Reconstruction of Slowness Distribution of a Medium between Two Boreholes from First Arrival Traveltime Data

DURSUN ÜSTÜNDAĞ,
The Department of Mathematics, Marmara University, Göztepe Campus, Istanbul, TURKEY.
dustundag@marmara.edu.tr

Abstract: We consider a problem of reconstruction of seismic wave speed distribution from a set of measured first arrival traveltimes in presence of strong velocity contrasts, which cause the problem to be highly non-linear. In this context, a stable iterative reconstruction algorithm, proposed by Berryman (1990) is studied. An attempt is therefore made to improve and also used it for a simple synthetic borehole test after writing its algorithm in a C++ code. The simulations results support the effectiveness of the method.

Keywords: Seismic tomography, Nonlinear Integral Equations, Inverse Problems, Nonlinear Programming and Convex Optimisation.

1 Introduction

Tomographic analyses have been applied in many fields, with tomography a term that has been used to represent a variety of analysis procedures, the most common being medical CAT (Computerized Axial Tomography) scanning (Herman 1980). In seismic tomography (Lee & Pereyra 1993), it refers to the measurements of arrival times of waves that pass through the subsurface medium. Variations in the arrival times of the waves are associated with medium velocity or structure. All tomographic techniques rely generally on the measurements of variations in some specific parameters in the medium between the sources and receivers. So, the result is an image of physical property variation denoted as a tomogram. The objective of this paper is therefore to reconstruct subsurface velocity structures using the principle of seismic wave propagation.

Let a seismic wave velocity \( v(\mathbf{x}) \) be a function of the position \( \mathbf{x} \) in a medium and let \( \varphi \) denote all paths connecting a given source and receiver in this medium. Fermat’s principle (Berryman 1989) states that the ray path is stationary with respect to the seismic wave velocity. Therefore the correct ray path between the source and receiver is the one which has the least traveltime among the others paths. Let us define \( \tau^\varphi \) to be the functional that yields the traveltime along the Fermat ray path \( p \in \varphi \):

\[
\tau^\varphi = \Psi(v(x), p) = \min_{p \in \varphi} \int_{p} dl^p \int_{p} \frac{d\mathbf{x}}{v(x)} \quad (1)
\]

where \( \Psi \) is a nonlinear operator and \( dl^p \) denotes the infinitesimal distance along the ray path \( p \). If more than one path produces the same minimum traveltime value, the ray path \( p \) denotes any particular member in this set. The task of tomography is thus to find a function \( v(x) \) given the integrals \( \tau^\varphi \) over a family of manifold \( \varphi \). The difficulty in performing this integration (1) is that the ray path taken by seismic energy depends on the velocity structure. Moreover, the ray path is required to be known in order to evaluate this integral.

Although the solution to the problem of how to reconstruct a function \( v(x) \) from the line integrals \( \tau^\varphi \) dates back to the paper by Random (1917), its applied importance has been made clear by Cormack and Housfield (1960). They developed an effective numerical and medical technique for exploring the interior of the human body for diagnostic purposes. Outside of the field of medicine, it has many uses including electron microscopy, acoustic and optical tomography and radio astronomy. Aki (1976) was first to use seismic data in their 3-D study of the earth's crust. After this study, seismic tomography has become an important geophysical tool for producing a two or three dimensional image of a region of the subsurface and it has been widely studied in the
Literature (Nolet 1993; Scales 1994; Bosch et al. 2005).

2 A Mathematical Problem

To reconstruct subsurface velocity structure between two boreholes shown in Figure 1, we first divide the rectangular region enclosed by our sources \((S_1, S_2, \ldots, S_m)\) and receivers \((R_1, R_2, \ldots, R_m)\) into rectangular cells of constant slowness \(s(x)\), which is a reciprocal of wave velocity \(v(x)\). Let \(t = \{t_1, t_2, \ldots, t_m\}\) be the observed \(m\) dimensional data vector whose elements \(\{t_i\}\) \((i = 1, 2, 3, \ldots, m)\) is the traveltime along the \(i^{th}\) ray path. By using Equation (1), the relationship between \(s(x)\) and \(t\) can be, in general, given by the following formula:

\[
t_i = \Psi(s(x), p_i) + e_i = \min_{p \in \Phi} \int s(x) dt^p + e_i,
\]

