# 一、问题重述

## 1.1问题背景

在传统的燃煤发电过程中，锅炉是一种重要的热能动力设备。它通过燃烧煤粉颗粒释放出热量，将水加热成的水蒸汽，去用来推动汽轮机旋转并驱动发电机发电。锅炉的主要受热部分是炉内的水冷壁，它通常由数排钢管组合而成，分布在锅炉炉膛的四周，其内部均为流水，用于吸收因高温燃烧而产生的辐射热量。

## 1.2问题概述

然而，在实际生产过程中，水冷壁的温度变化需要尽可能的平稳，水冷壁温度不宜过高，以保证安全。根据附件信息拟解决下列问题：

1.分析各个水冷壁管道的温度数据，并给出刻画这些温度时间序列数据变化情况的特征。

2.对附件1中10个水冷壁管道的温度数据曲线进行评价，确定其中的最优工作曲线和最差工作曲线。

3.请利用附件1和附件2中的数据，分别建立10个水冷壁管道温度变化规律的数学模型，并对模型效果进行评价。

4.第10个水冷壁管道温度曲线如图2所示，从图中可以看出，在第3172个样本点后水冷壁出现明显的超温现象，请基于给出的数据，分析并定位引发超温现象的主要操作变量。

5.请针对第10个水冷壁管道温度曲线超温段建立优化模型，给出该超温段从第3172个样本开始的最优调节策略，满足操控的变量数尽量少、操作变量的调控量尽量小、优化调节后的工作曲线与问题2中的最优工作曲线的特征尽量吻合。

# 二、问题分析

本文要研究对象是燃煤锅炉，研究内容为炉膛内的温度变化情况，该问题描述了该系统在不同的管道和操作、状态条件下温度的变化情况，并在不同的条件下提出了不同的要求。

针对问题一，问题给出管道的温度变化数据，需要我们利用时间序列分析，分析其平均温度，温度变化情况(方差)，最高温度，最低温度，以此用来表达温度的时间序列数据变化的特征情况。

针对问题二，问题给出附近1中所有的温度数据，需要我们画出相应时间序列图像，进行分析，同时根据题目要求，温度变化需要尽可能平稳且水冷壁温度不宜过高，可以使用第一问的方差和最大值。从数据大小的角度将这两个指标定义为负向指标，利用数据归一化，将数据归一化至[0,1]，评价算法用熵权法求得权重后乘以归一化数据求和得出评价分数。

针对问题三，附件一包含10个管道的5000组温度数据，附件二中包含了111个操作变量和42个状态变量的温度数据，利用皮尔森相关性系数，找出附近1中10个管道与附近2中153的变量间相关性系数，求取平均值后，找出平均相关性系数最高的2项，利用1stopt遍历下较符合的公式，随后用1stopt和MATLAB找出函数解析式及其图像，并对其做出分析。

针对问题四，题目已经绘制第10个水冷壁管道温度曲线，找出在第3172个样本点后温度明显高于445摄氏度，需要我们求第十个管道的3000-5000温度数据对附件二的111个操作变量的皮尔森相关性系数，找出平均相关性系数最高的几项。

针对问题五，题目要求在管道十超温段的第3172个样本开始，在题目中的3个要求，需要我们利用逐步回归的方式，求解出管道十的近似公式，随后建立方差，最大值，与最优曲线差距、操作数这四个变量的目标函数，利用遗传算法的多目标优化算法，寻找出Pareto最优解。

# 三、模型假设

1、假设温度的改变都是用附录2提到的变量，无其他变量

2、假设不考虑锅炉因使用年限而造成的磨损。

3、假设不考虑锅炉内部气流、产生气体的影响

4、假设不考虑操作时间对锅炉温度的影响

5、假设不考虑附件1的十个管道仅和附件2的变量相关

# 四、符号说明

符号 符号说明

x1温度的平均值

max 温度的最大值

min 温度的最小值

DX 温度的方差

第项的权重

第项的评价分

第项的信息冗余度

第项的熵值

皮尔森相关性系数

# 五、模型建立与求解

## 5.1问题一

利用附件1所给的十个管道的数据求出均值、方差、最大值、最小值。我们用x1作为均值，DX作为方差，max作为最大值，min作为最小值。利用MATLAB求出各项数值，如表1所示

表1 管道均值、最大值、最小值、方差记录表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **管道** | **平均值** | **最大值** | **最小值** | **方差** |
| 管道1温度 | 374.3719 | 420.2000 | 338.6000 | 460.1841 |
| 管道2温度 | 397.9088 | 426.6000 | 356.8000 | 284.0029 |
| 管道3温度 | 396.7092 | 420.2000 | 364.2000 | 113.9759 |
| 管道4温度 | 387.6063 | 420.9000 | 350.6000 | 268.1675 |
| 管道5温度 | 391.3618 | 419.9000 | 360.0000 | 168.4203 |
| 管道6温度 | 376.8434 | 409.5000 | 338.0000 | 277.4512 |
| 管道7温度 | 386.1856 | 416.5000 | 350.0000 | 256.6979 |
| 管道8温度 | 378.6845 | 421.3000 | 351.2000 | 296.2075 |
| 管道9温度 | 408.0366 | 454.3000 | 368.7000 | 313.5170 |
| 管道10温度 | 424.9091 | 465.4000 | 378.8000 | 349.4685 |

## 5.2问题二

### 5.2.1建模准备

5.2.1.1熵权法

在信息论中，熵是对不确定性的一种度量。不确定性越大，熵就越大，包含的信息量越大；不确定性越小，熵就越小，包含的信息量就越小。

根据熵的特性，可以通过计算熵值来判断一个事件的随机性及无序程度，也可以用熵值来判断某个指标的离散程度，指标的离散程度越大，该指标对综合评价的影响（权重）越大。比如样本数据在某指标下取值都相等，则该指标对总体评价的影响为0，权值为0.

熵权法是一种客观赋权法，因为它仅依赖于数据本身的离散性。

1、对个样本，个指标，则为第个样本的第个指标的数值

2、指标的标准处理:异质指标同质化。由于各项指标的计量单位并不统一,因此在用它们计算综合指标前，先要进行标准化处理，即把指标的绝对值转化为相对值，从而解决各项不同质指标值的同质化问题。此外，正向指标和负向指标数值代表的含义不同(正向指标数值越高越好,负向指标数值越低越好)，因此，对于正负向指标需要采用不同的算法进行数据标准化处理。

正向指标：

负向指标：

3、计算第项指标下第个样本值占该指标的比重：

4、计算第项指标的熵值：

其中,,满足;

5、计算信息熵冗余度(差异)：

6、计算各项指标的权重:

7、计算各项指标的评价得分：

### 5.2.1.2归一化法

对原始数据的线性变换，使结果值映射到 之间。实现方法是将变量值减去最小值并除以最大值和最小值的差。

转化函数：

### 5.2.2模型建立

根据问题一求解出的最大值和方差，我们发现这两者具有不同的值范围时，当梯度更新时，会来回震荡，经过较长的时间才能达到局部最优值或全局最优值。为了避免该问题，我们采用归一化，以此确保不同的特征具有相同的值范围，这样梯度下降可以很快的收敛。

