Wave-equation traveltime inversion using 2D cross-correlation function

Olga Savasta*, Anton Ziolkowski, and Robert G. Pearce; The University of Edinburgh

Summary

An automatic procedure for velocity estimation is set up using the 2D cross-correlation function. Focusing the objectives of our inversion method on the reflected or diffracted waveforms, we modify the wave-equation traveltime inversion by introducing the cross-correlation function in the time and space domains simultaneously. The velocity model is updated using the current velocity model, the observed pressure field and the fields computed by reverse time propagation of two pseudoresidual functions acting as sources in a particular location. Full wave inversion could be incorporated in a further step in order to improve the resolution of the velocity estimation.

Introduction

The principal aim of this method is to automate the estimation of velocity models using reflected or diffracted seismograms recorded in a seismic reflection survey. The travel time is the key parameter for velocity and depth estimation, and the traveltime residual that maximises the cross-correlation function between the real and calculated seismograms can be used as an error criterion (Luo and Schuster, 1991; Zhou, 1995). This method treats adjacent traces as independent, whereas moveout between traces is required to resolve the ambiguity between velocity and depth for a given arrival.

We introduce the 2D cross-correlation function to tackle the offset dependence that is ignored by the 1D function. Instead of moving a real trace over a calculated trace (in 1D), the 2D function moves the full observed shot gather over the full calculated shot gather, giving a shift in offset as well as in time as a similarity index. We develop the procedure that updates the velocity model according to the method of steepest descent and will combine this with the full-waveform inversion method (Gauthier et al., 1986; Pica et al., 1990; Zhou et al., 1995). We present the theory.

Theory

The following analysis describes the procedure for finding a new velocity model using the 2D cross-correlation function as a connective function. As in wave-equation traveltime inversion the connective and the misfit functions are defined, and the wave equation is used to derive the perturbation of the misfit function with respect to the velocity.

Let \( p(x_t, t; x_s)_{\text{obs}} \) and \( p(x_t, t; x_s)_{\text{calc}} \) denote, respectively, the observed and calculated pressure seismograms at receiver location \( x_t \) due to a line source at location \( x_s \). The calculated seismogram satisfies the acoustic wave equation:

\[
\frac{1}{c(x)^2} \frac{\partial^2 p(x_t, t; x_s)}{\partial t^2} = \rho(x) \nabla \cdot \left( \frac{1}{\rho(x)} \nabla p(x_t, t; x_s) \right) = s(t; x_s),
\]

(1)

where \( s(t; x) \), \( \rho(x) \), and \( c(x) \) correspond to the source function, the density function and the wave speed, respectively. The 2D cross-correlation function is defined as a function of a space shift \( \sigma \) and a time shift \( \tau \) by

\[
f(\sigma, \tau; x_s) = \int dx_t \int dx_s \frac{p(x_t + \sigma, t + \tau; x_s)_{\text{obs}} p(x_t, t; x_s)_{\text{calc}}}{A(x_s)_{\text{obs}}},
\]

(2)

where \( A(x_s)_{\text{obs}} \) is the maximum amplitude of the observed seismogram. The best match between the observed and calculated seismograms is reached when the function \( f \) has a maximum

\[
f(\Delta \sigma, \Delta \tau; x_s) = \max \{ f(\sigma, \tau; x_s) \} | \sigma \in [-T, T], \tau \in [-X, X] \}.
\]

(3)

The derivatives in both dimensions should be zero and the following equations have to be satisfied simultaneously

\[
\begin{align*}
\left[ \frac{\partial f(\sigma, \tau; x_s)}{\partial \tau} \right]_{\sigma = \Delta \sigma, \tau = \Delta \tau} &= j = 0, \\
\left[ \frac{\partial f(\sigma, \tau; x_s)}{\partial \sigma} \right]_{\sigma = \Delta \sigma, \tau = \Delta \tau} &= j' = 0,
\end{align*}
\]

(4)

where
\[
\begin{align*}
&\int^{l}_1 \, d\tau \, \frac{1}{A(x_0)} \rho(x_{r} + \Delta \sigma, t + \Delta \tau; x_0) \rho(x_{r} + \Delta \sigma, t + \Delta \tau; x_0)_{obs} p(x_r, t; x_{s} \text{cal}), \\
&f' = \int^{l}_1 \, d\tau \, \frac{1}{A(x_0)} \rho'(x_{r} + \Delta \sigma, t + \Delta \tau; x_0) \rho(x_{r} + \Delta \sigma, t + \Delta \tau; x_0)_{obs} p(x_r, t; x_{s} \text{cal}).
\end{align*}
\]

Here, dots and primes indicate time and space derivatives, respectively.

