Approximate Multi-Parameter Linear Inversion By Phase-Space Scaling and RTM

Rami Nammour

The Rice Inversion Project

December 9, 2010 / TRIP Review Meeting

Outline

- Linearization of the inverse problem
- Properties of the normal operator
- Proposed method
- Preliminary results:
  - Constant density acoustics
  - Layered variable density acoustics (2-parameters)
- Work in progress, future work

Let:

- $m(x)$: model (consists of p-parameters: impedance, velocity, density,...)
- $p(x,t)$: state (the solution of the system: pressure,...)

Then, if S is the Forward Map:

- The Forward Problem:

$$S[m] = p|_{surface}$$

- The Inverse Problem:

$$S[m] \approx S^{obs}$$

Given S^{obs}, get $m(x)$

Nonlinear and Large Scale !

Linearization

Solution depends nonlinearly on coefficients; if we have an approximation m_0 to the model, Linearization is advantageous:

- Write $m = m_0 + \delta m$
  m_0: Given reference model
  δm: First order perturbation about m_0

- Define Linearized Forward Map $F[m_0]$ (Born Modeling):

$$F[m_0]\delta m = \delta p$$

- Linear inverse problem:

$$F[m_0]\delta m \approx S^{obs} - S[m_0] := d$$

Normal Equations

Interpret as least squares problem: need to solve normal equations

$$N[m_0]\delta m := F^*[m_0]F[m_0]\delta m = F^*[m_0]d$$

$N := F^*[m_0]F[m_0]$: Normal Operator (Modeling + Migration),
$b := F^*d$: migrated image

- *Large Scale*: millions of equations/unknowns, also $\delta m \to N\,\delta m$ expensive
- Cannot use Gaussian elimination $\Rightarrow$ need rapidly convergent iteration $\Rightarrow$ good preconditioner
- Not narrow band (like Laplace in 2D/3D) $\Rightarrow$ matrix preconditioners ineffective

RICE

ΨDOs and their symbols

- Pseudodifferential (ΨDO) operators defined by symbols $a(x, \xi)$

$$Op(a)u(x) = \int\int a(x, \xi)u(y)e^{-i[(x-y).\xi]}\, d\xi\, dy,$$

- $a(x, \xi)$: Scalar function of position x and wavenumber ξ

$$|a(x, \xi)| = \mathcal{O}(|\xi|^m), \quad \text{as } |\xi| \to \infty;$$

$m = ord(a) := ord(Op(a))$

- Calculus of scalar symbols:
  1. $Op(\alpha_1 a_1 + \alpha_2 a_2) = \alpha_1 Op(a_1) + \alpha_2 Op(a_2), \quad \alpha_1, \alpha_2$ scalars
  2. $Op(a_1 a_2) \simeq Op(a_1)Op(a_2) \simeq Op(a_2)Op(a_1)$
  3. $ord(a_1 a_2) \leq ord(a_1) + ord(a_2)$

  ($\simeq$: difference is lower order ΨDO)

Properties of Normal Operator

- Normal operator is a matrix of pseudodifferential operators:
  - Smooth background model m_0 (Beylkin 1985)
  - Scalar wave fields
  - Polarized vector fields (P-P, P-S, S-S). (Beylkin and Burridge, 1989; De Hoop, 2003)
- $N = Op(A)$, $A = p \times p$ matrix of scalar symbols

$$Op(A)u(x) = \int \int A(x, \xi) u(y) e^{-i[(x-y).\xi]} \, d\xi \, dy,$$

Cramer's Rule for ΨDOs

- A is a matrix of symbols
- Adjugate of A defined by:

$$Adj(A)\,A = A\,Adj(A) = det(A)\,I, \tag{1}$$

- Symbol calculus $\Rightarrow$

$$Adj(N)\,N \approx N\,Adj(N) \approx det(N)\,I. \tag{2}$$

- Notation: $Adj(N) = Op(Adj(A))$, $det(N) = Op(det(A))$.

