数值计算方法

曲线拟合

张晓平

2019 年 11 月 17 日

武汉大学数学与统计学院

Table of contents

简介

简介

在科学研究与工程技术中，经常会获取到一些数据$\{(x_i, y_i)\}_{i=1}^n$，如何来确定 x 与 y 之间的函数关系呢?

简介

在科学研究与工程技术中，经常会获取到一些数据$\{(x_i, y_i)\}_{i=1}^n$，如何来确定 x 与 y 之间的函数关系呢?

虽然插值是函数逼近的一种重要方法，但存在以下缺陷:

简介

在科学研究与工程技术中，经常会获取到一些数据$\{(x_i, y_i)\}_{i=1}^{n}$，如何来确定 x 与 y 之间的函数关系呢?

虽然插值是函数逼近的一种重要方法，但存在以下缺陷:

1. 测量数据不可避免带有误差，而插值函数又经过所有数据点，从而使得插值函数保留了这些误差，从而影响了逼近精度;
2. 如果实验数据较多，则需要用到次数较高的插值多项式，而高次插值多项式数值不稳定，从而缺乏实用价值。

简介

在科学研究与工程技术中，经常会获取到一些数据$\{(x_i, y_i)\}_{i=1}^n$，如何来确定 x 与 y 之间的函数关系呢?

虽然插值是函数逼近的一种重要方法，但存在以下缺陷:

1. 测量数据不可避免带有误差，而插值函数又经过所有数据点，从而使得插值函数保留了这些误差，从而影响了逼近精度;
2. 如果实验数据较多，则需要用到次数较高的插值多项式，而高次插值多项式数值不稳定，从而缺乏实用价值。

因此，给定一组数据，如何构造一个逼近函数 $y = \varphi(x)$，使得它从总体来说与 y_i 的偏差按某种方法度量达到最小而又不一定过所有数据点，这就是我们将要介绍的最小二乘曲线拟合法。

用最小二乘法求解矛盾方程组

用最小二乘法求解矛盾方程组

最小二乘原理

问题

构造曲线 $y = \varphi(x)$ 来拟合数据$\{(x_i, y_i)\}_{i=1}^{n}$，“拟合得最好”的标准是什么呢?

显然，希望选择 $\varphi(x)$，使得 $\varphi(x_i)$ 与 y_i 相差都很小，即使得偏差

$$\varphi(x_i) - y_i, \quad i = 1, 2, \cdots, m$$

都很小。

最小二乘原理

1. 使偏差之和
$$\sum_{i=1}^{m}[\varphi(x_i)-y_i]$$
很小来保证每个偏差都很小。

最小二乘原理

1. 使偏差之和

$$\sum_{i=1}^{m} [\varphi(x_i) - y_i]$$

很小来保证每个偏差都很小。但由于偏差有正有负，求和时可能抵消。

最小二乘原理

1. 使偏差之和

$$\sum_{i=1}^{m}[\varphi(x_i) - y_i]$$

很小来保证每个偏差都很小。但由于偏差有正有负，求和时可能抵消。

2. 使偏差的绝对值之和

$$\sum_{i=1}^{m}|\varphi(x_i) - y_i|$$

最小。

最小二乘原理

1. 使偏差之和

$$\sum_{i=1}^{m} [\varphi(x_i) - y_i]$$

很小来保证每个偏差都很小。但由于偏差有正有负，求和时可能抵消。

2. 使偏差的绝对值之和

$$\sum_{i=1}^{m} |\varphi(x_i) - y_i|$$

最小。但涉及到绝对值，不便于分析。

最小二乘原理

1. 使偏差之和
$$\sum_{i=1}^{m}[\varphi(x_i)-y_i]$$
很小来保证每个偏差都很小。但由于偏差有正有负，求和时可能抵消。

2. 使偏差的绝对值之和
$$\sum_{i=1}^{m}|\varphi(x_i)-y_i|$$
最小。但涉及到绝对值，不便于分析。

3. 使偏差的平方和
$$\sum_{i=1}^{m}[\varphi(x_i)-y_i]^2$$
最小来保证每个偏差的绝对值都很小。

最小二乘原理

1. 使偏差之和
$$\sum_{i=1}^{m}[\varphi(x_i)-y_i]$$
很小来保证每个偏差都很小。但由于偏差有正有负，求和时可能抵消。

2. 使偏差的绝对值之和
$$\sum_{i=1}^{m}|\varphi(x_i)-y_i|$$
最小。但涉及到绝对值，不便于分析。

3. 使偏差的平方和
$$\sum_{i=1}^{m}[\varphi(x_i)-y_i]^2$$
最小来保证每个偏差的绝对值都很小。使“偏差平方和最小”的原则成为最小二乘原则，按最小二乘原则拟合曲线的方法称为最小二乘法或最小二乘曲线拟合法

