

一种工业过程时变参数估计新算法 ——修正目标函数法

张承慧

(山东大学控制科学与工程学院, 济南 250061)

[摘要] 针对某些工业过程控制参数高度时变的特点和 LS 拟合算法存在病态估计的现象, 提出了一类基于修正目标函数的时变参数估计算法, 并给出递推辅助变量算法。此类算法不仅具有较强的实时跟踪能力和较高的估计精度, 而且能克服病态估计, 摆脱了伪随机信号在工程中带来的麻烦。新算法没有增加计算量, 非常适合于工程应用, 仿真及实际应用结果都证明新算法是有效的。文中给出了有关新算法的定理及其证明。

[关键词] 参数估计; 计算方法; LS 算法; 时变系统; 工业生产

1 引言

自从 Gauss 在 1775 年提出 LS 算法以来, 对其进行改进和完善的报导几乎没有停止过^[1,2], 但在实际工程中大量使用的仍是普通 LS 算法。这主要是因为它计算简单, 便于理解和掌握。当然 LS 算法确实存在诸如不稳健、易出现病态估计等缺陷, 许多文献给出了值得称道的改进算法, 但多数是停留在理论或仿真研究上, 很少有应用于工程实践的。其主要原因是改进算法复杂难以理解, 且研究人员受实践条件的制约。另外, 这些算法成立的前提是系统要满足持续激励条件, 持续激励条件不仅影响系统的可估计性, 而且影响自适应控制系统的鲁棒性。而实际过程中许多系统的输入信号变化不大, 持续激励条件难以满足, 普通 LS 算法及其相关改进算法不能直接使用, 必须外加伪随机信号, 这将给工程实现带来许多麻烦。事实上, 许多工业过程的先验知识是可以利用的, 根据先验知识重新修正目标函数是克服 LS 算法病态估计的有效途径之一。

笔者立足于工程实践, 提出了基于修正目标函数的时变参数估计算法, 有效地克服了病态估计,

巧妙地避免使用伪随机信号。同时考虑到某些生产过程参数的高度时变性, 在新算法中引入加权矩阵, 强化当前观测数据对参数估计的作用, 提高跟踪时变参数的能力。新算法一方面摆脱了伪随机信号带来的麻烦; 另一方面避免了数据饱和现象, 提高了自适应控制系统的鲁棒性和控制精度。它已用于实际工业生产, 取得了良好效果。仿真结果及实际应用结果都证明: 新型估计算法简单易行、收敛速度快和估计精度高。

2 基于修正目标函数的时变参数估计算法

时变参数估计是工程实践中经常遇到的问题, 为了使所估计的参数能跟踪实际参数的变化, 需要人为地削弱旧数据的影响, 以增强新数据的权。常用的办法主要有两类: **a.** 采用矩形窗函数, 舍弃旧的数据, 只用一段新的数据进行参数估计的限定记忆法; **b.** 是通过遗忘因子对历史数据进行指数加权, 以逐渐消除旧数据影响的渐消记法, 本文的方法只涉及第二类。

2.1 基本原理

考虑单变量离散线性时变系统

[收稿日期] 2001-04-09; **修回日期** 2001-07-02

[基金项目] 山东省自然科学基金资助项目 (Z200G01)

[作者简介] 张承慧 (1963-), 男, 山东陵县人, 山东大学教授, 博士生导师

$$y_{(t)} = \boldsymbol{\lambda}_{(t)}^T \hat{\boldsymbol{\Xi}}_{(t)} + \hat{\omega}_{(t)}, \quad (1)$$

式中 $t=1, 2, 3, \dots$,

$$\boldsymbol{\lambda}_{(t)}^T = [-y(t-1), -y(t-2), \dots, -y(t-n_a), u(t-1), u(t-2), \dots, u(t-n_b)], \quad (1a)$$

$$\boldsymbol{\Xi}_{(t)}^T = [a_{1(t)}, a_{2(t)}, \dots, a_{n_a(t)}, b_{1(t)}, b_{2(t)}, \dots, b_{n_b(t)}], \quad (1b)$$

其中 $\{u_{(i)}\}, \{y_{(j)}\}$ 分别是系统的输入和输出测量序列, $\hat{\omega}_{(t)}$ 为随机噪声, n_a, n_b 为过程的结构参数, $a_{i(t)}, b_{j(t)}$ 是待识时变参数。

为跟踪参数 $\boldsymbol{\Xi}_{(t)}$ 的变化应引入遗忘因子。

定义 1 设遗忘因子为 δ , ($0 < \delta < 1$) 且定义加权矩阵为

$$\mathbf{W}_{(t)} = \begin{bmatrix} \delta^{2(t-1)} & & & & \\ & \delta^{2(t-2)} & & & \\ & & \ddots & & \\ & & & \delta^2 & \\ & & & & 1 \end{bmatrix}, \quad (2)$$

