A tomographic velocity inversion for unstacked data

William S. Harlan and Robert Burridge

Abstract

Tomography will resolve intermediate interval velocity changes better than amplitude inversions or root-mean-square methods. Velocity information from reflections off continuous beds appear over offset, h, information from diffractions over midpoint, y. Scattering surfaces with high curvature contain the most velocity information at depth; thus, the data reduction used should allow arbitrarily shaped scatters.

By performing local three-dimensional slant stacks in the seismic data cube, over midpoint, offset, and time, one may recognize segments of wavefronts as sufficiently coherent for signal. From the arrival angles, midpoints, offsets, and total travel-times of these segments, one may assign corresponding simple raypaths. These raypaths make no assumption of the shape of the scattering surface. Simple raypaths suffice, without later perturbation, for our resolution. The data cube reduces to a list of raypath parameters, $\bar{d}$. Each genuine event produces a family of raypaths, whose spatial distribution correspond to the accuracy with which one can estimate exit angles. This family of raypaths preserves resolution information through the non-linear conversion to model coordinates and allows a robust solution to the following least-squares inversion. We may emphasize diffraction information by first estimating and removing continuous bed reflections with local two-dimensional slant stacks.

One defines an earth model, $\bar{g}$, as a partition of interval slownesses. A transform, L, which sums this model along raypaths found in the data should reproduce the travel-times in $\bar{d}$. The best $\bar{g}$ minimizes $(L\bar{g} - \bar{d})^2$. We easily find the adjoint of L and a steepest descent algorithm for the best slowness model.

Introduction

The stacking of common-midpoint gathers and the migration of point scatterers indicate only very low-frequency velocity changes. Abrupt changes, on the order of meters, seem potentially best extracted from the amplitudes of reflections, by inverting residual scattering operators. Intermediate velocity changes, whose vertical and horizontal periods are greater than say 50 meters and less than 500 meters, do not suit either method, yet these changes span those most usable for geologic interpretation. Tomography methods, seem most effective for resolving these intermediate changes--just as they already resolve those medical structures too gradational to create reflections but too rapid for crude root-mean-square averaging.

Because direct transmission paths are not available, we must make use of primary scattered waves, those from reflections of continuous beds and those from diffractions off simple scatterers (laterally unpredictable) such as bed truncations and point scatterers. Single common-shot or midpoint gathers contain velocity information largely from continuous reflections--information which is weakened at great depths by limited offsets. Common offset gathers, however, contain velocity information only from simple scatterers--information which is equally strong at all depths because of the arbitrary range of midpoints available. Unfortunately simple scatterers appear less frequently than do continuous reflections. Between these two extremes lie events scattered from surfaces of varying curvature. Those with greatest curvature contain the most velocity information at depth. A tomography method should use both sources of information with perhaps a bias towards simple scatterers. The reduction of the data cube prior to the tomographic inversion must make no assumptions on the shapes of scatterers.

Reducing the data cube

Before applying the principles of tomography we must reduce the vast cube of seismic data to a statistical order appropriate to the resolution we can hope to achieve. Tomography requires a geometric raypath and a total travel-time for any event to be used. Each raypath individually describes some average or root-mean-square estimate of the velocity along the raypath; collectively they constrain perturbations of velocity where the raypaths cross. Let us then extract from the data cube all primary raypaths.

We may estimate the exit and arrival angle of raypaths by making three-dimensional slant stacks of small windows of the data cube. The original data, a function of geophone position, g, and shot position, s, are uncomfortably skewed. Let us sort the data $d(y,h,t)$ into midpoint, $y \equiv (g + s)/2$, versus offset, $h \equiv (g - s)/2$. We begin with a summing technique closely related to Zavalishin's C.D.R. imaging of reflectors (Sword, 1981). We select a narrow cylindrical window about some (y_0,h_0) and sum along planes with the transformation

$$d'(p_y,p_h,\tau,y_0,h_0) = \int_{y_0-\Delta y}^{y_0+\Delta y} \int_{h_0-\Delta h}^{h_0+\Delta h} d[y,h,t=\tau+p_y(y-y_0)+p_h(h-h_0)]\, dh dy \ , \qquad (1)$$

plus a "rho" filter.

