文章编号: 1000−8152(2009)01−0069−04

一种新的基于遗忘因子的递推子空间辨识算法

杨 华[1], 李少远[2]

(1. 中国海洋大学 信息科学与工程学院, 山东 青岛 266100; 2. 上海交通大学 自动化研究所, 上海 200240)

摘要: 针对工业系统中广泛存在的时变特性, 提出一种新的递推子空间辨识算法, 实现对系统状态空间模型的在线递推估计. 为更好地跟踪系统时变特性, 研究基于遗忘因子的输入输出数据矩阵构造机制, 以提高递推算法的收敛速度; 针对算法中奇异值分解的求解问题, 将梯度型算法引入基于遗忘因子的状态子空间跟踪中, 实现对广义能观测矩阵的估计, 避免了子空间近似带来的估计有偏性; 该算法计算简单有效, 且对初值具有更高的鲁棒性; 最后给出该递推算法的性能分析, 理论证明其收敛性, 并通过仿真实例验证算法的有效性.

关键词: 递推算法; 子空间方法; 在线辨识; 收敛性

中图分类号: TP273　　**文献标识码**: A

A novel recursive MOESP subspace identification algorithm based on forgetting factor

YANG Hua[1], LI Shao-yuan[2]

(1. College of Information Science and Engineering, Ocean University of China, Qingdao Shandong 266100, China;
2. Institute of Automation, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: A new recursive subspace identification algorithm is proposed for the recursive estimation of state space model of linear time-varying systems. A forgetting factor is introduced in the Hankel matrices of the input-output data to increase the convergent rate and improve the performance in tracking the time-varying information. In solving the singular value decomposition (SVD) problem, a gradient-type subspace tracking method is employed to update the state-space subspace based on forgetting factor, realizing the unbiased estimation of the extended observability matrix and improving the robustness to the uncertainty in initial values. The proposed method is simple and highly accurate in numerical computation. The convergence of the proposed method is also proved theoretically. Finally, the efficiency of this method is illustrated with a simulation example.

Key words: recursive identification; subspace method, online identification; convergence analysis

1 引言(Introduction)

在很多情况下, 需要在系统运行过程中在线获取系统模型, 特别是对于工业系统中广泛存在的时变特性, 必须利用当前获得的观测值更新辨识结果, 即在辨识过程中对可获得的输入输出数据进行序列、递推处理[1,2]. 因此, 在线递推辨识技术在自适应控制、信号处理以及系统性能监测与评估中都得到了广泛的应用.

作为一种新型的状态空间模型辨识方法, 子空间方法已成为控制和辨识领域的研究热点问题, 特别是对于递推算法的研究也有了一定的初步成果. 文献[1]基于矩阵逆引理提出了MOESP类子空间算法的递推实现, 但没有得到广义能观测矩阵的递推更新, 因此还没有解决关键的计算负担问题. 文献[3]为避免SVD分解的求解, 基于投影近似子空间跟踪(PAST)实现对奇异值分解的更新以递推获取扩展能观测矩阵, 但是由于投影的近似性, 估计值不一定能收敛到真值. 文献[4]考虑4SID类子空间辨识算法的递推实现形式, 但是系统矩阵$\{A, B, C, D\}$的估计均为离线算法. 特别是针对工业系统中广泛存在的时变特性, 子空间递推辨识算法的实现仍有很多有待解决的问题.

本文针对一类线性时变系统, 提出一种新的递推子空间辨识算法, 实现基于RLS-like遗忘因子的数据Hankle矩阵的构造, 更好地跟踪系统时变特性, 以提高递推算法的收敛速度. 将梯度型算法引入到基

收稿日期: 2007−07−19; 收修改稿日期: 2008−03−11.
基金项目: 国家自然科学基金资助项目(60774015, 60604018, 60534020); 国家863计划资助项目(2006AA04z173); 高校博士点基金资助项目(20060248001).

