Physics of the Earth and Planetary Interiors, 21 (1980) 120–125

COMPUTATIONAL METHODS FOR INVERSE PROBLEMS IN GEOPHYSICS: INVERSION OF TRAVEL TIME OBSERVATIONS

V. PEREYRA [1], H.B. KELLER [2] and W.H.K. LEE [3]

[1] *Universidad Central de Venezuela, Apartado 59002, Caracas 105 (Venezuela)*
[2] *Applied Mathematics Department, California Institute of Technology, 101-50 Pasadena, CA 91125 (U.S.A.)*
[3] *U.S. Geological Survey, Menlo Park, CA (U.S.A.)*

(Received December 5, 1978; accepted for publication March 13, 1979)

Pereyra, V., Keller, H.B. and Lee, W.H.K., 1980. Computational methods for inverse problems in geophysics: inversion of travel time observations. Phys. Earth Planet. Inter., 21: 120–125.

General ways of solving various inverse problems are studied for given travel time observations between sources and receivers. These problems are separated into three components: (a) the representation of the unknown quantities appearing in the model; (b) the nonlinear least-squares problem; (c) the direct, two-point ray-tracing problem used to compute travel time once the model parameters are given. Novel software is described for (b) and (c), and some ideas given on (a). Numerical results obtained with artificial data and an implementation of the algorithm are also presented.

1. Introduction

Most inverse problems in geophysics which use travel time observations between a source and a receiver can be stated as

$$\min_{\boldsymbol{\mu}} \sum_{i=1}^{I} [T_{i\,\mathrm{obs}} - T_{i\,\mathrm{comp}}(\mathbf{w}_i, \boldsymbol{\mu})]^2 \tag{1}$$

subject to

$$\begin{aligned} \mathbf{w}_i' &= f(s, \mathbf{w}_i; \boldsymbol{\mu}) \\ g_i(\mathbf{w}_i(0), \mathbf{w}_i(1); \boldsymbol{\mu}) &= 0 \end{aligned} \tag{2}$$

where the vector function $\mathbf{w}_i(s)$ describes the ray between the source ($s = 0$) and the receiver ($s = 1$). The vector function $\boldsymbol{\mu}$ may include a description of the medium, both in terms of geometry (boundary interfaces) and material (velocity of propagation of elastic waves). It may also contain the unknown location of the source (hypocenter) and the starting time of the event. Specifically, (2) represents the differential equations for two-point ray tracing in a very general medium.

Stated in this generality the problem is probably unsolvable. In particular, for the case of inversion, in obtaining velocity distributions (which are considered here to be completely general, piecewise smooth functions, i.e. the medium considered is totally nonhomogeneous with material interfaces) there may be a large amount of indeterminacy. The observed rays may not sample the area under study properly. In any case there will only be a finite number of observations to determine a general function of several variables. This points clearly to our first basic observation: Since we are only able to determine a finite number of parameters, problem (1–2) must be reformulated as an approximation problem in which the functions $\boldsymbol{\mu}$ are restricted to a finite-dimensional subspace, preferably of dimension much less than I, the number of observations.

Of course, this is what practitioners have been doing for a while now (Backus and Gilbert, 1968; Aki and Lee, 1976). We simply wish to center attention on what is feasible, and to connect this important problem with areas that either are well developed or may have to be developed once connection with the problem is established.

The most widely used technique in recent times seems to be the ACH method or a variant of it (Aki et al., 1977). This is just the first step indicated by our first observation. In the ACH method, the velocity field is appoximated by a piecewise constant function. In the most complex versions, blocks of different sizes may be used.

Once one recognizes the connection between this problem and that of approximating a function of several variables from, it should be emphasized, a very unusual kind of data, then it is possible to introduce powerful techniques developed for other purposes. For instance, it is possible to borrow valuable information from the following areas.

1.1. Finite elements or multidimensional splines

There are very elaborate techniques for generating nonuniform triangulations that adapt themselves to the (unknown) function being represented in order to economize in the number of parameters needed for a desired accuracy (Whiteman, 1976), and also a large variety of approximating spaces from which to choose. Standard software for one-dimensional problems is available (de Boor, 1971) and it can be extended easily to higher dimensions via tensor products (Pereyra and Scherer, 1973; de Boor, 1977).

