Predictive uncertainty quantification for Bayesian Physics-Informed Neural Network (PINN) in hypocentre estimation problem

Introduction

Physics-informed neural networks (PINNs) have appeared on the scene as a flexible and a versatile framework for solving partial differential equations (PDEs), along with any initial or boundary conditions (Raissi et al., 2019). For seismic applications, Song et al. (2021) and Alkhalifah et al. (2021) used PINNs as a solver for the wave equation, including inverting for the velocity model. Waheed et al. (2021b) utilised it to provide a framework to solve the eikonal equation and Waheed et al. (2021a) extended the framework for seismic tomography. An important component of these solutions, especially when using the data as a boundary condition, is our confidence in their accuracy. There has been little study of PINN accuracy as an inversion tool. We introduce an approximate Bayesian framework for estimating predictive uncertainties in Physics-Informed Neural Network (PINN).

Neural networks naturally embed stochasticity through the random realisation of weights and biases, thus, propagating uncertainties into its predictive solutions (Yang et al., 2021). This work investigates propagation of uncertainties from the random realisations of Eikonal PINN’s weights and biases using the Laplace approximation (Ritter et al., 2018; Daxberger et al., 2021). We specifically use a hypocenter estimation problem based on the eikonal equation to demonstrate the approach effectiveness in measuring the predictive uncertainty in the PINN hypocenter estimation. Laplace approximation is arguably the simplest family of approximations for the intractable posteriors of deep neural networks. We first train the Eikonal PINN to obtain the optimised weights for predicting hypocentre location. Next, we approximate the covariance matrix at the optimised Eikonal PINN’s weights for posterior sampling with the Laplace approximation. The posterior samples represent various realisations of PINN’s weights. Finally, we predict the locations of hypocentre associated with those weights’ realisations to investigate the uncertainty propagation that comes from the weights’ realisations. The uncertainties estimation from this approach is called predictive uncertainty or, simply, forward modelling uncertainty in the context of PINN.

Laplace Approximation for Bayesian PINNs

Bayesian framework for PINNs can be formulated through unnormalized Bayes’ Theorem (Yang et al., 2021) as

$$p(\theta|D) \propto p(D|\theta)p(\theta) \approx \exp\left(-\mathcal{L}(D; \theta)\right),$$

where $\theta$ denotes the learnable PINN’s parameters and $D$ represents the dataset associated with PINN’s training e.g., observed data collected by physical receivers. The last term in equation (1) is known as the Gibbs distribution. We can transform the Gibbs distribution in equation (1) into a similar representation of the loss function in a deterministic setting by reformulating it in the log-posterior as follows:

$$\log p(\theta|D) \propto -\log p(D|\theta) - \log p(\theta) \approx \mathcal{L}(D; \theta).$$

By minimising equation (2), we obtain the Maximum-A-Posteriori (MAP) solution that we consider as the centre of our Laplace approximation. Laplace approximation use a second-order expansion (Taylor expansion) of $\mathcal{L}(D; \theta)$ around $\theta_{\text{MAP}}$ to approximate $p(\theta|D)$. We consider

$$\mathcal{L}(D; \theta) \approx \mathcal{L}(D; \theta_{\text{MAP}}) + \frac{1}{2}(\theta - \theta_{\text{MAP}})^T \nabla^2_{\theta}\mathcal{L}(D; \theta_{\text{MAP}})(\theta - \theta_{\text{MAP}}),$$

and identify the Laplace approximation for $p(\theta|D)$ as

$$p(\theta|D) \approx \mathcal{N}(\theta_{\text{MAP}}, \Sigma), \quad \text{with} \quad \Sigma = -\left(\nabla^2_{\theta}\mathcal{L}(D; \theta_{\text{MAP}})\right)^{-1}.$$

