DOI: 10.7641/CTA.2017.60158

Wiener系统的变聚点样条逼近递推贝叶斯算法

景绍学[1,2†], 李正明[2]

(1. 淮安信息职业技术学院 电气工程系, 江苏 淮安 223003;

2. 江苏大学 电气信息工程学院, 江苏 镇江 212013)

摘要: 为了辨识过程噪声干扰的Wiener非线性系统, 提出了一种基于三样条函数逼近的递推贝叶斯算法. 众所周知, 传统的多项式逼近具有不能外推、高阶易震荡等缺点. 为了克服这些缺点, 首先利用三样条函数对Wiener系统的非线性反函数进行逼近, 在此基础上将待辨识系统参数化为伪线性回归系统. 然后把估计到的噪声方差融入算法, 接着使用递推贝叶斯算法对参数进行了估计. 为了提高三样条函数对非线性反函数的逼近能力, 一种基于均值的变聚点选择方法被应用于算法. 文中还对算法的收敛性进行了分析, 并用数值仿真和案例建模验证了算法的有效性.

关键词: 参数估计; Wiener系统; 过程噪声; 三样条函数; 递推贝叶斯算法; 可变聚点

中图分类号: TP273 **文献标识码**: A

Variable knots spline approximation recursive Bayesian algorithm for identification of Wiener systems

JING Shao-xue[1,2†], LI Zheng-ming[2]

(1. Department of Electrical Engineering, Huai'an College of Information and Technology, Huai'an Jiangsu 223003, China;
2. School of Electrical and Information Engineering, Jiangsu University, Zhenjiang Jiangsu 212013, China)

Abstract: To estimate the Wiener nonlinear systems with process noise, a recursive Bayesian algorithm based on cubic spline approximation is proposed. It's well known that the polynomial approximation does not extrapolate well and high degree polynomials have oscillatory behavior, etc. To overcome these drawbacks, a cubic spline function is used to approximate the inverse function of the output nonlinearity. And then the original Wiener system is parameterized to be a pseudo-linear regression model. The estimated variance of the noise is also integrated in the algorithm to estimate the parameters. In order to approximate the inverse nonlinearity, a mean-value based variable knot-selection method is employed. After the convergence is analyzed, a numerical simulation and a case study validate the algorithm.

Key words: parameter estimation; Wiener system; process noise; cubic spline function; recursive Bayesian algorithm; variable knots

1 引言(Introduction)

Wiener系统是模块化非线性系统的一种, 由一个动态线性模块和一个静态非线性模块串联而成. 这种结构已经在很多实际环境中获得了应用, 如化学过程[1–2]、生物系统[3]、蒸馏塔[4]、色谱分离过程[5]、混沌系统[6]等. 近年来, Wiener系统的辨识问题引起了研究人员的广泛关注[7–10].

为了辨识Wiener系统, 研究人员进行了大量工作, 取得了很多的研究成果. 例如: 基于迭代辨识方法, Wang和Ding提出了最小二乘迭代辨识算法和梯度迭代算法[12], Ding, Ma和Xiao提出了一种牛顿迭代辨识算法[13]. 基于最小二乘原理, Hu等人提出了两种计算量较小的递推辨识算法[14]. Bai和Reyland分析了辨识Wiener系统所需的最少先验知识[15]. 基于EM算法, Vanbeylen等人提出了一种仅用输出数据来辨识Wiener系统的盲辨识方法[16], Xiong等人提出了一种EM辨识算法来辨识输出数据不完整Wiener系统的参数[17]. 针对含有非均匀采样数据的Wiener系统, Zhou, Li和Pan提出了一种梯度迭代辨识算法[18]. 针对多输入多输出Wiener系统, Fan和Lo提出了一种递

收稿日期: 2016−03−22; 录用日期: 2016−08−11.
†通信作者. E-mail: jingsx3@126.com; Tel.: +86 15061411036.
本文责任编委: 邓飞其.
国家自然科学基金项目(51477070), 江苏大学研究生科研创新项目(KYXX_0003)资助.
Supported by National Natural Science Foundation of China (51477070) and Graduate Education Innovation Project of Jiangsu University (KYXX_0003).

推辨识算法, 并分析了算法的收敛性[19]. 针对二值输出Wiener系统, Zhao等人提出了一种参数辨识算法和一种非参数辨识算法[20]. 针对输出不连续分段线性函数Wiener系统, Chen提出了一种递推辨识算法[21]. 另外, 智能方法, 如粒子群优化算法[22–23]、神经网络算法[9,24]、遗传算法[25–26]和进化计算方法[27], 在Wiener系统的辨识问题中也获得了广泛应用. 由于本文考虑传统辨识算法而非智能方法来辨识Wiener系统参数, 故不对上述智能方法展开叙述. 虽然上述非智能算法针对不同的Wiener系统的辨识都获得了成功, 但也存在着一些不容忽视的问题:

1) 考虑的噪声大都是输出噪声, 这种噪声不随过程增益的变化而变化, 易于建模. 而很多实际情况却是: 过程增益越大, 噪声对输出的影响就越大, 过程增益变小, 其对输出的影响也会变小.

2) 非线性环节传递函数的反函数(简称为非线性反函数)的逼近方式大都采用了多项式逼近方法. 众所周知, 多项式具有以下不足[28]: a) 不能很好地外推; b) 高阶多项式具有震荡特性; c) 参数估计过程中经常出现病态现象.

3) 没有充分利用估计到的噪声信息.

针对这些问题, 本文提出了一种基于可变聚点样条函数逼近的递推贝叶斯算法. 本文的其他部分组织如下: 第2节介绍了待辨识的Wiener系统, 并引出了本文研究的问题. 在此基础上, 第3节推导了可变聚点三样条函数逼近递推贝叶斯算法, 在第4节里对算法的收敛性进行了分析. 随后的第5节里用一个数值仿真和一个实例对所提算法进行了验证. 最后一节给出了一些主要结论.

2 问题描述(Problem description)

考虑如图1所示的离散Wiener非线性系统. 图1中: $u(k)$和$y(k)$分别为系统的输入、输出信号, $v(k)$为均值为零的高斯白噪声信号, $B(Z^{-1})$为线性动态模块传递函数, $N(x_2)$为静态非线性传递函数. $x_1(k)$和$x_2(k)$为不可测中间变量.

