Comparing Objective Functions for Velocity Inversion

Mark S. Gockenbach
William W. Symes

December, 1993

TR93-46

Comparing Objective Functions for Velocity Inversion

Mark S. Gockenbach, William W. Symes*

Abstract

The success of automatic velocity inversion is highly dependent on the numerical tractibility of the optimization problem which defines the solution. The purpose of this paper is to compare the objective (cost) functions arising from two different formulations of the problem. It is shown by example that the output least–squares approach defines an objective function which can be highly nonconvex and which can have local, nonglobal minima. In contrast, the method of Differential Semblance Optimization defines a cost function which has a unique minimizer and which appears to be nearly convex. These two approaches are applied to the simple problem of determining the depth of a horizontal reflector and the velocity of the layer above it.

1 Introduction

Reflection seismologists commonly use seismic data to determine the structure of subsurface rock formations. Central to this task is the detection of reflectors — structures where rapid changes in rock properties, in particular the velocity at which waves are propagated, cause wave energy to be reflected. In order to correctly locate the reflectors, it is necessary to determine the velocity field of the region of the subsurface which is under study.

The process of determining a velocity field is typically an interactive exercise in which the (subjective) judgement of the seismologist comes into play. Consequently, as has been demonstrated by example (cf. [6]), different practitioners can produce markedly different velocity fields from the same seismic data. The goal of velocity inversion is to automate the determination of velocities by fitting the data to the predictions of a velocity model via an optimization routine.

Velocity inversion has been hampered by the fact that its most straightforward formulation as a best-fit problem — the output least–squares approach — can suffer from ill–conditioning and the existence of local, nonglobal minima. We give a simple example in which these properties of the output least–squares objective function can be clearly seen.

*Department of Computational and Applied Mathematics, Rice University, Houston TX 77251-1892

We then show that the *Differential Semblance Optimization (DSO)* objective function has much better properties. The DSO approach was introduced by Symes (see, for example, [3] or [4]).

We consider the simplest problem of reflection seismology, namely, the determination of the velocity of a layer and the depth of a horizontal reflector lying beneath it. Figure 1.1 shows the experiment which we have in mind: a source of seismic waves is located at the surface and a line of receivers record the vibrations produced by waves bouncing off the reflector. These recordings (the seismogram) comprise the seismic data used to determine both the velocity of the layer and the depth of the reflector.

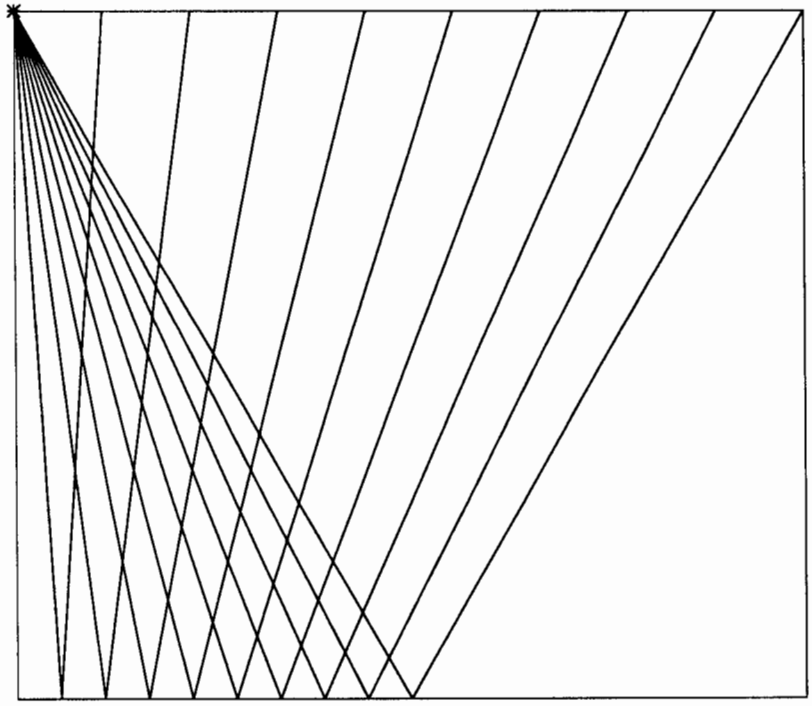

Figure 1.1: Seismic exeriment

Various aspects of this problem have been examined recently by Stork [2] and Lines [1]. The focus in these papers is the local conditioning of the problem, with the goal of describing how the solution is determined by the data. Our focus is more global in nature; we want to show how the formulation of the problem affects its tractability.

