

文章编号: 1001-0920(2010)03-0399-05

## 灰色局部支持向量回归机及应用

蒋 辉<sup>1,2</sup>, 王志忠<sup>1</sup>

(1. 中南大学 数学科学与计算技术学院, 长沙 410075; 2. 惠州学院 数学系, 广东 惠州 516007)

**摘 要:** 为了解决全局支持向量回归机 (Global-SVR) 在大样本数据集中计算效率低下的问题, 将局部支持向量回归机与灰色系统理论有机结合, 并利用灰色关联度作为局部邻域函数构造灰色局部支持向量回归机 (GL-SVR), 该做法具有一定的理论价值. 优化过程中采用留一法评估学习机的泛化性能, 利用模式搜索法选择模型参数. 真实的股价涨跌幅预测实验结果表明, 该方法既加快了运算速度, 又提高了预测精度.

**关键词:** 局部支持向量回归; 灰色关联度; 股价涨跌幅

中图分类号: O234

文献标识码: A

## Grey local support vector regression and its application

J I A N G H u i<sup>1,2</sup>, W A N G Z h i-z h o n g<sup>1</sup>

(1. School of Mathematical Sciences and Computing Technology, Central South University, Changsha 410075, China;  
2. Department of Mathematical Sciences, Huizhou University, Huizhou 516007, China. Correspondent: J I A N G H u i,  
E-mail: jh@hzu.edu.cn.)

**Abstract:** Due to much time consumption of global support vector regression (Global-SVR) for large sample size data, grey local SVR (GL-SVR) combined grey relational grade regarded as neighbourhood function with local support vector regression is considered. To optimize the machine, based on leave-one-out errors, pattern search method is adopted for model selection. Experiments are carried out on a real stock price change forecasting with GL-SVR and the results show that the proposed approach can not only speed up the computing speed, but also improve the prediction accuracy.

**Key words:** Local support vector regression; Grey relational grade; Stock price change

### 1 引 言

时间序列数据具有大样本、高噪声、不稳定和非线性等特点, 用标准的支持向量学习方法在整个样本数据集上进行学习, 构造能涵盖整个输入空间的全局学习机往往使计算效率低下, 计算成本和代价较高. 为了解决这个问题, 通常采用化整为零、分而治之的策略, 提出了只在兴趣点 (预测点) 附近某个邻域内考虑学习问题的局部方法<sup>[1,2]</sup>. 一般地, 训练集在输入空间的分布是不均匀的, 局部方法根据该特点将输入空间划分成多个子区间, 无需构造适合整个训练集的学习机而只需在兴趣点 (预测点) 的某个范围内进行学习, 不仅可以节约大量的计算时间, 而且能提高计算结果的准确性. 但是, 训练集不能随意而无原则地缩小, 对于具有有限 VC 维的非负完全有界函数  $L(x, y, f(x, a))$ , Vapnik 等在经验风险最小化和收敛界定理的基础上, 借助邻域函数, 提

出小样本下局部风险最小化估计子的界, 引入局部风险最小化模型来最小化局部风险<sup>[3]</sup>. 在具体应用中邻域函数种类繁多, 如矩形函数、三角形函数、四次函数和高斯函数等<sup>[4]</sup>. 在时间序列的预测中, 也有学者从 - 不敏感损失函数出发, 根据兴趣点 (预测点) 周围局部样本的离散程度不同, 采用标准差作为局部权重函数确定灵活变化的, 以替代全局支持向量机中固定不变的, 这种方法能取得理想的预测效果<sup>[5]</sup>. 同时, 多重核、局部核回归和  $K$ - 近邻支持向量机的有机结合也能提高时间序列数据的预测精度<sup>[6]</sup>.

本文对于兴趣点 (预测点) 的特定局部样本, 由于训练点对预测点影响的不确定性, 将其看作灰色系统. 结合灰色系统理论<sup>[7]</sup>, 运用灰色关联分析, 将灰色关联度作为局部邻域函数构造灰色局部支持向量回归机 (GL-SVR) 来提高计算效率. 真实的股票

收稿日期: 2009-03-29; 修回日期: 2009-06-03.

