文章编号: 1000−8152(2007)05−0851−05

基于自适应密度估计的系统参数辨识

衷路生, 宋执环

(工业控制技术国家重点实验室, 浙江大学工业控制技术研究所 浙江 杭州 310027)

摘要: 针对噪声分布未知的ARMAX系统, 提出了一种自适应非参数噪声密度估计方法, 由估计误差动态调整高斯核函数的全局带宽和局部带宽, 实现了未知噪声分布密度的自适应估计; 通过极小化似然函数, 给出了基于噪声密度估计的参数辨识迭代算法, 分析了算法的收敛性并给出了算法收敛的充分条件. 仿真结果表明本文提出的算法在系统噪声未知时具有较强的抗噪能力和良好的收敛性.

关键词: 自适应密度估计; 非参数密度估计; 参数估计; 极大似然函数

中图分类号: TP273 **文献标识码**: A

Parameter identification based on adaptive density estimation

ZHONG Lu-sheng, SONG Zhi-huan

(National Laboratory of Industrial Control Technology , Institute of Industrial Process Control, Zhejiang University, Hangzhou Zhejiang 310027, China)

Abstract: The problem of estimating parameters of ARMAX-model subject to unknown noise density distribution is addressed in this paper. Firstly, a method for adaptive nonparametric noise density estimation by using the Gaussian kernel is proposed. The global bandwidth and local bandwidth of the Gaussian kernel are dynamically computed according to the estimation error. Secondly, an iterative algorithm for parameter estimation is presented by minimizing the maximum likelihood function and the algorithm is based on the nonparametric estimate of the noise density. Thirdly, the convergence of the algorithm is analyzed and the convergent condition is given. Finally, simulation results show the proposed technique offers improved performance over existing methods when the noise density distribution is unknown.

Key words: adaptive density estimation; non-parametric density estimation; parameter estimation; maximum likelihood function

1 引言(Introduction)

当噪声概率密度分布已知时中外学者针对不同系统提出了许多性能良好的辨识方法. Huber基于Minimax方法, 对冲击噪声系统提出了鲁棒M估计方法[1~3], Fang详细介绍了噪声为高斯分布时的最小二乘算法及其改进算法[4]; Yohai给出了算法收敛的充分条件[5], Wang将Huber的估计方法应用于多用户检测[6]; Ding将递阶辨识的方法应用于高维、待估计参数多且含有均值为零、二阶矩有界的噪声的复杂大系统, 大大降低了计算量和存储量, 并用鞅超收敛定理证明了算法的收敛性[7]. 这一类方法称为基于噪声参数模型的辨识[8], 其共同的特点是系统的噪声分布可以用精确的数学模型描述, 在此基础上用优化、统计学和鞅超收敛性等数学方法进行参数估计.

然而, 由于电磁辐射、大型设备的机械振动等复杂环境的影响, 使得系统的实际噪声模型严重偏离预先假定的分布, 影响基于噪声参数模型的辨识效果[6,8,9]. 针对噪声分布密度未知的系统, 本文提出一种自适应非参数噪声密度估计方法, 基于极大似然原理, 给出了参数估计的迭代算法并分析了算法的收敛性.

2 辨识模型描述(System description)

为表述方便起见, 考虑SISO ARMAX模型[5]:

$$y(k)=-\sum_{i=1}^{n_a}a_iy(k-i)+\sum_{j=1}^{n_b}b_ju(k-j)+v(k)=$$
$$h^{\mathrm T}(k)\theta+v(k), k=1,2,\cdots,L-1,L, \tag{1}$$

其中L表示样本长度,

$$\begin{aligned}h(k)=&[-y(k-1),-y(k-2),\cdots,-y(k-n_a),\\&u(k-1),u(k-2),\cdots,u(k-n_b)]^{\mathrm T}=\\&[h_1(k),h_2(k),\cdots,h_{n_a}(k),h_{n_a+1}(k),\\&h_{n_a+2}(k),\cdots,h_{n_a+n_b}(k)]^{\mathrm T},\end{aligned}$$

收稿日期: 2005−10−09; 收修改稿日期: 2006−07−04.
基金项目: 国家973计划项目(2002CB312203-02).

$$\theta=[a_1,a_2,\cdots,a_{n_a},b_1,b_2,\cdots,b_{n_b}]^{\mathrm{T}}=$$
$$[\theta_1,\theta_2,\cdots,\theta_{n_a},\theta_{n_a+1},\theta_{n_a+2},\cdots,\theta_{n_a+n_b}]^{\mathrm{T}},$$

T是转置符号, $v(k)$是的观测噪声, 其分布密度函数$f(v)$未知.

