第 15 卷 第 4 期　　自 动 化 学 报　　Vol. 15, No. 4
1989 年 7 月　　ACTA AUTOMATICA SINICA　　July, 1989

正 交 寻 优 法

王子才　高而强　张家余
（哈尔滨工业大学）

摘　　要

本文提出的正交寻优法是适用于复杂非线性目标函数的近似寻优法．它基于逐点线性化后按梯度法进行寻优，在线性化及求梯度过程中借用正交试验的概念和正交表的性质，简化了矩阵求逆过程．该法具有计算量小、搜索速度快及对初值不敏感等特点．利用该法设计的某型惯导平台稳定回路，得到了理想结果．

关键词——优化，正交试验，梯度法，线性化．

一、引　言

到目前为止，优化问题实际上仍然没得到根本解决．比如，二次型的最优问题加权矩阵选择，还不得不采用寻优的方法来确定．

现有非线性目标函数的寻优方法，可分为两大类[1]：解析法和直接法．但在实际工程中遇到的优化问题目标函数是较复杂的非线性函数．因此，很难直接应用解析法进行寻优；而直接法不仅收敛速度慢，并将遇到高维问题的极大困难．显然，寻求一个收敛速度快且计算量小的新寻优法，是亟待研究的课题．

二、数 学 准 备

设 n 元目标函数 $y(z_1, z_2, \cdots z_n)$ 为较复杂的非线性函数．选择初始点为 $Z^0=(z_1^0, z_2^0, \cdots z_n^0)$ 及相对狭小区间为 $\Delta=(\delta_1, \delta_2, \cdots \delta_n)$．定义

$$z_j^h = z_j^0 + \delta_j, \quad z_j^l = z_j^0 - \delta_j. \tag{1),(2}$$

其中，(1)式表示变量 z_j 的上水平；(2)式表示其下水平．适当选择 z_j^h，z_j^l，代入回归方程，可根据 N 组数据确定回归系数[2]．

为简化计算，变量选择按正交试验的正交表安排．设 $x_{ij}(i=1, 2, \cdots 2^u, j=1, 2, \cdots q, q=2^u-1)$ 表示正交表中的元素，X 表示矩阵 $\{x_{ij}\}$．则正交表具有如下性质[3]：(1)每个元素各个水平在试验中出现相同的次数；(2)任何两个因素的各种不同水平的搭配在试验中都出现相同的次数；(3)具有正交性，即任一列的代数和为零

本文于 1986 年 6 月 13 日收到．

$$\sum_{\alpha} x_{\alpha j} = 0,$$

任两列的内积为零

$$\sum_{\alpha} x_{\alpha i} x_{\alpha j} = 0 (i \neq j).$$

由此可得

$$X^{\tau}X = \begin{bmatrix} \sum_i x_{i1}^2 & & 0 \\ & \sum_i x_{i2}^2 \ddots & \\ 0 & & \sum_i x_{iq}^2 \end{bmatrix} = \begin{bmatrix} 2^u & & \\ & 2^u & \\ & & \ddots \\ & & & 2^u \end{bmatrix}. \tag{3}$$

式中，u 是正交表的基本列数。此结果就是简化算法的基本关系式。

三、公 式 推 导

根据正交表对 n 个二水平自变量目标函数作 m 次计算得

$$\begin{cases} y_1 = \theta_1 z_{11} + \theta_2 z_{12} + \cdots + \theta_n z_{1n} + \varepsilon_1, \\ y_2 = \theta_1 z_{21} + \theta_2 z_{22} + \cdots + \theta_n z_{2n} + \varepsilon_2, \\ \vdots \\ y_m = \theta_1 z_{m1} + \theta_2 z_{m2} + \cdots + \theta_n z_{mn} + \varepsilon_m. \end{cases} \tag{4}$$

设 $\boldsymbol{y} = [y_1, y_2, \cdots, y_m]^{\tau}$, $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m]^{\tau}$, $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_n]^{\tau}$, (5),(6),(7)

