第18卷 第5期 自 动 化 学 报 Vol. 18, No. 5
1992年9月 ACTA AUTOMATICA SINICA Sep., 1992

短文

递推计算灵敏度的初值正交计算法[1)]

崔平远 吴瑶华 黄文虎 李乃宏
(哈尔滨工业大学空间科学与技术系，150006)

摘 要

本文对连续-离散系统的极大似然估计方法提出了一种灵敏度初值的正交计算法，它是对具有递推计算灵敏度的修正牛顿-拉夫森算法的改进. 文中通过借用正交试验的概念和正交表的性质，解决了原有算法的初始矩阵求逆所存在的困难.

关键词: 连续-离散系统，极大似然估计，灵敏度，正交试验.

一、引 言

应用极大似然法对连续-离散系统进行参数估计，实际上是一个似然函数的优化计算问题. 已有的 modified Newton-Raphson (MNR) 等算法，都是采用积分求解灵敏度方程计算输出变量关于未知参数的灵敏度[1]，这对于含有 n 个状态方程、p 个未知参数的连续-离散系统，灵敏度方程为 $n\times p$ 个，积分求解的困难是显然的. Murphy 在文献[2]中给出了一种递推计算灵敏度方法，避免了积分求解灵敏度方程. 然而，由于其灵敏度初值的计算问题没有解决而不能推广应用. 为此，本文给出一种借助正交试验的正交表计算灵敏度初值的方法，通过对未知参数初始摄动的合理选取，将高维矩阵的求逆运算转化为标量的求倒数运算，解决了灵敏度初值的计算问题.

二、递推计算灵敏度算法

考虑下述非线性连续-离散系统:

$$\dot{\boldsymbol{x}}(t)=\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\theta}),\quad \boldsymbol{x}(0)=\boldsymbol{x}_0, \tag{1}$$

$$\boldsymbol{y}(t)=\boldsymbol{g}(\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\theta}), \tag{2}$$

$$\boldsymbol{z}(i)=\boldsymbol{y}(i)+\boldsymbol{v}(i),\quad i=1,2,\cdots,N, \tag{3}$$

$$E\{\boldsymbol{v}(i)\}=0,\quad E\{\boldsymbol{v}_i\boldsymbol{v}_j^T\}=R\delta_{ij},$$

其中 $\boldsymbol{x}\in R^n$、$\boldsymbol{u}\in R^r$、$\boldsymbol{\theta}\in R^p$、$\boldsymbol{y}\in R^m$ 分别是状态、输入、参数和输出向量，$\boldsymbol{z}_i$ 和 $\boldsymbol{v}_i$ 分别是在 $t=t_i$ 时刻的测量向量和测量噪声向量，R 是测量噪声方差阵，N 是数据样本长度. 则参数向量的极大似然估计的指标函数为[1]

本文于1990年2月20日收到.
1) 国家自然科学基金资助项目.

$$J(\boldsymbol{\theta})=\frac{1}{2}\sum_{i=1}^{N}[\boldsymbol{z}(i)-\hat{\boldsymbol{y}}(i)]^{T}\hat{R}^{-1}[\boldsymbol{z}(i)-\hat{\boldsymbol{y}}(i)]+\frac{N}{2}\ln|\hat{R}|, \tag{4}$$

其中，$\hat{\boldsymbol{y}}$ 是 $\boldsymbol{y}$ 的预报值，$\hat{R}$ 是 R 的估计值.

以参数向量 $\boldsymbol{\theta}$ 的某一初值开始,使用 MNR 算法可以获得参数的校正估计[1]

$$\boldsymbol{\theta}_{k+1}=\boldsymbol{\theta}_{k}+\Delta\boldsymbol{\theta}_{k}, \tag{5}$$

$$\Delta\boldsymbol{\theta}_{k}=\left[\sum_{i=1}^{N}G_i^T\hat{R}^{-1}G_i\right]^{-1}\left[\sum_{i=1}^{N}G_i^T\hat{R}^{-1}(\boldsymbol{z}(i)-\boldsymbol{y}(i))\right]\Bigg|_{\theta=\theta_k}, \tag{6}$$

其中，$G_i=\left\{\frac{\partial y_k}{\partial\theta_j}\right\}_i$，称为灵敏度矩阵，$y_k$ 和 θ_j 分别是向量 $\boldsymbol{y}$ 和 $\boldsymbol{\theta}$ 的第 k 个和第 j 个元素. 递推计算灵敏度的基本思想和算法如下[2]:

在未知参数向量空间把输出向量 $\boldsymbol{y}(\boldsymbol{\theta})$ 拟合成未知参数的函数，并假设相应的灵敏度近似等于拟合曲面的斜率. 在 p 维参数空间的 $p+1$ 个样本点上拟合 $\boldsymbol{y}(\boldsymbol{\theta})$ 为 $\boldsymbol{\theta}$ 的一阶多项式

$$y_{ki}^{j}(\boldsymbol{\theta}^{j})=s_{k0}+s_{k1}\theta_1^j+\cdots+s_{kp}\theta_p^j, \tag{7}$$