(2)

where \(e_i\) represents observation errors in the measurements. If \(l_{ij}\) is denoted as the length of the \(i^{th}\) ray path passing through the \(j^{th}\) cell and defined by

\[
l_{ij} = \min_{p_i \cap \text{cell } j} \int dt^R,
\]

(3)

Equation (2) can be reduced to a system of equations in the following form:

\[
t_i = \sum_{j=1}^{n} l_{ij}s_j + e_i, \quad (i = 1, 2, \ldots, m).
\]

(4)

In the vector- matrix notation,

\[
t = Ms + e,
\]

(5)

where the matrix \(M\) is a \((m \times n)\) matrix whose entries \(l_{ij}\) are described by Equation (3). Note that for any given ray path \(i\), the ray path lengths \(l_{ij}\) are zero for most cells \(j\), as the given ray path will, in general, intersect only a few of the cells in the model. This is the basic equations of forward modelling for the ray equation analysis. In other words, it can be considered as a numerical approximation to Equation (2). The mathematical problem is simply to find the slowness \(s\) and the ray path matrix \(M\) from the measurements \(t\). More precisely, given the first arrival data, our aim is to reconstruct the slowness distribution of the medium between boreholes shown in Figure 1.

As it is very well known, this problem leads to mathematical model that belong to the family of ill-posed problem in the sense that the operator is unknown and depends of the function \(s(x)\) to be estimated (Penrose 1955; Tarantola 1987; Üstündağ et al. 1991; Scales & Snieder 2000). So, standard techniques for solving inverse problems cannot be applied. This is because \(\Psi\) is not continuously invertible, solutions of the inverse problem are unstable under data perturbations.
3 Ray-Tracing

The ray tracing (Bońa, & Slawinski 2002) is based on the concept that seismic energy follows a trajectory determined by tracing equations which physically describe how energy continues in the same direction until it is refracted by the velocity variations. This is an important step and it is also carried in the iterative process. A good choice of ray tracing algorithm would, therefore, be needed for the calculation of ray travel times between two known endpoints. Ray travel times are used because the endpoints of the ray lines remain projected on the same horizontal plane, hereafter referred to as a source-receiver pair respectively, are assumed to be known. The ith ray path \( p_i \) that connects the source - receiver pair can be found by using Fermat’s principle. If the horizontal distance between two vertical boreholes is \( L \), the initial ray path for the ith ray travelling from the source to the receiver is taken to be a straight line:

\[
y_i^0(x) = \left( \frac{y_R - y_S}{x_R - x_S} \right) (x - x_S) + y_S.\]

(6)

The perturbation is to be harmonic series of the form:

\[
\delta y(x; \mathbf{c}) = \sum_{k=1}^{K} c_k \sin \left( \frac{k \pi x}{L} \right).
\]

(7)

where \( \mathbf{c} = \{c_1, c_2, ..., c_K\} \) is a vector of the coefficients of the harmonic series. Only sine and not cosine terms are used because the end points of the ray remain unperturbed. The ith perturbed ray path, defined by

\[
y_i(x; \mathbf{c}) = y_i^0(x) + \delta y(x; \mathbf{c}),
\]

(8)

is given by the functional:

\[
\tau_i(\mathbf{c}) = \int_{x_L}^{x_R} s(x, y_i(x; \mathbf{c})) \sqrt{1 + \left( \frac{dy_i(x; \mathbf{c})}{dx} \right)^2} dx.
\]

(9)

The number of the coefficients in calculating the traveltime of the ray in Equation (10) depends on the resolution of our tomographic model. For a relatively low resolution, it is necessary to seek a general bend in rays, so we only need to determine a few coefficients of \( \mathbf{c} \). However, the determination of these coefficients is difficult since they are connected nonlinearly to the travel times \( \tau_i \). To simplify this problem, we chose to ignore the Snell’s law (Prothero et al. 1988) at the cell boundaries and assumed that \( K = 2 \). Then, the problem is simply reduced to minimising the traveltime functional given in Equation (9) with respect to two coefficients \( c_1 \) and \( c_2 \). In this context, we used a multidimensional search algorithm, which is known as the Simplex method (Nelder and Mead 1965; Press W.H et al. 1995). Basically, starting with three points whose corresponding travel times are \( \tau^1 \), \( \tau^2 \) and \( \tau^3 \), respectively the algorithm seeks to replace the point with the largest traveltime by a smaller one and 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 vertex is found, the vertices are relabelled and the process starts over for the new triangle. If no improvement (or improvement less than a preset threshold) is attained or some fixed number of iterations is executed, the process terminates for this ray path.

