系 统 科 学 与 数 学
J. Sys. Sci. & Math. Scis.
32(10) (2012, 10), 1180–1192

变量带误差系统的递推辨识：从线性到非线性系统*

陈翰馥 牟必强

(中国科学院数学与系统科学研究院系统科学研究所系统控制重点实验室， 北京 100190)

摘要 介绍变量带误差 (EIV) 系统的递推辨识方法. 在引言中扼要介绍了 EIV 系统辨识的现状后，分别对多变量线性 EIV 系统及 EIV Hammerstein 系统给出了递推辨识算法，并给出条件使这些估计以概率 1 收敛到真值. 最后提出了一些值得进一步研究的问题.

关键词 变量带误差系统、递推估计、随机逼近、强一致性.

MR(2000) 主题分类号 93B30, 93E12

1 引 言

对实际系统量测往往带有误差，不仅对系统输出的量测，甚至对系统的输入的量测，也可能伴有误差. 这类对系统输入输出量测都不能精确量测的系统，叫变量带误差 (EIV) 系统. 由于在实际中的重要性， EIV 系统的辨识一直受到控制界的关注，例如，文 [1–3] 等. EIV 系统的可辨识性问题在文 [4] 和 [5] 中分别对单变量及多变量的线性系统有所讨论.

可以想象，仍用精确量测下对线性系统辨识行之有效的最小二乘或推广的最小二乘法来辨识线性 EIV 系统，但这样通常会导致有偏的估计. 于是提出了多种方法，例如，偏差补偿方法、基于协方差矩阵的估计方法、频域估计法、极大似然法及预报误差估计法等，在文献 [6–7] 中有很好的综述. 但这些算法的强一致性不易保证，特别当得到新的量测数据时，要更新对系统参数的估计，就要用全部数据重新运算，很不方便.

本文关注对 EIV 系统的递推辨识，这类算法在得到新的数据时，很容易更新估计，计算方便. 但在以往的文献中，据作者所知，只有不多的工作 [8–11] 是用递推方法辨识 EIV 系统，并且主要针对线性系统.

本文结构如下. 在第 2 节中，我们介绍对线性多变量 EIV 系统的递推辨识方法，主要工具是扩张截断的随机逼近算法 (SAAWET)[12], 算法给出的估计以概率 1 收敛到系统的未知参数. 第 3 节介绍对 EIV Hammerstein 系统的递推辨识，所用方法不仅是 SAAWET, 由于还

* 国家自然科学基金项目 (61120106011, 61134013, 61273193) 和国家数学与交叉科学中心资助课题.
收稿日期： 2012-08-31.

要估计非线性函数，所以还要结合反卷积核函数，最终也得到递推估计算法，并且也以概率 1 收敛到真值．最后，在第 4 节给出一些总结性注记．

2　线性多变量 EIV 系统的递推辨识

设线性多变量系统的输入 $u_k^0 \in \mathbb{R}^m$ 与输出 $y_k^0 \in \mathbb{R}^n$ 的关系用下面的差分方程表示

$$A(z)y_k^0 = B(z)u_k^0, \tag{1}$$

其中

$$A(z) = I + A_1 z + \cdots + A_p z^p, \tag{2}$$

$$B(z) = B_1 z + B_2 z^2 + \cdots + B_q z^q, \tag{3}$$

z 表示后移算子： $zy_k = y_{k-1}$, $A(z)$ 和 $B(z)$ 分别是 $n \times n$ 和 $n \times m$ 阶系数未知但系统阶次 p, q 的已知矩阵多项式．真实的输入与输出信号不能精确量测，而只能量测到带误差的 u_k 和 y_k, 即

$$u_k = u_k^0 + \eta_k, \quad y_k = y_k^0 + \xi_k, \tag{4}$$

其中， η_k 和 ξ_k 分别是输入与输出端的量测噪声．对于这类系统，基于对系统输入及量测噪声不同假设条件下，我们分别地给出具有强一致收敛的辨识算法．确切地说，下述两类系统相比，第一类系统的输入较第二类系统的输入更一般，但量测噪声在第二类系统中更具一般形式．

2.1　第一类

这一部分内容主要基于文献 [9, 11].

对这类系统作如下假设

A1 系统的真实输入 u_k^0 是一个多元的 ARMA 过程

$$P(z)u_k^0 = Q(z)\varepsilon_k, \tag{5}$$

其中

$$P(z) = I + P_1 z + P_2 z^2 + \cdots + P_{n_p} z^{n_p}, \tag{6}$$

$$Q(z) = I + Q_1 z + Q_2 z^2 + \cdots + Q_{n_q} z^{n_q}. \tag{7}$$

系数阵 P_i, $i = 1, 2, \cdots, n_p$ 及 Q_j, $j = 1, 2, \cdots, n_q$ 是未知的．

A2 $A(z)$ 和 $P(z)$ 是稳定的矩阵多项式，即 $\det A(z)$ 和 $\det P(z)$ 的根都在单位闭圆外．

A3 $\Delta_k \triangleq [\xi_k^{\mathrm{T}}, \eta_k^{\mathrm{T}}, \varepsilon_k^{\mathrm{T}}]^{\mathrm{T}}$ 是一个独立同分布的随机变量序列，其中 $\mathrm{E}\Delta_k = 0$, $\mathrm{E}\|\Delta_k\|^{2+\delta} < \infty$, $\delta > 0$, 并且

$$\mathrm{E}\Delta_k {\Delta_k}^{\mathrm{T}} \triangleq R_\Delta = \begin{bmatrix} R_\xi & R_{\xi\eta} & R_{\xi\varepsilon} \\ R_{\eta\xi} & R_\eta & R_{\eta\varepsilon} \\ R_{\varepsilon\xi} & R_{\varepsilon\eta} & R_\varepsilon \end{bmatrix}. \tag{8}$$