利用归一化将方差和最大值化为[0-1]的结果，建立熵权法模型：

### 5.2.3模型求解

计算单位都为摄氏温度，因此不需要标准化，只需将方差和最大值将其设置为负向指标，利用归一化法化为[0-1]，再将其转化为正向指标利用熵权法求出熵值与冗余度，最后求出权重与评价分。

方差的权重为：0.5053，最大值的权重为0.4947

十个管道的评价分如下表

表2 十个管道的评价分表

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **管道1** | **管道2** | **管道3** | **管道4** | **管道5** | **管道6** | **管道7** | **管道8** | **管道9** | **管道10** |
| 39.9597 | 59.8905 | 90.1897 | 67.2019 | 82.5529 | 75.8829 | 72.7363 | 62.7819 | 31.2426 | 16.2627 |
|  |  |  |  |  |  |  |  |  |  |

根据评价分结果：

最优管道为3，评价分为90.1897;

最差管道为10，评价分为16.2627.

## 5.3问题三

### 5.3.1建模准备

皮尔森相关系数是线性相关的系数的一种，用来反映这两个量之间的线性相关程度的。通常使用小写字母r来表示系数值。其范围在-1到1之间，绝对值越接近于1，相关性越强。

其公式表示为:

其中，cov为两个变量的协方差，分母为2个变量标准差的乘积。是的平均值，是的平均值，E为期望。

### 5.3.2模型建立

因为该问题所求结果就为模型，因此我们需要利用皮尔森相关系数求解出153个变量中相关性最大的2个变量。

### 5.3.3模型求解

利用MATLAB求出附件1对附件2的皮尔森相关性系数，用Excel并求取平均值，利用降序取均值最大的两个

表3 皮尔森相关性系数记录表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **管道1** | **管道2** | **管道3** | **管道4** | **管道5** | **管道6** |
| 0.96386 | 0.81379 | 0.85809 | 0.92394 | 0.94486 | 0.92031 |
| 0.94735 | 0.73411 | 0.76633 | 0.89648 | 0.90362 | 0.90208 |
| **管道7** | **管道8** | **管道9** | **管道10** | **均值** | **第x列** |
| 0.93657 | 0.92358 | 0.54935 | 0.18898 | 0.802333 | 119 |
| 0.94086 | 0.94288 | 0.55244 | 0.19581 | 0.778196 | 112 |

由此可见对附件1所提到的十个管道影响最大的变量是第119个变量分离器出口汽温定值与第112个变量省煤器出口温度2。

利用1stopt分别遍历出十个管道相应的函数：

管道1：



管道2:



管道3:



管道4：



管道5：



管道6：

 管道7：



管道8：



管道9：



管道10：

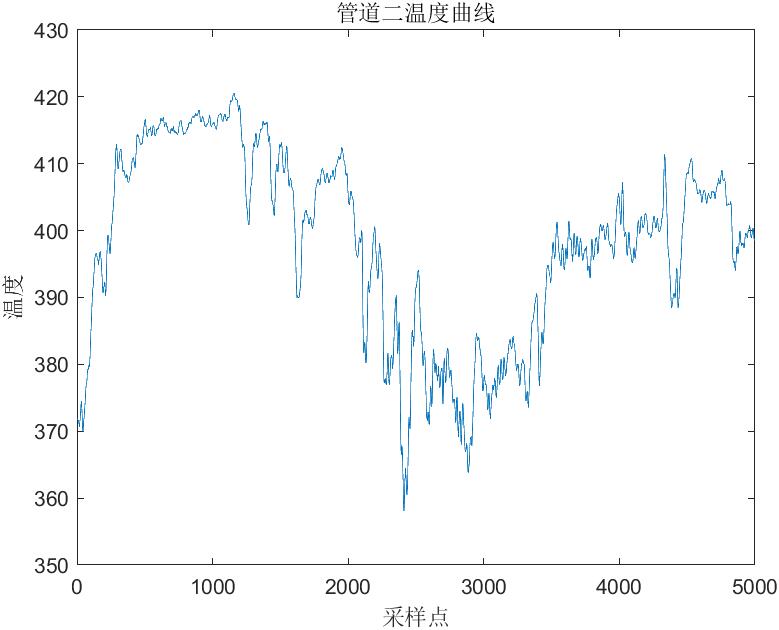
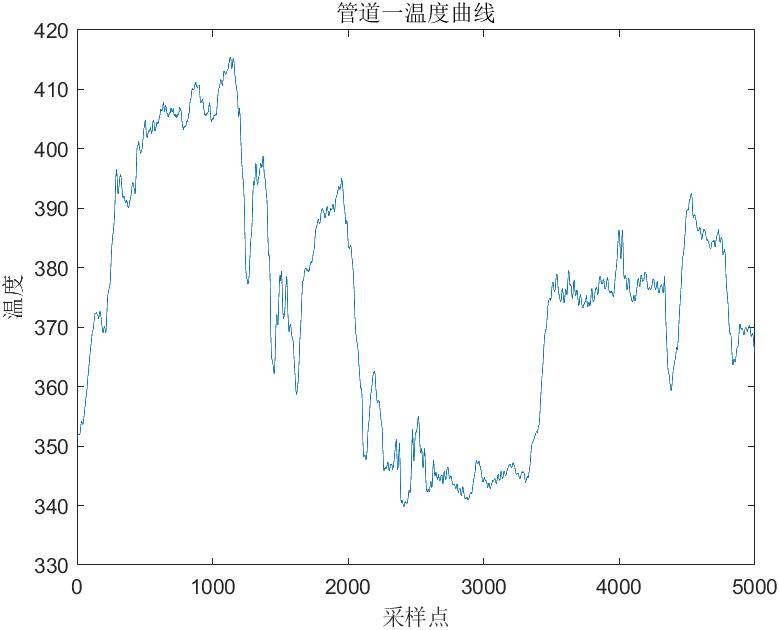


