

SOLUTIONS MANUAL

THIRD EDITION

Neural

Networks

and

Learning Machines

Simon Haykin

and

Yanbo Xue

McMaster University

Canada

CHAPTER 1

Rosenblatt's Perceptron

Problem 1.1

- (1) If $\mathbf{w}^T(n)\mathbf{x}(n) > 0$, then $y(n) = +1$.
 If also $\mathbf{x}(n)$ belongs to C_1 , then $d(n) = +1$.
 Under these conditions, the error signal is

$$e(n) = d(n) - y(n) = 0$$
 and from Eq. (1.22) of the text:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta e(n)\mathbf{x}(n) = \mathbf{w}(n)$$
 This result is the same as line 1 of Eq. (1.5) of the text.

- (2) If $\mathbf{w}^T(n)\mathbf{x}(n) < 0$, then $y(n) = -1$.
 If also $\mathbf{x}(n)$ belongs to C_2 , then $d(n) = -1$.
 Under these conditions, the error signal $e(n)$ remains zero, and so from Eq. (1.22) we have

$$\mathbf{w}(n+1) = \mathbf{w}(n)$$
 This result is the same as line 2 of Eq. (1.5).

- (3) If $\mathbf{w}^T(n)\mathbf{x}(n) > 0$ and $\mathbf{x}(n)$ belongs to C_2 we have

$$y(n) = +1$$

$$d(n) = -1$$
 The error signal $e(n)$ is -2 , and so Eq. (1.22) yields

$$\mathbf{w}(n+1) = \mathbf{w}(n) - 2\eta\mathbf{x}(n)$$
 which has the same form as the first line of Eq. (1.6), except for the scaling factor 2.

- (4) Finally if $\mathbf{w}^T(n)\mathbf{x}(n) < 0$ and $\mathbf{x}(n)$ belongs to C_1 , then

$$y(n) = -1$$

$$d(n) = +1$$
 In this case, the use of Eq. (1.22) yields

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\eta\mathbf{x}(n)$$
 which has the same mathematical form as line 2 of Eq. (1.6), except for the scaling factor 2.

Problem 1.2

The output signal is defined by

$$y = \tanh\left(\frac{v}{2}\right)$$

$$= \tanh\left(\frac{b}{2} + \frac{1}{2}\sum_i w_i x_i\right)$$

CHAPTER 4

Multilayer Perceptrons

Problem 4.1

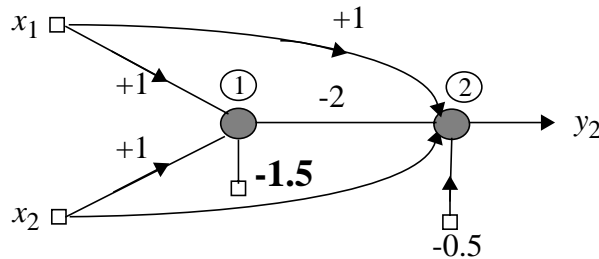


Figure 4: Problem 4.1

Assume that each neuron is represented by a McCulloch-Pitts model. Also assume that

$$x_i = \begin{cases} 1 & \text{if the input bit is 1} \\ 0 & \text{if the input bit is 0} \end{cases}$$

The induced local field of neuron 1 is

$$v_1 = x_1 + x_2 - 1.5$$

We may thus construct the following table:

x_1	0	0	1	1
x_2	0	1	0	1
v_1	-1.5	-0.5	-0.5	0.5
y_2	0	0	0	1

The induced local field of neuron ② is

$$v_2 = x_1 + x_2 - 2y_1 - 0.5$$

Accordingly, we may construct the following table:

x_1	0	0	1	1
x_2	0	1	0	1
y_1	0	0	0	1
v_2	-0.5	0.5	-0.5	-0.5
y_2	0	1	1	1

- (b) If $q(\mathbf{y}_i)$ happens to equal the source pdf $f_U(\mathbf{y}_i)$ for all i , we then find that $D_{\tilde{f}||q} = 0$. In such a case, (1) reduces to

$$h(\mathbf{Z}) = -D_{f||\tilde{f}}$$

That is, the entropy $h(\mathbf{Z})$ is equal to the negative of the Kullback-Leibler divergence between the pdf $f_{\mathbf{Y}}(\mathbf{y})$ and the corresponding factorial distribution $\tilde{f}_{\mathbf{Y}}(\mathbf{y})$.

