Artificial Intelligence and Robotics Research 人工智能与机器人研究, 2014, 3, 19-24
Published Online May 2014 in Hans. http://www.hanspub.org/journal/airr
http://dx.doi.org/10.12677/airr.2014.32004

Hans汉斯

Parameters Selection of SVM Based on Extended APSO Algorithm

Jingnan Li, Kaichun Ren, Jialing Yu, Fuguang Chen, Zhaoming Wu

Chongqing Communication Institute, Chongqing
Email: 260701887@qq.com

Received: Mar. 14th, 2014; revised: Apr. 10th, 2014; accepted: Apr. 22nd, 2014

Abstract

Support Vector Machine (SVM), a new mathematic modeling tool, has been widely used in many industry applications. The good generalization ability and estimation accuracy are impacted by parameters selection of SVM. Particle Swarm Optimization is improved by using active target. The active target particle swarm optimization was proposed to search the optimal combination of SVM parameters. Simulations show that active target particle swarm optimization is an effective way to search the SVM parameters and has good performance in classification.

Keywords

Support Vector Machines, Active Target Particle Swarm Optimization, Parameter Selection

基于活跃目标点粒子群算法的SVM参数选取

李景南，任开春，余佳玲，陈福光，吴钊铭

重庆通信学院，重庆
Email: 260701887@qq.com

收稿日期：2014年3月14日；修回日期：2014年4月10日；录用日期：2014年4月22日

摘　要

支持向量机是最近才兴起的一种分类工具，它广泛用于控制领域，但是其预测精度受到了其参数选取的

影响。使用活跃目标点改进粒子群优化算法,利用活跃目标点粒子群算法搜索支持向量机的最优参数组合。对比仿真实验表明：活跃目标点粒子群算法可以正确支持向量机的参数，能够进行较为准确的分类。

关键词

支持向量机，活跃目标点粒子群算法，参数选取

1. 引言

支持向量机(Support Vector Machines, SVMs)是二十世纪九十年代中期在统计学习理论基础上发展起来的一种新型机器学习方法[1]。传统统计学是样本数目趋于无穷大时的渐近理论，但在实际问题中，样本数往往是有限的，因此一些理论上很优秀的学习方法在实际中的表现却可能不尽人意[2]。统计学习理论是研究小样本统计估计和预测的理论，Vapnik 等人从二十世纪六、七十年代开始致力于此方面研究，统计学习理论的主要内容包括以下四个方面[3]：一是经验风险最小化准则(Empirieal Risk Minimization, ERM)下统计学习一致性条件；二是统计学习方法推广性的界；三是在推广性的界的基础上建立的结构风险最小化原则；四是实现这些准则的支持向量机方法。

支持向量机采用结构风险最小化准则(Structural Risk Minimization, SRM)训练学习机器,其主要优点有[4]：将学习问题归结为一个凸二次规划问题，从理论上说，得到的将是全局最优解，解决了在神经网络方法中无法避免的局部极值问题；通过非线性变换将数据映射到高维特征空间，使数据在高维空间中可以用线性判别函数分类；巧妙地解决了维数问题，算法复杂度与样本维数无关；具有简洁的数学形式和直观的几何解释，人为设定的参数少，便于理解和使用。支持向量机建立在严格的理论基础之上,较好地解决了非线性、高维数、局部极小点等问题，成为继神经网络研究之后机器学习领域新的研究热点。

针对 SVM 这种算法选取参数不完善的地方，选取参数的算法需具备较强的普遍性，收敛速度快，计算量小和全局搜索能力强的要求，本文利用张英男[5]等人提出的活跃目标点粒子群算法对 SVM 参数进行选取。通过仿真实验可以看出改进的粒子群算法能够选出较为有效的算法。

2. SVM 基本原理

支持向量机最初是用来处理两类数据的分类问题，它试图寻找一个能够分离两类点的一个超平面，并让这些点和超平面的距离最大化。如图 1 所示，三角形和圆圈分别表示两类训练样本，把它们投射到高维特征空间，寻找一个超平面 H，能够把这两类正确的分离开，并使两类距离超平面最近的点与超平面的距离最大，即 H_1 和 H_2 与 H 的距离最大化，H_1 与 H_2 之间的距离即为分类间隔。H 所表示的就是最优超平面。

