

再生核支持向量机在非线性系统中的应用

胡丹^{1,2}, 肖建¹, 车畅²

(1. 西南交通大学电气工程学院 成都 610031; 2. 西华大学机械与自动化学院 成都 610039)

【摘要】为了提高非线性系统辨识的精度, 提出用Walsh函数作为空间 V_0 的尺度函数, 构造出 $L^2(\mathbf{R})$ 空间的正交规范序列。结合小波多分辨分析, 将Hilbert空间分为一系列子空间, 并由可分Hilbert空间与 $L^2(\mathbf{R})$ 的等价性, 利用内积同构的线性算子, 可以把 V_0 子空间的尺度函数折算为Hilbert空间的子空间 \tilde{V}_0 的尺度函数, 构造出新的Walsh序列再生核。通过仿真实验, 与传统的RBF核函数、高斯核函数等比较, 该尺度再生核函数具有更高的辨识精度, 较少支持向量数目, 充分体现了支持向量机较好的推广性能。

关键词 Hilbert空间; 回归; 再生核; 尺度核; 支持向量机
中图分类号 TM301.2 **文献标识码** A

Reproducing Kernel Function Based on Walsh Series

HU Dan^{1,2}, XIAO Jian¹, CHE Chang²

(1. School of Electric Engineering of Southwest Jiaotong University Chengdu 610031;
2. School of Mechanical Engineering and Automation, Xihua University Chengdu 610039)

Abstract A new reproducing kernel function of least square support vector machines based on Walsh series is presented in this paper. The reproducing kernel is constructed in reproducing kernel Hilbert space (RKHS). Because the Hilbert space and the square integrable space are isomorphic, according to the wavelet multi-resolution analysis, Walsh consequence can be seen a set of orthogonal basis to construct the reproducing kernel. The simulation results are discussed to illustrate the proposed method.

Key words Hilbert space; regression; reproducing kernel; scaling kernel; support vector machine

支持向量回归的主要思想是通过“核映射”把输入样本空间映射到高维的特征空间, 在特征空间中进行线性回归来实现非线性处理。支持向量回归具有一些很好的特性, 如问题转换为凸优化问题, 支持向量的稀疏性, 良好的泛化能力等。在支持向量机的框架中融合小波技术, 可以充分发挥两者的优势, 近年来引起了许多研究者的兴趣。文献[1]将小波支持向量机用于非线性组合预测, 文献[2]将最小二乘方法与小波支持向量机相结合用于非线性系统辨识。文献[2-3]将小波变换与支持向量机相结合用于人脸辨识。本文在可分Hilbert空间与 $L^2(\mathbf{R})$ 等价的基础上, 尝试利用Walsh序列来构造一种新的多尺度核函数, 并将其用于非线性系统的辨识中。

1 可分Hilbert空间与 $L^2(\mathbf{R})$ 的等价性

由小波多分辨分析理论知, 从 V_0 的Reisz基 $\{g(x-k)\}_{k \in \mathbb{Z}}$ 出发, 经过规范正交化构造出 V_0 的规范

正交基 $\{\varphi(x-k)\}_{k \in \mathbb{Z}}$, 称 $\varphi(x)$ 为尺度函数(V_j 也称为尺度空间)。由Riesz表现定理, $\{V_j\}_{j \in \mathbb{Z}}$ 是 $L^2(\mathbf{R})$ 的一串子空间序列, 只要能构造出一组规范正交基即可。在文献[3]中已经证明了 $\{\varphi(x-k)\}_{k \in \mathbb{Z}}$ 构成规范正交系的充要条件, 并且得出: 若 $\{V_j\}_{j \in \mathbb{Z}}$ 为 $L^2(\mathbf{R})$ 的一个MRA, $\{\varphi(x-k)\}_{k \in \mathbb{Z}}$ 为 V_0 的规范正交基, 则对任意 $j \in \mathbb{Z}$, $\varphi_{j,k} = \{2^{j/2} \varphi(2^j x - k)\}$ 为 V_j 的一个规范正交基, 即 $V_j = \text{span}\{2^{j/2} \varphi(2^j x - k)\}_{k \in \mathbb{Z}}$ 。但Hilbert空间再生核是在Hilbert空间中定义和论述, 下面的定理1证实了把定义在 $L^2(\mathbf{R})$ 的一串子空间序列上的尺度函数的构造方法应用到Hilbert空间上。

