of the computation to nature, with experiment or experience, compare it directly with observation, to see if it works. If it disagrees with experiment it is wrong. In that simple statement is the key to science.
>
> — Richard P. Feynman, *The Character of Physical Law*

Nonlinear inverse problems are common in science and engineering. In fact the quotation from Feynman shows clearly that the process of discovering physical laws is itself an inverse problem.

What is an inverse problem? Subtraction is the inverse of addition. Division is the inverse of multiplication. Root is the inverse of power. Given the answer (say, the number 4) find the question ($2+2 = ?$ or $8/2 = ?$ or $\sqrt{16} = ?$). This last example (commonly seen in the game of Jeopardy) is most important since it is clear that the same answer (*i.e.*, the data) could come from many questions (*e.g.*, models and methods of analysis) — and therefore it is not surprising that a degree of ambiguity (sometimes a very high degree of ambiguity) is an inherent part of most realistic inverse problems. Physical scientists are used to thinking about situations that lead to equations with unique solutions. Because of the traditional training they receive, most scientists are very uncomfortable with mathematical problems that do not have unique solutions. Yet both quantum mechanics and chaos theory provide numerous examples of real experiments (*e.g.*, the double slit experiment and weather) where ambiguities are typically encountered. The subject of inverse problems is another realm where lack of uniqueness commonly occurs. Methods of dealing with the ambiguities therefore play a vital — if not quite central — role in our analysis.

How do we solve an inverse problem? In general, we use the prescription described by Feynman: guess, compute, compare. But one more element is added in the inverse problems we discuss: feedback. When searching for physical laws, we make a guess, compute the consequences, and compare with experiment. If the comparisons are unfavorable, then we have learned that our first guess is bad, but we may not have any constructive procedure for generating a better guess. In contrast, when trying to solve problems in nonlinear inversion and tomography, we often think we know the physical laws (*i.e.*, the general equations governing the processes), but we may not know the precise values of the parameters in the equations. In such circumstances, it may be possible to make use of the observed discrepancies between the measured and computed results to adjust the parameters and, thereby, obtain improved guesses in a systematic way. This process of feeding the errors back to help produce a better set of parameters for the starting equation then becomes the paradigm for nonlinear inversion algorithms. In some cases it may happen that one feedback step is sufficient; in others many iteration steps may be needed to achieve the desired agreement between model and data.

James G. Berryman

Danville, CA
January, 1994

Chapter 1

Introduction to the Traveltime Inversion Problem

Our main topic is seismic traveltime inversion in 2- and 3-dimensional heterogeneous media. A typical problem is to infer the (isotropic) compressional-wave slowness (reciprocal of velocity) distribution of a medium, given a set of observed first-arrival traveltimes between sources and receivers of known location within the medium. This problem is common for crosswell seismic transmission tomography imaging a 2-D region between vertical boreholes in oil field applications. We also consider the problem of inverting for wave slowness when the absolute traveltimes are not known, as is normally the case in earthquake seismology.

In this Introduction, we define most of the general terminology we will use throughout our analysis.

1.1 Wave Slowness Models

When a sound wave or seismic wave is launched into a medium, it takes time for the influence of the wave to progress from a point close to the source to a more distant point. The time taken by the wave to travel from one point of interest to the next is called the *traveltime.* For a medium that is not undergoing physical or chemical changes during the passage of the sound, the wave has a definite speed with which it always travels between any two points in the medium. We call this speed the *average wave speed* or *wave velocity.* We can also define a *local wave speed* associated with each point in the medium by considering the average wave speed for two points that are very closely spaced. The *local slowness* is the inverse of the local wave speed. It is most convenient to develop inversion and tomography formulas in terms of wave slowness models, because the pertinent equations are linear in slowness.

We consider three kinds of slowness models. Sometimes we allow the slowness to be a general function $s(\mathbf{x})$ of the position $\mathbf{x}$. However, we often make one of two more restrictive assumptions that (i) the model comprises homogeneous cells (in 2-D), or blocks (in 3-D), with s_j then denoting the slowness value of the jth cell, or blocks. Or (ii) the model is composed of a grid with values of slowness assigned at the grid points together with some interpolation scheme (bilinear, trilinear, spline, etc.) to specify the values between grid points. Of course, as cells/blocks become smaller and smaller (down to infinitesimal), we can think of cells/blocks of constant slowness as a special case of continuous models, or of continuous models as a limiting case of cells/blocks.

When it is not important which type of slowness model is involved, we refer to the model abstractly as a vector $\mathbf{s}$ in a vector space $\mathcal{S}$. For a block model with n blocks we have $\mathcal{S} = \mathbf{R}^n$, the n-dimensional Euclidean vector space. ($\mathbf{R}$ denotes the set of real numbers.) A continuous slowness model, on the other hand, is an element of a function space, e.g., $\mathcal{S} = C(\mathbf{R}^3)$, the set of continuous functions of three real variables. No matter how we parameterize the model, we should always keep in mind that, for real materials, our models necessarily have far fewer parameters than the actual medium they are intended to represent. Thus, our models are analogous to cartoon drawings of public figures, trying to capture the main features with the minimum of detail.

1.2 Fermat's Principle and Traveltime Functionals

The traveltime of a seismic wave is the integral of slowness along a ray path connecting the source and receiver. To make this more precise, we will define two functionals[1] for traveltime.

Let P denote an arbitrary path connecting a given source and receiver in a slowness model $\mathbf{s}$. We will refer to P as a *trial ray path*. We define a functional τ^P which yields the traveltime along path P. Letting $\mathbf{s}$ be the continuous slowness distribution $s(\mathbf{x})$, we have

$$\tau^P(\mathbf{s}) = \int_P s(\mathbf{x})\, dl^P, \tag{1.1}$$

where dl^P denotes the infinitesimal distance along the path P.

Fermat's principle [Fermat, 1891; Goldstein, 1950; Born and Wolf, 1980] states that the correct ray path between two points is the one of least overall traveltime, *i.e.*, it minimizes[2] $\tau^P(\mathbf{s})$ with respect to path P.

Let us define τ^* to be the functional that yields the traveltime along the Fermat (least-time) ray path. Fermat's principle then states

$$\tau^*(\mathbf{s}) = \min_{P \in Paths} \tau^P(\mathbf{s}), \tag{1.2}$$

where *Paths* denotes the set of all continuous paths connecting the given source and receiver.[3] The particular path that produces the minimum in (1.2) is denoted P^*. If more than one path produces the same minimum traveltime value, then P^* denotes any particular member in this set of minimizing paths.

Substituting (1.1) into (1.2), we have *Fermat's principle of least time*:

$$\boxed{\tau^*(\mathbf{s}) \quad = \quad \int_{P^*} s(\mathbf{x})\, dl^{P^*} \quad = \quad \min_P \int_P s(\mathbf{x})\, dl.^P} \tag{1.3}$$

[1] A *functional* is a function which maps a function space or a vector space to the set of real numbers.

[2] Fermat's principle is actually the weaker condition that the traveltime integral is *stationary* with respect to variations in the ray path, but for traveltime inversion using measured first arrivals it follows that the traveltimes must be *minima*.

[3] The notation $P \in Paths$ means that P is a *member* of the set *Paths*.

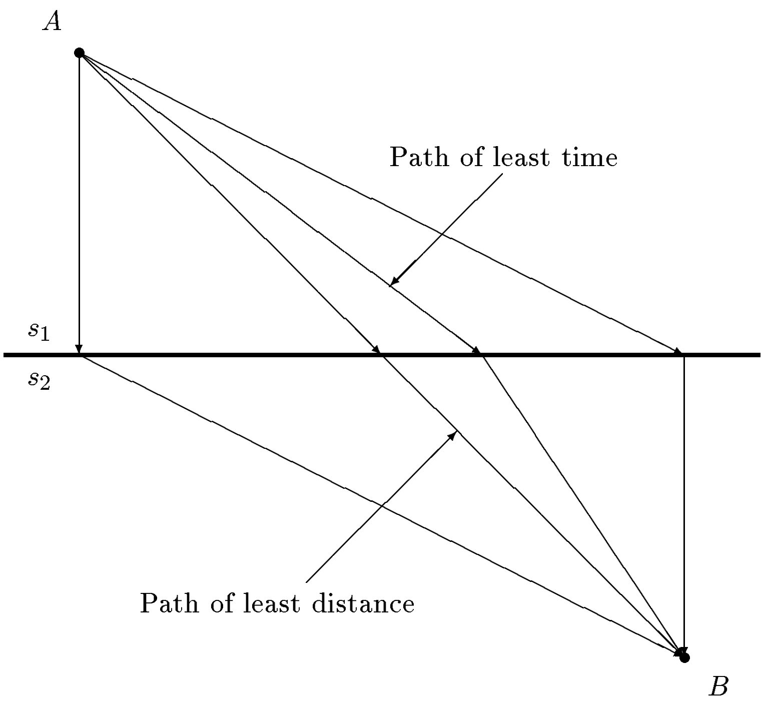


Figure 1.1: Snell's law gives the path of least traveltime from point A to point B. Other paths shown are: least distance through medium 1, least distance through medium 2, and least total distance.

The traveltime functional $\tau^*(\mathbf{s})$ is stationary with respect to small variations in the path $P^*(s)$.

1.3 Snell's Law

Snell's law is a consequence of Fermat's principle [Born and Wolf, 1980]. This result can be derived using a simple geometric argument based on stationarity of the traveltime functional, illustrated in Figures 5.1 and 5.2. The well-known result is

$$\boxed{s_1 \sin\theta_1 \quad = \quad s_2 \sin\theta_2, \qquad \text{(Snell's law)}} \tag{1.4}$$

where θ_1 and θ_2 denote the angles of the ray path from the normal to the boundary that separates the two regions.

A thorough discussion of the physical significance of Fermat's principle and its relation to Snell's law may be found in *The Feynman Lectures* [Feynman, Leighton, and Sands, 1963]. Relations to the *principle of least action* and *Hamilton-Jacobi theory* are discussed by Goldstein [1950], Boorse and Motz [1966], and Born and Wolf [1980]. An interesting and less technical account is given by Gleick [1993].

The main point to be made here is that Snell's law is special. There are various assumptions that go into the derivation such as: the points A and B are far from the boundary, the two media on either side of the boundary are homogeneous with constant isotropic slowness, etc. For general imaging problems, the underlying media may be very complex and it may not be convenient to apply Snell's law. A standard ray tracing method may fail in some circumstances, so it is preferable to consider more robust methods of determining approximate ray paths and traveltimes. Such methods will be discussed in some detail in Chapter 5.

1.4 Seismic Inversion and Tomography

Suppose we have a set of observed traveltimes, $t_1, \ldots, t_m$, from m source-receiver pairs in a medium of slowness $s(\mathbf{x})$. Let P_i be the Fermat ray path connecting the ith source-receiver pair. Neglecting observational errors, we can write

$$\int_{P_i} s(\mathbf{x})\, dl^{P_i} = t_i, \quad i = 1, \ldots, m. \tag{1.5}$$

Given a block model of slowness, let l_{ij} be the length of the ith ray path through the jth cell:

$$l_{ij} = \int_{P_i \cap \text{cell}_j} dl^{P_i}. \tag{1.6}$$

Given a model with n cells, Eq. (1.5) can then be written

$$\sum_{j=1}^{n} l_{ij} s_j = t_i, \quad i = 1, \ldots, m. \tag{1.7}$$

Note that for any given i, the ray-path lengths l_{ij} are zero for most cells j, as a given ray path will in general intersect only a few of the cells in the model. Figure 1.2 illustrates ray path segmentation for a 2-D cell model.

We can rewrite (1.7) in matrix notation by defining the column vectors $\mathbf{s}$ and $\mathbf{t}$ and the matrix $\mathbf{M}$ as follows:

$$\mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix}, \quad \mathbf{t} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_m \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} l_{11} & l_{12} & \cdots & l_{1n} \\ l_{21} & l_{22} & \cdots & l_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ l_{m1} & l_{m2} & \cdots & l_{mn} \end{pmatrix}. \tag{1.8}$$

Equation (1.7) then becomes the basic equation of forward modeling for ray equation analysis:

$$\boxed{\mathbf{Ms} \;=\; \mathbf{t}.} \tag{1.9}$$

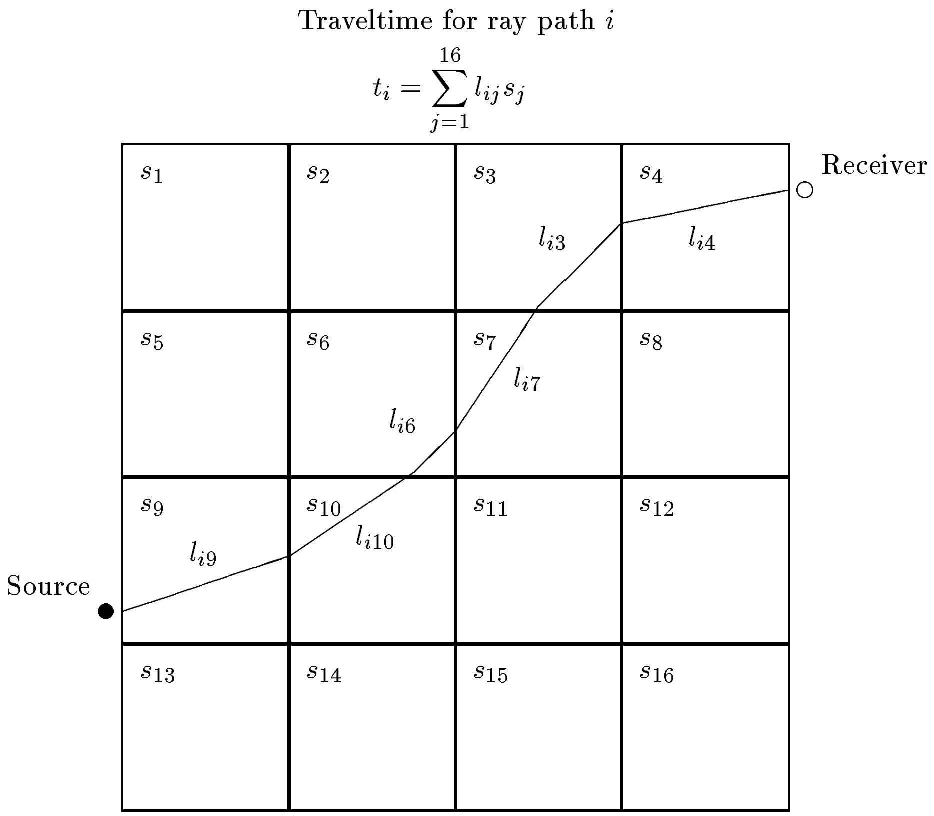


Figure 1.2: Schematic illustration of ray paths through a cell slowness model.

Note that equation (1.9) may be viewed as a numerical approximation to equation (1.3), *i.e.*, it is just a discretized form of the equation. We will study equation (1.9) at great length. Equation (1.9) may be used for any set of ray paths, whether those ray paths minimize (1.3) or not. If the ray paths used to form the matrix $\mathbf{M}$ actually are minimizing ray paths, then we should keep in mind that $\mathbf{M}$ is then implicitly a function of $\mathbf{s}$.

The methods developed apply to both two-dimensional and three-dimensional imaging applications. We use the term *inversion* for either 2-D or 3-D applications. When discussing only 2-D applications, we will use the term *tomography*. The prefix *tomo* is Greek for *slice* and therefore implies a 2-D reconstruction. Similarly, the cells in 2-D are sometimes called *pixels* since they are 2-D picture elements, while the cells or blocks in 3-D are sometimes called *voxels* since they are 3-D volume elements. Thus, traveltime tomography reconstructs the slowness values in model pixels (or cells), while traveltime inversion reconstructs the values in model voxels (or blocks or cells).

1.5 Backprojection for Bent Rays

The term *backprojection* will be used to mean a one-step approximate inversion scheme to solve (1.9) for the slowness vector $\mathbf{s}$.

The physical idea behind backprojection is this: If we have measured a traveltime t_i along the ith ray-path and we know the total path length along that path is $L_i = \sum_{j=1}^{n} l_{ij}$, then the path-average slowness along the ray path is

$$\langle s \rangle_i = \frac{t_i}{L_i} = \frac{\int_{P_i} s\, dl^{P_i}}{\int_{P_i} dl^{P_i}}. \tag{1.10}$$

The ith ray path passes through the jth cell if $l_{ij} > 0$ and misses the cell if $l_{ij} = 0$. An estimate of the slowness in cell j can be obtained by finding the mean of the path-average slownesses $\langle s \rangle_i$ for all the rays that do traverse the jth cell. This averaging process is *backprojection*: accumulating (summing) all the path-averages and then dividing by the total number of contributions.

We can formalize this procedure by introducing the sign function such that $\mathrm{sgn}(l_{ij}) = 1$ if $l_{ij} > 0$ and $\mathrm{sgn}(l_{ij}) = 0$ if $l_{ij} = 0$. Then, the total number of ray paths passing through the jth cell is $N_j = \sum_{i=1}^{m} \mathrm{sgn}(l_{ij})$, and the mean slowness is

$$s_j \simeq N_j^{-1} \sum_{i=1}^{m} \mathrm{sgn}(l_{ij}) \frac{t_i}{L_i}. \tag{1.11}$$

We call (1.11) the formula for *elementary backprojection*. Variations on this formula are explored in the PROBLEMS.

The formula (1.11) provides a fast but inaccurate estimate of the cell slowness, based on available data. The formula is so simple that it can easily be evaluated by hand or using a pocket calculator, whereas the more accurate methods of inverting for slowness (see Chapter 4) are not really practical unless modern computational facilities are available.

There are many possible modifications of the physical arguments for backprojection formulas. Each new choice seemingly leads to a new estimate, showing that the interpretation of these estimates is ambiguous and these methods should not be used for work requiring high accuracy reconstruction. For example, suppose that the slowness in cell j is determined by a weighted sum of the products $l_{ij}t_i$. This approach seems to be an improvement over the preceding one, since it still accounts for our expectation that the ith ray path should not contribute to the estimate if $l_{ij} = 0$ but in addition weights a ray path more heavily when it samples more of the cell. The formula then becomes

$$s_j \simeq \sum_{i=1}^{m} w_i l_{ij} t_i, \tag{1.12}$$

where some choice of the weights w_i must be made. Substituting (1.12) into (1.9), we find that

$$\sum_{j=1}^{n} l_{ij} s_j = \sum_{k=1}^{m} w_k \left(\sum_{j=1}^{n} l_{ij} l_{kj} \right) t_k \simeq t_i. \tag{1.13}$$

Our initial choice of the form of the weights was too simple to allow a rigorous solution to be developed this way, but an approximate solution is obtained by choosing

$$w_i = \left(\sum_{j=1}^{n} l_{ij}^2 \right)^{-1}. \tag{1.14}$$

Defining a diagonal matrix $\overline{\mathbf{D}}$ whose diagonal elements are given by

$$\overline{\mathbf{D}}_{ii} = (\mathbf{M}\mathbf{M}^T)_{ii} = \sum_{j=1}^{n} l_{ij}^2, \tag{1.15}$$

we see that (1.12) and (1.14) lead to the estimate

$$\mathbf{s} \simeq \overline{\mathbf{D}}^{-1}\mathbf{M}^T\mathbf{t}. \tag{1.16}$$

This result is not so simple as the formula (1.11) for elementary backprojection, but the implied computations are still manageable without using very sophisticated computers.

Formulas (1.11), (1.16), and numerous variations are all *backprojection* formulas. As we attempt to compute accurate inverses, these backprojection formulas will frequently reappear as the starting point of rigorous iteration schemes.

Problems

Problem 1.5.1 *Define the hit matrix* $\mathbf{H}$ *such that* $H_{ij} = sgn(l_{ij})$ *and the diagonal matrices* $\mathbf{N}$ *such that* $N_{jj} = \sum_{i=1}^{m} sgn(l_{ij})$*, and* $\mathbf{L}$*, such that* $L_{ii} = \sum_{j=1}^{n} l_{ij}$*. Show that (1.11) is equivalent to*

$$\mathbf{s} \simeq \mathbf{N}^{-1}\mathbf{H}^T\mathbf{L}^{-1}\mathbf{t}.$$

Problem 1.5.2 *Elementary backprojection can be applied either to wave slowness as in (1.11) or to wave velocity. Then,*

$$v_j = \frac{1}{s_j} \simeq N_j^{-1} \sum_{i=1}^{m} sgn(l_{ij}) \frac{L_i}{t_i}. \tag{1.17}$$

It is a general result that the harmonic mean of a set of numbers $\{x_i\}$ *(given by* $x_{harm}^{-1} = N^{-1}\sum_{i=1}^{N} x_i^{-1}$*) is always less than or equal to the mean (*$x_{mean} = N^{-1}\sum_{i=1}^{N} x_i$*), i.e.,*

$$x_{harm} \le x_{mean}. \tag{1.18}$$

Use this result to determine a general relation between the backprojection formulas (1.11) and (1.17).

PROBLEM 1.5.3 *Consider an elementary backprojection formula based on weighting the average ray slowness $\langle s \rangle_i = t_i/L_i$ with respect to the path length l_{ij} – instead of with the cell hit factor sgn(l_{ij}). Show that an alternative to (1.11) is then given by*

$$\mathbf{s} \simeq \mathbf{C}^{-1}\mathbf{M}^T\mathbf{L}^{-1}\mathbf{t}, \tag{1.19}$$

where $\mathbf{C}$ *is the diagonal matrix whose diagonal elements are given by* $C_{jj} = \sum_{i=1}^{m} l_{ij}$. *Define a corresponding estimate for the velocity, then use (1.18) to obtain a relation between these two estimates. Using (1.19), show that*

$$\sum_{i=1}^{m} (\mathbf{Ms})_i = \sum_{i=1}^{m} t_i,$$

*demonstrating that this backprojection formula is an unbiased estimator (*i.e.*, average predicted traveltime agrees with average measured traveltime). What can be said about the bias of the backprojection formula for velocity?*

PROBLEM 1.5.4 *Construct a backprojection formula by supposing that the slowness may be determined in the form*

$$s_j \simeq \sum_{i=1}^{m} w_i l_{ij} \frac{t_i}{L_i}$$

and by finding a useful set of weights w_i. *Compare the resulting formula to (1.16).*

1.6 Diffraction Tomography and Full Waveform Inversion

Geophysical diffraction tomography [Devaney, 1983; Harris, 1987; Wu and Toksöz, 1987; Lo, Duckworth, and Toksöz, 1990] consists of a collection of methods including Born [Born, 1926; Newton, 1966] and Rytov [Rytov, 1937; 1938; Keller, 1969; Born and Wolf, 1980] inversion that make use of full waveform information in seismic data. An example of real crosswell transmission data is shown in Figure 1.3. Successful inversion of real data has also been performed using both microwave and ultrasonic diffraction tomography [Tabbara, Duchêne, Pichot, Lesslier, Chommeloux, and Joachimowicz, 1988]. Instead of using only the first arrival traveltimes as the data in the inversion, amplitude and phase in the waveform following the first arrival are used. It is necessary to use full waveform information whenever the wavelengths of the probing waves are comparable in size to the anomalies present in the region to be imaged. The ray approximation is strictly valid only for very high frequencies or equivalently for wavelengths "small" compared with the size of the anomalies (there will be a discussion of the eikonal equation in a Chapter 5). The term "small" is subject to interpretation, but extensive experience with the asymptotic analysis of wave propagation problems [Bleistein, 1984] has shown that, if the largest wavelength found in bandlimited data is $\lambda_{\max}$, then the ray approximation is valid when the anomalies are of size $\simeq 3\lambda_{\max}$ or larger. If this relationship is violated by the tomographic experiment, then diffraction tomography should play an important role in the reconstruction.

Figure 1.3: Example of real data for crosswell seismic tomography showing the result of a single fan beam with a source at 700 feet in one borehole and receivers spaced 10 feet apart in the other hole. (Courtesy of CONOCO Inc.)

Diffraction tomography is both more and less ambitious than traveltime tomography. As it exists today, diffraction tomography is a strictly linear tomography method. A starting model is required. The usual starting model is a constant, because this method requires a comparison between predicted wave fields (planewaves for a constant background) and the measured wave fields. If a nonconstant starting model is used, then "distorted wave" diffraction tomography may be applied to the differences between the computed complex wave field and the measured wave field. In either case, it is possible to prove convergence of diffraction tomography to a solution of the inversion problem if the comparison wave field differs by a small enough amount from the measured wave field. Thus, diffraction tomography is one type of *linear tomography* — although in this case the "rays" may not be straight, it is still linear in the mathematical sense that the perturbations from the starting model must be very small in some sense. So diffraction tomography is less ambitious than traveltime tomography in the sense that it is inherently limited to be *linear tomography.*[4]

On the other hand, diffraction tomography is more ambitious than traveltime tomography, because it tries to make use of more of the information contained in the measured seismic waveforms. There are serious problems involved with this process, because amplitude information can be ambiguous. It is well known that wave attenuation, scattering, three-dimensional geometrical spreading, mode conversion, and reflection/transmission effects can all mimic each other — producing similar effects in the waveform. Thus, to be successful, diffraction tomography must achieve the ambitious goal of solving all of these problems simultaneously for real data. To date, most of the work in diffraction tomography has been limited to two-dimensional inversions and the most successful applications have used ultrasound for medical imaging or microwaves for imaging metallic reinforcements in concrete.

I view diffraction tomography and full waveform inversion as challenging long-term goals. The wave slowness results obtained from our traveltime tomography analysis may be used as the required starting model for "distorted wave" diffraction tomography. So the potential benefits of diffraction tomography provide an additional motivation for improving traveltime inversion and tomography.

1.7 Linear vs Nonlinear Inversion and Tomography

We now define three problems in the context of Eq. (1.9). Each of these problems will be studied at some length in this book.

In the *forward* problem, we are given $\mathbf{s}$; the goal is to determine $\mathbf{M}$ and $\mathbf{t}$. This entails computing the ray path between each source and receiver (*e.g.*, using a ray tracing algorithm) and then computing the traveltime integral along each path.

In *linear tomography or inversion* problems, we are given $\mathbf{M}$ and $\mathbf{t}$; the objective is to determine $\mathbf{s}$. The assumption here is that the ray paths are known *a priori*, which is justified under a linear approximation that ignores the dependence of the ray paths on the slowness distribution. Typically, the ray paths are assumed to be straight lines connecting

[4] An iterative method for diffraction tomography has been proposed recently by Ladas and Devaney [1991]; a nonlinear least-squares approach to full waveform inversion has been proposed by Tarantola and Valette [1982] and Tarantola [1984]. Such methods are "nonlinear" in the sense used here.

sources and receivers, adding a second connotation to the term *linear.* Linear tomography is commonly practiced in medical imaging and in many geophysical situations as well.

In *nonlinear tomography or inversion* problems, we are given only $\mathbf{t}$ (along with the source and receiver locations); the goal is to infer $\mathbf{s}$, and (for most of the methods considered) incidently $\mathbf{M}$. In this problem, the dependence of ray paths on the slowness distribution strongly influences the design of the inversion algorithm. Nonlinear inversion is required for problems with significant slowness variations across the region of interest, including many seismic inversion problems. The ray paths in such media will show large curvature (*i.e.*, be nonlinear) which cannot be known before the inversion process begins.

Linear tomography and inversion problems can be solved approximately using *backprojection* techniques (see Section 1.5). Linear inversion problems can also be solved more accurately using a variety of optimization techniques. In the standard least-squares method (see Section 3), for example, the normal solution for $\mathbf{s}$ is expressed analytically as

$$\hat{\mathbf{s}} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{t}, \tag{1.20}$$

assuming the matrix inverse exists. If the inverse does not exist, then (1.20) must be *regularized.* Typically, regularization is accomplished by adding a positive matrix to $\mathbf{M}^T\mathbf{M}$ and replacing the singular inverse in (1.20) by the inverse of the modified matrix.

For nonlinear inversion, an iterative algorithm is generally needed to find an approximate solution $\hat{\mathbf{s}}_b$. The basic structure of such an algorithm (see Figure 1.4) is as follows:

1. Set $\hat{\mathbf{s}}_b$ to a given initial model (a constant or the previously best-known geological model).

2. Compute the ray-path matrix $\mathbf{M}$ and traveltimes $\hat{\mathbf{t}}_b$ for $\hat{\mathbf{s}}_b$ and set $\Delta\mathbf{t} = \mathbf{t} - \hat{\mathbf{t}}_b$.

3. If $\Delta\mathbf{t}$ is sufficiently small, output $\hat{\mathbf{s}}_b$ and stop.

4. Find a model correction $\widehat{\Delta\mathbf{s}}$ as the solution to the linear inversion problem: $\mathbf{M}\widehat{\Delta\mathbf{s}} = \Delta\mathbf{t}$.

5. Update $\hat{\mathbf{s}}_b$ to the new model obtained by adding the model correction $\widehat{\Delta\mathbf{s}}$ to the previous model $\hat{\mathbf{s}_b}$.

6. Return to Step 2.

This algorithm looks very reasonable and in fact sometimes it actually works! But not always. For models with low slowness contrasts, the algorithm will converge to a sensible result. When the method fails, the failure mode is usually a divergence to a highly oscillatory model. *Ad hoc* procedures to reduce the possible range of slowness values and to guarantee a high degree of smoothness in the reconstructed model have commonly been introduced to deal with this instability. Such smoothness constraints come from external considerations (like the class of models in which we want the solution to lie), not from the data. But a really satisfactory method of stabilizing the iteration scheme based on information in the data itself has been lacking.

Analyzing the algorithm, we see that there are really only two significant calculations contained in it. Step 2 is just the solution of the *forward problem* for $\hat{\mathbf{s}}_b$. This step should not introduce any instability, since it can be performed essentially as accurately as desired (if the computing budget is large enough and the computers fast enough). Step 4, on the other hand, is a *linear inversion* step imbedded in a nonlinear algorithm. We should be skeptical of this step. Linear inversion implicitly assumes that the updated model (after adding the model correction) is not so different from the previous model that the ray-path matrix $\mathbf{M}$ should change significantly from one iteration to the next. If this implicit assumption is violated, then this step is not justified, and steps 4 and/or 5 in the algorithm must be modified.

Feasibility analysis supplies a set of rigorous physical constraints on the reconstruction process. Experience has shown that constraints on smoothness or limits on the maximum and minimum values of the model are generally not needed if feasibility constraints are applied.

In the Chapters that follow, these problems will be analyzed in some detail, and several methods of stabilizing the nonlinear inversion problem will be developed.

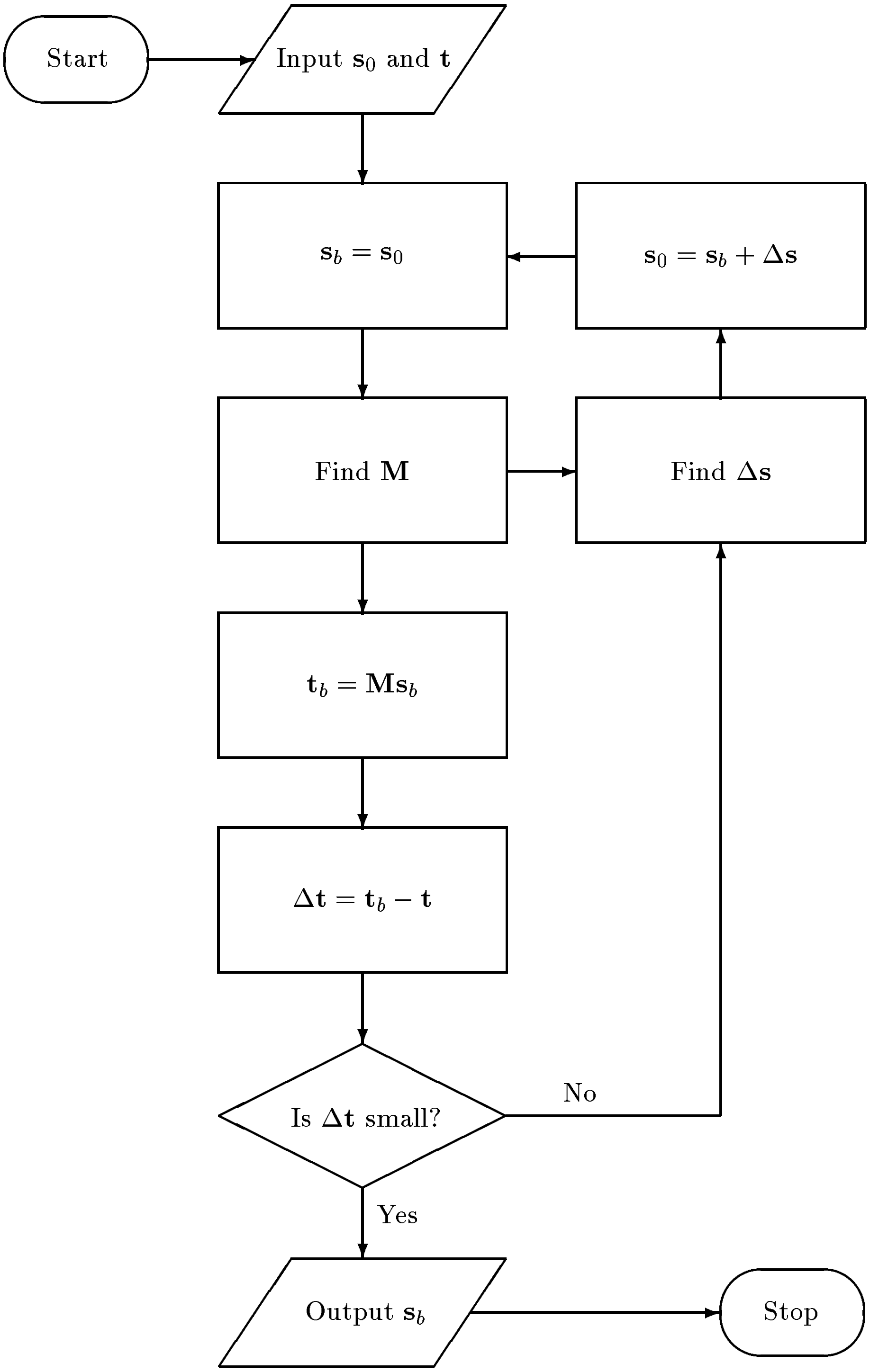


Figure 1.4: Iterative algorithm for traveltime inversion.

Chapter 2

Feasibility Analysis for Traveltime Inversion

The idea of using feasibility constraints in nonlinear programming problems is well established [Fiacco and McCormick, 1990]. However, it has only recently been realized that physical principles such as Fermat's principle actually lead to rigorous feasibility constraints for nonlinear inversion problems [Berryman, 1991]. The main practical difference between the standard analysis in nonlinear programming and the new analysis in nonlinear inversion is that, whereas the functions involved in nonlinear programming are often continuous, differentiable, and relatively easy to compute explicitly, the functionals in nonlinear inversion (*e.g.*, the traveltime functional) need not be continuous or differentiable and, furthermore, are very often comparatively difficult to compute. Feasibility constraints for inversion problems are implicit, rather than explicit.

We present the rigorous analysis here in a general setting, because it is actually quite easy to understand once we have introduced the concepts of convex function and convex set. This analysis is important because it will help to characterize the solution set for the inversion problem, and it will help to clarify questions about local and global minima of the inverson problem.

2.1 Feasibility Constraints Defined

Equation (1.5) assumes that P_i is a Fermat (least-time) path and leads to the equalities summarized in the vector-matrix equation $\mathbf{Ms} = \mathbf{t}$. Now let us suppose instead that P_i is a trial ray path which may or may not be the least-time path. Fermat's principle allows us to write

$$\int_{P_i} s(\mathbf{x}) dl^{P_i} \geq t_i, \tag{2.1}$$

where now t_i is the measured traveltime for source-receiver pair i. When we discretize (2.1) for cell or block models and all ray paths i, the resulting set of m inequalities may be written as

$$\mathbf{Ms} \geq \mathbf{t}. \tag{2.2}$$

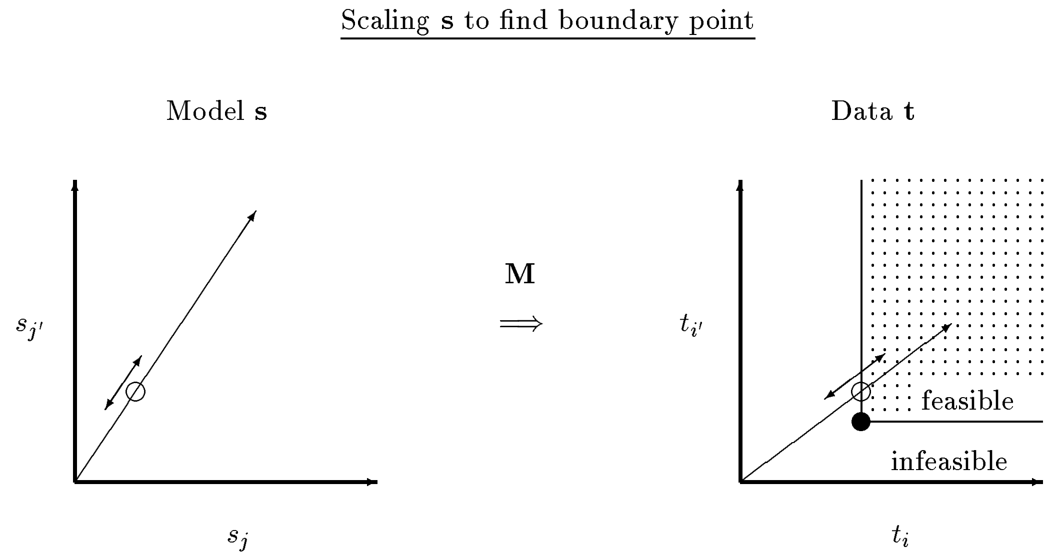


Figure 2.1: Feasible part of the model space is determined implicitly by the feasible part of the data space.

Equations (2.1) and (2.2) can be interpreted as a set of inequality constraints on the slowness model $\mathbf{s}$. When $\mathbf{s}$ obeys these m constraints, we say that $\mathbf{s}$ is *feasible*. When any of the constraints is violated, we say $\mathbf{s}$ is *infeasible*. The set of inequalities collectively will be called the *feasibility constraints*.

The concept of the feasibility constraint is quite straightforward in nonlinear programming problems [Fiacco and McCormick, 1990] whenever the constraints may be *explicitly* stated for the solution vector. However, in our inversion problems, an additional computation is required. Figure 2.1 shows that the feasibility constraints are *explicit* for the traveltime data vector, but they are only *implicit* (*i.e.*, they must be computed) for the slowness vector. This added degree of complication is unavoidable in the inversion problem, but nevertheless it is also very easily handled computationally with only very minor modifications of the usual nonlinear inversion algorithms.

2.2 Quick Review of Convexity

Here we define some mathematical concepts [Hardy, Littlewood, and Pólya, 1934] which will facilitate the discussion and analysis of feasible models. In the following, let $\mathcal{S}$ denote a linear vector space.

DEFINITION 2.2.1 (CONVEX SET) *A set $\mathcal{A} \subseteq \mathcal{S}$ is convex if, for every $\mathbf{s}_1, s_2 \in \mathcal{A}$ and every number $\lambda \in [0,1]$, we have $\lambda \mathbf{s}_1 + (1-\lambda)\mathbf{s}_2 \in \mathcal{A}$.*[1]

Examples of convex sets are

[1] $\mathcal{A} \subseteq \mathcal{S}$ means $\mathcal{A}$ is a *subset* of $\mathcal{S}$.

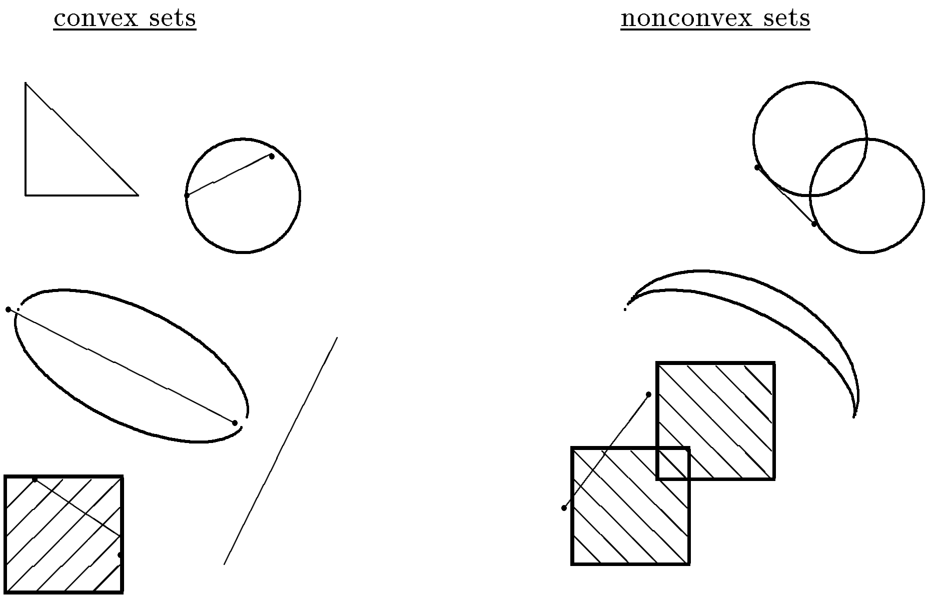


Figure 2.2: Examples of convex and nonconvex sets.

1. $\mathbf{R}$ (the real numbers).
2. $\mathbf{R}_+$ (the positive real numbers).
3. The positive n-tant $\mathbf{R}^n_+$; *i.e.*, the set of n-dimensional vectors whose components are all positive.
4. $C_+(\mathbf{R}^3)$ (the set of positive, continuous functions, $s(\mathbf{x}) > 0$, where $\mathbf{x} \in \mathbf{R}^3$).
5. A closed interval $[a, b]$ in $\mathbf{R}$.[2]
6. A hyperplane in $\mathbf{R}^n$; *i.e.*, vectors $\mathbf{s}$ obeying $\mathbf{c}^T\mathbf{s} = \gamma$ where $\mathbf{c}$ is a vector and γ is a scalar.[3]
7. The interior of a circular disk in 2-space; *i.e.*, points (x, y) obeying

$$(x-a)^2 + (y-b)^2 < c^2$$

   for real a, b and c.

[2] $[a, b]$ means the set of numbers x such that $a \le x \le b$.
[3] T used in a superscript means to take the transpose of a vector or a matrix.

We note that $\mathbf{R}_+^n$ (example 3) defines the set of n-dimensional block slowness models such that the slowness of each cell is a positive number. $C_+(\mathbf{R}^3)$ (example 4) is the space of positive, continuous 3-D slowness distributions.

PROPOSITION 2.2.1 *If $\mathcal{A}_1$ and $\mathcal{A}_2$ are convex sets, then $\mathcal{A}_1 \cap \mathcal{A}_2$ is a convex set.*[4]

Proof: If $\mathcal{A}_1 \cap \mathcal{A}_2$ is empty, it is convex by default (one cannot find $\mathbf{s}_1$, $\mathbf{s}_2$ and λ which disobey the definition).

Assume the intersection is not empty and let $\mathbf{s} = \lambda\mathbf{s}_1 + (1-\lambda)\mathbf{s}_2$ for some $0 \le \lambda \le 1$ and $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{A}_1 \cap \mathcal{A}_2$. Since $\mathcal{A}_1$ and $\mathcal{A}_2$ are each convex, we must have $\mathbf{s} \in \mathcal{A}_1$ and $\mathbf{s} \in \mathcal{A}_2$. Consequently, $\mathbf{s} \in \mathcal{A}_1 \cap \mathcal{A}_2$. ∎

DEFINITION 2.2.2 (CONE) *A set $\mathcal{A} \subseteq \mathcal{S}$ is a cone if, for every $\mathbf{s} \in \mathcal{A}$ and every number $\gamma > 0$, we have $\gamma\mathbf{s} \in \mathcal{A}$.*

Examples 1–4 of convex sets given above are also examples of cones. We infer that the set of positive slowness models (block or continuous) is convex and conical (a *convex cone*).

DEFINITION 2.2.3 (LINEAR FUNCTIONAL) *The functional $f\colon \mathcal{S} \to \mathbf{R}$ is linear if, for all $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{S}$ and real numbers λ_1, λ_2, we have*[5]

$$f(\lambda_1\mathbf{s}_1 + \lambda_2\mathbf{s}_2) = \lambda_1 f(\mathbf{s}_1) + \lambda_2 f(\mathbf{s}_2). \tag{2.3}$$

Considering $\lambda_1 = \lambda_2 = 0$, note that a linear functional necessarily vanishes at the origin. We will also need to consider the broader class of functionals that are linear except for a shift at the origin.

DEFINITION 2.2.4 (SHIFTED LINEAR FUNCTIONAL) *The functional $f\colon \mathcal{S} \to \mathbf{R}$ is shifted linear if the functional*

$$g(\mathbf{s}) \equiv f(\mathbf{s}) - f(0) \tag{2.4}$$

is linear.

DEFINITION 2.2.5 (CONVEX FUNCTIONAL) *Let $\mathcal{A}$ be a convex set in $\mathcal{S}$. A functional $f\colon \mathcal{A} \to \mathbf{R}$ is convex if, for every $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{A}$ and number $\lambda \in [0,1]$, we have*

$$f(\lambda\mathbf{s}_1 + (1-\lambda)\mathbf{s}_2) \le \lambda f(\mathbf{s}_1) + (1-\lambda) f(\mathbf{s}_2). \tag{2.5}$$

DEFINITION 2.2.6 (CONCAVE FUNCTIONAL) *A functional f is concave if $(-f)$ is convex.*

DEFINITION 2.2.7 (HOMOGENEOUS FUNCTIONAL) *Let $\mathcal{A}$ be a cone in $\mathcal{S}$. A functional $f\colon \mathcal{A} \to \mathbf{R}$ is homogeneous if, for every $\mathbf{s} \in \mathcal{A}$ and $\gamma > 0$, we have*

$$f(\gamma\mathbf{s}) = \gamma f(\mathbf{s}). \tag{2.6}$$

[4] $\mathcal{A}_1 \cap \mathcal{A}_2$ denotes the *intersection* of sets $\mathcal{A}_1$ and $\mathcal{A}_2$ (i.e., the set of elements common to both sets).

[5] $f\colon \mathcal{S} \to \mathbf{R}$ means: the function f which maps each element of the set $\mathcal{S}$ to an element of the set $\mathbf{R}$.

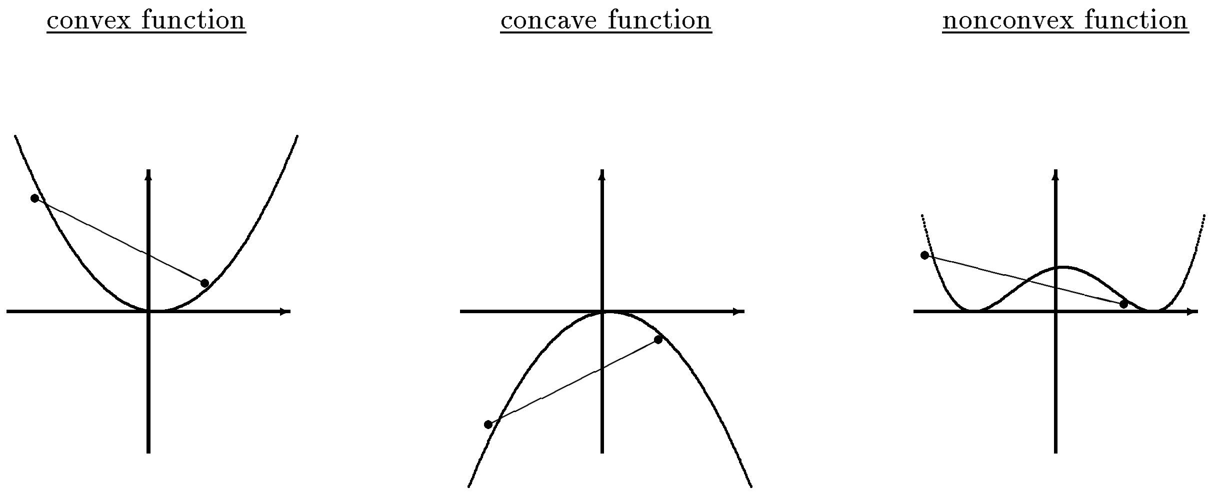


Figure 2.3: Examples of convex, concave, and nonconvex functionals.

It should be clear that every linear functional is also convex, concave, and homogeneous.