式中 t 为总测量次数。在 LS 算法目标函数中引入 $\mathbf{W}_{(t)}$ 是准确跟踪时变参数实现实时估计的重要途径。

某些工业过程的先验知识可用来克服输入信号不满足持续激励条件引起的病态估计。这里假设先验知识向量为

$$\mathbf{M} = [\bar{a}_1, \bar{a}_2, \dots, \bar{a}_{n_a}, \bar{b}_1, \bar{b}_2, \dots, \bar{b}_{n_b}], \quad (3)$$

式中 $\{\bar{a}_i\}, \{\bar{b}_j\}$ 是待估计过程的时变参数均值, 参数均值这一先验知识可以通过日常生产数据分析获得, 这个先验知识是指导选样的重要参考。

为克服病态估计, 在新算法中应尽量利用先验知识向量 \mathbf{M} , 以降低过程对持续激励条件的要求; 为适应时变过程的要求, 新算法还必须重视 $\mathbf{W}_{(t)}$ 的作用。考虑到上述两个要求, 根据最小二乘原理可定义下列二次型目标函数。

定义 2 基于修正目标函数的时变参数 LS 估计算法的目标函数为

$$J = [\mathbf{Y}_{(t)} - \boldsymbol{\Lambda}_{(t)} \hat{\boldsymbol{\Xi}}_{(t)}]^T \mathbf{W}_{(t)} [\mathbf{Y}_{(t)} - \boldsymbol{\Lambda}_{(t)} \hat{\boldsymbol{\Xi}}_{(t)}] + \alpha [\mathbf{M}_{(t)} - \frac{1}{t} \hat{\boldsymbol{\Xi}}_{(t)}]^T [\mathbf{M}_{(t)} - \frac{1}{t} \hat{\boldsymbol{\Xi}}_{(t)}], \quad (4)$$

式中,

$$\mathbf{Y}_{(t)} = [y_{(1)}, y_{(2)}, \dots, y_{(t)}]^T, \quad (4a)$$

$$\boldsymbol{\Lambda}_{(t)} = \begin{bmatrix} \boldsymbol{\lambda}_{(1)}^T \\ \boldsymbol{\lambda}_{(2)}^T \\ \vdots \\ \boldsymbol{\lambda}_{(t)}^T \end{bmatrix} \quad (4b)$$

$$\mathbf{M}_{(t)} = \mathbf{M} - \frac{1}{t} \sum_{i=1}^{t-1} \hat{\boldsymbol{\Xi}}_{(i)}, \quad (4c)$$

其中, α 为权因子, 它决定了对先验知识的重视程度, $\hat{\boldsymbol{\Xi}}_{(i)}$ 是由 i 个采样值得到的参数估计值组成的向量。

定理 1 使目标函数式 (4) 达到最小的参数估计值 $\hat{\boldsymbol{\Xi}}_{(t)}$ 应满足下列正规方程:

$$\boldsymbol{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \mathbf{Y}_{(t)} + \alpha \mathbf{M}_{(t)} = \left(\boldsymbol{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \boldsymbol{\Lambda}_{(t)} + \frac{\alpha}{t} \mathbf{I} \right) \hat{\boldsymbol{\Xi}}_{(t)}. \quad (5)$$

如果 $\left(\boldsymbol{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \boldsymbol{\Lambda}_{(t)} + \frac{\alpha}{t} \mathbf{I} \right)$ 满秩可得唯一解

$$\hat{\boldsymbol{\Xi}}_{(t)} = \left(\boldsymbol{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \mathbf{Y}_{(t)} + \alpha \mathbf{M}_{(t)} \right) \cdot \left(\boldsymbol{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \boldsymbol{\Lambda}_{(t)} + \frac{\alpha}{t} \mathbf{I} \right)^{-1}, \quad (6)$$

