Advances in Applied Mathematics 应用数学进展, 2021, 10(3), 694-700
Published Online March 2021 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2021.103076

计算机试验下Kriging模型选择的比较

李　涵[1]，赵建昕[2]，王　晓[1]，李新民[1]

[1]青岛大学数学与统计学院，山东 青岛
[2]海军潜艇学院基础部，山东 青岛
Email: xmliqd@163.com

收稿日期：2021年2月11日；录用日期：2021年3月8日；发布日期：2021年3月16日

摘　要

Kriging模型是计算机试验的一种常用模型，因具有良好的非线性拟合能力而被广泛使用。本文在一般Kriging模型相关研究的基础上，研究了Kriging模型的变量选择问题，并给出了Elastic Net变量选择方法。与Lasso和adaptive Lasso相比，数值模拟表明Elastic Net变量选择方法能够提高拟合模型的准确性和稳定性。

关键词

Kriging模型，模型选择，Elastic Net，Lasso

Comparison of Model Selection for Kriging Model in Computer Experiments

Han Li[1], Jianxin Zhao[2], Xiao Wang[1], Xinmin Li[1]

[1]School of Mathematics and Statistics, Qingdao University, Qingdao Shandong
[2]Basic Courses Department, Navy Submarine Academy, Qingdao Shandong
Email: xmliqd@163.com

Received: Feb. 11th, 2021; accepted: Mar. 8th, 2021; published: Mar. 16th, 2021

Abstract

Kriging model is a common model of computer experiment, which is widely used because of its good nonlinear fitting ability. Based on the research of Kriging model, this paper studies the variable selection of universal Kriging model, and gives the variable selection method of Elastic Net. Compared with Lasso and adaptive Lasso, numerical simulation shows that Elastic Net variable selection method can improve the accuracy and stability of the fitting model.

文章引用：李涵，赵建昕，王晓，李新民．计算机试验下 Kriging 模型选择的比较[J]．应用数学进展，2021，10(3): 694-700. DOI: 10.12677/aam.2021.103076

Keywords

Kriging Model, Model Selection, Elastic Net, Lasso

1. 引言

随着科技的不断进步与发展，计算机试验作为物理实验的替代和辅助正在变得越来越流行。计算机试验的一个主要目标是构建一个廉价的元模型。Kriging 模型最初是由南非的地质学家 Krige 在地质统计学中提出并发展的，Sacks 等[1]于 1989 年将 Kriging 模型引入到计算机试验中。从此，Kriging 模型作为一种重要的元模型，在农业、工业、化学、生物等领域中有着广泛的应用。

Kriging 模型包含平稳高斯过程和均值函数两部分。当均值函数只有一个常数时，称 Ordinary Kriging 模型；大部分的 Kriging 模型被假设均值函数部分已存在一些已知变量，这样的模型称为 Universal Kriging 模型(简称 UK)。当模型中包含大量的输入变量时，识别那些(相对较少的)对响应有显著影响的变量是很重要的。Welch 等人[2]提出可以通过计算机试验进行变量选择，并且指出在保持均值函数不变的情况下，选择对高斯过程有显著影响的变量是非常有意义的。

计算机试验中常用的变量选择方法为贝叶斯方法和惩罚似然方法。对于贝叶斯变量选择，Linkletter 等[3]提出了高斯过程的贝叶斯选择方法；Huang 等[4]提出一种盲 Kriging 模型下对均值函数的变量进行随机搜索的贝叶斯选择方法。对于惩罚似然方法，Li 等[5]提出高斯过程的惩罚似然变量选择方法，Hung [6]提出了对均值函数进行 Lasso 和 adaptive Lasso 惩罚的似然方法，Zhang 等[7]提出了一种均值函数的新的惩罚方法，并从贝叶斯观点出发证明了该方法的有效性。

本文研究均值函数的惩罚似然变量选择方法，在 Lasso 和 adaptive Lasso 惩罚似然变量的基础上，提出 Elastic Net (弹性网络)惩罚。本文的结构如下：第二部分介绍 Kriging 模型及其参数估计的算法；第三部分介绍了几种变量选择方法及其算法；第四部分通过数据模拟对 Kriging 模型的几种变量选择方法进行比较；第五部分实例分析，得出结论。