We examine two ways of using this data to solve the problem. The first technique is to identify the 'first-arrival times', that is, the time of arrival of the wavefront, at each receiver. These times are easily predicted from the velocity and depth, and so the velocity and depth can be found by choosing values which predict times most closely matching the

actual data. This problem is examined in Section 1. Given accurate first–arrival times, it is easy to determine the velocity and depth via the straightforward output least–squares approach. Nevertheless, it is in this context that we explain the idea behind the DSO approach, as even this simple problem has some undesirable characteristics which can be improved by a technique analogous to Differential Semblance Optimization.

The second way to use the seismogram to find the velocity and depth is to use the source waveform to predict the entire record at each receiver. In this way the problem of accurately identifying the first–arrival times can be avoided. However, it is necessary to know or find the source waveform. In section 2 we examine the problem of determining this waveform in addition to the velocity and depth. In this context the advantages of the DSO approach over that of output least–squares are seen clearly.

Throughout the paper (exact) synthetic data is used to compare the methods. Experience using the DSO approach with more realistic models can be found in [4].

2 Solution using First–Arrival Times

2.1 Least–squares approach

The following notation is used throughout the paper:

$$\begin{aligned} c &= \text{velocity of layer} \\ d &= \text{depth of reflector} \\ h &= \text{half-offset of receiver} \end{aligned}$$

We will assume that receivers are located at evenly–spaced offsets $0, 2h_2, 2h_3, \ldots, 2h_n$. The first-arrival time at the ith receiver is then given by

$$t_i = \frac{2\sqrt{d^2 + h_i^2}}{c},$$

and we can find c and d by solving:

$$\underset{c,d}{\text{minimize}} \quad J(c,d) \tag{2.1}$$

where

$$J(c,d) = \frac{1}{2}\sum_{i=1}^{n}\left(2\frac{\sqrt{d^2 + h_i^2}}{c} - t_i^{data}\right)^2$$

Figure 2.1 shows a contour plot of this objective function where the actual velocity and depth are $5m/ms$ and $500m$, respectively, and there are 51 receivers evenly spaced from $h = 0$ to $h = 500m$. The plot shows that the objective function has good global properties; there is a unique minimizer which could be easily found by any well–designed routine for

unconstrained minimization (there are actually two simple constraints, that c and d be nonnegative, but they are inactive at the solution and can be ignored). The function is mildly ill–conditioned; this ill–conditioning is not severe enough to cause difficulties in solving the problem, but as an illustration of the idea behind the Differential Semblance Optimization approach, we will discuss how this ill–conditioning can be alleviated.

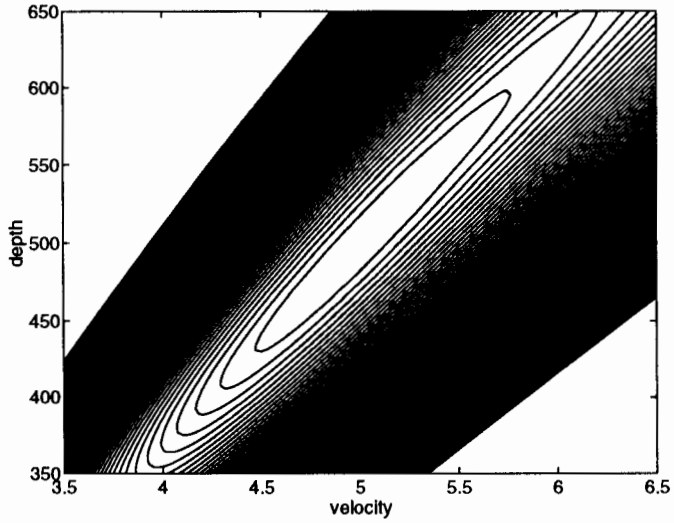

Figure 2.1: Contour plot of the least-squares objective function

2.2 Reformulating the Objective Function

The first step is to understand the shape of the level curves of the objective function J. The direction of the 'valley' on the contour plot shows that the value of the objective increases most slowly as (c, d) moves away from the solution in a direction which corresponds to increasing (or decreasing) both c and d. In fact, the eigenvalue–eigenvector pairs of the Hessian of J are

$$1.9005 \times 10^{-5} \quad , \quad \begin{bmatrix} 0.6058 \\ 0.7956 \end{bmatrix} ; 1.7120 \times 10^{-3} \quad , \quad \begin{bmatrix} -0.7956 \\ 0.6058 \end{bmatrix}$$

(Note: These values were computed after changing the units of time to hundreds of milliseconds so that the minimizer lies at (c,d) = (500,500). This change of units is necessary to avoid the nonessential ill–conditioning that arises from having variables of different orders of magnitude.) This behavior is easy to understand when we notice that the first–arrival time is increasing in d and decreasing in c; by changing c and d in the same direction, the two effects cancel each other out to some degree. On the other hand, if c is increased and d is decreased or vice versa, the two effects reinforce each other, accounting for the large eigenvalue in these directions and the steep wall of the valley on the contour plot.

