Derivation of FWI gradient using Adjoint State method for Elastic Media

Mrinal Sinha

KAUST

16/09/2014
Outline

- Gradient Derivation using Perturbation
- Adjoint State Method
- Elastic Wave Equation
- Derivation of Gradient using Adjoint Method
Elastic Green’s Tensor

The expression for Elastic Green’s Tensor is shown below:

\[
G(r|r_0) = \frac{-ik_\alpha}{12\pi(\lambda + 2\mu)} \left( I h_0^{(2)}(k_\alpha R) + (I - 3\hat{R}\hat{R}^T) h_2^{(2)}(k_\alpha R) \right)
\]

\[
+ \frac{ik_\beta}{12\pi\mu} \left( 2I h_0^{(2)}(k_\beta R) + (I - 3\hat{R}\hat{R}^T) h_2^{(2)}(k_\beta R) \right)
\]

Tensorial components of P wave  Tensorial components of S wave

\[
= \underbrace{G_\alpha}_{\text{Tensorial components of P wave}} + \underbrace{G_\beta}_{\text{Tensorial components of S wave}} \tag{1}
\]

where \( \hat{R} \) is the \( 3 \times 1 \) unit vector \( r - r_0 \) and the wave numbers \( k_\alpha \) and \( k_\beta \) are given by \( k_\alpha = \frac{\omega}{\alpha} \) and \( k_\beta = \frac{\omega}{\beta} \) where \( \alpha and \beta \) are P and S wave velocities.

The spherical Bessel functions are

\[
h_0^{(2)}(kR) = \frac{-e^{-ikR}}{ikR},
\]

\[
h_2^{(2)}(kR) = \left[ \frac{1}{ikR} + \frac{3}{(ikR)^2} + \frac{3}{(ikR)^3} \right] e^{-ikR}. \tag{2}
\]
Substituting the expression for Bessel functions we get

\[ G_\alpha = \frac{e^{-ik\alpha R}}{4\pi(\lambda + 2\mu)} \left( \frac{1}{R} \hat{R} \hat{R}^T - \frac{1}{ik\alpha R^2} (I - 3\hat{R} \hat{R}^T) \right. \]

\[ \left. - \frac{1}{(ik\alpha)^2 R^3} (I - 3\hat{R} \hat{R}^T) \right) , \]

\[ G_\beta = \frac{-e^{-ik\beta R}}{4\pi \mu} \left( \frac{1}{R} (I - 3\hat{R} \hat{R}^T) - \frac{1}{ik\beta R^2} (I - 3\hat{R} \hat{R}^T) \right. \]

\[ \left. - \frac{1}{(ik\beta)^2 R^3} (I - 3\hat{R} \hat{R}^T) \right) . \]  

(3)
Einstein notation

- $v_{i,j}$ stands for $\frac{\partial v_i}{\partial x_j}$
- A repetition of indices indicate a summation

$$v_i v_i = v_1^2 + v_2^2 + v_3^2$$
Derivation using Perturbation

- Elastic Wave equation can be written as

\[ \omega^2 \rho v_i + \partial_j (\lambda \partial_k v_k \delta_{ij} + \mu [\partial_i v_j + \partial_j v_i]) = -f_i \]  \hspace{1cm} (4)

- \( G_{im} \) is the solution to the wave equation for a point source at \( x_0 \) for a background medium (known \( \lambda, \mu, \) and \( \rho \))

Green's function

\[ \omega^2 \rho g_{im} + \partial_j (\lambda \partial_k G_{km} \delta_{ij} + \mu [\partial_i G_{jm} + \partial_j G_{im}]) = -\delta_{im} \delta(x-x_0) \]