式 (6) 便是基于修正目标函数的时变参数 LS 估计算法。

定理 1 证明: 对目标函数式 (4) 求导数得。

$$\begin{aligned} \frac{\partial J}{\partial \hat{\boldsymbol{\Xi}}_{(t)}} &= 2 \frac{\partial (\mathbf{Y}_{(t)} - \boldsymbol{\Lambda}_{(t)} \hat{\boldsymbol{\Xi}}_{(t)})^T}{\partial \hat{\boldsymbol{\Xi}}_{(t)}} \cdot \\ &\quad \mathbf{W}_{(t)} (\mathbf{Y}_{(t)} - \boldsymbol{\Lambda}_{(t)} \hat{\boldsymbol{\Xi}}_{(t)}) + \\ & 2\alpha t \frac{\partial \left(\mathbf{M}_{(t)} - \frac{1}{t} \hat{\boldsymbol{\Xi}}_{(t)} \right)}{\partial \hat{\boldsymbol{\Xi}}_{(t)}} \left(\mathbf{M}_{(t)} - \frac{1}{t} \hat{\boldsymbol{\Xi}}_{(t)} \right) = \\ & - 2 \boldsymbol{\Lambda}_{(t)}^T \mathbf{W}_{(t)} (\mathbf{Y}_{(t)} - \boldsymbol{\Lambda}_{(t)} \hat{\boldsymbol{\Xi}}_{(t)}) - \\ & 2\alpha t \frac{1}{t} \left(\mathbf{M}_{(t)} - \frac{1}{t} \hat{\boldsymbol{\Xi}}_{(t)} \right) = \\ & - 2 (\boldsymbol{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \mathbf{Y}_{(t)} - \boldsymbol{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \boldsymbol{\Lambda}_{(t)} \hat{\boldsymbol{\Xi}}_{(t)} + \\ & \quad \alpha \mathbf{M}_{(t)} - \frac{\alpha}{t} \hat{\boldsymbol{\Xi}}_{(t)}) = \\ & - 2 [(\boldsymbol{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \mathbf{Y}_{(t)} + \alpha \mathbf{M}_{(t)}) - \\ & \quad (\boldsymbol{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \boldsymbol{\Lambda}_{(t)} + \frac{\alpha}{t} \mathbf{I}) \hat{\boldsymbol{\Xi}}_{(t)}]. \quad (7) \end{aligned}$$

根据极值定理, 使目标函数式 (4) 最小的必要条件是 $\partial J / \partial \hat{\boldsymbol{\Xi}}_{(t)} = 0$, 即式 (5), 又 Hesse 矩阵为

$$\frac{\partial}{\partial \hat{\boldsymbol{\Xi}}_{(t)}} \left[\frac{\partial J}{\partial \hat{\boldsymbol{\Xi}}_{(t)}} \right]^T = 2 \left[\boldsymbol{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \boldsymbol{\Lambda}_{(t)} + \frac{\alpha}{t} \mathbf{I} \right]. \quad (8)$$

J 极小化的充分条件为

$$\frac{\partial}{\partial \hat{\Xi}_{(t)}} \left[\frac{\partial J}{\partial \hat{\Xi}_{(t)}} \right]^T = 2 \left[\mathbf{A}_{(t)}^T \mathbf{W}_{(t)} \mathbf{A}_{(t)} + \frac{\alpha}{t} \mathbf{I} \right] > 0. \quad (9)$$

由于权因子 α 可调整, 故即使矩阵 $\mathbf{A}_{(t)}^T$ 不满秩式 (9) 亦可成立且式 (6) 有唯一解。定理 1 得证。

推论 1 在式 (6) 中, 当先验知识权因子 $\alpha = 0$ 时, 则式 (6) 便是加权 LS 算法, 即:

$$\hat{\Xi}_{(t)} = (\mathbf{A}_{(t)}^T \mathbf{W}_{(t)} \mathbf{A}_{(t)})^{-1} \mathbf{A}_{(t)}^T \mathbf{W}_{(t)} \mathbf{Y}_{(t)}, \quad (10)$$

对应的目标函数为

$$J = [\mathbf{Y}_{(t)} - \mathbf{A}_{(t)} \hat{\Xi}_{(t)}]^T \mathbf{W}_{(t)} [\mathbf{Y}_{(t)} - \mathbf{A}_{(t)} \hat{\Xi}_{(t)}]. \quad (11)$$

推论 2 在式 (6) 中, 当 $\alpha = 0$ 且当 $\mathbf{W}_{(t)} = \mathbf{I}$ 时, 则式 (6) 便是普通 LS 算法, 即

$$\hat{\Xi}_{(t)} = (\mathbf{A}_{(t)}^T \mathbf{A}_{(t)})^{-1} \mathbf{A}_{(t)}^T \mathbf{Y}_{(t)}, \quad (12)$$

对应的目标函数为

$$J = [\mathbf{Y}_{(t)} - \mathbf{A}_{(t)} \hat{\Xi}_{(t)}]^T [\mathbf{Y}_{(t)} - \mathbf{A}_{(t)} \hat{\Xi}_{(t)}]. \quad (13)$$

从定理 1 中不难看出, 引入加权矩阵 $\mathbf{W}_{(t)}$ 和权因子 α 后, 为获得时变参数的估计值和克服病态估计提供了更多的灵活性。

定理 2 即使在过程式 (1) 持续激励条件不满足的情况下, 通过调整权因子 α 仍能保证式 (6) 有唯一解。

定理 2 说明: 从理论上讲, 若过程式 (1) 输入信号不满足持续激励条件, 在计算中表现为矩阵 $\mathbf{A}_{(t)}^T \mathbf{W}_{(t)} \mathbf{A}_{(t)}$ 不满秩, 而通过选择 α , 则能保证矩阵 $\left[\mathbf{A}_{(t)}^T \mathbf{W}_{(t)} \mathbf{A}_{(t)} + \frac{\alpha}{t} \mathbf{I} \right]$ 满秩, 仍能使式 (6) 有唯一的正确的解。显而易见, 若用普通算法就会出现多解而得不到正确的估算值。