The central issue is exposed even more clearly when the first-arrival time formula is written in terms of zero–offset time and Normal Moveout (NMO):

$$t_i = \frac{2d}{c} + \frac{2\left(\sqrt{d^2 + h_i^2} - d\right)}{c}$$

It is actually the zero–offset time which is preserved by changing c and d in the same direction, while a little analysis shows that the NMO is decreasing in both c and d. In other words, it is the problem of matching the zero–offset time which determines the conditioning of the problem in this formulation.

Now that the source of the ill–conditioning is exposed, we can deal with it as follows: since it is the difficulty of matching the zero-offset time which causes the ill–conditioning, we allow the zero-offset time to be fitted independently of c and d by introducing a new parameter T_0:

$$t_i = T_0 + \frac{2\left(\sqrt{d^2 + h_i^2} - d\right)}{c}$$

The correct variational problem would now be the following:

$$\begin{aligned} &\text{minimize} && \hat{J}(c, d, T_0) \\ &\text{subject to} && T_0 = \tfrac{2d}{c} \end{aligned} \tag{2.2}$$

where

$$\hat{J}(c, d, T_0) = \frac{1}{2} \sum_{i=1}^{n} \left(T_0 + \frac{2\left(\sqrt{d^2 + h_i^2} - d\right)}{c} - t_i^{data} \right)^2$$

This problem is completely equivalent to the original problem with one exception: formulation as a constrained minimization problem opens up the possibility of solution by an infeasible point method. In other words, the ability to have T_0 unequal to $2d/c$ while searching for the correct c and d may improve the conditioning of the problem. In fact, this is the case, as the examples below show.

The simplest infeasible point method is the penalty function method, in which a constrained optimization problem is turned into an unconstrained problem by incorporating the constraint into the objective function through the addition of a term which penalizes infeasibility. The above constrained problem could be expressed as

$$\begin{aligned} &\text{minimize} && J_r(c, d, T_0) \\ &c, d, T_0 \end{aligned} \tag{2.3}$$

where

$$J_r(c, d, T_0) = \hat{J}(c, d, T_0) + \frac{1}{2} r^2 \left(T_0 - \frac{2d}{c} \right)^2$$

Generally this unconstrained minimization problem would not have the same solution as the constrained problem, but rather a family of solutions parameterized by r would

be obtained, with the solution of (2.3) converging to the solution of (2.1) as $r \to \infty$. However, in this case, if we assume that the data is exact, the global solution of (2.3) must be the same as the global solution of (2.1). Therefore, (2.3) represents a different variational principle defining the same solution as the original formulation; moreover, r can be chosen to improve the conditioning of the problem.

Since the variable T_0 appears quadratically in the function J_r, we can solve for T_0 as a function of c and d, that is, we can reformulate (2.3) as

$$\min_{c,d} \min_{T_0} J_r(c,d,T_0) \tag{2.4}$$

We obtain

$$T_0(c,d) = \frac{1}{n+r^2}\left(\frac{2r^2 d}{c} + \sum_{i=1}^{n}\left(t_i^{data} - \frac{2\left(\sqrt{d^2+h_i^2}-d\right)}{c}\right)\right)$$

and the new variational problem

$$\min \tilde{J}_r(c,d) = J_r(c,d,T_0(c,d)) \tag{2.5}$$

We now examine the function $\tilde{J}_r$. Figure 2.2 shows its contour plots for various values of r. For even moderately large values of r, $\tilde{J}_r$ resembles the original objective function J, but as r is decreased, the conditioning of $\tilde{J}_r$ improves until, for $r \approx 0.75$ in this example, $\tilde{J}_r$ is almost perfectly conditioned near the solution. As r decreases further toward zero, the conditioning worsens until, at $r = 0$, the conditioning is worse than it was originally.

It is also interesting to note how the direction of the valley on the contour plot changes as r decreases. When r is large, that is, when (2.5) is essentially of the same character as (2.1), the direction of least change in the objective value is a direction in which c and d are changed at roughly the same rate. This, as explained above, is due to the fact that changing the zero–offset time affects the objective more than changing the NMO does. On the other hand, when $r = 0$, the direction of least change is a direction in which c and d are perturbed in opposite directions. This is because, when $r = 0$, there is complete freedom to match the zero–offset time independently of c and d; hence the direction of least change in the objective function is the direction in which the NMO is changed as slowly as possible. For r between 0 and ∞, the objective function $\tilde{J}_r$ interpolates between these two extremes. In particular, r can be chosen to give a perfectly well–conditioned problem.