若记:

$$Y_L=[y(1),y(2),\cdots,y(L-1),y(L)]^{\mathrm{T}},$$
$$H_L=[h(1),h(2),\cdots,h(L-1),h(L)]^{\mathrm{T}},$$
$$V_L=[v(1),v(2),\cdots,v(L-1),v(L)]^{\mathrm{T}},$$

则式(1)可表示为

$$Y_L=H_L\theta+V_L. \tag{2}$$

3 密度估计(Density estimation)

密度估计就是要从独立同分布观测序列$y(1),y(2),\cdots,y(L-1),y(L)$中获得噪声分布函数$f(v)$的估计. 常用的非参数密度估计方法有: Histograms, 样条函数法, 核(Kernal)估计法, 正交级数法, 小波变换法[10]等, 其中核估计由于能得到密度函数的封闭表达式而得到广泛应用.

3.1 自适应非参数密度估计(Adaptive non-parametric density estimation)

自适应核估计的本质是对每个观测值设定一个核来获得经验分布的点估计, 核的累加和作为噪声分布函数$f(v)$的估计, 其数学描述为

$$\hat{f}(v)=\frac{1}{L}\sum_{i=1}^{L}\frac{1}{h\lambda_i}K\left(\frac{v-v(i)}{h\lambda_i}\right), \tag{3}$$

其中: λ_i是局部带宽, 描述了分布函数的重尾、单双峰等局部信息; h是全局带宽, 用于设定分布函数$\hat{f}(v)$的整体平滑特性; L是观测点总数; $K(x)$是核函数, 满足可微、非负、$\int_{-\infty}^{+\infty}K(x)=1$等条件, 当核函数选为高斯核时, 核函数变为

$$\hat{f}(v)=\frac{1}{L}\sum_{i=1}^{L}\frac{1}{h\lambda_i}\frac{1}{\sqrt{\pi}}\mathrm{e}^{\left[-\left(\frac{v-v(i)}{h\lambda_i}^2\right)\right]}. \tag{4}$$

3.2 局部带宽和全局带宽的选择(Computation of local bandwidth and global bandwidth)

按下列步骤设定局部带宽λ_i和全局带宽h的值.

· 设定所有的$\lambda_i=1, i=1,2,\cdots,L-1,L$, 得到

$$\hat{f}_h(v)=\frac{1}{L}\sum_{i=1}^{L}\frac{1}{h}K\left(\frac{v-v(i)}{h}\right). \tag{5}$$

· 极小化以下目标函数, 得到h的最优值

$$\mathrm{MISE}=\mathrm{E}\left[\int_{-\infty}^{+\infty}\left(\hat{f}_h(v)-f(v)\right)^2\mathrm{d}v\right]. \tag{6}$$

当噪声密度分布$f(v)$满足高斯分布且$K(x)$选为高斯型核函数, 即:

$$f(v)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{\left[-\frac{(v-\mu)^2}{2\sigma^2}\right]},$$
$$\hat{f}_h(v)=\frac{1}{L}\sum_{i=1}^{L}\frac{1}{h}\mathrm{e}^{\left[-\left(\frac{v-v(i)}{h}\right)^2\right]},$$

通过极小化公式(6)可以得到公式(5)中h的最优估计为: $h_{\mathrm{opt}}=0.79RL^{-1/5}$, R为区间分位数, 其计算步骤如下:

给定一组观测值:

$$y(1),y(2),\cdots,y(L-1),y(L).$$

Step 1 对观测序列$y(1),y(2),\cdots,y(L-1),y(L)$由小到大排序, 假设排序后的序列为$y_s(1),y_s(2),\cdots,y_s(L-1),y_s(L)$;

