Model Extensions and Inverse Scattering: Inversion for Seismic Velocities

William W. Symes⋆

⋆Rice University

October 2007

Expanded version of presentation in “Recent Advances and Road Ahead” session, SEG 2007.

References to related SEG 07 talks in red - available via Scitation, www.seg.org.

PDF available at www.trip.caam.rice.edu.

Draft paper with references: www.caam.rice.edu, TR 07-05.

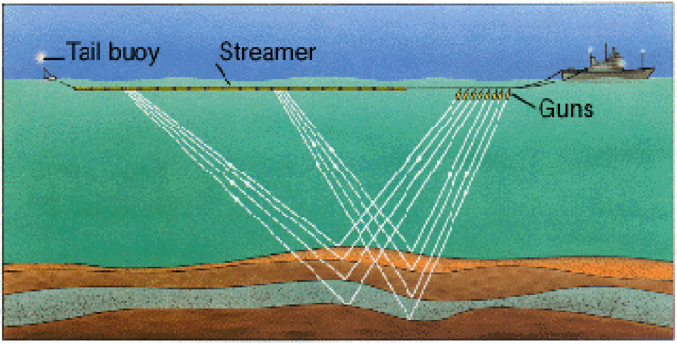
Introduction

Focus: recent developments in waveform inversion (WI) for *velocity*, and relation to migration velocity analysis (MVA).

Main topics:

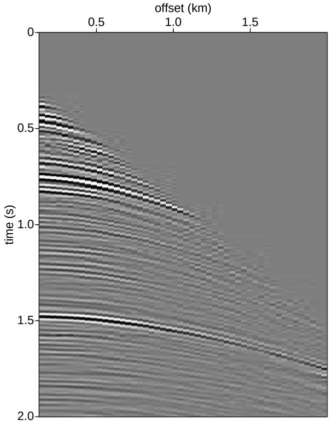
- Why inversion via least squares data fitting ("waveform inversion") doesn't work for exploration seismology;
- How migration is an approximate solution of the linearized inverse problem;
- How "Kirchhoff" and "Wave Equation" prestack depth migration differ, and what that means for migration velocity analysis;
- How to formulate migration velocity analysis via optimization, use all events;
- How to view migration velocity analysis as a solution of a "partly linear" waveform inversion problem;
- How nonlinear waveform inversion might be integrated with migration velocity analysis.

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.

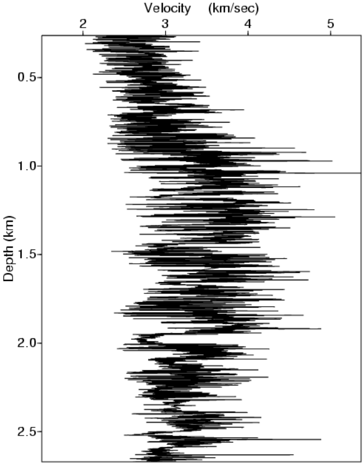
Typical Shot Record



CMP gather from North Sea Survey (thanks: Shell).
Processing applied:

- bandpass filter 3-8-25-35 Hz;
- 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 $= \Omega = \mathbf{R}^3$. 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 initial and boundary conditions.

Lions, late '60's: problem well posed for $v \in \mathcal{A}_0 = \{\log v \in L^\infty(\Omega)\}$, RHS in $L^2([0,T]\times\Omega)$.

Forward map: $\mathcal{F} : \mathcal{A}_0 \to L^2(\Sigma \times [0,T])$, $\Sigma \subset \{x_3 = 0\} \times \{x_3 = 0\}$ open, samples pressure:

$$\mathcal{F}[v](t,\mathbf{x}_r;\mathbf{x}_s) = \left(\phi\frac{\partial u}{\partial t}\right)(t,\mathbf{x}_r;\mathbf{x}_s),\ (t,\mathbf{x}_r,\mathbf{x}_s) \in [0,T]\times\Sigma,\ \phi \in C_0^\infty(\Sigma)$$

If $v = v_0$ known & constant in $\{x_3 < z\}$ for some $z > 0$, slight extension of Lions shows $\mathcal{F}$ well-defined. Stolk 2000: continuous, diffb'le "with loss of derivative".