2. Kriging 模型简述

已知 n 个观察点，记为 $X_n=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，$Y=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$，$y_i=y(x_i)$ 是点 x_i 对应的输出，其中 $y_i\in R$，$i=1,2,\cdots,n$。d 为设计变量 x 的维数。一般 Kriging 模型可写成如下形式：

$$y(x)=\sum_{j=1}^{d}f_j(x)\beta_j+z(x)=f^{\mathrm{T}}(x)\beta+z(x)=m(x)+z(x). \tag{2.1}$$

其中 $m(x)$ 是均值函数，$f_j(x),j=0,1,\cdots,d$ 是已知的基函数，$f(x)=(f_1(x),\cdots,f_d(x))^{\mathrm{T}}$ 是基向量，d 表示回归函数中基函数的个数，β_j 是未知系数，$\beta=(\beta_1,\cdots,\beta_d)^{\mathrm{T}}$ 是未知的回归系数向量，$z(x)$ 是一个高斯随机过程且满足均值为 0，即 $E(z(x))=0$，协方差函数为

$$\operatorname{cov}(z(x_i),z(x_j))=\sigma^2c(\psi;x_i,x_j),i,j=1,2,\cdots,n. \tag{2.2}$$

在许多计算机试验的文献中[8]，常用的相关函数为指数相关函数，Matern 相关函数及高斯相关函数。本

文仅考虑高斯相关函数

$$c\left(\psi;x_i,x_j\right)=\exp\left(-\sum_{i=0}^{d}\frac{\left|x_i-x_j\right|}{\psi_i^2}\right). \tag{2.3}$$

这里的ψ是自相关系数，$\psi=\left(\psi_1,\cdots,\psi_d\right)$。

Santner 等[8]表明最大似然估计优于交叉验证估计，所以本文使用最大似然估计来估计参数，对数似然函数为

$$l\left(\beta,\sigma^2,\psi\right)=-\frac{n}{2}\lg\left(2\pi\right)-\frac{n}{2}\lg\left(\sigma^2\right)-\frac{1}{2}\lg\left|C\left(\psi\right)\right|-\frac{1}{2\sigma^2}\left(y-F\beta\right)^{\mathrm{T}}C^{-1}\left(\psi\right)\left(y-F\beta\right) \tag{2.4}$$

这里的$F=\left(f\left(x_1\right),\cdots,f\left(x_n\right)\right)^{\mathrm{T}}$是$n\times d$的设计矩阵，$C\left(\psi\right)$是高斯随机过程$z\left(x\right)$的相关矩阵，第$ij$个元素为$c\left(x_i,x_j\right)$。在(2.4)中，$\left(\beta,\sigma^2,\psi\right)$均未知。首先假设$\psi$已知，则$\left(\beta,\sigma^2\right)$的估计可以通过最大化(2.4)得到：

$$\hat{\beta}=\left(F^{\mathrm{T}}C^{-1}\left(\psi\right)F\right)^{-1}F^{-1}C^{-1}\left(\psi\right)y, \tag{2.5}$$

$$\hat{\sigma}^2=\frac{1}{n}\left(y-F\beta\right)^{\mathrm{T}}C^{-1}\left(\psi\right)\left(y-F\beta\right). \tag{2.6}$$

此时，(2.4)的最大值为：

$$l_{\max}=-\frac{n}{2}\left(1+\lg\left(2\pi\right)\right)-\frac{1}{2}\left(n\log\hat{\sigma}^2+\log\left|C\left(\psi\right)\right|\right). \tag{2.7}$$

之后设定ψ未知，通过最大化(2.7)得到ψ的估计：

$$\hat{\psi}=\arg\min_{\psi}\left\{n\log\hat{\sigma}^2+\log\left|C\left(\psi\right)\right|\right\}. \tag{2.8}$$

与其它模型相比，Kriging 模型可以同时提供预测位置的预测值和预测误差，假定要预测x^*处的函数值，$\hat{y}\left(x^*\right)$表示该点的一个预测值，那么$\hat{y}\left(x^*\right)$的最佳线性无偏估计(BLUP)为

$$\hat{y}\left(x^*\right)=f^{\mathrm{T}}\left(x^*\right)\hat{\beta}+c\left(x^*,D\right)C^{-1}\left(y-F\hat{\beta}\right).$$