- After perturbing the above equation in \( \lambda \) and neglecting the perturbation terms which are higher than 2nd order we get

\[ \omega^2 \rho \delta v_i + \partial_j (\lambda \partial_k \delta v_k \delta_{ij} + \mu [\partial_i \delta v_j + \partial_j \delta v_i]) = -\partial_j [\delta \lambda \partial_k \delta v_k \delta_{ij}] \]

\[ \delta v_m = \int G_{im} \partial_j (\delta \lambda \partial_k v_k \delta_{ij}) \, dx^3 \]

\[ = \int \partial_j [G_{im} (\delta \lambda \partial_k v_k \delta_{ij})] \, dx^3 - \int [G_{im,j} (\delta \lambda \partial_k v_k \delta_{ij})] \, dx^3 \]
The first integral can be converted into a surface integral at infinity using Gauss-Divergence theorem and only considering outgoing waves the integral goes to zero and we arrive to the Born Approximation.

\[ v(g)_m = -\int [G_{im,j}(\delta \lambda \partial_k v_k \delta_{ij})] dx^3 \]

\[ = -\int \left[ G(g|x)_{im,j} \delta_{ij} \delta \lambda(x) \partial_k v(x)_k \right] dx^3 \]

Assuming a that one model point in the subsurface is perturbed i.e. \( \delta \lambda = \delta(x - x_0)\delta \lambda(x_0) \) and integrating over the model space and dividing throughout by \( \delta \lambda(x_0) \) we get the Frechét derivative

\[ \frac{\delta v(g)_m}{\delta \lambda(x_0)} = -\left( \frac{\partial G(g|x)_{xm}}{\partial x} + \frac{\partial G(g|x)_{zm}}{\partial z} \right) \left( \frac{\partial v(x)_x}{\partial x} + \frac{\partial v(x)_z}{\partial z} \right) \]
FWI misfit for 2D Elastic Media

- for a single shot gather the misfit can be defined as the sum of the data residuals of different data components

\[
\epsilon = \frac{1}{2} \sum_\omega \sum_g \left[ \Delta \nu(g)_x \ast \Delta \nu(g)_x + \Delta \nu(g)_z \ast \Delta \nu(g)_z \right]
\]  \hspace{1cm} (5)

- We want to minimize the objective function using a gradient based method. The gradient of the above objective is given by:

\[
\frac{\partial \epsilon(g|s)}{\partial \lambda(x)} = \sum_\omega \sum_g \text{Real} \left[ \frac{\partial \nu(g)_x}{\partial \lambda(x)} \Delta \nu(g)_x \ast + \frac{\partial \nu(g)_z}{\partial \lambda(x)} \Delta \nu(g)_z \ast \right]
\]  \hspace{1cm} (6)
using
\[
\frac{\delta v(g)_m}{\delta \lambda(x_0)} = -\left( \frac{\partial G(g|x)_{xm}}{\partial x} + \frac{\partial G(g|x)_{zm}}{\partial z} \right) \left( \frac{\partial v(x)_x}{\partial x} + \frac{\partial v(x)_z}{\partial z} \right)
\] (7)

Define \( \hat{v}(x)_x \) and \( \hat{v}(x)_z \) are defined as:

\[
\hat{v}(x)_x = \sum_g \left[ G(g|x)_{xx} \Delta v(g)_x^* + G(g|x)_{xz} \Delta v(g)_z^* \right]
\]

\( x \)-component of the receiver side wavefield

\[
\hat{v}(x)_z = \sum_g \left[ G(g|x)_{zx} \Delta v(g)_x^* + G(g|x)_{zz} \Delta v(g)_z^* \right]
\]

\( z \)-component of the receiver side wavefield

and substituting in equation derived before

\[
\frac{\partial \epsilon}{\partial \lambda(x)} = -\sum_\omega \text{Re}\left[ \left( \frac{\partial \hat{v}(x)_x}{\partial x} + \frac{\partial \hat{v}(x)_z}{\partial z} \right) \left( \frac{\partial v(x)_x}{\partial x} + \frac{\partial v(x)_z}{\partial z} \right) \right]
\] (8)
Outline

- Gradient Derivation using Perturbation
- Adjoint State Method
- Elastic Wave Equation
- Derivation of Gradient using Adjoint Method
Review of the Adjoint State Method

- $L_2$ misfit function which is the measure of the difference between the observed and predicted data.

\[ \epsilon(m) = \frac{1}{2} (w(m) - d, w(m) - d) \]  \hspace{1cm} (9)

where $w$ is the predicted data calculated using $Aw = f$, where $A$ is the forward modeling operator and $d$ is the observed data.