作者简介: 蒋辉 (1972 —), 男, 湖南岳阳人, 讲师, 博士生, 从事统计数据分析、机器学习等研究; 王志忠 (1963 —), 男, 湖南岳阳人, 教授, 博士生导师, 从事数据分析、统计理论及应用等研究.

数据预测实验结果表明,灰色局部支持向量回归机具有优良的性能.

## 2 GL-SVR

### 2.1 局部风险最小化

为实现局部方法,必须先定量地表现“局部”这个概念.为此,首先构造一个由局部参数所控制的邻域函数  $V(x, x_0; \tilde{\cdot})$ . 其中:  $x_0$  为预测点,  $\tilde{\cdot} \in (0, 1)$  为局部参数. 邻域函数要求满足条件

$$0 \leq V(x, x_0; \tilde{\cdot}) \leq 1, V(x_0, x_0; \tilde{\cdot}) = 1. \quad (1)$$

在邻域函数的基础上,局部风险泛函定义为

$$R(a, \tilde{\cdot}, x_0) =$$

$$L(x, y, f(x, a)) \frac{V(x, x_0; \tilde{\cdot})}{\bar{V}(x_0; \tilde{\cdot})} dF(x, y). \quad (2)$$

其中:  $L(x, y, f(x, a))$  为损失函数;  $V(x, x_0; \tilde{\cdot})$  为邻域函数,且有

$$\bar{V}(x_0; \tilde{\cdot}) = \int V(x, x_0; \tilde{\cdot}) dF(x). \quad (3)$$

在分布函数未知且样本  $D = \{(x_i, y_i)\}_{i=1}^l$  已知的情况下,最小化式(2),即基于经验数据最小化局部风险泛函. 局部经验风险泛函定义为

$$R_{\text{emp}}(a, \tilde{\cdot}, x_0) =$$

$$\frac{1}{l \bar{V}_{\text{emp}}(x_0; \tilde{\cdot})} \sum_{i=1}^l L(x_i, y_i, f(x_i, a)) V(x_i, x_0; \tilde{\cdot}). \quad (4)$$

其中:  $D = \{(x_i, y_i)\}_{i=1}^l \subset (X, Y)^l$  为给定的训练集,  $x_i \in X \subset \mathbb{R}^n$ ,  $y_i \in Y \subset \mathbb{R}$ , 且有

$$\bar{V}_{\text{emp}}(x_0; \tilde{\cdot}) = \frac{1}{l} \sum_{i=1}^l V(x_i, x_0; \tilde{\cdot}). \quad (5)$$

类似于全局学习,可利用结构风险最小化原则最小化基于经验数据的局部风险泛函.

**定理 1** 设  $0 \leq L(x, y, f(x, a)) \leq B < +\infty$  ( $a \in \mathbb{R}$ ) 为非负完全有界函数集, 函数  $L(x, y, f(x, a)) \times V(x, x_0; \tilde{\cdot})$  ( $a \in \mathbb{R}, \tilde{\cdot} \in (0, 1)$ ) 具有有限 VC 维  $h_1$ , 函数  $V(x, x_0; \tilde{\cdot})$ ,  $\tilde{\cdot} \in (0, 1)$  具有有限 VC 维  $h_2$ , 则对所有的  $a \in \mathbb{R}$  和  $\tilde{\cdot} \in (0, 1)$ , 不等式

$$R(a, \tilde{\cdot}, x_0) \leq \left( R_{\text{emp}}(a, \tilde{\cdot}, x_0) + \frac{B}{2} E_{h_1}(l) \left( 1 + \sqrt{\frac{4 R_{\text{emp}}(a, \tilde{\cdot}, x_0)}{B E_{h_1}(l)}} \right) \right) / (\bar{V}_{\text{emp}}(x_0; \tilde{\cdot}) - \sqrt{E_{h_2}(l)}) + \quad (6)$$

