

文章编号: 1001-0920(2009)07-1100-04

## 基于核岭回归的非线性内模控制

黄宴委

(福州大学 电气工程与自动化学院, 福州 350108)

**摘要:** 提出一种基于核岭回归(KRR)建模的内模控制策略. 该方法充分利用基于结构风险最小化为学习规则的回归方法的非线性拟合性能, 建立内模控制系统, 从理论上分析了内模控制系统的稳定性和稳态误差同逆模与内模估计误差的关系问题. 仿真表明, 在训练样本有限和有噪声污染情况下, 该系统较神经网络方法具有更好的控制性能.

**关键词:** 内模控制; 核岭回归; 神经网络; 稳态误差; 稳定性

中图分类号: TP278

文献标识码: A

## Kernel ridge regression-based nonlinear internal model control

HUANG Yanwei

(School of Electric Engineering and Automation, Fuzhou University, Fuzhou 350108, China. E-mail: sjtu\_huanghao@hotmail.com)

**Abstract:** A design of internal model control systems for process with delay using kernel ridge regression (KRR) is proposed, which fully uses the excellent nonlinear estimation performance of KRR based on structural risk minimization principle. Internal model control system is constructed. Closed-system stability and steady error are analyzed for the existence of modeling errors. The simulations show that the proposed control systems have the better control performance than neural networks in training samples with finite size and noises.

**Key words:** Internal model control; Kernel ridge regression; Neural networks; Steady error; Stability

### 1 引言

内模控制(IMC)是一种实用性和抗干扰性很强的控制策略,能够对不确定性、大纯滞后系统进行有效控制,主要分为基于数学模型和基于智能回归估计模型两种<sup>[1]</sup>.在实际应用中,由于被控对象往往是非线性的,且其数学模型难以获得,使得基于智能学习估计模型的非线性 IMC 得到了广泛应用,并取得很多成果,其中基于神经网络建模的非线性 IMC 系统尤为广泛<sup>[2-5]</sup>.然而由于基于经验风险最小化为学习规则的神经网络建模估计,存在着一些固有的缺点(如对学习样本数量和质量要求很高,网络结构难以确定,容易陷入局部最小解和过学习现象等),使得神经网络泛化能力极大地降低,增大了 IMC 系统中的内模和逆模的估计误差,进而影响 IMC 系统的性能<sup>[6,7]</sup>.因此,寻找一种新型的模型估计方法来实现非线性 IMC 系统是非常必要的.

核岭回归(KRR)是一种以结构风险最小化为学习规则的算法,与神经网络相比,具有更高的泛化

能力和非线性建模精度,特别是在样本有限和有噪声污染的情况下,效果更为明显.因此,本文提出了一种基于 KRR 建模的非线性 IMC 策略,以降低 IMC 系统稳态误差,提高系统的控制性能.

### 2 基于 KRR 建模

核回归方法通过将原始样本空间的数据映射到高维特征空间中,在构造最优决策函数时利用结构风险最小化原则,解决原来在样本输入空间中不能用线性方法求解的非线性问题.

设样本  $x \in R^n$  从输入空间到高维特征空间的非线性变换为  $\phi(x)$ , 则求解的回归方程形式为

$$f(x) = w^T \phi(x), \quad (1)$$

式中  $w$  为待定权值变量. 传统上,最小二乘回归问题主要是通过求解经验风险为最小化均方差  $R_{\text{erm}} = \frac{1}{l} \sum_{i=1}^l (y_i - f(x_i))^2$  为其优化目标来确定  $w$ . 然而,实际测得的样本往往存在多重共线性问题,这便极大地降低了建模的准确性. 针对回归中样本数据共线

收稿日期: 2008-07-22; 修回日期: 2008-11-01.

基金项目: 福州大学科技发展基金项目(2008-XQ-19).

作者简介: 黄宴委(1976—),男,福建闽侯人,讲师,博士,从事统计学习、支持向量机建模及智能控制等研究.

性问题, Hoerl 等<sup>[9]</sup>提出了 Ridge 回归方法. 该方法引入与样本数据点相对的回归曲线偏差, 构造出结构风险最小化函数  $R_{\text{svm}} = R_{\text{erm}} + w^2$ . 其中  $w$  为正则项参数, 用来减小数据共线性的影响, 可以控制函数  $f$  的复杂度. 因此, 可等价于对偶空间中的优化函数

$$\min_w J = w^2 + \sum_{i=1}^l \frac{1}{2} (y_i - w(x_i))^2, \quad (2)$$

s. t.  $y_i - w(x_i) = \epsilon_i, i = 1, 2, \dots, l$ .

