1 问题描述

由于钢水对材料的侵蚀作用，随着使用次数的增加，炼钢厂出钢时所用盛钢水的钢包的容积不断增大。现希望获得增大容积y与使用次数x之间的函数关系，实测得到如下数据：

使用次数x	2	3	4	5	6	7	8	9
增大容积y	6.7	8.2	9.58	9.5	9.7	10	9.96	9.99
使用次数x	10	11	12	13	14	15	16	
增大容积y	10.49	10.59	10.6	10.8	10.6	10.9	10.76	

（1）请用插值方法建立y与x之间的函数关系，画出散点图和插值函数曲线，从结果说明插值方法是否适合该问题。
（2）请分别按 从 ① $\frac{1}{y}=a+\frac{b}{x}$ 和② $y=ae^{\frac{b}{x}}$ 两种形式拟合建立y与x之间的函数关系，画出散点图和拟合函数曲线，并根据你选定的适当指标判断哪一种形式更好。

2 问题分析

本题是一个用插值和拟合方法来拟合函数的问题，其中

```
X=[2,3,4,5,6,7,8,9,10,11,12,13,14,15,16];
Y=
[6.7,8.2,9.58,9.5,9.7,10,9.96,9.99,10.49,10.59,10.6,10.8,10.6,10.9,10.76]
;
```

由于本题的现实需求，使最终得出的函数能够反映增大容积y与使用次数x之间的关系，那么就要求最终的函数满足：预测下一个值，函数连续且可导。

3 代码实现

3.1 插值方法

拉格朗日插值法

```
function f=lagrange(x,X,Y)
    n=length(X);
    f=0;
    for i=1:n
        up=1;
        down=1;
        for j=1:n
```

```
            if i~=j
                up=up*(x-X(j));%求基函数分子
            end
        end
        for j=1:n
            if i~=j
                down=down*(X(i)-X(j));%求基函数分母
            end
        end
        lk=up/down;%基函数
        f=f+Y(i)*lk;%累加
    end
end
```

牛顿插值法

```
function result=different(X,Y)%求差商的函数
    n=length(X);
    result=0;
    for i=1:n
        down=1;
        for j=1:n
            if i~=j
                down=down*(X(i)-X(j));%求分母
            end
        end
        lk=Y(i)/down;
        result=result+lk;
    end
end

function f=newton(x,X,Y)
    f=Y(1);
    n=length(X);
    for i=2:n
        k=1;
        result=different(X(1:i),Y(1:i));
        k=result;
        for j=1:i-1
            k=k*(x-X(j));
        end
        f=f+k;
    end
end
```

分段插值法

```
function f=cut(x,X,Y)
    f_set=[];%每个区间的函数的集合
```

```
    n=length(X);
    for i=1:n-1
        X_part=X(i:i+1);
        Y_part=Y(i:i+1);
        f_element=lagrange(x,X_part,Y_part);
        f_set=[f_set,f_element];
    end
    f=f_set(1).*(x>=X(1)&x<=X(2));
    disp(f);
    for i=2:length(f_set)
        f=f+f_set(i).*(x>X(i)&x<=X(i+1));%对应区间有对应的函数
    end
end
```

三次样条插值法

```
function f=spline(x,X,Y)
X=X(:);
Y=Y(:);
n=length(X);
h=diff(X);
A=zeros(n,n);
b=zeros(n,1);
A(1,1)=1;
A(n,n)=1;
for i=2:n-1
    A(i,i-1)=h(i-1);
    A(i,i)=2*(h(i-1)+h(i));
    A(i,i+1)=h(i);
    b(i)=3*((Y(i+1)-Y(i))/h(i)-(Y(i)-Y(i-1))/h(i-1));
end
M=A\b;% 求解线性方程组，得到二阶导数M
f=0;
for i=1:n-1 %构建分段符号函数
    segFunc=Y(i)+M(i)/2*(x-X(i))^2+(Y(i+1)-Y(i)-M(i)*h(i)^2/2)/h(i)*(x-
X(i))+(M(i+1)-M(i))/(6*h(i))*(x-X(i))^3;
    f=f+segFunc*heaviside(x-X(i))*heaviside(X(i+1)-x);
end
f=f+(Y(n)+M(n)/2*(x-X(n))^2)*heaviside(x-X(n));
end
```