定理 1(可分Hilbert空间与 $L^2(\mathbf{R})$ 的等价性) 设 H 是任意可分Hilbert空间, 如果它和 $L^2(\mathbf{R})$ 定义在一个数域上 \square 上, 则 H 与 $L^2(\mathbf{R})$ 等价。

证明 因 H 可分, 必有正交规范基 $\{e_n\}$ 在 H 是

收稿日期: 2006-11-09; 修回日期: 2007-04-03

基金项目: 四川省教育厅项目(2006ZD022); 四川省软件重点实验室项目(SCSL 05014、SCSL 06010)

作者简介: 胡丹(1970-), 男, 博士, 教授, 主要从事信息安全、分布式计算、并行计算方面的研究。

完全的。由Parseval关系式有:

$$\sum_{i=1}^n |\langle x, e_i \rangle|^2 = \|x\|^2 \quad \forall x \in H \quad (1)$$

定义映射 $T: H \rightarrow L^2$, 即:

$$Tx = \langle e_1, x \rangle, \langle e_2, x \rangle, \dots, \langle e_n, x \rangle \quad \forall x \in H \quad (2)$$

令 $\zeta = \langle e_i, x \rangle$, $\forall i \in N$, 则 $\zeta = (\zeta_1, \zeta_2, \dots)$ 满足:

$$\|\zeta\|^2 = \sum_{i=1}^{\infty} |\langle e_i, x \rangle|^2 < \infty \quad (3)$$

所以 $\zeta = Tx \in L^2$ 。其次, 对任意 $\alpha, \beta \in \square$, 任意 $x, y \in H$ 有:

$$T(\alpha x + \beta y) = \langle e_n, \alpha x + \beta y \rangle = \alpha \langle e_n, x \rangle + \beta \langle e_n, y \rangle = \alpha Tx + \beta Ty \quad (4)$$

所以 T 是线性算子。又因为Parseval关系式表明, 当 $x \neq y$ 时, $Tx = \zeta \neq \eta = Ty$, 即 T 是一对一的; 另外,

对任意 $\zeta \in L^2$, 令 $x_k = \sum_{n=1}^k \zeta_n e_n$, $\forall k \in N$, 则 $\{x_k\}$ 是 H 中的 Cauchy 序列, 由 H 的完备性知

$\lim_{k \rightarrow \infty} x_k = x \in H$, 而且 $\langle e_n, x \rangle = \lim_{k \rightarrow \infty} \langle e_n, \sum_{i=1}^k \zeta_i e_i \rangle = \zeta_n$,

$\forall n \in N$, 所以 $Tx = \zeta$, 即 T 是满秩的。由此知, T 是一一对应映射的。

最后, 对任意 $x, y \in H$ 有:

$$\begin{aligned} \langle Tx, Ty \rangle &= \sum_{n=1}^{\infty} \overline{\langle e_n, x \rangle} \langle e_n, y \rangle = \\ \lim_{k \rightarrow \infty} \left\langle \sum_{i=1}^k \langle e_n, x \rangle e_n, \sum_{m=1}^k \langle e_m, y \rangle e_m \right\rangle &= \\ \left\langle \lim_{k \rightarrow \infty} \sum_{i=1}^k \langle e_n, x \rangle e_n, \lim_{k \rightarrow \infty} \sum_{m=1}^k \langle e_m, y \rangle e_m \right\rangle &= \langle x, y \rangle \end{aligned} \quad (5)$$

所以 T 是内积同构的, 即 H 与 L^2 等价。证毕

推论 1 设 $\{V_j\}_{j \in \mathbb{Z}}$ 是构成 L^2 中的多尺度分析的闭子空间序列, 算子 T 作用于 $\{V_j\}_{j \in \mathbb{Z}}$ 上得到的空间记作 $\tilde{V}_{j \in \mathbb{Z}} = TV_{j \in \mathbb{Z}}$, 则闭子空间序列 $\{\tilde{V}_j\}_{j \in \mathbb{Z}}$ 满足关系:

$$(1) \dots \supset_{\infty \leftarrow j} \tilde{V}_j \supset \tilde{V}_{-1} \supset \tilde{V}_0;$$

$$(2) \dots \supset_{\infty \leftarrow j} \tilde{V}_j \supset \tilde{V}_{-1} \supset \tilde{V}_0。$$