Cramer's Rule for ΨDOs

- Want to solve: $N\delta m = b$
- Only given ability to apply N
- Apply Adjugate, in terms of N (later):

$$Adj(N)\, b = Adj(N)\, N\, \delta m \approx det(N)\, \delta m \tag{3}$$

- Have $det(N)\, \delta m$, want δm
- $det(N)\, \delta m \approx$ amplitude scaling of δm

RICE

Dividing by the determinant

- To undo $det(N)$, apply N to form:

$$N\, det(N)\, \delta m \approx det(N)\, N\, \delta m = det(N)\, b \tag{4}$$

- given b and $det(N)\, b$, approximate scaling factor c:

$$c = \underset{c \in \Psi DO}{\textit{argmin}} \| b - c\, det(N)\, b \|^2 \tag{5}$$

- Approximate solution:

$$\begin{aligned} \delta m = N^{-1} b \approx N^{-1}\, c\, det(N)\, b \approx c\, det(N)\, N^{-1} b \\ \approx c\, det(N)\, \delta m := \delta m_{inv} \end{aligned} \tag{6}$$

Approximation of ΨDO

- How to represent c?
- The action of the ΨDO in 2D (Bao and Symes, 1996)

$$Op(a)\,u(x,z) \approx \int\int a(x,z,\xi,\eta)\hat{u}(\xi,\eta)e^{i(x\xi+z\eta)}\,d\xi\,d\eta, \quad \hat{u} = \mathcal{F}[u]$$

- Direct Algorithm $O(N^4\log(N))$ complexity ($N = \mathcal{O}(10^3)$)!
- Finite Fourier series of length K:

$$a(x,z,\xi,\eta) \approx \sum_{l=-K/2}^{l=K/2} \hat{a}_l(x,z)e^{il\theta}, \quad \theta = \arctan\left(\frac{\eta}{\xi}\right)$$

  - Use FFT $\Rightarrow O(KN^2[\log(N)+\log(K)])$
  - K independent of N, depends on smoothness of a
  - θ captures dip-dependence

RICE

Recap

To solve $N\delta m = b$,

- Apply $Adj(N)$ on b : $Adj(N)\,b \approx det(N)\,\delta m$
- Apply N : $N\,det(N)\,\delta m \approx det(N)\,b$
- Represent scaling factor : $c = Q_m[q]$
- Compute c:

$$c = \underset{c \in \Psi DO}{\text{argmin}} \|b - c\,det(N)\,b\|^2.$$

- Approximate the inverse: $\delta m_{inv} := c\,det(N)\,\delta m \approx \delta m$

Pseudodifferential scaling method that resolves multiple dip events

RICE

One-Parameter Case: $p = 1$

- In this case: $det(N) \equiv N$, $Adj(N) \equiv I$
- Only need c:

$$c = \underset{c \in \Psi DO}{\textit{argmin}} \|b - c\,N\,b\|^2.$$

- Method reduces to usual scaling methods (Masters thesis)

Scaling Methods

- Hessian $\approx$ multiplication by a smooth function (Claerbout and Nichols, 1994; Rickett, 2003)
- Near Diagonal Approximation of Hessian (Guitton, 2004)
- Special case (well defined dip): normal operator $\approx$ multiplication by smooth function after composition with power of Laplacian (correction to Claerbout-Nichols - Symes, 2008)
- Herrmann et al. (2007) derive a scaling method using *curvelets* to approximate eigenvectors

RICE

Amplitude Versus Offset (AVO)

- AVO variation contains info about anomaly in physical parameters
- Uses simplification of nonlinear Zoeppritz equations (Aki Richards; Shuey, 85)
- Rutherford and Williams (1989): 3 classes of AVO behaviors
- Lortzer and Berkhout (1989): Statistical Bayesian approach to approximate anomaly

RICE

Linearized Multi-Parameter Inversion

- Santosa and Symes (1988): layered acoustic fluid, with conditioning study
- Bourgeois et al. (1989): impedance and velocity in variable density acoustics
- Virieux et al. (1992): P and S impedances in linear elasticity (geometric optics computations)
- Foss et al. (2005): anisotropic elasticity, asymptotic inverse
- Minkoff (1995): linear inversion successful if you take care of everything: source estimation, attenuation . . .
- Charara et al. (1996): P and S velocities and density in the linear elasticity

