

反应曲面回归分析

胡良平^{1,2*}

(1. 军事科学院研究生院, 北京 100850;

2. 世界中医药学会联合会临床科研统计学专业委员会, 北京 100029

* 通信作者: 胡良平, E-mail: lphu812@sina.com)

【摘要】 本文目的是介绍反应曲面回归分析的概念、作用以及用软件实现计算的方法。先介绍有关的基本概念, 再介绍基本原理, 最后通过一个实例并基于 SAS 软件演示如何实施此分析方法。结果表明: ①此法最适合用于需要确定多定量因素的最佳生产条件的试验研究场合; ②应用此法的前提条件是, 参与试验研究的定量因素已经过专业和统计学方法严格筛选并被保留下来, 而且它们之间的全部或大多数二次项和交叉乘积项都具有统计学意义。

【关键词】 反应曲面回归分析; 二次项; 交叉乘积项; 失拟; 等高线; 稳定点

中图分类号: R195.1

文献标识码: A

doi:10.11886/j.issn.1007-3256.2018.03.004

Response surface regression analysis

Hu Liangping^{1,2*}

(1. Graduate School, Academy of Military Sciences PLA China, Beijing 100850, China;

2. Specialty Committee of Clinical Scientific Research Statistics of World Federation of Chinese Medicine Societies, Beijing 100029, China

* Corresponding author: Hu Liangping, E-mail: lphu812@sina.com)

【Abstract】 The paper was aimed at introducing the concepts and functions and the calculation methods by using the statistical software of the response surface regression analysis. Firstly, the basic concepts of the regression analysis was introduced. Secondly, the basic principle of the regression analysis was given. Finally, the response surface regression analysis was demonstrated through one example by using the SAS software. The following two distinct conclusions could be drawn: ① The method mentioned above was suitable for the situation of determining the best production conditions of the multiple quantitative factors. ② The prerequisite conditions of using this approach were as follows. Firstly, the quantitative factors in the experimental research were kept by strictly screening in accordance with the specialty and statistics. Secondly, the whole or the most quadratic terms and cross product terms of the quantitative factors were significant in statistics.

【Keywords】 Response surface regression analysis; Quadratic term; Cross product term; Lack of fit; Contour; Stationary point

1 概 述

1.1 何为反应曲面回归分析

就是在构建二重线性回归模型时, 引入两个自变量的平方项和交叉乘积项, 所构建的二重线性回归模型实际上是三维空间 (x_1, x_2, y) 中的一个二次曲面。由于此曲面的纵轴代表的是因变量(常被称作“响应变量”), 故称此曲面为“响应曲面”或“反应曲面”。也就是说, y 是关于 (x_1, x_2) 的二次函数。若自变量的个数多于 2 个, 所形成的二次函数就应该被称为“超反应曲面回归模型”了。二次反应曲面回归模型如下面的式(1)和(2)所示:

$$y = f(x_1, x_2) \quad (1)$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \beta_5 x_1 x_2 + \varepsilon \quad (2)$$

在式(1)和(2)中, y 为定量的结果变量, 即“因变量”或“响应变量”; x_1, x_2 为两个原因变量, 即自

变量, 在实际的多因素试验研究中, 它们通常是对试验结果 y 有重要影响的试验条件。

1.2 反应曲面回归分析应用的场合

一般来说, 反应曲面回归分析应用在下列场合:

在一个化学或物理或生物学试验研究中, 涉及到两个或两个以上定量影响因素(即自变量), 当各定量因素分别取一个特定水平时就构成了一个特定的“试验条件”。设有一个定量评价指标 y , 则在任何一个特定条件下做试验, y 就会有一个取值。假定研究者有 n 个特定试验条件下分别做了 m 次独立重复试验(若 $m = 1$ 就代表各试验条件下只做了一次试验), 此时, 研究者最希望得到的结果是: ①用一个二次方程式来定量反映因变量 y 随自变量变化而变化的依赖关系; ②各自变量分别取什么值(或水平)时, 因变量 y 可以取得“最大值”或“最小值”。以前述的两点为分析目的的“回归分析”被称为“反应曲面回归分析”^[1]。

简而言之,反应曲面回归分析常应用于确定一个定量多因素的试验研究中的“最佳生产条件”或“最佳工艺配方”,即在各定量因素分别取什么水平时做试验,其试验结果的取值最好(高优指标时,希望结果取得最大值,如产量;低优指标时,希望结果取得最小值,如能源消耗量)。