4 Reconstruction Method

Following Berryman’s works, the forward problem in (5) can be replaced with the feasibility constraints:

\[
(M_s) \geq t_i, \quad (i = 1, 2, ..., m).
\]

(11)

This arises from Fermat’s principle and implies that the first arrival rays follow the path with a minimum traveltime for a given model \( s \). Thus, if \( s \) is a true model then any ray path matrix \( M \) must satisfy these constraints. Therefore, any model that violates the constraints in (11) along any ray path matrix \( M \) is called a nonfeasible set. Moreover, for the m-feasible constraints the limiting equality is an equation for the hyperplane in the n-dimensional model space. The feasible region is therefore bounded by these hyperplanes and by the planes determined by the positivity constraint,

\[
s_j > 0, \quad (j = 1, 2, 3, ..., n).
\]

(12)
It can easily be shown that the constraints in (11) and (12) imply that the feasible region of the model space is convex. Hence, for a fixed ray path matrix \( M \) the set of all feasible models includes models either inside of the feasible region or on the feasibility boundary determined by \( M \) and \( t \). For any combination of the ray-path matrix \( M \), slowness vector \( s \) and the measured traveltimes \( t \), the number of rays violating those constraints (11) can easily be calculated and it is called the feasibility violation number, determined by

\[
VN_M(s) = \sum_{i=1}^{m} \delta[t_i - (Ms)_i],
\]

(13)

where the step function \( \delta \) is defined by

\[
\delta(x) = \begin{cases} 
1 & x > 0 \\
0 & x \leq 0 .
\end{cases}
\]

Following Lanczos (1950), a generalised eigenvalue problem can be implemented in the form:

\[
\begin{bmatrix}
0 & M \\
M^T & 0
\end{bmatrix}
\begin{bmatrix}
u \\
v
\end{bmatrix}
= \begin{bmatrix}
L & 0 \\
0 & C
\end{bmatrix}
\begin{bmatrix}
u \\
v
\end{bmatrix},
\]

(15)

where \( u \) and \( v \) are \( m \) and \( n \) vectors of ones, respectively. The matrix on the right is defined in terms of diagonal matrices \( L \) and \( C \) whose diagonal elements are the row sum \( L_i \) and the column sum \( C_j \) of the matrix \( M \), respectively. The quantity \( L_i \) is the total length of the \( i \)th ray path. The quantity \( C_j \) is the total ray path segments passing through the \( j \)th cell. It is called the coverage of cell \( j \). Any cell with \( C_j = 0 \) is uncovered and therefore lies outside the span of the data for the current choice of ray paths. We retain only covered cells in the reduced slowness vector \( \tilde{s} \) with \( \tilde{n} \leq n \). By deleting the corresponding columns in the matrix \( M \), the size of the ray path matrix \( M \) is thus reduced. For the simplicity, we assumed \( \tilde{n} = n \) in the following discussions.

In agreement with Berryman (1990), an analogous to the eigenvalue problem (Penrose 1955) providing for high contrast reconstruction can be given in the following form:

\[
\begin{bmatrix}
0 & M \\
M^T & 0
\end{bmatrix}
\begin{bmatrix}
w_i \\
x_i
\end{bmatrix} = \lambda
\begin{bmatrix}
0 & D \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
w_i \\
x_i
\end{bmatrix},
\]

(16)

where for \( \lambda = 1, w_i = u, x_i = s \) and the diagonal weighted matrices are as follow

\[
T = Ls \quad \text{and} \quad D = Cs^{-1}.
\]

(17)