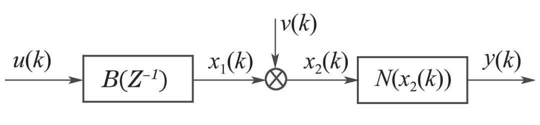

图 1 含过程噪声的Wiener系统框图

Fig. 1 Wiener systems with process noise

众所周知, 只要阶次足够高, 有限脉冲响应函数(finite impulse response, FIR)可以以任意精度逼近稳定的线性传递函数, 故设待辨识Wiener系统的动态线性模块的传递函数为FIR 形式(或者说可以由一个FIR函数逼近):

$$x_1(k) = B(z^{-1})u(k), \tag{1}$$

其中$B(z^{-1}) = b_1z^{-1} + b_2z^{-2} + \cdots + b_{n_\mathrm{b}}z^{-n_\mathrm{b}}$.

为了克服多项式逼近的缺点, 采用如下的三样条函数对输出非线性反函数$N^{-1}(y)$进行逼近,

$$\begin{aligned}&N^{-1}(y(k)) = x_2(k) = \\&\sum_{i=2}^{m-1}\gamma_i\,|y(k)-y_i|^3 + \gamma_m + y(k),\end{aligned} \tag{2}$$

其中: $y_i(i=1,2,\cdots,m)$为聚点集Y中的聚点, 并且满足$\min(y(j)) = y_1 < y_2 < \cdots < y_{m-1} < y_m = \max(y(j))$ $(j=1,2,\cdots,L)$, L为采样数据的个数. 当$i=2,\cdots,m-1$时的聚点y_i称为内部聚点, 内部聚点的取值由用户选择, 具体的选择方法将在第3.2节讨论. m表示三样条函数的阶次, 亦由用户选择, 其选择方法见第3.4节.

为了保证参数辨识能够进行, 给出如下的假设:

1) 非线性函数$N(\cdot)$连续、单调且存在反函数[29–30];

2) 过程噪声$v(k)$是均值为零方差为σ_v^2的高斯白噪声序列;

3) FIR阶次n_b已知.

在上述假设的前提下, 含过程噪声的Wiener系统的参数辨识问题可以描述为: 根据观测到的数据$D^{(k)} = \{u(i), y(i)\}_{i=1}^{k}$, 估计式(1)–(2)中的各多项式的系数.

3 可变聚点样条逼近递推贝叶斯算法(Cubic spline approximation recursive Bayesian algorithm with variable knots)

3.1 系统参数化(System parameterization)

根据式(1)和(2), 整理可得

$$\begin{aligned}&\sum_{i=2}^{m-1}\gamma_i\,|y(k)-y_i|^3 + \gamma_m + y(k) = \\&B(z^{-1})u(k) + v(k),\end{aligned} \tag{3}$$

于是待辨识Wiener系统可以参数化为

$$y(k) = \phi^\mathrm{T}(k)\theta + v(k), \tag{4}$$

其中:

$$\begin{cases}\phi(k) = [-|y(k)-y_2|^3, \cdots, -|y(k)-y_{m-1}|^3, \\ \qquad -1, u(k-1), \cdots, u(k-n_\mathrm{b})] \in \mathbb{R}^{n\times 1}, \\ \theta = [\gamma_2, \cdots, \gamma_m, a_1, \cdots, a_{n_a}] \in \mathbb{R}^{n\times 1}, \\ n = n_a + m - 1.\end{cases} \tag{5}$$

3.2 内部聚点的选择方法(Selection of internal knots)

从式(5)不难看出, 内部聚点y_i的选择对于非线性反函数的逼近效果有重要影响[32], 从而对整个系统的辨识效果产生直接影响. 鉴于内部聚点取值的重要性,

下面讨论内部聚点的选择方法.

1) 平均分布法(even spaced method, ESM).

顾名思义, 这种选择方法使得内部聚点取值区间里均匀分布, 即

$$y_i = \min y + \frac{i-1}{m-1}(\max y - \min y), \tag{6}$$

其中$i = 2, \cdots, m-1$.

2) 试凑法(trial and error method, TEM).

在这种方法中, 用户需首先定义一个代价函数, 然后依据代价函数的变化情况改变聚点的取值. 一个典型的代价函数为

$$J = \frac{1}{L}\sum_{k=1}^{L}[x_2(k) - \hat{x}_2(k)]^2, \tag{7}$$

其中: $\hat{x}_2(k) = \sum_{i=2}^{m-1}\hat{\gamma}_i(k)|y(k) - y_i| + \gamma_m + y(k)$, $x_2(k) = N^{-1}(y(k))$, 用户通过不断地改变y_i 的取值以获得较小的J.

上述方法中, 平均分布法具有易于操作、计算量小等特点. 对于均匀分布的y, 平均分布法可以获得比较好的内部聚点. 试凑法以逼近效果作为代价函数, 可以获得比较好的逼近效果, 对y的分布也无要求, 但是随意性较强, 并且计算量也比平均法大. 由于非线性模块的存在, 导致Wiener系统中的输出变量y一般都不是均匀分布. 为了获得比较好的内部聚点, 本文基于y的统计信息提出了一种聚点选择方法:

3) 均值法(mean value method, MVM).

该方法选择区间内y的平均值作为本区间内中心聚点的取值. 若样条函数阶次$m = 5$, 由于聚点集的端点聚点$y_1 = \min y$, $y_5 = \max y$是确定的, 于是其y_1, y_5的中心聚点$y_3 = \mathrm{mean}(y)$, y_1, y_3的中心聚点$y_2 = \mathrm{mean}(y_i\,|y_1 \leqslant y_i < y_3)$, 同理$y_3, y_5$的中心聚点$y_4 = \mathrm{mean}(y_i\,|y_3 \leqslant y_i < y_5)$.

该方法以均值作为内部聚点的选择依据, 符合在变量分布较多的地方选择聚点的原则. 和试凑法一样, 本方法对y的分布无特殊要求. 不同的是本方法无需计算代价函数. 如果变量y服从均匀分布, 则本方法与平均法得到的结果相同.