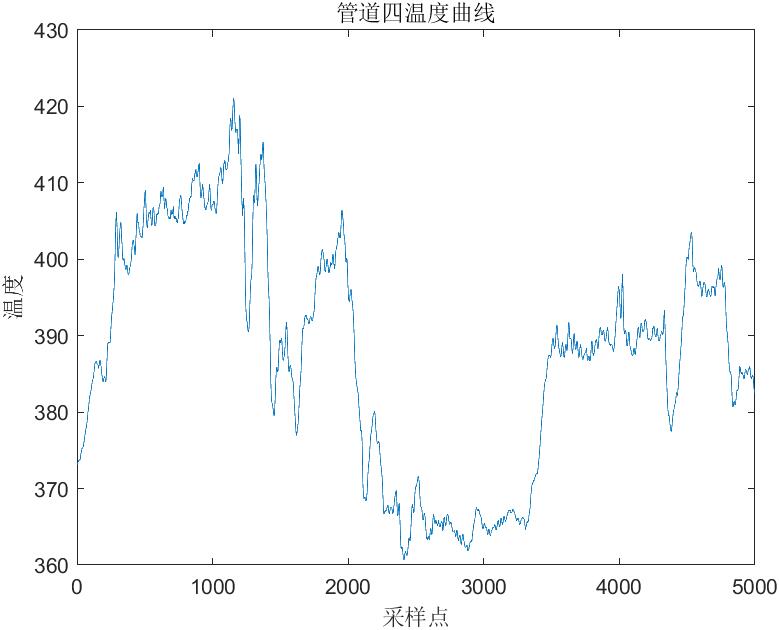
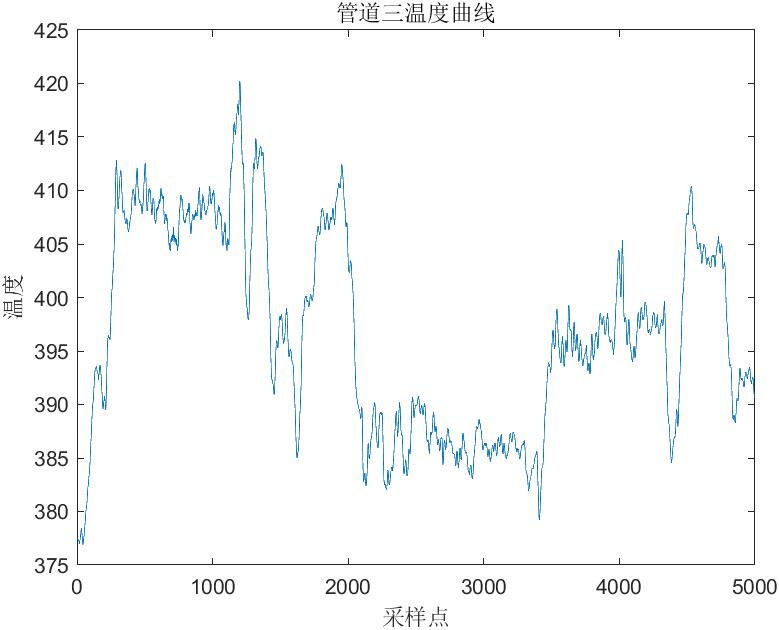
利用1stopt所求出各个参数的值：

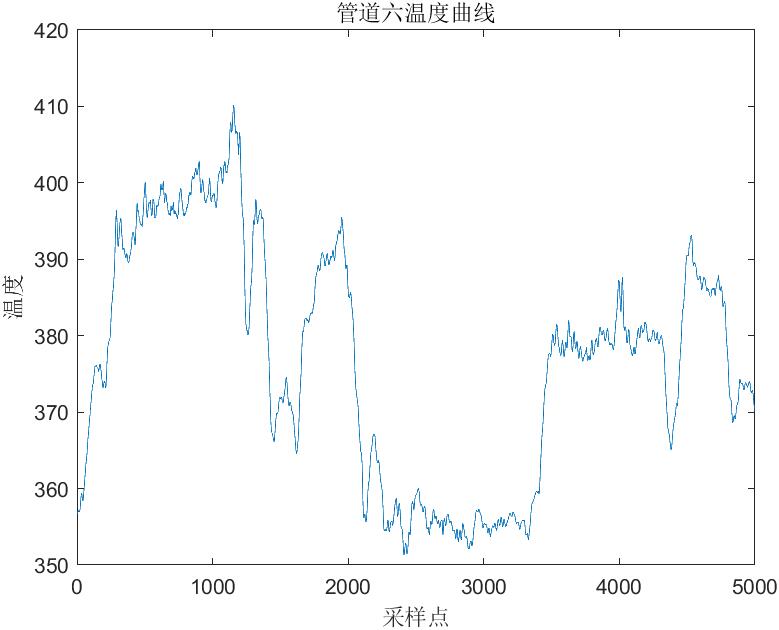
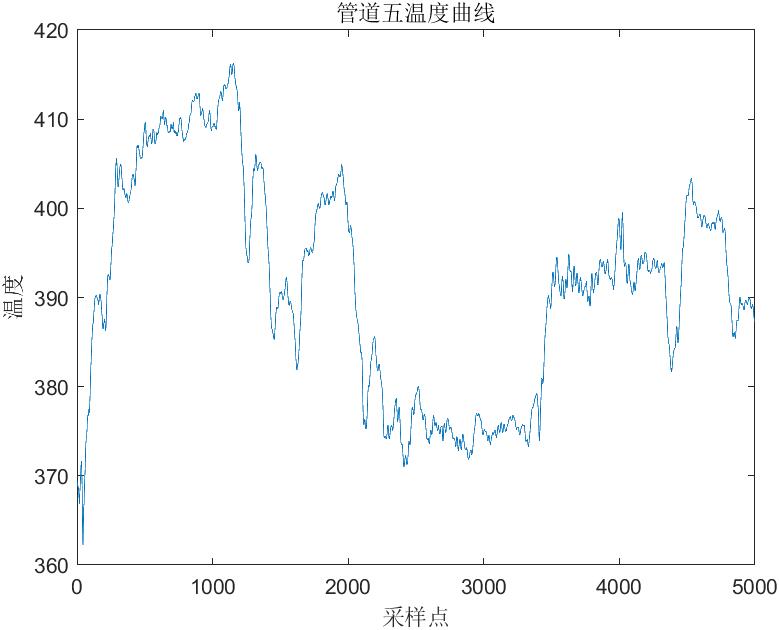
表4各管道拟合函数参数值

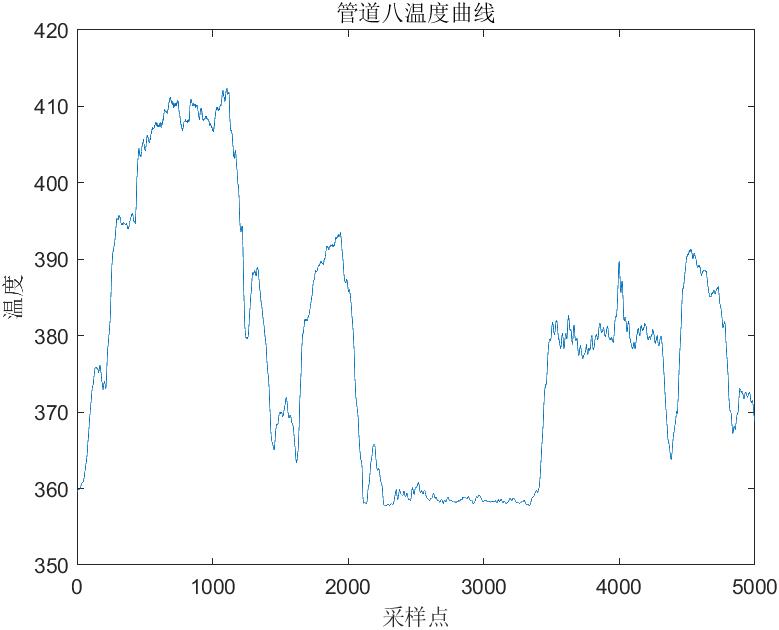
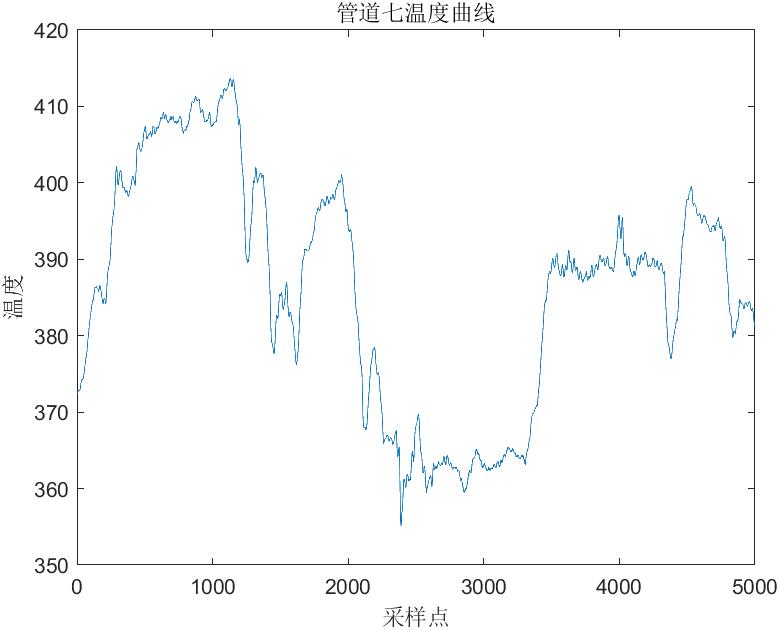