性质 1 先验知识向量 \mathbf{M} 对基于式 (4) 的最小二乘解的影响, 随着观测次数的增大而减小, 当时 $t \rightarrow \infty$ 时, 式 (6) 趋向式 (10), 即新算法收敛于加权 LS 算法。

性质 1 说明: 从 (6) 式可以看出, 先验知识向量 \mathbf{M} 对最小二乘解的影响, 不仅与权因子 α 有关, 而且与观测次数 t 有关。 t 较小时影响明显, 可有效地避免由于输入信号不持续激励而带来的问

题; t 太大时, 先验知识向量几乎不起作用, 保证收敛性。

性质 1 非常重要, 因为 t 较小时, 持续激励条件一般不易满足, 增大 \mathbf{M} 的作用显然对避免病态估计十分有利; t 太大时, 保证收敛性则是新算法成立的理论基础。

值得注意的是, 遗忘因子 δ 和权因子 α 要经过适当的选择。在进行估计的初期, 为了得到一个好的收敛速度, δ 可选得小些, 而 α 可选得大些; 在估计的后期, 为得到好的精度 (小方差), δ 可趋近于 1, 而 α 可趋近于 0。

2.2 实时加权递推算法

利用新算法对时变过程式 (1) 进行参数估计, 显然希望得到其递推形式。而新算法不能直接给出递推形式, 需作近似处理。当 t 相对于 α 较大时, 数量矩阵 $\frac{\alpha}{t} \mathbf{I}$ 相对于 $\mathbf{A}_{(t)}^T \mathbf{W}_{(t)} \mathbf{A}_{(t)}$ 很小, 可以忽略。这时, 式 (6) 演变为

$$\hat{\Xi}_{(t)} = (\mathbf{A}_{(t)}^T \mathbf{W}_{(t)} \mathbf{Y}_{(t)} + \alpha \mathbf{M}_{(t)}) (\mathbf{A}_{(t)}^T \mathbf{W}_{(t)} \mathbf{A}_{(t)})^{-1}. \quad (14)$$

在推导递推公式时要频繁引用矩阵求逆引理。

引理 1 设 $\mathbf{A} + \mathbf{B}\mathbf{C}^T$ 和 $\mathbf{I} + \mathbf{C}^T \mathbf{B}^{-1} \mathbf{B}$ 满秩, 则下面的矩阵恒等式成立

$$(\mathbf{A} + \mathbf{B}\mathbf{C}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{I} + \mathbf{C}^T \mathbf{A}^{-1} \mathbf{B})^{-1} \mathbf{C}^T \mathbf{A}^{-1}. \quad (15)$$

定理 3 未知变量 Ξ 的基于修正目标函数的

加权最小二乘估计 $\hat{\Xi}$ 的递推公式为

$$\hat{\Xi}_{(t+1)} = \hat{\Xi}_{(t)} + \Psi_{(t+1)} (y_{(t+1)} - \lambda_{(t+1)}^T \hat{\Xi}_{(t)}) + \alpha \Pi_{(t+1)} (\mathbf{M}_{(t+1)} - \delta^2 \mathbf{M}_{(t)}), \quad (16)$$

$$\Psi_{(t+1)} = \frac{\Pi_{(t)} \lambda_{(t+1)}}{\delta^2 + \lambda_{(t+1)}^T \Pi_{(t)} \lambda_{(t+1)}}, \quad (17)$$

$$\Pi_{(t+1)} = \frac{1}{\delta^2} (\mathbf{I} - \Psi_{(t+1)} \lambda_{(t+1)}^T) \Pi_{(t)}. \quad (18)$$

定理 3 证明: 令

$$\mathbf{A}_{(t+1)} = \begin{bmatrix} \mathbf{A}_{(t)} \\ \lambda_{(t+1)}^T \end{bmatrix}, \quad (19)$$

$$\mathbf{Y}_{(t+1)} = \begin{bmatrix} \mathbf{Y}_{(t)} \\ y_{(t+1)} \end{bmatrix}, \quad (20)$$

$$\Pi_{(t)} = (\mathbf{A}_{(t)}^T \mathbf{W}_{(t)} \mathbf{A}_{(t)})^{-1}, \quad (21)$$

$$\text{则 } \Pi_{(t+1)} = (\mathbf{A}_{(t+1)}^T \mathbf{W}_{(t+1)} \mathbf{A}_{(t+1)})^{-1}. \quad (22)$$