3.3 递推贝叶斯算法 (Recursive Bayesian algorithm)

根据贝叶斯理论, 参数的贝叶斯估计可以最大化该参数的后验概率密度得到, 即

$$\hat{\theta} = \arg\max_{\theta} p(\theta|D^{(k)}). \tag{8}$$

根据贝叶斯概率公式, 参数的后验概率密度函数表示为

$$p(\theta|D^{(k)}) = \frac{p(y(k)|\theta, D^{(k-1)})p(\theta|D^{(k-1)})}{p(y(k)|D^{(k-1)})}, \tag{9}$$

其中$p(y(k)\,|\theta, D^{(k-1)})$为在给定参数$\theta$和数据$D^{(k-1)}$条件下输出$y(k)$的先验概率密度.

假设$p(\theta\,|D^{(k-1)})$服从均值为$\hat{\theta}(k-1)$方差为$\boldsymbol{P}(k-1)$的正态分布, 并考虑到式(4)和噪声的概率分布, 可知输出也服从正态分布, 即, $y(k) \sim N(\phi^{\mathrm{T}}(k)\theta, \sigma_{\mathrm{v}}^2)$, 因此数据的先验概率密度函数为

$$p(y(k)|\theta, D^{(k-1)}) = \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{v}}}\exp(-\frac{1}{2}(\frac{y(k) - \phi^{\mathrm{T}}(k)\hat{\theta}(k-1)}{\sigma_{\mathrm{v}}})^2), \tag{10}$$

于是参数的后验密度函数$p(\theta\,|D^{(k)})$的表达式为

$$p(\theta|D^{(k)}) = K\exp(-\frac{1}{2}(\frac{y(k) - \phi^{\mathrm{T}}(k)\hat{\theta}(k-1)}{\sigma_{\mathrm{v}}}) - \frac{1}{2}(\theta - \hat{\theta}(k-1))^{\mathrm{T}}\boldsymbol{P}^{-1}(k-1)(\theta - \hat{\theta}(k-1))), \tag{11}$$

其中K是一个与参数θ不相关的常数.

令$\frac{\partial \log p(\theta\,|D^{(k)})}{\partial\theta}|_{\theta=\hat{\theta}(k)} = 0$, 求导后整理得到

$$[(\theta - \hat{\theta}(k))^{\mathrm{T}}\boldsymbol{P}^{-1}(k)(\theta - \hat{\theta}(k))] = 0, \tag{12}$$

其中:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{1}{\sigma_{\mathrm{v}}^2}\boldsymbol{P}(k)\phi(k)[y(k) - \phi^{\mathrm{T}}(k)\hat{\theta}(k-1)], \tag{13}$$

$$\boldsymbol{P}^{-1}(k) = \boldsymbol{P}^{-1}(k-1) + \frac{1}{\sigma_{\mathrm{v}}^2}\phi(k)\phi^{\mathrm{T}}(k). \tag{14}$$

式(13)–(14)为紧凑形式的递推贝叶斯算法, 这种形式常用于理论分析. 为避免矩阵求逆运算, 对式(14)运用矩阵求逆反演公式, 并用σ_{v}^2在$(k-1)T$时刻的估计值$\hat{\sigma}_{\mathrm{v}}^2(k-1)$代替未知的噪声方差, 得到递推贝叶斯(RB)公式:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \boldsymbol{L}(k)[y(k) - \phi^{\mathrm{T}}(k)\hat{\theta}(k-1)], \tag{15}$$

$$\boldsymbol{L}(k) = \frac{\boldsymbol{P}(k-1)\phi(k)}{\sigma_{\mathrm{v}}^2(k-1) + \phi^{\mathrm{T}}(k)\boldsymbol{P}(k-1)\phi(k)}, \tag{16}$$

$$\boldsymbol{P}(k) = [\boldsymbol{I} - \boldsymbol{L}(k)\phi^{\mathrm{T}}(k)]\boldsymbol{P}(k-1). \tag{17}$$

噪声方差可以用下式递推估计:

$$\begin{cases} \hat{v}(k) = y(k) - \phi^{\mathrm{T}}(k)\hat{\theta}(k), \\ \hat{\sigma}_{\mathrm{v}}^2(k) = \dfrac{1}{k-1}\left[(k-2)\hat{\sigma}_{\mathrm{v}}^2(k-1) + \hat{v}^2(k)\right]. \end{cases} \tag{18}$$

由于本算法利用三样条函数逼近输出非线性反函数, 然后递推贝叶斯算法进行参数估计, 所以所提出的算法被称为三样条逼近递推贝叶斯算法 (cubic spline approximation RB algorithm, CSA–RB). 算法的

步骤如下:

Step 1 初始化;

Step 2 收集输入输出数据;

Step 3 用样条函数逼近非线性函数的反函数;

Step 4 参数化Wiener系统;

Step 5 用RB算法估计参数乘积项;

Step 6 估计噪声方差;

Setp 7 令$k=k+1$, 如果$k<L$, 转Step 2; 否则结束.

3.4 样条阶次的选择方法(Selection of spline order)

模型定阶方法大致可分为3种[31−32], 第1种需要指定概率测度的置信区间作为阶次检验的标准, 如F检验定阶法; 第2种方法不需要指定概率测度, 如行列式比定阶法; 第3种方法需要确定一个准则函数来作为阶次检验的标准, 将定阶问题转化为最优问题, 如最终预报误差(FPE)准则定阶法[33]. 由于易于操作, 第3种方法被广泛采用. 针对本文的Wiener系统, FPE法的准则函数定义为

$$\mathrm{FPE}(m)=\frac{L+(n_{\mathrm{b}}+m-2)}{L-(n_{\mathrm{b}}+m-2)}\frac{1}{L}\sum_{i=1}^{L}[y(i)-\hat{y}(i)]^2. \tag{19}$$

3.5 变聚点CSA–RB算法(CSA–RB algorithm with variable knots)

为了提高样条函数对非线性反函数的逼近能力, 采用了变聚点策略. 其主要思想是按照一定的规则改变聚点的位置以减小逼近误差、提高逼近效果. 下面以聚点y_i的改变步骤来介绍变聚点的具体操作.

