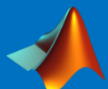


第4章 插值与拟合2

- MATLAB拟合函数
 - polyfit
 - regress
 - nlinfit



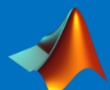
上讲内容

- interp1, pchip, spline的使用

1) 函数的功能是什么？

2) 函数要求的输入变量是什么？这些输入变量有什么特殊要求？

3) 函数返回的输出变量是什么？



MATLAB程序基本结构

```
function [Output]=FcnName (Input)
```

%Output和Input分别为输入和输出变量，有时并不需要，则函数声明语句可以更改为 `function FcnName`

%主函数的功能完成主要计算过程并进行输出

%Step 1: 选择合适的核心的求解函数

%Step 2: 根据求解函数输入变量，按顺序进行赋值

%Step 3: 在核心求解函数之后，输出求解结果

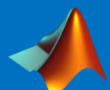
(`disp, fprintf, plot`)

```
function [Output2]=SubFunction (Input2)
```

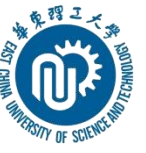
% 子函数主要功能是描述待求解问题，供主函数使用

% 子函数通常计算Input2指定处的目标函数值。

% 子函数与主函数是两个相对独立的空间，变量并不通用



MATLAB程序基本结构



```
function Work07_3
a=250;T0=450;b=10000;c=exp20;d=0.25;
x0=50;
x=fzero(@Tout,x0)
function y=Tout(x)
T = T0 + a*x;
y = c*(1-x)^2*d*exp(-b/T) - x;
disp('x')
```



MATLAB最小二乘法拟合函数

一元多项式函数



polyfit

一元线性函数



polyfit

多元线性函数



regress

非线性函数

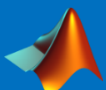


nlinfit

样条函数



csaps
spap2
spaps



polyfit

polyfit是MATLAB的多项式拟合函数

调用格式： $p = \text{polyfit}(x, y, n)$

输入变量： x , y 为待拟合数据，一般为向量；
 n 表示多项式的最高阶数；

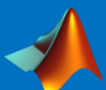
输出变量： p 为拟合所得多项式的系数向量（按降幂排列）

$$\text{length}(p) = n + 1$$

polyval可用于计算拟合多项式在指定点的值

调用格式为： $y = \text{polyval}(p, x)$ 。

输入变量： p 为多项式向量；
 x 为指定点

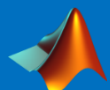


例题

已知x和y的数据由以下两条命令生成，试采用polyfit函数将此数据拟合为一个四次多项式。

```
x=[1.5:5.5,6:8]
```

```
y=0.02*x.^4+0.3*x.^3-5*x+0.1+20*rand(size(x))
```



例题

```
function PolyFitting
```

```
x=[1.5:5.5,6:8];
```

```
y=0.02*x.^4+0.3*x.^3-  
5*x+0.1+20*rand(size(x));
```

```
plot(x,y,'o'),hold on
```

```
p=polyfit(x,y,4);
```

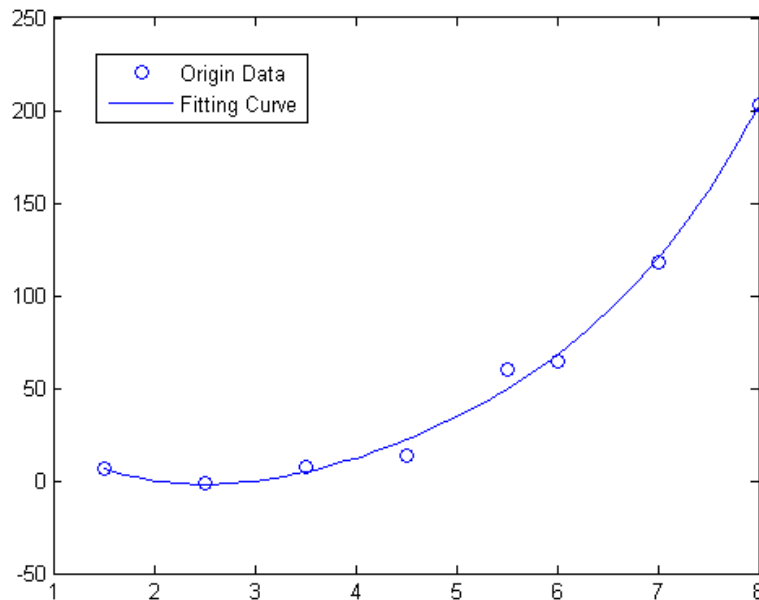
```
xx=1.5:0.1:8;
```

```
yy=polyval(p,xx);
```

```
plot(xx,yy,'-')
```

