The Rice Inversion Project 2012 Review

William Symes

Computational and Applied Mathematics, Rice University

April 2013

Agenda

Overview

Linearized Modeling and Inversion

Extended modeling and MVA

Extended Waveform Inversion

People

- Yin Huang (CAAM, 3rd year grad)
- Lei Fu (ESCI, 2nd year grad)
- Muhong Zhou (CAAM, 2nd year grad)
- Mario Bencomo (CAAM, 2nd year grad)
- Jie Hou (ESCI, 1st year grad)
- Papia Nandi (BP & ESCI, 1st year grad)
- Yujin Liu (visitor AY '12-'13, China U. of Petroleum)
- William W. Symes (CAAM & ESCI, 29th year faculty)

Sponsors

- BHP Billiton
- BP
- Chevron
- ConocoPhillips
- ExxonMobil
- Hess
- ION-GXT
- Landmark Graphics
- Shell
- Statoil
- Total
- WesternGeco

Goal

contribute mathematical and computational innovations to development of waveform inversion for seismic exploration

Projects

- Migration Velocity Analysis without picking (Yujin, Yin, Jie)
- Combining Full Waveform Inversion with MVA (Yujin, Yin, Lei, Papia)
- Fast modeling in heterogeneous media (Muhong, Mario)

Agenda

Overview

Linearized Modeling and Inversion

Extended modeling and MVA

Extended Waveform Inversion

M = model space = mechanical parameter fields (bulk modulus, density, $C_{ijkl}(\mathbf{x})$,...)

D = data space = $\{d(\mathbf{x}_r, \mathbf{x}_s, t)\}$

$F : M \to D$ modeling operator = forward map = solve wave equations for pressure, displacement,..., sample at $\mathbf{x}_r, t$ (RHS = function of $\mathbf{x}_s$)

Full Waveform Inversion problem:

given $d \in D$, find $m \in M$ so that

$$F[m] \simeq d$$

Least squares inversion = FWI: minimize

$$J_{LS}[m] = \|F[m] - d\|^2 [+ \text{ regularizing terms}]$$

Example: constant density acoustics
$M = \{\kappa(\mathbf{x}) = \rho c^2(\mathbf{x})\}$, isotropic point radiator

$$\left(\frac{\partial^2}{\partial t^2} - c^2(\mathbf{x})\nabla_{\mathbf{x}}^2\right) p(\mathbf{x}, \mathbf{x}_s, t) = \delta(\mathbf{x} - \mathbf{x}_s) w(t);$$

$$p \equiv 0, t << 0$$

$F[c^2] = \{p(\mathbf{x}_r, \mathbf{x}_s, t)\}$ (solve wave equation, sample pressure field)

Bandlimited source $\Rightarrow$ derivative DF makes sense (aka linearized fwd map, Born modeling operator, etc. etc. - Stolk 00, Blazek et al 13)

Least squares inversion for linearized map: Given d find $m, \delta m$ so that

$$\min_{m,\delta m} \|DF[m]\delta m - (d - F[m])\|^2 + \ldots$$

When m is included as inversion target, no easier than non-linearized inverse problem!

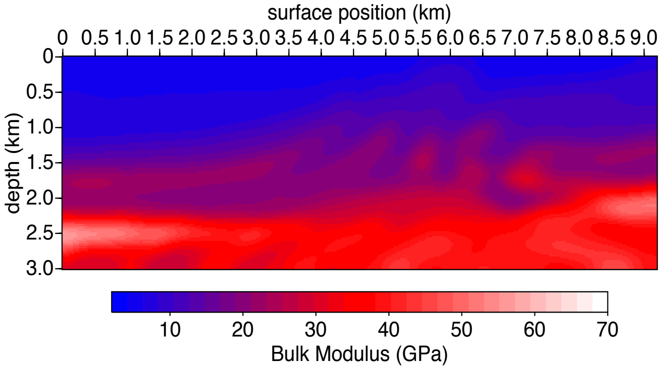
Example: constant density acoustics: background $c^2(\mathbf{x})$, perturbation $\delta c^2(\mathbf{x})$

$$\left(\frac{\partial^2}{\partial t^2} - c^2(\mathbf{x})\nabla^2\right) p(\mathbf{x}, \mathbf{x}_s, t) = \delta(\mathbf{x} - \mathbf{x}_s) w(t)$$

