① sigmoid:

$$y = w_1x_1 + w_2x_2 + \cdots + w_nx_n + b$$

$$y = x^T w + b$$

$$g(z) = \frac{1}{1+e^{-z}}$$

$$g(z) = h_\theta(x) = \frac{1}{1+e^{-\theta x}}$$

② 对数几率:

$$\ln\frac{h_\theta(x)}{1-h_\theta(x)} = \ln\frac{\frac{1}{1+e^{-\theta x}}}{1-\frac{1}{1+e^{-\theta x}}} = \ln\frac{1}{e^{-\theta x}} = \ln e^{\theta x} = \theta x$$

③ loss

$$P\{y=1 \mid x;\theta\} = h_\theta(x)$$

$$P\{y=0 \mid x;\theta\} = 1-h_\theta(x)$$

$$P\{y \mid x;\theta\} = h_\theta(x)^y[1-h_\theta(x)]^{1-y}$$

$$Loss(\theta) = \prod_{i=1}^{n} h_\theta(x_i)^{y_i}[1-h_\theta(x_i)]^{1-y_i}$$

$$\ln Loss(\theta) = \sum_{i=1}^{n} y_i \cdot \ln h_\theta(x_i) + (1-y_i)\cdot\ln[1-h_\theta(x_i)] \text{ (损失函数)}$$

但是利用最大似然求解是越大越好，但需要Loss越小越好

$$\therefore J(\theta) = -\frac{1}{n}\ln Loss(\theta) = -\frac{1}{n}\sum_{i=1}^{n} y_i\ln h_\theta(x_i) + (1-y_i)\ln[1-h_\theta(x_i)]$$

④ 求导

$$\frac{\partial J(\theta)}{\partial\theta} = -\frac{1}{n}\sum_{i=1}^{n} y_i\cdot\frac{1}{h_\theta(x_i)}\cdot\frac{0-1\cdot e^{-\theta x_i}\cdot(-x_i)}{(1+e^{-\theta x_i})^2} + (1-y_i)\cdot\frac{-1}{1-h_\theta(x_i)}\cdot\frac{0-1\cdot e^{-\theta x_i}\cdot(-x_i)}{(1+e^{-\theta x_i})^2}$$

$$= -\frac{1}{n}\sum_{i=1}^{n} y_i\cdot\frac{e^{-\theta x_i}x_i}{1+e^{-\theta x_i}} + (1-y_i)\cdot\frac{-1}{1-\frac{1}{1+e^{-\theta x_i}}}\cdot\frac{e^{-\theta x_i}\cdot x_i}{(1+e^{-\theta x_i})^2}$$

$$= -\frac{1}{n}\sum_{i=1}^{n} y_i\cdot\frac{e^{-\theta x_i}\cdot x_i}{1+e^{-\theta x_i}} + (y_i-1)\cdot\frac{1+e^{-\theta x_i}}{e^{-\theta x_i}}\cdot\frac{e^{-\theta x_i}x_i}{(1+e^{-\theta x_i})^2}$$

$$= -\frac{1}{n}\sum_{i=1}^{n} y_i\cdot\frac{e^{-\theta x_i}x_i}{1+e^{-\theta x_i}} + (y_i-1)\cdot\frac{x_i}{1+e^{-\theta x_i}}$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\left[y_i[1-h_\theta(x_i)]+(y_i-1)\cdot h_\theta(x)\right]\cdot x_i$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\cdot(y_i-h_\theta(x))\cdot x_i$$

$$= \frac{1}{n}\sum_{i=1}^{n}(h_\theta(x)-y_i)x_i$$

$$= \frac{1}{n}\cdot x^T\cdot[h_\theta(x)-y]$$

⑤利用梯度求导解出 w

$$w = w - lr\cdot\frac{1}{n}\cdot x^T\cdot[h_\theta(x)-y]$$

$$\because h_\theta(x)=\frac{1}{1+e^{-xw}}$$

$$w = w - lr\cdot\frac{1}{n}\cdot x^T\cdot[\mathrm{sigmoid}(x\cdot w)-y]$$

⑥带 l_2 正则化的逻辑回归

Loss: $J(\theta) = -\frac{1}{n}\sum_{i=1}^{n} y_i\ln h_\theta(x_i)+(1-y_i)\ln[1-h_\theta(x_i)]+\frac{\lambda}{2n}\sum_{j=1}^{m}\theta_j^2$

$$\frac{\partial J(\theta)}{\partial\theta} = \frac{1}{n}x^T\cdot[h_\theta(x)-y]+\frac{\lambda}{n}\theta$$

$$w = w - lr\cdot\frac{1}{n}\cdot x^T[\mathrm{sigmoid}(x\cdot w)-y]-\lambda\cdot\frac{1}{n}\cdot\theta$$