Problem 10.20

- (a) From Eq. (10.124) in the text,

$$\Phi = \log|\det(\mathbf{A})| + \log|\det(\mathbf{W})| + \sum_i \log\left(\frac{\partial z_i}{\partial y_i}\right)$$

The matrix \mathbf{A} of the linear mixer is fixed. Hence differentiating Φ with respect to \mathbf{W} :

$$\frac{\partial \Phi}{\partial \mathbf{W}} = \mathbf{W}^{-T} + \sum_i \frac{\partial}{\partial \mathbf{W}} \log\left(\frac{\partial z_i}{\partial y_i}\right) \quad (1)$$

- (b) From Eq. (10.126) of the text,

$$z_i = \frac{1}{1 + e^{-y_i}}$$

Differentiating z_i with respect to y_i :

$$\begin{aligned} \frac{\partial z_i}{\partial y_i} &= \frac{e^{-y_i}}{(1 + e^{-y_i})^2} \\ &= z_i - z_i^2 \end{aligned} \quad (2)$$

Hence, differentiating $\log\left(\frac{\partial z_i}{\partial y_i}\right)$ with respect to the demixing matrix \mathbf{W} , we get

$$\frac{\partial}{\partial \mathbf{W}} \log\left(\frac{\partial z_i}{\partial y_i}\right) = \frac{\partial}{\partial \mathbf{W}} \log(z_i - z_i^2)$$

$$\mathbf{B}' = [\mathbf{w}_0^T \mathbf{B}, b]$$

The output $y(n)$ is defined by

$$y(n) = \phi(x(n)) \quad (3)$$

Equations (2) and (3) define the state-space model of Fig. P15.7a, assuming that its linear dynamic component is described by (1).

(b) Consider next the local output feedback system of Fig. 15.7b. Let the linear dynamic component of this system be described by (1). The output of the whole system in Fig. 15.7b is then defined by

$$\begin{aligned} x(n) &= \phi(\mathbf{w}^T \mathbf{z}(n) + b) \\ &= \phi\left([w_1, \mathbf{w}_0^T] \begin{bmatrix} x(n-1) \\ \mathbf{B}\mathbf{u}(n) \end{bmatrix} + b\right) \\ &= \phi(w_1 x(n-1) + \mathbf{B}'\mathbf{u}'(n)) \end{aligned} \quad (4)$$

where w_1 , \mathbf{w}_0 , \mathbf{B}' , and $\mathbf{u}'(n)$ are all as defined previously. The output $y(n)$ of Fig. P15.7b is

$$y(n) = x(n) \quad (5)$$

Equations (4) and (5) define the state-space model of the local output feedback system of Fig. P15.7b, assuming that its linear dynamic component is described by (1).

The process (state) equation of the local feedback system of Fig. P15.7a is linear but its measurement equation is nonlinear, and conversely for the local feedback system of Fig. P15.7b. These two local feedback systems are controllable and observable, because they both satisfy the conditions for controllability and observability.

Problem 15.8

We start with the state equation

$$\mathbf{x}(n+1) = \phi(\mathbf{W}_a \mathbf{x}(n) + \mathbf{w}_b u(n))$$

Hence, we write

$$\begin{aligned} \mathbf{x}(n+2) &= \phi(\mathbf{W}_a \mathbf{x}(n+1) + \mathbf{w}_b u(n+1)) \\ &= \phi(\mathbf{W}_a \phi(\mathbf{W}_a \mathbf{x}(n) + \mathbf{w}_b u(n)) + \mathbf{w}_b u(n+1)) \end{aligned}$$