第 49 卷 第 5 期
2013 年 9 月

南京大学学报(自然科学)
JOURNAL OF NANJING UNIVERSITY
(NATURAL SCIENCES)

Vol. 49, No. 5
Sept. ,2013

粗糙 ε -支持向量回归模型*

张仕光[1,2],米据生[1,3**],胡清华[4]

(1. 河北师范大学数学与信息科学学院,石家庄,050024;2. 衡水学院数学与计算机学院,衡水,053000;
3. 河北省计算数学与应用重点实验室,石家庄,050024;4. 天津大学计算机科学与技术学院,天津,300072)

摘　要: 在 ε -支持向量回归和粗糙 v -支持向量回归模型的基础上,研究了新的粗糙 ε -支持向量回归模型. 利用固定对称边界粗糙 ε -不敏感损失函数,得到粗糙 ε -不敏感管,构造固定对称边界粗糙 ε -支持向量回归模型;利用固定非对称边界粗糙-不敏感损失函数,得到粗糙 $\varepsilon^u-\varepsilon^d$ -不敏感管,构造固定非对称边界粗糙 ε -支持向量回归模型. 通过引进 Lagrange 函数和根据 KKT 条件,处理粗糙 ε -支持向量回归模型的对偶问题.

关键词: ε -支持向量回归,粗糙边界,粗糙 ε -支持向量回归,粗糙集

Rough ε-support vector regression model

Zhang Shi-Guang[1,2], *Mi Ju-Sheng*[1,3], *Hu Qing-Hua*[4]

(1. College of Mathematics and Information Science, Hebei Normal University, Shijiazhuang, 050024, China;
2. College of Mathematics and Computer Science, Hengshui University, Hengshui, 053000, China;
3. Hebei Key Laboratory of Computational Mathematics and Applications, Shijiazhuang, 050024, China;
4. School of Computer Science and Technology, Tianjin University, Tianjin, 300072, China)

Abstract: Two new rough ε-support vector regression models are proposed based on ε-support vector regression, rough v-support vector regression and rough set theory. Firstly, a rough boundary ε-insensitive tube is defined with fixed symmetrical boundary rough ε-insensitive loss function, the method of optimization and ε-support vector regression model. Moreover we design a fixed symmetrical boundary rough ε-support vector regression model(RFSM ε-SVR). Secondly, we extend the model to the case that asymmetrical loss function is considered. Finally, according to Karush-Kuhn-Tucker(KKT) conditions, we derive their dual problems by introducing the Lagrange functional into rough ε-support vector regression models.

Key words: ε-support vector regression, rough boundary, rough ε-support vector regression, rough set

---

* 基金项目:国家自然科学基金(61170107,61222210),河北省高等学校科学研究计划(Z2010188)
收稿日期:2013－06－20

** 通讯联系人,E-mail:mijsh@263.net

支持向量回归方法已成功应用于研究和应用等各个领域,并吸引了人们的广泛关注[1~7]. 设给定训练集 $D_l=\{(x_1,y_1),(x_2,y_2),\cdots,(x_l,y_l)\}$,其中 $x_l\in R^n,y_i\in R,i=1,\cdots,l$. 假定 D_l 是某个概率分布 $P(x,y)$ 上选取的独立同分布的样本,给定损失函数为 $c(x,y,f(x))$.

Vapnik、Cortes[1,2] 提出 ε-不敏感损失函数:

$$c(x,y,f(x))=|y-f(x)|_\varepsilon \tag{1}$$

其中:

$$|y-f(x)|_\varepsilon=\begin{cases}0 & |y-f(x)|\leqslant\varepsilon\\ |y-f(x)|-\varepsilon & |y-f(x)|>\varepsilon\end{cases}$$

ε 是给定的正数.

Yang 等[6,7]提出固定非对称边界 ε-不敏感损失函数:

$$c(x_i,y_i,f(x_i))=\begin{cases}0 & -\varepsilon_i^d\leqslant y_i-f(x_i)\leqslant\varepsilon_i^u\\ y_i-f(x_i)-\varepsilon_i^u & y_i-f(x_i)\geqslant\varepsilon_i^u\\ f(x_i)-y_i-\varepsilon_i^d & f(x_i)-y_i\geqslant\varepsilon_i^d\end{cases} \tag{2}$$

其中 $\varepsilon_i^u,\varepsilon_i^d$ 分别对应第 i 个上边界和下边界. 当 $\varepsilon_i^u,\varepsilon_i^d(i=1,2,\cdots,l)$ 为同一个常数时,称为固定对称边界,式(2)即为式(1)的 ε-不敏感损失函数;当 $\varepsilon_i^d=\varepsilon^d$ 和 $\varepsilon_i^u=\varepsilon^u(i=1,\cdots,l)$ 且 $\varepsilon^u\neq\varepsilon^d$ 时,称为固定非对称边界.