依 1 - 2 的概率成立,  $\bar{V}_{\text{emp}}(x_0; \tilde{\cdot})$  由式(5) 确定, 且有

$$R_{\text{emp}}(a, \tilde{\cdot}, x_0) = \frac{1}{l} \sum_{i=1}^l L(x_i, y_i, f(x_i, a)) V(x_i, x_0; \tilde{\cdot}), \quad (7)$$

$$E_{h_1}(l) = 4 \frac{h_1 (\ln \frac{2l}{h_1} + 1) - \ln 4}{l}, \quad (8)$$

$$E_{h_2}(l) = \frac{h_2 (\ln \frac{2l}{h_2} + 1) - \ln 4}{l} + \frac{1}{l}. \quad (9)$$

在定理给出的收敛界的基础上,可以构造结构风险,实行结构风险最小化归纳原则. 由于是在局部上考虑学习,假设函数集并不需要像全局学习那样丰富,许多经典方法即可使简单函数在局部上取得不错的逼近效果.

### 2.2 局部核回归

考虑再生希尔伯特空间  $H$  中的核学习问题,回归(预测)问题归结为最小化如下正则化泛函<sup>[3]</sup>:

$$\min_{f \in H} R(f) = \frac{1}{l} \sum_{i=1}^l L(x_i, y_i, f(x_i)) + \|f\|^2. \quad (10)$$

引入邻域函数,并在预测点  $x_0$  邻域范围内实施结构风险最小化,则式(10) 为

$$\min_{f \in H} R(f, x_0) = \frac{1}{l \bar{V}_{\text{emp}}(x_0; \tilde{\cdot})} \sum_{i=1}^l L(x_i, y_i, f(x_i)) \times V(x_i, x_0; \tilde{\cdot}) + \|f\|^2. \quad (11)$$

定义如下新损失函数:

$$L(x_i, y_i, f(x_i)) = \frac{V(x_i, x_0; \tilde{\cdot})}{\bar{V}_{\text{emp}}(x_0; \tilde{\cdot})} L(x_i, y_i, f(x_i)), \quad (12)$$

则式(11) 可表述为

$$\min_{f \in H} R(f, x_0) = \frac{1}{l} \sum_{i=1}^l L(x_i, y_i, f(x_i)) + \|f\|^2. \quad (13)$$

由广义表示定理<sup>[3]</sup>,可以找到式(13) 的基函数(核)形式的解. 比较式(13) 与(10) 可知,局部方法等同于修正损失函数的全局学习方法,即用邻域函数对损失函数进行加权. 与最小化局部经验风险的经典局部方法相比,式(13) 引入了正则项,使得本文方法更具鲁棒性,且避免了过学习.

### 2.3 GL-SVR

灰色系统理论在实践中有着广泛的应用. 在局部支持向量回归中,对于某一预测点特定的局部样本,训练集数据的分布是不均匀的,各训练点对预测点的影响是不确定的,可以将其看作一个灰色系统,预测点同每一训练点的关系紧密程度可以用灰色关联度来描述. 灰色关联度是衡量数据点之间关系紧密程度的一个指标,其原理如下<sup>[8]</sup>:  $x_0 = (x_0(1), x_0(2), \dots, x_0(m))$  为参考序列,  $x_i = (x_i(1), x_i(2), \dots, x_i(m))$  为比较序列,  $i = 1, 2, \dots, n$ .

$\dots, x_i(m)) (i = 1, 2, \dots, n)$  为对比序列,  $x_0$  与  $x_i$  在第  $k(k = 1, 2, \dots, m)$  个分量的灰关联系数为

$$\begin{aligned} & (x_0(k), x_i(k)) = \\ & (\min_i \min_k |x_0(k) - x_i(k)| + \\ & \max_i \max_k |x_0(k) - x_i(k)|) / (|x_0(k) - \\ & x_i(k)| + \max_i \max_k |x_0(k) - x_i(k)|), \quad (14) \end{aligned}$$

