Bayesian First-Arrival Traveltime Tomography using Physics-Informed Neural Networks and Function-Space Particle-Based Variational Inference

Project Representative
Ryoichiro Agata Subduction Dynamics Research Center, Research Institute for Marine Geodynamics, Japan Agency for Marine-Earth Science and Technology

Authors
Ryoichiro Agata *[1], Kazuya Shiraishi *[1], Gou Fujie *[1]
*[1] Subduction Dynamics Research Center, Research Institute for Marine Geodynamics, Japan Agency for Marine-Earth Science and Technology

Keywords：Seismic tomography, physics-informed neural network (PINN), Bayesian neural network, function-space stein variation gradient descent (fSVGD)

1. Introduction

First-arrival traveltime tomography (FAT) using refraction seismic data is a crucial and useful technique for understanding the seismic wave velocity structure at depth. In order to ensure the reliability of the estimation, it is essential to quantify the uncertainty of seismic velocity estimates in tomography. For this purpose, Bayesian estimation has been introduced to seismic tomography to estimate the posterior probability distribution function (P-PDF) of the velocity structure based on errors in travel time data and prior information (e.g., [1]). All of these previous studies introduce a grid- or mesh-based discretization of the analysis domain for calculating the travel time using a numerical method and parametrizing the velocity for Bayesian estimation. Since this estimation problem is nonlinear, sampling methods such as Markov chain Monte Carlo (MCMC) are commonly used.

Physics-informed neural networks (PINN) [2], which solves partial differential equations and inverse problems with neural networks (NNs) constrained from the equations, has attracted much attention in recent years. It has been also applied to seismic tomography [3]. This is a mesh-free framework that leverages continuous functions represented by NNs, which are plausible and flexible for modeling the velocity structure. Considering this advantage, this study develops a novel and efficient Bayesian estimation framework for PINN-based seismic tomography [4].

2. Methods

In PINN-based seismic tomography, two NNs are used: one predicting seismic velocities from coordinates and the other predicting travel time from the source and receiver coordinates. The weight parameters of the two NNs are optimized so that the predicted travel time is close to the observed one, while the seismic velocity and travel time satisfy the Eikonal equation at random evaluation points. In Bayesian seismic tomography, the goal is to estimate the posterior predictive distribution of seismic wave velocities predicted by the NNs. Such an approach is classified as Bayesian neural network (BNN), in which the weight parameters that compose NNs are Bayesianly estimated using prior information. We introduce a novel and efficient Bayesian estimation method called particle-based variational inference (ParVI), best known by stein variational gradient descent (SVGD) [5], which has advantages in its high parallelism and good approximation efficiency. In SVGD, a functional expressing the difference between the PDF approximated by many particles and the true one is minimized by updating each particle sequentially using gradient information. However, when the network structure is complex, multimodality of the P-PDF of the weight parameters becomes stronger, and the BNN based on ordinary SVGD underestimates the uncertainty. Therefore, we formulate SVGD not in the space of weights, but in the space of continuous functions predicted by the NN [6], i.e., seismic wave velocities. This approach leverages the assumption that the shape of the PDF in the function space is simpler and smoother than in the weight space, and thus easier to approximate. We call this proposed method "velocity-space SVGD for PINN-based seismic tomography" (vSVGD-PINN-ST). The computer program was written by Python, using the deep learning framework of Pytorch. MPI parallelization is employed for particles introduced by SVGD using "MPI for Python".

3. Results of numerical experiments

We first verified our method via one-dimensional linear tomographic analyses with Gaussian prior and likelihood function, obtaining the posterior PDF that agrees well with analytical solutions. Such problems cannot be handled by naive baseline algorithms that perform SVGD in the weight space of both velocity and traveltime NNs. Then we conducted numerical experiments on the observation arrangement and velocity structure simulating refraction FAT (Figure 1). 512 SVGD particles were employed. The mean velocity model agrees well with the true one within the ray coverage. The uncertainty is larger outside of the coverage, while it is smaller but shows heterogeneous distribution elsewhere. The results of estimating the P-PDF were found to be reasonable for the observation arrangement. This result suggests that the proposed method is applicable to real seismic data. In addition, each particle (sample) of the velocity distribution that constitutes the approximate P-

PDF had a physically natural smooth distribution represented by NN (Figure 2). This is in contrast to some previous Bayesian seismic tomographic methods based on adaptive and low-dimensional parametrization, in which individual samples discontinuous velocity distributions due to Voronoi tessellation (e.g., [7]). This calculation used 512 CPU cores (64-core AMD EPYC 7742 x 8) in Earth Simulator and required 27.8 hours.

