Gradients and Hessians for extended Born Waveform Inversion

William Symes

The Rice Inversion Project - Annual Review 2014

RICE

Waveform Inversion

- efficiency ?
- reliability ?
- value ?

Waveform Inversion

- efficiency ?
- reliability ?
- value $\surd$

RICE

Waveform Inversion

- FWI $\times$ ("cycle skipping")
- extended FWI ("always fit data")
  - Born-based ("IVA")
  - Full-waveform based

Waveform Inversion

Theory $\Rightarrow$ EFWI

Practice $\Rightarrow$ FWI

What stands in the way of merging Theory with Practice: EFWI efficiency, reliability

RICE

Inversion Velocity Analysis

WEMVA, with ext'd LSM (Nemeth et al. 99,...) = ext'd linearized inversion

IVA Objective function = focusing measure

Reliability issue: gradient accuracy ("artifacts" - Fei & Williamson 10, Vyas & Tang 10, ...)

Inversion Velocity Analysis

Data $d \in$ Data Space D

Physical models $m \in$ physical model space M

Extended models (perturbational) $\delta \bar{m} \in$ ext'd model space $\bar{M}$

RICE

Inversion Velocity Analysis

Ext'd linearized fwd operator for $m \in M, : \bar{F}[m] : \bar{M} \rightarrow D$

Focus operator ("annihilator") $A : \bar{M} \rightarrow \ldots$

$A\delta\bar{m} = 0 \Leftrightarrow \delta\bar{m} \in M \subset \bar{M}$

RICE

Inversion Velocity Analysis

Can always fit data: $\bar{F}[m]$ "invertible"

$$J_{\mathrm{IVA}}[m] = \frac{1}{2}\|A\bar{F}[m]^{-1}d\|^2$$

$\|u\|$ = RMS of u, $\langle u, v\rangle$ = dot product of u, v

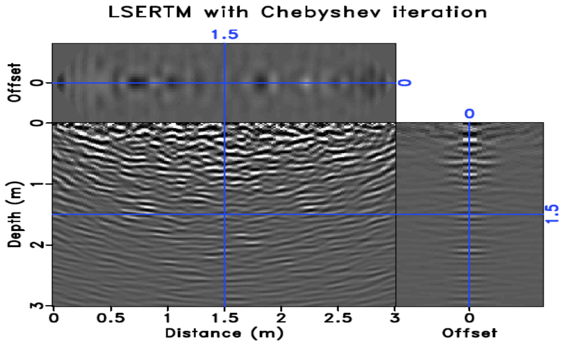
Inversion Velocity Analysis

Alternative: "traditional" WEMVA/DSO (Shen et al. 03,...):

$$J_{\mathrm{MVA}}[m] = \frac{1}{2}\|A\bar{F}[m]^T d\|^2$$

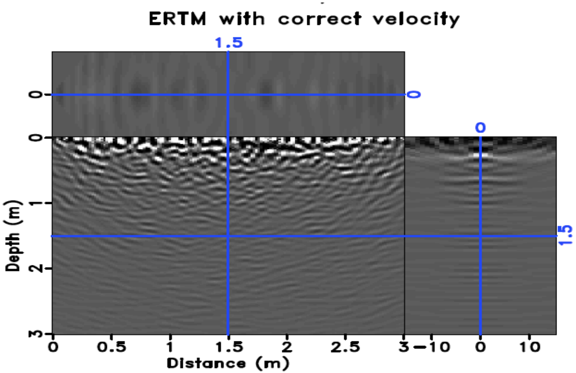
Example: 2D acoustics, subsurface offset extension (thanks: Y. Liu) ($A =$ mult. by h)

Inversion Velocity Analysis



$$\delta\bar{m} \approx \bar{F}[m]^{-1}d\text{: } m = 1.00 * m_{\text{true}}$$

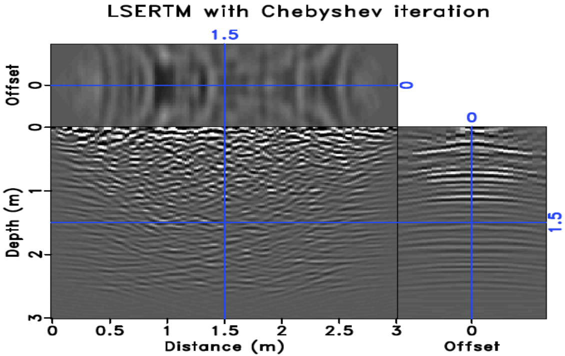
Inversion Velocity Analysis



$$\delta \bar{m} = \bar{F}[m]^T d: \; m = 1.00 * m_{\text{true}}$$

RICE

Inversion Velocity Analysis



$$\delta \bar{m} \approx \bar{F}[m]^{-1} d: \; m = 0.85 * m_{\text{true}}$$

Gradients

Gradient inaccuracy affects

- model resolution
- convergence rate of iterative optimization

Gradients

Computed gradient comparison, acoustics: Y. Liu, EAGE 14.

