基于支持向量机的非线性系统辨识研究*

张明光[1], 阎威武[2], 李战明[1]

(1. 兰州理工大学 电气与信息工程学院, 甘肃 兰州 730050; 2. 上海交通大学 自动化研究所, 上海 200030)

摘　要: 研究了基于支持向量机的非线性系统辨识方法并进行了仿真试验,用交叉验证的方法进行支持向量机参数选择。有效的仿真结果表明支持向量机是非线性系统辨识的一种非常有效的方法。

关键词: 支持向量机; 系统辨识; 非线性

中图法分类号: TP274　　**文献标识码**: A　　**文章编号**: 1001-3695(2006)05-0047-02

Study of Nonlinear System Identification Based on Support Vector Machine

ZHANG Ming-guang[1], YAN Wei-wu[2], LI Zhan-ming[1]

(1. College of Electric & Information Engineering, Lanzhou University of Science & Technology, Lanzhou Gansu 730050, China; 2. Institute of Automation, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract: Nonlinear system identification based on SVM(Support Vator Machine) was discussed and corresponding simulation was implemented. Cross validation method was used to select hyperparameter of SVM model. Good result indicated that SVM is effective tool for nonlinear system identification.

Key words: Support Vector Machine; System Identification; Nonlinear

非线性系统辨识是控制领域里一个非常重要的研究内容,也是控制理论能否成功地应用于实际的关键之一。根据辨识的系统模型可以设计各种合适的控制器,从而使非线性系统工作在所需要的状态;同时,利用系统模型也可以方便地实现对非线性系统的监控。非线性系统辨识在信号处理中同样具有很重要的作用。对线性、非时变和具有不确定参数的对象进行辨识和自适应控制的研究已取得了很大的进展,与线性系统的辨识相比,非线性系统的辨识有很多难点,如模型的结构和阶次很难确定。人工神经网络由于其可以逼近任意类型的非线性,所以给未知非线性系统的辨识带来了新的方法。然而人工神经网络存在着收敛速度慢和局部极小值的问题,且在一般的神经网络中还没有通用的方法确定隐层单元数目,存在过拟合现象。支持向量机是由 Vapnik[1]最初在统计学习理论的基础上建立起来的一种非常有力的机器学习方法。它是近几年发展起来的一种人工智能技术,支持向量机较好地解决了传统学习算法中存在的小样本、非线性、高维数、局部极小点等实际问题,具有很强的泛化能力。支持向量机为非线性系统辨识提供了一个非常有力的工具。

1　支持向量机非线性估计算法

支持向量机主要是基于如下思想: 通过事先选择的非线性映射将输入向量映射到高维特征空间,在这个空间中构造最优决策函数。在构造最优决策函数时,利用了结构风险最小化原则,最后利用原空间的核函数取代了高维特征空间中的点积运算。

收稿日期: 2004-10-01; 修返日期: 2005-06-15

基金项目: 国家科技攻关计划资助项目(2002BA901A28)

1.1　支持向量机估计算法[2]

设样本为 n 维向量,某区域的 l 个样本及其值表示为 $(x_1, y_1), \cdots, (x_l, y_l) \in R^n \times R$

首先用一非线性映射 $\psi(\cdot)$ 把样本从原空间 R^n 映射到特征空间 $\psi(x) = (\varphi(x_1), \varphi(x_2), \cdots, \varphi(x_l))$。在这个高维特征空间中构造最优决策函数 $y(x) = w \cdot \phi(x) + b$,这样非线性估计函数转换为高维特征空间中的线性估计函数。利用结构风险最小化原则,寻找 w, b 就是最小化 $R = \frac{1}{2} \cdot \| w \|^2 + c \cdot R_{emp}$,其中 $\| w \|^2$ 是控制模型的复杂度,c 是正规化参数,R_{emp} 为误差控制函数,也即 ε 不敏感损失函数。