Problems

PROBLEM 2.2.1 *Is the union of two convex sets convex? Give an example.*

PROBLEM 2.2.2 *Decide whether the following sets are convex:*

1. *the interior of a cube;*
2. *the interior of a tetrahedron;*
3. *the interior of a rectangular prism;*
4. *any compact region in n-dimensional vector space, all of whose boundaries are hyperplanes;*
5. *the interior of an ellipsoid;*
6. *the interior of an n-dimensional sphere;*
7. *the interior of two partially overlapping spheres;*
8. *the interior of a boomerang;*
9. *the cheesy part of a swiss cheese;*
10. *the interior of any object having a rough surface.*

PROBLEM 2.2.3 *If* $\mathbf{c}$ *is an arbitrary* n*-vector, which of the following functionals is linear in* $\mathbf{s}$*?*

1. $\mathbf{c}^T\mathbf{s}$;
2. $\mathbf{s}\mathbf{s}^T\mathbf{c}$;
3. $(\mathbf{s}-\mathbf{c})^T(\mathbf{s}-\mathbf{c})$.

PROBLEM 2.2.4 *Show that a linear functional is convex, concave, and homogeneous.*

PROBLEM 2.2.5 *Is a shifted linear functional convex, concave, and/or homogeneous?*

PROBLEM 2.2.6 *Are all cones convex? If not, give an example of a nonconvex cone.*

2.3 Properties of Traveltime Functionals

PROPOSITION 2.3.1 τ^P *is a linear functional.*

The proof of this stems from the fact that integration is a linear functional of the integrand. Since it is linear, it follows that τ^P is also convex, concave, and homogeneous.

PROPOSITION 2.3.2 τ^* *is a homogeneous functional.*

Proof: Given $\gamma > 0$ we have

$$\tau^*(\gamma\mathbf{s}) = \min_P \tau^P(\gamma\mathbf{s}). \tag{2.7}$$

Using the linearity of τ^P,

$$\tau^*(\gamma\mathbf{s}) = \min_P \gamma\tau^P(\mathbf{s}) = \gamma\min_P \tau^P(\mathbf{s}) = \gamma\tau^*(\mathbf{s}). \quad \blacksquare \tag{2.8}$$

PROPOSITION 2.3.3 τ^* *is a concave functional.*

Proof: Given slowness models $\mathbf{s}_1$ and $\mathbf{s}_2$ and $\lambda \in [0,1]$, let $\mathbf{s} = \lambda\mathbf{s}_1 + (1-\lambda)\mathbf{s}_2$. Letting $P^*(\mathbf{s})$ be the Fermat ray path for $\mathbf{s}$, we have

$$\tau^*(\mathbf{s}) = \tau^{P^*(\mathbf{s})}(\mathbf{s}). \tag{2.9}$$

The linearity of τ^P then implies

$$\tau^*(\mathbf{s}) = \lambda\tau^{P^*(\mathbf{s})}(\mathbf{s}_1) + (1-\lambda)\tau^{P^*(\mathbf{s})}(\mathbf{s}_2). \tag{2.10}$$

Since τ^* minimizes τ^P for any fixed model, it must be the case that $\tau^{P^*(\mathbf{s})}(\mathbf{s}_1) \geq \tau^*(\mathbf{s}_1)$ and similarly for $\mathbf{s}_2$. Further, λ and $(1-\lambda)$ are non-negative. Therefore, (2.10) implies

$$\tau^*(\mathbf{s}) \geq \lambda\tau^*(\mathbf{s}_1) + (1-\lambda)\tau^*(\mathbf{s}_2). \quad \blacksquare \tag{2.11}$$

2.4 Feasibility Sets

Given the set of observed traveltimes, t_i for $i = 1, \ldots, m$, we define two sets of models:

DEFINITION 2.4.1 (LOCAL FEASIBILITY SET) *The local feasibility set with respect to a set of trial ray paths* $\mathcal{P} = \{P_1, \ldots, P_m\}$ *and observed traveltimes* $t_1, \ldots, t_m$ *is*

$$\mathcal{F}^{\mathcal{P}} = \{\mathbf{s} \mid \tau_i^P(\mathbf{s}) \geq t_i, \quad \textit{for all} \quad i = 1, \ldots, m\}. \tag{2.12}$$

DEFINITION 2.4.2 (GLOBAL FEASIBILITY SET) *The global feasibility set with respect to the observed traveltimes* $t_1, \ldots, t_m$ *is*

$$\mathcal{F}^* = \{\mathbf{s} \mid \tau_i^*(\mathbf{s}) \geq t_i, \quad \textit{for all} \quad i = 1, \ldots, m\}. \tag{2.13}$$

Now we show that the concavity of τ_i^P and τ_i^* implies the convexity of $\mathcal{F}^{\mathcal{P}}$ and $\mathcal{F}^*$.

THEOREM 2.4.1 $\mathcal{F}^{\mathcal{P}}$ *is a convex set.*

Proof: Suppose $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{F}^{\mathcal{P}}$ and let $\mathbf{s}_\lambda = \lambda \mathbf{s}_1 + (1-\lambda)\mathbf{s}_2$ where $0 \leq \lambda \leq 1$. Since, for each i, τ_i^P is a concave (actually linear) functional, we have

$$\tau_i^P(\mathbf{s}_\lambda) \geq \lambda \tau_i^P(\mathbf{s}_1) + (1-\lambda)\tau_i^P(\mathbf{s}_2). \tag{2.14}$$

(Although equality applies in the present case, the "greater than or equal to" is important in the next proof.) But $\tau_i^P(\mathbf{s}_1), \tau_i^P(\mathbf{s}_2) \geq t_i$ and λ and $(1-\lambda)$ are non-negative. Therefore,

$$\tau_i^P(\mathbf{s}_\lambda) \geq \lambda t_i + (1-\lambda)t_i = t_i. \tag{2.15}$$

Thus, $\mathbf{s}_\lambda \in \mathcal{F}^{\mathcal{P}}$. ∎

THEOREM 2.4.2 $\mathcal{F}^*$ *is a convex set.*

The proof proceeds in analogy with the previous proof, with τ_i^* replacing τ_i^P, but the inequalities come into play this time.

The next theorem follows easily from an analysis of Figure 2.1.

THEOREM 2.4.3 *Given any model* $\mathbf{s}$, *there exists a finite scalalr* $\gamma^* > 0$ *such that* $\gamma\mathbf{s} \in \mathcal{F}^*$ *for all* $\gamma \geq \gamma^*$.

Proof: Let

$$\gamma^* = \max_{k \in \{1, \ldots, m\}} \frac{t_k}{\tau_k^*(\mathbf{s})}. \tag{2.16}$$

For any i, τ_i^* is homogeneous, implying

$$\tau_i^*(\gamma^*\mathbf{s}) = \gamma^*\tau_i^*(\mathbf{s}) = \tau_i^*(\mathbf{s}) \max_k \frac{t_k}{\tau_k^*(\mathbf{s})} \geq \tau_i^*(\mathbf{s}) \frac{t_i}{\tau_i^*(\mathbf{s})} = t_i. \tag{2.17}$$

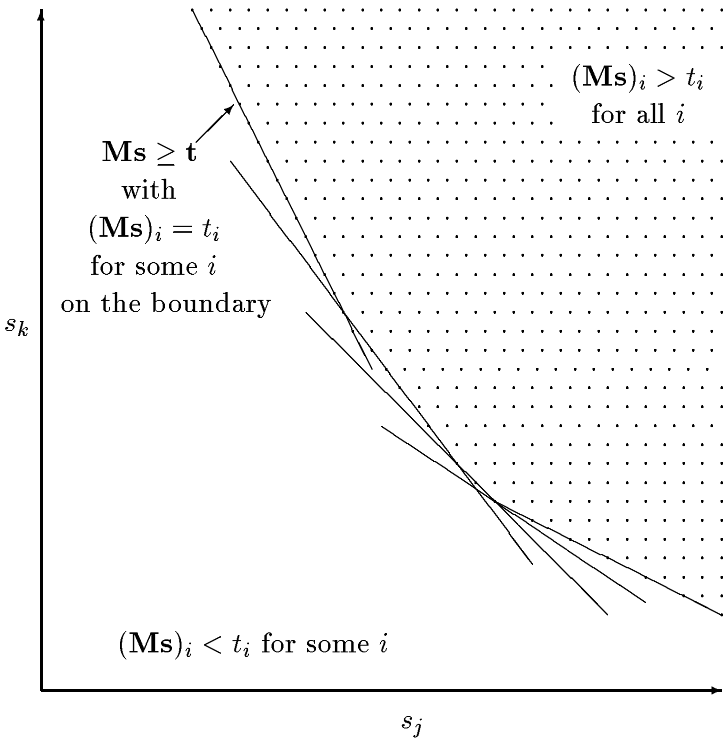


Figure 2.4: The defining conditions for the feasible and infeasible parts of the model space and the boundary separating them.

We see that $\gamma^* \mathbf{s}$ satisfies all the feasibility constraints, so it is in $\mathcal{F}^*$, and so is $\gamma \mathbf{s}$ for any $\gamma > \gamma^*$. ∎

We can decompose $\mathcal{F}^*$ into two parts: its *boundary* and its *interior*. The boundary of $\mathcal{F}^*$, denoted $\partial\mathcal{F}^*$, comprises feasible models $\mathbf{s}$ which satisfy some feasibility constraint with equality, *i.e.*,

$$\partial\mathcal{F}^* = \{\mathbf{s} \in \mathcal{F}^* \mid \tau_i^*(\mathbf{s}) = t_i, \text{ for some } i\}. \tag{2.18}$$

Models in the interior of $\mathcal{F}^*$, denoted $\operatorname{Int}\mathcal{F}^* = \mathcal{F}^* - \partial\mathcal{F}^*$, satisfy all constraints with inequality:

$$\operatorname{Int}\mathcal{F}^* = \{\mathbf{s} \in \mathcal{F}^* \mid \tau_i^*(\mathbf{s}) > t_i, \text{ for all } i\}. \tag{2.19}$$

These characteristics of the feasible set are illustrated in Figure 2.4.

Problems

Problem 2.4.1 *Prove that $\mathcal{F}^*$ is convex.*

PROBLEM 2.4.2 *Define another type of feasibility set* $\mathcal{R}$ *by*

$$\mathcal{R} = \{\mathbf{s} \mid \sum_{i=1}^{m} \tau_i^*(\mathbf{s}) \geq \sum_{i=1}^{m} t_i\}.$$

Is $\mathcal{R}$ *convex? What is the relationship between this set and the other ones defined in this section? (Larger or smaller set?) [relaxed constraints]*

PROBLEM 2.4.3 *Consider the velocity vector space related to the slowness vector space through the nonlinear transform* $\mathbf{v} = (1/s_1, 1/s_2, \ldots, 1/s_n)$. *Suppose there is some reason to expect that all models solving our inversion problem should lie in a part of the velocity vector space satisfying the hyperplane constraint given by*

$$\mathcal{V} = \{\mathbf{v} \mid \mathbf{d}^T\mathbf{v} \geq \beta\},$$

where β *is some positive constant and* $\mathbf{d}$ *is a nonnegative vector. Now, let* $\mathcal{V}'$ *be the set of slownesses corresponding to*

$$\mathcal{V}' = \{\mathbf{s} \mid \mathbf{v} \in \mathcal{V}\}$$

and define the slowness overlap set

$$\mathcal{D} = \mathcal{V}' \cap \mathcal{F}^*.$$

Assuming that the set $\mathcal{D}$ *is not empty, is it convex?*

2.5 Convex Programming for Inversion

We first define convex programming for first-arrival traveltime inversion. Then we present some basic theorems about convex programming in this context.

DEFINITION 2.5.1 *Let* $\Phi(\mathbf{s})$ *be any convex functional of* $\mathbf{s}$. *Then the convex nonlinear programming problem associated with* Φ *is to minimize* $\Phi(\mathbf{s})$ *subject to the global feasibility constraints* $\tau_i^*(\mathbf{s}) \geq t_i$, *for* $i = 1, \ldots, m$.

DEFINITION 2.5.2 *Let*

$$\Psi^{\mathcal{P}}(\mathbf{s}) = \sum_{i=1}^{m} w_i[\tau_i^{P_i}(\mathbf{s}) - t_i]^2 \tag{2.20}$$

for some positive weights $\{w_i\}$ *and some set of ray paths* $\mathcal{P} = \{P_1, \ldots, P_m\}$. *Then, the convex linear programming problem associated with* $\Psi^{\mathcal{P}}$ *is to minimize* $\Psi^{\mathcal{P}}(\mathbf{s})$ *subject to the local feasibility constraints* $\tau_i^{P_i}(\mathbf{s}) \geq t_i$, *for* $i = 1, \ldots, m$.

THEOREM 2.5.1 *Every local minimum* $\mathbf{s}^*$ *of the convex nonlinear programming problem associated with* $\Phi(\mathbf{s})$ *is a global minimum.*

THEOREM 2.5.2 *Every local minimum* $\mathbf{s}^*$ *of the convex linear programming problem associated with* $\Psi^{\mathcal{P}}(\mathbf{s})$ *is a global minimum.*

Proof: This proof follows one given by Fiacco and McCormick [1990]. Let $\mathbf{s}^*$ be a local minimum. Then, by definition, there is a compact set $\mathcal{C}$ such that $\mathbf{s}^*$ is in the interior of $\mathcal{C} \cap \mathcal{F}^*$ and

$$\Phi(\mathbf{s}^*) = \min_{\mathcal{C}\cap\mathcal{F}^*} \Phi(\mathbf{s}). \tag{2.21}$$

If $\mathbf{s}$ is any point in the feasible set $\mathcal{F}^*$ and $0 \le \lambda \le 1$ such that $\mathbf{s}^*_\lambda \equiv \lambda\mathbf{s}^* + (1-\lambda)\mathbf{s}$ is in $\mathcal{C} \cap \mathcal{F}^*$, then

$$\Phi(\mathbf{s}) \ge \frac{\Phi(\mathbf{s}^*_\lambda) - \lambda\Phi(\mathbf{s}^*)}{1-\lambda} \ge \frac{\Phi(\mathbf{s}^*) - \lambda\Phi(\mathbf{s}^*)}{1-\lambda} = \Phi(\mathbf{s}^*). \tag{2.22}$$

The first step of (2.22) follows from the convexity of Φ and the second from the fact that $\mathbf{s}^*$ is a minimum in $\mathcal{C} \cap \mathcal{F}^*$. Convexity of $\mathcal{F}^*$ guarantees that the convex combination $\mathbf{s}^*_\lambda$ lies in the feasible set. This completes the proof of the first theorem.

The proof of the second theorem follows that of the first once we have shown that the function $\Psi^{\mathcal{P}}$ is convex. Consider a term of $\Psi^{\mathcal{P}}$

$$\begin{aligned}
[\tau_i^{P_i}(\lambda\mathbf{s}_1 + (1-\lambda)\mathbf{s}_2) - t_i]^2 &= [\lambda\tau_i^{P_i}(\mathbf{s}_1) + (1-\lambda)\tau_i^{P_i}(\mathbf{s}_2) - t_i]^2 \\
&= \lambda[\tau_i^{P_i}(\mathbf{s}_1) - t_i]^2 + (1-\lambda)[\tau_i^{P_i}(\mathbf{s}_2) - t_i]^2 \\
&\quad -\lambda(1-\lambda)[\tau_i^{P_i}(\mathbf{s}_1) - \tau_i^{P_i}(\mathbf{s}_2)]^2 \\
&\le \lambda[\tau_i^{P_i}(\mathbf{s}_1) - t_i]^2 + (1-\lambda)[\tau_i^{P_i}(\mathbf{s}_2) - t_i]^2.
\end{aligned}$$

Then, if $\mathbf{s}_\lambda = \lambda\mathbf{s}_1 + (1-\lambda)\mathbf{s}_2$,

$$\Psi^{\mathcal{P}}(\mathbf{s}_\lambda) \le \lambda\Psi^{\mathcal{P}}(\mathbf{s}_1) + (1-\lambda)\Psi^{\mathcal{P}}(\mathbf{s}_2), \tag{2.23}$$

so $\Psi^{\mathcal{P}}$ is a convex function. ∎

Thus, linear inversion is a convex programming problem. These results show further that, if we could find a convex functional of slowness $\mathbf{s}$ pertinent to the nonlinear inversion problem, then the nonlinear programming problem would be easy (*i.e.*, proofs of convergence become trivial), because there would be no local minima. However, this analysis does not guarantee the existence of such a functional, nor do we know how to construct such a functional even if we suppose one exists. It remains an open question whether an appropriate convex functional for nonlinear seismic inversion can be found.

I expect this question to remain open for a long time, but nevertheless challenge the reader to prove me wrong in this prediction.

Chapter 3

Least-Squares Methods

We consider solutions to the inversion problem for block models. Given a set of weights $w_i > 0$, $i = 1, \ldots, m$, we define the functional $\Psi^*: \mathcal{S} \to \mathbf{R}$ by

$$\Psi^*(\mathbf{s}) = \sum_{i=1}^{m} w_i \left[\tau_i^*(\mathbf{s}) - t_i\right]^2 . \tag{3.1}$$

$\Psi^*(\mathbf{s})$ measures the degree of misfit between the observed data and traveltimes predicted by the model $\mathbf{s}$. Ψ^* is the nonlinear least-squares functional since it uses the full traveltime functional τ^* in the error calculation. The linear least-squares functional was defined previously in (2.20).

3.1 Normal Equations

The standard least-squares problem is a simplified version of (3.1) with all weights equal to unity and the traveltime functional replaced by its linear approximation $\mathbf{Ms}$ for a cell model. Then, the squared error functional is

$$\Psi(\mathbf{s}) = (\mathbf{t} - \mathbf{Ms})^T(\mathbf{t} - \mathbf{Ms}). \tag{3.2}$$

The minimum of this functional is found by differentiating with respect to the value of the slowness in each cell. At the minimum, all these derivatives must vanish, so

$$\frac{\partial \Psi}{\partial s_j} = 2\left[\mathbf{M}^T(\mathbf{t} - \mathbf{Ms})\right]_j = 0, \tag{3.3}$$

for all $j = 1, \ldots, n$. Thus, (3.3) implies the slowness at the minimum of (3.2) satisfies

$$\sum_{i=1}^{m}\sum_{k=1}^{n} l_{ij} l_{ik} \hat{s}_k = \sum_{i=1}^{m} l_{ij} t_i \quad \text{for} \quad j = 1, \ldots, n, \tag{3.4}$$

or equivalently that

$$\mathbf{M}^T\mathbf{M}\hat{\mathbf{s}} = \mathbf{M}^T\mathbf{t}. \tag{3.5}$$

There are n equations for the n unknowns s_j, since $\mathbf{M}^T\mathbf{M}$ is an $n \times n$ square and symmetric matrix.

These equations are known as the *normal equations* for the solution $\hat{\mathbf{s}}$ of the standard least-squares problem. If the number of data m exceeds the number of cells n in the discretized model so $m > n$, we say the discretized inversion problem is *overdetermined.*[1] If the the number of cells n exceeds the number of data m so $m < n$, we say the discretized inversion problem is *underdetermined.* The normal equations may be used in either case, but the form of the resulting solution is substantially different. We generally assume that the inversion problem is overdetermined, but there may still be situations where we want to use only a small part of the available data to make corrections to the slowness model; then the resulting problem is equivalent to the underdetermined version of the normal equations. General methods for solving (3.5) will be discussed in Chapter 4.

PROBLEMS

PROBLEM 3.1.1 *Use the chain rule to show that the minimum of a least-squares functional occurs at the same model whether we use slowness or velocity as the variable.*

PROBLEM 3.1.2 *An experimental configuration has m source-receiver pairs and the region to be reconstructed is modeled using n cells, so the ray-path matrix $\mathbf{M}$ is $m \times n$. Suppose that p independent measurements of the traveltimes have been made, resulting in p traveltime m-vectors $\mathbf{t}_1, \ldots, \mathbf{t}_p$. Then, the inversion problem can be formulated as*

$$\begin{pmatrix} \mathbf{M} \\ \mathbf{M} \\ \vdots \\ \mathbf{M} \end{pmatrix} \mathbf{s} = \begin{pmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \\ \vdots \\ \mathbf{t}_p \end{pmatrix}.$$

Show that the normal equations for this problem become

$$\mathbf{M}^T\mathbf{M}\mathbf{s} = \mathbf{M}^T \langle \mathbf{t} \rangle ,$$

where $\langle \mathbf{t} \rangle = \frac{1}{p}\sum_{q=1}^{p} \mathbf{t}_q$. Explain the significance of this result.

3.2 Scaled Least-Squares Model

DEFINITION 3.2.1 (SCALED LEAST-SQUARES MODEL) *The scaled least-squares model with respect to a given model $\mathbf{s}_0$, and set of weights w_i, is the model $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0]}$ minimizing Ψ^* subject to the constraint that $\mathbf{s} = \gamma\mathbf{s}_0$ for $\gamma > 0$. Thus*

$$\Psi^*(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0]}) = \min_{\gamma} \Psi^*(\gamma\mathbf{s}_0). \tag{3.6}$$

[1] Recall that the underlying physical problem is essentially the reconstruction of a continuous function from finite data, so this continuous reconstruction problem is always grossly *underdetermined.*

The scaled least-squares model associated with $\mathbf{s}_0$ is unique.

To solve for the scaled least-squares model, we expand $\Psi^*(\gamma\mathbf{s}_0)$ as

$$\Psi^*(\gamma\mathbf{s}_0) = \sum_i w_i[\tau_i^*(\gamma\mathbf{s}_0) - t_i]^2 \tag{3.7}$$

$$= \sum_i w_i\tau_i^*(\gamma\mathbf{s}_0)^2 - 2\sum_i w_i t_i\tau_i^*(\gamma\mathbf{s}_0) + \sum_i w_i t_i^2. \tag{3.8}$$

Using the homogeneity of τ_i^*, we can write

$$\Psi^*(\gamma\mathbf{s}_0) = \gamma^2\sum_i w_i\tau_i^*(\mathbf{s}_0)^2 - 2\gamma\sum_i w_i t_i\tau_i^*(\mathbf{s}_0) + \sum_i w_i t_i^2. \tag{3.9}$$

This is simply a second-order polynomial in γ and achieves its minimum at $\gamma = \gamma_{\mathrm{LS}[\mathbf{s}_0]}$, where

$$\gamma_{\mathrm{LS}[\mathbf{s}_0]} = \frac{\sum_i w_i t_i\tau_i^*(\mathbf{s}_0)}{\sum_i w_i\tau_i^*(\mathbf{s}_0)^2}. \tag{3.10}$$

Thus

$$\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0]} = \mathbf{s}_0\frac{\sum_i w_i t_i\tau_i^*(\mathbf{s}_0)}{\sum_i w_i\tau_i^*(\mathbf{s}_0)^2}. \tag{3.11}$$

THEOREM 3.2.1 *For any* $\mathbf{s}_0$, $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0]} \notin \mathrm{Int}\,\mathcal{F}^*$.

Proof: We have from (3.10)

$$\gamma_{\mathrm{LS}[\mathbf{s}_0]}\sum_{i=1}^m w_i\tau_i^*(\mathbf{s}_0)^2 = \sum_{i=1}^m w_i t_i\tau_i^*(\mathbf{s}_0), \tag{3.12}$$

or, given the homogeneity of τ_i^*,

$$\sum_{i=1}^m w_i\tau_i^*(\mathbf{s}_0)[\tau_i^*(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0]}) - t_i] = 0. \tag{3.13}$$

Since the w_i and values of τ_i^* are positive, this can only be true if either $\tau_i^*(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0]}) = t_i$ for all i (i.e., $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0]} \in \partial\mathcal{F}^*$ and is an exact solution to the inversion problem) or if $\tau_i^*(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0]}) < t_i$ for at least one i (*i.e.*, $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0]} \notin \mathcal{F}^*$). Thus, the scaled least-squares model cannot be in $\mathrm{Int}\,\mathcal{F}^*$. ∎

This important result shows that

> *a scaled least-squares slowness model can never be a strictly interior point of the global feasible set.*

The only way for a scaled least-squares point to be in the feasible set is for it to be on the boundary and then only if it solves the inversion problem.

We can write the scaled least-squares model in matrix notation as follows. Let $\mathbf{W}$ be the diagonal matrix formed from the positive weights w_i:

$$\mathbf{W} = \begin{pmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_m \end{pmatrix}. \tag{3.14}$$

Further, let $\mathbf{M}_0$ be the ray-path matrix computed from $\mathbf{s}_0$. Thus, $\tau_i^*(\mathbf{s}_0) = [\mathbf{M}_0\mathbf{s}_0]_i$. In matrix notation, (3.10) becomes

$$\gamma_{\mathrm{LS}[\mathbf{s}_0]} = \frac{\mathbf{s}_0^T\mathbf{M}_0^T\mathbf{W}\mathbf{t}}{\mathbf{s}_0^T\mathbf{M}_0^T\mathbf{W}\mathbf{M}_0\mathbf{s}_0}, \tag{3.15}$$

implying

$$\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0]} = \mathbf{s}_0\frac{\mathbf{s}_0^T\mathbf{M}_0^T\mathbf{W}\mathbf{t}}{\mathbf{s}_0^T\mathbf{M}_0^T\mathbf{W}\mathbf{M}_0\mathbf{s}_0}. \tag{3.16}$$

PROBLEM

PROBLEM 3.2.1 *If $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0]}$ is defined in terms of the linear least-squares functional $\Psi^{\mathcal{P}}$ instead of Ψ^*, is there a result corresponding to Theorem 3.2.1 for this model?*

3.3 Nonlinear Least-Squares Models

DEFINITION 3.3.1 (LEAST-SQUARES MODEL) *A least-squares model, with respect to weights w_i, is a vector $\hat{\mathbf{s}}_{\mathrm{LS}}$ which minimizes Ψ^*, i.e.,*

$$\Psi^*(\hat{\mathbf{s}}_{\mathrm{LS}}) = \min_{\mathbf{s}} \Psi^*(\mathbf{s}). \tag{3.17}$$

The least-squares model may be nonunique. Nonuniqueness is expected when $m < n$, *i.e.*, there are fewer traveltime data than model cells, or when $m > n$ and the ray-path matrix has a right null space containing ghosts $\mathbf{g}$. The most common method of picking the "best" least-squares solution [Penrose, 1955b] is to choose the one of minimum Euclidean norm. This "best" solution has some nice properties as we shall see when we discuss ghosts in tomography, but it may not represent the "best" solution to the inversion problem.

THEOREM 3.3.1 $\hat{\mathbf{s}}_{\mathrm{LS}} \notin \mathrm{Int}\,\mathcal{F}^*$.

Proof: This theorem follows from the fact that $\hat{\mathbf{s}}_{\mathrm{LS}} = \hat{\mathbf{s}}_{\mathrm{LS}[\hat{\mathbf{s}}_{\mathrm{LS}}]}$, *i.e.*, a least-squares model is the scaled least-squares model with respect to itself (or otherwise there would be a model yielding smaller Ψ^*). ▮

Any nonlinear least-squares solution is infeasible unless it solves the inversion problem, in which case it lies on the boundary of the feasible set.

The preceding proof is entirely adequate to establish the infeasibility of the least-squares point. However, it may be enlightening to present a second proof based on stationarity of the ray paths.

Consider the deviation of the least-squares functional induced by a small change in the model:

$$\delta\Psi^* = \Psi^*(s+\delta s) - \Psi^*(s) = \sum_{i=1}^{m} w_i \left[\tau_i^*(s+\delta s) - t_i\right]^2 - \sum_{i=1}^{m} w_i \left[\tau_i^*(s) - t_i\right]^2 . \quad (3.18)$$

This equation may be rearranged without approximation into the form

$$\delta\Psi^* = 2\sum_{i=1}^{m} w_i \left[\tau_i^*(s+\delta s) - \tau_i^*(s)\right]\left[(\tau_i^*(s+\delta s) + \tau_i^*(s))/2 - t_i\right] . \quad (3.19)$$

For small slowness perturbations δs, the first bracket in the sum of (3.19) is clearly of order δs, while any contributions of order δs in the second bracket are therefore of second order and may be neglected. If $dl_i^*[s]$ is the infinitesimal increment of the (or a) least-time ray along path i for s, then

$$\tau_i^*(s+\delta s) - \tau_i^*(s) = \int (s+\delta s)\, dl_i^*[s+\delta s] - \int s\, dl_i^*[s]. \quad (3.20)$$

Recall that stationarity of the ray paths near the one of least time implies that

$$\int s\, dl_i^*[s+\delta s] = \int s\left\{dl_i^*[s] + d\delta l_i^*\right\} \simeq \int s\, dl_i^*[s], \quad (3.21)$$

where $d\delta l_i^*$ is the perturbation in the infinitesimal increment $dl_i^*[s]$ of the ray path induced by the fact that $dl_i^*[s+\delta s]$ is the one for the perturbed model and therefore generally[2] only slightly different from that for s. Using (3.21) in (3.20), we find that

$$\tau_i^*(s+\delta s) - \tau_i^*(s) \simeq \int \delta s\, dl_i^*[s+\delta s] \simeq \int \delta s\, dl_i^*[s] \quad (3.22)$$

to lowest order in δs. Thus, (3.19) becomes

$$\delta\Psi^* = 2\sum_{i=1}^{m} w_i \left(\int \delta s\, dl_i^*[s]\right)\left[\tau_i^*(s) - t_i\right]. \quad (3.23)$$

Equation (3.23) is the expression needed to construct the functional (Frechét) derivative of Ψ^*. If s produces the minimum of $\Psi^*(s)$, then the functional derivative should vanish. We see that the weights w_i are positive, the coefficient of δs is the integral of the increment of the ray path itself in the regions of change which is strictly positive, and if the the traveltime function $\tau_i^*(s) - t_i \geq 0$ as is required for all i in order for the model s to be feasible, then the derivative cannot vanish and therefore s is not the minimum. This contradiction shows again that either the minimum of the traveltime function must be infeasible, or it must solve the inversion problem.

Problem

PROBLEM 3.3.1 *Determine whether there is a result analogous to Theorem 3.3.1 for the linear least-squares functional* $\Psi^{\mathcal{P}}$.

[2] There are pathological cases where a small change in the model s can induce a large change in the ray path, but we will ignore this possibility for the present argument.

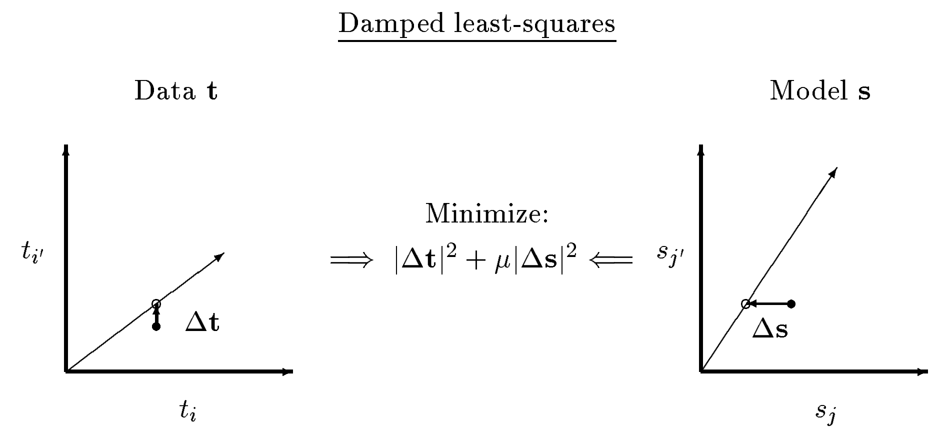


Figure 3.1: Schematic illustration of deamped least squares analysis.

3.4 Damped Least-Squares Model

Let $\mathbf{C}$ be a diagonal (coverage) matrix formed from the positive weights c_j:

$$\mathbf{C} = \begin{pmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_n \end{pmatrix}. \tag{3.24}$$

The c_js may be treated here as arbitrary positive weights, but a definite choice of the c_js will be found later.

Definition 3.4.1 (damped least-squares model) *The damped least-squares model with respect to a given model* $\mathbf{s}_0$, *and set of weights* w_i, c_j, *is the model* $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]}$ *minimizing*

$$\Psi^*(\mathbf{s}) + \mu(\mathbf{s} - \mathbf{s}_0)^T \mathbf{C}(\mathbf{s} - \mathbf{s}_0). \tag{3.25}$$

[Levenberg, 1944]

Like the scaled least-squares model, the damped least-squares model is unique.

We can solve for the damped least-squares model based on a linear approximation to the traveltime functionals. Given the model $\mathbf{s}_0$, let P_i^0 denote the least-time ray paths through $\mathbf{s}_0$. Then, to first order in $\mathbf{s} - \mathbf{s}_0$ we have

$$\tau_i^*(\mathbf{s}) \simeq \tau_i^{P_i^0}(\mathbf{s}). \tag{3.26}$$

This approximation yields

$$\Psi^{\mathcal{P}}(\mathbf{s}) = (\mathbf{t} - \mathbf{M}_0\mathbf{s})^T \mathbf{W}(\mathbf{t} - \mathbf{M}_0\mathbf{s}), \tag{3.27}$$

where $\mathbf{M}_0$ is the ray-path matrix obtained from the ray paths P_i^0.

Using the first-order approximation, the damped least-squares model becomes

$$\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]} = \mathbf{s}_0 + (\mathbf{M}_0^T\mathbf{W}\mathbf{M}_0 + \mu\mathbf{C})^{-1}\mathbf{M}_0^T\mathbf{W}(\mathbf{t} - \mathbf{M}_0\mathbf{s}_0). \tag{3.28}$$

This equation can be rearranged to show that

$$\mu\mathbf{C}(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]} - \mathbf{s}_0) = \mathbf{M}_0^T\mathbf{W}(\mathbf{t} - \mathbf{M}_0\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]}). \tag{3.29}$$

Then, we obtain the following two theorems:

THEOREM 3.4.1 *If $\mathbf{s}_0 \notin \mathcal{F}^{\mathcal{P}^0}$, then the model $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]}$ defined in (3.28) does not solve the inversion problem for any $\mu > 0$.*

THEOREM 3.4.2 *If $\mathbf{s}_0 \notin \mathcal{F}^{\mathcal{P}^0}$, then $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]} \notin \mathcal{F}^{\mathcal{P}^0}$.*

Proof: The proofs are by contradiction.

First, suppose that $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]}$ solves the inversion problem so $\mathbf{M}_0\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]} \equiv \mathbf{t}$. Then, (3.29) shows that $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]} = \mathbf{s}_0$ (if $\mu > 0$) so $\mathbf{M}_0\mathbf{s}_0 \equiv \mathbf{t}$. But this result contradicts the assumption that $\mathbf{s}_0 \notin \mathcal{F}^{\mathcal{P}^0}$ so $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]}$ does not solve the inversion problem if $\mathbf{s}_0$ is infeasible.

Second, suppose that $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]}$ is feasible (*i.e.*, $\mathbf{M}_0\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]} \geq \mathbf{t}$ and, using the previous theorem, we may exclude the possibility that it solves the inversion problem so in fact $\mathbf{M}_0\hat{\mathbf{s}}_s ls\mathbf{s}_0, \mu > \mathbf{t}$), then it follows from the positivity of all the matrix elements in (3.29) that $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]} \leq \mathbf{s}_0$. But, if $\mathbf{s}_0$ is infeasible so $(\mathbf{M}_0\mathbf{s}_0)_i < t_i$ for some i, then it also follows that $(\mathbf{M}_0\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]})_i < t_i$ for the same i so $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]}$ is infeasible, which contradicts the original feasibility supposition on $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]}$. ∎

To paraphrase the results, "we cannot get there from here." If we start our computation at any local infeasible point, we cannot get to a solution or to any local feasible point using the damped least-squares method. These results are very strong because they show the infeasibility of $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]}$ holds for any value of the *damping parameter* $\mu > 0$ and also for any choice of the weight matrix $\mathbf{C}$. In fact, damped least-squares always leads to a biased estimate:

> *unless the starting point $\mathbf{s}_0$ already solves the inversion problem, the damped least-squares solution never solves the inversion problem for any $\mu > 0$.*

Practical application of damped least-squares requires a definite choice of the damping parameter μ. Some methods for choosing its magnitude will be explored in the problems and also in the section discussing linear inversion algorithms (Section 4). Types of damping more general than the norm damping considered in (3.24) will also be discussed.

PROBLEMS

PROBLEM 3.4.1 *Assume that $\mathbf{s}_0$ solves the inversion problem except for a scale factor γ, i.e.,*

$$\mathbf{M}\gamma\mathbf{s}_0 = \mathbf{t}.$$

Show that the damped least-squares solution with $\mathbf{s}_0$ as the starting model does not solve the inversion problem for any $\mu > 0$ unless $\gamma = 1$.

PROBLEM 3.4.2 *Assume the starting point* $\mathbf{s}_0$ *is not a stationary point of* $\Psi^*(\mathbf{s})$. *Then, show that the damped least-squares point always gives a function value less than that of the starting point, i.e., that this point satisfies*

$$\Psi^*(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]}) < \Psi^*(\mathbf{s}_0).$$

[Hint: Show that $\Psi^*(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]}) + \mu(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]} - \mathbf{s}_0)^T\mathbf{C}(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_o,\mu]} - \mathbf{s}_0) \le \Psi^*(\mathbf{s}_0)$.*] Thus, the squares of the residuals will be reduced. [Levenberg, 1944]*

PROBLEM 3.4.3 *If* $\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,0]}$ *is the standard least-squares solution, show that*

$$(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]} - \mathbf{s}_0)^T\mathbf{C}(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]} - \mathbf{s}_0) < (\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,0]} - \mathbf{s}_0)^T\mathbf{C}(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,0]} - \mathbf{s}_0).$$

[Hint: Show that $\Psi^*(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,0]}) \le \Psi^*(\hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_0,\mu]})$.*] Thus, the weighted sums of the squares of the model corrections is less than that for the standard least-squares problem. [Levenberg, 1944]*

PROBLEM 3.4.4 *Consider the model correction* $\Delta\mathbf{s}_d = \mathbf{s} - \hat{\mathbf{s}}$ *as a function of the damping parameter* μ, *where*

$$(\mathbf{M}^T\mathbf{M} + \mu\mathbf{I})\Delta\mathbf{s}_d = \mathbf{M}^T(\mathbf{t} - \mathbf{M}\hat{\mathbf{s}}) \equiv \Delta\mathbf{s}_g. \tag{3.30}$$

The angle θ *between the damped least-squares solution* $\Delta\mathbf{s}_d$ *and the negative of the least-squares functional gradient* $\Delta\mathbf{s}_g$ *is determined by*

$$\cos\theta = \frac{\Delta\mathbf{s}_g^T\Delta\mathbf{s}_d}{||\Delta\mathbf{s}_g||\,||\Delta\mathbf{s}_d||}.$$

Show that

1. $\cos\theta \to 1$ *as* $\mu \to \infty$;
2. $\cos\theta \to 0$ *as* $\mu \to 0$ *if* $\mathbf{M}^T\mathbf{M}$ *is singular.*

Use these results to characterize (3.30) as an interpolation formula. [Marquardt, 1963]

PROBLEM 3.4.5 *Suppose that traveltime measurements have been repeated* K *times, resulting in the set of data vectors* $\mathbf{t}^{(k)}$ *for* $k = 1, \ldots, K$. *Form the data matrix*

$$\mathbf{T} = (\,\mathbf{t}^{(1)} \quad \mathbf{t}^{(2)} \quad \cdots \quad \mathbf{t}^{(K)}\,)$$

and the associated solution matrix

$$\mathbf{S} = (\,\mathbf{s}^{(1)} \quad \mathbf{s}^{(2)} \quad \cdots \quad \mathbf{s}^{(K)}\,).$$

Also, introduce the noise matrix $\mathbf{N}$ *defined by*

$$\mathbf{N} = \mathbf{T} - \mathbf{M}\mathbf{S} = \Delta\mathbf{T} - \mathbf{M}\Delta\mathbf{S},$$

where the correction matrices $\Delta\mathbf{T}$ *and* $\Delta\mathbf{S}$ *are defined similarly by column vectors* $\Delta\mathbf{s}^{(k)} \equiv \mathbf{s}^{(k)} - \bar{\mathbf{s}}$ *and* $\Delta\mathbf{t}^{(k)} \equiv \mathbf{t}^{(k)} - \mathbf{M}\bar{\mathbf{s}}$. *The correlation matrices are then defined by*

$$\mathcal{C}_{ss} = \Delta\mathbf{S}\Delta\mathbf{S}^T, \quad \mathcal{C}_{tt} = \Delta\mathbf{T}\Delta\mathbf{T}^T, \quad \mathcal{C}_{nn} = \mathbf{N}\mathbf{N}^T,$$

$$\mathcal{C}_{st} = \Delta\mathbf{S}\Delta\mathbf{T}^T, \quad \mathcal{C}_{nt} = \mathbf{N}\Delta\mathbf{T}^T, \quad \mathcal{C}_{ns} = \mathbf{N}\Delta\mathbf{S}^T,$$

etc. Suppose that $\mathcal{C}_{ns} \simeq 0 \simeq \mathcal{C}_{sn}$ *(i.e., these correlation matrices are essentially negligible compared with the others) so the noise is uncorrelated with the solution. Then, show the following:*

1. $\mathcal{C}_{nn} \simeq \mathcal{C}_{tn}$;
2. $\mathcal{C}_{nn} = \mathcal{C}_{tt} - \mathbf{M}\mathcal{C}_{st} - \mathcal{C}_{ts}\mathbf{M}^T + \mathbf{M}\mathcal{C}_{ss}\mathbf{M}^T$;
3. $\mathcal{C}_{ts} \simeq \mathbf{M}\mathcal{C}_{ss}$;
4. $\mathcal{C}_{tt} = \mathbf{M}\mathcal{C}_{st} + \mathcal{C}_{nt}$.

Show that the best linear unbiased estimate (BLUE) of the solution based on the data set k is

$$\hat{\mathbf{s}}^{(k)} = \mathcal{C}_{st}\mathcal{C}_{tt}^{-1}\mathbf{t}^{(k)} \tag{3.31}$$

and that

$$\mathcal{C}_{st}\mathcal{C}_{tt}^{-1} \simeq \mathcal{C}_{ss}\mathbf{M}^T\left(\mathbf{M}\mathcal{C}_{ss}\mathbf{M}^T + \mathcal{C}_{nn}\right)^{-1}.$$

Contrast this result with the damped least-squares method assuming that the correlation matrix $\mathcal{C}_{nn}$ *is diagonal. The result (3.31) is known as the "stochastic inverse" [Franklin, 1970; Jordan and Franklin, 1971].*

3.5 Physical Basis for Weighted Least-Squares

So far we have treated the weights w_i as if they are arbitrary positive constants. But are they arbitrary? If they are not arbitrary, then what physical or mathematical feature of the inversion problem determines the weights?

Our goal is ultimately to solve (if possible) the nonlinear inversion problem, so we must keep in mind that the arguments often given for determining the weights in weighted least-squares schemes in other contexts may not be relevant to our problem. In particular, these weights are very often chosen on the basis of statistical (uncorrelated) errors in the data. The assumption behind these choices may be very good indeed in some cases, but generally not in the nonlinear inversion problem. Our working hypothesis for this analysis is that the major source of error in nonlinear inversion is not the measurement error, but the error due to the erroneous choices of ray paths currently in use in the algorithm. The statistical errors in the data become a significant issue only after we have constructed a reliable set of ray paths so that the errors due to wrong ray paths are smaller than the errors in the traveltime data. In fact, for high contrast reconstructions, it may be the case that the errors in the traveltime data are only a small fraction of one percent while the errors introduced by erroneous choices of ray paths are on the order of several percent, or even more in pathological cases.

We envision a two step process. First, we solve the inversion problem iteratively to find a good set of ray paths. This step requires the weighting scheme described here. Second, once we have a reliable set of paths, the weighting scheme can be changed to take proper account of the statistical errors in the data.

Now we use physical arguments to construct a proper set of weights [Berryman, 1989]. Suppose that the traveltime data in our reconstruction actually come from a model that is homogeneous, *i.e.*, with constant slowness σ_0. What will be the characteristics of such data? Clearly, the rays will in fact be straight and the average wave slowness along each ray will be the same constant

$$\sigma_0 = \frac{t_1}{L_1} = \frac{t_2}{L_2} = \cdots = \frac{t_m}{L_m}, \tag{3.32}$$

where

$$L_i = \sum_{j=1}^{n} l_{ij}. \tag{3.33}$$

Furthermore, it follows that the constant value of slowness is also given by the formula

$$\sigma_0 = \frac{\sum_{i=1}^m t_i}{\sum_{i=1}^m L_i}. \tag{3.34}$$

This problem is an ideal use for the scaled least-squares approach presented earlier. We know the ray paths are straight, so we know the ray-path matrix $\mathbf{M}$. We also know that the slowness has the form $\mathbf{s} = \gamma\mathbf{v}$, where $\mathbf{v}^T = (1, \ldots, 1)$ is an n-vector of ones. We want to minimize the least-squares error

$$\Psi(\gamma\mathbf{v}) = (\mathbf{t} - \mathbf{M}\gamma\mathbf{v})^T\mathbf{W}(\mathbf{t} - \mathbf{M}\gamma\mathbf{v}) \tag{3.35}$$

with respect to the coefficient γ. The minimum of (3.35) occurs for

$$\mathbf{v}^T\mathbf{M}^T\mathbf{W}(\mathbf{t} - \mathbf{M}\gamma\mathbf{v}) = 0. \tag{3.36}$$

Solving for γ gives

$$\sigma_0 = \gamma = \frac{\mathbf{v}^T\mathbf{M}^T\mathbf{W}\mathbf{t}}{\mathbf{v}^T\mathbf{M}^T\mathbf{W}\mathbf{M}\mathbf{v}}. \tag{3.37}$$

For easier comparison of (3.37) and (3.34), we now introduce some more notation. Define the m-vector of ones $\mathbf{u}^T = (1, \ldots, 1)$. Then,

$$\mathbf{M}\mathbf{v} = \mathbf{L}\mathbf{u} \tag{3.38}$$

and

$$\mathbf{M}^T\mathbf{u} = \mathbf{C}\mathbf{v}, \tag{3.39}$$

where $\mathbf{L}$ is a diagonal $m \times m$ *length matrix* whose diagonal elements are the row sums of $\mathbf{M}$ given by (3.33) and $\mathbf{C}$ is a diagonal $n \times n$ matrix whose diagonal elements are the column sums of $\mathbf{M}$ given by

$$C_{jj} = \sum_{i=1}^{m} l_{ij}. \tag{3.40}$$

In our later analysis, we will see that the matrix $\mathbf{C}$ (which we call the *coverage matrix*) is a good choice for the second weight matrix in damped least-squares (3.24).

Now we see that (3.37) can be rewritten in this notation as

$$\sigma_0 = \frac{\mathbf{u}^T\mathbf{LWt}}{\mathbf{u}^T\mathbf{LWLu}}, \tag{3.41}$$

while (3.34) becomes

$$\sigma_0 = \frac{\mathbf{u}^T\mathbf{t}}{\mathbf{u}^T\mathbf{Lu}}. \tag{3.42}$$

Comparing (3.41) to (3.42), we see that these two equations would be identical if

$$\mathbf{WLu} = \mathbf{u}. \tag{3.43}$$

Equation (3.43) states that $\mathbf{u}$ is an eigenvector of the matrix $\mathbf{WL}$ with eigenvalue unity. Two choices for the product $\mathbf{WL}$ are

$$\mathbf{WL} = \mathbf{I}, \tag{3.44}$$

where $\mathbf{I}$ is the identity matrix and

$$\mathbf{WL} = \mathbf{L}^{-1}\mathbf{MC}^{-1}\mathbf{M}^T. \tag{3.45}$$

The choice (3.45) is undesirable because it leads to a weight matrix that is not positive definite which would lead to spurious zeroes of the least-squares functional. The choice (3.44) leads to

$$\mathbf{W} = \mathbf{L}^{-1}, \tag{3.46}$$

which is both positive definite and diagonal.

The full significance of the result (3.46) becomes more apparent when we consider that the traveltime data $\mathbf{t} = \bar{\mathbf{t}} + \Delta\mathbf{t}$ will generally include some experimental error $\Delta\mathbf{t}$. If we assume the data are unbiased and the number of source/receiver pairs is sufficiently large, then to a good approximation we should have $\mathbf{u}^T\Delta\mathbf{t} = 0$. The result (3.41) can be rewritten as

$$\gamma = \frac{\mathbf{a}^T\mathbf{L}^{-1}\mathbf{t}}{\mathbf{a}^T\mathbf{u}}, \tag{3.47}$$

where $\mathbf{a} = \mathbf{LWLu}$ may be treated for these purposes as an arbitrary weighting vector. For γ to be unbiased, we must have

$$\mathbf{a}^T\mathbf{L}^{-1}\Delta\mathbf{t} = \mathbf{u}^T\Delta\mathbf{t} = 0. \tag{3.48}$$

Since the $\Delta\mathbf{t}$s are otherwise arbitrary, we must have

$$\mathbf{u} = \mathbf{L}^{-1}\mathbf{a} = \mathbf{WLu}, \tag{3.49}$$

which is the same condition as that found in (3.43). Thus, the choice (3.46) produces the simplest weight matrix giving a linear unbiased estimator of the scale factor for a constant slowness model. In Section 7.2, we derive weights producing unbiased estimates for arbitrary slowness.