Note that a naive implementation of the covariance matrix in equation (4) is infeasible, and it scales quadratically with the number of learnable PINN’s parameters, $\theta$. This work focuses on the diagonal...
approximation for the covariance matrix. Interested readers may refer to Ritter et al. (2018) for a good review on the scalable Laplace approximation for Bayesian neural networks. The diagonal approximation of the covariance matrix based on the Fisher information matrix $F$ can be computed efficiently using automatic differentiation. It is simply the expectation of the squared gradients with respect to the network parameters $\theta$:

$$H \approx \text{diag}(F) = \text{diag} \left( \mathbb{E} \left[ \nabla_\theta \mathcal{L}(\mathbf{D}; \theta) \nabla_\theta \mathcal{L}(\mathbf{D}; \theta)^T \right] \right) = \text{diag} \left( \mathbb{E} \left[ (\nabla_\theta \mathcal{L}(\mathbf{D}; \theta))^2 \right] \right),$$

where "diag" extracts the diagonal of a matrix. Note that, even if the expansion in equation (3) is accurate, this approximation will unfortunately place probability mass in low probability regions of the true posterior if some of the PINN’s parameters $\theta$ exhibit high covariance. Despite the fact, it has been used successfully in neural network weights pruning and transfer learning (Kirkpatrick et al., 2017). Based on the diagonal approximation, we can approximate our covariance by

$$\Sigma \approx H^{-1} = \frac{1}{\text{diag}(F)}.$$  

(5)

To apply the Laplace approximation for uncertainty estimation, we first minimise equation (2) to obtain the $\theta_{\text{MAP}}$. Next, we approximate the covariance matrix at $\theta_{\text{MAP}}$ and construct the Laplace approximation of the posterior distribution as in equation (4). The posterior samples represent various realisations of PINN’s weights, $\theta$. Finally, we predict the solutions associated with those weights’ realisations $\theta$ to investigate the uncertainty propagation that comes from the weights’ realisations.

**PINN for the eikonal Equation**

Now, we discuss the PINN formulation for the eikonal equation. The eikonal equation is a non-linear, first-order, hyperbolic PDE of the form:

$$|\nabla T(x)|^2 = \frac{1}{v(x)^2}, \quad \forall x \in \Omega$$

(6)

where $\Omega$ is a domain in $\mathbb{R}^d$ with $d$ as the space dimension, $T(x)$ is the travel time from the point-source $x_s$ to any point $x$, $v(x)$ is the velocity defined in $\Omega$, and $\nabla$ denotes the spatial differential operator. PINN for the eikonal equation can be formulated as an optimisation problem to optimise the learnable PINN’s parameters $\theta$ in approximating the eikonal solution and estimating the hypocentre. The loss function for solving eikonal PINN can be constructed using a mean-squared error (MSE) loss as:

$$\mathcal{L}(\theta) = \frac{1}{N_\mathcal{I}} \sum_{x^* \in \mathcal{I}} ||\nabla T_{\theta}(x^*)||^2 - \frac{1}{v^2(x^*)} \|^2 + \frac{1}{N_\mathcal{D}} \sum_{x \in \mathcal{D}} ||T_{\theta}(\hat{x}) - T(\hat{x})||^2,$$

(7)

where $T_{\theta}(x)$ represents the neural networks for the eikonal solution $T(x)$. The first term on the right side of equation (7) imposes the validity of the eikonal equation as in equation (6) on a given set of training points $x^* \in \mathcal{I}$, with $N_\mathcal{I}$ as the number of sampling points. The second term acts as data loss on a given set of travel time data at the receiver locations $\hat{x} \in \mathcal{D}$, with $N_\mathcal{D}$ representing the number of receivers. We minimise the loss in equation (7) to obtain a good approximation of the eikonal solution and hypocentre location.

