DOI:10.3969/j.issn.1673-4785.201211026
网络出版地址:http://www.cnki.net/kcms/detail/23.1538.TP.20130515.0842.003.html

支持向量回归机多项式光滑函数的逼近精度

冯能山,熊金志
(东莞理工学院 计算机学院,广东 东莞 523808)

摘 要:为了解决支持向量回归机多项式光滑函数的逼近精度问题,根据该类光滑函数的复杂性,提出五步求的基本思路:首先把多项式光滑函数的逼近精度问题表示为一个求逼近函数的最大值问题,接着证明这个逼近函数是一个对称函数,然后分别求出逼近函数在 $[0,\varepsilon]$ 和 $(\varepsilon,+\infty)$ 上的最大值,最后对这2个最大值进行比较,得出光滑函数的逼近精度.通过实例计算,结果证明了该方法的有效性和正确性,解决了无穷多个多项式光滑函数的逼近精度问题,为光滑支持向量回归机提供了基本的理论支持.

关键词:支持向量回归机;多项式光滑函数;逼近精度;对称函数
中图分类号: TP18 **文献标志码**:A **文章编号**:1673-4785(2013)03-0266-05

中文引用格式:冯能山,熊金志.支持向量回归机多项式光滑函数的逼近精度[J].智能系统学报, 2013, 8(3): 266-270.
英文引用格式:**FENG Nengshan, XIONG Jinzhi. The approximation accuracies of polynomial smoothing functions for support vector regression[J].CAAI Transactions on Intelligent Systems**, 2013, 8(3): 266-270.

The approximation accuracies of polynomial smoothing functions for support vector regression

FENG Nengshan, XIONG Jinzhi
(Computer College, Dongguan University of Technology, Dongguan 523808, China)

Abstract: In order to solve the problem of polynomial smoothing function's approximation accuracies for support vector regression, the researcher proposes to solve the problem utilizing a five step approach, according to the complexity of smooth functions. The first step, examines the problem of the approximation accuracy in the polynomial smoothing function represented by solving the maximum value of an approximation function; second, the function was verified to be a symmetric function; third, the maximum values of the approximation function were derived respectively at the intervals $[0,\varepsilon]$ and $(\varepsilon,+\infty)$; fourth, the two maximum values were compared, and finally the approximation accuracy was obtained. Through the calculation with examples, the correctness and effectiveness of the method was validated, and the approximation accuracy problem of the infinite polynomial smoothing functions was systematically solved in this paper, which offers basic theoretical support for smooth support vector regression.
Keywords: support vector regression; polynomial smoothing function; approximation accuracies; symmetric function

光滑函数在支持向量机中得到了成功应用,特别是对分类问题获得了良好的效果[1-3].对回归问题,国内外学者进行了深入研究,如Lee等[4]找到一类 p_{ε}^{2}-函数,并作为光滑函数对回归问题中的目标函数进行光滑处理,提出 ε-不敏感的光滑支持向量回归机模型.其实验结果表明,这种光滑支持向量回归机在一定程度上改善了回归效果.为使支持向量回归机的效果得到进一步改善,针对原支持向量回归机中 ε-不敏感损失函数的平方项不光滑的问题,2008年笔者又提出了一类多项式光滑函数[5].这类多项式光滑函数的形式复杂,具有以下3个显著特点:1)这种多项式光滑函数有无穷多个;2)每个多项式光滑函数都是某个多项式函数的复合函数;3)每个多项式光滑函数都是三段函数[5-6].按照光滑技术的基本思路,用多项式光滑函数对原支持向量回

收稿日期:2012-11-02. **网络出版日期**:2013-05-15.
基金项目:广东省自然科学基金资助项目(9151170003000017);东莞市科技计划资助项目(2012108102027).
通信作者:熊金志. E-mail:dgxiongjinzhi@126.com.