其中  $[0, 1]$  是分辨系数.  $x_0$  与  $x_i$  的灰色关联度为

$$(x_0, x_i) = \frac{1}{m} \sum_{k=1}^m (x_0(k), x_i(k)), \quad (15)$$

其中  $(x_0, x_i)$  是参考序列  $x_0$  与比较序列  $x_i$  关系紧密程度的度量,  $(x_0, x_i)$  越大,  $x_i$  与  $x_0$  的相似程度越高. 根据灰色关联度的定义,  $(x_0, x_i) \in [0, 1]$ , 且  $x_i = x_0$  时,  $(x_0, x_i) = 1$ , 因此它具有邻域函数的特性. 考虑如下标准的支持向量机 (SVM) 数学形式:

$$\begin{aligned} \min_{w,b,\xi} & C \sum_{i=1}^l (\xi_i + \xi_i^*) + \frac{1}{2} \|w\|^2; \\ \text{s.t.} & y_i - ((w \cdot \phi(x_i)) + b) \leq \xi_i, \\ & ((w \cdot \phi(x_i)) + b) - y_i \leq \xi_i^*, \\ & \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, \dots, l. \end{aligned} \quad (16)$$

将灰色关联度作为邻域函数引入支持向量机, 构造 GL-SVR 如下:

$$\begin{aligned} \min_{w,b,\xi} & C \sum_{i=1}^l \frac{(x_0, x_i)}{l} (\xi_i + \xi_i^*) + \frac{1}{2} \|w\|^2; \\ \text{s.t.} & y_i - ((w \cdot \phi(x_i)) + b) \leq \xi_i, \\ & ((w \cdot \phi(x_i)) + b) - y_i \leq \xi_i^*, \\ & \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, \dots, l. \end{aligned} \quad (17)$$

其中  $\bar{C} = \frac{1}{l} \sum_{i=1}^l (x_0, x_i)$ . 以邻域函数修正参数  $C$ , 定义

$$C_{x_0,i} = (x_0, x_i) \bar{C}, \quad (18)$$

则式 (17) 可改写为

$$\begin{aligned} \min_{w,b,\xi} & C_{x_0,i} \sum_{i=1}^l (\xi_i + \xi_i^*) + \frac{1}{2} \|w\|^2; \\ \text{s.t.} & y_i - ((w \cdot \phi(x_i)) + b) \leq \xi_i, \\ & ((w \cdot \phi(x_i)) + b) - y_i \leq \xi_i^*, \\ & \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, \dots, l. \end{aligned} \quad (19)$$

考虑到样本数据分布可能不均匀, 同样的局部半径在不同预测点覆盖的样本容量可能相差很大, 影响预测效果. 因此实际应用中, 在训练集中选择与预测点  $x_0$  最近的  $k$  个样本点来构造支持向量回归机, 由此产生  $k$ -近邻支持向量回归机 (KNN-SVR). 控制局部性的参数  $k$  必须谨慎选择, 太小易受噪声影响, 出现过学习; 太大则难以捕捉目标函数局部性

特点, 并增加计算成本.

### 2.4 模型优化

局部方法考虑样本较少, 可采用留一法评估学习机的泛化性能, 优化算法采用计算效率较高的模式搜索法, 以实现高效预测的目标.

留一法的原理是在容量为  $l$  的数据  $T$  上, 剔除一个样本点; 然后在余下的  $l - 1$  个样本点上训练、构造学习机预测剔除的样本点. 如此循环, 让剔除样本点跑遍整个样本集, 共进行  $l$  次留一学习与预测. 学习机的泛化性能综合  $l$  次预测结果进行评价, 其评估度量指标采用 MSE (Mean square error) 方法. 整个数据集上的留一经验风险是对期望风险的几乎无偏估计<sup>[3,9]</sup>.