Agenda

Inversion Generalities

The usual set-up:

- $\mathcal{M}$ = a set of *models* ($v \in \mathcal{A}_0$);
- $\mathcal{D}$ = a Hilbert space of (potential) data ($L^2([0,T] \times \Sigma)$);
- $\mathcal{F} : \mathcal{M} \to \mathcal{D}$: modeling operator or "forward map".

Waveform inversion problem: given $d \in \mathcal{D}$, find $v \in \mathcal{M}$ so that $\mathcal{F}[v] \simeq d$.
$\mathcal{F}$ can incorporate *any physics* - acoustics, elasticity, anisotropy, attenuation,.... (and v may be more than velocity...).

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

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

Output Least Squares Inversion

Given $d \in \mathcal{D}$, find $v \in \mathcal{M}$ to minimize

$$J_{OLS}(v,d) = \frac{1}{2}\|d - \mathcal{F}[v]\|^2 \equiv \frac{1}{2}(d - \mathcal{F}[v])^T(d - \mathcal{F}[v])$$

Has long and productive history in geophysics - but not in reflection seismology.

Only Newton and relatives scale well - but these find only local minima. Unfortunately, J_{OLS} has lots of local minima having nothing to do with "truth", for typical length, time, and frequency scales of exploration seismology.

$\Rightarrow$ least squares waveform inversion with Newton-like iteration "doesn't work" (Gauthier 86, Kolb 86, Santosa & S. 89, Bunks 95, Shin 01, Shin and Min 06, many others - see Chung SI 2.4).

Output Least Squares Inversion

Simple but instructive example: 1D reflection, $v = v(z)$, wavefield is plane wave at normal incidence with wavelet $w(t)$.

At constant velocity $v(z) \equiv v_0$,

$$\nabla J_{OLS}[v_0](z) = \mathrm{const.} \left(\frac{d\check{w}}{dt} * (d - w) \right) \left(\frac{2z}{v_0} \right)$$

where $\check{w}(t) = w(-t)$.

If data (hence w) contains no energy at frequencies below $f_{\min}$, then gradient contains no energy at spatial wavelengths longer than $v_0/(2f_{\min})$ $\Rightarrow$ first step of Newton does not even begin to reconstruct nonzero mean deviations if $z_d > v_0/(4f_{\min})$.

Output Least Squares Inversion

Upgrade to layered (or near-layered) media via plane wave expansion, range of incidence angles θ: $v_0 \rightarrow v_0/\cos\theta$.

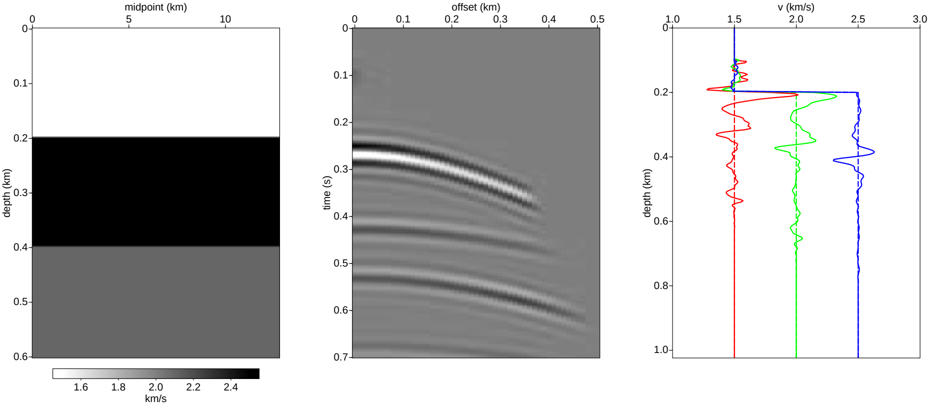
$\Rightarrow$ using reflection data with incidence angles $\leq 60^\circ$, gradient-based method cannot update mean velocity over depth interval $[0, z_d]$ if

$$f_{\rm min} > \frac{v_0}{2z_d}$$

Typical shallow sediment imaging: $v_0 \simeq 3$ km/s, $z_d = 5$ km $\Rightarrow$ to recover nonzero mean deviation from 3 km/s must have significant energy at $f_{\rm min} \simeq 0.3 Hz$ (cf. Bunks 95, Chung SI 2.4, Pillet SI 4.5)

