Wasserstein metric based full waveform inversion

Anna Karpova, 20.09.18
Contents

- Introduction
- The forward problem
- Different misfit functionals
- Minimisation methods
- Computational results
- Conclusion
Introduction

- the thesis deals with a simple application of the reflection method to the exploration of oil and gas fields

- full waveform inversion (FWI) is the technique used in reflection seismology to recover the information from observed seismograms

- FWI consists of an efficient solution of the forward problem and iterative improvement of a subsurface model by minimisation a misfit between observed and synthetic seismic waveforms
Introduction

- the least squares ($L^2$) norm is a classically used misfit functional in FWI, but it usually provides multiple local minima

- Wasserstein metric is another misfit functional we want to use in this work, because it has some desirable properties like convexity and insensitivity to noise

- the goal of this work is to create an example of a two-dimensional seismic problem and to solve it with FWI technique once with $L^2$ norm and once with Wasserstein metric
Introduction

Elastic waves in two-dimensional space can be modelled by the following wave equation in order to obtain the resulting wavefield \( u(x,z,t) \) for a given wave velocity \( c(x,z) \)

\[
    u_{tt}(x, z, t) - \nabla \cdot (c^2(x, z) \nabla u(x, z, t)) = s(x, z, t),
\]

where \( s(x,z,t) \) is the source function.

observed data:

\[
    g = u(x_r, 0, t)
\]

modelled data:

\[
    f(c) = u(x_r, 0, t).
\]

minimisation problem:

\[
    \tilde{c} = \arg\min_c d(f(c), g),
\]

where \( d(f,g) \) is a misfit functional for two signals
The forward problem
The seismic reflection experiment

Cartoon of the land seismic experiment
The forward problem
The problem statement

\[ \Omega = (0, a) \times (0, b), \quad a, b > 0 \]

\[ \Omega_\infty = \{(x, z) \in \mathbb{R}^2 \mid z \geq 0\} \]

\[ \begin{align*}
    u_{tt}(x, z, t) - \nabla \cdot (c^2(x, z) \nabla u(x, z, t)) &= s(x, z, t) \quad (x, z) \in \Omega, \quad t > 0, \tag{2.1a} \\
    u_{tt}(x, z, t) - c^2 \Delta u(x, z, t) &= 0 \quad (x, z) \in \Omega_\infty \setminus \Omega, \quad t > 0, \tag{2.1b} \\
    \frac{\partial}{\partial n} u(x, z, t) &= 0 \quad (x, z) \in \partial \Omega_\infty, \quad t > 0, \tag{2.1c} \\
    \lim_{(x, z) \to \infty} u(x, z, t) &= 0 \quad t > 0, \tag{2.1d} \\
    u(x, z, 0) &= u_0(x, z), \quad (x, z) \in \Omega, \tag{2.1e} \\
    u_t(x, z, 0) &= v_0(x, z) \quad (x, z) \in \Omega. \tag{2.1f}
\end{align*} \]

\( u(x, z, t) \) – wave field
\( s(x, z, t) \) – source
\( c(x, z) \) – wave velocity
\( u_0, \; v_0 \) – initial conditions
The forward problem

The problem statement

The PML modified wave equation (2.5) replace the conditions 2.1a, 2.1b and 2.1d.

\[
\begin{align*}
  u_{tt} + (\zeta + \eta) u_t + \zeta \eta u &= \nabla \cdot (c^2 \nabla u) + s + \nabla \cdot \begin{pmatrix} \phi \\ \psi \end{pmatrix}, \\
  \phi_t &= -\zeta \phi + c^2 (\eta - \zeta) \frac{\partial u}{\partial x}, \\
  \psi_t &= -\eta \psi + c^2 (\zeta - \eta) \frac{\partial u}{\partial z}.
\end{align*}
\]  

(2.5)