支持向量机估计算法的优化问题为

$$\min J(w, \xi) = \frac{1}{2} w \cdot w + c \sum_{i=1}^{l} (\xi_i + \xi_i^*) \tag{1}$$

$$s.t.\quad y_i - w \cdot \varphi(x_i) - b \le \varepsilon + \xi$$

$$w \cdot \varphi(x_i) + b - y_i \le \varepsilon + \xi^*,\ \xi^*,\ \xi \ge 0,\ i = 1, \cdots, l$$

这个优化问题的解可通过 Lagrang 函数的鞍点给出:

$$L(w, \xi^*, \xi\ a, a^*, c, \beta, \beta^*) = \frac{1}{2} w \cdot w + c \sum_{i=1}^{l} (\xi + \xi_i^*) - \sum_{i=1}^{l} a_i ((w \cdot \varphi(x_i)) - y_i + b + \varepsilon + \xi) - \sum_{i=1}^{l} a_i^* (y_i - (w \cdot \phi(x_i)) - b + \varepsilon + \xi^*) - \sum_{i=1}^{l} (\beta \xi + \beta^* \xi^*) \tag{2}$$

根据优化条件: $\frac{\partial L}{\partial w} = 0, \frac{\partial L}{\partial b} = 0, \frac{\partial L}{\partial \xi_i^*} = 0, \frac{\partial L}{\partial \xi_i} = 0$

可得: $w = \sum_{i=1}^{l} (a_i - a_i^*) \varphi(x_i)$, $\sum_{i=1}^{l} (a_i - a_i^*) = 0$,

$$c - a_i - \beta_i = 0,\ c - a_i^* - \beta_i^* = 0,\ i = 1, \cdots, l \tag{3}$$

根据 Mercer's 条件, 定义核

$$K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j) \tag{4}$$

由式(2)、式(3) 和式(4), 优化问题转换为如下优化问题:

$$\max W(a, a^*) = -\frac{1}{2}\sum_{i,j=1}^{l}(a_i - a_i^*)(a_j - a_j^*)K(x_i, x_j) + \sum_{i=1}^{l}(a_i - a_i^*)y_i - \sum_{i=1}^{l}(a_i - a_i^*)\varepsilon \quad (5)$$

$$s.t. \quad \sum(a_i - a_i^*) = 0,$$

$$0 \le a_i \le c,\ 0 \le a_i^* \le c,\ i = 1, \cdots, l$$

非线性估计函数为

$$f(x) = \sum_{支持向理}(a_i - a_i^*)K(x_i, x_j) + b \quad (6)$$

1.2 核函数

选择不同的核函数 $K(x, x_i)$ 可构造不同的支持向量积，常用的核包括：

(1) 生成多项式的核 $K(x, x_i) = [(x, x_i) + 1]^d$，$d$ 为阶数；

(2) 生成径向基函数的核 $K(x, x_i) = K(|x - x_i|)$；

(3) 生成神经网络的核 $K(x, x_i) = S(\upsilon(x, x_i) + c)$。

2 基于支持向量机的系统辨识

非线性系统辨识的基本方法就是利用已有的数据构造一个数学模型，即 $\bar{y} = f(u, y)$，使它能够对输出变量进行最佳估计和预测。其中 u 为系统已知的输入量，y 为系统已知的输出变量，$\bar{y}$ 为模型 f 对输出变量的最优估计。一般传统的系统辨识方法是根据系统的输入/输出数据对 u/y，利用数学方法提取系统的数学模型 f，所采用的训练原则为 $e = \|\bar{y} - y\| < \varepsilon$ 成立，其中 ε 为预定的辨识精度。这种传统的训练原则存在着缺点，即只对训练误差进行约束，而没有对模型的结构进行控制。所以有时最后得到的模型虽然有良好的训练精度，但泛化能力却不令人满意，如神经网络中的过拟合现象。而支持向量机算法采用了结构风险最小原则，达到了训练精度和模型复杂度之间的某种平衡，在保证训练精度的情况下，很好地控制了模型的复杂度，从而在算法上保证了模型具有良好的泛化性能。