Pawlak 等提出的粗糙集模型[8~13]是处理各种不完备信息的有效工具. Zhao[14]、Lingras 等[15]把粗糙集模型和支持向量回归模型结合起来,分别提出粗糙 ν-支持向量回归和粗糙支持向量回归模型,取得了较好的效果.

本文在 ε-支持向量回归和粗糙 ν-支持向量回归模型的基础上,研究了新的粗糙 ε-支持向量回归模型. 应用固定对称边界粗糙 ε-不敏感损失函数,构造固定对称边界粗糙 ε-支持向量回归模型;应用固定非对称边界粗糙 ε-不敏感损失函数,构造固定非对称边界粗糙 ε-支持向量回归模型.

1　粗糙集理论

粗糙集理论是利用集合的上近似、下近似与边界等概念而建立的,是处理不精确、不一致、不完整等各种不完备信息的有效工具,可以直接对数据进行分析和推理,从中发现隐含的知识,揭示潜在的规律,因此是一种天然的数据挖掘或者知识发现方法. 设非空有限论域 $U=\{x_1,x_2,\cdots,x_n\}$,A 为非空有限属性集,对任意属性子集 $B\subseteq A$,定义 U 上的等价关系(也称为不可辨识关系)R:

$$R=\{(x_i,x_j)\mid\forall a\in B,a(x_i)=a(x_j)\}$$

其中 $a(x_i)$表示属性 a 在点 x 的值. 等价关系 R 把 U 划分成一些等价类,表示为 U/R. 对于一个集合 $Y\subseteq U$,Y 的下近似定义为:

$$\underline{R}(Y)=\bigcup\{X\in U/R\mid X\subseteq Y\mid\} \tag{3}$$

Y 的上近似定义为:

$$\overline{R}(Y)=\bigcup\{X\in U/R\mid X\cap Y\neq\varnothing\} \tag{4}$$

由式(3)、(4),定义分类精度:

$$\rho_R=1-\frac{|\underline{R}(Y)|}{|\overline{R}(Y)|} \tag{5}$$

若 $0<\rho_R\leqslant1$,则 Y 是粗糙集;若 $\rho_R=1$,则 Y 是经典集.

2　ε-支持向量回归模型

Vapnik、Cortes[1,2] 利用 ε-不敏感损失函数得到 ε-支持向量回归模型(ε-support vector regression model,ε-SVR)的原问题为:

$$\min\{g_{P_\varepsilon}=\frac{1}{2}\tilde{\omega}^T\tilde{\omega}+C/l(\sum_{i=}^{l}(\xi_i+\xi_i^*))\}$$

$$\text{s.t.}\begin{cases}\tilde{\omega}^T\cdot\phi(x_i)+b-y_i\leqslant\varepsilon+\xi_i\\ y_i-\tilde{\omega}^T\cdot\phi(x_i)-b\leqslant\varepsilon+\xi_i^*\\ \xi_i,\xi_i^*\geqslant0,i=1,2,\cdots,l\end{cases} \tag{6}$$

其中 ε 是给定的,$\phi:R^n\to H$ 为核映射,H 为 Hilbert 空间.

引进 Lagrange 泛函,利用 KKT 条件可得 ε-SVR 的对偶问题[3,4]:

$$\max\{g_{D_\varepsilon}=-\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i^*-\alpha_i)(\alpha_j^*-\alpha_j)K(x_i,x_j)+\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)y_i-\varepsilon\sum_{i=1}^{l}(\alpha_i^*+\alpha_i)\}$$

$$\text{s.t.}\begin{cases}\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)=0\\ 0\leqslant\alpha_i,\alpha_i^*\leqslant C/l,i=1,\cdots,l\end{cases} \tag{7}$$

其中$(\phi(x_i),\phi(x_j))$为H空间中的内积,$K(x_i,x_j)=(\phi(x_i),\phi(x_j))$为核函数.

