The Inverse Problem of Seismic Velocities

William W. Symes

Computational and Applied Mathematics
Rice University

Agenda

Abstract Ruminations about Inverse Problems

The Seismic Reflection Experiment and the Acoustic Model

Solution by Optimization

Example

Agenda

Abstract Ruminations about Inverse Problems

The Seismic Reflection Experiment and the Acoustic Model

Solution by Optimization

Example

Practical Abstract Inverse Problems

The usual set-up:

- $\mathcal{M}$ = a set of *models*
- $\mathcal{D}$ = a Hilbert space of (potential) data
- $\mathcal{F} : \mathcal{M} \rightarrow \mathcal{D}$

[These things are collectively "the model".]

Inverse problem: given $d \in \mathcal{D}$, find $m \in \mathcal{M}$ so that $\mathcal{F}[m] \simeq d$.

Practical Abstract Inverse Problems

Chief requirement of a "solution": must be able to

1. Characterize - what problem does the "solution" solve? Does it exist? What degree of nonuniqueness? [*Interpretation of inaccurate, insufficient, and inconsistent data*, D. D. Jackson, Geophys. J. Royal Astr. Soc. **28** (1972), pp. 97-110.]
2. Find - does the characterizing problem admit an effective numerical solution?

Common pattern for 2.: solution is extremum of variational principle, for instance

$$m = \mathrm{argmin} \| \mathcal{F}[m] - d \|$$

RICE

Practical Abstract Inverse Problems

Example, topic of this talk: reflection seismology. Naturally formulated as inverse problem using various physical descriptions of seismic wave motion (acoustic, elastic, viscoelastic,...)

Typical problem size for adequately sampled 3D reflection seismic survey: $\mathrm{dim}(\mathcal{M}) \sim 10^9, \mathrm{dim}(\mathcal{D}) \sim 10^{12}$

$\Rightarrow$ any computational "solution" must admit algorithms that scale well with problem size - if iterative, then iteration count should be ess. independent of dimension.

Optimization $\Rightarrow$ Newton's method $\Rightarrow$ must be satisfied with *any* stationary point.

Practical Abstract Inverse Problems

Takeaway messages of this talk:

Straightforward data fitting (eg. by least squares) does not work well for this class of problems

"Relaxed" variational formulation via model extensions leads to effective numerical algorithms

For simplest cases, can show that all stationary points are approximate global minimizers.

Numerical evidence for more than this, but many open questions

Agenda

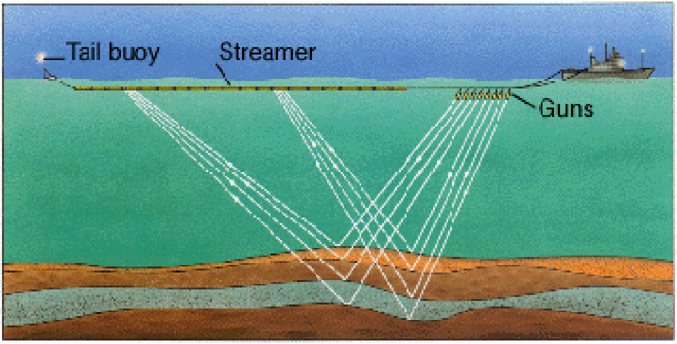
Abstract Ruminations about Inverse Problems

The Seismic Reflection Experiment and the Acoustic Model

Solution by Optimization

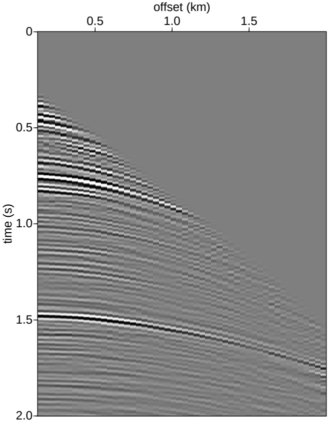
Example

Marine Seismic Reflection Experiment



Airguns = source of sound. Streamer consists of hydrophone receiver groups. Each group records a trace (time series of pressure) for each shot = excitation of source. Source-receiver distance = offset.