利用新息表达方法可以将量测过程 $z_k \triangleq [y_k^{\mathrm{T}}, u_k^{\mathrm{T}}]^{\mathrm{T}}$ 表示为 $n+m$ 维的 ARMA 形式. 定义 $\lambda \triangleq \max\{p, q, n_p, n_q\}$, 并且对 $i > p$, $j > q$, $s > n_p$ $t > n_q$, 设 $A_i \triangleq 0$, $B_j \triangleq 0$, $P_s \triangleq 0$, $Q_t \triangleq 0$.

令

$$G(z) \triangleq \begin{bmatrix} A(z) & -B(z) \\ 0 & P(z) \end{bmatrix} = I + G_1 z + \cdots + G_\lambda z^\lambda \tag{9}$$

和

$$S(z) \triangleq \begin{bmatrix} A(z) & -B(z) & 0 \\ 0 & P(z) & Q(z) \end{bmatrix} = S_0 + S_1 z + \cdots + S_\lambda z^\lambda, \tag{10}$$

其中

$$G_i = \begin{bmatrix} A_i & -B_i \\ 0 & P_i \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}, \quad S_i = \begin{bmatrix} A_i & -B_i & 0 \\ 0 & P_i & Q_i \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+2m)}. \tag{11}$$

根据式 (1), (4), 和 (5), 量测过程 $\{z_k\}_{k\geq 0}$ 满足下面的等式

$$G(z)z_k = S(z)\varDelta_k, \tag{12}$$

注意到 $\varDelta_k$ 的维数是 $n+2m$, 而 z_k 的维数是 $n+m$, 式 (12) 不是一个标准的 ARMA 过程. 然而，应用文 [9, 13] 中使用的方法，和式 (12) 对应的 Yule-Walker 方程仍然可以类似地导出.

为此，我们还需要如下的一个假设.

A4 $G(z)$ 和 $H(z)R_\varDelta H(z^{-1})$ 没有公共的左公因子并且 $\det A_\lambda \neq 0$ 与 $\det P_\lambda \neq 0$.

通过 A1, A2, 和 A3, $\{z_k\}_{k\geq 0}$ 是渐近平稳的和遍历的. 于是，对任意一个整数 τ, 记

$$R_\tau \triangleq \lim_{k\to\infty} \mathrm{E} z_k z_{k-\tau}^{\mathrm{T}}. \tag{13}$$

在式 (12) 的两边，同时乘以 z_{k-s}^{T}, $s \geq \lambda + 1$, 并取期望. 令 $k \to \infty$, 我们能得到下面的等式

$$[G_1, G_2, \cdots, G_\lambda] \begin{bmatrix} R_{s-1} \\ R_{s-2} \\ \vdots \\ R_{s-\lambda} \end{bmatrix} = -R_s, \quad s \geq \lambda + 1. \tag{14}$$

定义

$$W \triangleq [R_{\lambda+1}, R_{\lambda+2}, \cdots, R_{\lambda+(n+m)\lambda}], \quad \theta \triangleq [G_1, G_2, \cdots, G_\lambda]^{\mathrm{T}}, \tag{15}$$

在式 (14) 中，取 $\lambda + 1 \leq s \leq (n+m)\lambda$, 我们能得到下面的矩阵方程

$$\varGamma^{\mathrm{T}}\theta = -W^{\mathrm{T}}, \tag{16}$$

其中

$$\Gamma \triangleq \begin{pmatrix} R_\lambda & R_{\lambda+1} & \cdots & R_{\lambda+\lambda(n+m)-1} \\ R_{\lambda-1} & R_\lambda & \cdots & R_{\lambda+\lambda(n+m)-2} \\ \vdots & \vdots & \ddots & \vdots \\ R_1 & R_2 & \cdots & R_{\lambda(n+m)} \end{pmatrix}. \tag{17}$$

引理 2.1[9,13] 假设 A1–A4 成立，那么矩阵 Γ 是行满秩的.

下面我们给出矩阵系数 $A_i,\ i=1,2,\cdots,p$ 和 $B_j,\ j=1,2,\cdots,q$ 的递推估计. 因为

$$G_i = \begin{bmatrix} A_i & -B_i \\ 0 & P_i \end{bmatrix},\quad 1\le i\le \lambda,$$

所以只需要估计 θ 就得到了对 $A_i,\ i=1,2,\cdots,p$ 和 $B_j,\ j=1,2,\cdots,q$ 的估计. 对于任意的整数 $\tau>0$, R_τ 能用下面的算法递推地估计

$$R_{\tau,k} \triangleq \frac{1}{k}\sum_{i=1}^{k-\tau} z_{i+\tau}z_i^{\mathrm{T}} = R_{\tau,k-1} - \frac{1}{k}\big(R_{\tau,k-1} - z_k z_{k-\tau}^{\mathrm{T}}\big), \tag{18}$$

其中， $R_{\tau,k}$ 表示在时刻 k 对 R_τ 的估计.