Step 2 确定样本分组group1和group2, 分两种情况讨论:

i) 当L为偶数,

group1为

$$y_s(1),y_s(2),\cdots,y_s(\frac{L}{2}-1),y_s(\frac{L}{2}),$$

group2为

$$y_s(\frac{L}{2}+1),y_s(\frac{L}{2}+2),\cdots,y_s(L-1),y_s(L).$$

ii) 当L为奇数,

group1为

$$y_s(1),y_s(2),\cdots,y_s(\frac{L+1}{2}-1),y_s(\frac{L+1}{2}),$$

group2为

$$y_s(\frac{L+1}{2}),y_s(\frac{L+1}{2}+1),\cdots,y_s(L-1),y_s(L);$$

Step 3 确定group1和group2的分位数$Q1$和$Q2$;

Step 4 确定区间分位数R为: $Q2-Q1$.

当噪声分布密度$f(v)$未知时, 通常用交叉确认方法或者再抽样方法极小化公式(6) 而得到h得最优估计[11]. 在实际应用中, 由公式$h_{\mathrm{opt}}=0.79RL^{-1/5}$ 得到的值满足计算要求[11].

· 将h_{opt}代入式(5)得到

$$\hat{f}_0(v)=\frac{1}{L}\sum_{i=1}^{L}\frac{1}{h_{\mathrm{opt}}}K\left(\frac{v-v(i)}{h_{\mathrm{opt}}}\right). \tag{7}$$

· 更新λ_i $(i=1,2,\cdots,L-1,L)$的值, 计算公式为

$$\lambda_i=\frac{\hat{f}_0(v(i))}{\prod_{j=1}^{L}\hat{f}_0(v(j))}. \tag{8}$$

· 将λ_i $(i=1,2,\cdots,L-1,L)$的值和h_{opt}代入

式(3)得到噪声密度函数的估计

$$\hat{f}_h(v)=\frac{1}{L}\sum_{i=1}^{L}\frac{1}{h_{\rm opt}\lambda_i}K\left(\frac{v-v(i)}{h_{\rm opt}\lambda_i}\right), \tag{9}$$

经过适当的数学处理, 式(9)便可很好的描述具有重尾、对称、单峰分布特征的非高斯噪声; 文献[12]讨论了核估计的一致性条件.

4 参数估计迭代算法(Iterative algorithm for parameter estimation)

4.1 极大似然函数的推导(The maximum-likelihood function)

当噪声分布密度为$f(v)$时, 极大似然函数为

$$L(\theta/Y_L)=\prod_{k=1}^{L}f\left(y(k)-h^{\rm T}(k)\theta\right), \tag{10}$$

对式(10)两边取对数, 得到被估计参数的极大似然解为

$$\hat{\theta}_{ML}=\arg\min_{\theta}\sum_{k=1}^{L}-\log f\left(y(k)-h^{\rm T}(k)\theta\right). \tag{11}$$

式(11)对θ_i $(i=1,2,\cdots,(n_a+n_b))$求导, 并令导数为0, 可得(n_a+n_b)个方程:

$$\sum_{k=1}^{L}\varPsi\left(y(k)-\sum_{j=1}^{n_a+n_b}h_j(k)\theta_j\right)h_l(k)=0,\quad l=1,2,\cdots,n_a+n_b, \tag{12}$$

其中$\varPsi(x)=-f'(x)/f(x)$称为代价函数.

写成向量形式为

$$H_L^{\rm T}\varPsi\left(Y_L-H_L\theta\right)=0_{(n_a+n_b)}, \tag{13}$$

其中$0_{(n_a+n_b)}$表示(n_a+n_b)维列向量.