模式搜索法只需直接计算优化的目标函数 (学习问题中的风险), 不需要关于目标函数的梯度或其他先验信息, 与其他算法 (如遗传算法) 比较, 计算成本较低<sup>[10]</sup>. 该方法通过搜索当前点周围的一系列点 (网格 Mesh), 寻找使目标函数值低于当前点的对应点. 如此重复, 从而确定一个越来越接近于理性点的序列. 当前点是算法在上一步计算出来的, 算法通过将当前点加上一个称之为模式的固定向量的数乘来构成网格. 如果算法在网格中通过表决找到一个改善当前点对应目标函数值的新点, 则算法就在下一步将新点作为当前点. 为避免过学习、增加学习机的鲁棒性, 模式搜索法通常只要求所选超参数对应学习机的泛化性能在验证集上表现优良, 无需绝对的最优. 另外, 模式搜索法可以结合其他方法获得更好的表现, 也可以采用装袋法增强算法稳定性, 跳出局部最小值<sup>[11]</sup>. 对于局部学习机超参数的选择, 模式搜索法具有较好的效果.

### 3 实证分析

根据上海证券交易所一支股票价格的历史数据进行试验, 检验局部学习机 GL-SVR 的性能. 实验采用 Gaussian 核, 所有计算均在 Matlab 7.0.1 上编程实现. 为了比较, 构造全局学习机 (Global-SVR) 和以欧式距离的倒数为加权因子<sup>[12]</sup>的局部学习机 (DL-SVR), 为提高其计算效率, 程序中应用优化子程序 Optimiser 模块<sup>[13]</sup>.

数据集 SP-Data 共含 1 194 个数据点, 连续 1 194 个交易日的收盘价变化曲线和涨跌幅 (%) 曲线分别如图 1 和图 2 所示. 重构数据集如下<sup>[14]</sup>:

$$\hat{y}_i = f(x_i) = f(x_{i-4}, x_{i-3}, x_{i-2}, x_{i-1}). \quad (20)$$

其中  $f$  通过学习机获得,  $x_i$  和  $y_i$  分别为第  $i$  个交易日的收盘价和股价涨跌幅 (%). 由式 (20) 可知, 新数据集含 1 190 个数据输入模式 (4 维向量) 和数据输出模式 (1 维向量). 为便于比较, 此处采用 MSE

度量学习机泛化性能. 为检验学习机的性能,对于 3 种学习机在同一数据集 SP-Data 内进行实验. 实验设置如表 1 所示,模式搜索法用于模型优化、确定超参数值. 为提高效率,预先给定一个有限且足够宽泛的搜索空间,并利用  $\ln(\cdot)$  对空间适当压缩,有

$$u = (\ln C, \ln C, \ln C)$$

为待优化的超参数向量,搜索范围设置为  $[-2, 10], [-2, 10], [-2, 10]$ ,即  $0.1353, 22.026, 0.1353, C, 22.026, 0.1353, 7.3891$ .

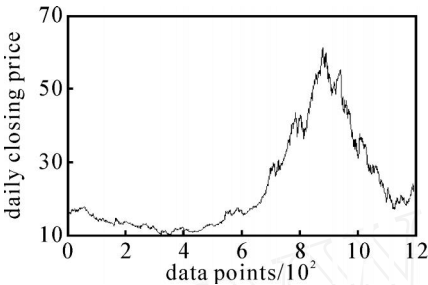


图 1 股价趋势曲线

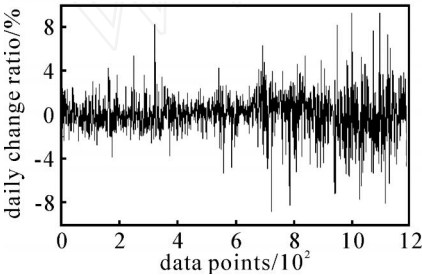


图 2 股价涨跌幅曲线

表 1 学习机实验设置

学习机	$k$ 值	模型选择	学习模式	训练集大小
Global-SVR	—	验证集误差	批量	980
DL-SVR	20	留一误差	在线	20
GL-SVR	20	留一误差	在线	20
	10 ~ 30	留一误差	在线	$k$