The misfit function defined as
\[
S = \frac{1}{2} \sum (\Delta \sigma(x_s))^2 + \beta \Delta \tau(x_s)^2,
\]
has a minimum when the predicted seismogram best matches the observed seismogram (the parameter \(\beta\) is introduced in order to keep the dimensions of this equation consistent).

The steepest descent method can be used to find the velocity model \(c(x)\) that minimises \(S\). The new velocity field is calculated using the following equation:
\[
c(x)_{k+1} = c(x)_k + \alpha_k \gamma(x)_k,
\]
where \(\gamma(x)\) is the steepest descent direction of the misfit function \(S\), given by
\[
\gamma(x) = -a \frac{\partial S}{\partial c} = -a \left( \sum c_{i=1}^N \frac{\partial \sigma}{\partial c} \Delta \sigma(x_s) - \beta \sum c_{i=1}^N \frac{\partial \tau}{\partial c} \Delta \tau(x_s) \right) = \gamma_1 + \beta \gamma_2,
\]
and \(\alpha_k\) is a constant scaling factor (either estimated analytically or chosen by trial and error). The rule for implicit differentiation leads to a system of equations that have to be solved simultaneously. Then, using equations (4) and (5), and assuming that mixed second order partial derivatives commute, we obtain:
\[
\frac{\partial \Delta \sigma}{\partial c(x)} = \frac{1}{d_1} \left( \int^{l}_1 \, d\tau \, \frac{\partial \rho_{obs}}{\partial c(x)} \right) - \int^{l}_1 \, d\tau \, \frac{\partial \rho_{cal}}{\partial c(x)} \int^{l}_1 \, d\tau \, \frac{\partial \rho_{obs}}{\partial c(x)} \int^{l}_1 \, d\tau \, \frac{\partial \rho_{cal}}{\partial c(x)}
\]
\[
\frac{\partial \Delta \tau}{\partial c(x)} = \frac{1}{d_2} \left( \int^{l}_1 \, d\tau \, \frac{\partial \rho_{obs}}{\partial c(x)} \right) - \int^{l}_1 \, d\tau \, \frac{\partial \rho_{cal}}{\partial c(x)} \int^{l}_1 \, d\tau \, \frac{\partial \rho_{obs}}{\partial c(x)} \int^{l}_1 \, d\tau \, \frac{\partial \rho_{cal}}{\partial c(x)}
\]

where \(d_1\) and \(d_2\) are, respectively:
\[
d_1 = A(x_0) \frac{\partial f'}{\partial \Delta \sigma} - A(x_0) \frac{\partial f}{\partial \Delta \sigma} \frac{\partial f}{\partial \Delta \tau} - \int^{l}_1 \, d\tau \, \frac{\partial \rho_{obs}}{\partial c(x)} \int^{l}_1 \, d\tau \, \frac{\partial \rho_{obs}}{\partial c(x)}
\]
\[
d_2 = A(x_0) \frac{\partial f'}{\partial \Delta \tau} - A(x_0) \frac{\partial f}{\partial \Delta \tau} \frac{\partial f}{\partial \Delta \sigma} - \int^{l}_1 \, d\tau \, \frac{\partial \rho_{obs}}{\partial c(x)} \int^{l}_1 \, d\tau \, \frac{\partial \rho_{obs}}{\partial c(x)}
\]