从上述对偶问题解得最优值α_i,α_i^* $(i=1,\cdots,l)$,可得ε-SVR决策函数:

$$f(x)=\tilde{\omega}^T\cdot\phi(x)+b=\sum_{i\in RSV}(\alpha_i^*-\alpha_i)K(x_i,x)+b \tag{8}$$

其中RSV表示$(\alpha_i^*-\alpha_i)\neq 0$所对应的向量,称为支持向量.

3　固定对称边界粗糙ε-支持向量回归模型

利用ε-不敏感损失函数(式(1)),构造固定对称边界粗糙ε-不敏感损失函数:

$$l_{R-\varepsilon}(x,y,f(x))=|y-f(x)|_{R-\varepsilon} \tag{9}$$

其中:

$$|y-f(x)|_{R-\varepsilon}=\begin{cases}0 & |y-f(x)|\leqslant\varepsilon_1\\ \varepsilon_2-|y-f(x)| & |y-f(x)|\in(\varepsilon_1,\varepsilon_2]\\ |y-f(x)|-\varepsilon_2 & |y-f(x)|\geqslant\varepsilon_2\end{cases}$$

这里ε_1、ε_2 $(\varepsilon_1<\varepsilon_2)$是给定的正数.

利用ε-不敏感损失函数构造的ε-SVR只考虑单个对称ε-不敏感管,落在ε-不敏感管内部的样本是非支持向量,落在ε-不敏感管的边界上或外边的样本是支持向量.根据粗糙集理论和最优化理论,利用ε_1、ε_2代替ε构造固定对称边界粗糙ε-不敏感损失函数,得到固定对称粗糙ε-不敏感管的边界,称之为粗糙ε-不敏感管.采用固定对称边界粗糙ε-不敏感损失函数,提出固定对称边界粗糙ε-支持向量回归模型(Fixed Symmetrical Boundary Rough ε-Support Vector Regression Model,RFSM ε-SVR).

RFSM ε-SVR的原问题为:

$$\min\{g_{P'_\varepsilon}=\frac{1}{2}\tilde{\omega}^T\tilde{\omega}+C/l((\sum_{i=}^{l}(\xi_i+\xi_i^*)+\delta\sum_{i=}^{l}(\eta_i+\eta_i^*))\}$$

$$\text{s. t.}\quad\begin{cases}\tilde{\omega}^T\cdot\phi(x_i)+b-y_i\leqslant\varepsilon_1+\xi_i+\eta_i\\ y_i-\tilde{\omega}^T\cdot\phi(x_i)-b\leqslant\varepsilon_1+\xi_i^*+\eta_i^*\\ 0\leqslant\xi_i,\xi_i^*\leqslant\varepsilon_2-\varepsilon_1\\ \eta_i,\eta_i^*\geqslant 0,i=1,\cdots,l\end{cases} \tag{10}$$

其中$\varepsilon_1,\varepsilon_2\geqslant 0,\varepsilon_2\geqslant\varepsilon_1,\delta\geqslant 1$是给定的.

定理1　RFSM ε-SVR对偶问题为:

$$\max\{g_{D'_\varepsilon}=-\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i^*-\alpha_i)(\alpha_j^*-\alpha_j)K(x_i,x_j)+\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)y_i-\varepsilon_2\sum_{i=1}^{l}(\alpha_i^*+\alpha_i)+2C/l(\varepsilon_2-\varepsilon_1)\}$$

$$\text{s. t.}\quad\begin{cases}\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)=0\\ 0\leqslant\alpha_i,\alpha_i^*\leqslant\delta C/l,i=1,\cdots,l\end{cases} \tag{11}$$

证明　引入Lagrange函数$L(\tilde{\omega},b,\xi,\xi^*,\eta,\eta^*)$:

$$L(\tilde{\omega},b,\xi,\xi^*,\eta,\eta^*)=\frac{1}{2}\tilde{\omega}^T\tilde{\omega}+C/l((\sum_{i=1}^{l}\xi_i+\xi_i^*)+\delta\sum_{i=1}^{l}(\eta_i+\eta_i^*))-\sum_{i=1}^{l}(\beta_i\eta_i+\beta_i^*\eta_i^*)-\sum_{i=1}^{l}\alpha_i(\varepsilon_1+\xi_i+\eta_i-\tilde{\omega}^T\phi(x_i)-b+y_i)-\sum_{i=1}^{l}\alpha_i^*(\varepsilon_1+\xi_i^*+\eta_i^*+\tilde{\omega}^T\phi(x_i)+b-y_i)-\sum_{i=1}^{l}\gamma_i(\varepsilon_2-\varepsilon_1-\xi_i)-\sum_{i=1}^{l}\gamma_i^*(\varepsilon_2-\varepsilon_1-\xi_i^*)$$