有了 R_τ 的估计 $R_{\tau,k}$ 之后，在矩阵 Γ 和 W 中，用 $R_{\tau,k}$ 代替 R_τ, Γ 和 W 的估计分别记为 Γ_k 和 W_k. 根据引理 2.1 可知，当 k 充分大的时候， Γ_k 是行满秩的. 因此根据式 (16), θ 的估计 θ_k 能很容易地给出

$$\theta_k = -(\Gamma_k\Gamma_k^{\mathrm{T}})^{-1}\Gamma_k W_k^{\mathrm{T}}, \tag{19}$$

但为了避免矩阵求逆，下面递推算法用来估计 θ_k[9]

$$\theta_{k+1} = \left[\theta_k - \frac{1}{k}\Gamma_k(\Gamma_k^{\mathrm{T}}\theta_k + W_k^{\mathrm{T}})\right]\cdot I_{\left[\|\theta_k - \frac{1}{k}\Gamma_k(\Gamma_k^{\mathrm{T}}\theta_k + W_k^{\mathrm{T}})\|\le M_{\delta_k}\right]}, \tag{20}$$

$$\delta_k = \sum_{j=1}^{k-1} I_{\left[\|\theta_j - \frac{1}{j}\Gamma_j(\Gamma_j^{\mathrm{T}}\theta_j + W_j^{\mathrm{T}})\| > M_{\delta_j}\right]}. \tag{21}$$

在文 [11] 中，根据式 (16), 提出了下面基于随机逼近的主成分分析算法. 对于式 (16), 可以通过简单的移项，得到下面的等式

$$M\Theta = 0, \tag{22}$$

其中

$$M \triangleq [\Gamma^{\mathrm{T}}\ W^{\mathrm{T}}] \in \mathbb{R}^{\lambda(n+m)^2\times(n+m)(\lambda+1)},\quad \Theta \triangleq [\theta^{\mathrm{T}}\ I]^{\mathrm{T}} \in \mathbb{R}^{(n+m)(\lambda+1)\times(n+m)}. \tag{23}$$

因为 W 与 Γ 是线性相关的，所以 M 的秩仍为 $(n+m)\lambda$, 并且 Θ 的各列是线性无关的，因此， Θ 的各列恰好是矩阵 $M^{\mathrm{T}}M$ 的零特征值对应的特征向量.

在 $M^{\mathrm{T}}M$ 中, 用 $R_{\tau,k}$ 代替 R_τ 后, $M^{\mathrm{T}}M$ 的估计记为 $M_k^{\mathrm{T}}M_k$. 假设对于对称矩阵 $M_k^{\mathrm{T}}M_k$, 应用基于随机逼近的主成分分析 [12,14] 得到的有次序的特征值估计分别为 $\bar{\lambda}_{1,k}, \bar{\lambda}_{2,k}, \cdots,$ $\bar{\lambda}_{(n+m)(\lambda+1),k}$, 并且 $\bar{\lambda}_{1,k} \leq \bar{\lambda}_{2,k} \leq \cdots \leq \bar{\lambda}_{(n+m)(\lambda+1),k}$, 其对应的特征向量分别为 $\bar{u}_{1,k}, \bar{u}_{2,k}, \cdots,$ $\bar{u}_{(n+m)(\lambda+1),k}$.

记 $\Theta_k \triangleq [\bar{u}_{1,k}, \bar{u}_{2,k}, \cdots, \bar{u}_{(n+m),k}] \triangleq \begin{bmatrix} \Theta_k^{(1)} \\ \Theta_k^{(2)} \end{bmatrix}$, 其中, $\Theta_k^{(1)} \in \mathbb{R}^{\lambda(n+m)\times(n+m)}$, $\Theta_k^{(2)} \in \mathbb{R}^{(n+m)\times(n+m)}$.

所以, θ 的估计 θ_k 可以由下式得到

$$\theta_k \triangleq \Theta_k^{(1)}\big(\Theta_k^{(2)}\big)^{-1}. \tag{24}$$

2.2 第二类

这一节的内容主要基于文献 [15], 和 "第一类" 相比, 这里量测噪声更一般, 但输入仅限于独立序列.

我们先给出对系统及其量测噪声的假设.

B1 $A(z)$ 和 $B(z)$ 没有左公因子, $[A_p, B_q]$ 是行满秩的, 并且 $A(z)$ 是稳定的, 即, $\det A(z) \neq 0, \forall |z| \leq 1$.

B2 量测噪声 η_k 和 ξ_k 是多元的回归滑动平均过程 ARMA 过程

$$P(z)\eta_k = Q(z)\omega_k, \tag{25}$$

$$C(z)\xi_k = D(z)\zeta_k, \tag{26}$$

其中

$$P(z) = I + P_1 z + P_2 z^2 + \cdots + P_{n_p} z^{n_p}, \tag{27}$$

$$Q(z) = I + Q_1 z + Q_2 z^2 + \cdots + Q_{n_q} z^{n_q}, \tag{28}$$

$$C(z) = I + C_1 z + C_2 z^2 + \cdots + C_{n_c} z^{n_c}, \tag{29}$$

$$D(z) = I + D_1 z + D_2 z^2 + \cdots + D_{n_d} z^{n_d}, \tag{30}$$

这里 $P(z)$ 和 $Q(z)$ 没有左公因子, 同样 $C(z)$ 和 $D(z)$ 也没有左公因子, 并且 $P(z)$ 和 $C(z)$ 都稳定. 驱动噪声 ω_k 和 ζ_k 存在密度函数, 并且他们是彼此独立的独立同分布的零均值随机向量. 此外, 对某个 $\varDelta > 2$, $\mathrm{E}(\|\omega_k\|^{\varDelta}) < \infty$ 和 $\mathrm{E}(\|\zeta_k\|^{\varDelta}) < \infty$.

