

随机系统的递推辨识: 从个例到一般框架

赵文斌[†], 陈翰馥

(中国科学院 数学与系统科学研究院 系统控制重点实验室, 北京 100190;

国家数学与交叉科学研究中心, 北京 100190)

摘要: 本文考察Hammerstein系统、Wiener系统和非线性带外源输入的自回归系统(autoregressive system with exogenous input, ARX)等常见的随机非线性系统的递推辨识和因特网PageRank的分布式、随机化算法. 对非线性系统分别构造递推辨识算法, 证明了估计的强一致性; 对因特网PageRank的分布式、随机化算法, 给出估计的强一致性和收敛速度; 在此基础上, 总结了这类问题的统一处理框架—将辨识(估计)问题转化为函数求根、进而基于随机逼近构造算法得到强一致的递推辨识; 最后, 通过数值例子来验证算法的有效性.

关键词: Hammerstein系统; Wiener系统; 非线性ARX系统; PageRank; 随机逼近

中图分类号: TP273 文献标识码: A

Recursive identification of stochastic systems: from individual system to a general framework

ZHAO Wen-xiao[†], CHEN Han-fu

(Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China;
National Center for Mathematics and Interdisciplinary Sciences, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: We consider the recursive identification algorithms for the stochastic nonlinear systems, such as Hammerstein system, Wiener system, and nonlinear autoregressive system with exogenous input (ARX) system and the distributed randomized PageRank algorithm (DRPA). The strong consistency of estimates given by the identification algorithms and DRPA is established. Based on the stochastic approximation algorithm with expanding truncations, a unified framework for solving these kinds of problems is introduced. Numerical examples are given to testify the performance of the proposed algorithms.

Key words: Hammerstein system; Wiener system; nonlinear ARX system; PageRank; stochastic approximation

1 引言(Introduction)

系统辨识利用系统的输入-输出数据来建立系统数学模型, 经过几十年的发展, 线性系统的辨识理论已日臻完善, 并在工程中得到成功应用^[1-4]. 实际系统的动态特性大多是时变的、非线性的, 同时系统本身往往存在随机因素或测量带有噪声, 如化工过程中的蒸馏塔、pH滴定、生物控制论中的视网膜成像过程、分子生物网络中的时变和开关现象、卫星通讯信道建模等^[5-9]. 因而从随机、非线性的角度来研究辨识领域的一些问题, 有科学意义和工程价值. 国际自动控制联合会(IFAC)在2008年和2011年举行的连续两届世界大会上, 均安排了非线性系统辨识方向的大会报告(plenary talk)^[10-11]. 当数据增多时, 需要适时地修正估计, 所以要考虑系统的递推辨识.

非线性的表示形式非常多样, 一种循序渐近的方法

是首先从具有较强工程背景的特殊非线性系统入手, 逐步深入, 进而考虑更为一般的非线性系统辨识.

由静态非线性环节和动态线性子系统串联而成的随机非线性系统(如Hammerstein系统、Wiener系统等)在工程中有广泛应用^[5, 8-9], 其辨识问题得到广泛关注, 很多国内外学者深入开展了这方面的研究工作, 理论成果越来越多^[12-18]. 一般说来, Hammerstein系统和Wiener系统中的非线性环节是静态的, 近年来, 对形式更为一般且包含动态非线性环节的随机系统, 如非线性ARX系统(nonlinear ARX, NARX)等, 其辨识问题开始得到关注^[19-20].

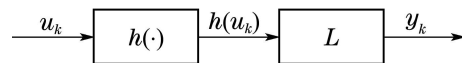


图1 Hammerstein系统

Fig. 1 Hammerstein system

收稿日期: 2014-02-24; 录用日期: 2014-05-20.

[†]通信作者. E-mail: wxzhao@amss.ac.cn; Tel.: +86 10-6261 1460.

基金项目: 国家重点基础研究发展计划“973”计划资助项目(2014CB845301); 国家自然科学基金资助项目(61104052, 61273193, 61227902, 61134013, 61120106011).

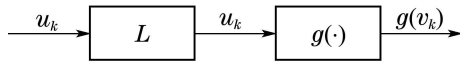


图2 Wiener系统

Fig. 2 Wiener system

从方法上看, Hammerstein 系统、Wiener 系统和 NARX 系统的辨识方法大致可分为迭代算法、极大似然算法、频率域算法、基于优化的算法和盲辨识算法等. 迭代辨识算法主要针对系统中的非线性函数为已知基函数的线性组合, 在给定目标函数及数据集的基础上, 对线性子系统和非线性函数中含有的两组未知参数分别迭代优化, 直至估计序列收敛^[13,21]; 极大似然算法主要借鉴统计、时间序列分析等领域的基本思想, 利用诸如极大化似然函数之类的方法来辨识系统^[14]; 频率域算法利用不同频率的三角函数作为系统输入, 并结合函数的傅里叶级数展开和信号的傅里叶变换来辨识线性子系统和非线性函数^[22-23]; 基于优化的辨识算法主要思想是将辨识转化为优化问题, 特别是凸优化, 然后利用优化领域许多成熟有效的数值算法来求解相关的最优化问题^[20]; 盲辨识主要针对一些实际系统只可以获得输出信号, 而输入信号未知的情形, 利用盲辨识的思想来对 Hammerstein, Wiener 系统进行辨识的文献见 [24-25] 等. 总体上, 这些方法是基于固定样本容量的优化算法, 不容易实现递推计算, 且算法是否具有强一致收敛性尚不清楚; 此外, 非线性系统辨识中所采用的方法尚未与线性系统辨识一样形成较为统一的研究框架和处理手段.

近年来, 一些应用性、交叉性的课题如主分量分析、因特网网页排序算法 PageRank、多自主体系统等, 得到越来越多不同领域研究人员的关注, 这些问题也吸引自动化领域的研究人员从系统与控制的角度进行探索. 文献 [26] 从分布式、随机化的角度研究了 PageRank 算法, 证明了估计序列在均方意义下收敛到真实的 PageRank 值. 关于算法的进一步性质, 如强一致收敛性、收敛速度等, 尚不完全清楚; 此外, 对非线性系统(如 Hammerstein 系统、Wiener 系统和 NARX 系统)的辨识和网络相关的建模与估计(如 PageRank), 看似差异很大的两类问题, 是否能找到建立起彼此联系的研究桥梁?

文献 [17, 27-29] 对单输入-单输出 (single input single output, SISO) Hammerstein 系统、Wiener 系统的辨识问题进行了研究, 对系统中的静态非线性环节为参数化建模和非参数建模都给出递推辨识算法并证明估计的强一致性; 文献 [30] 构造基于扩张截尾的随机逼近算法, 结合马氏链的平稳遍历理论首次给出非线性 ARX 系统强一致的递推辨识, 给这方面的进一步研究引入新思路; 经过简单推导可以发现^[26]中的分布式、随机化 PageRank 算法实为递推的, 文献 [31] 结合

随机逼近算法的相关理论进一步证明了分布式、随机化 PageRank 算法的强一致性、并得到收敛速度. 下面简要介绍上述论文中的主要成果, 基于此, 从递推的角度给出处理这类问题的较为统一的研究框架和处理手段: 首先将辨识、估计问题转化为求回归函数的零点, 然后构造随机逼近算法得到递推估计, 最后分析和处理噪声、进而得到强一致的递推辨识.

论文具体组织结构如下: 第 2 节分别介绍 Hammerstein 系统、Wiener 系统和 NARX 系统的递推辨识及相关结论, 第 3 节介绍分布式、随机化 PageRank 算法的强一致性, 第 4 节基于随机逼近算法, 给出上述问题的一个统一处理框架, 第 5 节做数值模拟以检验算法的有效性, 最后, 第 6 节作总结与讨论.

数学符号: 基本概率空间记为 $(\Omega, \mathcal{F}, \mathbb{P})$, $E(\cdot)$ 为期望算子, z 为后移算子, 即 $zx_k = x_{k-1}$. 取值于 $(\mathbb{R}^m, \mathcal{B}^m)$ 中的马氏链 $\{X_k\}_{k \geq 0}$ 第 k 步转移概率和边缘概率分布函数分别记为 $P_k(x, A) = \mathbb{P}\{X_k \in A | X_0 = x\}$ 及 $P_k(A) = \mathbb{P}\{X_k \in A\}$, $x \in \mathbb{R}^m$, $A \in \mathcal{B}^m$. 假若马氏链的不变概率分布存在, 则记为 $P_{IV}(\cdot)$. 对符号测度 $\nu(\cdot)$, 其全变差范数记为 $\|\nu\|_{\text{var}}$. 对向量 $x \in \mathbb{R}^m$, 其欧几里得范数和其他类型的范数分别记为 $\|x\|$ 和 $\|x\|_{\nu}$. $I_A(\cdot)$ 为集合 A 的示性函数, 即 $I_A(\omega) = 1$, $\omega \in A$; $I_A(\omega) = 0$, $\omega \notin A$. 矩阵 M 的转置记为 M^T , 向量 $\mathbf{1} = (1, \dots, 1)^T \in \mathbb{R}^n$. 对整数 p 和 q , 定义 $p \vee q = \max\{p, q\}$, $p \wedge q = \min\{p, q\}$.

2 典型随机非线性系统的递推辨识 (Recursive identification of several classes of stochastic nonlinear systems)

从对 Hammerstein 系统、Wiener 系统、NARX 系统中非线性函数的刻画来看, 辨识方法可分为参数化辨识^[14, 21-25, 28]和非参数化辨识^[15, 17, 19-20, 29-30, 32]. 在参数化辨识中, 非线性函数常表示为基函数的组合, 如多项式、样条, 也包括饱和、死区非线性等, 从而辨识转化为对未知参数的估计; 在非参数化辨识中, 通常直接估计非线性函数在给定点的函数值. 一般说来, 参数化辨识需要系统的某些先验信息, 在这些信息得不到或不容易验证的情况下, 非参数辨识是可以采用的手段, 此外, 非参数辨识也有直接的工程和物理意义^[33-34]. 本节内容属非参数辨识.

