

两种时变参数估计的自适应递推算法

王维平 潘德惠

(东北工学院管理工程系, 沈阳)

摘要 本文就参数呈快时变(包括阶跃变化)时系统辨识问题提出了两种在线自适应辨识算法. 它们是通过参数时变检验对最小二乘递推算法进行适时校正得来的算法, 并通过仿真与其它方法比较, 验证了它们的有效性.

关键词: 系统辨识; 时变参数; 检验方法

1. 引言

线性时变系统的辨识方法常常是由非时变系统辨识的递推算法改进而来, 如遗忘因子法、定长窗口法(限定记忆法)和振荡存贮方法等^{[1][2]}. 这些方法对时变参数的估计是有效的, 但当参数变化复杂(具有快时变或阶跃变化)时, 这些方法不能给出满意的结果.

为了辨识阶跃变化的参数, Andersson^[3]提出了一种 AFMM 算法, 其对参数的阶跃变化具有很快的跟踪速度; 但对呈连续时变的参数估计不理想, 且计算复杂, 用多个最小二乘递推算法. 就此, Chen^[4]提出了一种根据估计量的变差来构成对参数时变检验的跟踪算法 RT. 这种方法适应性强, 但需要对估计量逐个进行检验, 而其校正量的大小不能适时调整, 这会引起在误判的情况下参数估计发生较大波动. 实际中误判是不可避免的.

为了进一步得到实用的辨识算法, 本文从数据统计分析的角度提出一种更为简单的参数时变检验方法及两种对递推最小二乘法进行校正的算法, 得到了自适应递推算法. 它们能适用于参数变化复杂的辨识问题. 大量的仿真验证了它的有效性.

2. 算法的原理

考虑下列动态系统

$$\begin{aligned} y(k) = & -a_1 y(k-1) - a_2 y(k-2) \cdots - a_n y(k-n) \\ & + b_1 u(k-1) + \cdots + b_r u(k-r) + \varepsilon_k \end{aligned} \quad (1)$$

设系统的结构已知; $\{\varepsilon_k\}$ 为零均值正态白噪声, 方差有界; 输入 $u(k)$ 和输出 $y(k)$ 可由观测得到; 待辨识参数 $\{a_i\}$ 和 $\{b_i\}$ 可为时变的.

$$\text{令 } \theta(k) = [a_1(k), a_2(k), \cdots, a_n(k), b_1(k), b_2(k), \cdots, b_r(k)]^T, \quad (2)$$

$$X(k) = [-y(k-1), -y(k-2), \cdots, -y(k-n), u(k-1), u(k-2), \cdots, u(k-r)]^T, \quad (3)$$

则(1)式可简写为

$$y(k) = X^T(k)\theta(k) + \varepsilon_k. \quad (4)$$

用递推方法来估计 $\hat{\theta}(k)$, 有下列的一般形式

$$\hat{\theta}(k+1) = \hat{\theta}(k) + g(k+1)f(y(k+1), \hat{\theta}(k)), \quad (5)$$

其中 $\hat{\theta}(k)$ 表示对 $\theta(k)$ 的估计, $f(y, \hat{\theta})$ 表示新息; $g(k)$ 是对新息的增益. 递推算法的关键就是如何根据已知的信息来确定 $g(k)$.

对应用最广的最小二乘法, 如取指标为

$$\left. \begin{aligned} J_N &= \sum_{k=1}^N \beta(k)[y(k) - X^T(k)\theta(k)]^2, \\ \beta(k) &= \pi \prod_{j=k-1}^{N-1} \lambda(j), \quad 0 \leq j \leq N-1, \quad \beta(N) = 1, \\ 0 &< \lambda(k) \leq 1. \end{aligned} \right\} \quad (6)$$

则可得下列递推最小二乘公式 (RLS)

$$\hat{\theta}(k+1) = \hat{\theta}(k) + G(k+1)[y(k+1) - X^T(k+1)\hat{\theta}(k)], \quad (7a)$$

$$G(k+1) = P(k)X(k+1) / [\lambda(k+1) + X^T(k+1)P(k)X(k+1)], \quad (7b)$$

$$P(k+1) = [P(k) + G(k+1)X^T(k+1)P(k)]\lambda(k+1), \quad (7c)$$

这里称 $P(k)$ 为估计方差阵; $\lambda(k)$ 为权因子. 显然, 要估计时变参数, 可以通过选择 $\lambda(k)$ 来得到遗忘因子法. 但它们还不能适用于快变参数的估计^[4]. 通常的遗忘因子法就是在(7)中取 $\lambda(k)$ 为常数因子时的特例.

