

文章编号: 1001-0920(2004)02-0223-03

## 支持向量机的多层动态自适应参数优化

朱家元<sup>1</sup>, 杨云<sup>2</sup>, 张恒喜<sup>1</sup>, 任博<sup>1</sup>

(1. 空军工程大学 工程学院 飞机与发动机工程系, 陕西 西安 710038;

2. 北京航空航天大学 工程系统工程系, 北京 100083)

**摘要:** 首先提出了基于多层动态自适应搜索技术的最小二乘支持向量机参数优化方法, 然后采用最小二乘支持向量机对典型非线性控制系统的辨识进行了研究. 辨识结果表明, 最小二乘支持向量机可以用于非线性控制系统辨识, 多层动态自适应搜索方法确定了最优支持向量机参数, 从而获得精确的非线性控制系统辨识结果.

**关键词:** 机器学习; 神经网络; 支持向量机; 最小二乘支持向量机; 非线性控制系统

中图分类号: TP18

文献标识码: A

## Multi-layer adaptive parameters optimization approach for support vector machines

ZHU Jia-yuan<sup>1</sup>, YANG Yun<sup>2</sup>, ZHANG Heng-xi<sup>1</sup>, REN Bo<sup>1</sup>

(1. Department of Aircraft and Engine Engineering, Air Force Engineering University, Xi'an 710038, China;

2. Department of Engineering Systems Engineering, Beijing University of Aeronautics and Astronautics,

Beijing 100083, China. Correspondent: ZHU Jia-yuan, E-mail: ajiaman@yahoo.com.cn)

**Abstract:** A novel least squares support vector machines (LS-SVM) for function estimation is presented. And then, a hyper parameters and kernel parameters optimization approach called multi-layer adaptive best-fitting parameters search method is developed. According to different learning problem, the optimization approach can obtain appropriate LS-SVM parameters adaptively. Non-linear control system identification is studied using the improved LS-SVM. The results show that the optimization approach can induct best-optimized parameters for LS-SVM, and optimized LS-SVM provides excellent control system identification precision.

**Key words:** machine learning; neural networks; support vector machines; least square support vector machines; nonlinear control system

### 1 引言

支持向量机(SVM)是Vapnik<sup>[1]</sup>在统计学习理论基础上,提出的一种新型机器学习方法.由于支持向量机具有完备的统计学习理论基础和出色的学习能力及推广性能,迅速成为智能技术界研究的新热点.过去的几年,支持向量机的理论和算法得到了进一步发展和完善,提出了包括C-SVC,  $\nu$ -SVC, Single-class SVM,  $\epsilon$ -SVR,  $\nu$ -SVR 和 LS-SVM 等一系

列支持向量机算法<sup>[2]</sup>.对于这些不同类型支持向量机,如何确定最优支持向量机参数一直是提高支持向量机学习和泛化能力的主要研究问题之一.

本文首先阐述了用于函数估计的最小二乘支持向量机(LS-SVM)算法,然后提出基于多层动态自适应搜索技术的最小二乘支持向量机参数优化方法,最后采用最小二乘支持向量机对非线性控制系统的辨识进行了研究.结果表明最小二乘支持向量

收稿日期: 2002-12-26; 修回日期: 2003-02-25

基金项目: 空军重点型号工程资助项目

作者简介: 朱家元(1974—),男,湖南东安人,博士生,从事飞机型号发展工程、支持向量机等研究;张恒喜(1937—),男,江苏江堰人,教授,博士生导师,从事飞机型号工程、航空装备发展论证等研究

机可以辨识非线性控制系统, 多层动态自适应搜索方法可以确定最优支持向量机参数, 从而获得精确的非线性控制系统辨识结果

## 2 最小二乘支持向量机

Suykens J A K<sup>[3]</sup>提出的新型支持向量机——最小二乘支持向量机是通过将最小二乘线性系统引入支持向量机, 代替传统的支持向量机采用二次规划方法解决分类和函数估计问题。最小二乘支持向量机仍是基于多类核的机器学习, 即采用核函数, 根据mercer条件, 从原始空间中抽取特征, 将原始空间中的样本映射为高维特征空间中的一个向量, 以解决原始空间中线性不可分的问题