2.1 Hammerstein 系统的递推辨识 (Recursive identification of Hammerstein systems)

考虑线性子系统为无穷脉冲响应 (infinite impulse response, IIR) 的单输入单输出 (SISO) Hammerstein 系统:

$$y_{k+1} = a_1 y_k + \dots + a_p y_{k+1-p} + h(u_k) + b_1 h(u_{k-1}) + \dots + b_q h(u_{k-q}) + \varepsilon_{k+1}, \quad (1)$$

$$z_{k+1} = y_{k+1} + \xi_{k+1}, \quad (2)$$

其中: u_k, y_k 分别为系统的输入和输出信号, ξ_k 为系统输出端的量测噪声. 这类系统的非参数辨识就是基于测量数据 $\{u_k, z_k\}_{k \geq 0}$ 构造递推算法, 估计线性系统的参数 $(a_1, \dots, a_p, b_1, \dots, b_{q-1})$ 和非线性函数 $h(\cdot)$ 在给定点 u^* 的函数值.

线性子系统的特征多项式记为

$$A(z) \triangleq 1 - a_1 z - \dots - a_p z^p, \\ B(z) \triangleq 1 + b_1 z + \dots + b_q z^q.$$

选取输入信号 $\{u_k\}_{k \geq 0}$ 为有界的零均值独立同分布(independent and identically distributed, iid)随机序列, 即 $E u_k = 0$ 且存在常数 c_1 使得 $|u_k| \leq c_1$, u_k 有概率密度函数 $p(\cdot)$, $p(\cdot)$ 在 u^* 连续且 $p(u^*) > 0$.

定义

$$\mu \triangleq E h(u_k), R_1 \triangleq \text{var}\{h(u_k)\}.$$

作如下假设:

A1) $A(z)$ 稳定, 即 $A(z) = 0$ 的根在单位圆周之外, 并且系数 $a_p \neq 0$;

A2) $1 + d_1 + \dots + d_q \neq 0$;

A3) $\{\xi_k\}_{k \geq 0}$ 和 $\{\varepsilon_k\}_{k \geq 0}$ 均为 iid 序列, 且满足 $E \xi_k = E \varepsilon_k = 0, R_2 \triangleq E \xi_k^2 < \infty, R_3 \triangleq E \varepsilon_k^2 < \infty; \{u_k\}, \{\xi_k\}$ 和 $\{\varepsilon_k\}$ 相互独立, $E u_k h(u_k) \neq 0$;

A4) $B(z)B(z^{-1})R_1 + R_2$ 与 $A(z)A(z^{-1})$ 无公共零点;

A5) $h(\cdot)$ 可测、局部有界且在点 u^* 处连续.

记

$$a \triangleq \sum_{i=0}^p a_i, b \triangleq \sum_{i=0}^q b_i, a_0 \triangleq 1, b_0 \triangleq 1,$$

则系统(1)–(2)可以表示为

$$A(z)(z_{k+1} - \frac{b}{a}\mu - \varepsilon_{k+1}) = \\ B(z)(h(u_k) - \mu) + \xi_{k+1}. \quad (3)$$

在上述假设条件下, 系统是渐近平稳的, 所以有以下关系成立:

$$\mu^* \triangleq \lim_{k \rightarrow \infty} E z_k (= \frac{b}{a}\mu), \quad (4)$$

$$\gamma(\tau) \triangleq \lim_{k \rightarrow \infty} E(z_{k+\tau} - \mu^*)(z_k - \mu^*), \tau \geq 0. \quad (5)$$

在式(3)两边同时乘以 $(z_{k-s} - \frac{b}{a}\mu)$, $s \geq p \vee q$ 并取期望, 再令 $k \rightarrow \infty$ 可得

$$\gamma(s+1) + a_1 \gamma(s) + \dots + a_p \gamma(s+1-p) = 0. \quad (6)$$

令 $s = p \vee q, p \vee q + 1, \dots, p \vee q + p - 1$, 可得

$$\begin{bmatrix} \gamma(p \vee q) & \dots & \gamma(p \vee q + 1 - p) \\ \gamma(p \vee q + 1) & \dots & \gamma(p \vee q + 2 - p) \\ \vdots & \ddots & \vdots \\ \gamma(p \vee q + p - 1) & \dots & \gamma(p \vee q) \end{bmatrix}.$$

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} \gamma(p \vee q + 1) \\ \gamma(p \vee q + 2) \\ \vdots \\ \gamma(p \vee q + p) \end{bmatrix}. \quad (7)$$

定义

$$T \triangleq \begin{bmatrix} \gamma(p \vee q) & \dots & \gamma(p \vee q + 1 - p) \\ \gamma(p \vee q + 1) & \dots & \gamma(p \vee q + 2 - p) \\ \vdots & \ddots & \vdots \\ \gamma(p \vee q + p - 1) & \dots & \gamma(p \vee q) \end{bmatrix},$$

$$\Theta \triangleq [a_1 \ a_2 \ \dots \ a_p]^T,$$

$$\Gamma \triangleq [\gamma(p \vee q + 1) \ \gamma(p \vee q + 2) \ \dots \ \gamma(p \vee q + p)]^T.$$

注1 方程 $T\Theta = -\Gamma$ 称为广义 Yule-Walker 方程. 注意到相关函数可以利用测量数据的样本平均来估计, 从上面的分析可见, 假若矩阵 T 非奇异, 那么从线性方程组 $T\Theta = -\Gamma$ 或等价方程组 $T^T T \Theta = -T^T \Gamma$ 就可以得到线性子系统参数 (a_1, \dots, a_p) 的估计, 也即辨识问题转化为求函数的零点.

为行文方便, 后文统称含未知零点的函数为“回归函数”(例如广义 Yule-Walker 方程等).

引理1^[17] 假设A1)A3)–A4)和A5)成立, 则相关函数构成的矩阵 T 非奇异.

下面构造辨识算法, 分为3步: 首先估计线性子系统参数 (a_1, \dots, a_p) , 第2步估计参数 (b_1, \dots, b_q) , 最后估计函数 $h(\cdot)$ 在给定点 u^* 的函数值, 其中第1步是关键.

第1步 (a_1, \dots, a_p) 的估计.

首先给出均值 μ^* 和相关函数 $\gamma(\tau)$ 的递推估计算法:

$$\mu_k^* = (1 - \frac{1}{k})\mu_{k-1}^* + \frac{1}{k}z_k, \quad (8)$$

$$\gamma_k(\tau) = (1 - \frac{1}{k})\gamma_{k-1}(\tau) + \frac{1}{k}(z_k - \mu_{k-\tau}^*)(z_{k-\tau} - \mu_{k-\tau}^*). \quad (9)$$

将矩阵 T 和 Γ 中的相关函数 $\gamma(\tau)$ 用估计值 $\gamma_k(\tau)$ 代替, 相应的矩阵记为 T_k 和 Γ_k , Θ 的递推估计算法如下:

$$\Theta_{k+1} = (\Theta_k - \frac{1}{k}(T_k^T T_k \Theta_k + T_k^T \Gamma_k)) \cdot I_{\|\Theta_k - \frac{1}{k}(T_k^T T_k \Theta_k + T_k^T \Gamma_k)\| \leq M_{\delta_k}}, \quad (10)$$

$$\delta_k = \sum_{j=1}^{k-1} I_{\|\Theta_j - \frac{1}{j}(T_j^T T_j \Theta_j + T_j^T \Gamma_j)\| > M_{\delta_j}}, \quad (11)$$

其中: Θ_0 为任意初值, $\delta_0 = 0, \{M_k\}_{k \geq 0}$ 为实序列满足 $M_k > 0, M_{k+1} > M_k, M_k \xrightarrow[k \rightarrow \infty]{} \infty$.

注2 算法(10)–(11)称为扩展截尾的随机逼近算法, 递推地估计回归函数 $G(\Theta) = -(T^T T \Theta + T^T \Gamma)$ 的零点, 若算法中的示性函数取值恒为1, 则退化为经典的Robins-Monro 算法. 从算法的构造可见, 若 $\|\Theta_k - \frac{1}{k}(T_k^T T_k \Theta_k + T_k^T \Gamma_k)\| \leq$

M_{δ_k} , 则第 $k+1$ 步估计 $\Theta_{k+1} = (\Theta_k - \frac{1}{k}(T_k' T_k \Theta_k + T_k' \Gamma_k))$;
 若 $\|\Theta_k - \frac{1}{k}(T_k' T_k \Theta_k + T_k' \Gamma_k)\| > M_{\delta_k}$, 则第 $k+1$ 步估计 $\Theta_{k+1} = 0$ 并且 $\delta_{k+1} = \delta_k + 1$. 扩展截尾的引入, 实质性地减弱随机逼近算法收敛所需的条件^[36].

记 $\Theta_k = [\theta_k(1) \cdots \theta_k(p)]^T$.

定理 1^[17] 假设 A1), A3)–A5) 成立, 则由递推算法(10)–(11)得到的估计是强一致的,

$$\theta_k(i) \xrightarrow[k \rightarrow \infty]{} a_i \text{ a.s. } i = 1, \dots, p. \quad (12)$$

证 这里仅列出主要思路.