由于实际噪声分布函数$f(v)$未知, 在计算时用式(9)得到的密度估计函数$\hat{f}(v)$代替, 当核函数$K(x)$选为高斯核函数, 则:

$$\varPsi(v)=-\frac{\hat{f}'(v)}{\hat{f}(v)}=\frac{\frac{1}{N}\sum_{k=1}^{L}\frac{1}{\sqrt{\pi}}\frac{1}{h_{\rm opt}\lambda_k}\frac{-2}{h_{\rm opt}\lambda_k}\frac{v-v(k)}{h_{\rm opt}\lambda_k}\mathscr{A}}{\frac{1}{N}\sum_{k=1}^{L}\frac{1}{\sqrt{\pi}}\frac{1}{h_{\rm opt}\lambda_k}\mathscr{A}}=\frac{\sum_{k=1}^{L}\frac{1}{\lambda_k}\frac{-2}{h_{\rm opt}\lambda_k}\frac{v-v(k)}{h_{\rm opt}\lambda_k}\mathscr{A}}{\sum_{k=1}^{L}\frac{1}{\lambda_k}\mathscr{A}}, \tag{14}$$

其中$\mathscr{A}={\rm e}^{\left[-\left(\frac{v-v(k)}{h_{\rm opt}\lambda_k}\right)^2\right]}$. 可以看出式(14)的分母不为0, 确保了$\varPsi(v)$的存在性.

4.2 迭代算法(Iterative algorithm)

按以下步骤进行参数迭代计算:

Step 1　置$l=0$, 由最小二乘法得到参数的初始估计值:

$$\hat{\theta}^0=\hat{\theta}_{LS}=\left(H_L^{\rm T}H\right)^{-1}H_L^{\rm T}Y_L.$$

Step 2　计算观测残差向量

$$\hat{v}=Y_L-H_L\tilde{\theta}^l. \tag{15}$$

Step 3　将残差向量$\hat{v}$ 的各分量代入式(14)计算代价函数向量:

$$z=\varPsi(\hat{v})=-\frac{\hat{f}'(\hat{v})}{\hat{f}(\hat{v})}. \tag{16}$$

Step 4　更新参数向量:

$$\hat{\theta}^{l+1}=\hat{\theta}^l+\frac{1}{\mu}\left(H_L^{\rm T}H\right)^{-1}H_L^{\rm T}z, \tag{17}$$

其中$\mu=\max(|\varPsi^l(x)|)$.

Step 5　如果$\|\hat{\theta}^{l+1}-\hat{\theta}^l\|\leqslant\varepsilon$, 则停止迭代, 否则回到Step2继续.

4.3 算法收敛性分析(Convergence analysis of the algorithm)

定义目标函数:

$$c(\theta)=\sum_{k=1}^{L}-\log f\left(y(k)-h^{\rm T}(k)\theta\right), \tag{18}$$

则对以上迭代算法, 有以下收敛定理:

定理 1　如果$R=H_L^{\rm T}H_L$正定, 则由式(16)(17)所得的参数估计满足以下关系:

$$C(\hat{\theta}^l)-C(\hat{\theta}^{l+1})\geqslant\frac{\mu}{2}(\hat{\theta}^l-\hat{\theta}^{l+1})^{\rm T}R(\hat{\theta}^l-\hat{\theta}^{l+1})=\frac{1}{2\mu}z((\hat{\theta}^l)^{\rm T}SR^{-1}S^{\rm T}z(\hat{\theta}^l), \tag{19}$$

其中

$$z(\hat{\theta}^l)=[\varPsi((y(1)-h^{\rm T}(1)\hat{\theta}^l),\varPsi((y(2)-h^{\rm T}(2)\hat{\theta}^l),\cdots,\varPsi((y(L)-h^{\rm T}(L)\hat{\theta}^l)]^{\rm T}.$$

证　定义函数

$$d(\delta)=\frac{1}{\mu}[C(\hat{\theta}^l)-C(\hat{\theta}^l+\delta)\frac{1}{2}[\delta^{\rm T}H_L^{\rm T}H_L\delta-\frac{2}{\mu}\delta^{\rm T}H_L^{\rm T}z(\hat{\theta}^l)],\delta\in\mathbb{R}^{(n_a+n_b)}, \tag{20}$$

