第 26 卷第 4 期
2009 年 4 月

控 制 理 论 与 应 用
Control Theory & Applications

Vol. 26 No. 4
Apr. 2009

文章编号: 1000−8152(2009)04−0365−06

非持续激励条件下系统辨识递推最小二乘最小范数算法

李银国, 汤卓群, 黄 镭

(重庆邮电大学 自动化学院, 重庆 400065)

摘要: 系统辨识中广泛应用的最小二乘算法需要输入向量序列满足持续激励性条件(PE条件); 但在大多情况下这是难以满足的. 本文提出了一种不依赖于PE条件的递推最小二乘、最小范数辨识算法. 首先分析了最小二乘算法解空间的结构, 并运用罚函数方法, 将参数辨识问题转化为无约束优化问题. 然后, 提出了将步长、罚因子等过程控制参数统一的迭代−递推形式的辨识算法, 证明了算法在给定的控制参数约束下收敛于唯一的最小二乘、最小范数解向量. 仿真实验表明在非PE条件下算法的有效性.

关键词: 系统辨识; 最小二乘算法; 持续激励条件; 最小二乘最小范数解

中图分类号: TP13 **文献标识码**: A

Recursive least-squares and minimum-norm algorithm for system identification without persistent excitation condition

LI Yin-guo, TANG Zhuo-qun, HUANG Lei

(Department of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: The widely used least-squares algorithms for system identification rely on the assumption that the sequence of input vectors satisfies the persistent excitation conditions (PE conditions); however, this condition is difficult to be realized in most cases. For this purpose, a recursive least-squares and minimum-norm (RLS–MN) identification algorithm that does not depend upon PE conditions is proposed. First, the structure of the solution space of the least squares algorithm is analyzed and the parameter identification problem is converted to an unconstraint optimization problem by using the penalty function method; and then, an iteration-recursion-based identification algorithm is presented by unifying the process control parameters, such as step width and penalty factor. This algorithm is proved to converge to a unique least-squares and minimum-norm solution vector when the control parameters satisfy the given constraint conditions. Finally, simulation results are presented to confirm the validity of the algorithm without using PE conditions.

Key words: system identification; least square algorithm; persistent excitation conditions; recursive least-square and minimum-norm algorithm

1 引言(Introduction)

非线性系统辨识一直是系统科学与控制工程领域研究的重点和难点[1]. 为解决复杂多样的非线性系统结构的未知性问题, 人们寻求用一组确定的基函数(basis function)逼近于未知非线性结构, 继而对其中的线性参数进行辨识或估计, 以满足工程应用的需要. 例如, 基于径向基和小波基类神经网络的机器学习、基于Volterra 级数模型的系统辨识、基于NARMAX模型的非线性回归估计等[2~5], 都可归结为如下形式的离散MISO非线性系统中线性参数的辨识问题:

$$y(t) = W^{\mathrm{T}}(t) \cdot \phi(x(t)) + v(t). \tag{1}$$

其中, 输入向量$x(t) \in \mathbb{R}^n$, 系统输出$y(t) \in \mathbb{R}$, 非线性向量函数$\phi(x) = (\varphi_1(x), \varphi_2(x), \cdots, \varphi_N(x))^{\mathrm{T}}$, 待辨识的参数向量$W(t) \in \mathbb{R}^N$, $v(t)$为噪声. 因此, 研究式(1)所示的非线性系统的参数辨识问题具有广泛的代表性.

对于线性参数的辨识问题, 目前普遍应用的有最小二乘(LMS)或递推最小二乘(RLS)估计算法及其改进算法等[4,6,7], 这些方法均要求输入回归向量序列具有持续激励性条件(persistency excitation conditions, PE条件). 但是, 从机器学习的观点看, 系统(1)是由输入空间的向量序列$\{x(t)\}$, 通过一组非线性特征函数集$\{\varphi_i(x)\}_{i=1}^N$, 向特征空间映射而得

收稿日期: 2007−09−11; 收修改稿日期: 2008−10−31.
基金项目: 国家863计划资助课题(2006AA11A1C1-3); 国家973计划前期研究专项资助课题(2008CB317111).