算法(10)可改写为

$$\Theta_{k+1} = (\Theta_k + \frac{1}{k}(G(\Theta_k) + e_{k+1})).$$

$$\beta_{k+1}(i) = (\beta_k(i) - \frac{1}{k}(\beta_k(i) - u_k(z_{k+1+i} + \theta_{k-p}(1)z_{k+i} + \dots + \theta_{k-p}(p)z_{k+1+i-p}))) \cdot I_{\|\beta_k(i) - \frac{1}{k}(\beta_k(i) - u_k(z_{k+1+i} + \theta_{k-p}(1)z_{k+i} + \dots + \theta_{k-p}(p)z_{k+1+i-p}))\| \leq M_{\sigma_k(i)}}, \quad (14)$$

$$\sigma_k(i) = \sum_{j=1}^{k-1} I_{\|\beta_j(i) - \frac{1}{j}(\beta_j(i) - u_j(z_{j+1+i} + \theta_{j-p}(1)z_{j+i} + \dots + \theta_{j-p}(p)z_{j+1+i-p}))\| > M_{\sigma_j(i)}}. \quad (15)$$

定理 2^[17] 假设 A1)A3)–A5) 成立, 则由递推算法(14)–(15)得到的估计是强一致的,

$$\beta_k(i) \xrightarrow[k \rightarrow \infty]{} \rho b_i \text{ a.s. } i = 0, 1, \dots, q. \quad (16)$$

证 与定理 1 类似, 从略.

注意到假设条件 A3) 和 $b_0 = 1$, 从 $\beta_k(i)$ 可得 b_i 的强一致估计, $i = 0, 1, \dots, q$.

第 3 步 $h(u)$ 的估计.

首先定义核函数

$$\eta_{k+1}(u) = (\eta_k(u) - \frac{1}{k}(\eta_k(u) - (w_k(u^*) - 1)(z_{k+1} + \theta_{k-p}(1)z_k + \dots + \theta_{k-p}(p)z_{k+1-p}))) \cdot I_{\|\eta_k(u) - \frac{1}{k}(\eta_k(u) - (w_k(u^*) - 1)(z_{k+1} + \theta_{k-p}(1)z_k + \dots + \theta_{k-p}(p)z_{k+1-p}))\| \leq M_{\lambda_k(u)}}, \quad (19)$$

$$\lambda_k(u) = \sum_{j=1}^{k-1} I_{\|\eta_j(u) - \frac{1}{j}(\eta_j(u) - (w_j(u^*) - 1)(z_{j+1} + \theta_{j-p}(1)z_j + \dots + \theta_{j-p}(p)z_{j+1-p}))\| > M_{\lambda_j(u)}}. \quad (20)$$

定理 3^[17] 假设 A1)–A5) 成立, 则

$$\eta_k(u) \xrightarrow[k \rightarrow \infty]{} h(u^*) - Ef(u_1) \text{ a.s.} \quad (21)$$

$$\eta_k(u) + \mu_k \xrightarrow[k \rightarrow \infty]{} h(u^*) \text{ a.s.} \quad (22)$$

证 与定理 1 类似, 从略.

2.2 Wiener 系统的递推辨识(Recursive identification of Wiener systems)

考虑单输入单输出 Wiener 系统

$$v_{k+1} = c_1 v_k + \dots + c_p v_{k+1-p} + u_k + d_1 u_{k-1} + \dots + d_{q-1} u_{k+1-q}, \quad (23)$$

$$y_{k+1} = g(v_{k+1}) + \varepsilon_{k+1}, \quad (24)$$

其中: u_k, y_k 分别为系统的量测输入和量测输出,

$$I_{\|\Theta_k + \frac{1}{k}(G(\Theta_k) + e_{k+1})\| \leq M_{\delta_k}}, \quad (13)$$

其中:

$$G(\Theta_k) = -(T^T T \Theta_k + T^T \Gamma),$$

$$e_{k+1} = (T^T T - T_k^T T_k) \Theta_k + (T^T \Gamma - T_k^T T_k).$$

对回归函数 $G(\Theta)$ 和噪声 e_k 分别验证扩展截尾随机逼近算法的收敛性条件可得(12).

第 2 步 (b_1, \dots, b_q) 的估计.

经过简单推导可得 $E u_k(y_{k+1+i} + a_1 y_{k+i} + \dots + a_p y_{k+1+i-p}) = \rho b_i, i = 0, \dots, q, \rho \triangleq E u_k h(u_k)$.

以 $G_i(x) = -(x - \rho b_i)$ 为回归函数, 与式(10)–(11)类似地构造扩展截尾随机算法来估计 (b_1, \dots, b_q) , 具体算法由(14)–(15)给出:

$$w_k(u^*) \triangleq \frac{e^{-\left(\frac{u_k - u^*}{b_k}\right)^2}}{\int_{-c_1}^{c_1} e^{-\left(\frac{x - u}{b_k}\right)^2} p(x) dx}, \quad (17)$$

其中: $b_k = 1/k^\delta, \delta \in (0, 1/2)$. 可以验证

$$E w_k(u^*) = 1, E w_k(u^*) h(u_k) \xrightarrow[k \rightarrow \infty]{} h(u^*). \quad (18)$$

与式(10)–(11)(14)–(15)类似, 基于回归函数 $G(x) = -(x - h(u^*)), h(u^*)$ 的递推估计算法由式(19)–(20)给出:

ε_k 为输出端的量测噪声. 这类系统的非参数辨识就是基于测量数据 $\{u_k, y_k\}_{k \geq 0}$ 构造递推算法, 估计线性系统的参数 $(c_1, \dots, c_p, d_1, \dots, d_{q-1})$ 和函数 $g(\cdot)$ 在给定点 v^* 的函数值.

线性子系统的特征多项式记为

$$C(z) \triangleq 1 - c_1 z - \dots - c_p z^p,$$

$$D(z) \triangleq 1 + d_1 z + \dots + d_{q-1} z^{q-1}.$$

选取输入信号 $\{u_k\}_{k \geq 0}$ 为 iid 标准正态随机序列, 即 $u_k \in \mathcal{N}(0, 1)$. 作如下假设:

B1) $C(z)$ 稳定, 即 $C(z) \neq 0, \forall |z| \leq 1$;

B2) $C(z)$ 和 $D(z)$ 无公因子, 且 $c_p \neq 0, d_{q-1} \neq 0$;

B3) $\{\varepsilon_k\}_{k \geq 0}$ 为 iid 序列, $E \varepsilon_k = 0, E \varepsilon_k^2 < \infty$;

$\{\varepsilon_k\}_{k \geq 0}$ 和 $\{u_k\}_{k \geq 0}$ 相互独立;

B4) 初值 $[v_0 \ v_{-1} \ \dots \ v_{1-s} \ u_0 \ u_{-1} \ \dots \ u_{1-s}]^T = 0$, 其中 $s = \max\{p, q\}$;

B5) $|g(v)| \leq c|v|^r + c, \forall v \in \mathbb{R}$, 其中: $c > 0, r > 0$ 为常数.

定义矩阵

$$C \triangleq \begin{bmatrix} c_1 & 1 & & \\ c_2 & 0 & \ddots & \\ \vdots & & \ddots & 1 \\ c_s & \dots & \dots & 0 \end{bmatrix}, D \triangleq \begin{bmatrix} 1 \\ d_1 \\ \vdots \\ d_{s-1} \end{bmatrix}, H \triangleq \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T,$$

则线性子系统可以改写为

$$x_{k+1} = Cx_k + Du_k, \tag{25}$$

$$v_{k+1} = Hx_{k+1}, \tag{26}$$

$$v_{k+1} = HC^{k+1}x_0 + \sum_{j=0}^k HC^j Du_{k-j}. \tag{27}$$

定义 $h_j \triangleq HC^j D, j \geq 0, \sigma_k^2 = \sum_{j=0}^k h_j^2$. 注意到 $\{u_k\}_{k \geq 0}$ 为iid正态序列, 利用Bussgang定理(见文献[14-15]), 可得

$$E[u_{k-j}g(v_{k+1})] = \frac{h_j}{\sigma_k^2} \int_{\mathbb{R}} xg(x) \frac{1}{\sqrt{2\pi}\sigma_k} \exp\{-\frac{x^2}{2\sigma_k^2}\} dx, \tag{28}$$

$$\lim_{k \rightarrow \infty} E[u_{k-j}g(v_{k+1})] = h_j \rho, \tag{29}$$

其中:

$$\rho \triangleq \frac{1}{\sqrt{2\pi}\sigma^3} \int_{\mathbb{R}} xg(x) \exp\{-\frac{x^2}{2\sigma^2}\} dx,$$

$$\sigma^2 \triangleq \sum_{j=0}^{\infty} h_j^2.$$

式(28)和(29)两式意味着, 在正态输入下, 利用系统的输入和输出可以得到线性子系统的脉冲响应序列, 利用脉冲响应序列进一步就可得到线性子系统的各个未知参数. 脉冲响应序列与线性子系统参数之间的关系分析如下:

作如下假设:

B6) $\rho = \frac{1}{\sqrt{2\pi}\sigma^3} \int_{\mathbb{R}} xg(x) \exp\{-\frac{x^2}{2\sigma^2}\} dx \neq 0$.