推论 2 若 $\{\alpha_n\}_{n \in \mathbb{Z}}$ 是 L^2 中的标准正交基, 则 $\{T\alpha_n\}_{n \in \mathbb{Z}}$ 是 H 的标准正交基。

由推论1、推论2和上述理论得知, \tilde{V}_j 构成 H 的多尺度分析。

2 尺度核的支持向量机模型

选取 Walsh 函数作为 \tilde{V}_0 子空间的尺度函数, 在

区间 $[0, 1]$ 上的 Walsh 函数定义为:

$$\begin{cases} \phi^{(0)}(t) = \chi_0^{(0)}(t) & t \in [0, 1] \\ \phi^{(1)}(t) = \chi_0^{(1)}(t) & t \in [0, 1] \\ \phi^{(2)}(t) = \frac{1}{\sqrt{2}} [\chi_1^{(1)}(t) + \chi_1^{(2)}(t)] & t \in [0, 1] \\ \phi^{(3)}(t) = \frac{1}{\sqrt{2}} [\chi_1^{(1)}(t) - \chi_1^{(2)}(t)] & t \in [0, 1] \end{cases} \quad (6)$$

或写成:

$$\begin{bmatrix} \phi^{(1)}(t) \\ \phi^{(2)}(t) \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \chi_1^{(1)}(t) \\ \chi_1^{(2)}(t) \end{bmatrix} = W_2 \begin{bmatrix} \chi_1^{(1)}(t) \\ \chi_1^{(2)}(t) \end{bmatrix}$$

当 $n=3$ 时, 定义:

$$\begin{bmatrix} \phi^{(1)}(t) \\ \phi^{(2)}(t) \\ \phi^{(3)}(t) \\ \phi^{(4)}(t) \end{bmatrix} = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \chi_2^{(1)}(t) \\ \chi_2^{(2)}(t) \\ \chi_2^{(3)}(t) \\ \chi_2^{(4)}(t) \end{bmatrix} = W_3 \begin{bmatrix} \chi_2^{(1)}(t) \\ \chi_2^{(2)}(t) \\ \chi_2^{(3)}(t) \\ \chi_2^{(4)}(t) \end{bmatrix} \quad (7)$$

式中 W_2 和 W_3 都是正交对称阵, 称为 Walsh 阵。更一般的情况, 设: $\phi_n = [\phi_n^{(1)}, \phi_n^{(2)}, \dots, \phi_n^{(n-1)}]^T$, $\chi_{n-1} = [\chi_{n-1}^{(1)}, \chi_{n-1}^{(2)}, \dots, \chi_{n-1}^{(n-1)}]^T$, 则有 $\phi_n(t) = W_n \chi_{n-1}(t)$, $t \in [0, 1]$ 。其中, W_n 是 $2^{n-1} \times 2^{n-1}$ 正交对称阵, 即 Walsh 阵。此时 $\{\phi_n^{(i)}\}$ 是 Hilbert 空间的正交规范序列, 这是因为当 n 固定时, W_n 是一个规范的正交变换阵, 即令 $\omega_i^{(n)}$ 表示 W_n 的第 i 个行向量 ($i=1, 2, \dots, 2^{n-1}$) 满足:

$$\text{tr}\{\omega_i^{(n)} \omega_j^{(n)}\} = \begin{cases} 1, i=j \\ 0, i \neq j \end{cases} \quad i, j = 1, 2, \dots, 2^{n-1} \quad (8)$$