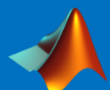
```
legend('Origin Data','Fitting Curve')
```

```
p=[0.1953, -3.1774, 24.0334, -72.5310, 71.1817]
```



注：1.拟合时，采用绘图的方法检验原始数据和拟合效果时常规做法，必须掌握。

2. rand函数生成的随机数每次运行时会有所不同，因此结果会略有不同。



例题

某化学反应其反应产物的浓度随时间变化的数据如下：

时间t	5	10	15	20	25	30	35	40	45	50	55
浓度y	1.27	2.16	2.86	3.44	3.87	4.15	4.37	4.51	4.58	4.62	4.64

试用最小二乘法关联y与t的关系。

1. 对实验点绘图，得：

由曲线形状特点，可用 $y = ae^{b/t}$
拟合实验数据

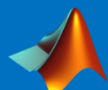
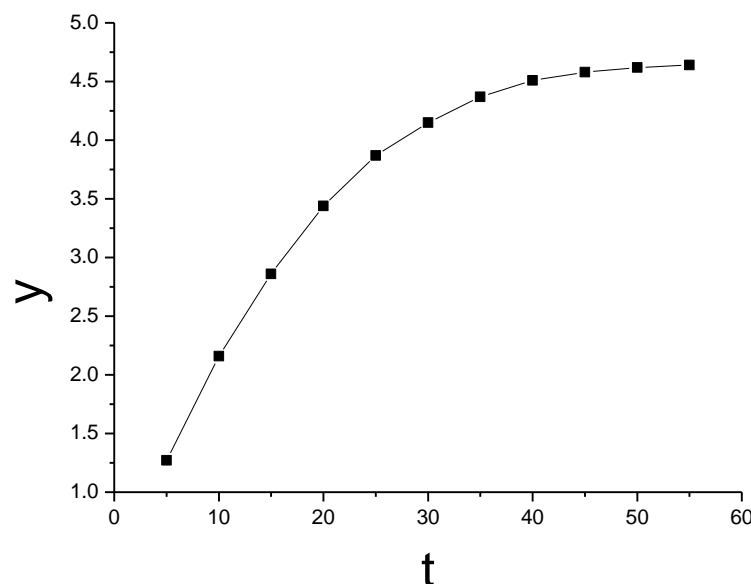
2. 线性化变换

令： $Y = \ln y$, $X = 1/t$

则 $y = ae^{b/t}$



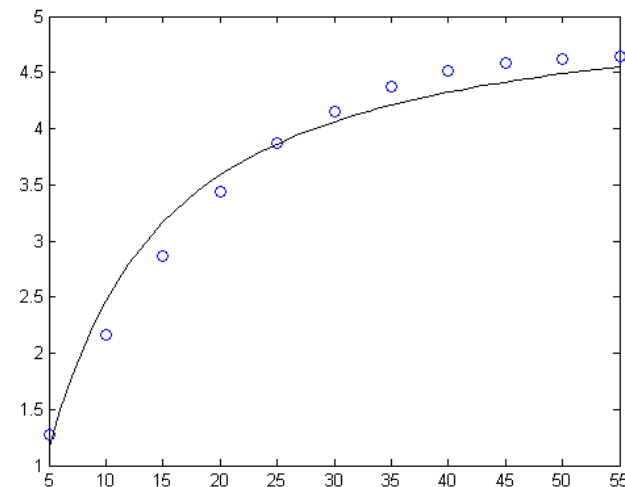
$Y = \ln a + bX$



例题

(3) 计算参数，程序如下：

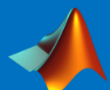
```
function FittingCon
t=[5 10 15 20 25 30 35 40 45
50 55];
y=[1.27 2.16 2.86 3.44 3.87
4.15 4.37 4.51 4.58 4.62
4.64];
plot(t,y,'bo'),hold on
Y=log(y);X=1./t;
p=polyfit(X,Y,1);
a=exp(p(2))
b=p(1)
f=@(x) a*exp(b./x);
plot(5:55,f(5:55),'k-')
xlabel('Time'),ylabel('Con')
```



