Neural Networks and Support Vector Machines

October 7, 2010

1. Instead of the activation function

$$g(x) = \frac{1}{1 + e^{-x}}$$

used in backpropagation, one could use $tanhx$.

   (a) Show that tanh(x) = 2g(2x) - 1

   (b) Given the previous result how should the backpropagartion formulas be changed with the tanh activation function?

2. Hand design a small multilayer network that can solve the XOR problem

x_1	x_2	y
1	1	1
-1	-1	1
1	-1	-1
-1	1	-1

   Draw the network showing the values of weights you have chosen.

3. Show that the separation distance between two classes can be expressed as

$$\rho = \frac{1}{||w_o||^2}$$

4. Bishop Problem 7.4

5. Bishop Problem 7.5

6. Show that the kernel $K(\mathbf{x}, \mathbf{x_i}) = (1 + \mathbf{x}^T\mathbf{x_i})^p$ has the unitary invariance property

$$K(\mathbf{x}, \mathbf{x_i}) = K(Q\mathbf{x}, Q\mathbf{x_i})$$

   where Q is the unitary matrix defined by

$$Q^{-1} = Q^T$$

7. For Support Vector machines what is the 'Kernel Trick'? Explain in detail.