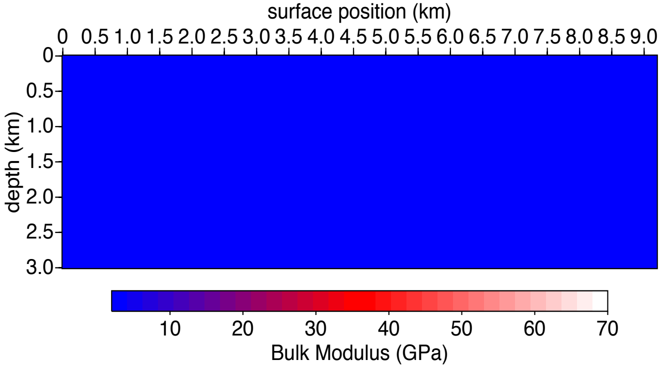
$$\left(\frac{\partial^2}{\partial t^2} - c^2(\mathbf{x})\nabla^2\right) \delta p(\mathbf{x}, \mathbf{x}_s, t) = \delta c^2(\mathbf{x})\nabla^2 p(\mathbf{x}, \mathbf{x}_s, t)$$

$$p, \delta p = 0, t << 0$$

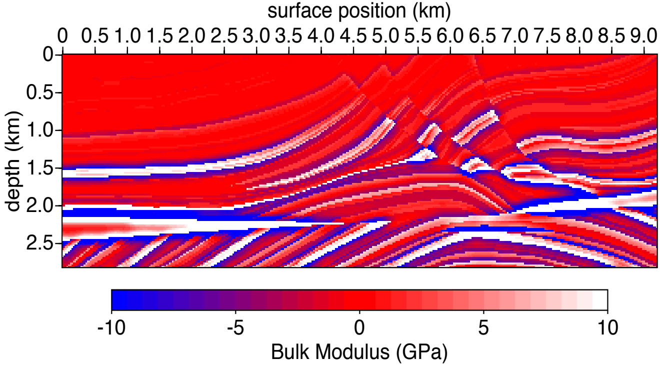
$$DF[c^2]\delta c^2 = \{\delta p(\mathbf{x}_r, \mathbf{x}_s, t)\}$$



Smoothed Marmousi model $= c_{\mathrm{sm}}^2 =$ "100%"
(240 m $\times$ 240 m bilinear hat)



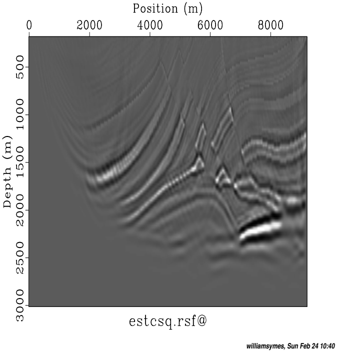
Deep blue sea $= c_0^2 =$ “0%”

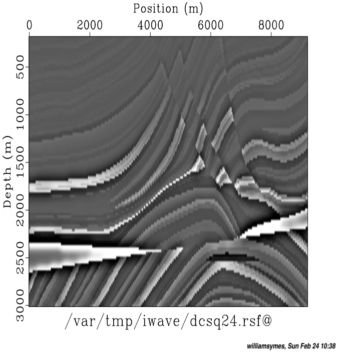


$$\delta c^2 = c_1^2 - c_{\mathrm{sm}}^2$$

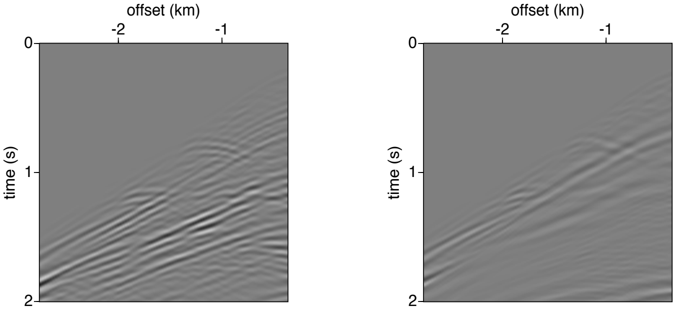
RICE

For noise-free data, correct m: recovery of correct δm, small data residual





Linearized inversion $m = c_{\rm sm}^2$ ("100%"), from 60 shot records 100m spacing. Left: estimated δc^2, 50 CG its. Right: target δc^2



