

文章编号: 1001-0920(2006)01-0111-03

## 基于混沌优化算法的支持向量机参数选取方法

袁小芳, 王耀南

(湖南大学 电气与信息工程学院, 长沙 410082)

**摘要:** 支持向量机(SVM)的参数取值决定了其学习性能和泛化能力。对此, 将SVM参数的选取看作参数的组合优化, 建立组合优化的目标函数, 采用变尺度混沌优化算法来搜索最优目标函数值。混沌优化算法是一种全局搜索方法, 在选取SVM参数时, 不必考虑模型的复杂度和变量维数。仿真表明, 混沌优化算法是选取SVM参数的有效方法, 应用到函数逼近时具有优良的性能。

**关键词:** 机器学习; 支持向量机; 混沌优化; 参数选取

**中图分类号:** TP18 **文献标识码:** A

## Selection of SVM Parameters Using Chaos Optimization Algorithm s

YUAN Xiao-fang, WANG Yao-nan

(College of Electrical and Information Engineering, Hu'nan University, Changsha 410082, China Correspondent: YUAN Xiao-fang, E-mail: yuanxiaof@21cn.com)

**Abstract** Appropriate parameters are very crucial to support vector machines (SVM) learning results and generalization ability. The selection problem of SVM parameters is considered as a compound optimization problem. Then objective function of optimization problem is set and a mutative scale chaos optimization algorithm is employed to search optimal objective function. Chaos optimization algorithm is global search method and it need not to consider SVM dimensionality and complexity. Simulations show that the proposed method is an effective approach for parameter selection and the good performance for function approximation is obtained.

**Key words:** Machine learning; Support vector machines (SVM); Chaos optimization; Parameters selection

### 1 引言

支持向量机(SVM)是Vapnik等学者在统计学习理论上提出的一种新型机器学习方法<sup>[1,2]</sup>。与神经网络等学习算法相比, SVM具有小样本学习、泛化能力强等特点, 能有效地避免过学习、局部极小点以及“维数灾难”等问题<sup>[1]</sup>。自1995年以来, SVM的实用算法研究、设计和实现已经取得了丰硕的成果<sup>[2]</sup>, 目前SVM已广泛地应用于文本分类、图像识别、函数估计、非线性建模等多个领域<sup>[3]</sup>。

SVM具有优良的学习能力和推广能力, 然而其性能依赖于参数的选取。针对参数选取问题, 一些学

者从不同的角度进行了研究和探讨<sup>[4~6]</sup>。本文将SVM的参数选取看作参数的组合优化, 对组合优化问题建立目标函数, 采用一种变尺度混沌优化算法<sup>[7]</sup>来搜索最优的目标函数值, 从而找到合适的参数取值。混沌优化算法是一种全局搜索方法, 用于选取SVM参数时, 无需考虑SVM模型的复杂度和变量维数, 可以作为一种普遍的选取参数方法。

本文首先介绍了SVM回归及其参数, 然后建立优化目标函数, 采用混沌优化算法来搜索优化值, 最后将其应用于非线性函数逼近, 取得了优良的效果, 进而验证了方法的有效性。

收稿日期: 2004-11-15; 修回日期: 2005-03-10

基金项目: 国家自然科学基金项目(60375001); 高校博士点基金项目(20030532004)。

作者简介: 袁小芳(1979-), 男, 湖南安仁人, 博士生, 从事智能控制、机器学习等研究; 王耀南(1957-), 男, 昆明人, 教授, 博士生导师, 从事智能控制、智能信息处理等研究。