有以下结果:

$$d(0)=0, \tag{21}$$

$$\frac{\partial(d(\delta))}{\partial\delta}\Big|_{\delta=0}=\frac{1}{\mu}\sum_{k=1}^{L}\varPsi\left(y(k)-h^{\rm T}(k)(\hat{\theta}^l+\delta)\right)h(k)+$$

$$H_L^{\mathrm{T}}H_L\delta - \frac{1}{\mu}H_L^{\mathrm{T}}z(\hat{\theta}^l)\Big|_{\delta=0} = \tag{22}$$

$$\frac{1}{\mu}H_L^{\mathrm{T}}z(\hat{\theta}^l+\delta)+H_L^{\mathrm{T}}H_L\delta-\frac{1}{\mu}H_L^{\mathrm{T}}z(\hat{\theta}^l)\Big|_{\delta=0}=0^{(n_a+n_b)}, \tag{23}$$

$$\frac{\partial^2(d(\delta))}{\partial\delta^2} = -\frac{1}{\mu}\sum_{k=1}^{L}\Psi'\left(y(k)-h^{\mathrm{T}}(k)(\hat{\theta}^l+\delta)\right)\cdot h(k)h^{\mathrm{T}}(k)+H_L^{\mathrm{T}}H_L, \tag{24}$$

因为 $\Psi'(x)\leqslant\mu$,

所以 $$\frac{\partial^2(d(\delta))}{\partial\delta^2}\geqslant-\sum_{k=1}^{L}h(k)h^{\mathrm{T}}(k)+H_L^{\mathrm{T}}H_L = 0^{(n_a+n_b)\times(n_a+n_b)}. \tag{25}$$

由式(21)~(25)可知: 对任何$\delta\in\mathbb{R}^{(n_a+n_b)}$, $d(\delta)\geqslant 0$.

设$\delta=\hat{\theta}^l-\hat{\theta}^{l+1}=\dfrac{1}{\mu}(H_L^{\mathrm{T}}H_L)^{-1}H_L^{\mathrm{T}}z(\hat{\theta}^l)$, 则由式(21)可得:

$$0\leqslant d(\delta)=\frac{1}{\mu}[C(\hat{\theta}^l)-C(\hat{\theta}^l+\delta)]-\frac{1}{2\mu^2}z^{\mathrm{T}}(\hat{\theta}^l)H_L(H_L^{\mathrm{T}}H_L)^{-1}H_L^{\mathrm{T}}z(\hat{\theta}^l), \tag{26}$$

式(26)等价于:

$$C(\hat{\theta}^l)-C(\hat{\theta}^{l+1})\leqslant\frac{\mu}{2}(\hat{\theta}^l-\hat{\theta}^{l+1})^{\mathrm{T}}R(\hat{\theta}^l-\hat{\theta}^{l+1})=\frac{1}{2\mu}z(\hat{\theta}^l)^{\mathrm{T}}SR^{-1}S^{\mathrm{T}}z(\hat{\theta}^l).$$

证毕.

推论 1 如果$\rho(x)=-\log(f(x))$是凸函数, 并且有下界, 则当$l\to\infty$, $\hat{\theta}^l\to\theta^*$.
其中: θ^*使$C(\theta)$的取值达到最小.

证 因为$-\log(f(x))$是凸函数, 并且有下界, 所以$C(\hat{\theta}^l)=\sum\limits_{k=1}^{L}-\log f(y(k)-h^{\mathrm{T}}(k)\hat{\theta}^l)$也是凸函数, 并且有下界.

由定理知$C(\hat{\theta}^l)$是一个递减序列, 所以$C(\hat{\theta}^l)$收敛.

由式(26)可知: 当$l\to\infty$时,

$$z^{\mathrm{T}}(\hat{\theta}^l)H_L(H_L^{\mathrm{T}}H_L)^{-1}H_L^{\mathrm{T}}z(\hat{\theta}^l)\leqslant 2\mu[C(\hat{\theta}^l)-C(\hat{\theta}^l+\delta)]\to 0. \tag{27}$$

因为$H_L(H_L^{\mathrm{T}}H_L)^{-1}H_L^{\mathrm{T}}$满秩, 所以$\lim\limits_{l\to\infty}z(\hat{\theta}^l)\to 0$.
由于$z(\hat{\theta}^l)$的第i $(i=1,2,\cdots,L)$个分量为:

$$\Psi(y(i)-h^{\mathrm{T}}(i)\hat{\theta}^l)=-\frac{f'(y(i)-h^{\mathrm{T}}(i)\hat{\theta}^l)}{f(y(i)-h^{\mathrm{T}}(i)\hat{\theta}^l)}= \tag{28}$$

$$-\frac{\partial\log f(y(i)-h^{\mathrm{T}}(i)\hat{\theta}^l)}{\partial f(y(i)-h^{\mathrm{T}}(i)\hat{\theta}^l)}, \tag{29}$$