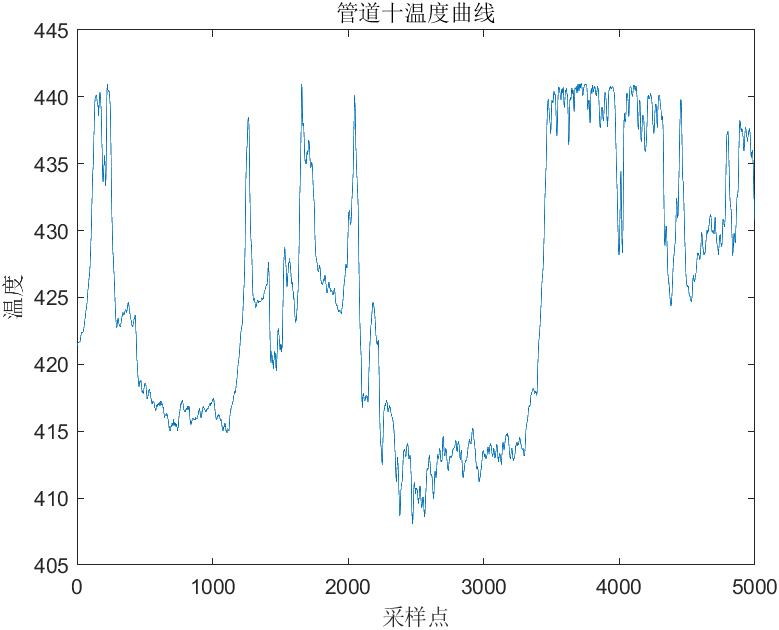
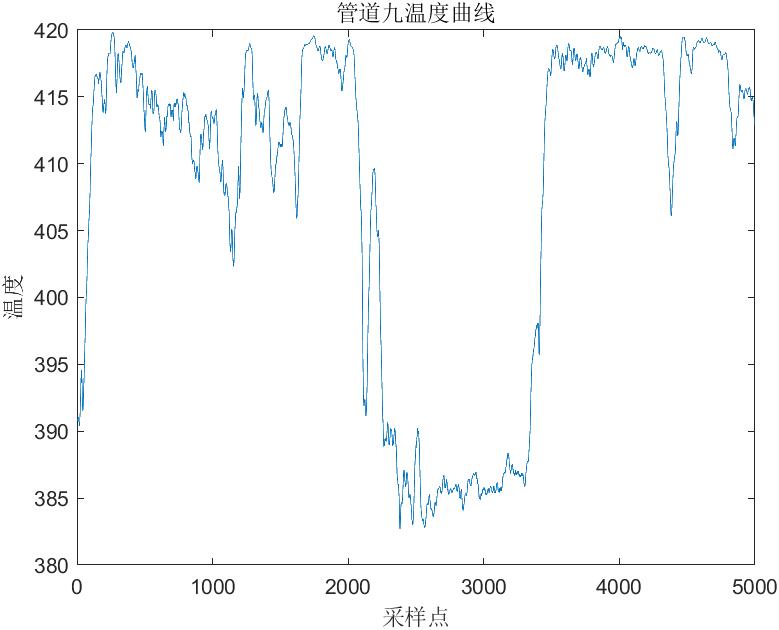
根据所求值代入函数中，利用MATLAB画出相应的图像:











## 5.4问题四

利用MATLAB求出管道十对附件2中111个操作变量的皮尔森相关性系数，利用降序取均值较高的操作变量。

表5 皮尔森相关性系数

|  |  |  |
| --- | --- | --- |
| **相关度** | **第x列** | **操作变量** |
| 0.700266 | 100 | D磨入口冷一次风电调挡板位置 |
| 0.534862 | 55 | 3角DD层二次风调节挡板位置反馈 |
| 0.534626 | 103 | E磨入口冷一次风电调挡板位置 |
| 0.509699 | 47 | 3角DE层二次风调节挡板位置反馈 |

由此可见，引发超温现象的主要操作变量是D磨入口冷一次风电调挡板位置、3角DD层二次风调节挡板位置反馈、E磨入口冷一次风电调挡板位置、3角DE层二次风调节挡板位置反馈。

## 5.5问题五

### 5.1建模准备

逐步回归原理：进行最小二乘回归拟合，然后基于系数和样本均数的标准误差公式计算得到t检验值，然后由特定公式计算出p值，如果各指标的值均<0.05就接纳，如果有几个指标的值大于0.05，那么会排除掉值最大的一列指标，然后重新计算和值，以此类推，直到所有指标的值均<0.05。逐步回归后，得到了真正显著的自变量集。