为了比较灰色局部学习机 LGSVR 的性能,此处同时考察 Global-SVR 和 DL-SVR 的性能. 对于 Global-SVR,模型优化采用交叉检验法,学习模式采用批量学习,将数据集 SP-Data 分成训练集、验证集与测试集 3 部分,分别包含 980,130,80 个数据模式,优化 3 个超参数,即  $u = (\ln C, \ln C, \ln C)$ ;对于 DL-SVR,采用欧氏距离的倒数作为加权因子对惩罚参数  $C$  进行加权. 学习机 GL-SVR 和 DL-SVR 均使用固定的  $k(k = 20)$  值表达局部性,模型选择均依据留一误差,学习模式均采用在线型式. 通过实验,发现局部方法能节省大量运算时间,比全局方法具有更高的精度,而且在局部方法中,GL-SVR 比

DL-SVR 误差小. 计算结果见表 2,比较曲线见图 3.

表 2 学习机在测试集上的 MSE 比较

	Global-SVR	DL-SVR	GL-SVR
MSE	8.1794	1.7610	1.4338

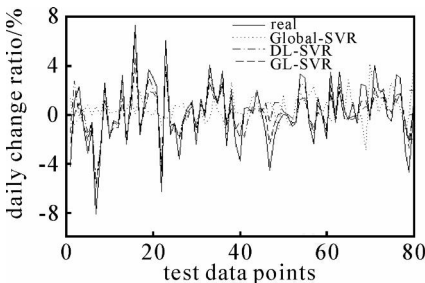


图 3 SP-Data 测试集上的预测曲线与真实曲线

以下考察局部性的设置对学习机 GL-SVR 性能的影响. 实验中,局部性由  $k$  表达,为分析  $k$  对实验结果的影响,设定 10  $k$  30 分别进行实验. 结果表明, $k$  值越小,计算时间越少,预测误差越小. 为了更清晰地了解  $k$  值的影响,分别取  $k = 15, 20, 25$ ,将测试集上 MSE 的计算结果列于表 3,预测曲线如图 4 所示. 上述分析表明,大的  $k$  值并不能保证更好的泛化性能(其极端情况就是全局学习机),小的  $k$  值会影响学习机的稳定性,导致过学习;另外,计算效率还与  $k$  值的选择有关. 因此,必须综合考虑学习机的泛化性能、稳定性及计算效率等因素,谨慎选择  $k$  值.

表 3 GL-SVR在测试集上的 MSE 比较

$k_{GL-SVR}$	$k = 15$	$k = 20$	$k = 25$
MSE	0.9023	1.4338	2.0402

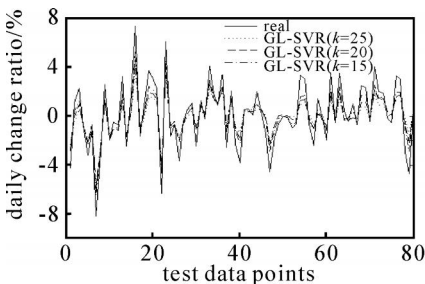


图 4 SP-Data 测试集上预测曲线与真实曲线(不同  $k$  值)

4 结 论

本文讨论了减少学习规模以提高学习效率的局部学习方法,重点研究了 GL-SVR 的构造及其应用. GL-SVR 引进灰色关联度作为局部邻域函数,并与标准支持向量回归机有机结合起来,具有与标准支持向量回归机类似的一般形式,等同于以表现局部性的邻域函数加权修正损失函数的标准支持向量回归机. 同时,为了提高学习效率,本文在方法上采用适合局部学习的留一精确误差评估学习机的泛化

性能,并应用模式搜索法对超参数进行选择,为大规模的学习问题提供了一条简明的新思路.

为检验实际问题中 GL-SVR 的性能,本文在真实时间序列 SP-Data 数据集上进行了实验,并与 Global-SVR 和 DL-SVR 的实验结果进行了比较.结果肯定了 GL-SVR 的性能,表明了局部回归机能以较小的计算成本取得泛化性能的改善,且模式搜索法计算效率高,性能稳定,可作为优化超参数的一种理想选择.