其中$\alpha_i\geqslant 0,\alpha_i^*\geqslant 0,\beta_i\geqslant 0,\beta_i^*\geqslant 0,\gamma_i\geqslant 0,\gamma_i^*\geqslant 0$,$i=1,\cdots,l$是Lagrange乘子.由KKT条件得RFSM ε-SVR对偶问题(式(11)).

从对偶问题(式(11))解得最优值α_i,α_i^* $(i=1,\cdots,l)$可得RFSM ε-SVR的决策函数:

$$f(x)=\tilde{\omega}^T\phi(x)+b=\sum_{i\in RSV}(\alpha_i^*-\alpha_i)K(x_i,x)+b \tag{12}$$

其中RSV表示$(\alpha_i^*-\alpha_i)\neq 0$所对应的样本集.

(1)当$\alpha_i\in(0,C/l)$,$\alpha_i^*=0$,或者$\alpha_i\in(C/l,\delta C/l)$,$\alpha_i^*=0$时:

$$b=y_j-\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)K(x_i,x_j)+\varepsilon_1$$

(2)当$\alpha_i^*\in(0,C/l)$,$\alpha_i=0$,或者$\alpha_i^*\in(C/l,\delta C/l)$,$\alpha_i=0$时:

$$b=y_k-\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)K(x_i,x_k)-\varepsilon_1$$

(3)当$j\in\{i:(\alpha_i^*-\alpha_i)\in(0,C/l)\}$,$k\in\{i:(\alpha_i^*-\alpha_i)\in(-C/l,0)\}$;或当$j\in\{i:(\alpha_i^*-\alpha_i)\in(C/l,\delta C/l)\}$,$k\in\{i:(\alpha_i^*-\alpha_i)\in(-\delta C/l,-C/l)\}$时:

$$b=\frac{1}{2}\{y_j+y_k-\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)K(x_i,x_j)-\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)K(x_i,x_k)\}$$

4 固定非对称边界粗糙 ε-支持向量回归模型

利用固定非对称边界 ε-不敏感损失函数(式(2)),构造固定非对称边界粗糙 ε-不敏感损失函数：

$$l_{R-\varepsilon}(x,y,f(x))=l_{-R-\varepsilon}(x,y,f(x))+l_{R-\varepsilon}^{-}(x,y,f(x)) \tag{13}$$

其中：

$$l_{-R-\varepsilon}(x_i,y_i,f(x_i))=\begin{cases}0 & f(x_i)-y_i\leqslant\varepsilon^d\\ \varepsilon^{dd}-(f(x_i)-y_i) & \varepsilon^d<f(x_i)-y_i\leqslant\varepsilon^{dd}\\ (f(x_i)-y_i)-\varepsilon^{dd} & f(x_i)-y_i\geqslant\varepsilon^{dd}\end{cases}$$

$$l_{R-\varepsilon}^{-}(x_i,y_i,f(x_i))=\begin{cases}0 & y_i-f(x_i)\leqslant\varepsilon^u\\ \varepsilon^{uu}-(y_i-f(x_i)) & \varepsilon^u<y_i-f(x_i)\leqslant\varepsilon^{uu}\\ (y_i-f(x_i))-\varepsilon^{uu} & y_i-f(x_i)\geqslant\varepsilon^{uu}\end{cases}$$

这里 $\varepsilon^d,\varepsilon^{dd},\varepsilon^u,\varepsilon^{uu}$ 是给定的正数，$\varepsilon^{dd}\geqslant\varepsilon^d,\varepsilon^{uu}\geqslant\varepsilon^u,\varepsilon^u,\varepsilon^d>0$.

根据粗糙集理论和最优化理论，利用 ε^{uu}、ε^{dd}、ε^u、ε^d 代替 ε^u、ε^d，构造固定非对称边界粗糙 ε-不敏感损失函数，得到固定非对称粗糙 ε^u-ε^d-不敏感管的边界，称之为粗糙 ε^u-ε^d-不敏感管. 采用固定非对称边界粗糙 ε-不敏感损失函数，我们提出固定非对称边界粗糙 ε-支持向量回归模型(Fixed Asymmetrical Boundary Rough ε-Support Vector Regression Model, RFAM ε-SVR).

