第 24 卷 第 2 期
1998 年 3 月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 24, No. 2
March., 1998

最差情况 H_∞ 辨识的时域设计方法[1)]

郑立辉 冯 珊
(华中理工大学系统工程研究所 武汉 430074)

潘德惠
(东北大学工商管理学院 沈阳 110006)

摘 要 针对最差情况 H_∞ 辨识中最常见的一类模型集,给出了一种基于时域数据的两步辨识算法.第一步通过信息一致性原理把辨识问题转化为受限凸规划问题,第二步利用一个多项式逼近定理在第一步结果基础上,对不确定集中的系统进行逼近,得到辨识出的名义模型.最后分析了辨识算法的局部误差、全局误差和算法的收敛性.

关键词 最差情况 H_∞ 辨识,时域数据,信息一致性,多项式逼近

1 引言

九十年代以来,鲁棒辨识作为针对鲁棒控制而进行的辨识越来越受到人们的重视.最差情况系统辨识是鲁棒辨识的重要组成部分.最差情况系统辨识就是先假定由先验信息已知被辨识系统包含于某一模型集内,然后利用关于系统输入输出的实验数据,估计出被辨识系统的名义模型及其与先验模型集中所有系统偏差的上界,并且这一界限应该与鲁棒控制理论中刻划扰动量大小的度量(如 H_∞ 范数,l^1 范数等)相一致.

本文针对最差情况 H_∞ 辨识中最常见的一类模型集,给出了一种基于时域数据的辨识算法.目前,该模型集的辨识算法基本上是针对频域数据的.在时域数据方面,Zames 等对无观测噪声的情况进行了研究,给出了该模型集的辨识 n-宽度[1].在有噪声情况下,该模型集的时域辨识算法直到最近才由 Chen 等给出[2].Chen 等的算法首先把辨识问题归结为受限凸规划问题,然后用 Schur 算法得到具有内插性质的辨识模型.本文提出的辨识算法也包含两个步骤:第一步的思路与 Chen 等算法的第一步相同.第二步基于一个逼近定理,在第一步结果的基础上对“不确定集”中的系统进行逼近,得到辨识模型.这种算法不仅比 Chen 等的算法简单,而且大大地减少了辨识误差的上界.当观测噪声足够小时本算法是最优的.

2 问题的描述

对于任意正实数 ρ,记 $D_\rho := \{z\in C: |z|<\rho\}$,并定义赋范线性空间

$$H_{\infty,\rho} := \{\hat{h}: \hat{h} \text{ 在 } D_\rho \text{ 内解析,且} \|\hat{h}\|_{\infty,\rho} := \sup_{|z|<\rho} |\hat{h}(z)| < \infty\}.$$

当 $\rho=1$ 时省略 ρ,即 $H_{\infty,1}=H_\infty$, $\|\cdot\|_{\infty,1}=\|\cdot\|_\infty$. 记 $H_+ := \bigcup_{\rho>1} H_{\infty,\rho}$. 对于 $M\geqslant 0$, $\rho\geqslant 1$,定义集合

$$\overline{B}H_{\infty,\rho}(M) := \{\hat{h}\in H_{\infty,\rho}: \|\hat{h}\|_{\infty,\rho}\leqslant M\}.$$

(1) 中国博士后科学基金资助项目.
收稿日期 1995-12-14

用小写字母(如 $\boldsymbol{x}$)表示集合 l^∞ 中的元素 $x(0),x(1),\cdots,x(n),\cdots$. $\boldsymbol{x}$ 的 Z-变换记为 $\hat{x}$,其前 n 项构成的向量记为 $\boldsymbol{x}_n:=[x(0),x(1),\cdots,x(n-1)]^T$,定义截断算子 $T_n:l^\infty\to R^n$ 为 $T_n\boldsymbol{x}:=\boldsymbol{x}_n$,$\boldsymbol{x}_n$ 的下三角 Toeplitz 矩阵定义为

$$X:=\begin{bmatrix} x(0) & 0 & \cdots & 0 \\ x(1) & x(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ x(n-1) & x(n-2) & \cdots & x(0) \end{bmatrix}. \tag{1}$$