用于函数估计的最小二乘支持向量机算法推导如下: 设训练样本集

$$D = \{(x_k, y_k) | k = 1, 2, \dots, N\},$$

$$x_k \in R^n, y_k \in R,$$

其中:  $x_k$  为输入数据,  $y_k$  为输出数据。在权  $w$  空间中的函数估计问题可以描述求解下面问题:

$$\min_{w, b, e} J(w, e) = \frac{1}{2} w^T w + \frac{1}{2} \sum_{k=1}^N e_k^2, \quad (1)$$

$$\text{s.t. } y_k = w^T \mathcal{Q}(x_k) + b + e_k$$

其中  $\mathcal{Q}(\cdot): R^n \rightarrow R^{n_h}$  为核空间映射函数,  $w \in R^{n_h}$  为权矢量,  $e_k \in R$  为误差变量,  $b$  为偏差量,  $\gamma$  为可调参数

根据式(1), 可定义拉格朗日函数

$$L(w, b, e; \alpha) = J(w, e) - \sum_{k=1}^N \alpha_k \{w^T \mathcal{Q}(x_k) + b + e_k - y_k\}, \quad (2)$$

其中: 拉格朗日乘子  $\alpha_k \in R$ 。通过  $L$  对  $w, b, e_k$  和  $\alpha_k$  求偏导等于零, 对式(2)进行优化, 消除变量  $w, e$ , 可得矩阵方程

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & \Omega + \frac{1}{\gamma} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}. \quad (3)$$

其中

$$x = [x_1, \dots, x_N], y = [y_1, \dots, y_N],$$

$$1_v = [1, \dots, 1], \alpha = [\alpha_1, \dots, \alpha_N],$$

$$\Omega_{kl} = \Psi(x_k, x_l) = \mathcal{Q}(x_k)^T \mathcal{Q}(x_l).$$

根据mercer条件, 存在映射函数  $\mathcal{Q}$  和核函数  $\Psi(\cdot, \cdot)$ , 使得

$$\Psi(x_k, x_l) = \mathcal{Q}(x_k)^T \mathcal{Q}(x_l). \quad (4)$$

最小二乘支持向量机的函数估计为

$$y(x) = \sum_{k=1}^N \alpha_k \Psi(x, x_k) + b \quad (5)$$

其中:  $a$  和  $b$  由式(3)求解出; 核函数  $\Psi(\cdot, \cdot)$  的形式可以是:  $\Psi(x, x_k) = x_k^T x$  (线性核);  $\Psi(x, x_k) = (x_k^T x + 1)^d$  ( $d$  阶多项式核);  $\Psi(x, x_k) = \exp\{-\frac{\|x - x_k\|^2}{2\sigma^2}\}$  (径向基核) 等

## 3 LS-SVM 参数的多层动态自适应优化

由上述算法推导可知, 对于采用径向基核的最小二乘支持向量机的主要参数是正则化参数  $\gamma$  和径向基核参数  $\sigma$ , 这两个参数在很大程度上决定了最小二乘支持向量机的学习能力和泛化能力。在此以采用径向基核函数的最小二乘支持向量机参数优化为例, 说明参数的多层动态自适应优化方法。具体步骤如下:

1) 确定参数  $\gamma$  和  $\sigma$  取值范围, 依据最小二乘支持向量机原理, 最大取值范围是  $\gamma \in [0.1, 10000]$ ,  $\sigma \in [0.1, 100]$

2) 在最大取值范围内选取参数值, 构建参数对  $(\gamma_i, \sigma_j)$  二维网格平面, 其中:  $i = 1, 2, \dots, m; j = 1, 2, \dots, n$ 。例如两个参数各选取 10 个数值, 则构成  $10 \times 10$  网格平面和 100 个  $(\gamma_i, \sigma_j)$  参数对。对于参数选取有两种方法: 第 1 种是首先确定两个参数的取值范围, 再根据所需参数对数进行均匀取值; 第 2 种是根据学习样本的特征和经验确定参数对值