RFAM ε-SVR 的原问题可描述为：

$$\min\{g_{P'_{\varepsilon^u,\varepsilon^d}}=\frac{1}{2}\tilde{\omega}^T\tilde{\omega}+C/l(\sum_{i=}^{l}(\xi_i+\xi_i^*)+\delta\sum_{i=}^{l}(\eta_i+\eta_i^*))\}$$

$$\text{s. t.}\begin{cases}\tilde{\omega}^T\cdot\phi(x_i)+b-y_i\leqslant\varepsilon^d+\xi_i+\eta_i\\ y_i-\tilde{\omega}^T\cdot\phi(x_i)-b\leqslant\varepsilon^u+\xi_i^*+\eta_i^*\\ 0\leqslant\xi_i\leqslant\varepsilon^{dd}-\varepsilon^d,\eta_i\geqslant0,i=1,\cdots,l\\ 0\leqslant\xi_i^*\leqslant\varepsilon^{uu}-\varepsilon^u,\eta_i^*\geqslant0,i=1,\cdots,l\end{cases} \tag{14}$$

其中 $\varepsilon^{dd}\geqslant\varepsilon^d,\varepsilon^{uu}\geqslant\varepsilon^u,\varepsilon^u,\varepsilon^d>0,\delta\geqslant1$ 是给定的.

定理 2　RFAM ε-SVR 的对偶问题为：

$$\max\{g_{D_\varepsilon}=-\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i^*-\alpha_i)(\alpha_j^*-\alpha_j)K(x_i,x_j)+\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)y_i-\varepsilon^{dd}\sum_{i=1}^{l}\alpha_i-\varepsilon^{uu}\sum_{i=1}^{l}\alpha_i^*+C/l(\varepsilon^{dd}-\varepsilon^d+\varepsilon^{uu}-\varepsilon^u)\}$$

$$\text{s. t.}\begin{cases}\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)=0\\ 0\leqslant\alpha_i,\alpha_i^*\leqslant\delta C/l,i=1,\cdots,l\end{cases} \tag{15}$$

证明　引入 Lagrange 函数 $L(w,b,\xi,\xi^*,\eta,\eta^*)$：

$$L(w,b,\xi,\xi^*,\eta,\eta^*)=\frac{1}{2}\tilde{\omega}^T\tilde{\omega}+C/l(\sum_{i=1}^{l}(\xi_i+\xi_i^*)+\delta\sum_{i=1}^{l}(\eta_i+\eta_i^*))-\sum_{i=1}^{l}(\beta_i\eta_i+\beta_i^*\eta_i^*)-\sum_{i=1}^{l}\alpha_i(\varepsilon^d+\xi_i+\eta_i-\tilde{\omega}^T\phi(x_i)-b+y_i)-\sum_{i=1}^{l}\alpha_i^*(\varepsilon^u+\xi_i^*+\eta_i^*+\tilde{\omega}^T\phi(x_i)+b-y_i)-\sum_{i=1}^{l}\gamma_i(\varepsilon^{dd}-\varepsilon^d-\xi_i)-\sum_{i=1}^{l}\gamma_i^*(\varepsilon^{uu}-\varepsilon^u-\xi_i^*)$$

其中：$\alpha_i\geqslant0,\alpha_i^*\geqslant0,\beta_i\geqslant0,\beta_i^*\geqslant0,\gamma_i\geqslant0,\gamma_i^*\geqslant0,i=1,\cdots,l$ 是 Lagrange 乘子. 由 KKT 条件得 RFAM ε-SVR 的对偶问题(式(15)).

从对偶问题(式(15))解得最优值 $\alpha_i,\alpha_i^*(i=1,\cdots,l)$，可得 RFAM ε-SVR 的决策函数：

$$f(x)=\tilde{\omega}^T\phi(x)+b=\sum_{i\in RSV}(\alpha_i^*-\alpha_i)K(x_i,x)+b \tag{16}$$

其中 RSV 表示 $(\alpha_i^*-\alpha_i)\neq0$ 所对应的样本集.