3.2 拟合方法

以$\frac{1}{y}=a+\frac{b}{x}$形式来拟合

$$\begin{bmatrix} m & \sum_{i=1}^{m} x_i & \cdots & \sum_{i=1}^{m} x_i^n \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 & \cdots & \sum_{i=1}^{m} x_i^{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{m} x_i^n & \sum_{i=1}^{m} x_i^{n+1} & \cdots & \sum_{i=1}^{m} x_i^{2n} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} y_i \\ \sum_{i=1}^{m} x_i y_i \\ \vdots \\ \sum_{i=1}^{m} x_i^n y_i \end{bmatrix}$$

以上图所示的原理，只有a，b两个参数，n=2；m为拟合点的数量，m=15。
所用的拟合点要经过处理，x每个值要取倒数，y每个值也取倒数。

```
function result=getadd(m,n,X)%求系数矩阵中每个值的通用函数
    result=0;
    for i=1:m
        result=result+power(X(i),n);
    end
end

function f=fitting1(x,X,Y)%第一种方式的拟合函数
    a=0;b=0;
    A=zeros(2,2);
    B=zeros(2,1);
    m=length(X);
    A(1,1)=m;
    A(1,2)=getadd(m,1,X);
    A(2,1)=A(1,2);
    A(2,2)=getadd(m,2,X);
    B(1,1)=getadd(m,1,Y);
    result=0;
    for i=1:m
        result=result+X(i)*Y(i);
    end
    B(2,1)=result;
    M=A\B;%求解a, b
    a=M(1);b=M(2);
    f=1/(a+b/x);
end

for i=1:15
    X_1=[X_1,1/X(i)];
    Y_1=[Y_1,1/Y(i)];
end
y5=fitting1(x,X_1,Y_1);
```

以$y=ae^{\frac{b}{x}}$形式来拟合
先对两侧取对数：$lny=lna+\frac{b}{x}$
所用的拟合点要经过处理，x每个值要取倒数，y每个值要取对数。而且求得的a值要再求指数。

```
function f=fitting2(x,X,Y)
```

```
    a=0;b=0;
    A=zeros(2,2);
    B=zeros(2,1);
    m=length(X);
    A(1,1)=m;
    A(1,2)=getadd(m,1,X);
    A(2,1)=A(1,2);
    A(2,2)=getadd(m,2,X);
    B(1,1)=getadd(m,1,Y);
    result=0;
    for i=1:m
        result=result+X(i)*Y(i);
    end
    B(2,1)=result;
    M=A\B;
    a=exp(M(1));b=M(2);
    f=a*exp(b/x);
end
for i=1:15
    X_2=[X_2,1/X(i)];
    Y_2=[Y_2,log(Y(i))];
end
y6=fitting2(x,X_2,Y_2);
```

拟合方式评判标准，选择用最大偏差与均方误差

```
y5_fit=y5_handle(X);
y5_fit_diff=Y-y5_fit;
[a_1,index_1]=max(abs(y5_fit_diff));
y5_fit_max=y5_fit_diff(index_1);
mes_1=mean((Y-y5_fit).^2);
fprintf("第一个拟合方式最大偏差%.16f\n",y5_fit_max);
fprintf("第一个拟合方式均方误差%.16f\n",mes_1);

y6_fit=y6_handle(X);
y6_fit_diff=Y-y6_fit;
[a_2,index_2]=max(abs(y6_fit_diff));
y6_fit_max=y6_fit_diff(index_2);
mes_2=mean((Y-y6_fit).^2);
fprintf("第二个拟合方式最大偏差%.16f\n",y6_fit_max);
fprintf("第二个拟合方式均方误差%.16f\n",mes_2);
```