由式(2)可构造相应的拉格朗日函数为

$$L(w, \lambda) = w^2 + \sum_{i=1}^l \frac{1}{2} (y_i - w(x_i) - \epsilon_i)^2 + \sum_{i=1}^l \lambda_i (y_i - w(x_i) - \epsilon_i), \quad (3)$$

式中  $\lambda_i$  为拉格朗日乘子. 由优化条件  $\partial L / \partial w = 0$  和  $\partial L / \partial \lambda_i = 0, i = 1, 2, \dots, l$ , 可得到线性方程组

$$2w + \sum_{i=1}^l \lambda_i (x_i)(x_j) + I_{l \times l}^{-1} y = 0$$

式中  $i, j = 1, 2, \dots, l$ . 而  $w = \frac{1}{2} \sum_{i=1}^l \lambda_i (x_i)$ ,  $\lambda_i = \epsilon_i / 2$ , 并定义内核函数  $K(x_i, x_j) = (x_i)(x_j)$ , 因此最终得到回归函数为

$$f(x) = y^T (K + I_{l \times l})^{-1} k.$$

式中:  $K$  为  $l \times l$  的 Gram 矩阵;  $k$  为新样本与训练样本在特征空间中的内积向量,  $k_i = K(x_i, x), i = 1, 2, \dots, l$ .

选择不同的核函数  $K$ , 可以构造不同的 KRR 形式, 常用的核有:

多项式核

$$K(x, x_i) = [ (x, x_i) + 1 ]^q, q \text{ 为阶数};$$

径向基(RBF)核

$$K(x, x_i) = \exp(-\|x - x_i\|^2 / 2\sigma^2);$$

Sigmoid 核

$$K(x, x_i) = S[v(x, x_i) + c].$$

本文采用 RBF 函数为回归内核.

### 3 基于 KRR 的 IMC 系统

#### 3.1 系统结构

基于 KRR 的 IMC 系统主要由 4 个部分组成, 如图 1 所示. 控制对象  $G$ , 对象模型  $G_m$  (不包含纯滞后环节) 和由逆模型  $G_m^{-1}$  与滤波器  $F$  构成的控制器为  $G_c(z) = F(z) G_m^{-1}(z)$ . 其中最小相位系统  $G_m$  为  $z(k$

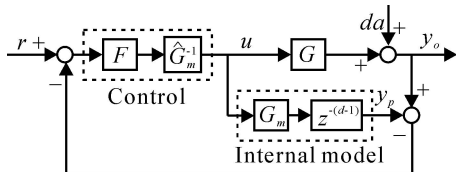


图 1 IMC 系统

+ 1) =  $G(z) \{ z^{k-n+1}, u(z^{k-m+d}) \}$ , 纯滞后环节  $z^{-(d-1)}$  为  $y_p(k+d) = z(k+1)$ ;  $G_m^{-1}$  为  $G_m$  逆模型估计  $u(k) = G_m^{-1}(z) \{ z^{k-n+1}, u(z^{k-m+d}) \}$ . 这里  $G_m$  和  $G_m^{-1}$  均由 KRR 来实现其非线性模型.

#### 3.2 鲁棒稳定性分析

设图 1 中, 稳定的被控对象离散传函为  $G(z)$ , 基于 KRR 的内模为  $G_m(z)$  和控制器为  $G_c(z)$ , 由于非线性建模误差的存在, 鲁棒性考虑的对象族为

$$I = \left\{ G_m(z) : \left| \frac{G(z) - G_m(z) z^{-(d-1)}}{G_m(z)} \right| \leq e_m \right\}, \quad (4)$$

式中  $e_m$  为 KRR 建模误差上界.