- The gradient is given by

\[ \frac{\partial \epsilon(m)}{\partial m_i} = \left( \frac{\partial w(m)}{\partial m_i}, w(m) - d \right) \] \hspace{1cm} (10)

- where $\frac{\partial w(m)}{\partial m_i}$ is the Frechét derivative.
Review Continued

Taking the Frechét derivative of $Aw = f$ and rearranging terms we get

$$\frac{\partial w(m)}{\partial m_i} = -A(m)^{-1} \frac{\partial A(m)}{\partial m_i} w(m) \quad (11)$$

Plugging equation 11 in place of $\frac{\partial w(m)}{\partial m_i}$ in equation 10.

$$\frac{\partial \epsilon(m)}{\partial m_i} = -(A(m)^{-1} \frac{\partial A(m)}{\partial m_i} w(m), \Delta d) \quad (12)$$

where $\Delta d$ is the data residual. Rearranging the terms the above equation can be written as:

$$\frac{\partial \epsilon(m)}{\partial m_i} = -\left(\frac{\partial A(m)}{\partial m_i} w(m), \right) \left(A(m)^{-1}\right)^\dagger \Delta d \quad (13)$$

$v-$Solve Adjoint Wave equation
Outline

- Gradient Derivation using Perturbation
- Adjoint State Method
- Elastic Wave Equation
- Derivation of Gradient using Adjoint Method
Isotropic wave equation in terms of Lamé parameters $\lambda$ and $\mu$ and the density $\rho$ can be written as a matrix operator as shown below:

\[
\begin{bmatrix}
\rho \frac{\partial}{\partial t} \\
0 \\
-(\lambda + 2\mu) \frac{\partial}{\partial x} \\
-\lambda \frac{\partial}{\partial x} \\
-\mu \frac{\partial}{\partial z}
\end{bmatrix}
\begin{bmatrix}
0 \\
0 \\
-\lambda \frac{\partial}{\partial z} \\
\frac{\partial}{\partial t} \\
-\mu \frac{\partial}{\partial x}
\end{bmatrix}
\begin{bmatrix}
\frac{\partial}{\partial x} \\
0 \\
\frac{\partial}{\partial z} \\
0 \\
0
\end{bmatrix}
\begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\begin{bmatrix}
0 \\
0 \\
-\frac{\partial}{\partial z} \\
-\frac{\partial}{\partial x} \\
-\frac{\partial}{\partial x}
\end{bmatrix}
\begin{bmatrix}
\dot{u}_x \\
\dot{u}_z \\
\tau_{xx} \\
\tau_{zz} \\
\tau_{xz}
\end{bmatrix}
= 
\begin{bmatrix}
f_{\dot{u}_x} \\
f_{\dot{u}_x} \\
f_{\tau_{xx}} \\
f_{\tau_{zz}} \\
f_{\tau_{xz}}
\end{bmatrix}
\]
Outline

- Gradient Derivation using Perturbation
- Adjoint State Method
- Elastic Wave Equation
- Derivation of Gradient using Adjoint Method
Derivation of Gradient

- Using the forward modeling operator we can evaluate the gradients with respect to the perturbations in different model parameters such as $\lambda$, $\mu$ and $\rho$.
- Gradient with respect to $\lambda$, $\mu$ and $\rho$ can be found out by calculating $\frac{\partial A}{\partial \lambda}$, $\frac{\partial A}{\partial \mu}$ and $\frac{\partial A}{\partial \rho}$ respectively.

