

# 数学实验



## Experiments in Mathematics

### 实验二 插值与拟合

清华大学数学科学系

什么是插值?

什么是拟合?



从查函数表和做物理实验说起



标准正态分布函数表  $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$

x	0	1	2	...
1.0	0.8413	0.8438	0.8461	...
1.1	0.8643	0.8665	0.8686	...
1.2	0.8849	0.8869	0.8888	...

求  $\Phi(1.114)$

$\Phi(1.114) = 0.8665 + (0.8686 - 0.8665) \times 0.4 = 0.8673$

插  
值

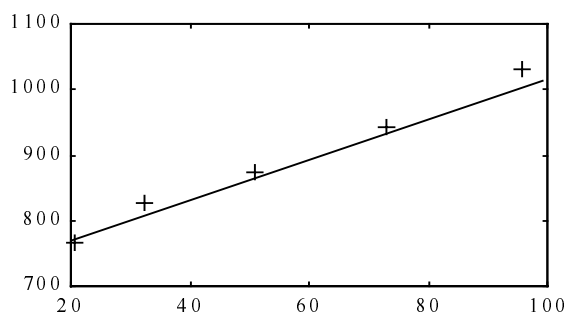


## 热敏电阻实验

测得一组  
实验数据

温度 $t(^{\circ}\text{C})$	20.5	32.7	51.0	73.0	95.7
电阻 $R(\Omega)$	765	826	873	942	1032

求  $60^{\circ}\text{C}$  时的电阻



dianzu1

设  $R = at + b$ ,  $a, b$   
由实验数据确定

拟合

## 实验二 插值与拟合 的基本内容

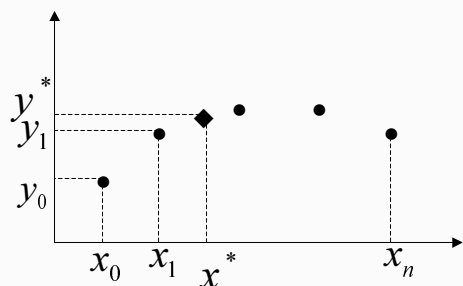


1. 插值的基本原理；三种插值方法：拉格朗日插值，分段线性插值，三次样条插值。
2. 拟合的基本原理；线性最小二乘拟合。
3. 插值与拟合的 **MATLAB** 实现。
4. 插值与拟合的应用（面对一个实际问题，应该用插值，还是拟合）。

### 插值的 基本原理

### 插值问题的提法

已知  $n+1$  个节点  $(x_j, y_j)$  ( $j = 0, 1, \dots, n$ , 其中  $x_j$  互不相同，不妨设  $a = x_0 < x_1 < \dots < x_n = b$ )，求任一插值点  $x^*$  ( $\neq x_j$ ) 处的插值  $y^*$ 。



节点可视为由  $y = g(x)$  产生， $g$  表达式复杂，甚至无表达式

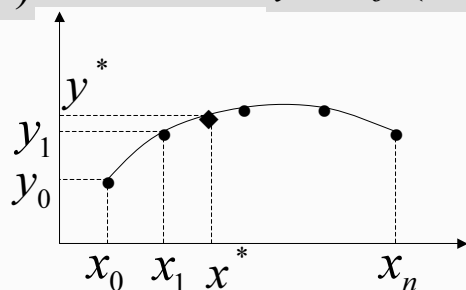
### 插值的 基本原理

### 求解插值问题的基本思路

构造一个(相对简单的)函数  $y = f(x)$ , 通过全部节点, 即

$$f(x_j) = y_j \quad (j = 0, 1, \dots, n)$$

再用  $f(x)$  计算插值, 即  $y^* = f(x^*)$ .



### 三种插值 方法

#### 1. 拉格朗日 (Lagrange) 多项式插值

##### 1.1 插值多项式

$$L_n(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \quad (1)$$

$$L_n(x_j) = y_j \quad (j = 0, 1, \dots, n) \quad \xrightarrow{\text{求 } a_i} \quad XA = Y \quad (2)$$

$$X = \begin{bmatrix} x_0^n & x_0^{n-1} & \dots & 1 \\ \dots & \dots & \dots & \dots \\ x_n^n & x_n^{n-1} & \dots & 1 \end{bmatrix}, \quad A = \begin{bmatrix} a_n \\ \vdots \\ a_0 \end{bmatrix}, \quad Y = \begin{bmatrix} y_0 \\ \vdots \\ y_n \end{bmatrix}$$