引理 1 对于单位负反馈控制系统, 具有建模误差的被控对象为  $G(z)$ , 其模型为  $G_m(z)$ , 控制器为  $C(z)$ , 则保证闭环系统稳定的充要条件为<sup>[8]</sup>

$$\left| \frac{C(z) G_m(z) z^{-(d-1)}}{1 + C(z) G_m(z) z^{-(d-1)}} \right| \leq e_m < 1. \quad (5)$$

定理 1 对于图 1 所描述的 IMC 系统, 保证闭环系统稳定的充要条件为

$$\left| \frac{G_c(z) G_m(z) z^{-(d-1)}}{1 + G_c(z) G_m(z) z^{-(d-1)}} \right| \leq e_m < 1. \quad (6)$$

证明 由图 1 可得 IMC 系统的闭环传函为

$$H(z) = \frac{G(z) G_c(z)}{1 + G_c(z) (G(z) - G_m(z) z^{-(d-1)})}, \quad (7)$$

其等效开环传递函数为

$$O(z) = \frac{H(z)}{1 - H(z)} = \frac{G(z) G_c(z)}{1 - G_c(z) G_m(z) z^{-(d-1)}}, \quad (8)$$

进而可得到等价的单位负反馈系统控制器的传函为

$$C(z) = \frac{G_c(z)}{1 - G_c(z) G_m(z) z^{-(d-1)}}. \quad (9)$$

将式(9)代入(5), 得

$$\left| \frac{\frac{G_c(z) G_m(z) z^{-(d-1)}}{1 - G_c(z) G_m(z) z^{-(d-1)}}}{1 - \frac{G_c(z) G_m(z) z^{-(d-1)}}{1 - G_c(z) G_m(z) z^{-(d-1)}}} \right| \leq e_m < 1. \quad (10)$$

显而易见, 式(6)成立. 同时, 考虑  $G_m^{-1}(z)$  与  $G_m(z)$  之间的模型误差, 定义

$$e_{cm} = \left\{ G_m^{-1}(z) \left| \frac{G_m^{-1}(z) - G_m(z)}{G_m^{-1}(z)} \right| \leq e_{cm} \right\},$$

有

$$\left| \frac{G_m^{-1}(z) G_m(z) - 1}{G_m(z) G_m^{-1}(z)} \right| \leq e_{cm}, \quad (11)$$

且得到

$$\max(G_m^{-1}(z) G_m(z)) = 1 + e_{cm}. \quad (12)$$

将式(12)代入(6), 得

$$\max \left( \left| \frac{F(z) G_m^{-1}(z) G_m(z) z^{-(d-1)}}{F(z) (1 + e_{cm})} \right| \leq e_m < 1. \quad (13)$$

由式(13)可见,在IMC系统中,对于一定的内模和控制器逆模估计误差界 $e_m$ 和 $e_{cm}$ ,可通过滤波器参数的选取,使式(13)成立,从而保证闭环系统稳定.

3.3 系统性能分析

如图1所示的IMC系统,在一定的建模误差情况下,系统误差方程为

$$E(k) = \frac{1 - F(z)G_m^{-1}(z)G_m(z)z^{-(d-1)}}{1 + F(z)G_m^{-1}(z)[G(z) - G_m(z)z^{-(d-1)}]} \times (R(k) - da(k)), \quad (14)$$

取 $F(z) = (1 - \alpha)z^{-1}/(1 - \alpha z^{-1}), 0 < \alpha < 1; R(k)$ 和 $da(k)$ 均为阶跃信号.则系统稳态误差为

$$\lim_{z \rightarrow 1} E(k) = \frac{1 - z^{-1} - (1 - \alpha)z^{-1}G_m^{-1}(z)}{1 - z^{-1} + (1 - \alpha)z^{-1}G_m^{-1}(z)G_m(z)} \times \frac{G_m(z)z^{-(d-1)}}{G_m^{-1}[G(z) - G_m(z)z^{-(d-1)}]} (R - da), \quad (15)$$

其中 $R$ 和 $da$ 为常值.

由式(4)和(11),分别有 $G_m^{-1}(z)G_m(z) = 1 \pm e_{cm}$ 和 $G_m^{-1}(z)[G(z) - G_m(z)z^{-(d-1)}] = \pm e_m$ ,代入式(15),同时省略二次项 $e_{cm}e_m$ ,得

$$\lim_{z \rightarrow 1} E(k) = \pm \frac{e_{cm}}{(1 \pm e_m)} (R - da). \quad (16)$$

由式(16)可见,IMC系统的跟踪阶跃信号的稳态误差受到逆模和内模的建模误差影响,其中逆模误差是主要产生因素.因此提高建模准确度有利于减小系统稳态误差,提高IMC系统控制精度.