\(\phi, \psi\) – auxiliary functions
\(\zeta, \eta\) – damping profiles

\(\zeta(x) = 0\) in \(\Omega\) and \(\zeta(x) > 0\) in \(\Omega^C\),

\(\eta(z) = 0\) in \(\Omega\) and \(\eta(z) > 0\) in \(\Omega^C\).
The forward problem
Discretisation

\[ x_i = i \cdot \Delta x \text{ and } z_j = j \cdot \Delta z, \]
where \( \Delta x = \frac{1}{k+1}, i = 0, 1, \ldots k + 1 \) and \( \Delta z = \frac{1}{l+1}, j = 0, 1, \ldots l + 1 \)

\[ w_{ij}^m \simeq u(x_i, z_j, t_m) \]

\[
\frac{w_{ij}^{m+1} - 2w_{ij}^m + w_{ij}^{m-1}}{(\Delta t)^2} = c_{ij}^2 \left( \frac{w_{i+1,j}^m - 2w_{ij}^m + w_{i-1,j}^m}{(\Delta x)^2} + \frac{w_{i,j+1}^m - 2w_{ij}^m + w_{i,j-1}^m}{(\Delta z)^2} \right)
\]

CFL stability condition for equal mesh size \( h = \Delta x = \Delta z \):

\[ c \left( \frac{\Delta t}{h} \right) \leq \frac{1}{\sqrt{2}}. \]
The forward problem
Numerical experiments

Snapshots of the numerical solutions at different times in $\Omega = [-0.5, 0.5]^2$ enclosed by a PML of width $L = 0.1$
The forward problem
Numerical experiments

Illustration of the two-layers model. The regular computation domain $\Omega$ is surrounded by absorbing layers in which the plane waves decay rapidly as they approach the boundary.
Numerical solution of the two-layers model in $\Omega = (0, 3) \times (0, 1)$ with wave velocities $c_1 = 1$ in the upper layer, $c_2 = 3$ in the lower layer and the depth of the upper layer $d_1 = 0.4$
Different misfit functionals

Least-squares ($L^2$) norm

For observed data $g$ and simulated data $f$ recorded in the time interval $(0, T)$ at receiver position $x_r$, the conventional full waveform inversion defines $L^2$ norm misfit functional as

$$L^2(f(m), g) = \int_0^T |f(x_r, t, m) - g(x_r, t)|^2 dt,$$

where $m$ is the model parameter.

For numerical approximation of the $L^2$ norm, we subdivide the interval $(0, T)$ in $N$ equal subintervals $[t_i, t_{i+1}]$ of the length $h$. In each subinterval we use the Simpson's formula

$$S[\tilde{f}] = \frac{h}{6} \left( \tilde{f}(t_i) + 4\tilde{f}\frac{t_i + t_{i+1}}{2} + \tilde{f}(t_{i+1}) \right) \approx \int_{t_i}^{t_{i+1}} \tilde{f}(t) dt,$$

where $\tilde{f}(t) = (f(x_r, t, m) - g(x_r, t))^2$. 
Different misfit functionals

Quadratic Wasserstein metric $W_2$

Wasserstein metric computes the lowest cost of rearranging one distribution into another given a cost function.

For two probability density functions $f,g$ the quadratic Wasserstein metric is given by

$$W_2^2(f, g) = \inf_{T \in M} \int_X |x - T(x)|^2 f(x) \, dx,$$

where $M$ is the set of all maps that rearrange $f$ into $g$.

$$F(x) = \int_{-\infty}^{x} f(t) \, dt, \quad G(y) = \int_{-\infty}^{y} g(t) \, dt$$

**Theorem 3.4** Let $0 < f, g < \infty$ be two probability density functions, each supported on a connected subset of $\mathbb{R}$. Then the optimal map from $f$ to $g$ is $T = G^{-1} \circ F$.

$$W_2^2(f, g) = \int_0^T |t - G^{-1}(F(t))|^2 f(t) \, dt.$$

Wasserstein metric based full waveform inversion, Anna Karpova, 20.09.18
Different misfit functionals
Quadratic Wasserstein metric $W_2$

Convexity of $W_2$ metric with respect to shift:

**Theorem 3.7** Let $T : X \subset \mathbb{R}^n \to Y \subset \mathbb{R}^n$ be an optimal transference plan rearranging $f$ into $g$, where $f$ and $g$ are two density functions. Then the optimal transference plan from a shifted density function $f_s(x) = f(x - s\eta)$, $\eta \in \mathbb{R}^n$ into $g$ is $T_s = T(x - s\eta)$. Moreover, $W_2^2(f_s, g)$ is convex with respect to the shift size $s$.