利用式(28)-(29), 在式(23)两边同时乘以 u_{k-j} 并取期望可得以下公式:

$$h_j - c_1 h_{j-1} - \dots - c_j h_0 = d_j, \tag{30}$$

$$0 \leq j \leq p \wedge (q-1),$$

$$h_j - c_1 h_{j-1} - \dots - c_p h_{j-p} = d_j, \tag{31}$$

$$p+1 \leq j \leq (q-1),$$

其中: $d_0 \triangleq 1$,

$$h_j - c_1 h_{j-1} - \dots - c_j h_0 = 0, \tag{32}$$

$$q \leq j \leq p,$$

$$h_j - c_1 h_{j-1} - \dots - c_p h_{j-p} = 0, \tag{33}$$

$$j \geq p \vee (q-1) + 1.$$

定义矩阵

$$M \triangleq \begin{bmatrix} h_{p \vee (q-1)} & \dots & h_{p \vee (q-1) + 1 - p} \\ h_{p \vee (q-1) + 1} & \dots & h_{p \vee (q-1) + 2 - p} \\ \vdots & \ddots & \vdots \\ h_{p \vee (q-1) + p - 1} & \dots & h_{p \vee (q-1)} \end{bmatrix}. \tag{34}$$

从式(32)和(33)可以得下面线性方程组:

$$M[c_1 \ \dots \ c_p]^T = -[h_{p \vee (q-1) + 1} \ \dots \ h_{p \vee (q-1) + p}]^T. \tag{35}$$

利用第2.1节的思路和式(28)-(29), 基于回归函数 $F_j(x) = -(x - \rho h_j)$, 可以构造随机逼近算法来估计 $\rho h_j, j \geq 0$; 进一步, 如果矩阵 M 又是非奇异的, 利用 $h_j, j \geq 0$ 的估计和式(30)-(35)就可以得到线性子系统未知参数 c_i 和 d_j 的估计.

关于矩阵 M 的非奇异性有以下结论:

引理 2^[29] 假设B1)-B2)成立, 则式(34)定义的矩阵 M 是非奇异的.

由上面的分析可见, 在B1)-B6)假设条件下, 可以首先得到线性子系统参数的估计, 利用这些估计值, 可以估计Wiener系统的内部信号 v_k , 进而利用核函数的思想来估计系统的非线性函数. 在这个思路下, 构造递推辨识算法:

第1步 $(c_1, \dots, c_p), (d_1, \dots, d_{q-1})$ 的估计.

以 $F_j(x) = -(x - \rho h_j)$ 为回归函数的扩张截尾随机逼近算法如下:

$$\alpha_{k+1}(i) = [\alpha_k(i) - \frac{1}{k+1}(\alpha_k(i) - u_{k-i}y_{k+1})] \cdot I_{[|\alpha_k(i) - \frac{1}{k+1}(\alpha_k(i) - u_{k-i}y_{k+1})| \leq M_{\Delta_k(i)}]}, \tag{36}$$

$$\Delta_{k+1}(i) = \sum_{j=0}^k I_{[|\alpha_j(i) - \frac{1}{j+1}(\alpha_j(i) - u_{j-i}y_{j+1})| > M_{\Delta_j(i)}]}, \tag{37}$$

其中: $\{M_k\}_{k \geq 0}$ 为单调递增的正序列, $M_k \rightarrow \infty, \Delta_0(i) = 0$, 初值 $\alpha_0(i), i = 0, \dots, (p \vee (q-1)) + p$ 任意给定.

可以证明, $\{\alpha_k(i)\}_{k \geq 1}$ 收敛到 $h_i \rho$, 因此, 由 $h_{i,k} \triangleq \alpha_k(i)/\alpha_k(0), i = 1, \dots, (p \vee (q-1)) + p$ 就得到 h_i 的估计, 进一步, 将式(30)-(35)中的 h_i 用其估计值 $h_{i,k}$ 代替, 就可以估计出 c_i 和 d_j .

第2步 估计系统内部信号 v_k .

第 k 步递推计算时 c_i 和 d_j 的估计值分别记为 $c_{i,k}$ 和 $d_{j,k}$, 矩阵 C 和 D 中的未知参数用其估计值代替, 相应的矩阵记为 C_k 和 D_k :

$$C_k \triangleq \begin{bmatrix} c_{1,k} & 1 & & \\ c_{2,k} & 0 & \ddots & \\ \vdots & & \ddots & 1 \\ c_{s,k} & \cdots & \cdots & 0 \end{bmatrix}, D_k \triangleq \begin{bmatrix} 1 \\ d_{1,k} \\ \vdots \\ d_{s-1,k} \end{bmatrix}.$$

利用 C_k 和 D_k 来估计内部信号 v_k , 估计值记为 \hat{v}_k :

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k u_k, \quad (38)$$

$$\hat{v}_{k+1} = H \hat{x}_{k+1}, \quad (39)$$

初值 $\hat{x}_0 = 0$.

第3步 估计函数值 $g(v^*)$.

与式(17)类似, 首先定义核函数

$$\hat{w}_k(v^*) \triangleq \frac{1}{\sqrt{2\pi}} \frac{1}{b_k} \exp\left\{-\frac{1}{2} \left(\frac{\hat{v}_k - v^*}{b_k}\right)^2\right\}, \quad (40)$$

其中: $b_k = \frac{1}{k^\alpha}$, $0 < \alpha < \frac{1}{10}$. 注意式(40)中, 系统内部信号 v_k 用估计值 \hat{v}_k 代替.

注意到输入信号 $\{u_k\}_{k \geq 0}$ 为标准正态分布, 经过简单推导可得

$$E w_k(v^*) y_k \xrightarrow[k \rightarrow \infty]{} \rho(v^*) g(v^*), \quad (41)$$

$$E w_k(v^*) \xrightarrow[k \rightarrow \infty]{} \rho(v^*), \quad (42)$$

其中 $\rho(v^*) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(v^*)^2}{2\sigma^2}}$.

与第2.1节类似, 基于回归函数 $-(x - \rho(v^*)g(v^*))$ 和 $-(x - \rho(v^*))$, $g(v^*)$ 的递推辨识算法如下:

$$\begin{aligned} \bar{\eta}_{k+1}(v^*) &= \\ &[\bar{\eta}_k(v^*) - \frac{1}{k+1}(\bar{\eta}_k(v^*) - \hat{w}_k(v^*)y_k)] \cdot \\ &I_{[|\bar{\eta}_k(v^*) - \frac{1}{k+1}(\bar{\eta}_k(v^*) - \hat{w}_k(v^*)y_k)| \leq M_{\bar{\lambda}_k}]}, \quad (43) \\ \bar{\lambda}_{k+1} &= \sum_{j=0}^k I_{[|\bar{\eta}_j(v^*) - \frac{1}{j+1}(\bar{\eta}_j(v^*) - \hat{w}_j(v^*)y_j)| > M_{\bar{\lambda}_j}]} \end{aligned} \quad (44)$$

和

$$\begin{aligned} \bar{\bar{\eta}}_{k+1}(v^*) &= \\ &[\bar{\bar{\eta}}_k(v^*) - \frac{1}{k+1}(\bar{\bar{\eta}}_k(v^*) - \hat{w}_k(v^*))] \cdot \\ &I_{[|\bar{\bar{\eta}}_k(v^*) - \frac{1}{k+1}(\bar{\bar{\eta}}_k(v^*) - \hat{w}_k(v^*))| \leq M_{\bar{\lambda}_k}]}, \quad (45) \\ \bar{\bar{\lambda}}_{k+1} &= \sum_{j=0}^k I_{[|\bar{\bar{\eta}}_j(v^*) - \frac{1}{j+1}(\bar{\bar{\eta}}_j(v^*) - \hat{w}_j(v^*))| > M_{\bar{\lambda}_j}]}, \quad (46) \end{aligned}$$

其中: $\bar{\lambda}_0 = \bar{\bar{\lambda}}_0 = 0$, $\bar{\eta}_{k+1}(v^*)$ 和 $\bar{\bar{\eta}}_{k+1}(v^*)$ 可为任意初值.

定理 4^[29] 假设B1)–B6)成立, 则上述辨识算法是强一致的, 即

$$\alpha_k(l) - h_l \rho = O\left(\frac{1}{k^{\frac{1}{2}-\varpi}}\right) \text{ a.s. } \forall l \geq 0, \quad (47)$$

$$|c_{i,k} - c_i| = O\left(\frac{1}{k^{\frac{1}{2}-\varpi}}\right) \text{ a.s. } i = 1, \dots, p, \quad (48)$$

$$|d_{j,k} - d_j| = O\left(\frac{1}{k^{\frac{1}{2}-\varpi}}\right) \text{ a.s. } j = 1, \dots, q-1, \quad (49)$$

$$|\hat{v}_k - v_k| = O\left(\frac{1}{k^{\frac{1}{2}-\varpi}}\right) \text{ a.s.}, \quad (50)$$

$$|\hat{w}_k(v^*) - w_k(v^*)| = O\left(\frac{1}{k^{\frac{1}{2}-3\alpha-\varpi}}\right) \text{ a.s.}, \quad (51)$$

$$\bar{\eta}_k(v^*) \xrightarrow[k \rightarrow \infty]{} g(v^*) \rho(v^*) \text{ a.s.}, \quad (52)$$

$$\bar{\bar{\eta}}_k(v^*) \xrightarrow[k \rightarrow \infty]{} \rho(v^*) \text{ a.s.}, \quad (53)$$

对任意 $\varpi \in (0, \frac{1}{2})$.

证 与定理1–3类似并结合扩张截尾随机逼近算法的收敛速度定理^[36], 详细证明从略.

由定理4可得Wiener系统参数 (c_1, \dots, c_p) , (d_1, \dots, d_{q-1}) 和函数值 $g(v^*)$ 的强一致估计.

2.3 NARX系统的递推辨识(Recursive identification of NARX systems)

考虑单输入单输出NARX系统

$$y_{k+1} = f(y_k, \dots, y_{k+1-p}, u_k, \dots, u_{k+1-q}) + \varepsilon_{k+1}. \quad (54)$$

对NARX系统(54), 要估计非线性函数 $f(\cdot)$ 在任意给定点 $\varphi^* = [y^{(1)} \dots y^{(p)} u^{(1)} \dots u^{(q)}]^T \in \mathbb{R}^{p+q}$ 的函数值.