$\therefore \det(X) \neq 0$  (在什么条件下)

$\therefore (2)$  有唯一解

三种插值  
方法

$$L_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \quad (1)$$

1.2 拉格朗日插值多项式

$$XA = Y \quad (2)$$

$$L_n(x) = \sum_{i=0}^n y_i l_i(x) \quad (3)$$

$$l_i(x) = \frac{(x-x_0) \cdots (x-x_{i-1})(x-x_{i+1}) \cdots (x-x_n)}{(x_i-x_0) \cdots (x_i-x_{i-1})(x_i-x_{i+1}) \cdots (x_i-x_n)}, \quad i=0,1,\dots,n$$

$$\because l_i(x_j) = \begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases} \therefore L_n(x_j) = y_j \quad \begin{matrix} l_i(x) \\ \text{称基函数} \end{matrix}$$

又 (2) 有唯一解, 故 (3) 与 (1) 相同。



1.3 误差估计

三种插值  
方法

$$R_n(x) = g(x) - L_n(x) = \frac{g^{(n+1)}(\xi)}{(n+1)!} \prod_{j=0}^n (x-x_j), \quad \xi \in (a,b)$$

$$\left| g^{(n+1)}(\xi) \right| \leq M_{n+1} \Rightarrow \left| R_n(x) \right| \leq \frac{M_{n+1}}{(n+1)!} \prod_{j=0}^n |x-x_j|$$

如何使误差  $|R_n(x)|$  减小 (粗略地看)

$x$  接近  $x_j$

$g$  平缓

$n$  增加

### 三种插值方法

### 思考



1) 对于 $n+1$ 个节点，拉格朗日插值为什么用 $n$ 次多项式，若用次数大于 $n$ 或小于 $n$ 的多项式作插值，结果如何？

2) 用 $n$ 次多项式作插值，得到的 $L_n(x)$ 次数会不会小于 $n$ ？可以用 $n=2$ 来说明。

3) 若产生 $n+1$ 个节点的 $g(x)$ 为 $m$ 次多项式，问 $L_n(x)$ 与 $g(x)$ 的关系如何（分 $m \leq n$ ,  $m > n$ 两种情况）？

### 三种插值方法

#### 1.4 拉格朗日插值多项式的振荡

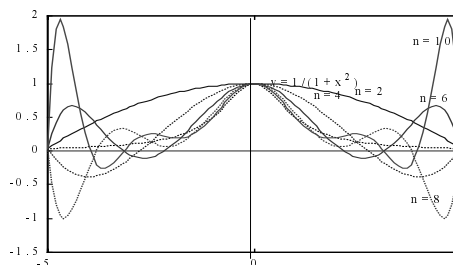
$$n \uparrow \Rightarrow L_n(x) ? \Rightarrow |R_n(x)| \downarrow ?$$

$$g(x) = \frac{1}{1+x^2}, \quad -5 \leq x \leq 5$$

 MATLAB 5.3.1ink **runge**

取 $n=2,4,6,8,10$ , 计算 $L_n(x)$ , 画出图形

**Runge现象**

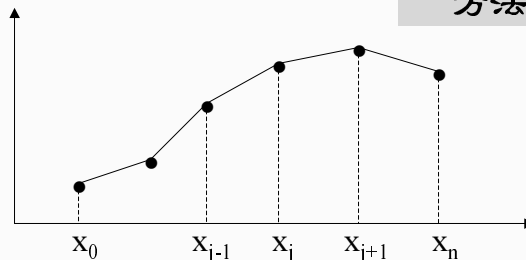


$$\lim_{n \rightarrow \infty} L_n(x) = g(x), \quad -3.63 \leq x \leq 3.63$$



## 2. 分段线性插值

三种插值  
方法



$$I_n(x) = \sum_{j=0}^n y_j l_j(x)$$

$$l_j(x) = \begin{cases} \frac{x - x_{j-1}}{x_j - x_{j-1}}, & x_{j-1} \leq x \leq x_j \\ \frac{x - x_{j+1}}{x_j - x_{j+1}}, & x_j \leq x \leq x_{j+1} \\ 0, & \text{其它} \end{cases}$$

计算量与n无关;  
n越大, 误差越小.

