

泛化的统一切比雪夫多项式核函数

赵金伟¹, 冯博琴¹, 闫桂荣²

(1. 西安交通大学计算机科学与技术系, 710049, 西安;

2. 西安交通大学机械结构强度与振动国家重点实验室, 710049, 西安)

摘要: 针对分布稀疏、特征不明显的小样本数据回归中的属性冗余问题, 基于统一切比雪夫多项式, 提出了一种向量形式输入的可变正交多项式核函数——泛化的统一切比雪夫多项式核函数. 新的核函数通过利用统一切比雪夫多项式的正交性和可变性扩大了函数的搜索空间, 通过调整多项式阶数有效地控制了特征空间维数, 从而解决了稀疏数据回归中的属性冗余问题. 另外, 利用 Mercer 定理证明了该核函数的有效性. 在多组标准数据集和实际工程数据集上对核函数的性能进行了实验对比, 结果证明新的核函数预测精度较高, 泛化能力较好, 在大多数标准数据集上的性能优于其他切比雪夫多项式核函数.

关键词: 统一切比雪夫多项式; 统一切比雪夫多项式核函数; 支持向量机; 回归问题

中图分类号: TP301.6 **文献标志码:** A **文章编号:** 0253-987X(2012)08-0043-06

Generalized Uniform Chebyshev Polynomial Kernel

ZHAO Jinwei¹, FENG Boqin¹, YAN Guirong²

(1. Department of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China;

2. State Key Laboratory of Strength and Vibration of Mechanical Structures, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Based on a group of unified Chebyshev polynomials (UCP), a new kernel for vector inputs, named generalized uniform Chebyshev polynomial kernel (GUCK), is proposed to solve the problem of redundant attributes in the regression analysis on small-scale data sets. The proposal kernel can extend the search space of optimal kernel function by the orthogonality and adaptivity of UCP and control the dimension of the feature space by adjusting the polynomial coefficient of UCP. The problem of redundant attributes is settled by this method. Moreover, the proposal kernel, GUCK, has been proved that it is a valid support vector machine (SVM) kernel. The simulation results and application results show that GUCK can lead to better generalization performance in comparison with other common kernels, and is well applicable to the practical dataset. The GUCK has an advantage over other Chebyshev kernels on the majority of benchmark data sets

Keywords: uniform Chebyshev polynomials; uniform Chebyshev kernel; support vector machine; regression problem

随着科学技术的发展, 利用支持向量机学习算法来解决小样本分类、回归和概率密度估计问题越来越广泛^[1]. 在工程中常常会遇到空间稀疏的数据集, 这类数据具有数量少、分布特征不明显、实测时

噪声大等特点, 会影响支持向量机的学习效果. 解决这种问题通常有两种方案: 一种是学习算法, 如最小二乘支持向量机(LSSVM)^[2]、近似支持向量机(PS-

收稿日期: 2011-11-22. 作者简介: 赵金伟(1974—), 男, 博士生; 冯博琴(通信作者), 男, 教授, 博士生导师. 基金项目: 国家自然科学基金资助项目(10776026).

网络出版时间: 2012-06-15

网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20120612.1044.007.html>

VM)^[3]、模糊支持向量机(FSVM)^[4]、概率支持向量机(PSVM)^[5]等;另一种是利用不变核函数融合来寻求解决方案。Rostamizadeh 等人^[6-13]将各种不同的核融合在一起,利用自适应优化策略来控制参数,但这种方法需要设计高效的寻参机制和核选择机制。

从本质上来看,解决小样本回归学习问题的关键是避免训练样本从原始空间映射到高维空间时带来的属性冗余问题^[14]。理论上已证明,切比雪夫多项式能很好地逼近任何一条曲线,可以有效解决高维空间中的属性冗余问题^[15-18]。然而,只使用第一类切比雪夫多项式来构造核函数,会因为多项式的有限逼近能力而导致数据学习效果出现时好时不好的情况^[14]。

针对上述问题,本文研究了可变系数的正交多项式——统一一切比雪夫多项式^[14],并基于此提出了一种新的向量形式输入的统一一切比雪夫多项式核函数——泛化的统一一切比雪夫多项式核函数。

