偏导数的最小二乘估计及其SAS程序实现

李东方

(许昌电气职业学院,河南 许昌 461000)

【摘 要】对于高等数学中求解多元函数偏导数的方法大家都很熟悉,这时可以由偏导数的定义得到一种估计形式,本文用统计中试验设计的方法给出了求偏导数的另外一种估计形式——最小二乘估计,并给出了相应的理论证明、算法和实例分析。

【关键词】偏导数;试验设计;最小二乘估计

【中图分类号】O212.6 【文献标识码】A 【文章编号】1673-1891(2014)01-0036-03

考虑多元函数$y=f(x_1,\cdots,x_m)$,其自变量$x=(x_1,\cdots,x_m)$的定义域Ω为如下多维区间$x_j\in[a_j,b_j]$,$j=1,2,\cdots,m$若此函数在定义域Ω上可导,则在给定的点$x_0=(x_{10},\cdots,x_{m0})\in\Omega$,存在偏导数$\frac{\partial f(x_0)}{\partial x_j}$,$j=1,\cdots,m$。按照高等数学中关于偏导数的定义知,在$\Delta x_j$比较小时,可由下式作为偏导数$\frac{\partial f(x_0)}{\partial x_j}$,$j=1,\cdots,m$的一个估计量:

$$\frac{\partial f(x_0)}{\partial x_j}=\frac{f(x_{10},\cdots,x_{(j-1)0},x_{j0}+\Delta x_j,x_{(j+1)0},\cdots,x_{m0})-f(x_0)}{\Delta x_j},\ j=1,\cdots,m。\quad(1)$$

本文用统计中试验设计的方法给出了偏导数的另一种估计。

预备知识

设一复杂系统的输出值y可用如下模型来表示:

$$y=f(x_1,\cdots,x_m)+\varepsilon,\quad(2)$$

其中$f(x_1,\cdots,x_m)$为定义在区域Ω上的多元函数,ε为试验误差,且$E(\varepsilon)=0$,$Var(\varepsilon)=\sigma^2(x)$,其中$\sigma^2(x)$表示此方差与试验点$x=(x_1,\cdots,x_m)$有关。

对给定的点$x_0=(x_{10},\cdots,x_{m0})\in\Omega$和试验容差$\Delta x=(\Delta x_1,\cdots,\Delta x_m)$,可以按照某个正交设计安排试验。

具体方法如下:

首先需要选取一个设计,即一个$n\times m$的矩阵:

$$H=(a_{ij})_{n\times m},$$

其中a_{ij}为整数,且$0\le a_{ij}\le p_j-1$,$i=1,\cdots,n$ $j=1,\cdots,m$。p_j为因子x_j的水平数。诸p_j可以相等,也可以不等。这里的H可以是某个正交表的全部,也可以是其中的几列。

对设计H中的a_{ij},定义$x_j(a_{ij})$为在区间$[a_j,b_j]$上选取的包括端点在内的等间距点,具体是

$$x_{ij}=x_j(a_{ij})=x_{j0}-\Delta x_j+a_{ij}\cdot\frac{2\Delta x_j}{p_j-1},x_{j0}=\frac{a_j+b_j}{2},\Delta x_j=\frac{b_j-a_j}{2}。\quad(4)$$

这样$(x_{i1},\cdots,x_{im})=(x_1(a_{i1}),\cdots,x_m(a_{im}))$为按设计矩阵$H$的第$i$行$(a_{i1},\cdots,a_{im})$选取的试验点,其中$i=1,\cdots,n$。并称$x_{ij}=x_j(a_{ij})$为设计$H$的水平变换,$X=(x_{ij})_{n\times m}$为设计$H$对应的设计矩阵。且按此设计矩阵做试验得到观测值$y_1,\cdots,y_n$,其中$y_i=f(x_{i1},\cdots,x_{im})+\varepsilon_i$,$i=1,\cdots,n$。基于此观测值$y_1,\cdots,y_n$,就可以得到偏导数$\frac{\partial f(x_0)}{\partial x_j}$,$j=1,\cdots,m$的最小二乘估计。

1 主要结果

为了给出此估计的具体形式,我们需要一些引理和相关记号:

对a_{ij}作正交变换:$C_{ij}=C_j(a_{ij})=a_{ij}\cdot\frac{2}{p_j-1}-1$,$i=1,\cdots,n$,$j=1,\cdots,m$,称$W(p_j)=\frac{1}{p_j}\sum_{k=0}^{p_j-1}C_j(k)^2$为因子$x_j$的试验权重。

因子x_j(或者H的第j列)水平k的示性集合记为$H_{jk}=\{i:a_{ij}=k,i=1,\cdots,n\}$,$k=0,\cdots,p_j-1$。集合$H_{jk}$内元素的个数记为$r_{jk}=|H_{jk}|$。

引理1[1] 假如设计H为强度$t\ge2$的正交设计,定义D为H中第j列的示性集合H_{jk}中的行,并除去第j列组成的设计,则D为强度$t-1$的正交设计。

定理1 在H为强度不小于2的正交设计中,记$C=(C_{ij})_{n\times m}$,C_{ij}为正交变换,x_{ij}为水平变换,则如下结论成立:

1. 列和

$$\frac{1}{n}\sum_{i=1}^{n}C_{ij}=\frac{1}{r_{jk}}\sum_{i\in H_{jk}}C_{ij'}=0,\forall k,j,j',(j\ne j')。\quad(5)$$

2. 正交变换与水平变换间的关系

$$x_{ij}-x_{j0}=C_{ij}\Delta x_j,\forall i,j\quad(6)$$

证明1. 由引理1可知,设计H为强度不小于2的正交设计时,D必是强度不小于 的正交设计,即设计H的第j列中各个数码即水平$k,0\le k\le p_j-1$,出现同样次数。记此次数为r_j,则$r_{jk}=r_j$,$n=p_jr_j$。直接计算知

收稿日期:2013-09-08

作者简介:李东方(1982-),男,助教,硕士研究生,研究方向:试验设计。

$$\frac{1}{n}\sum_{i=1}^{n}C_{ij}=\frac{r_j}{n}\sum_{k=0}^{p_j-1}C_j(k)=\frac{1}{p_j}\left[\frac{p_j(p_j-1)}{2}\cdot\frac{2}{p_j-1}-p_j\right]=0$$ 。

再次用引理1可知，H中相应于因子x_j水平k的示性集合H_{jk}中的行并除去第j列组成的设计D为强度$t-1\geq1$的正交设计，从而是强度1的正交设计，行数为r_{jk}，列数为$m-1$。用上述结论知：

$$\frac{1}{r_{jk}}\sum_{i\in H_{jk}}C_{ij'}=0,\forall j',j'\neq j$$ 。

2. 按因子x_j的水平变换x_{ij}的定义公式，计算知：

$$x_{ij}=x_j(a_{ij})=x_{j0}-\Delta x_j+a_{ij}\cdot\frac{2\Delta x_j}{p_j-1}=x_{j0}+\left[-1+a_{ij}\cdot\frac{2}{p_j-1}\right]\Delta x_j=x_{j0}+C_{ij}\Delta x_j$$

结论成立。

引理2[1] 设$f(x_1,\cdots,x_m)$是一个二阶可微的多元函数，对给定的中心点$x_0=(x_{10},\cdots,x_{m0})$，给定的试验容差$\Delta x=(\Delta x_1,\cdots,\Delta x_m)$，$\Delta=\max\limits_{1\leq j\leq m}\Delta x_j=\Delta x_j$，在多维区间$\prod_{j=1}^{m}[x_{j0}-\Delta x_j,x_{j0}+\Delta x_j]$上选取点$x=(x_1,\cdots,x_m)$，那么如下条件等价：

1. $f(x_1,\cdots,x_m)$的二阶偏导数在多维区间$\prod_{j=1}^{m}[x_{j0}-\Delta x_j,x_{j0}+\Delta x_j]$上有界。

2. 对$f(x_1,\cdots,x_m)$的一阶偏导数$\frac{\partial f(x)}{\partial x_j}$可以$x_0$为中心，在多维区间$\prod_{j=1}^{m}[x_{j0}-\Delta x_j,x_{j0}+\Delta x_j]$泰勒展开

$$\frac{\partial f(x)}{\partial x_j}=\frac{\partial f(x_0)}{\partial x_j}+O(\Delta),\qquad(7)$$