As mentioned above, there is really no need to reformulate this problem since the original problem is easily solved. In the next section, we turn to a more difficult problem in which there are significant difficulties with the straightforward least–squares approach. In that case, the idea of expanding the model space to allow extra degrees of freedom in matching the data, and then penalizing that freedom, will pay substantial dividends.

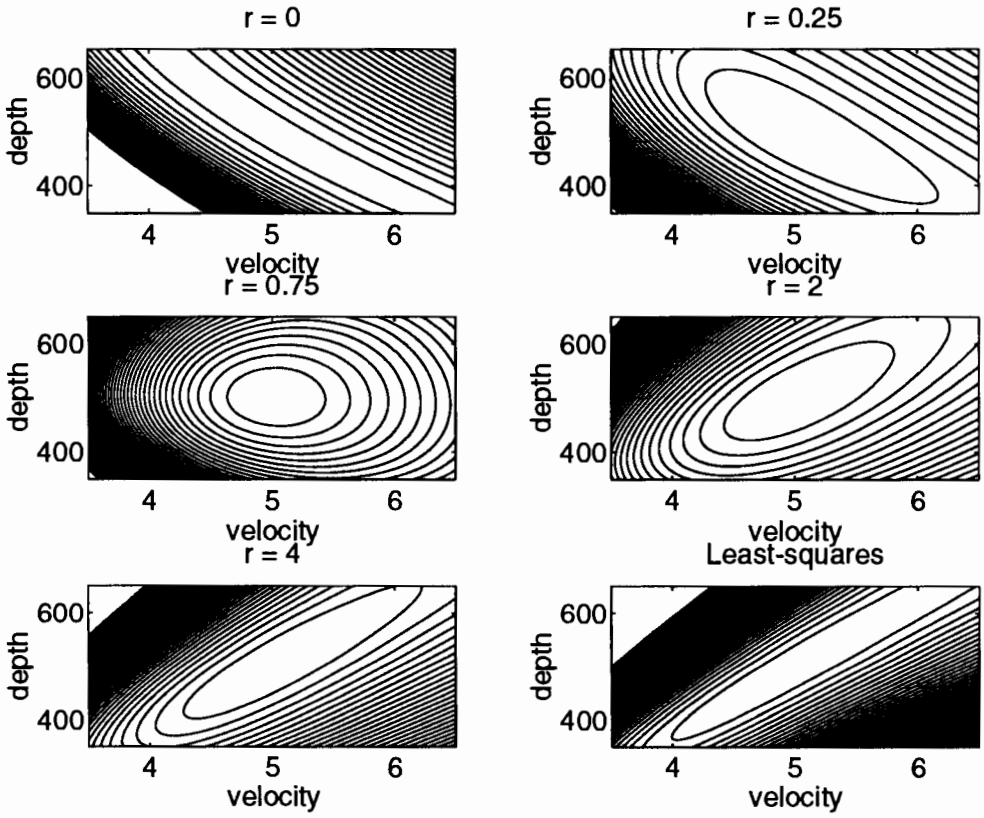

Figure 2.2: Comparison of reformulated objective functions

3 Solution using Full Seismogram

3.1 Output Least-squares: Difficulties and Examples

We now wish to determine the velocity c and the depth d by finding values for these parameters for which the predicted seismogram matches the actual seismogram. As the seismogram also depends on the source waveform $f(t)$, we will try to determine this from the data as well. We must construct the *forward map*, that is, the mapping which produces the seismogram $D(h,t)$ from c, d, and $f(t)$; in order to simplify the computations, we assume that the amplitudes of the waves do not decrease due to spherical spreading or reflection. The correct map would then be

$$S[(c,d),f](h,t) = f\left(t - \frac{2\sqrt{d^2+h^2}}{c}\right),$$

where $f \in L^2([0,T])$, and c and d are positive real numbers. The seismogram will belong to $L^2([0,H]\times[0,T])$, and the output least– squares variational principle is

$$\min \quad \|S[(c,d),f] - D\|^2_{L^2([0,H]\times[0,T])} \tag{3.1}$$

$$c, d, f$$

An obvious difficulty with this formulation is that the objective function is not in general differentiable with respect to c and d for an arbitrary L^2 function f. Moreover, suppose we require the source waveform f to lie in a more restrictive space, such as a Sobolev space $H^k([0,T])$ (where k is chosen large enough to induce the desired number of derivatives in the objective function). Then the problem immediately becomes ill–posed when the data D is not correspondingly smooth (that is, if $D(h,t) = \bar{f}\left(t - \frac{2\sqrt{d^2+h^2}}{c}\right)$ for some $\bar{f}$ which belongs to L^2 but not to H^k, then (3.1) clearly has no solution in H^k) .