由式 (2) 得

$$\mathbf{W}_{(t+1)} = \left[\begin{array}{c|c} \begin{bmatrix} \delta^{2(t-1)} & & & \\ & \delta^{2(t-2)} & & 0 \\ & & \ddots & \\ & 0 & & \delta^2 \\ \hline & & & & 1 \end{bmatrix} & \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \\ \hline \begin{bmatrix} \delta^2 \mathbf{W}_{(t)} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{1} \end{bmatrix} & \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \end{array} \right] = \begin{bmatrix} \delta^2 \mathbf{W}_{(t)} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{1} \end{bmatrix} \quad (23)$$

由式 (19)、式 (20) 及式 (23)，式 (22) 可变为

$$\begin{aligned}
 \mathbf{\Pi}_{(t+1)} &= \{ [\mathbf{\Lambda}_{(t)}^T \ ; \ \boldsymbol{\lambda}_{(t+1)}] \cdot \\
 &\left[\begin{array}{c|c} \delta^2 \mathbf{M}_{(t)} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{1} \end{array} \right] \left[\begin{array}{c} \mathbf{\Lambda}_{(t)} \\ \hline \boldsymbol{\lambda}_{(t+1)}^T \end{array} \right] \}^{-1} = \\
 &(\delta^2 \mathbf{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \mathbf{\Lambda}_{(t)} + \boldsymbol{\lambda}_{(t+1)} \boldsymbol{\lambda}_{(t+1)}^T)^{-1} = \\
 &[\delta^2 \mathbf{\Pi}_{(t)}^{-1} + \boldsymbol{\lambda}_{(t+1)} \boldsymbol{\lambda}_{(t+1)}^T]^{-1}. \quad (24)
 \end{aligned}$$

由引理 1，式 (24) 可化为

$$\mathbf{\Pi}_{(t+1)} = \frac{1}{\delta^2} (\mathbf{I} - \beta_{(t+1)} \mathbf{\Pi}_{(t)} \boldsymbol{\lambda}_{(t+1)} \boldsymbol{\lambda}_{(t+1)}^T) \mathbf{\Pi}_{(t)}, \quad (25)$$

式中 $\beta_{(t+1)} = (\delta^2 + \boldsymbol{\lambda}_{(t+1)}^T \mathbf{\Pi}_{(t)} \boldsymbol{\lambda}_{(t+1)})^{-1}$ ， $\beta(t+1)$ 是一个标量。

令：
$$\mathbf{\Psi}_{(t+1)} = \mathbf{\Pi}_{(t+1)} \boldsymbol{\lambda}_{(t+1)}, \quad (27)$$

将式 (25) 代入式 (26) 可得

$$\begin{aligned}
 \mathbf{\Psi}_{(t+1)} &= \frac{1}{\delta^2} \mathbf{\Pi}_{(t)} \boldsymbol{\lambda}_{(t+1)} \cdot \\
 (\mathbf{I} - \beta_{(t+1)} \mathbf{\Pi}_{(t)} \boldsymbol{\lambda}_{(t+1)} \boldsymbol{\lambda}_{(t+1)}^T) &= \\
 \frac{\mathbf{\Pi}_{(t+1)} \boldsymbol{\lambda}_{(t+1)}}{\delta^2 + \boldsymbol{\lambda}_{(t+1)}^T \mathbf{\Pi}_{(t)} \boldsymbol{\lambda}_{(t+1)}}. \quad (28)
 \end{aligned}$$

式 (17) 得证。

由式 (28)，式 (25) 可变为 $\mathbf{\Pi}_{(t+1)} = (1/\delta^2) (\mathbf{I} - \mathbf{\Psi}_{(t+1)} \boldsymbol{\lambda}_{(t+1)}^T) \mathbf{\Pi}_{(t)}$ ，式 (18) 得证。

另外，由式 (19)、式 (20) 及式 (23) 可得

$$\begin{aligned}
 \mathbf{\Lambda}_{(t+1)} \mathbf{W}_{(t+1)} \mathbf{Y}_{(t+1)} &= \\
 \{ [\mathbf{\Lambda}_{(t)}^T \ ; \ \boldsymbol{\lambda}_{(t+1)}] \left[\begin{array}{c|c} \delta^2 \mathbf{M}_{(t)} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{1} \end{array} \right] \left[\begin{array}{c} \mathbf{Y}_{(t)} \\ \hline \mathbf{y}_{(t+1)} \end{array} \right] \} = \\
 \delta^2 \mathbf{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \mathbf{Y}_{(t)} + \boldsymbol{\lambda}_{(t+1)} \mathbf{y}_{(t+1)}. \quad (29)
 \end{aligned}$$

