The Quadratic Wasserstein Metric for Earthquake Location

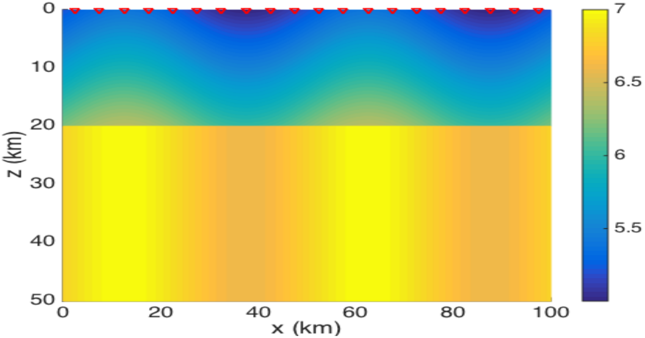
Yifan Chen
Tsinghua University

Collaborators: J. Chen, H. Wu, D.H. Yang

October 2017

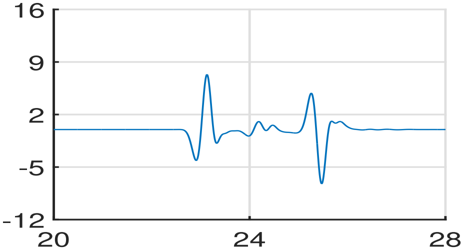
- twolayer model

$$c(x,z)=\begin{cases}5.2+0.05z+0.2\sin(\pi x/25), & 0km\le z\le 20km,\\ 6.8+0.2\sin(\pi x/25), & 20km< z\le 40km.\end{cases}$$



red triangles indicate the receivers; total number: r

- the earthquake happened at location $\boldsymbol{\xi}_T$ and time τ_T (unknown)
- seismic signals $g_i (1 \leq i \leq r)$ observed by the receivers (known)



- given $\boldsymbol{\xi}, \tau$, we could compute the synthetic signals using PDE-based model

$$f_i = \mathcal{L}_i(\boldsymbol{\xi}, \tau, c) \quad 1 \leq i \leq r$$

- optimization problem

$$(\boldsymbol{\xi}^*, \tau^*) = \underset{\boldsymbol{\xi}, \tau}{\operatorname{argmin}} \sum_{i=1}^{r} \mathbf{d}(f_i, g_i)$$

Problems

- existence of local minima
- observed signals contain noise
- expensive PDE solving

Previous work

- Prof. Engquist and Dr. Froese first used the Wasserstein metric to measure the misfit between seismic signals [1] in velocity structure inversion
- Dr. Métivier and collaborators proposed the KR norm based full waveform inversion [2]

Our solutions

- **model :** square and normalize the signal and use Wasserstein metric to compare them
- **optimization :** use LMF algorithm to optimize with high efficiency

[1] Engquist, Froese, and Yang [2016]
[2] Métivier, Brossier, Mérigot, Oudet, and Virieux [2017]

- for probability density functions $\tilde{f}$, $\tilde{g}$, the quadratic Wasserstein metric

$$W_2^2(\tilde{f},\tilde{g}) = \inf_{T\in\mathcal{M}} \int_{\mathbb{R}} |t - T(t)|^2 \tilde{f}(t)\mathrm{d}t.$$

  Set $\mathcal{M}$ contains all the rearrange maps from $\tilde{f}$ to $\tilde{g}$.

- for one dimension $\tilde{f}$, $\tilde{g} : \mathbb{R} \to \mathbb{R}^+$

$$W_2^2(\tilde{f},\tilde{g}) = \int_0^1 \left|F^{-1}(t) - G^{-1}(t)\right|^2 \mathrm{d}t,$$

  in which

$$F(t) = \int_{-\infty}^t \tilde{f}(\tau)\mathrm{d}\tau, \quad G(t) = \int_{-\infty}^t \tilde{g}(\tau)\mathrm{d}\tau.$$

- the optimal transport map

$$T(t) = G^{-1}(F(t)).$$

- our misfit function: for synthetic signal f and observed signal g supported in $[0, t_f]$

$$\mathbf{d}(f,g) = W_2^2\left(\frac{f^2}{\langle f^2\rangle}, \frac{g^2}{\langle g^2\rangle}\right)$$

- Fréchet gradient

$$\delta\mathbf{d} = \int_0^{t_f} 4\left(A(t) - B\right) f(t)\delta f(t)\mathrm{d}t$$

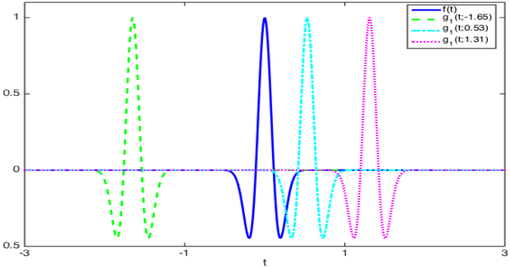
where

$$A(t) = \frac{\int_0^t (\tau - T(\tau))\mathrm{d}\tau}{\int_0^{t_f} f^2(t)\mathrm{d}t}, \quad B = \frac{\int_0^{t_f}\left(\int_0^t (\tau - T(\tau))\mathrm{d}\tau\right) f^2(t)\mathrm{d}t}{\left(\int_0^{t_f} f^2(t)\mathrm{d}t\right)^2}$$