1.2. Picture reconstruction and surface approximation from scattered data

This is perhaps an area more closely connected with our subject than finite elements and where many developments are occurring at present (Herman et al., 1978). The second observation we would like to make is: We can use readily available, high-quality, high-performance mathematical software for solving problem (1–2).

It is necessary to recognize the various modules into which this problem can be subdivided, and for which there is readily available, professionally produced software. We have in mind essentially three modules.

1.3. The representation module (I)

We have to decide how to approximate the functions $\boldsymbol{\mu}$ by functions restricted to a finite-dimensional space. Implicit in this choice is the need for minimizing the dimensionality of the approximating function space in a way compatible with the information present in the data. This process may have to be dynamic; i.e. the approximating subspace may change within an iteration for improving the model. This is probably the least developed module, but recognizing the problem may stimulate research in this direction.

It will be assumed that a representation has been chosen, and that an n-vector of parameters $\boldsymbol{\eta}$ describes the approximation to the desired functions $\boldsymbol{\mu}$ (for instance, in the case of velocity inversion by the ACH method, $\boldsymbol{\eta}$ would be the vector of the average velocities in the chosen blocks). We shall describe the approximate problem by equations (1–2) with $\boldsymbol{\mu}$ replaced by $\boldsymbol{\eta}$.

1.4. The nonlinear least-squares module (II)

Clearly, once approximated, problem (1) is a standard nonlinear least-squares problem of finding the optimum value of the parameters $\boldsymbol{\eta}$, given the I observations $T_{i\ obs}$, $i = 1, \dots I$. Since the $\boldsymbol{\eta}$ appear in the differential equations (2) that define implicitly the calculated quantities $T_{i\ \mathrm{comp}}$, this is more properly called a "parameter identification problem". Any of the available iterative techniques for solving nonlinear least-squares problems can be used. In particular, they will require the solution of the two-point boundary value problem (2) for a given $\boldsymbol{\eta}$ and for each observation. Many of these techniques will also require the partial derivatives of $T_{i\ \mathrm{comp}}$ with respect to the $\boldsymbol{\eta}$.

1.5. The direct problem (III)

As indicated in Section 1.4., it will be necessary to solve (2) for given $\boldsymbol{\eta}$ and also to calculate $\partial T_{i\ \mathrm{comp}}/\partial \boldsymbol{\eta}$. These tasks are crucial, since in most applications the number of observations I, i.e. the number of rays to be calculated, is fairly large.

We shall indicate our approach to solving these problems and describe the general software chosen, mainly for modules II and III. We shall also give some specific examples and some preliminary numerical results.

2. Some specific examples

2.1. Two dimensional velocity inversion of a general, non-homogeneous, isotropic, smooth medium

Equations (2) for two-point ray tracing in two dimensions have been discussed in detail by Pereyra et al. (1978). They are

$$\begin{aligned} w_1' &= w_4 \cos w_3 \\ w_2' &= w_4 \sin w_3 \\ w_3' &= w_4 u(w_1, w_2)[u_z(w_1, w_2) \cos w_3 \\ &\quad - u_x(w_1, w_2) \sin w_3] \\ w_4' &= 0 \end{aligned} \tag{3}$$

with the boundary conditions

$$\begin{aligned} w_1(0) &= x_0\,, \quad w_2(0) = z_0 \\ w_1(1) &= x_f\,, \quad w_2(1) = z_f \end{aligned} \tag{4}$$

where $\mathbf{w} = (w_1, w_2, w_3, w_4)^T = (x, z, \psi, S)^T$, $v(x, z)$ is the velocity of propagation of the elastic waves considered and $u(x, z) = 1/v(x, z)$ is the slowness. Thus (x, z) are the Cartesian coordinates of the ray, while ψ is the ray parameter. The independent variable t is the normalized arc length over the ray $t = s/S$, where $S \equiv w_4$ is the total arc length (unknown). The boundary conditions specify that the desired ray joins the two points (x_0, z_0) and (x_f, z_f) in the (x, z) plane. In using arc length as independent variable, rays that curve back on themselves present no special problem.