于遗忘因子的子空间跟踪中, 对初值具有更高的鲁棒性, 而且该算法在数值计算上具有更高的简化性和精确度. 在对已有离线算法做相应修正的基础上, 结合RLS算法实现对系统矩阵$\{A,B,C,D\}$的在线递推估计; 最后给出递推算法的收敛性分析, 并通过数值仿真验证算法的有效性.

2 问题描述(Problem description)

考虑如下线性时变系统

$$\bar{y}(t)=f(\theta(t),\bar{u}(t)), \tag{1}$$

其中: f为未知函数, $\bar{u}(t)\in\mathbb{R}^m$和$\bar{y}(t)\in\mathbb{R}^l$分别为$t$时刻的系统输入和输出, $\theta(t)$为时变的系统参数. 输入输出测量信号模型为

$$y(k)=\bar{y}(t)+p(t), \tag{2}$$

$$u(k)=\bar{u}(t)+v(t). \tag{3}$$

其中: $p\in\mathbb{R}^l$, $v\in\mathbb{R}^m$为系统噪声, 包含测量噪声、过程噪声等不确定性影响, 假定噪声p, v均为有界噪声信号, 且与过去时刻的输入信号u不相关.

本文研究的问题可以描述为: 在子空间辨识框架内, 提出一种新的递推子空间辨识方法, 基于在线获得的序列输入输出数据u和y, 引入带RLS-like遗忘因子的数据矩阵构造机制, 更高效地描述系统时变特性, 在线递推辨识如下状态空间模型:

$$\begin{aligned}&x(t+1)=Ax(t)+Bu(t)+w(t),\\&y(t)=Cx(t)+Du(t)+v(t).\end{aligned} \tag{4}$$

在对已有算法做相应修正的基础上, 结合RLS算法实现对系统矩阵$\{A,B,C,D\}$的在线递推估计, 满足递推算法的辨识精度和收敛性要求.

为保证系统的可辨识性以及收敛性分析需要, 假定系统满足下述条件:

A1) (A,B)能控, (A,C)能观测, 且系统为最小实现.

A2) 外部输入u与噪声w, v不相关, 且u满足充分激励条件.

3 递推子空间辨识算法(Recursive MOESP subspace identification algorithm)

针对系统(1), 子空间辨识算法的递推实现主要包括3个基本步骤: 数据Hankel矩阵的构造、广义能观测矩阵的更新和系统矩阵的递推估计.

3.1 数据Hankel矩阵构造(Construction of data Hankel matrices)

给定输入输出序列$\{u(i),y(i)\}(i=1,2,\cdots,j)$, 构造新的数据Hankel矩阵[5]如下:

$$U^{\beta}_{1,N,j}=\begin{bmatrix}\beta^{j-1}u(1) & \beta^{N-1}u(j-N+1)\\ \beta^{j-2}u(2) & \beta^{N-2}u(j-N+2)\\ \vdots & \vdots\\ \beta^{j-N}u(N) & u(j)\end{bmatrix}, \tag{5}$$

矩阵$Y^{\beta}_{1,N,j}$也可类似定义, 其中β为指数加权遗忘因子, 满足$0<\beta<1$, 则旧的数据在辨识中的权重比当前数据小, 因为对于时变系统而言, 新的数据能更好地描述当前系统特性. 则新的数据Hankel矩阵可以记作:

$$U^{\beta}_{1,N,j}=T_mU^{\mathrm{T}}_{1,N,j}, \tag{6}$$

$$Y^{\beta}_{1,N,j}=T_pY^{\mathrm{T}}_{1,N,j}. \tag{7}$$

其中:

$$\begin{aligned}&T_m=\mathrm{diag}\{\beta^{N-1}I_m,\beta^{N-2}I_m,\cdots,I_m\},\\&T_p=\mathrm{diag}\{\beta^{N-1}I_p,\beta^{N-2}I_p,\cdots,I_p\},\\&T=\mathrm{diag}\{\beta^{N-1},\beta^{N-2},\cdots,1\}.\end{aligned}$$

