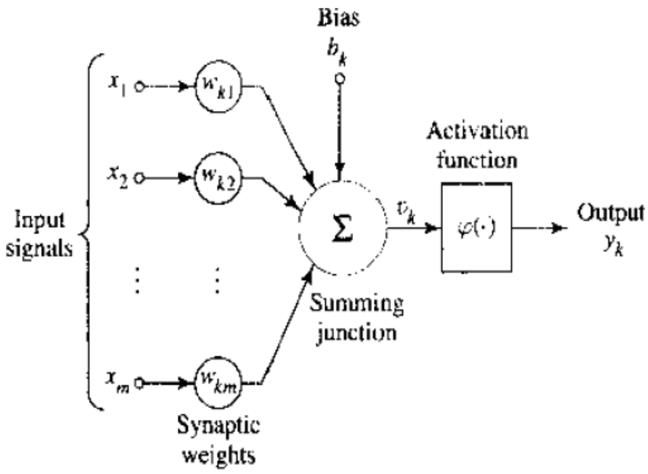
Homework #1

Only softcopy will be accepted. Please supply the code in your answer if computer experiment is involved.

Q1.



Suppose that in the signal-flow graph of the perceptron shown in the above figure, the activation function is the logistic function: $\varphi(v) = \frac{1}{1+e^{-av}}$, where v is the induced local field. The classification decision made by the perceptron is defined as follows:

Observation vector $x = [x_1 \quad x_2 \quad \cdots \quad x_m]^T$ *belongs to class C_1 if the output $y > \xi$,*

where ξ is a threshold, $0 < \xi < 1$*; otherwise, x belongs to class C_2.*

Show that the decision boundary so constructed is a hyper-plane.

Q2.

Consider the logic function, EXCLUSIVE OR (XOR).

Truth Table of XOR

x_1	0	1	0	1
x_2	0	0	1	1
y	0	1	1	0

It is well known that the XOR problem is not linearly separable. Please supply a rigorous mathematical proof for this statement.

Q3.
The perceptron could be used to perform numerous logic functions, such as AND, OR, COMPLEMENT, and EXCLUSIVE OR function, whose truth tables are tabulated as follows respectively.

x1	0	0	1	1
x2	0	1	0	1
y	0	0	0	1

AND

x1	0	0	1	1
x2	0	1	0	1
y	0	1	1	1

OR

x	0	1
y	1	0

COMPLEMENT

x1	0	0	1	1
x2	0	1	0	1
y	0	1	1	0

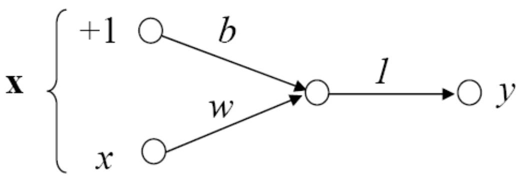
XOR

a). Demonstrate the implementation of the logic functions AND, OR, and COMPLEMENT with selection of weights by off-line calculations.

b). Demonstrate the implementation of the logic functions AND, OR, and COMPLEMENT with selection of weights by learning procedure. Suppose initial weights are chosen randomly and learning rate η is 1.0. Plot out the trajectories of the weights for each case. Compare the results with those obtained in (a).

c). What would happen if the perceptron is applied to implement the EXCLUSIVE OR function with selection of weights by learning procedure? Suppose initial weight is chosen randomly and learning rate η is 1.0. Do the computer experiment and explain your finding.

Q4.
Single layer perceptron can be used to fit a linear model to a set of input-output pairs. Suppose that we are given the following pairs:
{(0.8,-1),(1.6,-4),(3,-5),(4.0,-6),(5.0,-9)}
and a single linear neuron as shown in the following figure.



a). Find the solution of w and b using the standard linear least-squares (LLS) method. Plot out the fitting result.

b). Suppose that initial weight is chosen randomly and learning rate η is 0.02. Find the solution of w and b using the least-mean-square (LMS) algorithm for 200 epochs. Plot out the fitting result and the trajectories of the weights. Will the weights converge?

c). Compare the results obtained by LLS and the LMS methods.

Q5.

Consider that we are trying to fit a linear model to a set of input-output pairs (x(1), d(1)), (x(2), d(2)) …, (x(n), d(n)) observed in an interval of duration n, where input x is m-dimensional vector, $x = [x_1 \quad x_2 \quad \cdots \quad x_m]^T$. The linear model takes the following form:

$$y(x) = w_1 x_1 + w_2 x_2 + \cdots w_m x_m = w^T x.$$

Derive the formula to calculate the optimal parameter w* such that the following cost function J(w) is minimized.

$$J(w) = \sum_{i=1}^{n} r(i)e(i)^2 = \sum_{i=1}^{n} r(i)(d(i) - y(i))^2$$

where r(i)>0 are the weighting factors for each output error e(i).