

基函数下非线性系统的支持向量机辨识

王建宏, 许莺, 熊朝华, 徐欣

(中国电子科技集团公司第二十八研究所信息工程重点实验室, 南京 210007)

摘要: 对于非线性系统的辨识问题, 将非线性系统在某组基函数下展开成线性回归形式。当基函数已知时, 对于线性回归形式中的未知参数矢量, 在性能指标中增加未知参数的惩罚项, 采用规范化的最小二乘法来辨识未知参数矢量。当基函数未知时, 在性能指标函数中增加预测误差的约束条件。对于约束最优化问题, 分别从基-对偶角度考虑其最优解, 并分析最优解的统计有偏性。为避免基函数的先验信息, 利用最小二乘支持向量机中的核函数来替换线性回归矩阵的乘积运算, 直接利用核函数来表示此约束最优化问题中的对偶矢量, 并使用此对偶矢量和核函数来近似原非线性函数。最后利用仿真算例验证采用支持向量机的核函数来表示非线性系统的有效性。

关键词: 非线性系统; 支持向量机; 核函数; 非线性辨识; 最小二乘

中图分类号: TP273 **文献标志码:** A **文章编号:** 1671-637X(2016)05-0011-05

Support Vector Machine Identification of Nonlinear System Under Basis Function

WANG Jian-hong, XU Ying, XIONG Zhao-hua, XU Xin

(Science and Technology on Information Systems Engineering Laboratory, The 28th Research Institute of China Electronics Technology Group Corporation, Nanjing 210007, China)

Abstract: To the problem of nonlinear system identification, the nonlinear system is expanded in a linear regression form under one basis function set. When the basis function set is known, the problem of how to identify the nonlinear system is transformed to estimate the unknown parameter vector exists in the linear regression form. One penalty term about the unknown parameter vector is added in the cost function and the regularized least square method is applied to estimate the unknown parameter vector. When this basis function set is unknown, the constraint about prediction error is considered in the cost function. As to this constrain optimization problem, the optimal solution is solved from the primal-dual point and its bias is also analyzed. To avoid the priori information about the basis function, the multiply operation coming from the linear regression matrix is replaced by one kernel function defined in the least square support vector machine theory. Then the dual vector in the constrain optimization problem is denoted by this kernel function directly and further the formal nonlinear function can be also approximated by one weighted sum form which is consisted of the kernel function and dual vector. Finally the simulation example results confirm the effectiveness of approximating the nonlinear system based on kernel function from support vector machine theory.

Key words: nonlinear system; support vector machine; kernel function; nonlinear identification; least square

0 引言

系统辨识按照对象特征可分为线性系统与非线性

系统辨识。线性系统辨识的输入-输出间存在着线性关系, 从而简化辨识过程, 但线性系统是一种理想化形式, 在自然界所存在的系统都应是非线性系统。为实现非线性系统的辨识, 将线性系统辨识中的若干辨识方法、渐近性分析和最优输入信号设计等引入至非线性系统中, 近年来已陆续开展对非线性系统辨识的研究。

线性系统辨识已日趋成熟, 文献[1]从时域角度

收稿日期: 2015-04-17

修回日期: 2016-03-17

基金项目: 国家自然科学基金(61402426)

作者简介: 王建宏(1980—), 男, 江西吉安人, 博士, 副教授, 研究方向为系统辨识与自适应控制。

分析线性系统的辨识方法及渐近性、收敛性问题;文献[2]从频域角度分析线性系统的辨识,指出怎样避免做傅里叶变换而产生的混叠效应;文献[3]从算法运行速度和复杂度方面分析线性系统的在线递推辨识;文献[4]研究线性参数变化系统的辅助变量辨识;文献[5]研究切换线性回归系统的范数规范和辨识;文献[6]研究双线性块形式下非线性系统的状态空间辨识;文献[7]提出 Wiener 系统的频率加权极大似然辨识法;文献[8]提出 Hammerstein 系统的实时鲁棒集员估计。

本文研究非线性系统在某组基函数下的辨识。当基函数已知时,将非线性系统改写成线性回归式。在优化的性能指标中增加惩罚项,通过该项来惩罚那些远离球体的参数值,以将它们给投影回来。当基函数未知时,对性能指标函数增加预测误差的约束条件。对于约束最优化问题,分别从基-对偶角度考虑其最优解,并分析最优解的统计有偏性。为避免基函数的先验信息,将支持向量机引进非线性系统辨识中。文献[9]联合最小二乘法和支持向量机得到最小二乘支持向量机法;文献[10]将此联合法应用于多输入多输出 Hammerstein 模型的参数辨识;文献[11]将支持向量机与预测误差法相结合,解决复杂动态网络在闭环反馈条件下的参数辨识问题。