于是有:

$$\begin{aligned} \langle \phi_n^{(i)}, \phi_n^{(j)} \rangle &= \int_0^1 \phi_n^{(i)} \phi_n^{(j)} dt = \int_0^1 \chi_{n-1}^T(t) \omega_n^{(i)} \omega_n^{(j)T} \chi_{n-1} dt = \\ &= \int_0^1 \text{tr}\{\omega_i^{(n)} \omega_j^{(n)T} \chi_{n-1} \chi_{n-1}^T(t)\} dt = \\ &= \text{tr}\{\omega_i^{(n)} \omega_j^{(n)T} \int_0^1 \chi_{n-1}(t) \chi_{n-1}^T(t) dt\} = \\ &= \text{tr}\{\omega_i^{(n)} \omega_j^{(n)T}\} I_{2^{n-1}} = \text{tr}\{\omega_i^{(n)} \omega_j^{(n)T}\} \end{aligned} \quad (9)$$

所以有:

$$\langle \phi_n^{(i)}, \phi_n^{(j)} \rangle = \begin{cases} 1, i=j \\ 0, i \neq j \end{cases} \quad i, j = 1, 2, \dots, 2^{n-1} \quad (10)$$

这样, 由 $[\chi_{n-1}^{(1)}, \chi_{n-1}^{(2)}, \dots, \chi_{n-1}^{(2^{n-1})}]$ 的正交规范性决定了 $[\phi_n^{(1)}, \phi_n^{(2)}, \dots, \phi_n^{(2^{n-1})}]$ 的正交规范性; 其次当 $n \neq m$ 时, 因 $\chi_n^{(i)}$ 与 $\chi_{m-1}^{(j)}$ 正交, 从而决定了 $\phi_n^{(i)}$ 与 $\phi_m^{(j)}$ 正交。这样 $\{\phi_n^{(i)}\}$ 是 Hilbert 空间的正交规范序列。同时, 对任

意 n 固定, 因 W_n 满秩, 所以 $\text{span}\{\phi_n^{(i)}\} = \text{span}\{\chi_{n-1}^{(i)}\}$, 则由 $\{\chi_{n-1}^{(i)}\}$ 在 Hilbert 空间中的完备性决定了 $\phi_n^{(i)}$ 在 Hilbert 中的完备性。

借助小波多分辨分析的概念, $\{\tilde{V}_j\}_{j \in z}$ 的各个子空间均构成 RKHS, 由 Walsh 函数作为 \tilde{V}_0 子空间的尺度函数生成相应的再生核为:

$$K(t, s) = \sum_{j,k} \phi_n^{j,k}(s) \phi_n^{j,k}(t) \quad j, k \in z \quad (11)$$

定理 2 设 $\{\phi_n^{j,k}, j \in [j_{\min}, j_{\max}], k \in z\}$ 为 Hilbert 空间的正交基, $K(t, s) = \sum_{j,k} \phi_n^{j,k}(s) \phi_n^{j,k}(t)$ 定义在上 Hilbert, 那么 $\text{span}\{\phi_n^{j,k}, j \in [j_{\min}, j_{\max}], k \in z\}$ 赋予 Hilbert 内积构成一个 RKHS。

证明 由定理 2 知, $\text{span}\{\phi_n^{j,k}, j \in [j_{\min}, j_{\max}], k \in z\}$ 是 Hilbert 的闭子空间, 且赋予 Hilbert 内积, 故它是一个 Hilbert 空间。

(1) 验证 K 的正定性: 假定样本容量为 l , a_m 、 a_n 非负, 则:

$$\sum_{m,n=1}^l K(t_m, s_n) a_m a_n = \sum_{m,n=1}^l \phi_n^{j,k}(t_m) \phi_n^{j,k}(s_n) = \sum_{j,k} \left(\sum_{m=1}^l a_m \phi_n^{j,k}(s_m) \right)^2 > 0 \quad (12)$$

(2) 验证 K 的再生性:

$$\langle f(s), K(t, s) \rangle = \left\langle f(s), \sum_{j,k} \phi_n^{j,k}(s) \phi_n^{j,k}(t_m) \right\rangle \quad (13)$$

$$\sum_{j,k} \langle f(\cdot), \phi_n^{j,k}(\cdot) \rangle \phi_n^{j,k}(t) = f(t) \quad (14)$$