### 参考文献(References)

- [1] Bottou L, Vapnik V N. Local learning algorithm[J]. Neural Computation, 1992, 4(6): 888-900.
- [2] Fernandez R. Predicting time series with a local support vector regression machine [EB/OL]. <http://www.iit.demokritos.gr/skel/eetn/acai99/>, 2008-10-23.
- [3] Vapnik V, Statistical learning theory[M]. New york: Wiley, 1998.
- [4] Hardle W, Muller M, Sperlich S, et al. Nonparametric and semiparametric models [M]. Berlin: Springer, 2004: 40-145.
- [5] Kaizhu H, Haiqin Y, Irwin K, et al. Local support vector regression for financial time series prediction[C]. Int Joint Conf on Neural Networks Sheraton Vancouver Wall Centre Hotel. Vancouver, 2006:1622-1627.
- [6] He W, Wang Z. Optimized local kernel machines for fast time series forecasting[C]. Proc of the 3rd Int Conf on Natural Computation. Washington: IEEE, 2007: 620-627.
- [7] Deng J. Control problems of grey systems[J]. Systems and Control Letter, 1982, 1(5): 288-294.
- [8] Deng J. Gray prediction and gray decision making[M]. Wuhan: Central China University Science and Technology Press, 2002: 204-206.
- [9] Chapelle O, Vapnik V, Bousquet O, et al. Choosing multiple parameters for support vector machines [J]. Machine Learning, 2002, 46(1-3): 131-159.
- [10] Momma M, Bennett K P. A pattern search method for model selection of support vector regression[C]. Proc of the 2nd SIAM Int Conf on Data Mining. Philadelphia, 2002.
- [11] Payne J L, Eppstein M J. A hybrid genetic algorithm with pattern search for finding heavy atoms in protein crystals[C]. Proc of the 2005 Conf on Genetic and Evolutionary Computation. New York: ACM, 2005: 377-384.
- [12] Chris Atkeson, Andrew Moore, Stefan Shaal. Locally weighted learning [J]. Artificial Intelligence Review, 1997, 11(1-5): 11-73.
- [13] Steve Gunn. Matlab SVR toolbox [EB/OL]. <http://www.isis.ecs.soton.ac.uk/resources/svminfo/>, 2000-08-12.
- [14] Haiqin Yang King, Chan I Laiwan. Non-fixed and asymmetrical margin approach to stock market prediction using support vector regression[C]. Proc of the 9th Int Conf on Publication Date. 2002, 3: 1398-1402.
- [9] Albert H R Ko, Robert Sabourin, Alceu de Souza Britto Jr, et al. Pairwise fusion matrix for combining classifiers[J]. Pattern Recognition, 2007, 40(8): 2198-2210.
- [10] Khan S S, Ahmad A. Cluster center initialization algorithm for  $k$ -means clustering [J]. Pattern Recognition Letters, 2004, 25(11): 1293-1302.
- [11] Stephen J Redmond, Conor Heneghan. A method for initialising the K-means clustering algorithm using  $k$ -trees[J]. Pattern Recognition Letters, 2007, 28(8): 965-973.
- [12] Wang Y N, Li C S, Zuo Y. A selection model for optimal fuzzy clustering algorithm and number of clusters based on competitive comprehensive fuzzy evaluation[J]. IEEE Trans on Fuzzy Systems, 2009, 17(3): 568-577.
- [13] The UCI machine learning repository [EB/OL]. <http://www.ics.uci.edu/~mlearn/MLRepository.html>, 1993-07-01.
- [14] Chang C C, Lin C J. LIBSVM: A library for support vector machines 2001 [EB/OL]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2001-10-16.
- [15] Wei W, Mendel J M. Optimality tests for the fuzzy  $c$ -means algorithm [J]. Pattern Recognition Letters, 1994, 27(11): 1567-1573.
- [16] Campello R J GB. A fuzzy extension of the rand index and other related indexes for clustering and classification assessment [J]. Pattern Recognition Letters, 2007, 28(7): 833-841.

(上接第 398 页)