Data normalisation:

- $f^+ = \max\{f, 0\}$, $f^- = \max\{-f, 0\}$,
  \[
  \langle f \rangle = \int_X f(x) dx
  \]
  \[
  W_2^2(f, g) \approx W_2^2\left( \frac{f^+}{\langle f^+ \rangle}, \frac{g^+}{\langle g^+ \rangle} \right) + W_2^2\left( \frac{f^-}{\langle f^- \rangle}, \frac{g^-}{\langle g^- \rangle} \right)
  \]

- $W_2^2(f, g) \approx W_2^2\left( \frac{f + c}{\langle f + c \rangle}, \frac{g + c}{\langle g + c \rangle} \right)$
Different misfit functionals
Numerical experiments
Different misfit functionals

Numerical experiments
Minimisation methods
The Method of Steepest Descent

Algorithm 1: Method of Steepest Descent

Input: \( x_0 \equiv 0 \)

Output: \( x_k \in \mathbb{R}^n \)

1. Initialize \( k \leftarrow 0 \)

2. while \( \nabla f(p) \neq 0 \) do
   1. Determine a direction of steepest descent \( d_k \in \mathbb{R}^n \), so that \( \nabla f(x_k)^\top d_k < 0 \)
   2. Solve the problem \( \alpha_k \approx \arg\min_{\alpha \in \mathbb{R}} f(x_k + \alpha d_k) \)
   actualise \( x_{k+1} \leftarrow x_k + \alpha_k d_k \)
   set \( k \leftarrow k + 1 \)

7 return \( x_k \)

Armijo-Goldstein rule:

\[
f(p + \alpha \delta p) \leq f(p) + \alpha \sigma \nabla f(p)^\top \delta p
\]
Minimisation methods
Newton's Method

Algorithm 3: Newton algorithm

**Input:** function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and start approximation $x_0 \in \mathbb{R}^n$

**Output:** sequence of iterations $\{x_k\}_{k \in \mathbb{N}}$

1. Initialisation: set $k := 0$
2. Calculate the Newton-direction by solving the linear equation system

   $$\nabla^2 f(x_k) d_k = -\nabla f(x_k)$$

3. Set $x_{k+1} := x_k + d_k$
4. Increase $k := k + 1$ and go to 2.
Computational results
FWI with $L^2$ and $W_2$ misfit functionals

$\mathbf{d}_1 = 0.5$

$\mathbf{d}_1 = 1.0$

$\mathbf{d}_1 = 1.5$
Computational results
FWI with $L^2$ and $W_2$ misfit functionals

Misfit function in $L^2$ for varying layer thickness $d_i$
Computational results
FWI with $L^2$ and $W^2_2$ misfit functionals

Using the Method of Steepest Descent

Using the Newton's Method
Computational results
FWI with $L^2$ and $W_2$ misfit functionals

Lower value region of misfit function in $W_2$ for varying layer thickness $d_1$
Computational results
FWI with $L^2$ and $W_2$ misfit functionals

Using the Method of Steepest Descent
Computational results
Extension to many observation points

Misfit function in $L^2$

One observation point

Six observation points
Computational results
Extension to many observation points

Misfit function in $W_2$

One observation point

Six observation points
Computational results
Extension to three-layers model

Misfit function in $L^2$ norm
Computational results
Extension to three-layers model

Misfit function in $W_2$ norm
Conclusion

Wasserstein metric is a good candidate to replace the commonly used $L_2$ norm in the full waveform inversion.
Acknowledgements

- Prof. Dr. Markus Grote

- Jet Hoe Tang
Thank you
for your attention.