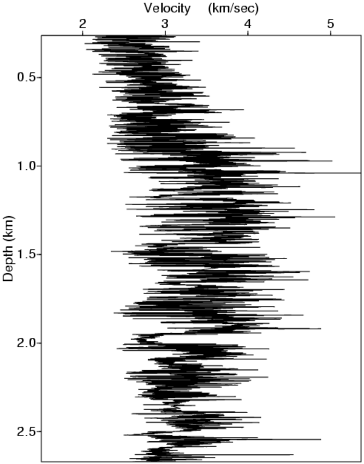
Typical Shot Record



North Sea Survey (thanks: Shell).
Processing applied:

- bandpass filter 3-8-25-35 Hz (data was oscillatory to begin with!);
- cutoff or mute to remove non-reflection energy (direct, diving, head waves);
- predictive deconvolution to suppress multiple reflections.

Mechanical properties of sedimentary rocks



Well (v_p) log from Texas borehole
(thanks: P. Janak, Total E&P, USA)

- v_p varies significantly.
- Heterogeneity at all scales - km to mm to μm.

Point Source Acoustics - the minimal model

Earth "=" $\Omega \subset \mathbf{R}^3$, wave velocity $v : \Omega \rightarrow \mathbf{R}, v > 0$.

Wave equation for acoustic potential response to isotropic point radiator at $\mathbf{x}_s$, time dependence $w(t)$:

$$\left(\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} - \nabla^2\right) u(t, \mathbf{x}; \mathbf{x}_s) = w(t)\delta(\mathbf{x} - \mathbf{x}_s)$$

plus appropriate initial and boundary conditions. NB: to model oscillatory nature of data, w must be oscillatory - $\hat{w}(\omega) = O(|\omega|^p)$, $p \geq 1$.

Lions, late '60's: proper notion of weak solution, well posed for $v \in \mathcal{M} = \{\log v \in L^\infty(\Omega)\}$, RHS in $L^2([0, T] \times \Omega)$

Point Source Acoustics - the minimal model

Forward map: $\mathcal{F} : \mathcal{M} \rightarrow \mathcal{D} = L^2([0, T] \times \Sigma)$,
$\Sigma \subset \{x_3 = 0\} \times \{x_3 = 0\}$ open, samples pressure in support of
$\phi \in C_0^\infty(\Sigma)$: for $(t, \mathbf{x}_r, \mathbf{x}_s) \in [0, T] \times \Sigma$,

$$\mathcal{F}[v](t, \mathbf{x}_r; \mathbf{x}_s) = \left(\phi \frac{\partial u}{\partial t}\right)(t, \mathbf{x}_r; \mathbf{x}_s)$$

If $v = v_0$ known & constant in $\{x_3 < z\}$ for some $z > 0$,
$w \in L^2(\mathbf{R})$, slight extension of Lions' argument shows $\mathcal{F}$
well-defined.

Stolk 2000: continuous, differentiable "with loss of derivative".

Agenda

Abstract Ruminations about Inverse Problems

The Seismic Reflection Experiment and the Acoustic Model

Solution by Optimization

Example

Least Squares Data Fitting

Natural formulation, in view of defn of $\mathcal{D}$: choose v by

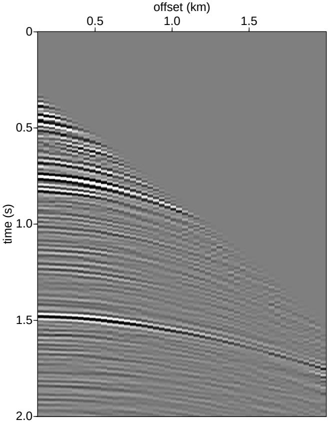
$$v = \mathrm{argmin}\,(\|\mathcal{F}[v] - d\|^2_{L^2([0,T]\times\Sigma)} + \mathcal{R}[v])$$

("mean square error") in which $\mathcal{R}$ (*regularization* functional) supplies additional stability.

Promoted heavily by Tarantola and others in the 1980's on grounds of Bayesian justification (maximum likelihood solution given Gaussian data error statistics)

Recently revived as major industry interest ("Full Waveform Inversion", FWI) - all-day workshop with attendance > 300 at SEG 09.