本文将由线性回归矢量组成的矩阵定义为 Gramian 矩阵,利用统计学习理论中的 Mercer 定理将 Gramian 矩阵表示成无需任何关于回归矢量先验信息下的核函数。将此核函数应用于原未知参数矢量的求解过程中,考虑带有约束条件的最优化问题。通过对该约束最优化问题的拉格朗日函数应用广义 KKT 最优性充要条件可知:原两非线性函数可表示为核函数下的基函数展开,其中核函数前的权值即为约束最优化问题中的对偶变量。

1 模型描述

设非线性系统为

$$y(t) = \sum_{i=1}^{n_a} f_i^0(y(t-i)) + \sum_{j=0}^{n_b} g_j^0(u(t-j)) + v_0(t) \quad (1)$$

式中: $u(t)$ 和 $y(t)$ 分别表示该输入-输出非线性系统在某 t 时刻的输入和输出观测数据; t 表示离散时刻; $f_i^0, g_j^0: R \rightarrow R$ 为在 R 上的充分光滑、有界的非线性函数,且满足 $f_i^0(0) = 0, g_j^0(0) = 0; v_0(t)$ 为零均值的随机噪声; n_a, n_b 分别表示两已知的模型阶次。式(1)的两非线性函数可改写成两基函数展开式

$$\begin{cases} f_i^0(y(t-i)) = \sum_{l=1}^{n_g} \theta_{i,l}^0 \psi_{i,l}(y(t-i)) \\ g_j^0(u(t-j)) = \sum_{l=1}^{n_g} \theta_{j+n_a+1,l}^0 \psi_{j+n_a+1,l}(u(t-i)) \end{cases} \quad (2)$$

式中: $\{\psi_{i,l}: R \rightarrow R\}_{i=1, l=1}^{n_g, n_g}$ 为函数空间中未知的零对称分布的基函数,因各个基函数均为标量形式,且 $n_g = n_a + n_b + 1, \{\theta_{i,l}^0 \in R\}_{i=1, l=1}^{n_g, n_g}$ 为未知的常参数。将式(2)代入式(1)可得

$$y(t) = \sum_{i=1}^{n_a} \sum_{l=1}^{n_g} \theta_{i,l}^0 \psi_{i,l}(y(t-i)) + \sum_{j=0}^{n_b} \sum_{l=1}^{n_g} \theta_{j+n_a+1,l}^0 \psi_{j+n_a+1,l}(u(t-i)) + v_0(t) \quad (3)$$

定义如下各个矢量

$$\begin{cases} \theta_i = [\theta_{i,1} \cdots \theta_{i,n_g}]^T, \\ \psi_i = [\psi_{i,1} \cdots \psi_{i,n_g}]^T \\ \theta = [\theta_1 \cdots \theta_{n_g}]^T \\ \varphi(t-1) = \begin{bmatrix} \psi_1^T(y(t-1)) \\ \vdots \\ \psi_{n_a}^T(y(t-n_a)) \\ \psi_{n_a+1}^T(u(t)) \\ \vdots \\ \psi_{n_g}^T(u(t-n_b)) \end{bmatrix} \end{cases} \quad (4)$$

采集输入-输出观测数据对序列 $\{u(t), y(t)\}_{t=1}, \varphi(t-1)$ 为线性回归矢量,存在一个参数矢量 $\theta_0 \in R^{n_g}$ 使得式(3)可改写成

$$y(t) = \varphi^T(t-1) \theta_0 + v_0(t) = \varphi_{t-1}^T \theta_0 + v_0(t) \quad (5)$$

因线性回归矢量 $\varphi(t)$ 是由各个基函数 ψ_i 定义,当零对称分布的基本函数 ϕ_i 已知时,式(5)中的参数矢量可采用预测误差法来辨识。

2 规范化递推最小二乘法

对于式(5)中的未知参数矢量 θ 可增加限制条件:假设 θ 是 n_g 维球 $S(\bar{\theta}, \varepsilon) = \{\theta \in R^{n_g}: \|\theta - \bar{\theta}\| \leq \varepsilon\}$ 中的一个内点。 n_g 维球体 $S(\bar{\theta}, \varepsilon)$ 的施加是保证对于所有的参数矢量 $\theta \in S(\bar{\theta}, \varepsilon)$ 时,对应所有模型都可控^[12]。球体的中心 $\bar{\theta}$ 可理解为一个标定值,中心 $\bar{\theta}$ 和球半径 ε 可作为先验已知量或通过离线辨识的方法来获取。为辨识式(5)中的未知参数矢量 θ ,同时保证 $\theta \in S(\bar{\theta}, \varepsilon)$,对性能指标函数进行调整,即

$$J_1(\theta) = \sum_{k=1}^t (y_k - \varphi_{k-1}^T \theta)^2 + \alpha_t \|\theta - \bar{\theta}\|^2 \quad (6)$$