1 稀疏数据回归问题及相关工作

用于机器学习的数据集一般分为相关性较强和相关性较弱的两类数据集。相关性较强的数据集的数据之间存在着自相关性和多元相关性,所以回归学习和分类学习的效果比较好。相关性较弱的数据集包括空间稀疏的数据集和具有大量缺失数据的数据集。前者数据的间隔较大,相邻数据之间相关性极弱,数据空间往往不连通;后者虽然数据量较大,但是由于操作的原因或现实条件的限制,每次记录不能完全涵盖全部属性,从而造成数据集中存在大量的空洞,如果将这些有“空洞”的记录剔除掉,完全可以将该数据集当成空间稀疏的数据集。

空间稀疏数据的回归学习问题本质上是挖掘数据集中有用信息的过程,这些有用的信息包括有用的特征和相关的属性。属性冗余是指,数据集中包含了多余、无用的特征或属性。对于分类问题中的属性冗余,已有较为详细的分析和描述^[19]。然而,在支持向量回归机中,由于利用核函数将原始空间的样本点映射到了高维特征空间,虽然改进了样本的可分性,但是高维特征空间维数的多少是未知的、无法控制的,如何解决高维特征空间中的属性冗余问题至今是一个难点。

本文通过研究认为,核函数是回归学习中控制特征空间维数和解决属性冗余问题的一种途径。根

据样本分布特征选择合适的核函数,可以构造具有最小属性冗余的高维特征空间,就能更有效地提高样本的可分性,进而提高回归机的预测精度及泛化性能。解决这一问题可从两个方面入手:①寻找一个具有广泛适应能力的核函数;②利用该核函数构造具有最小属性冗余的高维特征空间。于是,文献[14]提出一种可变、正交的切比雪夫多项式——统一一切比雪夫多项式(UCP),定义如下。

定义 1 若函数

$$\frac{1-xt}{1-axt} \frac{1}{1-2xt+t^2}$$

中 $|x| \leq 1, |t| \leq 1, a \geq 0$,该函数称为统一一切比雪夫多项式的母函数。

定义 2 函数

$$\frac{1-xt}{1-axt} \frac{1}{1-2xt+t^2} = \sum_{n=0}^{\infty} P_n(x) t^n \quad (1)$$

式中: n 为多项式的阶数; a 为多项式系数。如果式(1)同时满足条件 $x \in [-1, 1], t \in [-1, 1], n=0, 1, 2, 3, \dots$,则称 $P_n(x)$ 为统一一切比雪夫多项式。

两类传统的切比雪夫多项式仅是 UCP 的特例:当 $a=0$ 时, $P_n(x)$ 为第一类切比雪夫多项式;当 $a=1$ 时, $P_n(x)$ 为第二类切比雪夫多项式;当 $a \neq 0, 1$ 时,可以得到更多的多项式,它们都是关于权函数 $(1-x^2)^{1/2}$ 正交的。 $P_n(x)$ 的递推式为

$$P_n(x) = (a+2)xP_{n-1}(x) - (2ax^2 + 1)P_{n-2}(x) + axP_{n-3}(x) \quad (2)$$

基于式(2),文献[14]又提出了一种标准统一一切比雪夫多项式核函数(UCK)。

2 泛化的统一一切比雪夫多项式核函数

由于 UCP 的可变性,所以扩大 UCK 的选择空间可以提高 UCP 的适应能力,另外 UCP 具有正交性,所以基于它映射得到的高维特征空间的正交基与其阶数是密切相关的,利用阶数可以有效解决高维特征空间的属性冗余问题。然而,在现实工程中,输入的是向量形式的数据,而在 UCK 中并未考虑向量数据。在空间稀疏数据的回归学习中,数据集本身可能存在着缺失数据或非常小的数据,这时原空间中很多比较靠近的样本点会由于这种情况而使得样本点在核矩阵中的对应元素比较小,从而映射到高维特征空间后元素彼此之间的距离比较大,这会带来样本空间失真的现象。为了避免此问题,本文提出一种向量形式的核函数,称之为泛化的统一一切比

雪夫多项式核函数(GUCK),表达式为

$$K(\mathbf{x}, \mathbf{z}) = (M - \langle \mathbf{x}, \mathbf{z} \rangle)^{-1/2} \sum_{n=0}^d \mathbf{P}_n(\mathbf{x}) \mathbf{P}_n^T(\mathbf{z}) \quad (3)$$

式中: M 为样本维度; d 为 GUCK 的阶数.