到更高维的特征向量序列$\{\phi(x(t))\}$, 从而可用线性学习器学习更复杂的非线性关系, 其代价是待辨识的线性参数向量$W \in \mathbb{R}^N$的维数明显增大, 即$N >> n$, 这便是函数基类神经网络学习中常见的"维数灾"问题[2,3]. 低维的输入向量序列$\{x(t)\}$向高维空间非线性映射, 再加之样本数相对于较大的N来说数量通常不足, 使得$\{\phi(x(t))\}$的PE条件难以满足或难以验证, 从而使RLS等算法的应用失去了基础, 无偏性、稳定性难以保证. 人们提出了一些改进RLS算法, 如基于协方差阵$P(t)$奇异值分解(SVD)技术的递推辨识算法等[9], 但都存在高维矩阵复杂计算的新问题. 因此, 非线性系统(1)在最为一般的输入样本条件下的参数辨识问题, 在理论上和算法上并未能得到有效的解决.

本文在分析最小二乘算法解空间结构的基础上, 运用优化理论中的罚函数原理, 提出了不依赖于PE条件的递推最小二乘、最小范数辨识算法, 简称RLS–MN算法. 该算法将辨识过程的控制参数(如学习步长、罚因子等)的变化过程协调统一, 通过迭代计算可得到唯一的Moor-Penrose广义逆形式的解向量, 即最小二乘、最小范数解. 当系统加入新信息后, 以递推方式对有关矩阵和向量进行更新, 进而得到新的解向量. 本文理论证明了该算法的收敛性和最优解的唯一性, 给出了控制参数间应满足的约束条件. 最后给出了算法的仿真实验结果.

2 最小二乘算法解空间的结构(Structure of solution space for least squares algorithm)

设系统 (1) 的输入输出样本集为 $J = \{(x(p), y(p)), p = 1, 2, \cdots, t\}$, 则$t$时刻参数向量$W(t)$在最小二乘(LS)意义下的最优解集为

$$S(t) = \{W_{\rm LS}(t) | \varPhi(t) \cdot W_{\rm LS}(t) = P(t)\}, \quad (2)$$

其中$\varPhi(t)$和$P(t)$可取如下带遗忘因子的形式:

$$\varPhi(t) = \sum_{p=1}^{t} \beta^{t-p} \cdot \phi(x(p)) \cdot \phi^{\rm T}(x(p)), \quad (3)$$

$$P(t) = \sum_{p=1}^{t} \beta^{t-p} \cdot \phi(x(p)) \cdot y(p), \quad (4)$$

其中$0 \leqslant \beta < 1$为遗忘因子.

若向量序列$\{\phi(x(p))\}$满足PE条件, 则$\varPhi(t)$是非奇异阵, 此时可通过一次性算法、LMS算法、RLS算法等辨识算法得到解集$S(t)$中的唯一最优解

$$W^*_{\rm LS}(t) = \varPhi^{-1}(t) \cdot P(t). \quad (5)$$

但是, 当$\{\phi(x(p))\}$不满足PE条件时, $\varPhi(t)$为降秩矩阵, LS 类算法是病态的, 式(2)的LS解不唯一, 最优解向量$W^*_{\rm LS}(t)$不确定. 进一步地对LS解集$S(t)$在$\mathbb{R}^N$空间中作正交分解得

$$W_{\rm LS}(t) = W^{(0)}_{\rm LS}(t) + W^{(1)}_{\rm LS}(t). \quad (6)$$

其中$W^{(0)}_{\rm LS}(t) \in K[\varPhi(t)], W^{(1)}_{\rm LS}(t) \in K^{\perp}[\varPhi(t)], K[\varPhi(t)]$为实矩阵 $\varPhi(t)$ 的零子空间 (核空间), $K^{\perp}[\varPhi(t)]$ 为$K[\varPhi(t)]$的正交补子空间. 由线性空间理论知:

1) $W^{(1)}_{\rm LS}(t)$是式(2)所示的LS解空间中唯一的最小二乘、最小范数解.

2) 对于任何$W^{(0)}_{\rm LS}(t) \in K[\varPhi(t)]$, 则$W^{(0)}_{\rm LS}(t) + W^{(1)}_{\rm LS}(t)$均为最小二乘解.

正因为$W^{(0)}_{\rm LS}(t)$可在子空间$K[\varPhi(t)]$中任意取, 从而造成当$\varPhi(t)$奇异时最优参数向量$W^*(t)$不确定, 参数辨识(或网络学习)过程不稳定. 为保证算法过程收敛性和唯一性, 需剔除$W^{(0)}_{\rm LS}(t)$.