考虑稳定的单输入单输出、线性时不变分布参数系统的辨识问题. 此类系统的传递函数为 H_+ 中的元素,并可以写成

$$\hat{h}(z)=\sum\nolimits_{k=0}^{\infty}h(k)z^k$$

的形式,其中 $\{h(k)\}_{k=0}^{\infty}$ 是系统的单位脉冲响应. 假定由先验信息已知被辨识系统在模型集 $\overline{B}H_{\infty,\rho}(M)$ 内,观测噪声 $\boldsymbol{v}\in l^\infty$ 满足

$$\boldsymbol{v}\in\overline{B}l_\infty(\varepsilon):=\{\boldsymbol{v}\in l^\infty:|v(k)|\leqslant\varepsilon,k=0,1,\cdots\}.$$

给系统加入已知的输入信号 $\boldsymbol{u}\in l^\infty$,对系统输出 $\boldsymbol{y}\in l^\infty$ 观测 n 次,则有

$$y(t)=\sum\nolimits_{k=0}^{t}h(t-k)u(k)+v(t),\quad t=0,1,\cdots,n-1.$$

系统的这些时域输入输出实验数据可以写成

$$\boldsymbol{y}_n=U\boldsymbol{h}_n+T_n\boldsymbol{v},\qquad \boldsymbol{v}\in\overline{B}l_\infty(\varepsilon), \tag{2}$$

其中 U 是由式(1)定义的 $\boldsymbol{u}_n$ 的 Toeplitz 矩阵. 当输入 $\boldsymbol{u}$ 固定时,系统的所有可能输出构成集合

$$Y:=\{\boldsymbol{y}:\boldsymbol{y}_n=UT_n\boldsymbol{h}+T_n\boldsymbol{v},\hat{h}\in\overline{B}H_{\infty,\rho}(M),\boldsymbol{v}\in\overline{B}l_\infty(\varepsilon)\}.$$

在已知 $\boldsymbol{u}$ 和 $\boldsymbol{y}_n$ 时,定义如下不确定集.

$$\hat{P}(\boldsymbol{y}_n):=\{\hat{h}:\boldsymbol{y}_n=UT_n\boldsymbol{h}+T_n\boldsymbol{v},\hat{h}\in\overline{B}H_{\infty,\rho}(M),\boldsymbol{v}\in\overline{B}l_\infty(\varepsilon)\},$$

它是所有既与先验信息相符,又和后验观测数据相一致的被辨识系统的集合.
记 $P_n(\boldsymbol{y}_n):=\{T_n\boldsymbol{h}:\hat{h}\in\hat{P}(\boldsymbol{y}_n)\}$.

最差情况 H_∞ 辨识问题就是在已知输入 $\boldsymbol{u}$ 的条件下,求一辨识算法,即从实验数据 $\boldsymbol{y}_n$ 到 H_+ 空间的映射 $A_n(\boldsymbol{y}_n)\in H_+$,同时给出此辨识算法的最差情况辨识误差,并保证算法是收敛的.

定义 1. 对于任意输出 $\boldsymbol{y}_n\in Y$,辨识算法 A_n 关于 $\boldsymbol{y}_n$ 的最差情况局部误差定义为

$$e(A_n;\boldsymbol{y}_n;\varepsilon,M,\rho):=\sup_{\hat{h}\in\hat{P}(\boldsymbol{y}_n)}\|\hat{h}-A_n(\boldsymbol{y}_n)\|_\infty,$$

全局误差定义为

$$e(A_n;\varepsilon,M,\rho):=\sup_{\boldsymbol{y}_n\in Y}e(A_n;\boldsymbol{y}_n;\varepsilon,M,\rho)$$

定义 2. 若辨识算法 A_n 的全局辨识误差满足

$$\lim_{\substack{\varepsilon\to 0\\ n\to\infty}}e(A_n;\varepsilon,M,\rho)=0,$$

则称算法 A_n 是收敛的.

