



2019 年 兰 州 大 学 数 学 建 模 竞 赛

题目： 回归模型解决敏感材料温度控制问题

Hollow Man

成员姓名	蒋嵩林 肖锦恒 白舒睿
学 院	信息科学与工程学院
班 级	计算机科学与技术基地班(2018 级)
日 期	二〇一九年六月

摘 要:

对于工件烘烤上色的问题，工件模板本身固有的性质，工件的厚度，与烘烤温度，烘烤时间都会对工件的颜色产生影响。本题采用了一种特定的工件模板，要求探讨工件的厚度，烘烤温度，烘烤时间对工件的颜色所产生的影响。通过对数据的统计分析，我们采用了统计分析方法建立回归模型。

关键词:

回归模型，敏感材料温度控制

一、 问题重述

1.1 问题背景

工厂制作工件，完成后需要漆上涂料，然后进行烘烤，得到一定标准的颜色。然而，对与一个给定的工件模板，由于材料的差异，要烘烤出一个指定的颜色，烘烤温度和时间都是不同的，比较难以控制。因此，需要建立颜色，工件模板与烘烤温度，烘烤时间之间的数学模型。

1.2 问题概述

本给出了在 950-1300 摄氏度下间隔 30 度的 10-30um 下间隔 5um 厚度的一种敏感材料的色差值数据，并且分别提供 3 分钟，10 分钟和 30 分钟的烘烤模式的数据。问题一要求解出不规则厚度下得到此种颜色的温度值。问题二则要求在问题一的基础上，探讨烘烤时间和烘烤温度对模板厚度，烘烤后颜色值的关系。

二、 模型假设

本题探讨多元变量之间的关系，根据常识，这些变量都和色差有较大关系。于是很自然地，我们使用了完全多元二次多项式回归模型。

三、 问题分析

3.1 问题 1 的分析

本题要求解出不同烘烤时间的不规则厚度下得到此种颜色的温度值，在这种情况下，因为我们要预测出色差，所以我们分别对同一烘烤时间下地数据进行建模，使用 $y = \beta_0 + \beta_1 x_1^2 + \beta_2 x_1 + \beta_3 x_2^2 + \beta_4 x_2 + \beta_5 x_1 x_2 + \epsilon$ 来进行拟合数据，其中， y 为被解释变量（因变量）， x_1 ， x_2 为解释变量（回归变量，自变量）， β_0 ， β_1 ， β_2 ， β_3 ， β_4 ， β_5 为回归系数， ϵ 为随机误差（均值为零的正态分布随机变量）。我们将色差作为被解释变量（因变量） y 。剩下的两个变量，烘烤温度和工件的厚度，则应为解释变量（回归变量，自变量） x_1 ， x_2 。因为在该式中， x_1 ， x_2 是等效的，所以不做区分，将 x_1 作为工件厚度， x_2 作为烘烤温度。

3.2 问题 2 的分析

在第一问的基础上，这里另外加入了烘烤时间这个变量，因此，仿照第一问，我们采用 $y = \beta_0 + \beta_1 x_1^2 + \beta_2 x_1 + \beta_3 x_2^2 + \beta_4 x_2 + \beta_5 x_3^2 + \beta_6 x_3 + \beta_7 x_1 x_2 + \beta_8 x_1 x_3 + \beta_9 x_2 x_3 + \epsilon$ 公式来进行模拟求解，将色差作为被解释变量（因变量） y 。剩下的三个变量，工件的厚度，烘烤温度和烘烤时间，分别为解释变量（回归变量，自变量） x_1 ， x_2 ， x_3 。

四、 模型建立与求解

4.1 问题 1

我们采用 Matlab 统计工具箱，由数据 y, x_1, x_2 估计 β ，通过 Matlab 中的函数 `[b,bint,r,rint,stats]=regress(y,x,alpha)` 进行拟合与计算。

该函数中，对于输入， y 为 n 维数据向量，为色差数据； $x=[1 \ x_1 \ x_2 \ x_2^2]$ ，为 $n \times 4$ 的数据矩阵，第 1 列为全 1 向量，解释变量(回归变量，自变量) x_1, x_2 将按照方案对应进行计算。 α 为置信水平。这里我们采用默认置信度 0.05。