由式 (29) 及式 (14) 得

$$\begin{aligned}
 \hat{\mathbf{\Xi}}_{(t+1)} &= [\mathbf{\Lambda}_{(t+1)}^T \mathbf{W}_{(t+1)} \mathbf{\Lambda}_{(t+1)}]^{-1} \cdot \\
 [\mathbf{\Lambda}_{(t+1)}^T \mathbf{W}_{(t+1)} \mathbf{Y}_{(t+1)} + \alpha \mathbf{M}_{(t+1)}] &= \\
 \mathbf{\Pi}_{(t+1)} [\delta^2 \mathbf{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \mathbf{Y}_{(t)} + \boldsymbol{\lambda}_{(t+1)} \mathbf{y}_{(t+1)}] + \\
 \alpha \mathbf{\Pi}_{(t+1)} \mathbf{M}_{(t+1)}. \quad (30)
 \end{aligned}$$

由式 (14) 及式 (21) 得

$$\begin{aligned}
 \mathbf{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \mathbf{Y}_{(t)} &= (\mathbf{\Lambda}_{(t)}^T \mathbf{W}_{(t)} \mathbf{\Lambda}_{(t)}) \hat{\mathbf{\Xi}}_{(t)} - \\
 \alpha \mathbf{M}_{(t)} &= \mathbf{\Pi}_{(t+1)}^{-1} \hat{\mathbf{\Xi}}_{(t)} - \alpha \mathbf{M}_{(t)}. \quad (31)
 \end{aligned}$$

将式 (31) 代入式 (30) 可得

$$\begin{aligned}
 \hat{\mathbf{\Xi}}_{(t+1)} &= \mathbf{\Pi}_{(t+1)} (\delta^2 \mathbf{\Pi}_{(t)}^{-1} \hat{\mathbf{\Xi}}_{(t)} + \boldsymbol{\lambda}_{(t+1)} \mathbf{y}_{(t+1)}) + \\
 \alpha \mathbf{\Pi}_{(t+1)} (\mathbf{M}_{(t+1)} - \delta^2 \mathbf{M}_{(t)}) \quad (32)
 \end{aligned}$$

由式 (24) 得

$$\delta^2 \mathbf{\Pi}_{(t)}^{-1} = \mathbf{\Pi}_{(t+1)}^{-1} - \boldsymbol{\lambda}_{(t+1)} \boldsymbol{\lambda}_{(t+1)}^T. \quad (33)$$

将式 (33) 代入式 (32) 有

$$\begin{aligned}
 \hat{\mathbf{\Xi}}_{(t+1)} &= \mathbf{\Pi}_{(t+1)} [\mathbf{\Pi}_{(t+1)}^{-1} - \boldsymbol{\lambda}_{(t+1)} \boldsymbol{\lambda}_{(t+1)}^T] \hat{\mathbf{\Xi}}_{(t)} + \\
 \boldsymbol{\lambda}_{(t+1)} \mathbf{y}_{(t+1)}] + \alpha \mathbf{\Pi}_{(t+1)} (\mathbf{M}_{(t+1)} - \delta^2 \mathbf{M}_{(t)}) &= \\
 \hat{\mathbf{\Xi}}_{(t)} + \mathbf{\Pi}_{(t+1)} \boldsymbol{\lambda}_{(t+1)} (\mathbf{y}_{(t+1)} - \boldsymbol{\lambda}_{(t+1)}^T \hat{\mathbf{\Xi}}_{(t)}) + \\
 \alpha \mathbf{\Pi}_{(t+1)} (\mathbf{M}_{(t+1)} - \delta^2 \mathbf{M}_{(t)}). \quad (34)
 \end{aligned}$$

将式 (27) 代入式 (34) 得

$$\begin{aligned}
 \hat{\mathbf{\Xi}}_{(t+1)} &= \hat{\mathbf{\Xi}}_{(t)} + \mathbf{\Psi}_{(t+1)} (\mathbf{y}_{(t+1)} - \boldsymbol{\lambda}_{(t+1)}^T \hat{\mathbf{\Xi}}_{(t)}) + \\
 \alpha \mathbf{\Pi}_{(t+1)} (\mathbf{M}_{(t+1)} - \delta^2 \mathbf{M}_{(t)}). \quad (35)
 \end{aligned}$$

定理 3 即式 (16) 证毕。

在定理 3 中，新递推算法的初值一般可令 $\hat{\mathbf{\Xi}}_{(0)} = 0$ ， $\mathbf{\Pi}_{(0)} = \sigma^2 \mathbf{I}$ ，其中 σ 为充分大的正数。

新递推算法的步骤如下：

- Step 1 置初值 $\hat{\mathbf{\Xi}}_{(0)}$ ， $\mathbf{\Pi}_{(0)}$ ，输入初始数据；
- Step 2 采样当前输入和输出；
- Step 3 按式 (16)、式 (17) 及式 (18) 计算 $\hat{\mathbf{\Xi}}_{(t)}$ ， $\mathbf{\Psi}_{(t)}$ ， $\mathbf{\Pi}_{(t)}$ ；