4 仿真结果

考虑如下具有纯滞后环节( $d = 5$ )的非线性系统:

$$y(k+1) = \frac{y(k)}{1 + y^2(k)} - \frac{y(k-1)u^2(k-5)}{1 + 2u^2(k-5)} + 1.25u(k-4) + da(k+1).$$

对系统施加一伪随机信号 $u \in [-2.5, 2.5]$ 和均值为0及方差为0.001的白噪声 $da$ ,得到100组输入输出样本数据.同时选择 $x_1(k) = [y(k), y(k-1), u(k)]$ 和 $t_1(k) = y(k+1)$ ,以 $\{x_1(k), t_1(k)\}_{k=1}^{100}$ 形式来训练内模 $G_m(z)$ ,KRR参数的选择为: $\alpha = 1.2, \beta = 0.01$ ;再选择 $x_2 = [y(k+1), y(k), y(k-1)]$ 和 $t_2 = u(k)$ ,以 $\{x_2(k), t_2(k)\}_{k=1}^{100}$ 来训练逆模 $G_m^{-1}(z)$ ;KRR参数 $\alpha = 1.1, \beta = 0.01$ .由训练完成后的KRR模型分别作为内模和控制器逆模,并选择滤波器 $F(z) = 0.6z^{-1}/(1 - 0.4z^{-1})$ .在仿真中,施加干扰 $da(k) = -0.2(100 - k/140)$ 和 $da(k) = 0.2(180 - k/220)$ .图2为基于KRR建模的IMC系统,实现了对输入的设置跟踪.

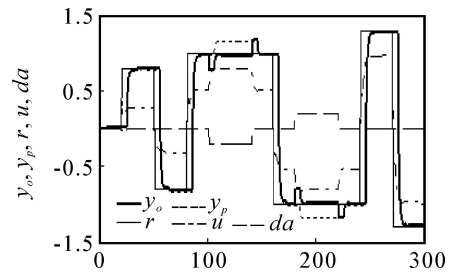


图2 基于KRR的IMC

由RBF神经网络来估计内模和控制器逆模型,在相同的训练样本条件下,内模参数为 $\alpha = 1.2$ ,训练误差 $e = 0.01$ ;逆模参数为 $\alpha = 1.1, e = 0.01$ .由RBF神经网络实现IMC系统的跟踪系统输出结果如图3所示.

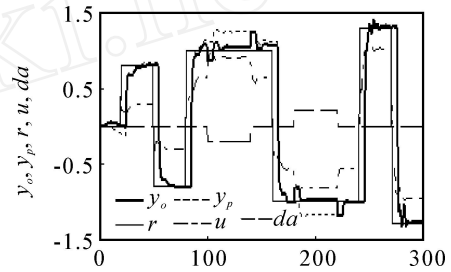


图3 RBF神经网络IMC

对比图2和图3可见,基于KRR方法的IMC系统响应快,缩短了过渡过程,动态特性好,稳态误差更小,跟踪精度更高,具有很强的抗干扰能力.表1为训练样本集大小与IMC最大稳态跟踪误差 $E$ 的关系,其中由式(16)得 $E = \max(|\lim_{z \rightarrow 1} E(k)|)$ .表1说明,基于KRR的最大跟踪误差受训练样本数量的变化影响较小,在样本数量较少时能获得较小的误差值;而基于神经网络IMC跟踪误差随着样本数量的增加而误差值大大减小,且逐渐接近于KRR的误差.这些都说明,基于结构风险最小化的KRR方法能够通过综合评价经验风险最小化和拟合函数复杂度来最终确定拟合函数的具体形式,更加接近于实际风险.优点是避免了样本存在噪声污染时,学习机出现过学习和泛化能力下降现象.因此,在小样本及有噪声存在的情况下,KRR方法更适合于建立IMC系统.

表1 样本数量对最大稳态误差E的影响

样本数	80	100	150	200	250
KRR	0.047	0.036	0.031	0.025	0.022
RBF networks	0.075	0.061	0.049	0.040	0.029

5 结论

基于KRR建模的非线性IMC系统具有传统神经网络对被控对象数学模型依赖性小的特点,同时

又能有效克服神经网络对样本要求高、噪声敏感的缺点. 同时从理论上也分析了存在建模误差下的 IMC 系统的鲁棒稳定性和稳态误差问题, 得到在一定的建模误差存在下, IMC 系统控制精度主要受到逆模型精度的影响. 仿真结果表明, 在小样本和有噪声时, 基于 KRR 的非线性 IMC 系统较 RBF 神经网络的 IMC 系统, 改善了系统的动态性能, 缩短了过渡过程, 减小了系统稳态误差, 提高了 IMC 系统的跟踪精度, 同时具有很强的抗干扰能力.