对于输出， b 为 β 的估计值； $bint$ 为 b 的置信区间； r 为残差向量 $y-xb$ ； $rint$ 为 r 的置信区间； $Stats$ 为检验统计量 R^2, F, p, s^2 ，分别为相关系数，F 统计，与统计量 F 对应的概率 P，估计误差方差。

我们使用 Excel 整理提供的数据导入到 data 变量中，并使用以下代码计算与绘图：

```
x1 = data(:,1);  
  
x2 = data(:,2);  
  
y = data(:,3);  
  
X = [ones(size(x1)) x1.*x1 x1 x2.*x2 x2 x1.*x2];  
  
[b,bint] = regress(y,X)  
  
scatter3(x1,x2,y,'filled')  
  
hold on  
  
x1fit = min(x1):0.5:max(x1);  
  
x2fit = min(x2):0.5:max(x2);  
  
[X1FIT,X2FIT] = meshgrid(x1fit,x2fit);  
  
YFIT = b(1)+ b(2)*X1FIT.*X1FIT+b(3)*X1FIT + b(4)*X2FIT.*X2FIT +  
b(5)*X2FIT + b(6)*X1FIT.*X2FIT;  
  
mesh(X1FIT,X2FIT,YFIT)  
  
xlabel('x1')
```

```
ylabel('x2')
```

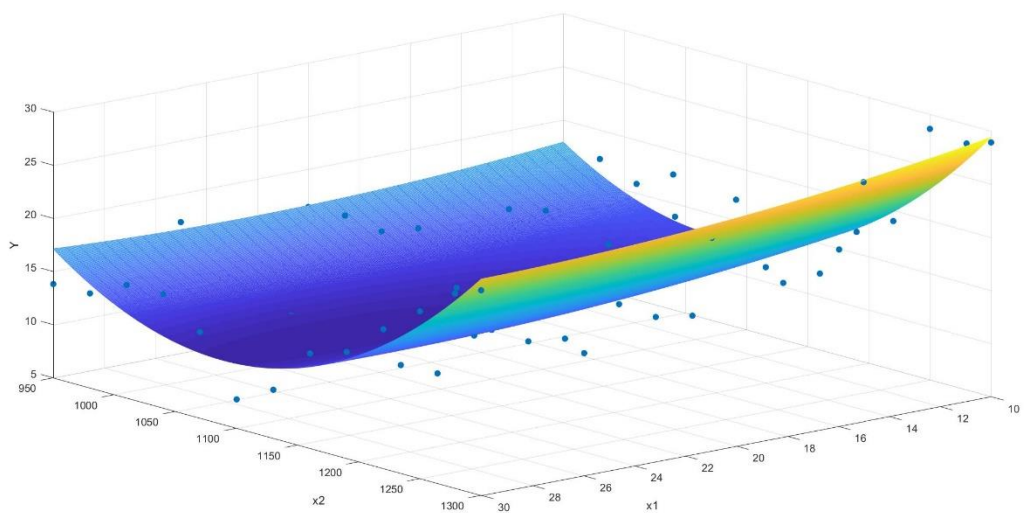
```
zlabel('Y')
```

```
view(140,30)
```

导入数据，拟合，得出以下拟合结果：

图+表 1： 3min 的结果

参数	参数估计值	置信区间
β_0	380.1905	[296.5150, 463.8660]
β_1	0.0094	[-0.0070, 0.0258]
β_2	0.0313	[-1.1591, 1.2217]
β_3	0.0003	[0.0003, 0.0004]
β_4	-0.6816	[-0.8275, 0.5358]
β_5	-0.0005	[-0.0013, 0.0004]
$R_2=0.7730$ $F=40.1802$ $p<0.00001$ $s_2=7.6618$		

