第 20 卷 第 1 期　　自 动 化 学 报　　Vol. 20, No. 1
1994 年 1 月　　ACTA AUTOMATICA SINICA　　Jan., 1994

简 报

极大似然估计的递推计算灵敏度算法[1)]

崔平远　吴瑶华　黄文虎　李乃宏
(哈尔滨工业大学航天工程与力学系　150001)

关键词: 连续-离散系统,极大似然估计,灵敏度,正交试验.

1 引言

采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，而其核心又是灵敏度的计算[1]. 为了避免已有解析求导法在计算灵敏度方面所存在的困难，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. 文献[3]进一步给出了基于正交表计算灵敏度初值的方法,解决了递推计算灵敏度的初值计算问题. 本文对灵敏度的递推算式做了进一步推导，给出了一个易于实现的递推计算灵敏度算法.

2 递推计算灵敏度算法构成

考虑下述非线性连续-离散系统

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

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

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

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

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. 则使用 MNR (modified Newton-Raphson) 算法可以获得参数向量的极大似然校正估计[3]

$$\boldsymbol{\theta}_{k+1}=\boldsymbol{\theta}_k+\Delta\boldsymbol{\theta}_k \tag{4}$$

$$\Delta\boldsymbol{\theta}_k=\left[\sum_{i=1}^{N}G_i^T\hat{R}^{-1}G_i\right]^{-1}\left[\sum_{i=1}^{N}G_i^T\hat{R}^{-1}(\boldsymbol{z}(i)-\hat{\boldsymbol{y}}(i))\right]\Bigg|_{\theta=\theta_k} \tag{5}$$

本文于 1992 年 2 月 14 日收到.
1) 国家自然科学基金资助项目.

其中，$\hat{\boldsymbol{y}}$ 是 $\boldsymbol{y}$ 的预报值，$\hat{R}$ 是 R 的估计值，$G_i = \left\{\frac{\partial y_k}{\partial \theta_j}\right\}_i$，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算法为[2]：在 p 维参数空间的 $p+1$ 个样本点上把输出向量 $\boldsymbol{y}(\boldsymbol{\theta})$ 拟合为 $\boldsymbol{\theta}$ 的一阶多项式

$$y^j_{k_i}(\boldsymbol{\theta}^j) = s_{k_0} + s_{k_1}\theta^j_1 + \cdots + s_{k_p}\theta^j_p \tag{6}$$

其中 i 表示时间上的第 i 时刻，k 表示输出向量的第 k 个分量，j 表示 p 维参数空间的 $p+1$ 个样本点之一，$s_{k_1},\cdots,s_{k_p}$ 是对应的灵敏度 $(\partial y_k/\partial \theta_j)_i$. 对 $\boldsymbol{y}$ 的第一个分量 y_1 在时刻 i 有

$$\boldsymbol{y}_{1i} = X_\theta \boldsymbol{s}^0_{1i} \tag{7}$$

矩阵 X_θ 有 $p+1$ 行对应 $p+1$ 个样本点，消去第一行得

$$\Delta \boldsymbol{y}_{1i} = \Delta X_\theta \boldsymbol{s}_{1i} \tag{8}$$

式中，$\boldsymbol{y}_{1i} = [y^0_{1i}, y^1_{1i}, \cdots, y^p_{1i}]^T$，$\boldsymbol{s}^0_{1i} = [s_{10}, s_{11}, \cdots, s_{1p}]^T_i$，$\Delta \boldsymbol{y}_{1i} = [y^1_{1i} - y^0_{1i}, y^2_{1i} - y^0_{1i}, \cdots, y^p_{1i} - y^0_{1i}]^T$，$\boldsymbol{s}_{1i} = [s_{11}, s_{12}, \cdots, s_{1p}]^T_i$，$\Delta\theta^j_k = \theta^j_k - \theta^0_k$，

$$X_\theta = \begin{bmatrix} 1 & \theta^0_1 \cdots \theta^0_p \\ 1 & \theta^1_1 \cdots \theta^1_p \\ \vdots & \vdots \quad\quad \vdots \\ 1 & \theta^p_1 \cdots \theta^p_p \end{bmatrix}, \quad \Delta X_\theta = \begin{bmatrix} \Delta\theta^1_1 & \Delta\theta^1_2 & \cdots \Delta\theta^1_p \\ \Delta\theta^2_1 & \Delta\theta^2_2 & \cdots \Delta\theta^2_p \\ \vdots & \vdots & \vdots \\ \Delta\theta^p_1 & \Delta\theta^p_2 & \cdots \Delta\theta^p_p \end{bmatrix}.$$

由此，在时刻 i 输出向量 $\boldsymbol{y}$ 的第一个分量关于参数向量的灵敏度由下式给出：

$$\boldsymbol{s}_{1i} = [\Delta X_\theta]^{-1} \Delta \boldsymbol{y}_{1i} \tag{9}$$

由上式得到灵敏度初值之后，可采用如下递推算法计算校正估计过程中所需的灵敏度[2].