执行结果：

a=5.2693

b=-7.7783



多元线性拟合函数regress

常用调用格式为：`b=regress(y, x)`

输入变量：

y：因变量数据，**m**行1列的数值向量；

x：自变量数据，**m**行**n**列或**m**行**n+1**列的矩阵；当拟合方程包括常数项时，需要在**x**的第一列前增加一列1，此时为**n+1**列；

$$x = \begin{bmatrix} 1 & x_{12} & \cdots & x_{1n} \\ 1 & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{m2} & \cdots & x_{mn} \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_m \end{bmatrix} \quad \begin{array}{l} n \text{ 为变量个数,} \\ m \text{ 为实验次数。} \end{array}$$

输出变量：

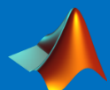
b：拟合的多元线性方程参数；`length(b)=n+1`



例题

已知 $x_1=1:6$; $x_2=1.5:0.2:2.5$; $y=[5.9512 \ 10.6730 \ 15.5676 \ 20.6234 \ 25.8599 \ 31.2622]$; 试采用以上数据拟合
 $y=a*x_1+b*x_2+c*x_1^2+d*x_1*x_2$ 中的系数。

将 x_1 , x_2 , x_1^2 和 x_1*x_2 视为自变量, 则以上的拟合为多元线性拟合, 可以采用`regress`函数求解。

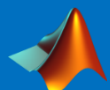


例题

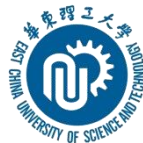
```
function LinearFitting
x1=1:6;
x2=1.5:0.2:2.5;
y=[5.9512 10.6730 15.5676 20.6234 25.8599
31.2622];
X=[x1',x2',x1'.^2,x1'.*x2'];
b=regress(y',X)
ycal=b(1)*x1+b(2)*x2+b(3)*x1.^2+b(4)*x1.*x2;
plot(y,ycal,'bo')
hold on
refline(1,0)
xlabel('Original Data'),ylabel('Fitting Data')
```

$$y=0*x1+1.0775x2-0.5687*x1^2+3.2694x1*x2$$

注意绘图语句的方法



例题



设某气体反应可表示为： $A+B+C \rightarrow D$

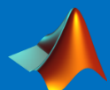
其反应动力学方程可用下列非线性方程表示：

$$V = KP_A^{n_1} P_B^{n_2} P_C^{n_3}$$

式中， V 为反应速度， K 为反应速度常数，

P_A P_B P_C 依次为A、B、C的分压，下表为实验测定的不同分压下的 V 值。试确定此气体反应动力学方程。

P_A	P_B	P_C	V	P_A	P_B	P_C	V
8.998	8.298	2.699	8.58	7.001	3.900	9.895	2.18
8.199	7.001	4.402	6.05	3.310	3.401	9.796	2.11
7.901	6.203	5.900	4.73	6.501	2.601	10.903	1.88
7.001	4.302	8.199	3.35	7.997	2.199	17.797	1.04



例题

解：（1）首先将动力学方程的两边取对数：

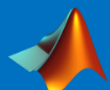
$$\ln V = \ln K + n_1 P_A + n_2 P_B + n_3 P_C$$

令：

$$y = \ln V, x_1 = \ln P_A, x_2 = \ln P_B, x_3 = \ln P_C$$

则上式变换为：

$$y = \ln K + n_1 x_1 + n_2 x_2 + n_3 x_3$$



例题

$$y = \ln K + n_1 x_1 + n_2 x_2 + n_3 x_3$$

function KineticsFitting

```
PA=[8.998, 8.199,7.901, 7.001, 7.001, 3.310, 6.501, 7.997];
```

```
PB=[8.298,7.001, 6.203, 4.302, 3.900, 3.401, 2.601, 2.199];
```

```
PC=[2.699, 4.402,5.900 ,8.199, 9.895 ,9.796,10.903,17.797];
```

```
V=[8.58, 6.05, 4.73, 3.35, 2.18, 2.11, 1.88 ,1.04];
```