则(4)式可写为矩阵方程形式

$$\boldsymbol{y} = Z\boldsymbol{\theta} + \boldsymbol{\varepsilon}. \tag{8}$$

式中，$\boldsymbol{\theta}$ 是待估计向量；$\boldsymbol{\varepsilon}$ 是误差向量；$\boldsymbol{y}$ 是目标函数向量，并有

$$Z = \begin{bmatrix} z_{11} & z_{12} \cdots z_{1n} \\ z_{21} & z_{22} \cdots z_{2n} \\ \vdots & \vdots \\ z_{m1} & z_{m2} \cdots z_{mn} \end{bmatrix}. \tag{9}$$

其中，z_{ij} 是根据正交表取值的，

$$z_{ij} = z_{ij}^0 + x_{ij}\delta_{ij}. \tag{10}$$

按照最小二乘法，所估计的 $\hat{\theta}$, $\hat{y}$ 值为

$$\hat{\boldsymbol{\theta}} = [Z^{\tau}Z]^{-1}Z^{\tau}\boldsymbol{y}, \quad \hat{\boldsymbol{y}} = \hat{\theta}_1 z_1 + \hat{\theta}_2 z_2 + \cdots + \hat{\theta}_n z_n. \quad (11),(12)$$

做变换

$$x_{ij} = \frac{z_{ij} - z_{ij}^0}{\delta_{ij}}, \quad i = 1, 2, \cdots, m; \ j = 1, 2, \cdots, n, \tag{13}$$

$$z_{ij} = x_{ij}\delta_{ij} + z_{ij}^0. \tag{14}$$

将(14)式代入(4)式

$$\begin{aligned} y_j &= \theta_1\delta_{j1}x_{j1} + \theta_1 z_{j1}^0 + \theta_2\delta_{j2}x_{j2} + \theta_2 z_{j2}^0 + \cdots + \theta_n\delta_{jn}x_{jn} + \theta_n z_{jn}^0 + \varepsilon_j \\ &= \theta_1^* x_{j1} + \theta_2^* x_{j2} + \cdots + \theta_n^* x_{jn} + y_{j0} + \varepsilon_j. \end{aligned} \tag{15}$$

式中，$\theta_i^* = \theta_i\delta_{ji}$，$y_{j0} = y(z_{j1}^0, z_{j2}^0, \cdots, z_{jn}^0)$.　　(16),(17)

根据(11)式可求得估值 $\hat{\boldsymbol{\theta}}^* = [X^\tau X]^{-1}X^\tau[\boldsymbol{y} - \boldsymbol{y}_0]$.　　(18)

经过这样处理后，在 Z^0 点处以 Δ 为邻域，即可用线性方程

$$\hat{y} = \hat{\theta}_1^* x_1 + \hat{\theta}_2^* x_2 + \cdots + \hat{\theta}_n^* x_n + y_0 \tag{19}$$

逼近原目标函数.

将(13)式代入(19)式得

$$\hat{y} = \frac{\hat{\theta}_1^*}{\delta_1} z_1 + \frac{\hat{\theta}_2^*}{\delta_2} z_2 + \cdots + \frac{\hat{\theta}_n^*}{\delta_n} z_n - \left(\frac{\hat{\theta}_1^*}{\delta_1} z_1^0 + \frac{\hat{\theta}_2^*}{\delta_2} z_2^0 + \cdots + \frac{\hat{\theta}_n^*}{\delta_n} z_n^0\right) + y_0. \tag{20}$$

考虑到(16)式及(17)式

$$\frac{\hat{\theta}_1^*}{\delta_1} z_1^0 + \frac{\hat{\theta}_2^*}{\delta_2} z_2^0 + \cdots + \frac{\hat{\theta}_n^*}{\delta_n} z_n^0 = y_0,$$

则(20)式可写成

$$\hat{y} = \frac{\hat{\theta}_1^*}{\delta_1} z_1 + \frac{\hat{\theta}_2^*}{\delta_2} z_2 + \cdots + \frac{\hat{\theta}_n^*}{\delta_n} z_n. \tag{21}$$

基于上述结果，将 $\hat{y}$ 对 z_i 求偏导即可得到目标函数在 Z^0 点处梯度的简单表示式

$$\nabla \hat{y} = \left\{\frac{\hat{\theta}_1^*}{\delta_1}, \frac{\hat{\theta}_2^*}{\delta_2}, \cdots, \frac{\hat{\theta}_n^*}{\delta_n}\right\}. \tag{22}$$