Weighting inversely with respect to the lengths of the ray paths can be justified on physical grounds using several different arguments [Frank and Balanis, 1989]. Signal-to-noise ratio is expected to be better on shorter paths than longer ones, since the overall attenuation will typically be smaller and the likelihood of missing the true first arrival therefore smaller. Shorter trial paths are more likely to correspond to real paths that remain completely in the image plane for two-dimensional reconstruction problems.

A disadvantage of using this weighting scheme is that sometimes the ray path is long because the source and receiver are far apart (*e.g.*, from the top of one borehole to the bottom of the other). Yet the information contained in the ray is important because such diagonal rays may help to determine the horizontal extent of some feature of interest, especially when the experimental view angles are severely limited as in crosswell tomography. Weighting inversely with respect to the ray-path length tends to reduce the possibly significant improvement in horizontal resolution that can come from inclusion of these rays. This disadvantage can be circumvented to some extent by using more of these diagonal rays, *i.e.*, using more closely spaced sources and receivers for the diagonal rays. Then, the weights of the individual rays are smaller, but their overall influence on the reconstruction can still be significant.

In Section 4.3, we show that an argument based on stability and regularization leads to the same choice of weight matrices.

3.6 Partial Corrections Using Backprojection

Suppose we have found a solution $\hat{\mathbf{s}}$ of the overdetermined ($m > n$) normal equations

$$\mathbf{M}^T\mathbf{M}\hat{\mathbf{s}} = \mathbf{M}^T\mathbf{t}, \tag{3.50}$$

but this solution does not satisfy the data exactly so

$$\mathbf{M}\hat{\mathbf{s}} \neq \mathbf{t}. \tag{3.51}$$

Then, we argue that a correction $\Delta\mathbf{s}$ could be added to $\hat{\mathbf{s}}$ and the correction should satisfy

$$\mathbf{M}\Delta\mathbf{s} = \Delta\mathbf{t} \equiv \mathbf{t} - \mathbf{M}\hat{\mathbf{s}}. \tag{3.52}$$

Now suppose further that for some subset of the ray paths either $\Delta t_i = 0$, or we are satisfied for some other reason with the agreement between the predicted and measured data (*e.g.*, $|\Delta t_i| \leq \epsilon$ for some small threshold ϵ, or ray path i corresponds to a feasible ray path with $\Delta t_i \leq 0$). Then, we may want to make corrections using only the ray paths that are considered unsatisfactory. We renumber the ray paths so the unsatisfactory ones are the first m' of the m total paths and suppose $m' < n$. Next we rewrite (3.52) as

$$\mathbf{M}'\Delta\mathbf{s} = \Delta\mathbf{t}', \tag{3.53}$$

where $\mathbf{M}'$ is an $m' \times n$ matrix and $\Delta\mathbf{t}'$ is the corresponding m'-vector of unsatisfactory traveltime errors. The problem of solving for the Δs_js is underdetermined as stated.

We can solve (3.53) using a type of backprojection. We argue that the correction vector component $\Delta\mathbf{s}_j$ should be a sum whose terms are proportional to l_{ij} (so that rays not passing through cell j make no contribution) and it should be a linear combination of the traveltime errors Δt_i. However, these corrections should also be made in a way that minimizes the overall effect on the agreement already attained in (3.50). One way to do this approximately is to weight inversely with respect to the cell coverage C_{jj}; then, the cells with the most coverage will change the least and therefore the result should have the smallest effect on (3.50). This argument results in a general form for the correction

$$\Delta s_j = C_{jj}^{-1} \sum_{k'k}^{m'} l_{k'j} w_{k'k} \Delta t_k, \tag{3.54}$$

where $w_{k'k}$ is some weight matrix to be determined by substituting (3.54) into (3.53). On making the substitution, we find that

$$\sum_{j=1}^{n} \sum_{kk'}^{m'} \left(l_{ij} C_{jj}^{-1} l_{k'j} \right) w_{k'k} \Delta t_k = \Delta t_i, \tag{3.55}$$

which implies that

$$\sum_{i=1}^{m'} \left(\sum_{j=1}^{n} l_{ij} C_{jj}^{-1} l_{k'j} \right) w_{k'k} = \delta_{ik}, \tag{3.56}$$

showing that $w_{k'k}$ is the inverse of the matrix

$$\left[\mathbf{M}'\mathbf{C}^{-1}(\mathbf{M}')^T \right]_{kk'} = \sum_{j=1}^{n} l_{kj} C_{jj}^{-1} l_{k'j}. \tag{3.57}$$

Thus, we have

$$\Delta\mathbf{s} = \mathbf{C}^{-1}(\mathbf{M}')^T \left[\mathbf{M}'\mathbf{C}^{-1}(\mathbf{M}')^T \right]^{-1} \Delta\mathbf{t}'. \tag{3.58}$$

It is straightforward to verify that (3.58) is a formal solution by substituting it into (3.53).

To use (3.58) in general requires that the inverse of the $m' \times m'$ matrix $\mathbf{M}'\mathbf{C}^{-1}(\mathbf{M}')^T$ must always exist. This may not always be true, but there is one particularly simple case where the formula can be evaluated: $m' = 1$. Then,

$$\Delta s_j = \frac{l_{1j} \Delta t_1 / C_{jj}}{\sum_{k=1}^{n} l_{1k}^2 / C_{kk}}. \tag{3.59}$$

Problems

PROBLEM 3.6.1 *Let* $\mathbf{D}$ *be an arbitrary* $n \times n$ *positive diagonal matrix. Show that another set of model corrections can be chosen to be*

$$\Delta\mathbf{s} = \mathbf{D}^{-1}(\mathbf{M}')^T \left[\mathbf{M}'\mathbf{D}^{-1}(\mathbf{M}')^T\right]^{-1} \Delta\mathbf{t}'. \tag{3.60}$$

Other than $\mathbf{C}$*, what are some physically relevant choices of the weight matrix* $\mathbf{D}$*?*

PROBLEM 3.6.2 *Consider the stochastic inverse (3.31) when the noise correlation is negligible so that* $C_{nn} \simeq 0$*. Compare the resulting formula with (3.58) and (3.60).*

PROBLEM 3.6.3 *Using (3.58), find an explicit expression for* Δs_j *when slownesses along only two ray paths need correction, i.e.,* $m' = 2$*. Show that the required matrix inverse exists for this problem if*

$$\left(\sum_{j=1}^{n} l_{1j}l_{2j}/C_{jj}\right)^2 < \left(\sum_{j=1}^{n} l_{1j}^2/C_{jj}\right)\left(\sum_{k=1}^{n} l_{2k}^2/C_{kk}\right). \tag{3.61}$$

Use Cauchy's inequality for sums to show that (3.61) is always satisfied unless $l_{1j} = \gamma l_{2j}$ *for all* j*, where* $\gamma > 0$ *is some scalar. Explain the physical significance of the special case when* $l_{1j} = \gamma l_{2j}$ *and suggest a method of solving the problem in this case.*

PROBLEM 3.6.4 *Using (3.58), find an explicit expression for* Δs_j *when slownesses along three ray paths need correction* ($m' = 3$)*. Determine conditions on the matrix elements necessary to guarantee that the required matrix inverse exists.*

Chapter 4

Algorithms for Linear Inversion

In Chapter 3, much of our effort was expended showing that least-squares methods generally produce *infeasible models* in traveltime inversion, *i.e.*, models that violate at least one and often many of the physical constraints imposed on the slowness model by the data through Fermat's principle. Having ruined the reputation of least-squares methods in this way, we try to recover and arrive at a new understanding of the true significance of least-squares methods for inversion problems in this section. Two main points should be stressed: (1) The least-squares methods and generalized inverses are intimately related and, in principle, lead to the same results. (2) Iterative methods for inversion based on least-squares criteria fall into the class of "exterior" methods for nonlinear programming, *i.e.*, at each step of the iteration sequence the "best estimate" of the solution is infeasible so this method approaches the solution (lying on the boundary) from outside the set of feasible models.

In linear inversion with block models, we must solve the linear system of equations given by

$$\mathbf{Ms} = \mathbf{t}, \tag{4.1}$$

where we recall that $\mathbf{M}$ is a known $m \times n$ ray-path matrix, $\mathbf{s}$ is an unknown n-vector of slowness values, and $\mathbf{t}$ is a known m-vector of traveltimes.

Three major difficulties arise in solving (4.1):

1. $\mathbf{M}$ is *not* a square matrix;

2. $\mathbf{M}$ is often rank deficient;

3. $\mathbf{M}$ is often poorly conditioned.

Because of these three difficulties, we cannot simply solve (4.1) in terms of an inverse matrix of $\mathbf{M}$, because such an inverse does not exist. The inverse of an $n \times n$ square matrix $\mathbf{A}$ is defined as the unique matrix $\mathbf{X}$ such that

$$\mathbf{XA} = \mathbf{I} = \mathbf{AX}, \tag{4.2}$$

where $\mathbf{I}$ is the $n \times n$ identity matrix. The standard notation for the matrix inverse is $\mathbf{X} = \mathbf{A}^{-1}$. It is clear from the definition of the inverse (4.2) that $\mathbf{X}$ must also be an

$n \times n$ square matrix. Thus, the fact that $\mathbf{M}$ is not square is sufficient to guarantee that the standard definition of an inverse cannot be applied to our problem. It might still be possible to generalize the concept of inverse so that the equation

$$\mathbf{XM} = \mathbf{I} \tag{4.3}$$

uniquely defines a meaningful $n \times m$ inverse matrix $\mathbf{X}$ associated with $\mathbf{M}$. When $n < m$ and the discretized mathematical problem is *overdetermined*, setting

$$\mathbf{X} = \left(\mathbf{M}^T\mathbf{M}\right)^{-1}\mathbf{M}^T \tag{4.4}$$

gives a formal solution to (4.3) if the inverse $(\mathbf{M}^T\mathbf{M})^{-1}$ exists. We will see that this approach may succeed, but its success is tempered by the second and third difficulties: $\mathbf{M}$ is usually rank deficient, or poorly conditioned, or both. The rank of a matrix is the dimension of the subspace spanned by its columns (or rows) and cannot exceed the smaller of the two dimensions of the matrix. Letting r be the rank of our m by n matrix $\mathbf{M}$, if $r = \min(m, n)$ we say $\mathbf{M}$ has *full rank.* If $r < m, n$ then $\mathbf{M}$ is *rank deficient.* If $\mathbf{M}$ is rank deficient, then $(\mathbf{M}^T\mathbf{M})^{-1}$ does not exist and more sophisticated solutions than (4.4) are required. A similar difficulty arises if $m < n$, so the discretized problem is *underdetermined.* A matrix is *poorly conditioned* if the ratio of largest to smallest nonzero eigenvalue $\lambda_1/\lambda_r >> 1$. For example, this ratio is commonly found equal to 100 or 1000, or even more. Computing an accurate pseudoinverse for a poorly conditioned matrix is difficult. If $\mathbf{M}$ is very poorly conditioned, it may be difficult to compute the smallest eigenvalues accurately enough to obtain satisfactory results from the SVD approach for computing $\mathbf{M}^\dagger$. Then, other numerical techniques for iteratively computing the solution of (4.1) may be preferred.

Two techniques for handling the first difficulty ($\mathbf{M}$ not square) are *completing the square* and *Moore-Penrose pseudoinverses* [Moore, 1920; Penrose, 1956a]. Two techniques for handling the second difficulty (rank deficiency) are *regularization* and *pseudoinverses.* Thus, the Moore-Penrose pseudoinverse is a common solution for both of these problems. Regularization is usually accomplished either (i) by altering the rank deficient square matrix $\mathbf{M}^T\mathbf{M}$ in a way that produces an invertible matrix, or (ii) by performing a singular value decomposition on $\mathbf{M}$ and using formulas for the SVD of $\mathbf{M}^\dagger$ to construct the pseudoinverse.

There are a number of numerical algorithms for solving the system (4.1) and some of these are especially useful when $\mathbf{M}$ is poorly conditioned. These methods include:

1. standard tomographic reconstruction methods (e.g., ART and SIRT),
2. iterative matrix methods (e.g., Gauss-Seidel and Jacobi's method),
3. conjugate direction/gradient methods,
4. simple iteration,
5. a "neural network" method.

These methods may be analyzed most conveniently in terms of their convergence to the pseudoinverse.

Since the pseudoinverse plays such a central role in all these problem/solution pairs, we begin our discussion by deriving and analyzing $\mathbf{M}^\dagger$. Then, we discuss regularization techniques and finally analyze various numerical techniques for solving (4.1).

4.1 Moore-Penrose Pseudoinverse and SVD

Any nonsymmetric (and/or nonsquare) matrix $\mathbf{M}$ of real numbers can be decomposed in terms of a set of positive eigenvalues and two sets of orthonormal eigenvectors. Let r be the rank of $\mathbf{M}$. There exist r solutions to the eigenvalue problem

$$\mathbf{M}\mathbf{z} = \lambda\mathbf{y}, \tag{4.5}$$

$$\mathbf{M}^T\mathbf{y} = \lambda\mathbf{z}, \tag{4.6}$$

such that $\lambda > 0$ and $\mathbf{y}^T\mathbf{y} = \mathbf{z}^T\mathbf{z} = 1$. Letting λ_i, $\mathbf{y}_i$, $\mathbf{z}_i$, $i = 1, \ldots, r$, denote the solutions, then

$$\mathbf{y}_i^T\mathbf{M}\mathbf{M}^T\mathbf{y}_j = (\lambda_i^2\mathbf{y}_i^T)\mathbf{y}_j = \mathbf{y}_i^T(\lambda_j^2\mathbf{y}_j), \tag{4.7}$$

and

$$\mathbf{z}_i^T\mathbf{M}^T\mathbf{M}\mathbf{z}_j = (\lambda_i^2\mathbf{z}_i^T)\mathbf{z}_j = \mathbf{z}_i^T(\lambda_j^T\mathbf{z}_j), \tag{4.8}$$

so that

$$(\lambda_j^2 - \lambda_i^2)\mathbf{y}_i^T\mathbf{y}_j = 0 = (\lambda_j^2 - \lambda_i^2)\mathbf{z}_i^T\mathbf{z}_j, \tag{4.9}$$

for all combinations of i, j. Furthermore,

$$\lambda_i\mathbf{y}_i^T\mathbf{y}_i = \mathbf{y}_i^T\mathbf{M}\mathbf{z}_i = \mathbf{z}_i^T\mathbf{M}^T\mathbf{y}_i = \lambda_i\mathbf{z}_i^T\mathbf{z}_i, \tag{4.10}$$

showing that

$$\mathbf{y}_i^T\mathbf{y}_i = \mathbf{z}_i^T\mathbf{z}_i, \tag{4.11}$$

for all $1 \le i \le r$. Then, after normalizing the eigenvectors, it follows from (4.9) and (4.11) that

$$\mathbf{y}_i^T\mathbf{y}_j = \mathbf{z}_i^T\mathbf{z}_j = \delta_{ij}. \tag{4.12}$$

The vectors $\mathbf{y}_i$ and $\mathbf{z}_i$, respectively, are left- and right-hand eigenvectors of $\mathbf{M}$ corresponding to the eigenvalue λ_i. Multiple eigenvectors associated with the same eigenvalue are not necessarily orthogonal to each other, but they do form a subspace that is orthogonal to all other eigenvectors with different eigenvalues.

4.1.1 Resolution and completeness

The model space is n-dimensional while the data space is m-dimensional. Any slowness vector in the model space can be expanded in terms of a complete orthonormal set of vectors $\{\mathbf{z}_j\}$ as

$$\mathbf{s} = \sum_{j=1}^{n}\sigma_j\mathbf{z}_j = \sum_{j=1}^{r}\sigma_j\mathbf{z}_j + \mathbf{s}_0, \tag{4.13}$$

and similarly for a data vector

$$\mathbf{t} = \sum_{i=1}^{m} \tau_i \mathbf{y}_i = \sum_{i=1}^{r} \tau_i \mathbf{y}_i + \mathbf{t}_0, \tag{4.14}$$

when expanded in the basis set $\{\mathbf{y}_i\}$. The vectors $\mathbf{s}_0$ and $\mathbf{t}_0$ are respectively arbitrary vectors from the right and left null spaces of $\mathbf{M}$. *Completeness* implies that the identity matrix can be represented in terms of these sets of vectors by taking sums of outer products according to

$$\mathbf{I}_n = \sum_{j=1}^{n} \mathbf{z}_j \mathbf{z}_j^T \tag{4.15}$$

and

$$\mathbf{I}_m = \sum_{i=1}^{m} \mathbf{y}_i \mathbf{y}_i^T. \tag{4.16}$$

Then, for example,

$$\mathbf{I}_n \mathbf{s} = \sum_{jk}^{n} \mathbf{z}_j (\mathbf{z}_j^T \sigma_k \mathbf{z}_k) = \sum_{j=1}^{n} \sigma_j \mathbf{z}_j = \mathbf{s}. \tag{4.17}$$

In each space, only the first r of these vectors can be eigenvectors of $\mathbf{M}$ with $\lambda > 0$. The remaining $n - r$ and $m - r$ vectors necessarily lie in the right and left null spaces of $\mathbf{M}$.

The completeness relation can be written using any complete set of vectors; using the eigenvectors of $\mathbf{M}$ is a choice made as a convenience for the SVD analysis of $\mathbf{M}$.

Now we define *resolution matrices* for the two vector spaces based on *partial sums* of the completeness relations

$$\mathcal{R}_n = \sum_{j=1}^{r} \mathbf{z}_j \mathbf{z}_j^T \tag{4.18}$$

and

$$\mathcal{R}_m = \sum_{i=1}^{r} \mathbf{y}_i \mathbf{y}_i^T. \tag{4.19}$$

Note that for a real, square, and symmetric matrix, there is only one resolution matrix.

Applying the resolution matrices to $\mathbf{s}$ and $\mathbf{t}$, we find

$$\mathcal{R}_n \mathbf{s} = \sum_{j=1}^{r} \sigma_j \mathbf{z}_j \tag{4.20}$$

and

$$\mathcal{R}_m \mathbf{t} = \sum_{i=1}^{r} \tau_i \mathbf{y}_i. \tag{4.21}$$

Thus, the resolution matrices strip off the parts of $\mathbf{s}$ and $\mathbf{t}$ lying in the right and left null spaces of $\mathbf{M}$, respectively.

The full significance of the resolution matrices will become apparent as we develop the generalized inverse, and particularly when we examine the relationship between least-squares methods and the pseudoinverse.

PROBLEMS

PROBLEM 4.1.1 *Show that* $\mathcal{R}_n^2 = \mathcal{R}_n$ *and* $\mathcal{R}_m^2 = \mathcal{R}_m$.

PROBLEM 4.1.2 *Show that* $Tr(\mathcal{R}_n) = Tr(\mathcal{R}_m)$.

4.1.2 Completing the square

These results are most easily derived and understood by using a technique of Lanczos [1961] for completing the square. We define a real, square, and symmetric $(m+n)\times(m+n)$ matrix

$$\mathbf{H} = \begin{pmatrix} 0 & \mathbf{M} \\ \mathbf{M}^T & 0 \end{pmatrix}. \tag{4.22}$$

Then, (4.5)–(4.6) becomes

$$\mathbf{H}\begin{pmatrix} \mathbf{y}_i \\ \mathbf{z}_i \end{pmatrix} = \lambda_i \begin{pmatrix} \mathbf{y}_i \\ \mathbf{z}_i \end{pmatrix}. \tag{4.23}$$

Clearly, for each positive eigenvalue λ_i with eigenvector $(\mathbf{y}_i^T, \mathbf{z}_i^T)^T$, there is a corresponding negative eigenvalue $-\lambda_i$ with eigenvector $(\mathbf{y}_i^T, -\mathbf{z}_i^T)^T$.

PROBLEM

PROBLEM 4.1.3 *Show that* $\mathbf{H}$ *is rank deficient if* $m+n$ *is odd.*

4.1.3 Finding the generalized inverse

The *singular value decomposition* (SVD) of $\mathbf{M}$ is given by

$$\mathbf{M} = \sum_{i=1}^{r} \lambda_i \mathbf{y}_i \mathbf{z}_i^T. \tag{4.24}$$

The Moore-Penrose pseudoinverse of $\mathbf{M}$ can be expressed as

$$\mathbf{M}^\dagger = \sum_{i=1}^{r} \lambda_i^{-1} \mathbf{z}_i \mathbf{y}_i^T. \tag{4.25}$$

Our goal in this section is to derive (4.25). Also, notice that this definition implies

$$\mathbf{M}^\dagger\mathbf{M} = \sum_{i=1}^{r} \mathbf{z}_i \mathbf{z}_i^T = \mathcal{R}_n \tag{4.26}$$

and

$$\mathbf{M}\mathbf{M}^{\dagger} = \sum_{i=1}^{r} \mathbf{y}_i \mathbf{y}_i^T = \mathcal{R}_m, \tag{4.27}$$

thus making the connection to the resolution matrices.

Completing the square permits us to find a simple and intuitive derivation of the uniqueness conditions required for a meaningful generalized inverse giving rise to the formula (4.25). First, we find the generalized inverse for the square matrices appearing in $\mathbf{H}^2$. Then, we use these results to derive (4.25).

Let $\mathbf{A} = \mathbf{M}^T\mathbf{M}$ so that

$$\mathbf{A} = \sum_{i=1}^{r} \lambda_i^2 \mathbf{z}_i \mathbf{z}_i^T. \tag{4.28}$$

Then, $\mathbf{A}$ is real symmetric and therefore has real eigenvalues. Since the $\mathbf{z}_i$s are assumed to be an orthonormal and complete set of vectors, any generalized inverse for $\mathbf{A}$ can be written in the form

$$\mathbf{A}^{\dagger} = \sum_{ij}^{r} \alpha_{ij} \mathbf{z}_i \mathbf{z}_j^T, \tag{4.29}$$

where the coefficients $\alpha_{ij} = \mathbf{z}_i^T \mathbf{A}^{\dagger} \mathbf{z}_j$ are to be determined and the upper limit on the sum has been taken as r for the sake of simplifying this derivation. Consistency conditions are

$$\mathbf{A}\mathbf{A}^{\dagger} = \mathbf{A}^{\dagger}\mathbf{A} = \sum_{i=1}^{r} \mathbf{z}_i \mathbf{z}_i^T = \mathcal{R}_n, \tag{4.30}$$

which are the conditions intuition[1] suggests are the right ones for the generalized inverse of a square matrix. The final expression in (4.30) is just the completeness relation within the subspace orthogonal to the null space of $\mathbf{A}$. Equation (4.30) implies that $\mathbf{A}^{\dagger}$ is the unique matrix satisfying the conditions

$$\mathbf{A}\mathbf{A}^{\dagger}\mathbf{A} = \mathbf{A}, \tag{4.31}$$

$$\mathbf{A}^{\dagger}\mathbf{A}\mathbf{A}^{\dagger} = \mathbf{A}^{\dagger}. \tag{4.32}$$

It follows easily from (4.28)–(4.30) that

$$\alpha_{ij} = \delta_{ij}/\lambda_i^2. \tag{4.33}$$

Thus, the generalized inverse of this symmetric square matrix is just

$$\mathbf{A}^{\dagger} = \sum_{i=1}^{r} \mathbf{z}_i \mathbf{z}_i^T/\lambda_i^2. \tag{4.34}$$

[1] Compare (4.2).

To find the needed relation for the nonsymmetric/nonsquare matrix $\mathbf{M}$, again consider the square matrix $\mathbf{H}$. We find easily that

$$\mathbf{H}^2 = \begin{pmatrix} \mathbf{M}\mathbf{M}^T & 0 \\ 0 & \mathbf{M}^T\mathbf{M} \end{pmatrix}. \tag{4.35}$$

Then, for consistency we suppose

$$\mathbf{H}^\dagger = \mathbf{H}(\mathbf{H}^2)^\dagger = (\mathbf{H}^2)^\dagger\mathbf{H}, \tag{4.36}$$

from which it follows that

$$\mathbf{H}^\dagger = \begin{pmatrix} 0 & \mathbf{M}(\mathbf{M}^T\mathbf{M})^\dagger \\ \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^\dagger & 0 \end{pmatrix} = \begin{pmatrix} 0 & (\mathbf{M}\mathbf{M}^T)^\dagger\mathbf{M} \\ (\mathbf{M}^T\mathbf{M})^\dagger\mathbf{M}^T & 0 \end{pmatrix}. \tag{4.37}$$

Equation (4.37) implies that

$$\mathbf{M}^\dagger = \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^\dagger = (\mathbf{M}^T\mathbf{M})^\dagger\mathbf{M}^T. \tag{4.38}$$

Using (4.34) in (4.38) then finally yields (4.25). Thus, we have completed one derivation of the pseudoinverse.

A more direct derivation comes from (4.34) by writing down the equivalent expansion for $\mathbf{H}^\dagger$. First, expand $\mathbf{H}$ in terms of the eigenvectors as

$$\mathbf{H} = \frac{1}{2}\sum_{i=1}^{r}\lambda_i\left[\begin{pmatrix} \mathbf{y}_i \\ \mathbf{z}_i \end{pmatrix}(\,\mathbf{y}_i^T \quad \mathbf{z}_i^T\,) - \begin{pmatrix} \mathbf{y}_i \\ -\mathbf{z}_i \end{pmatrix}(\,\mathbf{y}_i^T \quad -\mathbf{z}_i^T\,)\right] \tag{4.39}$$

$$= \sum_{i=1}^{r}\lambda_i\left[\begin{pmatrix} 0 \\ \mathbf{z}_i \end{pmatrix}(\,\mathbf{y}_i^T \quad 0\,) + \begin{pmatrix} \mathbf{y}_i \\ 0 \end{pmatrix}(\,0 \quad \mathbf{z}_i^T\,)\right]. \tag{4.40}$$

[The factor of one-half in (4.39) arises from the fact that the norm of the eigenvectors of $\mathbf{H}$ (as defined here) is 2.] Then, from (4.39) and (4.34), we obtain

$$\mathbf{H}^\dagger = \sum_{i=1}^{r}\lambda_i^{-1}\left[\begin{pmatrix} 0 \\ \mathbf{z}_i \end{pmatrix}(\,\mathbf{y}_i^T \quad 0\,) + \begin{pmatrix} \mathbf{y}_i \\ 0 \end{pmatrix}(\,0 \quad \mathbf{z}_i^T\,)\right] = \begin{pmatrix} 0 & (\mathbf{M}^T)^\dagger \\ \mathbf{M}^\dagger & 0 \end{pmatrix}, \tag{4.41}$$

and (4.25) again follows, thus completing another derivation.

We observe two special cases in which $\mathbf{M}$ is of full rank. If $r = n \leq m$ so the problem is either determined or overdetermined but $\mathbf{M}$ is of full rank, then $\mathbf{s}_0$ — the vector from the right null space — is necessarily zero. Further, we can write

$$\mathbf{M}^\dagger = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T = \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^\dagger. \tag{4.42}$$

Second, if $r = m \leq n$ so the problem is either determined or underdetermined but $\mathbf{M}$ is of full rank, then $\mathbf{t}_0$ — the vector from the left null space — vanishes and

$$\mathbf{M}^\dagger = \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1} = (\mathbf{M}^T\mathbf{M})^\dagger\mathbf{M}^T. \tag{4.43}$$

A subcase of both cases is $r = m = n$ so the problem is just determined and $\mathbf{M}$ is of full rank. $\mathbf{M}$ is then also square and invertible. Since in this case $(\mathbf{M}^T\mathbf{M})^{-1} = \mathbf{M}^{-1}(\mathbf{M}^T)^{-1}$ and $(\mathbf{M}\mathbf{M}^T)^{-1} = (\mathbf{M}^T)^{-1}\mathbf{M}^{-1}$, (4.42) and (4.43) reduce to

$$\mathbf{M}^\dagger = \mathbf{M}^{-1}, \tag{4.44}$$

consistent with the intuitive derivation of the generalized inverse given here.

Example 4.1.1 *Consider a 2×2 model with the layout*

s_1	s_2
s_3	s_4

Suppose the ray-path matrix is

$$\mathbf{M} = \begin{pmatrix} 1.00 & 1.00 & 0 & 0 \\ 1.05 & 1.05 & 0 & 0 \end{pmatrix},$$

corresponding to two ray paths going through cells 1 and 2 at slightly different angles. The SVD of $\mathbf{M}$ *is*

$$\mathbf{M} = \begin{pmatrix} 1.00 \\ 1.05 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix},$$

with the single nonzero eigenvalue $\lambda = \sqrt{4.205}$. *Then, the generalized inverse of* $\mathbf{M}$ *is*

$$\mathbf{M}^\dagger = \frac{1}{\sqrt{4.205}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1.00 & 1.05 \end{pmatrix}.$$

The corresponding resolution matrices are

$$\mathcal{R}_4 = \mathbf{M}^\dagger \mathbf{M} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{4.45}$$

and

$$\mathcal{R}_2 = \mathbf{M}\mathbf{M}^\dagger = \frac{1}{2.1025} \begin{pmatrix} 1.00 \\ 1.05 \end{pmatrix} \begin{pmatrix} 1.00 & 1.05 \end{pmatrix} = \frac{1}{2.1025} \begin{pmatrix} 1.00 & 1.05 \\ 1.05 & 1.1025 \end{pmatrix}. \tag{4.46}$$

Equation (4.45) shows that the two ray paths contain equivalent information about the two cells 1 and 2, but no information about cells 3 and 4. Equation (4.46) shows that, even though the two ray paths do in fact have the same information about the model, the longer ray path is treated as more reliable simply because it is longer. (It is this sort of contradictory result that leads us to consider using other weighting schemes in Section 3.5.*) Another way of displaying the information in the resolution matrix* R_4 *is to exhibit the diagonal values of the matrix on the grid of the slowness cells as*

$\frac{1}{2}$	$\frac{1}{2}$
0	0

This display has nothing to do with the actual slowness values *computed in the inversion, but it does have something to do with the relative reliability of the values computed. We should have more confidence in the computed values of slowness in those cells with the higher diagonal resolution.*

Example 4.1.2 *Using the same layout as the preceding example, consider the ray-path matrix*

$$\mathbf{M} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

This matrix corresponds to having one horizontal ray (through cells 1 and 2) and one vertical ray (through cells 1 and 3). The symmetric matrix

$$\mathbf{M}\mathbf{M}^T = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \frac{3}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\begin{pmatrix} 1 & -1 \end{pmatrix},$$

showing that the left-eigenvectors of $\mathbf{M}$ *are* $(1,1)$ *and* $(1,-1)$ *with eigenvalues* $\sqrt{3}$ *and 1, respectively. Thus,* $\mathbf{M}$ *is of full rank. Multiplying* $\mathbf{M}$ *on the left by the left-eigenvectors determines the right-eigenvectors and shows that*

$$\mathbf{M} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\begin{pmatrix} 0 & 1 & -1 & 0 \end{pmatrix},$$

which is easily verified. The generalized inverse is then

$$\mathbf{M}^\dagger = \frac{1}{6}\begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & -1 \end{pmatrix}.$$

The resolution matrices are therefore

$$\mathcal{R}_4 = \mathbf{M}^\dagger\mathbf{M} = \begin{pmatrix} 2/3 & 1/3 & 1/3 & 0 \\ 1/3 & 2/3 & -1/3 & 0 \\ 1/3 & -1/3 & 2/3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.47}$$

and

$$\mathcal{R}_2 = \mathbf{M}\mathbf{M}^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.48}$$

The diagonal values of the model resolution matrix are displayed in

$\frac{2}{3}$	$\frac{2}{3}$
$\frac{2}{3}$	0

.

The data resolution matrix is the identity because the ray-path matrix $\mathbf{M}$ *is of full rank, and* $r = m = 2$.

Example 4.1.3 *Consider a* 3×2 *model with the layout*

s_1	s_2
s_3	s_4
s_5	s_6

.

Consider the ray-path matrix

$$\mathbf{M} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & \sqrt{2} & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix},$$

corresponding to horizontal rays through the three horizontal pairs of cells and one diagonal ray cutting through cells 2 and 3. The symmetric matrix

$$\mathbf{M}\mathbf{M}^T = \begin{pmatrix} 2 & 0 & \sqrt{2} & 0 \\ 0 & 2 & \sqrt{2} & 0 \\ \sqrt{2} & \sqrt{2} & 4 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

has the eigenvalues $\lambda = 2$ *(twice) and* $\lambda = 3 \pm \sqrt{5}$*, so* $\mathbf{M}$ *is again of full rank. The corresponding left-eigenvectors of* $\mathbf{M}$ *are*

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ \frac{1}{\sqrt{2}}(1+\sqrt{5}) \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ \frac{1}{\sqrt{2}}(1-\sqrt{5}) \\ 0 \end{pmatrix}.$$

Except for normalization, the right-eigenvectors of $\mathbf{M}$ *for the nonzero eigenvalues are then found to be*

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2+\sqrt{5} \\ 2+\sqrt{5} \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2-\sqrt{5} \\ 2-\sqrt{5} \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

The resolution matrices are

$$\mathcal{R}_6 = \mathbf{M}^\dagger\mathbf{M} = \begin{pmatrix} 3/4 & 1/4 & -1/4 & 1/4 & 0 & 0 \\ 1/4 & 3/4 & 1/4 & -1/4 & 0 & 0 \\ -1/4 & 1/4 & 3/4 & 1/4 & 0 & 0 \\ 1/4 & -1/4 & 1/4 & 3/4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix} \tag{4.49}$$

and

$$\mathcal{R}_4 = \mathbf{M}\mathbf{M}^\dagger = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.50}$$

The diagonal elements of the model resolution matrix are displayed in

$\frac{3}{4}$	$\frac{3}{4}$
$\frac{3}{4}$	$\frac{3}{4}$
$\frac{1}{2}$	$\frac{1}{2}$

The data resolution matrix is again the identity matrix in this example because the ray-path matrix $\mathbf{M}$ *is of full rank and* $r = m = 4$*.*

Problems

Problem 4.1.4 *Using (4.26) and (4.27) show that* $Tr(\mathcal{R}_n) = Tr(\mathcal{R}_m) = rank(\mathbf{M})$*.*

Problem 4.1.5 *If* $\{\mathbf{z}_j\}_n$ *is an orthonormal, complete set of* n*-vectors and* $\{\mathbf{y}_i\}_m$ *is an orthonormal, complete set of* m*-vectors, define the* $n \times n$ *matrix*

$$\mathbf{Z} = (\,\mathbf{z}_1 \quad \mathbf{z}_2 \quad \cdots \quad \mathbf{z}_n\,)$$

and the $m \times m$ *matrix*

$$\mathbf{Y} = (\,\mathbf{y}_1 \quad \mathbf{y}_2 \quad \cdots \quad \mathbf{y}_m\,)\,.$$

Let $\mathbf{O}_{m,n}$ *be an* $m \times n$ *matrix of zeroes. Then, the matrix* $\mathbf{\Lambda}$ *is the* $m \times n$

$$\mathbf{\Lambda} = \left(\begin{array}{cccc|c} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \ddots & & \mathbf{O}_{r,n-r} \\ & & & \lambda_r & \\ \hline & & \mathbf{O}_{m-r,r} & & \mathbf{O}_{m-r,n-r} \end{array}\right).$$

where the first r *components of* $\mathbf{\Lambda}$ *along the diagonal are the first rp eigenvalues of* $\mathbf{M}$*. Then, if the first* r *vectors in the two sets are the right and left eigenvectors of* $\mathbf{M}$*, show that (4.24) can be rewritten as*

$$\mathbf{M} = \mathbf{Y}\mathbf{\Lambda}\mathbf{Z}^T.$$

Show that

$$\mathbf{M}^\dagger = \mathbf{Z}\mathbf{\Lambda}^\dagger\mathbf{Y}^T,$$

where the $n \times m$ *matrix* $\mathbf{\Lambda}^\dagger$ *is given by*

$$\mathbf{\Lambda}^\dagger = \left(\begin{array}{cccc|c} \lambda_1^{-1} & & & & \\ & \lambda_2^{-1} & & & \\ & & \ddots & & \mathbf{O}_{r,m-r} \\ & & & \lambda_r^{-1} & \\ \hline & & \mathbf{O}_{n-r,r} & & \mathbf{O}_{n-r,m-r} \end{array}\right).$$

Show that

$$\mathcal{R}_n = \mathbf{M}^\dagger \mathbf{M} = \mathbf{Z}\mathbf{Z}^T$$

and

$$\mathcal{R}_m = \mathbf{M}\mathbf{M}^\dagger = \mathbf{Y}\mathbf{Y}^T.$$

PROBLEM 4.1.6 *If* $\mathbf{A}$ *is real and symmetric, prove that its eigenvalues are real.*

PROBLEM 4.1.7 *Derive (4.25) from (4.24), (4.34), and (4.38).*

PROBLEM 4.1.8 *Show that (4.38) implies* $(\mathbf{M}\mathbf{M}^T)^\dagger = (\mathbf{M}^T)^\dagger \mathbf{M}^\dagger$.

PROBLEM 4.1.9 *It was implicitly assumed in (4.36) that*

$$(\mathbf{H}^2)^\dagger = (\mathbf{H}^\dagger)^2. \tag{4.51}$$

Show that (4.51) follows from (4.34). Then, verify that (4.38) is consistent with (4.51).

PROBLEM 4.1.10 *Find the pseudoinverse and the resolution matrix of*

$$\mathbf{A} = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}.$$

PROBLEM 4.1.11 *Find the pseudoinverse and the resolution matrix of*

$$\mathbf{A} = \begin{pmatrix} 5/4 & -1/2 & -3/4 \\ -1/2 & 1 & -1/2 \\ -3/4 & -1/2 & 5/4 \end{pmatrix}.$$

PROBLEM 4.1.12 *Find the pseudoinverse and the resolution matrices of*

$$\mathbf{M} = \begin{pmatrix} 3/5 & 6/5 & 3/5 & 0 \\ 3/5 & 0 & 3/5 & 6/5 \end{pmatrix}.$$

PROBLEM 4.1.13 *Suppose that* $\mathbf{A}$ *is nonsingular,* $\mathbf{w}$ *is some normalized vector, and* ϵ *is a positive scalar. Show that*

$$\left[\mathbf{A} + \epsilon \mathbf{w}\mathbf{w}^T\right]^{-1} = \mathbf{A}^{-1} - \epsilon \frac{\mathbf{A}^{-1}\mathbf{w}\mathbf{w}^T\mathbf{A}^{-1}}{1 + \epsilon \mathbf{w}^T \mathbf{A}^{-1}\mathbf{w}}. \tag{4.52}$$

Then, use (4.52) to show that, if

$$\mathbf{B} = \mathbf{A} - \lambda \mathbf{w}\mathbf{w}^T \quad \text{and} \quad \mathbf{B}\mathbf{w} = 0,$$

then

$$\mathbf{B}^\dagger = \mathbf{A}^{-1} - \frac{1}{\lambda}\mathbf{w}\mathbf{w}^T.$$

PROBLEM 4.1.14 *Suppose* $\mathbf{A}$ *is real and symmetric with SVD given by* $\mathbf{A} = \sum_{i=1}^{n} \rho_i \mathbf{z}_i \mathbf{z}_i^T$ *where* $\rho_i > 0$ *for* $1 \le i \le r$ *and* $\rho_i = 0$ *for* $r+1 \le i \le n$. *Show that*

$$\mathbf{A}^\dagger = \left[\mathbf{A} + \mu \sum_{i=r+1}^{n} \mathbf{z}_i \mathbf{z}_i^T\right]^{-1} - \frac{1}{\mu} \sum_{i=r+1}^{n} \mathbf{z}_i \mathbf{z}_i^T \tag{4.53}$$

for any $\mu > 0$.

PROBLEM 4.1.15 *Suppose that (4.29) is replaced by*

$$\mathbf{A}^\dagger \equiv \sum_{i,j}^{n} \alpha_{ij} \mathbf{z}_i \mathbf{z}_j^T,$$

where the upper limit on the sum is taken to be the size of the model vector space. Repeat the derivation of the pseudoinverse and explain the differing results.

PROBLEM 4.1.16 *Show that the four equations*

$$\boxed{\begin{array}{rcl} \mathbf{MXM} & = & \mathbf{M}, \\ \mathbf{XMX} & = & \mathbf{X}, \\ (\mathbf{MX})^T & = & \mathbf{MX}, \\ (\mathbf{XM})^T & = & \mathbf{XM}, \end{array}}$$

have a unique solution $\mathbf{X}$ *for any real matrix* $\mathbf{M}$. *Show that* $\mathbf{X} = \mathbf{M}^\dagger$. *Then show directly that* $(\mathbf{M}^T\mathbf{M})^\dagger = \mathbf{M}^\dagger(\mathbf{M}^\dagger)^T$. *[Penrose, 1955a]*

PROBLEM 4.1.17 *If* $\mathbf{A}$ *is real, square, and symmetric, then show that the equations*

$$\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger\mathbf{A} = \mathcal{R}_n$$

are equivalent to the set of uniqueness conditions in PROBLEM 4.1.16 *for* $\mathbf{A}^\dagger$.

PROBLEM 4.1.18 *Show that, if* $\mathbf{X}$ *is a real, square, and symmetric matrix satisfying* $\mathbf{X}^2 = \mathbf{X}$, *then* $\mathbf{X}^\dagger = \mathbf{X}$. *Does* $\mathbf{X}^\dagger = \mathbf{X}$ *imply* $\mathbf{X}^2 = \mathbf{X}$*?*

PROBLEM 4.1.19 *Use the defining equations for* $\mathbf{M}^\dagger$ *in* PROBLEM 4.1.16 *to show directly that* $\mathbf{M}^\dagger = \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^\dagger = (\mathbf{M}^T\mathbf{M})^\dagger\mathbf{M}$ *thus verifying (4.38). [Hint: Use the defining equations three times, once each for* $(\mathbf{M}^T\mathbf{M})^\dagger$, *for* $(\mathbf{M}\mathbf{M}^T)^\dagger$, *and for* $\mathbf{M}^\dagger$.*]*

PROBLEM 4.1.20 *If* r *is the rank of* $\mathbf{M}$, *show that* $\mathbf{M}$ *can be factored into a product of an* $m \times r$ *matrix* $\mathbf{L}$ *and an* $r \times n$ *matrix* $\mathbf{R}$ *such that*

$$\mathbf{M} = \mathbf{LR},$$

where $\mathbf{L}$ *and* $\mathbf{R}$ *are each of rank* r. *Show that the generalized inverse* $\mathbf{M}^\dagger$ *may be written*

$$\mathbf{M}^\dagger = \mathbf{R}^T(\mathbf{R}\mathbf{R}^T)^{-1}(\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T. \tag{4.54}$$

Use (4.54) to show that $\mathbf{M}^\dagger = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T$ *when* $r = n$ *and that* $\mathbf{M}^\dagger = \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}$ *when* $r = m$. *[Smith and Franklin, 1969]*

PROBLEM 4.1.21 *Any real matrix* $\mathbf{M}$ *can be partitioned into the form*

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \tag{4.55}$$

after some rearrangement of the rows and columns.

1. *If* $\mathbf{A}$ *is a nonsingular submatrix of* $\mathbf{M}$ *whose rank is equal to that of* $\mathbf{M}$*, then verify that*

$$\mathbf{M}^\dagger = \begin{pmatrix} \mathbf{A}^T\mathbf{P}\mathbf{A}^T & \mathbf{A}^T\mathbf{P}\mathbf{C}^T \\ \mathbf{B}^T\mathbf{P}\mathbf{A}^T & \mathbf{B}^T\mathbf{P}\mathbf{C}^T \end{pmatrix}, \tag{4.56}$$

where $\mathbf{P} = (\mathbf{A}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T)^{-1}\mathbf{A}(\mathbf{A}^T\mathbf{A} + \mathbf{C}^T\mathbf{C})^{-1}$*. [Hint:* $\mathbf{M} = \begin{pmatrix} \mathbf{I}_r \\ \mathbf{C}\mathbf{A}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix}$*]*

2. *How do we know the inverses in the definition of* $\mathbf{P}$ *exist?*

[Penrose, 1955b]

PROBLEM 4.1.22 *Rao and Mitra [1971] and Barnett [1990] discuss variations of the generalized inverse obtained by relaxing the constraints on its definition given in* PROBLEM 4.1.16.

1. *Give an example of a generalized inverse* $\mathbf{X}$ *satisfying*

$$\mathbf{M}\mathbf{X}\mathbf{M} = \mathbf{M}, \tag{4.57}$$

but not satisfying at least one of the remaining conditions.

2. *Show that all generalized inverses satisfying (4.57) may be expressed in the form*

$$\mathbf{X} = \mathbf{X}_0 + \mathbf{Y} - \mathbf{X}_0\mathbf{M}\mathbf{Y}\mathbf{M}\mathbf{X}_0,$$

where $\mathbf{X}_0$ *is a particular matrix satisfying condition (4.57) and* $\mathbf{Y}$ *is an arbitrary* $n \times m$ *matrix. Explain this result using singular value decomposition.*

3. *Show that, if* $\mathbf{M}$ *is rearranged as in (4.55), then one choice of* $\mathbf{X}$ *in (4.57) is*

$$\mathbf{X} = \begin{pmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{Z} \\ 0 & \mathbf{Z} \end{pmatrix},$$

where $\mathbf{Z}$ *is an arbitrary matrix having the proper dimensions.*

PROBLEM 4.1.23 *Consider the stochastic inverse*

$$\mathbf{X} = \mathcal{C}_{st}\mathcal{C}_{tt}^{-1}.$$

introduced in PROBLEM 3.4.5*. If the noise* $\mathbf{N}$ *is negligible so that* $\mathbf{M}\mathbf{S} = \mathbf{T}$ *but* $\mathcal{C}_{tt}$ *is still invertible, show that* $\mathbf{X}$ *satisfies the first three of the Moore-Penrose conditions (see* PROBLEM 4.1.16*) for the generalized inverse. Show that this matrix is a special case of*

$$\mathbf{X} = \mathbf{Z}\left[\mathbf{M}\mathbf{Z}\right]^{-1}, \tag{4.58}$$

where $\mathbf{Z}$ *is an arbitrary* $n \times m$ *matrix except that* $\mathbf{M}\mathbf{Z}$ *must be invertible. What are the conditions on* $\mathbf{Z}$ *needed to guarantee that the third Moore-Penrose condition is satisfied? What is the matrix* $\mathbf{Z}$ *if* $\mathbf{X}$ *satisfies all four conditions?*

4.1.4 Relation to least-squares

Now we can solve the least-squares problem using the SVD of $\mathbf{M}$. To see this, we will let $w_i = 1$ for simplicity. To begin, first recognize that $\mathbf{s}$ and $\mathbf{t}$ may be expanded in terms of the left- and right-eigenvectors:

$$\mathbf{t} = \sum_{i=1}^{r} \tau_i \mathbf{y}_i + \mathbf{t}_0, \tag{4.59}$$

$$\mathbf{s} = \sum_{i=1}^{r} \sigma_i \mathbf{z}_i + \mathbf{s}_0, \tag{4.60}$$

where

$$\mathbf{z}_i^T \mathbf{s}_0 = \mathbf{y}_i^T \mathbf{t}_0 = 0 \quad \text{for all} \quad i = 1, \ldots, r, \tag{4.61}$$

and

$$\tau_i = \mathbf{y}_i^T \mathbf{t}, \tag{4.62}$$

$$\sigma_i = \mathbf{z}_i^T \mathbf{s}. \tag{4.63}$$

In terms of the expansion coefficients and unit weights, we have

$$\Psi(\mathbf{s}) = (\mathbf{Ms} - \mathbf{t})^T (\mathbf{Ms} - \mathbf{t}) \tag{4.64}$$

$$= \mathbf{t}_0^T \mathbf{t}_0 + \sum_{i=1}^{r} (\lambda_i \sigma_i - \tau_i)^2. \tag{4.65}$$

For nonzero eigenvalues, setting

$$\lambda_i \sigma_i = \tau_i \tag{4.66}$$

minimizes Ψ by eliminating the sum in (4.65). Then,

$$\mathbf{s} = \mathbf{s}_0 + \sum_{i=1}^{r} \lambda_i^{-1} \tau_i \mathbf{z}_i. \tag{4.67}$$

The vector $\mathbf{s}_0$ is an arbitrary vector from the right null space of $\mathbf{M}$. We can minimize $\mathbf{s}^T\mathbf{s}$ by setting $\mathbf{s}_0 = 0$. Thus, we obtain the minimum-norm least-squares model:

$$\hat{\mathbf{s}}_{\mathrm{LS}} = \sum_{i=1}^{r} \lambda_i^{-1} \tau_i \mathbf{z}_i. \tag{4.68}$$

The reader can easily verify that

$$\hat{\mathbf{s}}_{\mathrm{LS}} = \mathbf{M}^{\dagger} \mathbf{t}. \tag{4.69}$$

It is a general result that the Moore-Penrose pseudoinverse solves the least-squares problem. We will make use of this fact later when we attempt to construct methods of solving the inversion problem that are both fast and easy to implement.