3) 将最小二乘、最小范数解$W^{(1)}_{\rm LS}(t)$一般地表示为$W^*(t)$, 则由矩阵理论知$W^*(t)$可表为

$$W^*(t) = \varPhi^{+}(t) \cdot P(t). \quad (7)$$

其中$\varPhi^{+}(t)$为矩阵$\varPhi(t)$的Moor-Penrose广义逆. 当向量序列$\{\phi(x(p))\}$不具备PE性时, $\varPhi(t)$为奇异阵, 此时的参数辨识目标转为求$\varPhi(t) \cdot W(t) = P(t)$的广义逆型最优解. 这不仅仅能保证辨识算法过程的收敛性、稳定性, 而且便于降低系统的复杂度, 提高算法效率和辨识效果.

但是, 计算式(7)中的高阶时变矩阵的广义逆$\varPhi^{+}(t)$并没有有效算法, 如Greville迭代法是针对低维的定常矩阵, 广义逆递推公式的自适应滤波算法[10]只针对线性自回归辨识问题, 且矩阵结构较简单. 因此, 求式(7)的最优解, 还需要研究更为有效的辨识算法.

3 递推最小二乘最小范数算法(Recursive least-squares and minimum-norm algorithm)

求式(7)所示的最小二乘、最小范数解, 等价于求解如下二次规划问题

$$\begin{aligned} &\min \|W(t)\|^2, \\ &\text{s.t.}\ \ \varPhi(t) \cdot W(t) = P(t), \\ &\qquad W(t) \in \mathbb{R}^n. \end{aligned} \quad (8)$$

而式(8)的约束优化问题又可运用罚函数(penalty function)原理, 将其转化为无约束优化问题. 即:

$$\min F[W(t)] = \|W(t)\|^2 + \mu \cdot \|\varPhi(t) \cdot W(t) - P(t)\|. \quad (9)$$

其中$\mu > 0$为罚因子(penalty factor). 设式(9)关于μ的最优解向量为$W^*(t, \mu)$, 则由罚函数定理可得式(8)的最优解为

$$W^*(t) = \lim_{\mu\to+\infty} W^*(t,\mu). \quad (10)$$

由于$F[W(t)]$关于$W(t)$的二次函数是严格凸的, 故优化问题(9)的解向量$W^*(t,\mu)$可由梯度算法得到, 其迭代方法为

$$\begin{aligned}&W(t,\mu,k+1) =\\&W(t,\mu,k) - \eta_k[(I+\mu\cdot\Phi^2(t))\cdot\\&W(t,\mu,k)-\mu\cdot\Phi(t)\cdot P(t)].\end{aligned} \quad (11)$$

其中$\eta_k > 0$为学习率.

根据上述分析知, 欲获得最优解向量$W^*(t)$, 需先对递增罚因子系列$\{\ \mu_i\ \}$中的 μ_i 按照式(11)对k进行迭代计算, 得到的解向量 $W^*(t,\mu_i)$, 然后再对于新的罚因子 μ_{i+1} 重复上述迭代, 得到解向量 $W^*(t,\mu_{i+1})$, 由此类推, 直到满足某种精度要求时停止, 此时有充分大的μ_I, 使得$W^*(t) \approx W^*(t,\mu_I)$. 显然, 这种多重迭代方法在实际工程应用中是难以办到的.

如果在时间点t将上述关于学习率序列$\{\eta_k\}$和罚因子序列$\{\mu_i\}$的双重迭代过程统一起来, 由式(11)整理后可得

$$\begin{aligned}&W(t,k+1) =\\&[(1-\eta_k)I - \eta_k\cdot\mu_k\Phi^2(t)]\ \cdot\\&W(t,k) + \eta_k\cdot\mu_k\Phi(t)\cdot P(t).\end{aligned} \quad (12)$$

其中 $k\to+\infty$ 时 $\eta_k > 0, \mu_k \to +\infty$. 从初始向量$W(t,0)$出发经迭代计算得最优辨识参数向量$W^*(t)$. 当有新样本加入, 较大地改变了$\Phi(t)$和$P(t)$时, 再按如下递推式更新:

$$\Phi(t+1) = \beta\cdot\Phi(t)+\phi(x(t+1))\cdot\phi^{\mathrm T}(x(t+1)), \quad (13)$$
$$P(t+1) = \beta\cdot P(t)+\phi(x(t+1))\cdot y(t+1). \quad (14)$$

然后再以$W^*(t)$为初始向量按式(12)进行迭代得$W^*(t+1)$. 由此类推.

本文称式(12)～(14)便是系统(1)关于线性参数辨识问题的递推最小二乘、最小范数算法(RLS–MN).