其中 i 表示时间上的第 i 时刻，k 表示输出向量的第 k 个元素，j 表示 p 维参数空间的 $p+1$ 个样本点之一. $s_{k1},\cdots,s_{kp}$ 是相应的斜率或灵敏度 $(\partial y_k/\partial\theta_j)_i$，$s_{k0}$ 是一常值. 在每一样本点上有

$$y_k^j(\boldsymbol{\theta}^j)=y_k(\boldsymbol{\theta}^j), \tag{8}$$

y_k 是输出向量 $\boldsymbol{y}$ 的第 k 个元素，y_k^j 是平面拟合向量 $\boldsymbol{y}_k$ 的第 j 个元素. 对 y_1 在时刻 i 有

$$\begin{bmatrix}y_{1i}^0\\ y_{1i}^1\\ \vdots\\ y_{1i}^p\end{bmatrix}=\begin{bmatrix}1 & \theta_1^0\cdots\theta_p^0\\ 1 & \theta_1^1\cdots\theta_p^1\\ \vdots & \vdots \quad\quad \vdots\\ 1 & \theta_1^p\cdots\theta_p^p\end{bmatrix}\begin{bmatrix}s_{10}\\ s_{11}\\ \vdots\\ s_{1p}\end{bmatrix}_i. \tag{9}$$

消去第一行得

$$\begin{bmatrix}y_{1i}^1-y_{1i}^0\\ y_{1i}^2-y_{1i}^0\\ \vdots\\ y_{1i}^p-y_{1i}^0\end{bmatrix}=\begin{bmatrix}\Delta\theta_1^1 & \Delta\theta_2^1\cdots\Delta\theta_p^1\\ \Delta\theta_1^2 & \Delta\theta_2^2\cdots\Delta\theta_p^2\\ \vdots & \vdots \quad\quad \vdots\\ \Delta\theta_1^p & \Delta\theta_2^p\cdots\Delta\theta_p^p\end{bmatrix}\begin{bmatrix}s_{11}\\ s_{12}\\ \vdots\\ s_{1p}\end{bmatrix}_i, \tag{10}$$

其中，$\Delta\theta_k^j=\theta_k^j-\theta_k^0$. 将(10)式记为

$$\Delta\boldsymbol{y}_{1i}=\Delta X\boldsymbol{s}_{1i}. \tag{11}$$

因此,在时刻 i 输出向量 $\boldsymbol{y}$ 的第一个元素(y_1)关于参数向量的灵敏度由下式给出:

$$\boldsymbol{s}_{1i}=[\Delta X]^{-1}\Delta\boldsymbol{y}_{1i}. \tag{12}$$

由上式得到灵敏度初值之后,进一步推导可以获得灵敏度的递推计算公式[2].

三、灵敏度初值的正交计算

虽然递推计算灵敏度算法解决了积分计算灵敏度方程和预先确定模型结构问题，但对未知参数数目较多的系统，方程(12)中的矩阵 ΔX 求逆运算难以实现. 本文借助安排正交试验的正交表给出如下灵敏度初值的正交计算法.

设 $a_{ml}(m=1,2,\cdots,2^u; l=1,2,\cdots,q; q=2^u-1$, u 是二水平正交表的基本列数)表示正交表的元素，A 表示矩阵 $\{a_{ml}\}$，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，

$$\sum_{\alpha} a_{\alpha l}=0.$$

任两列的内积为零，即

$$\sum_{\alpha} a_{\alpha m}a_{\alpha l}=0,\ (m\neq l).$$

由此可得

$$A^TA=\begin{bmatrix}\sum_m a_{m1}^2 & & & 0\\ & \sum_m a_{m2}^2 & & \\ & & \ddots & \\ 0 & & & \sum_m a_{mq}^2\end{bmatrix}=\begin{bmatrix}2^u & & 0\\ & 2^u & \\ 0 & & \ddots \\ & & & 2^u\end{bmatrix}. \tag{13}$$

根据正交表选择(9)式中矩阵 X 的元素 $\theta_i^j(i=1,2,\cdots,p;\ j=0,1,\cdots,M;\ M=2^u)$ 如下：

$$\theta_i^j=\theta_i^0+a_{ji}\Delta\theta_i. \tag{14}$$

当 $j=0$ 时，定义 $a_{ji}=0$，$\Delta\theta_i$ 是 θ_i 相对 θ_i^0 的摄动量。则方程(10)有如下形式：

$$\begin{bmatrix}y_{1i}^1-y_{1i}^0\\ y_{1i}^2-y_{1i}^0\\ \vdots\\ y_{1i}^M-y_{1i}^0\end{bmatrix}=\begin{bmatrix}a_{11}\Delta\theta_1 & a_{12}\Delta\theta_2 & \cdots a_{1p}\Delta\theta_p\\ a_{21}\Delta\theta_1 & a_{22}\Delta\theta_2 & \cdots a_{2p}\Delta\theta_p\\ \vdots & \vdots & \vdots\\ a_{M1}\Delta\theta_1 & a_{M2}\Delta\theta_2 & \cdots a_{Mp}\Delta\theta_p\end{bmatrix}\begin{bmatrix}s_{11}\\ s_{12}\\ \vdots\\ s_{1p}\end{bmatrix}_i$$

