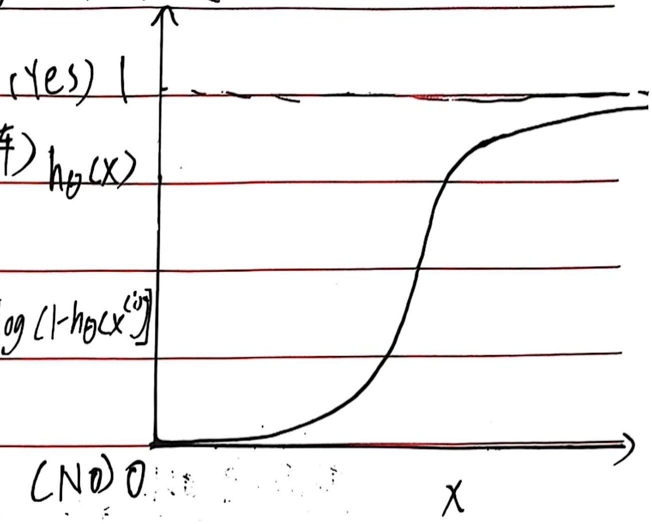
逻辑回归梯度下降

(分类) ~~f(x z =~~

假设函数为 sigmoid 函数



$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}} \qquad z = -\theta^T X \text{ (X是矩阵)}$$

单~~特~~特征(非矩阵~~型式~~形式):

损失函数 $L(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log(h_\theta(x^{(i)})) + (1-y^{(i)})\log(1-h_\theta(x^{(i)}))\right]$

$$\frac{\partial L(\theta)}{\partial\theta} = \frac{\partial L(\theta)}{\partial h_\theta}\cdot\frac{\partial h_\theta}{\partial z}\cdot\frac{\partial z}{\partial\theta}$$

$$\frac{\partial L(\theta)}{\partial h_\theta} = -\frac{1}{m}\sum_{i=1}^{m}\left[\frac{y^{(i)}}{h_\theta} - \frac{1-y^{(i)}}{1-h_\theta}\right]$$ ~~(常可忽略)~~

$$\frac{\partial h_\theta}{\partial z} = h_\theta\cdot(1-h_\theta)$$

$$\frac{\partial z}{\partial\theta} = x^{(i)}$$

往回代

$$\frac{\partial L(\theta)}{\partial\theta} = \cancel{(h_\theta - y)\cdot x}\ \frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})\,x^{(i)}$$

$$\theta_{new} = \theta_{old} - \frac{\alpha}{m}\cdot\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})\,x^{(i)}$$

矩阵形式: $\theta_{new} = \theta_{old} - \frac{\alpha}{m}X^T\cdot(h_{\vec{\theta}}(X) - \vec{y})$