If not present (energetics - Ziolkowski 93) and/or filtered from data, and if $\langle v \rangle \neq v_0$, then spurious minima must exist (and will be found by gradient-based optimization from $v_{\rm init} = v_0$)!

Output Least Squares Inversion



Left: Layered model. Middle: response to point source in center, 4-10-30-40 Hz bandpass wavelet. Right: OLS inversions, dashed=initial, solid=final. Quasi-Newton iteration terminated when gradient reduced by 10^{-2}.

Output Least Squares Inversion

Examples of successful waveform inversion from synthetic data containing very low frequencies ($<< 1$ Hz): Bunks 95, Shin and Min 06.

Another grand class of examples: basin inversion from earthquake data: target of several major efforts. QuakeShow (Ghattas), SpecFEM3D (Tromp, Komatisch), SPICE (Käser, Dumbser). Typical $z_d = 20$ km, $f_{\min} = 0.1Hz$, $\langle v_s \rangle = 4$ km/s - just OK! Will be done, in 3D, in near future.

Transmission waveform inversion less sensitive to lack of low frequencies (Gauthier 86, Mora 89) but still can fail in same way.

Another way to look at it: inversion will succeed if $v_{\rm init}$ gives accurate arrival times to within $0.5\lambda_{\min}$ - then in effect $f_{\min}$ replaced by $f_{\max}$.

Basis for very successful transmission waveform tomography of Pratt 99, Brenders TOM 1.5.

Agenda

The Seismologist's Standard Model

[Thanks: P. Lailly]. Because (1) $\mathcal{F}$ is hard to understand, (2) it's a lot simpler, and (3) it works sometimes, assume separation of scales: v [and other mechanical parameters] superposition of:

- smooth macromodel v: the long-scale component of velocity etc. (scales $\simeq$ 1 km and larger).
- oscillatory perturbation δv: high-frequency component of the velocity, scale $\simeq$ 10's of m (wavelength).

$$\left(\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} - \nabla^2\right)\delta u(t,\mathbf{x};\mathbf{x}_s) = \frac{2\delta v(\mathbf{x})}{v^3(\mathbf{x})}\frac{\partial^2 u}{\partial t^2}(t,\mathbf{x};\mathbf{x}_s),\ D\mathcal{F}[v]\delta v = \left.\frac{\partial \delta u}{\partial t}\right|_{[0,T]\times\Sigma}$$

Linearized Acoustic Inverse Problem

v smooth, r oscillatory (or even singular):

$$\left(\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} - \nabla^2\right)\delta u(t,\mathbf{x};\mathbf{x}_s) = \frac{2\delta v(\mathbf{x})}{v^3(\mathbf{x})}\frac{\partial^2 u}{\partial t^2}(t,\mathbf{x};\mathbf{x}_s),\ D\mathcal{F}[v]\delta v = \left.\frac{\partial \delta u}{\partial t}\right|_{[0,T]\times\Sigma}$$

Admissible sets of macromodels: bounded $\mathcal{A} \subset C^\infty(\Omega)$,...

Beylkin 85, Bleistein 87, Rakesh 88, Burridge 89, Nolan 97, de Hoop 97, ten Kroode 98, Stolk 00: under ever-weaker conditions, $D\mathcal{F}[v]^* D\mathcal{F}[v]$ is (microlocally) invertible *pseudodifferential operator*.

Means: $D\mathcal{F}[v]$ *almost unitary*, $D\mathcal{F}[v]^*(d - \mathcal{F}[v])$ has same (near-)singularities as δv, differs by scaling (S., SI 2.2) - image of δv.

$D\mathcal{F}[v]^* =$ prestack depth migration operator.