Problems

Problem 4.1.24 *Verify (4.69).*

Problem 4.1.25 *Define* $\mathbf{Z}_0$ *to be a "best approximate solution" of the matrix equation* $\mathbf{MZ} = \mathbf{Y}$ *if for all* $\mathbf{Z}$, *either*

1. $||\mathbf{MZ} - \mathbf{Y}|| > ||\mathbf{MZ}_0 - \mathbf{Y}||$, *or*

2. $||\mathbf{MZ} - \mathbf{Y}|| = ||\mathbf{MZ}_0 - \mathbf{Y}||$ *and* $||\mathbf{Z}|| = ||\mathbf{Z}_0||$,

where $||\mathbf{A}||^2 \equiv Tr(\mathbf{A}^T\mathbf{A})$. *Then, show that* $\mathbf{Z} = \mathbf{M}^\dagger\mathbf{Y}$ *is the unique best approximate solution of* $\mathbf{MZ} = \mathbf{Y}$. *[Penrose, 1955b]*

Problem 4.1.26 *Let* $\mathbf{Z}_0 = \mathbf{M}^\dagger\mathbf{Y}$ *and*

$$\mathbf{Z} = \mathbf{Z}_0 + \sum_{i=r+1}^{n} \alpha_i \mathbf{z}_i \mathbf{y}_i^T \mathbf{Y},$$

where the α_i*s are scalars and the* $\mathbf{z}_i$*s and* $\mathbf{y}_i$*s for* $r+1 \le i \le n$ *are vectors from the right and left null spaces of* $\mathbf{M}$. *Then, which condition in* Problem 4.1.25 *does* $\mathbf{Z}$ *violate? Under what circumstances is the "best approximate solution" defined in* Problem 4.1.25 *really the best?*

Problem 4.1.27 *Show that*

$$\hat{\mathbf{s}}_{\mathrm{LS}} = \mathcal{R}_n \mathbf{s},$$

where $\mathcal{R}_n = \mathbf{M}^\dagger\mathbf{M}$ *is the model space resolution matrix. [Backus and Gilbert, 1968; 1970; Jackson, 1972]*

Problem 4.1.28 *Show that*

$$\mathbf{M}\hat{\mathbf{s}}_{\mathrm{LS}} = \mathcal{R}_m \mathbf{t},$$

where $\mathcal{R}_m = \mathbf{M}\mathbf{M}^\dagger$ *is the data space resolution matrix. [Wiggins, 1972; Jackson, 1972]*

4.2 Scaling Methods

Given $\mathbf{M}$ we define two diagonal matrices based on row and column sums of its elements, l_{ij}. Let $\mathbf{L}$ and $\mathbf{C}$ be diagonal matrices such that

$$L_{ii} = \sum_{j=1}^{n} l_{ij}, \quad i = 1, \ldots m, \tag{4.70}$$

$$C_{jj} = \sum_{i=1}^{m} l_{ij}, \quad j = 1, \ldots, n. \tag{4.71}$$

L_{ii} is the length of the ith ray path, obtained by summing the lengths of its intersection with all cells. C_{jj}, on the other hand, is the total length of ray segments intersecting the jth cell. C_{jj} (or its minor variations) is known variously as the *illumination*, *hit parameter*, or *coverage* of cell j.

Let $\mathbf{v}$ be the n-vector whose components are each 1:

$$\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{4.72}$$

Similarly, let $\mathbf{u}$ be the analogous m-vector. Then $\mathbf{Mv}$ is the m-vector containing the ray lengths. We can also infer that $\mathbf{Lu}$ is the same vector. Analogously, $\mathbf{M}^T\mathbf{u}$ and $\mathbf{Cv}$ are both the n-vector containing the cell coverages. That is,

$$\mathbf{Mv} = \mathbf{Lu}, \tag{4.73}$$

$$\mathbf{M}^T\mathbf{u} = \mathbf{Cv}. \tag{4.74}$$

This implies that $\lambda = 1$, $\mathbf{y} = \mathbf{u}$, $\mathbf{z} = \mathbf{v}$ is a solution to the eigenvalue problem

$$\mathbf{Mz} = \lambda\mathbf{Ly}, \tag{4.75}$$

$$\mathbf{M}^T\mathbf{y} = \lambda\mathbf{Cz}. \tag{4.76}$$

This problem is a generalization of our earlier eigenvalue problem (4.5)–(4.6) in that it incorporates positive definite weighting matrices $\mathbf{L}$ and $\mathbf{C}$. In place of the orthonormality conditions (4.12), we require the *conjugacy* conditions

$$\mathbf{y}_i^T\mathbf{L}\mathbf{y}_j = \mathbf{z}_i^T\mathbf{C}\mathbf{z}_j = \delta_{ij}. \tag{4.77}$$

With these conditions, the generalized eigenvalue problem can be converted to the standard form of (4.5)–(4.6) using the (*preconditioning*) transformations

$$\mathbf{M}' = \mathbf{L}^{-1/2}\mathbf{M}\mathbf{C}^{-1/2}, \tag{4.78}$$

$$\mathbf{y}' = \mathbf{L}^{1/2}\mathbf{y}, \tag{4.79}$$

$$\mathbf{z}' = \mathbf{C}^{1/2}\mathbf{z}. \tag{4.80}$$

Whenever $\mathbf{L}^{-\frac{1}{2}}$ and $\mathbf{C}^{-\frac{1}{2}}$ appear in the formulas, we make the implicit assumption that all the diagonal elements of both these matrices are nonzero. A zero diagonal component of $\mathbf{L}$ would correspond to a ray path with no length, which is clearly unphysical. However, a zero diagonal component of $\mathbf{C}$ corresponds to a cell with no ray coverage, which clearly can and does happen in practice. If so, then we assume that this cell is removed from the inversion problem.[2]

By construction, the eigenvalues of $\mathbf{M}$ and $\mathbf{M}'$ are the same, but for different eigenvalue problems: (4.75)–(4.76) and

$$\mathbf{M}'\mathbf{z}' = \lambda\mathbf{y}', \tag{4.81}$$

$$\mathbf{M}'^T\mathbf{y}' = \lambda\mathbf{z}'. \tag{4.82}$$

[2]Some methods for doing so are discussed in Section 6.

PROPOSITION 4.2.1 *The eigenvalues of* $\mathbf{M}'$ *lie in the interval [-1,1].*

Proof: Recall that the eigenvalues come in pairs: if $\lambda, \mathbf{y}', \mathbf{z}'$ solves the eigenvalue problem, so does $-\lambda, \mathbf{y}', -\mathbf{z}'$. Then, we may (without loss of generality) restrict the discussion to eigenvalues satisfying $\lambda \geq 0$.

Let λ, $\mathbf{y}$, $\mathbf{z}$ be any solution to (4.75)–(4.76) with $\lambda > 0$. Then, in components,

$$\sum_j l_{ij} z_j = \lambda L_{ii} y_i, \tag{4.83}$$

$$\sum_i l_{ij} y_i = \lambda C_{jj} z_j. \tag{4.84}$$

Let $y_{\max}$ be the largest absolute component of $\mathbf{y}$, i.e., $y_{\max} = \max_i |y_i|$. Similarly, let $z_{\max} = \max_j |z_j|$. Since $l_{ij} \geq 0$, we can infer

$$z_{\max} \sum_j l_{ij} \geq \lambda L_{ii} |y_i|, \tag{4.85}$$

$$y_{\max} \sum_i l_{ij} \geq \lambda C_{jj} |z_j|. \tag{4.86}$$

Recalling the definitions of C_{jj} and L_{ii} given by (4.70) and (4.71), this implies

$$z_{\max} \geq \lambda |y_i|, \tag{4.87}$$

$$y_{\max} \geq \lambda |z_j|, \tag{4.88}$$

which must hold for all i and j; thus

$$z_{\max} \geq \lambda y_{\max}, \tag{4.89}$$

$$y_{\max} \geq \lambda z_{\max}. \tag{4.90}$$

Thus, we have $z_{\max} \geq \lambda^2 z_{\max}$, which implies $\lambda^2 \leq 1$ and therefore $-1 \leq \lambda \leq 1$. ∎

PROBLEMS

PROBLEM 4.2.1 *Show that the eigenvectors of* $\mathbf{M}'$ *having eigenvalue* $\lambda = 1$ *are* $\mathbf{y}' = \mathbf{L}^{\frac{1}{2}}\mathbf{u}$ *and* $\mathbf{z}' = \mathbf{C}^{\frac{1}{2}}\mathbf{v}$. *What are the eigenvectors of* $(\mathbf{M}')^\dagger$ *having unit eigenvalue?*

PROBLEM 4.2.2 *Using* $\mathbf{M}\mathbf{M}^\dagger\mathbf{M} = \mathbf{M}$ *and (4.70), demonstrate the general result*

$$\mathbf{M}\mathbf{M}^\dagger\mathbf{L}\mathbf{u} = \mathcal{R}_m\mathbf{L}\mathbf{u} = \mathbf{L}\mathbf{u}. \tag{4.91}$$

Thus, $\mathbf{Lu}$ *is an eigenvector of the data resolution matrix* $\mathcal{R}_m$, *having unit eigenvalue.*

PROBLEM 4.2.3 *Using* $\mathbf{M}\mathbf{M}^\dagger\mathbf{M} = \mathbf{M}$ *and (4.71), derive the general result that*

$$\mathbf{v}^T\mathbf{C}\mathbf{M}^\dagger\mathbf{M} = \mathbf{v}^T\mathbf{C}\mathcal{R}_n = \mathbf{v}^T\mathbf{C}. \tag{4.92}$$

Thus, $\mathbf{v}^T\mathbf{C}$ *is an eigenvector of the model resolution matrix* $\mathcal{R}_n$, *having unit eigenvalue.*

PROBLEM 4.2.4 *Three examples of data resolution matrices are presented in (4.46), (4.48), and (4.50). Check to see if they agree with (4.91).*

PROBLEM 4.2.5 *Three examples of model resolution matrices are presented in (4.45), (4.47), and (4.49). Check to see if they agree with (4.92).*

PROBLEM 4.2.6 *Use the definition of the pseudoinverse in* PROBLEM 4.1.16 *to show that, if*

$$\mathbf{M}' = \mathbf{L}^{-\frac{1}{2}}\mathbf{M}\mathbf{C}^{-\frac{1}{2}},$$

then

$$\mathbf{X} = \mathbf{C}^{-\frac{1}{2}}(\mathbf{M}')^{\dagger}\mathbf{L}^{-\frac{1}{2}}, \tag{4.93}$$

where $\mathbf{X}$ *is an approximate generalized inverse satisfying the first two conditions (*$\mathbf{MXM} = \mathbf{M}$ *and* $\mathbf{XMX} = \mathbf{X}$*). Use (4.93) to show that the SVDs of* $\mathbf{M}$ *and* $\mathbf{X}$ *have the form*

$$\mathbf{M} = \frac{\mathbf{L}\mathbf{u}\mathbf{v}^T\mathbf{C}}{\mathbf{u}^T\mathbf{L}\mathbf{u}} + \dots$$

and

$$\mathbf{X} = \frac{\mathbf{v}\mathbf{u}^T}{\mathbf{u}^T\mathbf{L}\mathbf{u}} + \dots, \tag{4.94}$$

where the terms not shown are for eigenvectors of $\mathbf{M}'$ *with eigenvalues* $\lambda < 1$*. Treat (4.94) as an approximate inverse, and compare the resulting estimate of* $\mathbf{s}$ *to (3.42). What can be said about the accuracy of this approximate inverse?*

PROBLEM 4.2.7 *Use (4.93) and (4.94) to show that*

$$\mathbf{u}^T\mathbf{M}\mathbf{X} = \mathbf{u}^T, \tag{4.95}$$

for a ray-path matrix $\mathbf{M}$*. Thus,* $\mathbf{u}$ *is a left eigenvector of the approximate data resolution matrix* $\mathbf{MX}$*, having unit eigenvalue. What restrictions (if any) are there on the validity of (4.95)?*

PROBLEM 4.2.8 *Use (4.93) and (4.94) to show that*

$$\mathbf{X}\mathbf{M}\mathbf{v} = \mathbf{v}, \tag{4.96}$$

for a ray-path matrix $\mathbf{M}$*. Thus,* $\mathbf{v}$ *is a right eigenvector of the approximate model resolution matrix* $\mathbf{XM}$*, having unit eigenvalue. What restrictions (if any) are there on the validity of (4.96)?*

4.3 Weighted Least-Squares, Regularization, and Effective Resolution

In weighted least-squares, a good choice of weighting matrix is $\mathbf{L}^{-1}$, that is, the inverse of the ray length matrix. In Section 3.5, we discussed the physical arguments for using such a weight matrix. Here we will show that mathematical arguments based on stability and regularization lead to the same choice of weight matrix.

4.3.1 General weights and objective functionals

There is an inherent arbitrariness to the choice of weight matrix in a least-squares minimization. Let $\mathbf{F}$ and $\mathbf{G}$ be two positive, diagonal weight matrices, $m \times m$ and $n \times n$ respectively. Then define the scaled inversion problem so that

$$\mathbf{M}' = \mathbf{F}^{-\frac{1}{2}}\mathbf{M}\mathbf{G}^{-\frac{1}{2}}, \quad \mathbf{s}' = \mathbf{G}^{\frac{1}{2}}\mathbf{s}, \quad \mathbf{t}' = \mathbf{F}^{-\frac{1}{2}}\mathbf{t}. \tag{4.97}$$

The (unweighted) damped least-squares minimization problem associated with (4.97) is to minimize the functional

$$\Psi'(\mathbf{s}') = (\mathbf{t}' - \mathbf{M}'\mathbf{s}')^T(\mathbf{t}' - \mathbf{M}'\mathbf{s}') + \mu(\mathbf{s}' - \mathbf{s}'_0)^T(\mathbf{s}' - \mathbf{s}'_0), \tag{4.98}$$

with respect to $\mathbf{s}'$. The normal equations resulting from (4.98) are

$$(\mathbf{M}'^T\mathbf{M}' + \mu\mathbf{I})(\mathbf{s}' - \mathbf{s}'_0) = \mathbf{M}'^T(\mathbf{t}' - \mathbf{M}'\mathbf{s}'_0). \tag{4.99}$$

The result for the untransformed $\mathbf{s}$ is exactly the same whether we use the functional (4.98) or the weighted least-squares functional

$$\Psi(\mathbf{s}) = (\mathbf{t} - \mathbf{M}\mathbf{s})^T\mathbf{F}^{-1}(\mathbf{t} - \mathbf{M}\mathbf{s}) + \mu(\mathbf{s} - \mathbf{s}_0)^T\mathbf{G}(\mathbf{s} - \mathbf{s}_0). \tag{4.100}$$

In either case, the result is

$$\mathbf{s} = \mathbf{s}_0 + (\mathbf{M}^T\mathbf{F}^{-1}\mathbf{M} + \mu\mathbf{G})^{-1}\mathbf{M}^T\mathbf{F}^{-1}(\mathbf{t} - \mathbf{M}\mathbf{s}_0). \tag{4.101}$$

In truth, every least-squares method is a special case of the general weighted least-squares method — the more common ones just have unit weights everywhere.

The minimum of (4.100) is achieved by the slowness model given in (4.101) as long as the matrix $\mathbf{M}^T\mathbf{F}^{-1}\mathbf{M} + \mu\mathbf{G}$ is invertible. Thus, some relaxation of the conditions placed on the weight matrix $\mathbf{G}$ is possible. One common choice is to make the regularization term correspond to minimizing the gradient or curvature of the model. Then, the matrix $\mathbf{G} = \mathbf{K}^T\mathbf{K}$, where $\mathbf{K}\mathbf{s}$ is either the gradient of the model or its Laplacian. Such a weight matrix is neither diagonal nor positive. In fact, a constant model vector lies in the null space of such a $\mathbf{G}$. The combined matrix $\mathbf{M}^T\mathbf{F}^{-1}\mathbf{M} + \mu\mathbf{G}$ may still be invertible however, since the null spaces of the two terms are generally orthogonal.

PROBLEMS

PROBLEM 4.3.1 *Find explicit expressions for the matrix* $\mathbf{K}$ *such that*

1. $\mathbf{Ks}$ *is the gradient of the slowness model;*

2. $\mathbf{Ks}$ *is the Laplacian of the slowness model.*

PROBLEM 4.3.2 *Suppose the desired weight matrices* $\mathbf{F}$ *and/or* $\mathbf{G}$ *are nonnegative diagonal, (i.e., have some zeroes along the diagonal). Generalize (4.97), replacing these definitions by*

$$\mathbf{M}' = (\mathbf{F}^\dagger)^{\frac{1}{2}}\mathbf{M}(\mathbf{G}^\dagger)^{\frac{1}{2}}, \quad \mathbf{s}' = \mathbf{G}^{\frac{1}{2}}\mathbf{s}, \quad \mathbf{t}' = (\mathbf{F}^\dagger)^{\frac{1}{2}}\mathbf{t},$$

where $\mathbf{F}^\dagger$ *and* $\mathbf{G}^\dagger$ *are the pseudoinverses of* $\mathbf{F}$ *and* $\mathbf{G}$ *respectively. If* $\mathbf{F} = \mathbf{L}$ *and* $\mathbf{G} = \mathbf{C}$ *with some of the cell coverages vanishing, compare the approach using generalized inverses to the usual method of deleting uncovered cells from the inversion problem.*

PROBLEM 4.3.3 *Suppose that the damping parameter* $\mu = 0$ *and the diagonal elements of* $\mathbf{F}$ *are given by*

$$F_{ii} = |(\mathbf{Ms})_i - t_i|,^{2-p}$$

where the term in the exponent $p \geq 1$ *(*$p = 2$ *for least-squares). Show that the resulting special case of (4.101) is the slowness minimizing*

$$\Psi_p(\mathbf{s}) = \sum_{i=1}^{m} |(\mathbf{Ms})_i - t_i|.^p$$

Assuming that some of the traveltime residuals vanish and $p = 1$*, use the result of* PROBLEM 4.3.2 *to provide an appropriate generalization of (4.101). This method is known as* iteratively reweighted least-squares. *[Claerbout and Muir, 1973; Claerbout, 1976; Scales, Gersztenkorn, and Treitel, 1988]*

4.3.2 Regularization

There are physical reasons for choosing particular weighting schemes and some of these reasons have been discussed in Section 3.5. A sound mathematical reason for choosing a particular scheme [Burkhard, 1980] might be either "convergence" or "regularization." It may be difficult or impossible to compute the result (4.101) unless appropriate weight matrices are used, since $\mathbf{M}^T\mathbf{F}^{-1}\mathbf{M}$ may be poorly conditioned or noninvertible. We will see in our discussion of *simple iteration* (Section 4.4.3) that this method converges if the eigenvalues of the matrix $\mathbf{M}$ (or equivalently $\mathbf{M}'$ here) lie in the range $-\sqrt{2} \leq \lambda_i \leq \sqrt{2}$. So how can we choose the weight matrices to guarantee that the eigenvalues fall in the desired range?

For the sake of argument, suppose that

$$\mathbf{Ms} = \lambda\mathbf{Fr}, \tag{4.102}$$
$$\mathbf{M}^T\mathbf{r} = \lambda\mathbf{Gs}. \tag{4.103}$$

Then, in terms of components, we have

$$\sum_j l_{ij}s_j = \lambda F_{ii}r_i, \tag{4.104}$$

$$\sum_i l_{ij}r_j = \lambda G_{jj}s_i. \tag{4.105}$$

Letting $s_{\max}$ be the magnitude of the largest component of $\mathbf{s}$ and $r_{\max}$ the magnitude of the largest component of $\mathbf{r}$, we have

$$s_{\max}L_{ii} \geq \lambda F_{ii}|r_i|, \tag{4.106}$$

$$r_{\max}C_{jj} \geq \lambda G_{jj}|s_j|. \tag{4.107}$$

It follows that

$$s_{\max} \geq \lambda\frac{F_{ii}}{L_{ii}}r_{\max} \geq \lambda^2\frac{F_{ii}G_{jj}}{L_{ii}C_{jj}}s_{\max}. \tag{4.108}$$

So, in general, we can guarantee that the eigenvalues λ will be bounded above by unity by requiring that

$$1 \geq \frac{L_{ii}C_{jj}}{F_{ii}G_{jj}}, \quad \text{for all} \quad i,j. \tag{4.109}$$

Many choices of $\mathbf{F}$ and $\mathbf{G}$ are permitted by (4.109), but perhaps the simplest choice is

$$\mathbf{F} = \mathbf{L} \quad \text{and} \quad \mathbf{G} = \mathbf{C}. \tag{4.110}$$

Thus, although the choice (4.110) is certainly not unique, it is nevertheless a good choice for the weight matrices in weighted least-squares, and guarantees that $\lambda^2 \leq 1$ as desired.

In Section 7.2, we find that another choice of weight matrices has the same constraining properties on the eigenvalues, yet has more useful properties in nonlinear tomography algorithms.

4.3.3 Effective resolution

Another way of understanding the significance of formulas such as (4.101) is to reconsider the fundamental relation

$$\mathbf{Ms} = \mathbf{t} \tag{4.111}$$

and its rearrangement

$$\mathbf{M}\Delta\mathbf{s} = \mathbf{t} - \mathbf{M}\mathbf{s}_0, \tag{4.112}$$

where $\Delta\mathbf{s} = \mathbf{s} - \mathbf{s}_0$. Now, view the matrix

$$\mathbf{X} = (\mathbf{M}^T\mathbf{F}^{-1}\mathbf{M} + \mu\mathbf{G})^{-1}\mathbf{M}^T\mathbf{F}^{-1} \tag{4.113}$$

as an approximate generalized inverse of $\mathbf{M}$. Multiplying (4.112) on the left by $\mathbf{X}$, we have

$$\mathbf{XM}\Delta\mathbf{s} = (\mathbf{M}^T\mathbf{F}^{-1}\mathbf{M} + \mu\mathbf{G})^{-1}\mathbf{M}^T\mathbf{F}^{-1}(\mathbf{t} - \mathbf{M}\mathbf{s}_0). \tag{4.114}$$

If $\mathbf{XM} = \mathbf{I}$ were true, then (4.114) and (4.101) would be identical; however, for any $\mu > 0$,

$$\mathbf{XM} = \mathbf{I} - \mu(\mathbf{M}^T\mathbf{F}^{-1}\mathbf{M} + \mu\mathbf{G})^{-1}\mathbf{G} \neq \mathbf{I}. \tag{4.115}$$

Thus, the product $\mathbf{XM}$ is analogous to the resolution matrix $\mathcal{R}_n = \mathbf{M}^\dagger\mathbf{M}$. We will call the product $\mathbf{XM}$ the $n \times n$ *effective resolution matrix*

$$\mathbf{XM} \equiv \mathcal{E}_n, \tag{4.116}$$

and similarly define the product

$$\mathbf{MX} \equiv \mathcal{E}_m, \tag{4.117}$$

as the $m \times m$ effective resolution matrix.

The generalized inverse $\mathbf{M}^\dagger$ gives optimal performance in the sense that no other choice of inverse can produce resolution matrices closer to the identity matrix than $\mathcal{R}_n$ and $\mathcal{R}_m$. To see how well the approximate inverse $\mathbf{X}$ does in this regard, we can compare the effective resolution matrices with the optimal ones. For simplicity, consider the case $\mathbf{F} = \mathbf{G} = \mathbf{I}$. Then, the SVD of $\mathcal{E}_n$ shows that

$$\mathcal{E}_n = (\mathbf{M}^T\mathbf{M} + \mu\mathbf{I})^{-1}\mathbf{M}^T\mathbf{M} = \sum_{j=1}^{r}\left(\frac{\lambda_j^2}{\lambda_j^2 + \mu}\right)\mathbf{z}_j\mathbf{z}_j^T. \tag{4.118}$$

Then, it is easy to see that the effective resolution matrix is closely related to the resolution matrix $\mathcal{R}_n$ by

$$\mathcal{E}_n = \mathcal{R}_n - \sum_{j=1}^{r}\frac{\mu}{\lambda_j^2 + \mu}\mathbf{z}_j\mathbf{z}_j^T. \tag{4.119}$$

If the damping parameter is sufficiently small but still positive ($\mu \to 0^+$), we expect $\mathcal{E}_n \to \mathcal{R}_n$. Similarly, the effective resolution matrix $\mathcal{E}_m$ satisfies

$$\mathcal{E}_m = \mathbf{M}(\mathbf{M}^T\mathbf{M} + \mu\mathbf{I})^{-1}\mathbf{M}^T = \mathcal{R}_m - \sum_{i=1}^{r}\frac{\mu}{\lambda_i^2 + \mu}\mathbf{y}_i\mathbf{y}_i^T. \tag{4.120}$$

and $\mathcal{E}_m \to \mathcal{R}_m$ as $\mu \to 0^+$. We see then that the effective resolution matrices are suboptimal, but approach optimal in the limit that $\mu \to 0^+$. This calculation shows that the approximate inverse $\mathbf{X}$ is biased, but not very strongly biased if very small values of μ are used.

Problem

PROBLEM 4.3.4 *Use (4.119) and (4.120) to show that* $Tr(\mathcal{E}_n) = Tr(\mathcal{E}_m) < rank(\mathbf{M})$ *if* $\mu > 0$.

4.4 Sequential and Iterative Methods

First consider the case where $r = n$. The least-squares solution is then given by

$$\hat{\mathbf{s}}_{\mathrm{LS}} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{t}. \tag{4.121}$$

We begin by summarizing the main ideas behind two matrix inversion methods that work if $\mathbf{M}^T\mathbf{M}$ is invertible. Then, we discuss other methods applicable to more realistic problems in tomography.

Our main focus in this discussion will be to elucidate the general principles behind these methods and to show how they relate to the Moore-Penrose pseudoinverse. A later section will be devoted to evaluating iterative methods and making some jugdments about which algorithms are best for tomography and inversion problems.

4.4.1 Series expansion method

Again letting $\mathbf{A} = \mathbf{M}^T\mathbf{M}$, observe that $\mathbf{A}$ is square and suppose it to be of full rank. In terms of the SVD of $\mathbf{M}$,

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i^2 \mathbf{z}_i \mathbf{z}_i^T. \tag{4.122}$$

Let $\rho_i = \lambda_i^2$. Then, since $\mathbf{A}$ satisfies its own characteristic polynomial, we have the following matrix identity:

$$(\mathbf{A} - \rho_1\mathbf{I})(\mathbf{A} - \rho_2\mathbf{I}) \ldots (\mathbf{A} - \rho_n\mathbf{I}) = 0. \tag{4.123}$$

The left-hand side of this equation is simply an nth order matrix polynomial in $\mathbf{A}$, which can be rewritten as

$$\mathbf{A}^n - (\rho_1 + \cdots + \rho_n)\mathbf{A}^{n-1} + \cdots + (-1)^n \rho_1 \cdots \rho_n \mathbf{I} = 0. \tag{4.124}$$

Multiplying through formally by $\mathbf{A}^{-1}$

$$\mathbf{A}^{n-1} - (\rho_1 + \cdots + \rho_n)\mathbf{A}^{n-2} + \cdots + (-1)^n \rho_1 \cdots \rho_n \mathbf{A}^{-1} = 0, \tag{4.125}$$

or

$$\mathbf{A}^{-1} = \frac{(-1)^{n+1}}{\rho_1 \cdots \rho_n}\left[\mathbf{A}^{n-1} - (\rho_1 + \cdots + \rho_n)\mathbf{A}^{n-2} + \cdots + \mathbf{I}\right]. \tag{4.126}$$

This gives a series expansion for $\mathbf{A}^{-1}$ in powers of $\mathbf{A}$ itself. Based on this series, $\mathbf{A}^{-1}\mathbf{M}^T\mathbf{t}$ may be computed recursively if the eigenvalues of $\mathbf{A}$ are known, or at least if the symmetric functions of the eigenvalues that appear in the formulas are known.

This approach clearly fails if $\mathbf{A}$ is not of full rank, since the multiplication leading to (4.125) cannot be performed. The final division by the product of the eigenvalues in (4.126) also cannot be performed.

PROBLEM

PROBLEM 4.4.1 *A real symmetric matrix* $\mathbf{B}$ *has a single vector* $\mathbf{w}$ *in its null space. Use (4.126) to find an expression for the pseudoinverse* $\mathbf{B}^\dagger$.

4.4.2 Conjugate directions and conjugate gradients

In the method of conjugate directions [Hestenes and Stiefel, 1952], a different expansion of $\mathbf{A}^{-1}$ is used. Let $\mathbf{p}_1, \ldots, \mathbf{p}_n$ be a set of vectors such that

$$\mathbf{p}_i^T\mathbf{A}\mathbf{p}_j = \delta_{ij}\mathbf{p}_i^T\mathbf{A}\mathbf{p}_i. \tag{4.127}$$

The vectors $\mathbf{p}_i$ are not necessarily orthogonal with respect to the usual vector dot product, but by construction they are orthogonal relative to the matrix $\mathbf{A}$. The vectors $\mathbf{p}_i$ are said to be conjugate relative to $\mathbf{A}$. Then consider

$$\mathbf{A}' \equiv \sum_{i=1}^{n} \frac{\mathbf{p}_i\mathbf{p}_i^T}{\mathbf{p}_i^T\mathbf{A}\mathbf{p}_i}. \tag{4.128}$$

It follows from (4.127) and (4.128) that

$$\mathbf{A}'(\mathbf{A}\mathbf{p}_j) = \sum_{i=1}^{n} \mathbf{p}_i\delta_{ij} = \mathbf{p}_j, \tag{4.129}$$

and

$$(\mathbf{p}_j^T\mathbf{A})\mathbf{A}' = \sum_{i=1}^{n} \mathbf{p}_i^T\delta_{ij} = \mathbf{p}_j^T, \tag{4.130}$$

which is also just the transpose of (4.129) since $\mathbf{A} = \mathbf{A}^T$. Thus, if the $\mathbf{p}_i$s span the entire vector space (*i.e.*, if they are complete), (4.129) and (4.130) show that

$$\mathbf{A}'\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}'. \tag{4.131}$$

Uniqueness of the inverse then implies that

$$\mathbf{A}^{-1} = \mathbf{A}'. \tag{4.132}$$

The completeness relation in terms of the $\mathbf{p}_i$s is therefore

$$\mathbf{I} = \sum_{i=1}^{n} \frac{\mathbf{p}_i\mathbf{p}_i^T\mathbf{A}}{\mathbf{p}_i^T\mathbf{A}\mathbf{p}_i} = \sum_{i=1}^{n} \frac{\mathbf{A}\mathbf{p}_i\mathbf{p}_i^T}{\mathbf{p}_i^T\mathbf{A}\mathbf{p}_i}. \tag{4.133}$$

This approach produces a valid and simple formula (4.128) for $\mathbf{A}^{-1}$ when $\mathbf{A}$ is of full rank, and furthermore it is guaranteed to converge in a finite number of steps (see PROBLEM 4.4.5). But, when $\mathbf{A}$ is rank deficient, it must happen that $\mathbf{p}_i^T\mathbf{A}\mathbf{p}_i = 0$ for some $\mathbf{p}_i$ and, therefore, this method also fails in the cases often of most interest in tomography.

Conjugate directions may still be useful for singular $\mathbf{A}$s if care is taken to choose only $\mathbf{p}_i$s orthogonal to the null space of $\mathbf{A}$. Then, this approach may be used to generate the generalized inverse of $\mathbf{A}$.

To see an example of how this works, consider the method of conjugate gradients [Hestenes and Stiefel, 1952; Golub and Van Loan, 1983; Ashby, Manteuffel, and Saylor, 1990] for solving $\mathbf{Ms} = \mathbf{t}$ in the least-squares sense. The normal equations take the form

$$\mathbf{M}^T\mathbf{Ms} = \mathbf{M}^T\mathbf{t}, \tag{4.134}$$

which may be rewritten for these purposes as

$$\mathbf{As} = \mathbf{b} \tag{4.135}$$

where $\mathbf{A} = \mathbf{M}^T\mathbf{M}$ and $\mathbf{b} = \mathbf{M}^T\mathbf{t}$. Then, starting with the error vector $\mathbf{r}^{(1)} = \mathbf{b} - \mathbf{As}^{(1)}$ equal to the first direction vector $\mathbf{p}^{(1)}$, the conjugate-gradient method uses the iteration scheme

$$\mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \frac{(\mathbf{p}^{(k)})^T\mathbf{r}^{(k)}}{(\mathbf{p}^{(k)})^T\mathbf{Ap}^{(k)}}\mathbf{p}^{(k)}, \tag{4.136}$$

$$\mathbf{r}^{(k+1)} = \mathbf{b} - \mathbf{As}^{(k+1)}, \tag{4.137}$$

$$\mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} - \frac{(\mathbf{p}^{(k)})^T\mathbf{Ar}^{(k+1)}}{(\mathbf{p}^{(k)})^T\mathbf{Ap}^{(k)}}\mathbf{p}^{(k)}. \tag{4.138}$$

The philosophy of this method is to generate a sequence of directions by taking the latest error vector (4.137) as the primary source and then orthogonalizing (4.138) relative to $\mathbf{A}$ with respect to all previous directions taken (see PROBLEM 4.4.4). Since the error vectors are all of the form $\mathbf{r} = \mathbf{M}^T \times$ *a vector*, this iteration sequence cannot generate vectors in the right null space of $\mathbf{M}$. Thus, in principle, this method can converge to the minimum-norm least-squares solution.

Nevertheless, finite but small eigenvalues can have a large effect through the influence of the denominators appearing in (4.136) and (4.138). Small computational errors get magnified under circumstances of poor conditioning. Regularization of this method can be achieved by terminating the process when the latest direction vector satisfies $(\mathbf{p}^{(k+1)})^T\mathbf{Ap}^{(k+1)} \leq \epsilon$ where the scalar ϵ is some preset threshold, or by adding a small positive constant μ to the diagonal elements of $\mathbf{A}$.

PROBLEMS

PROBLEM 4.4.2 *Using definition (4.128), show that*

$$\mathbf{A}'\mathbf{A}\mathbf{A}' = \mathbf{A}'. \tag{4.139}$$

Then, use the positivity of $\mathbf{A}$ *together with (4.128) to show that (4.139) implies*

$$\mathbf{A}'\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}'.$$

Prove the inverse of a matrix is unique and therefore that $\mathbf{A}' = \mathbf{A}^{-1}$.

PROBLEM 4.4.3 *Show that, if*

$$\mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \alpha^{(k)}\mathbf{p}^{(k)}$$

is one in a sequence of iterates to solve $\mathbf{As} = \mathbf{b}$ *for* $\mathbf{s}$ *and if*

$$(\mathbf{p}^{(k)})^T\mathbf{Ap}^{(i)} = 0 \quad \textit{for} \quad i = 1, \dots, k-1,$$

then a residual reducing choice of the scalar $\alpha^{(k)}$ *is*

$$\alpha^{(k)} = \frac{(\mathbf{p}^{(k)})^T[\mathbf{b} - \mathbf{As}^{(k)}]}{(\mathbf{p}^{(k)})^T\mathbf{Ap}^{(k)}}.$$

PROBLEM 4.4.4 *Show that* $\mathbf{p}^{(k+1)}$ *as defined in (4.138) is conjugate to* $\mathbf{p}^{(i)}$ *for* $1 \le i \le k$.

PROBLEM 4.4.5 *Define the matrix*

$$\mathbf{A}'_k = \sum_{i=1}^{k} \frac{\mathbf{p}^{(i)}(\mathbf{p}^{(i)})^T}{(\mathbf{p}^{(i)})^T \mathbf{A} \mathbf{p}^{(i)}}.$$

Show that $\mathbf{A}'_k \mathbf{A} \mathbf{A}'_k = \mathbf{A}'_k$. *Then, show that the iterates obtained in the conjugate-gradient method (4.136)–(4.138) satisfy*

$$\mathbf{s}^{(k+1)} = \mathbf{A}'_k \mathbf{b} + \left(\mathbf{I} - \mathbf{A}'_k \mathbf{A}\right) \mathbf{s}^{(1)}. \tag{4.140}$$

Use this expression to show that the iteration converges to the solution in n *steps if* $\mathbf{A}$ *is an* $n \times n$ *positive matrix. What conditions (if any) are required on the eigenvalues of* $\mathbf{A}$ *for this scheme to converge?*

PROBLEM 4.4.6 *Consider the set of vectors*

$$\mathbf{x}, \mathbf{A}\mathbf{x}, \mathbf{A}^2\mathbf{x}, \ldots, \mathbf{A}^n\mathbf{x},$$

where $\mathbf{x}$ *is not an eigenvector of* $\mathbf{A}$ *but has components along all eigenvectors with nonzero eigenvalues (i.e.,* $\mathbf{x}^T\mathbf{z}_i \ne 0$ *for* $1 \le i \le r$*). Then, use Gram-Schmidt orthogonalization to produce the new set of vectors*

$$\mathbf{x}^{(k+1)} = \left[\mathbf{I} - \sum_{j=1}^{k} \frac{\mathbf{x}^{(j)}(\mathbf{x}^{(j)})^T}{(\mathbf{x}^{(j)})^T \mathbf{x}^{(j)}}\right] \mathbf{A}^{k-1}\mathbf{x}, \tag{4.141}$$

where $\mathbf{x}^{(1)} = \mathbf{x}$, $\mathbf{x}^{(2)} = \mathbf{A}\mathbf{x} - \mathbf{x}(\mathbf{x}^T\mathbf{A}\mathbf{x}/\mathbf{x}^T\mathbf{x})$, ... *How many orthogonal vectors can be produced using this technique? If an arbitrary vector* $\mathbf{x}$ *is chosen, analyze the behavior of this procedure in terms of the eigenvectors of* $\mathbf{A}$.

PROBLEM 4.4.7 *What changes must be made in the conjugate-gradient method in order to solve a weighted least-squares problem?*

PROBLEM 4.4.8 *Use the defining relations of conjugate gradients (4.136)–(4.138) to show that*

$$\frac{(\mathbf{p}^{(k)})^T \mathbf{r}^{(k)}}{(\mathbf{p}^{(k)})^T \mathbf{A} \mathbf{p}^{(k)}} = \frac{(\mathbf{r}^{(k)})^T \mathbf{r}^{(k)}}{(\mathbf{p}^{(k)})^T \mathbf{A} \mathbf{p}^{(k)}} = \frac{(\mathbf{r}^{(k)})^T \mathbf{r}^{(k)}}{(\mathbf{r}^{(k)})^T \mathbf{A} \mathbf{p}^{(k)}} \tag{4.142}$$

and

$$\frac{(\mathbf{p}^{(k)})^T \mathbf{A} \mathbf{r}^{(k+1)}}{(\mathbf{p}^{(k)})^T \mathbf{A} \mathbf{p}^{(k)}} = -\frac{(\mathbf{r}^{(k+1)})^T \mathbf{r}^{(k+1)}}{(\mathbf{r}^{(k)})^T \mathbf{r}^{(k)}}. \tag{4.143}$$

[Hestenes and Stiefel, 1952]

PROBLEM 4.4.9 *Use the results of* PROBLEM 4.4.8 *to show that the following algorithm is a conjugate-gradients algorithm for the traveltime inversion problem:*

$$
\begin{array}{l}
\mathbf{s}^{(0)} = 0; \quad \mathbf{p}^{(0)} = \mathbf{r}^{(0)} = \mathbf{M}^T\mathbf{t}; \\
\textit{for} \quad k = 0, 1, 2, \ldots \\
\{ \\
\quad \alpha_k = \frac{(\mathbf{r}^{(k)})^T \mathbf{r}^{(k)}}{(\mathbf{p}^{(k)})^T \mathbf{M}^T \mathbf{M} \mathbf{p}^{(k)}}; \\
\quad \mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \alpha_k \mathbf{p}^{(k)}; \\
\quad \mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k \mathbf{M}^T \mathbf{M} \mathbf{p}^{(k)}; \\
\quad \textit{if} \quad |\mathbf{r}^{(k+1)}| \quad \textit{is below threshold then quit}; \\
\quad \beta_k = \frac{(\mathbf{r}^{(k+1)})^T \mathbf{r}^{(k+1)}}{(\mathbf{r}^{(k)})^T \mathbf{r}^{(k)}}; \\
\quad \mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k \mathbf{p}^{(k)}; \\
\}
\end{array}
$$

[Hestenes and Stiefel, 1952; van der Sluis and van der Vorst, 1987]

PROBLEM 4.4.10 *Use the result of* PROBLEM 4.4.8 *to show that the following algorithm is also a conjugate-gradients algorithm:*

$$
\begin{array}{l}
\mathbf{s}^{(0)} = 0; \quad \mathbf{t}^{(0)} = \mathbf{t}; \quad \mathbf{p}^{(0)} = \mathbf{r}^{(0)} = \mathbf{M}^T\mathbf{t}; \\
\textit{for} \quad k = 0, 1, 2, \ldots \\
\{ \\
\quad \mathbf{q}^{(k)} = \mathbf{M}\mathbf{p}^{(k)}; \\
\quad \alpha_k = \frac{(\mathbf{r}^{(k)})^T \mathbf{r}^{(k)}}{(\mathbf{q}^{(k)})^T \mathbf{q}^{(k)}}; \\
\quad \mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \alpha_k \mathbf{p}^{(k)}; \\
\quad \mathbf{t}^{(k+1)} = \mathbf{t}^{(k)} - \alpha_k \mathbf{q}^{(k)}; \\
\quad \mathbf{r}^{(k+1)} = \mathbf{M}^T \mathbf{t}^{(k+1)}; \\
\quad \textit{if} \quad |\mathbf{r}^{(k+1)}| \quad \textit{is below threshold then quit}; \\
\quad \beta_k = \frac{(\mathbf{r}^{(k+1)})^T \mathbf{r}^{(k+1)}}{(\mathbf{r}^{(k)})^T \mathbf{r}^{(k)}}; \\
\quad \mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k \mathbf{p}^{(k)}; \\
\}
\end{array}
$$

Note that this variation on conjugate gradients does not require the formation of the (generally) dense matrix $\mathbf{M}^T\mathbf{M}$. *[Björck and Elfving, 1979; Paige and Saunders, 1982; van der Sluis and van der Vorst, 1987]*

4.4.3 Simple iteration

In simple iteration,[3] we start with an initial model $\mathbf{s}^{(0)}$ and iteratively generate a sequence $\mathbf{s}^{(k)}$, $k = 1, 2, \ldots$ using

$$\mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \mathbf{M}^T(\mathbf{t} - \mathbf{M}\mathbf{s}^{(k)}). \tag{4.144}$$

In terms of eigenvector expansion coefficients, the iteration sequence becomes

$$\sigma_i^{(k+1)} = \sigma_i^{(k)} + \lambda_i(\tau_i - \lambda_i\sigma_i^{(k)}). \tag{4.145}$$

To solve this equation, note that it can be rewritten as

$$\sigma_i^{(k+1)} = \lambda_i\tau_i + (1-\lambda_i^2)\sigma_i^{(k)} = \lambda_i\tau_i + (1-\lambda_i^2)[\lambda_i\tau_i + (1-\lambda_i^2)\sigma_i^{(k-1)}]. \tag{4.146}$$

Rearranging the resulting series, we find

$$\sigma_i^{(k+1)} = [1 + (1-\lambda^2) + (1-\lambda^2)^2 + \ldots + (1-\lambda^2)^k]\lambda_i\tau_i + (1-\lambda^2)^{k+1}\sigma_i^{(0)}. \tag{4.147}$$

The series multiplying τ_i can be summed exactly for any value of $\lambda_i \neq 0$ as

$$[1 + (1-\lambda_i^2) + (1-\lambda_i^2)^2 + \ldots + (1-\lambda_i^2)^k] = \frac{1-(1-\lambda_i^2)^{k+1}}{1-(1-\lambda_i^2)} \tag{4.148}$$

from which it follows that

$$\sigma_i^{(k+1)} = \left[\frac{1-(1-\lambda_i^2)^{k+1}}{\lambda_i}\right]\tau_i + (1-\lambda_i^2)^{k+1}\sigma_i^{(0)}. \tag{4.149}$$

If $\lambda_i = 0$, (4.145) shows that $\sigma_i^{(k+1)} = \sigma_i^{(0)}$. If we assume that the eigenvalues, λ_i, are all between $-\sqrt{2}$ and $\sqrt{2}$, the iteration sequence converges. The condition $-\sqrt{2} < \lambda_i < \sqrt{2}$ thus implies that $\sigma_i^{(k)} \to \tau_i/\lambda_i$ as $k \to \infty$. That is, the iteration converges to a least-squares model.

We have already seen in Section 4.2 that the stronger condition $-1 \leq \lambda_i \leq 1$ can be guaranteed with an appropriate preconditioning (prescaling) of the matrix $\mathbf{M}$.

Simple iteration is a good method for solving linear tomography problems, and is much simpler to implement than other methods such as conjugate directions or conjugate gradients. This method has significant computational advantages when the dimensions of $\mathbf{M}$ are large. The method is also closely related to SIRT (Simultaneous Iterative Reconstruction Technique) which will be discussed in Section 4.4.5.

PROBLEMS

[3] Also known as Richardson iteration [Varga, 1962].

PROBLEM 4.4.11 *Repeat the analysis of simple iteration for a damped least-squares objective functional. Show that*

$$\mathbf{s}_\mu^{(k+1)} = \mathbf{s}_\mu^{(k)} + \mathbf{M}^T(\mathbf{t} - \mathbf{M}\mathbf{s}^{(0)}) - (\mathbf{M}^T\mathbf{M} + \mu\mathbf{I})(\mathbf{s}_\mu^{(k)} - \mathbf{s}^{(0)}) \tag{4.150}$$

is a valid iteration scheme. Show that

$$\sigma_i^{(k+1)} = \left[\frac{1-(1-\lambda_i^2-\mu)^{k+1}}{\lambda_i^2+\mu}\right](\lambda_i\tau_i + \mu\sigma_i^{(0)}) + (1-\lambda_i^2-\mu)^{k+1}\sigma_i^{(0)}. \tag{4.151}$$

What restrictions must be placed on the λ_is and μ to guarantee convergence of (4.151)? Find the differences between the asymptotic results for (4.151) and those for the undamped least-squares method.

PROBLEM 4.4.12 *Show that, if the maximum eigenvalue is $\lambda_1 = \sqrt{2}$, then convergence of simple iteration is improved by considering $\overline{\mathbf{s}}^{(k+1)} = \frac{1}{2}(\mathbf{s}^{(k+1)} + \mathbf{s}^{(k)})$. [Ivansson, 1983]*

4.4.4 Neural network method

Consider a sequence of models $\mathbf{s}(\eta)$ as a function of a continuous index variable η. We think of η as a measure of the iteration computation time, or as a continuous version of the iteration counter k used in the preceding discussion. The data misfit functional, Ψ, applied to this sequence then is also a function of η. We have

$$\frac{d\Psi}{d\eta} = 2\frac{d\mathbf{s}^T}{d\eta}\nabla_{\mathbf{s}^T}\Psi, \tag{4.152}$$

where

$$\nabla_{\mathbf{s}^T}\Psi = \mathbf{M}^T(\mathbf{M}\mathbf{s} - \mathbf{t}). \tag{4.153}$$

We would like $d\Psi/d\eta < 0$ so that $\mathbf{s}(\eta)$ converges to a model minimizing Ψ as $\eta \to \infty$. It is easy to verify that a negative derivative is achieved by requiring, for some positive scalar $\gamma > 0$,

$$\frac{d\mathbf{s}}{d\eta} = -\gamma\mathbf{M}^T(\mathbf{M}\mathbf{s} - \mathbf{t}). \tag{4.154}$$

Since the differential change in $\mathbf{s}$ is proportional to the local gradient of the objective functional, this choice produces a type of gradient descent method. We thus have a first-order differential equation for $\mathbf{s}(\eta)$. In terms of the expansion coefficients, σ_i, this becomes

$$\frac{d\sigma_i}{d\eta} = \gamma\lambda_i(\tau_i - \lambda_i\sigma_i). \tag{4.155}$$

Using $\sigma_i = 0$ as an initial condition, the solution to (4.155) is given by

$$\sigma_i(\eta) = \lambda_i^{-1}\tau_i\left[1 - e^{-\gamma\lambda_i^2\eta}\right]. \tag{4.156}$$

We see that $\mathbf{s}(\eta)$ does indeed converge to $\mathbf{s}_{\text{LS}} = \sum_{i=1}^{r} \lambda_i^{-1}\tau_i\mathbf{z}_i$, with its exponential convergence rate controlled by γ and the magnitudes of the positive eigenvalues λ_i for $i = 1, \ldots, r$.

