线性回归

数据集：$D\{(x_1, y_1), (x_2, y_2) \cdots\}$.

$f(x) = w_1 x_1 + w_2 x_2 + \cdots + w_d x_d + b$.

向量化： $f(x) = w^T \vec{x} + b$

$f(x_i) = w x_i + b$.

目标函数： $\sum_{i=1}^{m} (f(x_i) - y_i)^2$

$$(w^*, b^*) = \arg\min_{(w,b)} \sum_{i=1}^{m} (f(x_i) - y_i)^2$$

$$E(w, b) = \sum_{i=1}^{m} (y_i - (w x_i + b))^2$$

$$\frac{\partial E}{\partial w} = 2\left(w \cdot \sum_{i=1}^{m} x_i^2 - \sum_{i=1}^{m} (y_i - b) x_i\right)$$

$$\frac{\partial E}{\partial b} = 2\left(mb - \sum_{i=1}^{m} (y_i - w x_i)\right)$$

$$w = \frac{\sum_{i=1}^{m} y_i (x_i - \bar{x})}{\sum_{i=1}^{m} x_i^2 - \frac{1}{m} \left(\sum_{i=1}^{m} x_i\right)^2}$$

$$b = \frac{1}{m} \sum_{i=1}^{m} (y_i - w x_i)$$

对数几率回归

Sigmoid 函数：$y=\frac{1}{1+e^{-z}}$

$y=\frac{1}{1+e^{-(w^Tx+b)}}$ $\qquad \ln\frac{y}{1-y}=w^Tx+b$

y：x 正类

$1-y$ 反例

$y/1-y$：odds. 相对可能性.

决策函数：

$$p(y=1|x)=\frac{1}{1+e^{-(w^Tx+b)}}=\frac{e^{w^Tx+b}}{1+e^{w^Tx+b}}$$

$$p(y=0|x)=\frac{1}{1+e^{w^Tx+b}}$$

$$y=\frac{1}{1+e^{-(w^Tx+b)}}$$

$$w^Tx+b=\ln\frac{p(y=1|x)}{1-p(y=1|x)}$$

用线性回归模型逼近真实标记的对数几率

极大似然估计.

$$\ell(w,b) = \sum_{i=1}^{m} \ln p(y_i \mid x_i; w, b)$$