根据 Mercer 定理,当一个函数满足 Mercer 条件时,它所构建的核矩阵是半正定的,也就是说,必须满足下述定理,该函数为有效的核函数.

定理 1 GUCK($K(\mathbf{x}, \mathbf{z}), \mathbf{x}, \mathbf{z} \in \mathbf{X}$)的有效条件是

$$\int_{\mathbf{X} \times \mathbf{X}} K(\mathbf{x}, \mathbf{z}) f(\mathbf{x}) f(\mathbf{z}) d\mathbf{x} d\mathbf{z} = \int_{\mathbf{X} \times \mathbf{X}} (M - \mathbf{x} \cdot \mathbf{z})^{-1/2} \sum_{n=0}^d \mathbf{P}_n(\mathbf{x}) \mathbf{P}_n^T(\mathbf{z}) \cdot f(\mathbf{x}) f(\mathbf{z}) d\mathbf{x} d\mathbf{z} \geq 0$$

证明 由构造原则^[13]可知,两个有效核函数的乘积仍是一个有效核,即函数 $(M - \mathbf{x} \cdot \mathbf{z})^{-1/2}$ 和

$\sum_{n=0}^d \mathbf{P}_n(\mathbf{x}) \mathbf{P}_n^T(\mathbf{z})$ 满足 Mercer 条件.

由于 $(M - \mathbf{x} \cdot \mathbf{z})^{-1/2}$ 的麦克劳林表达式中所有系数是非负的,所以 $(M - \mathbf{x} \cdot \mathbf{z})^{-1/2}$ 是一个有效的核函数^[15-16].

函数 $\sum_{n=0}^d \mathbf{P}_n(\mathbf{x}) \mathbf{P}_n^T(\mathbf{z})$ 也满足 Mercer 条件,因为 $\mathbf{x}, \mathbf{z} \in \mathbf{X}$, 所以

$$\begin{aligned} \int_{\mathbf{X} \times \mathbf{X}} \sum_{n=0}^d \mathbf{P}_n(\mathbf{x}) \mathbf{P}_n^T(\mathbf{z}) f(\mathbf{x}) f(\mathbf{z}) d\mathbf{x} d\mathbf{z} &= \\ \sum_{n=0}^d \int_{\mathbf{X} \times \mathbf{X}} \mathbf{P}_n(\mathbf{x}) \mathbf{P}_n^T(\mathbf{z}) f(\mathbf{x}) f(\mathbf{z}) d\mathbf{x} d\mathbf{z} &= \\ \sum_{n=0}^d \left\| \int_{\mathbf{X}} \mathbf{P}_n(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} \right\|^2 &\geq 0 \end{aligned}$$

可见, $\sum_{n=0}^d \mathbf{P}_n(\mathbf{x}) \mathbf{P}_n^T(\mathbf{z})$ 满足 Mercer 条件,也是一个有效的核函数.

根据核函数的构造原则^[13], GUCK 必定是一个有效的核函数.

3 仿真实验

为了客观地评价 GUCK 对回归模型泛化性能的影响,在实验中将 GUCK 与 UCK 及其他 8 种核函数用于标准支持向量机算法 ϵ -SVM,进行了泛化性能对比.这 8 种核函数分别是:正交切比雪夫多项式核 (OCK)^[15]、泛化切比雪夫多项式核 (GCK)^[16-17]、改进的切比雪夫多项式核 (MCK)^[18]、普通多项式核 (GPK)^[1]、径向基核 (RBF)^[1]、sig-

