Anisotropic travel-time inversion with error bars: theory

Joe Dellinger

ABSTRACT

In most geophysical inversion problems, there is a bothersome null space which intrudes into the desired answer. If the space we invert in is visualizeable, like a 2D scalar velocity field, null-space artifacts may be recognized and ignored. Here I examine the problem of inverting for anisotropic elastic parameters in a layered medium, where the interplay of anisotropic elastic constants can create a large null space. Since the space of elastic parameters is hard to visualize, null-space artifacts are hard to recognize and can make the results worthless. The solution is to take a probabilistic approach to the problem, by calculating not only a set of best-fitting elastic constants but also error bars.

INTRODUCTION

Shooting a multi-offset multi-component VSP is expensive, so often only a few shot points are used. Given such a limited ray coverage, inverting for the complete velocity field around the well is not feasible. We can greatly reduce the number of parameters by assuming flat isotropic layers, but this is perhaps unnecessarily restrictive. If a wide enough range of propagation angles is available, it is reasonable to keep the flat-layer assumption but allow each layer to be anisotropic.

Several authors, most recently Jan Douma (1988), have performed travel-time inversions for similar experiments using synthetic data, and proven such an inversion to be possible in principle by recovering input synthetic elastic parameters with high accuracy. I claim there is no hope of such an ideal inversion working for real data, due to the ill-conditioning of the problem. The slightest noise in the input travel-time picks can drastically change the inverted elastic constants.

This does not mean the inverted results are useless; it means we must do more than just invert data for best-fitting model parameters. We must also recognize how well each element of the inverted result is constrained by the input data.

RAY TRACING IN FLAT ANISOTROPIC LAYERS

Why flat layers?

General ray tracing may seem difficult enough in isotropic media, but things get *really* messy in three-dimensional anisotropic media. Even for such a "simple" case as transverse isotropy in three dimensions, complicated effects such as shear wave coupling are significant (Chapman and Shearer, 1989). To avoid such complications in this paper, I will limit my discussion of ray tracing to the kinematics of direct rays in two-dimensional anisotropic media.

Phase and group

Imagine an ideal plane wave propagating in an anisotropic media. We will call the velocity of this plane wave the *phase* velocity, and the direction of propagation the *phase* direction. Because an ideal plane wave is homogeneous and infinite, we cannot tell if it is "slipping sideways" as it propagates. For this reason the phase velocity direction is necessarily perpendicular to the planar wavefront.

Now imagine a bit of energy riding along on the plane wave. We call the velocity of this bit of energy the *group* velocity, and the direction of propagation the *group* direction. In isotropy the phase and group velocity and direction are always the same, but in anisotropic media they may be quite different.

To ray trace in a layered anisotropic medium we need to keep track of both phase and group information in each layer. Phase velocities and directions are needed to match wavefronts across interfaces, while group velocities and directions are needed to track where the energy in a wavefront goes as it traverses a layer.

Figure 1 shows an example of ray tracing in a flat-layered medium with 2 layers. The lower layer is isotropic, and the upper layer is anisotropic. For rays traveling within about 45° of vertical, the upper medium appears to be faster than the lower medium, while for rays nearer the horizontal the upper medium appears to be slower. In fact both the phase and group velocities are faster in the upper anisotropic medium for any direction of propagation.

Across interfaces

The familiar statement of Snell's law in isotropic media, $\sin\theta/v$ = constant, is a consequence of the continuity of wave fields across interfaces. A more general way of expressing Snell's law for plane waves is ω = constant (the temporal frequency must be the same on both sides of the interface) and K_x = constant (the phase