4　递推最小二乘最小范数算法的收敛性分析(Analysis of convergence of RLS–MN algorithm)

引理 1　设X $\subset \mathbb{R}^n$为有界紧集, $\{x(p)\}$为以概率密度$\rho(x)$从X中独立抽取的输入样本集, 由此得系统(1)的样本集$J = \{(x(p),y(p)), p = 1,2,\cdots,t\}$, 则式(3)(4)定义的矩阵$\Phi(t)$和$P(t)$分别依概率1收敛于定常矩阵$\Phi$和$P$.

引理1可根据Kolmogrov强大数定理证得. 详细证明略.

RLS–MN算法的收敛性包括$W^*(t,k)$关于k的迭代过程的收敛性和$W^*(t)$关于时间t的收敛性. 引理1的结论说明, 在较一般的条件下$\Phi(t)$, $P(t)$依概率1收敛于常矩阵Φ和P, 显然$W^*(t)$必收敛于最优参数向量W^*. 因此, 下面只需证明$W^*(t,k)$关于k的迭代过程的收敛性. 为表述方便, 不妨省略式(12)中的时间记号t得:

$$\begin{aligned}W(k+1) = {}&[(1-\eta_k)I-\eta_k\cdot\mu_k\Phi^2]\cdot W(k)+\\&\eta_k\cdot\mu_k\Phi\cdot P.\end{aligned} \quad (15)$$

在下面的讨论中, 使用如下的向量范数及相应的矩阵导出范数: $\|x\| = (\sum\limits_{i=1}^{n} x_i^2)^{1/2}$, $\|A\| = (\lambda_{\max}(A^TA))^{1/2}$, 其中$\lambda_{\max}(A)$为矩阵的最大特征值. 不难证明, 当$A\in\mathbb{R}^{n\times n}$为实对称矩阵时, $\|A\| = \max\limits_{1\leqslant i\leqslant n}|\lambda_i(A)|$; 当$A\in\mathbb{R}^{n\times n}$为非负定实对称矩阵时, $\|A\| = \lambda_{\max}(A)$.

引理 2　设系统(1)的非线性向量函数$\phi(x(t))$有界, 则存在 $B > 0$ 和 $C > 0$, 使得 $\|\Phi(t)\| < B^2$, $\|P(t)\| < B\cdot C$.

证　由$\phi(x(t))$的有界性得知存在$b > 0$和$c > 0$, 使得$\|\phi(x(t))\| \leqslant b$, $|y(t)| \leqslant c$成立.

$$\|\Phi(t)\| \leqslant \sum_{p=1}^{t}\beta^{t-p}\cdot\|\phi(x(p))\cdot\phi^{\mathrm T}(x(p))\|, \quad (16)$$

设$\Phi_p = \phi(x(p))\cdot\phi^{\mathrm T}(x(p))$, 显然$\mathrm{rank}(\Phi_p) = 1$, 且$\lambda_1(\Phi_p) > 0, \lambda_2(\Phi_p) = \cdots = \lambda_n(\Phi_p) = 0$, 从而

$$\begin{aligned}&\|\Phi_p\| = \lambda_{\max}(\Phi_p) =\\&\mathrm{tr}[\phi(x(p))\cdot\phi^{\mathrm T}(x(p))] = \|\phi(x(p))\|^2 \leqslant b^2.\end{aligned}$$

再代入式(16)得$\|\Phi(t)\| \leqslant \sum\limits_{p=1}^{t}\beta^{t-p}\cdot b^2 = \dfrac{1-\beta^t}{1-\beta}\cdot b^2 < \dfrac{1}{1-\beta}\cdot b^2$, 设$B = \dfrac{b}{\sqrt{1-\beta}}$, 则$\|\Phi(t)\| < B^2$. 同理可证$\|P\| < B\cdot C$.

引理 3　设二次目标函数的优化问题

$$\begin{aligned}&\min\ \ \|W\|^2,\\&\text{s.t. } \Phi\cdot W = P,\\&W\in\mathbb{R}^n\end{aligned} \quad (17)$$

的最优解向量为W^*; 罚函数形式的无约束优化问题

$$\min F[W(\mu_k)] = \|W\|^2+\mu_k\cdot\|\Phi\cdot W-P\|^2 \quad (18)$$

的最优解为$W^*(\mu_k)$, 则

$$W^* = Q^{\mathrm T}\cdot\mathrm{diag}\{\frac{1}{\lambda_1},\frac{1}{\lambda_2},\cdots,\frac{1}{\lambda_r},0,\cdots,0\}\cdot Q\cdot P, \quad (19)$$

$$W^*(\mu_k)=Q^{\mathrm{T}}\cdot\mathrm{diag}\{\frac{\mu_k\lambda_1}{1+\mu_k\lambda_1^2},\frac{\mu_k\lambda_2}{1+\mu_k\lambda_2^2},\cdots,\frac{\mu_k\lambda_r}{1+\mu_k\lambda_r^2},0,\cdots,0\}\cdot Q\cdot P. \tag{20}$$

其中Q为正交矩阵, $\lambda_1,\cdots,\lambda_r$为$\Phi$的$r$个正惯性指数.