归机模型进行光滑处理，可以得到支持向量机的光滑模型，从而提高回归速度，而这个过程必须用到多项式光滑函数的逼近精度[4]．因此，多项式光滑函数的逼近精度是光滑支持向量回归机必须解决的一个重要问题．该问题是否存在一个统一的求解方法，多年来一直是一个尚待解决的问题[5,7]．对此，本文提出五步求的基本思路，并用二分法试图系统解决这个问题．

1　支持向量回归机及其光滑函数

1.1　回归问题

对数据集 $S=\{(\boldsymbol{x}_1,y_1),(\boldsymbol{x}_2,y_2),\cdots,(\boldsymbol{x}_m,y_m)\}\subseteq \mathbf{R}^n\times\mathbf{R}$，令矩阵 $\boldsymbol{A}=[\boldsymbol{x}_1\ \boldsymbol{x}_2\cdots\ \boldsymbol{x}_m]$，$\boldsymbol{x}_i$ 是一个 n 维向量，每个 $\boldsymbol{x}_i$ 对应有一个观测值 y_i．显然 $\boldsymbol{A}\in\mathbf{R}^{n\times m}$，$S=\{(\boldsymbol{A}_i,y_i)\mid\boldsymbol{A}_i\in\mathbf{R}^n,y_i\in\mathbf{R}\}$，$i=1,2,\cdots,m$，其中 $\boldsymbol{A}_i$ 是矩阵 $\boldsymbol{A}$ 的第 i 个向量．回归的目的是利用给定的数据集 S，训练出一个回归函数 $f(\boldsymbol{x})$，使 $f(\boldsymbol{x})$ 能对新的输入 $\boldsymbol{x}$ 较准确地预测出输出 y，这就是回归问题．采用的标准是 ε-不敏感损失函数：$|y-f(\boldsymbol{x})|_\varepsilon=\max\{0,|y-f(\boldsymbol{x})|-\varepsilon\}$．对于线性回归的情形，$f(\boldsymbol{x})=\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b$，其中 $\boldsymbol{w}\in\mathbf{R}^n$ 是一个待定向量，b 是一个待定常量[8-9]．

1.2　支持向量回归机

上述回归问题可表示为无约束最优化问题[5]：

$$\min_{(\boldsymbol{w},b)\in\mathbf{R}^{n+1}}\frac{1}{2}(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}+b^2)+\frac{C}{2}\sum_{i=1}^{m}|\boldsymbol{A}_i\boldsymbol{w}+b-y_i|_\varepsilon^2. \tag{1}$$

式中：C 为一个大于 0 的惩罚参数．式(1)称为无约束的支持向量回归机模型．由于式(1)中 $|\cdot|_\varepsilon^2$ 不光滑，导致此模型的目标函数不光滑．因此该支持向量回归机不能用快速的 Newton-Armijo 算法进行求解，使得模型的精度和效率受到影响．为此，用一类 d 阶光滑的多项式函数 $p_{d\varepsilon}^2(x,k)$ 作为光滑函数，逼近上述支持向量回归机模型的 $|\cdot|_\varepsilon^2$，得到

$$\min_{(\boldsymbol{w},b)\in\mathbf{R}^{n+1}}\varphi_{d\varepsilon,k}(\boldsymbol{w},b):=\min_{(\boldsymbol{w},b)\in\mathbf{R}^{n+1}}\frac{1}{2}(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}+b^2)+\frac{C}{2}\sum_{i=1}^{m}p_{d\varepsilon}^2(\boldsymbol{A}_i\boldsymbol{w}+b-y_i,k) \tag{2}$$

因此该支持向量回归机是光滑的，可用快速的 Newton-Armijo 算法进行求解，从而提高模型的精度和效率．称式(2)为 d 阶多项式光滑的支持向量回归机，多项式函数 $p_{d\varepsilon}^2(x,k)$ 称为支持向量回归机的多项式光滑函数[5]．

容易看出，在把原支持向量回归机模型变换为光滑的支持向量回归机的过程中，光滑函数起了关键作用，因此光滑函数与 $|\cdot|_\varepsilon^2$ 的逼近精度是光滑支持向量回归机的一个重要理论问题[5]，也是本文要研究的问题．为便于研究，下面列出支持向量回归机的多项式光滑函数．

1.3　光滑函数

文献[5]已对 ε-不敏感损失函数 $|x|_\varepsilon$ 及其平方函数 $|x|_\varepsilon^2$ 的多项式光滑函数进行了定义和说明，在此不再赘述．为便于分析，先给出支持向量回归机的多项式光滑函数．