Densely sampled src, rec near surface, 10 Hz Ricker source

$m_{\mathrm{true}} = 3$ km/s. $\delta m_{\mathrm{true}} = 3$ trunc reflectors

Gradients

Amsterdam '14

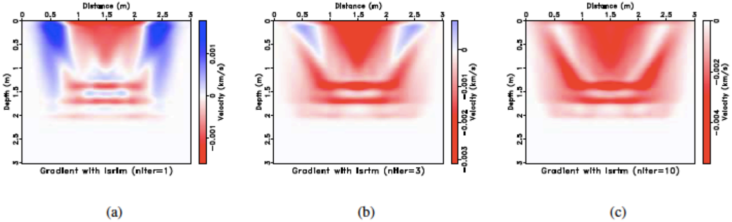


Figure 2 Background velocity gradient using extended reflectivity image by linearized inversion after different iteration number (a) iteration 1; (b) iteration 3; (c) iteration 10.

Left: $\nabla J_{\rm MVA}$ - artifacts! Right: $\nabla J_{\rm IVA}$, at $m =$ 2.5 km/s

RICE

Gradients

Rate of change: $m+ = \delta m \Rightarrow$

$$\delta J_{\mathrm{IVA}} = -\langle \bar{F}[m]^{-1} D\bar{F}[m](\delta m)\delta \bar{m}, A^T A \delta \bar{m}\rangle$$

$\delta \bar{m} = \bar{F}[m]^{-1} d \Rightarrow$ 2 solves (iterative!)

Gradients

Rate of change: $m+ = \delta m \Rightarrow$

$$\delta J_{\rm MVA} = \langle D\bar{F}[m]^T(\delta m)d, A^TA\bar{F}[m]^Td\rangle$$

$\Rightarrow$ no iteration, $\nabla J_{\rm MVA}$ is *exact* (except for FD error etc.) ! ???

Gradients

MVA: If "gradient artifacts" aren't errors, what are they?

IVA: how does iterative approx. of F^{-1} affect computed gradient accuracy?

Analysis

1st Key observation - "factorization lemma":

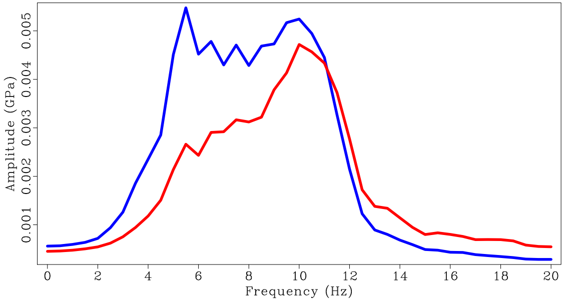
$$D\bar{F}[m]\delta m = \bar{F}[m]Q[m, \delta m]$$

Q is (1) 1st order, (2) linear in δm,
(3) skew-adjoint

(S. IPTA 14 & EAGE 15, ten Kroode IPTA 14)

RICE

Analysis



Avg spectra of $\bar{F}$ (blue) and $D\bar{F}$ (red), Marmousi-derived Born modeling, 2.5-5-10-12.5 filter of delta half-deriv wavelet, norm. inc. plane wave

Analysis

MVA Hessian at consistent data:
$d = \bar{F}[m]\delta\bar{m}_{\rm true}, A\delta\bar{m}_{\rm true} = 0$

$$\delta^2 J_{\rm MVA} = \langle \bar{F}^T\bar{F}\delta\bar{m}_{\rm true}, Q^2A^TA\bar{F}^T\bar{F}\delta\bar{m}_{\rm true}\rangle + \ldots$$

Lead term generically $\neq 0 \Rightarrow$ "true" model is not local min, gradient oscillates (Khoury 06)

RICE

Analysis

IVA Hessian at consistent data: many terms cancel,

$$\delta^2 J_{\mathrm{IVA}} = \|[Q, A]\delta \bar{m}_{\mathrm{true}}\|^2 + \text{ l. o. t.}$$

Positive semi-definite 0-order form, proportional to tomographic Hessian

Analysis

$$\delta J_{\mathrm{IVA}} = -\langle Q^T \delta \bar{m}, A^T A \delta \bar{m} \rangle$$

but

- can only approximate $\delta \bar{m}_{\mathrm{approx}} \approx F[m]^{-1} d$ RMS - no control over derivs! (Q!)
- can only approximate $Q[m, \delta m] \delta \bar{m}$

Analysis

2nd Key observation: (Hou, ten Kroode,...)
computable *asymptotic* inverse $\bar{F}^{\dagger}$

$$\delta_{\rm approx} J = \langle \bar{F}^{\dagger} D \bar{F} \delta \bar{m}_{\rm approx}, A^T A \delta \bar{m}_{\rm approx} \rangle$$

$\bar{F}^{\dagger} D \bar{F}$ also skew + 0-order $\Rightarrow \delta \bar{m}_{\rm approx} \rightarrow \bar{F}^{-1} d$
$\simeq$ error in $\delta \bar{m}_{\rm approx}$

$\bar{F}^{\dagger} D \bar{F} - Q$ is *smoothing* - $\bar{F}^{\dagger} D \bar{F} \rightarrow Q$
$\simeq O$(wavelength).

RICE

Conclusion

- ∇J_{MVA} "artifacts" = feature, not bug
- J_{IVA} locally "as convex as refl. tomography" for noise-free data
- ELSM + asymptotic inverse op $\Rightarrow$ error control for ∇J_{IVA}
- omitted: regularization, full waveform analog, elasticity

Thanks to

- Fons ten Kroode, Jie Hou, Yujin Liu, Yin Huang, Lei Fu, Hervé Chauris, René-Edouard Plessix, Maarten de Hoop, Biondo Biondi, Peng Shen, Uwe Albertin, Chris Stolk
- Sponsors of The Rice Inversion Project
- my very patient audience!

RICE