Fast extended waveform inversion using Morozov's discrepancy principle

Lei Fu

The Rice Inversion Project, Rice University

April 25, 2016

Extended waveform inversion (EWI)

Object function:

$$J_\alpha[m,r] = \frac{1}{2}\|F[m]r - d\|^2 + \frac{\alpha}{2}\|Ar\|^2$$

Role of α: controls amount of penalty applied for model extension.

$\alpha \to 0$,	weak constraint on r,	good data fit
$\alpha \to \infty$,	strong	poor

m - nonlinear model

α - weighting

r - linear extended model

A - annihilator

Variable projection method

$$J_\alpha[m, r] = \frac{1}{2}\|F[m]r - d\|^2 + \frac{\alpha}{2}\|Ar\|^2$$

Classified as separable least-squares inverse problems, solved with variable projection method.
[Golub and Pereyra, 1973, Golub and Pereyra, 2003]

Nested optimization:

- inner loop, given m, find r optimizes $J_\alpha[m, r]$.
- outer loop, find m optimizes reduced objective function $J_\alpha[m, r[m]]$.

Inner loop, optimize J over r

$\nabla_r J_\alpha[m, r] = 0 \Rightarrow$ normal equation:

$$(F^T F + \alpha A^T A) r = F^T d$$

Solved by linear iterative method, e.g., conjugate gradient (CG).

Solution depends on α

$$r_\alpha \approx (F^T F + \alpha A^T A)^{-1} F^T d$$

Outer loop, update m

Gradient of reduced objective function $J_\alpha[m, r_\alpha[m]]$ respect to m:

$$\nabla_m J_\alpha[m, r[m]] = DF^T\,(r_\alpha, F[m]r_\alpha - d)$$

DF^T tomographic or WEMVA operator.

DF, bilinear operator, linear in dm and r.

Weight α

Rewrite data misfit and penalty terms

$$\begin{aligned} e(\alpha) &= \frac{1}{2}\|Fr(\alpha) - d\|^2 \\ p(\alpha) &= \frac{1}{2}\|Ar(\alpha)\|^2 \end{aligned}$$

Differentiate e and p

$$2p \geq \frac{de}{d\alpha} \geq 0$$

$$\frac{dp}{d\alpha} \leq 0$$

Weight α

Basic bound:

$$\frac{de}{d\alpha} \leq 2p$$

Suppose current weight α_c, updated weight α_+

$$\alpha_+ \gtrapprox \alpha_c + \frac{e(\alpha_+) - e(\alpha_c)}{2p(\alpha_c)}$$

Morozov's discrepancy principle

Target data misfit X, $X_- < X < X_+$.

Adjust α *to keep* $e(\alpha)$ *between* $\frac{1}{2}X_-^2$ *and* $\frac{1}{2}X_+^2$*, i.e. "near"* $\frac{1}{2}X^2$*.*

Application

Suppose $e(\alpha_c) < \frac{1}{2}X_-^2$, e.g. m is updated, adjust α to keep

$e(\alpha) \in [\frac{1}{2}X_-^2, \frac{1}{2}X_+^2]$

$$\alpha_+^{\text{est}} = \alpha_c + \frac{\frac{1}{2}X_+^2 - e(\alpha_c)}{2p(\alpha_c)}$$

Algorithm to update α

1. Initial m, calculate $e(\alpha = 0)$ to estimate $X_- = 0.8 * e(0)$ and $X_+ = 1.2 * e(0)$

   $e(0) \neq 0$, determined by accuracy of solving normal equation

2. Update α

$$\alpha_+^{\text{est}} = \alpha_c + \frac{\frac{1}{2}X_+^2 - e(\alpha_c)}{2p(\alpha_c)}$$