最小二乘原理

一般地，拟合曲线 $\varphi(x)$ 是 n 个线性无关函数 $\varphi_1(x), \varphi_2(x), \cdots, \varphi_n(x)$ 的线性组合，即

$$\varphi(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \cdots + a_n\varphi_n(x), \quad m > n$$

其中 $a_1, a_2, \cdots, a_n$ 待定，线性无关函数组 $\varphi_1(x), \varphi_2(x), \cdots, \varphi_n(x)$ 称为基函数。

最小二乘原理

一般地，拟合曲线 $\varphi(x)$ 是 n 个线性无关函数 $\varphi_1(x), \varphi_2(x), \cdots, \varphi_n(x)$ 的线性组合，即

$$\varphi(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \cdots + a_n\varphi_n(x), \quad m > n$$

其中 $a_1, a_2, \cdots, a_n$ 待定，线性无关函数组 $\varphi_1(x), \varphi_2(x), \cdots, \varphi_n(x)$ 称为基函数。常用的基函数有

1 多项式：$1, x, x^2, \cdots, x^n$

2 三角函数：$\sin x, \sin 2x, \cdots, \sin nx$

3 指数函数：$e^{\lambda_1 x}, e^{\lambda_2 x}, \cdots, e^{\lambda_n x}$

用最小二乘法求解矛盾方程组

用最小二乘法求解矛盾方程组

用最小二乘法求解矛盾方程组

方程个数多于未知数个数的方程组称为矛盾方程组，此类方程组往往无解。而最小二乘法是用来求解矛盾方程组的一种常用方法。

用最小二乘法求解矛盾方程组

方程个数多于未知数个数的方程组称为矛盾方程组，此类方程组往往无解。而最小二乘法是用来求解矛盾方程组的一种常用方法。

设有矛盾方程组

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \cdots, m; \; m > n$$

令

$$\delta_i = \sum_{j=1}^{n} a_{ij}x_j - b_i, \quad i = 1, 2, \cdots, m, \quad \text{偏差}$$

按最小二乘原则，常采用偏差的平方和

$$Q = \sum_{i=1}^{m} \delta_i^2 = \sum_{i=1}^{m} \left[\sum_{j=1}^{n} a_{ij}x_j - b_i \right]^2$$

达到最小来衡量一个近似解近似程度的重要标志。

用最小二乘法求解矛盾方程组

方程个数多于未知数个数的方程组称为矛盾方程组，此类方程组往往无解。而最小二乘法是用来求解矛盾方程组的一种常用方法。

设有矛盾方程组

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \cdots, m; \;\; m > n$$

令

$$\delta_i = \sum_{j=1}^{n} a_{ij}x_j - b_i, \quad i = 1, 2, \cdots, m, \quad \text{偏差}$$

按最小二乘原则，常采用偏差的平方和

$$Q = \sum_{i=1}^{m} \delta_i^2 = \sum_{i=1}^{m} \left[\sum_{j=1}^{n} a_{ij}x_j - b_i \right]^2$$

达到最小来衡量一个近似解近似程度的重要标志。若 x_j 使得 Q 达到最小，则称它为矛盾方程组的最优近似解。

用最小二乘法求解矛盾方程组

偏差平方和 Q 可看作是 n 个自变量 x_j 的二次函数，因此求解矛盾方程组的问题归结为求二次函数 Q 的最小值问题。

用最小二乘法求解矛盾方程组

偏差平方和 Q 可看作是 n 个自变量 x_j 的二次函数，因此求解矛盾方程组的问题归结为求二次函数 Q 的最小值问题。Q 取得极值的必要条件是

$$\frac{\partial Q}{\partial x_k}=0, \quad k=1,2,\cdots,n$$

即

$$0=\frac{\partial Q}{\partial x_k}=\sum_{i=1}^{m} 2\left[\sum_{j=1}^{n} a_{ij}x_j-b_i\right] a_{ik}=2\sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}a_{ik}\right) x_j-2\sum_{i=1}^{m} a_{ik}b_i$$

