Neural Networks Homework #2.

Prof: Cheng-Yuan Liou, TA: Jiun-Wei Liou

The SOM perceptron

The detail and the meaning of notations for SOM perceptron can be seen in [1]. The dataset can be obtained from the homework page, the data format for hw2pt.dat is (input1, input2) indicates the input data. The hw2class.dat shows the similarity for each data pattern pairs including the self pairs. If the value for row a and column b is 1, it means data a and data b are in the same class, else the value will be 0. The data size is 100. The required algorithm and formula can be described as following:

Algorithm 1 SOM perceptron training

---

for $m = 1; m \leq L; m++$ **do**
  Insert W^m as m-th layer of SOM perceptron
  for limited epochs **do**
    Pick any two patterns in the same class (U_1 or U_2) satisies the condition
    $(x^p, x^q) = \text{argmax}_{\{(x^i,x^j)\in U_1 \text{ or } (x^i,x^j)\in U_2\}} \|y^{(i,m)} - y^{(j,m)}\|^2$.
    The pair patterns (x^p, x^q) have the longest distance in the output space of the m-th layer among all pair patterns within the same class.

    Find the pair patterns, x^r and x^s in different classes, which satisfy
    $(x^r, x^s) = \text{argmin}_{(x^i,x^j)\in V_{1,2}} \|y^{(i,m)} - y^{(j,m)}\|^2$.
    The pair patterns (x^r, x^s) have the shortest distance in the output space of the m-th layer among all pair patterns within different class.

    Adjust W^m by
    $W^m = W^m - (\eta^{att} \frac{\partial E^{att}(x^p,x^q)}{\partial W^m} + \eta^{rep} \frac{\partial E^{rep}(x^r,x^s)}{\partial W^m})$
  end for
end for

$$
\begin{aligned}
X &= \{x^p, p = 1, \cdots, P\} \\
U_{c_i} &= \{(x^p, x^q); C(x^p) = C(x^q) = c_i\} \\
V_{c_i,c_j} &= \{(x^r, x^s); C(x^r) = c_i \neq c_j = C(x^s)\} \\
Y^m &= \{y^{p,m}, p = 1, \cdots, P\} \\
E^{att}(x^p, x^q) &= \frac{1}{2}\|y^{(p,m)} - y^{(q,m)}\|^2 \\
E^{rep}(x^r, x^s) &= -\frac{1}{2}\|y^{(r,m)} - y^{(s,m)}\|^2
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial E^{att}(x^p,x^q)}{\partial W^m} &= +\begin{bmatrix} (y_1^{(p,m)}-y_1^{(q,m)})(1-(y_1^{(p,m)})^2) \\ \vdots \\ (y_{n_m}^{(p,m)}-y_{n_m}^{(q,m)})(1-(y_{n_m}^{(p,m)})^2) \end{bmatrix} \left[y_1^{(p,m-1)},\cdots,y_{n_{m-1}}^{(p,m-1)},-1\right] \\
&\quad -\begin{bmatrix} (y_1^{(p,m)}-y_1^{(q,m)})(1-(y_1^{(q,m)})^2) \\ \vdots \\ (y_{n_m}^{(p,m)}-y_{n_m}^{(q,m)})(1-(y_{n_m}^{(q,m)})^2) \end{bmatrix} \left[y_1^{(q,m-1)},\cdots,y_{n_{m-1}}^{(q,m-1)},-1\right] \\
\frac{\partial E^{rep}(x^p,x^q)}{\partial W^m} &= -\begin{bmatrix} (y_1^{(p,m)}-y_1^{(q,m)})(1-(y_1^{(p,m)})^2) \\ \vdots \\ (y_{n_m}^{(p,m)}-y_{n_m}^{(q,m)})(1-(y_{n_m}^{(p,m)})^2) \end{bmatrix} \left[y_1^{(p,m-1)},\cdots,y_{n_{m-1}}^{(p,m-1)},-1\right] \\
&\quad +\begin{bmatrix} (y_1^{(p,m)}-y_1^{(q,m)})(1-(y_1^{(q,m)})^2) \\ \vdots \\ (y_{n_m}^{(p,m)}-y_{n_m}^{(q,m)})(1-(y_{n_m}^{(q,m)})^2) \end{bmatrix} \left[y_1^{(q,m-1)},\cdots,y_{n_{m-1}}^{(q,m-1)},-1\right]
\end{aligned}
$$

The possible stop criteria:

$$
\begin{aligned}
\min_{(x^p,x^q)\in V_{1,2}} \|y^{(p,L)}-y^{(q,L)}\|^2 &\approx 2^2\times n_L \\
\max_{\{(x^p,x^q)\in U_1 \text{ or } (x^p,x^q)\in U_2\}} \|y^{(p,L)}-y^{(q,L)}\|^2 &\approx 0
\end{aligned}
$$

1. Implement SOM perceptron for recognizing dataset to corresponding classes and briefly states your findings. One suggested parameter set: $L = 5$, $n_0 = 2$, $n_1 = n_2 = n_3 = n_4 = n_5 = 5$, $\eta^{att} = 0.01$, $\eta^{rep} = 0.1$, and number of epochs is 5000.

2. For the first fraction of epochs(for example, first 500 epochs), replace the maximal selection by random selection from same and different classes, redo the SOM perceptron, and briefly states your findings.

Notes:

1. Suggested length of your homework report is no more than 6 pages.

2. According to the derivation of energy function, the activation function selected for SOM perceptron will be sigmoid function.

References

[1] Cheng-Yuan Liou and Wei-Chen Cheng (2011), Forced Accretion and Assimilation Based on Self-organizing Neural Network, Self Organizing Maps - Applications and Novel Algorithm Design, Chapter 35 in Book edited by: Josphat Igadwa Mwasiagi, page 683-702, ISBN: 978-953-307-546-4, Publisher: InTech, Publishing date: January 2011.