moid 核 (SK)^[1]、线性核 (LK)^[1] 和小波核 (WLK)^[20].泛化性能评价中使用了均方根误差 (MSE) 和相关性系数 (SCC):均方根误差 $e_{\text{MSE}} = (1/l) \sum_{i=1}^l |f(\mathbf{x}_i) - y_i|^2$,其中 l 表示样本数, $f(\mathbf{x}_i)$ 表示预测结果, y_i 表示真实结果.均方根误差是回归模型的泛化能力的表征量,其值越小,表示预测结果与真实结果越接近,回归模型泛化能力越强.相关性系数

$$\begin{aligned} C_{\text{SCC}} &= (l \sum_{i=1}^l (f(\mathbf{x}_i) y_i) - \sum_{i=1}^l f(\mathbf{x}_i) \sum_{i=1}^l y_i)^2 / \\ &((l \sum_{i=1}^l (f(\mathbf{x}_i) f(\mathbf{x}_i)) - \sum_{i=1}^l f(\mathbf{x}_i) \sum_{i=1}^l f(\mathbf{x}_i)) \cdot \\ &(l \sum_{i=1}^l y_i y_i - \sum_{i=1}^l y_i \sum_{i=1}^l y_i)) \end{aligned}$$

表征预测结果与真实结果的相关程度,也是泛化能力的表征量,其值越大,表明预测结果与真实结果的相关程度越大,回归模型泛化能力越强.本实验中学习算法为标准支持向量机算法 ϵ -SVM,同时利用粒子群算法 (PSO) 和十折交叉验证进行了参数优化. **实验 1** 在多个标准测试数据集 UCI^[21] 上,针对不同的核函数,利用 ϵ -SVM 进行了 10 次十折交叉验证,最后分别求出平均的 MSE 和 SCC.实验前,对所有数据进行归一化处理,实验中标准数据集描述如表 1 所示,实验结果如表 2 所示.

表 1 实验 1 中各种标准数据集描述

数据集名称	属性值	原样本数	实验样本数
Ozone	12	126	40
Concrete	9	1 030	20
Servo	4	167	30
Concrete Slump Test	10	103	20
Auto-Mpg	8	398	40
MGH10	2	16	16
Forest_Fires	13	517	20
Parkinsons	16	5 876	20
Wine_Quality	11	1 599	30
YearPrediction MSD	90	515 345	30

表 2 实验 1 中不同核函数在各种标准数据集上的性能比较

核函数	Ozone		Concrete		Servo		Concrete Slump Test		Auto-Mpg	
	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$
RBF	19.523	19.141	5.384	42.759	2.990	92.639	15.062	46.295	3.269	93.429
LK	22.460	9.717	5.100	46.754	16.517	61.140	12.434	58.250	3.526	92.908
SK	23.880	4.003	NAN	NAN	22.521	35.824	23.057	23.114	4.496	90.823
WLK	<u>18.600</u>	20.737	5.635	41.408	3.356	91.842	16.877	41.370	3.384	93.259
GPK	21.044	24.351	6.305	40.273	35.652	56.764	17.750	40.044	2.600	95.000
OCK	21.087	23.908	5.209	46.59	3.157	92.495	23.057	23.114	2.257	95.423
GCK	28.486	11.056	7.173	30.41	2.846	<u>93.276</u>	14.778	46.896	2.826	94.372
MCK	18.500	<u>24.405</u>	5.806	38.813	<u>2.740</u>	93.204	<u>11.290</u>	<u>60.619</u>	2.677	94.561
UCK	18.764/0.900	31.289	4.323/0.200	52.144	2.91/0.586	92.948	6.07/0.030	79.165	3.08/0.500	93.938
GUCK	26.50/0.160	3.540	<u>4.57/0.300</u>	<u>51.770</u>	2.49/0.810	94.100	14.78/0.480	4.460	2.75/0.500	<u>95.000</u>