We can transform the loss function in equation (7) into Bayesian framework as in equation (2) by reformulating it in the log-posterior form with a chosen log-prior distribution $\log p(\theta)$ that commonly acts as a regulariser in a deterministic setting. We perform the procedure described in the previous section to investigate the uncertainty propagation that comes from the PINN’s parameters realisations.
Numerical Example

This section demonstrates the proposed methodology on an educational example with a velocity model of $2 \times 3$ km that varies with depth, as illustrated in Figure 1 (a). The true Eikonal solution is computed analytically. We consider a neural network with 6 fully connected layers and 20 neurons per layer for Eikonal PINN to obtain the Eikonal solution and estimate the hypocentre. We consider \textit{swish} as the activation function. We randomly sample 2500 points in the domain and collect the travel time value at 11 receivers on the surface as data for training the Eikonal PINN. We minimise equation (2) with equation (7) as the log-likelihood term and Gaussian prior (Tikhonov regularisation). We perform the minimisation for 10,000 epochs and predict the Eikonal solution with the last epoch’s weights. In Figure 1, we observe that the Eikonal PINN solution is visually matches the analytical one. We can estimate the location of the hypocentre by taking the location of the Eikonal PINN solution where its value is minimum. Based on the results, we take the last epoch’s weights as our $\theta_{MAP}$ for the uncertainty analysis using the Laplace approximation.

To study the uncertainties propagations from the PINN’s weights to the predictive solution, we construct the Laplace approximation to the posterior distribution as in equations (4). This work considers the diagonal approximation of the covariance matrix as described in the previous section with 1,761 learnable network parameters. With this approximation at hand, we sample 1000 weights’ realisations and perform the Eikonal solution predictions by realising PINN with those respective weights. Based on those realisations, we have 1000 predicted Eikonal solutions, and from those solutions, we obtain the predicted hypocentre locations. The results for this predictive uncertainty is illustrated in Figure 2.

**Figure 1** Minimising equation (2) to obtain $\theta_{MAP}$. (a) Velocity model and real hypocentre location denoted by a black star. (b) The analytical Eikonal solution. (c) The Eikonal PINN solution with $\theta_{MAP}$.

**Figure 2** Predictive uncertainty of the locations of the hypocentre associated with weights’ realisations $\theta$ from Laplace approximation. (a) The locations of hypocentre associated with 1000 $\theta$ realisations from the Laplace approximation denoted by white stars. (b) The histogram of depth locations of the hypocentre realisations. (c) The histogram of lateral locations of hypocentre realisations.
Discussion

This work focuses on predictive uncertainties or, simply, forward modelling uncertainties in the context of PINN. This work is different from the physical model uncertainty, in which the quantity of interest is the physical quantity, e.g., hypocentre locations, velocity, etc. Our quantity of interest here is the PINN’s network parameters \( \theta \). In the previous section, we demonstrated the propagation of uncertainties from the PINN’s weights \( \theta \) to the predicted solutions. Based on Figure 2, we observe that the predictions of Eikonal solution and the hypocentre location vary significantly with different PINN’s weights realisation. This shows that the uncertainty in the PINN’s weights propagates into the predictive solution and significantly influences the prediction. In Figure 2, we also observe the uncertainty of the hypocentre is larger in the depth direction as we use surface recordings. However, this is not the case because the predictive uncertainty depends on the loss landscape of the PINN. A different loss landscape where the MAP solution lands on will give a significantly different predictive uncertainty. For example, we could see two of the predicted hypocentre locations are completely far away from the true location and no longer reflect the physical constraint. This shows us that the predictive uncertainty is sensitive to the PINN’s loss landscape. Interested readers may refer to Li et al. (2017) to learn more about the loss landscape of neural networks. The example shown here is meant to highlight critical features of this predictive uncertainty. We will show applications to realistic models in the presentation of this work.

Conclusions

In this work, we introduced an approximate Bayesian framework for estimating uncertainties in Eikonal Physics-Informed Neural Network (PINN). This work focuses on predictive uncertainty or, simply, forward modelling uncertainty in the context of PINN. We investigated the uncertainties propagation from the random realisations of Eikonal PINN’s weights and biases using the Laplace approximation to the predicted solutions. Our work shows that the predictions of Eikonal solution and hypocentre location vary significantly with different PINN’s weights realisation. This opens up new pathways in investigating the training dynamics of PINN, especially in the PINN’s weights initialisation to obtain a correct solution.

References