By writing (15) in the canonical form, we get

\[
\begin{bmatrix}
0 & A \\
A^T & 0
\end{bmatrix}
\begin{bmatrix}
u \\
v
\end{bmatrix} = \begin{bmatrix}
L & 0 \\
0 & C
\end{bmatrix}
\begin{bmatrix}
u \\
v
\end{bmatrix}
\]

\[
= \begin{bmatrix}
0 & M \\
M^T & 0
\end{bmatrix}
\begin{bmatrix}
u \\
v
\end{bmatrix},
\]

(18)

and

\[
\begin{bmatrix}
y \\
z
\end{bmatrix} = \begin{bmatrix}
T^2u \\
D^2s
\end{bmatrix}
\]

(19)

Thus, the problem given in (15) can be transformed into the following form:

\[
\begin{bmatrix}
0 & A \\
A^T & 0
\end{bmatrix}
\begin{bmatrix}
y \\
z
\end{bmatrix} = \begin{bmatrix}
y \\
z
\end{bmatrix}
\]

(20)

As we have seen above, with normalisation the current slowness model \( s \) gives rise to the unique eigenvector with the highest eigenvalue and that eigenvalue is unity. Given a set of transmitter-receiver pairs and any model slowness \( s \), Fermat’s principle can then be used to find the ray-path matrix \( M \) associated with \( s \) and with any slowness \( \gamma s \) (\( \gamma > 0 \)) in the same direction as \( s \). If the normalised data is given by

\[
y = T^{-1}t
\]

(21)

then the first problem is to find \( \gamma \) such that the function,

\[
\phi(\gamma) = (y - A\gamma z)^T(y - A\gamma z),
\]

(22)

achieves its minimum at

\[
\gamma = \frac{z^TA^Ty}{z^Tz}.
\]

(23)

Having found optimal slowness \( s_\gamma = \gamma s \) in the given direction \( s \), the second problem is to find another direction in the slowness vector space that gives better fit to the measured traveltimes by minimizing the functional:

\[
\phi(z) = (y - Az)^T(y - Az) + \mu(z - z_\gamma)^T(z - z_\gamma),
\]

(24)
where $\mu$ is a damping parameter. The minimum of Equation (24) occurs at $z = z_\mu$ where $z_\mu$ satisfies the following system of the linear equations:

$$\left(A^T A + \mu I\right)(z_\mu - z_\gamma) = A^T y - z_\gamma. \tag{25}$$

The matrix $(A^T A + \mu I)$ for any $\mu > 0$ is non-singular matrix so that the conjugate-gradient method, which is an all purpose optimizer and simultaneous equation solver (Press W.H et. al 1995; Atkinson 1997), can be used to solve Equation (25) for $z_\mu$.

Another point $s_\mu = D^{-\frac{1}{2}}z_\mu$ in the slowness vector space can thus be obtained. Although the point $s_\mu$ gives a better fit to travel time data, this fit is certainly spurious to some extent because it is based on the wrong ray path matrix $M$ used in the computation of $s_\mu$. Thus, both of the points we have found, lie in the nonfeasible part of the vector space. If the solution of Equation (5) exists, in agreement with Berryman, it must lie on the boundary of the feasible region. So $s_\gamma$ and $s_\mu$ may be used to find an optimum point on this boundary in the sense that it is as consistent as possible with the ray path matrix, with the travel time measurements and with the feasibility constraints. Because of the convex property of the feasible region, there exists a point $s$ between points $s_\gamma$ and $s_\mu$ that is closer to the feasible region than the either of two end points. This can easily be found by computing the feasibility violation number and choosing the model that gives a minimum violation number when we move in the direction $(s_\mu - s_\gamma)$ from $s_\gamma$ (Berryman 1991, 1993; Marquardt 1970). Then, we get

$$\hat{s} = s_\gamma + \alpha(s_\mu - s_\gamma). \tag{26}$$

As $\alpha$ gets smaller, it is expected that the inversion method is not providing any further improvement so that a threshold for $\alpha$ of 0.25 is used to stop searching. Once we find $\hat{s}$ and then scale it up to the point, denoted as $s_\mu$ in the same direction lying in the feasibility boundary. It is not hard to see that these three points $s_\gamma, s_\mu$ and $s_\mu$ are distinct unless we found the exact solution of the inverse problem in Equation (5). Otherwise, these three points form a triangle and its area gives us an estimate how far we are from the solution. An iterative reconstruction algorithm, coded in C++ programming language, that uses above ideas is tested for artificially generated traveltime data in the following section.