3　准备知识

引理 1[3]. 设 $\hat{p}_n=p(0)+p(1)z+\cdots+p(n-1)z^{n-1}$. 在模型集 $\overline{B}H_{\infty,\rho}(M)$ 中存在形如

$$\hat{h}(z)=\sum_{k=0}^{n-1}p(k)z^k+\sum_{k=n}^{\infty}h(k)z^k\in\overline{B}H_{\infty,\rho}(M)$$

的元素,即 $\hat{h}$ 的前 n 项系数与 $\hat{p}_n$ 的系数相同的充要条件是矩阵

$$\Gamma_\rho(\boldsymbol{p}_n):=\begin{bmatrix}p(n-1)\rho^{n-1} & p(n-2)\rho^{n-2} & \cdots & p(0)\\ p(n-2)\rho^{n-2} & p(n-3)\rho^{n-3} & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ p(0) & 0 & \cdots & 0\end{bmatrix}\tag{3}$$

的最大奇异值不超过 M,即 $\sigma_{\max}(\Gamma_\rho(\boldsymbol{p}_n))\leqslant M$.

引理 2[4]. 设 D_n 是所有 n 阶多项式构成的集合,且 $M\geqslant0,\rho\geqslant1$. 则

$$\sup_{\hat{h}\in\overline{B}H_{\infty,\rho}(M)}\ \inf_{\hat{p}_n\in D_n}\|\hat{h}-\hat{p}_n\|_\infty=M\rho^{-n}.$$

另外,对任意的 $\hat{h}\in\overline{B}H_{\infty,\rho}(M)$,若定义线性算子 B_ρ^n,

$$B_\rho^n\hat{h}:=\sum_{k=0}^{n-1}(1-\rho^{2(k-n)})h(k)z^k,\quad \hat{h}\in\overline{B}H_{\infty,\rho}(M),\tag{4}$$

则有

$$\sup_{\hat{h}\in\overline{B}H_{\infty,\rho}(M)}\|\hat{h}-B_\rho^n\hat{h}\|_\infty=M\rho^{-n}.$$

在最差情况系统辨识的误差分析中,IBC(Information-Based Complexity)理论占有十分重要的地位.下面的结果是经常用到的.

定义 3. 在 n 维赋范空间$(R^n,\|\cdot\|)$中,集合 $P_n(\boldsymbol{y}_n)$的局部直径定义为

$$d(P_n(\boldsymbol{y}_n),\|\cdot\|):=\sup_{\boldsymbol{x}_n,\boldsymbol{z}_n\in P_n(\boldsymbol{y}_n)}\|\boldsymbol{x}_n-\boldsymbol{z}_n\|,$$

全局直径定义为

$$d^*(\|\cdot\|):=\sup_{\boldsymbol{y}_n\in Y}d(P_n(\boldsymbol{y}_n),\|\cdot\|).$$

由下三角矩阵的性质,当输入 $\boldsymbol{u}$ 满足 $u(0)\neq0$ 时,$\boldsymbol{u}_n$ 的 Toeplitz 矩阵 U 的逆阵 U^{-1}可写成

$$U^{-1}=\begin{bmatrix}w(0) & 0 & \cdots & 0\\ w(1) & w(0) & \cdots & 0\\ \vdots & \vdots & \ddots & 0\\ w(n-1) & w(n-2) & \cdots & w(0)\end{bmatrix}\tag{5}$$

的形式,记

$$h_u(k):=\min\left\{\sum\nolimits_{j=0}^{k}[w(k-j)y(j)+|w(j)|\varepsilon],M\rho^{-k}\right\},$$

$$h_l(k):=\max\left\{\sum\nolimits_{j=0}^{k}[w(k-j)y(j)+|w(j)|\varepsilon],-M\rho^{-k}\right\},\quad k=0,1,\cdots,n-1.$$

