Observing PINNs behavior through multiple uniform random number of samples

Quy Le

Tampere University

24th January 2025

Introduction

- In my previous implementation of PINNs for simulating Damped Harmonic Oscillator, the model failed to capture the movement of the string.
- In this presentation, I address why it failed and what improvements I have done.
- I also trained PINNs with different random samples to demonstrate the dependence of PINNs (and neural network in general) on number of training samples.

Last week PINNs

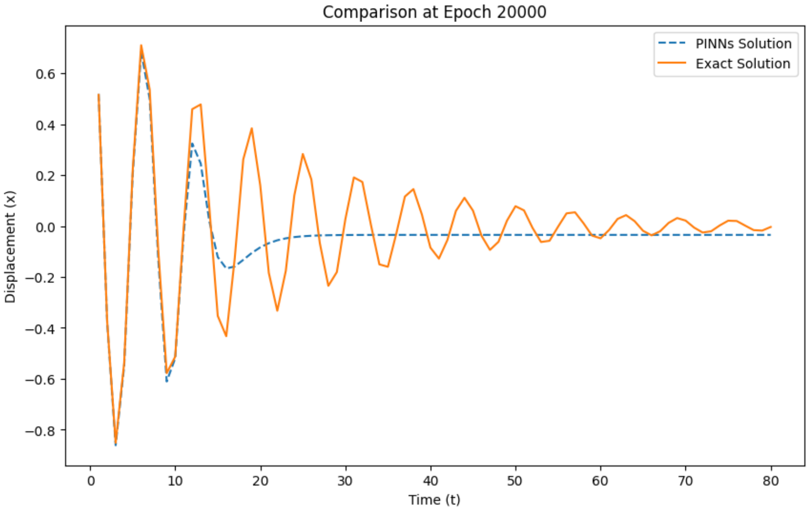

Figure 1: Last week's result.

Improvements

- Reduce the frequency of the oscillation
- Reduce the time frame
- Change optimizer to LFBGS

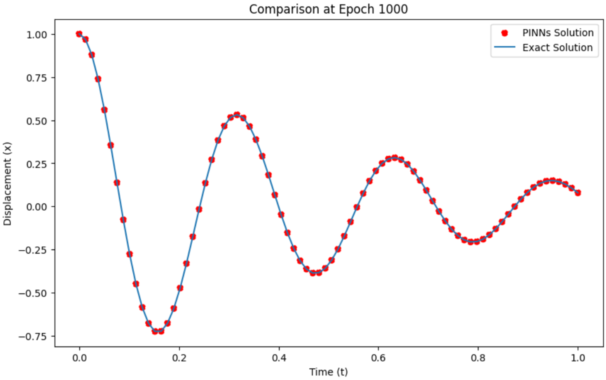

Figure 2: This week's result.

PINNs' dependence on number of samples

Hypothesis:

- The larger the number of samples, the better the model behaves
- Due to random samplings, a larger number of samples does not ensure good performance
- If points cover the time frame well, or to some extend linearly separated, the model will perform well.

Approach:

- 1. Use uniformly random distribution to sample points across the domain
- 2. Train the model for each number of samples and note the loss values E
- 3. Compare the loss E on different number of samples from 0 to 100.

Model training

Training Data:

- 1 point at t=0 for boundary condition
- n random points

Loss function:

$$\begin{aligned}\mathcal{L}(\theta) =& (u_{\mathrm{PINN}}(t=0;\theta)-1)^2 \\ & + \lambda_1 \left(\frac{d\, u_{\mathrm{PINN}}}{dt}(t=0;\theta)-0\right)^2 \\ & + \frac{\lambda_2}{N}\sum_i^N \left(\left[m\frac{d^2}{dt^2}+\mu\frac{d}{dt}+k\right] u_{\mathrm{PINN}}(t_i;\theta)\right)^2\end{aligned} \tag{1}$$

where $\lambda_1 = 0.1$ and $\lambda_2 = 0.01$.

Training loop:

- epochs = 5000
- learning rate = 0.1
- optimizer = LBFGS

Step 1: Uniformly random distribution

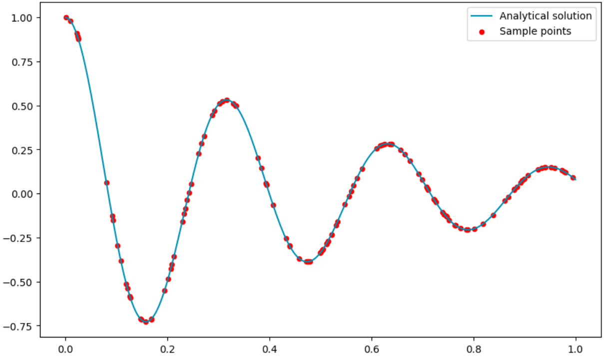

Figure 3: Example uniformly random sampling.