5 Computer Simulations

For a comparison with Berryman’ results, we took a similar model slowness structure shown in Figure 2(a). In according to Berryman, the slowness model consists of $(8\times16)$ cells that have normalised slowness. It is allowed that there is a low speed anomaly on bottom and a high-speed anomaly on the top. We then parameterised the model slowness by $\beta$, where the slow region had a slowness of $1+\beta$ and the fast region had a slowness of $1-\beta$ ($\beta = 0.2, 0.5$). Thus, variations in the value of $\beta$ will provide variations in the contrast of the model slowness. We first chose $\beta = 0.2$ for the low contrast and generate the traveltime data by using the ray tracing method described in the paper. As mentioned before, the simplex ray path is constrained by the use of two coefficients in the sine series expansions to be quite smooth, perhaps smoother than it should be for such high contrast media. Traveltime data, shown in Figures 2(b), consists of 320 rays, including 256 $(16\times16)$ rays travelling from left to right and 64 $(8\times8)$ rays travelling from bottom to top. Then our computer program, written in C++ programming language, was run on a personal computer. The results, shown in Figures 2(c)-(f) illustrate the convergence of the method. It can be seen that as the number of iterations increases, the fast anomaly is reconstructed very well while the slow anomaly is located well. As $\beta$ increases we get higher contrast and expect that the slow anomaly could always be harder to reconstruct because few or no first arrivals pass through this region. This is illustrated in Figure 3. For small contrast (less than 20%), the method produces uniformly excellent results even though the data contain zero-mean random noise with a variance less than 0.01. For large contrast, it becomes less accurate. The method converges quite rapidly to a definite result unless we force the algorithm to make a minimum percentage correction step per iteration as 1-10% of the distance along the direction $(z_\mu - z_\gamma)$. In an agreement with Berryman, it requires at least 10 iterations for getting a reasonable results. Variations with respect to the number of iterations in the stopping criterion used in the algorithm are shown in Figure 4. It is clearly seen that as the number of iterations increases, the area of triangle decreases monotonically.

ISSN: 1790-2769

and subsequently oscillates around a small number. On the other hand, a quantitative measure of the improvement appears in Figure 4, which plots reconstruction error$^1$ as a function of the number of iterations. It can be seen that the reconstruction error first decreases rapidly and then reduces gradually as the number of iteration increases.

![Figure 3](image-url)

**Figure 3** (a) a slowness model having contrast with 50% anomaly. (b) Traveltime data. (c) Reconstruction.

![Figure 2](image-url)

**Figure 2** (a) Slowness Model having contrast with 20% anomaly. (b) Traveltime data. Reconstruction at (c) the first, (d) 10th, (e) 20th and (f) 30th iteration.

![Figure 4](image-url)

**Figure 4.** Area of triangle as a stopping criterion for the algorithm and the root mean square error (rmse) versus with the number of iterations.

### 6 Conclusions

The results presented here are encouraging us for retrieving the slowness distribution of a medium from the first arrival traveltimes, which contain the errors made by neglecting ray bending effects far more significant than measurements errors. Even if the data contain errors less than 1%, it gives very stable reconstructions and avoids the large oscillations often found in traditional least squared methods. Although Fermat’s principle determines the ray path matrix once slowness is given, it also determines which slowness vectors are feasible and infeasible. Therefore, this plays an essential role in the reconstruction algorithm when the data have no noise. However, it requires a large consumption of computer time because of the computation of the ray path matrix. This can be reduced by using parallel processing techniques because each ray path may be computed independently of the others. If we incorporate into above analysis any other geophysical information as constraints to guide the imaging or inversion, we could get more

$^1$ \( rmse = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (t_i - \tilde{t}_i)^2} \)
improvement in the reconstruction. Therefore, it deserves further investigation.

Acknowledgment: This work is a part of the project, whose name is seismic tomography with number of BSE-040/020701, supported by the University of Marmara.

References