The IWave++ Inversion Framework

William Symes
Dong Sun

The Rice Inversion Project

Annual Review Meeting
Dec 9, 2010
**Concept of IWAVE++**

- **Goal:** make IWAVE capabilities - large-scale, parallel, extensible FD modeling - available for research on inversion/imaging driven by time-domain simulation

- **Method:** embed IWAVE in Rice Vector Library (RVL) Operator type, use RVL utilities (linear algebra, construction of least squares functions, etc.) and algorithms (CG, NLCG, LBFGS,...) to construct inversion applications.
Agenda

1. From Modeling to Inversion
2. Numerical Experiments & Discussion
3. Summary and Future Work
Simulation-driven Inversion

Waveform Inversion:

$$\min_{m \in \mathcal{M}} J[m] := \frac{1}{2} \| \mathcal{F}[m] - d \|^2_D$$

Gradient $\nabla J[m] = D\mathcal{F}[m]^T (\mathcal{F}[m] - d)$

- $\mathcal{M}$: Model Space, $\mathcal{D}$: Data Space
- $\mathcal{F}: \mathcal{M} \rightarrow \mathcal{D}$ modeling operator
  $(\mathcal{F} = Su, u$ wavefield, $S$ sampling operator)
- $D\mathcal{F}[m]$: derivative map of $\mathcal{F}$ at $m$ (Born operator)
- $D\mathcal{F}[m]^T$: adjoint operator of $D\mathcal{F}[m]$
Abstraction Discrepancy & Solution

Inversion involves two levels of abstraction:

- applications of $\mathcal{F}, D\mathcal{F}[m], D\mathcal{F}[m]^T$
  
  involve solution of PDE’s, depending on specific physics and numerical realization, ...

- optimization algorithm
  
  only involve generic linear algebra operations, independent of specific representation or modeling construction
Abstraction Discrepancy & Solution

Object-oriented programming offers a solution to share information between vector calculus & modeling software.

Both sides provide interfaces to reference the same common data structures for model grids, data traces, ..., which need not be part of either.

In our case, the external data objects are structured disk files (RSF file structure for model parameter, SEGY for seismic traces):

- Vector calculus (RVL) operations perform as disk-to-disk filters.
- IWAVE has its own i/o functions and is regulated by parameter table.
Abstraction Discrepancy & Solution

A solution (WWS, DS & ME,10; DS & WWS, TR10-05, TR10-06):

- wrap $\mathcal{F}$ as a C++ class (RVL operator) for vector valued functions, with methods to
  - do modeling $\mathcal{F}[m]$
  - apply first derivative (action of Born Map $D\mathcal{F}[m]$)
  - apply adjoint derivative (adjoint action of Born Map $D\mathcal{F}[m]^T$)

- auxiliary classes:
  - State classes for $(m,u), (m,u,\delta m, \delta u)$ with necessary methods
  - Sampler classes to link $u, \delta u$ with data $d, \delta d$
  - Stack class to manage dynamic wavefields at specific time levels
  - Algorithm classes to update fields, read/write $m, \delta m$, initialize $u, \delta u$

...
Forward, Born and Adjoint Simulation

Model problem:
\[
\frac{1}{\kappa(x)} \frac{\partial p}{\partial t} + \nabla \cdot v = f(x, t) \\
\frac{1}{b(x)} \frac{\partial v}{\partial t} + \nabla p = 0
\]

\[
d = F[m] := Su
\]
\[
u = (p, v), \ S \text{ sampling operator (} Sp := p(x_r, t))
\]

Simulation

**SIM (1)** \( u = 0 \)

**SIM (2)** For \( n = 0, \ldots, N - 1 \) do:

**SIM (2.1)** do: \( p += -\Delta t \kappa \nabla \cdot v := L_0[m, u] \)
\[
v += -\Delta t b \nabla p := L_1[m, u]
\]

**SIM (2.2)** \( u += Rf \)

**SIM (2.3)** \( d += Su \)
Forward, Born and Adjoint Simulation

**Born Simulation**

**BORN SIM (1)** \( u = 0, \, \delta u = 0; \)

**BORN SIM (2)** For \( n = 0, \ldots, N - 1 \) do:

**BORN SIM (2.1)** do:

\[
\delta p^+ = -\Delta t \kappa \nabla \cdot \delta v := L_0[m, \delta u] \\
\delta p^+ = -\Delta t \delta \kappa \nabla \cdot v := L_0[\delta m, u] \\
p^+ = -\Delta t \kappa \nabla \cdot v := L_0[m, u] \\
p^+ = -\Delta t \kappa \nabla \cdot v := L_0[m, u] \\
\delta v^+ = -\Delta t b \nabla \delta p := L_1[m, \delta u] \\
\delta v^+ = -\Delta t \delta b \nabla p := L_1[\delta m, u] \\
v^+ = -\Delta t b \nabla p := L_1[m, u] \\
v^+ = -\Delta t b \nabla p := L_1[m, u]
\]