式(28)中分母$f(y(i)-h^{\mathrm{T}}(i)\hat{\theta}^l)$是一个概率密度, 取值范围为$(0,1)$, 故$\lim\limits_{l\to\infty}\Psi(y(i)-h^{\mathrm{T}}(i)\hat{\theta}^l)\to 0$等价于$\lim\limits_{l\to\infty}f'(y(i)-h^{\mathrm{T}}(i)\hat{\theta}^l)\to 0$.

由推论条件: $\rho(x)=-\log(f(x))$是凸函数, 并且有下界及式(29)可知: 存在$\lim\limits_{l\to\infty}\hat{\theta}^l\to\theta^*$, 使$C(\theta)$的取值达到最小. 证毕.

由定理和推论可知, 对满足一定条件的系统, 由迭代算法可以得到参数的一致估计.

5 仿真与结果分析(Simulation study)

考虑如下由差分方程描述的辨识对象:

$$y(k)=1.5y(k-1)-0.7y(k-2)+u(k-1)+0.5u(k-2)+v(k),$$

输入信号采用幅度为4阶M序列[5], $v(k)$服从$f_v(v)=(1-\varepsilon)f_N(v)+\varepsilon f_C(v)$分布[6], $f_N(v)$是高斯分布$N(0,\sigma^2)$的概率密度, $f_C(v)$是高斯分布$N(0,\kappa\sigma^2)$的概率密度, $\kappa\gg 1$, 在本文中取$\sigma^2=1$, $\kappa=100$, $\varepsilon=0.1$, 数据长度$L=500$, 第35个输出点后, 每隔15个观测点, 输出值同期性地加上$\{-4,6,-8,3,-5,4,-7,6\}$序列, 噪声比$SNR=0.3387$, 参数估计向量$\theta=[a_1,a_2,b_1,b_2]^{\mathrm{T}}=[-1.5,0.7,1.0,0.5]^{\mathrm{T}}$ (T表示转置符号), 前十个数据所得的最小二乘用来作为M估计递推计算的初值, 经过500步计算, 由递推最小二乘算法、Wang[13]给出的算法以及本文递推算法的均方误差$MSE^l=(\hat{\theta}^l-\theta)^{\mathrm{T}}(\hat{\theta}^l-\theta)$随迭代次数的变化如图1所示:

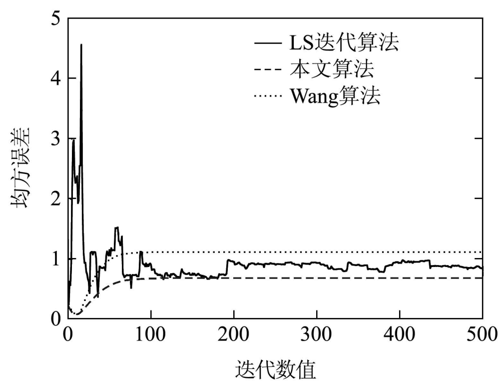

图 1 3种算法均方误差图

Fig. 1 Mean square errors of the three algorithms

从图1可以看出: Wang算法的均方误差最大, 这是因为在输出数据中加入了周期性的冲击信号, 偏离了Wang算法的最优性噪声条件; 本文的自适应密度估计算法的均方误差最小、收敛速度快、抗噪能力强; 最小二乘递推算法均方误差和抗噪能力介于本文算法和Wang算法之间. 另外, 随着迭代次数的

增加, 用本文算法获得的参数估计误差在$10^0 = 1$的数量级, 这是由于噪信比过大的缘故.