引理 1[5]　设 $p_d(x,k)$ 是正号函数 x_+ 的 d 阶多项式光滑函数，令函数 $p_{d\varepsilon}^2(x,k)$ 满足：

$$p_{d\varepsilon}^2(x,k)=p_d(x-\varepsilon,k)^2+p_d(-x-\varepsilon,k)^2, \tag{3}$$

则 $p_{d\varepsilon}^2(x,k)$ 是 $|x|_\varepsilon^2$ 的 d 阶多项式光滑函数．

可见，$|x|_\varepsilon^2$ 的光滑函数 $p_{d\varepsilon}^2(x,k)$ 是正号函数 x_+ 的光滑逼近 $p_d(x,k)$ 的复合函数，而 $p_d(x,k)$ 由文献[6]知有无穷多个，皆可由一个递推公式求出．

因此，由引理 1 可求得一类的 $|x|_\varepsilon^2$ 具有 d 阶光滑的多项式光滑函数：

$$\{p_{d\varepsilon}^2(x,k),d=1,2,\cdots\}. \tag{4}$$

显然，这种光滑函数有无穷多个，下面以 $d=1,2,3,4,5$ 为例分别求出 $|x|_\varepsilon^2$ 的前 5 个光滑函数．

1) 当 $d=1$ 时，$|x|_\varepsilon^2$ 的光滑函数为

$$p_{1\varepsilon}^2(x,k)=p_1(x-\varepsilon,k)^2+p_1(-x-\varepsilon,k)^2. \tag{5}$$

式中：

$$p_1(x,k)=\begin{cases}x, & x\geqslant\dfrac{1}{k};\\ \dfrac{k}{4}x^2+\dfrac{1}{2}x+\dfrac{1}{4k}, & -\dfrac{1}{k}<x<\dfrac{1}{k};\\ 0, & x\leqslant-\dfrac{1}{k}.\end{cases}$$

2) 当 $d=2$ 时，$|x|_\varepsilon^2$ 的光滑函数为

$$p_{2\varepsilon}^2(x,k)=p_2(x-\varepsilon,k)^2+p_2(-x-\varepsilon,k)^2. \tag{6}$$

式中：

$$p_2(x,k)=\begin{cases}x, & x\geqslant\dfrac{1}{k};\\ -\dfrac{1}{16k}(kx+1)^3(kx-3), & -\dfrac{1}{k}<x<\dfrac{1}{k};\\ 0, & x\leqslant-\dfrac{1}{k}.\end{cases}$$

3) 当 $d=3$ 时，$|x|_\varepsilon^2$ 的光滑函数为

$$p_{3\varepsilon}^2(x,k)=p_3(x-\varepsilon,k)^2+p_3(-x-\varepsilon,k)^2.$$

式中：

$$p_3(x,k)=$$
$$\begin{cases} x, & x \geqslant \dfrac{1}{k}; \\ \dfrac{1}{32k}(kx+1)^4(k^2x^2-4kx+5), & -\dfrac{1}{k}<x<\dfrac{1}{k}; \\ 0, & x \leqslant -\dfrac{1}{k}. \end{cases}$$

4)当 $d=4$ 时,$|x|_\varepsilon^2$ 的光滑函数为

$$p_{4\varepsilon}^2(x,k)=p_4(x-\varepsilon,k)^2+p_4(-x-\varepsilon,k)^2.$$

式中:

$$p_4(x,k)=$$
$$\begin{cases} x, & x \geqslant \dfrac{1}{k}; \\ -\dfrac{1}{256k}(kx+1)^5(5k^3x^3-25k^2x^2+47kx-35), & -\dfrac{1}{k}<x<\dfrac{1}{k}; \\ 0, & x \leqslant -\dfrac{1}{k}. \end{cases}$$

5)当 $d=5$ 时,$|x|_\varepsilon^2$ 的光滑函数为

$$p_{5\varepsilon}^2(x,k)=p_5(x-\varepsilon,k)^2+p_5(-x-\varepsilon,k)^2.$$

式中:

$$p_5(x,k)=$$
$$\begin{cases} x, & x \geqslant \dfrac{1}{k}; \\ \dfrac{1}{512k}(kx+1)^6(k^4x^4-42k^3x^3+102k^2x^2-122kx+63), & -\dfrac{1}{k}<x<\dfrac{1}{k}; \\ 0, & x \leqslant -\dfrac{1}{k}. \end{cases}$$