逐步回归通过stepwise函数实现，可以得出一个的经验公式，写出关系式带入预测的指标数据可以计算出相应的因变量结果。

### 5.2模型建立

根据假设和分析初始定义以下多次幂自变量集。

### 5.3模型求解

利用逐步回归对3172后的数据进行分析，对111个操作变量进行筛选。



筛选的结果为：1 2 3 4 6 12 13 31 34 37 38 39

相关度为0.6990

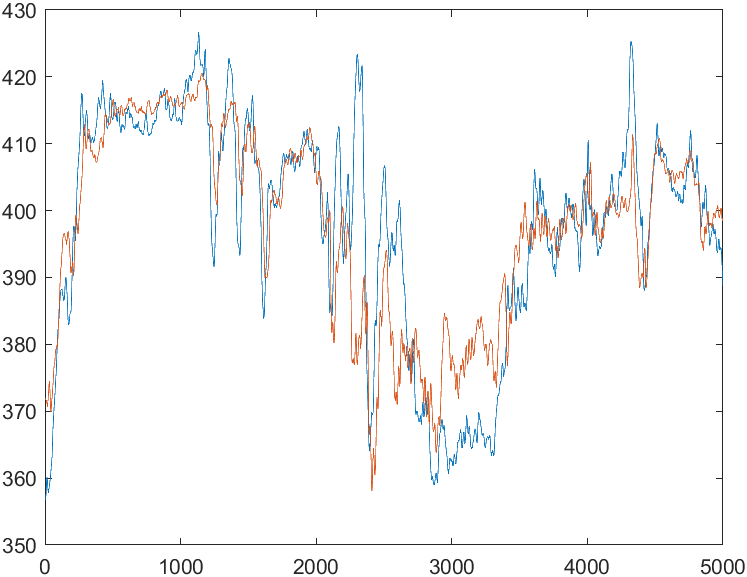
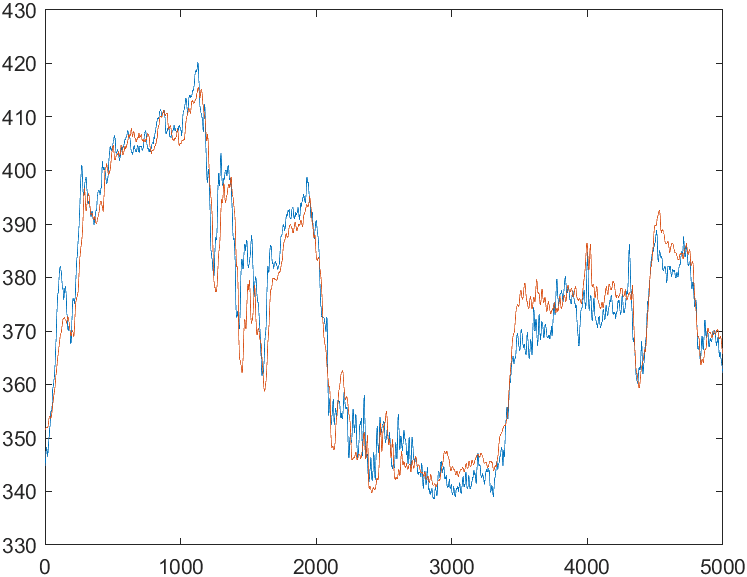
再次利用逐步回归，拟合出经验公式，经验公式图像如下图所示：

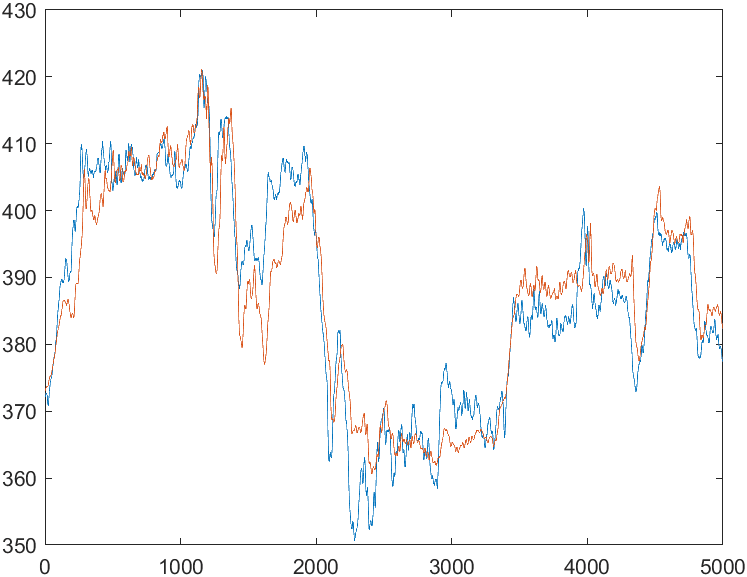
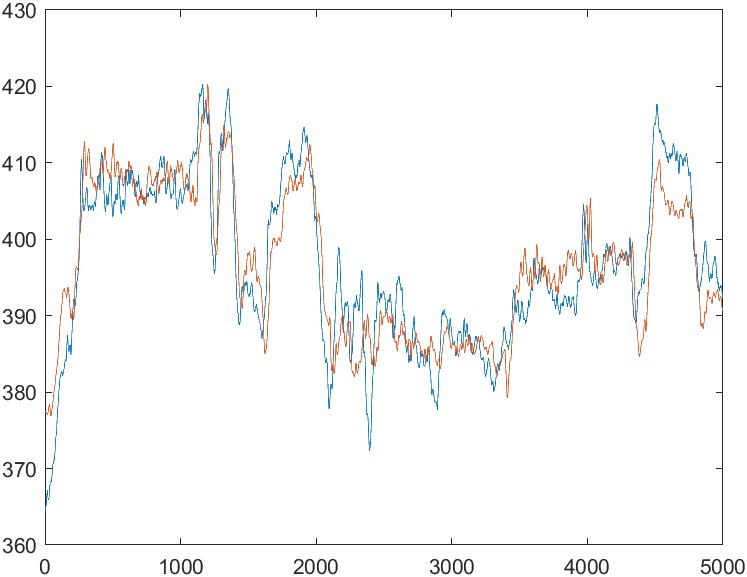


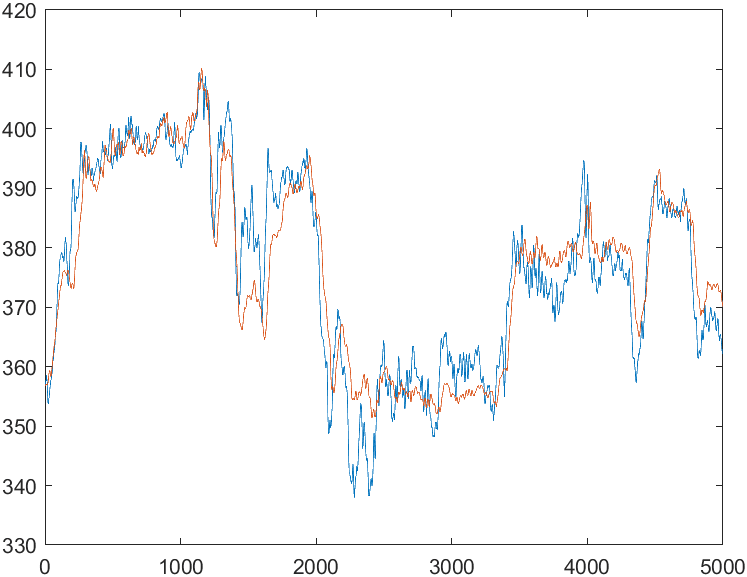
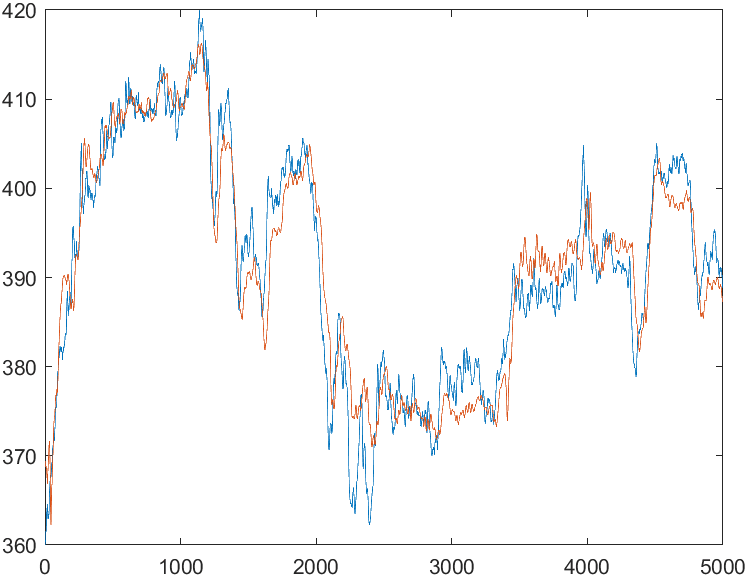
# 六、模型检验与优化

