Homework #2 solutions

Prayag
Neural networks and learning systems-I

March 4, 2019

Problem 4.3.

Solution. Let $\alpha = -\beta^2$ for some $0 < \beta \leq 1$. Substituting this in equation $\Delta w_{ji}(n) = -\eta \sum_{t=0}^{n} \alpha^{(n-t)} \frac{\partial E}{\partial w_{ji}(t)}$ we get

$$\Delta w_{ji}(n) = -\eta \sum_{t=0}^{n} \left(-\beta^2\right)^{(n-t)} \frac{\partial E}{\partial w_{ji}(t)} \tag{1}$$

The term $\left(-\beta^2\right)^{(n-t)}$ alternates in sign in consecutive iterations. Now there are two cases, if the sign of the gradient of E remains same ($+$ or $-$) equation (1) results in slower convergence. If the sign of the gradient alternates, equation (1) results in large positive or negative values resulting faster convergence. ∎

Problem 4.4.

Solution. With $0 < \alpha \leq 1$ we know that

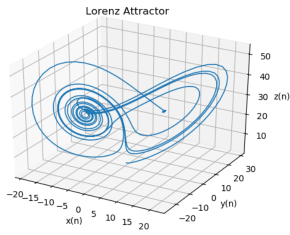
$$\Delta w_{ji}(n) = -\eta \sum_{t=0}^{n} \alpha^{(n-t)} \frac{\partial E}{\partial w_{ji}(t)}. \tag{2}$$

Now by differentiating $E(w) = k_1 \left(w - w_0\right)^2 + k_2$ with respect to w we get

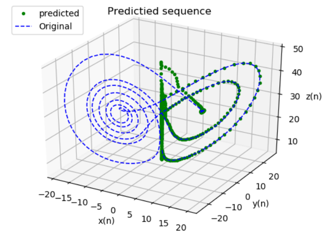
$$\Delta w_{ji}(n) = -\eta 2 k_1 \sum_{t=0}^{n} \alpha^{(n-t)} \left(w - w_0\right). \tag{3}$$

Now depending on the sign of the term $(w - w_0)$, one can comment on the convergence of the update equation. ∎

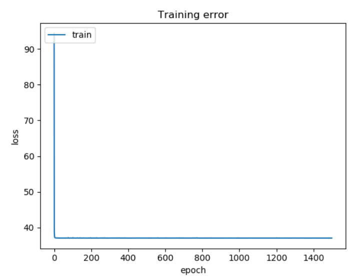
Problem 4.19.



(a)



(b)



(c)

Figure 1

Solution. Figure 1a shows the sequence generated by the solving the Lorenz dynamical system. The network parameters are as follows: Number of layers = 1, number of hidden neurons=200, number of epochs=1500, and activation function = ReLu. Figure 1b shows the predicted sequence along with the original sequence. Figure 1c shows the loss during training with epochs. ■

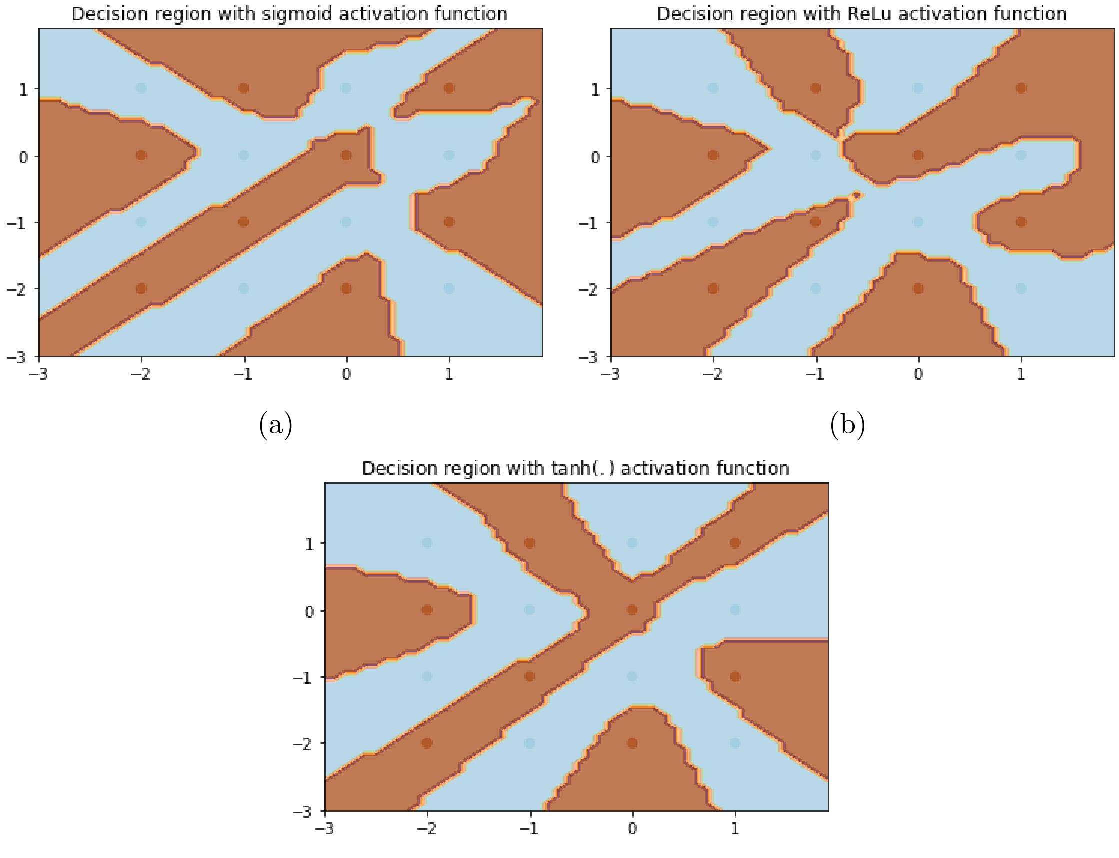
Problem 3.