Least Squares Data Fitting



A Sad Story:

Data is *oscillatory* - $O(100)$ wavelengths

Small changes in velocity $\Rightarrow$ small changes in data isosurfaces $\Rightarrow$ large changes in mean square error $\Rightarrow$ saturation $\Rightarrow$ many stationary points

Least Squares Data Fitting

Upshot:

- FWI via iterative optimization method recovers very detailed subsurface models, at least in numerical tests with model data, when starting model is sufficiently accurate (Tarantola and coworkers 80's, 90's; Bunks 95; much recent work)
- Fails when starting model is not sufficiently accurate (stalls at stationary point with poor data fit)
- Hard to tell what "sufficiently accurate" means - no *a priori* test
- Continuation from low to high frequency / depth permits convergence with less accurate starting model (Kolb et al 1986, Bunks 95, Pratt 2004, recent from Shin and coauthors) - however no guarantees

Solution via Model Extension

Extension of $\mathcal{F}$:

- $\chi : \mathcal{M} \rightarrow \bar{\mathcal{M}}$,
- $\bar{\mathcal{F}} : \bar{\mathcal{M}} \rightarrow \bar{\mathcal{D}}$
- $\phi : \bar{\mathcal{D}} \rightarrow \mathcal{D}$

so that

$$\begin{array}{cccccc} & & & \bar{\mathcal{F}} & & \\ & \bar{\mathcal{M}} & & \rightarrow & \bar{\mathcal{D}} & \\ \chi & \uparrow & & & \downarrow & \phi \\ & \mathcal{M} & & \rightarrow & \mathcal{D} & \\ & & & \mathcal{F} & & \end{array}$$

commutes - that is,

$$\phi[\bar{\mathcal{F}}[\chi[v]]] = \mathcal{F}[v],\ v \in \mathcal{M}$$

Solution via Model Extension

Example:

$\bar{\mathcal{M}} \subset$ self-adjoint positive definite bounded operators on $L^2(\Omega)$ [Remark: action-at-a-distance], $\bar{\mathcal{D}} \subset \mathcal{D}'(\mathbf{R} \times \Sigma)$. For $\bar{v} \in \bar{\mathcal{M}}$,

$$\bar{\mathcal{F}}[\bar{v}](t, \mathbf{x}_r; \mathbf{x}_s) = \left(\phi \frac{\partial u}{\partial t}\right)(t, \mathbf{x}_r; \mathbf{x}_s)$$

in which u is causal solution of

$$\left(\bar{v}^{-2}\frac{\partial^2 u}{\partial t^2} - \nabla^2\right) u(t, \mathbf{x}; \mathbf{x}_s) = \delta(t)\delta(\mathbf{x} - \mathbf{x}_s)$$

Minor modification of Lions' construction $\Rightarrow$ well-posed when $\bar{v}$ acts as multiple of identity on functions supported near $\mathbf{x}_s$.

$\chi : \mathcal{M} \subset L^\infty(\Omega) \rightarrow \bar{\mathcal{M}}$ multiplier: $\chi[v]u = vu$. $\phi : \bar{\mathcal{D}} \rightarrow \mathcal{D}$ by $\phi[d] = w *_t d$ - filters out low freqs

Range of $\chi =$ "physical" models

Solution via Model Extension

Invertible extension: $\bar{\mathcal{F}}$ has approximate left inverse $\bar{\mathcal{G}}$ (on $\mathcal{R}(\bar{\mathcal{F}})$)

NB: trivial extension - $\bar{\mathcal{M}} = \mathcal{M}, \bar{\mathcal{F}} = \mathcal{F}, \chi = \phi = id$ - virtually never invertible.

Example: considerable numerical evidence (but little theory, except for space dimn $= 1$) strongly suggests that example extension is invertible.

Solution via Model Extension

Reformulation of inverse problem: seek $\bar{v} \in \bar{\mathcal{M}}, \bar{d} \in \phi^{-1}[d]$ so that

$$\bar{\mathcal{F}}[\bar{v}] \simeq \bar{d}, \; \bar{v} \in \mathcal{R}(\chi)$$

Then $\bar{v} = \chi[v]$ and v is an (approx.) solution of original inverse problem

Only advantageous if $\bar{\mathcal{F}}$ is invertible, with approximate inverse $\bar{\mathcal{G}}$ - then problem becomes:

find $\bar{d} \in \phi^{-1}[d]$ so that $\bar{\mathcal{G}}[\bar{d}] \in \mathcal{R}(\chi)$

Solution via Model Extension

Practical importance - back to Main Example:

Range of χ consists of multiplication ops by L^∞ functions,

$\Rightarrow$ distribution kernels $v(\mathbf{x},\mathbf{y})$ supported on (near) $\subset$ diagonal $\mathbf{x}=\mathbf{y}$

$\Rightarrow$ VISUALLY OBVIOUS in plot of $v(\mathbf{x},\mathbf{y})$!!!!