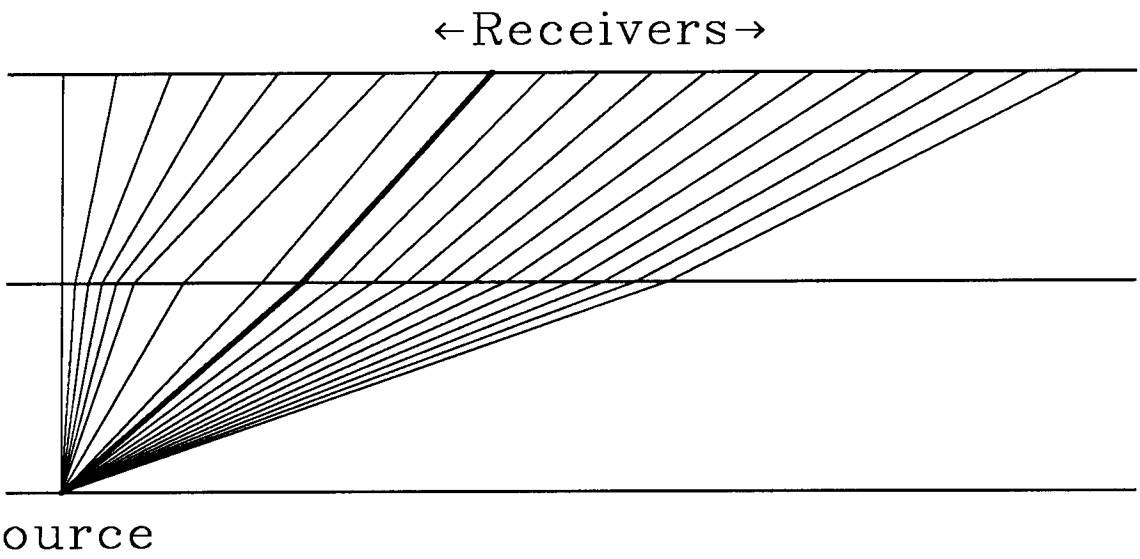

FIG. 1. An example of ray tracing in anisotropic media. The lower layer is isotropic, and the upper layer is anisotropic. Despite the way some of the rays refract upwards at the interface, the upper layer is faster for all directions of propagation. Figure 2 shows how this can happen.

velocity along the interface must be the same on both sides). To ray trace across an interface between anisotropic media, we have to find K_z on each side given K_x and ω.

For any set of elastic constants, the Christoffel equation gives ω as a function of K_x and K_z (Auld, 1973), which is not exactly what we want. Fortunately in the transverse isotropic case there is a reasonably simple analytical formula for K_z as a function of K_x and ω. The process of solving for the K_z's in two different media given a constant K_x is shown graphically in Figure 2.

Across layers

The group velocity and direction can be expressed parametrically in terms of the associated plane wave parameters K_x, K_z, and ω (Dellinger and Muir, 1985). Again for transverse isotropy there is a reasonably simple analytical solution. Graphically, the group direction for a particular ray is given by the normal to the dispersion relation curve, as shown in Figure 2. This figure shows how it is possible for the thick ray in Figure 1 to enter a faster medium but still have the *group* direction bend towards the layer normal.

TRAVEL-TIME INVERSION

Given a stack of layers with known elastic constants, a source at the bottom, and a group of receivers on the surface, we can determine travel times for each