In order to get around the objections raised in the previous paragraph, we can proceed as follows. We insist that the data $D(h,t)$ be smooth, say twice-continuously differentiable and write (3.1) as

$$\min_{c,d} \quad J(c,d) \tag{3.2}$$

where

$$J(c,d) = \min_{f} \quad \|S[(c,d),f] - D\|^2_{L^2([0,H]\times[0,T])} \tag{3.3}$$

It is not difficult to find the f that solves the inner minimization problem for a given c and d:

$$f_{cd}(t) = \frac{1}{H}\int_0^H D\left(h, t + \frac{2\sqrt{d^2+h^2}}{c}\right)\,\mathrm{dh}$$

Moreover, under the assumption that D is smooth, f_{cd} is also smooth, which gives us a smooth objective function J. Hence (3.2) can be solved using smooth optimization techniques.

Unfortunately, the fact that we have produced a smooth optimization problem does not guarantee that the problem can be solved easily. In fact, the objective function J tends to have certain unfavorable characteristics. As an example, the values of $J(c,d)$ were computed for exact data generated by a source waveform equal to the derivative of a Ricker wavelet (with peak frequency 50 Hz.), with velocity $c = 5m/ms$ and $d = 500m$. The receiver–source geometry was as in the previous examples. A contour plot of the objective function for this problem is shown in Figure 3.1. From this plot, we can see that the problem tends to be somewhat ill–conditioned and nonconvex. Furthermore, and more importantly, when we look outside of the valley which contains the solution, the objective function tends to be very flat. The one-dimensional scan of $J(c,d)$ in Figure 3.1 shows this clearly (as in all the scans shown in this paper, the depth is held fixed at the correct value and the velocity is varied).

This tendency of the output least–squares approach to produce objective functions which dip very sharply down to the solution and which are otherwise rather flat can produce local, nonglobal minimizers when the data is noisy or the source waveform is not compactly supported. For instance, we computed values of J for exact data generated by a

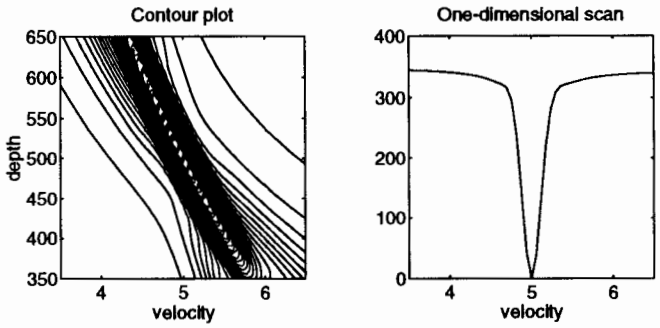

Figure 3.1: $J(c, d)$ for a Ricker source

periodic (sinusoidal) source with the same velocity and depth as in the previous example. A contour plot of the results is shown in Figure 3.2; this plot shows several stationary points besides the solution. A one–dimensional scan of J is also shown in Figure 3.2.

The existence of local, nonglobal minimizers is troublesome because the most efficient algorithms for finding minimizers are local techniques which are in no way guaranteed to find the global minimizer as opposed to local, non–global minizers. In fact, these algorithms tend to converge to the nearest (in some sense) minimizer, local or global. Much effort has been expended in recent years to develop algorithms capable of performing genuine global optimization, but no efficient method has been proven effective.

3.2 Reformulation using Differential Semblance Optimization

The foregoing discussion and examples have illustrated the short–comings of the output least–squares method. We now turn to a description of the method of Differential Semblance Optimization, which will be shown by example to alleviate some of the difficulties discussed above. The example in the previous section shows the outline of the method. First, we must discover exactly what it is in the nature of the problem that gives the objective function its undesirable characteristics. Next, we add (artificial) degrees of freedom to the model which allow a better fit of the data far from the solution. Finally, these degrees of freedom are penalized.

In our current problem, the various traces of the seismogram consist of the source waveform shifted by various amounts; these different shifts are determined by the hyperbolic Normal Moveout. (It should be noted, incidentally, that the problem of matching the NMO is responsible for the direction of the valleys on the contour plots of J. The freedom