This approach may be used with other objective functionals. For example, we may compare the results of this approach directly with those of simple iteration by considering the damped least-squares functional

$$\Psi_\mu = (\mathbf{t} - \mathbf{M}\mathbf{s})^T(\mathbf{t} - \mathbf{M}\mathbf{s}) + \mu(\mathbf{s} - \mathbf{s}_b)^T(\mathbf{s} - \mathbf{s}_b), \tag{4.157}$$

where $\mathbf{s}_b$ is the starting model $\mathbf{s}(0) = \mathbf{s}_b$. Then, the same analysis shows that a reasonable equation of motion for $\mathbf{s}(\eta)$ is

$$\frac{d\mathbf{s}}{d\eta} = -\gamma\left[(\mathbf{M}^T\mathbf{M} + \mu\mathbf{I})\mathbf{s} - \mathbf{M}^T\mathbf{t} - \mu\mathbf{s}_b\right], \tag{4.158}$$

where γ is again some positive scalar. Now the coefficients satisfy

$$\frac{d\sigma_i}{d\eta} = \gamma\left[\lambda_i\tau_i + \mu\sigma_i(0) - (\lambda_i^2 + \mu)\sigma_i(\eta)\right], \tag{4.159}$$

which yields upon integration

$$\sigma_i(\eta) = \frac{\lambda_i\tau_i + \mu\sigma_i(0)}{\lambda_i^2 + \mu}\left[1 - e^{-\gamma\lambda_i^2\eta}\right] + \sigma_i(0)e^{-\gamma\lambda_i^2\eta}. \tag{4.160}$$

In the presence of the damping term, the method does not converge to $\mathbf{s}_{\text{LS}}$. Instead, it converges exponentially to an approximation with the coefficients $\sigma_i(\infty)$ being weighted averages of the initial value $\sigma_i(0)$ and τ_i/λ_i. The coefficients of the eigenvectors in the null space do not change from their initial values.

Further discussion of this approach together with comparisons to other methods may found in Jeffrey and Rosner [1986a,b] and Lu and Berryman [1990].

PROBLEMS

PROBLEM 4.4.13 *Repeat the analysis of the simple iteration and neural network methods assuming the objective functionals are damped and weighted least-squares. Compare the asymptotic results.*

PROBLEM 4.4.14 *Compare the convergence rates of simple iteration and the neural network method.*

4.4.5 ART and SIRT

Probably the two best known methods of solving linear equations for tomographic applications in general geometries are ART and SIRT. ART is the Algebraic Reconstruction Technique [Gordon, Bender, and Herman, 1970; Tanabe, 1971; Herman, Lent, and Rowland, 1973; Natterer, 1986], while SIRT is the Simultaneous Iterative Reconstruction Technique [Gilbert, 1972; Dines and Lytle, 1979; Ivansson, 1983]. ART is closely related to Kaczmarz's iterative projection method of solving linear equations [Kaczmarz, 1937; Gordon, 1974; Guenther et al., 1974]. Our discussion will not distinguish between ART and

Kaczmarz's algorithm, although the term ART is often used to refer to any algebraic reconstruction method including the many variations on Kaczmarz's algorithm [Gordon, 1974]. The discussion of ART presented here is a simplified version of Tanabe's analysis of Kaczmarz's method. The discussion of SIRT follows Ivansson's [1983] analysis of the method developed by Dines and Lytle [1979].

First, we present definitions of two types of projection operators that will be important for the analysis of ART algorithms. Some of the properties of these operators are presented in PROBLEM 4.4.15.

DEFINITION 4.4.1 *A projection operator* $P(\mathbf{a}_i)$ *for a vector* $\mathbf{a}_i$ *is*

$$P(\mathbf{a}_i) \equiv \frac{\mathbf{a}_i\mathbf{a}_i^T}{\mathbf{a}_i^T\mathbf{a}_i}.$$

If a definite set of vectors $\{\mathbf{a}_i\}$ *is under consideration (so no confusion can arise), we shorten the notation to* $P_i = P(\mathbf{a}_i)$.

DEFINITION 4.4.2 *The orthogonal projection operator* $P_\perp(\mathbf{a}_i)$ *for a vector* $\mathbf{a}_i$ *is*

$$P_\perp(\mathbf{a}_i) \equiv \mathbf{I} - P(\mathbf{a}_i).$$

If a definite set of vectors $\{\mathbf{a}_i\}$ *is under consideration, the notation is shortened to* $Q_i = P_\perp(\mathbf{a}_i)$.

Once again, to solve for the slowness model $\mathbf{s}$ given a traveltime vector $\mathbf{t}$ and a ray-path matrix $\mathbf{M}$, we choose to solve $\mathbf{Ms} = \mathbf{t}$ in the least-squares sense (although ART does not require a square matrix) by solving $\mathbf{As} = \mathbf{b}$ where $\mathbf{A} = \mathbf{M}^T\mathbf{M}$ and $\mathbf{b} = \mathbf{M}^T\mathbf{t}$. Now write the matrix $\mathbf{A}$ as

$$\mathbf{A}^T = (\, \mathbf{a}_1 \quad \mathbf{a}_2 \quad \ldots \quad \mathbf{a}_n \,)\,, \tag{4.161}$$

where the $\mathbf{a}_i$s are the n column vectors of the transpose of $\mathbf{A}$. Now suppose that we have an estimate of the slowness vector $\mathbf{s} \simeq \bar{\mathbf{s}}$ and we want to improve the agreement between the data and the estimate of the data by adding an optimal correction in the direction of vector $\mathbf{a}_i$. Then, considering the relation

$$\mathbf{a}_i^T(\bar{\mathbf{s}} + \alpha_i\mathbf{a}_i) = b_i, \tag{4.162}$$

the optimal value of the coefficient α_i is

$$\alpha_i = \frac{b_i - \mathbf{a}_i^T\bar{\mathbf{s}}}{\mathbf{a}_i^T\mathbf{a}_i}. \tag{4.163}$$

Next, we can define an iterative sequence — starting from a guess $\mathbf{s}^{(0)}$ and making use of only one row of $\mathbf{A}$ at a time — given by

$$\mathbf{s}^{(i)} = \mathbf{s}^{(i-1)} + \frac{b_i - \mathbf{a}_i^T\mathbf{s}^{(i-1)}}{\mathbf{a}_i^T\mathbf{a}_i}\mathbf{a}_i. \tag{4.164}$$

Then it is easy to rearrange (4.164) into the form

$$\mathbf{s}^{(i)} = Q_i \mathbf{s}^{(i-1)} + \frac{b_i}{\mathbf{a}_i^T \mathbf{a}_i} \mathbf{a}_i \qquad \text{for} \qquad 1 \le i \le n. \tag{4.165}$$

The significance of (4.165) is that any component of $\mathbf{s}^{(i-1)}$ in the direction $\mathbf{a}_i$ is removed by the orthogonal projection operator Q_i and then the optimal component proportional to $b_i/||\mathbf{a}_i||$ in the direction of the unit vector $\mathbf{a}_i/||\mathbf{a}_i||$ is added in its place.

A single iteration of ART is completed when we have cycled once through all the column vectors of $\mathbf{A}^T$. Then, the slowness $\mathbf{s}^{(n)}$ is the result and this becomes the starting point for the next iteration. In fact, (4.165) may be used for further iterations with only the minor modification that the i subscripts on everything but the slowness estimate should be replaced by $i' = i \bmod n$, except that $i' = n$ if $i \bmod n = 0$.[4] Thus, we have

$$\mathbf{s}^{(i)} = Q_{i'} \mathbf{s}^{(i-1)} + \frac{b_{i'}}{\mathbf{a}_{i'}^T \mathbf{a}_{i'}} \mathbf{a}_{i'} \qquad \text{for} \qquad 1 \le i' \le n, \quad \text{and} \quad 1 \le i, \tag{4.166}$$

and the iteration number is $k = [i/n]$, where the bracket stands for the greatest integer of the argument.

Now it is straightforward to show that

$$\begin{aligned} \mathbf{s}^{(n)} = {} & Q_n Q_{n-1} \cdots Q_1 \mathbf{s}^{(0)} + \frac{b_1}{\mathbf{a}_1^T \mathbf{a}_1} Q_n Q_{n-1} \cdots Q_2 \mathbf{a}_1 \\ & + \frac{b_2}{\mathbf{a}_2^T \mathbf{a}_2} Q_n Q_{n-1} \cdots Q_3 \mathbf{a}_2 + \ldots + \frac{b_n}{\mathbf{a}_n^T \mathbf{a}_n} \mathbf{a}_n, \end{aligned} \tag{4.167}$$

which can be written more compactly as

$$\mathbf{s}^{(n)} = \mathbf{Q}\mathbf{s}^{(0)} + \mathbf{R}\mathbf{b} \tag{4.168}$$

by introducing the matrices

$$\mathbf{Q} = Q_n Q_{n-1} \cdots Q_1 \tag{4.169}$$

and

$$\mathbf{R} = \begin{pmatrix} \frac{Q_n Q_{n-1} \cdots Q_2 \mathbf{a}_1}{\mathbf{a}_1^T \mathbf{a}_1} & \frac{Q_n Q_{n-1} \cdots Q_3 \mathbf{a}_2}{\mathbf{a}_2^T \mathbf{a}_2} & \cdots & \frac{\mathbf{a}_n}{\mathbf{a}_n^T \mathbf{a}_n} \end{pmatrix}. \tag{4.170}$$

It follows easily from (4.168) that

$$\mathbf{s}^{(kn)} = \sum_{p=0}^{k-1} \mathbf{Q}^p \mathbf{R}\mathbf{b} + \mathbf{Q}^k \mathbf{s}^{(0)}, \tag{4.171}$$

[4] The choice of range $1 \le i \le n$ for subscripts made here is typical of Fortran programming conventions. An inversion code written in C would more naturally use the range $0 \le i \le n-1$ and thereby avoid the need for the exception when $i \bmod n = 0$.

where k is the iteration number. Thus, as $k \to \infty$, the iterates $\mathbf{s}^{(kn)} \to \mathbf{s}$, the slowness solving our linear equation if

$$\lim_{k\to\infty} \sum_{p=1}^{k-1} \mathbf{Q}^p\mathbf{R} = \mathbf{A}^\dagger \tag{4.172}$$

and

$$\lim_{k\to\infty} \mathbf{Q}^k\mathbf{s}^{(0)} = \mathbf{n} \quad \text{or} \quad 0, \tag{4.173}$$

where $\mathbf{n}$ is any vector from the right null space of $\mathbf{A}$. However, it has been found that ART may not converge in practice if the data are inconsistent and/or if $\mathbf{A}$ is singular or nearly so [Gilbert, 1972; Gordon, 1974; Dines and Lytle, 1979], which brings us to SIRT.

One common version of the SIRT algorithm [Dines and Lytle, 1979; Ivansson, 1983] may be written in component form as

$$s_j^{(k+1)} = s_j^{(k)} + N_{jj}^{-1} \sum_{i=1}^{m} \frac{(t_i - \sum_{p=1}^{n} l_{ip}s_p^{(k)})l_{ij}}{\sum_{q=1}^{m} l_{iq}^2}, \tag{4.174}$$

where N_{jj} is the number of rays passing through cell j (sometimes known as the hit parameter). In vector notation, this becomes

$$\mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \mathbf{N}^{-1}\mathbf{M}^T\mathbf{D}^{-1}(\mathbf{t} - \mathbf{M}\mathbf{s}^{(k)}), \tag{4.175}$$

where $\mathbf{N}$ is the diagonal matrix whose components are N_{jj} and $\mathbf{D}$ is the diagonal matrix whose components are

$$D_{ii} = (\mathbf{M}\mathbf{M}^T)_{ii} = \sum_{j=1}^{n} l_{ij}^2. \tag{4.176}$$

Convergence of this algorithm has been proven by Ivansson [1983]. We present a similar proof.

The analysis is very similar to that presented in Section 4.2 on scaling methods. Define

$$\mathbf{M}'' = \mathbf{D}^{-\frac{1}{2}}\mathbf{M}\mathbf{N}^{-\frac{1}{2}}, \tag{4.177}$$

$$\mathbf{y}'' = \mathbf{D}^{\frac{1}{2}}\mathbf{y}, \tag{4.178}$$

$$\mathbf{z}'' = \mathbf{N}^{\frac{1}{2}}\mathbf{z}. \tag{4.179}$$

Then, consider the eigenvalue problem

$$\mathbf{M}''\mathbf{z}'' = \lambda\mathbf{y}'', \tag{4.180}$$

$$\mathbf{M}''^T\mathbf{y}'' = \lambda\mathbf{z}'', \tag{4.181}$$

which is equivalent to

$$\mathbf{M}\mathbf{z} = \lambda \mathbf{D}\mathbf{y}, \tag{4.182}$$

$$\mathbf{M}^T\mathbf{y} = \lambda \mathbf{N}\mathbf{z}. \tag{4.183}$$

We see that, if $\mathbf{y}_i$ and $\mathbf{z}_j$ are eigenvectors for eigenvalues λ_i and λ_j,

$$\mathbf{y}_i^T\mathbf{M}\mathbf{z}_j = \lambda_i \mathbf{y}_i^T\mathbf{D}\mathbf{y}_j = \lambda_i \mathbf{z}_i^T\mathbf{N}\mathbf{z}_j, \tag{4.184}$$

so we are free to normalize the eigenvectors so that

$$(\mathbf{y}_i'')^T\mathbf{y}_j'' = (\mathbf{z}_i'')^T\mathbf{z}_j'' = \delta_{ij}.p \tag{4.185}$$

PROPOSITION 4.4.1 *The eigenvalues of* $\mathbf{M}''$ *lie in the interval [-1,1].*

Proof: The eigenvalue problem (4.180)–(4.181) can be written in components as

$$\sum_{j=1}^{n} \frac{l_{ij}}{(D_{ii}N_{jj})^{\frac{1}{2}}} z_j'' = \lambda y_i'', \tag{4.186}$$

$$\sum_{i=1}^{m} \frac{l_{ij}}{(N_{jj}D_{ii})^{\frac{1}{2}}} y_i'' = \lambda z_j''. \tag{4.187}$$

Now define the sign function as

$$\mathrm{sgn}(l_{ij}) = \begin{cases} +1 & \text{if} \quad l_{ij} > 0, \\ 0 & \text{if} \quad l_{ij} = 0, \\ -1 & \text{if} \quad l_{ij} < 0, \end{cases} \tag{4.188}$$

and note that

$$N_{jj} = \sum_{i=1}^{m} \mathrm{sgn}(l_{ij}), \tag{4.189}$$

since the path lengths are never negative. Considering (4.186) and using Cauchy's inequality for sums, we have

$$\lambda^2 (y_i'')^2 \leq \sum_{p=1}^{n} \frac{l_{ip}^2}{D_{ii}} \sum_{j=1}^{n} \frac{\mathrm{sgn}(l_{ij})(z_j'')^2}{N_{jj}}. \tag{4.190}$$

Using the definition (4.176) of D_{ii} and summing (4.190) over i, we find

$$\lambda^2 \sum_{i=1}^{m} (y_i'')^2 \leq \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{\mathrm{sgn}(l_{ij})(z_j'')^2}{N_{jj}} = \sum_{j=1}^{n} (z_j'')^2. \tag{4.191}$$

Then, the normalization condition (4.185) shows that (4.191) reduces to

$$\lambda^2 \leq 1. \tag{4.192}$$

∎

Thus, an analysis completely analogous to that given previously for simple iteration (see Section 4.4.3) shows that an iteration scheme of the form

$$\mathbf{N}^{\frac{1}{2}}\mathbf{s}^{(k+1)} = \mathbf{N}^{\frac{1}{2}}\mathbf{s}^{(k)} + \mathbf{M}''^{T}\left[\mathbf{D}^{-\frac{1}{2}}\mathbf{t} - \mathbf{M}''\mathbf{N}^{\frac{1}{2}}\mathbf{s}^{(k)}\right], \tag{4.193}$$

is guaranteed to converge to a solution of $\mathbf{M}^T\mathbf{M}\mathbf{s} = \mathbf{M}^T\mathbf{t}$.

Problems

Problem 4.4.15 *Show that*

1. $P(\mathbf{a})\mathbf{a} = \mathbf{a}$; $\mathbf{a}^T P(\mathbf{a}) = \mathbf{a}^T$;
2. $P_\perp(\mathbf{a})\mathbf{a} = 0$; $\mathbf{a}^T P_\perp(\mathbf{a}) = 0$;
3. $P^2(\mathbf{a}) = P(\mathbf{a})$; $P^T(\mathbf{a}) = P(\mathbf{a})$;
4. $P(-\mathbf{a}) = P(\mathbf{a})$; $P(\gamma\mathbf{a}) = P(\mathbf{a})$ *for any scalar* γ;
5. $P(\mathbf{a})P_\perp(\mathbf{a}) = 0 = P_\perp(\mathbf{a})P(\mathbf{a})$;
6. *if* $\mathbf{a}_2^T\mathbf{a}_1 = 0$, *then* $P_1P_2 = P_2P_1$;
7. *if* $\mathbf{a}_2^T\mathbf{a}_1 \neq 0$, *then* $P_1P_2 \neq P_2P_1$ *unless* $\mathbf{a}_2 = \gamma\mathbf{a}_1$ *for some scalar* γ;
8. $P^\dagger(\mathbf{a}) = P(\mathbf{a})$.

Problem 4.4.16 *A beam of light will not pass through a pair of polarizing filters if their axes are crossed at right angles. However, if a third filter is inserted between the first two with its polarizing axis at* $45°$, *then some of the light can get through. Use projection operators for the vectors* $\mathbf{a}_1 = \hat{x}$, $\mathbf{a}_2 = (\hat{x}+\hat{y})/\sqrt{2}$, *and* $\mathbf{a}_3 = \hat{y}$ *to explain this physical effect. Design a product of projection operators that will project a vector* $\hat{x}$ *onto the direction of its reflection* $-\hat{x}$. *What is the smallest number of projection operators that can be used to produce a reflection? [Feynman, Leighton, and Sands, 1963]*

Problem 4.4.17 *Verify (4.167).*

Problem 4.4.18 *Rewrite (4.18) and (4.19) in terms of projection operators.*

Problem 4.4.19 *Rewrite (4.128) in terms of projection operators.*

Problem 4.4.20 *Rewrite the Gram-Schmit orthogonalization procedure (4.141) in terms of projection operators.*

PROBLEM 4.4.21 *Reconsider the conjugate gradients approach (4.136)–(4.138). Show that the iteration scheme can be written as*

$$\mathbf{s}^{(k+1)} = [\mathbf{I} - Q_A^{(k)}]\mathbf{b} + Q_A^{(k)}\mathbf{s}^{(k)}$$

and

$$\mathbf{p}^{(k+1)} = Q_A^{(k)}\left[\mathbf{b} - \mathbf{A}\mathbf{s}^{(k+1)}\right],$$

where

$$Q_A^{(k)} = \mathbf{I} - \frac{(\mathbf{p}^{(k)})^T\mathbf{p}^{(k)}\mathbf{A}}{(\mathbf{p}^{(k)})^T\mathbf{A}\mathbf{p}^{(k)}}.$$

Show that

$$\mathbf{s}^{(k+1)} = [\mathbf{I} - Q_A^{(k)}Q_A^{(k-1)}\cdots Q_A^{(1)}]\mathbf{b} + Q_A^{(k)}Q_A^{(k-1)}\cdots Q_A^{(1)}\mathbf{s}^{(1)}.$$

Compare this result to (4.140).

PROBLEM 4.4.22 *Suppose that* $\mathbf{A}^T = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4)$. *Show that*

$$\mathbf{RA} = \mathbf{I} - Q_4Q_3Q_2Q_1$$

if

$$\mathbf{R} = \begin{pmatrix} \frac{Q_4Q_3Q_2\mathbf{a}_1}{\mathbf{a}_1^T\mathbf{a}_1} & \frac{Q_4Q_3\mathbf{a}_2}{\mathbf{a}_2^T\mathbf{a}_2} & \frac{Q_4\mathbf{a}_3}{\mathbf{a}_3^T\mathbf{a}_3} & \frac{\mathbf{a}_4}{\mathbf{a}_4^T\mathbf{a}_4} \end{pmatrix}.$$

Show that

$$\mathbf{A}^\dagger = \sum_{k=0}^{\infty}[Q_4Q_3Q_2Q_1]^k\mathbf{R}.$$

PROBLEM 4.4.23 *What changes must be made in ART and SIRT in order to solve a weighted least-squares problem? Or, a damped and weighted least-squares problem?*

PROBLEM 4.4.24 *Show that ART may be applied directly to the system* $\mathbf{Ms} = \mathbf{t}$ *when the* $m \times n$ *ray-path matrix* $\mathbf{M}$ *is not square by deriving the formula*

$$s_j^{(i)} = s_j^{(i-1)} + \frac{t_{i'} - \sum_{p=1}^n l_{i'p}s_p^{(i-1)}}{\sum_{q=1}^n l_{i'q}^2}l_{i'j} \tag{4.194}$$

for the iteration sequence in component form.

PROBLEM 4.4.25 *Since* $\mathbf{A} = \mathbf{M}^T\mathbf{M}$ *in the traveltime inversion problem, the components of* $\mathbf{A}$ *are given by*

$$A_{jj'} = \sum_{p=1}^{m} l_{pj}l_{pj'}$$

and the components of the column vector $\mathbf{a}_i$ *are given by*

$$\mathbf{a}_i = \begin{pmatrix} \sum_{p=1}^m l_{pi}l_{p1} \\ \sum_{p=1}^m l_{pi}l_{p2} \\ \vdots \\ \sum_{p=1}^m l_{pi}l_{pn} \end{pmatrix}.$$

Use these facts to write an expression for the ART iteration scheme (4.166) in component form.

PROBLEM 4.4.26 *Do the two forms of ART in* Problems 4.4.24 *and* 4.4.25 *converge to the same slowness value? [Hint: Consider the generalized inverses of* $\mathbf{M}$ *and* $\mathbf{A} = \mathbf{M}^T\mathbf{M}$*.]*

PROBLEM 4.4.27 *The form of SIRT presented in (4.175) is just one of many possibilities. Show that the following alternatives converge and determine their convergence rates:*

1. $\mathbf{X}^{\frac{1}{2}}\mathbf{s}^{(k+1)} = \mathbf{X}^{\frac{1}{2}}\mathbf{s}^{(k)} + \mathbf{X}^{-\frac{1}{2}}\mathbf{M}^T\mathbf{D}^{-1}(\mathbf{t} - \mathbf{M}\mathbf{s}^{(k)})$ *where* $\mathbf{X} = \mathbf{N}/\gamma$ *for* $0 < \gamma < 2$*;*

2. $\overline{\mathbf{D}}^{\frac{1}{2}}\mathbf{s}^{(k+1)} = \overline{\mathbf{D}}^{\frac{1}{2}}\mathbf{s}^{(k)} + \overline{\mathbf{D}}^{-\frac{1}{2}}\mathbf{M}^T\overline{\mathbf{N}}^{-1}(\mathbf{t} - \mathbf{M}\mathbf{s}^{(k)})$ *where* $\overline{\mathbf{D}} = diag(\mathbf{M}^T\mathbf{M})$ *and* $\overline{\mathbf{N}}_{ii} = \sum_{j=1}^n sgn(l_{ij})$ *is the number of cells traversed by the* i*th ray;*

3. $\overline{\mathbf{D}}^{\frac{1}{2}}\mathbf{s}^{(k+1)} = \overline{\mathbf{D}}^{\frac{1}{2}}\mathbf{s}^{(k)} + n^{-1}\overline{\mathbf{D}}^{-\frac{1}{2}}\mathbf{M}^T(\mathbf{t} - \mathbf{M}\mathbf{s}^{(k)})$.

PROBLEM 4.4.28 *Jacobi's method for solving* $\mathbf{A}\mathbf{s} = \mathbf{b}$ *for a square matrix* $\mathbf{A}$ *is*

$$\mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \overline{\mathbf{D}}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{s}^{(k)}),$$

where the diagonal matrix

$$\overline{\mathbf{D}} = diag\mathbf{A}.$$

If $\mathbf{s}^{(-1)} = 0$*, compare* $\mathbf{s}^{(0)}$ *to the backprojection estimate (1.16). Show that, in component form, the iterates of Jacobi's method are*

$$s_j^{(k+1)} = \frac{1}{A_{jj}}\left(b_j - \sum_{p=1}^{j-1} A_{jp}s_p^{(k)} - \sum_{p=j+1}^{n} A_{jp}s_p^{(k)}\right) \tag{4.195}$$

for $j = 1, \ldots, n$*. Does this method converge for the traveltime inversion problem? If not, can it be modified to guarantee convergence? Under what circumstances are SIRT and Jacobi's method equivalent?*

PROBLEM 4.4.29 *The Gauss-Seidel method for solving* $\mathbf{A}\mathbf{s} = \mathbf{b}$ *decomposes the square matrix* $\mathbf{A}$ *into*

$$\mathbf{A} = \overline{\mathbf{D}} + \overline{\mathbf{L}} + \overline{\mathbf{U}},$$

where $\overline{\mathbf{D}}$ *is the diagonal of* $\mathbf{A}$ *while* $\overline{\mathbf{L}}$ *and* $\overline{\mathbf{U}}$ *are the lower and upper triangular pieces of* $\mathbf{A}$*. Then, the iteration scheme is given by*

$$(\overline{\mathbf{D}} + \overline{\mathbf{L}})\mathbf{s}^{(k+1)} = \mathbf{b} - \overline{\mathbf{U}}\mathbf{s}^{(k)}.$$

Show that, in component form, the iterates of the Gauss-Seidel method are

$$s_j^{(k+1)} = \frac{1}{A_{jj}} \left(b_j - \sum_{p=1}^{j-1} A_{jp} s_p^{(k+1)} - \sum_{p=j+1}^{n} A_{jp} s_p^{(k)} \right) \tag{4.196}$$

for $j = 1, \ldots, n$*. Compare and contrast (4.195) and (4.196). Does Gauss-Seidel converge for the traveltime inversion problem? If not, can it be modified to guarantee convergence?*

PROBLEM 4.4.30 *To apply the Gauss-Seidel method to* $\mathbf{Ms} = \mathbf{t}$ *when* $\mathbf{M}$ *is not square, define the new* m*-vector* $\mathbf{q}$ *such that* $\mathbf{s} = \mathbf{M}^T\mathbf{q}$*. Then, the Gauss-Seidel approach may be applied directly to*

$$\mathbf{M}\mathbf{M}^T\mathbf{q} = \mathbf{t}.$$

The resulting iteration scheme is

$$q_i^{(k+1)} = \left(t_i - \sum_{p=1}^{i-1} (\mathbf{M}\mathbf{M}^T)_{ip} q_p^{(k+1)} - \sum_{p=i+1}^{m} (\mathbf{M}\mathbf{M}^T)_{ip} q_p^{(k)} \right) / (\mathbf{M}\mathbf{M}^T)_{ii}. \tag{4.197}$$

Write (4.197) in terms of components and compare the result to (4.194). Is ART equivalent to the Gauss-Seidel method?

Chapter 5

Fast Ray Tracing Methods

The most expensive step in any traveltime inversion or tomography algorithm is the forward modeling step associated with ray tracing through the current best estimate of the wave speed model. It is therefore essential to make a good choice of ray tracing algorithm for the particular application under consideration. Prior to choosing a ray tracing method, a method of representing the model must be chosen. Three typical choices are: cells or blocks of constant slowness, a rectangular grid with slowness values assigned to the grid points and linearly interpolated values between grid points, or a sum over a set of basis functions whose coefficients then determine the model. The ray tracing method should be designed to produce optimum results for the particular model representation chosen.

We will consider three approaches to ray tracing:

1. Shooting methods.

2. Bending methods.

3. Full wave equation methods.

These three methods are based respectively on Snell's law [Born and Wolf, 1980], Fermat's principle [Fermat, 1891], and Huygen's principle [Huygens, 1690]. We will find that shooting methods and wave equation methods should generally be used with smooth representations of the model such as linearly interpolated grids or spline function approximations, while bending methods are preferred for constant cell representations.

We will study each of these approaches in some detail in this section. But first we address a question commonly asked about the necessity of using bent rays in traveltime tomography.

5.1 Why Not Straight Rays?

Straight rays are used in x-ray tomography and the results obtained are very good, so why not use straight rays in seismic inversion and tomography? For x-rays traveling through the body, the index of refraction is essentially constant, so the ray paths are in fact nearly straight. Furthermore, the reconstruction in x-ray tomography is performed on the attenuation coefficient, not the wave speed, so the situation is not really comparable to that of

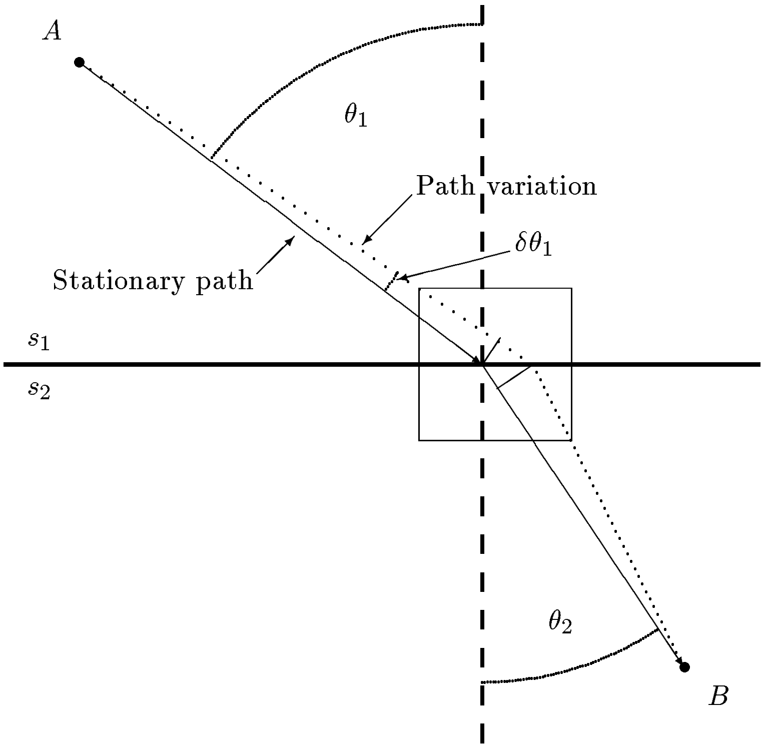


Figure 5.1: Snell's law is a consequence of the stationarity of the traveltime functional.

seismic tomography. Reconstructions in seismic inversion and tomography are most often performed on the wave speed or wave slowness. Since the earth is not homogeneous, the speed of sound varies significantly and the effective index of refraction is far from being constant. Thus, the rays in a seismic transmission experiment really do bend significantly and this fact should be taken into account in the reconstruction.

Suppose that we use straight rays in a tomographic reconstruction when in truth the rays whose traveltimes have been measured were actually bent according to Fermat's principle or Snell's law. In a region where the wave speed is quite low, the true rays will tend to go around the region, but the straight rays go through anyway. So the backprojection along a straight ray will naturally focus the effects of a slow region into a *smaller region* than it should. Similarly, in a region where the wave speed is quite high, the true rays will tend to accumulate in the fast region, whereas the straight rays are free to ignore this focusing effect. Thus, the backprojection along a straight ray will tend to defocus the effects of a fast region into a *larger region* than it should. If we could train our eyes to look for these effects in straight ray reconstructions, then it might not be essential to use bent rays. But until then, it is important to recognize that using straight rays has important effects on the resolution of the reconstruction. Regions of high wave speed will appear larger than true, so such regions are poorly resolved. Regions of low wave speed will appear smaller than

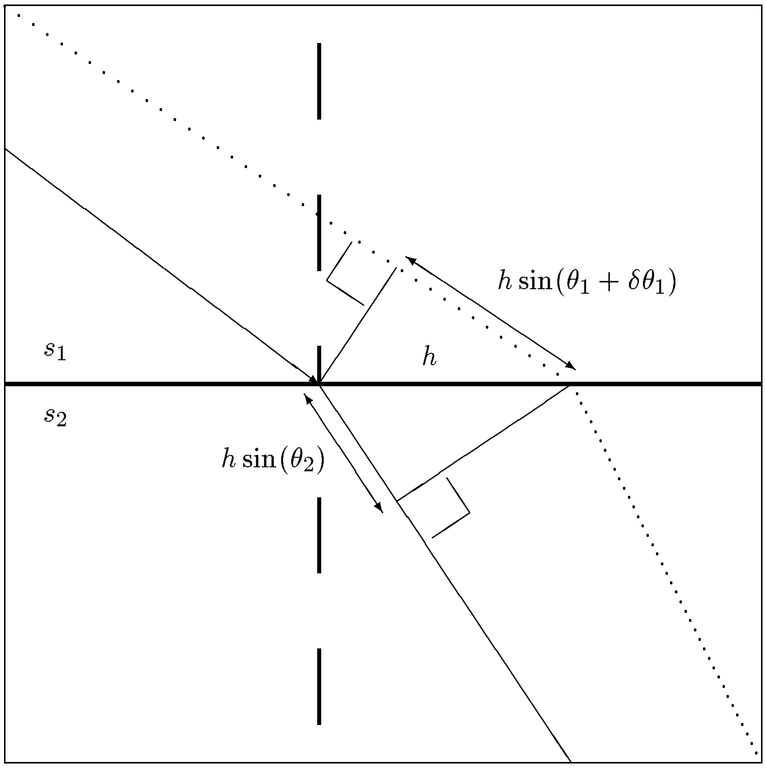


Figure 5.2: Detail of the stationarity calculation (see preceding Figure).

true, so such regions are poorly defined.

Having said all this, nevertheless there are circumstances where I would recommend using straight rays in the reconstruction. First, if the region to be imaged contains very high contrasts so that some of the assumptions normally made to speed up the ray tracing codes are expected to be violated (*e.g.*, rays double back on themselves), then stable reconstructions with bent rays may be impossible while a straight ray reconstruction can still give some useful information. Second, if the desired result is just a low resolution image showing whether or not an anomaly is present, then straight rays are entirely appropriate. Third, if a reconstruction for anisotropic wave speed is being attempted, then straight rays are recommended too, since the nonuniqueness expected in the reconstruction when bent rays are coupled with anisotropy in the model appears so overwhelming that I think little can be done to overcome the problem at the present time.[1]

Straight rays are always computed quickly since they depend only on the source and receiver locations. So if resolution is not an issue but speed of computation is, then of course straight rays can and probably should be used. However, using straight rays is limiting the reconstruction to be merely *linear inversion or tomography*, but — since our subject is *nonlinear inversion and tomography* — we will not consider straight rays further.

[1]See Jech and Pšenčík [1989; 1991].

5.2 Variational Derivation of Snell's law

For piece-wise constant slowness models, we consider Snell's law.

A medium with two regions of constant slowness s_1, s_2 is separated by a plane boundary. The ray path connecting two points, A and B, located on either side of the boundary is stationary, *i.e.*, small deviations from this path make only second order corrections to the traveltime. Referring to Figures 5.1 and 5.2, we let the solid line denote the ray path having stationary traveltime and let the dotted line be a perturbed ray path. Each path is assumed to comprise straight lines within each medium which then bend into a new direction upon crossing the boundary. We let θ_1 and θ_2 denote the angles of the stationary path from the normal to the boundary in the two regions, respectively. A simple geometrical argument can be used to infer the difference in length between the two paths to first order in h, the distance between the points where the paths intersect the boundary. We find that the segment of the perturbed path in region 1 is $h\sin(\theta_1+\delta\theta)$ units longer than the stationary path, while in region 2 the perturbed path is $h\sin\theta_2$ units shorter. Therefore, the traveltime along the perturbed ray differs from that along the stationary ray by Δt, given by

$$\Delta t = s_1 h \sin\theta_1 - s_2 h \sin\theta_2, \tag{5.1}$$

neglecting the second order effects due to finite $\delta\theta$ and due to the slight differences in the remainders of these two paths. Since the traveltime is stationary, we set $\Delta t = 0$ and find that

$$\boxed{s_1 \sin\theta_1 \quad = \quad s_2 \sin\theta_2. \qquad \text{(Snell's law)}} \tag{5.2}$$

5.3 Ray Equations and Shooting Methods

Let the ray path P between two points A and B be represented by a trajectory $\mathbf{x}(u)$, where u is a scalar parameter that increases monotonically along the ray. We can then write the traveltime along the path as

$$t = \int_P s(\mathbf{x}(u))\, dl(u) \tag{5.3}$$

$$= \int_{u(A)}^{u(B)} f(\mathbf{x}, \dot{\mathbf{x}})\, du, \tag{5.4}$$

where $\dot{\mathbf{x}} = d\mathbf{x}/du$ and

$$f(\mathbf{x}, \dot{\mathbf{x}}) = s(\mathbf{x})|\dot{\mathbf{x}}|. \tag{5.5}$$

Fermat's principle implies that the stationary variation [Whitham, 1974]

$$\delta t = \int_{u(A)}^{u(B)} [\nabla_{\mathbf{x}} f \cdot \delta\mathbf{x} + \nabla_{\dot{\mathbf{x}}} f \cdot \delta\dot{\mathbf{x}}]\, du = 0. \tag{5.6}$$

Integrating by parts

$$\delta t = \int_{u(A)}^{u(B)} \left[\nabla_{\mathbf{x}} f - \frac{d}{du} \nabla_{\dot{\mathbf{x}}} f \right] \cdot \delta \mathbf{x} du = 0. \tag{5.7}$$

Since this must be true for all $\delta \mathbf{x}$, we can infer

$$\nabla_{\mathbf{x}} f - \frac{d}{du} \nabla_{\dot{\mathbf{x}}} f = 0. \tag{5.8}$$

Now observe that

$$\nabla_{\mathbf{x}} f = |\dot{\mathbf{x}}| \nabla s, \tag{5.9}$$

$$\nabla_{\dot{\mathbf{x}}} f = s(\mathbf{x}) \frac{\dot{\mathbf{x}}}{|\dot{\mathbf{x}}|}. \tag{5.10}$$

Further, we have $dl = |\dot{\mathbf{x}}| du$, so stationarity of t implies

$$\boxed{\nabla s = \frac{d}{dl} \left(s \frac{d}{dl} \mathbf{x} \right).} \tag{5.11}$$

This is the *ray equation.*

In a 2-D application, the ray equation may be rewritten in terms of the angle θ of the ray from the x direction. First, note that

$$\frac{d}{dl} \mathbf{x} = \hat{\rho} = \cos\theta \hat{\mathbf{x}} + \sin\theta \hat{\mathbf{y}} \tag{5.12}$$

and

$$\frac{d}{dl} \hat{\rho} = \hat{\theta} \frac{d\theta}{dl} = (-\sin\theta \hat{\mathbf{x}} + \cos\theta \hat{\mathbf{y}}) \frac{d\theta}{dl}, \tag{5.13}$$

so that (5.11) may be rewritten as

$$\nabla s = \frac{ds}{dl} \hat{\rho} + s \hat{\theta} \frac{d\theta}{dl}, \tag{5.14}$$

which implies

$$\hat{\theta} \cdot \nabla s = s \frac{d\theta}{dl}. \tag{5.15}$$

Finally, we obtain

$$\boxed{\frac{d\theta}{dl} = \frac{1}{s} \left(\frac{\partial s}{\partial y} \cos\theta - \frac{\partial s}{\partial x} \sin\theta \right),} \tag{5.16}$$

giving an explicit differential equation for the ray angle θ along a 2-D path.

The ray equations form the basis for shooting methods of ray tracing. Starting at any source point, we initially choose a set of possible angles. An optimum initial span of angles can be determined if the range of wave-speed variation is known approximately. Then, we use the ray equations to trace the rays at each of these angles through the medium to the vicinity of the receiver of interest. Normally none of the initial angles turns out to be the correct one (*i.e.*, the one that produces a ray that hits the receiver), but often the receiver is bracketed by two of these rays. Then, by interpolation, we can find as accurate an approximation as we like: *i.e.*, choose a new set of angles between the pair that brackets the receiver, trace the rays for these angles, keep the two closest that bracket the receiver, and continue this process until some closeness objective has been achieved.

Shooting methods are very accurate, but also relatively expensive. We may have to shoot many rays to achieve the desired degree of accuracy. Furthermore, there can be pathological cases arising in inversion and tomography where it is difficult or impossible to trace a ray from the the source to receiver through the current best estimate of the slowness model. Such problems are most likely to occur for models containing regions with high contrasts. Then, there can exist shadow zones behind slow regions, where ray amplitude is small for first arrivals. Such problems can also arise due to poor choice of model parametrization. Shooting methods should normally be used with smooth models based on bilinear interpolation between grid points, or spline function approximations. If the desired model uses cells of constant slowness, shooting methods are not recommended.

Problems

Problem 5.3.1 *Verify (5.11).*

Problem 5.3.2 *Derive Snell's law (5.2) for the change in ray angle at a plane interface from the ray equation (5.11).*

Problem 5.3.3 *Can the ray equation be derived from Snell's law?*

Problem 5.3.4 *Consider a horizontally stratified medium with a sequence of layers having uniform slownesses s_1, s_2, s_3, ... Use Snell's law to show that a ray having angle θ_1 to the vertical in the first layer will have an associated invariant (called the ray parameter)*

$$\boxed{p = s_i \sin\theta_i} \tag{5.17}$$

in every layer i whose slowness satisfies $s_i > s_1 \sin\theta_1$. If the ray encounters a layer (say the nth layer) whose slowness satisfies $s_n \leq s_1 \sin\theta_1$, then what happens? Show that the constancy of the ray parameter is a direct consequence of (5.16).

5.4 The Eikonal Equation

Consider the wave equation for a field $\psi(\mathbf{x}, t)$ in a medium with slowness $s(\mathbf{x})$:

$$\nabla^2\psi = s^2(\mathbf{x})\frac{\partial^2\psi}{\partial t^2}. \tag{5.18}$$

Let us assume

$$\psi(\mathbf{x}, t) = e^{i\omega[\phi(\mathbf{x})-t]} = e^{-\omega\Im\phi(\mathbf{x})}e^{i\omega[\Re\phi(\mathbf{x})-t]}, \tag{5.19}$$

where $\phi(\mathbf{x}) = \Re\phi(\mathbf{x}) + i\Im\phi(\mathbf{x})$ is a complex phase. The imaginary part $\Im\phi$ determines the amplitude of ψ. Substituting into the wave equation, we get

$$\left[i\omega\nabla^2\phi - \omega^2\nabla\phi\cdot\nabla\phi + \omega^2 s^2(\mathbf{x})\right]\psi = 0. \tag{5.20}$$

In the limit $\omega \to \infty$, $\phi \to \Re\phi$, since (5.20) implies that

$$\nabla\Re\phi\cdot\nabla\Re\phi - \nabla\Im\phi\cdot\nabla\Im\phi = s^2(\mathbf{x}) \tag{5.21}$$

and

$$\nabla\Re\phi\cdot\nabla\Im\phi = 0, \tag{5.22}$$

and the wave equation reduces to the *eikonal*[2] *equation*

$$\boxed{|\nabla\phi| \quad = \quad s(\mathbf{x}).} \tag{5.23}$$

5.5 Vidale's Method

The method of Vidale (1988) uses a finite difference scheme to compute the traveltimes of waves in an arbitrary medium. The slowness of the medium is represented on the nodes of a rectilinear grid with bilinear (for 2-D media) interpolation assumed between nodes. The method approximates the wave field which propagates through a given element as a plane wave. This approximation is valid for the far field. (A different approach is used for the near field, but we will not cover this here.)

5.5.1 Algebraic derivation

Figure 5.3 shows one element of the grid. We number the nodes of the element in a counterclockwise manner, starting with the lower left node. Without loss of generality, we let the plane wave begin at node 0 with traveltime t_0, assumed known. The traveltime to the other nodes—t_1, t_2 and t_3—will then be greater than t_0 by an amount which depends

[2]The term *eikonal* (from the Greek $\varepsilon\iota\kappa\tilde{\omega}\nu$ meaning *icon* or *image*) was introduced by Bruns [1895].

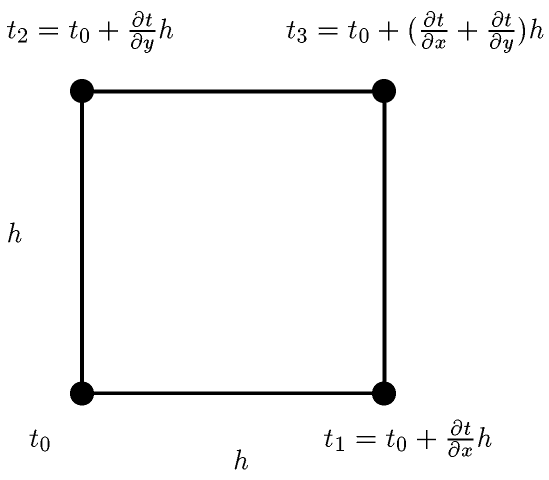


Figure 5.3: Diagram of a grid element used in Vidale's method.

on the direction of propagation and the grid element size h. In general we can write the Taylor series expansion

$$t_1 = t_0 + \frac{\partial t}{\partial x}h, \tag{5.24}$$

$$t_2 = t_0 + \frac{\partial t}{\partial y}h, \tag{5.25}$$

$$t_3 = t_0 + \left(\frac{\partial t}{\partial x} + \frac{\partial t}{\partial y}\right)h, \tag{5.26}$$

valid to first order in h. We can solve these equations for the gradient of t, obtaining

$$2h\frac{\partial t}{\partial x} = t_3 + t_1 - t_2 - t_0, \tag{5.27}$$

$$2h\frac{\partial t}{\partial y} = t_3 + t_2 - t_1 - t_0. \tag{5.28}$$

The eikonal equation implies that $|\nabla t|^2 = s^2(\mathbf{x})$. If we substitute from (5.27) and (5.28) for ∇t and an element average value of s, we get

$$(t_3 + t_1 - t_2 - t_0)^2 + (t_3 + t_2 - t_1 - t_0)^2 = 4\bar{s}^2h^2, \tag{5.29}$$

where

$$\bar{s} = \frac{1}{4}(s_0 + s_1 + s_2 + s_3). \tag{5.30}$$

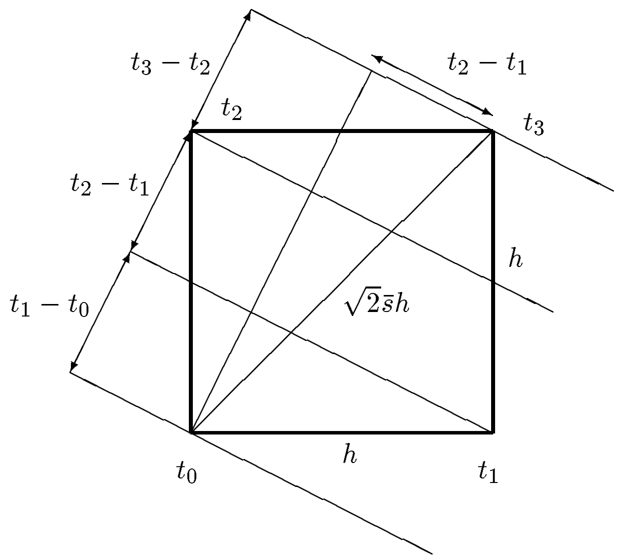


Figure 5.4: Geometry of plane wavefront incident on a grid element.

From (5.29), we find that the cross terms cancel so

$$(t_3 - t_0)^2 + (t_1 - t_2)^2 = 2\bar{s}^2 h^2. \tag{5.31}$$

Solving for t_3, we get *Vidale's formula*:

$$t_3 = t_0 + \sqrt{2\bar{s}^2 h^2 - (t_1 - t_2)^2}. \tag{5.32}$$

We can verify (5.32) for two limiting cases. First, for a wave traveling in the $+x$ direction, we must have $t_0 = t_2$ and, assuming s is constant, $t_1 = t_0 + \bar{s}h$. Substituting these into (5.32) then yields $t_3 = t_0 + \bar{s}h$, which is intuitively the correct answer. Similarly, Vidale's formula implies $t_3 = t_0 + \sqrt{2}\bar{s}h$ for a wave travel at 45 degrees to x, i.e., when $t_1 = t_2$.