证 因为Φ为非负定对称矩阵, 则对Φ作正交–对角分解得:

$$\Phi=Q^{\mathrm{T}}\cdot\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_r,0,\cdots,0)\cdot Q. \tag{21}$$

由Moore-Penrose广义逆的性质及式(7)可得:

$$W^*=\Phi^+\cdot P=Q^{\mathrm{T}}\cdot\mathrm{diag}(\frac{1}{\lambda_1},\frac{1}{\lambda_2},\cdots,\frac{1}{\lambda_r},0,\cdots0)\cdot Q\cdot P.$$

又根据优化问题(18)的目标函数的严格凸性知, 唯一最优解为$W^*(\mu_k)$, 可由$\dfrac{\partial F}{\partial W(\mu_k)}=0$得:

$$W^*(\mu_k)=(I+\mu_k\cdot\Phi^2)^{-1}\cdot\mu_k\cdot\Phi\cdot P. \tag{22}$$

再将矩阵Φ的正交–对角分解式式(21)代入式(22)便可得到式(20)成立.

引理 4 对于离散时变线性系统

$$W(k+1)=A(k)\cdot W(k)+B(k), \tag{23}$$

其中$W(k)\in\mathbb{R}^n, A(k)\in\mathbb{R}^{n\times n}, B(k)\in\mathbb{R}^n$. 若满足条件:

1) $\|A(k)\|<1-\eta_k$, 其中 $0<\eta_k\leqslant1$, 且$\sum\eta_k=\infty$;

2) $\|B(k)\|=o(\eta_k)$, 即$\dfrac{\|B(k)\|}{\eta_k}\to0$;

则系统(23)在外加摄动$B(k)$作用下大范围渐近鲁棒稳定, 即对于任何初态$W(0)$均有$\lim\limits_{k\to+\infty}\|W(k)\|=0$成立.

(引理4的证明见文[11].)

定理 1 (RLS–MN算法收敛性)若取学习率序列$\{\eta_k\}$和罚因子序列$\{\mu_k\}$当k充分大时满足:

1) $\eta_k>0,\eta_k\to0$;

2) $\overline{\lim\limits_{k\to\infty}}\dfrac{\eta_{k+1}}{\eta_k}=1$, 且$\sum\eta_k=\infty$;

3) 取$\mu_k=\dfrac{d}{\eta_k}$, 其中$0<d<\dfrac{2}{B^4}$.

则式(15)所示的RLS–MN算法使向量$W(k)$大范围收敛于最小二乘、最小范数解W^*. 即从任何初态$W(0)$迭代计算, 均有$W(k)\to W^*$.

证
$$\|W(k)-W^*\|\leqslant\|W(k)-W^*(\mu_k)\|+\|W^*(\mu_k)-W^*\|.$$

由引理3易证明$\|W^*(\mu_k)-W^*\|\to0$. 下面只需证明$\|W(k)-W^*(\mu_k)\|\to0$成立.

根据式(15)和式(22), 可将$\overline{W}(k)=W(k)-W^*(\mu_k)$表示为如下时变线性系统的状态方程形式:

$$\overline{W}(k+1)=A(k)\cdot\overline{W}(k)+B(k). \tag{24}$$

其中

$$A(k)=(1-\eta_k)I-\eta_k\mu_k\cdot\Phi^2, \tag{25}$$

$$B(k)=W^*(\mu_k)-W^*(\mu_{k+1}). \tag{26}$$

1) 由条件 3) 及 $\eta_k\to0$知, 存在 a, b, 满足取$\dfrac{a(1-\eta_k)}{\eta_k}<\mu_k<\dfrac{b(1-\eta_k)}{\eta_k}$, 其中

$$0<a<d<b<\frac{2}{B^4}.$$

因为 $\lambda_i[A(k)]=1-\eta_k-\eta_k\mu_k\cdot\lambda_i[\Phi^2]$, 易证$0\leqslant\lambda_i[\Phi^2]<B^4$, 从而有 $\lambda_i[A(k)]<1-\eta_k$.