\[
\frac{\partial A}{\partial \lambda} = \begin{bmatrix}
 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 \\
 -\frac{\partial}{\partial x} & -\frac{\partial}{\partial z} & 0 & 0 & 0 \\
 \frac{\partial}{\partial x} & \frac{\partial}{\partial z} & 0 & 0 & 0 \\
 \frac{\partial}{\partial t} & \frac{\partial}{\partial t} & 0 & 0 & 0 \\
\end{bmatrix}
\]

(15)

\[
\frac{\partial A}{\partial \mu} = \begin{bmatrix}
 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 \\
 -2\frac{\partial}{\partial x} & 0 & 0 & 0 & 0 \\
 0 & -2\frac{\partial}{\partial z} & 0 & 0 & 0 \\
 \frac{\partial}{\partial t} & \frac{\partial}{\partial t} & 0 & 0 & 0 \\
\end{bmatrix}
\]

(16)
Derivation of Gradient

Similarly $\frac{\partial A}{\partial \rho}$ is given by:

\[
\begin{bmatrix}
\rho \frac{\partial}{\partial t} & 0 & -\frac{\partial}{\partial x} & 0 & -\frac{\partial}{\partial z} \\
0 & \rho \frac{\partial}{\partial t} & 0 & -\frac{\partial}{\partial z} & -\frac{\partial}{\partial x} \\
-(\lambda + 2\mu) \frac{\partial}{\partial x} & -\lambda \frac{\partial}{\partial z} & \frac{\partial}{\partial t} & 0 & 0 \\
-\lambda \frac{\partial}{\partial x} & -(\lambda + 2\mu) \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial t} & 0 \\
-\mu \frac{\partial}{\partial z} & -\mu \frac{\partial}{\partial x} & 0 & 0 & \frac{\partial}{\partial t}
\end{bmatrix}
\]

\[\frac{\partial A}{\partial \rho} = \begin{bmatrix}
\frac{\partial}{\partial t} & 0 & 0 & 0 & 0 \\
0 & \frac{\partial}{\partial t} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix} \tag{17}\]

Using the equation derived before

\[
\frac{\partial \varepsilon(m)}{\partial m_i} = -\left(\frac{\partial A(m)}{\partial m_i} w(m), (A(m)^{-1})^\top \Delta d\right) \tag{18}
\]
Gradient with respect to the perturbation in $\lambda$

- Using the equation derived before

\[
\frac{\partial \epsilon(m)}{\partial \lambda} = \left( \frac{\partial A(m)}{\partial \lambda} \right) w(m), (A(m)^{-1})^\dagger \Delta d
\]  

(19)

- The above equation can be expanded into

\[
\frac{\partial \epsilon(m)}{\partial \lambda} = -\left( \begin{array}{c} 0 \\ 0 \\ \nabla \cdot \dot{u} \\ \nabla \cdot \dot{u} \\ 0 \end{array} \right), \left( \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{array} \right)
\]  

(20)

\[
= -\int_0^T \left( \frac{\partial \dot{u}_x}{\partial x} + \frac{\partial \dot{u}_z}{\partial z} \right) (v_3 + v_4) \, dt
\]
Gradient with respect to the perturbation in $\mu$ and $\rho$

Using a similar procedure as before we can show that the gradients with respect to the perturbations in $\mu$ and $\rho$ are given as.

$$\frac{\partial \epsilon(m)}{\partial \mu} = -2 \int_0^T \left( v_3 \frac{\partial \hat{u}_x}{\partial x} + v_4 \frac{\partial \hat{u}_z}{\partial z} + \frac{1}{2} v_5 \left\{ \frac{\partial \hat{u}_z}{\partial x} + \frac{\partial \hat{u}_x}{\partial z} \right\} \right) dt$$ (21)

$$\frac{\partial \epsilon(m)}{\partial \rho} = - \int_0^T \left( v_1 \frac{\partial \hat{u}_x}{\partial t} + v_2 \frac{\partial \hat{u}_z}{\partial t} \right) dt$$ (22)