EMS702P Statistical Thinking and Applied Machine Learning

Week 8 – PBL Solutions

Yunpeng Zhu

Solution 1:

(A) Based on the measured data, the linear classification model can be written into a matrix form as

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 0.5 & 1 \\ 1 & 0.5 \\ 2 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{1}$$

Denote

$$\mathbf{Y} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -2 \\ -2 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 0.5 & 1 \\ 1 & 0.5 \\ 2 & 2 \\ 2 & 1 \end{bmatrix}, \boldsymbol{\Theta} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{2}$$

Based on the LS method, the parameters $\boldsymbol{\Theta}$ can be evaluated as

$$\boldsymbol{\Theta} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y} \tag{3}$$

(B) In equation (3),

$$\mathbf{X}^{\mathrm{T}}\mathbf{X} = \begin{bmatrix} 0.5 & 1 & 2 & 2 \\ 1 & 0.5 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & 1 \\ 1 & 0.5 \\ 2 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 9.25 & 7 \\ 7 & 6.25 \end{bmatrix} \tag{4}$$

$$\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1} = \frac{1}{9.25 \times 6.25 - 7^2} \begin{bmatrix} 6.25 & -7 \\ -7 & 9.25 \end{bmatrix} = \frac{1}{8.8125} \begin{bmatrix} 6.25 & -7 \\ -7 & 9.25 \end{bmatrix} = \begin{bmatrix} 0.71 & -0.79 \\ -0.79 & 1.05 \end{bmatrix} \tag{5}$$

$$\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}} = \begin{bmatrix} 0.71 & -0.79 \\ -0.79 & 1.05 \end{bmatrix} \begin{bmatrix} 0.5 & 1 & 2 & 2 \\ 1 & 0.5 & 2 & 1 \end{bmatrix} = \begin{bmatrix} -0.435 & 0.315 & -0.16 & 0.63 \\ 0.655 & -0.265 & 0.52 & -0.53 \end{bmatrix} \tag{6}$$

Therefore,

$$\boldsymbol{\Theta} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y} = \begin{bmatrix} -0.435 & 0.315 & -0.16 & 0.63 \\ 0.655 & -0.265 & 0.52 & -0.53 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ -2 \\ -2 \end{bmatrix} = \begin{bmatrix} -0.82 \\ -0.37 \end{bmatrix} \tag{7}$$

(C) The evaluated second-order regression model is

$$y = 2 - 0.82x_1 - 0.37x_2 \tag{8}$$

The boundary of the classifier is achieved by $y = 0.5$, which is

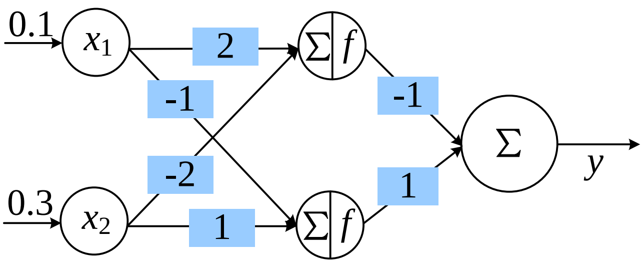
$$x_2 = (2 - 0.82x_1 - 0.5)/0.37 = 4.05 - 2.22x_1 \tag{9}$$

Solution 2:

$$x_1 + x_2 - 1 = 0 \Rightarrow \bar{\mathbf{x}}^{\mathrm{T}}\bar{\mathbf{w}} + w_0 = 0:\ [x_1 \quad x_2] \times \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 1 = 0$$

$$d = \frac{w_0 + \mathbf{x}_0{}^{\mathrm{T}}\bar{\mathbf{w}}}{\|\bar{\mathbf{w}}\|_2} = \frac{-1 + [1 \quad 1] \times \begin{bmatrix} 1 \\ 1 \end{bmatrix}}{\sqrt{1^2 + 1^2}} = \frac{\sqrt{2}}{2}$$

Solution 3:



$$\begin{cases} z_1^{[1]} = w_{1,1}^{[1]}x_1 + w_{2,1}^{[1]}x_2 = 2 \times 0.1 - 2 \times 0.3 = -0.4 \\ z_2^{[1]} = w_{1,2}^{[1]}x_1 + w_{2,2}^{[1]}x_2 = (-1) \times 0.1 + 1 \times 0.3 = 0.2 \end{cases}$$

$$\begin{cases} h_1 = f(z_1^{[1]}) = 0 \\ h_2 = f(z_2^{[1]}) = 0.2 \end{cases}$$

$$y = w_{1,1}^{[2]}h_1 + w_{2,1}^{[2]}h_2 = (-1) \times 0 + 1 \times 0.2 = 0.2$$