另一方面有

$$\lambda_i[A(k)]>1-\eta_k-\eta_k\mu_k\cdot B^4>1-\eta_k-\eta_k\cdot\frac{b(1-\eta_k)B^4}{\eta_k}>1-\eta_k-\frac{2}{B^4}(1-\eta_k)\cdot B^4=-(1-\eta_k),$$

所以有$|\lambda_i[A(k)]|<1-\eta_k$, 即

$$\|A(k)\|=\max_{1\leqslant i\leqslant n}|\lambda_i[A(k)]|<1-\eta_k.$$

再由$\eta_k>0,\eta_k\to0$和$\sum\eta_k=\infty$推知, 引理4的条件1)成立.

2) 根据引理3, 将$W^*(\mu_k)$的表达式式(20)代入$B(k)$, 利用$\mu_k>\dfrac{a(1-\eta_k)}{\eta_k}$条件可得:

$$\frac{\|B(k)\|}{\eta_k}<\frac{1}{a(1-\eta_k)}\cdot\max_{1\leqslant i\leqslant n}\left|\frac{(1-\frac{\mu_{k+1}}{\mu_k})\lambda_i}{(\frac{1}{\mu_k}+\lambda_i^2)(\frac{1}{\mu_k}+\frac{\mu_{k+1}}{\mu_k}\cdot\lambda_i^2)}\right|\cdot\|P\|.$$

因为$\eta_k\to0$时$\mu_k\to\infty$, 且已知

$$\overline{\lim_{k\to\infty}}\frac{\mu_{k+1}}{\mu_k}=\overline{\lim_{k\to\infty}}\frac{(1-\eta_{k+1})\cdot\eta_k}{(1-\eta_k)\cdot\eta_{k+1}}=1,$$

从而推知$\dfrac{\|B(k)\|}{\eta_k}\to0$.

根据引理4知$\|W(k)-W^*(\mu_k)\|\to0$成立. 定理得证.

推论 1 若取学习率序列$\{\eta_k\}$和罚因子序

列$\{\mu_k\}$当k充分大时满足:

1) $\eta_k = \dfrac{c}{k^{\alpha}}$, 其中$0 < \alpha \leqslant 1, c > 0$为常数;

2) 取$\mu_k = \dfrac{d}{\eta_k}$, 其中$0 < d < \dfrac{2}{B^4}$, $B = \max\limits_{t} \|\phi(x(t))\|$.

则式(15)所示的RLS–MN算法大范围收敛于最小二乘、最小范数解W^*. 证明略.

5　仿真实验结果(Simulation result)

为了验证上述RLS–MN算法在非持续激励条件下的有效性和最优解唯一等特性, 选择通常权系数很多而样本数不足的小波神经网络的学习问题, 将RLS–MN算法与应用较广泛的RLS算法进行比较实验. 设含有较大噪声的非线性信号为:

$$\begin{aligned} & y = f(x) + v = \\ & (x^2 - 2x + 1.5) \times (\sin(4\pi x) + 0.125 \cdot \cos(26\pi x)) + \\ & 0.125 \cdot \sin(2\pi x) + 2 + v. \end{aligned} \tag{27}$$

其中, $x \in [0.1,\ 1]$, 高斯噪声$v \sim N[0, 0.05^2]$, 随机抽取20对观察值作为样本, 按文献[8]中有关方法设计如下形式的小波网络:

$$\hat{Y}(t) = \sum_{i=1}^{N} w_i \cdot \varphi_i(x(t)). \tag{28}$$

其中$\varphi_i(x)(\varphi_1(x) \equiv 1)$为小波基函数, $N = 50$. 小波网络(28)的权系数$W = (w_1, w_2, \cdots, w_N)^{\mathrm{T}}$的学习问题, 便是式(1)所示非线性系统线性参数的辨识问题. 可以证明, 当学习样本个数小于向量$\phi(x) = (\varphi_1(x), \varphi_2(x), \cdots, \varphi_N(x))^{\mathrm{T}}$的维数$N$时, 非线性向量序列$\{\phi(x(p))\}$一定不满足PE条件. 分别用RLS算法和RLS–MN算法对式(28)权系数进行学习训练, 其效果比较见图1～4.