The observed quantity is the arrival time $\tau = T + T_o$, where T_o is the origin time and T is the travel time given by

$$T = w_4 \int_0^1 u(w_1, w_2)\,\mathrm{d}t \tag{5}$$

It will now be assumed that there are a number of stations $\{St_i\}_{i=1, \dots L}$, and a number of events $\{Ev_j\}_{j=1, \dots J}$ which are observed at the stations. Let $T_{obs\,ij}$ be the observed first P-wave arrival time at station i from source j. For simplicity it is assumed that the hypocenter and origin times are known already with sufficient accuracy, and that as a result of previous studies the velocity distribution in the area of interest can be accurately described by a linear function of position

$$v(x, z) = v_0 + ax + bz \tag{6}$$

It is intended to improve the given starting values of the parameters v_0, a, b by using the information $T_{obs\,ij}$.

In this case, the vector $\boldsymbol{\eta}$ is simply

$$\boldsymbol{\eta} = (v_0, a, b) \tag{7}$$

and it will influence both the differential equations for the rays and their travel times.

2.2. Recovery of a reflector in a piecewise continuous medium

It is now assumed that there is a curve of discontinuity in the medium, say $\theta(x, z) = 0$, and that the sources are now in the surface (explosions). Clearly in this case, the former assumption of known source positions and origin times holds. The arrival time observations $T_{obs\,ij}$ are now assumed to correspond to first arrivals of P-waves reflected (once) by the material discontinuity at $\theta(x, z) = 0$.

Let us suppose that $\theta(x, z) = 0$ can be accurately represented by a piecewise linear function (broken line). Let k be the number of linear segments of this approximation, and let (x_ν, z_ν), $\nu = 0, \dots k$, be the end-points of the segments

$$\begin{aligned} \theta_{ap}(x, z) &= (z - z_\nu) - (z_{\nu+1} - z_\nu)(x - x_\nu)/(x_{\nu+1} - x_\nu) \\ &= 0 \quad \text{if } x \in [x_\nu, x_{\nu+1}]\,. \end{aligned} \tag{8}$$

To simplify, we shall assume that the break-point abscissae x_ν are correctly positioned and that only the depths z_ν are to be adjusted. Pereyra et al. (1978) describe how to solve ray-tracing problems in media with very general geometries. In this particular instance we need only to consider eq. (3) in two parts: before and after the reflection. Each piece is smooth and satisfies all the requirements for the application of (3). The only difficulty is that the contact point with the curve $\theta_{ap}(x, z) = 0$ is not known. Let $\mathbf{w}_\mathrm{I}$, $\mathbf{w}_\mathrm{II}$ represent the ray before and after the reflection respectively. Equation (3) is written once for $\mathbf{w}_{\mathrm{I}}$ and once for $\mathbf{w}_{\mathrm{II}}$, giving eight differential equations and apparently only four boundary conditions (eq. 4). The four missing boundary conditions are obtained by imposing the continuity of the ray at the reflection point P_r, by applying Snell's law to the ray parameter, and by requiring that P_r be on the curve. These conditions are:

$$w_{\mathrm{I}1}(1) - w_{\mathrm{II}1}(0) = 0 ,$$

$$w_{\mathrm{I}2}(1) - w_{\mathrm{II}2}(0) = 0 ,$$

$$\begin{aligned} v(P_r)[\theta_x(P_r)(\sin w_{\mathrm{I}3}(1) - \sin w_{\mathrm{II}3}(0))] \\ - [\theta_z(P_r)(\cos w_{\mathrm{I}3}(1) - \cos w_{\mathrm{II}3}(0))] = 0 \end{aligned} \tag{9}$$

$$\theta(P_r) = 0$$

From (8),

$$\theta_x(P_r) = -(z_{\nu+1} - z_\nu)/(x_{\nu+1} - x_\nu) \quad \text{if } x \in [x_\nu, x_{\nu+1}] ,$$

$$\theta_z(P_r) = 1$$

The model parameters are in this case $\boldsymbol{\eta} = (z_0, z_1, \ldots z_k)$ and they appear only in the boundary conditions (9).