3) 输入每个网格结点的参数对  $(\gamma_i, \sigma_j)$  到最小二乘支持向量机中, 采用学习样本进行学习, 并输出学习误差。取最小误差对应的节点值  $(\gamma_i, \sigma_j)_{E_{\min}}$  为最优参数对

4) 如果学习精度没有达到所要求, 则以  $(\gamma_i, \sigma_j)_{E_{\min}}$  为中心, 构建新二维网络平面, 选取数值相近的参数值进一步学习, 从而获得更高精度的学习结果。这个新参数选取过程是自动执行的, 经验表明, 一般以  $(\gamma_i, \sigma_j)_{E_{\min}}$  值的 0.01 ~ 5 倍为一个扩展网格宽度, 构建新参数对  $(\gamma_i, \sigma_j)$  二维网格平面, 其中:  $i = 1, 2, \dots, k; j = 1, 2, \dots, l$ 。以此类推, 可构造多层参数优化网格平面, 不断优化最小二乘支持向量机参数, 直到达到需要的学习精度

## 4 非线性控制系统辨识

采用基于多层动态自适应参数优化技术的最小二乘支持向量机对典型的非线性控制系统<sup>[4]</sup> 辨识进行研究。辨识的对象满足差分方程

$$y(k+1) = 0.3y(k) + 0.6y(k-1) + f(u(k)), \quad (6)$$

其中未知函数

$$f(u) = 0.6 \sin(\pi u) + 0.3 \sin(3\pi u) +$$

$$0.1 \sin(57\pi u).$$

为了辨识这一系统的非线性部分, 采用如下串行-并行模型的差分方程:

$$\hat{y}(k+1) = 0.3y(k) + 0.6y(k-1) + \hat{F}(u(k)), \quad (7)$$

其中  $\hat{F}(\cdot)$  采用最小二乘支持向量机进行辨识. 非线性部分和系统的输入都为正弦曲线:  $u(k) = \sin(2\pi k/250)$ . 误差测度采用如下性能指标:

1) 辨识误差为

$$\frac{1}{k} \sum_{i=1}^k |T(i) - O(i)|,$$

其中:  $T(i)$  为期望输出,  $O(i)$  为辨识输出

2) 辨识过程: 确定最小二乘支持向量机辨识

时间过程. 设定

$$u(k) = \begin{cases} \sin(2\pi k/250), & k = 1 \sim 250 \text{ 时刻;} \\ 0.5 \sin(2\pi k/250) + \sin(2\pi k/250), & k = 250 \sim 500 \text{ 时刻;} \\ 0.5 \sin(2\pi k/250), & k = 500 \text{ 时刻;} \end{cases} \quad (8)$$

用于检验优化的最小二乘支持向量机对系统变化的适应能力

设置最小二乘支持向量机的参数  $\gamma$  和  $\sigma$  的取值范围. 采用两层网格平面优化, 根据非线性控制系统特征, 在第 1 参数优化层中, 取  $\gamma = [0.1, 1, 10, 50, 100, 500, 1000, 2500, 5000, 10000]$ ,  $\sigma = [0.1, 0.2, 0.5, 1, 5, 10, 15, 25, 50, 100]$ , 即采用  $10 \times 10$  网格结构. 在每一个节点参数对中, 采用最小二乘支持向量机对系统和非线性部分进行辨识, 获得的最优参数对  $(\gamma_i, \sigma_j)_{E_{min}}$  为  $(5000, 0.5)$ , 系统辨识误差为  $0.1642$ . 然后, 在第 2 参数优化层中, 以  $(\gamma_i, \sigma_j)_{E_{min}}$  为网格平面中心, 以  $(\gamma_i)_{E_{min}}$  值的  $\pm 0.1$  倍值为  $\gamma$  向扩展网格宽度, 以  $(\sigma_j)_{E_{min}}$  的  $\pm 0.05$  倍值为  $\sigma$  向扩展网格宽度, 构建  $10 \times 10$  网格平面, 再次采用最小二乘支持向量机对系统和非线性部分进行辨识, 获得的最优参数对为  $(4600, 0.425)$ , 辨识误差为  $0.0537$ , 即在采用此参数对时, 最小二乘支持向量机获得了十分理想的辨识结果.

