Report of Task1

Letian Yang
Shanghai Jiao Tong University

In this task we design a simple MLP to solve a regression task.

Problem Definition: The problem is defined as fitting a function using a statistic model. The targetted function is simply $f_0(x) = x^3$, but we would inject noises into the model, so the actual function our model can access is $f(x) = x^3 + N$ where N is the noise term.

In this task we assume $N \sim \mathcal{N}(0, \sigma^2)$ is a Gaussian, while σ is to be determined.

Solution and Performance: We use a simple *multi-layer perceptron* (MLP) with the architecture shown in Fig.1. ReLU is used for activation.

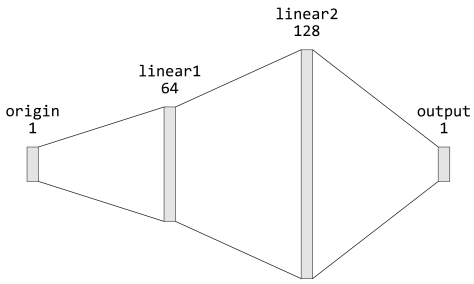
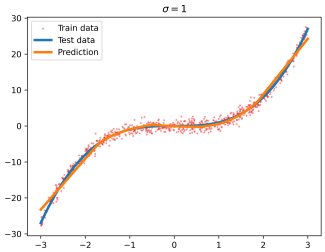


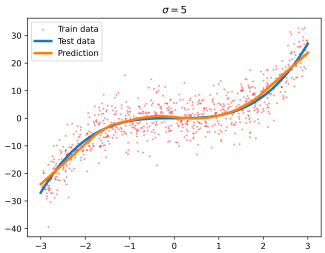
Figure 1. Architecture of MLP

In training, we use `jt.nn.SGD` as our optimizer, with learning rate 10^{-3} and weight decay of 10^{-4}. We train multiple models under different values of σ and demonstrate their performances. From Fig.2 it is displayed that the MLP as a statistic model is very robust to noises. Even if the σ was as high as 20, almost eliminating any obvious relationships between x and $f(x)$, the MLP succeeded in minimizing the mean square error, and depicts a rough trail of $f(x)$ with respect to x.

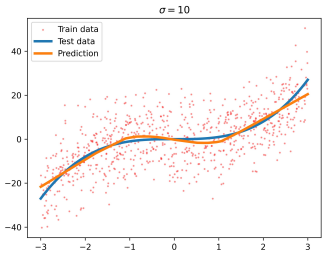
Since this task is simply a preparation for later tasks, we would like to stop early here. I got familiar with the process of training a neural network using `jittor`, which was a great gain.



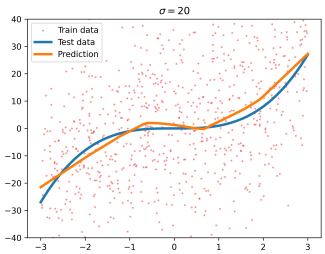
(a) $\sigma = 1$



(b) $\sigma = 5$



(c) $\sigma = 10$



(d) $\sigma = 20$

Figure 2. Performances under different σ values.