对时变参数估计算法来说, 需要考虑两方面的问题: 一是跟踪速度要快, 即当参数发生时变(快变或跃变)时, 估计要能迅速跟踪, 这就要求算法能及时抛弃旧数据而增强新息的作用. 二是估计的稳定性, 即为参数稳定或缓慢时变时, 要有高的估计精度和消除噪声干扰的能力, 这就要求算法纳入足够的数据量. 因而这两者之间是矛盾的. 要解决此问题, 一个有效途径是对参数的时变情况进行检验, 并据此对算法进行适时校正. 我们在遗忘因子法的基础上利用系统输出与模型输出的误差来得出一种检验参数估计值是否接近真值的方法. 当检验出估计值偏离了真值时, 则对(7)中的估计方差阵 $P(k)$ 进行校正, 并给出了两种校正方法.

3. 参数时变的检验

考虑系统输出与模型输出的误差

$$e(k) = y(k) - X^T(k)\hat{\theta}(k). \quad (8)$$

从理论上讲 $e(k)$ 是非常复杂的随机过程的一个实现, 难于作严格的统计分析. 我们这里从工程角度作近似讨论. 假设参数估计 $\hat{\theta}(k)$ 就等于其真值 $\theta(k)$, 则

$$e(k) = X^T(k)\hat{\theta}(k) + \varepsilon_k - X^T(k)\hat{\theta}(k) = \varepsilon_k \quad (9)$$

就是零均值同分布正态白噪声序列. 反之, $\{e(k)\}$ 序列的均值可反映出 $\hat{\theta}(k)$ 是否精确, 因而我们作假设检验

$$H_0: E\{e(k)\} = 0, \quad H_1: E\{e(k)\} \neq 0. \quad (10)$$

如取 m 组数据进行检验, 可构造统计量

在原假设 H_0 下有

$$\frac{e(k)}{s(K)}\sqrt{m-1} \quad \text{服从 } t(m-1) \text{ 分布,} \quad (12)$$

其中 $t(m-1)$ 是自由度为 $m-1$ 的 t 分布变量, 因而可进行 t 检验

$$P\left\{\left|\frac{\bar{e}(k)}{S(k)}\sqrt{m-1}\right| < k_\alpha\right\} = 1 - \alpha. \quad (13)$$

在此检验下, 若假设 H_0 为真, 则说明 $\hat{\theta}(k)$ 以置信度 $1 - \alpha$ 等于真值. 实际上在时变参数的估计中, $\{\hat{\theta}(k)\}$ 要等于 $\{\theta(k)\}$ 是不可能也不必要的. 但依据检验的观点, 我们可以用判据

$$\left|\frac{\bar{e}(k)}{S(k)}\right| < d_0 \quad (14)$$

来给出估计 $\{\hat{\theta}(k)\}$ 是否接近 $\{\theta(k)\}$ 的检验. 但对时变参数估计来说, 不能一味纳入所有的旧数据而应着重当前的情况, 且应适合递推计算. 利用遗忘因子法的思想, 我们作下列改进. 取 $\bar{e}(k)$ 和 $S(k)$ 为

$$\left. \begin{aligned} \bar{e}(k) &= \frac{1}{M} \sum_{l=1}^k e(l) \mu^{k-l}, \\ S^2(k) &= \frac{1}{M} \sum_{l=1}^k [e(l) - \bar{e}(l)]^2 \mu^{k-l}, \end{aligned} \right\} \quad 0 < \mu < 1, \quad (15)$$

其中 μ 为检验的遗忘因子, M 取值如下述, 如果 $\{e(l)\}$ 为一个均值为 $E(e)$ 的总体的统计量, 则

$$\begin{aligned} E(\bar{e}(k)) &= E\left(\frac{1}{M} \sum_{l=1}^k e(l) \mu^{k-l}\right) = \frac{1}{M} \sum_{l=1}^k E[e(l)] \mu^{k-l} \\ &= \frac{1}{M} \left(\sum_{l=1}^k \mu^{k-l}\right) E(e). \end{aligned} \quad (16)$$

因而取

$$M = \sum_{l=1}^k \mu^{k-l} = 1 + \mu + \cdots + \mu^{k-1} = \frac{1 - \mu^k}{1 - \mu}, \quad (17)$$

就能使 $E[\bar{e}(k)] = E(e)$, 通常可近似取 $M = \frac{1}{1 - \mu}$. 对 (15) 式有递推形式

$$\left. \begin{aligned} \bar{e}(k) &= \mu \bar{e}(k-1) + (1 - \mu)e(k), \\ S^2(k) &= \mu S^2(k-1) + (1 - \mu)[e(k) - \bar{e}(k)]^2. \end{aligned} \right\} \quad (18)$$

由此可得检验判据

$$D = \left|\frac{\bar{e}(k)}{S(k)}\right| < d_0, \quad (19)$$

其中门限 d_0 和 (18) 中的 μ 值可由实际问题的先验知识来给出, 它取决于噪声大小和对跟踪速度的要求. 值得指出的是, 结合后面的校正方法而得出的整个算法, 对这两个值的选取并不敏感.