B3 输入 $\{u_k^0, k \geq 1\}$ 是独立同分布, u_k^0 是有密度的零均值随机向量, 并且独立于 ω_k 和 ζ_k. 此外, $\mathrm{E}(u_k^0 u_k^{0\mathrm{T}}) = I$, 对某个 $\varDelta > 2$, $\mathrm{E}(\|u_k^0\|^{2\varDelta}) < \infty$.

由于 $A(z)$ 稳定, 则

$$H(z) \triangleq A^{-1}(z)B(z) = \sum_{i=1}^{\infty} H_i z^i, \tag{31}$$

这里 $H_1 = B_1$, $\|H_i\| = O(\mathrm{e}^{-ri})$, $r > 0$, $i > 1$.

假设 $u_k^0 = 0,\ \forall k < 0$, 则

$$y_k^0 = \sum_{i=1}^{k} H_i u_{k-i}^0. \tag{32}$$

根据假设 B2 和 B3, u_k^0, ω_k 和 ζ_k 是相互独立的，则

$$\mathrm{E}(y_k u_{k-i}^{\mathrm{T}}) = \mathrm{E}(y_k^0 + \varepsilon_k)(u_{k-i}^0 + \eta_{k-i})^{\mathrm{T}} = \mathrm{E}y_k^0 u_{k-i}^{0\mathrm{T}} = H_i. \tag{33}$$

这提示我们可以用基于带噪声的输入 u_k 和输出 y_k 来估计线性系统的脉冲响应 H_i, $i = 1, 2, \cdots$.

基于式 (33), 利用 SAAWET[12], 对 H_i, $i \geq 1$ 在时刻 k 的估计 $H_{i,k}$ 由下式递推地给出

$$\begin{aligned} H_{i,k+1} = & \left[H_{i,k} - \frac{1}{k}(H_{i,k} - y_{k+1} u_{k+1-i}^{\mathrm{T}})\right] \\ & \cdot I_{\left[\|H_{i,k} - \frac{1}{k}(H_{i,k} - y_{k+1} u_{k+1-i}^{\mathrm{T}})\| \leq M_{\delta_{i,k}}\right]}, \end{aligned} \tag{34}$$

$$\delta_{i,k} = \sum_{j=1}^{k-1} I_{\left[\|H_{i,j} - \frac{1}{j}(H_{i,j} - y_{j+1} u_{j+1-i}^{\mathrm{T}})\| > M_{\delta_{i,j}}\right]}, \tag{35}$$

这里 $\{M_k\}$ 是任意选择的递增发散到正无穷的正实数序列， $H_{i,0}$ 是任选的初值， I_A 表示集合 A 的示性函数.

一旦得到了 H_i 的估计值 $H_{i,k}$, 矩阵系数 $\{A_1, A_2, \cdots, A_p, B_1, B_2, \cdots, B_q\}$ 能够通过他们之间具有的卷积关系进行估计.

根据式 (31), 有下面关于 z 的矩阵多项式的等式

$$B_1 z + B_2 z + \cdots + B_q z^q = (I + A_1 z + \cdots + A_p z^p)(H_1 z + H_2 z + \cdots + H_i z^i + \cdots). \tag{36}$$

比较上式两边 z 的相同幂次，能得到

$$B_i = \sum_{j=0}^{i \wedge p} A_j H_{i-j}, \quad \forall 1 \leq i \leq q \tag{37}$$

和

$$H_i = -\sum_{j=1}^{i \wedge p} A_j H_{i-j}, \quad \forall i \geq q+1, \tag{38}$$

这里 $a \wedge b$ 表示 $\min(a, b)$.

定义

$$\varGamma \triangleq \begin{pmatrix} H_q & H_{q+1} & \cdots & H_{q+np-1} \\ H_{q-1} & H_q & \cdots & H_{q+np-2} \\ \vdots & \vdots & \ddots & \vdots \\ H_{q-p+1} & H_{q-p+2} & \cdots & H_{q+(n-1)p} \end{pmatrix}, \tag{39}$$

其中对于 $i \le 0$, $H_i \triangleq 0$.

在式 (38) 中选取 $q+1 \le i \le q+np$, 通过式 (38) 和 (39) 能得到下面的代数方程

$$[A_1, A_2, \cdots, A_p]\Gamma = -[H_{q+1}, H_{q+2}, \cdots, H_{q+np}]. \tag{40}$$

当 B1 成立时, 式 (39) 中的 Γ 是行满秩的, 并且在条件 B1–B3 下, 当 $k \to \infty$ 时, $H_{i,k}$ 几乎处处收敛到 H_i. 所以当 k 充分大时,

$$\Gamma_k \triangleq \begin{pmatrix} H_{q,k} & H_{q+1,k} & \cdots & H_{q+np-1,k} \\ H_{q-1,k} & H_{q,k} & \cdots & H_{q+np-2,k} \\ \vdots & \vdots & \ddots & \vdots \\ H_{q-p+1,k} & H_{q-p+2,k} & \cdots & H_{q+(n-1)p,k} \end{pmatrix} \tag{41}$$

也行满秩, 在 Γ_k 中, 对 $i \le 0$, $H_{i,k} = 0$.

