Resolution of 3D Elastic Full Waveform Inversion

Espen Birger Nilsen and Børge Arntsen

Norwegian University of Science and Technology (NTNU)
Department of Petroleum Engineering & Applied Geophysics
E-mail: espen.nilsen@ntnu.no

The ROSE meeting 2012
April 24th 2012
Outline

Introduction
  Full Waveform Inversion
  The Elastic Wave Equation

The Misfit Functionals
  Definitions
  Gradients
  Numerical Results

Conclusions
  Summary
  Acknowledgments
  References
Full Waveform Inversion

- Initial model
- Modeling
- Norm calculation
- Ending model
- New model
- Gradient calculation

- if $NORM < TOL$
- if $NORM > TOL$

The Elastic Wave Equation

Wave equation for the particle displacement ([Aki and Richards, 2002])

\[
\rho(x) \ddot{u}_i(x, t) = \partial_j \tau_{ij}(x, t) + f_i(x, t) \tag{1}
\]

\[
\tau_{ij}(x, t) = c_{ijkl}(x) \partial_l u_k(x, t) - I_{ij}(x, t) \tag{2}
\]

\(\rho\): density, \(u\): particle displacement, \(\tau_{ij}\): stress tensor, \(f\): body force, \(c_{ijkl}\): stiffness tensor, \(I_{ij}\): volume force.
Wave equation for the stress tensor

\[ s_{pqij}(x)\ddot{\tau}_{ij}(x, t) = \frac{1}{2} \left[ \partial_q \left( \frac{1}{\rho(x)} \partial_j \tau_{pj}(x, t) \right) + \partial_p \left( \frac{1}{\rho(x)} \partial_j \tau_{qj}(x, t) \right) \right] \]

\[ + \frac{1}{2} \left[ \partial_q \left( \frac{1}{\rho(x)} f_p(x, t) \right) + \partial_p \left( \frac{1}{\rho(x)} f_q(x, t) \right) \right] \]

\[ + \ddot{Q}_{ij}(x, t), \quad (3) \]

\[ e_{qp}(x, t) = s_{pqij}(x)\tau_{ij}(x, t) - Q_{pq}(x, t) \quad (4) \]

\( s_{pqij} \): compliance tensor, \( e_{qp} \): strain, \( Q_{pq}(x, t) = s_{pqij}(x)I_{ij}(x, t) \): volume source.
Modeling

- Full three dimensional elastic modeling.
- Staggered finite difference method described by Virieux ([Virieux, 1986]).
- Perfectly Matched Layer (PML) ([Zhen et al., 2009]) used as Absorbing Boundary Conditions (ABCs).
- Free surface modeled using approach by Mittet ([Mittet, 2002]).
- Problems: Large memory requirements, computer time, parallellization, etc.
The Misfit Functionals

General definition

\[ \Psi(d, f(m)) = C \int_T \int_{\Omega_r} W(d, f(m)) \, dS \, dt, \]  

Three cases investigated further;

\[ \Psi_{L1}(d, f(m)) = \int_T \int_{\Omega_r} |d - f(m)| \, dS \, dt, \]  
\[ \Psi_{L2}(d, f(m)) = \frac{1}{2} \int_T \int_{\Omega_r} (d - f(m))^2 \, dS \, dt, \]  
\[ \Psi_C(d, f(m)) = \frac{1}{2} \int_T \int_{\Omega_r} \ln \left(1 + (d - f(m))^2\right) \, dS \, dt. \]

\(\Omega_r\): receiver surface, \(T\): time interval, \(d\): measured field, \(f(m)\): modeled field.
Gradients

General formulation

\[ \nabla_m \Psi = \int_T \frac{\partial W(d, f(m))}{\partial f(m)} \frac{\partial f(m)}{\partial m} \, dt \quad (9) \]

See [Fichtner, 2011].

- Dependent on how the problem is parametrized; i.e. Lamé parameters, velocities, etc.
- Problem: Due to computer power “impossible” to find the Fréchet kernel.
- Solution: Tarantola ([Tarantola, 1984]) forward/backward wave field formulation.
The general forward/backward gradients:

\[
\nabla_\rho \Psi = - \int_T \int_\Omega \dot{u}_p(t) \left[ \hat{\Gamma}^0_{pij}(-t) * \frac{\partial W(d, f(m))}{\partial f(m)} \right] dV dt, \quad (10)
\]

\[
\nabla_\Lambda \Psi = - \int_T \int_\Omega \tau_{nn}(t) \left[ \Gamma^0_{ppij}(-t) * \frac{\partial W(d, f(m))}{\partial f(m)} \right] dV dt, \quad (11)
\]

\[
\nabla_M \Psi = - \int_T \int_\Omega \tau_{np}(t) \left[ \Gamma^0_{npj}(-t) * \frac{\partial W(d, f(m))}{\partial f(m)} \right] dV dt, \quad (12)
\]

where

\[
\Lambda = -\frac{\lambda}{2\mu(3\lambda + 2\mu)} \quad \text{and} \quad M = \frac{1}{4\mu}. \quad (13)
\]

Conversion between parameters is done by using differential calculus.
Principle formulas for gradients

\[ \nabla_{\rho} \Psi = C_{\rho} \int_{T} (\vec{v}_z + \vec{v}_x + \vec{v}_y) (\vec{v}_z + \vec{v}_x + \vec{v}_y) \, dt, \quad (14) \]

\[ \nabla_{v_p} \Psi = C_{v_p} \int_{T} (\tau_{zz} + \tau_{xx} + \tau_{yy}) (\dot{\tau}_{zz} + \dot{\tau}_{xx} + \dot{\tau}_{yy}) \, dt, \quad (15) \]

\[ \nabla_{v_s} \Psi = C_{v_s} \int_{T} (\tau_{zx} + \tau_{yx} + \tau_{zy}) (\dot{\tau}_{zx} + \dot{\tau}_{yx} + \dot{\tau}_{zy}) \, dt. \quad (16) \]

\( \vec{v} \): particle velocity, \( \tau_{ij} \): stress tensor, \( C_i \): constant dependent on parameter under consideration.
Numerical Results

- 500 meters grid in each direction; sampling 5 meters.
- Four source-receiver geometries: one-shot-many-receivers geometries.
Source-receiver geometries

**G1**: The receivers are placed in the whole receiver layer, and the source is in the middle of this layer.

**G2**: The receivers are placed in a square which is one quarter of the full layer. The source is placed in the middle of the full layer, i.e. on the corner of the square.