用最小二乘法求解矛盾方程组

偏差平方和 Q 可看作是 n 个自变量 x_j 的二次函数，因此求解矛盾方程组的问题归结为求二次函数 Q 的最小值问题。 Q 取得极值的必要条件是

$$\frac{\partial Q}{\partial x_k} = 0, \quad k = 1, 2, \cdots, n$$

即

$$0 = \frac{\partial Q}{\partial x_k} = \sum_{i=1}^{m} 2\left[\sum_{j=1}^{n} a_{ij}x_j - b_i\right] a_{ik} = 2\sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}a_{ik}\right) x_j - 2\sum_{i=1}^{m} a_{ik}b_i$$

其矩阵形式为

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}, \quad \rightarrow \quad \text{法方程组}$$

其中

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

用最小二乘法求解矛盾方程组

用最小二乘法求解矛盾方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

1. 计算 $\boldsymbol{A}^T\boldsymbol{A}$ 和 $\boldsymbol{A}^T\boldsymbol{b}$ 得法方程组;
2. 求解法方程组 $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$，得出矛盾方程组的最小二乘解。

用多项式作最小二乘曲线拟合

用多项式作最小二乘曲线拟合

设

$$\varphi(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \quad (m > n + 1)$$

依最小二乘原理，给定 $\{(x_i, y_i)\}_{i=1}^{m}$，确定 a_j 使得偏差平方和最小。

用多项式作最小二乘曲线拟合

设

$$\varphi(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \quad (m > n+1)$$

依最小二乘原理，给定 $\{(x_i, y_i)\}_{i=1}^m$，确定 a_j 使得偏差平方和最小。于是可得到一个 $m+1$ 个未知数 a_j 的 n 个方程的矛盾方程组

$$\begin{cases} a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_n x_1^n & = & y_1, \\ a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_n x_2^n & = & y_2, \\ \cdots & & \\ a_0 + a_1 x_m + a_2 x_m^2 + \cdots + a_n x_m^n & = & y_m, \end{cases}$$

用多项式作最小二乘曲线拟合

设

$$\varphi(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \quad (m > n+1)$$

依最小二乘原理，给定 $\{(x_i, y_i)\}_{i=1}^{m}$，确定 a_j 使得偏差平方和最小。于是可得到一个 $m+1$ 个未知数 a_j 的 n 个方程的矛盾方程组

$$\left\{\begin{array}{rcl} a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_n x_1^n & = & y_1, \\ a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_n x_2^n & = & y_2, \\ \cdots & & \\ a_0 + a_1 x_m + a_2 x_m^2 + \cdots + a_n x_m^n & = & y_m, \end{array}\right.$$

其矩阵形式为

$$\boldsymbol{A}\boldsymbol{\alpha} = \boldsymbol{y},$$

其中

$$\boldsymbol{A} = \begin{bmatrix} 1 & x_1 & \cdots & x_1^n \\ 1 & x_2 & \cdots & x_2^n \\ \vdots & \vdots & & \vdots \\ 1 & x_m & \cdots & x_m^n \end{bmatrix}, \quad \boldsymbol{\alpha} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

其法方程组为

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{A}^T\boldsymbol{y}$$

由于 $x_1, x_2, \cdots, x_m$ 互异，故 $\boldsymbol{A}$ 的 $n+1$ 个列向量线性无关，从而 $r(\boldsymbol{A})=n+1$。于是 $\forall \boldsymbol{x}\in\mathbb{R}^{n+1}$，都有 $\boldsymbol{A}\boldsymbol{x}\neq\mathbf{0}$，由

$$\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}=(\boldsymbol{A}\boldsymbol{x})^T(\boldsymbol{A}\boldsymbol{x})>0$$

知 $\boldsymbol{A}^T\boldsymbol{A}$ 对称正定，从而 $\boldsymbol{A}^T\boldsymbol{A}$ 非奇异。从而上述法方程组的解存在惟一。

用多项式作最小二乘曲线拟合

$$
\begin{aligned}
\boldsymbol{A}^T\boldsymbol{A} &= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_m \\ \vdots & \vdots & & \vdots \\ x_1^n & x_2^n & \cdots & x_m^n \end{bmatrix} \begin{bmatrix} 1 & x_1 & \cdots & x_1^n \\ 1 & x_2 & \cdots & x_2^n \\ \vdots & \vdots & & \vdots \\ 1 & x_m & \cdots & x_m^n \end{bmatrix} \\
&= \begin{bmatrix} m & \sum_{i=1}^m x_i & \cdots & \sum_{i=1}^m x_i^n \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 & \cdots & \sum_{i=1}^m x_i^{n+1} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^m x_i^n & \sum_{i=1}^m x_i^{n+1} & \cdots & \sum_{i=1}^m x_i^{2n} \end{bmatrix}
\end{aligned}
$$

