DOI：10.13195/j.cd.2004.05.2.xujh.001

第 19 卷 第 5 期 控 制 与 决 策 2004 年 5 月
Vol. 19 No. 5 *Control and Decision* May 2004

文章编号: 1001-0920(2004)05-0481-04

支持向量机的新发展

许建华[1], 张学工[2], 李衍达[2]
(1. 南京师范大学 数学与计算机学院, 江苏 南京 210097; 2. 清华大学 自动化系, 北京 100084)

摘 要: Vapnik 等学者首先提出了实现统计学习理论中结构风险最小化原则的实用算法—支持向量机,比较成功地解决了模式分类问题. 其后, 机器学习界兴起了研究统计学习理论和支持向量机的热潮, 引人瞩目的研究分支有从最优化技术出发改进或改造支持向量机, 依据统计学习理论和支持向量机的优点设计新的非线性机器学习算法等. 对此, 较为系统地回顾了近 10 年来算法研究领域的新发展.

关键词: 机器学习; 统计学习理论; 支持向量机
中图分类号: TP18 文献标识码: A

Advances in support vector machines

XU Jian-hua[1], *ZHANG Xue-gong*[2], *LI Yan-da*[2]
(1. School of Mathematical and Computer Science, Nanjing Normal University, Nanjing 210097, China; 2. Department of Automation, Tsinghua University, Beijing 100084, China. Correspondent: XU Jian-hua, E-mail: xujianhua@email.njnu.edu.cn)

Abstract: Vapnik and his collaborators proposed a useful algorithm: support vector machines, which can implement the structural risk minimization principle in statistical learning theory. This novel algorithm handles the classification problems successfully. Since then more attentions have been paid to statistical learning theory and support vector machines. The attractive research areas are to improve or modify support vector machines by optimization techniques, and to design the novel non-linear machine learning algorithms based on statistical learning theory and some ideas in support vector machines, etc. The advances in such algorithm studies in the last ten years are reviewed.

Key words: machine learning; statistical learning theory; support vector machines

1 引 言

Vapnik 等学者提出的支持向量机算法(SVM)[1~4]和机器学习的统计学分析专著——统计学习理论(SLT)[5~7], 是近 10 年来机器学习、模式识别以及神经网络界最有影响力的成果之一. 支持向量机分类算法具有 4 个显著特点: 1) 利用大间隔的思想降低分类器的 VC 维, 实现结构风险最小化原则, 控制分类器的推广能力; 2) 利用 Mercer 核实现线性算法的非线性化; 3) 稀疏性, 即少量样本(支持向量)的系数不为零, 就推广性而言, 较少的支持向量数在统计意义上对应好的推广能力, 从计算角度看, 支持向量减少了核形式判别式的计算量; 4) 算法设计成凸二次规划问题, 避免了多解性.

自 1995 年以来, 在实用算法研究、设计和实现方面已取得丰硕的成果, 可归结为几个大的研究方向: 1) 提高 SVM 的计算速度, 以便于处理大规模问题, 如序列最小化算法[8]等; 2) 利用最优化技术改进或改造支持向量机形式, 简化计算过程, 如线性

收稿日期: 2003-03-31; 修回日期: 2003-06-09.
基金项目: 国家自然科学基金资助项目(60275007).
作者简介: 许建华(1962—), 男, 浙江长兴人, 高级工程师, 博士, 从事模式识别、机器学习等研究; 李衍达(1936—), 男, 广东南海人, 教授, 中国科学院院士, 从事信号处理、智能控制和生物信息学等研究.

SVM[7], LS-SVM[9]等; 3) 依据结构风险最小化原则和支持向量机的某些思想提出新的算法, 如 ν-SVM[10], 广义 SVM[11] 等算法; 4) 利用结构风险最小原则、核思想和正则化技术等改造传统的线性算法, 构造出相应的核形式, 如核主成份分析[12]等.

本文主要介绍上述与支持向量机二次规划形式密切相关的 2) 和 3) 两个研究方向的新进展.

2 支持向量机分类和回归算法

依据统计学习理论中的结构风险最小化原则, Vapnik 等人首先提出了模式分类算法[1, 2] 以及回归分析算法[3, 4]. 假设训练样本集为

$$(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \cdots, (\boldsymbol{x}_l, y_l), \tag{1}$$

其中: $\boldsymbol{x}_i \in R^n$, $i = 1, 2, \cdots, l$ (R 为实数域); 对于两类的分类问题, $y_i \in \{+1, -1\}$; 对于回归问题, $y_i \in R$.