5.5.2 Geometric derivation

We can gain more insight into the significance of Vidale's method by deriving the result another way. Now consider Fig. 5.4. We assume that to a first approximation it is satisfactory to treat the slowness in the cell as constant. The constant we choose is the average of the four grid slownesses at the corners of the cell $\bar{s} = 0.25(s_1 + s_2 + s_3 + s_4)$. If the planewave impinges on the cell from the lower left, making angle θ with the x-axis, then the simple geometrical construction in the figure shows that the following identities must hold:

$$t_1 - t_0 = \bar{s}h\cos\theta, \tag{5.33}$$

$$t_3 - t_2 = \bar{s}h\cos\theta, \tag{5.34}$$

$$t_2 - t_0 = \bar{s}h \sin\theta, \tag{5.35}$$

$$t_3 - t_1 = \bar{s}h \sin\theta. \tag{5.36}$$

We see directly from Fig. 5.4 that the right triangle whose hypotenuse is the diagonal of the cell and whose longest side is proportional to the time difference $t_3 - t_0$ has its short side proportional to $t_2 - t_1$. The Pythagorean theorem then tells us that

$$(t_3 - t_0)^2 + (t_2 - t_1)^2 = 2\bar{s}^2h^2, \tag{5.37}$$

in agreement with (5.31) and (5.32). Alternatively, we see that (5.33)–(5.36) show

$$(t_3 - t_0)^2 + (t_2 - t_1)^2 = \bar{s}^2h^2(\cos\theta + \sin\theta)^2 + \bar{s}^2h^2(\sin\theta - \cos\theta)^2 = 2\bar{s}^2h^2. \tag{5.38}$$

From our examination of the geometry for planewaves, we get a bonus. Now we can also find a simple estimate of the angle θ if we know the traveltimes. Clearly,

$$\tan\theta = \frac{t_2 - t_0}{t_1 - t_0} = \frac{t_3 - t_1}{t_3 - t_2} \tag{5.39}$$

follows from (5.33)–(5.36). It also follows from (5.27) and (5.28) that

$$\tan\theta = \frac{\partial t/\partial y}{\partial t/\partial x} = \frac{t_3 + t_2 - t_1 - t_0}{t_3 + t_1 - t_2 - t_0}, \tag{5.40}$$

a result that we may also infer from (5.33)–(5.36). Thus, it is possible to determine the angle θ to first order just by knowing the traveltimes at the corners of the cell. This fact suggests several alternatives for adding ray tracing to Vidale's finite difference traveltime computation, but we will not pursue that subject here.

Finally, note that (5.33)-(5.36) show that

$$t_3 = t_2 + t_1 - t_0. \tag{5.41}$$

Why is this *not* a useful identity for computing the traveltimes?

5.6 Bending Methods

Although in principle they can be, in practice bending methods are generally not as systematic or as accurate as shooting methods. However, they are also much less prone to convergence failures in the presence of pathological models with high relative contrasts (which can result in shadow zones occurring behind very slow regions). Bending methods start with some connected path between the source and receiver (generally a straight line for borehole-to-borehole tomography) and then use some method to reshape or bend that path to reduce and (we hope) minimize the overall traveltime along the path. Bending methods are conceptually based on Fermat's principle of least time; the minimization over paths in (1.2) is being performed now essentially using trial and error. This method is just as legitimate as the others discussed previously and can be just as accurate if the search routine is sufficiently sophisticated. Also, bending methods are the only ones that I recommend using when the model is composed of cells of constant slowness. Other methods such as shooting take the cell boundaries in these models too seriously — trying to satisfy Snell's law exactly at these artificial boundaries while the approximate satisfaction of Snell's law achieved by the bending method using Fermat's principle is more consistent with the approximation to the physics embodied in the model.

5.6.1 The method of Prothero, Taylor, and Eickemeyer

We summarize the bending method of Prothero, Taylor, and Eickemeyer (1988) for the case of 2-D ray paths.

Let (x_S, y_S) and (x_R, y_R) be the given endpoints of the ray. We seek the least time path between the two points, which we can describe with the function $y(x)$ or $x(y)$. [It is assumed that one of these functions is single valued.] Let us use $y(x)$ and, with no loss of generality, we take $x_S = 0$, $x_R = L$.

In ray bending, we begin with an initial ray $y_0(x)$ and seek a perturbation $\delta y(x)$ to the initial ray such that the traveltime along the perturbed ray is reduced. Typically the initial ray is taken to be a straight line:

$$y_0(x) = y_S(1 - \frac{x}{L}) + y_R\frac{x}{L}. \tag{5.42}$$

The perturbed ray is taken to be a harmonic series of the form

$$\delta y(x) = \sum_{k=1}^{K} a_k \sin\frac{k\pi x}{L}. \tag{5.43}$$

The order of the series is usually kept small (e.g., $K = 2$). Note that only sine, and not cosine, terms are used so that the endpoints of the ray remain unperturbed.

In terms of the $y(x)$, the traveltime is given by

$$t = \int_0^L s(x, y(x))\sqrt{1 + (dy/dx)^2}\, dx. \tag{5.44}$$

Prothero, Taylor, and Eickemeyer (1988) use the Nelder-Mead search procedure [Nelder and Mead, 1965; Press, Flannery, Teukolsky, and Vetterling, 1988] to find coefficients a_k such that the traveltime is reduced. The Nelder-Mead approach may be used in any number of dimensions to seek the minimum of a complicated function, especially when local gradients of the function are difficult or expensive to compute. The main idea is to perform a sequence of operations on an n-dimensional simplex, so that the vertices of the simplex converge on the point where the function is minimum. In 2-D, the simplex is a triangle. The complicated function to be minimized in our problem is the traveltime functional. Using this approach, the traveltimes associated with three choices of the ordered pairs (a_1, a_2) are compared—for example, the origin $(0, 0)$ and two other points in the a_1a_2-plane. The point with the largest traveltime is then replaced with a new point found as the mirror reflection of the point about a line passing through the other two points.

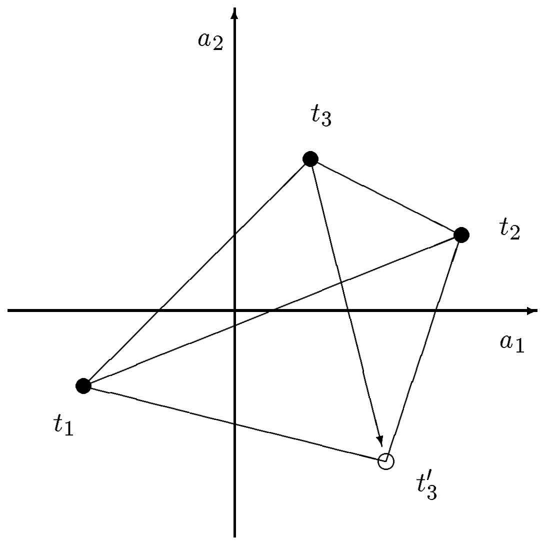


Figure 5.5: Illustration of the Nelder-Mead method.

Figure 5.5 illustrates the method. Starting with three points whose corresponding traveltimes are respectively $t_1 < t_2 < t_3$, the algorithm seeks to replace the point with the largest traveltime by a smaller traveltime t_3'. The figure shows the first attempt in such a process, which is usually reflection of the triangle across the line determined by the other two vertices of the triangle. If the traveltime associated with this point satisfies $t_3' < t_2$, then this point becomes a point of the new triangle. If $t_3' > t_2$, then other moves are made such as checking values between the original vertex and the reflected vertex or expansion/contraction of the triangle. When an improved (smaller traveltime) vertex is found, the vertices are relabelled and the process starts over for the new triangle. If no improvement (or improvement less than some preset threshold) is attained or some fixed number of iterations is exceeded, the process terminates for this ray path.

5.6.2 Getting started

One potential pitfall of this method occurs when attempting to choose a set of vertices for the starting triangle that avoids biasing the final results. Bias in this context means a tendency to choose rays that bow away from the straight path in the same direction. I recommend always choosing the origin $(a_1, a_2) = (0, 0)$ as one of the initial vertices, since this choice corresponds to a straight ray path and is clearly unbiased by definition. The straight path may be a good approximation to the true path whenever the wave speed constrasts in the model are low. Then, how should the other two vertices be chosen?

One rather obvious pairing can be excluded immediately: Suppose that we choose the point $(a_1, a_2) = (\alpha, \beta)$. Then the mirror image of the path across the source/receiver line is given by the point $(a_1, a_2) = (-\alpha, -\beta)$. However, rather than determining a triangle, these three points $(-\alpha, -\beta), (0, 0), (\alpha, \beta)$ form a straight line in the a_1a_2-plane. Thus, although pairing (α, β) with $(-\alpha, -\beta)$ is desirable from the point of view of minimizing bias, this pairing produces an undesirable degenerate version of the triangle needed in the Nelder-Mead algorithm. Therefore, we should exclude this possibility.

In general, we should expect the ray bending effect to be dominated by the coefficient a_1. Thus, although there clearly may be exceptions, we generally expect $|a_1| > |a_2|$ and very often $|a_1| >> |a_2|$. So we try to minimize the bias in the initial choice of vertices by pairing (α, β) with $(-\alpha, \beta)$, where $|\beta|$ is about an order of magnitude smaller than $|\alpha|$. This choice of pairing eliminates the major source of bias in the initial simplex while still producing a usable triangle for the Nelder-Mead algorithm. The precise value to be used for α depends on the expected range of variation (or contrast) in the wave speed in the region being imaged. In fact, the initial choice of α for this approach is closely related to the optimum choice of the maximum initial span of angles needed to start the shooting methods described earlier.

5.7 Comparison

On average, the method of Prothero *et al.* (1988) has been found to be as fast and as accurate as Vidale's method when 100 times fewer cells are used than in Vidale's modelization. So the bending method is considerably more accurate on a coarser grid, but also corresponding slower to compute. Vidale's method is not as accurate as the bending method for regions that are very slow compared to the background, due to limitations it has in or near shadow zones. The bending method is not quite as accurate as Vidale's method for regions of high wave speed relative to background and comparable computing time, apparently due to limitations of the ray parameterization embodied in (5.43). The hybrid approach of using the best (smallest) traveltimes found by either method as the "true" traveltime has been tested and gives better results than either method alone.

Chapter 6

Ghosts in Traveltime Inversion and Tomography

A ghost in seismic traveltime inversion is a model perturbation that does not affect the agreement between the predicted and measured first arrival traveltimes. For example, if

$$\mathbf{M}\mathbf{s} = \mathbf{t}, \quad \mathbf{M}(\mathbf{s}+\mathbf{g}) = \mathbf{t}, \tag{6.1}$$

then subtracting shows that

$$\boxed{\mathbf{M}\mathbf{g} \quad = \quad 0,} \tag{6.2}$$

so $\mathbf{g}$ lies in the null space of $\mathbf{M}$, *i.e.*, in the null space of the traveltime functional. A careful analysis of the ghosts shows that, while some are unavoidable due to the limited view angles used when the data were collected, others are caused by unfortunate choices made when discretizing the model. Thus, some ghosts may be eliminated by making unusual choices for the model parametrization.

It is important to realize from the outset that it may not be either possible or even desirable to eliminate all the ghosts. In fact, the normal solution to the least-squares problem cannot be found if $\mathbf{M}^T\mathbf{M}$ is not invertible. Lack of invertibility is caused by the presence of a right null space for $\mathbf{M}$ and the members of that null space we call *ghosts*. In some cases, simple tricks can be developed to eliminate the ghosts, but not always.

One reference on this topic is Ivansson [1986].

6.1 Feasibility Constraints and Ghosts

Because feasibility constraints depend on the traveltime data while ghosts are independent of the traveltime data (depending instead only on the ray path matrix), feasibility constraints are always orthogonal to all ghost vectors. This fact is illustrated in Figure 6.1. Actually, this statement is somewhat oversimplified in the context of nonlinear inversion algorithms, where we may want to consider many ray-path matrices simultaneously, but

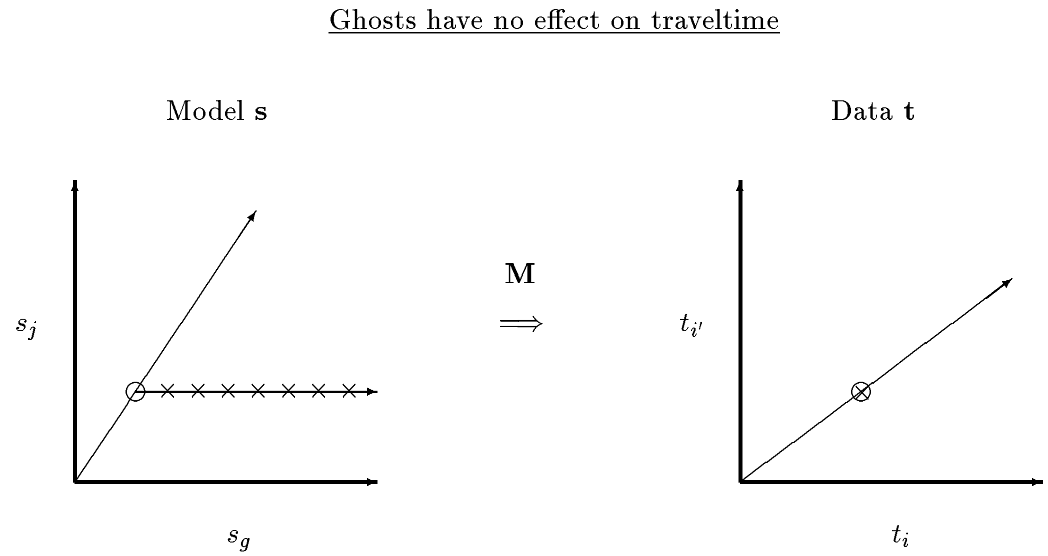


Figure 6.1: Feasibility constraints are orthogonal to all ghost vectors.

we will ignore this complication here. Thus, the remainder of this Chapter will describe various kinds of ghost and other means of dealing with them.

6.2 Types of Ghosts

We now carry out an in-depth analysis of a few common ghosts occurring in seismic inverse problems.

6.2.1 Single cell ghost

A single cell ghost occurs when no ray passes through a certain cell. That cell is uncovered, not illuminated, not hit by any of the rays in the data set (at least for the current choice of ray paths). Thus, the slowness of that cell is arbitrary, as it has no effect on determining any of the traveltimes in the data set.

The proper way to deal with such a cell is to assign it some arbitrary value, like the average slowness of all cells or the average of all contiguous cells.

6.2.2 Two cells with only one ray

When any two cells are covered by one and only one ray, a ghost arises because the increment of traveltime δt_i through these cells is invariant to a perturbation of the form

$$\mathbf{g}^T = (0, \dots, 0, l_{ik}, 0, \dots, 0, -l_{ij}, 0, \dots, 0), \tag{6.3}$$

since

$$\delta t_i = l_{ij} s_j + l_{ik} s_k = l_{ij}(s_j + \alpha l_{ik}) + l_{ik}(s_k - \alpha l_{ij}), \tag{6.4}$$

where α is an almost arbitrary scalar. The one constraint on α is that the perturbed slowness vector

$$\mathbf{s}' = \mathbf{s} + \alpha\mathbf{g} \tag{6.5}$$

must be positive. Note that there is no ghost associated with a single cell having only one covering ray.

The proper way to deal with such pairs of cells (especially if they are contiguous) is to treat them as if they were combined into one larger cell, *i.e.*, assign the same value of slowness to both cells. This approach has the effect of eliminating the ghost while simultaneously reducing the size of the model space by one dimension.

If more than two cells are covered by one and only one ray, then there will be multiple ghosts (for p cells there will be $p-1$ ghosts). Again, one way to eliminate this problem is to treat all such cells as a single cell. This approach may not be the best one if the cells are not contiguous. Other approaches will be discussed in the section on eliminating ghosts.

6.2.3 Underdetermined cells in an overdetermined problem

The preceding discussion is a special case of a more general problem: underdetermined cells imbedded in an overdetermined inversion problem. Underdetermination means having fewer equations than unknowns. The example of two cells with only one ray is a common example of this effect. Others would be three cells with two rays, 20 cells with 15 rays, etc. The existence of underdetermined cells may be the result of poor experimental design, of physical limitations at the experimental site that reduce the possible range of view angles significantly (as in crosswell geometry), or they may be caused by severe ray bending effects when high contrasts in the slowness values are present. In the latter situation, we expect that rays will tend to avoid very slow regions (Fermat's principle says to take the fastest path, which may mean to go around the slow region). Since experiments will normally be planned to achieve the desired resolution assuming straight-ray coverage, the actual coverage in slow regions is smaller than planned and may be so reduced by these ray bending effects to the extent of causing underdetermination.

This problem with ghosts can now be reduced to

$$\mathbf{M}'\mathbf{s}' = \delta\mathbf{t}', \tag{6.6}$$

where $\mathbf{M}'$ is an $m' \times n'$ matrix with $m' < n'$, $\mathbf{s}'$ is the subvector of the slowness model of length n', and $\delta\mathbf{t}'$ is the subvector of the traveltime increment of length m'. A particular solution of (6.6) is given by

$$\mathbf{s}' = \mathbf{M}'^T(\mathbf{M}'\mathbf{M}'^T)^{-1}\delta\mathbf{t}', \tag{6.7}$$

if the matrix $\mathbf{M}'\mathbf{M}'^T$ is invertible. But the general solution of (6.6) is a vector of the form

$$\mathbf{s}' = \mathbf{M}'^T(\mathbf{M}'\mathbf{M}'^T)^{-1}\delta\mathbf{t}' + \mathbf{g}', \tag{6.8}$$

where $\mathbf{g}'$ is any vector from the right null space of $\mathbf{M}'$. This null space must have dimension at least $n' - m'$.

The preferred solution to this problem is again to combine contiguous cells until the number of equations is at least equal to the number of unknowns. Then $n' - m' = 0$, and the null space is eliminated. Another method of dealing with the problem if the cells are not contiguous is to assign a slowness value to $n' - m'$ of those cells that have the least coverage, thus removing them from the inversion problem.

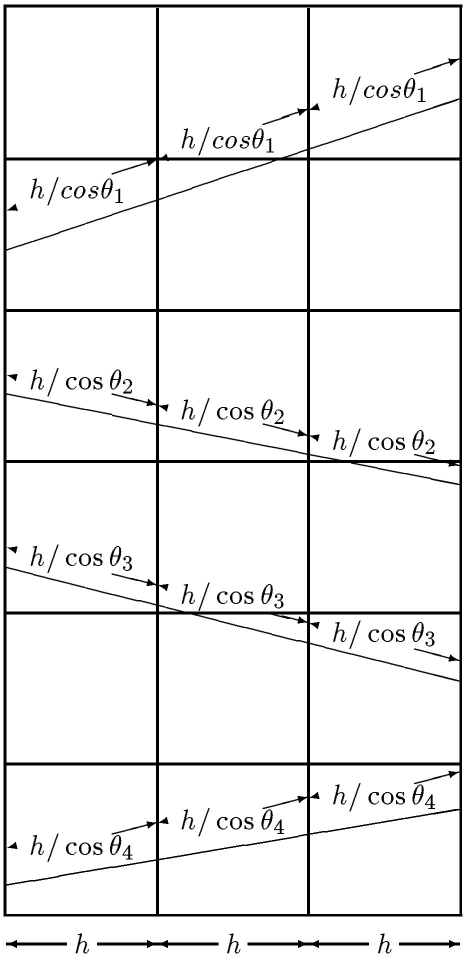


Figure 6.2: Stripes are caused by straight rays used in crosswell geometry.

6.2.4 Stripes

One of the most common types of ghosts in borehole-to-borehole tomography is the vertical stripe. Stripes are ghosts caused by an unfortunate resonance of the model parametrization, the limited set of view angles possible in the crosswell geometry, and the use of straight rays in the reconstruction (see Figure 6.2).

To see how the problem arises, consider the geometry of two vertical boreholes with square cells of dimension h in the image plane. For purposes of illustration, suppose that the borehole separation is just three cell widths and the borehole depth is two cell heights. To get from one borehole to the other, the rays must cross the lines forming the vertical boundaries between cells. Assuming straight rays, each ray is characterized by a total ray-path length $L_i = 3h/\cos\theta_i$, where θ_i is the angle the ray makes with the horizontal. So the total path length in each vertical column is just $h/\cos\theta_i$. This part of the path length is shared among the cells in a column differently for each ray path, but that is not so important. What is important is that the sum is constant in each column for every ray.

The ray-path matrix takes the form

$$\mathbf{M} = \begin{pmatrix} \frac{d_{11}h}{\cos\theta_1} & \frac{d_{12}h}{\cos\theta_1} & \frac{e_{13}h}{\cos\theta_1} & \frac{e_{14}h}{\cos\theta_1} & \frac{f_{15}h}{\cos\theta_1} & \frac{f_{16}h}{\cos\theta_1} \\ \frac{d_{21}h}{\cos\theta_2} & \frac{d_{22}h}{\cos\theta_2} & \frac{e_{23}h}{\cos\theta_2} & \frac{e_{24}h}{\cos\theta_2} & \frac{f_{25}h}{\cos\theta_2} & \frac{f_{26}h}{\cos\theta_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{d_{m1}h}{\cos\theta_m} & \frac{d_{m2}h}{\cos\theta_m} & \frac{e_{m3}h}{\cos\theta_m} & \frac{e_{m4}h}{\cos\theta_m} & \frac{f_{m5}h}{\cos\theta_m} & \frac{f_{m6}h}{\cos\theta_m} \end{pmatrix}, \tag{6.9}$$

where the ds, es, and fs are nonnegative fractions satisfying

$$\sum_{j=1}^{2} d_{ij} = \sum_{j=3}^{4} e_{ij} = \sum_{j=5}^{6} f_{ij} = 1, \tag{6.10}$$

for every ray path $1 \le i \le m$. The ds are associated with the cells in the first column; the es with the second column; and the fs with the third column. Then it is clear that these three vectors

$$\mathbf{g}_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{g}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \quad \mathbf{g}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ -1 \\ -1 \end{pmatrix}, \tag{6.11}$$

are ghosts for this problem, since in each case we find that

$$\mathbf{Mg} = \begin{pmatrix} \frac{h}{\cos\theta_1} - \frac{h}{\cos\theta_1} \\ \frac{h}{\cos\theta_2} - \frac{h}{\cos\theta_2} \\ \vdots \\ \frac{h}{\cos\theta_m} - \frac{h}{\cos\theta_m} \end{pmatrix} = 0 \tag{6.12}$$

follows from (6.10).

These ghosts show up in the reconstructed slowness as vertical stripes — a constant slowness perturbation is subtracted from one column and added to any other column.

To eliminate these ghosts, we need to break the unfortunate symmetry that has caused this artifact to arise. These ghosts would not exist if the cells were not lined up perfectly with both of the vertical boreholes. So one solution is to use cells that are not square or rectangular, *i.e.*, odd shapes like hexagons, triangles, etc. Using a rectangular but staggered grid would also remove the degeneracy. Or, combining a few of the poorly covered cells near the top and bottom would also break the symmetry. A still simpler method of eliminating the problem (at least conceptually) is to use bent rays, rather than the artificially straight rays that are often incorrectly assumed to be adequate.

Problems

Format for Problems 6.2.1–6.2.4: *The next four problems consider a* 3×3 *model with the layout*

s_1	s_2	s_3
s_4	s_5	s_6
s_7	s_8	s_9

Problem 6.2.1 *A ray-path matrix for a* 3×3 *slowness model is*

$$\mathbf{M} = \begin{pmatrix} 1.414 & 0 & 0 & 0 & 1.414 & 0 & 0 & 0 & 1.414 \\ 0 & 0.9 & 1.118 & 1.118 & 0.218 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.054 & 1.054 & 1.054 \end{pmatrix}.$$

Find the ghosts for this set of ray paths. If the measured traveltime vector is

$$\mathbf{t} = \begin{pmatrix} 4.666 \\ 3.488 \\ 3.794 \end{pmatrix},$$

which of the following statements is true of the slowness model?

1. *The model is not constant.*

2. *The average slowness of cells 7, 8, & 9 is 1.2.*

3. *The model is*

0.9	*1.2*	*0.9*
1.0	*1.3*	*1.2*
1.1	*1.4*	*1.1*

Problem 6.2.2 *A ray-path matrix for a* 3×3 *slowness model is*

$$\mathbf{M} = \begin{pmatrix} 0 & 0 & 0 & 1.045 & 1.045 & 1.045 & 0 & 0 & 0 \\ 0 & 0 & 0.403 & 1.052 & 1.052 & 0.649 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.021 & 1.057 & 1.057 & 1.036 & 0 & 0 \\ 1.011 & 0 & 0 & 0.101 & 1.112 & 0.101 & 0 & 0 & 1.011 \\ 0 & 0 & 0 & 0 & 0.502 & 1.004 & 1.004 & 0.502 & 0 \end{pmatrix}.$$

Find the ghosts for this set of ray paths.

PROBLEM 6.2.3 *A ray-path matrix for a* 3×3 *slowness model is*

$$\mathbf{M} = \begin{pmatrix} 1.012 & 0.502 & 0 & 0 & 0.510 & 1.012 & 0 & 0 & 0 \\ 0 & 0 & 0.444 & 1.012 & 1.012 & 0.568 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.005 & 1.005 & 1.005 \\ 0.503 & 1.103 & 1.103 & 0.600 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.001 & 0 & 0 & 0.009 & 1.010 & 1.010 \end{pmatrix}.$$

Find the ghosts for this set of ray paths.

PROBLEM 6.2.4 *A ray-path matrix for a* 3×3 *slowness model is*

$$\mathbf{M} = \begin{pmatrix} 1.033 & 1.033 & 1.033 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.049 & 1.049 & 1.049 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.101 & 1.101 & 1.101 \\ 0 & 0 & 1.414 & 0 & 1.414 & 0 & 1.414 & 0 & 0 \\ 1.118 & 0 & 0 & 0 & 1.118 & 1.118 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.052 & 0 & 0 & 0.025 & 1.077 & 1.077 \\ 0 & 0 & 0 & 0 & 1.044 & 1.055 & 1.055 & 0.011 & 0 \\ 0 & 0 & 0.352 & 1.068 & 1.068 & 0.716 & 0 & 0 & 0 \\ 0.022 & 0 & 0 & 1.053 & 1.075 & 1.075 & 0 & 0 & 0 \end{pmatrix}.$$

Find the ghosts for this set of ray paths, if any.

PROBLEM 6.2.5 *Consider an* $m \times n$ *ray-path matrix* $\mathbf{M}$ *for straight rays through a rectangular model with* q *columns and* r *rows, so* $n = q \times r$. *Suppose that all the rays considered are crosswell. Show that the diagonal elements of the model resolution matrix* $\mathcal{R} = \mathbf{M}^{\dagger}\mathbf{M}$ *satisfy*

$$\mathcal{R}_{jj} \leq \frac{n+1-q}{n}. \tag{6.13}$$

Make a table of these resolution bounds for different choices of the number of columns and rows. What strategy for model design leads to the best resolution when the rays are straight and crosswell?

PROBLEM 6.2.6 *Consider a slowness model composed of vertical stripes (see Figure 6.3). Suppose that all slowness values and thicknesses of the stripes are known, but the spatial order is not known. Use Snell's law to show that the spatial order of the stripes cannot be determined from vertical crosswell transmission traveltime data. [Hint: Consider the ray parameter (5.17).]*

6.2.5 Linear dependence

Ghosts arise from the linear dependence of $\mathbf{M}$ or of submatrices such as $\mathbf{M}'$. The example of stripes arises from a gross linear dependence of all the rows of the full matrix $\mathbf{M}$. The examples of underdetermined group of cells arise because of poor coupling (or coverage)

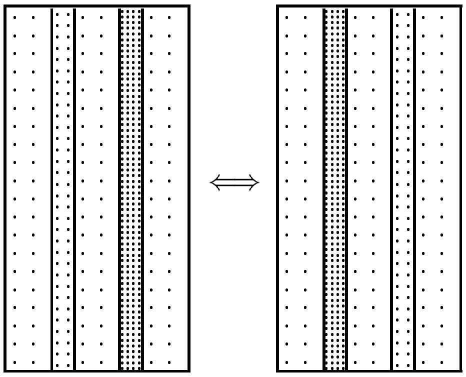


Figure 6.3: Two indistinguishable models: the spatial ordering of the stripes cannot be determined from crosswell transmission data using straight rays.

between the rays and the cells. Other, more subtle and complex, linear dependencies may also occur.

Indeed the defining equation for a ghost

$$\mathbf{M}\mathbf{g} = 0 \tag{6.14}$$

is a statement of linear dependence of the rows of $\mathbf{M}$. Equation (6.14) contains m equations for the n unknown components of $\mathbf{g}$, with $m > n$. Any n of these m equations are sufficient to determine $\mathbf{g}$ and the remaining $m-n$ equations can be determined from these n equations. An exception to this statement occurs when the ghost is caused by complete decoupling as in the case of the single cell ghost.

6.3 Eliminating Ghosts (Ghostbusting)

Ghosts may be removed by using a variety of techniques, some of which have already been described. Although ray bending can introduce ghosts in situations where very slow regions are avoided by most rays, it can also provide a simple solution to some of the problems created by the limited view angles available in crosswell measurements. Methods of improving the coupling between rays and cells include implementations using fat rays [Kak, 1984; Michelena and Harris, 1991].

6.3.1 Fat rays

The fact that ray paths are stationary (*i.e.*, that small variations in the ray path have no effect on the traveltime to first order) means that each ray actually has a bundle of rays close to it, all with virtually the same traveltime. We can improve the coupling between the rays and cells in the model by taking advantage of this fact. One possible approach is to use more than one ray between each source and receiver pair: for example, during the computation of the ray paths, we could save not only the ray we found with the least traveltime, but also save several other trial rays that were close to the best one. Then, in place of the single row of the ray-path matrix for the ith ray, we now insert multiple rows

$$\mathbf{M}_i = \begin{pmatrix} l_{11} & l_{12} & \cdots & l_{1n} \\ l_{21} & l_{22} & \cdots & l_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ l_{\mu 1} & l_{\mu 2} & \cdots & l_{\mu n} \end{pmatrix}, \tag{6.15}$$

where μ is the multiplicity of the the ith ray path. This approach is relatively easy to implement either for bending methods or for shooting methods of ray tracing, and has the effect of multiplying the size of the data set by the number of rays (μ) saved per source/receiver pair. The disadvantage of using multiple ray paths per source/receiver pair is that the data storage problem also gets multiplied by the number of rays saved per pair.

The method of fat rays is an alternative using the same underlying physics without increasing the size of our matrices. In this approach, we treat each ray as if it has a finite thickness. Then, instead of measuring the linear increment of the ray that has passed through a cell, in 2-D the ray now has an area associated with it and we measure the overlap of the ray area with the cell area. In three dimensions, these areas all become volumes. If the ray width in 2-D is Δw and the ray cross section in 3-D has area Δa, then the ray-path matrix becomes

$$\mathbf{M} = \frac{1}{\Delta w} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \tag{6.16}$$

with the a_{ij}s being overlap areas in two dimensions and

$$\mathbf{M} = \frac{1}{\Delta a} \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{m1} & v_{m2} & \cdots & v_{mn} \end{pmatrix}, \tag{6.17}$$

with the v_{ij}s being overlap volumes in three dimensions. For traveltime tomography, it is still important that the sums of the rows of $\mathbf{M}$ result in sensible ray-path lengths, otherwise the reconstructed slowness values will not be meaningful. The disadvantage of the method just outlined is that the overlap areas and volumes are often tedious to compute.

Another method that has the advantages of both of the previous methods is first to obtain the set of near-ray-path lengths shown in (6.15) and then average them according to

$$\bar{l}_{ij} = \frac{1}{\mu} \sum_{i'=1}^{\mu} l_{i'j}. \tag{6.18}$$

With this approach, we end up with a single effective ray path and so do not increase the ray storage problem, but we have the advantage that the individual contributions $l_{i'j}$ leading to $\bar{l}_{ij}$ through (6.18) are comparatively easy to compute. The resulting ray-path matrix is

$$\overline{\mathbf{M}} = \begin{pmatrix} \bar{l}_{11} & \bar{l}_{12} & \cdots & \bar{l}_{1n} \\ \bar{l}_{21} & \bar{l}_{22} & \cdots & \bar{l}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{l}_{m1} & \bar{l}_{m2} & \cdots & \bar{l}_{mn} \end{pmatrix}. \tag{6.19}$$

It is clear that fat rays will accomplish the goal of improving the coupling between the rays and cells. The matrices in (6.16) and (6.17) will certainly be significantly less sparse than the the usual $\mathbf{M}$ based on skinny rays. Whether this change will be sufficient to make a significant improvement in the reconstructions will, of course, depend on the particular application. In general, fat rays should be used in addition to (not instead of) the other methods of ghostbusting described in this section.

6.3.2 Damping

The general damped least-squares solution $\mathbf{s}$ to the inversion problem, including general weight matrices, was shown in (4.100) to satisfy

$$\left(\mathbf{M}^T\mathbf{F}^{-1}\mathbf{M} + \mu\mathbf{G}\right)(\mathbf{s} - \mathbf{s}_0) = \mathbf{M}^T\mathbf{F}^{-1}(\mathbf{t} - \mathbf{M}\mathbf{s}_0). \tag{6.20}$$

Since damping is designed to reduce the effects of contributions from eigenvectors with low eigenvalues, we should check to see if this approach eliminates ghosts. Multiplying (6.20) on the right by $\mathbf{g}^T$ and using the definition $\mathbf{M}\mathbf{g} = 0$, we find that

$$\mathbf{g}^T\mathbf{G}(\mathbf{s} - \mathbf{s}_0) = 0. \tag{6.21}$$

When the weight matrix $\mathbf{G} = \mathbf{I}$, (6.21) shows that no ghosts can contribute to $\mathbf{s}$ unless $\mathbf{s}_0$ already contains such terms. Thus, norm damping successfully eliminates ghosts in this one case. However, if $\mathbf{G} \neq \mathbf{I}$, (6.21) is a conjugacy condition, showing that $\mathbf{g}$ is orthogonal to the correction vector $\mathbf{s} - \mathbf{s}_0$ relative to the weight matrix $\mathbf{G}$. Although (6.21) is consistent with the absence of ghosts in the correction vector, it does not guarantee they are eliminated in all cases. When $\mathbf{G} = \mathbf{C}$ (the coverage matrix), the interpretation of (6.21) is almost as simple as that for the identity matrix; when $\mathbf{G} = \mathbf{K}^T\mathbf{K}$ where $\mathbf{K}$ is based on gradients or Laplacians of the the model vector, its interpretation becomes more complex.

6.3.3 Summary

Methods of eliminating ghosts can be divided into two main categories: (*i*) experimental design and (*ii*) model design together with analytical tricks.

No amount of analysis can salvage a badly designed experiment. When designing a tomographic experiment, it is important to gather data from as many view angles as possible. It is also important to gather enough data so that the cells we can resolve from our data analysis are about the same size as the anomalies we want to detect. A rule of thumb is that the number of source/receiver pairs should be about *twice* the number of cells we want to resolve in our experiment. Another useful rule of thumb is to choose the average cell size to be about $3\lambda_{\max}$, where $\lambda_{\max} = 1/f_{\min}s_{\min}$ is the maximum expected wavelength associated with the minimum frequency $f_{\min}$ in the pulse propagation data and the minimum slowness $s_{\min}$ expected in the region to be imaged. This rule arises from extensive experience with the asymptotic analysis of wave propagation which we will not present here.

The analyst must design the model to take optimum advantage of the data gathered, while accounting for any prior knowledge of the medium to be imaged. The shapes and sizes of the model cells are ours to choose, and should be used to advantage to solve any problems that cannot be eliminated through good experimental design. We are always free to choose cells larger than the expected resolution of the traveltime data. We may delete some cells if they have poor ray coverage, or some contiguous cells with poor coverage may be combined into a single cell for purposes of reconstruction. Cells can be of any shape we choose; the choice of square or rectangular cells is often made for ease of display and for ease of computation of ray paths, but other considerations may drive us to use odd shapes for cells in some applications. Analytical tricks can be applied during the reconstruction process once we have the data at home. Smoothing and clipping the slowness model values can be done to force the reconstructed values to lie within reasonable limits. Fat rays are a last resort if the other methods are not sufficient to eliminate the ghosts.

6.4 Significance of Ghosts

It is important to recognize that elimination of all ghosts may be neither possible nor desirable. In our efforts to solve the inverse problem

$$\mathbf{Ms} = \mathbf{t} \tag{6.22}$$

for the slowness model $\mathbf{s}$, we should keep in mind that there are really three stages in the inversion process. The first stage is to find, if possible, a particular model $\mathbf{s}$ that satisfies the data. The second stage is to analyze the null space of the operator $\mathbf{M}$. We may use standard numerical techniques at this point in the analysis to perform a singular value decomposition of $\mathbf{M}$ and obtain a full characterization of the null space. Having finished both of these steps, we can finally provide the complete solution to the inversion problem. In fact, it may be that we need perturbations from the null space to satisfy various physical or geological boundary conditions present at the site where the tomographic data were gathered. This process is entirely analogous to the process of solving an ordinary differential equation by finding a particular solution, computing a set of homogeneous solutions, and finally producing a linear combination that satisfies the initial or boundary conditions.

Problems

Problem 6.4.1 *An inhomogeneous linear differential equation of first order is*

$$\frac{dx}{dt} + \lambda x = f(t),$$

with the initial condition $x(t) = x(0)$ for $t = 0$. The solution of this equation is well known to be

$$x(t) = x(0)\exp(-\lambda t) + \exp(-\lambda t)\int_0^t f(t')\exp(\lambda t')dt'.$$

Which term of this solution is analogous to $\mathbf{s}_{\mathrm{LS}} = \mathbf{M}^\dagger\mathbf{t}$*? Which term is analogous to a ghost satisfying* $\mathbf{Mg} = 0$*? Which term is analogous to the scalar* γ *in the full solution* $\mathbf{s} = \mathbf{s}_{\mathrm{LS}} + \gamma\mathbf{g}$*?*

PROBLEM 6.4.2 *Solve* $\mathbf{Ms} = \mathbf{t}$ *for* $\mathbf{s}$ *when*

$$\mathbf{M} = \begin{pmatrix} 13/12 & 13/12 & 5/6 \\ 13/12 & 13/12 & 5/6 \\ 5/6 & 5/6 & 4/3 \end{pmatrix}$$

and $\mathbf{t}^T = (49/16, 49/16, 23/8)$*, subject to the constraint* $s_2 = s_1 + 1/4$*. If the constraint is changed to* $s_2 = s_1 + 3$*, does the problem still have a physical solution (*$\mathbf{s} > 0$*)? What is the range of permissible values of* σ *in the constraint* $s_2 = s_1 + \sigma$*?*

PROBLEM 6.4.3 *Reconsider* PROBLEM 6.2.1 *in light of the following additional information:*

1. *The total variation in the model from top to bottom is less than 20%.*
2. *The model is probably horizontally stratified, or nearly so.*
3. *The given ray-path matrix is only a straight ray approximation to the true ray path.*

First, solve the problem by considering only the first two of the new constraints while using the straight ray approximation. Then, try to solve the full problem using one-step backprojection based on bent rays.

PROBLEM 6.4.4 *Formulate a definition of the "best approximate solution" of the matrix equation* $\mathbf{MZ} = \mathbf{Y}$ *when constraints on* $\mathbf{Z}$ *are given. Compare and contrast this definition with the one given in* PROBLEM 4.1.25*. Is there a unique best approximate solution of* $\mathbf{MZ} = \mathbf{Y}$ *that satisfies the new definition?*

Chapter 7

Nonlinear Seismic Inversion

The introduction of feasibility constraints into the traveltime inversion problem offers a unique opportunity to develop a variety of new reconstruction algorithms. A few of the ones that have been explored so far will be discussed in this Chapter.

7.1 Linear and Nonlinear Programming

We will see that linear tomography maps easily onto linear programming, and nonlinear tomography onto nonlinear programming [Strang, 1986; Fiacco and McCormick, 1990].

Recall that, if $\mathbf{u}^T = (1, \ldots, 1)$ is an m-vector of ones and $\mathbf{v}^T = (1, \ldots, 1)$ is an n-vector of ones, then

$$\mathbf{u}^T\mathbf{M} = \mathbf{v}^T\mathbf{C}, \tag{7.1}$$

where $\mathbf{C}$ is the coverage *matrix*, *i.e.*, the diagonal matrix whose diagonal elements are the column sums of the ray-path matrix. We will now define the coverage *vector* as

$$\mathbf{c} = \mathbf{C}\mathbf{v}. \tag{7.2}$$

7.1.1 Duality

The concept of duality in linear programming leads to some useful ideas both for linear and nonlinear traveltime inversion. (Actually it is even more useful for electrical impedance tomography as we will see in Part II.) We will first define the following:

DEFINITION 7.1.1 *The primal problem for traveltime inversion is to find the minimum of* $\mathbf{c}^T\mathbf{s}$ *subject to* $\mathbf{M}\mathbf{s} \geq \mathbf{t}$ *and* $\mathbf{s} \geq 0$.

DEFINITION 7.1.2 *The dual problem associated with the primal is to maximize* $\mathbf{w}^T\mathbf{t}$ *subject to* $\mathbf{w}^T\mathbf{M} \leq \mathbf{c}^T$ *and* $\mathbf{w} \geq 0$.

The m-vector $\mathbf{w}$ has no physical significance, but plays the role of a nonnegative weight vector. One of the first consequences of this formulation is that, if we multiply the primal

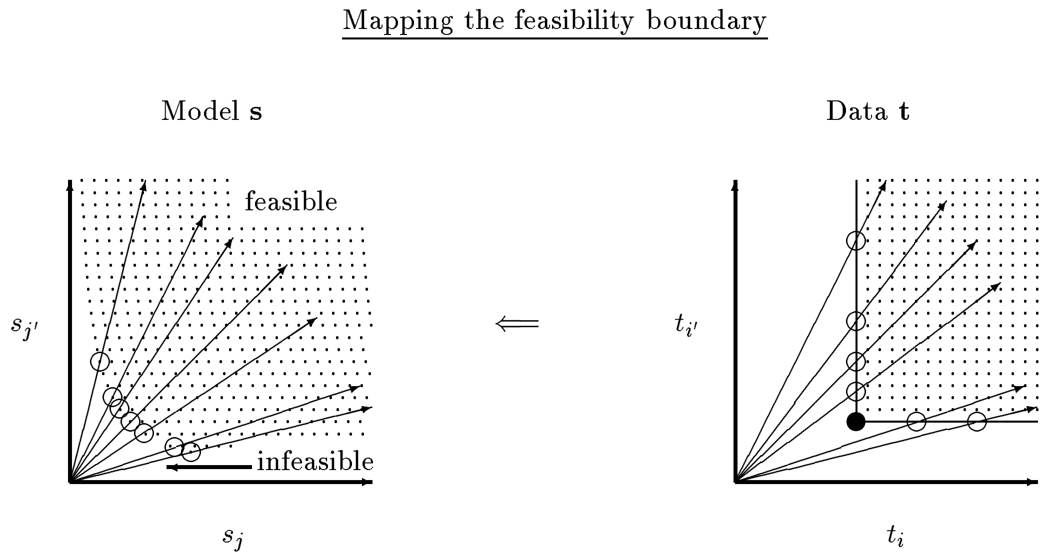


Figure 7.1: Mapping the feasibility boundary.

inequality on the right by $\mathbf{w}^T$ and the dual inequality on the left by $\mathbf{s}$ for feasible $\mathbf{s}$ and $\mathbf{w}$, then

$$\mathbf{c}^T\mathbf{s} \geq \mathbf{w}^T\mathbf{M}\mathbf{s} \geq \mathbf{w}^T\mathbf{t}. \tag{7.3}$$

We introduce a Lagrangian functional

$$\mathcal{L}(\mathbf{s}, \mathbf{w}) = \mathbf{c}^T\mathbf{s} + \mathbf{w}^T(\mathbf{t} - \mathbf{M}\mathbf{s}) \tag{7.4}$$
$$= (\mathbf{c}^T - \mathbf{w}^T\mathbf{M})\mathbf{s} + \mathbf{w}^T\mathbf{t}. \tag{7.5}$$

An admissible (feasible) weight vector is $\mathbf{w} = \mathbf{u}$. In fact, this is the only weight vector we need to consider because it saturates the dual inequality, producing equality in all components following (7.1) and (7.2). Thus, the dual problem in traveltime inversion is completely trivial. We introduced it here because, despite its apparent triviality, there is one interesting feature.

In problems with nontrivial duality structure, it is possible to obtain useful bounds with inequalities equivalent to (7.3). Here we are left with just the condition

$$\boxed{\mathbf{c}^T\mathbf{s} \quad \geq \quad \mathbf{u}^T\mathbf{t} \quad = \quad T,} \tag{7.6}$$

defining the hyperplane of constant total traveltime. Equation (7.6) could have been derived directly from the feasibility conditions $\mathbf{M}\mathbf{s} \geq \mathbf{t}$ for $\mathbf{s}$. The ease of its derivation should not, however, lead us to think that this equation is trivial. This hyperplane can play an important role in linear and nonlinear programming algorithms for traveltime inversion.

7.1.2 Relaxed feasibility constraints

Given the set of observed traveltimes, t_i for $i = 1, \ldots, m$, we define two more types of feasibility sets.

DEFINITION 7.1.3 (RELAXED LOCAL FEASIBILITY SET) *The relaxed local feasibility set with respect to a set of trial ray paths $\mathcal{P} = \{P_1, \ldots, P_m\}$ and observed traveltimes $t_1, \ldots, t_m$ is*

$$\mathcal{R}^{\mathcal{P}} = \{\mathbf{s} \mid \sum_{i=1}^{m} \tau_i^P(\mathbf{s}) \geq \sum_{i=1}^{m} t_i\}. \tag{7.7}$$

DEFINITION 7.1.4 (RELAXED GLOBAL FEASIBILITY SET) *The relaxed global feasibility set with respect to a set of observed traveltimes $t_1, \ldots, t_m$ is*

$$\mathcal{R}^{*} = \{\mathbf{s} \mid \sum_{i=1}^{m} \tau_i^*(\mathbf{s}) \geq \sum_{i=1}^{m} t_i\}. \tag{7.8}$$

PROPOSITION 7.1.1 (SUM OF CONCAVE FUNCTIONS) *A (nonnegatively) weighted sum of concave functions is concave.*

Proof: Let $\tau_i(\mathbf{s})$ for $i = 1, \ldots, m$ be a set of concave functions and let w_i be a set of nonnegative weights. Then,

$$\sum_{i=1}^{m} w_i \tau_i(\lambda \mathbf{s}_1 + (1-\lambda)\mathbf{s}_2) \geq \sum_{i=1}^{m} w_i [\lambda \tau_i(\mathbf{s}_1) + (1-\lambda)\tau_i(\mathbf{s}_2)] \tag{7.9}$$

$$= \lambda \sum_{i=1}^{m} w_i \tau_i(\mathbf{s}_1) + (1-\lambda) \sum_{i=1}^{m} w_i \tau_i(\mathbf{s}_2), \tag{7.10}$$

so the weighted sum is concave. ∎

THEOREM 7.1.1 *$\mathcal{R}^{\mathcal{P}}$ is a convex set.*

THEOREM 7.1.2 *$\mathcal{R}^{*}$ is a convex set.*

Proof: Both theorems follow from the proposition and the fact that the unit-weighted sums in the definitions of the sets $\mathcal{R}^{\mathcal{P}}$ and $\mathcal{R}^{*}$ are respectively sums of the concave functions $\tau_i^P(\mathbf{s})$ and $\tau_i^*(\mathbf{s})$. ∎

THEOREM 7.1.3 *Any point $\mathbf{s}^*$ that lies simultaneously on the boundary of both $\mathcal{F}^{\mathcal{P}}$ and $\mathcal{R}^{\mathcal{P}}$ solves the inversion problem.*

THEOREM 7.1.4 *Any point $\mathbf{s}^*$ that lies simultaneously on the boundary of both $\mathcal{F}^{*}$ and $\mathcal{R}^{*}$ solves the inversion problem.*

Proof: The boundary of $\mathcal{R}^{\mathcal{P}}$ is determined by the single equality constraint

$$\sum_{i=1}^{m} \tau_i^P(\mathbf{s}) = \sum_{i=1}^{m} t_i = T. \tag{7.11}$$

The boundary of $\mathcal{F}^{\mathcal{P}}$ is determined by the set of inequality constraints

$$\tau_i^P(\mathbf{s}) \geq t_i, \quad \text{for all} \quad i = 1, \ldots, m, \tag{7.12}$$

with equality holding for at least one of the constraints. Summing (7.12) gives

$$\sum_{i=1}^{m} \tau_i^P(\mathbf{s}) \geq T, \tag{7.13}$$

where the equality applies if and only if $\tau_i^P(\mathbf{s}) = t_i$ for all i. Therefore, any model $\mathbf{s}^*$ that satisfies both (7.11) and (7.12) must solve the inversion problem.

The proof of the second theorem follows the proof of the first, with $\tau^*(\mathbf{s})$ replacing $\tau^P(\mathbf{s})$ everywhere. ▮

The boundaries of relaxed feasibility sets (either local or global) are easier to compute than those for unrelaxed feasibility sets. The difference is, for example, a single hyperplane boundary for a relaxed local feasibility set versus up to m (the number of traveltime measurements) hyperplanes composing the boundary of an unrelaxed local feasibility set. Yet, the characteristics of the relaxed feasibility sets are very similar to the unrelaxed ones in other ways.