1.3 反应曲面回归分析的原理

1.3.1 相关基本概念

1.3.1.1 等高线

顾名思义,等高线就是在距离某个水平面相等高度上绘出的一条线。可以设想:一个技术非常高超的飞行员在距离某岛屿一万里的高空中绕着此岛屿飞行一周,飞机尾部喷出的白色雾所形成的“图案”就是一条等高线(高度为一万里)。

用数学语言可描述如下:在一个三维直角坐标系中,把一个二次反应曲面想像成一个“圆顶草帽”,若用一把锋利的锯子在距离坐标平面 k 个单位的高度上去平行地切割“圆顶草帽”(即具有高低起伏的曲面),其切口就是一个环形的曲线,它就是该曲面在 $y = k$ 时的等高线;当 k 在 y 的取值范围内取一系列数值时,就形成了一系列的等高线。研究者根据这些等高线的形状,就可比较清楚地看出:反应曲面在二维平面上呈现出来的且在各个方向上的变化情况。

1.3.1.2 稳定点

反应曲面上的稳定点包括极大值点、极小值点和鞍点。

何为极大值点?若 (x_{10}, x_{20}) 为极大值点,在包含 (x_{10}, x_{20}) 的任何一个小区域内的任何一点 (x_{1i}, x_{2i}) ,都满足 $f(x_{10}, x_{20}) > f(x_{1i}, x_{2i})$,这里, $y = f(x_{10}, x_{20})$ 就是 $(x_1, x_2) = (x_{10}, x_{20})$ 时 y 能取到的极大值;若前面所提及的“小区域”是两个自变量变化的整个区域,则此极大值就是最大值。

同理,若 (x_{10}, x_{20}) 为极小值点,在包含 (x_{10}, x_{20}) 的任何一个小区域内的任何一点 (x_{1i}, x_{2i}) ,都满足 $f(x_{10}, x_{20}) < f(x_{1i}, x_{2i})$,这里, $y = f(x_{10}, x_{20})$ 就是 $(x_1, x_2) = (x_{10}, x_{20})$ 时 y 能取到的极小值;若前面所提及的“小区域”是两个自变量变化的整个区域,则此极小值就是最小值。何为鞍点?若 (x_{10}, x_{20}) 为鞍点,在包含 (x_{10}, x_{20}) 的任何一个小区域内的任何一点 (x_{1i}, x_{2i}) ,都不一定满足 $f(x_{10}, x_{20}) > f(x_{1i}, x_{2i})$ 或 $f(x_{10}, x_{20}) < f(x_{1i}, x_{2i})$ 。也就是说, $y = f(x_{10}, x_{20})$ 的

数值与其邻近点上的数值谁大谁小是不确定的,可能存在方向性,即在某些方向上,鞍点可能属于极大值点;而在另一个方向上,鞍点可能属于极小值点。

1.3.2 计算原理

反应曲面回归模型的构建,若采用 SAS 中 REG 过程来实现,需要引入全部自变量的二次项和交叉乘积项,参见文献[2];若采用 SAS 中 RSREG 过程来实现,则非常简单(参见表 1 后面的 SAS 程序)。仍可以采用最小平方原理推导出正规方程组,通过求解此方程组可以获得模型(1)中参数(即截距项和回归系数)的估计值。

进行反应曲面回归分析的关键是如何对所构建的二次反应曲面回归模型进行分析,它涉及到“等高线”“稳定点”等的计算,因篇幅所限,需要时可查阅有关文献[3]。

2 基于 SAS 进行反应曲面回归分析

2.1 问题与数据结构

设有一个化学试验,涉及到反应温度(temp)和作用时间(time),受试对象为用于化学试验的“样品”。当温度和时间分别取某特定值时,就构成了一个特定的试验条件,试验之后,定量的试验结果“巯基苯并噻唑”(MBT)就会有一个具体的取值。某研究者考虑了“4.0、6.3、12.0、17.7 和 20.0”(h) 5 个不同的反应时间,又考虑了“220、229、250、271 和 280”(°C)5 个不同的温度。两个试验因素全部水平组合共有 25 种,某研究者只选取了其中一部分试验条件进行试验,其因素水平组合及其试验结果见表 1。