支持向量机分类算法的原始形式可归结为下列二次规划问题:

$$\min \frac{1}{2}(\boldsymbol{w}, \boldsymbol{w}) + C\sum_{i=1}^{l}\xi_i,$$

$$\text{s.t. } y_i((\boldsymbol{w}, \boldsymbol{x}_i) + b) - 1 + \xi_i \geq 0, \tag{2}$$

其中: (,) 为两向量之间的内积[13]; $\xi_i \geq 0$ 为松弛项, 表示错分样本的惩罚程度; C 为常数, 用于控制对错分样本惩罚的程度, 实现在错分样本数与模型复杂性之间的折衷; $\boldsymbol{w}$ 和 b 为判决函数 $f(\boldsymbol{x}) = (\boldsymbol{w}, \boldsymbol{x}) + b$ 中的权向量和阈值. 当无错分样本时, 最小化目标函数的第一项等价于最大化两类间的间隔, 可降低分类器的 VC 维, 实现结构风险最小化原则.

上述二次规划的对偶形式为

$$\max \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{l}\alpha_i\alpha_j y_i y_j(\boldsymbol{x}_i, \boldsymbol{x}_j),$$

$$\text{s.t. } \sum_{i=1}^{l}\alpha_i y_i = 0, 0 \leq \alpha_i \leq C, i = 1, 2, \cdots, l, \tag{3}$$

其中 α_i 为 Lagrange 乘子. 根据最优化理论中的 KKT 条件, 只有少量样本(判决函数值等于 ± 1 的样本和错分样本) 的 α_i 值不为零, Vapnik 等人称之为支持向量, 这便是支持向量机名称的由来.

由于对偶形式(3) 中只出现两向量间的内积运算, Vapnik 等人用满足 Mercer 条件的核函数 $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$ 来代替内积运算$(\boldsymbol{x}_i, \boldsymbol{x}_j)$, 实现线性算法的非线性化. 常用的核函数包括: 多项式核, 径向基核以及二层神经网[5~7]. 核形式的判别函数为

$$f(\boldsymbol{x}) = \sum_{i=1}^{l}\alpha_i y_i k(\boldsymbol{x}_i, \boldsymbol{x}) + b.$$

在 SVM 分类算法取得成功后, Vapnik 等人将分类算法的思想推广到回归分析, 提出了支持向量机回归算法, 采用 ϵ 线性不敏感损失函数、ϵ 平方不敏感损失函数和 Huber 损失函数代替经典回归分析中的平方损失函数, 从而使支持向量机回归算法仍可用二次规划来表达. 其中采用 ϵ 线性不敏感损失函数的支持向量机回归算法的对偶形式为

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

$$\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(\boldsymbol{x}_i, \boldsymbol{x}_j),$$

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

$$0 \leq \alpha_i, \alpha_i^* \leq C, \ i = 1, 2, \cdots, l. \tag{4}$$

这时, 位于 ϵ 管道边缘和外侧的样本便成为支持向量, ϵ 的大小可控制支持向量的数目.

与分类算法类似, 两向量间的内积运算可用核函数来代替, 从而实现算法的非线性化. 采用其他损失函数的支持向量机回归算法的形式参见文献[3～7]. 基于核的回归函数形式为

$$f(\boldsymbol{x}) = \sum_{i=1}^{l}(\alpha_i^* - \alpha_i)k(\boldsymbol{x}_i, \boldsymbol{x}) + b.$$

对于支持向量机分类和回归算法可参考两篇有影响的指导性文献[14, 15].

3 支持向量机的变形算法

支持向量机的成功促使众多研究人员投入到算法的设计和研究, 其中一个研究方向是对 SVM 的形式作某些改动, 简化求解过程或使算法参数具有一定的物理意义.