核函数	MGH10		Forest_Fires		Parkinsons		Wine_Quality		YearPrediction MSD	
	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$
RBF	0.000	100.000	<u>19.686</u>	<u>36.720</u>	44.749	1.044	21.884	<u>16.350</u>	<u>21.783</u>	0.040
LK	5.382	83.027	20.211	35.013	39.586	0.596	45.494	10.778	21.944	0.003
SK	7.656	76.411	29.516	16.213	86.109	2.923	<u>21.230</u>	15.071	NAN	NAN
WLK	0.626	98.074	20.439	35.038	43.142	0.050	23.763	13.989	21.786	0.030
GPK	0.292	99.190	20.184	35.951	39.586	0.596	51.619	6.048	22.296	0.520
OCK	0.000	90.000	25.662	29.007	44.900	5.289	22.265	2.887	NAN	NAN
GCK	0.000	100.000	22.348	31.539	40.059	<u>7.144</u>	NAN	NAN	21.815	0.923
MCK	0.490	98.576	26.919	21.160	39.094	40.237	24.659	7.760	23.071	<u>7.606</u>
UCK	0.000/0.400	100.000	18.750/1.040	51.230	<u>38.410/0.300</u>	0.072	22.010/0.600	8.532	21.850/0.400	0.019
GUCK	0.000/0.500	100.000	26.770/0.340	0.532	36.970/0.500	4.830	19.100/0.400	17.340	14.710/0.500	17.110

注:加粗数值表示最优值;下划线数值表示次优结果;斜线后的数值表示 UCK 和 GUCK 的多项式系数值。

从表 2 发现:对于 10 种标准数据集,当选择 UCK 和 GCK 时, ϵ -SVM 算法通常会得到比较小的 MSE 和比较大的 SCC;在 Ozone、Wine Quality、YearPrediction MSD 数据集上,UCK 的 MSE 和 SCC 并不理想;在 Ozone、Concrete Slump Test、Forest_Fires 数据集上,GUCK 的 MSE 和 SCC 也不理想. 尽管如此,这 2 个核函数的 MSE 和 SCC 也远远优于其他切比雪夫多项式核函数. 表 2 还给出了 UCK 和 GUCK 的多项式系数 a 值,可以看出,在很多情况下, a 不等于 0 和 1,证明了 UCP 核函数形式的多样性. 表 2 出现的 NAN 是由算法不收敛引起的.

实验 2 采用 holdout 方法分别将 5 个 UCI 数据集^[21]随机划分为训练集和验证集. 每组实验进行

10 次,最后求出最小预测 MSE 和相对应的 SCC,并进行了比较. 实验前,对所有数据进行归一化处理. 数据集详细描述如表 3 所示,实验结果如表 4 所示.

表 3 实验 2 中各种标准数据集描述

数据集名称	属性值	原样本数	训练样本数	测试样本数
Pyrim	27	30	20	10
Mg	6	57	27	30
Abalone	8	30	20	10
Space	6	30	20	10
Wine Quality	11	1 599	20	19
Mpg	7	392	20	10

表 4 实验 2 中不同核函数在各种标准数据集上的性能比较

核函数	Pyrim		Mg		Abalone		Space		Wine_Quality	
	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$	$e_{MSE}/\%$	$C_{SCC}/\%$
RBF	27.744	14.000	10.144	73.900	31.169	18.200	9.990	2.517	14.528	5.880
LK	30.825	42.719	10.176	74.130	32.230	13.816	9.550	7.118	9.552	28.254
SK	32.881	4.203	10.515	68.543	47.179	<u>23.440</u>	7.780	10.200	9.380	<u>28.514</u>
WLK	30.717	0.709	10.285	72.165	30.886	18.287	10.100	4.547	9.551	28.253
GPK	33.689	12.211	10.305	72.518	25.794	25.908	40.302	39.978	9.552	28.254
OCK	59.772	0.203	7.880	77.900	46.900	11.900	9.885	2.122	10.900	0.935
GCK	23.022	62.500	10.808	71.960	66.307	11.498	9.160	1.670	10.124	24.629
MCK	38.860	0.207	<u>4.234</u>	88.615	<u>9.678</u>	11.350	10.100	0.516	7.439	5.957
UCK	<u>17.110/0.010</u>	55.700	3.990/0.400	<u>82.000</u>	16.339/0.600	14.000	7.120/0.500	<u>15.000</u>	9.618/1.000	20.000
GUCK	15.248/0.900	<u>59.426</u>	10.361/0.500	72.168	5.251/0.010	14.000	<u>8.980/0.000</u>	0.627	<u>8.140/0.500</u>	40.758