2 多项式光滑函数的逼近精度及其求解步骤

2.1 任意阶多项式光滑函数的逼近精度

支持向量回归机的任意阶多项式光滑函数 $p_{d\varepsilon}^2(x,k)$,其具体形式皆可由式(4)推出,这种光滑函数显然有无穷多个.这无穷多个光滑函数逼近 ε-不敏感损失函数的平方项 $|x|_\varepsilon^2$ 的精度问题,显然皆可描述为求 $p_{d\varepsilon}^2(x,k)-|x|_\varepsilon^2$ 的最大值问题.记

$$F_{d\varepsilon}(x,k)=p_{d\varepsilon}^2(x,k)-|x|_\varepsilon^2, \tag{7}$$

称 $F_{d\varepsilon}(x,k)$ 为光滑函数 $p_{d\varepsilon}^2(x,k)$ 的误差函数,这种误差函数显然也有无穷多个,且与光滑函数 $p_{d\varepsilon}^2(x,k)$ 一一对应.

定理 1 误差函数 $F_{d\varepsilon}(x,k)$ 由式(7)给出,则

1)$F_{d\varepsilon}(x,k)$ 是 x 的对称函数;

2)$F_{d\varepsilon}(x,k)\leqslant r_d/k^2$.

其中 r_d 为一个与 $p_{d\varepsilon}^2(x,k)$ 的光滑阶数 d 相关的误差系数.

证明

1)将 $-x$ 代入式(7),结合式(5),可得

$$F_{d\varepsilon}(-x,k)=p_{d\varepsilon}^2(-x,k)-|-x|_\varepsilon^2=p_{d\varepsilon}^2(x,k)-|x|_\varepsilon^2=F_{d\varepsilon}(x,k),$$

可见 $F_{d\varepsilon}(x,k)$ 是对称函数.因此在整个 x 轴上求 $F_{d\varepsilon}(x,k)$ 的最大值问题可转换成在 x 轴的右半轴上求 $F_{d\varepsilon}(x,k)$ 的最大值问题.

2)把 x 轴的右半轴分成 2 个区间进行证明:

①当 $0\leqslant x\leqslant\varepsilon$ 时,显然 $|x|_\varepsilon=0$.因此

$$F_{d\varepsilon}(x,k)=p_{d\varepsilon}^2(x,k)-|x|_\varepsilon^2=p_d(x-\varepsilon,k)^2+p_d(-x-\varepsilon,k)^2.$$

此时显然有 $x-\varepsilon\leqslant0$ 和 $-x-\varepsilon\leqslant0$,因 $p_d(x,k)^2$ 是单调增函数[6],可得

$$F_{d\varepsilon}(x,k)\leqslant 2p_d(0,k)^2. \tag{8}$$

②当 $x>\varepsilon$ 时,$-x+\varepsilon<0$,显然有 $-x-\varepsilon<-x+\varepsilon<0$,因此 $(-x-\varepsilon)_+=0$.因 $p_d(x,k)$ 是单调增函数[6],可知 $p_d(-x-\varepsilon,k)\leqslant p_d(0,k)$.由二分法可算出 $p_d(x,k)^2$ 与正号函数 x_+ 的逼近精度:$p_d(x,k)^2-x_+^2\leqslant s_d/k^2$,其中 s_d 为一个与 $p_{d\varepsilon}^2(x,k)$ 的光滑阶数 d 相关的系数.由文献[4]知,$|x|_\varepsilon^2=(x-\varepsilon)_+^2+(-x-\varepsilon)_+^2$,因此

$$F_{d\varepsilon}(x,k)=p_{d\varepsilon}^2(x,k)-|x|_\varepsilon^2=p_d(x-\varepsilon,k)^2-(x-\varepsilon)_+^2+p_d(-x-\varepsilon,k)^2,$$

即

$$F_{d\varepsilon}(x,k)\leqslant s_d/k^2+p_d(0,k)^2. \tag{9}$$

综合①、②,取式(8)和(9)中右端的最大值,令

$$r_d=\max\{2k^2p_d(0,k)^2,s_d+k^2p_d(0,k)^2\},$$

可得

$$F_{d\varepsilon}(x,k)\leqslant r_d/k^2,x\in\mathbf{R}. \tag{10}$$

证毕.