Linearized Acoustic Inverse Problem

Partially linearized inverse problem = Migration Velocity Analysis: given $d \in \mathcal{D} \equiv L^2([0,T] \times \Sigma)$, find $v \in \mathcal{A}, \delta v \in \mathcal{E}'(\Omega)$ so that

$$D\mathcal{F}[v]\delta v \simeq d - \mathcal{F}[v].$$

Least squares approach no more successful than for basic IP. Instead, industry has developed migration velocity analysis methods.

Based on image volume I output by prestack depth migration - function of subsurface position $\mathbf{x}$ and other (redundant) parameters. Two major variants: *surface-oriented* and *depth-oriented*.

Recent advance: understanding the difference.

Migration and Imaging Conditions

(I) Surface oriented: diffraction sum representation of image volume

$$I_S(\mathbf{x},\mathbf{h}) = \sum_{\mathbf{m}} (...) d(\mathbf{m},\mathbf{h}, T[v](\mathbf{x},\mathbf{m}-\mathbf{h}) + T[v](\mathbf{x},\mathbf{m}+\mathbf{h}))$$

where $\mathbf{h} = 0.5(\mathbf{x}_r - \mathbf{x}_s)$ = half-offset, $\mathbf{m} = 0.5(\mathbf{x}_r + \mathbf{x}_s)$ = midpoint, $T(\mathbf{x},\mathbf{y})$ = one-way time from $\mathbf{x}$ to $\mathbf{y}$. $(...)$ = optional amplitude, antialiasing,.. Data with same $\mathbf{h}$ = *offset bin*.

Relation is *binwise*: offset bin of image depends only on corresponding offset bin of data, hence "common offset". Diffraction sum is only comp. feasible implementation, hence "Kirchhoff" migration.

Other binwise migrations: common shot, common receiver, common scattering angle...

Migration and Imaging Conditions

(II) Depth oriented image volume: also has diffraction sum representation

$$I_D(\mathbf{x},\mathbf{H}) = \sum_{\mathbf{m}}\sum_{\mathbf{h}}(...)d(\mathbf{m},\mathbf{h},\, T[v](\mathbf{x}-\mathbf{H},\mathbf{m}-\mathbf{h}) + T[v](\mathbf{x}+\mathbf{H},\mathbf{m}+\mathbf{h}))$$

H is *space shift* or depth offset vector - unrelated to acquisition geometry.

Note extra summation over **h**: every image value depends on *all* traces.

Usual implementation via one-way WE (shot profile or DSR, Claerbout 85) or two-way RTM (Biondi-Shan 02, S. 02) (hence "wave equation" migration).

Transform to scattering angle available - Prucha 99, Sava and Fomel 01. Time shift variant - Sava and Fomel 05.

Migration and Imaging Conditions

Imaging conditions: how to extract image from image volume.

(I) Surface oriented: stack over offset

$$I(\mathbf{x}) = \sum_{\mathbf{h}} I_S(\mathbf{x}, \mathbf{h})$$

(II) Depth oriented: extract zero (depth) offset section

$$I(\mathbf{x}) = I_D(\mathbf{x}, \mathbf{0})$$

NB: These really are the same! In both cases, $I \simeq D\mathcal{F}[v]^*(d - \mathcal{F}[v])$ (high freq asymptotic approximation).

So both variants produce same image...

Migration and Imaging Conditions

But not the same image volume!

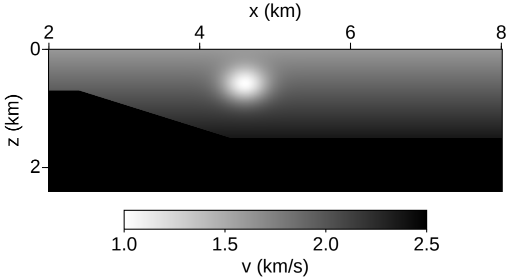
Nolan & S. 97, Stolk & S. 04, deHoop & Brandsberg-Dahl 03: multipathing (multiple rays connecting source, receiver, and image points, caustics) leads to artifacts in surface oriented image volume.