将(6)式略去角标 θ 和 i 并写成矩阵形式有

$$\boldsymbol{y}_k = X \boldsymbol{s}_k \tag{10}$$

则第 k 个分量的灵敏度向量 $\boldsymbol{s}_k$ 的最小二乘解是

$$\boldsymbol{s}_k = [X^T X]^{-1} X^T \boldsymbol{y}_k \tag{11}$$

对于第 $r+1$ 次迭代，定义 $P_{r+1} = [X^T_r X_r]^{-1}$，并略去角标 k，则信息矩阵和灵敏度的递推算式为

$$P_{r+1} = [P^{-1}_r + \boldsymbol{\varphi}^T_r \boldsymbol{\varphi}_r - (\boldsymbol{\varphi}^0_r)^T \boldsymbol{\varphi}^0_r]^{-1} \tag{12}$$

$$\boldsymbol{s}_{r+1} = \boldsymbol{s}_r - P_{r+1}\{(\boldsymbol{\varphi}^0_r)^T y^0_r - \boldsymbol{\varphi}^T_r y_r + [\boldsymbol{\varphi}^T_r \boldsymbol{\varphi}_r - (\boldsymbol{\varphi}^0_r)^T \boldsymbol{\varphi}^0_r] \boldsymbol{s}_r\} \tag{13}$$

式中，$\boldsymbol{\varphi}_r$ 是对应于一组新的 $\boldsymbol{\theta}$ 值而引入 X_r 的行向量，$\boldsymbol{\varphi}^0_r$ 是对应于一组产生最大指标函数的 $\boldsymbol{\theta}$ 值而从 X_r 中剔出的行向量，y^0_r 和 y_r 分别是在第 $r+1$ 次迭代相应地从 $\boldsymbol{y}_k$ 中剔出和引入 $\boldsymbol{y}_k$ 的分量.

3　递推计算灵敏度算法实现

为了避免方程(9)中的矩阵求逆运算，基于正交表选择 (7) 式中矩阵 X_θ 的元素 θ^j_i 如下[3]：

$$\theta^j_i = \theta^0_i + a_{ji}\Delta\theta_i \tag{14}$$

a_{ji} 表示正交表的元素，$\Delta\theta_i$ 是 θ_i 相对 θ_i^0 的摄动量．则灵敏度初值由下式给出[3]：

$$\boldsymbol{s}_{ki} = [s_{k_1}, s_{k_2}, \cdots, s_{kp}]_i^T \tag{15}$$

$$s_{kl} = s_{kl}^* / \Delta\theta_l, \quad l = 1, 2, \cdots, p \tag{16}$$

$$s_{kl}^* = \frac{1}{M}\sum_{j=1}^{M} a_{jl}(y_{ki}^j - y_{ki}^0), \quad l = 1, 2, \cdots, p \tag{17}$$

利用矩阵求逆引理可以给出方程(12)的非求逆递推算式．为此定义矩阵

$$Q_{r+1} = [P_r^{-1} + \boldsymbol{\varphi}_r^T \boldsymbol{\varphi}_r]^{-1} \tag{18}$$

则有

$$P_{r+1} = [Q_{r+1}^{-1} - (\boldsymbol{\varphi}_r^0)^T \boldsymbol{\varphi}_r^0]^{-1} \tag{19}$$

根据矩阵求逆引理可以推得

$$Q_{r+1} = P_r - P_r \boldsymbol{\varphi}_r^T (1 + \boldsymbol{\varphi}_r P_r \boldsymbol{\varphi}_r^T)^{-1} \boldsymbol{\varphi}_r P_r \tag{20}$$

$$P_{r+1} = Q_{r+1} + Q_{r+1}(\boldsymbol{\varphi}_r^0)^T [1 - \boldsymbol{\varphi}_r^0 Q_{r+1} (\boldsymbol{\varphi}_r^0)^T]^{-1} \boldsymbol{\varphi}_r^0 Q_{r+1} \tag{21}$$

由方程(13)、(20)和(21)构成的递推计算灵敏度公式避开了高阶矩阵求逆运算．

参 考 文 献

[1] Maine, R. E. and Iliff, K. W., Identification of Dynamic System, AGARD AG-300, Vol. 2, Jan. 1985.

[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] 崔平远，吴瑶华，黄文虎，李乃宏，递推计算灵敏度的初值正交计算法，自动化学报 **18**(1992)，(5)，619—622.

AN ALGORITHM OF RECURSIVE COMPUTING SENSITIVITIES FOR MAXIMUM LIKELIHOOD ESTIMATION

Cui Pingyuan Wu Yaohua Huang Wenhu Li Naihong

(Dept. of Aerospace Engineering and Mechanics, Harbin Institute of Technology, Harbin 150001)

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