We follow (1) with a rho filter which multiplies by ω^2, the square of the angular frequency for τ. We implement the above in two passes, over y then h, each transformation being just a slant-stack transformation. Events should appear locally planer in the original data and should then map roughly to points after transformation. The inverse transformation for this window takes the form of the adjoint of (1), without the rho filter.

$$d(y,h,t) = \int\int d'[p_y,p_h,\tau=t-p_y(y-y_0)-p_h(h-h_0),y_0,h_0] dp_y\, dp_h \ .$$

Two points then invert into two additively superimposed planes. The sampling of p_y and p_h should be sufficient to avoid aliasing of events at the edge of the window. Thus the narrower the window, the more sparsely one may sample dips. An inverse transform for the entire data set may take a simpler form.

$$d(y,h,t) = \int\int d'(p_y,p_h,\tau=t,y_0=y,h_0=h)\, dp_y\, dp_h \ .$$

(1) decomposes the data set into cubes containing narrow ranges of dips. These cubes add together to reconstruct the original data. For a more extensive treatment of slant stacks see the accompanying paper on signal/noise separation, this volume.

Because noise-spikes map to planes, noise should diffuse overall, becoming more gaussian and more "defocused." (We use this latter term to mean that samples become more statistically dependent because of this event.) Signal, on the other hand, becomes more non-gaussian and better focused. By examining local statistics of the data before and after transformation, one may recognize and extract those transformed samples containing the highest concentration of signal. (See the accompanying paper.) The power of distinguishing signal from noise by such summing accounts for the noise-free appearance of Zavalishin's imaged reflectors. Zavalishin preferred to assign one event to each data window. We prefer to allow an arbitrary number, so long as little of the energy can result from the summing of incoherent noise. Note that one need make no assumptions about the shape of scattering events. Point scatterers and arbitrarily dipping and curving beds all qualify; events may even overlap.

Our extraction must preserve information on resolution in order that non-linear mapping and least-squares processing may follow. The width of our window fixes the resolution of transformed events. Points will show a resolution corresponding to the angles over which planes add constructively within the chosen window, analogous to the Fresnel zone. The sampling of p's should avoid the aliasing of traces and allow inversion. The extraction must capture a family of nearby points for each event. Later least-squares processing will attempt to fit each of these nearby points as an independent event. We shall also see that each point may represent a raypath, a highly non-linear sort of mapping. Thus, we must retain this family of points to preserve the information on resolution and allow for a robust least-squares solution.

One may emphasize diffraction information by first removing continuous bed reflections. These reflections appear roughly linear in common offset sections. The slopes of diffractions and noise change rapidly. Thus, those events focusing best after local two-dimensional slant stacks over offset will represent the strongest bed reflections. The extracted events may be inverted and subtracted from the original data, leaving only weak bed reflections and all diffractions and noise (as described in the accompanying paper).

We may assign a raypath to each extracted event. All events in one cylindrical window are assumed to have been generated at $s = y - h$ and received at $g = y + h$. The (p_y, p_h, t) at which an event focuses indicates the total travel-time and the first derivatives of travel-time with respect to shot and geophone positions. That is

$$t_s = (t_y - t_h)/2$$

$$t_g = (t_y + t_h)/2$$

Subscripts designate partial derivatives. These derivatives constrain the entrance and exit angles of the corresponding raypath. One may trade resolution in (s,g) for resolution of the derivatives (t_s, t_g) by increasing Δy and Δh. Assume the scattering or reflecting point to have the lateral position and depth of (x,z). Assuming straight raypaths, one may calculate this point and an average velocity, v, for each event. The geometry of the event requires

$$\frac{(g - x)^2}{z^2} = \frac{v^2 t_g^2}{1 - v^2 t_g^2} \quad ; \quad \frac{(s - x)^2}{z^2} = \frac{v^2 t_s^2}{1 - v^2 t_s^2} ; \tag{2}$$

$$(s - x)^2 + (g - x)^2 + 2z^2 = v^2 t^2 . \tag{3}$$

These three equations uniquely determine the three unknown parameters (x, z, v). (3), however, constrains the velocity globally to fit travel-time perturbations due to local velocity anomalies. In order not to disguise the anomalies we hope to uncover, we should instead assume a background velocity which varies more slowly, and invert for the raypath constrained by (2) or its depth variable equivalent. Perturbations must then account for inconsistencies in travel-times.

The reduced data will be a list of sufficiently strong events made of raypath parameters, total travel-times, and weights. Because unrelated events deserve equal emphasis, a smooth automatic gain control should be applied to the extracted points. A smooth gain preserves local weights between the family of points which describe a single event. We save the gained amplitude, w, as a weight for future least-squares fitting. Hereafter, "data" refers to the list of parameters--for straight raypaths (s, g, x, z, t, w).

Iterative estimation of interval slownesses

We define the inverted earth cross-section as a partition of interval slownesses--slowness being the reciprocal of the interval velocity times the height of the bin. One may require bins to be coarse enough for some raypaths to overlap in each slowness bin, or one may choose a finer partition and require smooth variations. For the present we shall assume no *a priori* statistical dependence between the bins. The up- and down-going raypaths may each be assumed to cross one bin at any depth level (at most a linearly weighted combination of two).