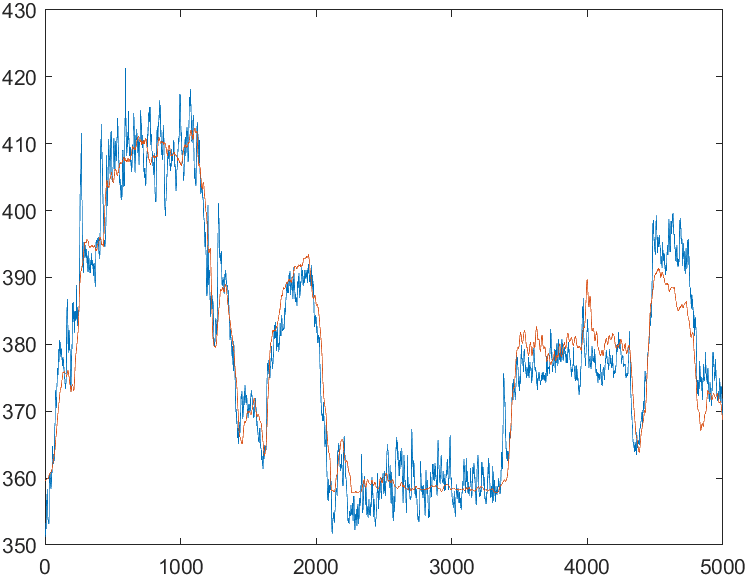
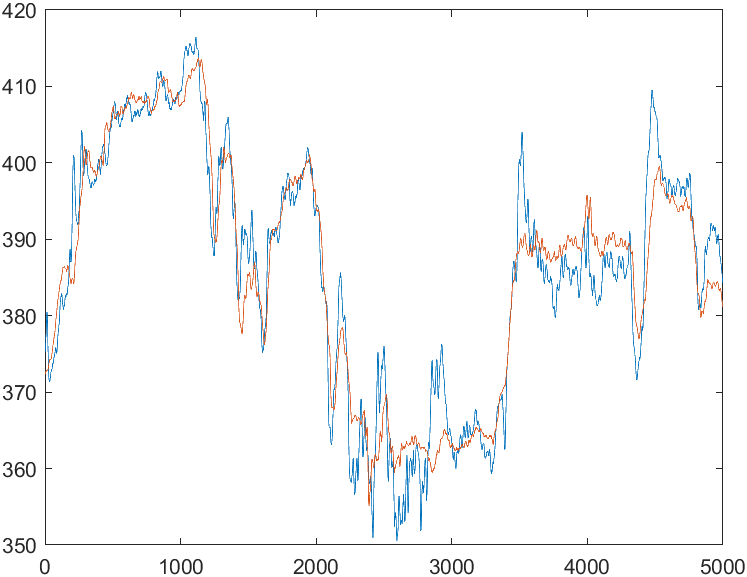
## 6.1问题三检验

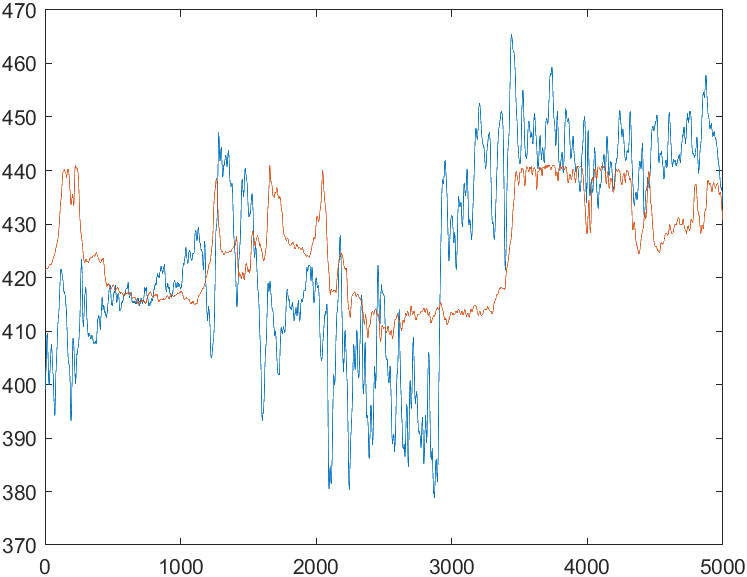
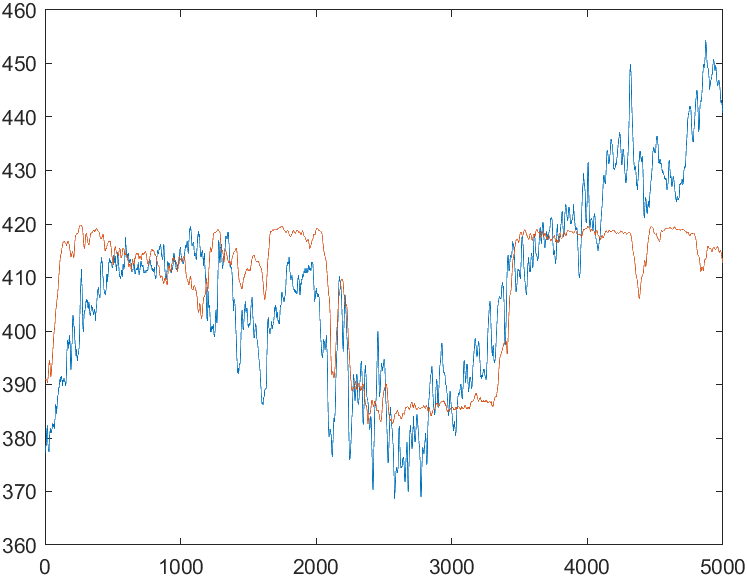
我们再将原有的真实温度放入图像中进行比较：











