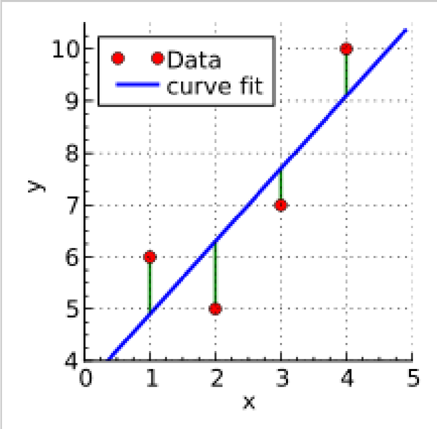
**最小二乘法**（又称**最小平方法**）是一种数学优化技术。它通过最小化误差的平方和寻找数据的最佳函数匹配。

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。

其他一些优化问题也可通过最小化能量或最大化熵用最小二乘法来表达。

示例[编辑]



数据点（红色）、使用最小二乘法求得的最佳解（蓝色）、误差（绿色

某次实验得到了四个数据点 $(x, y)$：$(1, 6)$、$(2, 5)$、$(3, 7)$、$(4, 10)$（右图中红色的点）。我们希望找出一条和这四个点最匹配的直线 $y = \beta_1 + \beta_2 x$，即找出在某种“最佳情况”下能够大致符合如下超定线性方程组的 $\beta_1$ 和 $\beta_2$：

$$\begin{aligned}\beta_1 + 1\beta_2 &= 6\\ \beta_1 + 2\beta_2 &= 5\\ \beta_1 + 3\beta_2 &= 7\\ \beta_1 + 4\beta_2 &= 10\end{aligned}$$

最小二乘法采用的手段是尽量使得等号两边的方差最小，也就是找出这个函数的最小值：

$$\begin{aligned}S(\beta_1, \beta_2) = &\left[6 - (\beta_1 + 1\beta_2)\right]^2 + \left[5 - (\beta_1 + 2\beta_2)\right]^2\\ &+ \left[7 - (\beta_1 + 3\beta_2)\right]^2 + \left[10 - (\beta_1 + 4\beta_2)\right]^2.\end{aligned}$$

最小值可以通过对 $S(\beta_1, \beta_2)$ 分别求 $\beta_1$ 和 $\beta_2$ 的偏导数，然后使它们等于零得到。

$$\frac{\partial S}{\partial \beta_1} = 0 = 8\beta_1 + 20\beta_2 - 56$$
$$\frac{\partial S}{\partial \beta_2} = 0 = 20\beta_1 + 60\beta_2 - 154.$$

如此就得到了一个只有两个未知数的方程组，很容易就可以解出：

$$\beta_1 = 3.5$$
$$\beta_2 = 1.4$$

也就是说直线 $y = 3.5 + 1.4x$ 是最佳的。

人们对由某一变量$t$ 或多个变量$t_1$……$t_n$ 构成的相关变量$y$感兴趣。如弹簧的形变与所用的力相关，一个企业的盈利
营业额，投资收益和原始资本有关。为了得到这些变量同$y$之间的关系，便用不相关变量去构建$y$，使用如下函数模

$$y_m = f(t_1, \dots, t_q; b_1, \dots, b_p),$$

$q$个独立变量或$p$个系数去拟合。

通常人们将一个可能的、对不相关变量t的构成都无困难的函数类型称作函数模型（如抛物线函数或指数函数）。参
是为了使所选择的函数模型同观测值y相匹配。（如在测量弹簧形变时，必须将所用的力与弹簧的膨胀系数联系起来
其目标是合适地选择参数，使函数模型最好的拟合观测值。**一般情况下，观测值远多于所选择的参数。**

其次的问题是怎样判断不同拟合的质量。高斯和勒让德的方法是，**假设测量误差的平均值为0**。令每一个测量误差对个变量并与其它**测量误差不相关（随机无关）**。人们假设，在测量误差中绝对不含系统误差，它们应该是**纯偶然误差固定的变异数)，围绕真值波动**。除此之外，**测量误差符合**正态分布，这保证了偏差值在最后的结果y上忽略不计。

确定拟合的标准应该被重视，并小心选择，较大误差的测量值应被赋予较小的权。并建立如下规则：被选择的参数，该使算出的**函数曲线与观测值之差的平方和最小**。用函数表示为：

$$\min_{\vec{b}} \sum_{i=1}^{n} (y_m - y_i)^2.$$

用欧几里得度量表达为：

$$\min_{\vec{b}} \|\vec{y}_m(\vec{b}) - \vec{y}\|_2 \ .$$

最小化问题的精度，依赖于所选择的函数模型。

# 线性函数模型[编辑]

典型的一类函数模型是线性函数模型。最简单的线性式是$y = b_0 + b_1 t$，写成矩阵式，为

$$\min_{b_0, b_1} \left\| \begin{pmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\|_2 = \min_b \|Ab - Y\|_2.$$

直接给出该式的参数解：

$$b_1 = \frac{\sum_{i=1}^n t_i y_i - n \cdot \bar{t}\bar{y}}{\sum_{i=1}^n t_i^2 - n \cdot (\bar{t})^2} \text{和 } b_0 = \bar{y} - b_1 \bar{t}$$

其中$\bar{t} = \frac{1}{n} \sum_{i=1}^{n} t_i$，为t值的算术平均值。也可解得如下形式：

$$b_1 = \frac{\sum_{i=1}^n (t_i - \bar{t})(y_i - \bar{y})}{\sum_{i=1}^n (t_i - \bar{t})^2}$$