引理 3[5]. 取 n 维欧氏空间中的范数为 $\|\cdot\|_1$,$\|\boldsymbol{x}_n\|_1=\sum\nolimits_{k=0}^{n-1}|x(k)|$. 在$u(0)\neq0$的条件下,对任意 $\boldsymbol{y}_n\in Y$,集合 $P_n(\boldsymbol{y}_n)$ 的局部直径为

$$\mathrm{d}(P_n(\boldsymbol{y}_n),\|\cdot\|_1)=\sum\nolimits_{k=0}^{n-1}[h_u(k)-h_l(k)],$$

全局直径为

$$\mathrm{d}^*(\|\cdot\|_1)=2\sum\nolimits_{k=0}^{n-1}\min\left[\varepsilon\sum\nolimits_{j=0}^{k}|w(j)|,M\rho^{-k}\right].$$

4 主要结果

在最差情况系统辨识中,如果不确定集 $\hat{P}(\boldsymbol{y}_n)\neq\varnothing$,即存在与先验信息和后验观测数据相符的系统,则称先验信息 M,ρ,ε 与后验观测数据 $\boldsymbol{y}_n$ 是一致的.

定理 1. 在输入 $\boldsymbol{u}$ 满足 $u(0)\neq 0$ 的条件下,系统的先验信息 M,ρ,ε 与后验观测数据 $\boldsymbol{y}_n$ 一致的充要条件是,存在向量 $\boldsymbol{h}_n$ 使

$$a)\quad \sigma_{\max}(\Gamma_\rho(\boldsymbol{h}_n))\leqslant M;\quad b)\quad (\boldsymbol{y}_n-U\boldsymbol{h}_n)_k\leqslant\varepsilon,\quad k=0,1,\cdots,n-1$$

成立. 其中 $\Gamma_\rho(\boldsymbol{h}_n)$ 由式(3)定义,$(\boldsymbol{x}_n)_k$ 表示向量 $\boldsymbol{x}_n$ 的第 k 个分量.

证明. 根据引理 1 及不确定集的定义知定理成立.

基于这一定理及引理 2,构造如下时域 H_∞ 辨识两步算法(简称 TDHII):

第一步. 求解受限凸规划问题

$$\min_{\boldsymbol{h}_n}\sigma_{\max}(\Gamma_\rho(\boldsymbol{h}_n)),\qquad \text{s.t.}\ (\boldsymbol{y}_n-U\boldsymbol{h}_n)_k\leqslant\varepsilon,\quad k=0,1,\cdots,n-1.$$

如果所得最优解 $\boldsymbol{h}_n^*$ 对应的目标函数值大于 M,则说明先验信息不正确,需要重新确定 M,ρ,ε 的值. 否则进入下一步.

第二步. 用(4)式定义的线性算子 B_ρ^n 作用 $\boldsymbol{h}_n^*$ 对应的多项式,得到辨识出的名义模型

$$\hat{h}^{id}=B_\rho^n\hat{h}_n^*=\sum_{k=0}^{n-1}(1-\rho^{2(k-n)})h^*(k)z^k.$$

定理 2. 在 $u(0)\neq 0$ 的条件下,对任意 $\boldsymbol{y}_n\in Y$,TDHⅡ 算法的局部误差满足

$$e(A_n;\boldsymbol{y}_n;M,\rho,\varepsilon)\leqslant\sum_{k=0}^{n-1}(1-\rho^{2(k-n)})[h_u(k)-h_l(k)]+M\rho^{-n}.$$

证明. 由引理 2、引理 3 及 H_∞ 与 l^1 空间范数之间的关系 $\|\cdot\|_\infty\leqslant\|\cdot\|_l$,得

$$\begin{aligned}e(A_n;\boldsymbol{y}_n;M,\rho,\varepsilon)&=\sup_{\hat{h}\in\hat{P}(\boldsymbol{y}_n)}\|\hat{h}^{id}-\hat{h}\|_\infty\leqslant\sup_{\hat{h}\in\hat{P}(\boldsymbol{y}_n)}[\|\hat{h}^{id}-B_\rho^n\hat{h}\|_\infty+\|B_\rho^n\hat{h}-\hat{h}\|_\infty]\\&\leqslant\sup_{\hat{h}\in\hat{P}(\boldsymbol{y}_n)}\|B_\rho^n(\hat{h}_n^*-\hat{h})\|_1+M\rho^{-n}\leqslant\sum_{k=0}^{n-1}(1-\rho^{2(k-n)})[h_u(k)-h_l(k)]\\&\quad+M\rho^{-n}.\end{aligned}$$