Artifact = coherent event in wrong place, of strength comparable to correct events.

Stolk & deHoop 01, S. 02, Stolk 05: depth-oriented image volume generally free of artifacts, even with strong multipathing.

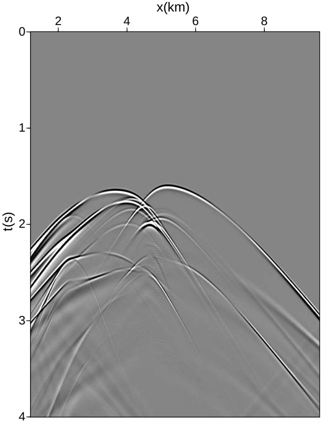
So the two types of image volume are not even kinematically equivalent! Accounts for perceived superiority of "wave equation migration". Consequences for velociy analysis: Nolan and S. 97, Xu TOM 1.4.

Migration and Imaging Conditions



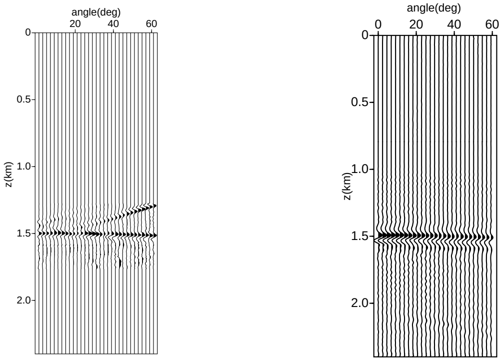
Velocity model after Valhall field, North Sea. Note sloping reflector at left, large low-velocity lens (modeling gas accumulation) in center. Both tend to produce multipathing. (Thanks: M. de Hoop, A. Malcolm)

Migration and Imaging Conditions



Typical shot gather over center of model, exhibiting extensive multipathing.

Migration and Imaging Conditions



Angle common image gathers at same horizontal position from surface-oriented (Kirchhoff) and depth-oriented (DSR) migrated image volumes. Left: ACIG from Kirchhoff migration: kinematic artifacts clearly visible. Right: ACIG from DSR migration: no artifacts!

Agenda

Semblance

Semblance condition - complementary to imaging condition.

Expresses consistency between data, velocity model in terms of image volume.

(I) Surface oriented: velocity-data consistency when $I_S(\mathbf{x}, \mathbf{h})$ independent of $\mathbf{h}$ (at least in terms of phase), i.e. image gathers are flat.

(II) Depth oriented: velocity-data consistency when $I_D(\mathbf{x}, \mathbf{H})$ concentrated near $\mathbf{H} = 0$, i.e. image gathers are focused [or flat, when converted to scattering angle].

Semblance

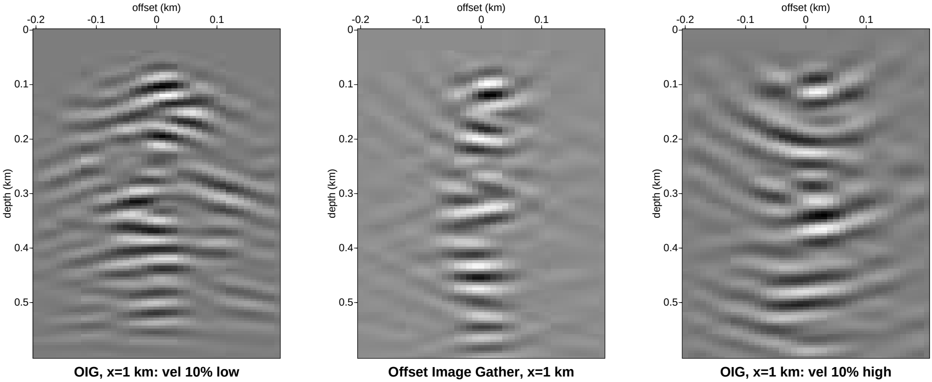


Image gathers $\{I_D(\mathbf{x}, \mathbf{H}) : x, y \,\mathrm{fixed}, \mathbf{H} = (0, h, 0)\}$, amplitude vs. (z, h), from velo model $v_0 + \delta v$, $v_0 =$ const., $\delta v =$ randomly distributed point diffractors. Left to Right: use $v = 90\%$, 100%, 110% of true velocity v_0.