3. If $e(\alpha_+^{\text{est}})$ is not in an acceptable range,

   $e(\alpha_+) < X_-, \alpha_+ * = 2;$

   $e(\alpha_+) > X+, \alpha_+ / = 1.5;$

Overthrust model

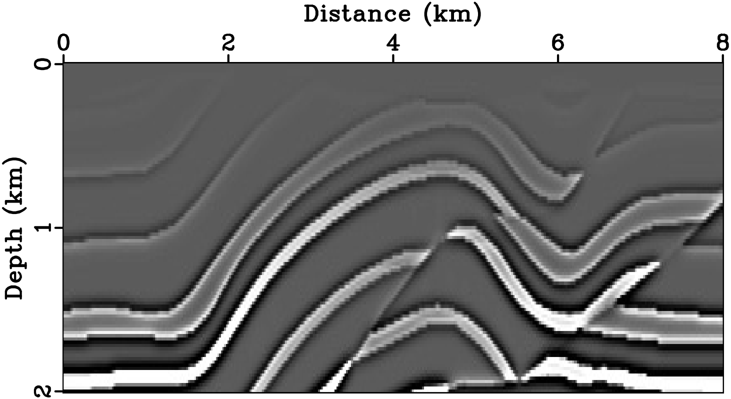
2D constant density acoustic 2-8 order finite difference code

Subsurface offset extended Born approximation.

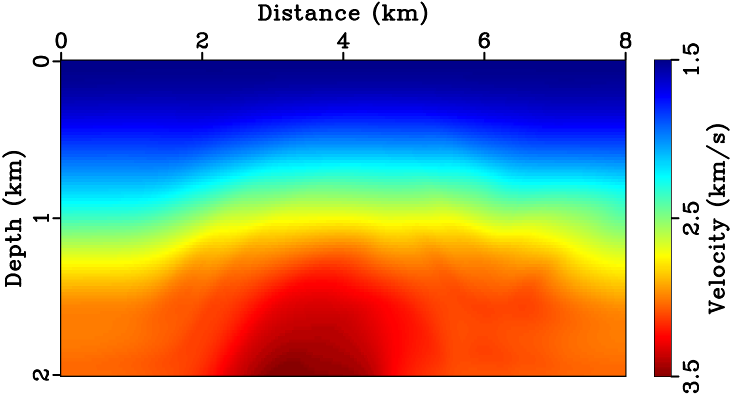
Adaptive multiscale method (2 refinement stages)

Parameter	Measurements
Source wavelet	bandpass $5 - 20\,Hz$
Initial velocity	$v = 1.5\,km/s$
Max iter inner loop	20
Source position $\mathbf{x_s}$	$x : 1 - 7\,km$ every $40\,m$, $z = 20\,m$
Receiver position $\mathbf{x_r}$	$x : 0 - 8\,km$ every $40\,m$, $z = 0\,m$
Space and time	$x = 8\,km$, $z = 2\,km$, $t = 3\,s$
Grid size	$dx = dh = dz = 20\,m$, $dt = 2\,ms$