The derivative of \(p(x_r, t; x_{s} \text{cal})\) with respect to \(c(x)\) is the Frechet derivative (Tarantola, 1987)
\[
\frac{\partial p(x_r, t; x_{s} \text{cal})}{\partial c(x)} = -\frac{2}{c(x)^2} \delta(x, t; x, 0) * \rho(x, t; x_{s} \text{cal})
\]
with the asterisk representing time convolution. \(p(x, t; x_{s} \text{cal})\) satisfies equation (1) and the corresponding Green’s function obeys
\[
\frac{1}{c(x)^2} \frac{\partial^2 g(x, t; x', t')}{\partial t'^2} - \rho(x) \nabla \cdot \left( \frac{1}{\rho(x)} \nabla g(x, t; x', t') \right) = \delta(x - x') \delta(t - t');
\]
\[
g(x, t; x', t') = 0; \quad \gamma(x, t; x', t') = 0; \quad t \leq t'.
\]
Defining the pseudoresiduals $\delta \sigma$ and $\delta \tau$ as

$$\delta \sigma(x_r, t; x_s) = -\frac{2}{d_t} p'(x_r + \Delta \sigma, t + \Delta \tau; x_s) \Delta \sigma(x_s), \quad (15)$$

$$\delta \tau(x_r, t; x_s) = -\frac{2}{d_t} p(x_r + \Delta \sigma, t + AT; x_s) \Delta \tau(x_s), \quad (16)$$

and using the properties of the Green’s function we can write

$$\gamma_1 = -\frac{1}{c(x)^3} \sum_s \int dt \int dx_r \hat{p}(x_r, t; x_s) \hat{Q}(x_r, t; x_s)$$

$$- \int dt \int dx_r \hat{p}(x_r, t; x_s) \hat{R}(x_r, t; x_s) \int dt \int dx_r \hat{p}_{obsPcal} \frac{\Delta \sigma}{AT}, \quad (17)$$

$$\gamma_2 = -\frac{1}{c(x)^3} \sum_s \int dt \int dx_r \hat{p}(x_r, t; x_s) \hat{R}(x_r, t; x_s)$$

$$- \int dt \int dx_r \hat{p}(x_r, t; x_s) \hat{Q}(x_r, t; x_s) \int dt \int dx_r \hat{p}_{obsPcal} \frac{\Delta \tau}{\Delta \sigma}, \quad (18)$$

with

$$\hat{Q}(x_r, t; x_s) = \delta \sigma(x_r, t; x_s), \quad (19)$$

$$\hat{R}(x_r, t; x_s) = \delta \tau(x_r, t; x_s), \quad (20)$$

and

$$Q(x_r, t; x_s) = g(x_r, t; x_s), \quad (21)$$

$$R(x_r, t; x_s) = g(x_r, t; x_s), \quad (22)$$

Finally, the steepest descent direction of the misfit function $S$, equation (8) is given by

$$\gamma = \gamma_1 + \beta \gamma_2 =$$

$$-\frac{1}{c(x)^3} \sum_s \int dt \int dx_r \hat{p}(x_r, t; x_s) \hat{Q}(x_r, t; x_s) \left[ 1 - \beta \int dt \int dx_r \hat{p}_{obsPcal} \frac{\Delta \sigma}{AT} \right]$$

$$+ \int dt \int dx_r \hat{p}(x_r, t; x_s) \hat{R}(x_r, t; x_s) \left[ \beta - \int dt \int dx_r \hat{p}_{obsPcal} \frac{\Delta \tau}{\Delta \sigma} \right] \quad (23)$$

where $Q(x_r, t; x_s)$ and $R(x_r, t; x_s)$ represent the pressure fields calculated when the pseudoresiduals $AT$ and $\Delta \sigma$ act as sources at the location $x_r$.

We remark on the similarity between the two terms of this expression and the analogous equation in the wave-equation traveltime inversion method. This fact simplifies the implementation of the new inversion code.

Examples and Conclusions

Using the example of a single point diffracon, we will show that the 1D function fails to resolve the velocity/depth ambiguity and converges to a local minimum. We will show how this failure is resolved using the 2D function, and we will present additional synthetic and real examples.

Acknowledgments

Olga Savasta thanks Intevep, S.A. for funding her Ph.D. research at the University of Edinburgh. We thank Dr. Sergey Zatsepin for checking the mathematics of this work.
References


Physical criteria for adaptive seismic tomography
Aldo Vesnaver and Gualtiero Böhm, OGS

Introduction

Traveltime inversion is a cost-effective tool for estimating the velocity model for complex geological structures. However, such estimates may be non-unique, especially if the velocity field is inverted by regular grids. The quality of the tomographic images may be degraded by the poor local ray coverage and by the shape mismatch between the velocity anomalies and the grid (Vesnaver 1994). We can master these mathematical and geometrical drawbacks by adapting the grid shape and its parameter number to the available rays (Böhm et al. 1996).