式(6)增加第2项 $\alpha_t \|\theta - \bar{\theta}\|^2$ 的目的在于惩罚那些离已知的标定值球心 $\bar{\theta}$ 较远的参数值。标量函数 α_t 的作用在于平衡性能指标 $J_1(\theta)$ 中的惩罚项和最小二乘项。当参数估计远离标定值球心时,标量函数 α_t 快速增长以达到惩罚的效果,反之,惩罚项 $\alpha_t \|\theta - \bar{\theta}\|^2$ 快速减少以避免破坏最小二乘算法的闭环特性。根据

上面阐述,可得出定理1。

定理1 构造如下性能指标函数

$$J_2(\boldsymbol{\theta}) = \sum_{k=1}^i (y_k - \boldsymbol{\varphi}_{k-1}^T \boldsymbol{\theta})^2 + (\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})^T Q (\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}) + \alpha_i \|\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}\|^2 \quad (7)$$

式中, $\sum_{k=1}^i (y_k - \boldsymbol{\varphi}_{k-1}^T \boldsymbol{\theta})^2 + (\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})^T Q (\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})$ 为带有规范项 $(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})^T Q (\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})$ 的标准最小二乘性能指标,且

$$\alpha_i = (\ln(\sum_{k=1}^i \|\boldsymbol{\varphi}_k\|^2 + \text{tr} Q))^{1+\delta} \quad \delta > 0 \quad (8)$$

最小化 $J_2(\boldsymbol{\theta})$ 即 $\min_{\boldsymbol{\theta} \in S(\bar{\boldsymbol{\theta}}, \varepsilon)} J_2(\boldsymbol{\theta})$ 的解为如下的递推过程,

$$\text{令 } T_0 = P_{i-1}, i=1 \rightarrow n_\theta, \text{取 } \boldsymbol{\phi}_i = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}。$$

递推计算

$$T_i = T_{i-1} - \frac{(\alpha_i - \alpha_{i-1}) T_{i-1} \boldsymbol{\phi}_i \boldsymbol{\phi}_i^T T_{i-1}}{1 + (\alpha_i - \alpha_{i-1}) \boldsymbol{\phi}_i^T T_{i-1} \boldsymbol{\phi}_i} \quad (9)$$

计算

$$P_i = T_n - \frac{T_n \boldsymbol{\varphi}_{i-1} \boldsymbol{\varphi}_{i-1}^T T_n}{1 + \boldsymbol{\varphi}_{i-1}^T T_n \boldsymbol{\varphi}_{i-1}} \quad (10)$$

未知参数矢量的递推最小二乘估计值为

$$\hat{\boldsymbol{\theta}}_i = \hat{\boldsymbol{\theta}}_{i-1} + P_i \boldsymbol{\varphi}_{i-1} (y_i - \boldsymbol{\varphi}_{i-1}^T \hat{\boldsymbol{\theta}}_{i-1}) + P_i (\alpha_i - \alpha_{i-1}) (\bar{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_{i-1}) \quad (11)$$

式中

$$r_i = r_{i-1} + \|\boldsymbol{\varphi}_{i-1}\|^2, \alpha_i = (\ln(r_i))^{1+\delta} \quad \delta > 0 \quad (12)$$

若 $\hat{\boldsymbol{\theta}}_i \notin S(\bar{\boldsymbol{\theta}}, \varepsilon)$, 则做估计值 $\hat{\boldsymbol{\theta}}_i$ 在球体 $S(\bar{\boldsymbol{\theta}}, \varepsilon)$ 上的投影点, 即做如下的运算

$$\bar{\boldsymbol{\theta}}_i = \begin{cases} \hat{\boldsymbol{\theta}}_i & \hat{\boldsymbol{\theta}}_i \in S(\bar{\boldsymbol{\theta}}, \varepsilon) \\ \frac{\hat{\boldsymbol{\theta}}_i - \bar{\boldsymbol{\theta}}}{\|\hat{\boldsymbol{\theta}}_i - \bar{\boldsymbol{\theta}}\|^2} \varepsilon + \bar{\boldsymbol{\theta}} & \hat{\boldsymbol{\theta}}_i \notin S(\bar{\boldsymbol{\theta}}, \varepsilon) \end{cases} \quad (13)$$

证明 因为 $J_2(\boldsymbol{\theta})$ 是关于未知参数矢量 $\boldsymbol{\theta}$ 的二次函数, 对 $J_2(\boldsymbol{\theta})$ 关于 $\boldsymbol{\theta}$ 求导后取值为0, 整理可得

$$\hat{\boldsymbol{\theta}}_i = \left[\sum_{k=1}^i \boldsymbol{\varphi}_{k-1} \boldsymbol{\varphi}_{k-1}^T + Q + \alpha_i \mathbf{I} \right]^{-1} \times \left[\sum_{k=1}^i y_k \boldsymbol{\varphi}_{k-1} + (Q + \alpha_i \mathbf{I}) \bar{\boldsymbol{\theta}} \right] \quad (14)$$

