Report of Programming assignment

Method

- Data

  Sample 2000 equally spaced points on $[-1,1]$ and compute targets by $y = 1/(1 + 25x^2)$ and its derivative $dy = -50x/(1 + 25x^2)^2$. Split the dataset into 80% training and 20% validation.

- Network architecture

  Use a connected feedforward network: two hidden layers with 20 units each, tanh activation, input dimension of 1 and an output dimension of 2, corresponding to the function value and the derivative value, respectively.

- Loss function

  The loss is calculated using mean squared error (MSE).

- Training setup

  Use the Adam optimizer with a learning rate of 0.001 and train for a total of 600 iterations.

Figures and results

- Figure 1 shows both the true Runge function and its derivative and the network prediction nearly overlapping across [-1,1].
- Figure 2 presents the total training and validation losses as a function of epoch. The losses decrease smoothly and no substantial overfitting is observed throughout training.
- Figure 3 separates the components of the total loss into function loss

and derivative loss. Both components decrease steadily, which demonstrates that the model simultaneously learns to approximate the target function and its derivative with increasing accuracy.

Figure 1

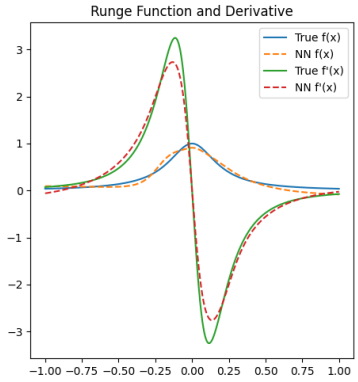


Figure 2

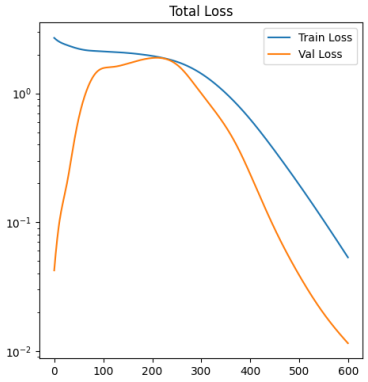
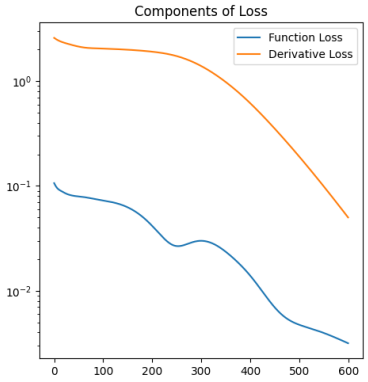


Figure 3



MSE of f(x): 1.52398e-01, Max Error of f(x): 9.66087e-01

MSE of f'(x): 3.40671e+00, Max Error of f'(x): 5.63267e+00