计算系数矩阵时，只需计算

$$
m, \quad \sum_{i=1}^m x_i, \quad \sum_{i=1}^m x_i^2, \quad \cdots, \quad \sum_{i=1}^m x_i^n, \quad \sum_{i=1}^m x_i^{n+1}, \quad \cdots, \quad \sum_{i=1}^m x_i^{2n}
$$

用多项式作最小二乘曲线拟合

例

通过实验获得数据如下

x_i	1	2	3	4	6	7	8
y_i	2	3	6	7	5	3	2

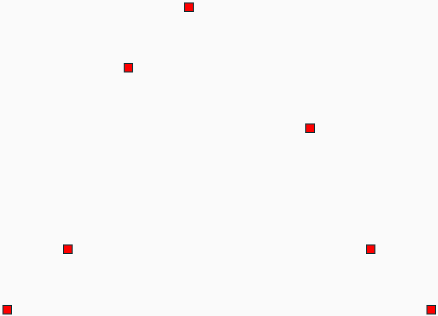
试用最小二乘法求多项式曲线，使与此数据相拟合。

例

通过实验获得数据如下

x_i	1	2	3	4	6	7	8
y_i	2	3	6	7	5	3	2

试用最小二乘法求多项式曲线，使与此数据相拟合。



用多项式作最小二乘曲线拟合

解

1 确定近似表达式

$$y = \varphi(x) = a_0 + a_1 x + a_2 x^2$$

用多项式作最小二乘曲线拟合

解

1 确定近似表达式

$$y = \varphi(x) = a_0 + a_1 x + a_2 x^2$$

2 建立法方程组

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 6 & 36 \\ 1 & 7 & 49 \\ 1 & 8 & 64 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 6 \\ 7 \\ 5 \\ 3 \\ 2 \end{pmatrix}$$

用多项式作最小二乘曲线拟合

解

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 6 & 7 & 8 \\ 1 & 4 & 9 & 16 & 36 & 49 & 64 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 6 & 36 \\ 1 & 7 & 49 \\ 1 & 8 & 64 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 6 & 7 & 8 \\ 1 & 4 & 9 & 16 & 36 & 49 & 64 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 6 \\ 7 \\ 5 \\ 3 \\ 2 \end{pmatrix}$$

用多项式作最小二乘曲线拟合

解

3 求解法方程组

$$\begin{pmatrix} 7 & 31 & 179 \\ 31 & 179 & 1171 \\ 179 & 1171 & 8147 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 28 \\ 121 \\ 635 \end{pmatrix}$$

得

解

3 求解法方程组

$$\begin{pmatrix} 7 & 31 & 179 \\ 31 & 179 & 1171 \\ 179 & 1171 & 8147 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 28 \\ 121 \\ 635 \end{pmatrix}$$

得

$$a_0 = -1.3185, \quad a_1 = 3.4321, \quad a_2 = -0.3864$$

解

3 求解法方程组

$$\begin{pmatrix} 7 & 31 & 179 \\ 31 & 179 & 1171 \\ 179 & 1171 & 8147 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 28 \\ 121 \\ 635 \end{pmatrix}$$

