逻辑回归（Logistic Regression）

逻辑回归是二分类模型，而非回归模型

与线性回归模型的区别在于计算出线性组合的结果之后，连接了 sigmoid() 激活函数，故用于二分类任务

$$sigmoid(x) = \frac{1}{1+e^{-x}},\quad x \in (-\infty, +\infty),\quad sigmoid(x) \in (0,1).$$

sigmoid() 函数有个很好的特性：

$$sigmoid'(x) = sigmoid(x) \cdot (1 - sigmoid(x)).$$

在模型中，令 $h(\vec{x}) = sigmoid(\vec{w}^T\vec{x}) = \frac{1}{1+e^{-\vec{w}^T\vec{x}}}$，$\vec{x}$ 在数据中对应的标签为 y，模型输出的标签为 y'，y' 从 0 和 1 中取值

$$y' = \begin{cases} 1, & \text{if } h(\vec{x}) > 0.5 \\ 0, & \text{otherwise} \end{cases}$$

〈预测阶段〉

$h(\vec{x})$ 表示数据为接近 1 的概率，因此有下列表达式

$$p(y=1 \mid \vec{x}; \vec{w}) = h(\vec{x})$$

$$p(y=0 \mid \vec{x}; \vec{w}) = 1 - h(\vec{x})$$

极大似然估计形式的损失函数推导

上述两个表达式综合起来可以写成：

$$p(y|\vec{x};\vec{w})=(h(\vec{x}))^{y}(1-h(\vec{x}))^{1-y}$$

取似然函数为：

$$L(\vec{w})=\prod_{i=1}^{m}p(y^{(i)}|\vec{x}^{(i)};\vec{w})=\prod_{i=1}^{m}(h(\vec{x}^{(i)}))^{y^{(i)}}(1-h(\vec{x}^{(i)}))^{1-y^{(i)}}$$

让它最大

对数似然函数为：

$$l(\vec{w})=\log L(\vec{w})=\sum_{i=1}^{m}(y^{(i)}\log h(\vec{x}^{(i)})+(1-y^{(i)})\log(1-h(\vec{x}^{(i)})))$$

最大似然估计就是要求得使得 $l(\vec{w})$ 取最大值时的 $\vec{w}$，使用梯度下降法求解，所以在对数似然函数 $l(\vec{w})$ 的前边乘了一个负的系数 $-\frac{1}{m}$

最终的损失函数为 $J(\vec{w})=-\frac{1}{m}l(\vec{w})$

$$=-\frac{1}{m}\sum_{i=1}^{m}(y^{(i)}\log(h(\vec{x}^{(i)}))+(1-y^{(i)})\cdot\log(1-h(\vec{x}^{(i)})))$$

拓展：

什么是似然函数？

- 在数理统计中，似然函数是一种关于统计模型中参数的函数，表示模型参数中的似然性。
- 在统计学中，"概率"表示在已知一些参数的情况下，预

测接下来在观测上的结果."似然性"用于在已知某些观测所得的结果时,对参数进行估值.