Proposed method:

- Not iterative
- Uses wave equation migration (Reverse Time Migration)
- No geometric optics computations
- Relies only on application of normal operator
- Novel for $p > 1$: Few applications of $N \rightarrow$ approximate inverse

Constant Density Acoustics: $p = 1$

On Marmousi 2D data:

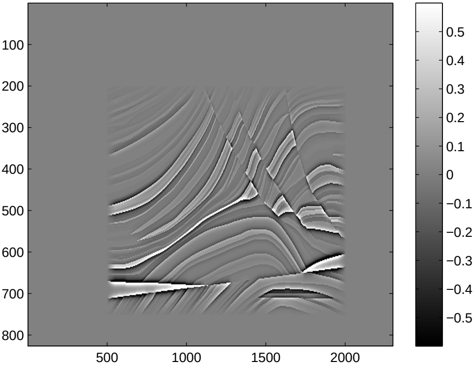

Figure: δm_{true}

Migration Vs Inversion

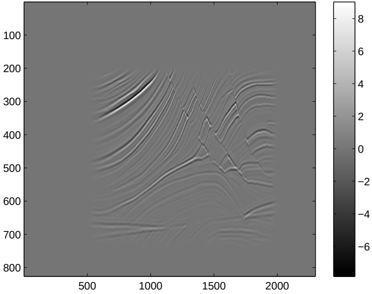

Figure: $\delta m_{mig} = F^* d$

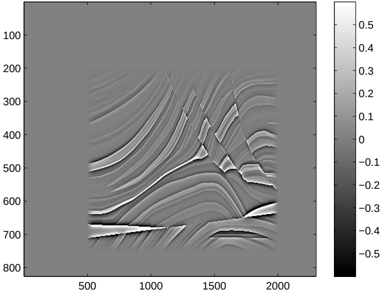

Figure: $\delta m_{true} = (F^* F)^{\dagger} F^* d$

Lessons Learned

- Discontinuities preserved
- Amplitudes distorted
- Similar relationship between $det(N)\,\delta m$ and δm
- Multi-parameter problem reduced to one parameter problem
- Solution: amplitude correction