1) 首先用第3.2节介绍的一种方法在y_i的左右两侧分别选择两个临时聚点$y_{i\mathrm{L}}$和$y_{i\mathrm{R}}$;

2) 分别用y_i, $y_{i\mathrm{L}}$和$y_{i\mathrm{R}}$作为第i个聚点, 按照式(7)计算得到3个代价函数值J_i, $J_{i\mathrm{L}}$和$J_{i\mathrm{R}}$;

3) 用如下的规则改变第i个聚点的取值:

$$y_i=\begin{cases}y_{i\mathrm{L}},\ J_{i\mathrm{L}}=\min(J_{i\mathrm{L}},J_{i\mathrm{R}})<J_i,\\ y_i,\quad J_i<\min(J_{i\mathrm{L}},J_{i\mathrm{R}}),\\ y_{i\mathrm{R}},\ J_{i\mathrm{R}}=\min(J_{i\mathrm{L}},J_{i\mathrm{R}})<J_i.\end{cases} \tag{20}$$

按照从y_2到y_{m-1}或者相反的顺序依次改变各聚点的位置, 从而完成一次迭代. 当准则函数小于设定值J_{s}或迭代次数大于$i\mathrm{tr}_{\max}$时, 结束迭代.

变聚点CSA–RB算法(VKCSA–RB)的流程图如图2所示.

4 收敛性分析(Convergence analysis)

仿照式(13)–(14), 将所提算法转化为如下的等效形式:

$$\hat{\theta}(k)=\hat{\theta}(k-1)+\frac{1}{\sigma_{\mathrm{v}}^2(k-1)}\boldsymbol{P}(k)\phi(k)\times [y(k)-\phi^{\mathrm{T}}(k)\hat{\theta}(k-1)], \tag{21}$$

$$\boldsymbol{P}^{-1}(k)=\boldsymbol{P}^{-1}(k-1)+\frac{1}{\sigma_{\mathrm{v}}^2(k-1)}\phi(k)\phi^{\mathrm{T}}(k). \tag{22}$$

于是得到下面的收敛性定理:

定理 对于系统(4)和CSA–RB算法(21)和(22), 假设噪声$v(k)$为均值为零方差为σ_{v}^2的高斯白噪声, 且与信息向量$\phi(k)$无关, 若存在正数α,β和正整数k_0使得下面的持续激励条件成立:

$$\alpha\boldsymbol{I}_n\leqslant\frac{1}{k}\sum_{j=1}^{k}\phi(j)\phi^{\mathrm{T}}(j)\leqslant\beta\boldsymbol{I}_n,\ \text{a.s.},\ k\geqslant k_0, \tag{23}$$

那么算法给出的参数估计值收敛于真值, 即

$$\lim_{k\to\infty}\hat{\theta}(k)=\theta_0, \tag{24}$$

其中θ_0为参数真值.

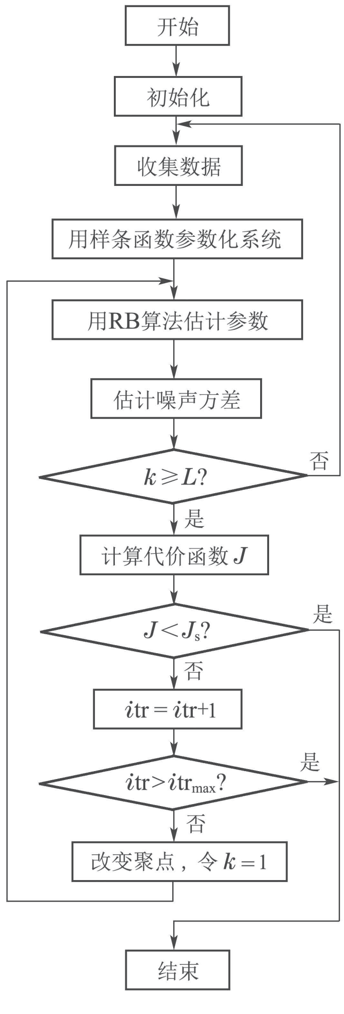

图 2 VKCSA–RB算法流程图

Fig. 2 Flowchart of VKCSA–RB algorithm

证 在式(21)两边同时减去θ_0并用$\phi^{\mathrm{T}}(k)\theta_0+v(k)$代替$y(k)$, 得到

$$\tilde{\theta}(k)=\boldsymbol{Q}(k)\tilde{\theta}(k-1)+\boldsymbol{S}(k)v(k), \tag{25}$$

其中:

$$\begin{cases}\widetilde{\theta}(k)=\theta(k)-\theta_0, \widetilde{\theta}(k-1)=\theta(k-1)-\theta_0,\\ \boldsymbol{Q}(k)=[\boldsymbol{I}-\dfrac{1}{\sigma_{\rm v}^2(k-1)}\boldsymbol{P}(k)\phi(k)\phi^{\rm T}(k)],\\ \boldsymbol{S}(k)=\dfrac{1}{\sigma_{\rm v}^2(k-1)}\boldsymbol{P}(k)\phi(k).\end{cases}$$

众所周知, 协方差矩阵$\boldsymbol{P}(k)$满足$\lim\limits_{k\to\infty}\boldsymbol{P}(k)=0$. 考虑到$v(k)$为白噪声且与$\phi(k)$无关, 可知$\{\tilde{\theta}^*=\boldsymbol{0}\}$为方程(25)的不变集. 那么问题(24)转化为差分方差(25)的稳定性问题. 在(22)两边同乘以$\boldsymbol{P}(k)$, 则$\boldsymbol{Q}(k)$写为

$$\boldsymbol{Q}(k)=\boldsymbol{I}-\frac{1}{\sigma_{\rm v}^2(k-1)}\boldsymbol{P}(k)\phi(k)\phi^{\rm T}(k)= \boldsymbol{P}(k)\boldsymbol{P}^{-1}(k-1). \tag{26}$$

设λ为矩阵$\boldsymbol{Q}(k)$的一个特征值, $\boldsymbol{p}$为其对应的非零特征向量, 即

$$\boldsymbol{Q}(k)\boldsymbol{p}=\lambda\boldsymbol{p}. \tag{27}$$

再考虑到式(26), 有

$$\boldsymbol{P}(k)\boldsymbol{P}^{-1}(k-1)\boldsymbol{p}=\lambda\boldsymbol{p}. \tag{28}$$

在上式两边同乘以$\boldsymbol{P}^{-1}(k)$, 考虑到式(22), 并整理得

$$(1-\lambda)\boldsymbol{P}^{-1}(k-1)\boldsymbol{p}=\lambda\frac{1}{\sigma_{\rm v}^2(k-1)}\phi(k)\phi^{\rm T}(k)\boldsymbol{p}. \tag{29}$$