If the correct ray-path matrix for the inversion problem has been found and the data are noise free, then we expect that the hyperplane defined by $\mathbf{c}^T\mathbf{s} = T$ will intersect the feasibility boundary exactly at the point or points that solve the inversion problem. If the correct ray-path matrix has not been found or there is uncorrelated noise in our data $\mathbf{t}$, then there will be a *splitting* between the hyperplane of constant total traveltime and the feasible region. The point (or points) of closest approach between the convex feasible set and the hyperplane may then be defined as the set of points *solving* the linear programming problem for fixed $\mathbf{M}$. An iterative nonlinear programming algorithm may then be constructed wherein the updated $\mathbf{M}$ is determined based on the solution of the last linear programming problem. This procedure converges if the degree of splitting (Euclidean distance) between the feasible set and the hyperplane of constant traveltime tends to zero from one iteration to the next.

Problem

PROBLEM 7.1.1 *Consider the following backprojection formulas (see* PROBLEMS *1.5.1 and 1.5.3):*

1. $\mathbf{s} = \mathbf{N}^{-1}\mathbf{H}^T\mathbf{L}^{-1}\mathbf{t}$;

2. $\mathbf{s} = \mathbf{C}^{-1}\mathbf{M}^T\mathbf{L}^{-1}\mathbf{t}$.

Does either formula always satisfy the constraint $\mathbf{c}^T\mathbf{s} = T$ *? Find another backprojection formula that does satisfy the constraint.*

7.2 More about Weighted Least-Squares

We learned in Sections 3.5 and 4.3 that a good set of weights for use with weighted least-squares was $\mathbf{L}^{-1}$ for the traveltime errors and $\mathbf{C}$ for the smoothing or regularization term in a damped least-squares method. The arguments were based on assumptions of small deviations from a constant background or on the desire to precondition the ray-path matrix so its eigenvalues λ were normalized to the range $-1 \leq \lambda \leq 1$.

In a sense the methods used to choose the weights previously were based on ideas of *linear inversion.* We should now try to see if these ideas need to be modified for *nonlinear inversion.* Let $\mathbf{s}$ be the latest estimate of the slowness model vector in an iterative inversion scheme. Then, if $\mathbf{u}^T = (1, \ldots, 1)$ is an m-vector of ones and $\mathbf{v}^T = (1, \ldots, 1)$ is an n-vector of ones,

$$\mathbf{Ms} = \mathbf{Tu}, \tag{7.14}$$

$$\mathbf{M}^T\mathbf{u} = \mathbf{Cv} \equiv \mathbf{Ds}, \tag{7.15}$$

where $\mathbf{C}$ is the coverage matrix (diagonal matrix containing the column sums of $\mathbf{M}$) defined previously and the two new matrices ($\mathbf{T}$ and $\mathbf{D}$) are diagonal matrices whose diagonal elements are T_{ii}, the estimated traveltime for the ith ray path through the model $\mathbf{s}$,

$$T_{ii} = \sum_{j=1}^{n} l_{ij} s_j, \tag{7.16}$$

and D_{jj} where

$$D_{jj} \equiv C_{jj}/s_j = \sum_{i=1}^{m} l_{ij}/s_j. \tag{7.17}$$

For the sake of argument, let the inverse of the diagonal traveltime matrix $\mathbf{T}^{-1}$ be the weight matrix, and compute the scaled least-squares point. The least-squares functional takes the form

$$\psi(\gamma) = (\mathbf{t} - \mathbf{M}\gamma\mathbf{s})^T\mathbf{T}^{-1}(\mathbf{t} - \mathbf{M}\gamma\mathbf{s}), \tag{7.18}$$

which has its minimum at

$$\gamma = \frac{\mathbf{s}^T\mathbf{M}^T\mathbf{T}^{-1}\mathbf{t}}{\mathbf{s}^T\mathbf{M}^T\mathbf{T}^{-1}\mathbf{Ms}}. \tag{7.19}$$

Equation (7.19) can be rewritten using (7.14) as

$$\gamma = \frac{\mathbf{u}^T\mathbf{t}}{\mathbf{u}^T\mathbf{Tu}}. \tag{7.20}$$

The factor γ that minimizes the least-squares error is therefore the one that either increases or decreases the total traveltime of the model $\mathbf{s}$ so it equals that of the data. If we assume that the measurement errors in the traveltime data $\mathbf{t}$ are unbiased, then it is very reasonable to choose models that have this property, because the total traveltime $\mathbf{u}^T\mathbf{t} = T$ will tend to have smaller error (by a factor of $m^{-1/2}$) than the individual measurements.

We see that requiring the models $\mathbf{s}$ to have the same total traveltime as the data is equivalent to requiring that the models all lie in the hyperplane defined by

$$\mathbf{u}^T\mathbf{M}\mathbf{s} = \mathbf{v}^T\mathbf{C}\mathbf{s} = \mathbf{c}^T\mathbf{s} = T. \tag{7.21}$$

But this is precisely the same hyperplane (7.6) that arose naturally in the earlier discussion of linear and nonlinear programming.

To carry this analysis one step further, consider the weighted least-squares problem

$$\phi_\mu(\mathbf{s}) = (\mathbf{t} - \mathbf{M}\mathbf{s})^T\mathbf{T}^{-1}(\mathbf{t} - \mathbf{M}\mathbf{s}) + \mu(\mathbf{s} - \mathbf{s}_0)^T\mathbf{D}(\mathbf{s} - \mathbf{s}_0), \tag{7.22}$$

where we assume that the starting model $\mathbf{s}_0$ satisfies $\mathbf{c}^T\mathbf{s}_0 = T$. Then, the minimum of (7.22) occurs for $\mathbf{s}_\mu$ satisfying

$$(\mathbf{M}^T\mathbf{T}^{-1}\mathbf{M} + \mu\mathbf{D})(\mathbf{s}_\mu - \mathbf{s}_0) = \mathbf{M}^T\mathbf{T}^{-1}(\mathbf{t} - \mathbf{M}\mathbf{s}_0). \tag{7.23}$$

Multiplying (7.23) on the left by $\mathbf{s}_0^T$, we find that

$$(1 + \mu)\mathbf{c}^T(\mathbf{s}_\mu - \mathbf{s}_0) = \mathbf{u}^T(\mathbf{t} - \mathbf{M}\mathbf{s}_0) = 0, \tag{7.24}$$

so the solution of the weighted least-squares problem (7.23) also has the property that its estimated total traveltime for all rays is equal to that of the data

$$\mathbf{c}^T\mathbf{s}_\mu = \mathbf{c}^T\mathbf{s}_0 = T. \tag{7.25}$$

Our conclusion is that the particular choice of weighted least-squares problem (7.23) has the *unique* property of holding the total estimated traveltime equal to the total of the measured traveltimes, *i.e.*, it constrains the least-squares solution to lie in the hyperplane $\mathbf{c}^T\mathbf{s} = T$. Assuming that the traveltime data are themselves unbiased (*i.e.*, $\mathbf{u}^T\Delta\mathbf{t} = 0$ where $\Delta\mathbf{t}$ is the measurement error vector), the result $\mathbf{s}$ is an unbiased estimator of the slowness. Moreover, this property is maintained for *any* value of the damping parameter μ. This result provides a connection between the linear programming approach and weighted linear least-squares. We can now use weighted least-squares and the formula (7.23) in a linear program as a means of moving around in the hyperplane $\mathbf{c}^T\mathbf{s} = T$.

Now, from our general analysis of the eigenvalue structure of weighted least-squares, recall that (4.108) shows, for $\mathbf{F} = \mathbf{T}$ and $\mathbf{G} = \mathbf{D}$, that we have

$$\frac{L_{ii}C_{jj}}{T_{ii}D_{jj}} \geq \lambda^2, \tag{7.26}$$

which must hold true for all values of i, j. From (7.17), we have $C_{jj}/D_{jj} = s_j$ so

$$\frac{L_{ii}s_j}{T_{ii}} \geq \frac{L_{ii}s_{\min}}{T_{ii}} \geq \lambda^2, \tag{7.27}$$

and from the definition of T_{ii} we have

$$T_{ii} = \sum_{j=1}^{n} l_{ij}s_j \geq L_{ii}s_{\min}. \tag{7.28}$$

We conclude that this choice of weight matrices also constrains the eigenvalues to be bounded above by unity, *i.e.*, $1 \geq \lambda^2$.

If the matrix $\mathbf{M}$ is very large, it may be impractical to solve (7.23) by inverting the matrix $(\mathbf{M}^T\mathbf{T}^{-1}\mathbf{M} + \mu\mathbf{D})$. Instead, we may choose to use some version of the method we called "simple iteration" in Section 4.4.3. For example, suppose that the kth iteration yields the model vector $\mathbf{s}_\mu^{(k)}$. Then, one choice of iteration scheme for finding the next iterate is

$$\mathbf{D}\mathbf{s}_\mu^{(k+1)} = \mathbf{D}\mathbf{s}_\mu^{(k)} + \mathbf{M}^T\mathbf{T}^{-1}(\mathbf{t} - \mathbf{M}\mathbf{s}_0) - (\mathbf{M}^T\mathbf{T}^{-1}\mathbf{M} + \mu\mathbf{D})(\mathbf{s}_\mu^{(k)} - \mathbf{s}_0). \tag{7.29}$$

It is not hard to show that this iteration scheme converges as long as the damping parameter is chosen so that $0 < \mu < 1$. [1] Furthermore, if we multiply (7.29) on the left by $\mathbf{s}_0^T$, we find that

$$\mathbf{c}^T(\mathbf{s}_\mu^{(k+1)} - \mathbf{s}_\mu^{(k)}) = (1 + \mu)\mathbf{c}^T(\mathbf{s}_0 - \mathbf{s}_\mu^{(k)}). \tag{7.30}$$

It follows from (7.30) that, if $\mathbf{c}^T\mathbf{s}_0 = T$ and if $\mathbf{s}_\mu^{(0)} = \mathbf{s}_0$, then

$$\mathbf{c}^T\mathbf{s}_\mu^{(k)} = T, \tag{7.31}$$

for all k. Thus, all the iterates stay in the hyperplane of constant total traveltime. If we choose not to iterate to convergence, then this desirable feature of the exact solution $\mathbf{s}_\mu$ proven in (7.25) is still shared by every iterate $\mathbf{s}_\mu^{(k)}$ obtained using this scheme.

Problem

Problem 7.2.1 *Use the definition of the pseudoinverse in* Problem 4.1.16 *to show that, if*

$$\mathbf{M}' = \mathbf{T}^{-\frac{1}{2}}\mathbf{M}\mathbf{D}^{-\frac{1}{2}},$$

then

$$\mathbf{X} = \mathbf{D}^{-\frac{1}{2}}(\mathbf{M}')^\dagger\mathbf{T}^{-\frac{1}{2}}, \tag{7.32}$$

where $\mathbf{X}$ *is an approximate generalized inverse satisfying the first two conditions (*$\mathbf{MXM} = \mathbf{M}$ *and* $\mathbf{XMX} = \mathbf{X}$*). Use (7.32) to show that the SVD of* $\mathbf{X}$ *has the form*

$$\mathbf{X} = \frac{\mathbf{s}_0\mathbf{u}^T}{\mathbf{u}^T\mathbf{T}\mathbf{u}} + \dots,$$

where the terms not shown are for eigenvectors of $\mathbf{M}'$ *with eigenvalues* $\lambda < 1$*. Apply this result to the inversion problem to show that*

$$\mathbf{s} \simeq \mathbf{X}\mathbf{t} = \mathbf{s}_0 + \dots,$$

where in this case the terms not shown are contributions orthogonal to $\mathbf{s}_0$*.*

[1] The reader may want to check this result using the methods of this section and the ones developed in Section 4.4.3 on simple iteration.

7.3 Stable Algorithm for Nonlinear Crosswell Tomography

Here we combine several ideas from the previous sections into an algorithm for nonlinear traveltime tomography. We recall that such algorithms are inherently iterative. In the general iterative algorithm posed earlier, the questionable step was how to update the current model $\hat{\mathbf{s}}$ to obtain an improved model. Here we propose a method for choosing this step [Berryman, 1989b; 1990].

Let $\mathbf{s}^{(k)}$ be the current model. An algorithm (see Figure 7.2) for generating the updated model $\mathbf{s}^{(k+1)}$ is as follows:

1. Set $\mathbf{s}_1$ to the scaled least-squares model:
$$\mathbf{s}_1 = \hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}^{(k)}]}.$$

2. Set $\mathbf{s}_2$ to the damped least-squares model with respect to $\mathbf{s}_1$:
$$\mathbf{s}_2 = \hat{\mathbf{s}}_{\mathrm{LS}[\mathbf{s}_1,\mu]}.$$

3. Define the family of models
$$\mathbf{s}(\lambda) = (1-\lambda)\mathbf{s}_1 + \lambda\mathbf{s}_2,$$
where $\lambda \in [0,1]$.

4. Solve for λ^*, defined so that $\mathbf{s}(\lambda^*)$ yields the fewest number of feasibility violations. The number of feasibility violations is defined as the number of ray paths for which $t_i > \tau^*(\mathbf{s}(\lambda))$.

5. If λ^* is less than some preset threshold (say 0.05 or 0.1), reset it to the threshold value.

6. Set $\mathbf{s}^{(k+1)} = \mathbf{s}(\lambda^*)$.

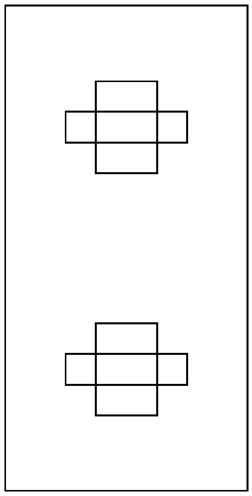
The feasibility structure of the algorithm is illustrated in Fig. 7.3. The model labeled $\mathbf{s}_3$ is a scaled version of $\mathbf{s}(\lambda^*)$, scaled so that $\mathbf{s}_3$ is on the boundary of the feasible region ($\mathcal{F}^*$). The iteration sequence stops when the perimeter of the triangle formed by $\mathbf{s}_1$, $\mathbf{s}_2$ and $\mathbf{s}_3$ drops below a prescribed threshold.

This algorithm has been tested on several problems both with real and with synthetic data and compared with a traditional damped least-squares algorithm (i.e., setting $\lambda^* = 1$ on each iteration). The new algorithm was found to be very stable and avoids the large oscillations in slowness often found in traditional least-squares methods.

Example 7.3.1 *Reconstructions were performed on models having* 16×8 *cells using 320 rays —including 256 rays (16 sources* $\times$ *16 receivers) from left to right and 64 rays (8 sources* $\times$ *8 receivers) from top to bottom. This measurement configuration was chosen to minimize the effects of ghosts, since the main focus of the exercise is to evaluate the usefulness of the feasibility constraints in stabilizing the algorithm.*

The traveltime data were generated with a bending method using the simplex search routine [Prothero et al., 1988; Nelder and Mead, 1965].

Three examples of typical results from the nonlinear algorithm BST (Borehole Seismic Tomography) are displayed on the following pages. The basic structure of the test problems has this form:



These slowness models have a low speed anomaly on top and a high speed anomaly on the bottom in each case. The first example has 20% *anomalies; the second has* 50% *anomalies; the third* 100% *anomalies.* EXAMPLE 7.3.1.1A *shows the target model, i.e., the model that used to generate the traveltime data;* EXAMPLE 7.3.1.1B *shows the reconstructed values using bent rays and feasibility constraints. The other two examples are presented similarly, progressing from smaller anomalies to larger ones.*

The reconstructions were found to converge after 15 or 20 iterations, and did not vary significantly if the iteration sequence was continued. In general, we expect slow anomalies to be harder to reconstruct than fast anomalies, because rays tend to avoid the slow regions in favor of fast regions. Thus, the coverage of slow anomalies tends to be much less than for fast anomalies, and therefore the resolution of such regions tends to poorer than for the fast regions. This effect is observed in all the examples. The reconstructions for 20% *and* 50% *contrasts were quite good, while that for* 100% *was noticeably worse than the other two. The main reason for this difference is that in the presence of high contrasts the ray paths tend to seek out the fastest regions; thus, even the background in the vicinity of a very fast anomaly can become poorly resolved, since all the rays close to the fast anomaly go through it and therefore do not sample the surrounding region well. To improve the resolution of the reconstruction in the presence of high contrasts requires a substantial increase in the density of ray-path sampling.*

1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.20	1.20	1.00	1.00	1.00
1.00	1.00	1.20	1.20	1.20	1.20	1.00	1.00
1.00	1.00	1.00	1.20	1.20	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	0.83	0.83	1.00	1.00	1.00
1.00	1.00	0.83	0.83	0.83	0.83	1.00	1.00
1.00	1.00	1.00	0.83	0.83	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

EXAMPLE 7.3.1.1A. The double-cross slowness model with 20% contrast.

0.98	0.97	1.02	1.06	1.04	1.00	0.98	0.96
1.02	1.00	0.98	1.01	1.02	1.00	1.01	1.02
1.00	1.01	0.98	0.98	0.98	0.98	1.02	1.05
1.01	1.00	1.03	1.15	1.14	1.04	1.02	1.03
0.99	0.99	1.12	1.20	1.18	1.14	1.02	1.02
1.00	1.00	1.00	1.15	1.18	1.06	1.00	1.02
0.98	1.01	1.00	1.02	1.03	0.96	1.02	1.01
1.00	1.01	1.00	1.01	1.00	0.99	1.01	1.01
1.00	1.00	1.00	1.01	1.00	1.01	1.00	0.99
0.99	1.00	0.99	0.98	0.99	1.01	1.00	0.99
0.99	0.99	0.99	0.89	0.87	0.95	1.04	0.97
1.00	0.98	0.89	0.85	0.84	0.87	0.94	0.97
1.03	0.96	0.95	0.87	0.85	0.95	0.94	0.99
1.04	1.00	0.99	0.97	0.98	0.95	1.02	1.00
1.01	1.01	1.03	0.98	0.97	1.00	1.01	0.98
1.00	1.04	0.97	0.98	0.98	0.99	1.01	1.02

EXAMPLE 7.3.1.1B. Reconstruction of the double-cross slowness model with 20% contrast using the BST code with feasibility constraints and noisy data (after 41 iterations).

1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.50	1.50	1.00	1.00	1.00
1.00	1.00	1.50	1.50	1.50	1.50	1.00	1.00
1.00	1.00	1.00	1.50	1.50	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	0.67	0.67	1.00	1.00	1.00
1.00	1.00	0.67	0.67	0.67	0.67	1.00	1.00
1.00	1.00	1.00	0.67	0.67	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

EXAMPLE 7.3.1.2A. The double-cross slowness model with 50% contrast.

1.04	0.94	0.97	1.20	1.08	1.02	0.99	0.94
1.05	1.05	0.92	1.06	1.02	0.98	1.00	1.06
1.04	0.93	1.07	0.99	1.04	0.90	0.96	1.14
1.02	1.01	0.99	1.10	1.23	1.33	1.04	1.02
0.95	1.00	1.42	2.20	1.28	1.23	1.05	1.04
0.97	1.03	0.96	1.21	1.21	1.10	1.07	1.05
0.99	1.06	0.98	1.08	0.98	0.92	1.11	1.03
0.97	1.08	0.98	1.01	0.95	1.03	1.03	1.04
0.99	1.00	1.01	1.04	0.95	1.05	1.03	1.01
0.98	1.00	0.94	0.96	0.91	1.04	1.00	0.94
0.97	1.03	0.93	0.81	0.73	0.90	1.06	0.95
1.00	0.98	0.75	0.68	0.68	0.74	0.88	0.96
1.09	0.91	0.88	0.78	0.71	0.92	0.93	0.94
1.06	1.02	0.92	1.05	0.93	0.93	1.02	0.96
1.01	1.03	1.00	0.98	0.97	0.98	1.04	0.95
0.98	1.14	0.92	0.94	0.96	0.96	1.06	1.01

EXAMPLE 7.3.1.2B. Reconstruction of the double-cross slowness model with 50% contrast using the BST code with feasibility constraints and noisy data (after 41 iterations).

1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	2.00	2.00	1.00	1.00	1.00
1.00	1.00	2.00	2.00	2.00	2.00	1.00	1.00
1.00	1.00	1.00	2.00	2.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	0.50	0.50	1.00	1.00	1.00
1.00	1.00	0.50	0.50	0.50	0.50	1.00	1.00
1.00	1.00	1.00	0.50	0.50	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

EXAMPLE 7.3.1.3A. The double-cross slowness model with 100% contrast.

1.04	0.98	1.03	1.04	1.18	1.07	1.00	1.02
1.12	1.05	0.84	1.00	1.10	1.01	1.04	1.16
1.08	0.94	1.01	1.00	1.02	0.81	0.93	1.39
1.05	0.99	1.00	1.09	1.38	1.53	1.08	1.19
0.90	1.00	1.70	2.38	1.46	1.23	1.04	1.12
0.91	1.09	0.97	1.28	1.40	1.10	1.02	1.11
1.08	1.16	0.99	1.17	0.93	0.86	1.15	1.09
1.07	1.13	0.99	0.98	0.91	0.97	1.05	1.09
1.09	0.97	1.00	1.02	0.94	1.03	0.98	1.04
0.99	0.95	0.94	0.86	0.88	1.02	1.00	0.94
0.99	1.03	0.89	0.66	0.62	0.82	1.08	0.90
1.08	0.86	0.66	0.56	0.57	0.62	0.80	0.94
1.05	0.83	0.81	0.69	0.62	0.77	0.84	0.84
1.12	1.00	1.01	0.91	0.81	0.90	0.93	1.02
1.01	0.96	1.04	0.97	0.86	0.99	0.97	0.96
0.97	1.12	0.99	0.95	0.92	0.90	1.08	1.01

EXAMPLE 7.3.1.3B. Reconstruction of the double-cross slowness model with 100% contrast using the BST code with feasibility constraints and noisy data (after 41 iterations).

7.4 Using Relative Traveltimes

When we do not have control over the seismic source location and timing as in the case of earthquakes, the absolute traveltimes are not accurately known and it is important to understand how relative traveltimes may be used in seismic tomography [Aki, Christoffersson, and Husebye, 1977].

Rigorous application of the feasibility constraints $\mathbf{Ms} \geq \mathbf{t}$ requires fairly accurate knowledge of the absolute traveltimes. When such information is sparse or unavailable, we can use the information known about gross geological structure of the region to estimate the mean traveltime. Then we remove the physically meaningless mean of the relative data T/m and add back in the geological mean τ_0.

The *remove-the-mean operator* $\mathbf{R}$ for an m-dimensional vector space is defined as

$$\mathbf{R} = \mathbf{I} - \mathbf{u}\frac{1}{m}\mathbf{u}^T, \tag{7.33}$$

where $\mathbf{u}^T = (1, \ldots, 1)$ is an m-vector of ones. Note that $\mathbf{RR} = \mathbf{R}$ so $\mathbf{R}$ is a projection operator. Then, we see that $\mathbf{R}$ applied to the traveltime vector $\mathbf{t}$ gives

$$\mathbf{Rt} = \mathbf{t} - \frac{T}{m}\mathbf{u}, \tag{7.34}$$

where $T/m = \mathbf{u}^T\mathbf{t}/m$ is the mean traveltime of the data set. Applying $\mathbf{R}$ to the ray-path matrix, we have

$$\mathbf{RM} = \mathbf{M} - \mathbf{u}\frac{T}{m}\mathbf{v}^T\mathbf{C} = \mathbf{M} - \mathbf{u}\frac{T}{m}\mathbf{c}^T. \tag{7.35}$$

The standard procedure for this problem is to solve for $\mathbf{s}$ in the equation

$$\mathbf{M}'\mathbf{s} = \mathbf{t}', \tag{7.36}$$

where $\mathbf{M}' = \mathbf{RM}$ and $\mathbf{t}' = \mathbf{Rt}$. To apply the feasibility constraints, we must modify the problem to

$$\mathbf{Ms} \geq \mathbf{Rt} + \tau_0\mathbf{I}_m. \tag{7.37}$$

Hidden in this analysis is the fact that the earthquake sources are often far from the region to be imaged, so the "effective" source locations may be placed at the boundaries of the region to be imaged.

If we have predetermined the mean for the traveltime data, then it is clearly desirable to use an inversion procedure that preserves this mean, *i.e.*, choosing $\Delta\mathbf{s}$ so that

$$\frac{\mathbf{u}^T\mathbf{M}(\mathbf{s} + \Delta\mathbf{s})}{m} = \tau_0 \tag{7.38}$$

for all $\Delta\mathbf{s}$. Preserving the mean is equivalent to preserving the total traveltime along all ray paths, so

$$\mathbf{c}^T(\mathbf{s} + \Delta\mathbf{s}) = m\tau_0. \tag{7.39}$$

In other words, vary $\mathbf{s}$ so it stays in the hyperplane determined by (7.39). But we have studied exactly this mathematical problem using linear programming in (7.6) and also using weighted least-squares in (7.25). So we do not need to develop any new inversion methods for this special case.

7.5 Parallel Computation

Traveltime inversion algorithms tend to be parallelizable in a variety of ways. The use of the feasibility constraints only increases the degree of parallelism that is achievable by these algorithms.

First, the *forward modeling* may be parallelized. If the forward problem is solved using either shooting or bending methods, then it is straightforward to parallelize the code because each ray may be computed independently of the others, and therefore in parallel. If the forward problem is solved using a finite difference algorithm or a full wave equation method, then whether the algorithm is parallelizable or not depends on the details of the particular algorithm. For example, Vidale's method is not parallelizable, but another related method by van Trier and Symes [1991] is parallelizable.

Second, the use of the feasibility constraints in inversion algorithms suggests that it might be advantageous to map the feasibility boundary and then use the information gained to search for improved agreement between the model and the data. Mapping the feasibility boundary can be done completely in parallel. Each model $\mathbf{s}$ may be treated in isolation, computing the best ray-path matrix for that model, and then finding the scaled model in the direction of $\mathbf{s}$ that intersects the feasibility boundary. The difficulty with this method is that it requires a figure of merit (in real problems) to help us determine whether (and to what degree) one point on the feasibility boundary is better than another. In ideal circumstances (no data error and infinite precision in our computers), the figure of merit would be the number of ray paths that achieve equality while still satisfying all the feasibility constraints

$$\mathbf{Ms} \geq \mathbf{t}. \tag{7.40}$$

When that number equals the number of ray paths, we have found an exact solution and, as the number increases towards this maximum value during an iterative procedure, the trial models $\mathbf{s}$ must be converging towards this solution. But in real problems, a figure of merit based on the number of equalities in (7.40) is not useful.

In a series of numerical experiments [joint work with A. J. DeGroot], we have found that a useful figure of merit for real problems is the nonlinear least-squares functional

$$\Psi(\mathbf{s}) = \sum_{i=1}^{n} w_i[\tau_i^*(\mathbf{s}) - \mathbf{t}_i]^2. \tag{7.41}$$

If we have found an exact solution $\mathbf{s}^*$ to the inversion problem, (7.41) will vanish at that point on the feasibility boundary. As we approach this global minimum, (7.41) is evaluated at an arbitrary point on the feasibility boundary and the values in a cluster of such points are compared, our analysis of convex programming shows that the points with the smallest values of (7.41) form a convex set. The smallest value we find may not actually vanish, in which case there is no exact solution to our inversion problem. This procedure has been implemented on a parallel processing machine, and the results obtained using this algorithm with the figure of merit (7.41) are comparable to those of the stable algorithm discussed earlier.

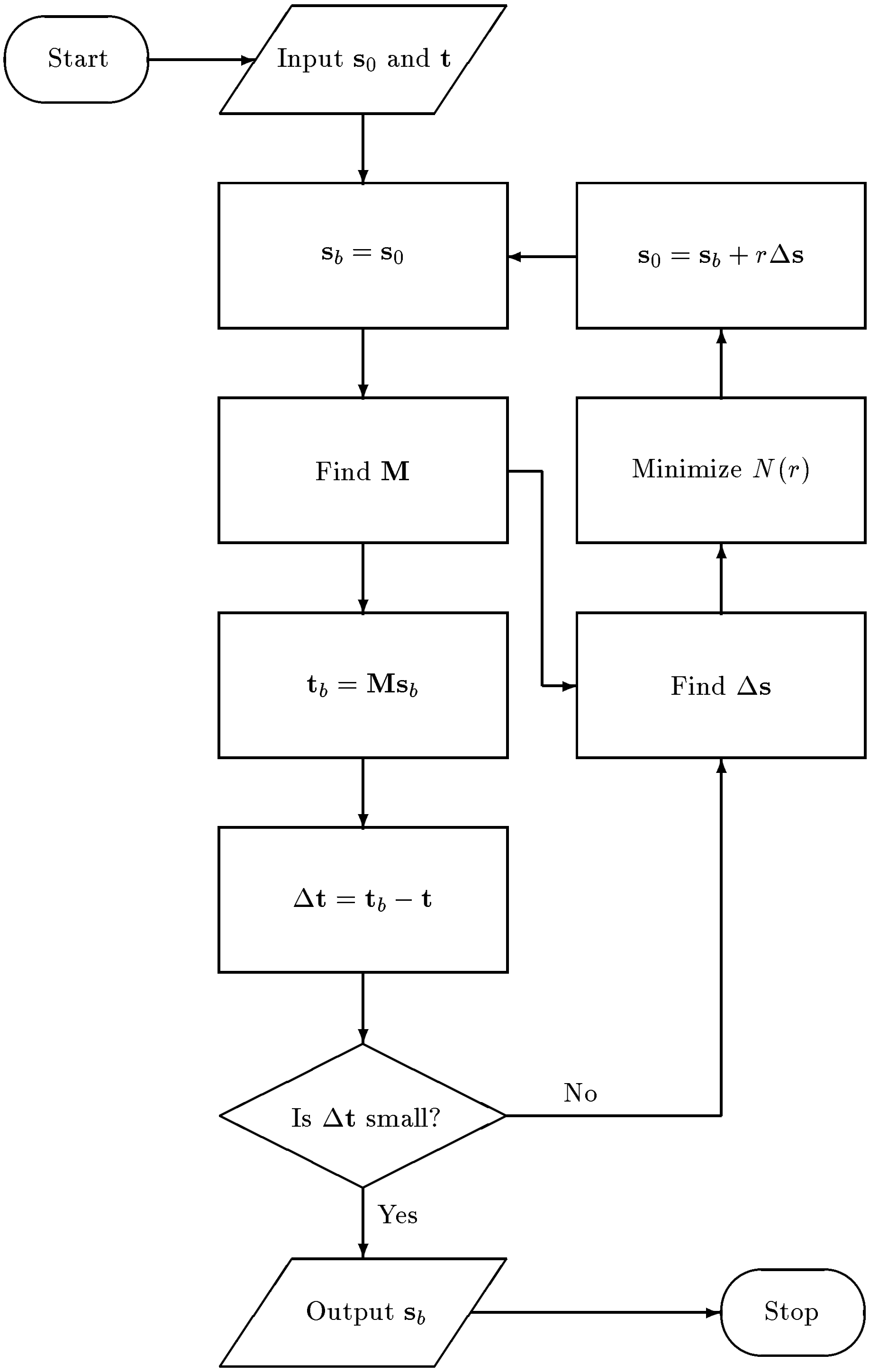


Figure 7.2: Modified iterative algorithm for traveltime inversion using feasibility constraints.

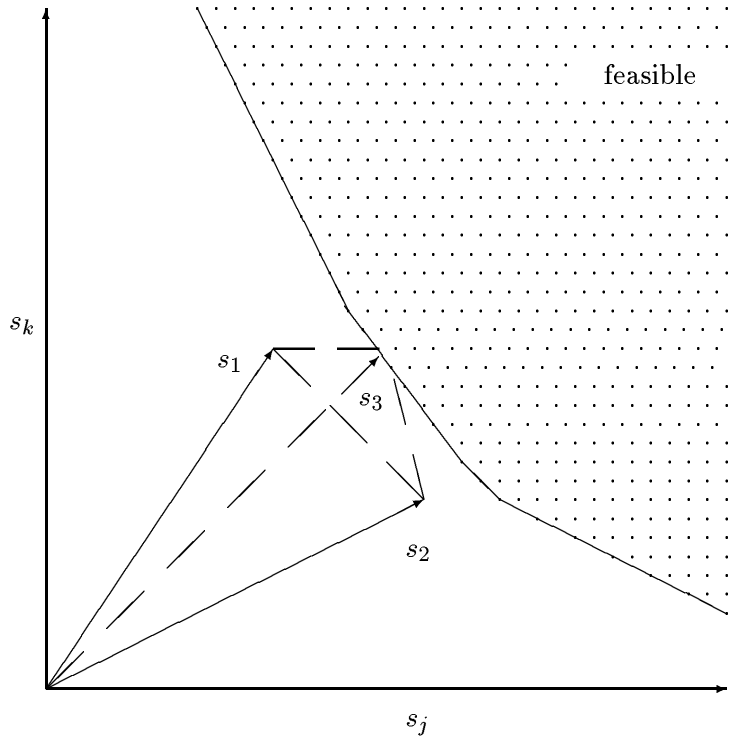


Figure 7.3: Snapshot of one iteration in a nonlinear inversion algorithm based on feasibility constraints.

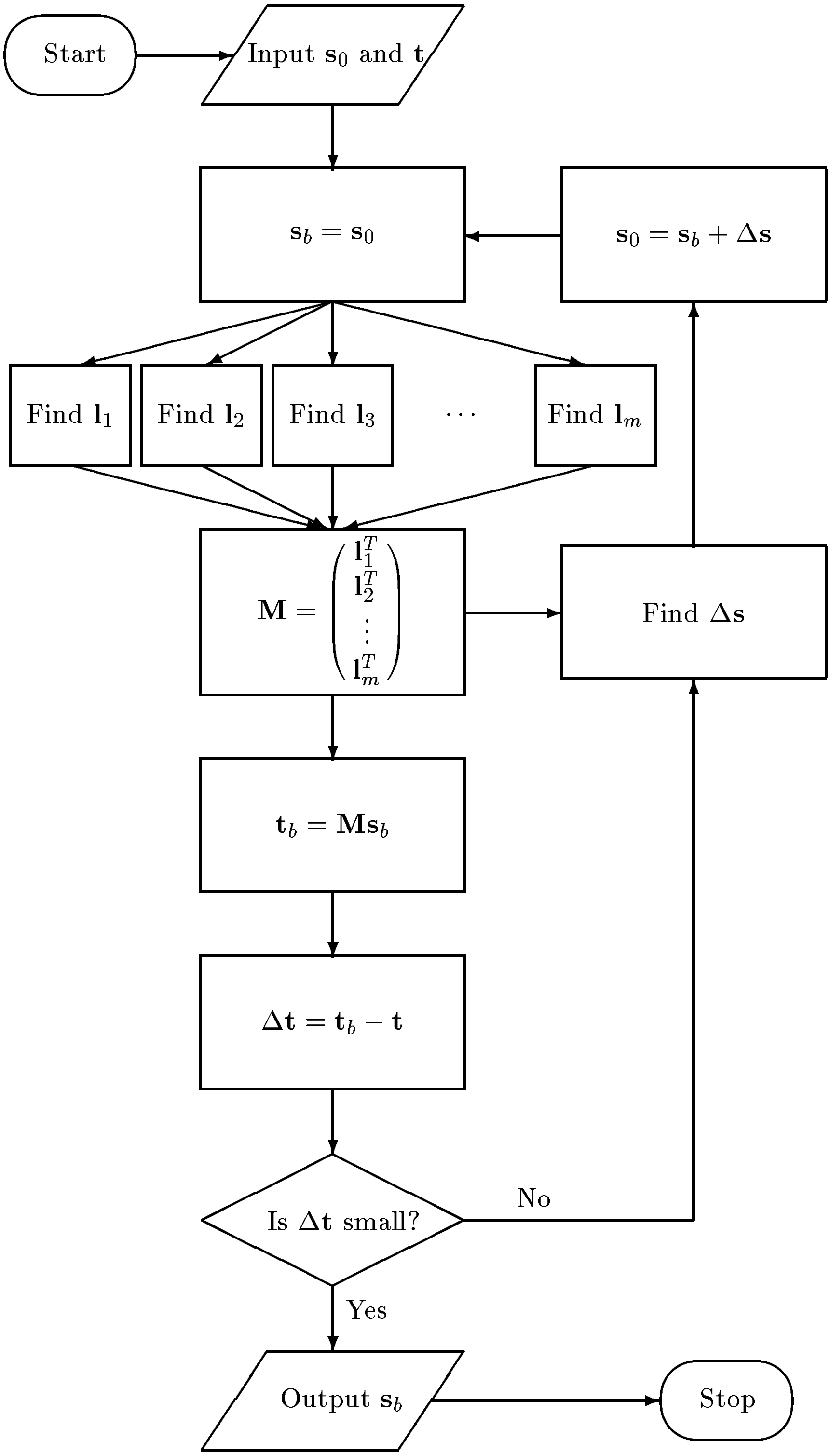


Figure 7.4: Computing ray paths for different source/receiver paairs in parallel.

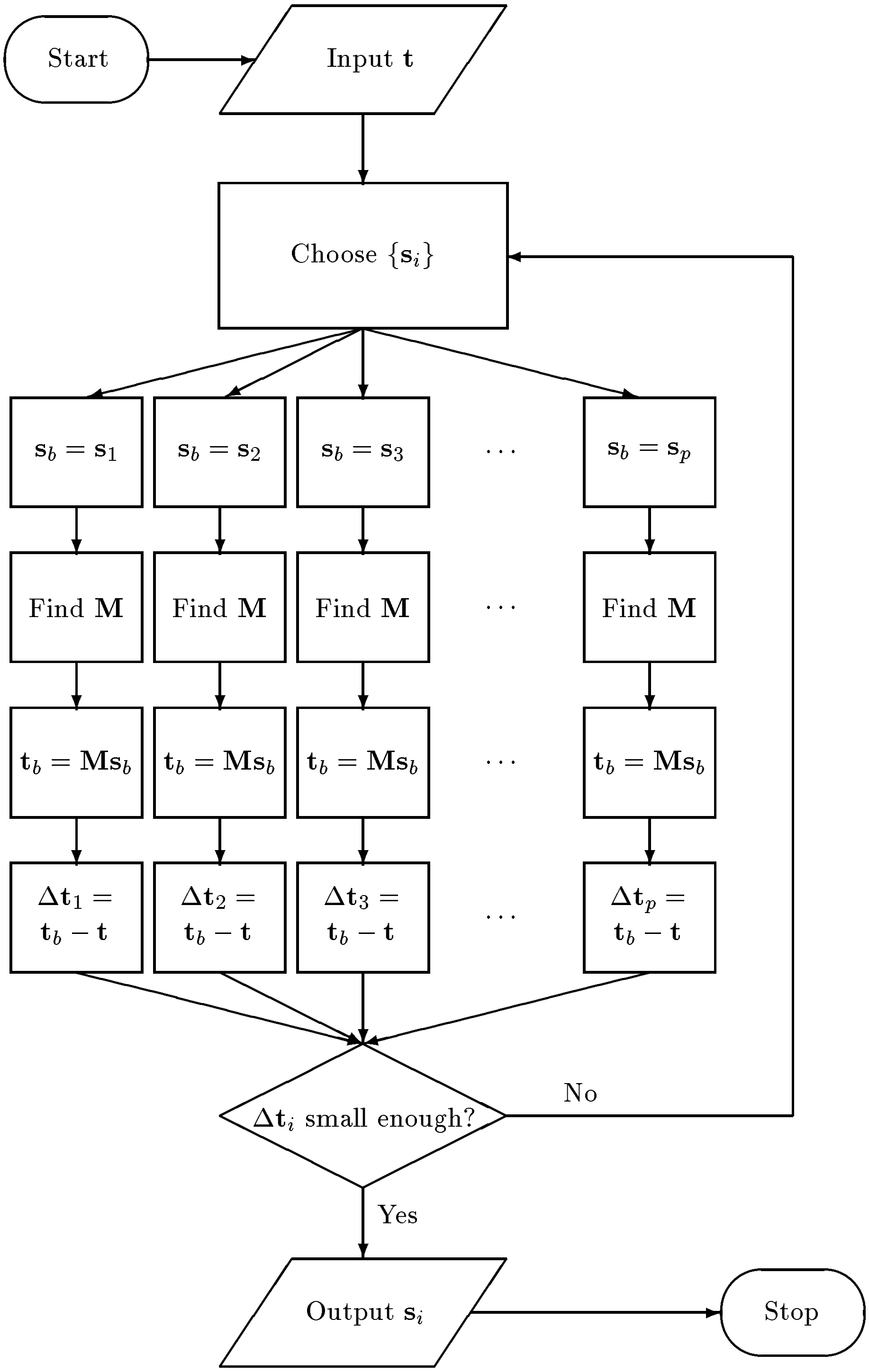


Figure 7.5: Monte Carlo method of mapping feasibility boundary.

Chapter 8

Other Nonlinear Inversion Problems

Although traveltime inversion has been the main thrust of these lecture notes, I want to make it clear that the ideas involving the feasibility constraints are very general. In fact, they apply to any inversion problem where the data are the minima of one of the variational problems of mathematical physics.

So in this final section of these notes, I present two other inversion problems that lead to convex feasible sets and then show the general structure needed to guarantee convex global feasibility. Finally, I present another example that leads to a nonconvex feasibility set and discuss the consequences of this difference for computing the solution to the inverse problem.

8.1 Electrical Impedance Tomography

Electrical impedance tomography [Dines and Lytle, 1981; Berryman and Kohn, 1990] attempts to image the electrical impedance (or just the conductivity) distribution inside a body using electrical measurements on its boundary. See Fig. 8.1. The method has been used successfully in both biomedical [Barber and Brown, 1986] and geophysical applications [Wexler, Fry, and Neuman, 1985; Daily, Lin, and Buscheck, 1987; Daily and Owen, 1991], but the analysis of optimal reconstruction algorithms is still progressing [Yorkey, Webster, and Tompkins, 1987; Kohn and McKenney, 1990]. The most common application is monitoring the influx or efflux of a conducting fluid (such as brine in a porous rock) through the body whose conductivity is being imaged. This method does not have high resolving power like radiological methods, but it is comparatively inexpensive and it therefore provides a valuable alternative when continuous monitoring is desired.

First, we review some facts about this problem that play an important role in the analysis that follows. Recall that the power dissipated into heat is [Jackson, 1962]

$$P = \int \mathbf{J}(\mathbf{x}) \cdot \mathbf{E}(\mathbf{x})\, d^3x, \tag{8.1}$$

Figure 8.1: Experimental setup for electrical impedance tomography.

where

$$\mathbf{J}(\mathbf{x}) = \sigma(\mathbf{x})\mathbf{E}(\mathbf{x}), \tag{8.2}$$

$$\mathbf{E}(\mathbf{x}) = -\nabla\Phi(\mathbf{x}), \tag{8.3}$$

and the current distribution satisfies

$$\nabla \cdot \mathbf{J}(\mathbf{x}) = 0 \tag{8.4}$$

away from all current sources. The quantities displayed are the current distribution $\mathbf{J}$, the isotropic conductivity σ, the electric field $\mathbf{E}$, and the potential Φ. Substituting (8.2) and (8.3) into (8.4) gives Poisson's equation

$$\nabla \cdot (\sigma\nabla\Phi) = 0. \tag{8.5}$$

Substituting (8.3) into (8.1) and using (8.4), we have

$$P = -\int \mathbf{J} \cdot \nabla\Phi \, d^3x = -\int \nabla \cdot (\Phi\mathbf{J}) \, d^3x. \tag{8.6}$$

Then, the divergence theorem shows that

$$P = -\int \Phi \mathbf{J} \cdot \hat{n}\, da, \tag{8.7}$$

where $\hat{n}$ is a unit outward normal vector and da is the infinitesimal surface area on the boundary. If current is injected through metallic electrodes, the potential takes a constant value Φ_k on the kth electrode of surface area a_k. If there are K electrodes, then (8.7) becomes

$$P = \sum_{k=1}^{K} \Phi_k I_k, \tag{8.8}$$

where

$$I_k = -\int_{a_k} \mathbf{J} \cdot \hat{n}\, da \tag{8.9}$$

is the total current injected ($I_k > 0$) or withdrawn ($I_k < 0$) at the kth electrode. Since these are the only sources and sinks, we also have the sumrule

$$\sum_{k=1}^{K} I_k = 0. \tag{8.10}$$

If there are only two injection electrodes, then (8.8) reduces to

$$P = (\Phi_1 - \Phi_2) I_1 = \Delta\Phi I, \tag{8.11}$$

so the power is the product of the measured potential difference $\Delta\Phi$ across the injection electrodes and the injected current I.

The data for electrical impedance tomography have most often been gathered by injecting a measured current between two electrodes while simultaneously measuring the voltage differences between pairs of other electrodes placed around the boundary of the body being imaged. This process is then repeated, injecting current between all possible (generally adjacent) pairs of electrodes, and recording the set of voltage differences for each injection pair i. This data set has normally not included the voltage difference across the injection electrodes, because these voltages cannot be measured as reliably. A substantial contact impedance develops at the interface between the body and the injection electrodes when large currents are present. This problem can be reduced by using large electrodes or small currents. In this lecture, we will assume that voltage differences (and therefore the powers dissipated) across the injection electrodes are known, but it is not necessary that they be known to high accuracy.

Dirichlet's principle [Courant, 1950; Courant and Hilbert, 1953] states that, given a conductivity distribution $\sigma(\mathbf{x})$ and a potential distribution $\Phi(\mathbf{x})$, the power dissipation p_i realized for the ith current injection configuration is the one that minimizes the integral $\int \sigma |\nabla\Phi|^2\, d^3x$ so that

$$p_i(\sigma) = \int \sigma(\mathbf{x}) |\nabla \Phi_i^*(\mathbf{x})|^2\, d^3x = \min_{\Phi_i} \int \sigma(\mathbf{x}) |\nabla \Phi_i(\mathbf{x})|^2\, d^3x. \tag{8.12}$$

The trial potential field for the ith injection pair is $\Phi_i(\mathbf{x})$, while the particular potential field that actually minimizes the the power is $\Phi_i^*(\mathbf{x})$, and this one also satisfies Poisson's equation $\nabla \cdot (\sigma \nabla \Phi_i^*) = 0$ within the body. Furthermore, if the effective power dissipation associated with the trial potential $\Phi_i(\mathbf{x})$ is defined as

$$\bar{p}_i^{(\Phi_i)}(\sigma) \equiv \int \sigma(\mathbf{x}) |\nabla \Phi_i(\mathbf{x})|^2 \, d^3x, \tag{8.13}$$

then the measured powers P_i must satisfy

$$P_i = p_i(\sigma^*) \leq \bar{p}_i^{(\Phi_i)}(\sigma^*), \tag{8.14}$$

if $\sigma^*(\mathbf{x})$ is the true conductivity distribution. Note that if we vary the trial power dissipation (8.13) with respect to the trial potential, we find

$$2 \int \sigma \nabla \Phi \cdot \nabla \delta \Phi \, d^3x = -2 \int \nabla \cdot (\sigma \nabla \Phi) \delta \Phi \, d^3x = 0 \tag{8.15}$$

at a stationary point. We integrated once by parts to obtain (8.15). Since the volume variation $\delta\Phi$ is arbitrary, its coefficient inside the integral must vanish, so we just recover Poisson's equation, as expected.

Now we begin to see the analogy developing between the seismic traveltime tomography problem and the electrical impedance tomography problem. If we consider the following set of correspondences:

$$s(\mathbf{x}) \rightarrow \sigma(\mathbf{x}),$$

$$t_i(s) \rightarrow p_i(\sigma),$$

$$\tau_i^P(s) \rightarrow \bar{p}_i^{(\Phi_i)}(\sigma),$$

$$T_i \rightarrow P_i,$$

$$dl_i^P \rightarrow |\nabla \Phi_i(\mathbf{x})|^2 \, d^3x,$$

$$dl_i^{P^*} \rightarrow |\nabla \Phi_i^*(\mathbf{x})|^2 \, d^3x,$$

then we see that the analysis of convex functionals and feasibility sets presented for seismic traveltime tomography carries over directly to the electrical impedance tomography problem when it is formulated this way. For example, the scale invariance property holds for electrical impedance tomography, so multiplying σ by a scalar γ does not change the optimum potential distribution.

The feasibility constraints for electrical impedance tomography now take the form

$$\boxed{\mathbf{K}\hat{\sigma} \quad \geq \quad \mathbf{p},} \tag{8.16}$$

where $\hat{\sigma}^T = (\sigma_1, \ldots, \sigma_n)$, $\mathbf{p}^T = (p_1, \ldots, p_m)$, and the E-squared matrix is given by

$$K_{ij} = \int_{\text{cell}_j} |\nabla \Phi_i|^2 \, d^3x. \tag{8.17}$$

Least-squares methods may be applied to this problem in much the same fashion as in traveltime tomography [Kallman and Berryman, 1992].