定理 3. 在 $u(0)\neq 0$ 的条件下,TDHII 算法的全局误差满足

$$e(A_n;M,\rho,\varepsilon)\leqslant 2\sum_{k=0}^{n-1}(1-\rho^{2(k-n)})\min\Big[\varepsilon\sum_{j=0}^{k}|w(j)|,M\rho^{-k}\Big]+M\rho^{-n},\tag{6}$$

从而算法是收敛的.

证明. 利用引理 3,由定理 2 的证明过程及定义 2 知定理成立.

由式(6)看出,最差情况辨识误差的上界(记为 b^*)由前后两项构成,前一项(记为 b_1)是由观测噪声所引起的不确定性,后一项(记为 b_2)是由于对无穷维辨识对象只进行了有限次观测所造成的截断误差. 当 $\varepsilon\to 0$ 时,$b_1\to 0$,辨识误差的上界等于 $M\rho^{-n}$,这与 Zames 等[1]在无观测噪声假设下得到的辨识 n-宽度相一致. 系统的辨识 n-宽度是辨识算法基于 n 次观测所能达到的最小辨识误差,因此当噪声足够小时本算法是最优的.

在有噪声情况下,本算法的误差上界也明显小于 Chen 等[2]的算法. 设 l 是使 $\varepsilon\sum_{j=0}^{k}|w(j)|<M\rho^{-k}$ 成立的整数 k 中的最大值,取 $l^*=\min\{l,n-1\}$,记

$$q(k) = (1/\rho^{2n})\sum_{j=k}^{l^*}\rho^{2j}, \qquad k = 0,1,\cdots,l^*.$$

则全局误差上界可化为

$$\begin{aligned} b^* &= 2\varepsilon\sum_{k=0}^{l^*}(1-\rho^{2(k-n)})\sum_{j=0}^{k}|w(j)| + 2\sum_{k=l^*+1}^{n-1}(1-\rho^{2(k-n)})M\rho^{-k} + M\rho^{-n} \\ &= 2\varepsilon\sum_{j=0}^{l^*}\sum_{k=j}^{l^*}(1-\rho^{2(k-n)})|w(j)| + 2M\left(\sum_{k=l^*+1}^{n-1}\rho^{-k} - \frac{1}{\rho^{2n}}\sum_{k=l^*+1}^{n-1}\rho^{k} + \frac{1}{2}\rho^{-n}\right) \\ &= 2\varepsilon\sum_{k=0}^{l^*}\sum_{k=j}^{l^*}[(l^*+1-k)-q(k)]|w(k)| \\ &\quad + 2\left[\frac{M}{\rho^{l^*}(\rho-1)} - \left(\frac{3+\rho}{2} - \rho^{l^*+1-n}\right)\frac{M}{\rho^n(\rho-1)}\right]. \end{aligned}$$

由于 $q(k)\geqslant 0,\rho\geqslant 1$，上式中前后两项分别小于等于文献[2]中(4.5)式误差上界的前后两项，所以本算法具有更加紧的误差上界.

