P171

AUTOMATIC VELOCITY ANALYSIS WITH TRANSMITTED WAVES IN 3D INHOMOGENEOUS MEDIA: A SYNTHETIC EXAMPLE

ANDREY MOROZOV[1], BORIS KASTHAN[1] and WIM A. MULDER[2]
[1]*Saint Petersburg State University, Russia*
[2]*Shell International E & P, Rijswijk, The Netherlands*

Summary

A method for tomographic velocity updating of transmission traveltimes is presented. The approach is based on a gradient-type optimisation procedure, using the limited-memory BFGS algorithm. The functional to be minimised is the least-squares error between picked and modelled traveltimes. The traveltimes were computed by a variant of the well-known wave front construction method. This provides traveltimes on an irregular grid, which were subsequently interpolated to the receiver positions. The optimisation procedure requires the gradient of the functional with respect to the parameters of the velocity model. The gradient was computed by an approximation of the adjoint-state method, which is much more efficient than techniques based on Fréchet derivatives and the solution of large sparse linear systems. The algorithm was tested on synthetic VSP data obtained for an isotropic inhomogeneous velocity model.

Introduction

Traveltime tomography is a well-known method for determining velocities from picked traveltimes, as for instance described in [1]. Minimisation of the error between observed and modelled traveltimes requires the gradient (derivatives) of the error with respect to the velocity model parameters. This gradient may be computed by determining Fréchet derivatives and solving large sparse linear systems. A considerably more efficient method can be obtained by using adjoint states. In [2], the adjoint-state method was applied to the discretised ray equations. In [3], the discrete set was replaced their continuum limit, which turned out to be identical to the perturbed traveltime equation as, for instance, derived in [4]. The same approach is applied here.

Traveltimes can be computed by ray tracing, using two-point ray tracing or the more efficient wave front construction method [5]. Here, a variant of the latter is used in which infill rays are not determined by interpolation but are recomputed from the source.

Method

The ray-trace equations are given by $d\xi/d\ell = f(\mathbf{p}, \sigma(\mathbf{x}))$, $\xi = (\mathbf{x} \quad \mathbf{p} \quad \tau)^T$, where σ is slowness, $\mathbf{x}$ position, $\mathbf{p}$ the slowness vector, and τ traveltime. The ray is parametrised by ℓ. We want to minimise $J = \frac{1}{2}\left(\tau_r - \tau_r^{\mathrm{obs}}\right)^2$, where τ_r^{obs} is the observed traveltime at the receiver $\mathbf{x}_r$, and τ_r is a solution of the ray equations for a shot at position $\mathbf{x}_r$. With two-point ray tracing, the constraints for this minimisation problem are $\tau_s = \tau(\ell_s)$, $\tau_r = \tau(\ell_r)$, $\mathbf{x}(\ell_r) = \mathbf{x}_r$, and $\mathbf{x}(\ell_s) = \mathbf{x}_s$. We may set $\ell_s = 0$ at the source without loss of generality. The ℓ_r at the receiver is implicitly defined by $\mathbf{x}(\ell_r) = \mathbf{x}_r$.

Recall that the minimisation of a multi-variate function $f(x_1, x_2, ..., x_N)$ subject to the constraint $g(x_1, x_2, ..., x_N) = 0$ can be accomplished by a lagrange multiplier λ, the requirement being that

$\partial f/\partial x_k = \lambda \partial g/\partial x_k$ for $k = 1, ..., N$. With multiple constraints $g_m = 0$, $m = 1, ..., M$, the condition is that $\partial f/\partial x_k = \sum_{m=1}^{M} \lambda_m \partial g_m/\partial x_k$, for $k = 1, ..., N$. The lagrangian in that case is $\mathcal{L} = f - \sum_{m=1}^{M} \lambda_m g_m$. Here we can use the same method to find the gradient of J constrained by the ray equations.

Instead of spelling out the details of the adjoint-state method for kinematic ray tracing in inhomogeneous isotropic media, we will explain the basic idea for the simpler ODE (ordinary differential equation)

$$\frac{dx}{dt} = f(x, \sigma(x)), \quad 0 \le t \le T,$$

with fixed initial condition $x(0) = x_0$ and fixed T (this is different from ray tracing where T also depends on σ). We may also write

$$x(T) = x(0) + \int_0^T f dt.$$

Let a functional J be given by $J = \frac{1}{2}[x(T) - y]^2$ for a measured y. To determine $dJ/d\sigma(x)$, we can use two approaches. Recall that the minimisation of a multi-variate function $f(x_1, x_2, ..., x_N)$ subject to the constraint $g(x_1, x_2, ..., x_N) = 0$ can be accomplished by a lagrange multiplier λ, the requirement being that $\partial f/\partial x_k = \lambda \partial g/\partial x_k$ for $k = 1, ..., N$. With multiple constraints $g_m = 0$, $m = 1, ..., M$, the condition is that $\partial f/\partial x_k = \sum_{m=1}^{M} \lambda_m \partial g_m/\partial x_k$, for $k = 1, ..., N$. The lagrangian in that case is $\mathcal{L} = f - \sum_{m=1}^{M} \lambda_m g_m$.