2.2 求逼近精度的一般步骤

这无穷多个多项式光滑函数有一个共同特征:都是正号函数 x_+ 的光滑逼近 $p_d(x,k)$ 形如式(5)的复合函数,而 $p_d(x,k)$ 皆可由文献[6]的递推公式求出.上文求出了任意阶光滑函数的逼近精度的表达式(10),下面对这些求解过程进行归纳和总结,可发现求解的一般规律.根据这个规律,提出五步求的基本思路,即求逼近精度的一般步骤.

1)把求光滑函数 $p_{d\varepsilon}^2(x,k)$ 的逼近精度问题表示为一个在整个 x 轴上求误差函数 $F_{d\varepsilon}(x,k)$ 的最大值问题:

$$F_{d\varepsilon}(x,k)=p_{d\varepsilon}^2(x,k)-|x|_\varepsilon^2;$$

2)证明 $F_{d\varepsilon}(x,k)$ 是对称函数,把在整个 x 轴上求 $F_{d\varepsilon}(x,k)$ 的最大值问题转换成在 x 轴的右半轴上求 $F_{d\varepsilon}(x,k)$ 的最大值问题;

3)求当 $0\leqslant x\leqslant\varepsilon$ 时误差函数 $F_{d\varepsilon}(x,k)$ 的最大值;

4)求当 $x>\varepsilon$ 时误差函数 $F_{d\varepsilon}(x,k)$ 的最大值;

5)对这 2 个最大值进行比较,最大者即为光滑函数的逼近精度.

2.3　一阶多项式光滑函数的逼近精度

仿照定理 1 及求任意阶多项式光滑函数的逼近精度的基本思想和步骤,下面具体求出一阶多项式光滑函数的逼近精度.

推论 1　一阶多项式光滑函数的逼近精度为 $0.153\ 4/k^2$.

证明

1)把求光滑函数 $p_{1\varepsilon}^2(x,k)$ 的逼近精度问题表示为一个在整个 x 轴上求误差函数 $F_{1\varepsilon}(x,k)$ 的最大值问题.支持向量回归机的一阶多项式光滑函数的形式如式(5).把多项式光滑函数应用于支持向量回归机,首先必须解决多项式光滑函数逼近 ε-不敏感损失函数的平方项 $|x|_\varepsilon^2$ 的精度问题.该问题可描述为一个求 $p_{1\varepsilon}^2(x,k)-|x|_\varepsilon^2$ 的最大值问题,记

$$F_{1\varepsilon}(x,k)=p_{1\varepsilon}^2(x,k)-|x|_\varepsilon^2,\qquad(11)$$

称 $F_{1\varepsilon}(x,k)$ 为光滑函数 $p_{1\varepsilon}^2(x,k)$ 的误差函数.因此,求多项式光滑函数逼近 ε-不敏感损失函数的平方项 $|x|_\varepsilon^2$ 的精度问题,就可转换成求 $F_{1\varepsilon}(x,k)$ 的最大值问题.

2)证明 $F_{1\varepsilon}(x,k)$ 是对称函数,把在整个 x 轴上求 $F_{1\varepsilon}(x,k)$ 的最大值问题转换成在 x 轴的右半轴上求 $F_{1\varepsilon}(x,k)$ 的最大值问题.将 $-x$ 代入式(11),结合式(5),可得

$$F_{1\varepsilon}(-x,k)=p_{1\varepsilon}^2(-x,k)-|-x|_\varepsilon^2=p_{1\varepsilon}^2(x,k)-|x|_\varepsilon^2=F_{1\varepsilon}(x,k).$$

可见 $F_{1\varepsilon}(x,k)$ 是对称函数.因此在整个 x 轴上求 $F_{1\varepsilon}(x,k)$ 的最大值问题就转换成了在 x 轴的右半轴上求 $F_{1\varepsilon}(x,k)$ 的最大值问题.