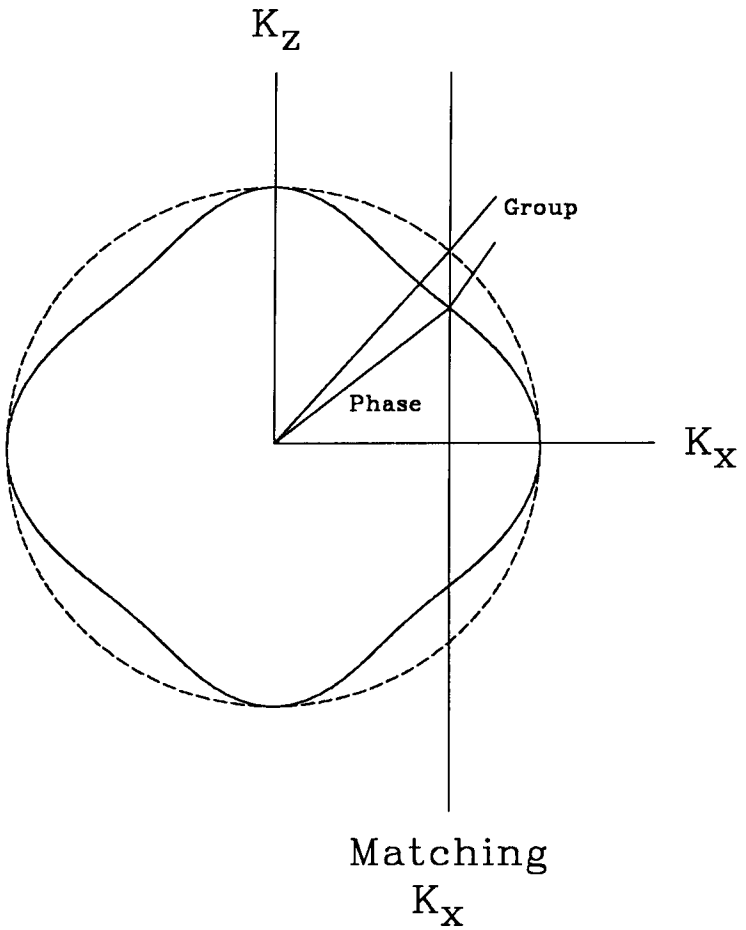

FIG. 2. Dispersion relations for the two media in Figure 1. The dashed circle is for the lower, isotropic media, and the solid "inflated diamond" is for the upper, anisotropic media. Dispersion relations relate K_z to K_x for fixed ω; they can also be thought of as plots of phase slowness versus phase direction. The line of constant K_x corresponds to the thick ray in Figure 1. The line segments labeled "Phase" and "Group" show the phase and group directions for this ray in the two media.

of the direct wave modes {P, SV, SH} using the ray tracing methods derived in the previous section. The inversion problem is to find the set of elastic constants that minimizes the "total error" between these predicted travel times and the times picked off real data.

Typically the "total error" is considered to be the sum of the squares of the travel-time errors for each ray:

$$\text{Error}_{\text{total}} = \sum_n (\text{Time}_{\text{actual}} - \text{Time}_{\text{predicted}})^2. \tag{1}$$

This is the square of the L_2 norm of the travel-time errors.

Travel-time error to pdf

How can we cast equation (1) in terms of probability theory? Assume each travel-time pick T_0 is contaminated by errors that have a Gaussian distribution. The probability distribution function (pdf) giving the chance that a given modeled travel-time T_M is consistent with a corresponding picked T_0 is then

$$\text{pdf}(T_M) = k\, e^{-\lambda(T_M - T_0)^2}, \tag{2}$$

where λ is a parameter controlling the magnitude of the Gaussian noise, and k is a normalization factor required to make sure the total probability integrates to unity.

Now assume each travel-time pick is independent, so that the individual probabilities multiply. Then the probability density function giving the chance that a set of modeled travel times $\{T_{M_n}\}$ is consistent with the set of picked travel times $\{T_{0_n}\}$ is

$$\text{pdf}(\{T_{M_n}\}) = \prod_n k_n\, e^{-\lambda_n (T_{M_n} - T_{0_n})^2}. \tag{3}$$

This can be rewritten

$$-\log(\text{pdf}(\{T_{M_n}\})) = -\sum_n \log k_n + \sum_n \lambda_n (T_{M_n} - T_{0_n})^2, \tag{4}$$

which just happens to be a generalized form of equation (1). The L_2 norm inversion is thus equivalent to finding the most likely set of elastic constants assuming that each travel-time pick is contaminated with independent Gaussian errors.

Input pdf to output pdf

What does an error function like the one defined in equation (1) look like? Figure 3 shows rays through a two-layered model, with the lower layer isotropic ($C_{11} = 1.$, $C_{33} = 1.$, $C_{44} = .3$, $C_{66} = .3$, $C_{13} = .4$) and the upper layer normalized Greenhorn shale (Jones and Wang, 1981) ($C_{11} = 1.5$, $C_{33} = 1.$, $C_{44} = .24$, $C_{66} = .47$, $C_{13} = .47$). The input data set consists of the travel times for the 30 rays shown. Figure 4 shows two different two-dimensional cross sections through the ten-dimensional log probability density function for this model.

Note that for both plots in Figure 4 the contours are evenly spaced, which means the minimum can be approximated very well by a quadratic surface. This is not surprising; the first derivative has to be zero because we are looking at a minimum (Dellinger and Muir, 1986). The general form of the minimum should be

$$\text{F}(X_{\min} + \Delta X) = \Delta X^T H \Delta X + \text{F}(X_{\min}), \tag{5}$$

where X is the vector of the most likely elastic constants and ΔX is a perturbation away. H is the second derivative matrix; it must be positive definite since we are at a minimum.