The lagrangian for the problem of minimising J subject to $x_t = f$ is given by

$$\mathcal{L}(\sigma, x; \lambda, \mu) = J + \int_0^T \lambda \left(\frac{dx}{dt} - f \right) dt + \mu[x(0) - x_0].$$

Here the λ's and μ are lagrange multipliers. Minimisation of $J(\sigma)$ is equivalent to finding a stationary point of $\mathcal{L}$. Stationarity is obtained for $\frac{\partial \mathcal{L}}{\partial x} = 0$ and $\frac{\partial \mathcal{L}}{\partial \sigma} = 0$. To determine the stationary point, it is convenient to perform partial integration:

$$\mathcal{L} = J + \lambda(T)x(T) - \lambda(0)x(0) - \int_0^T \left(x\frac{d\lambda}{dt} + \lambda f \right) dt + \mu[x(0) - x_0].$$

The stationarity requirements are:

$$\frac{\partial \mathcal{L}}{\partial x(T)} = \frac{dJ}{dx(T)} + \lambda(T) = 0, \quad \frac{\partial \mathcal{L}}{\partial x(0)} = -\lambda(0) + \mu = 0,$$

$$\frac{\partial \mathcal{L}}{\partial x(t)} = -\lambda_t - \lambda f_x = 0, \quad \text{for } 0 < t < T, \quad \frac{\partial \mathcal{L}}{\partial \sigma} = -\int_0^T \lambda f_\sigma dt.$$

In this way, we obtain the adjoint-state equations

$$\lambda_t = -\lambda f_x, \quad \text{for } 0 \le t \le T, \tag{1}$$

with *final* condition $\lambda(T) = -dJ/dx(T)$, so that λ has to be found by tracing *backward* along the ray. Finally, the gradient of the functional is found to be

$$\frac{\partial J}{\partial \sigma} = -\int_0^T \lambda f_\sigma dt, \tag{2}$$

with λ given by the stationarity conditions.

For the original problem, the adjoint-state equation can be determined from the discretised ray equations, as in [2]. By reverting back to the continuum form of these equations, it turns out that λ can be found in closed form and that eq. (2) reduces to the perturbed ray equation [4].

In practice, the velocity model is represented by 3D B-splines on a regular grid of spline nodes. The velocity in a point $\mathbf{x}$ is determined by interpolation. Interpolation is also required to obtain the traveltime at a given receiver position $\mathbf{x}_r$. Figure 1 shows a ray tube as occurring in the wavefront construction method. Linear interpolation is used to determine $\tau(\mathbf{x}_r)$ from the traveltimes at positions A1, A2, A3, B1, B2, and B3 [6].

The traveltime error is given by

$$J = \sum_{i=1}^{N_s} \sum_{j=1}^{N_r} J_{i,j}, \quad J_{i,j} = \frac{1}{2}(\tau_r^{\mathrm{obs}}{}_{i,j} - \tau_{i,j})^2, \tag{3}$$

where $\tau_{i,j}$ is the traveltime from source i to receiver j. To determine the gradient of J with respect to the velocity model, we take a source-receiver pair (i, j), compute the derivative of $J_{i,j}$ with respect to $\tau_{i,j}$, use adjoint interpolation to determine the gradients with respect to τ in the points A1, A2, etc. and then perform backward ray tracing with the perturbed traveltime equation to find the gradient with respect to the velocities. As a last step, adjoint interpolation to the spline nodes is required to determine the gradient of J with respect to the nodal values. The computational cost of this method is of the same order of magnitude as the forward ray tracing.

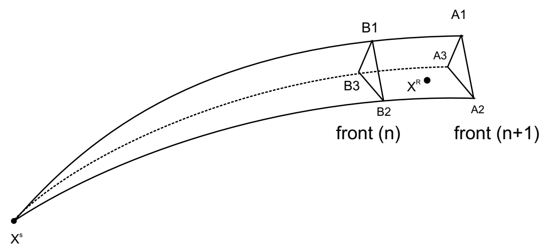

Figure 1: Three rays in the wavefront construction method constitute a ray tube for a given source position $\mathbf{x}_s$. Linear interpolation is used to find the traveltime at the receiver $\mathbf{x}_r$ from traveltimes at A1, A2, A3 and B1, B2, B3.