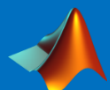
```
x1=log(PA);x2=log(PB);x3=log(PC);y=log(V);
```

```
x=[ones(8,1), x1', x2', x3'];
```

```
b=regress(y',x);
```

```
fprintf('K=%f\tn1=% .4f\tn2=% .4f\tn3=% .4f\n'...
```

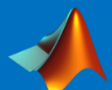
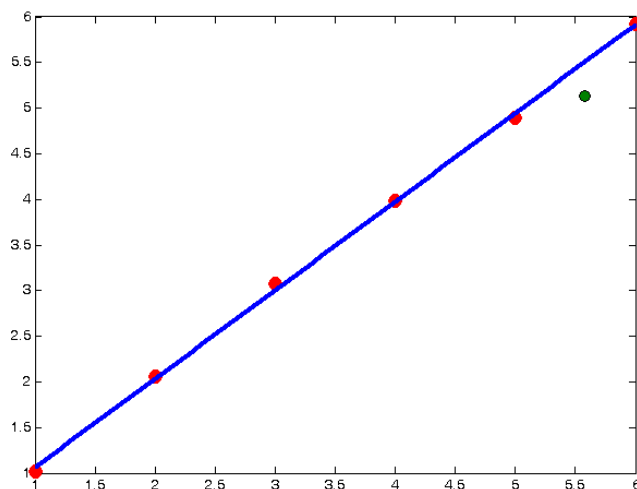
```
,exp(b(1)),b(2),b(3),b(4))
```



拟合效果的判断与检验

1. 图形直观判断（这是最基本的检验，如何可能必须使用！）

- 数据点是否有异常，对异常数据点应检查是否为输入有误，或实验不正常所致
- 拟合曲线是否符合原始数据点趋势
- 数据点是否均匀分布与拟合曲线两侧

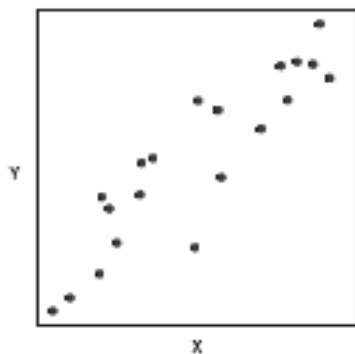


拟合效果的判断与检验*

2. 相关系数

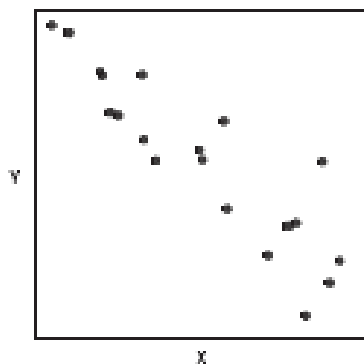
相关系数越接近1越好

$$R^2 = \sum (y_i^* - \bar{y})^2 / \sum (y_i - \bar{y})^2$$



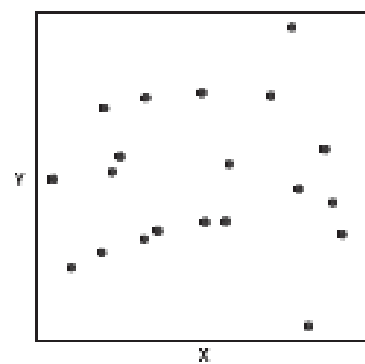
$r=0.9$

正线性相关



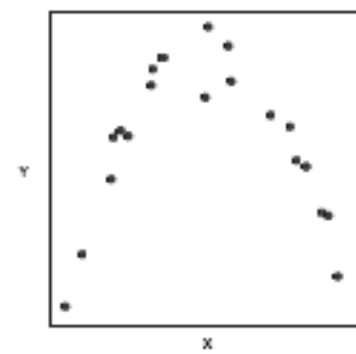
$r=-0.9$

负线性相关



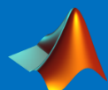
$r=0$

线性无关



$r=-0.09$

非线性相关



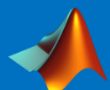
3. 置信区间

- 置信区间越小越好，至少不能大于拟合所得参数的数量级

残差的标准偏差

$$\hat{s} = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (y_i - y_i^*)^2}$$