状态序列X^{β}_i定义为

$$X^{\beta}_i=[\beta^{t-i+1}x_0\cdots\beta x_{t-i}\,x_{t-i+1}]. \tag{8}$$

根据上述矩阵的定义, 系统输入输出等式为:

$$Y^{\beta}_{0,i,t}=\varGamma^{\beta}_iX^{\beta}_i+H^{\beta}_iU^{\beta}_{0,i,t}+N^{\beta}_{0,i,t}. \tag{9}$$

这一等式作为递推子空间辨识的基础, 在辨识中起着重要作用其中. 其中

$$\varGamma^{\beta}_i=[(\beta^{i-1}C)^{\mathrm{T}}\,(\beta^{i-2}CA)^{\mathrm{T}}\cdots(CA^{i-1})^{\mathrm{T}}]^{\mathrm{T}}. \tag{10}$$

同理, 其他矩阵可作类似定义. 子空间辨识的重要步骤即实现对上述矩阵子空间的重构, 并基于这一重构实现对系统矩阵的估计.

为实现广义能观测矩阵的递推更新, 需要在每一步确定一个观测向量$\phi^{\beta}_i(t+1)$的更新. 假定在$t+1$时刻, 新的输入输出对$\{u(t+1),y(t+1)\}$已知, 构造下述数据向量

$$\phi^{\beta}_u(t+1)=\begin{bmatrix}\beta^{N-1}u(j-N+2)\\ \beta^{N-2}u(j-N+3)\\ \vdots\\ u(t+1)\end{bmatrix}. \tag{11}$$

$\phi^{\beta}_y(t+1)$类似定义. 则更新的数据Hankel矩阵为$\begin{bmatrix}U^{\beta}_{0,i,t}\\ Y^{\beta}_{0,i,t}\end{bmatrix}=\begin{bmatrix}\beta U^{\beta}_{0,i,t} & \phi^{\beta}_u(t+1)\\ \beta Y^{\beta}_{0,i,t} & \phi^{\beta}_y(t+1)\end{bmatrix}$, 这一新的数据矩阵的构造中包含了不同时刻数据对辨识的不同影响.

过程中存在时变特性是对系统模型进行在线更新一个重要原因, 为使数据能够更好地跟踪系统信息变化, 在数据矩阵的构造中引入RLS-like 指数遗

忘因子机制, 数据对系统特性的更好描述, 也将使得递推算法的收敛速度得到一定的提高.

3.2 梯度法更新广义能观测矩阵(Gradient type matrix subspace tracking)

本节利用上述基于遗忘因子的数据Hankle矩阵, 采用梯度型算法实现对广义能观测矩阵Γ_i^β的递推估计. 在t时刻对数据矩阵进行如下QR分解:

$$\begin{bmatrix} U_{0,i,t}^\beta \\ Y_{0,i,t}^\beta \end{bmatrix} = \begin{bmatrix} R_{11}(t) & 0 \\ R_{12}(t) & R_{22}(t) \end{bmatrix} \begin{bmatrix} Q_1(t) \\ Q_2(t) \end{bmatrix}. \quad (12)$$

则当$t+1$时刻获取新的数据后, 对更新的增广数据矩阵做QR分解得

$$\begin{bmatrix} U_{0,i,t+1}^\beta \\ Y_{0,i,t+1}^\beta \end{bmatrix} = \begin{bmatrix} \beta R_{11}(t) & 0 & \phi_u^\beta(t+1) \\ \beta R_{21}(t) & \beta R_{22}(t) & \phi_y^\beta(t+1) \end{bmatrix} \times \begin{bmatrix} Q_1(t) & 0 \\ Q_2(t) & 0 \\ 0 & 1 \end{bmatrix}.$$

通过一系列Givens旋转变换, 上式R部分变为

$$R = \begin{bmatrix} \beta R_{11}(t+1) & 0 & 0 \\ \beta R_{21}(t+1) & \beta R_2(t) & \phi^\beta(t+1) \end{bmatrix}. \quad (13)$$

其中矩阵$R_{22}(t+1)$可通过利用新数据更新QR分解获得, 表示如下:

$$R_{22}(t+1) = \begin{bmatrix} \beta R_{22}(t) & \phi^\beta(t+1) \end{bmatrix}. \quad (14)$$