$\Rightarrow$ industry standard algorithms: tweak (mostly by hand) parameters of $v(\mathbf{x},\mathbf{y})$ until support focuses on diagonal

RICE

Automation

Suppose $W : \bar{\mathcal{M}} \rightarrow \bar{\mathcal{M}}$ annihilates range of χ:

$$\mathcal{M} \xrightarrow{\chi} \bar{\mathcal{M}} \xrightarrow{W} \bar{\mathcal{M}} \rightarrow 0$$

Define

$$A \equiv W \circ \bar{\mathcal{G}} : \bar{\mathcal{D}} \rightarrow \bar{\mathcal{M}}$$

Then for $\bar{d} \in \phi^{-1}(d)$,

$$A[\bar{d}] = 0 \Rightarrow \bar{\mathcal{G}}[\bar{d}] = \chi v \Rightarrow d \simeq \phi[\bar{\mathcal{F}}[\bar{\mathcal{G}}[\bar{d}]]] = \bar{\mathcal{F}}[\chi[v]] = \mathcal{F}[v]$$

Thus inverse problem equivalent to:

find $\bar{d} \in \phi^{-1}[d]$ so that $A[\bar{d}] = 0$

Automation

Back to the main example: range of χ consists of multiplication ops by L^∞ functions, which commute with other multiplication ops - so can choose

$$W[\bar{v}] = [\bar{v}, \mathbf{x}]$$

in which $\mathbf{x}$ represents multiplication by coordinate vector.

Write $\bar{v}$ formally as integral operator with kernel $\bar{v}(\mathbf{x}, \mathbf{y})$. Then

$$W[\bar{v}]u(\mathbf{x}) = \int_\Omega d\mathbf{y}\,\bar{v}(\mathbf{x}, \mathbf{y})(\mathbf{x} - \mathbf{y})u(\mathbf{y})$$

multiplication of $\bar{v}$ by *offset* $\mathbf{x} - \mathbf{y}$

Automation

Why should you care?

For very simple model problem using drastic approximations to $\mathcal{F}$ and so on:

- least squares data fitting has stationary points unrelated to solution of inverse problem, even for noiseless data
- for proper choice of W, hence A, parametrization of $\phi^{-1}[d]$, and Hilbert norm $\|\cdot\|$ in $\bar{\mathcal{M}}$, all stationary points of

  $$\bar{d} \mapsto \|A[\bar{d}]\|^2$$

  are approximate global minimizers (WWS).
- numerical experiments with synthetic, field data suggest that same is true in some generality

Linearized Extension and Migration Velocity Analysis

Conventional simplification: replace $\bar{\mathcal{F}}$ with *linearization at smooth physical models*, in terms of $\bar{r} = \delta\bar{v}^{-2}$:

$\bar{\mathcal{M}}_1 = \{(v, \bar{r}) : \log v \in C^\infty(\Omega), \bar{r} \in \mathcal{B}_{\rm symm}(L^2(\Omega))\}$,
$\bar{\mathcal{D}}_1 = \{v : \log v \in C^\infty(\Omega)\} \times \mathcal{D}$

define linearization $D\bar{\mathcal{F}} : \bar{\mathcal{M}}_1 \rightarrow \bar{\mathcal{D}}_1$ by

$$D\bar{\mathcal{F}}[v, \bar{r}] = \left(v, \phi\frac{\partial \delta u}{\partial t}\right)$$

in which

$$\left(\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} - \nabla^2\right)\delta u(t, \mathbf{x}; \mathbf{x}_s) = -\bar{r}\left[\frac{\partial^2 u}{\partial t^2}(t, \cdot; \mathbf{x}_s)\right]$$

Linearized Extension and Migration Velocity Analysis

Notes:

- geometric optics analysis $\Rightarrow$ inclusion of low frequencies in extended data $\Leftrightarrow$ inclusion of smooth "background" velocity model in extended data
- $\phi[v, d] = d$, so search in $\phi^{-1}[d]$ is search over smooth background velocities
- Adjoint of $\bar{r} \mapsto D\bar{\mathcal{F}}[v, \bar{r}]$ is *shot-geophone migration operator*
- with H-S norm, $\bar{r} \mapsto D\bar{\mathcal{F}}[v, \bar{r}]$ is "nearly unitary", so adjoint is closely related to inverse, often used instead

Linearized Extension and Migration Velocity Analysis

Recall annihilator of physical model perturbations: $W[\bar{r}] = [\bar{r}, \mathbf{x}]$ - in terms of kernel.