$$\lim_{n \rightarrow \infty} I_n(x) = g(x), x_0 \leq x \leq x_n$$

## 3. 三次样条插值

样条函数的由来

三种插值  
方法

$$S(x) = \{s_i(x), x \in [x_{i-1}, x_i], i = 1, \dots, n\}$$

$$1) s_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \quad (i = 1, \dots, n)$$

$$2) S(x_i) = y_i \quad (i = 0, 1, \dots, n)$$

$$3) S(x) \in C^2[x_0, x_n]$$

4n个待定系数  
 $a_i, b_i, c_i, d_i$

$$3) \Rightarrow \begin{cases} s_i(x_i) = s_{i+1}(x_i), s'_i(x_i) = s'_{i+1}(x_i) \\ s''_i(x_i) = s''_{i+1}(x_i) \end{cases} \quad (i = 1, n-1) \quad 3')$$

2), 3') 共 4n-2 个方程

### 三种插值方法

三次样条插值确定 $4n$ 个系数需增加 2个条件

$$4) S''(x_0) = S''(x_n) = 0 \text{ (自然边界条件)}$$

$$2) 3) 4) \Rightarrow a_i, b_i, c_i, d_i \Rightarrow S(x)$$

$$\lim_{n \rightarrow \infty} S(x) = g(x)$$

### 思考

1) 自然边界条件的几何意义是什么

2) 样条插值为什么普遍用3次多项式，而不是2或4次？



### 三种插值方法小结

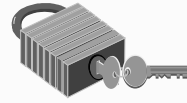
- 拉格朗日插值（高次多项式插值）：  
曲线光滑；误差估计有表达式；收敛性不能保证（振荡现象）。

用于理论分析，实际意义不大。

- 分段线性和三次样条插值（低次多项式插值）：  
曲线不光滑（三次样条插值已大有改进）；误差估计较难（对三次样条插值）；收敛性有保证。  
简单实用，应用广泛。



## 用MATLAB作插值计算



1. 拉格朗日插值: 自编程序, 如名为 `lagrange.m` 的M文件,  
第一行为  
输入: 节点 `x` 插值点 `xi` (均为数组, 长度自定义);  
输出: 插值 `yi` (与 `xi` 同长度数组)。  
应用时输入 `lagrange(x, xi)` 后, 运行

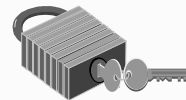
2. 分段线性插值: 已有程序

3. 三次样条插值: 已有程序

或

注: 程序可参考课本第60页

## 用MATLAB作插值计算



以  $g(x) = \frac{1}{1+x^2}$ ,  $-5 \leq x \leq 5$  为例, 作三种插值的比较

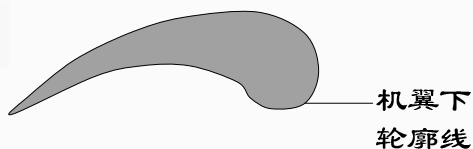
用n=11个节点, m=21个插值点, 三种方法作插值, 画图。



x	y	y1	y2	y3
0	1.0000	1.0000	1.0000	1.0000
0.5000	0.8000	0.8434	0.7500	0.8205
1.0000	0.5000	0.5000	0.5000	0.5000
1.5000	0.3077	0.2353	0.3500	0.2973
2.0000	0.2000	0.2000	0.2000	0.2000
2.5000	0.1379	0.2538	0.1500	0.1401
3.0000	0.1000	0.1000	0.1000	0.1000
3.5000	0.0755	-0.2262	0.0794	0.0745
4.0000	0.0588	0.0588	0.0588	0.0588
4.5000	0.0471	1.5787	0.0486	0.0484
5.0000	0.0385	0.0385	0.0385	0.0385