**BORN SIM (2.2)** \( u^+ = Rf \)

**BORN SIM (2.3)** \( \delta d^+ = S \delta u \)
Forward, Born and Adjoint Simulation

**Adjoint Simulation**

**ADJ SIM (1)** \( \lambda u = (\lambda p, \lambda v) = 0 \)

**ADJ SIM (2)** For \( n = N, \ldots, 1 \) do:

1. Set time in \( u, d \) to \( n - 1 \);
2. **ADJ SIM (2.1)** \( \lambda u \leftarrow S^T \lambda d; \)
3. **ADJ SIM (2.2)** do:

   \[
   \begin{align*}
   &\lambda b \leftarrow -\Delta t \lambda v \nabla p := L_1[\lambda u, u] \\
   &\lambda v \leftarrow \Delta t b \nabla \lambda p := L_1[m, u] \\
   &\lambda \kappa \leftarrow -\Delta t \lambda p \nabla \cdot v := L_0[\lambda u, u] \\
   &\lambda p \leftarrow \Delta t \kappa \nabla \cdot \lambda v := L_0[m, \lambda u]
   \end{align*}
   \]
Agenda

1. From Modeling to Inversion

2. Numerical Experiments & Discussion

3. Summary and Future Work
Numerical Verifications with the “Camembert”

Principle example in O. Gauthier, J. Virieux and A. Tarantola (1986): 
*Two-dimensional Nonlinear Inversion of Seismic Waveform: Numerical Results* 
(first published exploration of iterative FWI with multi-D data and multi-D models)

- 1 km × 1 km domain
- background: $v = 2.5$ km/s, $\rho = 4$ g/cm$^3$
- peak frequency $\sim 50$ Hz
- circular bulk modulus anomaly, diam = 0.5 km, “small” = 2% or “large” = 20%
- “nonlinear saturation” at 10% - full wavelength traveltime perturbation
- 5 m grid - 10 gridpts / (peak) wavelength
- absorbing BCs on all sides

Fig. 5. The model is now a circular inclusion (the “Camembert”) in a homogeneous medium. The size of the Camembert is about 10 wavelengths. The model is numerically defined in a grid with 200 × 200 points, so the model contains $10^4$ parameters (unknowns for the inversion).
Numerical Verifications with the “Camembert”

Numerical Verification of $D\mathcal{F}[m]$ with reflection configuration:

- 100 receivers (fixed spread), $z_r = 80$, $x_r = 0, 10, \ldots, 990$ m
- 8 sources, $z_s = 40$, $x_s = 110, 220, \ldots, 880$ m

\[
\frac{\mathcal{F}(m+h\delta m)-\mathcal{F}(m-h\delta m)}{2h} - D\mathcal{F}[m]\delta m \parallel / \parallel D\mathcal{F}[m]\delta m \parallel
\]

<table>
<thead>
<tr>
<th>h</th>
<th>relative error</th>
<th>conv rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.275</td>
<td>—</td>
</tr>
<tr>
<td>0.9</td>
<td>0.233</td>
<td>1.547</td>
</tr>
<tr>
<td>0.8</td>
<td>0.193</td>
<td>1.631</td>
</tr>
<tr>
<td>0.7</td>
<td>0.153</td>
<td>1.708</td>
</tr>
<tr>
<td>0.6</td>
<td>0.117</td>
<td>1.779</td>
</tr>
<tr>
<td>0.5</td>
<td>0.083</td>
<td>1.839</td>
</tr>
<tr>
<td>0.4</td>
<td>0.055</td>
<td>1.891</td>
</tr>
<tr>
<td>0.3</td>
<td>0.031</td>
<td>1.934</td>
</tr>
<tr>
<td>0.2</td>
<td>0.014</td>
<td>1.965</td>
</tr>
<tr>
<td>0.1</td>
<td>0.004</td>
<td>1.974</td>
</tr>
</tbody>
</table>