注:加粗数值表示最优值;下划线数值表示次优结果;斜线后的数值表示 UCK 和 GUCK 的多项式系数值。

从表 4 发现,在 5 个标准数据集的回归学习中,UCK 和 GUCK 的 MSE 总是小于其他核函数,而 SCC 也比较大,表明 UCK 和 GUCK 在回归学习中能提高学习算法的泛化性能.通过实验发现,在高维数据的回归学习中,GUCK 的 MSE 小于 UCK,GUCK 的 SCC 大于 UCK,说明相对于 UCK,GUCK 更适合于高维数据集.

4 工程应用

以圆柱壳仿真模型(见图 1)为研究对象,利用不同的夹具分别模拟了由不同边界条件(加盖与不加盖 2 种情况)构成的动力学系统.首先,对这 2 个系统壳体的同一位置分别施加相同的冲击激励,见图 1 中的“+”处.当施加了 30 组不同脉宽激励时,由同一节点(如图 1 所示的 A 点)可以得到 30 组振动响应,每一组振动响应的采样数为 1 400,包括 2 种不同边界的振动响应数据.然后,随机取其中的 29 组数据作为训练样本,另 1 组数据作为测试样本.最后,对每一个采样点分别进行训练学习,经过预测最终得到 1 400 个采样点的预测结果,由此计算出相应的 MSE 和 SCC.学习算法仍然采用 ϵ -SVM,核函数与上节实验相同,预测结果如表 5 所示.

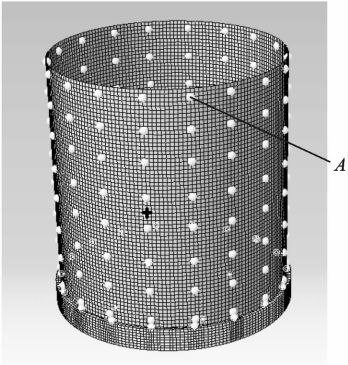


图 1 圆柱壳有限元模型

从表 5 发现,基于 UCK 和 GUCK 的回归模型的 MSE 最小且 SCC 最大,表明 UCK 和 GUCK 均适用于圆柱壳结构响应数据集的回归预测,同时显示预测结果与真实结果更为一致.

5 结 论

在解决稀疏数据回归问题时,UCK 和 GUCK 的 MSE 较小,SCC 较大,表明该系列核函数的泛化性能较好.这是因为,统一切比雪夫多项式是一类正交多项式,基于它映射得到的高维特征空间中的正

表 5 频域结构振动响应 PSD 谱的预测比较

核函数	RFB	WLK	SK	LK	GPK	OCK	GCK	MCK	UCK	GUCK
$e_{MSE}/\%$	11.520	15.790	2.140	1.740	3.560	2.480	9.380	9.000	<u>1.710</u>	1.700
$C_{SCC}/\%$	89.829	85.892	98.064	98.447	96.334	97.916	91.447	92.486	<u>98.427</u>	98.320

注:加粗数值表示最优值;下划线数值表示次优结果.

交基与多项式的阶数是密切相关的,所以通过调整阶数可以控制特征空间的维数,以便于解决属性冗余问题.由于统一比雪夫多项式依据不同的系数可以获得不同的数学表达式,所以在学习中根据训练数据的分布特点可以获得多种形式的核函数来参与学习,扩大核函数的选择空间,提高核函数的适应能力.虽然这2种核函数与组合核函数相似,但它们的核参数较少,所以应用价值更高.这2种核函数的不同之处在于,UCK适用于低维稀疏数据集,而GUCK适用于高维稀疏数据集.这是因为,对于高维的稀疏数据集,UCK会出现病态问题,从而影响了学习精度,而GUCK能有效地解决这个问题.