r at $h = 0\,m$



True background velocity



Initial background velocity v_0

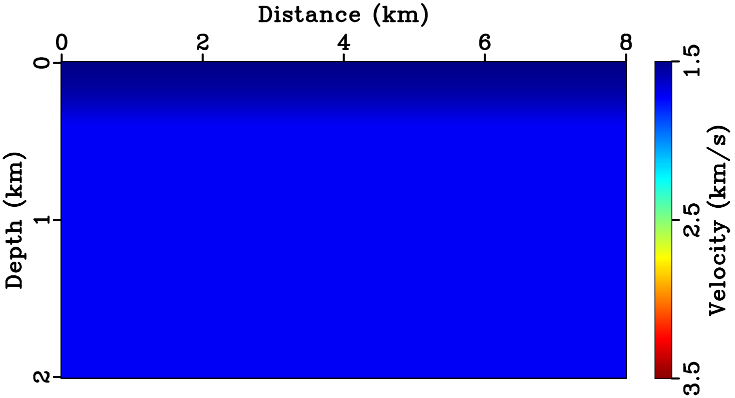
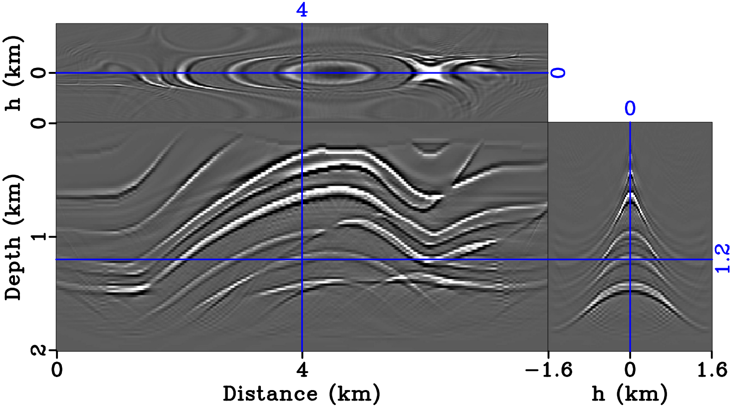


Image r_0 with initial velocity v_0



Inverted background velocity v_{20}

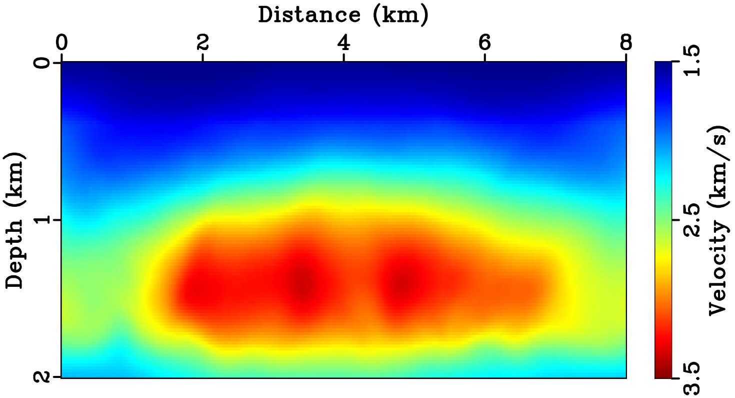
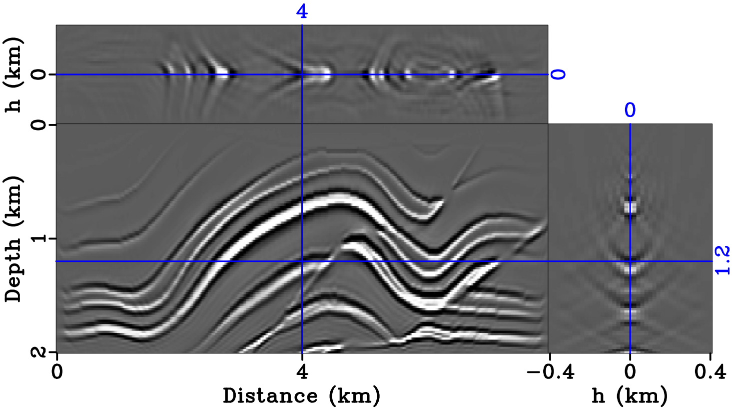
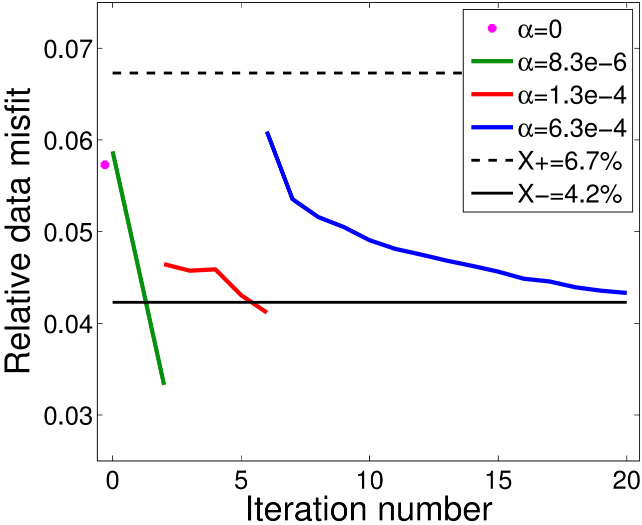