"prestack imaging operator": approximate inverse $\mathcal{I}[v]$ of $\bar{r} \mapsto D\bar{\mathcal{F}}[v, \bar{r}]$

Idealized extended inversion algorithm boils down to: minimize (over v) operator norm of $W[\mathcal{I}[v]d]$.

All implementations so far: take advantage of smoothness of numerical approximations to replace operator norm with H-S norm:

$$J_{DS}[v, d] = \frac{1}{2} \int d\mathbf{x} \int d\mathbf{y} |\mathcal{I}[v]d(\mathbf{x}, \mathbf{y})(\mathbf{x} - \mathbf{y})|^2$$

Estimate v by minimizing J_{DS}: "differential semblance", "annihilator-based waveform tomography",... (Stolk-de Hoop 01, Shen et al 03, 05, Kabir et al. 06, Shen & WWS 08,...)

Agenda

Abstract Ruminations about Inverse Problems

The Seismic Reflection Experiment and the Acoustic Model

Solution by Optimization

Example

Field Data Example

Shen & WWS *Geophysics* 08 (similar expl: Kabir et al. 07 SEG)

Based on penalty version of differential semblance: compute $\mathcal{I}[v]$ by fitting extended data in least-squares sense, with $J_{DS}[v, d]$ as penalty.

Various approximations -

- use adjoint in place of inverse of $D\bar{\mathcal{F}}$
- approximate adjoint by solving wave equation as evolution in depth (see Stolk-de Hoop *Wave Motion* 05, 06)
- compute gradient of J_{DS} by adjoint state method (Shen's thesis)
- iterative quasi-Newton optimization algorithm - limited memory BFGS with adjoint state gradients

Field Data Example

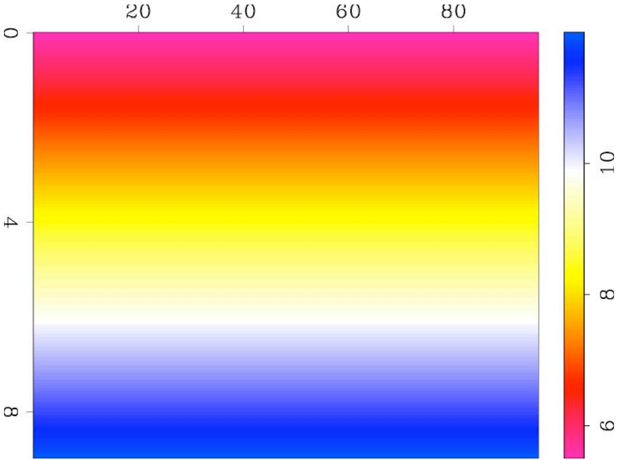
Gas chimney example - thanks Shell

Marine 2D line - preliminary imaging with regional velocity model shows gas-induced distortion ("sag").

Reflection tomography (traveltime inversion) partially removes sag effect, but interpreters not happy.

Differential Semblance to rescue - 20 iterations of Newton-like optimization algorithm produces more "geological" velocity (v), image (diagonal of $\mathcal{I}[v]d$) - interpreters happier.