定理 2 说明 $\{\phi_n^{j,k}, j \in [j_{\min}, j_{\max}], k \in z\}$ 构成一个 Hilbert 基, 由这组基可以生成再生核 $K(t, s) = \sum_{j,k} \phi_n^{j,k}(s) \phi_n^{j,k}(t)$, 称这个核函数为尺度再生核 (SR-SVR), $K(t, s)$ 实际上是一个 Mercer 核。

3 非线性系统的仿真

下面利用前述方法构造的尺度函数式 (11) 和最小二乘支持向量机相结合来进行一维函数的逼近。

已知函数 $\text{sinc}(x)$ 在均匀分布点 $x_k = -10 + 20/100k, k = 0, 1, \dots, 100$ 上的函数值为 $y = \text{sinc}(x_k), k = 0, 1, \dots, 100$ 。设训练集:

$$T = \{(x_1, y_1), (x_2, y_2), \dots, (x_{100}, y_{100})\} \in (\mathcal{X}^l)^{100} \quad (15)$$

是根据单变量函数 $\text{sinc}(x)$ 在受到噪声的干扰下产生的, 也就是说 x_1, x_2, \dots, x_{100} 是在 $[-10, 10]$ 上均匀分布的 100 个点, 而 $y_k = \text{sinc}(x_k) + \xi_k, k = 0, 1, \dots, 100$ 。其中噪声 ξ_k 服从正态分布, 且 $E\xi_k^2 = \sigma^2, E\xi_i = 0$,

这里 $\sigma^2 = 0.2$ 。为了和其他核函数方法进行比较, 表 1 中列出了三种核函数的均方误差和支持向量个数。仿真结果表明, 本文构造的核函数支持向量机模型方法不但逼近精度高, 而且支持向量数目少, 泛化能力强。这说明了式 (11) 核函数的有效性和可行性。

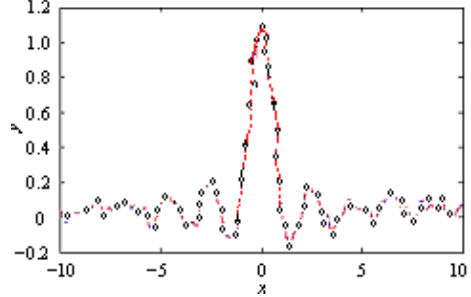


图 1 本文核函数的仿真结果 (“ \circ ” 为原函数, “-” 为 LP-WSVR2 的辨识结果)

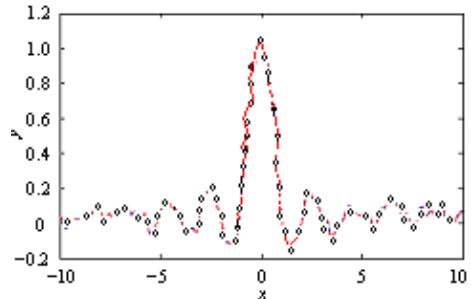


图 2 RBF 核函数的仿真结果 (“ \circ ” 为原函数, “-” 为高斯核函数的辨识结果)

表 1 性能比较

核函数	支持向量个数	性能指标(MSE)
高斯核函数	45	9.478 5e-004
RBF核函数	38	8.167 2e-004
本文核函数	29	2.843 2e-004

4 结论

本文在对核函数研究的基础上, 利用 Walsh 序列构造出再生核函数, 并将这种核函数用于一维函数的回归。仿真试验显示, 基于 Walsh 序列的核函数同 RBF 等核函数相比, 不仅提高了辨识速度, 而且增加了对非线性系统的辨识能力和相应的抗干扰能力, 减少了支持向量数目。同时由 VC 维理论知道, 支持向量的数目越少, 泛化能力越强。

参考文献

[1] 李元诚, 李波, 方廷健. 基于小波支持向量机的非线性组合预测方法研究[J]. 信息与控制, 2004, 33(3): 303-306.
 [2] 崔万照, 朱长纯, 保文星, 等. 最小二乘小波支持向量机在非线性系统辨识中的应用[J]. 西安交通大学学报, 2004, 38(6): 563-565.