Step 2: Train with different number of samples

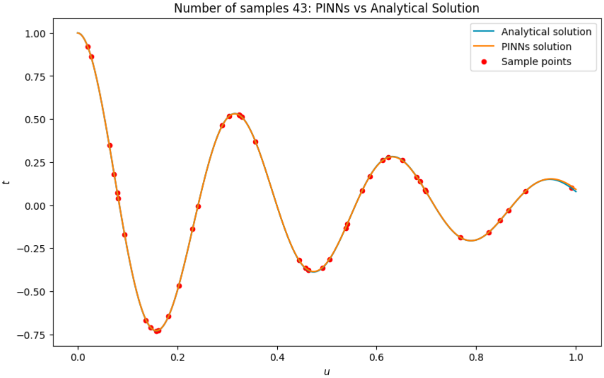

Figure 4: Model output for 43 data points.

Step 3: Observed error term E over number of samples

This verifies the first hypothesis.

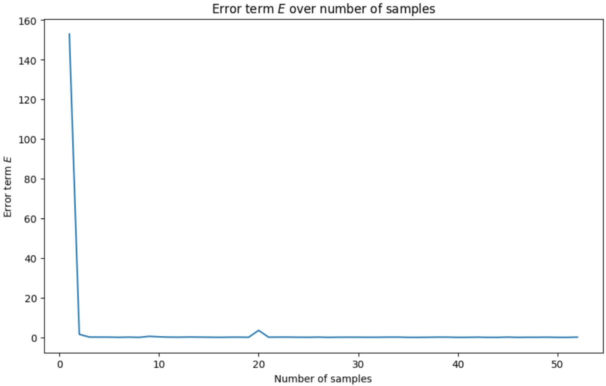

Figure 5: Error term E over number of samples.

Verify second hypothesis

This verifies the second hypothesis: Due to random samplings, a larger number of samples does not ensure good performance.

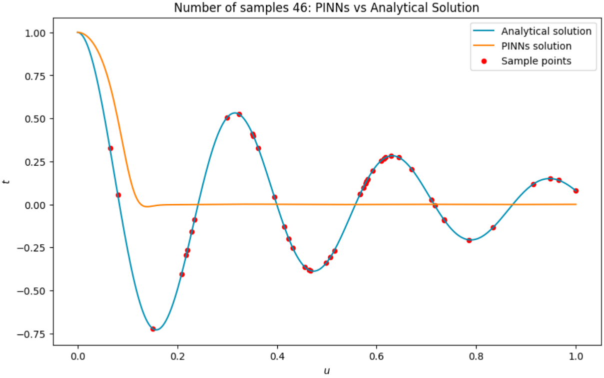

Figure 6: Model output for 46 data points.

Verify third hypothesis

This verifies the third hypothesis: If points cover the time frame well, or to some extend linearly separated, the model will perform well.

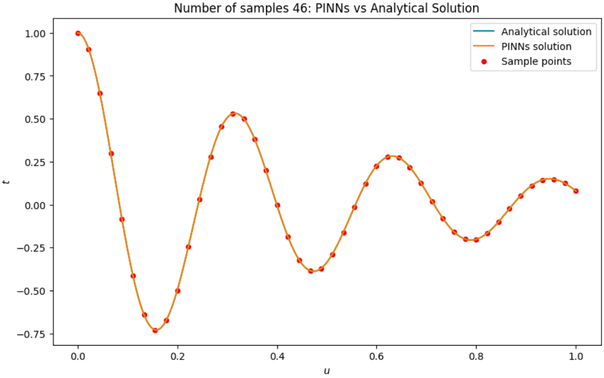

Figure 7: Model output for 46 data points.

Conclusion

Conclusion

- With increasing number of samples, the model gets better
- Distribution of data points plays huge impact on PINNs
- PINNs is powerful: they perform well even only train on physics loss from PDEs and boundary condition.

What I learn this week

- How to identify what go wrong in a neural network model.
- Train neural network on GPU
- Practicing writing on LaTex
- Work load: 12 hours.