$$=\begin{bmatrix}a_{11} & a_{12}\cdots a_{1p}\\ a_{21} & a_{22}\cdots a_{2p}\\ \vdots & \vdots\quad \vdots\\ a_{M1} & a_{M2}\cdots a_{Mp}\end{bmatrix}\begin{bmatrix}\Delta\theta_1 s_{11}\\ \Delta\theta_2 s_{12}\\ \vdots\\ \Delta\theta_p s_{1p}\end{bmatrix}_i. \tag{15}$$

记

$$\Delta \boldsymbol{y}_{1i}=\Delta X^*\boldsymbol{s}_{1i}^*, \tag{16}$$

则矩阵 ΔX^* 就是(13)式中的矩阵 A。由此可得

$$\boldsymbol{s}_{1i}^*=[(\Delta X^*)^T\Delta X^*]^{-1}(\Delta X^*)^T\Delta\boldsymbol{y}_{1i}, \tag{17}$$

$$s_{1k}^*=\frac{1}{M}\sum_{j=1}^{M}a_{jk}(y_{1i}^j-y_{1i}^0),\quad k=1,2,\cdots,p. \tag{18}$$

据此，灵敏度初始值由下式给出：

$$s_{1i}=[s_{11},s_{12},\cdots,s_{1p}]^T, \tag{19}$$

$$s_{1k}=s_{1k}^*/\Delta\theta_k, k=1,2,\cdots,p. \tag{20}$$

四、结束语

正交计算法通过适当增加参数空间的样本点(由 $p+1$ 个样本点增加为 $M+1$ 个样本点),将(12)式中的 $p\times p$ 维矩阵求逆运算转化为标量运算。与原算法相比,该算法有如下特点:

1) 在计算量方面,原算法在 p 维参数空间选择 $p+1$ 个样本点对 $y_{ki}^{j}(j=0,1,\cdots,p)$ 进行计算,然后通过(12)式计算灵敏度,当 p 较大时矩阵求逆运算难以实现;本文算法是根据正交表构成 $M+1(p\leqslant M<2p)$ 个样本点对 $y_{ki}^{j}(j=0,1,\cdots,M)$ 进行计算,然后由(17)—(20)式通过一组标量运算求得灵敏度,不存在任何数值计算困难。

2) 对参数初始摄动幅度的选择需基于参数变化的灵敏度进行。原算法对 $\Delta\theta_k^j(k=1,2,\cdots,p;\ j=1,2,\cdots,p)$ 的选择较困难,容易导致矩阵 ΔX 病态而无法进行求逆运算;本文算法对每一参数只需根据其灵敏程度选取一个合适的 $\Delta\theta_i(i=1,2,\cdots,p)$,然后根据正交表构成 ΔX,这样选择的 $\Delta\theta_i^j(i=1,2,\cdots,p;\ j=1,2,\cdots,M)$ 既具有正交试验的最优性,又不需进行矩阵求逆运算。

总之,正交计算法在保留原算法三个优点[2]的基础上,既解决了参数初始摄动选择问题,又用标量运算代替了矩阵求逆,所付出的代价是将参数空间的样本点由 $p+1$ 个增至 $M+1$ 个。

参 考 文 献

[1] Iliff, K. W. and Maine, R. E., More than You May Want to Know About Maximum Likelihood Estimation, AIAA (1984), 84-2070, 1—24.

[2] Murphy, P. C. and Klein, V., Maximum Likelihood Algorithm Using an Efficient Scheme for Computing Sensitivities and Parameter Confidence Intervals, AIAA (1984), 84-2084, 131—139.

[3] 北大概率统计组,正交设计,人民教育出版社,(1976).

AN ORTHOGONAL COMPUTATION METHOD OF INITIAL VALUE FOR RECURSIVE COMPUTING SENSITIVITIES

Cui Pingyuan Wu Yaohua Huang Wenhu Li Naihong

(Harbin Institute of Technology, 150006)

Abstract

This paper presents an orthogonal computation method of initial value of sensitivity for maximum likelihood estimation method to he continuous-discrete system. This method is an improvement to the modified Newton-Raphson scheme with estimated sensitivities (MNRES) method. The difficulty to invert the initial matrix in MNRES method is solved by the concept of orthogonal test and orthogonal table

Key words: Continuous-discrete system; maximum likelihood estimation; sensitivity; orthogonal test.