其中，$c\left(x^*,D\right)$是一个$1\times n$的向量，第i个元素是$c\left(x^*,x_i\right)$。

3. Kriging 模型的变量选择方法和算法

考虑 Kriging 模型的变量选择时，设候选基变量为$f\left(x\right)=\left(1,f_1\left(x\right),\cdots,f_p\left(x\right)\right)^{\mathrm{T}}$，$\beta=\left(\beta_0,\beta_1,\cdots,\beta_p\right)^{\mathrm{T}}$，$Z\left(x\right)$是均值为 0，相关矩阵为$C\left(\cdot,\cdot\right)$的高斯过程。为了选出$F\left(x\right)$中的重要变量，参数$\beta$需要通过最大化惩罚下面的似然函数估计得到

$$Q\left(\beta,\sigma^2,\psi\right)=l\left(\beta,\sigma^2,\psi\right)-\sum_{i=1}^{t}P_\lambda\left(\left|\beta_i\right|\right).$$

这里的$l\left(\beta,\sigma^2,\psi\right)$是(2.4)，$P_\lambda$是惩罚函数。

本节首先给出 Hung [6]提出的 Lasso 方法和 adaptive Lasso 方法[9]，然后提出 Elastic Net 方法。

3.1. Lasso 惩罚似然方法

为了估计惩罚 Kriging 模型中的参数，Hung [6]提出了重加权最小角度回归算法(IRLARS)，此算法比较容易实施，因此本文所使用的三种惩罚似然方法都要在此算法的框架下进行。

下面详述 Lasso 惩罚的 IRLARS 算法，而 adaptive Lasso 惩罚在 Lasso 惩罚算法的基础上稍加修改。Lasso 惩罚的 IRLARS 算法如下：

算法 1. Kriging 模型 Lasso 惩罚的 IRLARS 算法

Step 1：设置初值($\hat{\sigma}_0^2$，$\hat{\psi}_0$)，代入(2.3)得到相关矩阵 $C(\hat{\psi}_0)$，再将 $C(\hat{\psi}_0)$ 代入(2.5)得到 $\hat{\beta}_0$；

Step 2：分解 $C(\hat{\psi}_{(l)})^{-1}/\hat{\sigma}_{(l)}^2 = R'R$，更新数据 $y^* = Ry$，$X^* = RX$，利用交叉验证得到 Lasso 惩罚下的最优惩罚参数 λ，求解 Lasso 问题，

$$\hat{\beta}_{(l+1)} = \arg\min \left\| y^* - F^*\beta \right\|^2 + \lambda \sum_{j=1}^{p} \left| \beta_j \right|;$$

Step 3：利用 Step 2 中估计出的 $\hat{\beta}_{(l+1)}$ 代入到(2.6) (2.8)得到($\hat{\sigma}_{l+1}^2$，$\hat{\psi}_{l+1}^2$)；

Step 4：重复 Step 2, Step 3，直至 ψ 收敛；

对于 adaptive Lasso 惩罚的变量选择方法[9]，是在 Lasso 的基础上，引入一个已知的权重向量 w_j，则 adaptive Lasso 惩罚下参数 β 的参数估计为：

$$\hat{\beta}_{al}^{(n)} = \arg\min_{\beta} \left\| y - \sum_{j=1}^{p} x_j \beta_j \right\|^2 + \lambda_n \sum_{j=1}^{p} \hat{w}_j \left| \beta_j \right|. \tag{3.1}$$

这里的 $\hat{w} = 1/\left|\hat{\beta}\right|^{\gamma}$，$\gamma > 0$。在本文研究中，$\hat{w}$ 中的 $\hat{\beta}$ 使用最小二乘估计，即 $\hat{\beta} = \hat{\beta}_{ols}$，$\gamma = 1$。对应的算法只需对算法 1 中的 Step 2 稍加修改：

Step 2*：分解 $C(\hat{\psi}_{(l)})^{-1}/\hat{\sigma}_{(l)}^2 = R'R$，得到 $y^* = Ry$，$X^* = RX$，$F^{**} = F^*/v = RF/v$，利用交叉验证得到 Lasso 惩罚下的最优惩罚参数 λ，求解 Lasso 问题，

$$\hat{\beta}_{(l+1)} = \arg\min \left\| y^* - F^*\beta \right\|^2 + \lambda \sum_{j=1}^{p} \left| \beta_j \right|. \tag{3.2}$$

迭代 $\hat{\beta}_j^{(l+1)} = \hat{\beta}_j^*/v_j$，这里的 $v = \left|\hat{\beta}_{ols}\right|$。

