Neural Network Theory and Applications
Homework Assignment 1

February 27, 2017
Due at March 12, 2017

Problem one

One variation of the perceptron learning rule is

$$\begin{aligned} \mathbf{W}^{new} &= \mathbf{W}^{old} + \alpha \mathbf{e}\mathbf{p}^T \\ \mathbf{b}^{new} &= \mathbf{b}^{old} + \alpha \mathbf{e} \end{aligned}$$

where α is called the learning rate. Prove convergence of this algorithm. Does the proof require a limit on the learning rate? Explain.

Problem two

We have a classification problem with three classes of input vectors. The three classes are

$$\begin{aligned} \text{class 1} \quad &: \quad \left\{ \mathbf{p}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \mathbf{p}_3 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\} \\ \text{class 2} \quad &: \quad \left\{ \mathbf{p}_4 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \mathbf{p}_5 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \mathbf{p}_6 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\} \\ \text{class 3} \quad &: \quad \left\{ \mathbf{p}_7 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \mathbf{p}_8 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \mathbf{p}_9 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\} \end{aligned}$$

Implement the perceptron network based on the learning rule of problem one to solve this problem. Run your problem at different learning rate ($\alpha = 1$, 0.8, 0.5, 0.2), compare and discuss the results.

Problem three

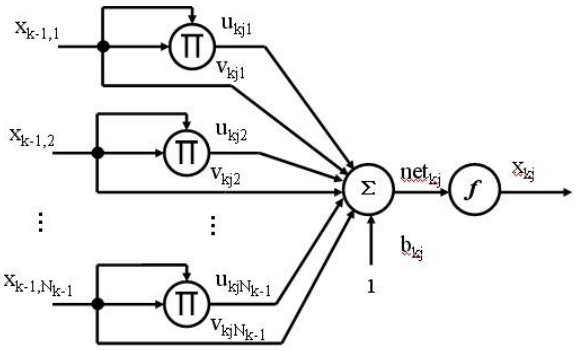
Suppose the output of each neuron in a multilayer quadratic perceptron (MLQP) network is

$$x_{kj} = f\left(\sum_{i=1}^{N_{k-1}} \left(u_{kji} x_{k-1,i}^2 + v_{kji} x_{k-1,i} \right) + b_{kj} \right) \tag{1}$$

$$\text{for } k = 2, 3, \cdots, M \text{ and } j = 1, 2, \cdots, N_k \tag{2}$$

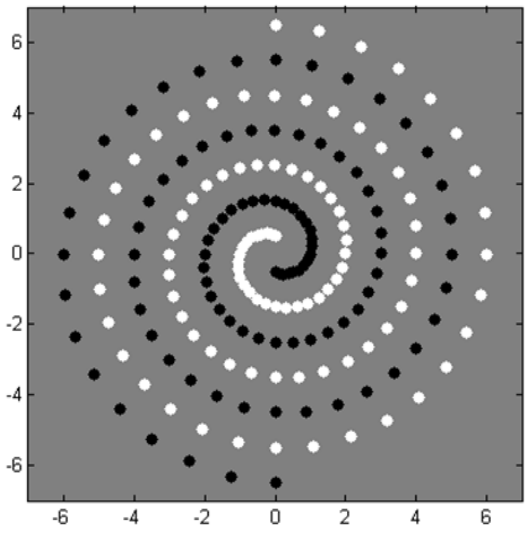
where both u_{kji} and v_{kji} are the weights connecting the ith unit in the layer $k-1$ to the jth unit in the layer k, b_{kj} is the bias of the jth unit in the layer k, N_k is the number of units in the k ($1 \leq k \leq M$), and $f(.)$ is the sigmoidal activation function.

The structure of the unit is shown as the following figure.



1. Please derive the back-propagation algorithms for MLQPs in both on-line learning and batch learning ways.

2. Write a program to implement the on-line learning back-propagation algorithms for training MLQPs with one hidden layer.



3. Run your program to classify the two-spiral problem and compare the training time and generalization performance of different learning rates (the number of hidden units can be set to 10).

The format of the attached files of two-spiral problem is as follows,

dimension 1	dimension 2	class
−0.0978	0.5476	1
0.0978	−0.5476	0

Training set: two_spiral_train, Test set:two_spiral_test

Submission

Format

All the homework should be uploaded to
ftp://bcmi.sjtu.edu.cn:2122/homework/hk1/

The submission file name should be unified. For example, if your name is Bill and your student ID is 1060339001, then you should pack all your submission files, including solutions and source code, in one .rar file with the file name "Bill_1060339001_hw1.rar" (_ is a underline). Chinese student can use Chinese name in that file name.

Please carefully check the final version of your homework. **Once you submit it, you can not modify it anymore**.

Grading policy

The basic rule is, **if any cheating activities are identified, both participants will be given ZERO scores for that homework**.

Grading

- Submission with excellent codes and report, will be given A+.
- Submission with both good codes and report, will be given A.
- Submission with some deficiencies in codes or report, and basically no big problems, will be given B+ or A-.
- Thought submission is not good enough, it can be given at least B, if we can identify that you have tried your best with independent and hard work.

Late submission

- Homework is scored with full credit on the due date.
- It is scored with one less grade for late submission, i.e., your late submission will degraded from A to B, or from B to C, if you miss the due date.