**线性回归**

## 相关概念和公式

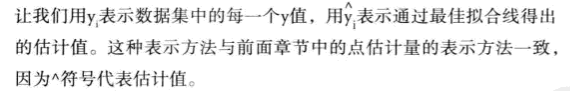
### 最小二乘法

一种数学方法，可用一条最佳拟合线将一组二变量数据拟合，通过将公式为 y = ax + b的一条直线与一组数值拟合，使得误差平方和最小。

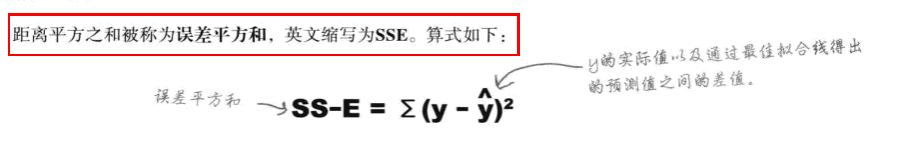
它通过最小化误差的平方和寻找数据的最佳[函数](https://baike.baidu.com/item/%E5%87%BD%E6%95%B0)匹配。利用最小二乘法可以简便地求得未知的数据，并使得这些求得的数据与实际数据之间误差的平方和为最小。最小二乘法还可用于[曲线拟合](https://baike.baidu.com/item/%E6%9B%B2%E7%BA%BF%E6%8B%9F%E5%90%88)。

### 使用说明

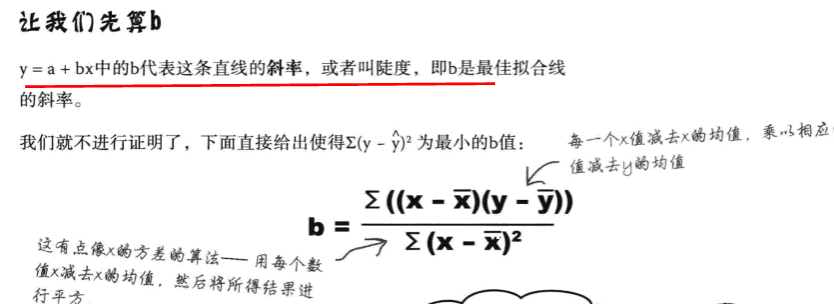
* 使用散点图，将数据显示，可以初略的看出数据的表述意思，但这只是初略的看出，通过散点图可以绘制出一条拟合线，可以用来进行预测数据。通过这条直线来得到的数据仍然是猜测，这时可以得出直线公式：y = bx + a。
* 需要求解出最佳拟合线，即将误差最小化，使用 最小二乘法



**误差平方和**



即通过最小二乘法对 求出的拟合线 y = bx + a 进行修正。以下公式其中 x和y减去的都是通过计算得到的均值。



**求解完成了斜率后，求Y轴上的值**

y = bx + a 由于这里已经将斜率b计算得到了，在求解a时，将计算出的x和y的均值代入方差，就可以计算到a。

于是计算得到的直线 y = bx + a 称为回归线。