4. Conclusion

In this study, we developed the vSVGD-PINN-ST algorithm, which performs PINN-based Bayesian seismic tomography using SVGD, the best-known particle-based variational inference method, applied in the velocity space predicted by NN and enhanced with several mathematical and numerical techniques. The vSVGD-PINN-ST performance was tested in the one- and two-dimensional Bayesian seismic tomography synthetic tests. The success of the 2D synthetic test adopting a realistic observation geometry suggests that our method is ready for application to actual observational data first-arrival traveltime tomography, which we consider as the next task.

Acknowledgement

This research was supported by JSPS KAKENHI Grant Number 21K14024.

References

[1] T. Ryberg and C. Haberland, “Bayesian inversion of refraction seismic traveltime data,” Geophysical Journal International, vol. 212, no. 3, pp. 1645–1656, 2018.

[2] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” Journal of Computational physics, vol. 378, pp. 686–707, 2019.

[3] U. B. Waheed, T. Alkhalifah, E. Haghighat, C. Song, and J. Virieux, “PINNtomo: Seismic tomography using physics-informed neural networks,” arXiv preprint arXiv:2104.01588, 2021.

[4] R. Agata, K. Shiraishi, and G. Fujie, “Bayesian seismic tomography based on velocity-space Stein variation gradient descent for physics-informed neural network,” arXiv preprint arXiv:2301.07901, 2023.

[5] Q. Liu and D. Wang, “Stein variational gradient descent: A general purpose bayesian inference algorithm,” Advances in neural information processing systems, vol. 29, 2016.

[6] Z. Wang, T. Ren, J. Zhu, and B. Zhang, “Function Space Particle Optimization for Bayesian Neural Networks,” in International Conference on Learning Representations, 2019.

[7] T. Bodin and M. Sambridge, “Seismic tomography with the reversible jump algorithm,” Geophysical Journal International, vol. 178, no. 3, pp. 1411–1436, 2009

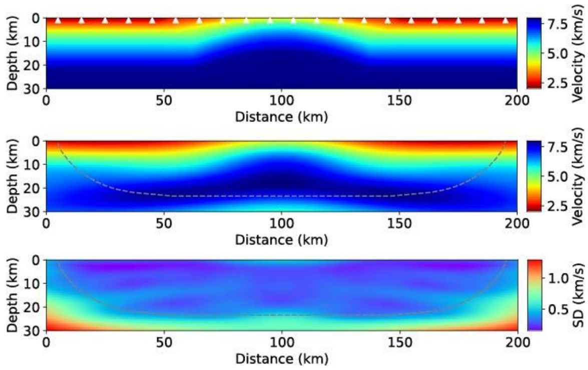

Figure 1. Numerical experiment on the observation arrangement and velocity structure simulating refraction FAT. (Upper) The true velocity model. (Middle) The estimated mean velocity model. The gray dashed line denotes the lower limit of ray coverage. (Bottom) The estimated standard deviation (SD).

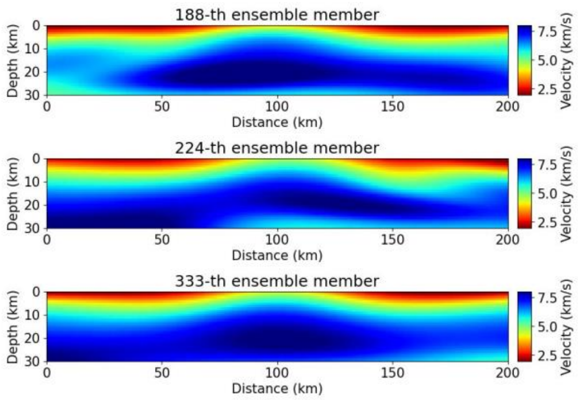

Figure 2. Three examples of individual particles (ensemble members).