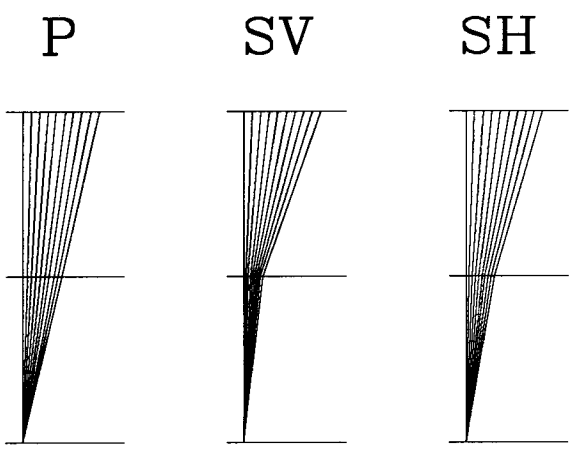

FIG. 3. Rays for P, SV, and SH waves through a simple two-layered model. To simulate typical VSP geometry with insufficient offset for good resolution of anisotropy, the maximum offset is one fourth of the total depth.

Output pdf to error bars

If we interpret equation (5) in probabilistic terms, we get

$$-\log(\text{pdf}(\{T_{M_n}\})) = \Delta X^T H \Delta X + \text{Constant} \tag{6}$$

or

$$\text{pdf}(\{T_{M_n}\}) = \text{pdf}(\{T_{0_n}\})\, e^{-\Delta X^T H \Delta X}. \tag{7}$$

This defines a Gaussian probability distribution function. The eigenvectors of H determine the independent "eigenparameters" of the solution and the eigenvalues of H determine their associated uncertainties.

In general the "eigenparameters" of H will be complicated linear combinations of all the inversion parameters. To get an error bar on a specific parameter such as "C_{13} in layer 14", we must project the uncertainty contributions from all the eigenvectors of H onto the desired inversion parameter.

Errors in the errors

Equation (7) is only a good approximation to equation (4) in the region of maximum likelihood. How valid is this approximation globally? If the pdf decays rapidly in amplitude away from its peak, the total integrated probability over the low probability "tails" of the pdf should be small enough that the chance of the solution actually lying in this region can be neglected. Furthermore, the pdf in equation (4) is probably only valid in the region of the solution anyway. The major difficulty is probably not the Gaussian approximation of the output pdf, but coming up with reasonable pdf's for the input travel-time picks in the first place.