两边同乘以$\boldsymbol{p}^{\rm T}$并整理得

$$(1-\lambda)\boldsymbol{p}^{\rm T}\boldsymbol{P}^{-1}(k-1)\boldsymbol{p}=\lambda\frac{1}{\sigma_{\rm v}^2(k-1)}\boldsymbol{p}^{\rm T}\phi(k)\phi^{\rm T}(k)\boldsymbol{p}. \tag{30}$$

考虑到$\boldsymbol{P}(k)$, $\boldsymbol{P}^{-1}(k-1)$和$\phi(k)\phi^{\rm T}(k)$的正定性, 知$\boldsymbol{p}^{\rm T}\boldsymbol{P}^{-1}(k-1)p$和$\dfrac{1}{\sigma_{\rm v}^2(k-1)}p^{\rm T}\phi(k)\phi^{\rm T}(k)\boldsymbol{p}$均大于0. 从而式(30)中的项$(1-\lambda)$和$\lambda$是正负同号的, 即$\boldsymbol{Q}(k)$的特征值$\lambda$满足$0<\lambda<1$. 根据差分方程稳定性理论, 差分方差(25)是稳定的, 即: $\lim\limits_{k\to\infty}\widetilde{\theta}(k)=0$, 因此定理1成立.

5　模型验证(Model validation)

5.1　数值仿真(Numerical simulation)

考虑如下的离散Wiener非线性系统:

$$\begin{cases}y(k)=N\left[B(z^{-1})u(k)+v(k)\right],\\ N(x_2)=x_2+2x_2^3,\\ B(z^{-1})=0.1057z^{-1}+0.2615z^{-2}+\\ \qquad 0.3066z^{-3}+0.2751z^{-4},\end{cases}$$

其中: $u(k)$是范围为$[0,0.4]$的平均分布随机数, 噪声取均值为零的高斯白噪声.

1) 三样条函数阶次m的估计.

在不同方差下运用最终预报误差准则(FPE)方法对三样条函数的阶次进行了估计. 表1显示了不同噪声情况下预报误差随三样条函数阶次的变化情况, 可见当$m=8$时预报误差最小. 表中的聚点用ESM方法选取.

表 1　预报误差随m变化情况

Table 1　FPEs versus m using FKSA–RB–ESM

m	$\sigma_{\rm v}^2=0.005^2$	$\sigma_{\rm v}^2=0.01^2$	$\sigma_{\rm v}^2=0.02^2$
3	1.63E−04	5.25E−04	2.21E−02
4	2.15E−05	1.00E−03	6.57E−05
5	1.30E−05	3.74E−06	3.60E−07
6	5.50E−08	5.72E−09	1.89E−08
7	7.71E−10	9.91E−10	1.79E−09
8	**1.42E−10**	**1.42E−10**	**4.62E−10**
9	3.75E−10	1.27E−08	6.39E−09
10	1.13E−08	5.49E−09	1.19E−09
11	4.94E−09	1.42E−09	5.94E−08

2) 固定聚点ESM和MVM方法的估计结果的比较(FKCSA–RB–ESM和FKCSA–RB–MVM).

为了比较ESM和MVM这两种聚点选择方法的优劣, 表2比较了使用这两种方法的FKSA–RB算法的估计结果. 其中线性部分估计误差$\delta_{\rm L}$定义为

$$\delta_{\rm L}=\frac{\text{norm}([\hat{b}_1,\hat{b}_2,\hat{b}_3,\hat{b}_4]-[b_1,b_2,b_3,b_4])}{\text{norm}([b_1,b_2,b_3,b_4])}\times 100.$$

非线性逼近误差J用式(7)计算. 综合误差CE定义为

$$\text{CE}_i=\frac{n_{\rm b}\delta_{{\rm L},i}}{\text{mean}(\delta_{\rm L})}+\frac{(m-1)J_i}{\text{mean}(J)},$$

其中i表示不同的辨识结果, $n_{\rm b}$和m分别为线性和非线性模块的阶次. 辨识时噪声的方差为0.01^2.

表 2　使用ESM和MVM的FKCSA–RB算法辨识结果

Table 2　Results using FKCSA–RB with ESM and MVM

	参数	ESM	MVM
内部聚点	y_2	0.0719	0.1070
	y_3	0.1361	0.1557
	y_4	0.2004	0.2064
	y_5	0.2646	0.2533
	y_6	0.3289	0.3115
	y_7	0.3931	0.3468
线性环节	b_1	0.1441	0.1520
	b_2	0.2521	0.2461
	b_3	0.2646	0.2954
	b_4	0.3111	0.2831
	$\delta_{\rm L}$/%	13.69	10.36
非线性环节	J	8.96E − 04	4.11E − 04
综合误差	CE	14.17	7.85

从表2不难看出, 不论是线性环节的估计误差、非线性反函数逼近误差还是综合估计误差, 所提出

的MVM方法都比传统的ESM方法的要小, 这表明所提的聚点选择算法可以获得比后者更好的辨识结果.

3) 固定聚点和可变聚点CSA–RB–MVM的估计结果的比较.

为了显示可变聚点方法的优越性, 作者比较了固定聚点和可变聚点CSA–RB–MVM算法的估计结果, 如表3所示, 表中噪声方差为0.01^2, 可变聚点迭代次数为3.

表 3 固定聚点和可变聚点的CSA–RB–MVM算法辨识结果

Table 3 Results using CSA–RB–MVM with fixed and variable knots

	变量	固定聚点	可变聚点
内部聚点	y_2	0.1070	0.1587
	y_3	0.1557	0.1956
	y_4	0.2064	0.1991
	y_5	0.2533	0.2138
	y_6	0.3115	0.2649
	y_7	0.3468	0.3468
线性环节	b_1	0.1520	0.1184
	b_2	0.2461	0.2539
	b_3	0.2954	0.3082
	b_4	0.2831	0.2787
	$\delta_{\mathrm{L}}/\%$	10.36	3.47
非线性环节	J	$4.11\mathrm{E}-04$	$7.6\mathrm{E}-05$
综合误差	CE	17.80	4.20

从表3可以看出, 可变聚点的线性环节参数估计误差、非线性环节逼近误差、综合误差分别为固定聚点时的33.49%, 18.54%和23.60%, 3个指标均有大幅度的减小, 表明可变聚点可以有效地提高辨识的精度.

