对数几率回归(Logistic Regression)

- 对数几率回归是一个分类问题
  - 分类问题：离散输出，比如二分类输出0或1。
  - 回归问题：预测一个连续的输出。

Logistic回归假设函数

- if h(x) ≥ 0.5，predict = 1
- if h(x) ≤ 0.5，predict = 0

$$h\theta(x) = g(\theta^T x)$$

sigmoid函数是

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

代入，得

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

- 因为θ^T * X 是一个线性函数，所以发生于不发生概率是

$$P(y=1|x) = \frac{e^{\theta^T x}}{1+e^{\theta^T x}} = \frac{1}{1+e^{-\theta^T x}} = h_\theta(x)$$

$$P(y=0|x) = \frac{1}{1+e^{\theta^T x}} = 1 - h_\theta(x)$$

- 我们可以令y=h(x)，则

$$\theta^T x = ln\frac{y}{1-y}$$

$\frac{y}{1-y}$称为几率，反映了x作为正例的相对可能性.，取对数得到

$$ln\frac{y}{1-y}$$

$$ln\frac{p(y=1|x)}{p(y=0|x)}=\theta^T x$$

- 通过极大似然函数法估计θ，得到似然函数

$$L(\theta)=\prod_{i=1}^{m}p(y_i|x_i:\theta)$$

  对数回归模型最大化对数似然：

$$\iota(\theta)=\sum_{i=1}^{m}lnp(y_i|x_i:\theta)$$

  令

$$p_1(x:\theta)=p(y=1|x:\theta)$$

$$p_0(x:\theta)=p(y=0|x:\theta)$$

上式的似然项可重写为

$$p(y_i|x_i:\theta)=y_ip_1(x_i:\theta)+(1-y_i)p_0(x_i:\theta)$$

代入对数几率回归模型中可得

$$\iota(\theta)=\sum_{i=1}^{m}ln[y_ip_1(x_i:\theta)+(1-y_i)p_0(x_i:\theta)]$$

通过求对数似然函数的最大值来估计模型参数。

上式对数似然函数的最大化等价于下式的最小化

$$\theta^*=argmax\iota(\theta)=argmin[-\iota(\theta)]=argmin\sum_{i=1}^{m}[ln(1+e^{\theta^T x})-y_i\theta^T x_i])$$

- 另一方面，令似然项

$$p(y|x:\theta)=(h_\theta(x))^y(1-h_\theta(x))^{1-y}$$

  则

$$L(\theta)=\prod_{i=1}^{m}p(y^i|x^i:\theta)=\prod_{i=1}^{m}(h_\theta(x^i))^y(1-h_\theta(x^i))^{1-y^{(i)}}$$

  得到对数似然如下

$$\iota(\theta)=\sum_{i=1}^{m}[y^i\theta^T x_i-ln(1+e^{\theta^T x_i})]$$

代价函数

$$J(\theta) = -\frac{1}{m}[\sum_{i=1}^{m} y_i ln(h(\theta(x^i))) + (1-y^i)ln(1-h_\theta(x^i))]$$

运用梯度下降，更新参数θ

$$\frac{\partial J(\theta)}{\partial \theta} = \frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^i)-y^i)x_j^i$$

其余的和线性回归一样，求导计算。