文献[2] 建议采用 ξ_i^k ($k \geq 0$) 形式的松弛项, 为便于计算, 支持向量机算法取 $k = 1$. Keerthi 等[16] 将 SVM 分类算法的松弛项改成平方型, 即 ξ_i^2, 此时其对偶形式的目标函数为

$$\max \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{l}\alpha_i\alpha_j\left(y_i y_j(\boldsymbol{x}_i, \boldsymbol{x}_j) + \frac{\delta_{ij}}{C}\right), \tag{5}$$

这使得对偶形式(5) 中的矩阵为正定阵(一般核矩阵至少是半正定的). 然而对于可分问题, 它的解只有在 $C \to \infty$ 时才是最大间隔分类面[17]. 文献[9] 将不等式约束条件改成等式约束, 提出了 SVM 的最小二乘形式(LS-SVM), 其对偶问题仅为一线性方程组, 即

$$\begin{bmatrix} \Omega + \boldsymbol{I}/C & \boldsymbol{y} \\ \boldsymbol{y}^{\mathrm{T}} & 0 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \boldsymbol{u} \\ 0 \end{bmatrix}. \tag{6}$$

其中: $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_l]^{\mathrm{T}}$, $\boldsymbol{y} = [y_1, y_2, \cdots, y_l]^{\mathrm{T}}$, $\boldsymbol{u} =$

$[1, 1, \cdots, 1]^T$, 矩阵 Ω的元素为 $\Omega_{ij} = y_i y_j k(\boldsymbol{x}_i, \boldsymbol{x}_j)$. 每个样本的 α 值与误差 ξ_i 成正比, 没有支持向量. 为获得稀疏表达, 文献[18] 提出训练后去掉一些小 α 值再训练的思路, 使多数样本的 α 值为零.

为避开求解二次规划问题, 文献[7] 用线性规划来表达SVM 回归算法的基本思想, 其形式为

$$
\begin{aligned}
&\min \left(\sum_{i=1}^{l}\alpha_i + \sum_{i=1}^{l}\alpha_i^*\right) + C\left(\sum_{i=1}^{l}\xi_i + \sum_{i=1}^{l}\xi_i^*\right),\\
&\text{s.t.}\\
&y_i - \sum_{j=1}^{l}(\alpha_j^* - \alpha_j)k(\boldsymbol{x}_i, \boldsymbol{x}_j) - b \le \epsilon - \xi_i^*,\\
&\sum_{j=1}^{l}(\alpha_j^* - \alpha_j)k(\boldsymbol{x}_i, \boldsymbol{x}_j) + b - y_i \le \epsilon - \xi_i,\\
&\alpha_i, \alpha_i^*, \xi_i, \xi_i^* \ge 0, i = 1, 2, \cdots, l.
\end{aligned}
\tag{7}
$$

这是用核形式回归函数直接构成的线性规划问题, 可采用熟悉的单纯形方法来求解.

由于支持向量机中的 C 和 ϵ 等参数缺乏一定的物理含义, Scholkopf 等人[10] 用 ν 代替原来的这些参数, 提出了 ν-SVM 分类和回归算法. 其中分类算法的对偶形式为

$$
\begin{aligned}
&\max -\frac{1}{2}\sum_{i,j=1}^{l}\alpha_i\alpha_j y_i y_j k(\boldsymbol{x}_i, \boldsymbol{x}_j),\\
&\text{s.t} \sum_{i=1}^{l}\alpha_i y_i = 0,\ 0 \le \alpha_i \le \frac{1}{l},\\
&\sum_{i=1}^{l}\alpha_i \ge \nu,\ i = 1, 2, \cdots, l.
\end{aligned}
\tag{8}
$$

其中 ν 的物理含义是: 错分样本数占总样本数的上界和支持向量数占总样本数的下界. 这种改造是非常合理的, 因为原始 SVM 中支持向量由边界上的样本和错分样本组成, 支持向量数一定大于错分样本数, ν 正好表示二者之间的某个中间值.

原始SVM 的解主要取决于两类分界面附近的样本点, 实际数据不可避免地带有噪声或存在野值. 文献[19] 提出了中心SVM, 以减小随机噪声的影响; [20] 则提出了中值 SVM, 以减小野值点的影响. 从有限的实验结果看, 它们都能有效地提高算法的鲁棒性.

