Tik-61.261 Principles of Neural Computing

Raivio, Venna

Exercise 11

1. The weights of the neurons of a Self-organizing map (SOM) are updated according to the following learning rule:
$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \eta(n)h_{j,i(\mathbf{x})}(n)(\mathbf{x} - \mathbf{w}_j(n)),$$
where j is the index of the neuron to be updated, $\eta(n)$ is the learning-rate parameter, $h_{j,i(\mathbf{x})}$ is the neighborhood function, and $i(\mathbf{x})$ is the index of the winning neuron for the given input vector $\mathbf{x}$. Consider an example where scalar values are inputted to a SOM consisting of three neurons. The initial values of the weights are
$$w_1(0) = 0.5,\ w_2(0) = 1.5,\ w_3(0) = 2.5$$
and the inputs are randomly selected from the set:
$$X = \{0.5, 1.5, 2.0, 2.5, 2.75, 3.0, 3.25, 3.5, 3.75, 4.0, 4.25, 4.5\}.$$
The Kronecker delta function is used as a neighborhood function. The learning-rate parameter has a constant value 0.02. Calculate a few iteration steps with the SOM learning algorithm. Do the weights converge? Assume that some of the initial weight values are so far from the input values that they are never updated. How such a situation could be avoided?

2. Consider a situation in which scalar inputs of a one-dimensional SOM are distibuted according to the probability distribution function $p(x)$. A stationary state of the SOM is reached when the expected changes in the weight values become zero:
$$E[h_{j,i(x)}(x - w_j)] = 0$$
What are the stationary weight values in the following cases:
   (a) $h_{j,i(x)}$ is a constant for all j and $i(x)$, and
   (b) $h_{j,i(x)}$ is the Kronecker delta function?

3. Assume that the input and weight vectors of a SOM consisting of $N \times N$ units are d-dimensional and they are compared by using Euclidean metric. How many multiplication and adding operations are required for finding the winning neuron. Calculate also how many operations are required in the updating phase as a function of the width parameter σ of the neighborhood. Assume then that of $N = 15$, $d = 64$, and $\sigma = 3$. Is it computationally more demanding to find the winning neuron or update the weights?

4. The function $g(y_j)$ denotes a nonlinear function of the response y_j, which is used in the SOM algorithm as described in Equation (9.9):
$$\Delta\mathbf{w}_j = \eta y_j \mathbf{x} - g(y_j)\mathbf{w}_j.$$
Discuss the implications of what could happen if the constant term in the Taylor series of $g(y_j)$ is nonzero. (Haykin, Problem 9.1)