利用判据 (19), 当 $D \geq d_0$ 时就认为估计 $\hat{\theta}(k)$ 已偏离了真值, 需要对算法进行校

正. 下面提出两种校正方法.

4. 校正方法

对于 RLS 算法(7), 其估计方差阵 $P(k)$ 的某个主对角元素值的大小反映了当前其对应估计值的不确定性程度. 它的值大就会导致增加增益向量 $G(k)$ 的相应项, 同时对它项也有些影响. 因而增大 $P(k)$ 的主对角元素对加强新息的作用是迅速而有效的. 基于此, 我们提出下列两种校正方法.

4.1 统一校正

假设按遗忘因子法计算到第 N 步时判据(19)不成立, 则按下式对 $P(N+1)$ 作校正

$$P(N+1) = P(N+1) + v|\bar{e}(N)| \cdot I, \quad v > 0 \quad (20)$$

其中等号表示赋值; I 为单位矩阵; v 为校正系数. 对 v 的选取后面再加以讨论.

这样校正比象 [4] 中那样单纯给 P 的主对角元加一常数效果好. 因为 $|\bar{e}(N)|$ 的大小反映估计偏离真值的大小. 另外, 判据 (19) 不成立也可能是 $S(N)$ 很小所致, 此时如 $|\bar{e}(N)|$ 较小, 则说明估计模型的误差并不大, 不需要大的校正来提高跟踪速度而损失稳定性, 采用 (20) 式的校正正好满足这一点.

4.2 分别校正

系统在一段时间内通常只有一个或某几个参数可能发生时变, 那么在校正时理想的方法是只对发生变化的参数增加其新息增益, 以减小对其它参数估计的影响. 这里, 我们提出对 $P(k)$ 的主对角元给以不同校正量的方法.

当实际系统的某个参数发生了变化, 一般会导致其对应估计量按相应的趋势变化, 据此我们将估计量 $\hat{\theta}(k)$ 与其前一步 $\hat{\theta}(k-1)$ 的变差 $\Delta\hat{\theta}(k)$ 进行低通滤波

$$R(k) = \beta R(k-1) + (1-\beta)\Delta\hat{\theta}(k), \quad R(1) = 0, \quad (21)$$

其中 $0 < \beta < 1$, $\Delta\hat{\theta}(k) = \hat{\theta}(k) - \hat{\theta}(k-1)$, $R(k)$ 为 $\hat{\theta}(k)$ 的同维向量.

这样得到的 $R(k)$ 滤除了 $\Delta\hat{\theta}(k)$ 的不确定波动. 向量 $R(k)$ 的分量 $r_i(k)$ 的绝对值越大, 表明其对应估计量 $\hat{\theta}_i(k)$ 发生时变的可能性越大, 因而应对其给予较大的校正. 由此, 当判据 (19) 不成立时引入下列校正

$$P(k+1) = P(k-1) + v\bar{R}(k), \quad (22)$$

其中 v 仍为校正系数; $\bar{R}(k) = \text{diag}[|r_i(k)|]$ 这种校正法比统一校正更有针对性, 效果更好, 但增加了一点计算量. 算法结构如图 1 所示.

4.3 系数的确定

在我们提出的检验判据 (19) 和校正方法 (20)(21) 或 (22) 中涉及几个系数的确定, 即 μ , d_0 , β 和 v . 通常它们可由先验知识确定. 仿真实验证实整个算法对这些系数的选取不敏感, 即有较好的鲁棒性. 但要更好的确定这些系数也可采用实验寻优.

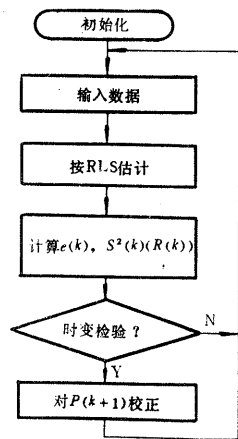


图 1 算法工作框图

例如在算法中系数 v 的影响较大, 假定已先确定了 μ , d_0 和 β , 那么 v 的选取可以用辨识实验数据来优化, 取优化指标为

$$J = \sum_{k=1}^N [y(k-1) - \hat{y}(k)]^2, \quad \hat{y}(k) = X^T(k) \hat{\theta}(k). \quad (23)$$

这样, 优化问题就为确定 v 使 J 达最小. 这可用优化算法求出, 如用黄金分割法. 实际上, 经过几次试探即可选定相对满意的系数.