SVM 的主要思想可以概括为两点：1) 它是针对线性可分情况进行分析，对于线性不可分的情况，通过使用非线性映射算法将低维输入空间线性不可分的样本转化为高维特征空间使其线性可分，从而使得高维特征空间采用线性算法对样本的非线性特征进行线性分析成为可能；2) 它基于结构风险最小化的原则，构造了一个目标函数，使得两类模式尽量正确的区分开。设训练样本为 $\left\{x_i, y_i \middle| x_i, y_i \subset R^n \times R, i=1,2,\cdots m\right\}$，将这个输入向量映射到高维特征空间，并在该特征空间构造最优分类面，将 SVM 回归表示为：

$$f(x)=\sum \omega \Phi(x)+b \tag{1}$$

其中： $\omega \in R^n, b \in R$ 。

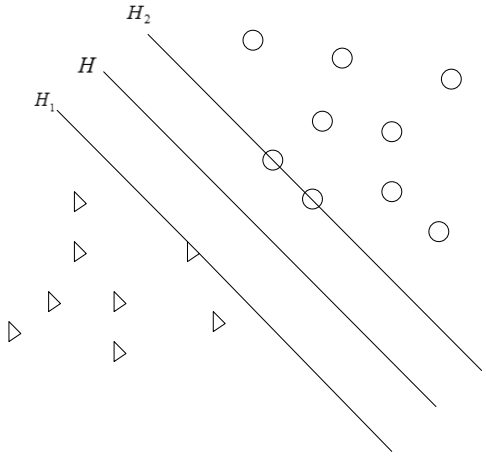

Figure 1. The optimal hyperplane
图 1. 最优超平面示意图

将其转化为解凸二次优化问题，即：

$$\min \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{m}\xi_i \tag{2}$$

满足约束条件：

$$y_i\left(w\cdot x_i + b\right) \geq 1-\xi_i \quad \xi_i \geq 0 \quad i=1,2,\cdots,m \tag{3}$$

为解决最优化问题，其对偶问题为：

$$\max W\left(\alpha\right) = \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{l}\alpha_i\alpha_j y_i y_j k\left(x_i, x_j\right) \tag{4}$$

$$s.t.\ \sum_{i=1}^{l}\alpha_i y_i = 0 \quad 0 \leq \alpha_i \leq C \quad i=1,2,\cdots,l \tag{5}$$

决策函数为：

$$f\left(x\right) = \operatorname{sgn}\left(\sum_{i=1}^{l} y_i\alpha_i k\left(x_i, x\right) + b\right) \tag{6}$$

$k\left(x_i, x\right)$ 为满足 Mercer 条件的核函数。

本文采用 RBF 核函数，$K\left(x_i, x\right) = \exp\left(\dfrac{-\left(x-x_i\right)^2}{\sigma^2}\right)$，$nv$ 为支持向量个数。

3. 活跃目标点粒子群算法

Kennedy 和 Eberhart 于 1995 年基于生物界中鸟群和鱼群等生物群体的捕食行为，提出的粒子群优化(PSO)算法。在自然界中，鸟搜索食物是通过群体合作实现的，即搜索离食物最近的鸟的周围的区域。我们称这些鸟为“粒子”，粒子具有各自的速度和位置，算法中还有一个适应度函数，在每一次迭代中寻找这一代的最优解和历史最优解以获得下一次迭代所需要的速度，从而来对“粒子”的状态进行改变，直到达到全局最优解或者满足终止条件。