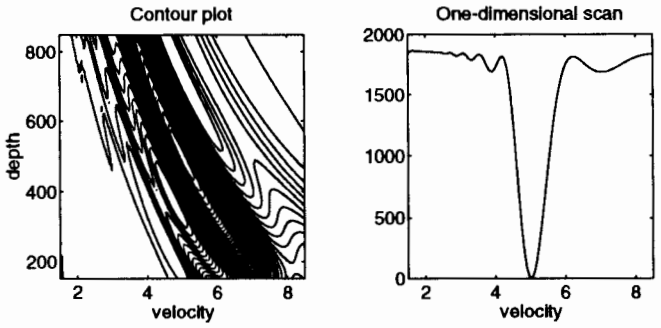

Figure 3.2: $J(c,d)$ for a periodic source

to choose the source waveform f includes implicitly the freedom to choose the zero-offset time because this source waveform can itself be shifted.) When the velocity and depth are wrong, so is the Normal Moveout, and the incorrect shifting of the source waveform means that the data is poorly fitted. With a localized source, as in the first example, very much error at all in the shifts means that the traces of the predicted seismograms will be almost completely orthogonal to the traces of the actual seismogram, and therefore the data misfit, that is, the least–squares objective function, is uniformly large outside of the narrow valley around the solution. With the periodic source, as the shift increases, the sine wave moves in and out of phase with the data; over many traces, when the NMO is wrong, the total error must be large (since many of the traces must be out of phase with the data), but local minima appear, apparently due to the fact that some of the traces of the predicted seismogram move in and out of phase with the data as the shift changes.

In considering how to improve the characteristics of the objective function, we must keep in mind that we would like the objective function to increase in a controlled manner as (c,d) moves away from the solution. We would like to avoid the situation in which we find ourselves, namely, that the function increases suddenly as (c,d) is moved a small distance from the solution, and then does not change much. The key is to 'cheat' by making it easier to match the data when c and d are wrong. We accomplish this by allowing the unknown source waveform f to depend on the receiver location; that is, for each trace we choose a different source to explain the data. Mathematically, this means that f will be a function of the half–offset h as well as of t: $f = f(h,t)$. With this additional freedom, it is easy to match the data; in fact, the data can be perfectly fitted regardless of the values of c and d by taking

$$f(h,t) = D\left(h, t + \frac{2\sqrt{d^2+h^2}}{c}\right)$$

(where now

$$\hat{S}[(c,d),f](h,t) = f\left(h, t - \frac{2\sqrt{d^2+h^2}}{c}\right)).$$

Next, we penalize the dependence of f on h by adding to the objective function the following term:

$$J_r(c,d,f) = \frac{1}{2}\|\hat{S}[(c,d),f] - D\|^2_{L^2([0,H]\times[0,T])} + \frac{1}{2}r^2\|\frac{\partial f}{\partial h}\|^2_{L^2([0,H]\times[0,T])}$$

(Here is the reason for the name Differential Semblance Optimization — the differential penalty term requires that the sources which produce the various traces on the seismogram resemble each other).

The final step in the DSO formulation for this problem is a technical one. As before, we wish to take advantage of the fact that f has a quadratic influence on the objective function to first minimize over f and thus to reduce the objective to a function of the two variables c and d. In order that the normal equations for minimizing over f be an elliptic boundary value problem, we change the forward map once again:

$$\tilde{S}[(c,d),f](h,t) = \frac{\partial f}{\partial t}\left(h, t - \frac{2\sqrt{d^2+h^2}}{c}\right)$$

The (reduced) DSO objective function is then

$$\tilde{J}_r(c,d) = \min_f \frac{1}{2}\|\tilde{S}[(c,d),f] - D\|^2_{L^2([0,H]\times[0,T])} + \frac{1}{2}r^2\|\frac{\partial f}{\partial h}\|^2_{L^2([0,H]\times[0,T])} \tag{3.4}$$

Note that in introducing the partial derivative with respect to t into the forward map, we have effectively changed variables: f no longer represents the seismic source but rather its anti-derivative. The natural space in which to find f would now be some subspace of $H^1([0,H]\times[0,T])$. With an appropriate choice of this subspace, the normal equations for minimizing over f can then be shown to be (formally) equivalent to the following elliptic boundary value problem:

$$\frac{\partial^2 f}{\partial t^2} + r^2\frac{\partial^2 f}{\partial h^2} = \frac{\partial D}{\partial t}\left(h, t + \frac{2\sqrt{d^2+h^2}}{c}\right) \tag{3.5}$$

$$f(h,0) \equiv f(h,T)$$

$$\frac{\partial f}{\partial t}(h,0) \equiv \frac{\partial f}{\partial t}(h,T)$$

$$\frac{\partial f}{\partial h}(0,t) \equiv \frac{\partial f}{\partial h}(H,t) \equiv 0$$

$$\int_0^H \int_0^T f(h,t)\ \mathrm{dt}\ \mathrm{dh} = 0$$

(One important aspect of this formulation is that it apparently induces smoothness in the reduced objective function. From results in analogous settings (cf. [5]), we can expect