为得到 $\hat{\boldsymbol{\theta}}_i$ 的递推表达式, 定义矩阵 P_i 为

$$P_i = \left[\sum_{k=1}^i \boldsymbol{\varphi}_{k-1} \boldsymbol{\varphi}_{k-1}^T + Q + \alpha_i \mathbf{I} \right]^{-1} \quad (15)$$

从而 $\hat{\boldsymbol{\theta}}_i$ 可表示为

$$\hat{\boldsymbol{\theta}}_i = P_i \left[\sum_{k=1}^i y_k \boldsymbol{\varphi}_{k-1} + (Q + \alpha_i \mathbf{I}) \bar{\boldsymbol{\theta}} \right] \quad (16)$$

对式(16)右边的项 $\sum_{k=1}^i y_k \boldsymbol{\varphi}_{k-1}$ 进行展开为

$$\sum_{k=1}^i y_k \boldsymbol{\varphi}_{k-1} = y_i \boldsymbol{\varphi}_{i-1} + P_{i-1}^{-1} \hat{\boldsymbol{\theta}}_{i-1} - (Q + \alpha_i \mathbf{I}) \bar{\boldsymbol{\theta}} \quad (17)$$

将式(17)代入 P_i^{-1} 的表达式中可得

$$P_i^{-1} = P_{i-1}^{-1} + \boldsymbol{\varphi}_{i-1} \boldsymbol{\varphi}_{i-1}^T + (\alpha_i - \alpha_{i-1}) \mathbf{I} \quad (18)$$

将式(17)再次代入 $\hat{\boldsymbol{\theta}}_i$ 的表达式(16)中可得 $\hat{\boldsymbol{\theta}}_i = \hat{\boldsymbol{\theta}}_{i-1} + P_i \boldsymbol{\varphi}_{i-1} (y_i - \boldsymbol{\varphi}_{i-1}^T \hat{\boldsymbol{\theta}}_{i-1}) + P_i (\alpha_i - \alpha_{i-1}) (\bar{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_{i-1})$ 即为定理1中的式(11)。

对式(18)两边同时取求逆运算, 其中使用矩阵求逆引理公式可得

$$P_i = [P_{i-1}^{-1} + (\alpha_i - \alpha_{i-1}) \mathbf{I}]^{-1} - [P_{i-1}^{-1} + (\alpha_i - \alpha_{i-1}) \mathbf{I}]^{-1} \times \frac{\boldsymbol{\varphi}_{i-1} \boldsymbol{\varphi}_{i-1}^T}{1 + \boldsymbol{\varphi}_{i-1}^T [P_{i-1}^{-1} + (\alpha_i - \alpha_{i-1}) \mathbf{I}]^{-1} \boldsymbol{\varphi}_{i-1}} \times [P_{i-1}^{-1} + (\alpha_i - \alpha_{i-1}) \mathbf{I}]^{-1}, \text{简记为 } T_n = [P_{i-1}^{-1} + (\alpha_i - \alpha_{i-1}) \mathbf{I}]^{-1}, P_i \text{ 的表达式可简写为 } P_i = T_n - \frac{T_n \boldsymbol{\varphi}_{i-1} \boldsymbol{\varphi}_{i-1}^T T_n}{1 + \boldsymbol{\varphi}_{i-1}^T T_n \boldsymbol{\varphi}_{i-1}}, \text{即为定理1中的式(10), 对 } T_n \text{ 再次使用矩阵求逆引理可得 } T_n \text{ 的递推计算式(9)}。$$

3 最小二乘支持向量机

根据线性回归关系式(5)在规范化最小二乘性能准则式(6)中取球心 $\bar{\boldsymbol{\theta}}=0$, 标量函数 α_i 选择一个常规参数 γ , 所考虑性能指标函数为

$$J_3(\boldsymbol{\theta}, e) = \frac{1}{2N} \sum_{t=1}^N e(t)^2 + \frac{\gamma}{2} \boldsymbol{\theta}^T \boldsymbol{\theta} = \frac{1}{2N} \|e(t)\|^2 + \frac{\gamma}{2} \|\boldsymbol{\theta}\|^2 \quad (19)$$

式中: $e(t)$ 为预测误差; 标量值 $\gamma > 0$ 为常规参数; N 表示观测数据总个数; 未知参数矢量 $\boldsymbol{\theta}$ 可通过如下最小化问题来求解。

$$\min_{\boldsymbol{\theta}, e} J_3(\boldsymbol{\theta}, e) \quad (20)$$

subject to $e(t) = y(t) - \boldsymbol{\varphi}^T(t-1) \boldsymbol{\theta} \quad t = 1, \dots, N$
对式(20)的求解采用最优性必要条件, 将预测误差 $e(t)$ 的表达式代入式(19)可得

$$J_3(\boldsymbol{\theta}) = \frac{1}{2N} \sum_{t=1}^N (y(t) - \boldsymbol{\varphi}^T(t-1) \boldsymbol{\theta})^2 + \frac{\gamma}{2} \boldsymbol{\theta}^T \boldsymbol{\theta} \quad (21)$$