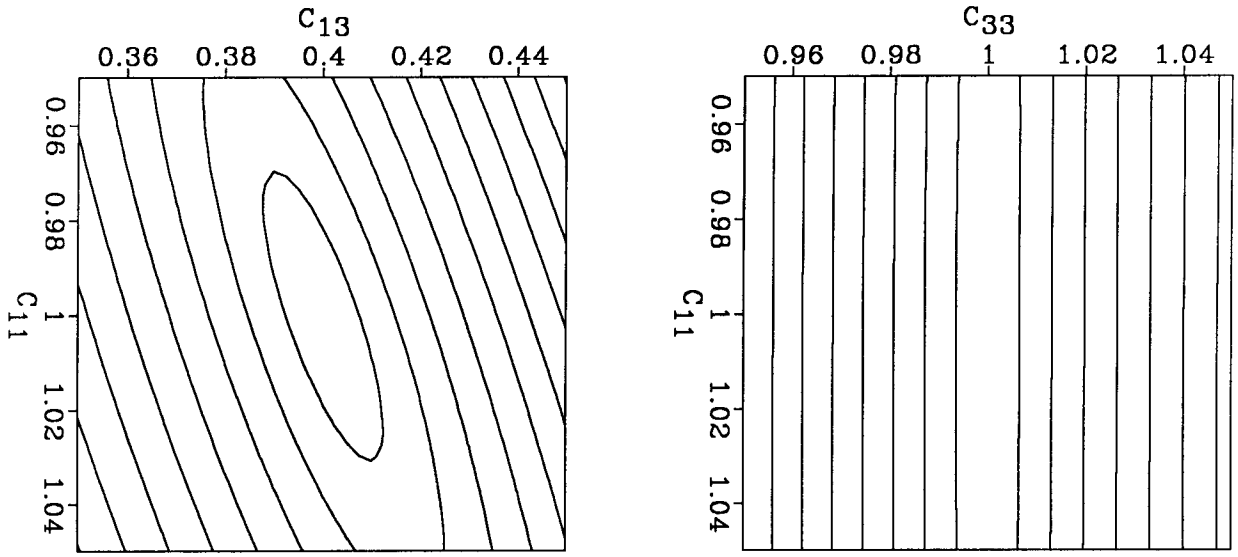

FIG. 4. Contour plots of the *square root* of the "Total Error" function as defined in equation (1) for the model in Figure 3 as two elastic constants in the lower isotropic layer are varied and all other parameters are held constant. The contours on the right plot are at 10 times the interval of the contours on the left plot.

AN EXAMPLE

What do the eigenvalues and eigenvectors of H look like for the example in Figures 3 and 4? To keep the problem small, we will assume that all the elastic constants in the top layer are perfectly known, and only allow the elastic constants in the lower isotropic layer to vary. For this problem we obtain

Eigenvalue	C_{11}	C_{33}	C_{44}	C_{66}	C_{13}
1353.27	0.	-.06	1.00	0.	0.
150.22	-.02	1.00	.06	-.04	-.03
16.98	-.22	0.	0.	-.54	.81
11.70	-.54	.03	0.	.76	.37
1.77	.81	.04	0.	.35	.46

(Note that while all the eigenvectors were normalized to full accuracy, only two digits after the decimal are shown.)

The eigenvector associated with the largest eigenvalue is almost pure C_{44}, which corresponds to the vertical velocity of both SV and SH waves. The eigenvector associated with the next largest eigenvalue is almost pure C_{33}, which corresponds to the vertical velocity of P waves. The difference in magnitude between the first two eigenvalues is accounted for by there being two S wave input data picks for each P one.

The eigenvectors for the next largest eigenvalues are dominated by C_{13} and C_{66}, respectively, which are the primary components controlling how the P and SH

velocities vary for small offsets. The SV picks give no information about C_{66}, so both eigenvalues are controlled by about the same number of data points. The difference in magnitude between the two eigenvalues is probably accounted for by the difference in aperture; in Figure 3 we can see the range of angles sampled by the P waves is somewhat greater than that for the SH waves.

The final smallest eigenvalue is a factor of 800 down from the largest one; the associated eigenvector is mostly C_{11}, which corresponds to the horizontal P velocity. Strangely enough, this elastic constant makes no contribution to the variation of P velocity for small offsets, which is why this eigenvalue is so small.

If we tried to invert for the 5 elastic parameters in the lower layer using a noisy version of this dataset, the inevitably huge errors in the value of the least-constrained eigenvector component would destroy our results for C_{11}, C_{66}, and C_{13}, even though some linear combinations of these elastic constants are reasonably well determined. A standard solution is to apply damping of some sort to keep unresolved components from getting out of control; this should be equivalent to incorporating *a priori* probability information into the inversion.

CONCLUSIONS

Standard least-squares travel-time inversion can be put on a more solid probability theory footing, allowing determination of error bars on the inverted parameters. The choice of parameters to represent the final result is very important, lest one extremely poorly resolved eigenvector component of the inversion result appears to destroy the reliability of many output parameters.

REFERENCES

Auld, B. A., 1973, Acoustic fields and waves in solids: **1**, John Wiley and Sons.

Chapman, C. H., and Shearer, P. M., 1989, Ray tracing in azimuthally anisotropic media—II. Quasi-shear wave coupling: Geophys. J., **96**, 65–83.

Dellinger, J., and Muir, F., 1985, Two domains of anisotropy: SEP–**44**, 59–62.

Dellinger, J., and Muir, F., 1986, NMO and beating the central limit theorem: SEP–**48**, 261–268.

Douma, J., 1988, Crack-induced anisotropy and its effect on vertical seismic profiling: Ph.D. thesis, Rijksuniversiteit te Utrecht.

Jones, E. A., and Wang, H. F., 1981, Ultrasonic velocities in Cretaceous shales from the Williston basin: Geophysics, **46**, 288–297.