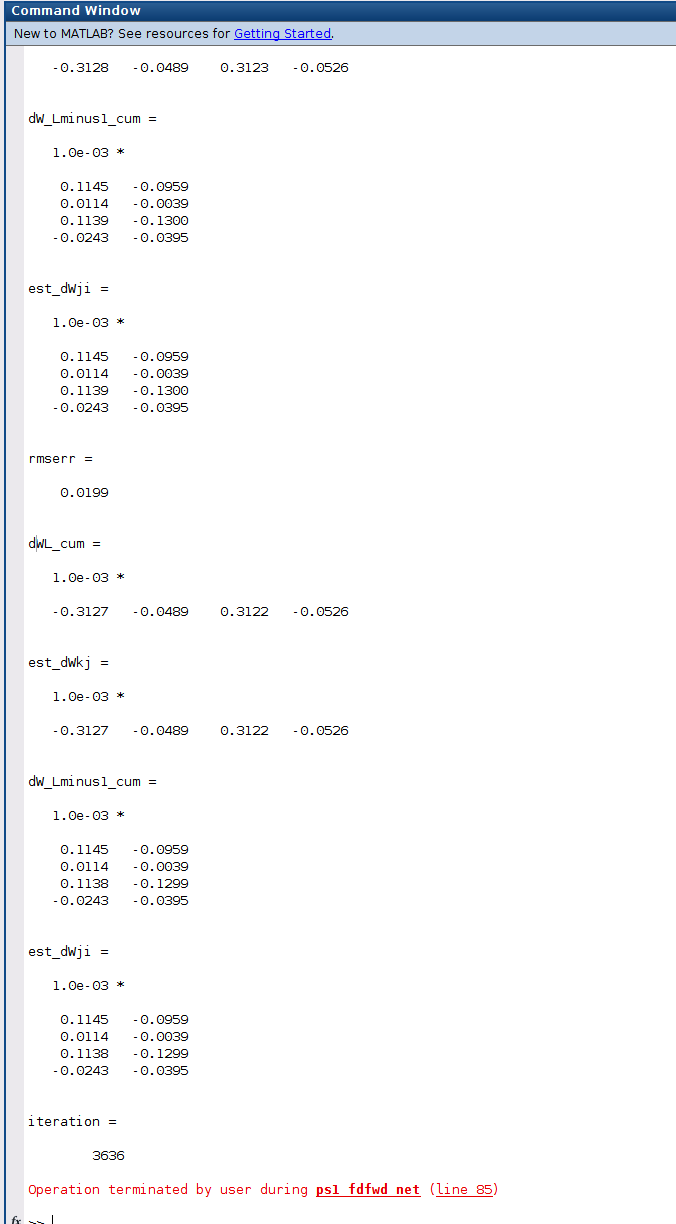
# Prove that your analytic computation matches the values of these estimates:



# Functional Code:



# Plots

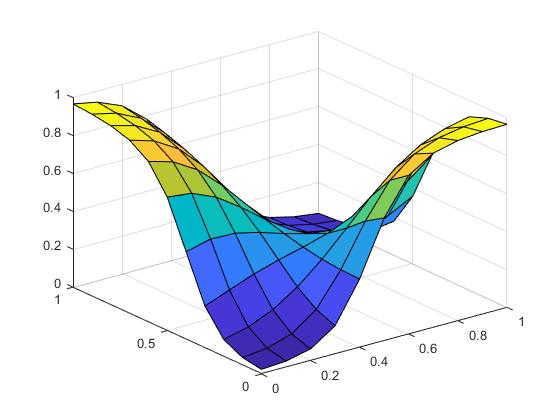


Figure 1 EPS=1; Layer =10; Iteration1515

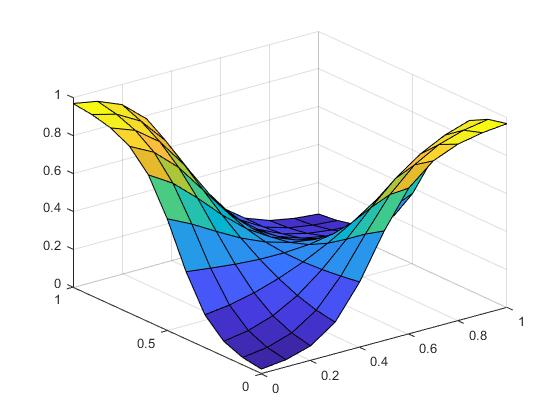
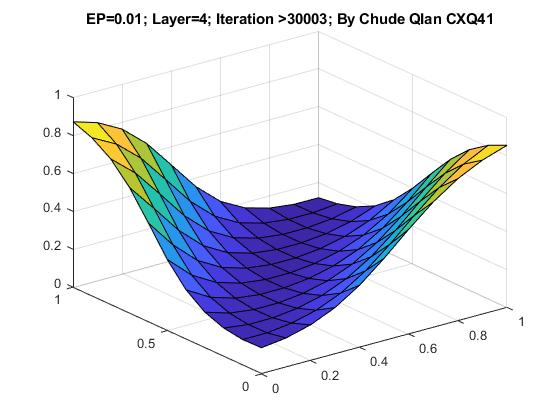
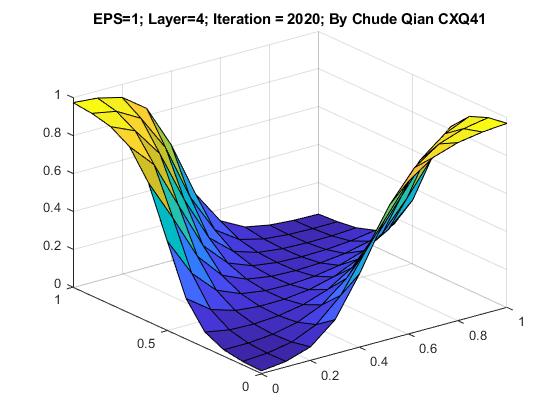
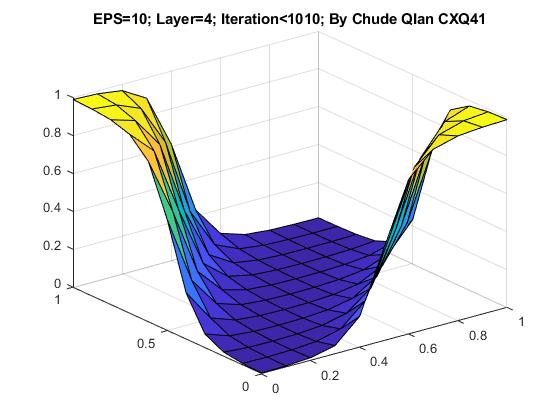


Figure 2EPS=1; Layer =40; Iteration1212



# Evaluate the influence of number of interneurons and choice of epsilon for gradient-descent computations:

I found that when Eps is small, it will cause the iteration needs to be hugely increase. As above example showed, the decrease of 100 times will result more than 100 times the iteration runs. However, then EPS is large, the iteration runs fast, but if we don’t stop it on time, we will run into issue of over learning like above EPS=10 Example.

Furthermore, the increase of neurons made the graph smoother even without running a lot of iterations. However, over exaggerated of neurons will cause a reverse effect of not able to “learn” at all. (use example of neurons of 400).