**G3**: The receivers consist of eight streamers that are separated by 50 meters. The streamers are placed in the middle of the receiver layer. The source is placed in front of the middle streamers.

**G4**: The receivers consist of a single streamer, which is placed in the middle of the receiver layer. The source is in front of the streamer.
Perturbation layer

(a): model 1, (b): model 2, (c): model 3.
Resolution matrix

In matrix notation

\[ \Delta \hat{v}_p = c \nabla v_p \Phi(d, f(m)) = c \mathbf{J}^T \mathbf{J} \Delta v_p, \quad (17) \]
\[ \Delta \hat{v}_p = c \mathbf{R} \Delta v_p, \quad (18) \]

\( \mathbf{R} = \mathbf{J}^T \mathbf{J} \): resolution matrix, \( c \): gradient constant.

\[
R \Delta v_p = \begin{bmatrix}
R_{11} & R_{12} & \ldots & R_{1n} \\
R_{21} & R_{22} & \ldots & R_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
R_{(m-1)1} & R_{(m-1)2} & \ldots & R_{(m-1)n} \\
R_{m1} & R_{m2} & \ldots & R_{mn}
\end{bmatrix}
\begin{bmatrix}
0 \\
0 \\
\vdots \\
\Delta v_{p,k} \\
0
\end{bmatrix} \quad (19)
\]
$L^2$ gradient images: Model 1

Geometry 1; \textit{top}: Horizontal slice at 250 m depth, \textit{bottom}: vertical slice at 250 m offset.
$L^2$ gradient images: Model 1

Geometry 2; top: Horizontal slice at 250 m depth, bottom: vertical slice at 250 m offset.
$L^2$ gradient images: Model 1

Geometry 3; top: Horizontal slice at 250 m depth, bottom: vertical slice at 250 m offset.
$L^2$ gradient images: Model 1

Geometry 4; *top*: Horizontal slice at 250 m depth, *bottom*: vertical slice at 250 m offset.
**$L^1$ gradient images: Model 1**

Geometry 1; *top*: Horizontal slice at 250 m depth, *bottom*: vertical slice at 250 m offset.
Gradient images: Model 2

Geometry 1, horizontal slices at 250 m depth; top: $L^2$-norm, bottom: $L^1$-norm.
Gradient images: Model 3

$L^2$-norm, horizontal slices at 250 m depth; top: Geometry 1, bottom: Geometry 3.
Summary

• The source-receiver geometry has major impact on the gradient.
  • Denser receiver grid gives more focusing of the gradient.
  • Many receivers compared to few receivers give better focusing. Conclusion: Use as many receivers as possible and put the source in the middle of the receiver grid.
• Coupled gradients.
• Different numerical artifacts for the gradients: $L^1$ seems to be the worst.
• The Cauchy and $L^2$ gradients have the same properties.
Summary

- The source-receiver geometry has *major* impact on the gradient.
- Denser receiver grid gives more focusing of the gradient.
- Many receivers compared to few receivers give better focusing. Conclusion: Use as many receivers as possible and put the source in the middle of the receiver grid.
- Coupled gradients.
- Different numerical artifacts for the gradients: $L^1$ seems to be the worst.
- The Cauchy and $L^2$ gradients have the same properties.
Summary

- The source-receiver geometry has *major* impact on the gradient.
- Denser receiver grid gives more focusing of the gradient.
- Many receivers compared to few receivers give better focusing. Conclusion: Use as many receivers as possible and put the source in the middle of the receiver grid.
- Coupled gradients.
- Different numerical artifacts for the gradients: $L^1$ seems to be the worst.
- The Cauchy and $L^2$ gradients have the same properties.
Summary

- The source-receiver geometry has *major* impact on the gradient.
- Denser receiver grid gives more focusing of the gradient.
- Many receivers compared to few receivers give better focusing. Conclusion: Use as many receivers as possible and put the source in the middle of the receiver grid.
- Coupled gradients.
  - Different numerical artifacts for the gradients: $L^1$ seems to be the worst.
  - The Cauchy and $L^2$ gradients have the same properties.
Summary

- The source-receiver geometry has *major* impact on the gradient.
- Denser receiver grid gives more focusing of the gradient.
- Many receivers compared to few receivers give better focusing. Conclusion: Use as many receivers as possible and put the source in the middle of the receiver grid.
- Coupled gradients.
- Different numerical artifacts for the gradients: $L^1$ seems to be the worst.
- The Cauchy and $L^2$ gradients have the same properties.
Summary

- The source-receiver geometry has *major* impact on the gradient.
- Denser receiver grid gives more focusing of the gradient.
- Many receivers compared to few receivers give better focusing. Conclusion: Use as many receivers as possible and put the source in the middle of the receiver grid.
- Coupled gradients.
- Different numerical artifacts for the gradients: $L^1$ seems to be the worst.
- The Cauchy and $L^2$ gradients have the same properties.
Acknowledgments

We acknowledge the BIGCCS Centre and the sponsors of the ROSE Consortium for financing this research.


References II