为实现对广义能观测矩阵子空间的重构, 对$R_{22}(t)$进行SVD分解[1]得

$$R_{22}(t)=[U_1 \ \ U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^{\mathrm{T}} \\ V_2^{\mathrm{T}} \end{bmatrix} = U_1 \Sigma_1 V_1^{\mathrm{T}}. \quad (15)$$

则广义能观测矩阵Γ_i^β可由下式计算得到:

$$\Gamma_i^\beta = U_1 \Sigma_1^{1/2}.$$

为实现对Γ_i^β的递推更新, 把这一子空间跟踪问题转换为如下无约束最优化问题的求解[6]:

$$J(W) = \mathrm{E}\left\|x - WW^{\mathrm{H}}x\right\|^2. \quad (16)$$

则有$J(W) = \mathrm{tr}(\Xi) - 2\mathrm{tr}(W^{\mathrm{H}}\Xi W) + \mathrm{tr}(W^{\mathrm{H}}\Xi W \times W^{\mathrm{H}}W)$, 其中$\Xi = \mathrm{E}\{xx^{\mathrm{H}}\}$, x为矩阵$R_{22}(t)$, 由于$J(W)$只有一个全局最小值, 此时权重W的张成等于信号子空间, 而且没有其他局部极小值, 因此可保证其全局收敛.

在数据中引入遗忘因子机制, 即对$\Xi = \mathrm{E}\{xx^{\mathrm{H}}\} = \sum\limits_{k=1}^{t} \beta^{t-k} x(k) x^{\mathrm{H}}(k)$, 应用梯度下降法进行信号子空间即系统的广义能观测矩阵Γ_i^β的提取. 类似于自适应滤波器的LMS算法, 用单位阵I代替$W(t-1)W(t-1)^{\mathrm{H}}$, 子空间更新公式为

$$y(t) = W^{\mathrm{H}}(t-1)x(t), \quad (17)$$

$$W(t) = W(t-1) - \mu[-2x(t) + x(t) + W(t-1)y(t)]y^{\mathrm{H}}(t). \quad (18)$$

利用子空间辨识与信号处理技术的相关性, 将子空间跟踪问题转化为一个优化问题的求解, 实现对于信号子空间的递推更新, 即广义能观测矩阵Γ_i^β的递推估计.

3.3 系统矩阵递推实现(Recursive implementation of system matrices)

在获得系统广义能观测矩阵Γ_i^β的估计值之后, 基于最小二乘算法实现对系统矩阵的递推估计. 矩阵C可以直接从$\hat{\Gamma}_i^\beta(t+1)$提取

$$\hat{C}(t+1) = \beta^{-i+1} \hat{\Gamma}_i^\beta(t+1)(1:p, 1:n). \quad (19)$$

矩阵A可通过下述线性方程求解获得

$$\hat{\Gamma}_i^\beta(t+1)(1:p(i-1), 1:n)\hat{A}(t+1) = \hat{\Gamma}_i^\beta(t+1)(p+1:pi, 1:n). \quad (20)$$

因此矩阵A可通过递推最小二乘法求解, 递推算法如下:

$$\begin{cases} \hat{A}(t) = \hat{A}(t-1) + \gamma(t) P^{-1}(t) \hat{\Gamma}_i^{\beta(1)} \varepsilon(t), \\ \varepsilon(t) = \hat{A}(t-1)\hat{\Gamma}_i^{\beta(1)} - \beta \hat{\Gamma}_i^{\beta(2)}, \\ P(t) = P(t-1) + \gamma(t)[\hat{\Gamma}_i^{\beta(1)}(\hat{\Gamma}_i^{\beta(1)})^{\mathrm{T}} - P(t-1)]. \end{cases} \quad (21)$$