与Hammerstein系统、Wiener系统相比, NARX系统中的非线性函数本身包含动态随机项, 因而与前两者相比, 辨识的难度更大. 这里采用如下思路: 首先定义向量

$$\begin{aligned} V_{k+1} &\triangleq [y_{k+1} \ \cdots \ y_{k+2-p} \ u_{k+1} \ \cdots \ u_{k+2-q}]^T, \\ \varphi_1(V_k) &\triangleq [f(y_k, \dots, y_{k+1-p}, u_k, \dots, u_{k+1-q}) \\ &\quad y_k \ \cdots \ y_{k+2-p}]^T, \\ \varphi(V_k) &\triangleq [\varphi_1(V_k)^T \ 0 \ u_k \ \cdots \ u_{k+2-q}]^T, \\ \xi_{k+1} &\triangleq [\varepsilon_{k+1} \ 0 \ \cdots \ 0 \ u_{k+1} \ 0 \ \cdots \ 0]^T, \end{aligned}$$

可将NARX系统表示成下面状态空间方程的形式:

$$V_{k+1} = \varphi(V_k) + \xi_{k+1}. \quad (55)$$

由式(55)可见, $\{V_k\}_{k \geq 0}$ 构成以 $(\mathbb{R}^{p+q}, \mathcal{B}^{p+q})$ 为状态空间的马氏链, 因而可用马氏链的理论来分析 $\{V_k\}_{k \geq 0}$ 的相关性质. 为此, 选取输入信号 $\{u_k\}_{k \geq 0}$

为 iid 序列, $E u_k = 0, E u_k^2 < \infty$, 有概率密度函数 $f_u(\cdot), f_u(\cdot)$ 在 \mathbb{R} 上连续且为正, 并作如下假设:

C1) $\{\varepsilon_k\}_{k \geq 0}$ 也为 iid 序列, $E \varepsilon_k = 0, E \varepsilon_k^2 < \infty$, 有概率密度函数 $f_\varepsilon(\cdot), f_\varepsilon(\cdot)$ 在 \mathbb{R} 上一致连续且为正;

C2) $\{\varepsilon_k\}_{k \geq 0}$ 和 $\{u_k\}_{k \geq 0}$ 相互独立;

C3) $f(\cdot)$ 在 \mathbb{R}^{p+q} 上连续, 存在 \mathbb{R}^p 上的向量范数 $\|\cdot\|_\nu$, 以及常数 $0 < \lambda < 1, c_1 > 0, c_2 > 0$ 和 $l \geq 1$ 使得对 $\forall x \in \mathbb{R}^{p+q}$,

$$\|\varphi_1(x)\|_\nu \leq \lambda \|s\|_\nu + c_1 \sum_{i=1}^q |t_i|^l + c_2, \quad (56)$$

其中: $s \triangleq [s_1 \ \cdots \ s_p]^T \in \mathbb{R}^p, t \triangleq [t_1 \ \cdots \ t_q]^T \in \mathbb{R}^q, x \triangleq [s^T \ t^T]^T \in \mathbb{R}^{p+q}, \varphi_1(\cdot)$ 由式(55)定义;

C4) $E|u_k|^l < \infty, Y_0 \triangleq [y_0 \ y_{-1} \ \cdots \ y_{1-p}]^T$ 满足 $E\|Y_0\| < \infty$.

注 2 条件(56)为系统稳定性条件, 可以证明式(56)包含稳定的线性系统、Hammerstein 系统等为特例^[30].

式(55)所定义的马氏链有以下性质:

引理 3^[30] 假设 C1)–C4) 成立, 则

i) $\{V_k\}_{k \geq 0}$ 是几何遍历的马氏链, 即存在 $(\mathbb{R}^{p+q}, \mathcal{B}^{p+q})$ 上的概率分布函数 $P_{IV}(\cdot)$, 非负可测函数 $M(x)$ 及常数 $M_1 > 0, 0 < \rho_1 < 1$ 使得 $\forall x \in \mathbb{R}^{p+q}$,

$$\|P_k(x, \cdot) - P_{IV}(\cdot)\|_{\text{var}} \leq M(x) \rho_1^k, \quad (57)$$

$$\|P_k(\cdot) - P_{IV}(\cdot)\|_{\text{var}} \leq M_1 \rho_1^k. \quad (58)$$

ii) $P_{IV}(\cdot)$ 有概率密度 $f_{IV}(\cdot), f_{IV}(\cdot)$ 在 \mathbb{R}^{p+q} 上为正.

iii) $\{V_k\}_{k \geq 0}$ 是 α -混合相依的, 混合系数记为 $\{\alpha(k)\}_{k \geq 0}$, 满足 $\alpha(k) \leq M_2 \rho_2^k$, 其中常数 $M_2 > 0, 0 < \rho_2 < 1$.

为估计函数 $f(\cdot)$ 在给定点的函数值 φ^* , 与式(17)(40)类似地构造核函数

$$w_k(\varphi^*) = \frac{1}{(2\pi)^{\frac{p+q}{2}}} \frac{1}{b_k^{p+q}} \exp\left\{-\frac{1}{2} \sum_{i=1}^p \left(\frac{y_{k+1-i} - y^{(i)}}{b_k}\right)^2 - \frac{1}{2} \sum_{i=1}^q \left(\frac{u_{k+1-i} - u^{(i)}}{b_k}\right)^2\right\}, \quad (59)$$

其中: $b_k = \frac{1}{k^\delta}, \delta \in (0, \frac{1}{2(p+q+1)})$.

利用引理 3 所给出的系统马氏性的相关结论, 可以证明:

引理 4^[30] 假设 C1)–C4) 成立, 则上面定义的核函数有下面关系成立:

$$E w_k(\varphi^*) f(V_k) \xrightarrow[k \rightarrow \infty]{} f_{IV}(\varphi^*) f(\varphi^*), \quad (60)$$

$$E w_k(\varphi^*) \xrightarrow[k \rightarrow \infty]{} f_{IV}(\varphi^*). \quad (61)$$

引理 4 启发笔者以 $-(x - f_{IV}(\varphi^*))f(\varphi^*)$ 和 $-(x - f(\varphi^*))$ 为回归函数来构造辨识算法, 与第 2.1 节和第 2.2 节类似, 递推辨识算法构造如下:

$$\begin{aligned} \bar{\theta}_{k+1}(\varphi^*) &= \\ &[\bar{\theta}_k(\varphi^*) - \frac{1}{k+1}(\bar{\theta}_k(\varphi^*) - w_k(\varphi^*)y_{k+1})] \cdot \\ &I_{[|\bar{\theta}_k(\varphi^*) - \frac{1}{k+1}(\bar{\theta}_k(\varphi^*) - w_k(\varphi^*)y_{k+1})| \leq M\bar{\sigma}_k]}, \end{aligned} \quad (62)$$

$$\bar{\sigma}_k = \sum_{j=1}^{k-1} I_{[|\bar{\theta}_j(\varphi^*) - \frac{1}{j+1}(\bar{\theta}_j(\varphi^*) - w_j(\varphi^*)y_{j+1})| > M\bar{\sigma}_j]} \quad (63)$$

和

$$\begin{aligned} \bar{\bar{\theta}}_{k+1}(\varphi^*) &= \\ &[\bar{\bar{\theta}}_k(\varphi^*) - \frac{1}{k+1}(\bar{\bar{\theta}}_k(\varphi^*) - w_k(\varphi^*))] \cdot \\ &I_{[|\bar{\bar{\theta}}_k(\varphi^*) - \frac{1}{k+1}(\bar{\bar{\theta}}_k(\varphi^*) - w_k(\varphi^*))| \leq M\bar{\bar{\sigma}}_k]}, \end{aligned} \quad (64)$$

$$\bar{\bar{\sigma}}_k = \sum_{j=1}^{k-1} I_{[|\bar{\bar{\theta}}_j(\varphi^*) - \frac{1}{j+1}(\bar{\bar{\theta}}_j(\varphi^*) - w_j(\varphi^*))| > M\bar{\bar{\sigma}}_j]} \quad (65)$$

其中: $\{M_k\}_{k \geq 0}$ 为单调递增的正序列, $M_k \rightarrow \infty, \bar{\sigma}_0 = \bar{\bar{\sigma}}_0 = 0$.

定理 5^[30] 假设 C1)–C4) 成立, 则

$$\bar{\theta}_k(\varphi^*) \xrightarrow[k \rightarrow \infty]{} f(\varphi^*) f_{IV}(\varphi^*) \text{ a.s.} \quad (66)$$

$$\bar{\bar{\theta}}_k(\varphi^*) \xrightarrow[k \rightarrow \infty]{} f_{IV}(\varphi^*) \text{ a.s.} \quad (67)$$

以及

$$\theta_k(\varphi^*) \triangleq \bar{\theta}_k(\varphi^*) / \bar{\bar{\theta}}_k(\varphi^*) \xrightarrow[k \rightarrow \infty]{} f(\varphi^*) \text{ a.s.} \quad (68)$$

证 以式(62)–(63)为例, 简述证明思路, 细节可见文献[30].