## 2 SVM 及其参数

### 2.1 SVM 介绍

SVM 包括分类和回归两种基本形式, 这里以 SVM 回归为例来说明参数的选取. 学习样本为  $\{(x_1, y_1), \dots, (x_n, y_n)\} \subset R^n \times R$ , SVM 回归用一个非线性映射  $\Phi$  将数据映射到高维特征空间, 再在高维特征空间进行线性回归. 将 SVM 回归表示为

$$f(x) = (w \Phi(x)) + b, \quad (1)$$

其中:  $w \in R^n, b \in R$ , SVM 依据结构风险最小化原则, 将其学习过程转化为优化问题, 即

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i + \xi_i^*, \quad (2)$$

$$\text{s.t. } y_i - f(x_i) \leq \epsilon + \xi_i,$$

$$f(x_i) - y_i \leq \epsilon + \xi_i^*,$$

$$\xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, n. \quad (3)$$

其中:  $\xi_i$  和  $\xi_i^*$  为松弛变量,  $\epsilon$  用于定义  $\epsilon$  线性不敏感损失函数, 即

$$L^\epsilon(x, y, f) = \max(0, |y - f(x)| - \epsilon). \quad (4)$$

方程(2)通过对偶形式的 Lagrange 多项式, 可转化为

$$\max \sum_{i=1}^n y_i (\alpha_i^* - \alpha_i) - \epsilon \sum_{i=1}^n (\alpha_i^* + \alpha_i) -$$

$$\frac{1}{2} \sum_{i,j=1}^n (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) K(x_i, x_j), \quad (5)$$

$$\text{s.t. } \sum_{i=1}^n \alpha_i^* = \sum_{i=1}^n \alpha_i,$$

$$C \alpha_i^*, \alpha_i \geq 0, i = 1, 2, \dots, n. \quad (6)$$

其中:  $\alpha_i^*$  和  $\alpha_i$  为 Lagrange 乘子,  $K(x_i, x_j)$  为满足 Mercer 条件<sup>[1]</sup> 的核函数, 这里采用 RBF 核函数

$$K(x_i, x_j) = \exp(-|x_i - x_j|^2 / \sigma^2). \quad (7)$$

由于 SVM 的稀疏性, 二次规划(5)的解中只有少量样本的系数  $(\alpha_i^* - \alpha_i)$  不为 0, 系数  $(\alpha_i^* - \alpha_i)$  不为 0 所对应的样本  $(x_i, y_i)$  就是支持向量. 设支持向量个数为  $g$ , 因而回归函数形式为

$$f(x) = \sum_{i=1}^g (\alpha_i^* - \alpha_i) K(x, x_i) + b \quad (8)$$

### 2.2 SVM 参数

SVM 参数的取值对于其学习能力和推广能力都有很大的影响, 确定参数取值也是 SVM 的一个重要研究内容. 对于 RBF 核函数的 SVM, 参数包括边界参数  $C$ , 核宽度  $\sigma$  和不敏感参数  $\epsilon$ .

一般而言, 参数  $C, \sigma$  和  $\epsilon$  的取值与学习样本和实际问题相关. 参数  $C$  是在结构风险和样本误差之间作出折衷, 参数  $C$  的取值与可容忍的误差相关, 较大的  $C$  值允许较小的误差, 较小的  $C$  值则允许较大的误差. 核宽度  $\sigma$  与学习样本的输入空间范围或宽

度相关, 样本输入空间范围越大,  $\sigma$  取值越大, 反之, 样本输入空间范围越小, 则  $\sigma$  取值越小. 不敏感参数  $\epsilon$  则与噪声水平相关, 其取值一般与噪声水平成比例关系, 实际应用中常常根据预先估计的噪声水平来确定参数  $\epsilon$ .

### 3 基于混沌优化算法选取 SVM 参数

混沌运动具有遍历性、随机性、“规律性”等特点, 混沌运动能在一定范围内按其自身的“规律”不重复地遍历所有状态, 因而可用于优化搜索. 本文将 SVM 的参数选取看作参数的组合优化, 对组合优化问题建立目标函数, 采用一种变尺度混沌优化算法来搜索最优的目标函数值, 从而找到合适的参数取值.

这里采用如下 Logistic 映射:

$$t_{k+1} = \mu t_k (1 - t_k), \quad (9)$$

取  $\mu = 4$ , 系统完全处于混沌状态. 以方均误差 (MSE)

$$\frac{1}{n} \sum_{i=1}^n (y_i - f(x_i))^2$$

来描述 SVM 回归与参考模型之间的偏差, 其中:  $n$  为样本个数,  $y$  为参考模型,  $f(x)$  为 SVM 回归.