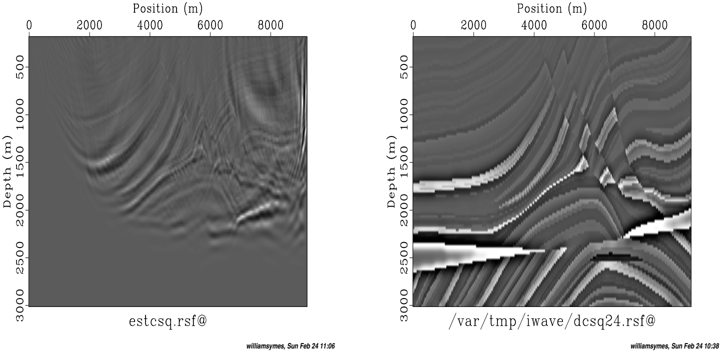
Relative RMS residual = 0.2. Left: predicted data, Right, residual, shot position 6 km

However, small errors in $m \Rightarrow$ large residual at optimal δm

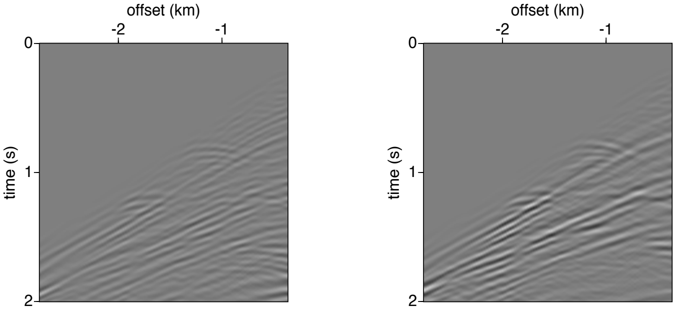
Away from optimum, small changes in m have little effect on data fit - *data fit does not encode m*

If m is substantially incorrect, *no δm is consistent with data*

RICE



Linearized inversion $m = 0.8c_{\rm sm}^2 + 0.2c_0^2$ ("80%"), from 60 shot records 100m spacing. Left: estimated δc^2, 50 CG its. Right: target δc^2



Relative RMS residual = 0.8. Left: predicted data, Right, residual, shot position 6 km

Agenda

Overview

Linearized Modeling and Inversion

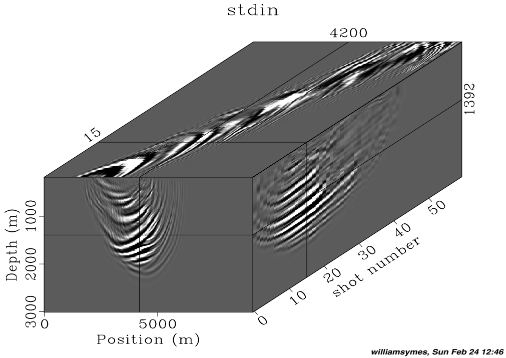
Extended modeling and MVA

Extended Waveform Inversion

Empirical observation (eg. Taner & Koehler 69):

May fit data subsets with perturbation δm, even with erroneous background m

Example of "fitable" subsets: *shot records* (data for single $\mathbf{x}_s$)



Inversion of δm ($= \delta c^2$) for each shot record independently, "80%" m ($= 0.8 c_{\rm sm}^2 + 0.2 c_0^2$) - Relative RMS Error = 0.10

M = physical model space

$\bar{M}$ = *bigger* extended model space

$\bar{F} : \bar{M} \to D$ extended modeling operator

Extension property: $M \subset \bar{M}$,
$m \in M \Rightarrow \bar{F}[m] = F[m]$

Linearized extension: extended perturbation about physical background model: $D\bar{F}[m]\overline{\delta m}$

Linearized acoustics, *shot record extension*:

$\bar{M} = \{c^2(\mathbf{x}), \overline{\delta c^2}(\mathbf{x}, \mathbf{x}_s)\}$

$$\left(\frac{\partial^2}{\partial t^2} - c^2(\mathbf{x})\nabla^2\right) p(\mathbf{x}, \mathbf{x}_s, t) = \delta(\mathbf{x} - \mathbf{x}_s)w(t)$$

$$\left(\frac{\partial^2}{\partial t^2} - c^2(\mathbf{x})\nabla^2\right) \overline{\delta p}(\mathbf{x}, \mathbf{x}_s, t) = \overline{\delta c^2}(\mathbf{x}, \mathbf{x}_s)\nabla^2 p(\mathbf{x}, \mathbf{x}_s, t)$$