Step 4 返回 Step 2 直到收敛满足要求为止。

2.3 辅助变量递推算法

根据数理统计及 LS 算法的知识，不难发现，当过程式 (1) 中的 $\hat{\omega}_{(t)}$ 为有色噪声时，定理 1 及定理 3 所给出的算法同普通 LS 算法一样，其估计值 $\hat{\mathbf{\Xi}}_{(t)}$ 不能保证是一致估计。为此，这里仿照普通最小二乘辅助变量算法可以得到基于修正目标函数的时变参数辅助变量算法。

定理 4 引入辅助变量 $\hat{\Sigma}_{(t)}$ ，并选择加权矩阵 $\mathbf{W}_{(t)}$ 使其满足

$$\mathbf{\Sigma}_{(t)} = \mathbf{W}_{(t)} \mathbf{\Lambda}_{(t)}, \quad (36)$$

如果 $\mathbf{\Sigma}_{(t)}$ 满足以下两个条件：

$$\mathbf{E} \{ \mathbf{\Sigma}_{(t)}^T \mathbf{W}_{(t)} \} = \mathbf{0}, \quad (37)$$

$$\mathbf{E} \{ \mathbf{\Sigma}_{(t)}^T \mathbf{\Lambda}_{(t)} \} = \mathbf{\Gamma}, \mathbf{\Gamma} \text{ 非奇异}, \quad (38)$$

那么，基于修正目标函数的时变参数辅助变量算法

可表示为

$$\hat{\Xi}_{(t)} = \left[\Sigma_{(t)}^T \mathbf{A}_{(t)} + \frac{\alpha}{t} \mathbf{I} \right]^{-1} \left[\Sigma_{(t)}^T \mathbf{Y}_{(t)} + \alpha \mathbf{M}_{(t)} \right]. \quad (39)$$

定理 4 证明从略。

同样地将式 (39) 稍作近似处理便可得到基于修正目标函数的辅助变量递推算法。

定理 5 未知变量 Ξ 的基于修正目标函数的辅助变量递推公式为

$$\hat{\Xi}_{(t+1)} = \hat{\Xi}_{(t)} + \Psi_{(t+1)} [y_{(t+1)} - \lambda_{(t+1)}^T \hat{\Xi}_{(t)}] + \alpha \Pi_{(t+1)} [M_{(t+1)} - \delta^2 M_{(t)}], \quad (40)$$

$$\Psi_{(t+1)} = \frac{\Pi_{(t)} \zeta_{(t+1)}}{\delta^2 + \lambda_{(t+1)}^T \Pi_{(t)} \zeta_{(t+1)}}, \quad (41)$$

$$\Pi_{(t+1)} = \frac{1}{\delta^2} [\mathbf{I} - \Psi_{(t+1)} \lambda_{(t+1)}^T] \Pi_{(t)}. \quad (42)$$

式中 $\zeta_{(t)} = [-v_{(t-1)}, -v_{(t-2)}, \dots, -v_{(t-na)}, u_{(t-1)}, u_{(t-2)}, \dots, u_{(t-nb)}]$, 可取 $v_{(t)} = \Sigma_{(t)}^T \hat{\Xi}_{(t)}$ 。

定理 5 证明从略。

另外, 新算法也可推广到广义 LS 算法, 在此暂不作研究。

3 仿真结果及实用情况

3.1 仿真模型

分别采用 LS 算法和基于修正目标函数的 LS 算法来估计时变系统, 采用的时变系统是水泥生料混合过程。

水泥生料混合控制历来是水泥生产的关键问题。图 1 示出水泥混合系统的流程。计算机根据 X 荧光分析仪检测数据分析计算后, 与给定值比较, 控制电子皮带秤, 调节原料的配比。

如果假设系统的工作点稳定, 且混合过程理想, 没有死区时间, 那么水泥生粉中氧化物的成分与生料重量的关系如下:

$$\begin{pmatrix} O_{CaO} \\ O_{SiO_2} \\ O_{Al_2O_3} \\ O_{Fe_2O_3} \end{pmatrix} = \begin{pmatrix} C_1 & C_2 & O & O \\ S_1 & S_2 & S_3 & O \\ A_1 & A_2 & O & O \\ F_1 & F_2 & O & F_4 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \end{pmatrix}, \quad (43)$$

或用下列向量形式表示:

$$\mathbf{O} = \mathbf{C}\boldsymbol{\omega} \quad (44)$$

式中, 矩阵 \mathbf{C} 为生料成分矩阵; \mathbf{O} 为生粉中氧化物的含量; $\boldsymbol{\omega}$ 表示生料中石灰石、粘土、石英砂