Migration - Remigration

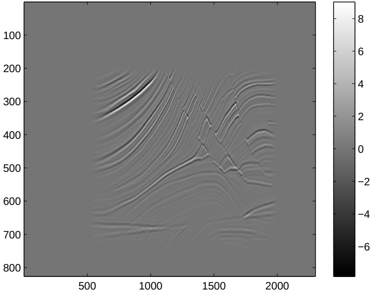

Figure: $\delta m_{mig} = F^* d$

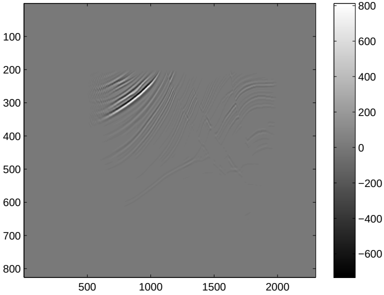

Figure: $\delta m_{remig} = F^* F \delta m_{mig}$

RICE

Scaling $K = 1$

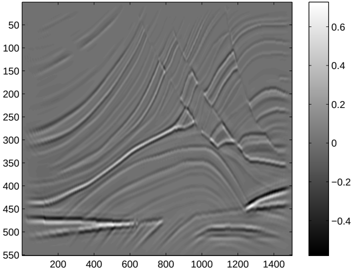

Figure: δm_{inv} with $K = 1$

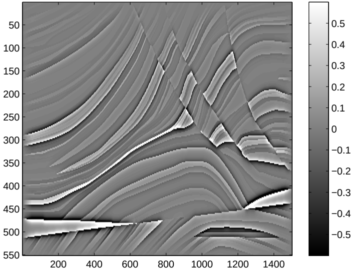

Figure: δm_{true}

Scaling $K = 5$

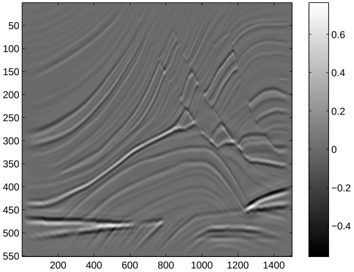

Figure: δm_{inv} with $K = 5$

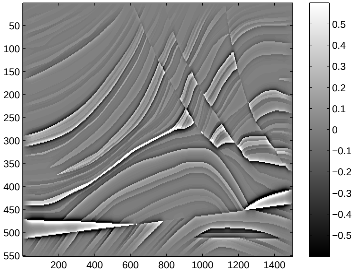

Figure: δm_{true}

RICE

Difference between $K = 1$ and $K = 5$

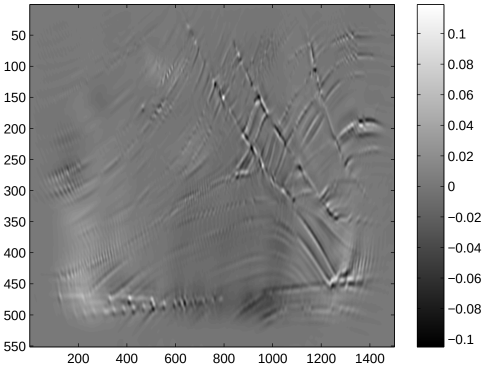

Figure: Difference between $K = 1$ and $K = 5$

Two-Parameter Case: $p = 2$

In this case:

$$N = \begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix}.$$

Adjugate given by:

$$Adj(N) = \begin{pmatrix} N_{22} & -N_{12} \\ -N_{12} & N_{11} \end{pmatrix}.$$

$$Adj(N)\, b = J^T N J\, b \tag{7}$$

Where,

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Explicitly,

$$N \begin{pmatrix} -b_2 \\ b_1 \end{pmatrix} = (N_{11}N_{22} - N_{12}^2) \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \tag{8}$$

- Normal operator on specific combination of migrated images $\rightarrow$ amplitude scaling of inverse
- Advantage: Only apply N to permutations of entries of b.
- Cost of method: 1+1=2 applications of N

Application: Layered Variable Density Acoustics

- Two parameter inverse problem: density and velocity
- Homogeneous background fields
- True model:

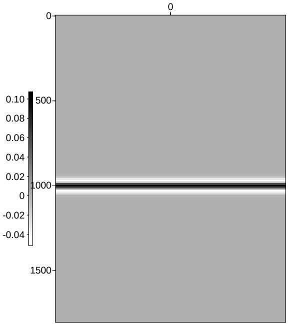

Figure: *vp*

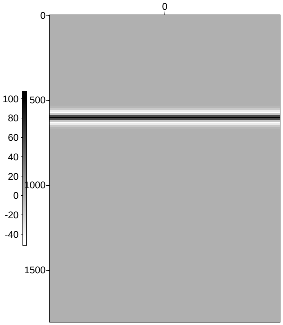

Figure: *dn*

Migration: Mix

Migration mixes effects of reflectors

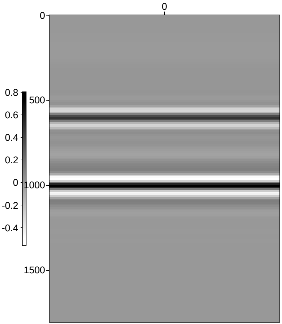

Figure: b_1

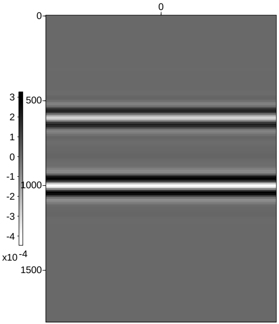

Figure: b_2

Adjugate: Un-Mix

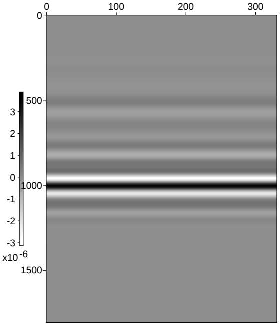

Figure: $(J^T NJ\, b)_1 \approx det(N)\, x_1$

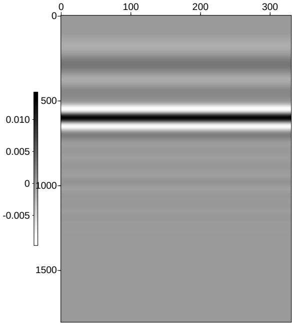

Figure: $(J^T NJ\, b)_2 \approx det(N)\, x_2$

RICE

Apply N again

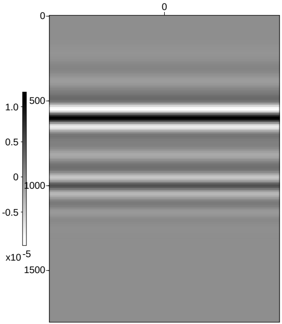

Figure: $det(N)\, b_1$

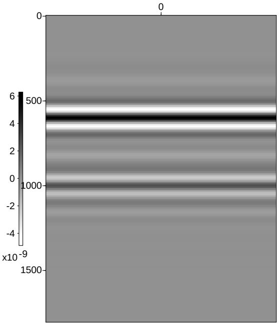

Figure: $det(N)\, b_2$

Approximate Inverse

Compute c and approximate the inverse:

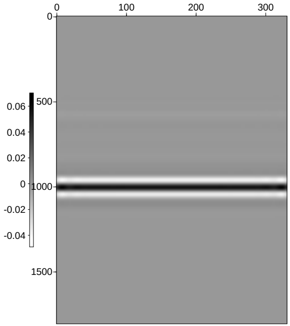

Figure: inv_{vp}

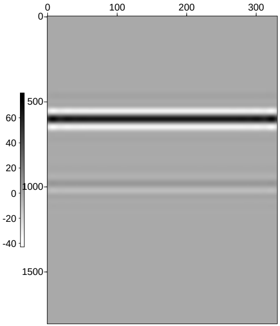

Figure: inv_{dn}

Work In Progress, Future Work

- In progress:
  - Apply to complex models: variable density acoustics Marmousi
  - Extend to 3D
- Future work
  - Precondition the nonlinear problem: Full Waveform Inversion (FWI)
  - Three-parameter case: Linear elasticity

Extension to 3D

- Only need to extend PsiDO algorithm
- Truncated spherical harmonics expansion of $\tilde{q}$:

$$\tilde{q}(x,y,z,\theta,\phi) = \sum_{l=0}^{K} \sum_{n=-l}^{l} c_{ln}(x,y,z) Y_l^n(\theta,\phi), \quad (9)$$

- Y_l^n spherical harmonics, expressed in terms of associated Legendre Polynomials
- $$Y_l^n(\theta,\phi) = \sqrt{\frac{(2l+1)(l-n)!}{4\pi(l+n)!}} P_l^n(cos(\theta)) e^{in\phi}. \quad (10)$$
- Express $Y_l^n(\theta,\phi) = Y_l^n(\xi,\zeta,\eta)$, using spherical coordinates transformation

Extension to 3D

- Plug into the action of a PsiDO

$$Q_m u(x,y,z) = \sum_{l=0}^{K} \sum_{n=-l}^{l} c_l(x,y,z)\mathcal{F}^{-1}\left\{\omega^m Y_l^n(\xi,\zeta,\eta)\hat{u}(\xi,\zeta,\eta)\right\} \tag{11}$$