4) FKCSA–RB–MVM与FKCSA–SG–MVM的估计结果的比较.

为了显示本文所用的递推贝叶斯参数辨识算法在固定聚点情况下的优越性, 将该方法和随机梯度辨识算法的辨识结果进行了比较, 结果见表4. 图中, 聚点与表2中"MVM"列相同.

从表4可以看出, 本文所用RB算法得到的辨识结果的估计误差要小于随机梯度算法的估计误差, 特别是对线性环节的估计误差, 本文算法显著地小于对比算法.

5) VKCSA–RB–MVM与VKCSA–SG–MVM的估计结果的对比.

为了显示本文所用RB算法在变聚点情况下的性能, 将其辨识结果与SG算法得到的辨识结果进行了比较, 如表5所示. 对比表4和表5的SG列可以发现, 变聚点对SG算法的估计结果没有明显影响, 而对比两表的RB列则完全不同: 变聚点RB算法的估计精度要远远高于固定聚点. 对比表5中的两组数据不难看出, 变聚点RB算法的估计误差要远远小于SG算法, 前者的估计误差仅为后者的15%, 显示出明显的优势.

表 4 FKCSA–RB–MVM与FKCSA–SG–MVM的估计结果

Table 4 Results using FKCSA–RB–MVM and FKCSA–SG–MVM

	变量	SG	RB
线性环节	b_1	0.1896	0.1520
	b_2	0.1840	0.2461
	b_3	0.3034	0.2954
	b_4	0.2988	0.2831
	$\delta_{\mathrm{L}}/\%$	23.77	10.36
非线性环节	J	$4.96\mathrm{E}-04$	$4.11\mathrm{E}-04$
综合误差	CE	13.23	8.77

表 5 VKCSA–RB–MVM与VKCSA–SG–MVM的估计结果

Table 5 Results using VKCSA–RB–MVM and VKCSA–SG–MVM

	变量	SG	RB
线性环节	b_1	0.1896	0.1184
	b_2	0.1840	0.2539
	b_3	0.3034	0.3082
	b_4	0.2988	0.2787
	$\delta_{\mathrm{L}}/\%$	23.77	3.47
非线性环节	J	$4.96\mathrm{E}{-}04$	$7.60\mathrm{E}{-}05$
综合误差	CE	19.12	2.88

6) 不同噪声水平下VKCSA–RB–MVM算法的辨识结果.

为了验证所提的算法在不同噪声水平下的辨识性能, 本文选择噪声方差分别为0.01^2, 0.03^2和0.05^2时进行了仿真, 表6显示了FKCSA–RB–MVM算法在这3个噪声水平时的线性环节参数的估计值, 表7为相应的参数估计值的均值和标准差. 由于不同噪声水平下聚点的位置不同, 导致得到的非线性环节的参数不同, 它们之间没有可比性, 故表6–7中没有列出这部分参数的信息.

从表6–7可以看出, 随着噪声方差的增大, 线性部分、非线性部分和整个系统的参数估计误差随之变大, 表7中各参数估计值的标准差也随之增大. 综上所述, 本文所提出的变聚点CSA–RB–MVM算法在与固定

聚点CSA–RB–MVM算法、变聚点CSA–SG–MVM算法和变聚点CSA–RB–ESM算法的比较中, 都显示了该算法的具有良好辨识能力.

表 6　VKCSA–RB–MVM算法在不同噪声水平时的辨识结果

Table 6　Results using VKCSA–RB–MVM under different noise levels

	变量	0.01^2	0.03^2	0.05^2
线性环节	b_1	0.1184	0.1171	0.0969
	b_2	0.2539	0.2880	0.3033
	b_3	0.3082	0.3061	0.2737
	b_4	0.2787	0.2465	0.2736
	δ_L/%	3.47	7.67	10.21
非线性环节	J	7.60E−05	2.80E−04	4.80E−04
综合误差	CE	3.85	11.34	17.80

表 7　VKCSA–RB–MVM算法在不同噪声水平时的辨识结果的统计信息

Table 7　Statistical data of estimates using VKCSA–RB–MVM under different noise levels

变量	0.01^2	0.03^2	0.05^2
b_1	0.1109 ± 0.0092	0.1472 ± 0.0567	0.1064 ± 0.0728
b_2	0.2581 ± 0.0060	0.2981 ± 0.0617	0.2951 ± 0.0779
b_3	0.3074 ± 0.0083	0.2886 ± 0.0336	0.2387 ± 0.0682
b_4	0.2750 ± 0.0077	0.2421 ± 0.0346	0.2783 ± 0.0482

5.2　应用实例(Case study)

本节使用文献[34]提供的tank模型对所提算法进行验证. 建模对象为下层水箱的水位. 系统的输入为施加在水泵上的电压, 输出为Tank 2水箱的水位. 采样周期为4.0 s, 采样数据长度为1500. 前1000组数据用于辨识参数, 后面500组数据用于模型验证. 输入输出数据波形分别见图3–4.

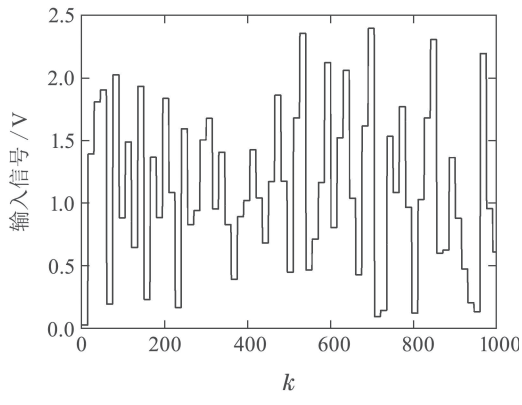

图 3　输入数据波形

Fig. 3　Input signal of Tank 2

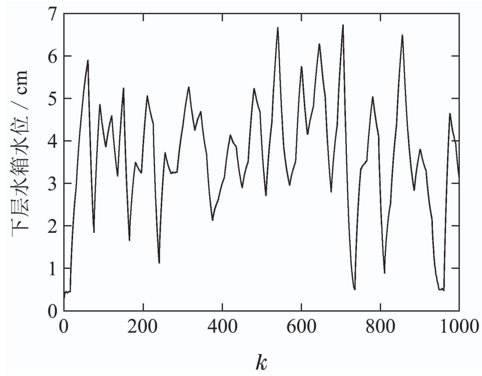

图 4　输出数据波形

Fig. 4　Output signal of Tank 2

用带过程噪声的Wiener系统进行辨识. 设待辨识系统的模型如图1所示, 其中:

$$\begin{cases} x_2(k) = B(z^{-1})u(k) + v(k), \\ x_2(k) = \sum\limits_{i=2}^{m-1} \gamma_i \left|y(k) - y_i\right|^3 + \gamma_m + y(k), \\ B(z^{-1}) = b_1 z^{-1} + \cdots + b_{n_b} z^{-n_b}. \end{cases}$$