3.2. Elastic Net 惩罚似然方法

对于 Elastic Net 惩罚的变量选择方法[10]，对于固定的非负 λ_1，λ_2，参数 β 的 Elastic Net 估计可通过下式求解，

$$\hat{\beta}_{Enet} = \arg\min_{\beta} \left\{ \left\| \tilde{Y} - \tilde{X}^{\mathrm{T}}\beta \right\|^2 + \lambda_2 \left\| \beta \right\|^2 + \lambda_1 \left\| \beta \right\|_1 \right\}. \tag{3.3}$$

其中，$\left\|\beta\right\|^2 = \sum_{j=1}^{p} \beta_i^2$，$\left\|\beta\right\|_1 = \sum_{j=1}^{p} \left|\beta_i\right|$。由此可见 Elastic Net 的罚函数是岭回归与 Lasso 的凸组合。本文研究中，最优惩罚参数 λ 依旧通过交叉验证得到，$\alpha = 0.7$。

由 Elastic Net 方法的定义可以看出，当(3.3)式中 $\lambda_2 = 0$ 时，Elastic Net 方法就变成了 Lasso 方法，因此 Elastic Net 方法结合岭回归与 Lasso 的优点，既能达到变量选择目的，又具有很好的群组效应。于是只需对算法 1 中的 Step 2 稍加修改：

Step 2**：分解 $C(\hat{\psi}_{(l)})^{-1}/\hat{\sigma}_{(l)}^2 = R'R$，得到 $y^* = Ry$，$X^* = RX$，$F^{**} = F^*/v = RF/v$，利用交叉验证得到 Elastic Net 惩罚下的最优惩罚参数 λ，求解 Elastic Net 问题，

$$\hat{\beta}_{(l+1)} = \arg\min \left\| y^* - F^*\beta \right\|^2 + \lambda_2 \left\| \beta \right\|^2 + \lambda_1 \left\| \beta \right\|_1 .$$

4. 数据模拟

本节对函数模型做 UK 模型拟合(UK)，对 UK 做 Lasso 惩罚(LUK)，对 UK 做 adaptive Lasso 惩罚(ALUK)，和对 UK 做 Elastic Net 惩罚(ENUK)，并对这些方法进行比较，最终从变量识别的准确性和预测精度这两

方面进行评价。变量识别的准确性用以下三个指标：积极变量识别率的平均(AEIR)；消极变量识别率的平均(IEIR)；积极变量个数的平均(MEAN)。预测精度用以下两个指标来评估：均方根预测误差(RMSPE)的平均值 MRMSPE 及标准差(sd(RMSPE))。假设有 n_0 个预测点 $x_1^*,\cdots,x_{n_0}^*$，$\hat{y}\left(x_j^*\right)$ 是 x_j^* 的 BLUP，则 RMSPE 的计算公式如下 RMSPE：

$$\mathrm{RMSPE}=\sqrt{\frac{1}{n_0}\sum_{j=1}^{n_0}\left(y\left(x_j^*\right)-\hat{y}\left(x_j^*\right)\right)^2}$$

显然，AEIR 越大越好，IEIR，RMSPE，sd(RMSPE)越小越好，MEAN 越接近真模型越好。

4.1. 数据模拟 1：已知函数

为了评估模型分别加入惩罚前后在预测精度及变量识别方面的性能，我们考虑一个已知线性函数模型[4]。这个已知函数模型被定义在 12 维($p = 12$)的输入空间 $[0,1]^{12}$ 上，其模型的前六个变量 $x_1,\cdots,x_6$ 对计算机试验输出的影响逐渐变小，其余变量 $x_7,\cdots,x_{12}$ 与输出无关(即零系数)。真模型为：

$$y(x)=0.4x_1+0.3x_2+0.2x_3+0.1x_4+0.05x_5+0.01x_6+\varepsilon. \tag{4.1}$$

这里的 $\varepsilon\sim N\left(0,\sigma_\varepsilon^2\right)$，$\sigma_\varepsilon=0.05$。响应值 y 使用(4.1)独立产生，在 Matlab 上，通过拉丁超立方抽样生成维数 p=12，样本量分别为 N=50,80,100 的样本 D，候选变量 C 由所有的一阶主效应组成。用计算机随机生成的 1000 个样本点作为测试集 G。基于 500 次的重复试验，模拟结果在表 1 中给出。