优化算法的目标函数是: 选取最佳的参数组合, 使得 SVM 回归与参考模型具有最小的方均误差, 即

$$\min f(z_1, \dots, z_i) = \min \text{MSE},$$

$$\alpha_i \in [z_i, b_i], i = 1, 2, 3. \quad (10)$$

这里优化变量  $z_i$  总共为 3 个, 对应于参数  $C, \sigma$  和  $\epsilon$ .  $[a_i, b_i]$  为各个变量  $z_i$  的定义域, 对于不同的参考模型和样本数据, 变量的定义域空间也有所不同.

变尺度混沌优化算法步骤如下 ( $i = 1, 2, 3$ ):

**Step1** 初始化  $k = 0$ , 细搜索标志  $r = 0$ ,  $t_i^k = t_i(0)$ ,  $t_i^* = t_i(0)$ ,  $a_i^r = a_i$ ,  $b_i^r = b_i$ , 当前最优目标函数值  $f^*$  初始化为一个较大正数.

**Step2** 将  $t_i^k$  映射到优化变量取值区间成为  $z_i^k$

$$z_i^k = a_i^r + (b_i^r - a_i^r) t_i^k. \quad (11)$$

**Step3** 优化搜索. 若  $f(z_i^k) < f^*$ , 则  $f^* = f(z_i^k)$ ,  $t_i^* = t_i^k$ ; 否则继续.

**Step4**  $k = k + 1$ ,  $t_i^k = \mu t_i^{k-1} (1 - t_i^{k-1})$ .

**Step5** 重复 Step2~4, 直到一定步数内  $f^*$  保持不变为止, 然后进行以下步骤.

**Step6** 缩小变量的搜索范围, 即

$$a_i^{r+1} = z_i^* - \Phi(b_i^r - a_i^r),$$

$$b_i^{r+1} = z_i^* + \Phi(b_i^r - a_i^r), \quad (12)$$

其中:  $\Phi \in (0, 0.5)$ ,  $z_i^* = a_i^r + t_i^* (b_i^r - a_i^r)$  为当前最优解. 为保证新范围不越界, 作以下处理:

若  $a_i^{r+1} < a_i^r$ , 则  $a_i^{r+1} = a_i^r$ ; 若  $b_i^{r+1} > b_i^r$ , 则  $b_i^{r+1} = b_i^r$ .

$b_i^k$  同时, 依据新范围将  $z_i^*$  映射到  $t_i^*$ .

**Step7** 用式 (13) 确定新的混沌变量  $m_i^k$

$$m_i^k = (1 - \alpha) t_i^* + \alpha t_i^k, \quad (13)$$

其中  $\alpha$  为一较小正数 重复 Step2~ 4 操作直到一定步数内  $f^*$  保持不变为止, 然后进行以下步骤

**Step8**  $r = r + 1$ , 减小  $\alpha$  值重复 Step6~ 7 操作

重复 Step8 若干次后, 结束寻优计算,  $z_i^*$  即为搜索到的最优参数取值

### 4 仿真研究

**实验 1** 非线性函数模型为 Hermite 函数

$$y = 1.1 * (1 - x + 2x^2) e^{-\frac{x^2}{2}},$$

其中  $x \in [0, 6]$ , 在此区间上随机取 100 对  $(x_i, y_i)$  为学习样本, 并等间隔取 40 对  $(x_i, y_i)$  为测试样本 采用变尺度混沌优化算法来确定 SVM 参数值, 以最小化方均误差为目标, 搜索范围  $C \in [0.5, 15]$ ,  $\sigma \in [0.5, 5]$ ,  $\epsilon \in [0.01, 0.2]$  混沌优化算法得到的参数值为  $C = 6.4, \sigma = 3.9, \epsilon = 0.026$

为了说明方法的有效性, 这里将本文方法与文献[4~ 6]中的 SVM 参数取值方法进行了对比 表 1 描述了 4 种不同参数取值方法时, SVM 应用于函数逼近的测试样本误差 仿真结果表明, 基于混沌优化方法选取参数的 SVM 具有优良的性能, 函数逼近效果好, 测试误差比较小