如果试验的随机误差服从正态分布，则实验值落在区间 $[y_i \pm 2\hat{s}]$ 之内的概率为95%



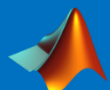
拟合效果的判断与检验*

4. F检验

- F检验实际就是方差分析，检验模型计算值是否与自变量存在函数关系。

$$F = \frac{SS_R}{df_R} / \frac{SS_e}{df_e} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{1} / \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2}$$

在给定的显著性水平 α 下，从F分布表查得 $F_{\alpha}(1, n-2)$ ， α 一般取0.01和0.05，当 $F < F_{0.05}$ 时则称x与y没有明显的关系，拟合方程不可信， $F_{0.05} < F < F_{0.01}$ ，则x, y有显著关系， $F > F_{0.01}$ 时，x, y有十分显著的关系

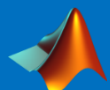


regress回归效果的判断与检验*

```
alpha=0.05;
```

```
[b,bint,r,rint,stats]=regress(y,x,alpha)
```

- b,bint为回归系数估计值及他们的置信区间;
- r,rint为残差及其置信区间;
- stats为检验回归模型的统计量, 包括四个元素
 1. 第一个元素为相关系数 R^2 ;
 2. 第二个元素为F统计量的值;
 3. 第三个为相对于所得F统计量值的概率p, 当 $p < \alpha$ 认为回归有意义;



非线性直接拟合原理

非线性方程建议采用直接非线性最小二乘法拟合

理论基础：泰勒展开

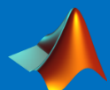
对于非线性函数 $y_k = f(x_k, b_1, b_2, \dots, b_n)$

若 b_i 的近似值为 $b_i^{(0)}$ ，误差为 Δ_i ，则 $b_i = b_i^{(0)} + \Delta_i$

当初值给定时对非线性函数在初值 $b_i^{(0)}$ 附近作泰勒展开，并略去 Δ_i 的二次以上的高次项，可以得到：

$$y_k = f(x_k, b_1, b_2, \dots, b_n) \approx f_{k,0} + \frac{\partial f_{k,0}}{\partial b_1} \Delta_1 + \frac{\partial f_{k,0}}{\partial b_2} \Delta_2 + \dots + \frac{\partial f_{k,0}}{\partial b_n} \Delta_n$$

其中： $f_{k,0} = f(x_k, b_1^{(0)}, b_2^{(0)}, \dots, b_n^{(0)})$



欲使Q最小，按极值的必要条件，应满足：





Δ_i



$$y_k = f(x_k, b_1, b_2, \dots, b_n) \approx f_{k,0} + \frac{\partial f_{k,0}}{\partial b_1} \Delta_1 + \frac{\partial f_{k,0}}{\partial b_2} \Delta_2 + \dots + \frac{\partial f_{k,0}}{\partial b_n} \Delta_n$$

直接非线性拟合函数

MATLAB求解非线性拟合的函数较多，多采用最优化方法，如lsqcurvefit, lsqnonlin等。

nlinfit函数的简单调用格式为：

```
beta = nlinfit(x,y,fun,beta0)
```

- x , y 分别为 $m \times n$ 阶矩阵和 m 阶列向量（ n 为变量个数， m 为实验次数）；
- fun 是表示待拟合非线性函数的自定义函数，函数声明语句如下：