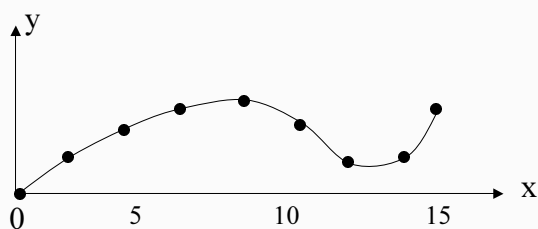
## 插值的应用



已知机翼下轮廓线上数据如下

X	0	3	5	7	9	11	12	13	14	15
Y	0	1.2	1.7	2.0	2.1	2.0	1.8	1.2	1.0	1.6

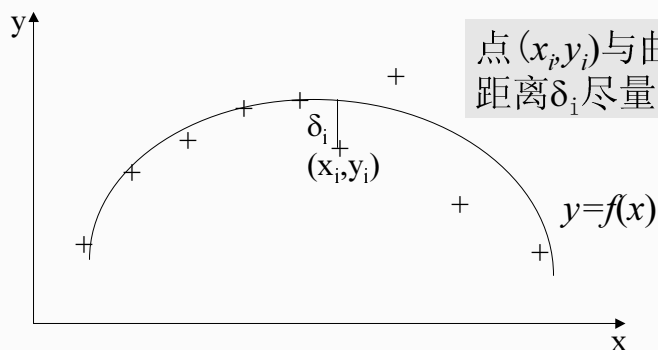
加工时需要x每  
改变0.1时的y值



## 拟合的 基本原理

### 曲线拟合问题的提法

已知一组（二维）数据，即平面上  $n$  个点  $(x_i, y_i)$ ,  $i=1, \dots, n$ , 寻求一个函数（曲线） $y=f(x)$ , 使  $f(x)$  在某种准则下与所有数据点最为接近，即曲线拟合得最好。



点  $(x_i, y_i)$  与曲线  $y=f(x)$  的距离  $\delta_i$  尽量小,  $i=1, \dots, n$

### 曲线拟合常用解法——线性最小二乘法的基本思路

先选定一组函数  $r_1(x), r_2(x), \dots, r_m(x)$ ,  $m < n$ , 令

$$f(x) = a_1 r_1(x) + a_2 r_2(x) + \dots + a_m r_m(x) \quad (1)$$

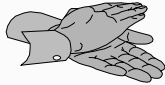
其中  $a_1, a_2, \dots, a_m$  为待定系数。

确定  $a_1, a_2, \dots, a_m$  的准则（最小二乘准则）：

使  $n$  个点  $(x_i, y_i)$  与曲线  $y = f(x)$  的距离  $\delta_i$  的平方和最小。

$$\begin{aligned} \text{记 } J(a_1, a_2, \dots, a_m) &= \sum_{i=1}^n \delta_i^2 = \sum_{i=1}^n [f(x_i) - y_i]^2 \\ &= \sum_{i=1}^n \left[ \sum_{k=1}^m a_k r_k(x_i) - y_i \right]^2 \end{aligned} \quad (2)$$

问题归结为：求  $a_1, a_2, \dots, a_m$  使  $J(a_1, a_2, \dots, a_m)$  最小。



### 线性最小二乘法的求解

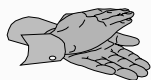
$$J(a_1, a_2, \dots, a_m) = \sum_{i=1}^n \left[ \sum_{k=1}^m a_k r_k(x_i) - y_i \right]^2 \quad (2)$$

$$\frac{\partial J}{\partial a_k} = 0 \Rightarrow \begin{cases} \sum_{i=1}^n r_1(x_i) \left[ \sum_{k=1}^m a_k r_k(x_i) - y_i \right] = 0 \\ \dots \dots \\ \sum_{i=1}^n r_m(x_i) \left[ \sum_{k=1}^m a_k r_k(x_i) - y_i \right] = 0 \end{cases} \quad (3)$$

$(k = 1, \dots, m)$

$$\text{记 } R = \begin{bmatrix} r_1(x_1) & \dots & r_m(x_1) \\ \dots & \dots & \dots \\ r_1(x_n) & \dots & r_m(x_n) \end{bmatrix}, \quad a = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$(3) \Rightarrow (R^T R)a = R^T y \quad (4)$$



## 线性最小二乘法的求解

$$(R^T R)a = R^T y \quad (4)$$

当  $R^T R$  可逆时(4)有唯一解

$$a = (R^T R)^{-1} R^T y \quad (5)$$

$$R = \begin{bmatrix} r_1(x_1) & \cdots & r_m(x_1) \\ \cdots & \cdots & \cdots \\ r_1(x_n) & \cdots & r_m(x_n) \end{bmatrix}_{n \times m}$$

### 问题

1. 怎样选择  $\{r_1(x), \dots, r_m(x)\}$ , 以保证系数  $\{a_1, \dots, a_m\}$  有唯一解?