Semblance

Leads to two methods for velocity updating:

(I) Depth domain reflection traveltime tomography:

- (auto)pick events in migrated image volume
- backproject inconsistency (eg. residual moveout of angle gather events) to construct velocity update as in standard traveltime tomography.

Used with both surface oriented and depth oriented image volume formation.

Drawback: uses only small fraction of events in typical image volume.

Semblance

(II) Depth domain reflection waveform tomography ("differential semblance"):

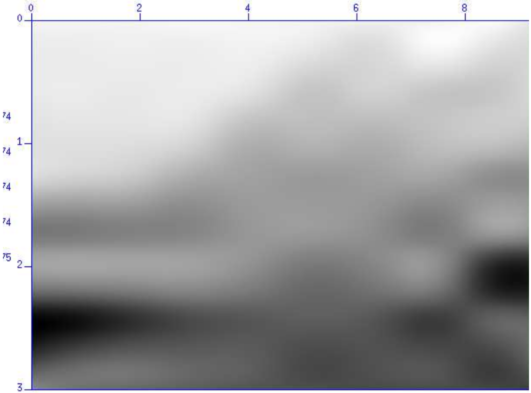
- form measure of deviation of image volume from semblance condition - function of velocity model; all energy not conforming to semblance condition contributes.
- optimize numerically: gradient = backprojection of semblance-inconsistent energy into velocity update.

Also used with both surface and depth oriented image volumes. Recent contributions: Shen 03, 05, Li & S. 05, Foss 06, Albertin 06, Khoury 06, Verm 06, Kabir SVIP 2.3.

Inherently uses all events in data, weighted by strength.

Example: minimize $J[v] = \sum |\mathbf{H} I_D(\mathbf{x}, \mathbf{H})|^2$ - penalizes energy at $\mathbf{H} \neq 0$.

Synthetic Example (Shen SEG 05)



Starting velocity model for waveform tomography. Data: Born version of Marmousi, fixed receiver spread across surface.

Synthetic Example (Shen SEG 05)

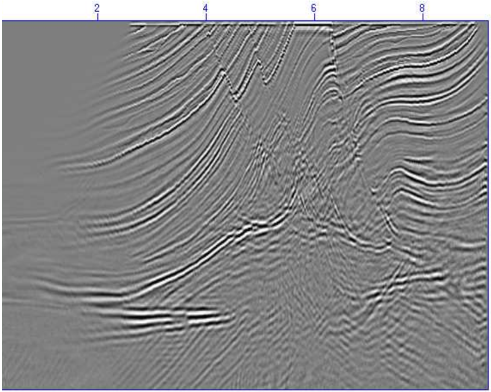
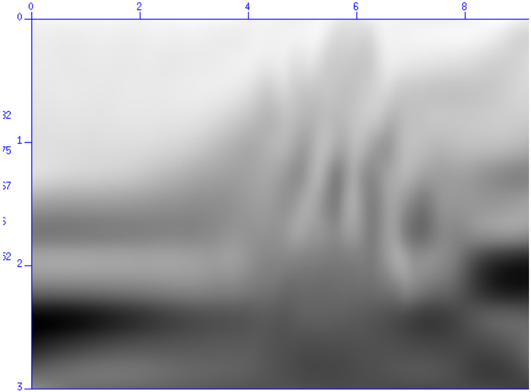


Image ($I_D(\mathbf{x}, \mathbf{H} = 0)$) at initial velocity.

Synthetic Example (Shen SEG 05)



Final velocity (47 iterations of descent method). Note appearance of high velocity fault blocks.