### 参考文献(References)

- [1] 赵曜. 内模控制发展综述[J]. 信息与控制, 2000, 29(6): 526-531.  
(Zhao Y. A survey of development of internal model control[J]. Information and Control, 2000, 29(6): 526-531.)
- [2] Wang Q G. Partial internal model control[J]. IEEE Trans on Industrial Electronics, 2001, 48(5): 976-983.
- [3] Hunt K, Sbarbaro D. Neural networks for nonlinear internal model control[J]. IEE Proceedings-control Theory and Applications, 1991, 138(5): 431-438.
- [4] 徐湘元, 毛宗源. 基于径向基函数神经网络的内模控制[J]. 电路与系统学报, 1999, 4(2): 86-92.  
(Xu X Y, Mao Z Y. Internal model control based on radial basis function neural networks[J]. J of Circuits and Systems, 1999, 4(2): 86-92.)
- [5] Isabelle R, Leon P. Nonlinear internal model control using neural networks: Application to processes with delay and design issues[J]. IEEE Trans on Neural Networks, 2000, 11(1): 80-90.
- [6] Vapnik V N. The nature of statistical learning theory [M]. New York: Springer-Verlag, 1999.
- [7] Johan A K Suykens. Nonlinear modeling and support vector machines [C]. IEEE Instrumentation and Measurement Technology Conf. Hungary, 2001, 3: 287-294.
- [8] 陈庆伟, 吕朝霞, 胡维礼, 等. 基于逆系统方法的非线性内模控制[J]. 自动化学报, 2002, 28(5): 715-721.  
(Chen Q W, Lv Z X, Hu W L, et al. Nonlinear internal model control based on inverse system method[J]. Acta Automatica Sinica, 2002, 28(5): 715-721.)
- [9] Hoerl A E, Kennard R W. Ridge regression biased estimation for nonorthogonal problems [J]. Technometrics, 1970, 12(4): 55-67.
- [5] Bell M R, Zadeh L A. Decision making in a fuzzy environment[J]. Management Science, 1970, 17(4): 141-146.
- [6] 陈守煜, 胡吉敏. 可变模糊方法及其在工件识别中的应用[J]. 系统工程与电子技术, 2006, 28(9): 1325-1328.  
(Chen S Y, Hu J M. Variable fuzzy method and its application in parts recognition[J]. Systems Engineering and Electronics, 2006, 28(9): 1325-1328.)
- [7] 刘建书, 李人厚, 常宏. 基于相关性函数和最小二乘的多传感器数据融合[J]. 控制与决策, 2006, 21(5): 714-716.  
(Liu J S, Li R H, Chang H. Multi-sensor data fusion based on correlation function and least square [J]. Control and Decision, 2006, 21(5): 714-716.)
- [8] 万树平. 多传感器数据的 Fisher 判别法[J]. 传感器与微系统, 2006, 25(8): 61-63.  
(Wan S P. Method of fisher discrimination for multi-sensors data fusion [J]. Transducer and Microsystem Technologies, 2006, 25(8): 61-63.)
- [9] 王毅, 雷英杰. 基于 Vague 集的多传感器信息融合方法[J]. 计算机技术与发展, 2007, 17(4): 232-235.  
(Wang Y, Lei Y J. Method for multisensor information fusion based on Vague sets[J]. Computer Technology and Development, 2007, 17(4): 232-235.)
- [10] Gau W L, Buehrer D J. Vague sets[J]. IEEE Trans on Systems, Man and Cybernetic, 1993, 23(2): 610-614.
- [11] 王建宏, 李新. 基于 Vague 集的模糊信息融合研究[J]. 华中科技大学学报, 2004, 32(8): 54-56.  
(Wang J H, Li X. Fuzzy data fusion based on Vague set [J]. J of Huazhong University of Science & Technology, 2004, 32(8): 54-56.)
- [12] Chen S M. Measures of similarity between Vague sets [J]. Fuzzy Sets and Systems, 1995, 74(2): 217-223.

(上接第 1099 页)