算法(62)可以改写为

$$\begin{aligned} \bar{\theta}_{k+1}(\varphi^*) &= \\ &[\bar{\theta}_k(\varphi^*) + \frac{1}{k+1}(F(\bar{\theta}_k(\varphi^*)) - e_{k+1})] \cdot \\ &I_{[|\bar{\theta}_k(\varphi^*) + \frac{1}{k+1}(F(\bar{\theta}_k(\varphi^*)) - e_{k+1})| \leq M\bar{\sigma}_k]}, \end{aligned} \quad (69)$$

其中:

$$F(\bar{\theta}_k(\varphi^*)) = -(\bar{\theta}_k(\varphi^*) - f(\varphi^*))f_{IV}(\varphi^*), \quad (70)$$

$$e_{k+1} = f(\varphi^*)f_{IV}(\varphi^*) - w_k(\varphi^*)y_{k+1}. \quad (71)$$

可以验证回归函数 $F(\cdot)$ 和噪声 e_{k+1} 分别满足扩张截尾随机逼近算法收敛所需的假设条件(参考附录或文献[36]), 从而就得到式(66), 式(67)可以类似地证明, 而式(68)是式(66)和(67)的直接结论.

3 分布式随机化PageRank算法的强一致收敛性(Strong consistency of distributed randomized PageRank algorithms)

PageRank算法利用网络节点之间的链接结构给出节点重要性排序. 下面首先简述基本的PageRank算法:

考察有 n 个节点($n \geq 2$)的网络 $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, 其中: $\mathcal{V} = \{1, 2, \dots, n\}$ 是节点的指标集, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ 为有向边集, 刻画节点之间的链接结构: 若 $(i, j) \in \mathcal{E}$, 则认为节点 i 指向节点 j .

记 S_j 为从节点 j 出发所指向的节点构成的集合, S_j 中节点的数目记为 n_j . 对网络 \mathcal{G} 定义链接矩阵如下:

$$A = [a_{ij}]_{n \times n}, a_{ij} = \begin{cases} \frac{1}{n_j}, & j \in \mathcal{L}_i, \\ 0, & \text{其他}, \end{cases} \quad (72)$$

其中 $\mathcal{L}_i = \{j : (j, i) \in \mathcal{E}\}$. 由式(72)可知 $\sum_{i=1}^n a_{ij} = 1$ 或 0 . $\sum_{i=1}^n a_{ij} = 0$ 的情形对应着一类节点, 这类节点不存在指向其他节点的边. 从实际应用的角度, 这类节点很少, 另外也可以在网络中人为地加入链接来改变这种情况. 从而, 对链接矩阵 A , 不妨作如下假设:

D1) 矩阵 A 满足 $\sum_{i=1}^n a_{ij} = 1, j = 1, \dots, n$.

第 i 个节点的PageRank值记为 x_i^* , 假设 $x_i^* \in [0, 1], i \in \mathcal{V}, \sum_{i=1}^n x_i^* = 1$.

PageRank算法的基本思想是节点的重要性用指向它的那些节点的重要性来刻画, 具体的数学模型如下:

$$x_i^* = \sum_{j \in \mathcal{L}_i} \frac{x_j^*}{n_j}. \quad (73)$$

定义向量 $x^* \triangleq (x_1^*, \dots, x_n^*)^T$, 则式(73)可以等价地表示为

$$x^* = Ax^*, x_i^* \in [0, 1]. \quad (74)$$

由式(74)可见, 网络的PageRank值实为链接矩阵特征值1对应的特征向量. 为保证特征值1对应特征子空间的维数为1, 文献[35]基于链接矩阵 A 定义矩阵 M 如下:

$$M \triangleq (1 - \alpha)A + \alpha \frac{S}{n}, \quad (75)$$

其中: $\alpha \in (0, 1)$, S 为 $n \times n$ 维矩阵、各元均为1.

引理 5^[35] 假设D1)成立, 则矩阵 M 有特征值1, 对应的特征子空间维数也为1.

定义 1^[35] 网络 \mathcal{G} 的PageRank值 x^* 定义为

$$x^* = Mx^*, x_i^* \in [0, 1], \sum_{i=1}^n x_i^* = 1. \quad (76)$$

一般说来, 由于网络的节点数目很大, 从定义出发直接计算 x^* 费时费力、很难在容许时间内得到有意义的计算结果, 因而如何计算PageRank值引起许多学者的关注^[26, 37-39]. 文献[26]从链接矩阵 A 出发, 通过分布式、随机化的思想构造“稀疏”的链接矩阵, 然后计算PageRank值. 具体算法如下:

基于矩阵 A , 定义 $A_i, i = 1, \dots, n$

$$(A_i)_{jl} \triangleq \begin{cases} a_{jl}, & \text{如果 } j = i \text{ 或者 } l = i, \\ 1 - a_{il}, & \text{如果 } j = l \neq i, \\ 0, & \text{其他}. \end{cases} \quad (77)$$

取 $\{\theta(k)\}_{k \geq 0}$ 为iid序列, 概率分布满足

$$P\{\theta(k) = i\} = \frac{1}{n}, i = 1, \dots, n. \quad (78)$$

基于 $A_i, i = 1, \dots, n$ 和 $\{\theta(k)\}_{k \geq 0}$, 构造算法递推地估计 x^*

$$x_{k+1} = (1 - \alpha_1)A_{\theta(k)}x_k + \frac{\alpha_1}{n}\mathbf{1}, \quad (79)$$

$$\bar{x}_{k+1} = \frac{1}{k+1} \sum_{l=0}^k x_l, \quad (80)$$

其中: 参数 $\alpha_1 \in (0, 1)$, 初值 x_0 满足各分量之和为1.

上述算法称为分布式、随机化的PageRank算法(distributed randomized PageRank algorithm, DRPA)^[26].

定理 6^[31] 假若D1)成立, 并且参数 $\alpha_1 = \frac{2\alpha}{n - \alpha(n - 2)}$, 则 \bar{x}_k 强一致地收敛到PageRank值, 即

$$\bar{x}_{1,k} - x^* \xrightarrow[k \rightarrow \infty]{} 0 \text{ a.s.} \quad (81)$$

证 算法(80)可以改写为

$$\begin{aligned} \bar{x}_{k+1} &= \bar{x}_k - \frac{1}{k+1}(\bar{x}_k - x_k) = \\ &= \bar{x}_k - \frac{1}{k+1}(\bar{x}_k - x^* + x^* - x_k). \end{aligned} \quad (82)$$

定义 $F(x) = -(x - x^*)$ 和 $e_{k+1} = -(x^* - x_k)$. 可见, 式(82)是以 $F(\cdot)$ 为回归函数、以 e_{k+1} 为噪声的随机逼近算法, 应用随机逼近算法收敛性的相关结论(见文献[36])就可以证明式(81).

4 辨识与参数估计问题的一个可行框架 (A feasible framework for identification and parameter estimation)

第2节和第3节给出了Hammerstein系统、Wiener系统和NARX系统强一致的递推辨识, 以及DRPA的收敛性分析. 从形式上看, Hammerstein系

统、Wiener系统、NARX系统和网络的PageRank问题各不相同,彼此存在很大差异,但从处理方法和算法上看,又有很大共性:首先,这几类系统的辨识、估计问题最终都转化成求根问题,比如,Hammerstein系统辨识的关键之处在于求解Yule-Walker方程,Wiener系统、NARX系统和DRPA的辨识与估计最终转化为求线性方程的零点等,这恰为利用“随机逼近”这一求函数零点的有力工具创造了条件。

另一方面,如何适当地构造回归函数并分析噪声,往往是随机逼近算法能否有效解决问题的关键。一般说来,选取适当的回归函数可以便于分析噪声和证明算法的收敛性,若不当,则效果可能相反;如何选取回归函数并分析噪声,需要根据问题的特点来具体分析。

扩张截尾的随机逼近算法在一定意义下,对收敛性所需的假设条件为最弱^[36]。基于扩张截尾随机逼近算法,笔者初步为Hammerstein系统、Wiener系统、NARX系统的辨识以及网络相关的DRPA算法收敛性分析建立了一个统一的处理框架:“构造回归函数”+“处理噪声”。应用这个框架,已得到了一批后续研究成果,比如Wiener-Hammerstein系统的递推辨识^[40]、变量带误差Wiener系统^[41]和变量带误差Hammerstein系统的递推辨识^[42]、递推的主分量分析^[43]、多智能体一致性控制^[44]等。由于许多辨识和估计问题本质上可转化为求某些回归函数的零点,相信本文所提框架以及所采用的算法将在相关领域中得到广泛应用。

为方便阅读,扩张截尾随机逼近算法的主要结论可参考附录。

5 仿真例子(Examples)

例1 考察下面的NARX系统:

$$y_{k+1} = f(y_k, u_k) + \varepsilon_{k+1} = \exp\{-y_k^2 - u_k^2\} + \varepsilon_{k+1},$$

其中: $\{\varepsilon_k\}$ 为i.i.d.高斯序列, $\varepsilon_k \in \mathcal{N}(0, 1)$ 。

取 $\{u_k\}$ 为i.i.d.高斯序列, $u_k \in \mathcal{N}(0, 1)$, $\{u_k\}$ 与 $\{\varepsilon_k\}$ 相互独立。核函数

$$w_k(y, u) \triangleq \frac{1}{2\pi b_k^2} \exp\left\{-\frac{1}{2}\left(\frac{y_k - y}{b_k}\right)^2 - \frac{1}{2}\left(\frac{u_k - u}{b_k}\right)^2\right\},$$

其中 $b_k = 1/k^{0.15}$ 。

将区间 $u \in [-1, 1]$ 等分成20个小区间,小区间长度为0.1,区间 $y \in [-1, 1]$ 等也分成20个小区间,从而将 $(y, u) \in [-1, 1] \times [-1, 1]$ 等分成400个小正方形,然后估计每一个小正方形顶点的函数值,再将估计值连起来和真实函数曲面对比,以检验算法效果。