对式(21)关于未知参数矢量 $\boldsymbol{\theta}$ 求偏导数, 并令其等于0, 可得

$$\frac{\partial J_3(\boldsymbol{\theta})}{\partial(\boldsymbol{\theta})} = 0 \quad (22)$$

将上式的偏导关系式展开, 整理可得最小值处的参数估计值为

$$\hat{\boldsymbol{\theta}}_p = \left[\gamma \mathbf{I} + \frac{1}{N} \sum_{t=1}^N \boldsymbol{\varphi}(t-1) \boldsymbol{\varphi}^T(t-1) \right]^{-1} \cdot \left[\frac{1}{N} \sum_{t=1}^N \boldsymbol{\varphi}(t-1) y(t) \right] \quad (23)$$

式中, I 表示维数为 n_θ 的单位矩阵。为简化式(23)中的连加和运算, 定义如下的各个矢量分别为 $Y = [y(1) \cdots y(N)]^T$, $\Phi = [\varphi(0) \cdots \varphi(N-1)]^T$, 从而式(23)可简记为

$$\hat{\theta}_p = \left(\gamma I + \frac{1}{N} \Phi^T \Phi \right)^{-1} \frac{1}{N} \Phi^T Y \quad (24)$$

对于由式(19)所得到的参数估计矢量 $\hat{\theta}_p$, 其统计特性可整理成定理2。

定理2 性能指标函数式(19)的最小二乘解 $\hat{\theta}_p$ 与真实值 θ_0 间的偏差记为

$$B_{\hat{\theta}} = E(\hat{\theta}_p - \theta_0) \quad (25)$$

式中, E 表示取数学期望运算。偏差 $B_{\hat{\theta}}$ 可表示为

$$\begin{cases} B_{\hat{\theta}} = \left(\gamma I + \frac{1}{N} \Phi^T \Phi \right)^{-1} \gamma \theta_0 + \left(\gamma I + \frac{1}{N} \Phi^T \Phi \right)^{-1} \frac{1}{N} \Phi^T V_0 \\ V_0 = [v_0(1) \cdots v_0(N)] \end{cases} \quad (26)$$

为减少偏差的大小, 要使得式(25)的第二项为0, 可采用辅助变量 $\xi(t)$ 实现。此时辅助变量 $\xi(t)$ 的元素要选择得与观测噪声 v_0 不相关, 即要保证有

$$E(\xi(t)v_0(t)) = 0 \quad t=1, \dots, N \quad (27)$$

在约束最优化理论中, 对于约束条件需采用对偶变量来定义对应的拉格朗日函数^[13], 即对于约束最优化问题式(20), 构造其对应的拉格朗日函数为

$$L(\theta, e, \alpha) = J_3(\theta, e) - \sum_{t=1}^N \alpha_t (\varphi^T(t-1)\theta + e(t) - y(t)) \quad (28)$$

式中, $\alpha_t \in \mathbf{R}$ 为拉格朗日乘子变量。对于式(28)的拉格朗日函数应用 KKT 最优充要条件得

$$\begin{cases} \frac{\partial L}{\partial e} = \frac{1}{N} e(t) - \alpha_t = 0 \\ \frac{\partial L}{\partial \alpha_t} = y(t) - \varphi^T(t-1)\theta - e(t) = 0 \\ \frac{\partial L}{\partial \theta} = \gamma \theta - \sum_{t=1}^N \alpha_t \varphi(t-1) = 0 \end{cases} \quad (29)$$

对式(29)中的3个等式整理可得

$$\begin{cases} \alpha_t = \frac{1}{N} e(t) \\ y(t) = \varphi^T(t-1)\theta + e(t) \quad t=1, \dots, N \\ \theta = \frac{1}{\gamma} \sum_{t=1}^N \alpha_t \varphi(t-1) \end{cases} \quad (30)$$

将式(30)中的第1个和第3个等式都代入第2个等式中可得

$$y(t) = \left(N + \frac{1}{\gamma} \sum_{i=1}^N \varphi^T(t-1)\varphi(t-1) \right) \alpha_t \quad (31)$$

联合所有的 $t=1, \dots, N$ 组成矢量形式为

$$Y = \left(NI + \frac{1}{\gamma} \Phi^T \Phi \right) \alpha \quad \alpha = [\alpha_1 \cdots \alpha_N]^T \in \mathbf{R}^N \quad (32)$$

从式(32)中可解出对偶矢量 α 为

$$\alpha = \left(NI + \frac{1}{\gamma} \Phi^T \Phi \right)^{-1} Y = \left(I + \frac{1}{N\gamma} \Phi^T \Phi \right)^{-1} \frac{Y}{N} \quad (33)$$

