Advances in Applied Mathematics 应用数学进展, 2019, 8(7), 1239-1242
Published Online July 2019 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2019.87143

Hans汉斯

D-Optimal Design for Duality Quadratic Polynomial Regression Models in Circle Region

Qinghai Kong

School of Science, Northeastern University, Shenyang Liaoning
Email: shenyanglaok@sina.com

Received: July 1st, 2019; accepted: July 16th, 2019; published: July 23rd, 2019

Abstract

For duality quadratic polynomial regression models in circle region, D-op designs were given and proved according to the criterion "Symmetry + Uniform" with four vertexes of special square and central angle in Circle, and the least squares estimates were given, It is a very useful attempt to apply optimal design theory to polynomial regression models.

Keywords

Regression Model, *D*-Optimal Design, Measure, Saturation Design, The Least Squares Estimates

圆域上的二元二次多项式回归模型的*D*-最优设计

孔庆海

东北大学理学院，辽宁 沈阳
Email: shenyanglaok@sina.com

收稿日期：2019年7月1日；录用日期：2019年7月16日；发布日期：2019年7月23日

摘　要

对圆域上的二元二次多项式回归模型在遵循设计点“对称 + 均匀”的前提下，给出并证明其特定的内接正方形的顶点和圆心组成的饱和设计是***D***-最优设计，并给出了相应设计的最小二乘估计，是把最优设

文章引用: 孔庆海. 圆域上的二元二次多项式回归模型的 *D*-最优设计[J]. 应用数学进展, 2019, 8(7): 1239-1242.
DOI: 10.12677/aam.2019.87143

计理论应用到多元多项式回归模型的又一有益的尝试。

关键词

回归模型，*D*-最优设计，测度，饱和设计，最小二乘估计

1. 引言

在试验设计中，令 R 表示设计利益区域，X 是区域 R 内任一点，回归模型一般形为

$$y = f^{\mathrm{T}}(X)\beta + \varepsilon$$

其中 $f^{\mathrm{T}}(X) = \left(f_1(X), f_2(X), \cdots, f_p(X)\right)$ 是由模型决定的函数向量，y 是响应观测值，而 $\beta^{\mathrm{T}} = \left(\beta_1, \beta_2, \cdots, \beta_p\right)$ 是模型中的待估参数，ε 是误差，通常假设 $E(\varepsilon) = 0$，$D(\varepsilon) = \sigma^2$，$\sigma$ 是已知的。

如果用 $M(\xi)$ 表示测度设计 ξ 的信息矩阵，所谓的 D-最优准则就是使得 $M(\xi)$ 的行列式达到最大，而且测度设计 ξ^* 是 D-最优设计的充分必要条件[1] [2]是方差函数 $d = f^{\mathrm{T}}(X)M^{-1}(\xi^*)f(X) \le p$ (模型中待估参数的个数)。

2. 二元二次多项式回归模型的最优设计

多元多项式回归模型可以用来处理一大类非线性问题，在应用数理统计学中占有重要地位。这里讨论的二元二次多项式回归模型

$$y = \beta_0 + \beta_1 x + \beta_2 y + \beta_{11} x^2 + \beta_{22} y^2 + \varepsilon \tag{1}$$

这里因子空间是圆域 $\Re = \left\{(x, y) \middle| x^2 + y^2 \le 1\right\}$，在模型中含有五个待估参数，我们尝试用圆域上的五个设计点去进行饱和设计，同时使得这些设计点的结构既要对称[3]，同时分布又很均匀[4]，即所谓的“对称 + 均匀”原则。

对模型(1)的饱和设计 ξ 的基本思想是：先取圆周的任意内接正方形的四个顶点和圆心组成的五点设计，不妨取圆心为坐标原点，两个互相垂直的直径分别为 x 轴，y 轴，设五个设计点分别是 (a,b)，$(-a,-b)$，$(b,-a)$，$(-b,a)$，$(0,0)$。

而且采用测度设计，记每个顶点的测度为 α，圆心的测度为 β，即测度矩阵 $D(\xi) = diag(\alpha,\alpha,\alpha,\alpha,\beta)$，其中 $4\alpha + \beta = 1, \alpha > 0, \beta > 0$。

可将模型(1)的函数向量改写为

$$f^{\mathrm{T}}(X) = \left(1, x^2, y^2, x, y\right)$$

设饱和设计 ξ 的结构矩阵为 X，

$$X = \begin{bmatrix} 1 & a^2 & b^2 & a & b \\ 1 & a^2 & b^2 & -a & -b \\ 1 & b^2 & a^2 & b & -a \\ 1 & b^2 & a^2 & -b & a \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix},$$

相应设计 ξ 的信息矩阵为 $M = X^{\mathrm{T}} D(\xi) X$ ，利用 $a^2 + b^2 = 1$，则

$$M = \begin{pmatrix} 1 & 2\alpha & 2\alpha & 0 & 0 \\ 2\alpha & 2\left(a^4+b^4\right)\alpha & 4a^2b^2\alpha & 0 & 0 \\ 2\alpha & 4a^2b^2\alpha & 2\left(a^4+b^4\right)\alpha & 0 & 0 \\ 0 & 0 & 0 & 2\alpha & 0 \\ 0 & 0 & 0 & 0 & 2\alpha \end{pmatrix}$$

