第二次期末大作业

本次作业是期末第二次大作业，该作业包含两道大题，请大家运用所学的知识独立完成。提交作业的时候需要提交一份报告（**PDF** 格式）和代码，禁止抄袭。

题目

Several exercises will make use of the following three-dimensional data sampled from three categories, denoted ω_i.

sample	ω_1			ω_2			ω_3		
	x_1	x_2	x_3	x_1	x_2	x_3	x_1	x_2	x_3
1	0.28	1.31	-6.2	0.011	1.03	-0.21	1.36	2.17	0.14
2	0.07	0.58	-0.78	1.27	1.28	0.08	1.41	1.45	-0.38
3	1.54	2.01	-1.63	0.13	3.12	0.16	1.22	0.99	0.69
4	-0.44	1.18	-4.32	-0.21	1.23	-0.11	2.46	2.19	1.31
5	-0.81	0.21	5.73	-2.18	1.39	-0.19	0.68	0.79	0.87
6	1.52	3.16	2.77	0.34	1.96	-0.16	2.51	3.22	1.35
7	2.20	2.42	-0.19	-1.38	0.94	0.45	0.60	2.44	0.92
8	0.91	1.94	6.21	-0.12	0.82	0.17	0.64	0.13	0.97
9	0.65	1.93	4.38	-1.44	2.31	0.14	0.85	0.58	0.99
10	-0.26	0.82	-0.96	0.26	1.94	0.08	0.66	0.51	0.88

1. Consider a 2-2-1 network with bias, where the activation function at the hidden units and the output unit is a sigmoid $y_j = a \tanh(b\, net_j)$ for $a = 1.716$ and $b = 2/3$.

(a) Suppose the matrices describing the input-to-hidden weights (w_{ji} for $j = 1, 2$ and $i = 0, 1, 2$) and the hidden-to-output weights (w_{kj} for $k = 1$ and $j = 0, 1, 2$) are, respectively,

$$\begin{pmatrix} 0.5 & -0.5 \\ 0.3 & -0.4 \\ -0.1 & 1.0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1.0 \\ -2.0 \\ 0.5 \end{pmatrix}.$$

The network is to be used to place patterns into one of two categories, based on the sign of the output unit signal. Shade a two-dimensional $x_1\, x_2$ input space ($-5 \le x_1, x_2 \le +5$) black or white according to the category given by the network.

(b) Repeat part(a) with the following weight matrices:

$$\begin{pmatrix} -1.0 & 1.0 \\ -0.5 & 1.5 \\ 1.5 & -0.5 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0.5 \\ -1.0 \\ 1.0 \end{pmatrix}.$$

2. Create a 3-1-1 sigmoidal network with bias to be trained to classify patterns from ω_1 and ω_2 in the table above. Use stochastic backpropagation to (Algorithm 1) with learning rate η = 0.1 and sigmoid as described in Eq. 1.

(a) Initialize all weights randomly in the range $-1 \le w \le +1$. Plot a learning curve — the training error as a function of epoch.

(b) Now repeat (a) but with weights initialized to be the *same* throughout each level. In particular, let all input-to-hidden weights be initialized with $w_{ji} = 0.5$ and all hidden-to-output weights with $w_{kj} = -0.5$.

(c) Explain the source of the differences between your learning curves.

$$f(net) = a \cdot \tanh(b \cdot net) = a \cdot [\frac{1-e^{b \cdot net}}{1+e^{b \cdot net}}] = \frac{2a}{1+e^{-b \cdot net}} - a \qquad \text{Eq. 1}$$

$$a = 1.716 \; and \; b = 2/3$$

Algorithm 1 (Stochastic backpropagation)

1. **begin initialize** network topology(# hidden units),**w**, criterion θ, η, $m \leftarrow 0$
2. **do** $m \leftarrow m + 1$
3. $\mathbf{x}^m \leftarrow$ randomly chosen pattern
4. $w_{ij} \leftarrow w_{ij} + \eta \, \delta_j \, x_i$; $\quad w_{jk} \leftarrow w_{jk} + \eta \, \delta_k \, y_j$
5. **until** $\nabla J(\mathbf{w}) < \theta$
6. **return w**
7. **end**

作业要求

1) 编程语言不限。
2) 作业包含一份报告（**PDF** 格式）和代码，并打包到.zip，其中 zip 文件的命名格式为学号_姓名。
3) 禁止使用深度学习与梯度下降相关的库。
4) 禁止抄袭。