表 1 在不同反应时间和温度条件下进行某化学试验得到某种物质的产率 MBT 的结果

反应时间(h)	温度(°C)	定量结果 MBT (%)
4.0	250	83.8
20.0	250	81.7
12.0	250	82.4
12.0	250	82.9
12.0	220	84.7
12.0	280	57.9
12.0	250	81.2
6.3	229	81.3
6.3	271	83.1
17.7	229	85.3
17.7	271	72.7
4.0	250	82.0

注:表 1 对应的试验及资料摘自 SAS 9.3 软件 RSREG 过程的第 1 个“样例”

【说明】在表 1 的每一行中,反应时间与温度分别取不同数值时,构成一个特定的试验条件(被称为试验点或设计点),完全不同的试验条件(试验点或设计点)只有 9 个,因为(4.0,250)出现了两次、(12.0,250)出现了三次。

2.2 研究目的与内容

研究目的:试通过分析表 1 中的资料,求出在反应时间和温度分别取什么数值条件下,所得到的产率 MBT 最高。这样的研究目的常被称为“最优生产条件的确定”问题。

研究内容:以结果变量 MBT 为因变量,以反应温度(temp)和作用时间(time)为两个试验因素(或自变量),构建“二次反应曲面回归模型”;通过此曲面模型,洞察其表现,即绘制出“等高线”图、找出其“稳定点”;若“稳定点”为“极值点”,再进一步求出具体的“极大值”或“极小值”。

2.3 所需要的 SAS 程序

所需要的 SAS 程序如下:

```
data a;
input Time Temp MBT;
label Time = "Reaction Time (Hours)"
Temp = "Temperature (Degrees Centigrade)"
MBT = "Percent Yield Mercaptobenzothiazole";
datalines;
(此处输入表 1 中 12 行 3 列数据)
;
run;
ods graphics on;
proc rsreg data = a plots = (ridge surface);
model MBT = Time Temp / lackfit;
```

残差	自由度	平方和	均方	F 值	Pr > F
缺少拟合	3	124.696053	41.565351	39.63	0.0065
纯误差	3	3.146667	1.048889		
总误差	6	127.842720	21.307120		

以上是关于“失拟检验”的结果:线性不能描述的部分(即拟合失败部分)具有统计学意义($F = 39.63$,

参数	自由度	估计值	标准误差	t 值	Pr > t	代码数据的参数估计值
截距	1	-545.867976	277.145373	-1.97	0.0964	82.173110
Time	1	6.872863	5.004928	1.37	0.2188	-1.014287
Temp	1	4.989743	2.165839	2.30	0.0608	-8.676768
Time * Time	1	0.021631	0.056784	0.38	0.7164	1.384394
Temp * Time	1	-0.030075	0.019281	-1.56	0.1698	-7.218045
Temp * Temp	1	-0.009836	0.004304	-2.29	0.0623	-8.852519

万方数据

```
ridge max;
run;
ods graphics off;
```

【SAS 程序说明】调用 RSREG 过程进行“反应曲面回归分析”;“plots = (ridge surface);”要求系统对反应曲面进行“岭分析”,即呈现出“等高线图”和“稳定点”;“lackfit”要求系统进行“失拟检验”,即检验可否用“多重线性回归模型”取代“二次反应曲面回归模型”;“ridge max”要求系统求出响应曲面上因变量的“最大值”。

2.4 SAS 输出结果及其解释

变量“MBT”的响应曲面:Percent Yield Mercaptobenzothiazole

响应均值	79.916667
均方根误差	4.615964
R ²	0.8003
偏差系数	5.7760

以上是关于因变量 MBT 的有关计算结果。

回归	自由度	I 型平方和	R ²	F 值	Pr > F
线性	2	313.585803	0.4899	7.36	0.0243
二次	2	146.768144	0.2293	3.44	0.1009
叉积	1	51.840000	0.0810	2.43	0.1698
总模型	5	512.193947	0.8003	4.81	0.0410

以上是关于整个二次反应曲面回归模型和其中各部分的假设检验结果:总模型具有统计学意义($F = 4.81, P = 0.041$),线性部分(即 Temp 和 Time 的一次项)具有统计学意义,两个平方项和一个交叉乘积项均无统计学意义。

$P = 0.0065$),说明有必要构建二次反应曲面回归模型。

以上呈现出二次反应曲面回归模型中各项的假设

检验结果,各项对应的 P 均 >0.05 。这个结果并不理想。

因子	自由度	平方和	均方	F 值	Pr > F	标签
Time	3	61.290957	20.430319	0.96	0.4704	Reaction Time (Hours)
Temp	3	461.250925	153.750308	7.22	0.0205	Temperature (Degrees Centigrade)