矩阵B和D包含在矩阵H_i^β中, 为得到其估计值, 须从式(9)中消去$\Gamma_i^\beta X_i^\beta$项, 定义矩阵$\Pi_{\Gamma^\perp}$为子空间Γ_i^β的正交补, 式(9)两边同时左乘$\Pi_{\Gamma^\perp}$右乘$(U_{0,i,t}^\beta)^\dagger$得

$$\Pi_{\Gamma^\perp} Y_{0,i,t}^\beta (U_{0,i,t}^\beta)^\dagger = \Pi_{\Gamma^\perp} H_i^\beta + \Pi_{\Gamma^\perp} N_{0,i,t}^\beta (U_{0,i,t}^\beta)^\dagger.$$

在每一时刻用最近的估计值代替Γ_i^β, 则$\hat{B}$和$\hat{D}$的估计可以类似于式(21)递推求解.

可以看出, 在传统的离线算法中, 矩阵A,B和D的估计需要通过求解一组超定方程组的最小二乘解获得. 本文使用递推最小二乘实现对系统矩阵的估计, 进一步降低在线计算的复杂度.

3.4 收敛性分析(Convergency analysis)

定理 1 为将该递推辨识方法应用于实际工业过程中, 需要对其性能进行分析. 有如下结论: 若条件A1), A2)成立, 当$t \to \infty$时, 由式(16)估计得到$\hat{\Gamma}_i^\beta(t)$几乎肯定(a.s.)收敛到Γ_i^β, 即

$$\lim_{t\to\infty} \hat{\Gamma}_i^\beta(t) \to \Gamma_i^\beta \quad \text{a.s.}$$

证 对$R_{22}(t)$进行SVD分解如式(15)所示, 其中U_1为主奇异值, 即$\Xi = R_{22}R_{22}^{\mathrm{T}}$的主特征向量, 由

子空间跟踪算法可知, U_1满足

$$\left.\frac{\partial J}{\partial W}\right|_{W=U_1}=(-2C+\varXi U_1U_1^{\mathrm{H}}+U_1U_1^{\mathrm{H}}\varXi)U_1=0.$$

定义参数估计误差

$$\tilde{U}_1(t)\triangleq\hat{U}_1(t)-U_1,$$

和随机Lyapunov函数

$$T(t)\triangleq\|\tilde{U}_1(t)\|^2=\tilde{U}_1(t)\tilde{U}_1^{\mathrm{T}}(t).$$

将参数估计误差$\tilde{U}_1(t)$带入式(16)得

$$\begin{aligned}\tilde{U}_1(t)=&\tilde{U}_1(t-1)-\mu[-2x(t)+x(t)+\\&U(t-1)U^{\mathrm{H}}(t-1)x(t)]U(t-1)x^{\mathrm{H}}(t).\end{aligned}$$

根据文献[4]定理5, 当步长μ满足$0<\mu<1$时, $T(t)$几乎肯定(a.s.)收敛于一有界随机变量T_0. 当持续激励条件A2)满足时, 梯度算法的参数估计误差$\tilde{U}_1(t)$一致收敛于零, 则有

$$\lim_{t\to\infty}\hat{\varGamma}_i^{\beta}(t)\to\varGamma_i^{\beta}\quad\text{a.s.}$$

证毕.

由上述分析可知, 广义能观测矩阵$\varGamma_i$的更新作为递推子空间算法的一个重要步骤, 使用该算法满足收敛性要求.

4 仿真算例(Simulation example)

通过一仿真算例验证算法的有效性, 考虑如下状态空间方程描述系统的辨识问题:

$$x(t+1)=0.6\begin{bmatrix}\cos\alpha(t)-\sin\alpha(t)\\\sin\alpha(t)\ \ \cos\alpha(t)\end{bmatrix}x(t)+B(t)u(t),$$
$$y(t)=C(t)x(t)+D(t)u(t).$$

其中: $\alpha(t)=-\pi/3$, $C(t)=[1\,0]$, $D(t)=0$.

仿真中辨识参数选择如下: $N=3$, $\beta=0.95$, $\mu=0.99$, $\gamma=1/\log(t+1)$, 辨识结果如图1所示. 因为辨识模型参数与原系统矩阵满足在相似变换内等价性, 因此本文使用状态矩阵A的特征值作为比较标准, 以验证算法的有效性. 如图1所示, 随着数据的不断增加, 递推子空间辨识的精度也不断提高, 并在数据量达到400左右实现了对原系统参数的高精度估计.