Synthetic Example (Shen SEG 05)

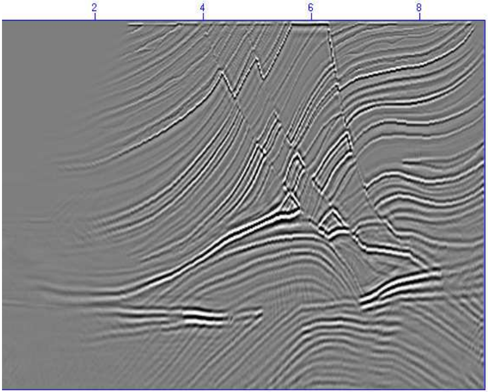


Image ($I_D(\mathbf{x}, \mathbf{H} = 0)$) at final velocity.

Field Example - Trinidad (Kabir SVIP 2.3)

[see Expanded Abstract, SEG 07.]

Agenda

1. Waveform Inversion
2. Migration Velocity Analysis
3. Semblance and Optimization
4. Extended Modeling: MVA + WI
5. Conclusions and Prospects

Extended Modeling

Where we are:

(i) WI lets you model any physics at all, and use all of the data, but doesn't work (spurious local minima);

(ii) MVA works - can even be made into an optimization without spurious local minima ("waveform tomography", differential semblance) - but only produces velocity, and assumes linearized model (single scattering, Born approximation, primaries-only data,...).

Can the two be combined somehow, while retaining their advantages?

Extended Modeling

Partial Answer: MVA solves a version of WI! To see this, need *extended modeling* concept, plus true amplitude migration or *linear inversion*.

Extended model $\bar{\mathcal{F}} : \bar{\mathcal{M}} \rightarrow \mathcal{D}$, where $\bar{\mathcal{M}}$ is a *bigger model space*. Physical model space $\mathcal{M}$ in 1-1 correspondence with a subset of $\bar{\mathcal{M}}$, via *extension map* χ.

For surface-oriented extended modeling, extended models depend on $\mathbf{h}$, and $\chi[v](\mathbf{x}, \mathbf{h}) = v(\mathbf{x})$, i.e. χ produces models not depending on $\mathbf{h}$.

For depth-oriented extended modeling, extended models depend on $\mathbf{H}$, and $\chi[v](\mathbf{x}, \mathbf{H}) = v(\mathbf{x})\delta(\mathbf{H})$, i.e. χ produces models focused at zero offset.

In either case, output of χ is an "image volume" satisfying the semblance condition, and vis-versa - which "explains" semblance condition.

Extended Modeling

Lailly, Tarantola, Claerbout (80's): migration operator (producing image) is *adjoint* or transpose $D\mathcal{F}[v]^*$. *True amplitude* migration is (pseudo)inverse $D\mathcal{F}[v]^{-1}$.

Same relation with extended modeling: migration operator (producing image *volume*) is adjoint $D\bar{\mathcal{F}}[\chi[v]]^*$ of linearized extended modeling operator. True amplitude migration defines (pseudo)inverse $D\bar{\mathcal{F}}[\chi[v]]^{-1}$.

For depth orientation, diffraction sum representation is

$$D\bar{\mathcal{F}}[\chi[v]]\delta\bar{v}(\mathbf{m},\mathbf{h},t)$$

$$= \int d\mathbf{x}d\mathbf{H}\delta\bar{v}(\mathbf{x},\mathbf{H})\delta(t - T(\mathbf{x}-\mathbf{H},\mathbf{m}-\mathbf{h}) - T(\mathbf{x}+\mathbf{H},\mathbf{m}+\mathbf{h}))$$

Easy to check: $D\bar{\mathcal{F}}[\chi[v]]^T d(\mathbf{x},\mathbf{H}) = I_D(\mathbf{x},\mathbf{H})$.

Extended Modeling

Claim: MVA (with true amplitude) solves "partially linearized" problem: find reference velocity v and perturbation δv so that $D\mathcal{F}[v]\delta v \simeq d - \mathcal{F}[v]$.

Proof: successful MVA produces image volume satisfying imaging condition, i.e. $I_D = \chi[\delta v]$.