- [3] 张莉, 周伟达, 焦李成. 尺度核函数支撑向量机[J]. 电子学报, 2002, 30(4): 527-529.
- [4] 杜平, 张燕昆, 刘重庆. 基于不变矩的人脸识别方法的研究[J]. 计算机仿真, 2002, 19(3): 78-81.
- [5] 祝海龙, 屈梁生, 张海军. 基于小波变换和支持向量机的人脸检测系统[J]. 西安交通大学学报, 2002, 36(9): 947-950.
- [6] VAPNIK V N. The nature of statistical learning theory[M]. Beijing: Tsinghua University Press, 2000.
- [7] ZHANG L, ZHOU W D, JIAO L C. Scaling Kernel function support vector machines[J]. Acta Electronica Sinica, 2002, 4(4): 527-529.
- [8] YAN H, ZHANG X G, MA Y Q, et al. The parameter estimation of RBF Kernel function based on verigram[J]. Acta Automatica Sinica, 2002, 28(3): 450-455.
- [9] OSUNA E, FREUND R, GIROSI F. Training support vector machines: an application to face detection[C]//Comput Vis Pattern Recogn. [S.l.]: IEEE, 1997: 130-136.
- [10] GENOV R. Massively parallel mixed-signal VLSI Kernel machines[D]. Baltimore: Johns Hopkins Univ, 2002.
- [11] LASKOV P. Feasible direction decomposition algorithms for training support vector machine[J]. Machine Learning, 2002, 46(1): 315-349.
- [12] LIN C F, WANG S D. Fuzzy support vector machines[J]. IEEE Transaction on Neural Networks, 2002, 48(1): 85-105.
- [13] VAPNIK V, CHAPPELL O. Bounds on error expectation for support vector machine[M]. Advance in Large Margin Classifiers: MIT Press, 1999.

编辑 漆蓉

(上接第83页)

由表2可知, 网格数越大, SOR并行迭代算法可以得到更好的加速比和效率, 因为计算时间占总的消耗时间比例越大。

该实验是在HUB连接的LAN下完成的, 并且整个局域网与广域网是连通的, 如果采用更加高速的网络连接, 可以得到更好的加速比和效率。

表2 SOR并行算法的加速比和效率与网格大小的关系

网格数	P=2		P=4	
	加速比	效率	加速比	效率
256×256	1.58	0.79	2.35	0.59
512×512	1.71	0.86	2.80	0.70
1 024×1 024	1.81	0.90	3.39	0.85

4 结束语

在二维CHIPIC软件中, 无论实际所需模拟的区域是否规则, 都会采用矩形的模拟区域, 每完成一次迭代都要进行相应的边界处理。因此, 对于复杂的图形结构, 只是边界条件的处理和设置不同, 可以得到与上述规则结构近似的结果。从而CHIPIC软件的泊松模块宜采用SOR的并行迭代算法实现。

参 考 文 献

- [1] 狄隽, 祝大军, 刘盛纲. CHIPIC软件电磁场计算方法[J]. 电子科技大学学报, 2005, 34(4): 484-488.
- [2] BIRDSALL K G, LANGDON A B. Plasma physics via computer simulation[M]. New York: McGraw-Hill, Inc., 1985.
- [3] 王秉中. 计算电磁学[M]. 北京: 科学出版社, 2002.
- [4] 钟尔杰, 黄廷祝. 数值分析[M]. 北京: 高等教育出版社, 2004.
- [5] 孙世新, 卢光辉, 张艳, 等. 并行算法及其应用[M]. 北京: 机械工业出版社, 2005.
- [6] GOPLEN B, LUDEKING L, SMITHE D, et al. User-configurable MAGIC for electromagnetic PIC calculations[J]. Comput. Phys Commun, 1995, 87: 54-86.
- [7] 刘大刚, 祝大军, 刘盛纲. 爆炸式发射二极管的粒子模拟研究[J]. 电子科技大学学报, 2005, 34(4): 481-484.
- [8] 莫则尧, 袁国兴. 消息传递并行编程环境[M]. 北京: 科学出版社, 2001.
- [9] QUINN M J. Parallel programming in C with MPI and openMP[M]. New York: McGraw-Hill, Inc, 2004.
- [10] 陈国良. 并行计算[M]. 北京: 高等教育出版社, 1999.

编辑 漆蓉