5. 仿真实例

对本文提出的两种算法, 我们曾进行了大量的仿真来验证它的有效性. 为便于比较, 下面给出与文献[3][4]完全相同的仿真实验所得的结果, 系统模型取为

$$\begin{aligned} y(k) &= a_1 y(k-1) + a_2 y(k-2) + b_1 u(k-1) + b_2 u(k-2) + \varepsilon_k \\ &= a_1 y(k-1) - 0.7y(k-2) + u(k-1) + 0.5u(k-2) + \varepsilon_k. \end{aligned} \quad (24)$$

输入序列 $\{u(k)\}$ 采用幅值为 ± 1 的伪随机二进制序列; 噪声 $\varepsilon_k \sim N(0, 0.1)$ 仅有参数 a_1 是时变的, 且分为两种情况:

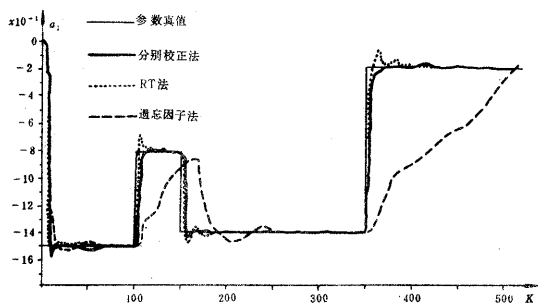


图 2 参数 a_1 呈阶跃时变的估计

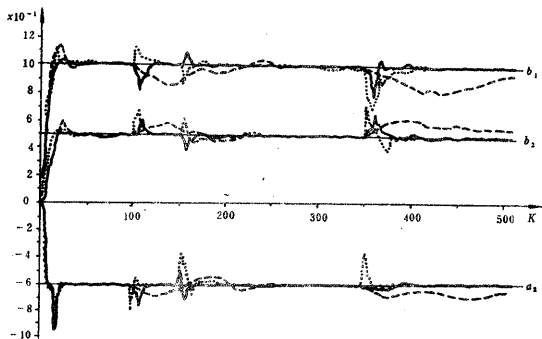


图 3 对其它参数的估计

1) a_1 呈阶跃时变 (图 1 和图 2)

图中只给出了采用遗忘因子法, Chen 的 RT 法和本文的分别校正法的辨识结果. 由图可见本文方法与 RT 的跟踪速度相近, 而估计稳定性方面却优于 RT 法. 至于与 AFMM 法的比较可参见文献 [4].

2) a_1 呈渐变 (图 3)

这时只给出了对 a_1 辨识的情况, 对其它几个参数的辨识受 a_1 时变的影响不大, 但用本文方法辨识的稳定性更好. 由图中可见, 除了遗忘因子法, 本文的方法与 RT 法均能给出满意的结果.

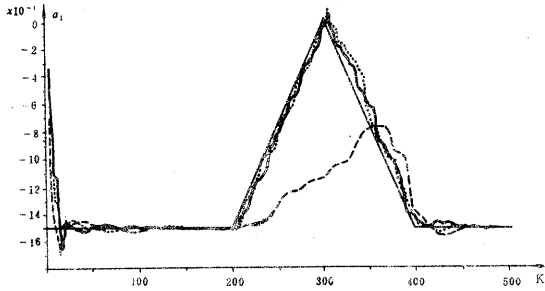


图 4 参数 a_1 呈渐变时的估计

另外, 我们对算法中系数选取的影响也进行了仿真实验, 以验证算法的鲁棒性, 实验结果表明, 当算法中的系数 μ , d_0 , β 和 ν 在一个较大范围取不同值时, 本文的两种方法都能给出比较满意的结果.

6. 结 论

本文提出的两种自适应递推辨识算法由 RLS 算法(遗忘因子法)、参数时变检验和两种校正方法构成, 算法简单且鲁棒性好, 计算量基本与普通的 RLS 算法相同, 它们适用于时变系统的参数在线辨识.

参 考 文 献

- (1) Goodwin, G. C., et. al, Dynamic System Identification: Experiment Design and Data Analysis, Academic Press, New York, (1977).
- (2) 潘德惠, 数学模型的统计方法, 辽宁科学技术出版社, 沈阳, (1986).
- (3) Andersson, P., Adaptive Forgetting in Recursive Identification Through Multiple Models, Int. J. Control, 42, 5, (1985).
- (4) Chen, M.J., et. al., Estimation Technique for Tracking Rapid Parameter Changes, Int. J. Control, 45, 4, (1987), 1387-1398.

Two Adaptive Recursive Algorithms for Estimating Time-varying Parameters

Wang Weiping, Pan Dehui

(Department of Management, Northeast University
of Technology, Shenyang)

Abstract: This paper is about two simple adaptive identification algorithms for estimating time-varying parameters (including abrupt and smooth changes). A parameter variation detector and two kinds of adjustment to the estimate covariance matrix $P(k)$ are proposed to embed in the Recursive Least Squares algorithm (RLS). A number of simulations are carried out to prove their validity and superiority in comparison with some other methods.

Key words: system identification; time-varying parameters; detection technique