Solution. (a) The parameters used in the experiment are as follows: $D = 1$, activation function is sigmoid, number of hidden neurons=10, and the maximum number of iterations = 50000. The decision boundary is shown in Figure 2a.

(b) Experiments with different activation functions.

- ReLU: $D = 1$, number of hidden neurons=30, and the maximum number of iterations = 50000. The decision boundary is shown in Figure 2b.
- tanh(.): $D = 1$, number of layers = 2, number of hidden neurons in layer 1 = 20, number of hidden neurons in layer 2=10, and the maximum number of iterations = 5000. The decision boundary is shown in Figure 2c.

(c) The value of D is varied from 0.1 to 0.9 in steps of 0.1. The number of hidden neurons is also noted and Figure 3a shows the variation of number of hidden neurons with respect to D. The decision boundaries corresponding to different D are shown in Figures 3b-3j. An important observation is that as the value of D decreases the decision boundary



(a) (b)

(c)

Figure 2

should have a large bend (curvature) indicating increasing complexity of the decision boundary. This results in more number of hidden neurons in the hidden layer.

■

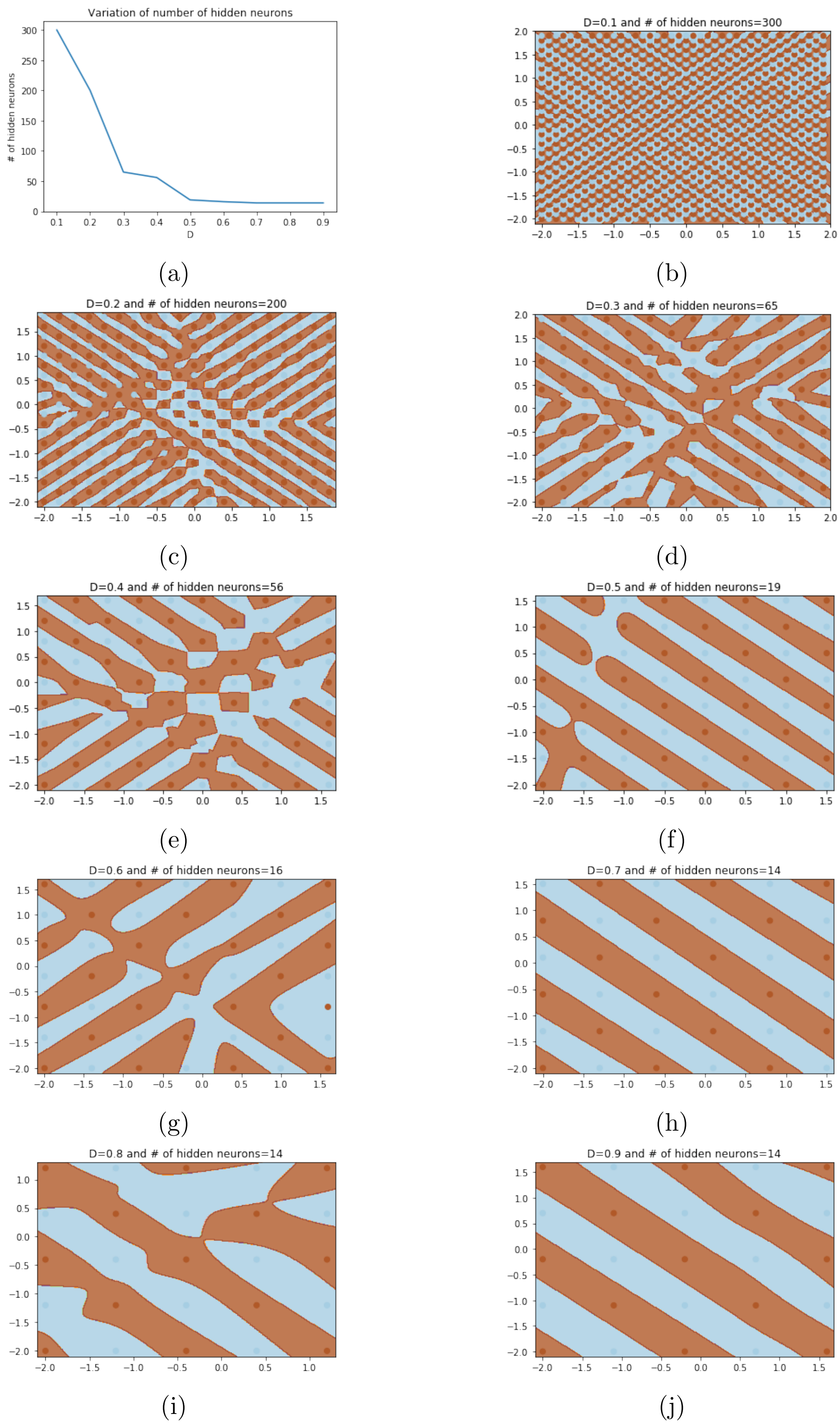

Variation of number of hidden neurons
# of hidden neurons
D
D=0.1 and # of hidden neurons=300
D=0.2 and # of hidden neurons=200
D=0.3 and # of hidden neurons=65
D=0.4 and # of hidden neurons=56
D=0.5 and # of hidden neurons=19
D=0.6 and # of hidden neurons=16
D=0.7 and # of hidden neurons=14
D=0.8 and # of hidden neurons=14
D=0.9 and # of hidden neurons=14
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)

Figure 3