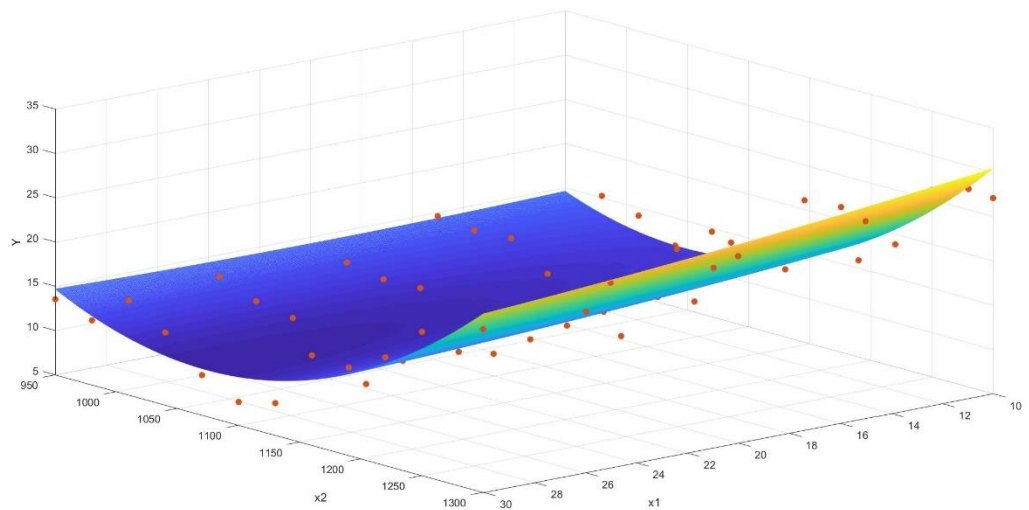


由此得到色差在烘烤时间为 3min 时关于烘烤温度和工件的厚度的回归公式：

$$y=380.1905+0.0094*x_1^2+0.0313*x_1+0.0003*x_2^2-0.6816*x_2-0.0005*x_1x_2$$

图+表 2： 10min 的结果

参数	参数估计值	置信区间
β_0	324.6056	[256.5902, 392.6209]
β_1	0.0040	[-0.0093, 0.0174]
β_2	0.5728	[-0.3948, 1.5404]
β_3	0.0003	[0.0002, 0.0003]
β_4	-0.6013	[-0.7199, -0.4828]
β_5	-0.0008	[-0.0015, -0.0001]
$R_2=0.8662$ $F=76.3922$ $p<0.00001$ $s_2=5.0623$		

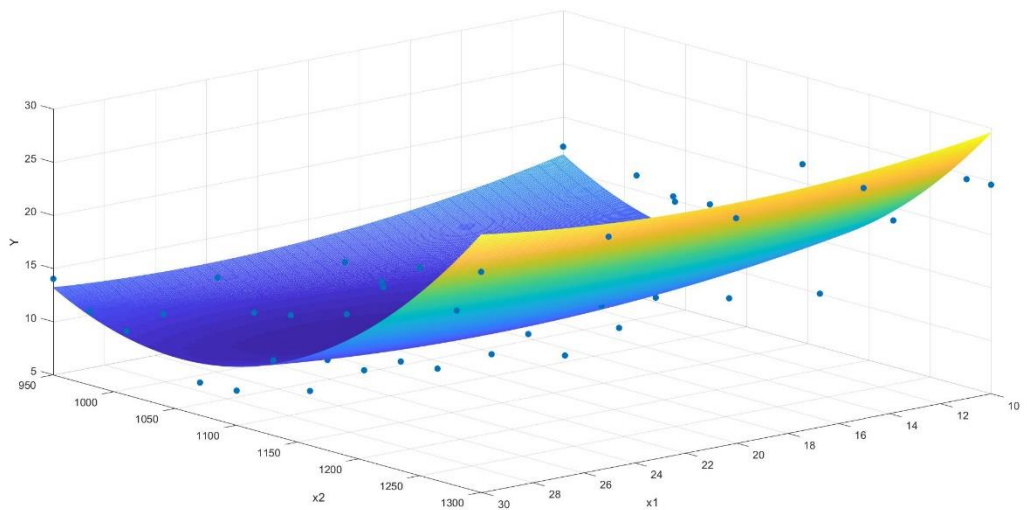


由此得到色差在烘烤时间为 10min 时关于烘烤温度和工件的厚度的回归公式：

$$y=324.6056+0.0040*x_1^2+0.5728*x_1+0.0003*x_2^2-0.6013*x_2-0.0008*x_1x_2$$

图+表 3： 30min 的结果

参数	参数估计值	置信区间
β_0	385.5961	[308.7328, 462.4594]
β_1	0.0183	[0.0032, 0.0334]
β_2	-1.2870	[-2.3806, 0.1935]
β_3	0.0003	[0.0003, 0.0004]
β_4	-0.6840	[-0.8180, 0.5500]
β_5	0.0004	[-0.0004, 0.0012]
$R_2=0.8591$ $F=71.9339$ $p<0.00001$ $s_2=6.4650$		