Field Data Example



Initial velocity model - regional trends with depth

Field Data Example

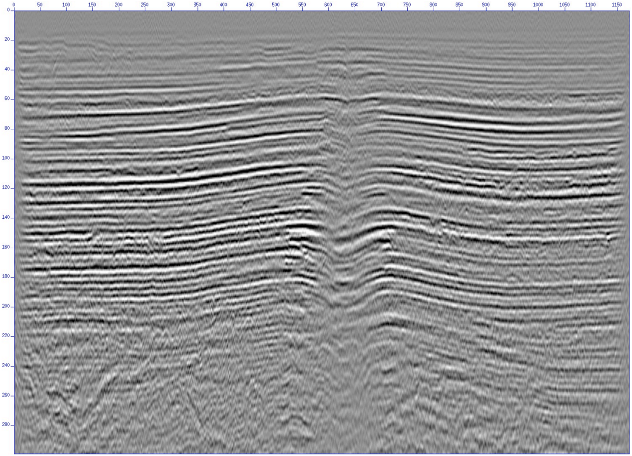
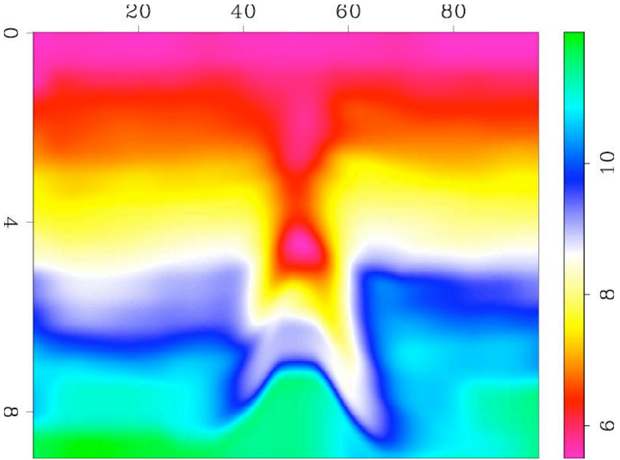


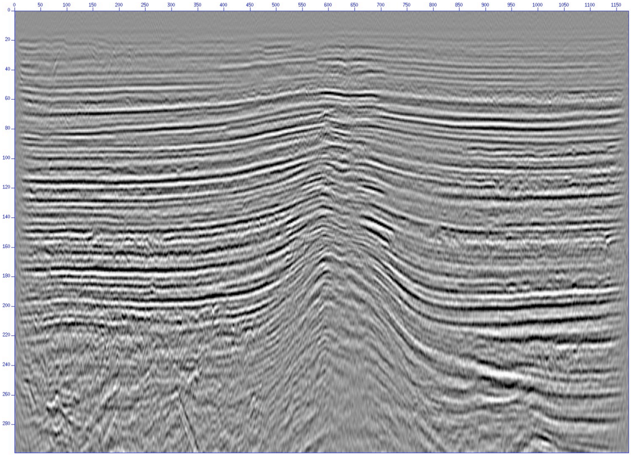
Image at initial model

Field Data Example



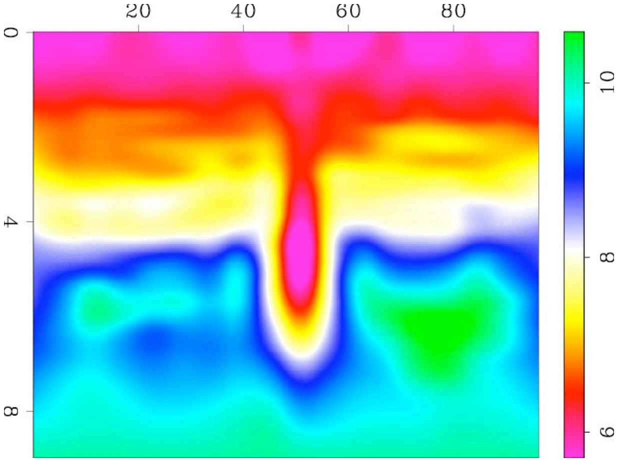
Model produced by Reflection Tomography

Field Data Example



Reflection Tomography image

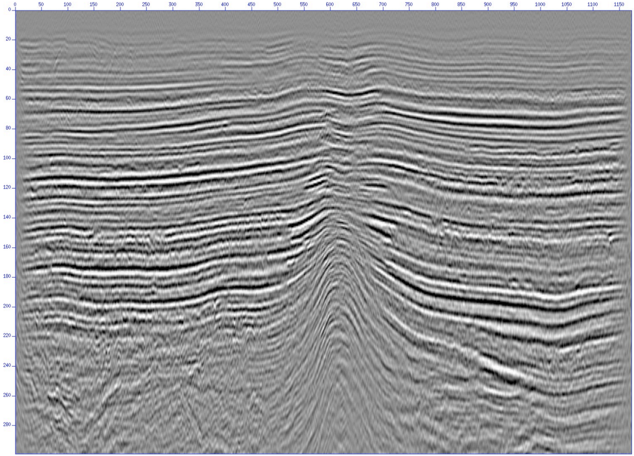
Field Data Example



Model produced by diff'l semblance (20 LBFGS iterations)

RICE

Field Data Example



Diffl semblance image (diagonal of $\mathcal{I}[v]d$)