$\{A_1, A_2, \cdots, A_p, B_1, B_2, \cdots, B_q\}$ 的估计很自然地可由下面的式子给出

$$[A_{1,k}, A_{2,k}, \cdots, A_{p,k}] = -[H_{q+1,k}, H_{q+2,k}, \cdots, H_{q+np,k}]\Gamma_k^{\mathrm{T}}(\Gamma_k \Gamma_k^{\mathrm{T}})^{-1}, \tag{42}$$

$$B_{i,k} = \sum_{j=0}^{i \wedge p} A_{j,k} H_{i-j,k}, \ \forall\, 1 \le i \le q. \tag{43}$$

3 EIV Hammerstein 系统的递推辨识

这部分是新结果, 详细证明将另文发表.

我们考察单变量的 EIV Hammerstein 系统

$$v_k^0 = f(u_k^0), \tag{44}$$

$$C(z)y_k^0 = D(z)v_k^0 + \xi_k, \tag{45}$$

这里

$$C(z) = 1 + c_1 z + \cdots + c_p z^p, \tag{46}$$

$$D(z) = z + d_2 z^2 + \cdots + d_q z^q, \tag{47}$$

ξ_k 是系统的内部噪声. 假设线性子系统的真实的阶次 p, q 已知. v_k^0 是不可量测的内部信号, u_k^0 是 y_k^0 分别是系统的输入与输出, 但不能得到其精确值. 我们能得到的仅仅是带有噪声 η_k 和 ε_k 干扰的量测输入 u_k 与输出 y_k

$$u_k = u_k^0 + \eta_k, \quad y_k = y_k^0 + \varepsilon_k. \tag{48}$$

3.1 线性子系统的辨识算法

我们首先给出对线性子系统的条件.

C1 输入 $\{u_k^0\}_{k\geq 0}$ 是零均值、独立同分布的，并具有密度函数的随机变量序列. 记 u_k^0 的密度函数为 $p(\cdot)$, 它的支撑是有界的，即 $|u_k^0| \leq U$, 这里 $U > 0$ 是一个常数.

C2 $C(z)$ 和 $z^{-1}D(z)$ 互素，并且都是稳定的，即： $C(z) \neq 0$ 和 $z^{-1}D(z) \neq 0, \forall |z| \leq 1$.

C3 量测噪声 η_k 和 ε_k 都是自回归滑动平均过程 (ARMA)

$$P(z)\eta_k = Q(z)\zeta_k, \tag{49}$$

$$E(z)\varepsilon_k = F(z)\varsigma_k, \tag{50}$$

这里

$$P(z) = 1 + p_1 z + p_2 z^2 + \cdots + p_{n_p} z^{n_p}, \tag{51}$$

$$Q(z) = 1 + q_1 z + q_2 z^2 + \cdots + q_{n_q} z^{n_q}, \tag{52}$$

$$E(z) = 1 + e_1 z + e_2 z^2 + \cdots + e_{n_e} z^{n_e}, \tag{53}$$

$$F(z) = 1 + f_1 z + f_2 z^2 + \cdots + f_{n_f} z^{n_f}, \tag{54}$$

$P(z)$ 和 $Q(z)$ 没有公共子， $E(z)$ 和 $F(z)$ 也没有公因子， $P(z)$ 和 $E(z)$ 都稳定. 驱动噪声 ζ_k 与 ς_k 和内部噪声 ξ_k 都是零均值、独立同分布的随机变量序列并具有密度函数. 此外， ζ_k, ς_k 和 ξ_k 之间相互独立，并对某个 $\varDelta > 2$, $\mathrm{E}(\|\xi_k\|^{\varDelta}) < \infty$, $\mathrm{E}(\|\zeta_k\|^{\varDelta+1}) < \infty$ 且 $\mathrm{E}(\|\varsigma_k\|^{\varDelta+1}) < \infty$, 并且他们独立于 u_k^0.

C4 函数 $f(\cdot)$ 可测，局部有界、并且在估计点是连续的. 此外，常数 $\rho \triangleq \mathrm{E}(f(u_k^0)u_k^0) \neq 0$.

因为 $C(z)$ 稳定，根据 C2, 我们得到

$$H(z) \triangleq \frac{D(z)}{C(z)} = \sum_{i=1}^{\infty} h_i z^i, \tag{55}$$

这里 $|h_i| = O(\mathrm{e}^{-ri}), r > 0,\ i \geq 2$, 并且 $h_1 = 1$.