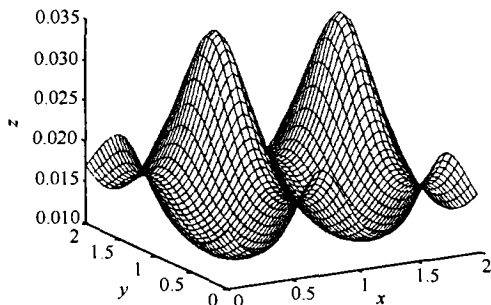
表 1 不同 SVM 参数取值方法时的测试对比

参数取值方法	方均误差	最大正误差	最大负误差
文献[4]	0.052	0.171	- 0.061
文献[5]	0.049	0.097	- 0.084
文献[6]	0.063	0.134	- 0.090
本文方法	0.037	0.078	- 0.043

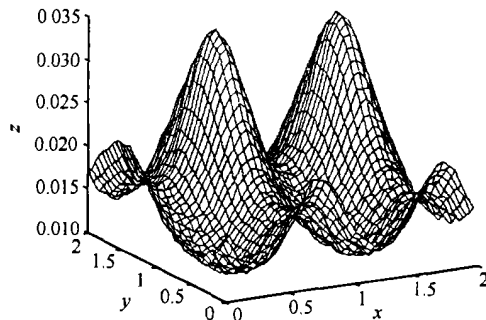
**实验 2** 设二维函数为

$$z = \frac{1 + \sin(xy)}{4 + \sin(2\pi x) + \sin(\pi y)},$$

其中:  $x \in [0, 2], y \in [0, 2]$ , 仍然采用变尺度混沌优化算法优化 SVM 参数 仿真结果如图 1 所示, 图 1 (a)



(a) 二维函数参考模型



(b) SVM 逼近模型

图 1 SVM 逼近二维函数

和 (b) 分别表示二维函数的参考模型和 SVM 逼近模型, 方均误差  $MSE = 0.00051$ . 结果表明, 混沌优化参数的 SVM 对于二维非线性函数也具有很好的逼近效果

### 5 结论

针对 SVM 参数选取问题, 本文将其看作参数的组合优化问题, 建立组合优化的目标函数, 采用混沌优化方法来搜索最优的目标函数 研究表明, 混沌优化算法是确定 SVM 参数取值的一种有效方法, 在非线形函数逼近的仿真中得到了验证 对于其他类型或应用场合的 SVM, 也可以采用该方法选取参数值

### 参考文献 (References)

- [1] Vapnik V. An Overview of Statistical Learning Theory [J]. *IEEE Trans on Neural Networks*, 1999, 10 (5): 988-999.
- [2] 许建华, 张学工, 李衍达. 支持向量机的新发展[J]. *控制与决策*, 2004, 19(5): 481-484. (Xu J J, Zhang X G, Li Y D. Advances in Support Vector Machines [J]. *Control and Decision*, 2004, 19 (5): 481-484.)
- [3] Christopher J C Burges. A Tutorial on Support Vector Machines for Pattern Recognition [J]. *Data Mining and Knowledge Discovery*, 1998, 2 (2): 121-167.
- [4] Vladimir Cherkassky, Yunqian Ma. Practical Selection of SVM Parameters and Noise Estimation for SVM Regression [J]. *Neural Networks*, 2004, 17 (1): 113-126.
- [5] 朱家元, 杨云, 张恒喜, 等. 支持向量机的多层动态自适应参数优化[J]. *控制与决策*, 2004, 19(2): 223-225. (Zhu J Y, Yang Y, Zhang H X, et al. Multi-layer Adaptive Parameters Optimization Approach for Support Vector Machines [J]. *Control and Decision*, 2004, 19(2): 223-225.)