(1) 当 $\alpha_i\in(0,C/l),\alpha_i^*=0$，或者 $\alpha_i\in(C/l,\delta C/l),\alpha_i^*=0$ 时：

$$b=y_j-\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)K(x_i,x_j)+\varepsilon_d$$

(2) 当 $\alpha_i^*\in(0,C/l),\alpha_i=0$，或者 $\alpha_i^*\in(C/l,\delta C/l),\alpha_i=0$ 时：

$$b=y_k-\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)K(x_i,x_k)-\varepsilon_u$$

5 结 论

根据粗糙集理论和最优化理论,利用固定对称边界粗糙 ε -不敏感损失函数,得到 ε -不敏感管的粗糙边界,称之为粗糙 ε -不敏感管,进而构造固定对称边界粗糙 ε -支持向量回归模型(RFSM ε-SVR). 当 $\varepsilon_2=\varepsilon_1=\varepsilon$ 时,RFSM ε-SVR 即为 ε-SVR. 利用固定非对称边界粗糙 ε -不敏感损失函数,得到 $\varepsilon^u-\varepsilon^d$ -不敏感管的粗糙边界,称之为粗糙 $\varepsilon^u-\varepsilon^d$ -不敏感管,进而构造固定非对称边界粗糙 ε -支持向量回归模型(RFAM ε-SVR). 当 $\varepsilon^{uu}=\varepsilon^u$, $\varepsilon^{dd}=\varepsilon^d$ 时,RFAM ε-SVR 即为 RFSM ε-SVR.

类似的,可以利用上述方法得到固定对称边界粗糙 ε -支持向量分类与固定非对称边界粗糙 ε -支持向量分类模型.

References

[1] Vapnik V. The nature of statistical learning theory. New York:Springer,1995,188.

[2] Cortes C, Vapnik V. Support vector networks. Machine Learning,1995,20:273～297.

[3] Bai J W,Wang W J,Guo H S. A novel support vector machine active learning strategy. Journal of Nanjing University(Natural Sciences),2012,48(2):182～189.(白龙飞,王文剑,郭虎升. 一种新的支持向量机主动学习策略. 南京大学学报(自然科学),2012,48(2):182～189).

[4] Deng N Y,Tian Y J. New method of data mining: support vector machine. Beijing:Science Press,2006,408.(邓乃杨、田英杰. 数据挖掘中的新方法:支持向量机. 北京:科学出版社,2006,408).

[5] A. Smola B S. A tutorial on support vector regression. Statistics and Computing, 2004, 14(3):199～222.

[6] Yang H,Chan L,King I. Support vector machine regression for volatile stock market prediction. Yin H,Allinson N,Freeman R,*et al*. Intelligent Data Engineering and Automated Learning. Springer,Lecture Notes in Computer Science (LNCS),2002,2412:391～396.

[7] Yang H. Margin variations in support vector regression for the stock market prediction. Phil M. Dissertation,Department of Computer Science and Engineering. The Chinese University of Hong Kong. http://www.svms.org/finance/Yang2003.6.pdf, 2013－06.

[8] Pawlak Z. Rough sets. International Journal of Computer and Information Science. 1982,11(5): 341～356.

[9] Pawlak Z. Rough sets: Theoretical aspects of reasoning about data. Dordrecht: Kluwer Academic Publishers,1991,229.

[10] Pawlak Z,Skoworn A. Rudiments of rough sets. Information Science,2007,177:3～27.

[11] Zhang W X,Liang Y,Wei W Z. Information systems and knowledge discovery. Beijing: Science Press, 2003,244.(张文修,梁怡,吴伟志. 信息系统与知识发现. 北京:科学出版社,2003,244).

[12] Hu Q H, Yu D R. Application of rough set calculation. Beijing:Science Press,2012,184.(胡清华,于达仁. 应用粗糙集计算. 北京:科学出版社,2012,184).

[13] Wang G Y,Yao Y Y,Yu H. A survey on rough set theory and applications. Chinese Journal of Computers,2009,32(7):1229～1247.(王国胤,姚一豫,于 洪. 粗糙集理论与应用研究综述. 计算机学报,2009,32(7):1229～1247).

[14] Zhao Y P, Sun J G. Rough ν-support vector regression. Expert Systems with Applications, 2009,36(6):9793～9798.

[15] Lingras P, Butz C J. Rough support vector regression. European Journal of Operational Research,2010,206(2):445～455.