We may safely omit raypath perturbations and thereby keep each iteration linear. Fermat's principle of least time allows one to say that small local perturbations in velocity, or slowness, will not alter the raypath. For reasonable finite perturbations in slowness, raypaths move less than we can hope to resolve. In fact, straight raypaths, or those allowing only for linear velocity gradations, suffice (see Toldi, 1983). Thus, the raypaths described by the reduced data require only the roughest prior knowledge of interval velocities -- r.m.s. velocities as available from (2) and (3) or from a stack over hyperbolas as in conventional velocity analyses.

We may now pose a simple linear system to be solved for slowness perturbations. Define L as the transformation which sums the slowness model along the raypaths specified by the data, producing total travel-times. Though allowing for quite arbitrary geometries, such a transformation should cost quite little: it need only accommodate those raypaths actually found in the data. Define $\bar{g}$ as the earth slowness model and $\bar{d}$ as the reduced data. We define the ideal slowness model as that $\bar{g}$ which minimizes

$$\min_{\bar{g}} \sum_r (w_r (L\bar{g}_r - d_r))^2 \ . \tag{4}$$

A sample of reduced data, d_r, assigns a travel-time perturbation to each raypath r, that is, the time not accounted for by the background velocity. A sample from the earth model, $g_{x,z}$, assigns a slowness perturbation (actual minus background) to each bin with coordinates (x,z). Additionally we preserve the weights of the extracted events (based on local amplitude variations as previously explained) as some function w_r. Subscripts here are discrete. We designate a raypath r by the function $f_{x,z}^r$, equal to one for bins lying on the raypath and zero elsewhere. $f_{x,z}^r$ should contain any weighting of nearby bins.

We now write the forward transform as

$$d_r = \sum_x \sum_z f_{x,z}^r a_{x,z}^r g_{x,z} = L\bar{g} \ .$$

The function $f_{x,z}^r$ selects for summing those bins lying on the raypath. $a_{x,z}^r$ adjusts for the angle (obliquity) with which a ray passes through a bin (e.g. $a_{x,z}^r \approx 1/\cos\theta_{x,z}^r$ where the ray deviates an angle $\theta_{x,z}^r$ from the vertical). For straight raypaths only two such angles exist.

To find the adjoint of L we first write the scaler product as a weighted sum over r.

$$<\bar{d} \,|\, L\bar{g}> = \sum_r w_r d_r \left(\sum_x \sum_z f_{x,z}^r a_{x,z}^r g_{x,z}\right)$$

$$= \sum_x \sum_z g_{x,z} \left(\sum_r w_r f_{x,z}^r a_{x,z}^r d_r\right) = <L^*\bar{d} \,|\, \bar{g}>$$

$$\rightarrow L^*\bar{d} = \sum_r w_r f_{x,z}^r a_{x,z}^r d_r$$

The adjoint operation, L^*, distributes a total travel-time along a raypath according to the weights--treating each raypath independently of others.

Steepest descent iterations proceed as follows.

1. Use the results of some r.m.s velocity analysis (those found in reducing the data) and Dix's equations to produce the first estimate of the model, $\bar{g}_1$.

2. Calculate the gradient of expression (4) at the previous best $\bar{g}_i$.

$$\nabla\bar{g}_i = L^*L\bar{g}_i - L^*\bar{d} \ .$$

3. Update the model.

$$\bar{g}_{i+1} = \bar{g}_i - \alpha_i \nabla\bar{g}_i$$

$$where\ \alpha_i = \frac{<\bar{g}_i \,|\, \bar{g}_i>}{<L\bar{g}_i \,|\, L\bar{g}_i>}$$

4. Repeat 2. and 3. until convergence is reached.

REFERENCES

Aki, K., and P. Richards, 1980, Quantitative seismology: San Francisco, W.H. Freeman and Co.

Harlan, W., Jon Claerbout, and Fabio Rocca, 1983, Extracting velocities from diffractions, SEP-35, p.107-126.

Harlan, W., 1983, Signal/noise separation and velocity estimation, SEP-37.

Luenberger, D., 1973, Introduction to linear and nonlinear programming: Reading, Mass., Addison-Wesley Pub. Co.

Sword, C., 1981, Controlled directional receptivity--a Russian method of pre-stack migration, SEP-26, p.289-295.

Toldi, J., 1983, Lateral velocity anomalies--a model study, SEP-35, p.3-17.