图 1 是  $u(k)$  随时间变化过程, 基于最优参数的最小二乘支持向量机对非线性部分辨识的结果如图 2 和图 3 所示, 系统实际输出与辨识输出如图 4 所示, 其中实线为实际输出, 虚线为辨识输出.

由图 2~ 图 4 可见, 采用参数优化后的最小二乘支持向量机具有优秀的非线性控制系统辨识能力.

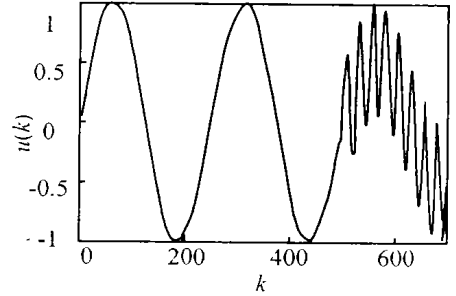


图 1  $u(k)$  曲线

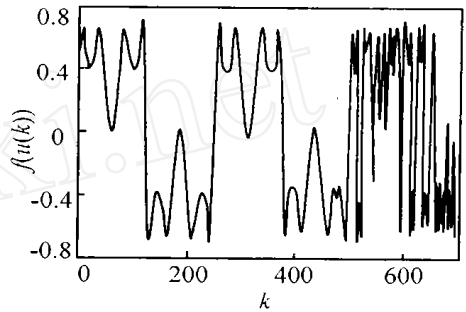


图 2  $k = 1 \sim 700$  非线性部分辨识

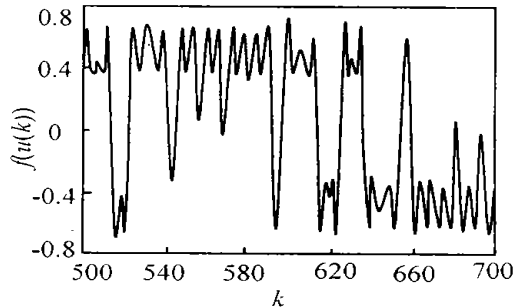


图 3  $k = 500 \sim 700$  的非线性部分辨识

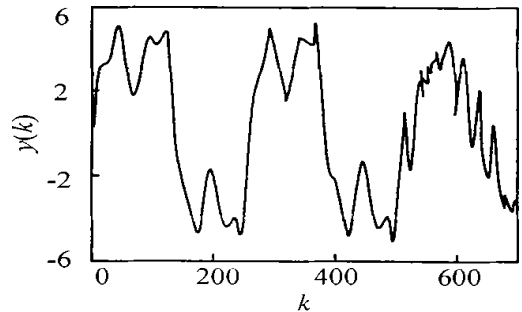


图 4 系统实际输出与辨识输出

## 5 结 论

最小二乘支持向量机是引入最小二乘线性系统的基于多类核学习的新型支持向量机. 本文提出了最小二乘支持向量机参数的多层动态自适应搜索优化方法, 并采用最小二乘支持向量机对典型非线性控制系统的辨识进行了研究.