运用本文所提出的CSA–RB–MVM算法, 用FPE定阶法, 得到线性环节和样条函数的阶次分别为2和6. 得到的辨识结果为

$$[b_1, b_2, \gamma_2] = [-0.003246, -0.001056, -0.06693],$$
$$[\gamma_3, \gamma_4, \gamma_5, \gamma_6] =$$
$$[0.1345, -0.1449, 0.07106, -3.612].$$

各参数随数据长度变化情况如图5所示.

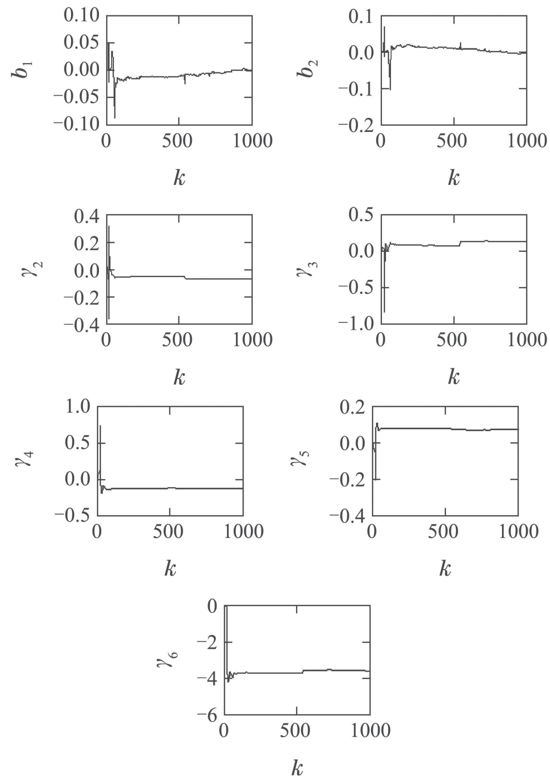

图 5　Tank 2模型参数估计值随k变化情况

Fig. 5　Estimates of Tank 2 versus k using CSR–RB

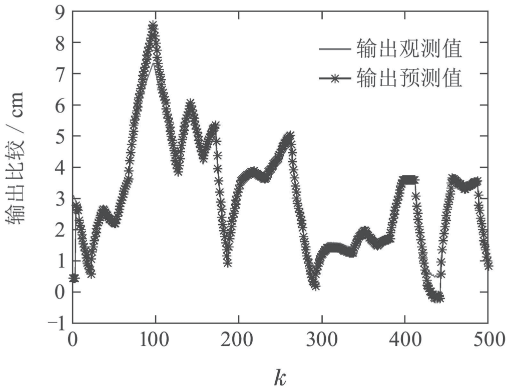

图 6 输出观测值和预测值的比较

Fig. 6 Comparison of observed output and predicted output

利用估计的参数, 和后500组数据的输入, 对输出进行预测, 得到的输出预测值如图6中"*"所示, 图中实线为输出观测值. 此次预测的均方误差和相对均方误差分别为

$$\mathrm{MSE}=\frac{1}{500}\sum_{i=1}^{500}[y_2(i)-\hat{y}_2(i)]^2=0.1167,$$

$$\mathrm{RMSE}=\frac{\mathrm{MSE}}{\mathrm{mean}(y_2(i))}\times 100=3.77\%,$$

其中$y_2(i)$和$\hat{y}_2(i)$分别为交叉验证部分的输出观测值和输出预测值. 从两者对比和两个误差可以看出, 输出预测值与输出观测值吻合的很好, 表明获得的辨识模型是令人满意的.

6 结论(Conclusions)

为了辨识含过程噪声的Wiener非线性系统, 出了一种基于三样条函数逼近的递推贝叶斯算法. 在参数化过程中, 为了避免多项式逼近的弊端, 利用三样条函数对Wiener系统的非线性反函数进行逼近. 在此基础上, 将过程噪声模型参数化为伪线性回归模型. 为了提高辨识精度, 将估计到的噪声方差融入算法, 并提出了一种基于均值原理的变聚点选择方法. 数值仿真结果表明, 与固定聚点算法和平均分布聚点方法先比, 所提方法可以显著地提高辨识精度. 真实数据建模验证了算法的有效性. 该算法可应用于其他模块化非线性系统的可逆输出非线性函数的辨识中.

参考文献(References):

[1] SHAFIEE G, AREFI M M, JAHED-MOTLAGH M R, et al. Nonlinear predictive control of a polymerization reactor based on piecewise linear Wiener model [J]. *Chemical Engineering Journal*, 2008, 143(1): 282 – 292.

[2] KALAFATIS A D, WANG L, CLUETT W R. Linearizing feedforward–feedback control of pH processes based on the Wiener model [J]. *Journal of Process Control*, 2005, 15(1): 103 – 112.

[3] SILVA M M, WIGREN T, MENDONCA T. Nonlinear identification of a minimal neuromuscular blockade model in anesthesia [J]. *IEEE Transactions on Control Systems Technology,* 2012, 20(1): 181 – 188.

[4] BLOEMEN H H J, CHOU C T, VAN DEN BOOM T J J, et al. Wiener model identification and predictive control for dual composition control of a distillation column [J]. *Journal of Process Control*, 2001, 11(6): 601 – 620.

[5] ARTO V, HANNU P, HALME A. Modeling of chromatographic separation process with Wiener-MLP representation [J]. *Journal of Process Control*, 2001, 11(5): 443 – 458.

[6] CHEN G, CHEN Y, OGMEN H. Identifying chaotic systems via a Wiener-type cascade model [J]. *IEEE Control Systems*, 1997, 17(5): 29 – 36.

[7] HAGENBLAD A, LJUNG L, WILLS A. Maximum likelihood identification of Wiener models [J]. *Automatica*, 2008, 44(11): 2697 – 2705.