Small anomaly (2%)
Numerical Verifications with the “Camembert”

Numerical Verification of $D\mathcal{F}[m]^T$ with reflection configuration

Let $\mathcal{J} := D\mathcal{F}[m]$, adjoint relation holds if

$$\text{AdjErr} := \frac{|(\mathcal{J}\delta m, \delta p) - (\delta m, \mathcal{J}^T\delta p)|}{\|\mathcal{J}\delta m\|\|\delta p\|} < 100*\text{macheps}$$

$$
\begin{array}{|c|}
\hline
(\mathcal{J}\delta m, \delta p) & 3.89868e+06 \\
(\delta m, \mathcal{J}^T\delta p) & 3.89868e+06 \\
\|\mathcal{J}\delta m\|\|\delta p\| & 3.89867525e+06 \\
\text{AdjErr} & 5.32232070e-06 \\
100*\text{macheps} & 1.19209290e-05 \\
\hline
\end{array}
$$
**Figure:** Born modeling acoustic synthetic shot gather ($DF[m]\delta m$) resulting from homogeneous background $m$ and random $\delta m$.

**Figure:** Migrated bulk modulus (component of $DF[m]^T\delta d$) with random $\delta d$. 
Dot-product Test with Random Vector

<table>
<thead>
<tr>
<th>Expression</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle DF[m]\delta m, \delta d \rangle_D$</td>
<td>$-2.91666225e+06$</td>
</tr>
<tr>
<td>$\langle \delta m, DF[m]^T \delta d \rangle_D$</td>
<td>$-2.94833250e+06$</td>
</tr>
<tr>
<td>$| DF[m]\delta m |_M | \delta d |_D$</td>
<td>$3.05501747e+09$</td>
</tr>
<tr>
<td>$\langle</td>
<td>DF[m]\delta m, \delta d \rangle_D - \langle \delta m, DF[m]^T \delta d \rangle_M</td>
</tr>
<tr>
<td>$100 \times \text{macheps}$</td>
<td>$1.19209290e-05$</td>
</tr>
</tbody>
</table>

**Table:** Standard RVL test for accuracy of adjoint operator pair: adequate quality if model space and data space inner products differ by less than a modest multiple of machine precision, relative to data-space norms of input and output data perturbations.
Various Inversions with the “Camembert”

inversion - reflection - band-limited data with central frequency \(~ 50 \text{ Hz}~

small anomaly (2%)

Initial MS resid = 3629; Final after 5 LBFGS steps = 254

Bulk modulus: Left, true model; Right, inverted model
Various Inversions with the “Camembert”

transmission configuration:
- 400 receivers - each side like top in reflection
- 8 sources at corners and side midpoints

small anomaly (2%)- band-limited data

Initial MS resid = $2.56 \times 10^7$; Final after 5 LBFGS steps = $2.6 \times 10^5$

Bulk modulus: Left, true model; Right, inverted model
Various Inversions with the “Camembert”

Large anomaly (20%) - transmission – band-limited data
Initial MS resid = $2.14 \times 10^8$; Final after 5 LBFGS steps = $1.44 \times 10^8$

Bulk modulus: Left, true model; Right, inverted model
Various Inversions with the “Camembert”

Impulsive Inversion:

**Large anomaly (20%)- reflection - impulsive data, 0-60 Hz**

Initial MS resid = $3.45 \times 10^{13}$; Final after 5 LBFGS steps = $4.78 \times 10^{11}$

Bulk modulus: Left, true model; Right, inverted model
Agenda

1. From Modeling to Inversion
2. Numerical Experiments & Discussion
3. Summary and Future Work
**Summary and Future Work**

**IWA VE++ Inversion Framework:**
- demonstrates a design principle for straightforward incorporation of sophisticated modeling techniques in inversion software
- provides a platform for various migration and inversion research and applications
- separates simulation from optimization, which facilitate other research projects
Summary and Future Work

Future Work

- complement the implementation of IWAVE++ for extended inversion and documentation
- explore LS inversion performance enhancing techniques and add more built-in options to IWAVE++
  - various regularization strategies
  - adaptive scaling
- explore different optimization methods
  - trust-region ...
  - beyond $L_2$ norm (methods from other communities)
Acknowledgements

Great thanks to

- Present and former TRIP team members especially Marco Enriquez, Xin Wang, Igor Terentyev, Tanya Vdovina
- Sponsors of The Rice Inversion Project
- NSF DMS 0620821

Thank you!