若对偶矢量 α 由式(33)求出后, 将式(33)代入式(30)的第3个等式可得参数估计值为

$$\hat{\theta}_d = \Phi^T \left(\gamma I + \frac{1}{N} \Phi \Phi^T \right)^{-1} \frac{1}{N} Y \quad (34)$$

通过比较式(24)中的 $\hat{\theta}_p$ 与式(34)中的 $\hat{\theta}_d$ 可知: 虽然两者在形式上有很大的差异, 但在凸二次线性等式约束条件下, 由强对偶性可得 $\hat{\theta}_p = \hat{\theta}_d$ 。

当非线性函数在基函数展开下, 若能估计出 $\hat{\theta}_d$ 或 $\hat{\theta}_p$ 后, 将参数估计值代入式(2), 可实现非线性系统的辨识。但由式(24)和式(34)可知, 要得到参数估计值需线性回归矩阵 Φ 或基函数的先验知识。若基函数 ψ_i 已知, 还需计算线性回归矢量 φ , 再构造线性回归矩阵 Φ , 最后代入参数估计值中。此过程增大辨识计算复杂度。

因对偶矢量 α 的表达式中出现线性回归矩阵 Φ 的内积 $\Phi \Phi^T$, 定义 Gramian 矩阵 $G = \Phi \Phi^T$ 。

根据 Mercer 定理^[14], 矩阵 G 在无 Φ 的显式信息下可用某核函数来定义, 且 G 分解为

$$[G]_{j,k} = \sum_{i=1}^n [G^{(i)}]_{j,k}, \text{ 其中, 每个 } G^{(i)} \text{ 都可表示为一种内积运算}$$

$$[G^{(i)}]_{j,k} = \langle \psi_i(x_i(j)), \psi_i(x_i(k)) \rangle = K^{(i)}(x_i(j), x_i(k)) \quad (35)$$

式中, $K^{(i)}$ 表示一个正定的核函数。且

$$\begin{cases} x_i(t) = y(t-i) & i=1, \dots, n_a \\ x_{n_a+1+j}(t) = u(t-j) & j=0, \dots, n_b \end{cases} \quad (36)$$

利用核函数 $K^{(i)}$ 来替换 G , 对偶矢量可表示为

$$\alpha = \left(I + \frac{1}{N\gamma} G \right)^{-1} \frac{Y}{N} \quad (37)$$

对于式(2)中的第1个非线性函数 f_i 可写成

$$f_i(y(t-i)) = \psi_i^T(y(t-i)) \theta_i \quad (38)$$

再将式(30)的第3个等式代入式(38), 并利用矢量中的各个分量可得

$$f_i(y(t-i)) = \frac{1}{\gamma} \sum_{i=1}^N \alpha_i \psi_i^T(y(t-i)) \psi_i(y(t-i)) \quad (39)$$

再将式(35)中的核函数代入式(39)中可得

$$f_i(y(t-i)) = \frac{1}{\gamma} \sum_{i=1}^N \alpha_i K^{(i)}(y(t-i)) \quad (40)$$

类似可将式(2)中的第2个非线性函数 g_j 改写成

$$g_j(u(t-j)) = \frac{1}{\gamma} \sum_{i=1}^N \alpha_i K^{(n_a+1+j)}(u(t-j)) \quad (41)$$

因核函数可直接由输入-输出观测数据对序列 $\{u(t), y(t)\}$ 构成, 加权和形式中起逼近作用的是对

偶变量 α 。对偶变量 α 由式(37)表示,且式(37)中的矩阵 G 同样可由所定义的核函数来替换。整个非线性系统辨识过程无需任何关于原来基函数的信息,只需依靠式(37)得到对偶矢量,其具体最小二乘支持向量机辨识可归纳如下。

给定模型阶次 n_a 和 n_b ,数据集 $\{u(t), y(t)\}_{t=1}^N$, 规范参数 γ , 径向基函数核 $\{K^{(i)}\}_{i=1}^{n_g}, n_g = n_a + n_b + 1$:

- 1) 令 $\tau \leftarrow 0$;
- 2) 根据式(35)经核函数 $K^{(i)}$ 计算矩阵 $G^{(i)}$;
- 3) 经式(35)估计对偶矢量 $\alpha^{(0)}$;
- 4) 重复过程;
- 5) $\tau \leftarrow \tau + 1$;

6) 再次利用额外采集到的数据集 $\{u(t), y(t)\}$, 以5)计算出来的对偶矢量估计值作为初始值,经式(37)估计对偶矢量 $\alpha^{(\tau)}$;

- 7) 重复6),直至 $\alpha^{(\tau)}$ 收敛或变化的范围很小;

8) 联合7)得到的收敛矢量 α 和给定的径向基函数核 $\{K^{(i)}\}_{i=1}^{n_g}$, 经式(40)和式(41)确定两非线性函数的核函数加权形式。