和铁矿石的重量比。

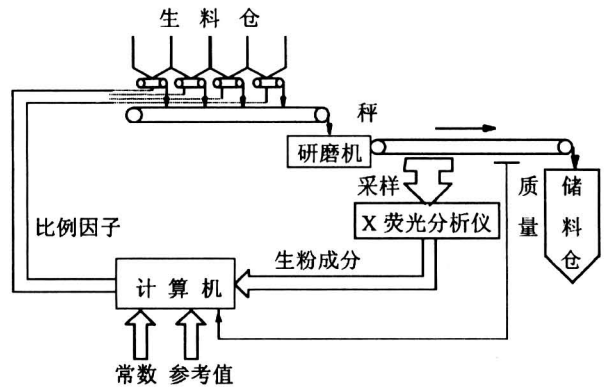


图 1 水泥生料磨机成分控制系统

Fig.1 Control system of cement composition at raw mill

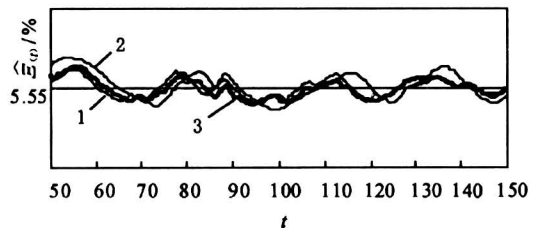
由于水泥原料(石灰石、粘土、石英砂和铁矿石等)大多呈块状、分散性大, 直接实时测量其成分是不可能的, 只能根据磨机输出的矩阵 \mathbf{O} 的生粉化学成分和原料进料量 $\boldsymbol{\omega}$ 来估计生料成分 \mathbf{C} , 从而为控制算法提供重要依据。这里以成分波动较大的石灰石中的 SiO_2 含量为例给出了其对比仿真结果。

3.2 仿真结果

图 2 表明: 基于修正目标函数的整匹估计算法的估计精度优于普通 LS 算法, 引入遗忘因子 δ 以后, 新算法对时变参数的跟踪能力大为提高, 精度也有所改善。图 3 表明: 基于修正目标函数的递推算法比普通递推 LS 算法对时变参数的估计精度大为提高。

3.3 应用情况

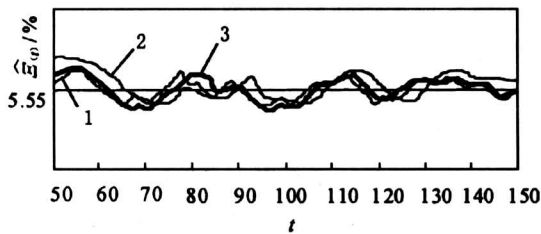
本算法是多年以来在水泥生料自适应控制系统



1—实际值; 2—普通 LS 估计值 ($\delta=1, \alpha=0$); 3—新算法估计值 ($\delta=0.985, \alpha=5.8$)

图 2 整匹算法对比仿真结果

Fig.2 Comparison of the new and old algorithm with bitch processing format



1—实际值；2—普通 LS 估计值 ($\delta=1, \alpha=0$)；
3—新算法估计值 ($\delta=0.985, \alpha=5.8$)

图 3 递推算法对比仿真结果

Fig.3 Comparison of the new and old algorithm with recursive format

研究的基础上经不断修正总结而来的。该算法结合自校正控制技术用于工程实际，已将磨机输出生粉成分方差降低 30 % 左右。工业现场使用结果表明：所提出的系统具有较高的鲁棒性和控制精度，可将磨机输出生粉成分方差降低到 1.58，而且长期运行稳定。

Modified Performance Index Method for Parameter Estimation for Industrial Process with Time-varying

Zhang Chenghui

(School of Control Science & Engineering, Shandong University, Jinan 250061, China)

[Abstract] A new estimation method based on modified performance index is proposed for time-varying process to avoid the ill-condition of least squares method. Based on the new method, a series of new recursive algorithms such as the new recursive least squares algorithm, the instrumental variable algorithm are constructed. The new identification method can ease the ill-condition and improve the identification accuracy without using the additional signals. Moreover, without additional computation, it is suitable for engineering practice very much. Both simulation results and practical operation show the new identification algorithm is effective. In this paper, the justification of this modified algorithm is proved.

[Key words] parameter estimation; method of calculation; least square method; time-varying system; industrial production

4 结论

导出了一类基于修正目标函数的时变参数 LS 算法，克服普通算法使病态程度加剧的缺陷，并提高对时变参数的跟踪能力；避免数据饱和现象，可提高自适应控制系统的鲁棒性。这种算法还具有运算量小、收敛性好的优点，十分便于工程实现和应用。已取得良好的控制效果。此外，以所提出的算法为基础，还易于构成相应的递推辅助变量算法、广义 LS 算法等，还可以构成对数据加权处理的限定记忆估计算法。

参考文献

- [1] 徐嗣鑫, 陈里得. 自适应向量遗忘因子辨识算法 [J]. 自动化学报, 1994, 20(3): 29~299
- [2] Davies W D T. System identification for self-adaptive control [M]. John Wiley, 1970