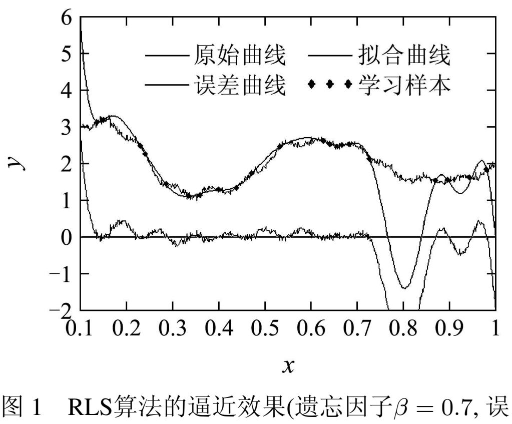


图 1　RLS算法的逼近效果(遗忘因子$\beta = 0.7$, 误差MSSR<0.035)

Fig. 1　Approximation effect of RLS algorithm (forgetting factor $\beta = 0.7$, error MSSR<0.035)

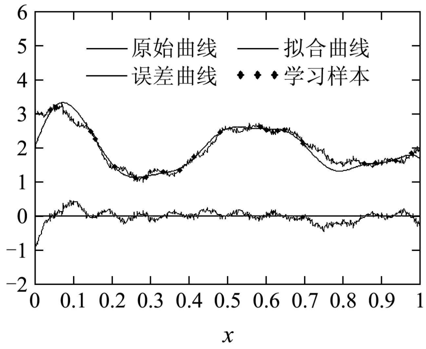


图 2　RLS–MN算法的逼近效果(控制参数$\alpha = 1, c = 1.2$, $a = 0.003$, 误差MSSR<0. 035)

Fig. 2　Approximation effect of RLS–MN algorithm(control parameters $\alpha = 1$, $c = 1.2$, $a = 0.003$, error MSSR<0.035)

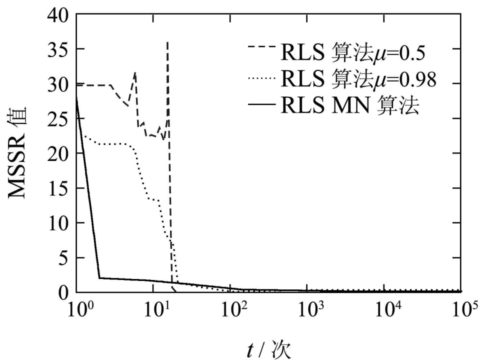
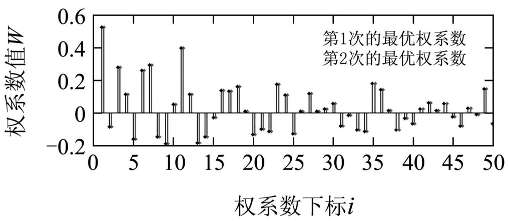
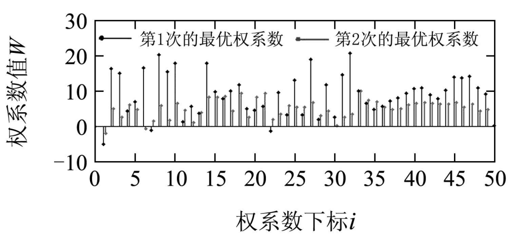


图 3　RLS算法和RLS–MN算法拟合误差比较

Fig. 3　Comparison of approximation errors between RLS algorithm and RLS–MN algorithm



(a)　RLS–MN算法权最优系数



(b)　RLS算法权最优系数

图 4　取不同初始值时RLS算法和RLS–MN算法获得最优权系数比较

Fig. 4　Comparison of approximation weights with different original values between RLS algorithm and RLS–MN algorithm.

仿真实验表明, 当PE条件不满足时, RLS–MS算法不但能很好逼近学习样本, 而且拟合曲线平缓, 在样本空缺或稀疏区不会出现大幅振荡(图1、2), 能较好避免过度拟合问题, 对于提高外推能力十分有效; RLS–MS算法逼近速度稍慢于RLS算法, 但逼近过程稳定, 不会出现对学习样本和控制参数过于敏感使误差大幅反弹现象(见图3); 对于不同的初始权值, RLS–MS算法都能保证学习过程收敛于同一最优权系数和相同的拟合曲线, 而RLS算法无法做到这一点(图4), 这证实了RLS–MN算法大范围收敛性和最优解的唯一性.

6 结论(Conclusion)

1) 当向量序列不满足PE条件时, RLS–MS算法仍能保证辨识过程从任意初态收敛于唯一的Moor-Penrose广义逆形式的最优解(正交投影解), 辨识过程稳定, 结果更为确定.