参考文献:

- [1] VAPNIK V N. The nature of statistical learning theory [M]. Dusseldorf, Germany: Springer-Verlag, 1995.
- [2] SUYKENS J A K, VANDEWALLE J. Least squares support vector machine classifiers [J]. Neural Processing Letters, 1999, 9(3):293-300.
- [3] FUNG G, MANGASARIAN O L. Proximal support vector machine classifiers [C]//Proc of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM, 2001:77-86.
- [4] HUANG H P, LIU Y H. Fuzzy support vector machines for pattern recognition and data mining [J]. International Journal of Fuzzy Systems, 2002, 4(3): 826-835.
- [5] LEE K Y, DAE-WON K. Possibilistic support vector machines [J]. Pattern Recognition, 2005, 38(8): 1325-1327.
- [6] ROSTAMIZADEH A. Theoretical foundations and algorithms for learning with multiple kernels [D]. New York, USA: New York University, 2010.
- [7] KLOFT M, BREFELD U, SONNENBURG S, et al. Non-sparse regularization for multiple kernel learning, machine learning group [J]. Franklinstr, 2010, 28(29):6-9.
- [8] CORTES C, MOHRI M, ROSTAMIZADEH A. Two-stage learning kernel algorithms [C]//Proc of the 27th International Conference on Machine Learning. Corvallis, USA. ICML, 2010:239-246.
- [9] CRISTIANINI N, KANDOLA J, ELISSEEF A, et al. On kernel target alignment [C]//Proc of the Neural Information Processing Systems. Cambridge, MA, USA: MIT Press, 2001: 367-373.
- [10] KANDOLA J, TAYLOR J S, CRISTIANINI N. Optimizing kernel alignment over combinations of kernels, Neuro COLT Technical Report NC-TR-02-121 [R]. London, UK: University of London, 2002.
- [11] HOLMES D E, JAIN L C. Innovations in machine learning [M]. Berlin, Germany: Springer-Verlag, 2006:205-256.
- [12] SONNENBURG S, RAETSCH G, SCHAEFER C, et al. Large scale multiple kernel learning [J]. Journal of Machine Learning Research, 2006, 7(7):1531-1565.
- [13] RAKOTOMAMONJY A, BACH F, CANU S, et al. More efficiency in multiple kernel learning [C]//Proc of the 24th International Conference on Machine Learning. Corvallis, USA: ICML, 2007:775-782.
- [14] ZHAO Jinwei, FENG Boqin, YAN Guirong, et al. The unified Chebyshev polynomial kernel function for support vector regression machine [C]//Proc of International Conference of Automatic Control and Artificial Intelligence. Herts SG12AY, UK: The Institution of Engineering and Technology, 2012:2364-2368.
- [15] YE N, SUN R, LIU Y, et al. Support vector machine with orthogonal Chebyshev kernel [C]//Proc. of the 18th International Conference on Pattern Recognition. Los Alamitos, CA, USA: IEEE Computer Society, 2006:752-755.
- [16] OZER S, CHEN C H. Generalized Chebyshev kernels for support vector classification [C]//Proc of the 19th International Conference on Pattern Recognition. Los Alamitos, CA, USA: IEEE Computer Society, 2008:1-4.
- [17] OZER S. On the classification performance of support vector machines using Chebyshev kernel functions [D]. Dartmouth, MA, USA: University of Massachusetts, 2007.
- [18] OZER S, CHEN C H, CIRPAN H A. A set of new Chebyshev kernel functions for support vector machine pattern classification [J]. Pattern Recognition, 2011, 44(7):1435-1447.
- [19] LANCKRIET G, BIE T D, CRISTIANINI N, et al. A statistical framework for genomic data fusion [J]. Bioinformatics, 2004, 20(16):2626-2635.
- [20] ZHANG L, ZHOU W, JIAO L. Wavelet support vector machine [J]. IEEE Trans Systems, Man, and Cybernetics; Part B Cybernetics, 2004, 34(1): 34-39.
- [21] ASUNCION A, NEWMAN D J. UCI machine learning repository [R/OL]. [2011-05-11]. http://www.ics.uci.edu/_mlearn/MLRepository.html, 2007.

(编辑 苗凌)