3.1. 标准 POS 算法

Kelmedy 和 Eberhart 的标准全局版 PSO 算法可概述如下[6]：假设在一个 d 维的目标搜索空间中，由

M 个粒子组成一个种群，其中第 i 个粒子位置表示为一个 d 维的向量 $X_i=(x_{i1},x_{i2},\cdots,x_{id})$，速度表示为 d 维向量 $V_i=(v_{i1},v_{i2},\cdots,v_{id})$。$pbest_{id}$ 为第 i 个粒子的历史最优位置，$gbest_d$ 为这一代粒子的全局最优位置，则第 i 个粒子速度和位置的更新公式为：

$$v_i^{k+1}=v_i^k+c_1rand_1^k\left(pbest_i^k-x_i^k\right)+c_2rand_2^k\left(gbest^k-x_i^k\right) \tag{7}$$

$$x_i^{k+1}=x_i^k+v_i^{k+1} \tag{8}$$

其中，v_i^k 是粒子 i 在第 k 次迭代中的速度向量，$v_i^k\in(-v_{\max},v_{\max})$，$v_{\max}$ 用于限制粒子速度，使之在一定的范围内飞行，不会应为速度过小导致陷入局部最优或者速度过大飞离最优解区域；c_1，c_2 是学习因子，c_1 为个体历史最优粒子对第 i 个粒子飞行速度的影响，c_2 为全局最优粒子对第 i 个粒子飞行速度的影响，$rand_1$，$rand_2$ 是[0,1]之间的随机数。

SVM 参数取值对于最终能否进行正确的分类有着重要的关系。其参数包括惩罚因子 c 和核参数 σ。因此我选用了改进的 PSO 算法，即 APSO，来对 c 和 σ 进行选取。

3.2. APSO 算法

为避免标准 PSO 早熟收敛陷入局部最优解，活跃目标点粒子群算法(APSO)在标准 PSO 的速度更新中引入了第三个目标点 p_i^k，区别于种群当代最优和粒子历史最优的位置，使粒子能够跳出局部最优解，所引入的第三个点 p_i^k 即为活跃目标点[7]。该点是利用当前粒子 i 的邻域中按一定规则随机搜索一点 $p_i^{k'}$，与 $pbest$ 和 $gbest$ 构成复合型，进而求得一个该点 p_i^k。APSO 可以较好地保持种群多样性，跳出局部最优能力较强，并且进入搜索后期也可以保持较快的收敛速度，但缺点是每次迭代增加了一次随机点的计算。

APSO 的速度更新公式为：

$$v_i^{k+1}=v_i^k+c_1rand_1^k\left(pbest_i^k-x_i^k\right)+c_2rand_2^k\left(gbest^k-x_i^k\right)+c_3rand_3^k\left(p_i^k-x_i^k\right) \tag{9}$$

其中，$c_1\sim c_3$ 为学习因子；$rand_1\sim rand_3$ 为[0,1]之间的随机数；p_i^k 为活跃目标点。下面给出求解活跃目标点 p_i^k 的步骤：

1) 在 x_i^k 邻域范围内搜索一点作为试用活跃目标点 $p_i^{k'}$：

$$p_i^{k'}=x_i^k+rand_4^k x_{\max} \tag{10}$$

其中，$rand_4^k$ 是 $[-0.1,0.1]$ 之间的随机数；$x_{\max}$ 为最大可行域范围。

2) 在粒子历史最优点 $pbest_i^k$，种群历史最优点 $gbest^k$ 和试用目标点 $p_i^{k'}$ 三点中找出最坏点

$$x^H:f\left(x^H\right)=\max\left\{f\left(p_i^{k'}\right),f\left(pbest_i^k\right),f\left(gbest^k\right)\right\} \tag{11}$$

设最坏点为 $p_i^{k'}$，进行最坏点映射：

$$x^R=\frac{pbest_i^k+gbest^k}{2}+\alpha\left(\frac{pbest_i^k+gbest^k}{2}-x^H\right) \tag{12}$$

其中，α 为映射系数，$\alpha=1.3\sim0.5$，呈递减趋势。

最后，将该映射点 x^R 作为活跃目标点 p_i^k 的位置。即 $p_i^k=x^R$。

4. 仿真实验——SVM 参数选取

由文献[8]给出的某柴油机燃油喷射系统的柱塞在正常和磨损状态下经过多次采用并经 AR 时序建模而得到的压力波形特征参数，表 1 为柱塞不同磨损状况时的 9 个标准样本。