(下转第 229 页)

因此, 本文方法是一种合理、可靠和系统的最优参数确定方法, 而且可推广应用于其他优化算法

另外, 由表 1 可得到关于最优参数组合的一些特点, 即较高的初温可保证算法具备足够的概率突跳能力, 较慢的退温速率可保证算法的退温历程接近稳态。同时, 在有限计算量条件下, 即总评价次数 ( $N$ ) 固定的情况, 算法需要在同一温度下的抽样数 ( $L$ ) 和退温次数 ( $N/L$ ) 两者间作出权衡, 仿真结果显示, 一般  $L$  不能取值过小以保证在温度稍高时能使 Metropolis 抽样接近稳态

## 6 结 语

本文将 SA 的最优算法参数确定问题表述为一个随机优化问题, 并提出基于序优化和最优计算量分配技术的系统性方法。基于典型算例的仿真结果和比较, 验证了该方法的有效性和可靠性。进一步的工作是将该方法推广到其他优化算法的参数与操作择优的问题

### 参考文献(References):

- [1] 王凌. 智能优化算法及其应用[M]. 北京: 清华大学出版社, 2001.
- [2] Park M W, Kim Y D. A systematic procedure for setting parameters in simulated annealing algorithms[J]. *Computers and Operations Research*, 1998, 25 (3): 207-217.
- [3] Wang T Y, Wu K B. A parameter set design procedure for the simulated annealing algorithm under computational time constraint[J]. *Computers and Operations Research*, 1999, 26(7): 665-678.
- [4] Bolte A, Thonemann U W. Optimizing simulated annealing schedules with genetic programming[J]. *European J of Operational Research*, 1996, 92(2): 402-416.
- [5] Ho Y C, Sreenivas R, Vakili P. Ordinal optimization of discrete event dynamic systems[J]. *Discrete Event Dynamic Systems*, 1992, 2(2): 61-88.
- [6] Dai L. Convergence properties of ordinal comparison in the simulation of discrete event dynamic systems[J]. *J of Optimization Theory and Applications*, 1996, 91(2): 363-388.
- [7] Chen C H, Lin J, Yucesan E, et al. Simulation budget allocation for further enhancing the efficiency of ordinal optimization[J]. *Discrete Event Dynamic Systems*, 2000, 10(3): 251-270.
- [8] Mitra D, Romeo F, Sangiovanni-Vicentelli A. Convergence and finite-time behavior of simulated annealing[J]. *Advances in Applied Probability*, 1986, 18(4): 747-771.
- [9] Reeves C R. A genetic algorithm for flow shop sequencing[J]. *Computers and Operations Research*, 1995, 22(1): 5-13.

(上接第 225 页)

结果表明, 最小二乘支持向量机可以用于非线性控制系统辨识, 多层动态自适应搜索技术确定了最优支持向量机参数, 从而获得精确的非线性控制系统辨识结果。同时, 多层动态自适应搜索方法也可适当推广到其他类型支持向量机参数优化中。

### 参考文献(References):

- [1] Vapnik V N. *Statistical Learning Theory* [M]. New York: John Wiley, 1998.
- [2] Chang C-C, Hsu C-W, Lin C-J. The analysis of decomposition methods for support vector machines[J]. *IEEE Trans on Neural Networks*, 2000, 11(4): 1003-1008.
- [3] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers[J]. *Neural Processing Letters*, 1999, 9(3): 293-300.
- [4] 王立新. 自适应模糊系统与控制——设计与稳定性分析[M]. 北京: 国防工业出版社, 1995: 39-56.
- [5] 朱家元, 张恒喜. 基于支持向量机的 R&D 项目中止决策研究[J]. *计算机科学*, 2002, 29(9): 440-442.  
(Zhu Jia-yuan, Zhang Heng-xi. R&D project termination decision based on support vector machines[J]. *Computer Science*, 2002, 29(9): 440-442.)
- [6] Zhu Jia-yuan, Ren Bo, Zhang Heng-xi, et al. Time series prediction via new support vector machines[A]. *IEEE in Proc of ICMLC 2002* [C]. Beijing, 2002: 364-366.
- [7] Zhu Jia-yuan, Zhang Heng-xi, Guo Ji-lian, et al. Data distributions automatic identification based on SOM and support vector machines[A]. *IEEE in Proc of ICMLC 2002* [C]. Beijing, 2002: 340-344.