(下转第 117 页)

$$F_{f \max} = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 2/9 & 1/9 \\ 1/9 & 1/9 & 3/9 \end{bmatrix},$$

$$R_x = 1, I = 0, u = X, y = Y.$$

为了使实验结果的比较具有公正性,均采取了 50 步的迭代过程 下面是对两种不同白细胞(单核、中性干状)的血液涂片进行白细胞检出的结果 在第 1 组图片中给出色标,对应其二值化输出

从图 1 中可以发现经 IFCNN 处理后的分割图虽然其边缘并没有百分之百的二值化,但对比经 FCNN 处理后的分割图,其边缘更接近真实边缘,而所得白细胞更加完整 为了更好的量化所得实验结果,在本文中,笔者定义了边缘检全率

$$\text{Binary rate} = \left( \frac{E1}{E2} \right) \times 100\%.$$

式中  $E1$  表示已二值化边缘像素点数,  $E2$  表示本应二值化边缘像素点数 对应于 6 类白细胞图像的量化值如表 1 所示,此表很清楚地显示了 IFCNN 的优势

表 1 边缘检全率比较 (%)

边缘 检全率	单核	淋巴	嗜碱	嗜酸	中性 干状	中性 分叶
FCNN	2.0	2.3	1.9	2.4	5.0	5.8
IFCNN	98.4	97.6	98.4	97.9	81.8	88.5

#### 4 结 语

血液显微图像分割的复杂性使得其分割成为一个有难度但又不失挑战意义的课题 尽管中外学者已经提出了许多的分割技术,但在很多情况下,为了降低图像分割误差给后续识别处理带来的影响,单个白细胞依然通过手工方式检出 本文就 FCNN 在细胞显微图像分割中的不足,提出了一种新模型 IFCNN,该模型在对血液显微图像白细胞自动识别

中的白细胞检出有很好的适应性,其检准率、边缘检全性高,实时性能强 实验结果证明了其优良性能,为后续处理奠定了良好的基础 并且由于模糊状态模板的引入,使得参数模板调节有更大的灵活性<sup>[6]</sup>,并有可能得出适用于医学图像处理的参数模板,这是笔者下一步研究的工作

#### 参考文献(References)

- [1] Chua L O, Yang L. Cellular Neural Networks: Applications [J]. *IEEE Trans on Circuits System*, 1988, 32(10): 1273-1290
- [2] Yang T, Yang L B. Fuzzy Cellular Neural Network: Theory [A]. *Proc Int'l Workshop on Cellular Neural Networks and Their Applications* [C]. New York: IEEE, 1996: 225-230
- [3] 王敏, 王士同. 基于MAT 理论和新型模糊细胞神经网络的医学细胞识别研究[D]. 镇江: 华东船舶工业学院, 2002: 22-33  
(Wang M, Wang S T. Study on Medical Cell Recognition Based on MAT and new fuzzy cellular Neural Network [D]. Zhenjiang: East China Shipbuilding Institute, 2002: 22-33)
- [4] Tao Y, Yang L B. The Global Stability of Fuzzy Cellular Neural Network [J]. *IEEE Trans on Circuits System*, 1996, (43): 880-883
- [5] 陈安平, 袁顺莲. 模糊细胞神经网络的全局稳定性[J]. *柳州师范高等专科学校学报*, 2001, 22(5): 6-9  
(Cheng A P, Yuan S L. The Global Stability of Fuzzy Cellular Neural Network [J]. *J of Binzhou Normal Educational College*, 2001, 22(5): 6-9).
- [6] Lin C T, Chang C L, Cheng W C. A Recurrent Fuzzy Cellular Neural Network System with Automatic Structure and Template Learning [J]. *IEEE Trans on Circuits System*, 2004, 51(2): 1057-1122

(上接第 113 页)

[6] 董春曦, 饶鲜, 杨绍全, 等. 支持向量机参数选择方法研究[J]. *系统工程与电子技术*, 2004, 26(8): 1117-1120  
(Dong C X, Rao X, Yang S Q, et al. Method for Selecting the Parameters of Support Vector Machines [J]. *Systems Engineering and Electronics*, 2004, 26(8): 1117-1120)

[7] 张彤, 王宏伟, 王子才. 变尺度混沌优化方法及其应用[J]. *控制与决策*, 1999, 14(3): 285-287.  
(Zhang T, Wang H W, Wang Z C. Mutative Scale Chaos Optimization Algorithm and Its Application [J]. *Control and Decision*, 1999, 14(3): 285-287.)