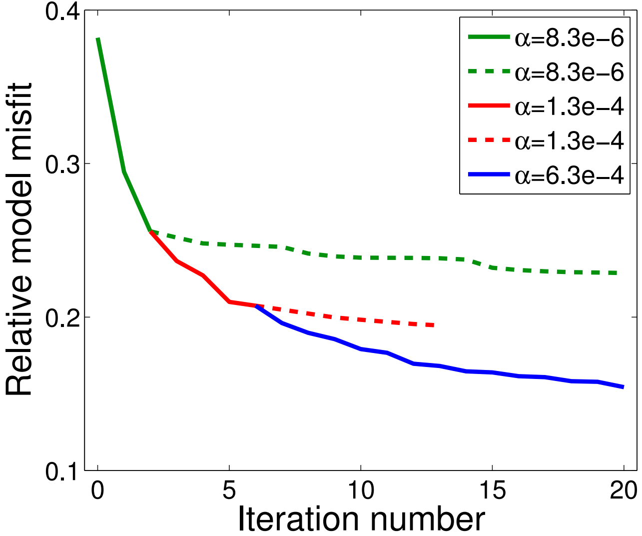
Image r_{20}



Relative data misfit



Relative model misfit



Overthrust model, scan test

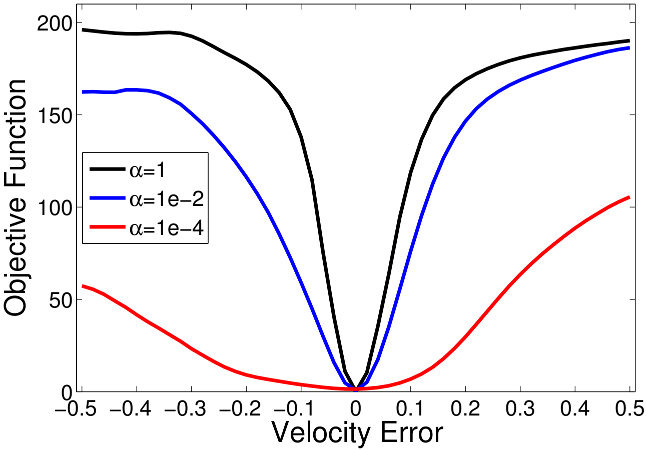


Figure : Scan test: objective function with different values of α. Background velocity error from -50% to $+50\%$

Summary

1. Update rule for α, significantly improves convergence rate.

$$\alpha_{+}^{\text{est}} = \alpha_c + \frac{\frac{1}{2}X_{+}^{2} - e(\alpha_c)}{2p(\alpha_c)}$$

2. Morozov discrepancy principle: adjust α to keep

$$e(\alpha) \in [\frac{1}{2}X_{-}^{2}, \frac{1}{2}X_{+}^{2}]$$

Future work

Consider physics errors, explore with field data.

Use variable density acoustics.

Accelerate inner loop and outer loop (preconditioner, optimization methods)

Acknowledgments

The Rice Inversion Project (TRIP) sponsors

Rice University Research Computing Support Group (RCSG)

Texas Advanced Computing Center (TACC)

Madagascar, Seismic Unix

TRIP members

Golub, G., and V. Pereyra, 1973, The differentiation of pseudoinverses and nonlinear least squares problems whose variables separate: SIAM Journal on Numerical Analysis, **10**, 413–432.

———, 2003, Separable nonlinear least squares: the variable projection method and its applications: Inverse Problems, **19**, R1–R26.