A thorough analysis of the electrical impedance tomography problem would require another set of lectures. Lucky for you, I will not try to present them here. However, to excite your curiosity, I will mention another feature of the electrical impedance tomography problem not shared by the seismic tomography problem. So far we have discussed only Dirichlet's principle (8.12). In fact, there are two distinct variational principles for the conductivity problem: Dirichlet's principle and its dual, known as Thomson's principle [Thomson, 1848; Thomson, 1884; Maxwell, 1891; Courant and Hilbert, 1953] The second variational principle takes the form

$$P_i \le \int |\mathbf{J}_i(\mathbf{x})|^2/\sigma(\mathbf{x}) \, d^3x, \tag{8.18}$$

where $\mathbf{J}_i(\mathbf{x})$ is a trial current distribution vector for the ith current injection pair that satisfies the continuity equation $\nabla \cdot \mathbf{J}_i = 0$. The trial current distribution $\mathbf{J}_i(\mathbf{x})$ and the trial gradient of the potential $\nabla\Phi_i(\mathbf{x})$ are generally unrelated except that, when the minimum of both variational functionals is attained, then $\mathbf{J}_i^*(\mathbf{x}) = -\sigma\nabla\Phi_i^*(\mathbf{x})$. Then, of course, the current equals the conductivity times the electric field.

The existence of dual variational principles is a general property whenever the primal variational principle is a true minimum principle. Fermat's principle is only a stationary (not a minimum) principle, and so traveltime tomography does not possess this dual property. (If we attempt to formulate a dual for Fermat's principle as we did in the lecture on linear and nonlinear programming, we find the content of the dual results are essentially trivial.) The existence of the dual variational principles for electrical impedance tomography is important because it means that there are two independent sets of feasibility constraints for the conductivity model $\sigma(\mathbf{x})$. Furthermore, as illustrated in Fig. 8.2, these two sets of constraints allow us (in some sense) to obtain upper and lower bounds on the region of the conductivity model space that contains the solution to the inversion problem. See Berryman and Kohn [1990] for more discussion of this point.

8.2 Inverse Eigenvalue Problems

Inverse eigenvalue problems arise in the earth sciences during attempts to deduce earth structure from knowledge of the modes of vibration of the earth [Dahlen, 1968; Wiggins, 1972; Jordan and Anderson, 1974; Hald, 1980; Hald, 1983; Anderson and Dziewonski, 1984; McLaughlin, 1986; Dziewonski and Woodhouse, 1987; Lay, Ahrens, Olson, Smyth, and Loper, 1990; Snieder, 1993].

Consider the typical forward problem associated with the inverse eigenvalue problem

$$-\nabla^2 u(\mathbf{x}) + q(\mathbf{x})u(\mathbf{x}) = \lambda u(\mathbf{x}) \tag{8.19}$$

Figure 8.2: Dirichlet's principle and Thomson's principle provide upper and lower bounds on the dual feasibility region for electrical impedance tomography.

on a finite domain with some boundary conditions on u. This is known as a Sturm-Liouville equation to mathematicians and as the Schroedinger equation to physicists. In quantum mechanics, the time-independent wave function is given by $u(\mathbf{x})$ and $q(\mathbf{x})$ is the potential. The eigenvalue is λ.

Now it is well-known that a Rayleigh-Ritz procedure may be used to approximate the eigenvalues λ [Courant and Hilbert, 1953]. In particular, the lowest eigenvalue is given in general by

$$\lambda_0 = \min_u \frac{\int (|\nabla u|^2 + qu^2)\, d^3x}{\int u^2\, d^3x}, \tag{8.20}$$

where admissible us satisfy the boundary conditions of (8.19) and have no other constraints, except being twice differentiable. The ratio on the right to be minimized is known as the Rayleigh quotient, and the denominator $\int u^2\, d^3x$ serves to normalize the wave function u.

Define the Rayleigh-quotient functional as

$$\Lambda(q, u_i) = \frac{\int (|\nabla u_i|^2 + qu_i^2)\, d^3x}{\int u_i^2\, d^3x}, \tag{8.21}$$

where u_i is a trial wave function subject to i constraints. Taking the variation of Λ with respect to u_i, we find that the stationary points of Λ satisfy

$$\frac{\int[-\nabla^2 u_i + qu_i - \Lambda(q,u_i)u_i]\delta u_i\, d^3x}{\int u_i^2\, d^3x} = 0. \tag{8.22}$$

We integrated once by parts to obtain (8.22) using the fact that the variations of δu vanish on the boundary. So, since the variations δu within the domain may be arbitrary, the term in brackets must vanish and the stationary points of the Rayleigh quotient therefore occur for u_is that satisfy (8.19) with $\lambda_i = \Lambda(q, u_i)$.

Clearly, we may define feasibility constraints for this problem in a manner analogous to that for the traveltime tomography problem and for the electrical impedance tomography problem. If the eigenvalues λ_i are our data, then for the correct potential q^* we must have

$$\lambda_i \equiv \Lambda(q^*, u_i^*[q^*]) \le \Lambda(q^*, u_i), \tag{8.23}$$

where $u_i^*[q]$ is the eigenfunction associated with eigenvalue λ_i of the potential q. Thus, feasible qs satisfy

$$\boxed{\lambda_i \quad \le \quad \Lambda(q, u_i)} \tag{8.24}$$

for all admissible u_is.

To show that this problem leads to a convex feasibility set, consider two potentials that satisfy the feasibility constraints for some fixed choice of u_i. Then,

$$\lambda_i \le \Lambda(q_1, u_i) \quad \text{and} \quad \lambda_i \le \Lambda(q_2, u_i) \tag{8.25}$$

and

$$\lambda_i \le \epsilon\Lambda(q_1, u_i) + (1-\epsilon)\Lambda(q_2, u_i) \tag{8.26}$$

$$= \frac{\int(|\nabla u_i|^2 + [\epsilon q_1 + (1-\epsilon)q_2]u_i^2)\, d^3x}{\int u_i^2\, d^3x} \tag{8.27}$$

$$= \Lambda(q_\epsilon, u_i), \tag{8.28}$$

where the convex combination $q_\epsilon \equiv \epsilon q_1 + (1-\epsilon)q_2$. Thus, local (fixed u_is) feasibility follows simply from the linearity of the Rayleigh quotient (except for the shift at the origin) with respect to the potential q. Global feasibility follows from the variational properties of Λ with respect to u_i. (See the next section for the proof.)

Note that there is no scale invariance property for Λ similar to the one for the traveltime functional. However, it is true that wave functions are invariant to a constant shift in the potential, since it is easy to see that

$$\Lambda(q+\gamma, u) = \Lambda(q, u) + \gamma. \tag{8.29}$$

In our analysis, we can also make use of other members of the invariance group of (8.19) [Ames, 1972].

This inverse eigenvalue problem can be reformulated in terms of a different set of variational functionals. In particular, one such set of generalized Rayleigh-Ritz quotients has been constructed by Berryman [1988]; however, these functionals have a more complicated dependence on the potential q. Without linearity or shifted linearity in q, we cannot prove the convexity of the feasibility set and the structure of the inversion problem becomes less certain and possibly more complex.

8.3 General Structure for Convex Inversion Problems

The feasibility analysis presented in these lectures applies to a wide class of inverse problems that can be formulated so the data are minima of an appropriate variational problem. To see the general structure, consider a set of functionals $\Gamma_i(q,u)$ of two variables q and u. Then, if each functional is linear in one variable so that

$$\Gamma_i(aq_1 + bq_2, u) = a\Gamma_i(q_1,u) + b\Gamma_i(q_2,u), \tag{8.30}$$

and if the data γ_i bound $\Gamma_i(q_1,u)$ and $\Gamma_i(q_2,u)$ from below for any second argument u, then

$$\gamma_i \le \Gamma_i(q_1,u) \quad \text{and} \quad \gamma_i \le \Gamma_i(q_2,u) \quad \text{for all} \quad i = 1,\ldots,m, \tag{8.31}$$

and we have

$$\gamma_i \le \lambda\Gamma_i(q_1,u) + (1-\lambda)\Gamma_i(q_2,u) = \Gamma_i(\lambda q_1 + (1-\lambda)q_2, u). \tag{8.32}$$

Therefore, Γ_i evaluated at the convex combination $q_\lambda = \lambda q_1 + (1-\lambda)q_2$ is also bounded below by the data. Thus, linearity for fixed u is sufficient to prove that feasible qs for the linear problem form a convex set. We call this the *local convex feasibility* property.

Then, when we consider variations of the second argument and assume that the data are minima of the variational functional over all possible us, we have

$$\gamma_i \equiv \Gamma_i(q^*, u^*[q^*]) \le \Gamma_i(q^*, u), \tag{8.33}$$

where $u^*[q]$ is the particular function that minimizes the the functional Γ_i when q is the first argument. Then, we have

$$\gamma_i \le \Gamma_i(q_1, u^*[q_1]) \le \Gamma_i(q_1, u^*[\cdot]), \tag{8.34}$$
$$\gamma_i \le \Gamma_i(q_2, u^*[q_2]) \le \Gamma_i(q_2, u^*[\cdot]), \tag{8.35}$$

where $u^*[\cdot]$ is the correct (minimizing) u for some yet to be specified q. Combining (8.34) and (8.35) using the linearity property of Γ_i for its first argument, we have

$$\begin{aligned} \gamma_i &\le \lambda\Gamma_i(q_1, u^*[q_1]) + (1-\lambda)\Gamma_i(q_2, u^*[q_2]) && (8.36)\\ &\le \lambda\Gamma_i(q_1, u^*[\cdot]) + (1-\lambda)\Gamma_i(q_2, u^*[\cdot]) && (8.37)\\ &= \Gamma_i(q_\lambda, u^*[\cdot]), && (8.38)\end{aligned}$$

where $q_\lambda = \lambda q_1 + (1-\lambda)q_2$ is again the convex combination of q_1 and q_2. Now we are free to choose the $\cdot$ to be any permissible q, so we choose it for convenience to be q_λ. Then, we have the final result that

$$\boxed{\gamma_i \quad \le \quad \Gamma_i(q_\lambda, u^*[q_\lambda]).} \tag{8.39}$$

The conclusion from (8.32) is that there are *local convex feasibility sets* and from (8.39) that there is a *global convex feasibility set* for the full nonlinear inversion problem, just as in the case for traveltime tomography.

The only properties used in the derivation were the linearity of the variational functional Γ_i for fixed u and the concavity of the functional that results from its variational nature.

The preceding proof is appropriate for Fermat's, Dirichlet's, and Thomson's principles. However, the proof must be modified for the inverse eigenvalue problem because the Rayleigh quotient is a shifted linear functional of the potential q. We can fix this minor difficulty by considering

$$\Delta\Lambda(q, u_i) = \Lambda(q, u_i) - \Lambda(0, u_i) = \frac{\int qu^2\, d^3x}{\int u^2\, d^3x}, \tag{8.40}$$

which is linear in q. If

$$\lambda_i \le \Lambda(0, u_i) + \Delta\Lambda(q_1, u_i), \tag{8.41}$$

$$\lambda_i \le \Lambda(0, u_i) + \Delta\Lambda(q_2, u_i), \tag{8.42}$$

then we carry through the analysis as before and conclude that

$$\lambda_i \le \Lambda(0, u_i) + \Delta\Lambda(\epsilon q_1 + (1-\epsilon)q_2, u_i) = \Lambda(q_\epsilon, u_i). \tag{8.43}$$

This proves the *local convex feasibility* property for problems with variational functionals linear in the first argument except for a constant. The proof of *global convex feasibility* follows the proof already presented step by step and will be left as an exercise.

PROBLEM

PROBLEM 8.3.1 *Prove global convex feasibility for the inverse eigenvalue problem.*

8.4 Nonconvex Inversion Problems with Feasibility Constraints

Although we expect the idea of using feasibility constraints in inversion problems with variational structure to be a very general method, it may not always be true that the variational functional is a concave functional of its arguments. If not, then the resulting nonlinear programming problem will not be convex.

As an example, consider the electrical impedance tomography problem again, but this time for complex (still isotropic) conductivity $\sigma = \sigma_R + i\sigma_I$. The dissipative part of σ is the real part σ_R, while the reactive part (proportional to the dielectric constant) is the imaginary part σ_I.

The current is proportional to the conductivity and the electric field, but now all quantities are complex so

$$\mathbf{J} = \sigma \mathbf{E} \tag{8.44}$$

becomes

$$\mathbf{j}_R + i\mathbf{j}_I = (\sigma_R + i\sigma_I)(\mathbf{e}_R + i\mathbf{e}_I). \tag{8.45}$$

The power dissipation for this problem is given by

$$\begin{aligned} P &= \frac{1}{2}\int (\mathbf{J}\cdot\mathbf{E}^* + \mathbf{J}^*\cdot\mathbf{E})\, d^3x && (8.46)\\ &= \int (\mathbf{j}_R\cdot\mathbf{e}_R + \mathbf{j}_I\cdot\mathbf{e}_I)\, d^3x && (8.47)\\ &= \int \sigma_R(\mathbf{e}_R\cdot\mathbf{e}_R + \mathbf{e}_I\cdot\mathbf{e}_I)\, d^3x. && (8.48) \end{aligned}$$

Rewriting (8.45) in matrix notation we have

$$\begin{pmatrix}\mathbf{j}_R\\ \mathbf{j}_I\end{pmatrix} = \begin{pmatrix}\sigma_R & -\sigma_I\\ \sigma_I & \sigma_R\end{pmatrix}\begin{pmatrix}\mathbf{e}_R\\ \mathbf{e}_I\end{pmatrix}. \tag{8.49}$$

Now we want to reformulate this problem as a variational principle in order to apply the ideas of feasibility constraints, but to do so we need a positive scalar functional. The power dissipation is a good choice again, but (8.49) is inconvenient for this purpose since the matrix is not positive definite [Milton, 1990; Cherkaev and Gibiansky, 1994]. Performing a Legendre transform on (8.49), we find that an alternative equation is

$$\begin{pmatrix}\mathbf{j}_R\\ \mathbf{e}_I\end{pmatrix} = \begin{pmatrix}\sigma_R + \frac{\sigma_I^2}{\sigma_R} & -\frac{\sigma_I}{\sigma_R}\\ -\frac{\sigma_I}{\sigma_R} & \frac{1}{\sigma_R}\end{pmatrix}\begin{pmatrix}\mathbf{e}_R\\ \mathbf{j}_I\end{pmatrix} \equiv \mathbf{\Sigma}\begin{pmatrix}\mathbf{e}_R\\ \mathbf{j}_I\end{pmatrix}. \tag{8.50}$$

Then, the matrix $\mathbf{\Sigma}$ is positive definite (for $\sigma_R > 0$), since

$$\mathbf{\Sigma}\begin{pmatrix}\mathbf{e}_R\\ \mathbf{j}_I\end{pmatrix} = \lambda\begin{pmatrix}\sigma_R & 0\\ 0 & 1/\sigma_R\end{pmatrix}\begin{pmatrix}\mathbf{e}_R\\ \mathbf{j}_I\end{pmatrix} \tag{8.51}$$

implies that

$$\lambda + \frac{1}{\lambda} = 2 + \frac{\sigma_I^2}{\sigma_R^2}, \tag{8.52}$$

which guarantees that the eigenvalues λ and $1/\lambda$ are positive.

So now the power is given by

$$P = \int (\mathbf{j}_R \cdot \mathbf{e}_R + \mathbf{j}_I \cdot \mathbf{e}_I)\, d^3x \tag{8.53}$$

$$= \int (\,\mathbf{e}_R \quad \mathbf{j}_I\,)\, \mathbf{\Sigma} \begin{pmatrix} \mathbf{e}_R \\ \mathbf{j}_I \end{pmatrix} d^3x \tag{8.54}$$

$$= \int [\sigma_R |\mathbf{e}_R|^2 + \frac{1}{\sigma_R} |\mathbf{j}_I - \sigma_I \mathbf{e}_R|^2]\, d^3x. \tag{8.55}$$

This is the final expression for the power. In this form, we have a valid variational principle. Also, note that the term $\mathbf{j}_I - \sigma_I \mathbf{e}_R = \sigma_R \mathbf{e}_I$ so the second term in the final expression for P is just $\sigma_R |\mathbf{e}_I|^2$.

To check the conditions for stationarity of this integral, we find that, if we vary with respect to $\mathbf{e}_R$, then

$$2 \int [\sigma_R \mathbf{e}_R - \frac{\sigma_I}{\sigma_R} (\mathbf{j}_I - \sigma_I \mathbf{e}_R)] \cdot \delta \mathbf{e}_R\, d^3x = 0. \tag{8.56}$$

If we vary with respect to $\mathbf{j}_I$, we find that

$$2 \int [\frac{1}{\sigma_R} (\mathbf{j}_I - \sigma_I \mathbf{e}_R)] \cdot \delta \mathbf{j}_I\, d^3x = 0. \tag{8.57}$$

Since the electric field is the gradient of a potential, (8.56) implies that

$$\nabla \cdot [\sigma_R \mathbf{e}_R - \frac{\sigma_I}{\sigma_R} (\mathbf{j}_I - \sigma_I \mathbf{e}_R)] = 0. \tag{8.58}$$

Similarly, since the current distribution is divergence free, (8.57) implies that

$$\frac{1}{\sigma_R} (\mathbf{j}_I - \sigma_I \mathbf{e}_R) = -\nabla \Phi \tag{8.59}$$

for some scalar potential function Φ. Thus, the expression in (8.59) acts like an electric field (in fact, it is $\mathbf{e}_I$) at the stationary point, while the quantity whose divergence is zero in (8.58) acts like a current distribution (in fact, it is $\mathbf{j}_R$). This completes the proof that (8.55) is a legitimate variational principle for the complex conductivity problem.

We can still talk about feasibility constraints for this problem, since

$$P_i \equiv \bar{p}_i(\sigma_R^*, \sigma_I^*, \mathbf{e}_R^*, \mathbf{j}_I^*) \leq \bar{p}_i(\sigma_R^*, \sigma_I^*, \mathbf{e}_R, \mathbf{j}_I) \tag{8.60}$$

with the trial power dissipation given by

$$\bar{p}_i(\sigma_R, \sigma_I, \mathbf{e}_R, \mathbf{j}_I) = \int [\sigma_R |\mathbf{e}_R|^2 + \frac{1}{\sigma_R} |\mathbf{j}_I - \sigma_I \mathbf{e}_R|^2]\, d^3x. \tag{8.61}$$

The starred quantities in (8.60) are the true ones for the experimental configuration. If we can find σs that violate the constraints implied by (8.60), then those σs are infeasible and the rest form the feasible set. However, $\bar{p}$ is not linear in its dependence on σ, so we cannot prove that this functional is concave. [1] Therefore, we lack a proof of the convexity of the feasible set.

For fixed σ_I, $\mathbf{j}_I$, and $\mathbf{e}_R$, the minimum of (8.61) is achieved, for a model of constant conductivity cells, when the real conductivity in the jth cell is given by

$$\sigma_R^2 = \frac{\int_{\text{cell}_j} |\mathbf{j}_I - \sigma_I \mathbf{e}_R|^2\, d^3x}{\int_{\text{cell}_j} |\mathbf{e}_R|^2\, d^3x}. \tag{8.62}$$

This minimum value is

$$\min_{\sigma_R} \bar{p}_i = 2\sum_{j=1}^{n} [\int_{\text{cell}_j} |\mathbf{e}_R|^2\, d^3x \int_{\text{cell}_j} |\mathbf{j}_I - \sigma_I \mathbf{e}_R|^2\, d^3x]^{\frac{1}{2}}. \tag{8.63}$$

Since the imaginary part of the conductivity may still be viewed as a variable, we can further minimize (8.63) by finding the minimum with respect to σ_I. This minimum occurs when

$$\sigma_I = \frac{\int_{\text{cell}_j} \mathbf{j}_I \cdot \mathbf{e}_R\, d^3x}{\int_{\text{cell}_j} \mathbf{e}_R \cdot \mathbf{e}_R\, d^3x} \tag{8.64}$$

for the imaginary part of the conductivity in the jth cell. Substituting into (8.63), we have the minimum power

$$\min_{\sigma_R,\sigma_I} \bar{p}_i = 2\sum_{j=1}^{n} [\int_{\text{cell}_j} \mathbf{e}_R \cdot \mathbf{e}_R\, d^3x \int_{\text{cell}_j} \mathbf{j}_I \cdot \mathbf{j}_I\, d^3x - (\int_{\text{cell}_j} \mathbf{j}_I \cdot \mathbf{e}_R\, d^3x)^2]^{\frac{1}{2}}. \tag{8.65}$$

It follows from the Schwartz inequality for integrals that

$$(\int \mathbf{a} \cdot \mathbf{b}\, d^3x)^2 \leq \int \mathbf{a} \cdot \mathbf{a}\, d^3x \int \mathbf{b} \cdot \mathbf{b}\, d^3x \tag{8.66}$$

with equality applying *only* when $\mathbf{b}$ is proportional to $\mathbf{a}$, that each bracket in (8.65) is positive unless there is an exact solution such that

$$\mathbf{j}_I = \gamma \mathbf{e}_R, \tag{8.67}$$

for some scalar γ.

If the nonlinear programming problem is nonconvex but feasibility constraints are still applicable, what are the consequences for numerical solution of the inversion problem? For convex feasibility sets, the convex combination of any two points on the feasibility boundary is also feasible and therefore either lies in the interior or on the boundary of the feasible set. This property implies a certain degree of smoothness for the boundary itself. Clearly, if the

[1] Looking at (8.54) we see that the power is a linear functional of the matrix elements of $\mathbf{\Sigma}$. However, this apparent linearity unfortunately does not help the analysis, because a physical constraint on the matrix elements is that $\det \mathbf{\Sigma} \equiv 1$. It is not difficult to show that the convex combination of two matrices with unit determinant does not preserve this property. So the nonlinearity cannot be avoided by the trick of considering convex combinations of the matrix elements.

feasible set is nonconvex, then the convex combination of two points on the boundary may or may not lie in the feasible set; thus, the boundary itself may be jagged. Since the solution of the inversion problem still lies on the boundary (just as it did in the convex case), the lack of smoothness of the boundary may have important computational consequences: the boundary is still expected to be continuous, of course, but sharp local jumps could occur that might make convergence of an iterative method difficult to achieve.

As an iterative scheme progresses, the absolute minimum of the trial power (8.65) decreases towards zero. Thus, the feasibility constraints become *more* important for this problem as the scheme progresses to convergence.

Bibliography

Cited References

Aki, K., A. Christoffersson, and E. S. Husebye, 1976, Determination of the three-dimensional seismic structure of the lithosphere, *J. Geophys. Res.* **82**, 277–296.

Ames, W. F., 1972, *Nonlinear Partial Differential Equations in Engineering, Vol. II,* Academic Press, New York, Chapter 2, 87–145.

Anderson, D. L., and A. M. Dziewonski, 1984, Seismic tomography, *Scientific American* **251**, *number 10*, 60–68.

Backus, G., and F. Gilbert, 1968, The resolving power of gross earth data, *Geophys. J. R. Astr. Soc.* **16**, 169–205.

Backus, G., and F. Gilbert, 1970, Uniqueness in the inversion of inaccurate gross earth data, *Philos. Trans. R. Soc. London* **266A**, 123–192.

Barnett, S., 1990, *Matrices — Methods and Applications*, Clarendon Press, Oxford, Chapters 6 and 10.

Barber, D. C., and B. H. Brown, 1986, Recent developments in applied potential tomography — APT, in *Information Processing in Medical Imaging*, S. L. Bascarach (ed.), Martinus Nijhoff, Dordrecht, 106–121.

Berryman, J. G., 1988, Bounds on decay constants for diffusion through inhomogeneous media, *J. Phys. A: Math. Gen.* **21**, 4423–4441.

Berryman, J. G., 1989a, Weighted least-squares criteria for seismic traveltime tomography, *IEEE Trans. Geosci. Remote Sensing* **27**, 302–309.

Berryman, J. G., 1989b, Fermat's principle and nonlinear traveltime tomography, *Phys. Rev. Lett.* **62**, 2953–2956.

Berryman, J. G., 1990, Stable iterative reconstruction algorithm for nonlinear traveltime tomography, *Inverse Problems* **6**, 21–42.

Berryman, J. G., 1991, Convexity properties of inverse problems with variational constraints, *J. Franklin Inst.* **328**, 1–13.

Berryman, J. G. and R. V. Kohn, 1990, Variational constraints for electrical impedance tomography, *Phys. Rev. Lett.* **65**, 325–328.

Bleistein, N., 1984, *Mathematical Methods for Wave Phenomena*, Academic Press, New York, p. 18.

Bloomfield, P., 1976, *Fourier Analysis of Time Series: An Introduction*, Wiley, New York, 26–27.

Boorse, H. A. and L. Motz, 1966, The principle of least action, *The World of the Atom*, Basic Books, New York, vol. II, Chapter 62.

Born, M., 1926, Quantenmechanik der Strossvorgänge (Quantum mechanics of impact processes), *Z. Phys.* **38**, 803–827.

Born, M., and E. Wolf, 1980, *Principles of Optics* – Electromagnetic Theory of Propagation, Interference, and Diffraction of Light, Pergamon Press, Oxford, p. 453 (Born approximation); pp. xxi-xii, 112, 128-130, 719, 732, 740, 742 (Fermat's principle); pp. 112, 724-725 (Hamilton-Jacobi theory); pp. 110, 119 (Rytov); p. xxi (Snell).

Bruns, H., 1895, *Abh. Kgl. Sächs. Ges. Wiss., math-phys. Kl.*, **21**, 323.

Burkhard, N. R., 1980, Resolution and error of the back projection technique algorithm for geophysical tomography, Lawrence Livermore National Laboratory preprint, UCRL-52984.

Cherkaev, A. V., and L. V. Gibiansky, Variational principles for complex conductivity, viscoelasticity, and similar problems in media with complex moduli, *J. Math. Phys.* **35**, 127–145, 1994.

Claerbout, J. F., and F. Muir, 1973, Robust modeling with erratic data, *Geophysics* **38**, 826–844.

Courant, R., 1950, *Dirichlet's Principle, Conformal Mapping, and Minimal Surfaces*, Interscience, New York.

Courant, R., and D. Hilbert, 1953, *Methods of Mathematical Physics*, Vol. 1, Wiley, New York, pp. 240–242 (Dirichlet's principle); pp. 132–134 (Rayleigh-Ritz); pp. 267–268 (Thomson's principle).

Dahlen, F. A., 1968, The normal modes of a rotating elliptical earth, *Geophys. J. R. Astron. Soc.* **16**, 329–367.

Daily, W., W. Lin, and T. Buscheck, 1987, Hydrological properties of Topopah Spring tuff: Laboratory measurements, *J. Geophys. Res.* **92**, 7854–7864.

Daily, W., and E. Owen, 1991, Cross-borehole resistivity tomography, *Geophysics* **56**, 1228–1235.

Devaney, A. J., 1984, Geophysical diffraction tomography, *IEEE Trans. Geosci. Remote Sensing* **22**, 3–13.

Dines, K. A., and R. J. Lytle, 1979, Computerized geophysical tomography, *Proc. IEEE* **67**, 1065–1073.

Dines, K. A., and R. J. Lytle, 1981, Analysis of electrical conductivity imaging, *Geophysics* **46**, 1025–1036.

Dziewonski, A. M., and J. H. Woodhouse, 1987, Global images of the earth's interior, *Science* **236**, 37–40.

Fermat, P. de, 1891, *Oeuvres de Fermat*, Paris, Vol. 2, p. 354.

Feynman, R. P., 1965, *The Character of Physical Law*, MIT Press, Cambridge, Massachusetts, p. 156.

Feynman, R. P., R. B. Leighton, and M. Sands, 1963, *The Feynman Lectures on Physics*, Addison-Wesley, Reading, Massachusetts, Vol. I, Chapter 26 and Vol. II, Chapter 19.

Fiacco, A. V. and G. P. McCormick, 1990, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, SIAM, Philadelphia, Chapter 6, 86–112.

Frank, M., and C. A. Balanis, 1989, Methods for improving the stability of electromagnetic geophysical inversions, *IEEE Trans. Geosci. Remote Sens.* **27**, 339–343.

Franklin, J. N., 1970, Well-posed stochastic extensions of ill-posed linear problems, *J. Math. Anal. Appl.* **31**, 682–716.

Freund, R. W., and N. M. Nachtigal, 1991, QMR: a quasi-minimal residual method for non-Hermitian linear systems, *Numer. Math.* **60**, 315–339.

Gilbert, P., 1972, Iterative methods for the three-dimensional reconstruction of an object from projections, *J. Theor. Biol.* **36**, 105–117.

Glasko, V. B., 1988, *Inverse Problems of Mathematical Physics*, American Institute of Physics, New York.

Gleick, J., 1992, *Genius: The Life and Science of Richard Feynman*, Vintage, New York, pp. 57–61.

Goldstein, H., 1950, *Classical Mechanics*, Addison-Wesley, Reading, Massachusetts, pp. 231, 312.

Gordon, R., 1974, A tutorial on ART (Algebraic Reconstruction Techniques), *IEEE Trans. Nucl. Sci.* **NS-21**, 78–93.

Gordon, R., R. Bender, and G. T. Herman, 1970, Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography, *J. Theor. Biol.* **29**, 471–481.

Guenther, R. B., C. W. Kerber, E. K. Killian, K. T. Smith, and S. L. Wagner, 1974, Reconstruction of objects from radiographs and the location of brain tumors, *Proc. Nat. Acad. Sci.* **71**, 4884–4886.

Hald, O. H., 1980, Inverse eigenvalue problems for the mantle, *Geophys. J. R. Astron. Soc.* **62**, 41–48.

Hald, O. H., 1983, Inverse eigenvalue problems for the mantle — II, *Geophys. J. R. Astron. Soc.* **72**, 139–164.

Hardy, G. H., J. E. Littlewood, and G. Pólya, 1934, *Inequalities*, Cambridge University Press, Cambridge, 70–101.

Harris, J. M., 1987, Diffraction tomography with arrays of discrete sources and receivers, *IEEE Trans. Geosci. Remote Sensing* **GE-25**, 448–455.

Herman, G. T., 1980, *Image Reconstruction from Projections – The Fundamentals of Computerized Tomography*, Academic, New York, Chapter 6, 100–107.

Herman, G. T., A. Lent, and S. Rowland, 1973, ART: Mathematics and applications. A report on the mathematical foundations and on the applicability to real data of the Algebraic Reconstruction Techniques, *J. Theor. Biol.* **42**, 1–32.

Hestenes, M. R., and E. Stiefel, 1952, Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Stan.* **B 49**, 409–436.

Huygens, Ch., 1690, *Traité de la lumière*, Leyden.

Ivansson, S., 1983, Remark on an earlier iterative tomographic algorithm, *Geophys. J. R. Astr. Soc.* **75**, 855–860.

Ivansson, S., 1986, Seismic borehole tomography — Theory and computational methods, *Proc. IEEE* **74**, 328–338.

Jackson, D. D., 1972, Interpretation of inaccurate, insufficient and inconsistent data, *Geophys. J. R. Astr. Soc.* **28**, 97–109.

Jackson, J. D., 1962, *Classical Electrodynamics*, Wiley, New York, 189–190.

Jech, J., and I. Pšenčík, 1989, First-order perturbation method for anisotropic media, *Geophys. J. Int.* **99**, 369–376.

Jech, J., and I. Pšenčík, 1991, Kinematic inversion for qP and qS waves in inhomogeneous hexagonally symmetric structures, *Geophys. J. Int.*, in press.

Jeffrey, W., and R. Rosner, 1986a, On strategies for inverting remote sensing data, *Astrophys. J.* **310**, 463–472.

Jeffrey, W., and R. Rosner, 1986b, Optimization algorithms: Simulated annealing and neural network processing, *Astrophys. J.* **310**, 473–481.

Jordan, T. H., and D. L. Anderson, 1974, Earth structure from free oscillations and travel times, *Geophys. J. R. Astr. Soc.* **36**, 411–459.

Jordan, T. H., and J. N. Franklin, 1971, Optimal solutions to a linear inverse problem in geophysics, *Proc. Nat. Acad. Sci.* **68**, 291–293.

Kaczmarz, S., 1937, Angenäherte Auflösung von Systemen linearer Gleichungen, *Bull. Acad. Polon. Sci. Lett.* **A**, 355–357.

Kak, A. C., 1984, Image reconstruction from projections, in *Digital Image Processing Techniques*, M. P. Ekstrom (ed.), Academic, New York, Chapter 4, 111–170.

Kallman, J. S., and J. G. Berryman, 1992, Weighted least-squares criteria for electrical impedance tomography, *IEEE Trans. Med. Imaging* **11**, 284–292.

Keller, J. B., 1969, Accuracy and validity of the Born and Rytov approximations, *J. Opt. Soc. Am.* **59**, 1003–1004.

Kohn, R. V., and A. McKenney, 1990, Numerical implementation of a variational method for electrical impedance tomography, *Inverse Problems* **6**, 389–414.

Ladas, K. T., and A. J. Devaney, 1991, Generalized ART algorithm for diffraction tomography, *Inverse Problems* **7**, 109–125.

Ladas, K. T., and A. J. Devaney, 1992, Iterative methods in geophysical diffraction tomography, *Inverse Problems* **8**, 119–132.

Lanczos, C., 1961, *Linear Differential Operators*, Van Nostrand, New York, Chapter 3, pp. 100–162.

Lanczos, C., 1970, *The Variational Principles of Mechanics*, Dover, New York, Chapter V, pp. 111–160.

Lay, T., T. J. Ahrens, P. Olson, J. Smyth, and D. Loper, 1990, Studies of the earth's deep interior: Goals and trends, *Phys. Today* **43**, *number 10*, 44–52.

Levenberg, K., 1944, A method for the solution of certain non-linear problems in least squares, *Quart. Appl. Math.* **2**, 164–168.

Lo, T.-W., G. L. Duckworth, and M. N. Toksöz, 1990, Minimum cross entropy seismic diffraction tomography, *J. Acoust. Soc. Am.* **87**, 748–756.

Lu, S.-Y., and J. G. Berryman, 1991, Inverse scattering, seismic traveltime tomography, and neural networks, *Intern. J. Imaging Sys. Tech.* **3**, to appear February, 1991.

Luneburg, R. K., 1964, *Mathematical Theory of Optics*, University of California Press, Berkeley.

Lytle, R. J., and K. A. Dines, 1980, Iterative ray tracing between boreholes for underground image reconstruction, *IEEE Trans. Geosci. Remote Sens.* **18**, 234–240.

Marquardt, D. W., 1963, An algorithm for least-squares estimation of nonlinear parameters, *SIAM J. Appl. Math.* **11**, 431–441.

Marquardt, D. W., 1970, Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation, *Technometrics* **12**, 591–612.

Maxwell, J. C., 1891, *A Treatise on Electricity and Magnetism*, Dover, New York, Vol. I, pp. 138–142, 423–424.

McLaughlin, J. R., 1986, Analytical methods for recovering coefficients in differential equations from spectral data, *SIAM Rev.* **28**, 53–72.

Michelena, R. J., and J. M. Harris, 1991, Tomographic traveltime inversion using natural pixels, *Geophysics* **56**, 635–644.

Milton, G. W., 1990, On characterizing the set of possible effective tensors of composites: The variational method and the translation method, *Commun. Pure Appl. Math.* **43**, 63–125 (see Section 16).

Moore, E. H., 1920, *Bull. Amer. Math. Soc.* **26**, 394–395.

Natterer, F., 1986, *The Mathematics of Computerized Tomography*, Wiley, New York, Chapter V.

Nelder, J. A., and R. Mead, 1965, A simplex method for function minimization, *Computer J.* **7**, 308–313.

Newton, R. G., 1966, *Scattering Theory of Waves and Particles*, McGraw-Hill, New York, pp. 233–246.

Parlett, B. N., D. R. Taylor, and Z. A. Liu, 1985, A look-ahead Lanczos algorithm for unsymmetric matrices, *Math. Comput.* **44**, 105–124.

Penrose, R., 1955a, A generalized inverse for matrices, *Proc. Cambridge Philos. Soc.* **51**, 406–413.

Penrose, R., 1955b, On best approximation solutions of linear matrix equations, *Proc. Cambridge Philos. Soc.* **52**, 17–19.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1988, *Numerical Recipes in C* – The Art of Scientific Computing, Cambridge Univeristy Press, Cambridge, pp. 292, 305–309.

Prothero, W. A., W. J. Taylor, and J. A. Eickemeyer, 1988, A fast, two-point, three-dimensional raytracing algorithm using a simple step search method, *Bull. Seismol. Soc. Am.* **78**, 1190–1198.

Rao, C. R., and S. K. Mitra, 1971, *Generalized Inverse of Matrices and Its Applications*, Wiley, New York.

Rytov, S. M., 1937, *Izv. Akad. Nauk SSSR* **2**, 223.

Rytov, S. M., 1938, *Compt. Rend. (Doklady) Acad. Sci., URSS* **18**, 263.

Scales, J. A., A. Gersztenkorn, and S. Treitel, 1988, Fast l_p solution of large, sparse, linear systems: Application to seismic travel time tomography, *J. Comput. Phys.* **75**, 314–333.

Smith, M. L., and J. N. Franklin, 1969, Geophysical application of generalized inverse theory, *J. Geophys. Res.* **74**, 2783–2785.

Snieder, R., 1993, Global inversions using normal modes and long-period surface waves, in *Seismic Tomography: Theory and Practice*, H. M. Iyer and K. Hirahara (eds.), Chapman & Hall, London, pp. 23–63.

Strang, G., 1986, *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, MA, Chapter 8, 665–734.

Tabbara, W., B. Duchêne, Ch. Pichot, D. Lesselier, L. Chommeloux, and N. Joachimowicz, 1988, Diffraction tomography: Contribution to the analysis of some applications in microwaves and ultrasonics, *Inverse Problems* **4**, 305–331.

Tanabe, K., 1971, Projection method for solving a singular system of linear equations and its applications, *Numer. Math.* **17**, 203–214.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation, *Geophysics* **49**, 1259–1266.

Tarantola, A., and B. Valette, 1982, Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys. Space Phys.* **20**, 219–232.

Thomson, W. (Lord Kelvin), 1848, *Cambridge and Dublin Mathematical Journal*, Feb.

Thomson, W. (Lord Kelvin), 1884, *Mathematical and Physical Papers*, Cambridge University Press, Cambridge, Vol. I.

van Trier, J., and W. W. Symes, 1991, Upwind finite-difference calculation of traveltimes, *Geophysics* **56**, 812–821.

Varga, R. S., 1962, *Matrix Iterative Analysis*, Prentice-Hall, Englewood, Cliffs, NJ, pp. 16–17, 141, 159.

Vidale, J. E., 1988, Finite-difference calculation of travel time, *Bull. Seismol. Soc. Am.* **78**, 2062–2076.

Vidale, J. E., 1990, Finite-difference calculation of travel time in 3-D, *Geophysics* **55**, 521–526.

Wexler, A., B. Fry, and M. R. Neuman, 1985, Impedance-computed tomography algorithm and system, *Appl. Opt.* **24**, 3985–3992.

Whitham, G. B., 1974, *Linear and Nolinear Waves,* Wiley, New York, Chapters 7, 11 and 14.

Wiggins, R. A., 1972, The general linear inverse problem: Implications of surface waves and free oscillations for Earth structure, *Rev. Geophys. Space Phys.* **10**, 251–285.

Wu, R. S., and M. N. Toksöz, 1987, Diffraction tomography and multi-source holography applied to seismic imaging, *Geophysics* **52**, 11–25.

Yorkey, T. J., J. G. Webster, and W. J. Tompkins, 1987, Comparing reconstruction algorithms for electrical impedance tomography, *IEEE Trans. Biomed. Engng.* **34**, 843–852.

General References

Aki, K., and P. G. Richards, *Quantitative Seismology: Theory and Methods,* Vol. II, Freeman, New York, Chapter 12.

Ammon, C. J., G. E. Randall, and G. Zandt, 1990, On the nonuniqueness of receiver function inversions, *J. Geophys. Res.* **95**, 15303–15318.

Anderson, D. L., 1989, *Theory of the Earth*, Blackwell, Boston.

Bates, R. H. T., V. A. Smith, and R. D. Murch, 1991, Manageable multidimensional inverse scattering theory, *Phys. Repts.* **201**, 185–277.

Beylkin, G., 1984, The inversion problem and applications of the generalized Radon transform, *Commun. Pure Appl. Math.* **37**, 579–599.

Bois, P., M. La Porte, M. Lavergne, and G. Thomas, 1972, Well-to-well seismic measurements, *Geophysics* **37**, 471–480.

Bording, R. P., A. Gersztenkorn, L. R. Lines, J. A. Scales, and S. Treitel, 1987, Applications of seismic travel-time tomography, *Geophys. J. R. Astr. Soc.* **90**, 285–303.

Chen, S. T., L. J. Zimmerman, and J. K. Tugnait, 1990, Subsurface imaging using reversed vertical seismic profiling and crosshole tomographic methods, *Geophysics* **55**, 1478–1487.

Claerbout, J. F., 1976, *Fundamentals of Geophysical Data Processing: With Applications to Petroleum Prospecting*, McGraw-Hill, New York.

Colton, D., R. Ewing, and W. Rundell (eds.), 1990, *Inverse Problems in Partial Differential Equations*, SIAM, Philadelphia.

Cottle, R. W., and C. E. Lemke (eds.), 1976, *Nonlinear Programming*, SIAM-AMS Proceedings, Volume IX, Am. Math. Soc., Providence, RI.

Daudt, C. R., L. W. Braile, R. L. Nowack, and C. S. Chiang, 1989, A comparison of finite-difference and Fourier method calculation of synthetic seismograms, *Bull. Seismol. Soc. Am.* **79**, 1210–1230.

Dudgeon, D. E., and R. M. Mersereau, 1984, *Multidimensional Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey.

Evans, J. R., and J. J. Zucca, 1988, Active high-resolution seismic tomography of compressional wave velocity and attenuation structure at Medicine Lake Volcano, Northern California Cascade Range, *J. Geophys. Res.* **93**, 15016–15036.

Ewing, W. M., W. S. Jardetzky, and F. Press, 1957, *Elastic Waves in Layered Media,* McGraw-Hill, New York.

Gisser, D. G., D. Isaacson, and J. C. Newell, 1988, Theory and performance of an adaptive current tomography system, *Clin. Phys. Physiol. Meas.* **9A**, 35–41.

Goulty, N. R., 1993, Controlled-source tomography for mining and engineering applications, in *Seismic Tomography: Theory and Practice*, H. M. Iyer and K. Hirahara (eds.), Chapman and Hall, London, Chapter 29, pp. 797–813.

Grünbaum, F. A., 1980, A study of Fourier space methods for "limited angle" image reconstruction, *Numer. Funct. Anal. Optimiz.* **2**, 31-42.

Henderson, R. P., and J. G. Webster, 1978, An impedance camera for spatially specific measurements of the thorax, *IEEE Trans. Biomed. Engng.* **25**, 250–254.

Iyer, H. M., and K. Hirahara (eds.), 1993, *Seismic Tomography: Theory and Practice*, Chapman & Hall, London.

Justice, J. H., A. A. Vassiliou, S. Singh, J. D. Logel, P. A. Hansen, B. R. Hall, P. R. Hutt, and J. J. Solanski, 1989, Tomographic imaging in hydrocarbon reservoirs, *J. Imaging Sys. Tech.* **1**, 62–72.

Louis, A. K., 1981, Ghosts in tomography — The null space of the Radon transform, *Math. Meth. Appl. Sci.* **3**, 1-10.

Luenberger, D. G., 1969, *Optimization by Vector Space Methods*, Wiley, New York, Chapters 6 and 10, 160–168, 283–297.

Luenberger, D. G., 1973, *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading Massachusetts, Chapters 7 and 8, 148–155, 168–186.

McMechan, G. A., J. M. Harris, and L. M. Anderson, 1987, Crosshole tomography for strongly variable media with applications to scale model data, *Bull. Seismol. Soc. Am.* **77**, 1945–1960.

Mersereau, R. M., and A. V. Oppenheim, 1974, Digital reconstruction of multidimensional signals from their projections, *IEEE Proc.* **62**, 1319–1338.

Nelson, G. D., and J. E. Vidale, 1990, Earthquake locations by 3-D finite-difference travel times, *Bull. Seismol. Soc. Am.* **80**, 395–410.

Nolet, G. (ed.), 1987, *Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics,* Reidel, Dordrecht.

Smith, K. T., D. C. Solmon, and S. L. Wagner, 1977, Practical and mathematical aspects of the problem of reconstructing objects from radiographs, *Bull. Am. Math. Soc.* **83**, 1227–1270.

Tarantola, A., and A. Nercessian, 1984, Three-dimensional inversion without blocks, *Geophys. J. R. Astron. Soc.* **76**, 299–306.

Vidale, J. E., and H. Houston, 1990, Rapid calculation of seismic amplitudes, *Geophysics* **55**, 1504–1507.

Witten, A. J., and E. Long, 1986, Shallow applications of geophysical diffraction tomography, *IEEE Trans. Geosci. Remote Sens.* **24**, 654–662.

Zandt, G., 1981, Seismic images of the deep structure of the San Andreas fault system, Central Coast Ranges, California, *J. Geophys. Res.* **86**, 5039–5052.

Zucca, J. J., G. S. Fuis, B. Milkereit, W. D. Mooney, and R. D. Catchings, 1986, Crustal structure of Northeastern California, *J. Geophys. Res.* **91**, 7359–7382.

Glossary

Index

acoustic – pertaining to sound waves in a gas or fluid (such as air or water), generally limited to compressional waves.

backprojection – a one-step, approximate reconstruction method.

block – an element of a three-dimensional region whose properties are to be reconstructed. Usually, the properties are assumed to constant within the block.

cell – an element of a two-dimensional or three-dimensional region whose properties are to be reconstructed. Usually, the properties are assumed to be constant within the cell.

consistent – a system of equations with at least one solution satisfying all the physical constraints on a model.

determined – a linear system with as many equations as unknowns (assuming that the equations are linearly independent). If the equations are consistent, there is generally a unique physical solution to such a systeme.

elastic – pertaining to sound waves in a solid, and explicitly including both compressional and shear waves.

feasible – pertaining to a part of a set (especially the set of all possible models) that satisfies all known physical constraints, such as positivity. Any model that is not feasible is infeasible.

homogeneous – constant, that is a physical property constant on the scale of investigation.

image – a picture showing qualitative differences in a physical property of some region.

imaging – the process of producing an image.

inconsistent – a linear system with no physical (or feasible) solution. For example, the system

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

has the unique solution

$$\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix},$$

but this solution is unphysical because it fails to satisfy the positivity constraint. The cause of inconsistency is usually a major error in data collection, but more subtle interactions between forward modeling and the data can also produce inconsistency.

infeasible – pertaining to a part of a set (especially the set of all possible models) that fails to satisfy any of the physical constraints, such as positivity. This set is complementary to the feasible set.

inhomogeneous – not constant, that is a physical property varying on the scale of investigation.

inverse – the opposite rule. For example, subtraction is the opposite of addition, while division is the opposite of multiplication.

inversion – the process of reconstructing a two-dimensional image or three-dimensional map of some physical property in a selected region.

konoscope – a device for reconstructing the properties of a three-dimensional region using tomography or inversion.

map – a picture or volume representation often showing quantitative differences in a physical property of some region. A map is generally quantitative whereas an image is qualitative. A map is two- or three-dimensional whereas an image is two-dimensional.

migration – the process of reconstructing an image of earth reflectivity from seismic reflection data. Also known as **wave equation migration**.

nonfeasible – same as infeasible.

overdetermined – any linear system with more equations than unknowns. (Caveat: if many of the equations are linearly dependent, then the reduced system may actually be either determined or underdetermined; however, it generally requires much computation to decide if this is so.) Generally no exact solution to such a system exists, so approximate methods of solution such as least-squares are used to find "best" approximate solutions. Inverting for local averages of physical properties may produce an overdetermined mathematical inversion problem.

pixel – a picture element, or two-dimensional cell.

reconstruction – the act of constructing again from pieces that have been disassembled, as in a puzzle.

seismic – pertaining to sound waves in the earth, and explicitly including both compressional and shear waves.

seismogram – a record of seismic signals. Seismogram is to seismograph as photograph is to camera.

seismograph – a device for measuring seismograms.

seismography – the study or observation of seismic signals.

tomogram – the reconstructed image of some physical property produced by tomography. Tomogram is to konoscope as photograph is to camera, or as micrograph is to microscope.

tomograph – same as tomogram.

tomography – the study of cross sections; the process of reconstructing a two-dimensional image of some physical property of a selected plane region.

underdetermined – any linear system with fewer equations than unknowns. There are generally many solutions (often an infinite number) to such a system. Inhomogeneous physical systems whose properties may be described as essentially continuous functions of position may be considerd to have an infinite number of unknowns; therefore, any attempt to reconstruct the continuous system from finite data leads to an underdetermined physical inversion problem.

voxel – a volume element, or three-dimensional cell.

Author Index

Index

Subject Index

Index