3)求当 $0\leqslant x\leqslant\varepsilon$ 时误差函数 $F_{1\varepsilon}(x,k)$ 的最大值.此时显然 $|x|_\varepsilon=0$.因此

$$F_{1\varepsilon}(x,k)=p_{1\varepsilon}^2(x,k)-|x|_\varepsilon^2=p_1(x-\varepsilon,k)^2+p_1(-x-\varepsilon,k)^2,$$

此时显然有 $x-\varepsilon\leqslant0$ 和 $-x-\varepsilon\leqslant0$,因 $p_1(x,k)^2$ 是单调增函数[3,6],可得

$$F_{1\varepsilon}(x,k)\leqslant 2p_1(0,k)^2=2(\frac{1}{4k})^2=0.125/k^2.$$

4)求当 $x>\varepsilon$ 时误差函数 $F_{1\varepsilon}(x,k)$ 的最大值.此时 $-x+\varepsilon<0$,显然有 $-x-\varepsilon<-x+\varepsilon<0$,因此 $(-x-\varepsilon)_+=0$.因 $p_1(x,k)$ 是单调增函数[3,6],可知 $p_1(-x-\varepsilon,k)\leqslant p_1(0,k)$.由二分法[10]可算出 $p_1(x,k)$ 与正号函数 x_+ 的逼近精度[3]:$p_1(x,k)^2-x_+^2\leqslant\frac{1}{11k^2}$.由文献[5]知,$|x|_\varepsilon^2=(x-\varepsilon)_+^2+(-x-\varepsilon)_+^2$,因此

$$F_{1\varepsilon}(x,k)=p_{1\varepsilon}^2(x,k)-|x|_\varepsilon^2=p_1(x-\varepsilon,k)^2-(x-\varepsilon)_+^2+p_1(-x-\varepsilon,k)^2\leqslant\frac{1}{11k^2}+p_1(0,k)^2=0.153\ 4/k^2.$$

5)对这 2 个最大值进行比较,最大者即为光滑函数的逼近精度.综合 4)和 5)的结果易知,$r_1=0.153\ 4$,$F_{1\varepsilon}(x,k)\leqslant0.153\ 4/k^2$. $0.153\ 4/k^2$ 即为一阶多项式光滑函数的逼近精度.

证毕.

2.4　二阶至五阶多项式光滑函数的逼近精度

支持向量回归机的二阶多项式光滑函数的形式如式(6).仿照上面求一阶多项式光滑函数的逼近精度的基本思想和步骤,可把二阶多项式光滑函数逼近 ε-不敏感损失函数的平方项 $|x|_\varepsilon^2$ 的精度问题描述为求 $p_{2\varepsilon}^2(x,k)-|x|_\varepsilon^2$ 的最大值问题.记

$$F_{2\varepsilon}(x,k)=p_{2\varepsilon}^2(x,k)-|x|_\varepsilon^2,$$

称 $F_{2\varepsilon}(x,k)$ 为光滑函数 $p_{2\varepsilon}^2(x,k)$ 的误差函数.

推论 2　二阶多项式光滑函数的逼近精度为 $0.087\ 79/k^2$.

证明过程与定理 1 及推论 1 类似,在此省略.

以上内容求出了一阶和二阶光滑函数的逼近精度,同理以同样的步骤,可求出三阶、四阶和五阶光滑函数的逼近精度,求解过程省略.表 1 列出一阶至五阶多项式光滑函数的逼近精度.

表 1　多项式光滑函数的逼近精度

Table 1　The approximation accuracies of polynomial smoothing functions

光滑函数	光滑阶数 d	误差系数 r_d	逼近精度
$p_{1\varepsilon}^2(x,k)$	1	0.153 40	$0.153\ 40/k^2$
$p_{2\varepsilon}^2(x,k)$	2	0.087 78	$0.087\ 78/k^2$
$p_{3\varepsilon}^2(x,k)$	3	0.035 78	$0.035\ 78/k^2$
$p_{4\varepsilon}^2(x,k)$	4	0.027 63	$0.027\ 63/k^2$
$p_{5\varepsilon}^2(x,k)$	5	0.024 33	$0.024\ 33/k^2$

因此,依照上述五步求的基本思路和步骤,可求出支持向量回归机无穷多个光滑函数的逼近精度.