首先随机生成粒子的初始位置和初始速度，并以 SVM 计算的准确率为其适应度，然后寻找其个体历

Table 1. Parameters of pressure wave of plungers that have abrasion and not have abrasion
表 1. 柱塞磨损与不磨损时的压力波特征参数

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	n_j
正常	−1.078	0.015	0.082	−0.039	−0.059	−0.117	−0.014	0.088	0.112	0.021	6.215
正常	−1.247	0.168	0.195	−0.064	−0.125	−0.121	0.035	0.062	0.137	−0.026	5.243
故障	−1.332	0.385	0.104	−0.101	−0.143	−0.129	0.101	0.006	0.072	0.051	6.101
故障	−0.111	−0.069	0.015	0.148	−0.05	0.023	−0.026	−0.033	0.092	0.018	5.001
故障	−1.266	0.110	0.206	0.083	−0.118	−0.046	0.014	−0.128	−0.032	0.188	6.284
故障	−1.237	0.106	0.184	0.022	−0.101	−0.048	−0.087	0.071	0.074	0.029	5.301
正常	−1.105	−0.076	0.274	0.058	−0.143	−0.178	0.012	0.011	0.165	0.006	3.329
正常	−1.136	−0.094	0.279	0.035	−0.106	−0.152	0.038	0.010	0.187	−0.007	6.469
故障	−1.114	−0.170	0.206	0.042	−0.051	−0.08	0.022	0.128	0.197	−0.171	2.968

Table 2. c, σ and its accuracy under different classification mode
表 2. 不同分类模式下的 c、σ 和其准确率

	bestc	*best*σ	*Vaccuarcy*
GA-SVM	0.1661	333.0730	55.5556
PSO-SVM	9.8551	11.5547	66.6667
APSO-SVM	1.0376	190.0354	88.8889

史最优值和群体最优值，之后进行迭代，更新个体历史最优值和群体最优值，直到满足终止条件。最优位置即为所需要的 c，σ，经由 SVM 计算出其准确率。

在采用 APSO 来优化 SVM 时，其分类的准确度要远远高于 GA 和 PSO 优化 SVM 后分类的准确度，表 2 为不同寻优方式下 SVM 分类的准确率和对应的 c 和 σ。

结果表明，基于 APSO 的 SVM 参数选取在分类准确率上要优于其他两种优化方法，在非线性可分情况下有较高的准确性。

5. 总结

SVM 被广泛的用来进行非线性分类，模式识别与故障诊断。有神经网络不具有的优越性。而对于 SVM 需要设定一定的参数才能进行准确的分类，其参数是根据不同的分类情况进行选定的。本文利用 APSO 来优化 SVM 参数的选取，通过仿真实验表明，经过 APSO 优化的 SVM 拥有更好的分类效果。该方法的优点是分类准确率高，缺点是计算量高，可以通过 RS 理论对条件进行化简，使得计算量减少，这是以后下一步准备做的。

参考文献 (References)

[1] Cortes, C. and Vapnik, V. (1995) Support-vector networks. *Machine Learning*, **20**, 273-297.

[2] 张学工 (2000) 关于统计学习理论与支持向量机. *自动化学报*, **1**, 32-42.

[3] Vapnik, V. (2000) 统计学习理论的本质. 张学工, 译. 清华大学出版社, 北京.

[4] Burges, C.J.C. (1998) A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, **2**, 121-167.

[5] Zhang, Y.N., Hu, Q.N. and Teng, H.F. (2008) Active target particle swam optimization. *Concurrency and Computation*: *practice and Experience*, **20**, 29-40.

[6] Kennedy, J. and Eberhart, R.C. (1995) Partiele swarm optimization. *IEEE International Conference on Neural Networks*, 27 November-1 December 1995, 1942-1948.

[7] 张英男 (2008) 改进的粒子群优化算法(APSO 和 DPSO)研究. 大连理工大学, 大连.

[8] 曹龙汉, 曹长修 (2002) 基于粗糙集理论的柴油机神经网络故障诊断研究. *内燃机学报*, **4**, 357-361.