$p, \overline{\delta p} = 0, t << 0$

$D\bar{F}[c^2]\overline{\delta c^2} = \{\overline{\delta p}(\mathbf{x}_r, \mathbf{x}_s, t)\}$

$D\bar{F}[c^2]^T =$ *prestack RTM* operator for shot record images

One way to deal with underdetermination of extended inversion, extract velocity information:

Semblance operator ("annihilator") $A : \bar{M} \to Z$ s.t.

$$A\overline{\delta m} = 0 \;\Leftrightarrow\; \overline{\delta m} \in M$$

Expl: for acoustics, shot record extension, possible choice is $A = P\nabla_{\mathbf{x}_s}$, $P =$ any ΨDO. Order P = -1 $\Rightarrow$ order A = 0

MVA by Semblance optimization: minimize over $m, \overline{\delta m}$

$$J[m, \overline{\delta m}] = \frac{1}{2}\|D\bar{F}[m]\overline{\delta m} - \delta d\|^2 + \frac{\lambda^2}{2}\|A\overline{\delta m}\|^2$$

- small λ limit: *inversion velocity analysis* = migration velocity analysis with LS migration
- large λ limit: least squares inversion for linearized modeling

Semblance optimization via shot record & similar acquisition gather extensions:

- Kern & Symes 94, Mulder & ten Kroode 02, Chauris & Noble 01, Brandsberg-Dahl et al 03
- extension to full nonlinear propagation: S. 08, Sun & S SEG 12, Sun thesis
- recent: Chauris, Perrone, Almomin,...

Semblance optimization via shot record & similar acquisition gather extensions:

+ computational cost essentially same as FWI
- restrictive geometric conditions: effective only for mild lateral heterogeneity - no multipathing, caustics (Nolan & S. 97, Stolk & S. 04)

Today: Yin (1015), Papia (1050)

Semblance optimization via survey-sinking (or shot-geophone or space-shift or ...) extension

- Claerbout 85, Sava & Fomel
- semblance optimization: Shen et al 03, 05, Kabir et al. 06, Khoury et al. 06, Vyas et al 10, Fei & Wiliamson 10, Albertin 10, Almomin & Biondi 12, Weibull & Arntsen 12, Yang & Sava 12, Biondi & Zhang 12,...
- extension to full nonlinear propagation: S. 08, Biondi & Almomin 12

Semblance optimization via survey-sinking extension: compared to surface-gather extensions,

+ less restrictive ray geometric conditions: effective for strong lateral heterogeneity, some multipathing (de Hoop & Stolk 01, de Hoop et al 09, Biondi & S. 04)
- straightforward implementation has high computational complexity

Today: Yujin (0945, 1500), Lei (1400, 1520), Jie (1420)

Beyond semblance optimization: Peng (1315)

- choose vector field V ("image residual") on extended model space $\bar{M}$: limit points $=$ physical models M
- given data d and background model m, compute $\overline{\delta m}$ by migration (or inversion)
- pull back $V(\overline{\delta m})$ to background model update by least squares fit

Contraction vector field $h\partial_h \Rightarrow$ (modified) Fei-Williamson (SEG 10) update

Related to Biondi-Sava 04, other "warping" methods, but no picking

A whole new world of MVA algorithms, NOT optimizations (updates not gradients)!

Implementing MVA by Semblance optimization: minimize over $m, \overline{\delta m}$

$$J[m, \overline{\delta m}] = \frac{1}{2}\|D\bar{F}[m]\overline{\delta m} - \delta d\|^2 + \frac{\lambda^2}{2}\|A\overline{\delta m}\|^2$$

Not an improvement: in acoustic expl, $w = \phi_\epsilon * \delta^{(\alpha)}$

$\Rightarrow \|D_m J[m, \overline{\delta m}]\| = O(\epsilon^{-1})$ (cf. Almomin-Biondi SEG 12)

Reduced objective (Yin, 1015):

$$\tilde{J}[m] = \min_{\overline{\delta m}} J[m, \overline{\delta m}]$$

1st order condition:

$$N[m]\overline{\delta m} \equiv (D\bar{F}[m]^T D\bar{F}[m] + \lambda^2 A^T A)\overline{\delta m} = D\bar{F}[m]^T \delta d$$

