HW5 ECE542

Ming Dai
unityID: mdai3

March 25, 2017

Q1

Q2

data set setting:
dist = 5.0, width = 6, radius =10
train samples = 2000, test samples = 2000

experiments

- training method
  - gradient descent
  - conjugate descent
  - Levenverg-Marquart
- hidden neuron numbers
  - 5
  - 20

Result

- testing error
- training time (both epochs and real clock time)
- Repeat measure time = 5
- convergence criterion

Q3

Repeat mearsure number = 5

(a)

hidden number =5

plot:

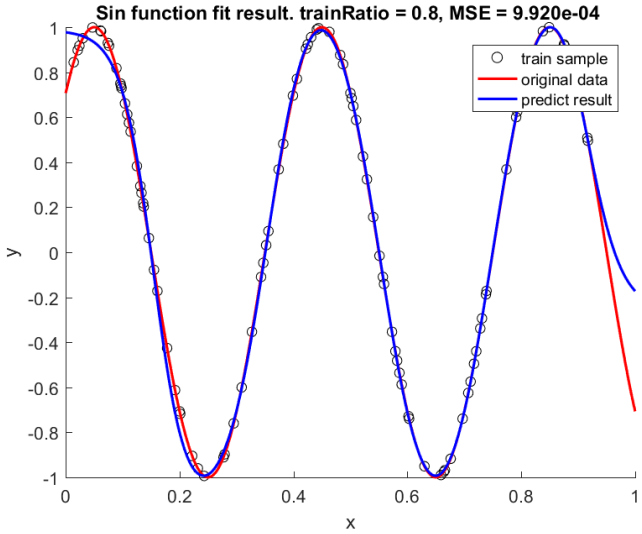

Figure 1: hidden number =5

convergence and termination criteria:

The gradient is less than the threshold 1e-7.

(b)

hidden number =20

plot:

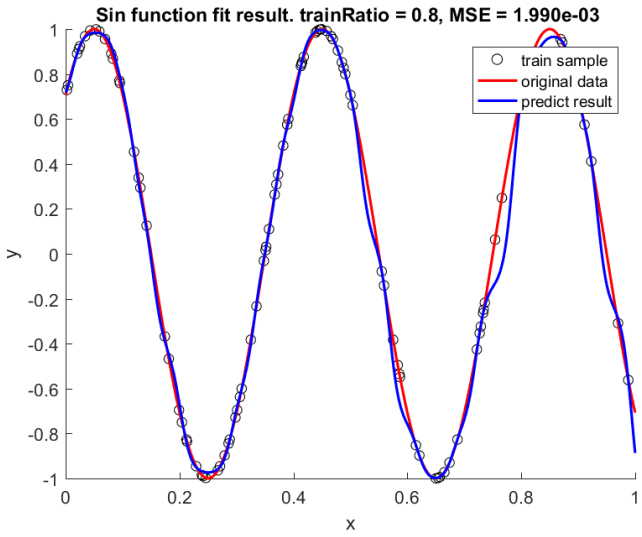

Figure 2: hidden number =20

convergence and termination criteria:

Reach the gaol of minimum mse of the **training** error.

(c) Comment

Apparently, with more hidden neurons, the fitting of the sin function is better. And with more training samples, the training error will be smaller , but the testing error may be larger due to the overffiting effect.

So when the training sample is not enough, increase the complexity of model can not always give us a better result.

Q4

Q5

Q6

Q7