Table 1. Data simulation results of function model (4.1)
表 1. 函数模型(4.1)的数据模拟结果

样本容量	方法	AEIR (%)	IEIR (%)	MEAN	MRMSPE	sd (RMSPE)
N = 50	UK	100	100	12.00	0.0658	0.0057
	LUK	70.97	12.27	4.99	0.0636	0.0095
	ALUK	70.13	12.13	4.94	0.0638	0.0098
	ENUK	71.07	13.30	5.06	0.0638	0.0096
N = 80	UK	100	100	12.00	0.0601	0.0039
	LUK	80.27	11.60	5.51	0.0578	0.0060
	ALUK	81.00	12.10	5.59	0.0576	0.0057
	ENUK	81.40	13.00	5.66	0.0574	0.0055
N = 100	UK	100	100	12.00	0.0580	0.0030
	LUK	82.27	8.57	5.45	0.0563	0.0040
	ALUK	82.20	8.90	5.47	0.0564	0.0040
	ENUK	82.37	9.73	5.53	0.0563	0.0039

通过表 1 数据可以发现，从 MRMSPE 上分析，UK 模型都高于 LUK，ALUK 和 ENUK 模型，这说明对线性模型进行变量选择可以一定程度上提高模型的预测精度，但 LUK，ALUK 和 ENUK 模型的预测精度相差不大。从识别率方面分析，ENUK 的积极变量识别率 AEIR 和识别个数 MEAN 均高于 LUK 和 ALUK。但 LUK 的消极变量识别率 IEIR 低于 LUK 和 ENUK。对于均值函数为多项式模型，模拟结果说明变量选择能够提高预测精度，并且 ENUK 方法的积极变量识别的准确性方面有明显优势，但消极变量

识别率没有优势。

4.2. 数据模拟 2：钻孔函数

考虑钻孔函数[11]如下：

$$y(x)=2\pi x_3(x_4-x_6)\left\{\log\left(\frac{x_2}{x_1}\right)\left(1+2\frac{x_3x_4}{\log(x_2/x_1)x_1^2x_8}+\frac{x_3}{x_5}\right)\right\}^{-1} \tag{4.2}$$

输入空间是一个矩形区间[0.05, 0.015] × [100, 5000] × [63,070, 115,600] × [990, 1110] × [63.1, 116] × [700, 820] × [1120, 1680] × [9855, 12,045]，拟合模型形式为 $\hat{y}(x)=m(x)+z(x)$，设定均值函数形式为 $m(x)=\beta_0+\beta_1x_1+\cdots+\beta_8x_8$。在 Matlab 上，通过拉丁超立方抽样生成 $n=100$，$d=8$ 的样本 D，计算机随机生成 1000 个样本点作为测试集 G，计算均方根预测误差 RMSPE。同样地，对(4.2)做 UK, LUK, ALUK, ENUK 几种方法的拟合。重复进行 500 次试验，并计算每种情况下的 MRMSPE 以及标准差 sd (RMSPE)得到表 2。

Table 2. Data simulation results of the functional model (4.2)
表 2. 函数模型(4.2)的数据模拟结果

指标 \ 方法	UK	LUK	ALUK	ENUK
AEIR (%)	100.00	98.42	98.36	98.57
MEAN	8.00	7.87	7.87	7.89
MRMSPE	1.0963	0.0115	0.0113	0.0118
sd (RMSPE)	2.9280	0.0028	0.0027	0.0029

从表 2 数据可以看出，从 MRMSPE 和 sd (RMSPE)的数据分析，UK 模型的 MRMSPE 和 sd (RMSPE)远高于 LUK, ALUK 和 ENUK 模型，这说明对非线性模型进行变量选择也可以一定程度上提高模型的预测精度，但 LUK, ALUK 和 ENUK 模型的预测精度相差不大。从变量选择方面分析，ENUK 的识别个数 MEAN 高于 LUK 和 ALUK。对于均值函数为多项式模型，模拟结果表明变量选择能够提高预测精度，并且 ENUK 方法的变量识别有优势。这说明在真实模型为非线性模型的条件下，变量选择大大提高了参数估计的准确性与稳定性。