Use true amplitude migration, and you get $D\bar{\mathcal{F}}[\chi[v]]^{-1}(d - \mathcal{F}[v]) \simeq \chi[\delta v]$, whence

$$D\mathcal{F}[v]\delta v = D\bar{\mathcal{F}}[\chi[v]]\chi[\delta v]$$

$$\simeq D\bar{\mathcal{F}}[\chi[v]]D\bar{\mathcal{F}}[\chi[v]]^{-1}(d - \mathcal{F}[v]) \simeq d - \mathcal{F}[v]$$

Q.E.D.

Extended Modeling

Linearization of what?

Exercise for reader: for $V =$ SAPD op on appropriate Hilbert space, define $\bar{\mathcal{F}}[V] \equiv \frac{\partial \bar{u}}{\partial t}\Big|_{[0,T]\times\Sigma}$, where

$$V^{-2}\frac{\partial^2 \bar{u}}{\partial t}^2 - \nabla^2 \bar{u} = w(t)\delta(\mathbf{x} - \mathbf{x}_s)$$

Suppose distribution kernel of V is $v(\mathbf{x})\delta\left(\frac{\mathbf{x}-\bar{\mathbf{x}}}{2}\right) + \delta\bar{v}\left(\frac{\mathbf{x}+\bar{\mathbf{x}}}{2}, \frac{\mathbf{x}-\bar{\mathbf{x}}}{2}\right)$, then $D\bar{\mathcal{F}}[\chi[v]]\delta\bar{v}$ has diffraction sum representation given above: its adjoint is depth-oriented prestack migration! ($\mathbf{H} \sim \frac{\mathbf{x}-\bar{\mathbf{x}}}{2}$)

Existence theory for symmetric hyperbolic systems with operator coefficients after Lions 68.

Extended Modeling

So what? Well,

- In this scheme, $\mathcal{F}$ can be *any modeling* operator - acoustic, elastic, ... - known how to do true amplitude, thanks to Beylkin, Burridge, Bleistein, de Hoop,... So: MVA extended to elastic (Born) modeling, for instance.
- For depth-oriented extension, $\bar{\mathcal{F}}$ expresses action at a distance: elastic moduli are nonlocal, stress at $\mathbf{x}+\mathbf{H}$ results from strain at $\mathbf{x}-\mathbf{H}$. So Claerbout's semblance principle is actually Cauchy's no-action-at-a-distance hypothesis! [Thanks: Scott Morton]
- Nonlinear MVA via enforcing semblance = no-action-at-distance on elastic moduli, treated as operators - MVA incorporating multiple scattering = WI with extended modeling.

Agenda

Conclusions

Takeaway messages of this talk:

- Least squares WI prone to get trapped in useless local minima - avoidance requires either initial velocity estimates good to 0.5 (shortest) wavelength, or longest wavelength exceeding the survey depth.
- MVA: "Kirchhoff" (surface-oriented) and "Wave Equation" (depth-oriented) prestack migrations have different kinematic properties.
- MVA via waveform tomography ("differential semblance"), based on semblance condition and numerical optimization, uses all events to constrain velocity updates, much less tendency towards local minima than least squares WI.
- MVA solves a "partially linearized" WI problem based on *extended modeling* - nonphysical degrees of freedom.
- Nonlinear extended scattering = framework for uniting MVA and waveform inversion.

Prospects

- Two kinematically inequivalent extensions: surface-oriented and depth-oriented. Classify all extensions by microlocal equivalence.
- Waveform MVA via Reverse Time Migration (= full-blown computation of $D\bar{\mathcal{F}}[\chi[v]]^*$) and differential semblance - kinematic accuracy, fast linear inversion (SI 2.2, Moghaddam SPMI 3.2).
- Concepts other than differential semblance, least squares: van Leeuwen SI 2.8.
- Nonlinear inversion via model extension ("nonlinear MVA") including multiple scattering $\Rightarrow$ sparse representation of operator coefficients, introduction of "control" ($\sim$ migration velocity), integration of source estimation (Minkoff & S. 97).

Thanks to

- Total E&P USA for permission to show Shen SEG 05
- BP and Uwe Albertin for Trinidad gas sag example
- Sponsors of The Rice Inversion Project for their long-term support of my work
- Organizers, for inviting me to speak
- All of you, for listening