得

$$a_0 = -1.3185, \quad a_1 = 3.4321, \quad a_2 = -0.3864$$

故所求拟合曲线为

$$y = \varphi(x) = -1.3185 + 3.4321x - 0.3864x^2.$$

用多项式作最小二乘曲线拟合

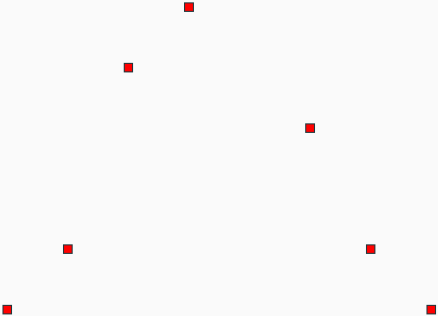


图 1: 曲线拟合

用多项式作最小二乘曲线拟合

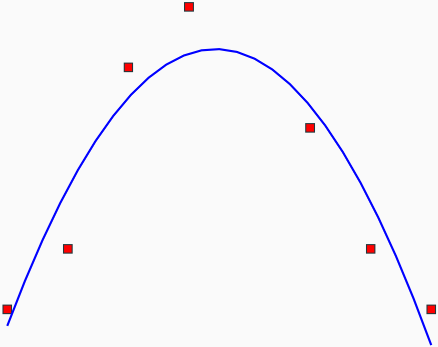


图 **1:** 曲线拟合

例

在一物理实验中，电压 V 与电流 I 的一组数据如下

V	1	2	3	4	5	6	7	8
I	1.53	2.05	2.74	3.66	4.91	6.56	8.78	11.76

试用最小二乘法求最佳拟合函数。

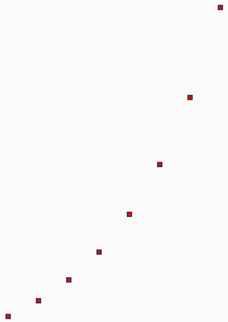
用多项式作最小二乘曲线拟合

例

在一物理实验中，电压 V 与电流 I 的一组数据如下

V	1	2	3	4	5	6	7	8
I	1.53	2.05	2.74	3.66	4.91	6.56	8.78	11.76

试用最小二乘法求最佳拟合函数。



用多项式作最小二乘曲线拟合

解

1 确定近似表达式

$$I = ae^{bV} \quad \Rightarrow \quad \ln I = \ln a + bV$$

V	1	2	3	4	5	6	7	8
I	1.53	2.05	2.74	3.66	4.91	6.56	8.78	11.76

用多项式作最小二乘曲线拟合

解

1 确定近似表达式

$$I = ae^{bV} \quad \Rightarrow \quad \ln I = \ln a + bV$$

V	1	2	3	4	5	6	7	8
I	1.53	2.05	2.74	3.66	4.91	6.56	8.78	11.76

V	1	2	3	4	5	6	7	8
$\ln I$	0.43	0.72	1.01	1.30	1.59	1.88	2.17	2.46

解

2 建立法方程组

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \\ 1 & 7 \\ 1 & 8 \end{pmatrix} \begin{pmatrix} \ln a \\ b \end{pmatrix} = \begin{pmatrix} 0.43 \\ 0.72 \\ 1.01 \\ 1.30 \\ 1.59 \\ 1.88 \\ 2.17 \\ 2.46 \end{pmatrix}$$

用多项式作最小二乘曲线拟合

解

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 5 & 4 & 6 & 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 6 \\ 1 & 7 \\ 1 & 8 \end{pmatrix} \begin{pmatrix} \ln a \\ b \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 5 & 4 & 6 & 7 & 8 \end{pmatrix} \begin{pmatrix} 2.7 \\ 3.0 \\ 3.3 \\ 3.6 \\ 3.9 \\ 4.2 \\ 4.5 \\ 4.8 \end{pmatrix}$$

用多项式作最小二乘曲线拟合

解

3 求解法方程组

$$\begin{pmatrix} 8 & 36 \\ 36 & 204 \end{pmatrix} \begin{pmatrix} \ln a \\ b \end{pmatrix} = \begin{pmatrix} 29.9787 \\ 147.1350 \end{pmatrix}$$

得

解

3 求解法方程组

$$\begin{pmatrix} 8 & 36 \\ 36 & 204 \end{pmatrix} \begin{pmatrix} \ln a \\ b \end{pmatrix} = \begin{pmatrix} 29.9787 \\ 147.1350 \end{pmatrix}$$

得

$$\ln a = 0.1343 \quad \Rightarrow \quad a = 1.14393$$

$$b = 0.2912$$

用多项式作最小二乘曲线拟合

解

3 求解法方程组

$$\begin{pmatrix} 8 & 36 \\ 36 & 204 \end{pmatrix} \begin{pmatrix} \ln a \\ b \end{pmatrix} = \begin{pmatrix} 29.9787 \\ 147.1350 \end{pmatrix}$$

得

$$\ln a = 0.1343 \quad \Rightarrow \quad a = 1.14393$$

$$b = 0.2912$$

故所求拟合曲线为

$$I = 1.14393 e^{0.2912\,V}$$

用多项式作最小二乘曲线拟合

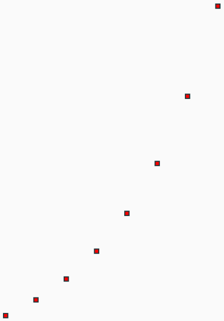


图 2: 曲线拟合

用多项式作最小二乘曲线拟合

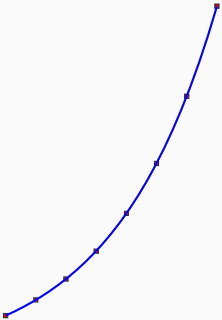


图 **2:** 曲线拟合