4 仿真算例

考虑如下利用差分形式表示的数据生成系统 $y(k) = -0.8y(k-1) + f^0(y(k-2)) + g^0(u(k)) - 0.6u(k-1) + v_0(k)$, 其中的非线性函数 f 和 g 分别定义为

$$\begin{cases} f^0(x) = 2x^2, \\ g^0(x) = \begin{cases} 1 & x \geq 1 \\ x & -1 \leq x \leq 1 \\ -1 & x \leq -1 \end{cases} \end{cases} \quad (42)$$

$v_0(k)$ 为有色噪声,其由白噪声 $e_0(k) \sim N(0, \sigma_e^2)$ 通过某滤波器得到^[15]。其中滤波器为

$$v_0(k) = b_0 e_0(k) + b_1 e_0(k-1) \quad (43)$$

式中,系数分别为 $b_0 = 1, b_1 = -0.6$ 。

采集观测数据集 $\{u(k), y(k)\}_{k=1}^N$, 系统利用 $N = 1000$ 个白色伪随机二元输入序列来激励,激励的初始条件取为零值。参数化原数据生成系统为

$$y(k) = \theta_1 y(k-1) + \theta_2 \varphi_2^T(y(k-2)) + \theta_3 \varphi_3^T(u(k)) + \theta_4 u(k-1) + e(k) \quad (44)$$

式中: $e(k)$ 表示残差; 函数 $\varphi_1(y(k-1))$ 和 $\varphi_4(u(k-1))$ 显式地定义为 $y(k-1)$ 和 $u(k-1)$; 特征映射 φ_2 和 φ_3 先验未知,采用径向基函数核来代表非线性。因仿真的作用是表征核函数下非线性系统的辨识精度,对于参数化的数据生成系统,采用1)~8)的辨识算法。在算法的实施过程中,径向基函数与其中的两参数 σ_2, σ_3 可用于估计基函数 φ_2 和 φ_3 。此时可令两参数分别为 $\sigma_2 = 1.5, \sigma_3 = 0.6$ 。

仿真过程中,因基函数 φ_1 和 φ_4 是显式地固定的,参数 θ_1 和 θ_4 可直接辨识,而参数 θ_2 和 θ_3 却无法直接估计,且仅函数 $f = \theta_2 \varphi_2^T$ 或 $g = \theta_3 \varphi_3^T$ 可计算出来。利用径向基函数和由7)得到的对偶矢量直接将两非线性函数 $f(x)$ 和 $g(x)$ 表示成核函数加权形式。其中,核函数加权形式与其实际形式间的比较如图1~图2所示。

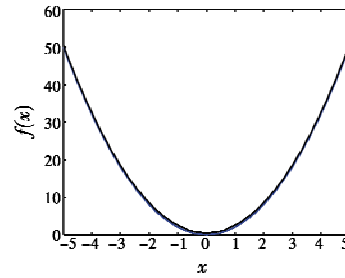


图1 非线性函数 $f(x)$ 的估计

Fig. 1 The estimation of nonlinear function $f(x)$

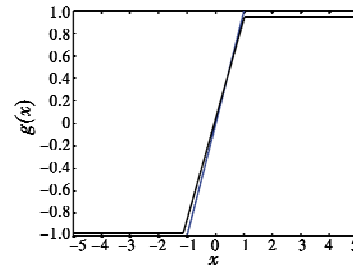


图2 非线性函数 $g(x)$ 的辨识估计

Fig. 2 The estimation of nonlinear function $g(x)$

图中:蓝色曲线表示真实实际的两非线性函数对应曲线;黑色曲线表示辨识估计的非线性函数曲线。由图1、图2可知,蓝色辨识曲线紧紧靠拢于蓝色曲线,即辨识曲线可近似地逼近于原实际曲线。此即表明,辨识估计的核函数加权形式可无限地逼近原两非线性函数。

5 结语

对于非线性系统在基函数的展开式:当基函数已知时,提出规范化递推最小二乘辨识法求解基函数展开式中的未知参数估计值;当基函数未知时,将由基函数乘积组成的矩阵表示成核函数形式,联合利用核函数可构成原非线性函数的核加权形式。为减小偏差的存在,可采用辅助变量最小二乘法实现无偏的非线性系统辨识,此分析思路将是以后的研究主题。

参考文献

[1] LJUNG L. System identification; theory for the user [M]. New Jersey: Prentice Hall Press, 1999.
 [2] PINTELON R, SCHOUKENS J. System identification; a frequency domain approach [M]. New York: IEEE Press, 2001.