图3中实线构成的是真实函数曲面,而虚线则是

将递推步数等于5000时得到函数估计值连起来构成的曲面,图4是递推步数等于5000时的估计误差曲面。

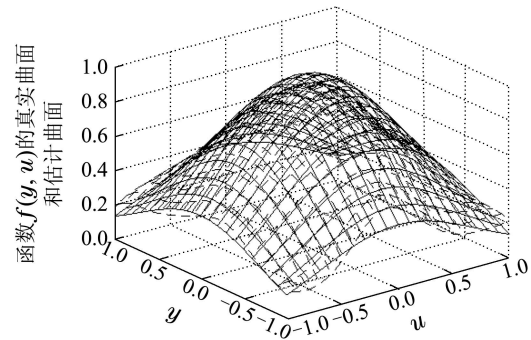


图3 函数曲面的拟合, $k = 5000$

Fig. 3 Estimates for function curve at $k = 5000$

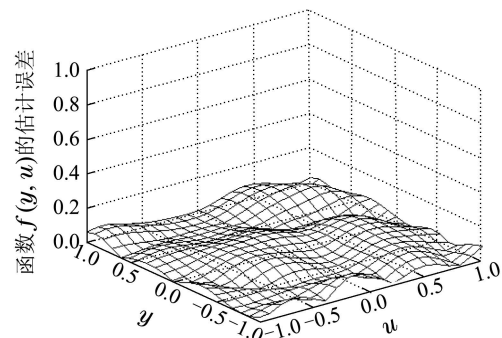


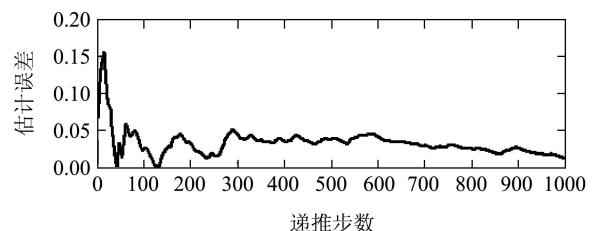
图4 估计误差曲面, $k = 5000$

Fig.4 Estimation error at $k = 5000$

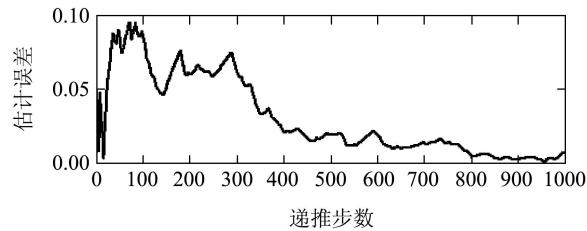
例2 设6节点网络的链接矩阵如下:

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{4} & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{3} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & 0 \end{bmatrix}.$$

图5-7绘制了DRPA算法得到的各个分量的估计误差序列。



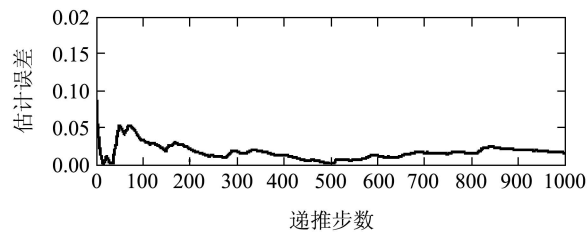
(a) PageRank向量第1个分量的估计误差



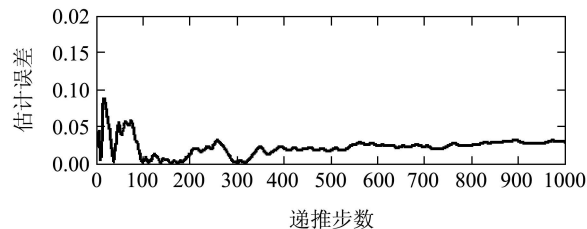
(b) PageRank向量第2个分量的估计误差

图5 PageRank第1、第2分量的估计误差

Fig. 5 Estimation error of 1st and 2nd elements



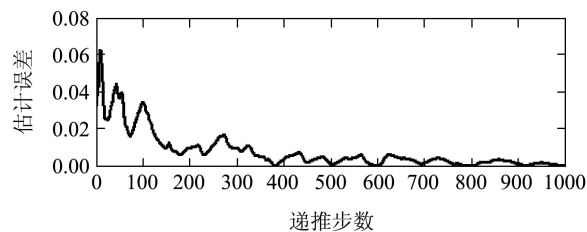
(a) PageRank向量第3个分量的估计误差



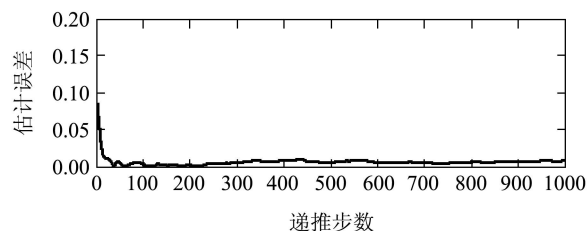
(b) PageRank向量第4个分量的估计误差

图6 PageRank第3、第4分量的估计误差

Fig. 6 Estimation error of 3rd and 4th elements



(a) PageRank向量第5个分量的估计误差



(b) PageRank向量第6个分量的估计误差

图7 PageRank第5、第6分量的估计误差

Fig. 7 Estimation error of 5th and 6th elements

6 小结(Concluding remarks)

本文考虑了几类典型随机非线性系统(Hammerstein系统、Wiener系统和NARX系统)的递推辨识和一类网络相关的估计问题(PageRank), 基于求回归函数零点的思想, 上述问题在统一框架下得到解决。

系统控制、信号处理、系统生物学等领域的许多辨识、估计和控制问题本质上是函数求根, 因此, 本文的思路及框架对于解决相关问题或有裨益。此外, 时变参数非线性系统的辨识、基于分布式测量的系统辨识和网络建模与估计也是热点, 值得关注。

参考文献(References):

- [1] CHEN H F, GUO L. *Identification and Stochastic Adaptive Control* [M]. Boston, MA: Birkhauser, 1991.
- [2] LJUNG L. *System Identification* [M]. Upper Saddle River, NJ: Prentice Hall, 1987.
- [3] VAN OVERCHEE P, DE MOOR B. *Subspace Identification for Linear Systems: Theory-Implementation-Applications* [M]. Boston, MA: Kluwer Academic Publisher, 1996.
- [4] PINTELON R, SCHOUKENS J. *System Identification: a Frequency Domain Approach* [M]. New York: IEEE, 2001.
- [5] BENEDETTO S, BIGLIERI E. Nonlinear equalization of digital satellite channels [J]. *IEEE Journal on Selected Areas in Communications*, 1983, 1(1): 57 – 62.
- [6] CINQUEMANI E, PORRECA R, FERRARI-TRECCATE G, et al. Subtilin production by bacillus subtilis: stochastic hybrid models and parameter identification [J]. *IEEE Transactions on Automatic Control, Special Issue on Systems Biology*, 2008, 53(1): 38 – 50.
- [7] FAN J Q, YAO Q. *Nonlinear Time Series: Nonparametric and Parametric Approach* [M]. New York: Springer-Verlag, 2003.
- [8] KIM J, KONSTANTINOOU K. Digital pre-distortion of wideband signals based on power amplifier model with memory [J]. *IEE Electronics Letters*, 2001, 37(23): 1417 – 1418.
- [9] ZHU Y. Distillation column identification for control using Wiener model [C] // *Proceedings of American Control Conference*. New York: IEEE, 1999, 5: 3462 – 3466.
- [10] LJUNG L. Perspectives on system identification [C] // *Proceedings of The 17th IFAC World Congress, Presented as Plenary Lecture*. London: Elsevier, 2008.
- [11] NINNESS B, WILLS A, HENRIKSEN S J, et al. Computational system identification [C] // *Proceedings of The 18th IFAC World Congress, Presented as Plenary Lecture*. London: Elsevier, 2011.
- [12] 贾立, 李训龙. Hammerstein模型辨识的回顾及展望 [J]. *控制理论与应用*, 2014, 31(1): 1 – 10.
(JIA Li, LI Xunlong. Identification of Hammerstein systems: review and prospect [J]. *Control Theory & Applications*, 2014, 31(1): 1 – 10.)
- [13] BAI E W, LI D. Convergence of the iterative Hammerstein system identification algorithm [J]. *IEEE Transactions on Automatic Control*, 2004, 49(11): 1929 – 1940.
- [14] HAGENBLAD A, LJUNG L, WILLS A. Maximum likelihood identification of Wiener models [J]. *Automatica*, 2008, 44(11): 2697 – 2705.