基于支持向量机的系统模型属于黑盒子一类的模型，所以只关心对象的输入与输出，而不必关心对象的具体结构，输入与输出的映射关系由支持向量机来实现。基于支持向量机的系统辨识过程如图 1 所示。在具体进行系统模型的辨识时，首先利用输入变量 u 和输出变量 y 的样本对支持向量机进行训练，然后利用训练好的模型对输出变量进行估计。基于支持向量机的系统模型如图 2 所示，其中中间层为支持向量。最后在非线性系统模型中，输入变量与输出变量之间的非线性函数关系由支持向量机 $f(x) = \sum_{支持向量}(a_i - a_i^*)K(x_i, x_j) + b$ 来实现。

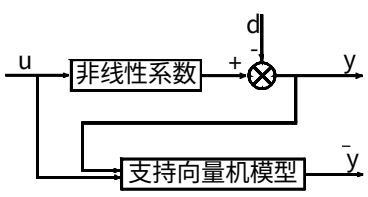

图 1 支持向量机系统辨识过程

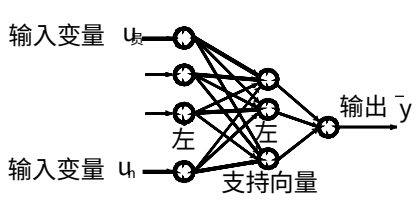

图 2 支持向量机系统模型

在进行支持向量机支持向量机建模时有两类参数，即正规化参数 c 和核参数(如生成多项式的核的阶数 d)，它们是非常重要的参数。寻找最佳正规化参数和核参数的问题也就是最佳模型选择问题。本文用交叉验证的方法进行参数的选取，首先确定正规化参数集 $S_c = \{c_1, \cdots, c_N\}$ 和核参数 $S_\theta = \{\theta_1, \cdots, \theta_N\}$，然后从参数集中选取参数分别进行组合，对支持向量机进行训练。用确认集检验，选出最佳的参数组合作为模型的最终参数。

基于支持向量机的系统辨识步骤如下：

(1) 选取样本，并对样本数据进行校正和预处理。

(2) 确定正规化参数集 S_c 和核参数集 S_θ。

(3) 从这两个参数集中选取参数分别进行组合。

(4) 利用所选参数进行支持向量机训练。

(5) 用确认集进行检验，返回步骤(3)直至组合结束。

(6) 选出最佳参数组合，建立支持向量机模型。

(7) 用建立好的模型进行预测。

3 仿真

双输入单输出的非线性静态系统[3]表示如下：

$$z = (1 + x^{-2} + y^{-1.5})^2 \quad 1 \le x, y \le 5 \quad (7)$$

采用支持向量机的方法对系统进行建模。

3.1 数据处理

首先对变量进行归一化：$\bar{x}_i = \dfrac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)}$ (8)

归一化后的属性值 $\bar{x}_i \in [0, 1]$。

核函数选取径向基函数：

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (9)$$

其中，$\|x - x_i\| = \sqrt{\sum_{k=1}^{n}(x^k - x_i^k)^2}$，$\sigma$ 为核宽度。

用测试集的泛化均方根误差(RMSE)判断建模效果的优劣：

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (10)$$

式中，n 为校验样本数目，y_i，$\bar{y}_i$ 分别为实际值与估计值。

正规化参数集 $S_c = \{50, 25, 10, 5, 0.75, 0.5, 0.25, 0.1, 0.05, 0.01\}$ 和核参数 $S_\sigma = \{5, 1, 0.75, 0.5, 0.25, 0.1, 0.075, 0.05, 0.025, 0.01\}$。