通过改变噪声的幅值, 控制噪信比, 获得的均方误差$MSE^l = (\hat{\theta}^l - \theta)^{\mathrm{T}}(\hat{\theta}^l - \theta)$ 和迭代次数的关系如图2所示:

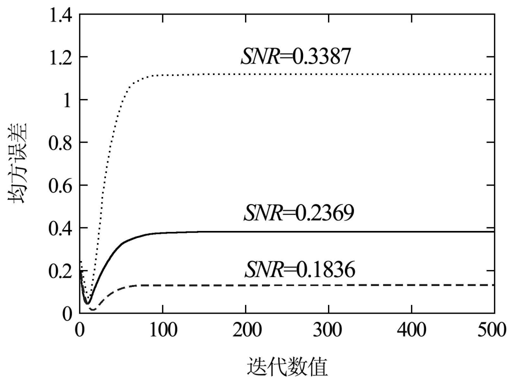

图 2　不同噪信比下均方误差图

Fig. 2　Mean square errors with different algorithm

从图2可以看出: 参数估计均方误差随着噪信比的减小呈指数下降的关系.

6　结束语(Conclusion)

本文提出了一种自适应密度估计方法, 适用于噪声分布未知的系统参数辨识, 通过极小化似然函数, 给出了参数估计的迭代算法, 分析了算法的收敛性, 仿真结果表明算法具有较强的抗噪能力和良好的收敛性.

另外, 本文选择的是高斯核函数, 在参数的每步迭代计算中都要估计全局带宽和局部带宽, 这在观测值数量很大时将严重增加计算开销, 影响数据处理的速度. 因此, 设计逼近效果好, 计算量小的密度估计器将是很有意义的研究方向.

参考文献(References):

[1] HUBER P J. Robust estimation of local parameter[J]. *Annals of Mathematical Statistics*, 1964, 35(1): 73 – 101.

[2] HUBER P J. Robust regression: Asymptotics, conjectures and monte carlo[J]. *Annals of Statistics*, 1973, 1(5): 799 – 821.

[3] HUBER P J. *Robust Statistics*[M]. New York: Wiley, 1981.

[4] 方崇智, 萧德云. 过程辨识[M]. 北京: 清华大学出版社, 1987. (FANG Congzhi, XIAO Deyun. *Process Identification*[M]. Beijing: Press of Tsinghua University, 1987.)

[5] YOHAI V, MARONNA R. Asymptotic behavior of M-estimators for the linear model[J]. *Annals of Statistics*, 1979, 7(2): 258 – 268.

[6] WANG X, POOR H V. Robust multiuser detector in non-Gaussian channels[J]. *IEEE Trans on Signal Processing*, 1999, 47(2): 289 – 305.

[7] 丁锋, 杨家本. 大系统的递阶辨识[J]. 自动化学报, 1999, 25(5): 647 – 654.
(DING Feng, YANG Jiaben. Hierarchical identification of large scale systems[J]. *Acta Automatica Sinica*, 1999, 25(5): 647 – 654.)

[8] ZABIN S, WRIGHT G. Nonparametric density estimation and detection in impulsive interference channels-part Ⅰ: estimation[J]. *IEEE Trans on Communication*, 1994, 42(2/3/4): 1684 – 1697.

[9] ZOUBIN A M, BRCICH R F. Multiuser detection in heavy tailed noise[J]. *Digital Signal Processing*, 2002, 12(2): 262 – 273.

[10] MORAD K, SVRCEK W Y, MCKAY I. Probability density estimation using incomplete data[J]. *ISA Transactions*, 2000, 39(4): 379 – 399.

[11] SILVERMAN B. *Density Estimation for Statistics and Data Analysis*[M]. London: Chapman and Hall, 1986.

[12] DEVROYE L. *A Course in Density Estimation*[M]. Basel: Birkhäuser, 1987.

[13] PHAM D, ZOUBIR A, BRCICH R. Adaptive ML signal detection in non-gaussian channels[C] // *Proc of IEEE Worshop on Statistical Signal Processing(SSP'03)*. Piscataway, NJ, USA: IEEE Press, 2003: 45 – 48.

作者简介:

衷路生　(1979—), 男, 博士研究生, 目前研究方向为系统辨识、信号处理、鲁棒估计, E-mail: lszhongzju@sina.com;

宋执环　(1962—), 男, 博士, 教授, 博士生导师, 目前研究方向为信号处理、故障诊断、预测控制, E-mail: zhsong@iipc.zju.edu.cn.