要寻求模型的 D-最优设计，即使得行列式 $|M|$ 取最大值的设计，而

$$|M| = 16\alpha^4 (1-4\alpha)\left(2a^2-1\right)^2$$

令 $\begin{cases} \dfrac{\partial |M|}{\partial \alpha} = 0 \\ \dfrac{\partial |M|}{\partial a} = 0 \end{cases}$ ，则有效驻点是 $\begin{cases} \alpha = \dfrac{1}{5} \\ a = 0 \end{cases}$， $\begin{cases} \alpha = \dfrac{1}{5} \\ a = \pm\dfrac{\sqrt{2}}{2} \end{cases}$，

这两组驻点能使得 $|M|$ 取最大值的，只有在驻点 $\alpha = \dfrac{1}{5}$，$a = 0$ 时达到，此时 $|M|_{\max} = \dfrac{16}{5^5}$，此时 $\alpha = \beta = \dfrac{1}{5}$，说明测度设计是均匀的，设计点呈对称结构，它们是 $(0,1)$， $(0,-1)$， $(1,0)$， $(-1,0)$， $(0,0)$。

下面证明上述五点设计对模型(1)是 D-最优设计，由前文知，只要能验证

$$d = f^{\mathrm{T}}(X) M^{-1}\left(\xi^*\right) f(X) \le 5。$$

此时

$$M^{-1}\left(\xi^*\right) = \begin{pmatrix} 5 & -5 & -5 & 0 & 0 \\ -5 & \dfrac{15}{2} & 5 & 0 & 0 \\ -5 & 5 & \dfrac{15}{2} & 0 & 0 \\ 0 & 0 & 0 & \dfrac{5}{2} & 0 \\ 0 & 0 & 0 & 0 & \dfrac{5}{2} \end{pmatrix},$$

$$d = 5 - \frac{15}{2}\left(x^2+y^2\right) + \frac{15}{2}\left(x^4+y^4\right) + 10x^2y^2$$

在区域 $\Re = \left\{(x,y) \middle| x^2+y^2 \le 1\right\}$ 的极大值情况为

① 在 $\Re$ 的内部 $x^2+y^2<1$，由 $\begin{cases} \dfrac{\partial d}{\partial x} = 30x^3 + 20xy^2 - 15x = 0 \\ \dfrac{\partial d}{\partial y} = 30y^3 + 20x^2y - 15y = 0 \end{cases}$，则

$$\begin{cases} x = 0 \\ y = 0 \\ d = 5 \end{cases},\quad \begin{cases} x = 0 \\ y^2 = \dfrac{1}{2} \\ d = \dfrac{25}{8} \end{cases},\quad \begin{cases} x^2 = \dfrac{1}{2} \\ y = 0 \\ d = \dfrac{25}{8} \end{cases},\quad \begin{cases} x^2 = \dfrac{3}{10} \\ y^2 = \dfrac{3}{10} \\ d = \dfrac{11}{4} \end{cases}$$

② 在 $\Re$ 的边界 $x^2+y^2=1$，有 $d=5-5x^2\left(1-x^2\right)\le 5$，当且仅当在 $x=0,\pm 1$ 的时候取最大值。

由①②知道方差函数，即正定二次型 d 的最大值是 5 (等于模型中的待估参数的个数)，而且当且仅当在上述设计点处才取得最大值，这就证明了所采用的饱和均匀的等测度设计是 D 最优设计。

3. 模型参数的最小二乘估计

由于模型(1)的待估参数向量为 $\vec{\beta}=\left(\beta_0,\beta_{11},\beta_{22},\beta_1,\beta_2\right)^{\mathrm{T}}$，模型(1)可以记为

$$y=f^{\mathrm{T}}(X)\vec{\beta}+\varepsilon$$

由最小二乘估计公式[5]

$$\hat{\vec{\beta}}=\left(X^{\mathrm{T}}X\right)^{-1}X^{\mathrm{T}}Y=\frac{1}{5}M^{-1}X^{\mathrm{T}}Y \tag{2}$$

这里列向量 $Y=\left(y_1,y_2,y_3,y_4,y_5\right)^{\mathrm{T}}$，其中 $y_i\left(i=1,2,3,4,5\right)$ 表示第 i 次试验的响应观测值。把上述的 M^{-1},X^{T},Y 的结果代入式(2)，则得到 $\vec{\beta}$ 的最小二乘估计为

$$\hat{\vec{\beta}}=\begin{pmatrix}\hat{\beta}_0\\ \hat{\beta}_{11}\\ \hat{\beta}_{22}\\ \hat{\vec{\beta}}_1\\ \hat{\vec{\beta}}_2\end{pmatrix}=\begin{pmatrix}y_5\\ \frac{1}{2}\left(y_1+y_2\right)-y_5\\ \frac{1}{2}\left(y_3+y_4\right)-y_5\\ \frac{1}{2}\left(y_1-y_2\right)\\ \frac{1}{2}\left(y_3-y_4\right)\end{pmatrix}。$$

参考文献

[1] Silvey, S.D. (1980) Optimal Design. Chapman and Hall, London, 11-13, 52-53.

[2] 关颖男. 最优回归设计[J]. 数理统计与应用概率, 1987, 2(4): 477-492.

[3] 朱伟勇, 段晓东. *D*-最优设计的对称性及其对称构造法[J]. 应用数学学报, 1991, 14(3): 360-367.

[4] Wang, Y. and Fang, K. (1996) Uniform Design of Experiments with Mixtures. *Science in China*, *Ser. A*, No. 3, 42-53.

[5] 茆诗松, 丁元, 周纪芗, 吕乃刚. 回归分析及其试验设计[M]. 上海: 华东师范大学出版社, 1981: 297-302.