### 分析

$\{a_1, \dots, a_m\}$  有唯一解  $\leftarrow R^T R$  可逆  $\leftarrow \text{Rank}(R^T R) = m$   
 $\leftarrow \text{Rank}(R) = m \quad \leftarrow R$  列满秩  
 $\leftarrow \{r_1(x), \dots, r_m(x)\}$  在  $x_i$  点线性无关 ( $i=1, \dots, n$ )

### 问题

2. 为什么要规定  $m < n$ ? 若  $m = n$  或  $m > n$ , 会如何?

### 分析

强令  $f(x_i) = a_1 r_1(x_i) + \dots + a_m r_m(x_i) = y_i$  ( $i=1, \dots, n$ )  
 (即曲线  $f(x)$  过全部数据点, 此时  $J=0$ ) 得

$$Ra = y, \quad R = \begin{bmatrix} r_1(x_1) & \cdots & r_m(x_1) \\ \cdots & \cdots & \cdots \\ r_1(x_n) & \cdots & r_m(x_n) \end{bmatrix}_{n \times m} \quad \begin{aligned} a &= [a_1, \dots, a_m]^T \\ y &= [y_1, \dots, y_n]^T \end{aligned}$$

若  $m > n$ ,  $a$  有无穷多解

若  $m = n$ ,  $R$  可逆时  $a$  有唯一解

若  $m < n$ , 超定方程组,  $a$  无解; 求最小二乘解

### 问题

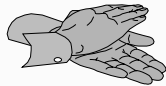
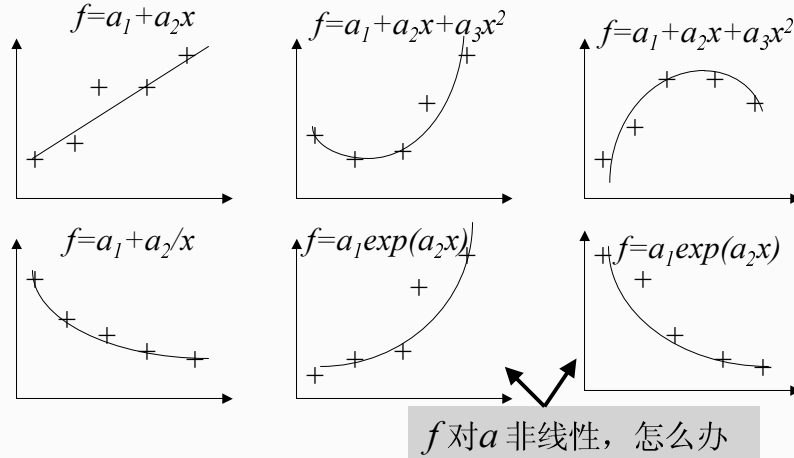
3. 线性最小二乘中的“线性”指的是什么

$f(x)$  对  $a$  线性, 于是求解线性方程组  $(R^T R)a = R^T y$

### 线性最小二乘拟合中函数 $\{r_1(x), \dots, r_m(x)\}$ 的选取

1. 通过机理分析建立数学模型来确定  $f(x)$

2. 将数据  $(x_i, y_i) \ i=1, \dots, n$  作图，通过直观判断确定  $f(x)$



### 用 MATLAB 作线性最小二乘拟合

1. 作多项式  $f(x)=a_1x^m+\dots+a_mx+a_{m+1}$  拟合，可利用已有程序：

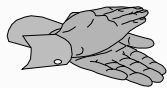
输入：数据 (同长度数组)； (拟合多项式次数)

输出：系数 (数组)。

多项式在 点的值：

2. 对超定方程组  $R_{n \times m} a_{m \times 1} = y_{n \times 1} \ (m < n)$

仍用  $a = R \setminus y$  可得最小二乘意义下的解



## 用MATLAB作线性最小二乘拟合

例. 由数据

温度 $t(^{\circ}\text{C})$	20.5	32.7	51.0	73.0	95.7
电阻 $R(\Omega)$	765	826	873	942	1032

拟合  $R=a_1t+a_2$



1. 用命令

得到  $a_1=3.3940$ ,  $a_2=702.4918$

2. 直接用  $a = R \backslash y$ , 结果相同。



## 给药方案 —— 拟合问题实例

问题

1. 在快速静脉注射的给药方式下, 研究血药浓度 (单位体积血液中的药物含量) 的变化规律。
2. 给定药物的最小有效浓度和最大治疗浓度, 设计给药方案 (每次注射剂量, 间隔时间)。

分析

实验: 血药浓度数据  $c(t)$  ( $t=0$  注射 300mg)