4 结果分析

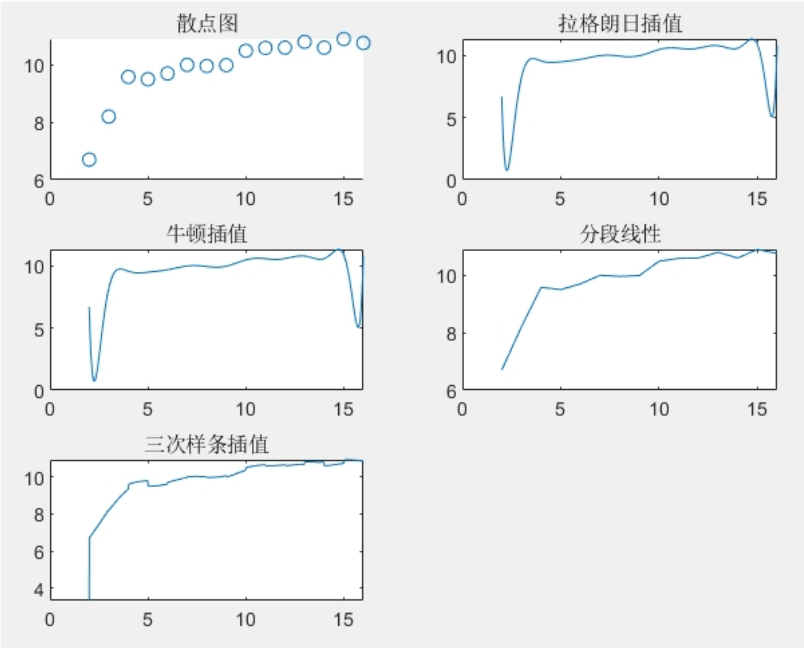
4.1 插值问题

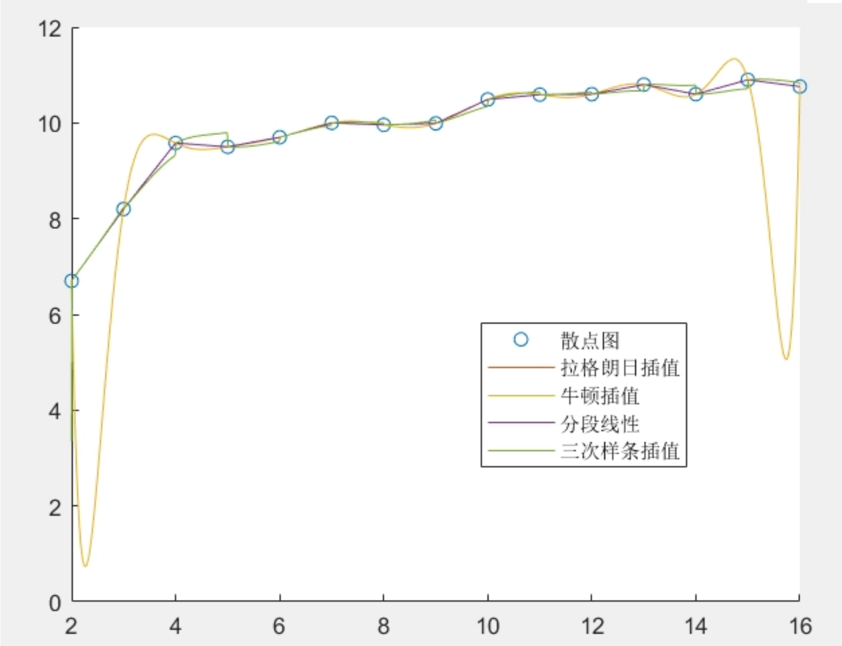
```
y1=lagrange(x,X,Y);
y2=newton(x,X,Y);
y3=cut(x,X,Y);
y4=spline(x,X,Y);

y1_handle=matlabFunction(y1);
y2_handle=matlabFunction(y2);
y3_handle=matlabFunction(y3);
y4_handle=matlabFunction(y4);

u=2:0.01:16;
y1_y=y1_handle(u);
y2_y=y2_handle(u);
y3_y=y3_handle(u);
y4_y=y4_handle(u);

figure;
subplot(3,2,1);
scatter(X,Y);
title("散点图");
subplot(3,2,2);
plot(u,y1_y);
title("拉格朗日插值");
subplot(3,2,3);
plot(u,y2_y);
title("牛顿插值");
subplot(3,2,4);
plot(u,y3_y);
title("分段线性");
subplot(3,2,5);
plot(u,y4_y);
title("三次样条插值");
```