其中$\frac{\partial f(x_0)}{\partial x_j}=\frac{\partial f(x)}{\partial x_j}|_{x=x_0}$，$j=1,\cdots,m$。而$O(\Delta)$为$\Delta$的同阶无穷小。

3. 对$f(x_1,\cdots,x_m)$可以x_0为中心，在多维区间$\prod_{j=1}^{m}[x_{j0}-\Delta x_j,x_{j0}+\Delta x_j]$上线性泰勒展开

$$f(x_1,\cdots,x_m)=f(x_0)+\frac{\partial f(x_0)}{\partial x_1}(x_1-x_{10})+\cdots+\frac{\partial f(x_0)}{\partial x_m}(x_m-x_{m0})+o(\Delta),\qquad(8)$$

其中$\frac{\partial f(x_0)}{\partial x_j}=\frac{\partial f(x)}{\partial x_j}|_{x=x_0}$，$j=1,\cdots,m$。而$o(\Delta)$为$\Delta$的高阶无穷小。

定理2 设f是二阶可导函数，且第二阶偏导数有阶，x_0是试验中心点，$H=(a_{ij})_{n\times m}$为正交设计，$y_1,\cdots,y_n$为试验数据或观测值，那么偏导数$\frac{\partial f(x_0)}{\partial x_j}$的最小二乘估计为

$$\frac{\partial \hat{f}(x_0)}{\partial x_j}=\frac{\frac{1}{p_j}\sum_{k=0}^{p_j-1}C_j(k)\hat{\mu}_{jk}}{W(p_j)\Delta x_j}\qquad(9)$$

其中$\hat{\mu}_{jk}=\frac{1}{r_{jk}}\sum_{i\in H_{jk}}y_i$为水平均值$\mu_{jk}=\frac{1}{r_{jk}}\sum_{i\in H_{jk}}f_i$的估计，$C_j(k)$为正交变换，$W(p_j)$为试验权重。

证明 在模型(2)的条件下，对系统函数f应用引理2可得：

$$y_i=f(x_0)+\frac{\partial f(x_0)}{\partial x_1}(x_{i1}-x_{10})+\cdots+\frac{\partial f(x_0)}{\partial x_m}(x_{im}-x_{m0})+o(\Delta)+\varepsilon_i\qquad(10)$$

其中 $x_{ij}=x_j(a_{ij})=x_{j0}-\Delta x_j+a_{ij}\cdot\frac{2\Delta x_j}{p_j-1}$

将(10)中的$f(x_0)$和诸偏导数看作如下回归模型

$$y_i=\beta_0+\beta_1(x_{i1}-x_{10})+\cdots+\beta_m(x_{im}-x_{m0})+\varepsilon,\ i=1,\cdots,n,$$

中的系数，可得$f(x_0)$和诸偏导数的估计量。

首先记

$$S_{x_{j_1},x_{j_2}}=\sum_{i=1}^{n}(x_{ij_1}-x_{j_10})(x_{ij_2}-x_{j_20}),\ S_{x_j,y}=\sum_{i=1}^{n}(x_{ij}-x_{j0})(y_i-\bar{y})$$

用定理1简化这两类统计量，可以得到：

$$S_{x_{j_1},x_{j_2}}=\sum_{i=1}^{n}C_{ij_1}C_{ij_2}\Delta x_{j_1}\Delta x_{j_2}=n\cdot W(p_{j_1})\cdot\Delta x_{j_1}^2\cdot\delta_{j_1,j_2},\quad \delta_{j_1,j_2}=\begin{cases}0,j_1\neq j_2\\1,j_1=j_2\end{cases},$$

$$S_{x_j,y}=\sum_{i=1}^{n}C_{ij}\Delta x_j(y_i-\bar{y})=\sum_{i=1}^{n}C_{ij}\Delta x_jy_i=\sum_{k=0}^{p_j-1}r_{jk}\left(\frac{1}{r_{jk}}\sum_{i\in H_{jk}}C_{ij}y_i\right)\Delta x_j=\frac{n}{p_j}\sum_{k=0}^{p_j-1}C_j(k)\hat{\mu}_{jk}\Delta x_j$$ 将这些结论用于对最小二乘估计的正规方程的计算就比较方便.正规方程变为：

$$\begin{pmatrix}n&0&\cdots&0\\0&S_{x_1,x_1}&\cdots&0\\ \vdots&\vdots&\vdots&\vdots\\0&0&0&S_{x_m,x_m}\end{pmatrix}\begin{pmatrix}f(x_0)\\ \frac{\partial f(x_0)}{\partial x_1}\\ \cdots\\ \frac{\partial f(x_0)}{\partial x_m}\end{pmatrix}=\begin{pmatrix}n\bar{y}\\S_{x_1,y}\\ \cdots\\S_{x_m,y}\end{pmatrix}$$