3　结束语

为求支持向量回归机光滑函数的逼近精度,首

先把求光滑函数的逼近精度问题转化成求误差函数的最大值问题;然后证明该误差函数是对称函数,把这个最大值问题转换成在 x 轴的右半轴上求最大值的问题;而后分别求当 $0 \leqslant x \leqslant \varepsilon$ 时和 $x > \varepsilon$ 时误差函数的最大值;最后对这 2 个最大值进行比较,求出光滑函数的逼近精度.以多个多项式光滑函数为例,解决了它们的逼近精度问题.在此基础上,总结出求此类精度问题的一般规律,对无穷多个光滑函数的精度问题提出了五步求的基本思路.研究结果表明:对支持向量回归机的无穷多个多项式光滑函数,都可按照本文的思路和步骤,分 5 个步骤用二分法解决其逼近精度问题.该方法成功解决了支持向量回归机中一个尚待解决的问题,即支持向量回归机无穷多个光滑函数的逼近精度问题,为光滑支持向量回归机的进一步研究工作提供了基本的理论支持.

参考文献:

[1]LEE Y J, MANGASARIAN O L. SSVM: a smooth support vector machine for classification[J]. Computational Optimization and Applications, 2001, 22(1): 5-21.

[2]CHANG C C, CHIEN L J, LEE Y J. A novel framework for multi-class classification via ternary smooth support vector machine[J]. Pattern Recognition, 2011, 44(6): 1235-1244.

[3]钱晓山,阳春华.基于 GEP 的最小二乘支持向量机模型参数选择[J].智能系统学报, 2012, 7(3): 225-229.
QIAN Xiaoshan, YANG Chunhua. A parameter selection method of a least squares support vector machine based on gene expression programming[J]. CAAI Transactions on Intelligent Systems, 2012, 7(3): 225-229.

[4]LEE Y J, HSIEH W F, HUANG C M. ε-SSVR: a smooth support vector machine for ε-insensitive regression[J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(5): 5-22.

[5]熊金志,胡金莲,袁华强,等.支持向量回归机的光滑函数研究[J].模式识别与人工智能, 2008, 21(3): 273-279.
XIONG Jinzhi, HU Jinlian, YUAN Huaqiang, et al. Smoothing functions for support vector regressions[J]. Pattern Recognition and Artificial Intelligence, 2008, 21(3): 273-279.

[6]熊金志,胡金莲,袁华强,等.一类光滑支持向量机新函数的研究[J].电子学报, 2007, 35(2): 366-370.
XIONG Jinzhi, HU Jinlian, YUAN Huaqiang, et al. Research on a new class of functions for smoothing support vector machines[J]. Acta Electronica Sinica, 2007, 35(2): 366-370.

[7]熊金志,徐建敏,袁华强.多项式光滑的支持向量回归机一般模型的收敛性研究[J].计算机研究与发展, 2011, 48(3): 464-470.
XIONG Jinzhi, XU Jianmin, YUAN Huaqiang. Convergenceness of a general formulation for polynomial smooth support vector regressions[J]. Journal of Computer Research and Development, 2011, 48(3): 464-470.

[8]李国正,王猛,曾华军.支持向量机导论[M].北京:电子工业出版社, 2004: 89-99.

[9]邓乃扬,田英杰.数据挖掘的新方法—支持向量机[M].北京:科学出版社, 2008: 111-122.

[10]李庆扬,王能超,易大义.数值分析[M].5 版.北京:清华大学出版社, 2011: 125-136.

作者简介:

冯能山,男,1972 年生,讲师,博士研究生,主要研究方向为人工智能、软件工程.

熊金志,男,1964 年生,教授,中国计算机学会高级会员,东莞市科技创新团队核心成员,主要研究方向为人工智能.作为核心成员承担国家自然科学基金项目 2 项,主持广东省科研项目 3 项、东莞市科研项目 1 项.获广东省科技进步二等奖和三等奖各 1 项、东莞市科技进步一等奖 2 项.发表学术论文 40 余篇,被 SCI、EI、ISTP 检索 10 余篇.