3. Solution of the nonlinear least-squares problem

The approximate version of problem (1–2), where all control functions that are included in the vector $\boldsymbol{\mu}$ have been discretized and represented by a finite-dimensional vector $\boldsymbol{\eta}$, can be stated as

$$\min_{\boldsymbol{\eta}} \sum_{i=1}^{I} [T_{i\,\mathrm{obs}} - T_{i\,\mathrm{comp}}(\mathbf{w}_i; \boldsymbol{\eta})]^2 \tag{10}$$

subject to

$$\begin{aligned} \mathbf{w}_i' = f(t, \mathbf{w}_i; \boldsymbol{\eta}) , \\ g_i(\mathbf{w}_i(0), \mathbf{w}_i(1); \boldsymbol{\eta}) = 0 \end{aligned} \tag{11}$$

Here (11) represents the appropriate ray equations and boundary conditions, of which two examples have been given. Pereyra et al. (1978) also present the treatment for three-dimensional ray tracing. The most favoured technique for solving nonlinear least-squares problems of this type is based upon the Levenberg–Marquardt (LM) algorithm. Let us represent (10) for simplicity as

$$\min_{\boldsymbol{\eta}} \|\boldsymbol{\tau}(\boldsymbol{\eta})\|_2^2 \tag{12}$$

where the vector $\boldsymbol{\tau}(\boldsymbol{\eta}) \equiv (T_{i\,\mathrm{obs}} - T_{i\,\mathrm{comp}})_{i=1,\ldots I}$, and $\|\ldots\|_2$ is the Euclidean norm.

Let $\boldsymbol{\tau}_{\boldsymbol{\eta}}(\boldsymbol{\eta})$ be the Jacobian matrix of the vector function $\boldsymbol{\tau}$:

$$\boldsymbol{\tau}_{\boldsymbol{\eta}}(\boldsymbol{\eta}) = \left(\frac{\partial \tau_i(\boldsymbol{\eta})}{\partial \eta_j}\right)_{\substack{i=1,\ldots I \\ j=1,\ldots n}} \tag{13}$$

The LM algorithm can be briefly described as the following iterative process. Given a starting guess for the unknown vector $\boldsymbol{\eta}$, say $\boldsymbol{\eta}^0$, one iterates according to

$$\boldsymbol{\eta}^{i+1} = \boldsymbol{\eta}^i + \boldsymbol{\gamma}^i \tag{14}$$

where $\boldsymbol{\gamma}^i$ is the least-squares solution to the linear least-squares problem

$$\min_{\boldsymbol{\gamma}} \left\| \begin{pmatrix} \boldsymbol{\tau}_{\boldsymbol{\eta}}(\boldsymbol{\eta}^i) \\ L_i \end{pmatrix} \boldsymbol{\gamma} - \boldsymbol{\tau}(\boldsymbol{\eta}^i) \right\|_2^2 \tag{15}$$

The $n \times n$ matrix L_i is the Choleski factor of a positive definite matrix K_i; i.e. $K_i = L_i^T L_i$. Usually this matrix is simply chosen to be a multiple of the identity matrix:

$$K_i = \lambda_i I \tag{16}$$

Other choices allow for some natural dynamical scalings.

The vector $\boldsymbol{\gamma}^i$ can be considered as a direction of search which can be proven to be of descent, i.e.

$$\|\boldsymbol{\tau}(\boldsymbol{\eta}^i + t_i \boldsymbol{\gamma}^i)\| \leqslant \|\boldsymbol{\tau}(\boldsymbol{\eta}^i)\| \tag{17}$$

if $0 < t_i$ is sufficiently small. In some implementattions of this method a parameter t_i is introduced to control the step and usually improve the convergence. A very good implementation of this method has been written by Moré (1978) as part of a subroutine package for minimization.

For this and other similar methods it is necessary to calculate the Jacobian matrix (13). The present matrix has the form

$$\boldsymbol{\tau}_{\boldsymbol{\eta}} = -\left(\frac{\partial T_{i\,\mathrm{comp}}}{\partial \eta_j}(\mathbf{w}_i, \boldsymbol{\eta})\right)_{\substack{i=1,\ldots I \\ j=1,\ldots n}}$$