Key Observation: under suitable ray-geometry conditions, $N[m]$ invertible and smooth in m

$\Rightarrow$ computable gradient approximation with controlled accuracy

Understanding least squares inversion and reduced objective - Normal operator = Hessian:

$$D\bar{F}[m]^T D\bar{F}[m]$$

Mapping properties understood for various sub cases of elasticity,

- m smooth
- minor geometric restrictions

ray-theoretic restrictions depending on extension: $\Rightarrow D\bar{F}[m]^T D\bar{F}[m]$ is *pseudodifferential operator* ("ΨDO ")

Definition for singular sources, *order*, uniform approximation with band limited sources

Composition, inverses ("micro local" = local in phase space)

pseudo local $\Rightarrow$ smooth in m, uniformly wrt band limited source approximation

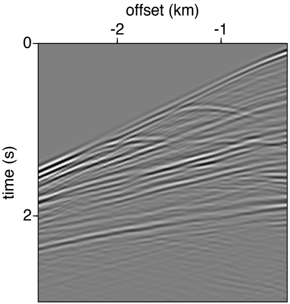
Acoustics, shot record extension:

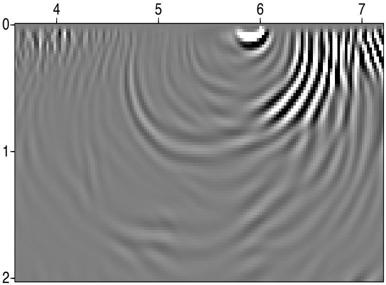
- rays from source, receivers to scattering points (supp $\overline{\delta c^2}$) may have *no conjugate points* ("no multipathing", no caustics) - fails otherwise (Nolan & S 97, Stolk & S 04)
- singular source $w = \delta^{(\alpha)}$, $\alpha = -1/2$ (2D) or $= -1$ (3D), then order $(D\bar{F}[c^2]^T D\bar{F}[c^2]) =$ zero (bounded op on L^2)
- $\{\phi_\epsilon\}_{\epsilon>0}$ Dirac family, $w = \phi_\epsilon * \delta^{(\alpha)} \Rightarrow$ L^2-bounded, uniformly in $\epsilon \to 0$ 0.5cm

Contributors:

Cohen and Bleistein 77, Beylkin 85, Rakesh 86, Bleistein 87, Beylkin & Burridge 88, Nolan & S 97, de Hoop & Bleistein 97, Burridge et al. 98, Smit et al 98, Stolk 00, de Hoop & Stolk 01, de Hoop, Stolk & S. 09,...

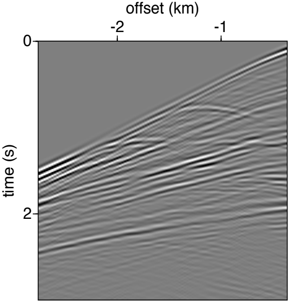
Illustration of pseudo-local property

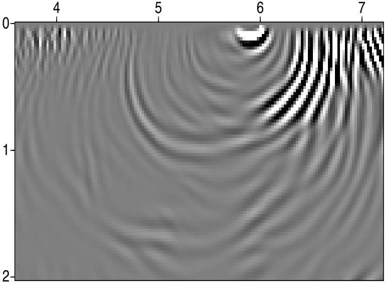




"100%": $D\bar{F}[m]\delta m$ (left), $D\bar{F}[m]^T D\bar{F}[m]\delta m$ (right)

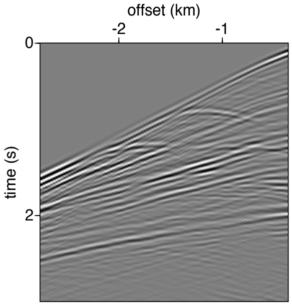
Illustration of pseudo-local property

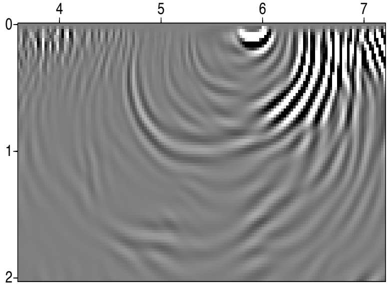




"95%": $D\bar{F}[m]\delta m$ (left), $D\bar{F}[m]^T D\bar{F}[m]\delta m$ (right)

