Indian Institute of Science

CCE: Neural Networks for Signal Processing-1

Instructor: Shayan Srinivasa Garani
TA: Prayag Gowgi
Home Work #4, Spring 2017

Late submission policy: Points scored = Correct points scored $\times e^{-d}$, d = # days late

Assigned date: April 13^{th} 2017 **Due date:** April 21^{st} 2017 in class

---

PROBLEM 1: Consider the case of a hyperplane for linearly separable patterns, which is defined by the equation

$$\mathbf{w}^{\mathrm{T}}\mathbf{x} + b = 0$$

where $\mathbf{w}$ denotes the weight vector, b denotes the bias, and $\mathbf{x}$ denotes the input vector. The hyperplane is said to correspond to a canonical pair $(\mathbf{w}, b)$ if, for the set of input patterns $\{\mathbf{x}_i\}_{i=1}^{N}$, the additional requirement

$$\min_{i=1,\ldots,N} |\mathbf{w}^{\mathrm{T}}\mathbf{x_i} + b| = 1$$

is satisfied. Show that this requirment leads to a margin of seperation between the two classes equal to $\frac{2}{\|\mathbf{w}\|}$

(20 pts.)

PROBLEM 2: The Mercer kernel $k(x_i, \mathbf{x}_j)$ is evaluated over a training sample $\mathcal{T}$ of size N, yielding the $N \times N$ matrix

$$\mathbf{K} = \{k_{ij}\}_{i,j=1}^{N}$$

where $k_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$. Assume that the matrix $\mathbf{K}$ is positive in that all of its elements have positive values. Using the similarity transformation

$$\mathbf{K} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{T}}$$

where $\mathbf{\Lambda}$ is a diagonal matrix made up of eigenvalues and $\mathbf{Q}$ is a matrix made up of the corresponding eigenvectors, formulate an expression for the Mercer kernel $k(\mathbf{x}_i, \mathbf{x}_j)$ in terms of the eigenvalues and eigenvectors of matrix $\mathbf{K}$. What conclusions can you draw from the representation?

(20 pts.)

PROBLEM 3:

(a) Demonstrate that all three Mercer kernels described below

$$k(\mathbf{x}, \mathbf{x_j}) = \left(\mathbf{x}^{\mathrm{T}}\mathbf{x}_i + 1\right)^p$$
$$k(\mathbf{x}, \mathbf{x_j}) = \exp\left(-\frac{1}{2\sigma^2}||\mathbf{x} - \mathbf{x}_i||^2\right)$$
$$k(\mathbf{x}, \mathbf{x_j}) = \tanh\left(\beta_0\mathbf{x}^T\mathbf{x}_i + \beta_1\right)$$

satisfy the *unitary invariance property*:

$$k(\mathbf{x}, \mathbf{x}_i) = k(\mathbf{Q}\mathbf{x}, \mathbf{Q}\mathbf{x}_i)$$

where $\mathbf{Q}$ is a unitary matrix defined by

$$\mathbf{Q}^{-1} = \mathbf{Q}^{\mathrm{T}}$$

(b) Does this property hold in general?

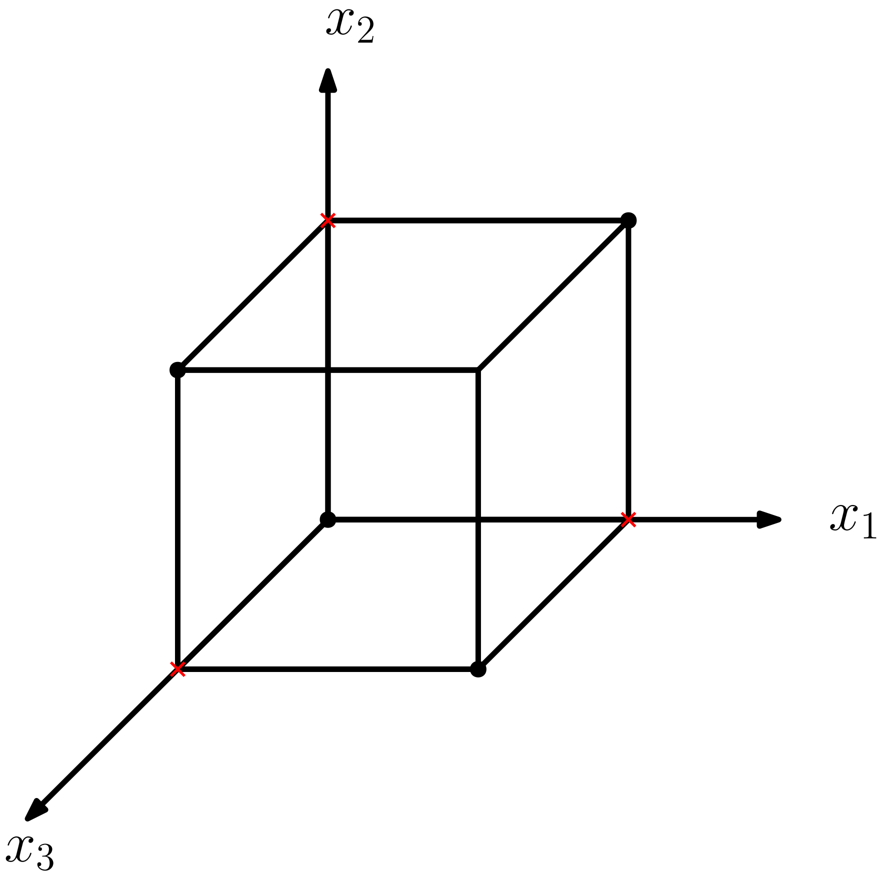

FIGURE 1. XOR operation

(10 pts.)

PROBLEM 4: Figure 1 shows the XOR function operating on a three dimensional pattern $\mathbf{x}$ described by the relationship

$$\text{XOR}(x_1, x_2, x_3) = x_1 \oplus x_2 \oplus x_3$$

where $\oplus$ denotes the XOR operation. Design a polynomial learning machine to separate the two classes. (20 pts.)

PROBLEM 5: This experiment investigates the scenario where the two moons in Figure (please refer to Figure 1.8 in Simon Haykin textbook 3rd edition, page number 61) overlap and are therefore nonseparable.

(a) Repeat the second part of the experiment in Figure 6.7 (refer to page number 290), for which the vertical separation between the two moons was fixed at $d = -6.5$. Experimentally, determine the value of parameter C for which the classification error rate is reduced to a minimum.

(b) Reduce the vertical separation between the two moons further by setting $d = -6.75$, for which the classification error rate is expected to be higher than that for $d = -6.5$. Experimentally, determine the value of parameter C for which the error rate is reduced to minimum. Comment on the results obtained for both parts of the experiment.

(20 pts.)

PROBLEM 6: Consider the extended XOR problem (refer to problem 3, Figure 1 in HW2). Implement SVM for the extended XOR problem to classify the two classes. (30 pts.)

Note: The problems 1-5 are from Simon Haykin textbook 3rd edition.