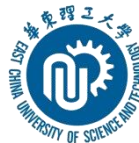
```
function y=fun(b,x)
```

其返回值 y 是待拟合的非线性函数在 x 处的函数值；

- $beta0$ 是回归系数的初值；
- $beta$ 是拟合所得非线性函数未知参数的值。



例题

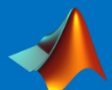


已知数据：

X1	1	2	3	4	5	6	7	8
Y1	15.3	20.5	27.4	36.6	49.1	65.6	87.8	117.6

拟合如下函数形式的曲线

$$y = a_1 e^{a_2 x}$$

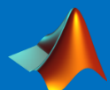
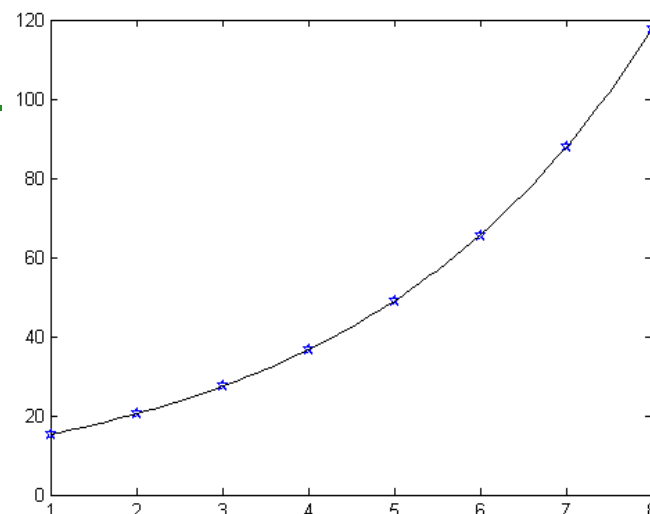


例题

```
function Cha4demoNL
x=1:8;
y=[15.3 20.5 27.4 36.6 49.1 65.6 87.8
117.6];
a0=[1,1];
a1=nlinfit(x,y,@objfun,a0)
plot(x,y,'bp',1:0.1:8,objfun(a1,1:0.1:8
),'k-')
```

%-----

```
function f=objfun(a,x)
f=a(1)*exp(a(2)*x);
```



例题

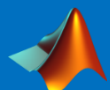
某催化剂活性Y与工作持续时间t的关系为：

$$Y = Ae^{(Bt+Ct^2)}$$

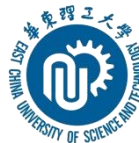
将下表所列的实验数据通过曲线拟合求系数A、B、C。

t(h)	0	27	40	52	70	89	106
Y(%)	100	82.2	76.3	71.8	66.4	63.3	61.3

本题先用线性变换法将模型转化为线性形式，采用polyfit 函数得参数初始值，再用非线性拟合函数nlinfit求解。



例题



```
function NlinFitDemo
t=[0 27 40 52 70 89 106]';
y=[100 82.2 76.3 71.8 66.4 63.3 61.3]';
Y=log(y);
p=polyfit(t,Y,2);
beta0=[exp(p(3)),p(2),p(1)];
format long
beta=nlinfit(t,y,@fun,beta0)
plot(t,y,'o',0:110,fun(beta,0:110),'-')
legend('Experiment','Fitting curve')
function y=fun(beta,t)
    y=beta(1)*exp(beta(2)*t+beta(3)*t.^2);
```



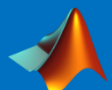
例题

已知某反应的动力学可以采用下式表示：

$$r = kC_A^a C_B^b$$

实验测得一组实验数据如下表所示。试采用nlinfit函数回归式中的k，a，b的值。

CA	0.13	0.2	0.33	0.4	0.53	0.62
CB	1.05	1.26	1.44	1.52	1.63	1.70
r	0.0076	0.0135	0.0213	0.0277	0.0360	0.0420



例题

function KineticFitting

CA=[0.13 0.2 0.33 0.4 0.53 0.62];

CB=[1.05 1.26 1.44 1.52 1.63 1.70];

r=[0.0076 0.0135 0.0213 0.0277 0.0360
0.0420];

k0=0.01;a0=1;b0=1;

beta=nlinfit([CA',CB'],r',@Rate,[k0,a0,b0])

rca1=Rate(beta,[CA',CB']);plot(r,rca1,'bo')

hold on, reffline(1,0)

xlabel('Experiment'),ylabel('Modeling')