由此得到色差在烘烤时间为 30min 时关于烘烤温度和工件的厚度的回归公式：

$$y=385.5961+0.0183*x_1^2-1.2870*x_1+0.0003*x_2^2-0.6840*x_2+0.0004*x_1x_2$$

由上面三个模型，用 Excel 进行单变量求解，得到结果：

3min				
厚度	L*	a*	b*	温度
12um	50	1.5	10	823.0924081
18um	50	1.5	10	820.1098203
23um	50	1.5	10	820.2751345
27um	50	1.5	10	822.0768133
10min				
厚度	L*	a*	b*	温度
12um	50	1.5	10	1048.581129
18um	50	1.5	10	1028.050208
23um	50	1.5	10	1020.689614
27um	50	1.5	10	1032.197755
30min				
厚度	L*	a*	b*	温度
12um	50	1.5	10	813.9844244
18um	50	1.5	10	801.1148467
23um	50	1.5	10	795.5921859
27um	50	1.5	10	794.4959499

4.2 问题 2

同问题一，用 Excel 整理提供的数据导入到 data 变量中，并使用以下代码计算回归模型：

```

x1 = data(:,1);

x2 = data(:,2);

x3=data(:,3);

y = data(:,4);

X = [ones(size(x1)) x1.*x1 x1 x2.*x2 x2 x3.*x3 x3 x1.*x2 x1.*x3 x2.*x3];

[b,bint,r,rint,stats] = regress(y,X)

scatter3(x1,x2,x3,y,'filled')

```


hold on

x1fit = min(x1):0.5:max(x1);

x2fit = min(x2):0.5:max(x2);

x3fit = min(x3):0.5:max(x3);

导入数据，拟合，得出以下拟合结果：

图+表 1： 3min 的结果

参数	参数估计值	置信区间
β_0	371.5322	[327.5192, 415.5452]
β_1	0.0106	[0.0020, 0.0192]
β_2	-0.2496	[-0.8752, 0.3760]
β_3	0.0003	[0.0003, 0.0003]
β_4	-0.6619	[-0.7383, -0.5855]
β_5	0.0015	[-0.0041, 0.0071]
β_6	-0.5981	[-0.9840, 0.2123]
β_7	-0.0003	[-0.0007, 0.0002]
β_8	0.0015	[-0.0029, 0.0060]
β_9	0.0004	[0.0002, 0.0007]
$R_2=0.8279$ $F=98.8881$ $p<0.00001$ $s_2=6.4624$		

由此得到色差关于烘烤时间，烘烤温度和工件的厚度的回归公式：

$$y=371.5322+0.0106*x_1^2-0.2496*x_1+0.0003*x_2^2-0.6619*x_2+0.0015*x_3^2-0.5981*x_3-0.0003*x_1x_2+0.0015*x_1x_3+0.0004*x_2x_3$$

由表中的模型评价数据可知， $s_2=6.4624$ 是比较小的，因而该模型所做的预测较为准确。

由表格中回归系数置信区间， β_2 ， β_5 ， β_7 ， β_8 置信区间含零点，所以得出结论： x_1 ， x_3^2 ， x_1x_2 ， x_1x_3 这几项对回归方程的影响较小。其中，工件的厚度为 x_1 ，烘烤温度为 x_2 ，烘烤时间为 x_3 。

五、 模型的评价

通过完全多元二次多项式回归模型,我们使用了题目中所给的数据,较好地拟合出了回归方程,计算简单,结果准确,符合实际,具有可行性,用图式的方式表达出来,更生动形象,各项检验统计量都表现良好。

关于模型的改进,我们可以增加模型的次数,使用多元三次多项式回归模型甚至更高次来进行拟合。虽然这样会增加模型的拟合计算量与复杂度,但同时模型获得了更高的表现力,将会在某种程度上更精确一些。