From (5) it follows that

$$\frac{\partial T_{i\,\mathrm{comp}}}{\partial \eta_j} = \frac{\partial w_{i4}}{\partial \eta_j} \int_0^1 u(w_{i1}, w_{i2})\,\mathrm{d}t$$

$$+ w_{i4} \int_0^1 \left(u_{w_{i1}} \frac{\partial w_{i1}}{\partial \eta_j} + u_{w_{i2}} \frac{\partial w_{i2}}{\partial \eta_j} \right) dt \qquad (18)$$

Therefore, it is necessary to calculate the partial derivatives of the solution to the ray equations with respect to the parameters $\boldsymbol{\eta}$. It is known (Coddington and Levinson, 1955) that those derivatives satisfy a linear differential system. If $\boldsymbol{\alpha}(t) = \partial \mathbf{w}/\partial \eta_j$, then

$$\begin{aligned} \boldsymbol{\alpha}' &= f_{\mathbf{w}}(t, \mathbf{w}; \boldsymbol{\eta})\, \boldsymbol{\alpha} + \partial f/\partial \eta_j \\ g_{\mathbf{w}(0)}\boldsymbol{\alpha}(0) + g_{\mathbf{w}(1)}\boldsymbol{\alpha}(1) &= -\partial g/\partial \eta_j \end{aligned} \qquad (19)$$

where $f_{\mathbf{w}}, f_{\eta}, g_{\mathbf{w}(0)}, g_{\mathbf{w}(1)}, g_{\eta}$ are the respective Jacobian matrices. Thus, in order to obtain the columns of the Jacobian matrix $\mathbf{w}_{\eta}$, it is necessary to solve n systems of linear ordinary differential equations of the form (19), one for each of the parameters $\eta_j, j = 1, \dots n$. This seems a formidable task but it proves to be fairly trivial and inexpensive, as will now be indicated.

4. The direct problem. Solving the two-point ray-tracing equations

Since Pereyra et al. (1978) discussed this problem in great detail we shall only highlight the main points pertinent to the present concern. Lentini and Pereyra (1977) have shown how to solve numerically general nonlinear two-point boundary value problems of the form (2). In Pereyra et al. (1978) the present authors have described the application of that general technique to two- and three-dimensional ray tracing between two points, for both smooth and piecewise smooth velocity distributions. We emphasize that general isotropic, heterogeneous media are discussed here.

General software is now available for these tasks. Furthermore, the software is considerably more robust and efficient than the traditional simple shooting techniques. These general programs use complex finite difference approximations on general, non-uniform meshes. They are adaptive, i.e. the order of the method and the mesh points are dynamically adjusted in the course of the computation to produce an approximate solution of a requested precision with a minimum of effort and user intervention.

For media with piecewise constant-velocity fields, eq. 2 reduces to a system of nonlinear transcendental equations. These equations are as many as the number of crossings or reflections of the ray on material interfaces. Within regions of constant velocity there is no need to integrate eqs. 11 since the rays are straight-line segments. An efficient and very general program for solving that type of problem has been developed by Perozzi and Keller (1978). This program, as opposed to those mentioned before, includes calculation of amplitudes which can be useful for other types of inversion process.

Our program for the general problem requires an initial ray to start the iterations. The program will usually produce a converged ray from fairly inaccurate initial guesses but, as is usual in iterative processes, the better the initial guess the faster and more stable will be the convergence. Thus careful consideration of the generation of appropriate initial guesses has been beneficial. On smooth velocity fields a very simple starting guess, the segment joining the two given end-points, has been used. This corresponds to a constant initial velocity field. For problems with interfaces, this simple guess will usually not make a great deal of sense.

Another possibility, which has not been implemented, is to use the Perozzi–Keller program on a simplified, piecewise constant model, in order to produce starting rays for our more general program. This will certainly be better than taking more or less arbitrary piecewise linear approximations as we have done so far, although it is not clear that it will necessarily save computer time.

Another idea is to save the discrete rays from one iteration of the LM procedure to the next. If the model parameters have not changed radically those should be fairly good approximations and therefore will make all iterations after the first considerably cheaper. Unfortunately the storage requirements may be prohibitive.