3.2 结果及讨论

图 3 为支持向量机模型的估计值与实际值的比较曲线。

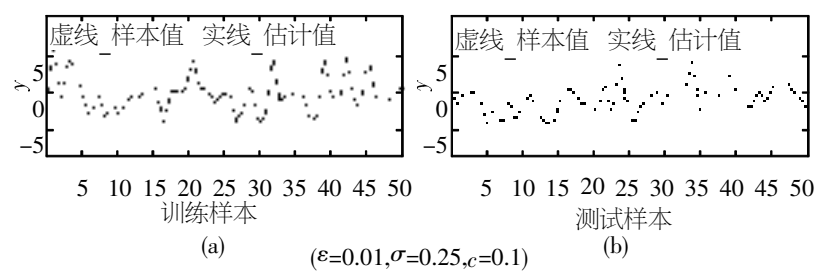

图 3 支持向量机模型的估计值与是及分析值的比较曲线

从图 3 可以看出，支持向量机模型的估计值与实际分析值吻合得非常好，估计结果的精度很高，较好地跟踪了系统的变化趋势。将试验结果与其他建模方法，如 Sugeno 使用的基于模糊聚类方法[4]、Gomez-Skarmeta 使用的 EST 算法及多模型方法[5]进行比较，比较结果如表 1 所示。从表 1 可以看出，支持向量机模型表现出了良好的泛化能力。

表 1 非线性系统建模误差

模型	Sugeno	EST1/EST2/EST3	支持向量机
RMSE	0.564	0.300/0.282/0.266	0.1047

通过仿真研究可以发现基于支持向量机的非线性系统辨识方法有如下优点：①只需要非线性系统较少的先验知识。②不需要确定系统的具体结构，而且与普通神经网络相比，其模型结构是自动确定的，没有局部极值问题。③由于 (下转第 82 页)

一些机制,允许将 Schema 组件设计在不同的模式文档之中,由这些模式文档共同组成一个完整的模式。这此机制包括: Include,允许在模式文档中包含具有相同命名空间的另一个模式文档; Redefine,如果被重定义的模式文档与重定义它的模式文档具有相同的目标命名空间,则可对模式文档进行重定义; Import,允许引用来自其他命名空间的组件。如果需要设计的模式数量很大,而且需要重用的 Schema 组件也很多,则可选择在多个模式文档间进行分层,具体的示例可参考文献[2]。

以下代码便是依据这种思想设计的一个简单 W3C XML Schema 模式"bookstore4. xsd"。在此模式中,共有三个模式文档:"bookelement. xsd","book. xsd"和"bookStore. xsd"。其中"book. xsd"包含"bookelement. xsd",而"bookstore4. xsd"又包含"book. xsd",形成了一个在多个模式文档间层层嵌套的关系。

(1) 模式文档 bookelement. xsd

```
< xsd: schema… >
< xsd: element name = " Title" type = " xsd: string" / >
< xsd: element name = " Author" type = " xsd: string" / >
< xsd: element name = " Data" type = " xsd: string" / >
< xsd: element name = " ISBN" type = " xsd: string" / >
< xsd: element name = " Publisher" type = " xsd: string" / >
< /xsd: schema >
```

(2) 模式文档 book. xsd: < xsd: schema… >

```
< xsd: include schemaLocation = " . \bookelement. xsd" / >
< xsd: element name = " Book" >
< xsd: complexType >
< xsd: sequence >
< xsd: element ref = " Title" / >
< xsd: element ref = " Author" / >
< xsd: element ref = " Date" / >
< xsd: element ref = " ISBN" / >
< xsd: element ref = " Publisher" / >
< /xsd: sequence >
< /xsd: complexType >
< /xsd: element >
< /xsd: schema >
```

(3) 模式文档 bookstore4. xsd: < xsd: schema… >

```
< xsd: include schemaLocation = " . \book. xsd" / >
< xsd: element name = " BookStore" >
< xsd: complexType >
< xsd: sequence >
< xsd: element ref = " Book" maxOccurs = " unbounded" / >
< /xsd: sequence >
< /xsd: complexType >
< /xsd: element >
< /xsd: schema >
```

1.3 综合结构化程序设计方法和面向对象的设计方法

这种设计方式实际上是综合了结构化程序设计方法和面向对象的设计方法。在此种设计方式中,全局组件(包括元素、类型等)和局部组件(包括元素、类型等)的设计是随意的,并被随意地组合在一起。此种方式是最差的一种模式设计方法,应该尽量避免。运用这种设计思想的最典型的 W3C XML Schema 设计模式为博洛尼亚(Bologna 意大利城市),具体内容可参考文献[1]。