定义 $\bar{\xi}_k = C(z)^{-1}\xi_k$, 假设 $u_k^0 = 0,\ \forall k < 0$, 很容易得到

$$y_k^0 = \sum_{i=1}^{k} h_i v_{k-i}^0 + C(z)^{-1}\xi_k = \sum_{i=1}^{k} h_i v_{k-i}^0 + \bar{\xi}_k. \tag{56}$$

注意到 $\xi_k, \zeta_k, \varsigma_k$ 与 u_k^0 之间是相互独立的，我们有下面的等式

$$\begin{aligned}
\mathrm{E}(y_k u_{k-i}) = \mathrm{E}(y_k^0 + \varepsilon_k)(u_{k-i}^0 + \eta_{k-i}) &= \mathrm{E}y_k^0 u_{k-i}^0 \\
&= \mathrm{E}\left(\sum_{j=1}^{k} h_j v_{k-j}^0 + \bar{\xi}_k\right) u_{k-i}^0 = \sum_{j=1}^{k} h_j \mathrm{E}(v_{k-j}^0) u_{k-i}^0 \\
&= \mathrm{E}(v_{k-i}^0 u_{k-i}^0) h_i = \rho h_i, \quad \forall\ i \geq 1,
\end{aligned} \tag{57}$$

其中 $\rho = \mathrm{E}(f(u_k^0)u_k^0)$.

在估计 h_i 之前，我们首先基于式 (57) 应用 SAAWET 递推地估计 ρh_i

$$\theta_{k+1}^{(i)} = \left[\theta_k^{(i)} - \frac{1}{k}\left(\theta_k^{(i)} - y_{k+1}u_{k+1-i}\right)\right] \cdot I_{\left[\left|\theta_k^{(i)} - \frac{1}{k}\left(\theta_k^{(i)} - y_{k+1}u_{k+1-i}\right)\right| \le M_{\delta_k^{(i)}}\right]}, \tag{58}$$

$$\delta_k^{(i)} = \sum_{j=1}^{k-1} I_{\left[\left|\theta_j^{(i)} - \frac{1}{j}\left(\theta_j^{(i)} - y_{j+1}u_{j+1-i}\right)\right| > M_{\delta_j^{(i)}}\right]}, \quad i \ge 1, \tag{59}$$

这里 $\{M_k\}$ 是任选的递增发散到正无穷的正实数序列， $\theta_0^{(i)}$ 是任选的初值.

因为 $h_1 = 1$, 所以 h_i, $i \ge 2$ 能用如下的方式估计

$$h_{i,k+1} \triangleq \begin{cases} \dfrac{\theta_{k+1}^{(i)}}{\theta_{k+1}^{(1)}}, & \text{若 } \theta_{k+1}^{(1)} \ne 0, \\ 0, & \text{若 } \theta_{k+1}^{(1)} \ne 0. \end{cases} \tag{60}$$

这里 $h_{i,k+1}$ 是对 h_i 在时刻 $k+1$ 的估计.

有了 h_i 的估计 $h_{i,k}$, 线性子系统的参数 $\{c_1, c_2, \cdots, c_p, d_2, d_3, \cdots, d_q\}$ 可以很容易地利用其与脉冲响应 $h_i, i \ge 1$ 的卷积关系得到.

根据式 (60), 有下面的关于 z 的多项式等式

$$z + d_2 z^2 \cdots + d_q z^q = (1 + c_1 z + \cdots + c_p z^p)(z + \cdots + h_i z^i + \cdots),$$

比较上式两边 z 的相同幂次，有

$$d_i = \sum_{j=0}^{i \wedge p} c_j h_{i-j}, \quad \forall\ 2 \le i \le q \tag{61}$$

和

$$h_i = -\sum_{j=1}^{i \wedge p} c_j h_{i-j}, \quad \forall\ i \ge q+1. \tag{62}$$

定义

$$\varGamma \triangleq \begin{pmatrix} h_q & h_{q-1} & \cdots & h_{q-p+1} \\ h_{q+1} & h_q & \cdots & h_{q-p+2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{q+p-1} & h_{q+p-2} & \cdots & h_q \end{pmatrix}, \tag{63}$$

其中，对于 $i \le 0$, 令 $h_i \triangleq 0$.

在式 (62) 中，取 $q+1 \le i \le q+p$, 能得到下面的线性方程

$$\varGamma[c_1, c_2, \cdots, c_p]^{\mathrm{T}} = -[h_{q+1}, h_{q+2}, \cdots, h_{q+p}]^{\mathrm{T}}. \tag{64}$$

可以证明，在条件 C2 下，由式 (63) 定义的矩阵 Γ 是非奇异的，并且在条件 C1–C4 下，当 k 趋向于无穷时， $h_{i,k}$ 几乎处处收敛于 h_i. 因此，当 k 充分大时，

$$\Gamma_k \triangleq \begin{pmatrix} h_{q,k} & h_{q-1,k} & \cdots & h_{q-p+1,k} \\ h_{q+1,k} & h_{q,k} & \cdots & h_{q-p+2,k} \\ \vdots & \vdots & \ddots & \vdots \\ h_{q+p-1,k} & h_{q+p-2,k} & \cdots & h_{q,k} \end{pmatrix} \tag{65}$$

也是非奇异的，这里当 $i \leq 0$ 时，设 $h_{i,k} = 0$.