[8] WU Dehui. Identification method for nonlinear dynamic system using Wiener neural network [J]. *Control Theory & Applications*, 2009, 26(11): 1192 – 1196.
(吴德会. 非线性动态系统的Wiener神经网络辨识法 [J]. 控制理论与应用, 2009, 26(11): 1192 – 1196.)

[9] ZHANG Y, LI Shaoyuan, WANG Xiaobo, et al, Particle swarm optimal identification of Wiener model and a case study [J]. *Control Theory & Applications*, 2006, 23(6): 991 – 995.
(张艳, 李少远, 王笑波, 等. 基于粒子群优化的Wiener模型辨识与实例研究 [J]. 控制理论与应用, 2006, 23(6): 991 – 995.)

[10] LI Shihua, WU Fubao, LI Qi. Identification of Wiener model using dynamic artificial neural networks [J]. *Control Theory & Applications*, 2000, 17(1): 92 – 103.
(李世华, 吴福保, 李奇. 一种基于动态人工神经网络的Wiener模型辨识 [J]. 控制理论与应用, 2000, 17(1): 92 – 103.)

[11] TIELS K, SCHOUKENS J. Wiener system identification with generalized orthonormal basis functions [J]. *Automatica*, 2014, 50(12): 3147 – 3154.

[12] WANG D, DING F. Least squares based and gradient based iterative identification for Wiener nonlinear systems [J]. *Signal Processing*, 2011, 91(5): 1182 – 1189.

[13] DING F, MA J, XIAO Y. Newton iterative identification for a class of output nonlinear systems with moving average noises [J]. *Nonlinear Dynamics*, 2013, 74(1/2): 21 – 30.

[14] HU Y, LIU B, ZHOU Q, et al. Recursive extended least squares parameter estimation for Wiener nonlinear systems with moving average noises [J]. *Circuits, Systems, and Signal Processing*, 2014, 33(2): 655 – 664.

[15] BAI E W, REYLAND J. Towards identification of Wiener systems with the least amount of a priori information on the nonlinearity [J]. *Automatica*, 2008, 44(4): 910 – 919.

[16] VANBEYLEN L, PINTELON R, SCHOUKENS J. Blind maximum-likelihood identification of Wiener systems [J]. *IEEE Transactions on Signal Processing*, 2009, 57(8): 3017 – 3029.

[17] XIONG W, YANG X, KE L, et al. EM algorithm-based identification of a class of nonlinear Wiener systems with missing output data [J]. *Nonlinear Dynamics*, 2015, 80(1/2): 329 – 339.

[18] ZHOU L, LI X, PAN F. Gradient-based iterative identification for Wiener nonlinear systems with non-uniform sampling [J]. *Nonlinear Dynamics*, 2014, 76(1): 627 – 634.

[19] FAN D, LO K. Identification for disturbed MIMO Wiener systems [J]. *Nonlinear Dynamics*, 2009, 55(1/2): 31 – 42.

[20] ZHAO Y, WANG L Y, YIN G G, et al. Identification of Wiener systems with binary-valued output observations [J]. *Automatica*, 2007, 43(10): 1752 – 1765.

[21] CHEN H F. Recursive identification for Wiener model with discontinuous piece-wise linear function [J]. *IEEE Transactions on Automatic Control*, 2006, 51(3): 390 – 400.

[22] WU Dehui. Identification method for nonlinear dynamic system using Wiener neural network [J]. *Control Theory & Applications*, 2009, 26(11): 1192 – 1196.

(吴德会. 非线性动态系统的Wiener神经网络辨识法 [J]. 控制理论与应用, 2009, 26(11): 1192 – 1196.)

[23] LI Shihua, WU Fubao, LI Qi. Identification of Wiener model using dynamic artificial neural networks [J]. *Control Theory & Applications*, 2000, 17(1): 92 – 103.
(李世华, 吴福保, 李奇. 一种基于动态人工神经网络的Wiener模型辨识 [J]. 控制理论与应用, 2000, 17(1): 92 – 103.)

[24] TANG Y, QIAO L, GUAN X. Identification of Wiener model using step signals and particle swarm optimization [J]. *Expert Systems with Applications*, 2010, 37(4): 3398 – 3404.

[25] TAN W W, LU F, LOH A P, et al. Modeling and control of a pilot pH plant using genetic algorithm [J]. *Engineering Applications of Artificial Intelligence*, 2005, 18(4): 485 – 494.

[26] AL-DUWAISH H N. *Identification of Wiener Model Using Genetic Algorithms* [M]. 2009.

[27] HATANAKA T, UOSAKI K, KOGA M. Evolutionary computation approach to block oriented nonlinear model identification [C] //*The 5th Asian Control Conference*. Melbourne: IEEE, 2004, 1: 90 – 96.

[28] LANCASTER P, SALKAUSKAS K. *Curve and Surface Fitting — An Introduction* [M]. London: Academic Press, 1986.

[29] SALHI H, KAMOUN S, ESSOUNBOULI N, et al. Adaptive discrete-time sliding-mode control of nonlinear systems described by Wiener models [J]. *International Journal of Control*, 2015, 89(3): 1 – 21.

[30] RAICH R, ZHOU G T, VIBERG M. Subspace based approaches for Wiener system identification [J]. *IEEE Transactions on Automatic Control*, 2005, 50(10): 1629 – 1634.

[31] CARL DE B. *A Practical Guide To Splines* [M]. Shanghai: World Publishing Corporation, 2008.

[32] FANG Chongzhi. *Process Identification* [M]. Beijing: Tsinghua University Press, 1988.
(方崇智. 过程辨识 [M]. 北京: 清华大学出版社, 1988.)

[33] LJUNG L. *System Identification: Theory for the User* [M]. 2nd ed. New Jersey: Prentice Hall PTR, 1999.

[34] WIGREN T. Input-output data sets for development and benchmarking in nonlinear identification [J]. *Technical Reports from the Department of Information Technology*, 2010, 20: 2010 – 020.

作者简介:

景绍学 (1976–), 男, 博士研究生, 目前研究方向为系统辨识、参数估计, E-mail: jingsx3@126.com;

李正明 (1958–), 男, 教授, 博士生导师, 目前研究方向为过程控制、参数辨识等, E-mail: lzming@ujs.edu.cn.