4　广义 SVM 及其特例

从最优化理论和方法的角度出发, Mangasarian 等人提出了广义SVM 的框架(GSVM)[11], 并从中派生出若干特例. 在广义 SVM 框架中, 直接以 α, β 和核矩阵 $\boldsymbol{K}$(其元素为 $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$) 构造出一个不等式约束的非线性优化问题, 即

$$
\begin{aligned}
&\min C\boldsymbol{u}^T\xi + f(\alpha),\\
&\text{s.t. } \boldsymbol{D}(\boldsymbol{KD}\alpha - \boldsymbol{u}\beta) \ge \boldsymbol{u} - \xi.
\end{aligned}
\tag{9}
$$

其中: $f(\alpha)$ 是一凸函数, 如某一范数或半范数; $\xi = [\xi_1, \xi_2, \cdots, \xi_l]^T$, $\boldsymbol{D} = \mathrm{diag}(y_1, y_2, \cdots, y_l)$. 他们证明了式(9) 的对偶形式与SVM 对偶形式(3) 之间的等价关系, 所以称为广义支持向量机. 但 GSVM 并不是直接求解式(9) 或对偶形式, 而是构造出若干特例:

1) 将广义 SVM 中目标函数的第 1 项改成 $C\xi^T\xi$, 第 2 项取为$(\alpha^T\alpha + \beta^2)/2$, 并将其转化为无约束问题, 用二次收敛的 Newton-Armijo 求解, 称为光滑 SVM[21];

2) 如果从光滑 SVM 问题的 Lagrange 函数出发, 直接构造一迭代算法求解, 这样的算法称为 Lagrange SVM[22];

3) 简化 SVM[23] 则利用分块思想, 从样本集中随机选取非常少的样本, 求解光滑 SVM 的无约束目标函数;

4) 近似SVM[24] 采用光滑SVM 的目标函数, 将不等式约束改为等式约束, 其解满足一线性方程组, 并用递推公式求解;

5) 增量SVM[25] 求解线性近似SVM 问题, 通过去除旧样本加入新样本来修正线性分类器, 同时完成矩阵更新.

广义 SVM 算法的特殊形式都采用迭代方法求解, 可处理不同规模样本集的分类问题.

5　分位数估计与新奇检测算法

文献[26] 提出了形式上类似于 SVM 算法的新奇检测和高维分位数估计方法, 本质上是一种无监督模式分类算法. 其目的是找到一个非线性的判别函数, 在包含多数样本的小区域内取值为 +1, 而在区域外取值为 -1. 算法实现分为两步: 1) 利用核函数将样本数据映射到一特征空间; 2) 使原点和训练样本之间的间隔最大. 算法的对偶形式为

$$
\begin{aligned}
&\max \frac{1}{2}\sum_{i,j=1}^{l}\alpha_i\alpha_j k(\boldsymbol{x}_i, \boldsymbol{x}_j),\\
&\text{s.t. } \sum_{i=1}^{l}\alpha_i = 1, 0 \le \alpha_i \le \frac{1}{\nu l}, i = 1, 2, \cdots, l,
\end{aligned}
\tag{10}
$$

其中 $\nu \in (0, 1]$ 具有明确的物理含义: 当 $\rho \ne 0$ 时, 它既是野值样本数占总样本数的上界, 也是支持向量数占总样本数的下界. 最终的判别函数为

$$
f(\boldsymbol{x}) = \mathrm{sign}\left(\sum_{i=1}^{l}\alpha_i k(\boldsymbol{x}, \boldsymbol{x}_i) + \rho\right),
$$

其中 ρ 的含义见文献[26].

6 讨 论

自 20 世纪 90 年代中期以来, 综合统计学习理论、支持向量机算法的优点和最优化技术, 设计实现新算法成为机器学习领域的一个研究热点, 并取得了丰硕的成果. 本文主要综述了从 SVM 算法发展而来的一些算法, 算法形式可分为 4 类: 1) 二次规划问题, 如 SVM 的一些变型算法和新奇检测算法; 2) 线性方程组, 如 LS-SVM; 3) 线性规划问题, 如线性 SVM; 4) 迭代求解的广义 SVM 特例.

在算法设计思路上, 多数算法的目标函数折衷考虑数据的拟合程度(或错分样本数目) 与模型的复杂程度, 通过正则化参数将二者线性组合起来, 可有效地控制算法的推广能力; 然后将优化问题转换为对偶形式, 使其只出现两样本向量的内积运算; 最后用核函数来代替内积, 实现算法的非线性化. 另外, 采用满足 Mercer 条件的核函数, 可保证核矩阵是半正定的, 从而确保解的唯一性. 对于不等式约束, 支持向量的出现是 KKT 条件的直接结果.

