非平稳随机系统两步最小二乘算法研究

戴华平　董嘉文　钱积新　孙优贤
(浙江大学工业控制技术研究所・杭州, 310027)

提要：本文针对输入端具有慢时变干扰的非平稳动态系统，通过理论推导，提出一种新型的两步最小二乘算法(TS-LS)，并说明 TS-LS 是增广矩阵方法的一个特例. 仿真结果表明了增广矩阵方法的可行性.

关键词：增广矩阵；两步最小二乘算法；慢时变干扰；相关性；系统辨识

1 引 言

许多实际应用中，诸如通信、信号处理、系统辨识等领域，往往在滤波器的主输入端或系统输入信号端存在一个不同程度相关的慢时变的低频干扰，此干扰信号通过一个动态系统输出后，可能成为快变化干扰影响输出结果，传统的最小二乘法着重于讨论高频噪声的影响，很少研究低频干扰. 在有色噪声干扰下，人们提出了一系列推广的最小二乘法来解决[1~3]，包括辅助变量法(IV)、广义最小二乘法(GLS)，增广最小二乘法(ELS)、相关分析最小二乘法(COR-LS)等. 若简单地把慢变化干扰归作非平稳噪声处理，必将导致滤波器权阵系数或系统模型系数的估计偏差，特别是干扰和输入信号的频带处于系统频带内，且两者存在相关性，高频噪声在系统频带外时. 因此，需要针对输入端慢时变干扰提出新的改进最小二乘法.

文[4]初步研究两步最小二乘法的原理和算法，本文对 TS-LS 的各种性质进行详细地分析，着重说增广矩阵的性质和应用，仿真结果表明此算法的普遍性和可行性.

2 两步最小二乘算法(TS-LS)的导出及实现

考虑如下线性差分方程描述的非平稳动态系统

$$y(k)=\sum_{i=1}^{n}-a_i y(k-i)+\sum_{i=1}^{m} b_i[u(k-i)+f(k-i)]+n(k). \tag{1}$$

式中 $y(k)$, $u(k)$ 分别为系统受扰输出和输入信号. $f(k)$ 为输入端慢时变干扰，$n(k)$ 为高斯白噪声. 研究的辨识系统如图 1 所示.

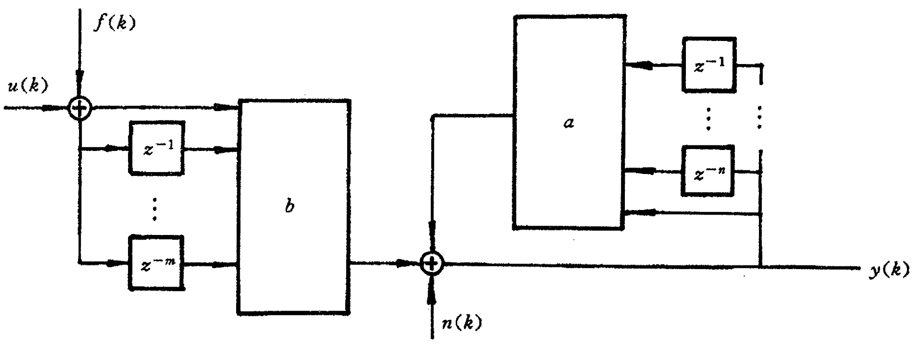

图 1　带输入端干扰辨识系统

* 国家重点自然科学基金资助项目.
本文于 1994 年 7 月 18 日收到. 1996 年 12 月 30 日收到修改稿.

若简单地把慢时变干扰归作噪声项处理，辨识精度将会很差．一般把慢时变干扰近似成时间的多项函数[3]，误差项作为高频噪声归入 $n(k)$ 中．

$$f(k)=g_0+\sum_{i=1}^{p}g_ik^i. \tag{2}$$

记 $g=[g_0,g_1,g_2,\cdots,g_p]^{\mathrm{T}}$，用一步最小二乘法估计 a，b 和 g 是一个非线性最小二乘问题，其计算复杂，计算最大．对于 $k=1,2,\cdots,L$，方程(1)写为

$$Y_L=H_L\theta+F_Lb+n_L. \tag{3}$$

其中

$$\theta=[a_1,\cdots,a_n,b_1,\cdots,b_m]^{\mathrm{T}},\quad b=[b_1,b_2,\cdots,b_m];$$

$$n_L=[n(1),n(2),\cdots,n(L)]^{\mathrm{T}},\quad Y_L=[y(1),y(2),\cdots,y(L)]^{\mathrm{T}};$$

$$H_L=\begin{bmatrix}-y(0)&\cdots&-y(1-n)&u(0)&\cdots&u(1-m)\\-y(1)&\cdots&-y(2-n)&u(1)&\cdots&u(2-m)\\\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\-y(L-1)&\cdots&-y(L-n)&u(L-1)&\cdots&u(L-m)\end{bmatrix}_{L\times(n+m)};$$

$$F_L=\begin{bmatrix}f(0)&f(-1)&\cdots&f(1-m)\\f(1)&f(0)&\cdots&f(2-m)\\\vdots&\vdots&\vdots&\vdots\\f(L-1)&f(L-2)&\cdots&f(L-m)\end{bmatrix}_{L\times m}.$$

