HW1

1. Goal:

   Use a neural network to approximate Runge function.

2. Method:

   (1) Dataset: Random uniform sampling in [-1, 1]. Training : Validation = 80:20.

   (2) Model architecture:

      (a) Numbers of layer: 4

      (b) Hidden size: [64, 64]

      (c) Activation: tanh

   (3) Training:

      (a) Loss function: MSE Loss

      (b) Optimizer: Adam optimizer

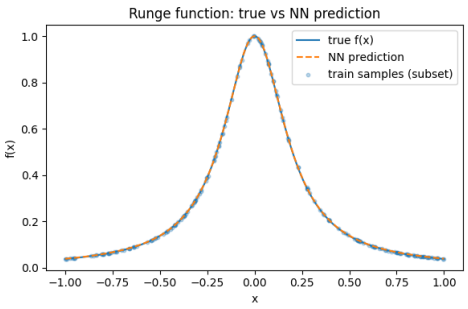
      (c) Batch size: 64

      (d) Epochs: 500

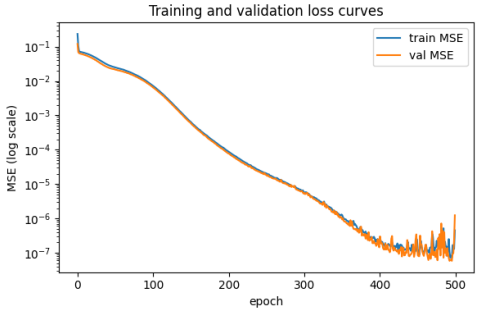
      (e) Learning rate: 1e-4

3. Result:

   (1) The true function and the neural network prediction:



   (2) The training/validation loss curves:



(3) Compute and report errors (MSE or max error):

```
ers/user/Desktop/機器學習/模型訓練/training.py
Epoch 1/500 train MSE=0.234463 val MSE=0.119959
Epoch 25/500 train MSE=0.047770 val MSE=0.041003
Epoch 50/500 train MSE=0.025237 val MSE=0.021999
Epoch 75/500 train MSE=0.016324 val MSE=0.014578
Epoch 100/500 train MSE=0.007498 val MSE=0.006778
Epoch 125/500 train MSE=0.002404 val MSE=0.002200
Epoch 150/500 train MSE=0.000665 val MSE=0.000627
Epoch 175/500 train MSE=0.000227 val MSE=0.000203
Epoch 200/500 train MSE=0.000091 val MSE=0.000079
Epoch 225/500 train MSE=0.000040 val MSE=0.000036
Epoch 250/500 train MSE=0.000021 val MSE=0.000019
Epoch 275/500 train MSE=0.000011 val MSE=0.000010
Epoch 300/500 train MSE=0.000006 val MSE=0.000005
Epoch 325/500 train MSE=0.000003 val MSE=0.000002
Epoch 350/500 train MSE=0.000001 val MSE=0.000001
Epoch 375/500 train MSE=0.000001 val MSE=0.000001
Epoch 400/500 train MSE=0.000000 val MSE=0.000000
Epoch 425/500 train MSE=0.000000 val MSE=0.000000
Epoch 450/500 train MSE=0.000000 val MSE=0.000000
Epoch 475/500 train MSE=0.000000 val MSE=0.000000
Epoch 500/500 train MSE=0.000000 val MSE=0.000001

Final validation results:
Validation MSE: 0.00000124
Validation max absolute error: 0.00164586
```

4. Discussion:

The result has shown that the model is completely approximate the Runge function. In the first plot, the NN prediction line is similar to the true Runge function line. In the loss curve, since the validation loss closely tracked the training loss, there is no obvious overfitting. Both of them decrease steadily from 0 to 400, and fluctuate from 400 to 500 with small amplitude, showing that the optimizer and learning rate are appropriate. In the last picture, MSE=0.00000124, indicating that the network achieves a good fit.