2 小结

本文从结构化程序设计和面向对象程序设计的视角共介绍了六种比较典型的 W3C XML Schema 模式设计方法,但是,具体选择哪一种模式设计方法,则取决于业务的实际需要。如果模式文档的大小是要考虑的主要问题,则结构化程序设计方法比较合适;如果名字空间的复杂性是要考虑的主要问题,则结构化程序设计方法和"封装的 Schema 组件类型侧重于类型"的设计方法比较合适;如果可扩展性是要考虑的主要问题,则面向对象的设计方法比较合适。

参考文献:

[1] edBiquitous. XML Schema Design Guidelines, version 1. 2[EB/OL]. http://www. medbiq. org/technology/tech_architecture/xmldesignguidelines. pdf, 2004-04-30.

[2] Ma Yongheng, Xiong Qianxing, Wu Yefu, *et al*. Building W3C XML Schema for UN/EDIFACT Messages with Multiplayer and Modular Architecture[C]. Proceedings of the International Conference on Web Service, Las Vegas, Nevada, USA: CSREA Press, 2003. 515-519.

[3] 李浩,孙统风,孟现飞,等. 基于面向对象思想构建 XML Schema [J]. 微机发展, 2003, 13(6): 59-64.

[4] 胡庆彬,黄石生. 结构化和面向对象程序设计方法的探讨[J]. 计算机应用研究, 1998, 15(1) 56-58.

[5] 王茹,宋瀚涛. XML 文档结构定义规范——XML Schema[J]. 计算机应用研究, 2002, 19(1): 127-129.

作者简介:

马永恒(1975-),男,湖北云梦人,硕士生,研究方向为计算机支撑技术;熊前兴(1943-),男,湖北洪湖人,博士生导师,研究方向为计算机支撑技术;杨金娥(1976-),女,湖北孝感人,助理研究员,硕士。

(上接第 48 页)采用了核函数来处理输入输出之间的非线性关系,且核的选择灵活多样,使之非常适合于非线性系统的辨识。④基于支持向量机的模型具有非常好的泛化能力。

4 结论

支持向量机是一种新的通用学习机器,它较好地解决了传统学习方法中存在的小样本、非线性、高维数、局部极小点等实际问题,具有很强的泛化能力。本文研究了基于支持向量机的非线性系统辨识方法,利用支持向量机进行非线性系统辨识。有效的仿真结果表明支持向量机为非线性系统辨识提供了一种十分有效的新方法,在非线性系统辨识中有良好的应用前景。

参考文献:

[1] apnik V N. The Nature of Statistical Learning Theory[M]. New York: Springer-Verlag, 1999.

[2] Suykens J A K. Nonlinear Modeling and Support Vector Machines [C]. Proc. of the 18th IEEE Conference on Instrumentation and Measurement Technology, 2001. 287-294.

[3] Nakanishi H, Turksen I B, Sugeno M. A Review and Comparison of Six Reasoning Method[J]. Fuzzy Sets and Systems, 1992, 57: 257-294.

[4] Sugeno M, Yasukawa T. A Fuzzy-Logic-based Approach to Qualitative Modeling[J]. IEEE Trans. Fuzzy Systems, 1993, 1(1): 7-31.

[5] Gomez-Skarmeta A F, Delgado M, Vila M A. About the Use of Fuzzy Clustering Techniques for Fuzzy Model Identification[J]. Fuzzy Sets and Systems, 1999, 106: 180-188.

作者简介:

张明光(1971-),男,副教授,硕士,主要研究领域为人工智能、计算机控制及其在过程控制中的应用;阎威武(1972-),男,讲师,博士,主要研究领域为人工智能、智能控制、数据挖掘;李战明(1962-),男,院长,教授,硕士,主要研究领域为人工智能和智能控制。