线性回归

线性回归是回归问题中的一种，线性回归假设目标值与特征之间线性相关，即满足一个多元一次方程。通过构建损失函数，来求解损失函数最小时的参数 w 和 b。通长我们可以表达成如下公式：

$$\hat{y} = wx + b$$

目标/损失函数：

$$L = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$$

即预测值与真实值之间的平均的平方距离，统计中一般称其为 MAE(mean square error)均方误差。把之前的函数式代入损失函数，并且将需要求解的参数 w 和 b 看做是函数 L 的自变量，可得

$$L(w,b) = \frac{1}{n}\sum_{i=1}^{n}(wx_i + b - y_i)^2$$

现在的任务是求解最小化 L 时 w 和 b 的值，
即核心目标优化式为

$$(w^*, b^*) = \arg\min_{(w,b)}\sum_{i=1}^{n}(wx_i + b - y_i)^2$$

求解方式有两种：

1）最小二乘法(least square method)

求解 w 和 b 是使损失函数最小化的过程，在统计中，称为线性回归模型的最小二乘"参数估计"(parameter estimation)。我们可以将 L(w,b) 分别对 w 和 b 求导，得到

$$\frac{\partial L}{\partial w} = 2\left(w\sum_{i=1}^{n}x^2 - \sum_{i=1}^{n}x_i(y_i - b)\right)$$

$$\frac{\partial L}{\partial b} = 2\left(nb - \sum_{i=1}^{n}(y_i - wx_i)\right)$$

令上述两式为 0，可得到 w 和 b 最优解的闭式(closed-form)解：

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

$$b = \frac{1}{n}\sum_{i=1}^{n}(y_i - wx_i)$$

2）梯度下降(gradient descent)

梯度下降核心内容是对自变量进行不断的更新（针对 w 和 b 求偏导），使得目标函数不断逼近最小值的过程

$$w \leftarrow w - \alpha\frac{\partial L}{\partial w}$$

$$b \leftarrow b - \alpha\frac{\partial L}{\partial b}$$