线性 SVM 和广义 SVM 则利用核形式的判决函数或回归函数直接构造算法, 可以体现支持向量机算法的主要思想(即控制模型复杂性和核思想), 但不会自动出现支持向量. 这是另一条应引起人们关注的算法研究思路.

大量仿真实验结果表明, 本文介绍的算法具有较好的推广能力. 但实际应用中将面临核函数类型及参数选取问题, 这是一个既有实际价值又有理论意义的研究课题. 因此, 利用问题的启发式知识(即背景知识) 来研究核参数的选择问题是值得重视的.

参考文献(References) :

[1] Boser B E, Guyon I M, Vapnik V N. A training algorithm for optimal margin classifiers [A]. *The 5th Annual ACM Workshop on COLT* [C]. Pittsburgh: ACM Press, 1992. 144-152.

[2] Cortes C, Vapnik V N. Support vector networks[J]. *Machine Learning*, 1995, 20(3) : 273-297.

[3] Drucker H, Burges C J C, Kaufman L, et al. Support vector regression machines[A]. *Advances in Neural Information Processing Systems*[C]. Cambridge: MIT Press, 1997. 155-161.

[4] Vapnik V N, Golowich S, Smola A. Support vector method for function approximation, regression estimation and signal processing [A]. *Advances in Neural Information Processing Systems* [C]. Cambridge: MIT Press, 1997. 281-287.

[5] Vapnik V N. *The Nature of Statistical Learning Theory*[M]. New York: Springer-Verlag, 1995.

[6] Vapnik V N. *Statistical Learning Theory* [M]. New York: Wiley, 1998.

[7] Vapnik V N. *The Nature of Statistical Learning Theory* [M]. 2nd edition. New York: Springer-Verlag, 1999.

[8] Platt J. Fast training of support vector machines using sequential minimal optimization [A]. *Advances in Kernel Methods — Support Vector Learning* [C]. Cambridge: MIT Press, 1999. 185-208.

[9] Suykens J A K, Vandewalle J. Least squares support vector machines[J]. *Neural Processing Letters*, 1999, 9(3): 293-300.

[10] Scholkopf B, Smola A J, Williamson R C, et al. New support vector algorithms[J]. *Neural Computation*, 2000, 12(5): 1207-1245.

[11] Mangasarian O L. Generalized support vector machines[A]. *Advances in Large Margin Classifiers* [C]. Cambridge: MIT Press, 2000. 135-146.

[12] Scholkopf B, Smola A, Muller K R. Nonlinear component analysis as a kernel eigenvalue problem[J]. *Neural Computation*, 1998, 10(5): 1299-1319.

[13] 张恭庆, 林源渠. 泛函分析讲义[M]. 北京: 北京大学出版社, 1987.

[14] Burges C J C. A tutorial on support vector machines for pattern recognition [J]. *Data Mining and Knowledge Discovery*, 1998, 2(2): 1-43.

[15] Smola J, Schölkopf B. A tutorial on support vector regression[R]. London: University of London, 1998.

[16] Keerthi S S, Shevade S K, Bhattacharyya C, et al. A fast iterative nearest point algorithm for support vector machine classifier design[R]. Bangalore: Indian Institute of Science, 1999.

[17] Lin C J. Foundations of support vector machines: A note from an optimization point view [J]. *Neural Computation*, 2001, 13(2): 307-317.

[18] Suykens J A K, Lukas L, Vandewalle J. Sparse least squares support vector machines classifiers[A]. *The 8th European Symposium on Artificial Neural Networks*[C]. Brugers, 2000. 37-42.

[19] Zhang X. Using class center vectors to build support vector machines [A]. *Neural Networks for Signal Processing*[C]. New York: IEEE Press, 1999. 3-11.

[20] Kou Z, Xu J, Zhang X, et al. An improved support vector machine using class-median vectors[A]. *Proc of 8th Int Conf on Neural Information Processing*[C]. Shanghai: Fudan University Press, 2001. 2: 883-887.

(下转第 495 页)

对 x_1 和 x_2 分别施加幅值为 0.000 1 和 0.01 的白噪声. 图 2 和图 3 分别显示了采用 FPE 和 PDC 控制方法得到的控制效果. 图中, 虚线为设定值, 实线为跟踪值.