比较定理 2 和文献[2]的定理 4.2 的证明过程可以看出，造成辨识误差上界减小的原因在于：1) 本算法第二步在凸规划所得结果 h_n^* 的基础上，对不确定集 $\hat{P}(y_n)$ 中以 $\hat{h}_n^*$ 为前 n 项的系统进行了“全局”逼近，使有限次观测所引起的最差情况截断误差 b_2 等于 $M\rho^{-n}$，而内插算法的最差情况偏差上界为 $M/[\rho^{n-1}(\rho-1)]$；2) 误差上界中由观测噪声引起的不确定性 b_1 取决于集合 $P_n(y_n)$ 的直径，本算法第二步在 $\hat{h}_n^*$ 的系数上分别乘了一个小于 1 的数，从而使这种不确定性由比 $P_n(y_n)$ 直径更小的集合来确定，进而减小了观测噪声所引起的不确定性. 本算法第二步的作用在 ρ 接近 1 时表现得最为明显. 当 $\rho\to 1$ 时，Chen 等算法的误差上界是发散的. 由于 $\lim\limits_{\rho\to 1} q(k) = l^*+1-k$，因而

$$\lim_{\rho\to 1} b^* = 2\lim_{\rho\to 1}\left[\frac{M}{\rho^{l^*}(\rho-1)} - \left(\frac{3+\rho}{2} - \rho^{l^*+1-n}\right)\frac{M}{\rho^n(\rho-1)}\right] = M.$$

可见，本算法解决了 ρ 接近 1 时文献[2]中辨识算法误差上界无限增大的问题.

5 结语

本文讨论了一类模型集基于时域数据的最差情况 H_∞ 辨识问题. 本文提出的辨识算法与其它基于凸规划的算法一样，存在着计算量较大的问题，寻找快速有效的凸规划解法是具有重要意义的研究课题.

参 考 文 献

1 Zames G, Lin L, Wang L Y. Fast identification n-widths and uncertainty principles for LTI and slowly varying systems. *IEEE Trans. Automatic Control*, 1994, **39**(9): 1827—1838

2 Chen J, Nett C N. The Caratheodory-Fejer problem and H^∞/l^1 identification: a time domain approach. *IEEE Trans. Autom. Control*, 1995, **40**(4): 729—735

3 Adamyn V M, Arov D Z, Krein M G. Analytic properties of schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem. *Math. Sbornik.*, 1971, **15**(1): 31—73

4 Pinkus A. n-Widths in Approximation Theory. Berlin: Springer-Verlag, 1985

5 Chen J, Nett C N, Fan M K H. Optimal non-parametric identification from arbitrary corrupt finite time series. *IEEE Trans. Autom. Control*, 1995, **40**(4): 768—776

A TIME DOMAIN APPROACH TO THE WORST CASE SYSTEM IDENTIFICATION IN H_∞

ZHENG Lihui　FENG Shan

(*Institute of Systems Engineering, Huazhong Univ. of Science and Technology, Wuhan* 430074)

PAN Dehui

(*Faculty of Business Administration, Northeastern University, Shenyang* 110006)

Abstract This paper presents a two-step algorithm for the worst case H_∞ identification of a class of well-known model set with time domain experimental data. Using the information consistency principle, the first step of the algorithm transforms the identification problem into a constrained convex programming, the result of which is then used, in the second step, to approximate systems in the uncertainty set to obtain the identified nominal model based on a polynomial approximation theorem. Discussions on the local and global identification errors and the convergence of the algorithm are also carried out respectively.

Key words Worst case identification in H_∞, time domain experimental data, information consistency, polynomial approximation

郑立辉　男,1968 年生. 1992 年于沈阳大学获经济学学士学位. 同年考入东北大学工商管理学院攻读硕士学位,1994 年直读该校自动控制系博士学位,并于 1997 年获得工学博士学位. 现为华中理工大学系统工程研究所博士后研究人员. 主要研究兴趣是鲁棒控制与辨识及其在经济和金融建模中的应用.

冯　珊　女, 1933 年生. 1955 年毕业于清华大学. 现为华中理工大学教授、博士生导师. 现在的研究领域是经济系统的智能建模与仿真.

潘德惠　男,1928 年生. 1949 年毕业于东北大学理学院. 现为东北大学工商管理学院教授,自动控制系控制理论与应用专业博士生导师. 中国数学学会理事. 曾多年从事应用数学的研究,现在的研究领域是分布参数系统的模型辨识与最优控制.