根据1stopt所遍历出的公式结果，找出拟合函数的均方差，残差平方和，相关系数，确定系数。如下表链接所示

表 拟合函数评价表链接



根据上表绘制评价参数折线图

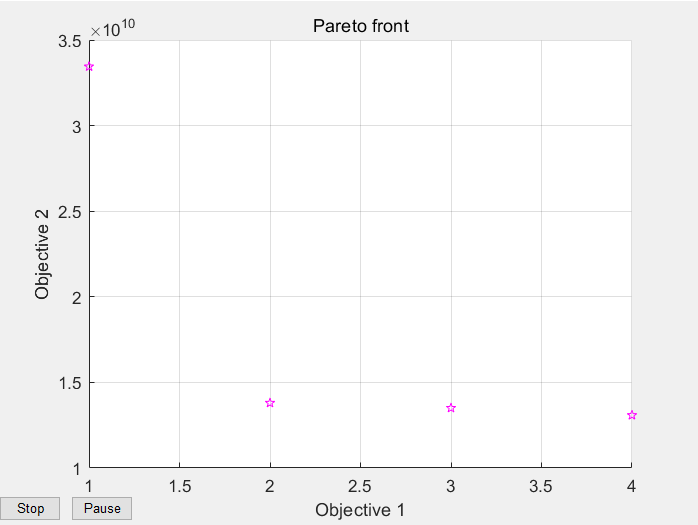
根据和原图像的比较以及4个评价量，可知管道1、8的拟合效果较好，管道6的拟合效果略微低。

## 6.2问题五模型优化

利用gamultiobj函数对模型进行优化求出Pareto最优解。

建立以操作数、与最优曲线差距、方差、最大值的目标函数。

绘制Pareto最优解图像如下图所示：



方法的不足：

由于受到计算机影响Pareto最优解图像200个点最终只能显示4个点。

# 附录：

问题一代码：

Y=xlsread('C:\Users\86152\Desktop\附件1.xlsx');

X=Y(:,[2:11]);

x1=mean(X,1) %求十个管道温度平均值

[max\_X,index]=max(X) %求十个管道温度最大值

[min\_X,index]=min(X) %求十个管道温度最小值

DX=var(X) %求十个管道温度方差

问题二代码:

function y=guiyi(x,type,ymin,ymax)

%实现正向或负向指标归一化，返回归一化后的数据矩阵

%x为原始数据矩阵, 一行代表一个样本, 每列对应一个指标

%type设定正向指标1,负向指标2

%ymin,ymax为归一化的区间端点

[n,m]=size(x);

y=zeros(n,m);

xmin=min(x);

xmax=max(x);

switch type

case 1

for j=1:m

y(:,j)=(ymax-ymin)\*(x(:,j)-xmin(j))/(xmax(j)-xmin(j))+ymin;

end

case 2

for j=1:m

y(:,j)=(ymax-ymin)\*(xmax(j)-x(:,j))/(xmax(j)-xmin(j))+ymin;

end

end

clc

clear all

X=xlsread('附件1.xlsx');

data=X(:,2:end);

[max\_data,index1]=max(data);

[min\_data,index]=min(data);

fangcha=var(data);

[max\_data,index1]=max(data);

D=[fangcha;max\_data]';

temp=[2 2];

[pj w]=shang(D,temp);

pj%pj 评价分

w%w 权重

function [s,w]=shang(x,ind)

%实现用熵值法求各指标(列）的权重及各数据行的得分

%x为原始数据矩阵, 一行代表一个样本, 每列对应一个指标

%ind指示向量，指示各列正向指标还是负向指标，1表示正向指标，2表示负向指标

%s返回各行（样本）得分，w返回各列权重

[n,m]=size(x); % n个样本, m个指标

%%数据的归一化处理

for i=1:m

if ind(i)==1 %正向指标归一化

X(:,i)=guiyi(x(:,i),1,0.002,0.996); %若归一化到[0,1], 0会出问题

else %负向指标归一化

X(:,i)=guiyi(x(:,i),2,0.002,0.996);

end

end

%%计算第j个指标下，第i个样本占该指标的比重p(i,j)

for i=1:n

for j=1:m

p(i,j)=X(i,j)/sum(X(:,j));

end

end

%%计算第j个指标的熵值e(j)

k=1/log(n);

for j=1:m

e(j)=-k\*sum(p(:,j).\*log(p(:,j)));

end

d=ones(1,m)-e; %计算信息熵冗余度

w=d./sum(d); %求权值w

s=100\*w\*X'; %求综合得分

问题三代码:

皮尔森相关性系数：

clc

clear all

X1=xlsread('附件1.xlsx');

data1=X1(:,2:11);

X2=xlsread('附件2.xlsx');

data2=X2(:,2:end);

data=[data1,data2];

%data=zscore(data);%对变量进行标准化

temp=corrcoef(data);% 生成皮尔森相关性系数计算公式

pershen=temp(1:10,11:end);%求附件一十个管道对附件二111个操作变量和53个状态变量的皮尔森相关性系数

pershen=abs(pershen);%将皮尔森相关性系数求取绝对值

csvwrite('wenti3.csv',pershen);

%经过Excel筛选皮尔森相关性系数最大列数分别为119 112 114 5 115 121 91 92 1 73

result=[data2(:,119),data2(:,112)];

csvwrite('data10.csv',result);

拟合函数图像代码：

clc;

clear

subplot(5,2,1)

x1=(1:1:5000);

M=csvread('C:\Users\86152\Desktop\data\data1.csv');

x=M(:,1);

y=M(:,2);

p1=-366.623574952674;

p2=-6309.81591583794;

p3=-57878.3062818374;

p4=15903.1406517026;

p5=-53.4927163299915;

p6=0.0889609164030065;

p7=542.236136002024;

p8=-218.986158913513;

p9=17.338434284516;

z=(p1+p2.\*log(x)+p3.\*(log(x)).^2+p4\*y+p5\*y.^2+p6\*y.^3)./(1+p7\*log(x)+p8\*(log(x)).^2+p9.\*y);

plot(x1,z);

xlabel('采样点');

ylabel('温度');

title('管道一温度曲线')

subplot(5,2,2)

x1=(1:1:5000);

M=xlsread('C:\Users\86152\Desktop\data\data2.xlsx');

x=M(:,1);

y=M(:,2);

p1=-760228.816854935;

p2=9.40803736929389;

p3=-0.0110468670816608;

p4=2.36049470120192E-7;

p5=10924.026332869;

p6=-58.9224433132279;

p7=0.141102010823114;

p8=-0.000126582641617285;

z=p1+p2.\*x+p3.\*x.^2+p4.\*x.^3+p5.\*y+p6.\*y.^2+p7.\*y.^3+p8.\*y.^4;

plot(x1,z);

xlabel('采样点');

ylabel('温度');

title('管道二温度曲线')

subplot(5,2,3)

x1=(1:1:5000);

M=csvread('C:\Users\86152\Desktop\data\data3.csv');

x=M(:,1);

y=M(:,2);

p1=-278364.765528611;

p2=-8053.13293300719;

p3=702.700979973376;

p4=4525.26042721373;

p5=-25.4017726090204;

p6=0.063270352619814;

p7=-5.9005816668393E-5;

z=p1+p2.\*log(x)+p3.\*(log(x)).^2+p4.\*y+p5.\*y.^2+p6.\*y.^3+p7.\*y.^4;

plot(x1,z);

xlabel('采样点');

ylabel('温度');

title('管道三温度曲线')

subplot(5,2,4)

x1=(1:1:5000);