5. A numerical example

We give now a numerical example of the type described in Section 2.2. Shots are set on the surface of the Earth ($z = 0$) at positions $x_{si} = -4 + 2i$, $i = 0, 1, \dots 4$, and their effect is registered at stations on the surface with abscissae $x_{\text{st}\,j} = -4 + j, j =$

TABLE I

Results of numerical example

x	z_{guess}	z_1	z_2
−5	4.00	3.69	3.72
−1.2	2.25	2.22	2.18
0	1.80	1.75	1.74
0.34	1.65	1.61	1.61
0.5	1.58	1.56	1.53
0.6	1.53	1.45	1.45
0.7	1.48	1.50	1.47
0.8	1.44	1.38	1.31
1.0	1.35	1.50	1.57
5.0	1.35	1.51	1.47
Residue	0.046	0.046	0.0042

0, ... 8. The units are kilometres. Thus there is a total of 45 data points. The data are generated artificially by running the ray-tracing program with a reflector $\theta(x, z) = 0$ given by

$$\theta(x,z) = \begin{cases} z - 1.5 & , \quad 1 \leqslant x \\ z - 1.75 + 0.25 \sin(\pi/2\, x) , & 0 \leqslant x \leqslant 1 \\ z - 1.75 + \pi/8\, x & , \quad x \leqslant 0 \end{cases}$$

The velocity distribution was taken to be $v(x, z) = 5 + 0.25(x + z)$ km s^{-1}.

Once the travel data had been generated we "forgot" $\theta(x, z)$ (but not $v(x, z)$), and tried to reconstruct it approximately as a piecewise linear function of the form (8). Table I presents the abscissae of the break points, x (fixed), the guessed values of the ordinates of the break points, z_{guess}, and the two first iterates of the Marquardt procedure, z_1, z_2, together with the residual sum of squares. The number of linear segments was ten. These are just preliminary results and we hope to obtain more comprehensive ones, both with artificial and real data.

6. Conclusions

We have discussed a large class of inversion problems in geophysics in a unified fashion. We have identified three main components in the numerical solution of such problems: representation, optimization, and the direct problem.

We have exemplified with some familiar problems what we intend for the representation of the unknown functions in the inverse problem, while indicating some existing powerful software for the remaining two components. Using such software we have shown promising numerical results.

Acknowledgment

This work has been supported, in part, under Contract 14-08-0001-16777 with the U.S. Geological Survey.

References

Aki, K., Christofferson, A. and Husebye, E.S., 1977. Determination of the three dimensional seismic structure of the lithosphere. J. Geophys. Res, 82: 277–296.

Aki, K. and Lee, W.H.K., 1976. Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes. 1. A homogeneous initial model. J. Geophys. Res., 81: 4381–4399.

Backus, G.E. and Gilbert, J.F., 1968. The resolving power of gross earth data. Geophys. J.R. Astron. Soc., 16: 169–205.

de Boor, C., 1971. A Subroutine Package for Calculating with B-Splines. Los Alamos Report LA4728-MS, NM.

de Boor, C., 1977. Efficient Computer Manipulation of Tensor Products. MRC Technical Report, Mathematics Research Center, University of Wisconsin, Madison, U.S.A. 17 pp.

Coddington, E. and Levinson, N., 1955. Theory of Ordinary Differential Equations. McGraw-Hill, New York, NY.

Herman, G.T., Lent, A. and Lutz, P.H., 1978. Relaxation methods for image reconstruction. Commun. ACM, 21: 152–158.

Lentini, M. and Pereyra, V., 1977. An adaptive finite difference solver for nonlinear two point boundary value problems with mild boundary layers. SIAM J. Numer. Anal., 14: 91–111.

Moré, J., 1978. MINPACK: A subroutine package for minimization. Argonne National Laboratory.

Pereyra, V. and Scherer, G., 1973. Efficient computer manipulation of tensor products with applications to multidimensional approximation. Math. Comput., 27: 595–605.

Pereyra, V., Lee, W.H.K. and Keller, H.B., 1978. Solving two-point seismic ray-tracing problems in a heterogeneous medium. Part 1: A general numerical method based on adaptive differences. Bull. Seismol. Soc. Am., in preparation.

Perozzi, D. and Keller, H.B., 1978. Ray tracing in heterogeneous media. Caltech, Pasadena, CA, in preparation.

Whiteman, J.R. (Editor), 1976. The Mathematics of Finite Elements and Applications II, MAFELAP 1975. Academic Press, New York, NY.