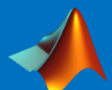
function y=Rate(beta,x)

CA=x(:,1);

CB=x(:,2);

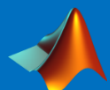
y=beta(1)*CA.^beta(2).*CB.^beta(3);

$$r = kC_A^a C_B^b$$

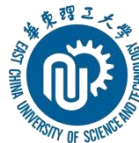


非线性拟合的注意事项

- 非线性拟合后一定要进行检验，检测拟合效果和可靠性；
- 对于高度非线性模型的拟合，结果与初始值密切相关：
 - 利用已知专业信息；
 - 采用线性化；
 - 顺序拟合：每次拟合一个参数
- 拟合效果取决于原始实验点的分布、质量和数量



nlinfit函数的求解选项



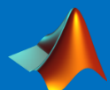
```
[b,r,J,COVB,mse] = nlinfit(X,y,fun,b0,options)
```

输入参数中的options包括：

选项	意义	默认值
MaxIter	最大迭代次数	100
TolFun	程序终止要求的残差平方和	1e-8
TolX	程序终止时要求的估计参数b的变化	1e-8
Display	是否显示求解中间过程	'off'

修改可以采用如下语句：

```
options=statset('TolFun',1e-10,'Display','iter')
```



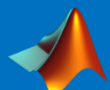
nlinfit函数拟合的检验*

nlinfit函数拟合的检验可以由以下命令实现：

```
[beta,r,J] = nlinfit(X,y,fun,beta0)
[yp,ci]=nlpredci(fun,X,beta,r,J)
ci=nlparci(beta,r,'jacobian',J,' alpha' ,alpha)
```

拟合方程预测值
的检验

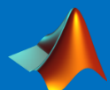
拟合方程参数的
检验



用最小二乘法拟合生成样条曲线

- 与样条插值不同，样条拟合并不要求曲线通过全部的数据点。
- 在有些情况下，拟合只是为以后的计算提供方便，不需要得到一个明确的有意义的函数，此时可采用样条拟合
- MATLAB样条工具箱提供的函数较多，常用的三个拟合函数如下：

函数名	曲线类型	拟合准则	是否平滑处理
csaps()	三次样条曲线	最小二乘法	是
spap2()	B样条曲线	最小二乘法	否
spaps()	B样条曲线	最小二乘法	是



函数csaps()的用法

功能：平滑生成三次样条函数，即对于数据 (x_i, y_i) ，所求的三次样条函数 $y=f(x)$ 满足

$$\min p \cdot \sum_i w_i (y_i - f(x_i))^2 + (1-p) \int \lambda(t) (D^2 f)^2 dt$$

调用格式： `sp=csaps(x,y,p)`

`ys=csaps(x,y,p,xx,w)`

输入参数： `x,y` --- 要处理的离散数据 (x_i, y_i) ;

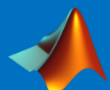
`p` ---- 平滑参数，取值区间为 $[0,1]$ 。当 $p=0$ 时，相当于最小二乘直线拟合；当 $p=1$ 时，相当于“自然的”三次样条插值，即相当于`csapi()`或`spline()`;

`xx` --- 用于指定在给定点`xx`上计算其三次样条函数值`ys`;

`w` --- 权值（权重），默认为1；

输出参数： `sp` --- 拟合得到的样条函数；

`ys` --- 在给定点`xx`上的三次样条函数值。

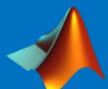


练习

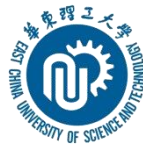
已知 $x=1:6$; $y=[2.29 \quad 5.70 \quad 11.03 \quad 18.28 \quad 27.45 \quad 38.54]$; a 和 b 满足以下关系式 $y=a+b*x+c*x^2$

试编写一个MATLAB函数，完成以上计算：

- 1) 采用polyfit函数求参数 a , b , c 的值
- 2) 采用regress函数求参数 a , b , c 的值
- 3) 采用nlinfit函数求参数 a , b , c 的值
- 4) 绘制曲线表示nlinfit函数的拟合效果



作业



公共邮箱下载文档：work09.pdf，直接打印
、完成后上交