5. 实例分析：活塞拍击噪声

本节应用本文提出的对 Kriging 模型做变量选择的各种方法，对一个活塞拍击噪声的实例进行分析。活塞拍击是由活塞二次运动引起的一种不必要的发动噪声。通过计算机试验，改变六种因素来减少排挤噪声。这六种因素分别为活塞与缸套之间的间隙下 x_1，峰值压力的位置 x_2，活塞裙部长度 x_3，裙部型线 x_4，裙部椭圆轮廓 x_5，活塞销偏置 x_6。实例的相关数据集来源于 Huang [4]，数据包括 100 个观测值，6 个输入变量，候选变量集 C 包括，所有的线性主效应，二次主效应，正交多项式编码下的两因素交叉效应[12]，因此 C 一共包含 72 个基变量。本研究进行 5 折交叉验证，每次将 100 个观测值随机选出 80 个数据作为训练集用于构建模型，剩下的 20 个作为预测集用于计算 RMSPE，比较 UK 模型，OK 模型，以及分别作 Lasso，adaptive Lasso，Elastic Net 惩罚前后的 MRMSPE，sd (RMSPE)，MEAN。比较 UK 模型，以及分别做 Lasso，adaptive Lasso，Elastic Net 惩罚后得到的 LUK，ALUK，ENUK 的 MRMSPE，sd(RMSPE)，MEAN。模拟结果如表 3 所示。

Table 3. Data simulation results of a piston slap noise example
表 3. 活塞拍击噪声实例的数据模拟结果

指标 \ 方法	UK	LUK	ALUK	ENUK
MEAN	72.00	14.00	14.20	15.80
MRMSPE	0.3187	0.1459	0.1637	0.1264
sd (RMSPE)	0.2470	0.0259	0.0491	0.0214

从表 3 的数据可以看出，ENUK 的 MRMSPE 和 sd (RMSPE)远比 UK, LUK 和 ALUK 模型小；模型变量的个数远低于 UK，而且 ENUK 模型的变量个数与 LUK 和 ALUK 相差不大。通过这个实例结果显示，使用 ENUK 方法可以有效简化模型的同时还可以降低均方预测误差。

6. 结论

本文在一般 Kriging 模型相关研究的基础上，研究了 Kriging 模型的变量选择问题，并给出了 Elastic Net 变量选择方法。模拟结果和实例验证表明，Elastic Net 变量选择方法与 Lasso 和 adaptive Lasso 相比，Elastic Net 能够提高拟合模型的准确性和稳定性。

参考文献

[1] Sacks, J., Welch, W.J., Mitchell, T.J. and Wynn, H.P. (1989) Design and Analysis of Computer Experiments. *Statistical Science*, **4**, 409-423. https://doi.org/10.1214/ss/1177012413

[2] Welch, W.J., Buck, R.J. and Sacks, J. (1992) Screening, Predicting, and Computer Experiments. *Technometrics*, **34**, 15-25. https://doi.org/10.2307/1269548

[3] Linkletter, C., Bingham, D. and Hengartner, N. (2006) Variable Selection for Gaussian Process Models in Computer Experiments. *Technometrics*, **48**, 478-490. https://doi.org/10.1198/004017006000000228

[4] Huang, H., Lin, D.K.J., Liu, M.Q. and Zhang, Q. (2019) Variable Selection for Kriging in Computer Experiments. *Journal of Quality Technology*, **52**, 1-14. https://doi.org/10.1080/00224065.2019.1569959

[5] Li, R. and Sudjianto, A. (2005) Analysis of Computer Experiments Using Penalized Likelihood in Gaussian Kriging Models. *Tecnometrics*, **47**, 111-120. https://doi.org/10.1198/004017004000000671

[6] Hung, Y. (2011) Penalized Blind Kriging in Computer Experiments. *Statistica Sinica*, **21**, 1171-1190. https://doi.org/10.5705/ss.2009.226

[7] Zhang, Y., Yao, W., Ye, S. and Chen, X. (2019) A Regularization Method for Constructing Trend Function in Kriging Model. *Structural and Multidisciplinary Optimization*, **59**, 1221-1239. https://doi.org/10.1007/s00158-018-2127-8

[8] Santner, T.J., Williams, B.J. and Notz, W.I. (2003) The Design and Analysis of Computer Experiments. Springer, New York. https://doi.org/10.1007/978-1-4757-3799-8

[9] Zou, H. (2006) The Adaptive Lasso and Its Oracle Properties. *Journal of the American Statistical Association*, **101**, 1418-1429. https://doi.org/10.1198/016214506000000735

[10] Zou, H. and Hastie, T. (2005) Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society*, **67**, 301-320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

[11] Worley, B.A. (1987) Deterministic Uncertainty Analysis. Oak Ridge National Lab, TN, USA.

[12] Altland, H.W. (2001) Experiments: Planning, Analysis, and Parameter Design Optimization. *Technometrics*, **43**, 368-369. https://doi.org/10.1198/004017001316975952