Field Data Example

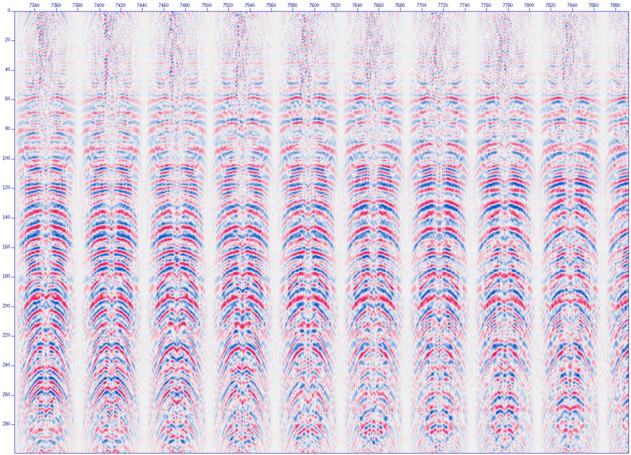
Angle domain common image gathers ("ADCIGs" - Sava & Fomel 03) - Radon transform of 2D slice ("depth-offset") of $\mathcal{I}[v]d$, should be flat at correct velocity - internal measure of consistency between v, d

Initial velocity - dramatic failure to flatten.

Reflection tomography - much better, but still not flat at larger depths.

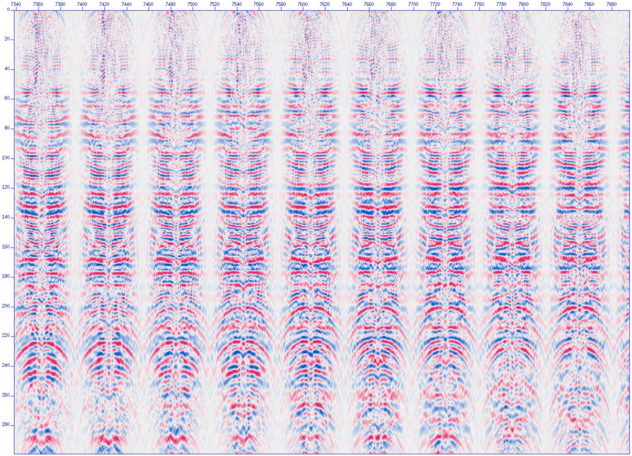
Differential semblance - better yet

Field Data Example



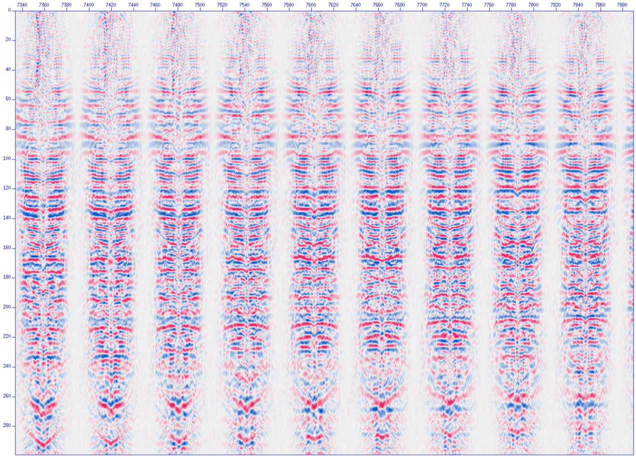
ADCIGs, initial model

Field Data Example



ADCIGs, reflection tomography

Field Data Example



ADCIGs, differential semblance

Summary

- Because of oscillatory/wave character of data, seismic inverse problem for wave velocity poorly suited for data fitting formulation - data misfit norm has many stationary points far from solution
- Model extensions permit reformulation in larger model domain ("relaxation", infeasible point method)
- Example extension: velocity coefficient in wave equation as possibly nondiagonal SPD operator
- Simplest examples of extended variational principle ("differential semblance") yield quasi-convex optimization problem - all stationary points are global mins
- With sufficiently many layers of approximation, applicable at field scale
- Almost every mathematical question is open

Acknowledgments

Thanks to:

- National Science Foundation
- The Sponsors of The Rice Inversion Project (TRIP).
- Schlumberger, Shell, Total E&P USA.
- Many colleagues: M. de Hoop, C. Stolk, B. Biondi,....
- All of you for listening...

RICE