Here we show that physical reasons support the use of irregular grids too, as the wave length and Fresnel radius variations due to velocity changes and absorption effects. We relate the sampling rate to the SVD truncation technique and to the noise. Also, we prove that the local resolution is not symmetrical with respect to sources and receivers, if we allow for the frequency band width of the seismic signals.

Experimental errors and SVD analysis

In the linearized inversion, we related the vector $t$ of traveltimes to the vector $u$ of model slowness by the tomographic matrix $A$:

$$ t = A u = U W V^T u, \quad (1) $$

where the matrices $U$, $W$ and $V$ are the singular value decomposition (SVD) of $A$. These matrices provide the solution for the linear system (1):

$$ u = V W^{-1} U^T t, \quad (2) $$

A popular trick to stabilise the inversion is to zero the elements of the diagonal matrix $W^1$ corresponding to the smallest or null singular values in $W$. The cutoff threshold is chosen often arbitrarily; however, it should be related to the experimental errors. The traveltime errors are larger or equal to the sampling rate, in the most favourable cases, i.e. some milliseconds. Since the traveltimes are of the order of a few seconds, the vector $t$ in (2) bears about 3 significant digits. The same is true for matrix $A$ (and so for $U$, $W$ and $V$). Its elements are the length-of-ray paths in the tomographic grid; the distances are measured with the metre precision, and are of the order of kilometres. Therefore, it is vain to consider in the solution (2) the singular values that are 4 or 5 orders of magnitude smaller than their maximum one.

Figure 1 is a synthetic model, composed by a layered medium with a central salt body. The velocity ranges from 1500 m/s in the upper layer to 3 100 m/s in the lowest one, and reaches 5300 m/s in the salt body. We placed sources and receivers at the surface and in two wells at the model sides (Fig. 2). The interval between adjacent sources or receivers is 25 m both at the surface and in the wells. We traced 1520 rays, simulating a combination of two VSP’s and a cross-well survey. We obtained a quite good image by a SIRT algorithm (Fig. 3) in a regular grid, composed by 560 squared pixels (i.e. 20 x 28) with a side length of 25 m. The normalised logarithmic curve of the singular values (Fig. 4) decays sharply below -2 at nearly the 520th term: so the solution depends mainly on the singular values of the first three orders of magnitude with respect to their maximum one. Often, one set the threshold for the SVD truncation at this point. Figure 5 shows that this is a good choice only in absence of noise. By extending the range of the considered singular values from 1 to 4 orders of magnitude, we increase the resolution but also the instability. A good tradeoff could be 3 orders in absence of noise (left) and 2 orders with noisy traveltimes (right). We remark that these values are the numbers of significant digits in the input data.

Geometrical spreading and absorption

Traveltime inversion is more cost-effective than diffraction tomography, because it relies on ray tracing. This forward modelling approach is much cheaper, but cannot allow for the limited bandwidth of the seismic pulses, the anelastic absorption of the Earth and the local variations of resolution due to velocity anomalies.

We can distinguish two interfaces along a ray when their distance is larger than a quarter of wave length of the dominant frequency of the picked seismic signals. To control the lateral resolution physically available, we can extend the use of the Fresnel ray $R$. This is the resolution limit in a homogeneous medium as a function of the depth $d$ and the dominant wavelength $A$:

$$ R = (d \lambda / 2)^{1/6}, \quad (3) $$

We substituted the depth with the distance from the source measured along the ray path. Furthermore, we assumed that the dominant frequency of the seismic pulses (and so the corresponding wavelength) decreases linearly, due to the anelastic absorption as a function of the distance $d$:

$$ f(d) = f(0) - d K / (Q v), \quad (4) $$

where the constant $K$ depends on the signal band width, $Q$ is the Q factor and $v$ is the local velocity. Naturally, both (3) and (4) are based on simplifying assumptions, but they can provide us a better physical insight. Figure 6 is a plot of the average Fresnel radius in the regular grid, for a dominant frequency of 60 Hz and the ray distribution in Figure 2. We assumed here that no absorption occurs. Since local velocity and wave length...