5 结论(Conclusion)

本文针对工业系统中广泛存在的时变特性, 将梯度型算法引入到基于遗忘因子的状态子空间跟踪中, 提出一类新的递推子空间辨识算法, 实现对系统状态空间模型的在线递推估. 基于RLS-like遗忘因子的数据矩阵构造机制, 更清晰地描述了不同时刻数据对辨识的影响, 更好地跟踪系统时变特性, 提高了递推估计的收敛速度; 梯度法子空间跟踪技术的应用, 降低了初值选取对参数估计的影响, 避免了收敛的有偏性, 算法计算简单有效; 并理论分析了递推子空间辨识算法的收敛性; 最后通过仿真实例验证了算法的有效性.

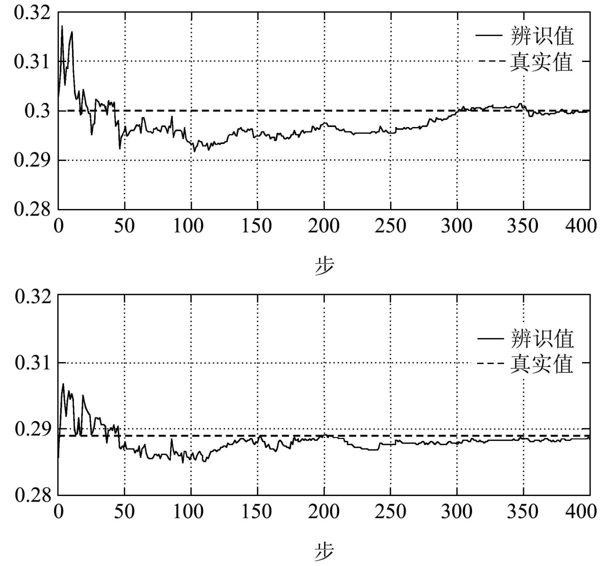

图 1 辨识参数A的特征值

Fig. 1 The eigenvalue of identified matrix A

参考文献(References):

[1] MERCERE G, LECOEUCHE S, LOVERA M. Recursive subspace identification based on instrumental variable unconstrained quadratic optimization[J]. *International Journal of Adapt Control Signal Process*, 2004, 18(4): 771 – 797.

[2] JIANG Y, FANG H. Recursive subspace identification based on principal component analysis[C] // *Proceedings of the 25th Chinese Control Conference*. Harbin, China: [s.n.], 2006.

[3] LOVERA M, GUSTAFSSON T, VERHAEGEN M. Recursive subspace identification of linear and non-linear Wiener state-space models[J]. *Automatica*, 2000, 36(11): 1639 – 1650.

[4] OKU H, KIMURA H. Recursive 4SID algorithms using gradient type subspace tracking[J]. *Automatica*, 2002, 38(6): 1035 – 1043.

[5] ZHANG C, BITMEAD R R. Subspace system identification for training-based MIMO channel estimation[J]. *Automatica*, 2005, 41(9): 1623 – 1632.

[6] YANG B. Asymptotic convergence analysis of the projection approximation subspace tracking algorithms[J]. *Signal Processing*, 1996, 50(1/2): 123 – 136.

作者简介:

杨 华 (1978—), 女, 2002年于山东大学获硕士学位, 2007年7月于上海交通大学自动化研究所取得博士学位, 现为中国海洋大学信息科学与工程学院讲师, 主要研究方向为空间方法系统辨识、预测控制、自适应控制等, E-mail: hyang@ouc.edu.cn;

李少远 (1965—), 男, 1997 年于南开大学获得博士学位, 1998年3月至2000年2月在上海交通大学从事博士后研究工作, 现为上海交通大学自动化系教授, 博士生导师, 研究领域为自适应预测控制、优化方法与优化控制等, E-mail: syli@sjtu.edu.cn.