2) RLS–MS算法获得的最优解不但使拟合样本的误差最小, 而且使最优参数解向量的范数最小, 从而能使拟合曲线(面)尽可能地平缓. 在机器学习问题中可省去系数为零或近似为零基函数部分, 使系统结构可进一步简化. 按照统计学习理论, 这利于提高学习系统的泛化能力.

3) RLS–MS算法计算简单, 控制参数选取直观, 在辨识参数数量庞大的情况下, 其特点更为突出.

但是, RLS–MS算法在理论上还有待于深入的研究, 其收敛速度等性能指标还可以进一步改善.

参考文献(References):

[1] NOWAK R D. Nonlinear system identification[J]. *Systems and Signal Processing*, 2002, 21(1): 109 – 122.

[2] JUN Z, GILBERT G W, YUBO M, et al. Wavelet neural networks for function learning [J]. *IEEE Transactions on Signal Processing*, 1995, 43(6): 1485 – 1497.

[3] 石宏理, 蔡远利, 邱祖廉. 一种基于小波分解的非线性系统辨识的新方法[J]. 信息与控制, 2004, 33(5): 554 – 559.
(SHI Hongli, CAI Yuanli, QIU Zulian. A new approach of nonlinear system identification based on wavelet decomposition[J]. *Information and Control*, 2004, 33(5): 554 – 559.)

[4] CHIKOUCHE D, BEKKA R E, TOUBEL Y. Performance study of LMS and RLS algorithms in the identification of nonlinear volterra systems[C] //*Proceedings of Modelling and Simulation MS'2004*. France: Association for Modeling and Simulation in Enterprise, 2004, AMSE: 10.9 – 10.12.

[5] SINGLA P, HENDERSON T, JUNKINS J L. A robust nonlinear system identification algorithm using orthogonal polynomial network[C] //*Proceedings of the AAS/ALAA Space Flight Mechanics Meeting*. San Diego, US: Univelt Inc, 2005, 120(1): 983 – 1002.

[6] 方崇智, 萧德云. 过程辨识[M]. 北京: 清华大学出版社, 1988.
(FANG Chongzhi, XIAO Deyun. *Prpcess Identification* [M]. Beijing: Tsinghua University Press, 1988.)

[7] FENG D Z, BAO Z, ZHANG X D. Modified RLS algorithm for unbiased estimation of FIR system with input and output noise[J]. *Electronics Letters*, 2000, 36(3): 273 – 274.

[8] 李银国, 张邦礼, 曹长修. 小波神经网络及其结构设计方法[J]. 模式识别与人工智能, 1997, 10(3): 197 – 205.
(LI Yinguo, ZHANG Bangli, CAO Changxiu. Wavelet neural networks and its structural design[J]. *Pattern Recognition and Artificial Intelligence*, 1997, 10(3): 197 – 205.)

[9] 张友民, 李庆国, 戴冠中, 等. 基于奇异值分解的递推辨识方法[J]. 控制理论与应用, 1995, 12(2): 224 – 229.
(ZHANG Youmin, LI Qingguo, DAI Guanzhong, et al. A new recursive identification method based on singular value decomposition[J]. *Control Theory & Applications*, 1995, 12(2): 224 – 229.)

[10] 高鹰, 谢胜利. 基于矩阵广义逆递推的自适应滤波算法[J]. 电子学报, 2002, 30(7): 1032 – 1034.
(GAO Ying, XIE Shengli. An adaptive filtering algorithm based on recursion of generalized inverse matrix[J]. *Acta Electronica Sinica*, 2002, 30(7): 1032 – 1034.)

[11] 李银国, 唐小我, 曹长修. 离散线性时变系统鲁棒稳定性分析[J]. 控制与决策, 1998, 13(5): 608 – 611.
(LI Yinguo, TANG Xiaowo, CAO Changxiu. Analysis of robust stability for discrete linear time-varying systems[J]. *Control and Decision*, 1998, 13(5): 608 – 611.)

作者简介:

李银国 (1955—), 男, 教授, 博士生导师, 主要研究方向为智能控制理论及应用、模式识别技术、汽车电子控制系统等, E-mail: liyg@cqupt.edu.cn;

汤卓群 (1981—), 女, 硕士研究生, 主要研究方向汽车电子控制系统, E-mail: tangzhuoqun@sohu.com;

黄 镭 (1981—), 男, 硕士研究生, 主要研究方向汽车电子智能控制理论及应用, E-mail: eatcaravan@gmail.com.