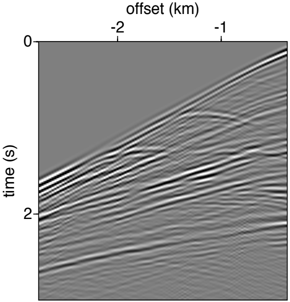
Illustration of pseudo-local property

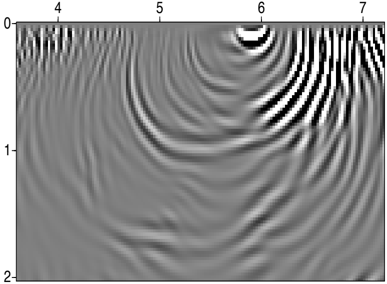




"90%": $D\bar{F}[m]\delta m$ (left), $D\bar{F}[m]^T D\bar{F}[m]\delta m$ (right)

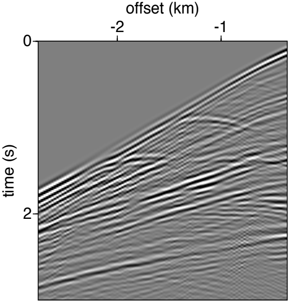
Illustration of pseudo-local property

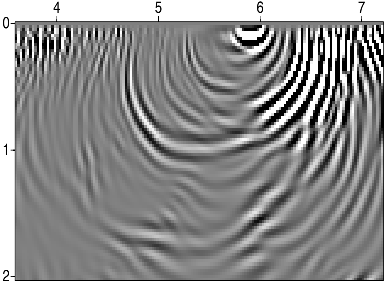




"80%": $D\bar{F}[m]\delta m$ (left), $D\bar{F}[m]^T D\bar{F}[m]\delta m$ (right)

Illustration of pseudo-local property





"70%": $D\bar{F}[m]\delta m$ (left), $D\bar{F}[m]^T D\bar{F}[m]\delta m$ (right)

$N[m]$ = invertible ΨDO order 0 $\Rightarrow$

$$\tilde{J}[m] = \frac{1}{2}\|(D\bar{F}[m]N[m]^{-1}D\bar{F}[m]^T - I)\delta d\|^2$$

$$+\frac{\lambda^2}{2}\langle \delta d, D\bar{F}[m]N[m]^{-1}A^T A N[m]^{-1}D\bar{F}[m]^T\delta d\rangle$$

ΨDO calculus $\Rightarrow$ under extension-dependent ray-theoretic conditions operators in inner products are ΨDOs, order 0, w symbols depending smoothly on m...

$\Rightarrow \tilde{J}$ is smooth in m, uniformly in ϵ: $D_m\tilde{J} = O_\epsilon(1)$

provided that $A^T A$ is a ΨDO of order 0.

In fact, if & only if (Stolk & S. 03)

Domain of convexity contains ball of radius $O_\epsilon(1)$

Algorithms for function value, gradient (1) requiring only solutions of wave equations, (2) convergent uniformly in ϵ with error controlled by normal residual (Kern & S. 94, Yin 1015)

Any computation requires iterative solution of normal equations. Preconditioning essential (Jie 1420)

and fast solution of wave equation (Muhong 1120, Mario 1140, Lei 1520)

RICE

Agenda

FWI+MVA

Extended nonlinear ("full") waveform inversion - sure why not - $D\bar{F}$ is derivative of a full waveform modeling operator $\bar{F}$

Shot-record & similar extensions: just let c^2 depend on $\mathbf{x}_s$ (S. 91, Sun & S. 12)

Survey-sinking: c^2 becomes *operator* (action at a distance) - S. 08, Biondi & Almomin 12

FWI+MVA

Semblance objective:

$$J[\bar{m}] = \frac{1}{2}\|\bar{F}[\bar{m}] - d\|^2 + \frac{\lambda^2}{2}\|A\bar{m}\|^2$$

As for Born case, fewer local mins than FWI but very ill-conditioned (cf Biondi-Almomin SEG 12)

D. Sun thesis: how to formulate reduced objective

Current projects (Yin, Yujin, Lei, Papia): analyze, extend

FWI+MVA

On the horizon: combine reduced objective construction for FWI+MVA with Peng's generalization of semblance

Thanks to...

- Many students and colleagues
- NSF
- Sponsors of The Rice Inversion Project
- The Distinguished and Very Patient Audience