$t$ (h)	0.25	0.5	1	1.5	2	3	4	6	8
$c$ ( $\mu\text{g/ml}$ )	19.21	18.15	15.36	14.10	12.89	9.32	7.45	5.24	3.01

## 给药方案 —— 拟合问题实例

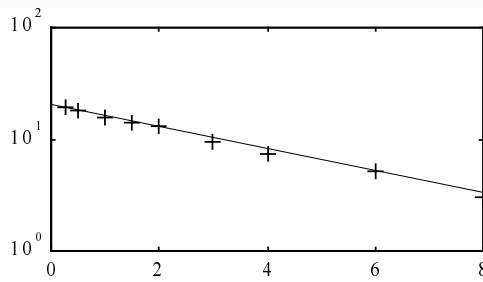
### 实验数据作图



xueyao1

### 半对数坐标系

### 下 $c(t)$ 的图形

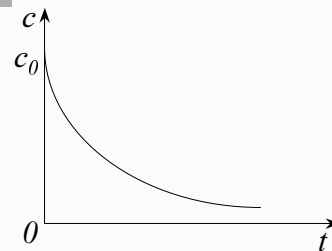


实验数据

### 负指数规律

$$c(t) = c_0 e^{-kt}$$

$c, k$  为待定系数



理论：用一室模型研究血药浓度变化规律

## 模型假设

1. 机体看作一个房室，室内血药浓度均匀——一室模型
2. 药物排除速率与血药浓度成正比，比例系数  $k(>0)$
3. 血液容积  $v$ ， $t=0$  注射剂量  $d$ ，血药浓度立即为  $d/v$

## 模型建立

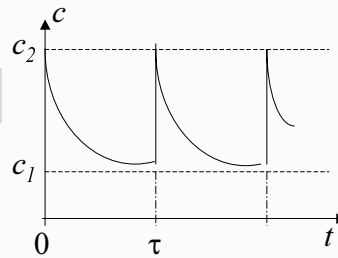
由假设2  $\frac{dc}{dt} = -kc$

由假设3  $c(0) = d/v$

$$c(t) = \frac{d}{v} e^{-kt}$$

### 给药方案设计

- 设每次注射剂量 $D$ , 间隔时间 $\tau$
- 血药浓度 $c(t)$  应 $c_1 \leq c(t) \leq c_2$
- 初次剂量 $D_0$  应加大



给药方案记作  $\{D_0, D, \tau\}$       $D_0 = vc_2, D = v(c_2 - c_1)$

$$c_1 = c_2 e^{-k\tau} \Rightarrow \tau = \frac{1}{k} \ln \frac{c_2}{c_1}$$

给定 $c_1=10, c_2=25$ , 为确定  $\{D_0, D, \tau\}$  只需确定参数  $k, v$

### 参数估计

由实验数据拟合曲线 $c(t)$ 以估计 $k, v$

$$c(t) = \frac{d}{v} e^{-kt} \quad \text{参数线性化} \quad \Rightarrow \quad \ln c = \ln(d/v) - kt$$

$$y = \ln c, \quad a_1 = -k, \quad a_2 = \ln(d/v) \quad \Rightarrow \quad y = a_1 t + a_2$$

用实验数据作线性最小二乘拟合

$$a_1 = -0.2347, a_2 = 2.9943$$

$$(d = 300) \quad \Rightarrow \quad k = 0.2347 (1/h), \quad v = 15.02 (l)$$



xueyao2



### 给药方案设计

$$k = 0.2347 (1/h), \quad v = 15.02 (l) \quad c_1=10, c_2=25$$

$$D_0 = vc_2, \quad D = v(c_2 - c_1) \quad \tau = \frac{1}{k} \ln \frac{c_2}{c_1}$$

$$\Rightarrow D_0 = 375.5, D = 225.3, \tau = 3.9$$

$$D_0 = 375 (mg), \quad D = 225 (mg), \quad \tau = 4 (h)$$

$$c(t) = \frac{d}{v} e^{-kt} \Rightarrow \ln c = \ln(d/v) - kt$$

思考：取对数化为线性最小二乘, 对结果有影响吗?



### 布置“插值与拟合”实验

#### 目的

1. 掌握用MATLAB计算三种插值的方法, 并对结果作初步分析;
2. 掌握用MATLAB作线性最小二乘的方法;
3. 用插值或拟合方法解决实际问题, 注意二者的联系和区别。

#### 内容

1. d; 8; 9。