以上是对涉及“Time”的三项合并的总评价(无统计学意义, $F=0.96, P=0.4704$)、对涉及“Temp”的三项合并的总评价($F=7.22, P=0.0205$),此结果表明:仅温度(Temp)对试验结果(MBT)的影响具有统计学意义。

在图 1 中出现了 9 个“圆圈”,它们代表表 1 中 9 个不同的“设计点(或试验点)”。图中的“弧线”反映了某些“设计点”及其附近“未做试验的点”对应的结果变量(MBT)的等高线。现以图 1 左上角标注“90”的那条“弧线”为例,说明“等高线”的含义:“90”代表“结果变量(MBT)的取值为 90”之意,“弧线”代表横坐标轴上的“温度”的变化范围大约在(220,235)之间,而纵坐标轴上的“反应时间”的变化范围大约在 20℃ 以上。读者可以尝试去解释标注“80”的两条等高线的含义,此处不再赘述。

自变量不同水平组合下因变量 MBT 的最大值及其标准误差见表 2。

【说明】在上面的输出结果中有这样一个结果:“两个平方项和一个交叉乘积项均无统计学意义”,这个结果提示:本例的试验研究中所涉及的两个定量因素对定量结果的影响不存在“二次项”效应,换句话说,基于本例资料,不需要采用反应曲面回归模型,而只需要采用“一元二重线性回归模型(用几何学来理解,就是一个二维平面,而不是一个三维曲面)”。若果真这样做,就不存在“反应曲面”了,也就不存在因变量的“最大值”或“最小值”了。由此得出如下推论:

表 2 自变量不同水平组合下因变量 MBT 的最大值及其标准误差

编码半径	估计响应	标准误差	未编码因子值	
			Time	Temp
0.0	82.173110	2.665023	12.000000	250.000000
0.1	82.952909	2.648671	11.964493	247.002956
0.2	83.558260	2.602270	12.142790	244.023941
0.3	84.037098	2.533296	12.704153	241.396084
0.4	84.470454	2.457836	13.517555	239.435227
0.5	84.914099	2.404616	14.370977	237.919138
0.6	85.390012	2.410981	15.212247	236.624811
0.7	85.906767	2.516619	16.037822	235.449230
0.8	86.468277	2.752355	16.850813	234.344204
0.9	87.076587	3.130961	17.654321	233.284652
1.0	87.732874	3.648568	18.450682	232.256238

(1) 反应曲面回归分析最适合用于需要确定多定量因素的最佳生产条件的试验研究场合。

由表 2 可知:当 Time = 18.451、Temp = 232.256 时,因变量 MBT = 87.733 为各种水平组合条件下的“最大值”。

(2) 应用此法的前提条件是:参与试验研究的定量因素已经过专业和统计学方法严格筛选并被保留下来,并且它们之间的全部或大多数二次项和交叉乘积项都具有统计学意义。

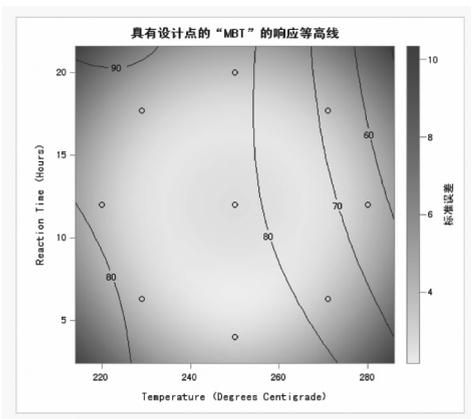


图 1 具有设计点的“MBT”的响应等高线

注:图 1 左边纵坐标轴为“Reaction Time (h)”,即“反应时间(h)”;右边纵坐标轴为“标准误差”;横坐标轴为“Temperature (Degrees Centigrade)”,即“温度(摄氏度)”

参考文献

[1] 胡良平. Windows SAS 6.12 & 8.0 实用统计分析教程[M]. 北京:军事医学科学出版社,2001:574-583.
 [2] 胡良平. 主成分分析应用(I)——主成分回归分析[J]. 四川精神卫生,2018,31(2):128-132.
 [3] 王万中. 试验的设计与分析[M]. 北京:高等教育出版社,2007:318-326.

(收稿日期:2018-05-03)

(本文编辑:唐雪莉)