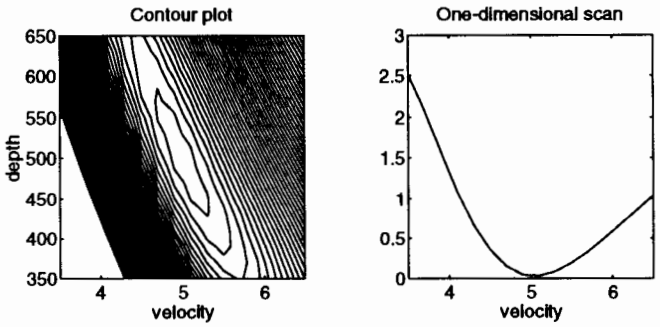

Figure 3.3: $\tilde{J}_1(c,d)$ for Ricker source

that $\tilde{J}_r$ will be smooth even if the data D is a general L^2 function. This has not yet been proven for the current problem, and we will ignore the question by always considering smooth data.)

The DSO formulation has the fundamental property that, given exact data, the desired solution is the global minimizer of the DSO objective function. This is easy to see, since at the solution, the data misfit term is zero (because the solution generates the data) and the penalty term is zero (because the source waveform does not depend on h). The numerical experiments to be described below suggest that the DSO approach works very well, alleviating the difficulties encountered in the least–squares approach; the only difficulty lies in knowing how to choose the penalty weight r.

For our first example, we use the seismic data generated by using the derivative of a Ricker wavelet as the source. This data produced the output least–squares objective function shown in Figure 3.1. Using the DSO formulation with $r = 1$, we computed $\tilde{J}_r$, with the results shown in Figure 3.3. It should be noted that the ill–conditioning is lessened (as is demonstrated by the fact that the valley on the contour plot is no longer as long and narrow), and that the DSO objective function does not have the property of being large and relatively flat away from the solution (as can be seen from the one-dimensional scan of the function).

The effect of the penalty parameter is to interpolate between two extreme problems. If $r = \infty$, the problem is the output least-squares formulation, while when $r = 0$, the function $\tilde{J}_r$ is identically zero (with no penalty on h-dependence, the data can be fitted perfectly for any c and d). We can expect, therefore, that the DSO formulation will yield a 'good' problem for some range of r, bounded away from zero and also bounded above. In Figure 3.4, scans of $\tilde{J}_{500}$ and $\tilde{J}_{2500}$ are shown; these graphs show how the character of

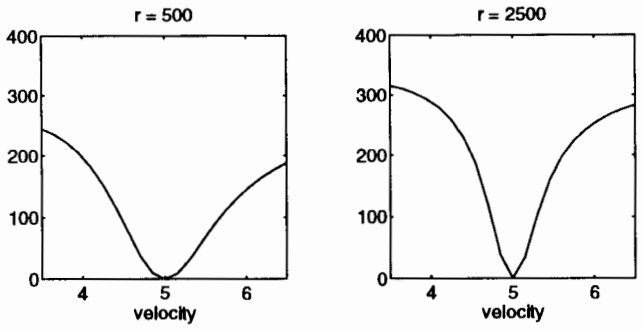

Figure 3.4: $\tilde{J}_r(c,d)$ for Ricker source

the output least–squares problem is recovered as r increases.

As another example, the DSO method was applied to the seismic data generated by the sinusoidal source; this data produced the output least–squares objective function illustrated in Figure 3.2. The DSO objective function $\tilde{J}_r$ was computed for various values of r, and one–dimensional scans of the results are shown in Figure 3.5. Together with the objective function are shown its components: the data misfit and DSO penalty terms. In these graphs, the effect of the penalty term can be clearly seen. Its influence is to convexify the objective function by smoothing out the bumps in the data misfit term. The penalty term appears to be convex for intermediate values of r; if r is chosen large enough, then the penalty term is weighted heavily enough to overcome the extreme non-convexity of the data misfit term. However, r can not be chosen too large, since the plot of $\tilde{J}_{128}$ shows that the penalty term can lose its convexity as r gets large enough. It should be appreciated that this convexifying influence of the penalty term is so effective because the penalty term has a unique global minimizer that occurs precisely at the solution of the problem (for exact data). Such a procedure — adding a convex function whose minimizer is in the right place — would improve any nonconvex problem; the difficulty, of course, is that in general we do not know how to construct such a convex function. In the present setting, the penalty term has its minimizer in the right place for the following reason: the term penalizes the use of an h–dependent source to fit the data; there is no penalty at the solution because the solution is the only point where taking advantage of h–dependent sources does not help in matching the data (the data can be matched perfectly at the solution already).