足无人作战飞机引导的在线解算需求。本文的研究成果对于提升无人作战飞机空战引导能力具有一定的工程应用价值。

参考文献

- [1] 陈冰,蔡满意,朱凡,等. 进入角约束下的攻击机空间远距引导研究[J]. 电光与控制, 2010, 17(9): 17-19. (CHEN B, CAI M Y, ZHU F, et al. Study on a 3D long-distance navigation method with entrance angle constraints for attackers[J]. Electronics Optics & Control, 2010, 17(9): 17-19.)
- [2] 张洪波,李国英,丁全心,等. 超视距空战下的态势评估技术研究[J]. 电光与控制, 2010, 17(4): 9-13. (ZHANG H B, LI G Y, DING Q X, et al. Research on situation assessment in BVR air combat [J]. Electronics Optics & Control, 2010, 17(4): 9-13.)
- [3] 付昭旺,于雷,李战武,等. 战斗机隐蔽接敌轨迹优化方法[J]. 国防科技大学学报, 2013, 35(5): 52-58. (FU Z W, YU L, LI Z W, et al. The trajectory optimization method for fighter stealthy approach [J]. Journal of National University of Defense Technology, 2013, 35(5): 52-58.)
- [4] 包勇,符小卫,高晓光. 势场理论的多无人机协同路径规划方法[J]. 火力与指挥控制, 2012, 37(3): 10-12. (BAO Y, FU X W, GAO X G. A method of cooperative path planning of multiple UAVs on potential field theory [J]. Fire Control & Command Control, 2012, 37(3): 10-12.)
- [5] 张涛,于雷,周忠良,等. 基于人工势场启发粒子群算法的空战机动决策[J]. 电光与控制, 2013, 20(1): 77-82. (ZHANG T, YU L, ZHOU Z L, et al. Decision-making of air combat maneuvering based on APF and PSO [J]. Electronics Optics & Control, 2013, 20(1): 77-82.)
- [6] 高永,向锦武. 超视距多机协同空战目标分配算法[J]. 北京航空航天大学学报, 2007, 33(3): 286-289. (GAO Y, XIANG J W. Target assignment in BVR air combat [J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(3): 286-289.)
- [7] DEMENGELE G, POUGET J P. Mathematics des courbes et des surfaces [M]. Paris: Ellipses Marketing S. A., 1998: 74-141.
- [8] BRINKER J S. Autonomous steering of the Joint Unmanned Combat Air Systems (J-UCAS) X-45A [C]//AIAA 3rd "Unmanned Unlimited" Technical Conference, Chicago, Illinois, 2004. doi:10.2514/6.2004-6575.
- (上接第15页)
- [3] VANDENBERGHE B S L. Convex optimization [M]. Cambridge: Cambridge University Press, 2004.
- [4] LAURAIN V, GILSON M, TOTH R, et al. Refined instrumental variable methods for identification of LPV BOX Jenkins models [J]. Automatica, 2010, 46(6): 959-967.
- [5] OHLSSON H, LJUNG L. Identification of switched linear regression models using sum of norms regularization [J]. Automatica, 2013, 49(4): 1045-1050.
- [6] VAN MULDER A, SCHOUKENS J, VANBEYLEN L. Identification of systems with localised nonlinearity: from statespace to block-structured models [J]. Automatica, 2013, 49(5): 1392-1396.
- [7] PILLONETTO G, DE NICOLA G. A new kernel based approach for linear system identification [J]. Automatica, 2010, 46(1): 81-93.
- [8] ZEILINGER M N, TONES C N, MORARI M. Real time suboptimal model predictive control using a combination of explicit MPC and online optimization [J]. IEEE Transactions of Automatic Control, 2011, 56(7): 1524-1534.
- [9] PILLONETTO G, DINUZZO F, CHEN T S, et al. Kernel methods in system identification, machine learning and function estimation: a survey [J]. Automatica, 2014, 50(3): 657-682.
- [10] GOETHALS I, PELCKMANS K, SUYKENS J A K, et al. Identification of MIMO Hammerstein models using least squares support vector machines [J]. Automatica, 2005, 41(7): 1263-1272.
- [11] VANDEN HOF P M J, DANKERS A, HEUBERGER P S C, et al. Identification of dynamic models in complex networks with prediction error methods [J]. Automatica, 2013, 49(10): 2994-3006.
- [12] FELLER C, JOHANSEN T A, OLARU S. An improved algorithm for combinatorial multi-parameter quadratic programming [J]. Automatica, 2013, 49(5): 1370-1376.
- [13] ZEILINGER M N, RAIMONDO D M, DOMAHIDI A, et al. On real time robust model predictive control [J]. Automatica, 2014, 50(3): 683-694.
- [14] BAI E W, LI K, ZHAO W X, et al. Kernel based approaches to local nonlinear nonparametric variable selection [J]. Automatica, 2014, 50(1): 100-113.
- [15] SCHOUKENS M, PINTELON R, ROLAIN Y. Identification of Wiener Hammerstein systems by a nonparametric separation of the best linear approximation [J]. Automatica, 2014, 50(2): 628-634.