负梯度方向为函数值下降最快的方向. 于是，就可以在选定的起始点 Z^0 的基础上，选取合适的步长 λ 进行搜索，迭代公式为

$$z_i^{k+1} = z_i^k - \lambda \frac{\|\Delta\|}{\|\nabla \hat{y}\|} \frac{\hat{\theta}_i^*}{\delta_i}. \tag{23}$$

式中，$\|\Delta\|/\|\nabla\hat{y}\|$ 是考虑 Δ 大小及梯度大小引入的修正系数. 由于 $[X^\tau X]$ 是对角阵，故简化了求逆计算. 据(18)式得

$$\hat{\theta}_i^* = \frac{1}{m}\sum_{i=1}^{m} x_{ij}(y_i - y_0),\ j = 1, 2, \cdots, n. \tag{24}$$

式中，$m = 2^u$ 是正交表的基本列数.

四、讨　　论

正交寻优法适合于解决无法求导或很难求导的非线性函数寻优问题. 表 1 给出了该法与单纯形加速法的比较结果. 该法具有如下特点:

(1) 对初值选择不敏感. 对于大多数算法，若初值选得偏离极值点太远，将使计算量增加. 但如何选得接近极值点又很困难. 正交寻优法对初值选择不敏感，解决了这个困难.

(2) 算法简单、计算工作量小、搜索速度较快. 目标函数越复杂，维数越高，这点越明显.

表　1

初值 $x_j^{(0)}$	y^*（极小值）		计算目标函数次数		目标函数 y
	单纯形加速法	正交寻优法	单纯形加速法	正交寻优法	
0,0	8.00001	8.00005	76	56	$x_1^2+x_2^2-x_1x_2-10x_1-4x_2+60$
10,3	8.00001	8.000019	64	48	
20,1	8.00001	8.00001	69	56	
100,4	8	8	317	100	
1000,4	8	8	2567	100	
1000,1000	8	8	4967	100	
10,20,30,40	8.99692E-7	1.49049E-7	103	47	$(3x_1x_3)^2/[x_1x_4+x_2^2+x_3^2+3x_1x_3+(x_1x_3-x_4^2)]$
62,31,24,18	3.28446E-3	1.01755E-3	223	125	$\lvert x_1x_2+[\ln x_2^2]^2[\ln x_2^2+x_1x_2]+\sin(x_1x_3-x_4^2)\rvert$
10,20,30,40	9.39846E-4	2.12491E-5	291	97	$\lvert x_1x_4+[\ln x_2^2][1+x_1x_3]+x_2+\cos(x_1x_3-x_2^2)\rvert$
5,3,4,8	7.18146E-8	7.21651E-7	199	42	$\left[\left(\frac{1}{x_1}+\frac{1}{x_2}\right)^2-\frac{1}{x_3^2}+x_3x_4\right]^2$

(3) 若目标函数有多个极值点，可能一次寻优得到的是局部极值点．则可改变初始值多次寻优，将所得结果进行比较，找出极值点；或在得到局部极值点后，加大步长进行搜索．

(4) 收敛速度与步长有关．若对步长选择加以改进，收敛速度将加快．

五、应 用 实 例

应用正交寻优法设计某型惯导平台的稳定回路，其状态方程为

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\end{bmatrix}=\begin{bmatrix}0 & 1895 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & -11250 & -2 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix}+\begin{bmatrix}0\\ 0\\ 1/250\\ 0\end{bmatrix}u.$$