所以，

$$\hat{f}(x_0)=\bar{y},\ \frac{\partial \hat{f}(x_0)}{\partial x_j}=\frac{S_{x_j,y}}{S_{x_j,x_j}},\ j=1,\cdots,m$$

从而有，

$$\frac{\partial \hat{f}(x_0)}{\partial x_j}=\frac{\frac{1}{p_j}\sum_{k=0}^{p_j-1}C_j(k)\hat{\mu}_{jk}}{W(p_j)\Delta x_j}$$

2 SAS算法实例分析

考虑多元函数$f(x_1,\cdots,x_6)=e^{x_1x_2}\ln(x_3^2+1)+x_1\cos(3x_4)\sqrt{x_5^2+x_6^2}$，在点$x_0=(1,2,1,2,1,2)$处的偏导数。

SAS主程序如下：%window ask

```
#4 @12 "偏导数的计算" color=blue
#6 @12 "输入函数y=f(x1,x2,...)"
#8 @12 gettext1 100 attr=underline A
=REV_VIDEO
#6 @44 "为几元函数"
@60 getdata1 2 attr=underline A=REV_VIDEO
#10 @20 "(所用符号为+,–,#,/,##,sin,cos,exp,
log,...)"
#16 @12 "所求点(x1,x2,...)可以不写括号"
#18 @12 gettext2 30attr=underline A
=REV_VIDEO
#30 @12 "press ENTER to continue.";
%display ask;
```

我们只需要点击run在ask窗口的相应黑色区域输入：6　exp(x1#x2)#log(x3##2+1)+x1#cos(3#x4)#(x5##2+x6##2)##(1/2)　1,2,1,2,1,2，可以很快得到

$$\left(\frac{\partial f(x_0)}{\partial x_1},\cdots,\frac{\partial f(x_0)}{\partial x_6}\right)=(12.390413,5.1217003,7.3890557,1.8743763,0.4294015,0.8588025)$$

3 结语

由(9)式可知，本文所给出的估计只与所选的设计有关，所以只要所选的设计合适，其估计就应该不错。模拟发现，定理2中(9)式关于偏导数的估计是比较拟合和稳定的。这说明本文用试验设计的思想求偏导数的方法在实际应用中是值得推荐的。

注释及参考文献：

[1]张应山.正交表的数据分析及其构造[D].华东师范大学,2006.

[2]茆诗松,周纪芗,陈颖. 试验设计[M]. 北京：中国统计出版社,2004.

[3]魏木生.广义最小二乘问题的理论与计算[M]. 北京：科学出版社,2006.

[4]黄燕,吴平.SAS统计分析及应用[M].北京:机械工业出版社,2007.

Least Square Estimate of the Partial Derivatives and its SAS Procedure

LI Dong-fang

(Xuchang Electric Vocational College, Xuchang, Henan 461000)

Abstract: We are familiar with the method of the partial derivatives from the Advanced Mathematics of course, a kind of estimate can be obtained from the definition of the partial derivatives. This paper gives another method of the partial derivatives using the designs of experiments, i.e. least square estimate. And also provides its theories, gives its algorithm and some examples.

Key words: partial derivatives; designs of experiments; least square estimate

(上接35页)

New Knowledge about the Mathematical Basis of Finite and Infinite

HE Jiang-lin

(College of Urban and Rural Construction，Chengdu University，Chengdu, Sichuan 610106)

Abstract: With new thinking real number is divided into finite and infinite. The boundary of finite portion and some properties of infinite are offered in this paper. Using the part of the infinitely small and infinitely large determine the nature of the convergence of positive series redefines the limits and definitions. Several assumptions are the proof of infinite. The real number was quantified，the function is more flexible ；a new form of expression of function is given.

Key words: limited ；infinity ；limit ；function of the infinitesimal expression.