- [15] HU X L, CHEN H F. Strong consistence of recursive identification for Wiener systems [J]. *Automatica*, 2005, 41(11): 1905 – 1916.
- [16] VERHAEGEN M, WESTWICK D. Identifying MIMO Wiener systems in the context of subspace model identification methods [J]. *International Journal of Control*, 1996, 63(2): 331 – 349.
- [17] ZHAO W X, CHEN H F. Recursive identification for Hammerstein systems with ARX subsystem [J]. *IEEE Transactions on Automatic Control*, 2006, 51(12): 1966 – 1974.
- [18] ZHAO Y L, WANG L Y, YIN G G, et al. Identification of Wiener systems with binary-valued output observations [J]. *Automatica*, 2007, 43(10): 1752 – 1765.
- [19] BAI E W. Non-parametric nonlinear system identification: an asymptotic minimum mean squared error estimator [J]. *IEEE Transactions on Automatic Control*, 2010, 55(7): 1615 – 1626.
- [20] ROLL J, NAZIN A, LJUNG L. Nonlinear system identification via direct weight optimization [J]. *Automatica*, 2005, 41(3): 475 – 490.
- [21] NARENDRA K S, GALLMAN P G. An iterative method for the identification of nonlinear systems using a Hammerstein model [J]. *IEEE Transactions on Automatic Control*, 1966, 11(7): 546 – 550.
- [22] BAI E W. Frequency domain identification of Hammerstein models [J]. *IEEE Transactions on Automatic Control*, 2003, 48(4): 530 – 542.
- [23] GIRI F, ROCHDI Y, GNING J B, CHAOUI F Z. An analytic geometry approach to Wiener system frequency identification [J]. *IEEE Transactions on Automatic Control*, 2009, 54(4): 683 – 696.
- [24] VANBEYLEN L, PINTELOON R, SCHOUKENS J. Blind maximum likelihood identification of Hammerstein systems [J]. *Automatica*, 2008, 44(12): 3139 – 3146.
- [25] VANBEYLEN L, PINTELOON R, SCHOUKENS J. Blind maximum likelihood identification of Wiener systems [J]. *IEEE Transactions on Signal Processing*, 2009, 57(8): 3017 – 3029.
- [26] ISHII H, TEMPO R. Distributed randomized algorithms for the PageRank computation [J]. *IEEE Transactions on Automatic Control*, 2010, 55(9): 1987 – 2002.
- [27] CHEN H F, ZHAO W X. Identification of several classes of stochastic nonlinear systems [J]. *Journal of Systems Science and Mathematics Sciences*, 2011, 31(9): 1019 – 1044.
- [28] ZHAO W X. Parametric identification of Hammerstein systems with consistency results using stochastic inputs [J]. *IEEE Transactions on Automatic Control*, 2010, 55(2): 474 – 480.
- [29] ZHAO W X, CHEN H F. Markov chain approach to identifying Wiener systems [J]. *Science China: Information Science*, 2012, 55(5): 1201 – 1217.
- [30] ZHAO W X, CHEN H F, ZHENG W X. Recursive identification for nonlinear ARX systems based on stochastic approximation algorithm [J]. *IEEE Transactions on Automatic Control*, 2010, 55(6): 1287 – 1299.
- [31] ZHAO W X, CHEN H F, FANG H T. Convergence of distributed randomized PageRank algorithms [J]. *IEEE Transactions on Automatic Control*, 2013, 58(12): 3255 – 3259.
- [32] ZHAO W X, CHEN H F. Identification of Wiener, Hammerstein, and NARX systems as Markov chains with improved estimates for their nonlinearities [J]. *Systems and Control Letters*, 2012, 61(12): 1175 – 1186.
- [33] ESKINAT E, JPHNSON S, LUYBEN W L. Use of Hammerstein models in identification of nonlinear systems. *AIChE Journal*, 1991, 37(2): 255 – 268.
- [34] KALAFATIS A, ARIFIN N, WANG L, et al. A new approach to the identification of pH processes based on the Wiener model. *Chemical Engineering Science*, 1995, 50(23): 3693 – 3701.
- [35] BRIN S, PAGE L. The anatomy of a large-scale hypertextual Web search engine [J]. *Computer Networks ISDN Systems*, 1998, 30(1): 107 – 117.
- [36] CHEN H F. *Stochastic Approximation and Its Applications* [M]. Dordrecht, The Netherlands: Kluwer, 2002.
- [37] DE JAGER D V, BRADLEY J T. Asynchronous iterative solution for state-based performance metrics [C] // *Proceedings of ACM SIGMETRICS/SI. I.*: [s.n.], 2007: 373 – 374.
- [38] KAMVAR S, HAVELIWALA T, GOLUB G. Adaptive methods for the computation of PageRank [J]. *Linear Algebra and its Applications*, 2004, 386: 51 – 65.
- [39] KOLLIAS G, GALLOPOULOS E, SZYLD D B. Asynchronous iterative computations with Web information retrieval structures: The PageRank case [M] // JOUBERT G R, Ed, et al. *Parallel Computing: Current and Future Issues of High-End Computing*. Julich, Germany: John von Neumann-Institut for Computing, 2006, 33: 309 – 316.
- [40] MU B Q, CHEN H F. Recursive Identification of Wiener-Hammerstein Systems [J]. *SIAM Journal of Control Optimization*, 2012, 50(5): 2621 – 2658.
- [41] MU B Q, CHEN H F. Recursive identification of errors-in-variables Wiener systems [J]. *Automatica*, 2013, 49(9): 2744 – 2753.
- [42] MU B Q. *Recursive identification of block-oriented nonlinear systems* [D]. Beijing: The University of Chinese Academy of Sciences, 2013.
- [43] CHEN H F, FANG H T, ZHANG L L. Recursive estimation for ordered eigenvectors of symmetric matrix with observation noise [J]. *Journal of Mathematical Analysis and Applications*, 2011, 382(2): 882 – 842.
- [44] FANG H T, CHEN H F, WEN L. On control of strong consensus for networked agents with noisy observations [J]. *Journal of Systems Science and Complexity*, 2012, 25(1): 1 – 12.

附录 扩张截尾随机逼近算法(Appendix Stochastic approximation algorithm with expanding truncations)

设回归函数函数 $f(\cdot) : \mathbb{R}^l \rightarrow \mathbb{R}^l$, $f(\cdot)$ 的零点集记为 J , 即 $\forall x^0 \in J, f(x^0) = 0$. 设 $\{M_k\}_{k \geq 0}$ 为单调递增的正序列, $M_k \rightarrow \infty$. 任意给定初值 $x_0 \in \mathbb{R}^l$, $\{x_k\}_{k \geq 1}$ 由下面的扩张截尾随机逼近算法递推产生,

$$x_{k+1} = (x_k + a_k y_{k+1}) I_{\|x_k + a_k y_{k+1}\| \leq M_{\sigma_k}} + x^* I_{\|x_k + a_k y_{k+1}\| > M_{\sigma_k}}, \quad (A1)$$

$$\sigma_k = \sum_{j=0}^{k-1} I_{\|x_j + a_j y_{j+1}\| > M_{\sigma_j}}, \quad \sigma_0 = 0, \quad (A2)$$

其中

$$y_{k+1} = f(x_k) + \varepsilon_{k+1}, \quad (A3)$$

y_{k+1} 为 $f(\cdot)$ 在点 x_k 的量测, ε_{k+1} 为噪声, a_k 为算法步长, x_k 是对 $f(\cdot)$ 零点集 J 的第 k 步逼近.

下面给出基本假设条件:

$$H1) \quad a_k > 0, \quad a_k \xrightarrow{k \rightarrow \infty} 0, \quad \sum_{k=1}^{\infty} a_k = \infty;$$

H2) 存在连续可微的Lyapunov函数 $V(\cdot) : \mathbb{R}^l \rightarrow \mathbb{R}$, 满足下面条件:

- i) $\sup_{\delta \leq d(x, J) \leq \Delta} \nabla V(x)^\top f(x) < 0, \quad \forall 0 < \delta < \Delta,$
- ii) 集合 $V(J) \triangleq \{V(x), x \in J\}$ 无处稠密(nowhere dense),
- iii) 对式(A1)中的 $x^* \in \mathbb{R}^l$, 存在 $c_0 > 0$, 有 $\|x^*\| < c_0$, 并且 $V(x^*) < \inf_{\|x\|=c_0} V(x),$

H3) 在样本轨道 ω 上, 对算法(A1)–(A2)产生的任意收敛子列 $\{x_{n_k}\}_{k \geq 1}$, 成立下面条件:

$$\lim_{T \rightarrow 0} \limsup_{k \rightarrow \infty} \frac{1}{T} \left\| \sum_{i=n_k}^{m(n_k, T_k)} a_i \varepsilon_{k+1} \right\| = 0, \quad \forall T_k \in (0, T], \tag{A4}$$

其中

$$m(n_k, T_k) \triangleq \max \left\{ m : \sum_{i=n_k}^m a_i \leq T_k \right\}.$$

H4) 未知函数 $f(\cdot)$ 可测、局部有界.

定理 A^[36] 对随机逼近算法(A1)–(A2), 假若H1)–H4)成立, 则在H3)成立的样本轨道 ω 上, 有

$$d(x_k, J) \xrightarrow[k \rightarrow \infty]{} 0. \tag{A5}$$

将 J 的闭包记为 \bar{J} , $\{x_k\}$ 的极限点构成的集合记为 J^* . 进一步有

$$d(x_k, J^*) \xrightarrow[k \rightarrow \infty]{} 0, \tag{A6}$$

且 J^* 为 \bar{J} 的闭连通子集, 其中 $d(x_k, J) \triangleq \inf_{y \in J} \|x_k - y\|.$

显然, 若 $f(x)$ 单零点, 即 $J = \{x^0\}$, 则在定理7假设条件下有 $d(x_k, x^0) \xrightarrow[k \rightarrow \infty]{} 0, J^* = J = \{x^0\}.$

作者简介:

赵文虢 (1981–), 男, 第19届(2013年)《关肇直奖》获奖论文作者, 博士, 副研究员, 曾赴澳大利亚西悉尼大学进行学术访问, 研究方向为系统辨识、网络建模, 以及系统生物学中的参数估计等, E-mail: wxzhao@amss.ac.cn;

陈翰馥 (1937–), 男, 研究员, 博士生导师, 研究领域包括随机系统的辨识、适应控制、参数及状态估计、随机逼近和优化及其系统控制、信号处理等的应用, 发表期刊论文200余篇, 专著7本. 1993年当选中国科学院院士, 1996年当选IEEE Fellow, 2005年当选第三世界科学院院士, 2007年当选IFAC Fellow, 2014年当选国际系统与控制科学院院士, E-mail: hfchen@iss.ac.cn.