式中，x_2 为平台绕稳定轴的偏差角．为保证平台在受到干扰时，要求 x_2 尽量小且过渡过

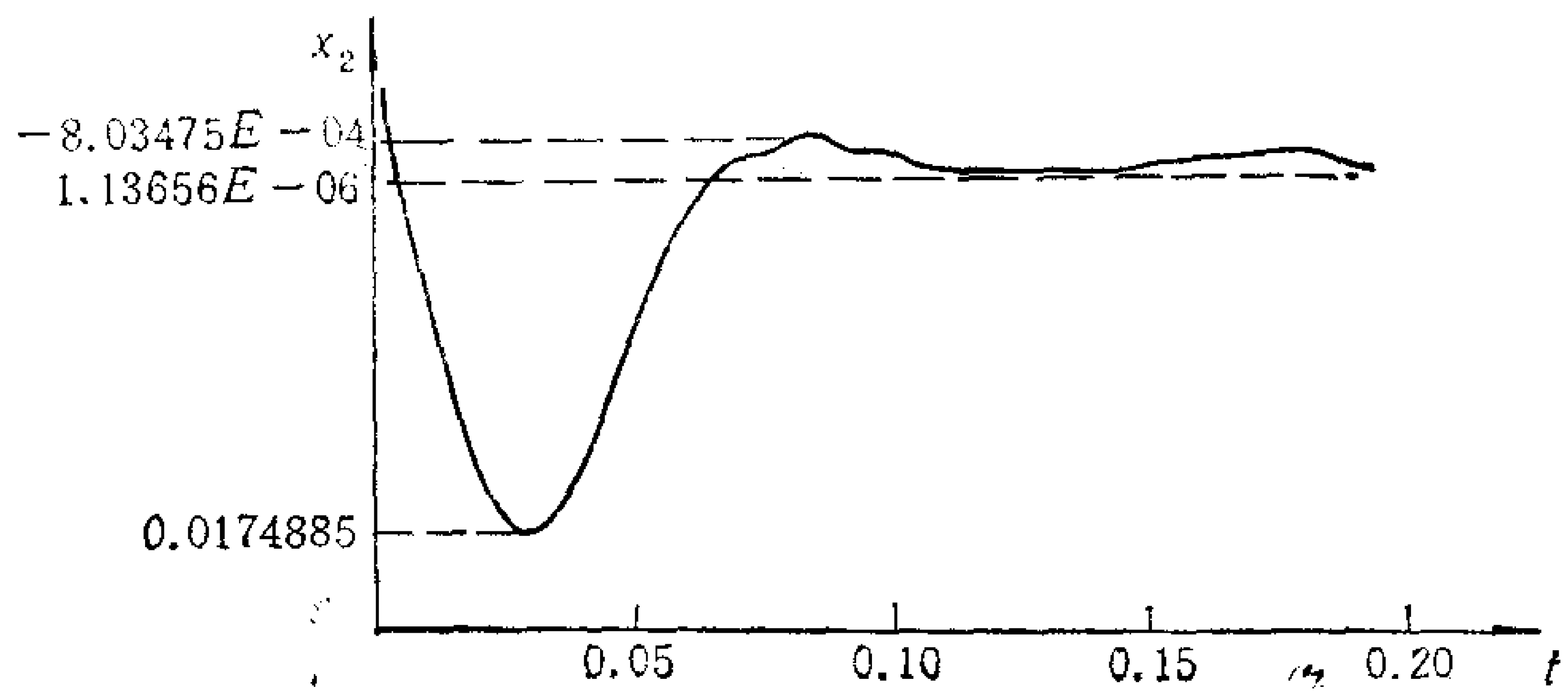

图　1

程最短，可设定目标函数为

$$J=\int_0^\infty t|x_2|\,dt\rightarrow \min.$$

应用正交寻优法求得的最优反馈矩阵为

$$K=[215600,11640000,159920,4120000].$$

x_2 的过渡过程曲线如图 1 所示.

参　考　文　献

[1] 范鸣玉等，最优化技术基础，清华大学出版社 (1982 年).
[2] 上海师大，回归分析及其试验设计，上海教育出版社(1978 年).
[3] 北大概率统计组，正交设计，人民教育出版社(1976 年).

AN ORTHOGONAL OPTIMIZATION METHOD

Wang Zicai　Gao Erqiang　Zhang Jiayu

(Harbin Institute of Technology)

Abstract

The method presented in this paper is an approximate optimization method for nonlinear objective function. Based on point to point linearization, it is optimized by the gradient method. Inverse matrix calculation is simplified by the concept of orthogonal test and orthogonal table in this process. This method possesses the advantages of calculation in small amounts, rapid search and insensitive to initial value. With this method, satisfactory result has been obtained in the design of the stationary system for an inertial guidance platform.

Key words ——optimization; orthogonal test; gradient method; linearization.