Synthetic example

As an illustration of the method, a synthetic example was considered. The 3D model has spline nodes on a $12 \times 12 \times 15$ grid, so the error functional depends on 2160 unknowns. The “observed” traveltimes were computed for a given model. Its geometry is shown in Fig. 2. Seven sources are used in example, indicated by stars. The receivers are placed on a 50×50 grid inside the black rectangle. Next, the minimisation of the traveltime error functional was carried out starting from a different model using a limited-memory BFGS method. The initial model was obtained from the orginal one by adding random values between 0% and 100% of the original node values. The decrease of the functional is shown in Fig. 3 as a function of the iteration count. The maximum error in the resulting spline coefficients was about 15%. The maximum error in the (interpolated) velocity was 12%. These errors are fairly large because not all parts of the model were illuminated. Figure 4 shows the error in the most interesting part of the model which is better illuminated. There, the errors were about 3%.

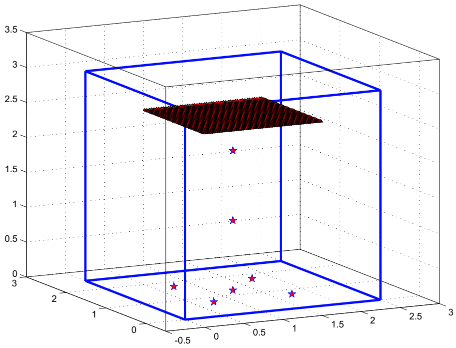

Figure 2: Geometry of the synthetic model. The source positions are indicated by stars. The receivers are placed on a 50×50 grid inside the black rectangle.

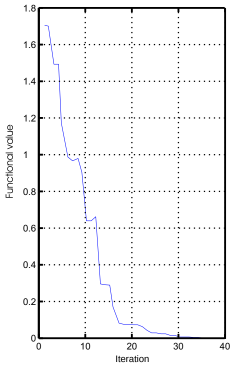

Figure 3: Convergence history of the error.

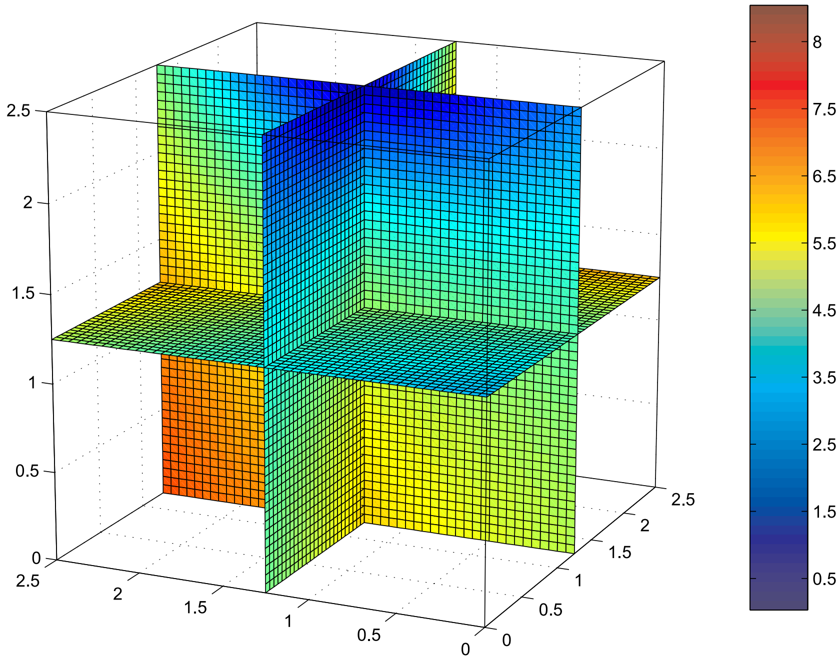

Figure 4: Errors in a subset of the model as a percentage.

References

[1] Bishop, T., Bube, K., Cutler, R., Langan, R., Love, P., Resnick, J., Shuey, R., Spindler, D., and Wyld, H., 1985, *Tomographic determination of velocity and depth in laterally varying media*, Geophysics **50**, 903–923.

[2] R.-E. Plessix, W.A. Mulder, and A.P.E. ten Kroode, 2002, *Automatic crosswell tomography by semblance and differential semblance optimization: theory and gradient computation*, Geophysical Prospecting **48**, 913–935.

[3] W.A. Mulder and A.P.E. ten Kroode, 2002, *Automatic Velocity Analysis by Differential Semblance Optimization*, Geophysics **67**, 1184–1191.

[4] R. Schnieder and M. Sambridge, 1992, *Ray perturbation theory for traveltimes and ray paths in 3-D heterogeneous media*, Geophys. J. Int. **109**, 294–322.

[5] Vinje, V., Iversen, E., and Gjøystdal, H., 1993, *Traveltime and amplitude estimation using wavefront construction*, Geophysics **58**, 1157–1166.

[6] Bulant P. and Klimes L., 1999, *Interpolation of ray theory traveltimes within ray cells*, Geophys. J. Int. **139**, 273–282.