Determine Logistic model

Consider the cost function as

$$\begin{aligned} J(\mathbf{W}) &= -\ln\prod_{i=1}^{N}\bar{y}(i)^{y(i)}\left[1-\bar{y}(i)\right]^{1-y(i)} \\ &= -\sum_{i=1}^{N}\left\{y(i)\ln\bar{y}(i)+\left[1-y(i)\right]\ln\left[1-\bar{y}(i)\right]\right\} \end{aligned}$$

where

$$\bar{y}(i)=\frac{1}{1+\exp\left(-\mathbf{x}_i{}^{\mathrm{T}}\mathbf{W}\right)};\ y(i)=0,1$$

Denote $J_i(\mathbf{W})=-\left\{y(i)\ln\bar{y}(i)+\left[1-y(i)\right]\ln\left[1-\bar{y}(i)\right]\right\},\ i=1,\ldots,N$, thus

$$J(\mathbf{W})=\sum_{i=1}^{N}J_i(\mathbf{W})$$

The gradient $\partial J_i(\mathbf{W})/\partial\mathbf{W}$ can be obtained by

$$\frac{\partial J_i(\mathbf{W})}{\partial\mathbf{W}}=\frac{\partial J_i(\mathbf{W})}{\partial\bar{y}_i}\frac{\partial\bar{y}_i}{\partial z_i}\frac{\partial z_i}{\mathbf{W}}$$

where $z=\mathbf{x}^{\mathrm{T}}\mathbf{W}$,

$$\frac{\partial J_i(\mathbf{W})}{\partial\bar{y}_i}=-\left(\frac{y(i)}{\bar{y}(i)}-\frac{1-y(i)}{1-\bar{y}(i)}\right)$$

$$\frac{\partial\bar{y}_i}{\partial z_i}=\bar{y}(i)\left[1-\bar{y}(i)\right]$$

and

$$\frac{\partial z_i}{\partial\mathbf{W}}=\mathbf{x}_i{}^{\mathrm{T}}$$

Therefore,

$$\begin{aligned} \frac{\partial J_i(\mathbf{W})}{\partial\mathbf{W}} &= -\left(\frac{y(i)}{\bar{y}(i)}-\frac{1-y(i)}{1-\bar{y}(i)}\right)\bar{y}(i)\left[1-\bar{y}(i)\right]\mathbf{x}_i{}^{\mathrm{T}} \\ &= \left[\bar{y}(i)-y(i)\right]\mathbf{x}_i{}^{\mathrm{T}} \end{aligned}$$

and

$$\frac{\partial J(\mathbf{W})}{\partial\mathbf{W}}=\sum_{i=1}^{N}\frac{\partial J_i(\mathbf{W})}{\partial\mathbf{W}}=\sum_{i=1}^{N}\left[\bar{y}(i)-y(i)\right]\mathbf{x}_i{}^{\mathrm{T}}=\mathbf{X}^{\mathrm{T}}\left(\bar{\mathbf{Y}}-\mathbf{Y}\right)$$

where

$$\mathbf{X}=\begin{bmatrix}\mathbf{x}_1{}^{\mathrm{T}}\\ \vdots\\ \mathbf{x}_N{}^{\mathrm{T}}\end{bmatrix}$$

In consequence, the model parameters can be updated as

$$\mathbf{W}=\mathbf{W}-\lambda\mathbf{X}^{\mathrm{T}}(\bar{\mathbf{Y}}-\mathbf{Y})$$