- Cost: Use FFT $\Rightarrow$ $(K+1)^2 N^3 \log(N)$

Dip Filtering With PsiDOs

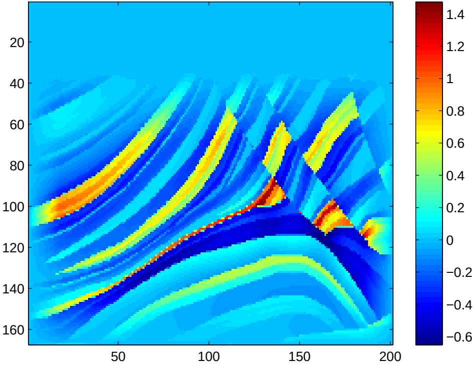

Figure: Marmousi model

Dip Filtering With PsiDOs: Pick Your Symbol

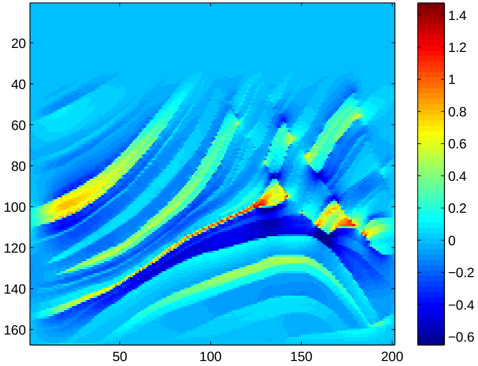

Figure: Vertical dips filtered out of Marmousi

Summary

- One-parameter case: Pseudodifferential Scaling
  - Fast and reliable solution if m_0 is a good reference model
  - Preconditioning iterative methods when m_0 is not a good reference model
  - Extension to 3D
- Multi-parameter case: Cramer's Rule
  - Reduced to one-parameter case
  - Applied to layered variable density acoustics ($p = 2$)
  - Testing on more complex models
  - Formulated for $p = 3$

THANK YOU !

- Dr William Symes
- Dr Eric Dussaud, Dr Fuchun Gao, Dr Uwe Albertin

Full Waveform Inversion

Back to the nonlinear forward model:

$$S[m] = d.$$

Least squares objective function:

$$J = \frac{1}{2}\|S[m] - d\|^2.$$

Gradient (migrated image):

$$g = F^*(S[m] - d).$$

Hessian:

$$H = F^*F + \frac{\partial F^*}{\partial m}(S[m] - d) \approx F^*F.$$

RICE

- Gauss Newton step:

$$m_{k+1} = m_k - H^\dagger g.$$

- $H^\dagger$ is too expensive, need approximation.
- Split:

$$m_k = m_{k0} + \delta m_k$$

- Compute:

$$c = \underset{c \in \Psi DO}{\text{argmin}} \|\delta m_k - c\, F^*[m_{k0}] F[m_{k0}] \delta m_k\|^2$$

- Approximate: $H^\dagger \approx c$
- The scaling factor c preconditions FWI:

$$m_{k+1} = m_k - c\, g.$$

- Similar work: Herrmann et al. (2008), Jang et al. (2009) (different approximations of $H^\dagger$)

Three-Parameter Case

$$N = \begin{pmatrix} N_{11} & N_{12} & N_{13} \\ N_{12} & N_{22} & N_{23} \\ N_{13} & N_{23} & N_{33} \end{pmatrix}.$$

Its adjugate is,

$$Adj(N) = \begin{pmatrix} (N_{22}N_{33} - N_{23}^2) & -(N_{12}N_{33} - N_{13}N_{23}) & (N_{12}N_{23} - N_{13}N_{22}) \\ -(N_{12}N_{33} - N_{23}N_{13}) & (N_{11}N_{33} - N_{13}^2) & -(N_{11}N_{23} - N_{13}N_{12}) \\ (N_{12}N_{23} - N_{22}N_{13}) & -(N_{11}N_{23} - N_{13}N_{12}) & (N_{11}N_{22} - N_{12}^2) \end{pmatrix}.$$

Applying the Adjugate

1. Form $N(e_2^T e_1 - e_1^T e_2)b$, $N(e_1^T e_3 - e_3^T e_1)b$ and $N(e_3^T e_2 - e_2^T e_3)b$
2. Form $(e_2^T e_1 - e_1^T e_2)N[e_3^T e_3 N(e_2^T e_1 - e_1^T e_2) + e_2^T e_3 N(e_1^T e_3 - e_3^T e_1) + e_1^T e_3 N(e_3^T e_2 - e_2^T e_3)]b$
3. Form $-(e_3^T e_1)N[e_3^T e_2 N(e_2^T e_1 - e_1^T e_2) + e_2^T e_2 N(e_1^T e_3 - e_3^T e_1) + e_1^T e_2 N(e_3^T e_2 - e_2^T e_3)]b$
4. Sum the last two images to obtain $Adj(N)\, b \approx det(N)\, x$

- Cost: 5+1 = 6 applications of N
- Reduce the cost? (special cases: linear elasticity)