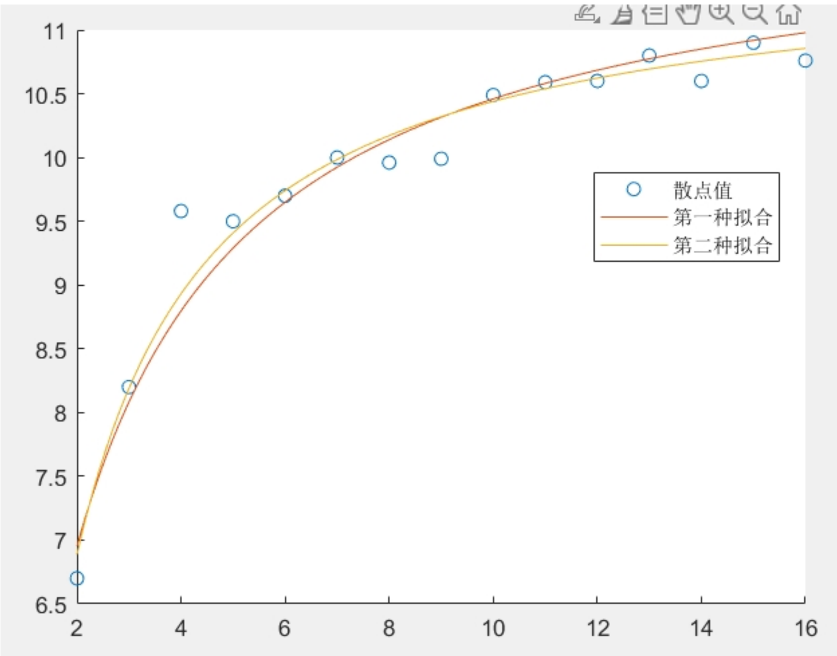
对拉格朗日插值法与牛顿插值法来说，出现Runge现象，函数两端点出波动较大。虽然符合过所有数据点，但是对预测下一个点不适合，而且整体趋势不符合。

对分段线性插值，其虽然趋势较符合，但是存在不可导点，而且由于他是分段函数，完全不能预测。

对三次样条插值，大题趋势也较符合，但是相邻节点之间的函数段波动较无规律，也不适合预测

综上，插值问题不适合这个问题。

4.2 拟合问题



第一个拟合方式最大偏差0.7832335121848395

第一个拟合方式均方误差0.0668841567767085

第二个拟合方式最大偏差0.6487198132331873

第二个拟合方式均方误差0.0451919476202879

可见第二种方法的拟合方式最大偏差更小而且均方误差也更小，故第二种拟合方式最好。
且无论从适合预测还是函数的连续性与可导性来看，拟合方法更合适

输出对应的参数为 11.5850031223202912
-1.0406067392164584

```
(3260888484047981*exp(-4686476122974469/(4503599627370496*x)))/281474976710656
```

故$y = \frac{3260888484047981}{281474976710656} e^{\frac{-4686476122974469}{4503599627370496}\frac{1}{x}}$

5 个人心得

对于实际问题上，要求拟合的函数满足能够预测下一个值、函数连续可导的要求，这种问题一般的插值法难以胜任，对于拉格朗日和牛顿插值法会有runge现象，对于分段插值不能预测，对于样条插值预测不准，如果没有要过所有数据点的硬性要求，尽量不用插值法，只在节点处吻合而函数整体误差很大。而拟合法虽然不过所有点，但是整体上反映了数据的趋势。适合在实际问题中用于预测等，且拟合函数形式的选择非常重要，不同的拟合形式会影响拟合的精度，找到最适合的拟合形式非常重要。