根据上面的推导， $\{c_1, c_2, \cdots, c_p, d_2, d_3, \cdots, d_q\}$ 的估计可以由下式给出

$$[c_{1,k}, c_{2,k}, \cdots, c_{p,k}]^{\mathrm{T}} \triangleq -\Gamma_k^{-1}[h_{q+1,k}, l_{q+2,k}, \cdots, l_{q+p,k}]^{\mathrm{T}}, \tag{66}$$

$$d_{i,k} \triangleq h_{i,k} + \sum_{j=1}^{i\wedge p} c_{j,k} h_{i-j,k}, \quad i = 2, 3, \cdots, q. \tag{67}$$

3.2　$f(\cdot)$ 的递推估计算法

估计非线性函数 $f(\cdot)$, 最重要的是要得到 $f(\cdot)$ 的输入及相对应的输出序列，对本文考虑的系统来说，就是需要 u_k^0 和 $f(u_k^0)$. 在有了线性部分的估计之后， $f(u_k^0)$ 可以利用线性系统的状态空间表示法得到．然而由于不知道 u_k^0 的确切值，所以通常的核函数估计不再有效．为了处理这个问题，我们利用反卷积核函数 [16] 来估计 $f(\cdot)$.

为了估计 $f(\cdot)$, 还需要下面两个假设.

C5 u_k^0 的方差 ϑ^2 是已知的.

C6 驱动噪声 $\{\zeta_k\}_{k\geq 0}$ 是零均值的独立同分布的高斯随机变量序列，但并不需要知道它的方差.

首先，我们利用状态空间给出 $f(u_k^0)$ 的一个合理的估计

定义

$$D_{k+1} \triangleq \begin{pmatrix} -d_{2,k+1} & 1 & & \\ \vdots & & \ddots & \\ \vdots & & & 1 \\ -d_{s+1,k+1} & 0 & \cdots & 0 \end{pmatrix}, \quad C_{k+1} \triangleq \begin{pmatrix} 1 \\ c_{1,k+1} \\ \vdots \\ c_{s-1,k+1} \end{pmatrix}, \quad H \triangleq \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

这里 $s \triangleq \max(p+1, q-1)$, D_{k+1}, C_{k+1} 和 H 分别是 $s \times s$, $s \times 1$ 和 $s \times 1$ 的矩阵.

在时刻 k, 对任一初值 $\widehat{x}_0$, $f(u_k^0)$ 可由下式进行估计

$$\widehat{x}_{k+1} = D_{k+1}\widehat{x}_k + C_{k+1}y_{k+1},\ \widehat{\varphi}_k = H^{\mathrm{T}}\widehat{x}_{k+1}, \tag{68}$$

这里， $\widehat{\varphi}_k$ 代表对 $f(u_k^0)$ 的估计.

辛克核函数 (Sinc)[17−18]

$$K(x) = \frac{\sin(x)}{\pi x} \tag{69}$$

的傅里叶变换为

$$\Phi_K(t) \triangleq \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tx} K(x)\mathrm{d}x = I_{[|t|\leq 1]}.$$

在条件 C6 下， η_k 也是一个零均值的高斯随机变量，其特征函数为

$$\Phi_\eta(t) \triangleq \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}tx} \frac{1}{2\pi\sigma} \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \mathrm{d}t = \mathrm{e}^{-\frac{\sigma^2 t^2}{2}},$$

这里 σ^2 是 η_k 的方差.

$K(x)$ 的反卷积核函数为

$$K_k(x) \triangleq \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}tx} \frac{\Phi_K(t)}{\Phi_\eta\left(\frac{t}{b_k}\right)} \mathrm{d}t = \frac{1}{2\pi} \int_{-1}^{1} \mathrm{e}^{-\mathrm{i}tx} \mathrm{e}^{\frac{\sigma^2 t^2}{2b_k^2}} \mathrm{d}t \tag{70}$$

与

$$\begin{aligned} w_k(x) \triangleq \frac{1}{b_k} K_k\left(\frac{u_k - x}{b_k}\right) &= \frac{1}{2\pi b_k} \int_{\mathbb{R}} \mathrm{e}^{\frac{-\mathrm{i}t(u_k - x)}{b_k}} \frac{\Phi_K(t)}{\Phi_\eta\left(\frac{t}{b_k}\right)} \mathrm{d}t \\ &= \frac{1}{\pi b_k} \int_0^1 \mathrm{e}^{\frac{-\mathrm{i}t(u_k - x)}{b_k}} \mathrm{e}^{\frac{\sigma^2 t^2}{2b_k^2}} \mathrm{d}t \\ &= \frac{1}{\pi} \int_0^{\frac{1}{b_k}} \cos[(u_k - x)t] \mathrm{e}^{\frac{\sigma^2 t^2}{2}} \mathrm{d}t, \end{aligned} \tag{71}$$

这里 $b_k = \left(\frac{1}{3\sigma^2} \log k\right)^{-\frac{1}{2}}$ 表示窗宽.

虽然 σ^2 是不知道的，但基于量测输入 u_k 可以估计

$$\sigma_{k+1}^2 = \frac{1}{k+1} \left(\sum_{i=1}^{k+1} u_i^2\right) - \vartheta^2,$$

上式可以容易地写成如下的递推形式

$$\sigma_{k+1}^2 = \sigma_k^2 - \frac{1}{k+1}(\sigma_k^2 + \vartheta^2 - u_{k+1}^2). \tag{72}$$

因此，当有了 σ^2 的估计 σ_k^2 之后，在时刻 k, $w_k(x)$ 可以由下式近似

$$\widehat{w}_k(x) \triangleq \frac{1}{\pi} \int_0^{\frac{1}{\widehat{b}_k}} \cos[(u_k - x)t] \mathrm{e}^{\frac{\sigma_k^2 t^2}{2}} \mathrm{d}t, \tag{73}$$

这里 $\widehat{b}_k = \left(\frac{1}{3\sigma_k^2} \log k\right)^{-\frac{1}{2}}$. 在实际计算中，式 (73) 需要借助于数值积分.