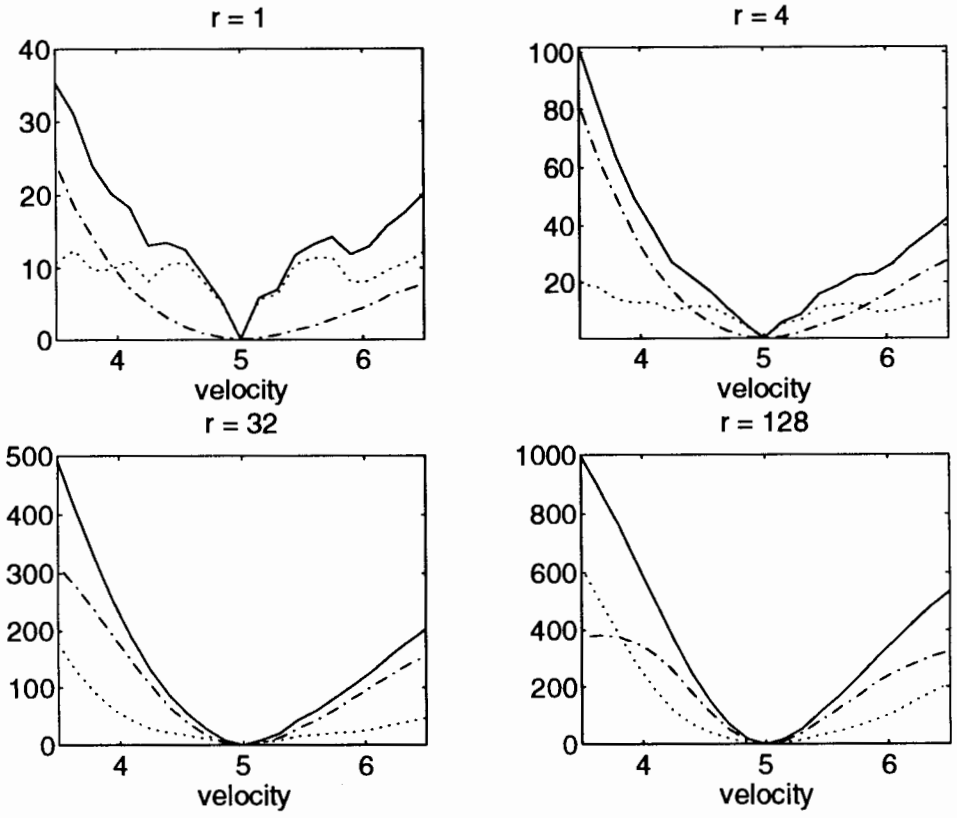

Figure 3.5: $\tilde{J}_r(c,d)$ for periodic source (solid line); data misfit term (dotted line); penalty term (broken line)

4 Discussion

The foregoing examples show that the technique embodied in the Differential Semblance Optimization approach to velocity inversion can improve the the optimization problem which must be solved. It is possible to reduce the ill–conditioning of the problem as well as eliminate local, nonglobal minimizers. We intend to follow these numerical experiments with related research in two directions. First, the analytical properties of DSO as applied to the problem discussed in this paper must be examined. Questions about convexity and smoothness of the DSO objective function need answers; analysis will hopefully lead to further insight into the role of the penalty parameter and how to choose it. Second, on a larger front, the idea behind DSO — expanding the space in which to search for the solution and penalizing the extra degrees of freedom — needs to be clarified. It is to be expected that such an approach would be fruitful in approaching other optimization problems; hopefully further research will lead to the uncovering of a general principle which will be useful in other settings.

Acknowledgements

The author wishes to thank Professor W. W. Symes for suggesting this problem, and both Professor Symes and Dr. R. Versteeg for helpful discussions regarding the implementation of DSO.

This paper is a contribution to The Rice Inversion Project. TRIP sponsors for 1993 are Advance Geophysical, Amoco Research, Conoco Inc., Mobil Research and Development Corp., Exxon Production Research Co., Earth Modeling Systems, and Cray Research Inc.

References

[1] LINES, L. [1993]. Ambiguity in Analysis of Velocity and Depth, *Geophysics 58*, 596-597.

[2] STORK, C. [1992]. Singular Value Decomposition of the Velocity–Reflector Depth Tradeoff, Part 1: Introduction Using a Two–parameter Model, *Geophysics 57*, 927-932.

[3] SYMES, W. W. [1990]. Velocity Inversion: a Case Study in Infinite-dimensional Optimization, *Mathematical Programming 48*, pp. 71–102.

[4] SYMES, W. W. and CARAZZONE, J. [1991]. Velocity Inversion by Differential Semblance Optimization, *Geophysics 56*, 654–663.

[5] SYMES, W. W. [1992]. The Plane-Wave Detection Problem, Technical Report 92-22, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77251

[6] VERSTEEG, R., and GRAU, G. eds. [1991]. The Marmousi Experience: Proceeding of the EAEG Workshop on Practical Aspects of Inversion, IFP/Technip