M=csvread('C:\Users\86152\Desktop\data\data4.csv');

x=M(:,1);

y=M(:,2);

p1=-72.9520473758939;

p2=-556.535840804353;

p3=-6289.99299391888;

p4=38088.2044118192;

p5=-187.688895088328;

p6=-0.314791034083831;

p7=-36.6632213229309;

p8=-7693.38639121209;

p9=65.3375954195133;

p10=135.981090347525;

p11=4418.78480395924;

z=(p1+p3.\*x+p5.\*log(y)+p7\*x.^2+p9.\*(log(y)).^2+p11.\*x.\*log(y))./(1+p2.\*x+p4.\*log(y)+p6.\*x.^2+p8.\*(log(y)).^2+p10.\*x.\*log(y));

plot(x1,z);

xlabel('采样点');

ylabel('温度');

title('管道四温度曲线')

subplot(5,2,5)

x1=(1:1:5000);

M=csvread('C:\Users\86152\Desktop\data\data5.csv');

x=M(:,1);

y=M(:,2);

p1=6636.72503715224;

p2=826099.455060947;

p3=-498.995296492192;

p4=1.80238545972406;

p5=-1143812.65641828;

p6=3030.85414746531;

p7=-4932.54637234801;

p8=3.5943846329167;

z=(p1+p2.\*x+p3.\*x.^2+p4.\*x.^3+p5.\*y)./(1+p6.\*x+p7.\*y+p8.\*y.^2);

plot(x1,z);

xlabel('采样点');

ylabel('温度');

title('管道五温度曲线')

subplot(5,2,6)

x1=(1:1:5000);

M=csvread('C:\Users\86152\Desktop\data\data6.csv');

x=M(:,1);

y=M(:,2);

p1=-5883.284880402;

p2=60.424921963928;

p3=-0.164225401010505;

p4=0.000149967919524129;

p5=61332.3184515964;

p6=79523.3753333665;

p7=-54368.4125968335;

p8=10405.5019030654;

p9=-639.850116036068;

z=p1+p2.\*x+p3.\*x.^2+p4.\*x.^3+p5.\*log(y)+p6.\*(log(y)).^2+p7.\*(log(y)).^3+p8.\*(log(y)).^4+p9.\*(log(y)).^5;

plot(x1,z);

xlabel('采样点');

ylabel('温度');

title('管道六温度曲线')

subplot(5,2,7)

x1=(1:1:5000);

M=csvread('C:\Users\86152\Desktop\data\data7.csv');

x=M(:,1);

y=M(:,2);

p1=201478.39441699;

p2=372414.059548135;

p3=-147089.218042618;

p4=12485.7289898091;

p5=-224.561025015429;

p6=3.32272031088369;

p7=-173.088899267814;

p8=4.01866601526878;

z = (p1+p2.\*log(x)+p3.\*(log(x)).^2+p4.\*(log(x)).^3+p5.\*y+p6.\*y.^2)./(1+p7.\*log(x)+p8.\*y);

plot(x1,z);

xlabel('采样点');

ylabel('温度');

title('管道七温度曲线')

subplot(5,2,8)

x1=(1:1:5000);

M=csvread('C:\Users\86152\Desktop\data\data8.csv');

x=M(:,1);

y=M(:,2);

p1=682051.099193563;

p2=-166.30866016075;

p3=0.443241999640146;

p4=-0.000392578722960921;

p5=-340091.23447923;

p6=58193.7202898162;

p7=-3310.5722076251;

z=p1+p2.\*x+p3.\*x.^2+p4.\*x.^3+p5.\*log(y)+p6.\*(log(y)).^2+p7.\*(log(y)).^3;

plot(x1,z);

xlabel('采样点');

ylabel('温度');

title('管道八温度曲线')

subplot(5,2,9)

x1=(1:1:5000);

M=csvread('C:\Users\86152\Desktop\data\data9.csv');

x=M(:,1);

y=M(:,2);

p1=421.576820618631;

p2=386.777802412627;

p3=119.426378501925;

p4=481.83371005309;

p5=254.888468073613;

p6=5.98063835171261;

p7=-525.072127917344;

z=p1./(1+((x-p2)./p3).^2)+p4./(1+((y-p5)./p6).^2)+p7./((1+((x-p2)./p3).^2).\*(1+((y-p5)./p6).^2));

plot(x1,z);

xlabel('采样点');

ylabel('温度');

title('管道九温度曲线')

subplot(5,2,10)

x1=(1:1:5000);

M=csvread('C:\Users\86152\Desktop\data\data10.csv');

x=M(:,1);

y=M(:,2);

p1=-1217.63395897089;

p2=-27430.5826029483;

p3=380.735591383903;

p4=-8.70371508967982;

p5=1639.83545496426;

p6=276.790380488085;

p7=423.356786189341;

p8=27449.3757790585;

z = p1+p2./(1+((x-p3)./p4).^2)+p5./(1+((y-p6)./p7).^2)+p8./((1+((x-p3)./p4).^2).\*(1+((y-p6)./p7).^2));

plot(x1,z);

xlabel('采样点');

ylabel('温度');

title('管道十温度曲线')

问题三Excel筛选表格链接：



问题四代码：

%第四题

clc

clear all

X1=xlsread('附件1.xlsx');

data1=X1(:,11);

X2=xlsread('附件2.xlsx');

data2=X2(:,44:end);

data=[data1,data2];

%data=zscore(data);%对变量标准化

temp=corrcoef(data);%生成皮尔森相关性系数计算公式

pershen=temp(1,2:end);%求附件一管道十对附件二111个操作变量的皮尔森相关性系数

pershen=abs(pershen); %将皮尔森相关性系数求取绝对值

csvwrite('wenti4.csv',pershen);

% 相关度最大的几列为

问题四Excel筛选表格链接：



问题五代码：

两次逐步回归：

clc

clear all

X1=xlsread('附件1.xlsx');%因变量

data1=X1(3172:end,11);

X2=xlsread('附件2.xlsx');%自变量

data2=X2(3172:end,2:50);

%1 2 3 4 6 12 13 31 34 37 38 39

%temp=[data2,data2.^2,data2.^3,log10(data2)];

%stepwise(temp,data1);%0.6990

%temp1=data2;

temp1=[data2(:,1),data2(:,2),data2(:,3),data2(:,4),data2(:,6),data2(:,12),data2(:,13),data2(:,31),data2(:,34),data2(:,37),data2(:,38),data2(:,39)];

temp2=[temp1,temp1.^2,temp1.^3,temp1.^4,temp1.^5,log10(temp1),1./temp1];

%stepwise(temp2,data1);%0.6990

%stepwise(temp2,data1);

%12列拟合相关度54%

%stats beta

%temp3=[x1,x2,x3,x4,x6,x12,x13,x31,x34,x37,x38,x39];