有了 $\widehat{w}_k(x)$ 的定义之后，结合 SAAWET 可以递推地估计 $f(x)$

$$\mu_{k+1}(x) = \left[\mu_k(x) - \frac{1}{k}\widehat{w}_k(x)(\mu_k(x) - \widehat{\varphi}_k)\right] I_{\left[|\mu_k(x) - \frac{1}{k}\widehat{w}_k(x)(\mu_k(x) - \widehat{\varphi}_k)| \leq m_{\Delta_k(x)}\right]}, \tag{74}$$

$$\Delta_k(x) = \sum_{j=1}^{k-1} I_{[|\mu_j(x) - \frac{1}{j}\widehat{w}_j(x)(\mu_j(x) - \widehat{\varphi}_j)| > m_{\Delta_j(x)}]}, \tag{75}$$

这里， $m_k = \frac{1}{k^b}, 0 < b < \frac{1}{6}$.

4　结　论

在文中，我们给出了多变量线性 EIV 系统及 EIV Hammerstein 系统的递推辨识算法，并给出了使它们收敛到真值的条件．在 EIV 系统辨识这个领域有许多问题值得进一步研究，不仅要研究多种类型的 EIV 非线性系统的辨识，例如 EIV Wiener 系统的辨识， EIV NARX 系统的辨识，以及更一般类型的 EIV 系统的辨识，即使对于线性的 EIV 系统，当系统输入及量测噪声均为 ARMA 过程时，如何递推辨识，也尚未解决．

参　考　文　献

[1] Anderson B D O. Identification of scalar errors-in-variables models with dynamics. *Automatica*, 1985, **21**: 709–716.

[2] Scherrer W, Deistler M. A structure theory for linear dynamic errors-in-variables models. *SIAM Journal of Control and Optimization*, 1998, **36**: 2148–2175.

[3] Söderström T, Mahata K, Soverini U. Identification of dynamic errors-in-variables model: Approaches based on two-dimensional ARMA modeling of the data. *Automatica*, 2003, **39**: 929–935.

[4] Agüero J C, Goodwin G C. Identifiability of errors in variables dynamic systems. *Automatica*, 2008, **44**: 371–382.

[5] Rojas C R. Identifiability of multivariable dynamic errors-in-variables systems. 2011 9th IEEE International Conference on Control and Automation (ICCA), 2011.

[6] Söderström T, Soverini U, Mahata K. Perspectives on errors-in-variables estimation for dynamic systems. *Signal Processing*, 2002, **82**: 1139–1154.

[7] Söderström T. Errors-in-variables methods in system identification. *Automatica*, 2007, **43**: 939–958.

[8] Chen H F, Yang J M. Strongly consistent coefficient estimate for errors-in-variables models. *Automatica*, 2005, **43**: 1025–1033.

[9] Chen H F. Recursive identification for multivariate errors-in-variables systems. *Automatica*, 2007, **43**: 1234–1242.

[10] Song Q J, Chen H F. Identification of errors-in-variables systems with ARMA observation noises. *Systems & Control Letters*, 2008, **57**: 420–424.

[11] Zhao W X, Chen H F. Stochastic approximation based PCA and its application to identification of EIV systems. Proceeding of the 10th World Congress on Intelligent Control and Automation, 2012.

[12] Chen H F. Stochastic Approximation and Its Applications. Dordrecht, The Netherlands: Kluwer, 2002.

[13] Stoica P. Generalized Yule-Walker equations and testing the orders of multivariate time series. *Int. J. Control*, 1983, **37**: 1159–1166.

[14] Chen H F, Fang H T, Zhang L L. Recursive estimation for ordered eigenvectors of symmetric matrix with observation noise. *Journal of Mathematical Analysis and Applications*, 2011, **382**: 822–842.

[15] Mu B Q, Chen H F. Recursive identification for multivariate EIV linear systems. Proceedings of the 31st Chinese Control Conference, 2012.

[16] Fan J, Truong Y K. Nonparametric regression with errors in variables. *The Annals of Statistics*, 1993, **21**: 1900–1925.

[17] Stefanski L, Carroll R J. Deconvoluting kernel density estimators. *Statistics*, 1990, **21**: 169–184.

[18] Davis K B. Mean square error properties of density estimates. *The Annals of Statistics*, 1975, **3**: 1025–1030.

RECURSIVE IDENTIFICATION FOR ERRORS-IN-VARIABLES SYSTEMS: FROM LINEAR TO NONLINEAR SYSTEMS

CHEN Han-Fu MU Biqiang

(*Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing* 100190)

Abstract The recursive identification methods for errors-in-variables (EIV) systems are described in the paper. After briefly introducing the current status of identification for EIV systems, the recursive identification algorithms for the multivariate linear EIV system as well as for the EIV Hammerstein system are presented. The conditions guaranteeing the strong consistency of the estimates given by the algorithms are demonstrated. Finally, some remarks on problems for future research are pointed out.

Key Words Errors-in-variables, recursive estimates, stochastic approximation, strong consistency.