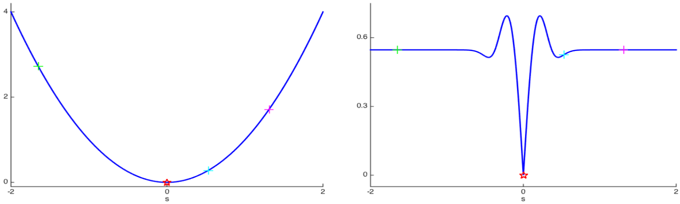
- apply adjoint method to obtain

$$\delta\mathbf{d} = K^{\boldsymbol{\xi}} \cdot \delta\boldsymbol{\xi} + K^{\tau}\cdot\delta\tau$$

- time shift convexity

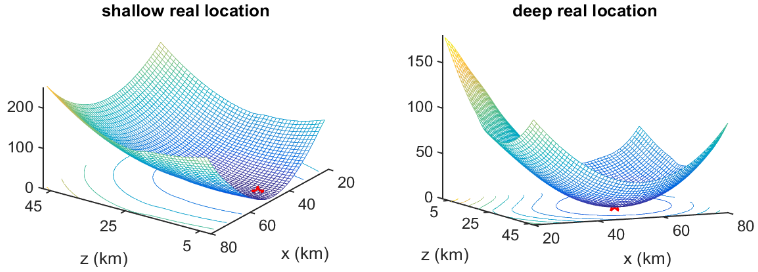


the signals $f(t),\ g_i(t;s)$



distance $\mathbf{d}_i(s)$(Left) and L^2 difference $\|f(t) - g_i(t;s)\|_2$ (Right)

- convexity of our $\sum_{i=1}^{r} \mathbf{d}(f_i(\boldsymbol{\xi}, \tau_T, c), g_i)$ with respect to $\boldsymbol{\xi}$



- quadratic structure and least-square formulation $\rightarrow$ LMF algorithm

- inversion result

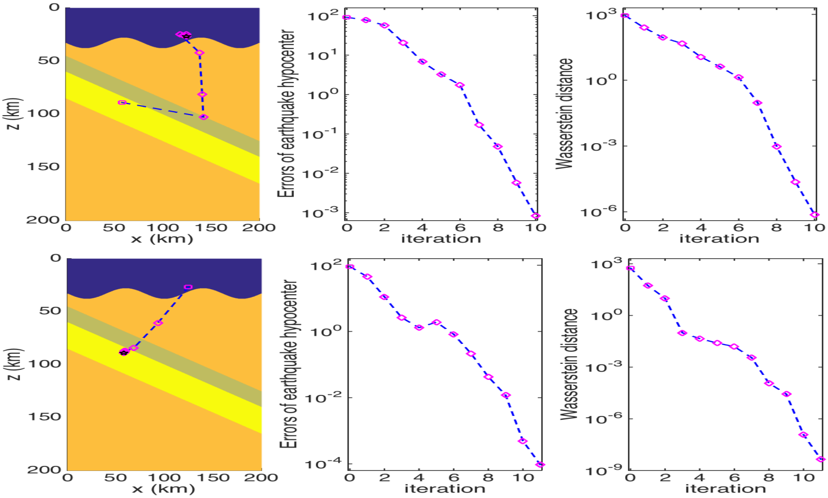


Figure: Convergence history of the subduction plate model

- noisy observed signal $g_N(t)$ and real signal g on $[0, t_f]$

$$g_N(t) = g(t) + r_N(t),\ r_N(t) = r_j,\ t \in (\frac{(j-1)t_f}{N}, \frac{jt_f}{N}],\ (1 \le j \le N)$$

where r_j are i.i.d. random variables with $\mathbb{E}r_j = 0,\ \mathbb{D}r_j = \sigma^2$

- the redefined distance function:

$$\mathbf{d}_{\lambda(t)}(f, g_N) = W_2^2\left(\frac{f^2+\lambda}{\langle f^2+\lambda\rangle}, \frac{g_N^2}{\langle g_N^2\rangle}\right)$$

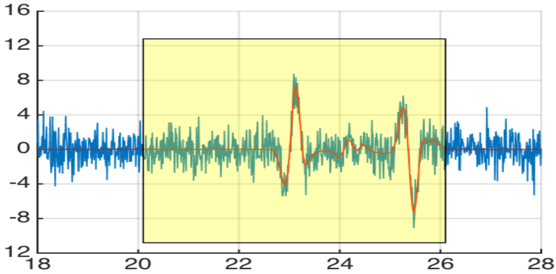
ideal choice $\lambda(t) = \sigma^2$, which leads to

$$\mathbb{E}\mathbf{d}_\lambda(g, g_N) = O(\frac{1}{N})$$

- Comparison

$$\mathbb{E}\|g - g_N\|_{L^2} = O(1)$$

- noise



advantage of W_2 metric in earthquake location problem

- promising convexity
- impact of data noise reduced
- fast convergence!

References:

Bjorn Engquist, Brittany D Froese, and Yunan Yang. Optimal transport for seismic full waveform inversion. 14(8), 2016.

L. Métivier, R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux. Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion. *Geophysical Journal International*, 205(1):332–364, 2017.

Thanks for your attention!