通过对比图 2 和图 3 可以看出, 本文给出的基于性能评估器的设计方法具有更好的控制性能和鲁棒性.

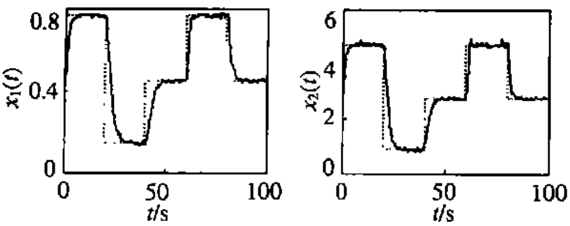

图 **2**　本文方法跟踪设定曲线

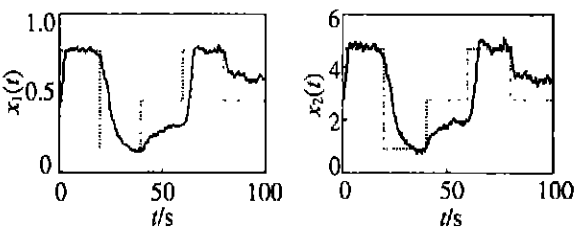

图 **3**　文献[**3**] 方法跟踪设定曲线

7　结　　语

通过对一类非线性时滞系统的分析, 提出了一种基于模糊性能评估器的鲁棒控制器设计方法. 该方法利用模糊性能评估器实现了 3 个功能:

1) 验证模糊模型的有效性;

2) 为控制器提供干扰抑制量信息;

3) 为模糊模型以及部分控制器参数调整提供一种无损调试的手段.

理论分析和仿真结果都证明了该方法具有良好的鲁棒性和实用性.

参考文献(References):

[1] Brierley S D, Chiasson J N, Lee E B, et al. On stability independent of delay for linear systems[J]. *IEEE Trans on Automatic Control*, 1982, 27(2): 252-254.

[2] Mahmoud M S, Al-Muthairi N F. Design of robust controller for time-delay systems[J]. *IEEE Trans on Automatic Control*, 1994, 39(8): 995-999.

[3] Cao Y Y, Frank P M. Analysis and synthesis of nonlinear time-delay systems via fuzzy control approach [J]. *IEEE Trans on Fuzzy Systems*, 2000, 8(2): 200-211.

[4] 张化光, 黎明. 基于 H_∞ 观测器原理的模糊自适应控制器设计[J]. 自动化学报, 2002, 28(1): 27-33.
(Zhang H G, Li M. Adaptive fuzzy controller design based on the principle of H_∞ observer [J]. *Acta Automatica Sinica*, 2002, 28(1): 27-33.)

[5] Li M, Zhang H G, He X Q. Adaptive fuzzy controller design based on H_∞ performance evaluator[A]. *Proc of the* 2002 *Int Conf on Control and Automation* [C]. Xiamen, 2002. 55-59.

[6] Lehman B, Bentsman J, Lunel S V, et al. Vibrational control of nonlinear time lag systems with bounded delay: Averaging theory, stability, and transient behavior [J]. *IEEE Trans on Automatic Control*, 1994, 39(7): 898-912.

[7] Xu L D, Cheng C W, Tang B Y. A linear matrix inequality approach for robust control of systems with delayed states[J]. *European J of Operational Research*, 2000, 124(2): 332-341.

(上接第 484 页)

[21] Lee Y J, Mangasarian O L. SSVM: A smooth support vector machines [R]. Wisconsin: University of Wisconsin, 1999.

[22] Mangasarian O L, Musicant D R. Lagrangian support vector machines [R]. Wisconsin: University of Wisconsin, 2000.

[23] Lee Y J, Mangasarian O L. RSVM: Reduced support vector machines [R]. Wisconsin: University of Wisconsin, 2000.

[24] Fung G, Mangasarian O L. Proximal support vector machine classifiers [R]. Wisconsin: University of Wisconsin, 2001.

[25] Fung G, Mangasarian O L. Increment support vector machine classification [R]. Wisconsin: University of Wisconsin, 2001.

[26] Scholkopf B, Platt J, Shawe-Taylor J, et al. Estimating the support of high dimensional distribution [R]. Redmond: Microsoft Research, 1999.