由文[4]可得：
$$Y_L=[H_L\quad\lambda]\begin{bmatrix}\theta\\\varphi\end{bmatrix}+n_L.$$

其中
$$\lambda=\begin{bmatrix}1&0&0&\cdots&0\\1&1&1&\cdots&1\\1&2&2^2&\cdots&2^p\\\vdots&\vdots&\vdots&\vdots&\vdots\\1&(L-1)&(L-1)^2&\cdots&(L-1)^p\end{bmatrix};\quad \varphi=[\varphi_0,\varphi_1,\cdots,\varphi_p]^{\mathrm{T}}.$$

最小二乘解为

$$\begin{bmatrix}\hat\theta\\\hat\varphi\end{bmatrix}=[[H_L\quad\lambda]^{\mathrm{T}}[H_L\quad\lambda]]^{-1}[H_L\quad\lambda]^{\mathrm{T}}Y_L.$$

3 两步最小二乘算法(TS-LS)的特性分析

3.1 干扰与输入信号的相关性对 TS-LS 精度的影响

TS-LS 算法简单，但从随机性角度考虑问题时，干扰与输入信号的相关性影响辨识精度．

前述假设 $n(k)$ 是均值为 0，方差为 σ^2 的高斯白噪声，考虑最小无偏估计量 $\begin{bmatrix}\hat\theta\\\hat\varphi\end{bmatrix}$，即有效估计量，下式成立

$$\mathrm{cov}\begin{bmatrix}\hat\theta\\\hat\varphi\end{bmatrix}=\left[\begin{bmatrix}H_L^{\mathrm{T}}\\\lambda^{\mathrm{T}}\end{bmatrix}[H_L\quad\lambda]\right]^{-1}\sigma^2=M^{-1}.$$

其中 M 为 Fisher 信息阵[2]

$$M=\mathrm{E}\left[\left[\frac{\mathrm{d}\log(p(Y_L|\beta))}{\mathrm{d}\beta}\right]^{\mathrm{T}}\left[\frac{\mathrm{d}\log(p(Y_L|\beta))}{\mathrm{d}\beta}\right]\right],\quad \beta\triangleq\begin{bmatrix}\theta\\\varphi\end{bmatrix}.$$

考查 $J=-\log(\det M)=-\log\det\begin{bmatrix} H_L^{\mathrm{T}}H_L & H_L\lambda \\ \lambda^{\mathrm{T}}H_L & \lambda^{\mathrm{T}}\lambda \end{bmatrix}+\log\sigma^2$，可见在估计 θ 的同时，因估计 φ 引入矩阵 λ，有可能使 J 达到较大数值，影响辨识精度.

定理[1] 对于 $B_1=(\lambda^{\mathrm{T}}H_L)_1, B_2=(\lambda^{\mathrm{T}}H_L)_2$，若 $B_1^{\mathrm{T}}(H_L^{\mathrm{T}}H_L)^{-1}B_1\geqslant B_2^{\mathrm{T}}(H_L^{\mathrm{T}}H_L)^{-1}B_2$，则有 $J|_{B_1}\geqslant J|_{B_2}$.

定理说明 H_L 与 λ 相关性越强，辨识结果精度越差，这是指无干扰 $f(k)$ 而引入增广矩阵的情况下. 因此慢时变干扰 $f(k)$ 估计阶次并非越来越好，在考查的时间范围内，方程(2) 能近似干扰慢变化的趋势就够了.

3.2 λ 对可辨识性的影响

保证输入信号充分激励，使 $[H_L \quad \lambda]$ 满秩，不可与近似的输入干扰属于同一线性子空间内；如不，导致方程(4)病态解，使系统不可辨识.

3.3 提高辨识精度的途径

从上面分析可知，矩阵 H_L 与 Vandermonde 矩阵 λ 正交时，辨识精度最高. 因此输入信号在近似的输入干扰所撑开的空间(满足方程(2) 的近似的干扰所撑开的空间为 $\mathbb{R}^{p+1}$) 上投影越小，辨识精度越高.

一般情况下，取 $p<n$.

4 增广矩阵

对于矩阵 $[H_L \quad \lambda]$ 来说，我们形式地认为是 H_L 的增广矩阵. TS-LS 的成功之处在于，把干扰近似为时间的多项式函数(多项式函数所撑开的空间为 $\mathbb{R}^{p+1}$)，经过变量替换(φ 引入)，把一个非线性最小二乘辨识问题转化成一个线性最小二乘辨识问题，最终引入增广矩阵. 这个增广矩阵反映了干扰的近似方法，更深地反映了这种转化方法. 这里我们仅考虑输入干扰问题，无论采用什么样的线性近似方法(转化方法)，最终的结果必为一个增广矩阵形式. 因此，我们得出如下结论：

结论 1 对于方程(1)，取一定形式的 λ，可以消除与之对应形式的干扰输入对辨识结果的影响，λ 反应了干扰的动态特性.

取 λ 为 Vamdermonde 矩阵，则可以消除时间多项式干扰的影响.

仿真结果也验证结论 1.

5 仿真研究

辨识二阶系统

$$\begin{aligned} y(k) = {} & 1.723y(k-1)-0.741y(k-2)+0.9057[u(k-1)-f(k-1)] \\ & -0.8203[u(k-2)-f(k-2)], \end{aligned}$$

$$\lambda_{L\times(P+1)} \text{ 的第一行 } = \begin{cases} [1\ 0\ 0\ 0], & P+1=4; \\ [1\ 0\ 0], & P+1=3. \end{cases}$$

$$u(k)=\sin(k/50), \quad F(k)=0.01k.$$

得到如下辨识结果(见表 1).

6 结 论

本文针对输入端具有慢时变干扰的非平稳动态系统，提出了代替非线性最小二乘法的简单易行的两步最小二乘算法(TS-LS)，严格的理论推导和仿真研究表明，TS-LS 方法的可行性，增广矩阵方法的普遍性.

表1 λ对辨识精度的影响

辨识次数	增广矩阵 λ_{ij}		辨识结果			
	λ第二行开始	$P+1$	a_1	a_2	b_1	b_2
1	$2(i-1)(j-1)+1$	3	1.81825	−0.81835	0.90576	−0.90655
		4	1.81822	−0.81832	0.90574	−0.90651
2	$[2(i-1)]^{j-1}$	3	1.72292	−0.74093	0.90574	−0.82022
		4	1.72292	−0.74093	0.90574	−0.82023
3	$(i+1)^{j-1}$	3	1.86848	−0.85910	0.90572	−0.96287
		4	1.87061	−0.86083	0.90673	−0.95477
4	$(i-1)^{j-1}$	3	1.72292	−0.74093	0.90574	−0.82022
		4	1.72293	−0.74094	0.90574	−0.92024

参考文献

1 G.C.哥德温,R.L.潘恩.动态和系统辨识试验设计与数据分析.北京:科学出版社,1983
2 徐宁寿.系统辨识技术及应用.北京:机械工业出版社,1986
3 方崇智,萧德云.过程辨识.北京:清华大学出版社,1988
4 戴华平,董嘉文,钱积新.一种新型的非平稳随机系统参数辨识算法.电子学报,1996,24(1):68—72

Study on the Two-Stage Least Square Algorithm and Its Performance for Nonstationary Stochastic System

DAI Huaping, DONG Jiawen, QIAN Jixin and SUN Youxian
(Institute of Industrial Process Control, Zhejiang University, Hangzhou, 310027)

Abstract: In this paper, a novel two-stage least square (TS-LS) algorithm is proposed by theoretical derivation for nonstationary dynamic systems with input slow time-varying disturbance, and we explain that TS-LS is an example of extended-matrix method. Numerical simulation results show the efficiency of extended-matrix methodis.

Key words: two-stage least square; slow time-varying disturbance; correlativity; system identification

本文作者简介

戴华平 1969年生.1991年获浙江大学学士学位,1994年获中科院自动化研究所硕士学位,现是浙江大学工业控制技术研究所博士生.主要研究兴趣是辨识,图象处理.

董嘉文 1969年生.1991年和1993年分别获得浙江大学工程学系学士和硕士学位,现为浙江大学工业控制技术研究所博士生.主要研究兴趣是复杂工业过程建模与控制,神经网络及其在智能辨识中的应用.

钱积新 1962年毕业于清华大学电机工程学系,1963—1985年在兰州石油化工设计院工作,1985年调入浙江大学,现为工业控制技术研究所教授,博士生导师.主要从事复杂工业过程的建模、控制与优化的研究工作.

孙优贤 见本刊1997年第1期第41页.