

一种基于向量的鲁棒 ARMA 系统辨识器^{*}

胡学骏, 赵慧民, 陈广义

(佛山科学技术学院机电学院, 广东 佛山 528000)

摘要: 提出一种新的基于向量方法的自回归和运动平均 (ARMA) 模型系统辨识器, 并给出了其参数的统计分析模型。应用结果表明, 向量 ARMA 算法和最小二乘法 LS 算法相比, 在一定条件下, 其预测误差精度提高了约 1.2 dB; 且该系统模型不受分离向量参数的影响。使用非线性函数核, 系统将会成为一个鲁棒的非线性辨识过程。

关键词: ARMA 建模; 支持向量方法; 系统辨识器; 时间序列

中图分类号: TK32; P42 **文献标识码:** A **文章编号:** 0529-6579 (2005) 02-0039-04

数字信号处理通常的模型是线性时不变系统 (LTIS), 它通过有理传输函数进行分析估计, 如运动平均 (MA), 自回归 (AR) 和零极点 ARMA 系统。ARMA 系统建模对系统辨识和时间序列分析及谱分析是非常重要的。由于传统的 ARMA 系统估计方法有很多限制^[1], 本文提出一种基于支持向量方法 SVM (support vector method) 的 ARMA 模型, 即 SVM-ARMA 模型。SVM 首先在归类问题上提出以获得最大边缘分离超平面^[2], 但这种技术已成为通用的学习理论^[3]。这种方法对于分类和回归问题的综合描述见文献 [4, 5]。我们所提出的 SVM 主要为了研究 ARMA 系统的鲁棒性, 鲁棒性正是 SVM 主要的控制特性。另一方面, 研究了本方法的参数统计分析。该系统主要的优点是: ①提供一个惟一的系统解决方案; ②具有超强的调整性能; ③尽管统计分配无知, 但它从获得的样本中可获得最大信息; ④减少了传统 ARMA 系统识别方法的限制。

1 SVM-ARMA 定理

考虑一个有理线性时不变系统 LTIS, 其输入/输出离散时间序列为 $\{x_n\}$, $\{y_n\}$ 。对应的差分方程为:

$$y_n = \sum_{i=1}^P a_i y_{n-i} + \sum_{j=1}^Q b_j x_{n-j+1} + e_n \quad (1)$$

这里 $\{a_n\}$, $\{b_n\}$ 分别表示了 P 序列自回归 AR 和 Q 序列运动平均 MA 的系数, $\{e_n\}$ 代表了离散时间序列的误差影响。

也考虑在相同时间瞬间观测到的 $\{x_n\}$, $\{y_n\}$ N 个的连续样本。差分方程可以设置说明它们的相关性、估计参数和误差。考虑初始化条件, (1) 式仅需要时间延迟 $n = k_0, \dots, N$, 这里 $k_0 = \max(P+1, Q)$ 。假设 P 和 Q 能被合理选择, ARMA(P, Q) 系数通过减少价值函数进行估计, 并取决于余项。

一般地, 对于线性分类和回归问题, SVM 算法把余项称为非灵敏- ϵ 损失函数, 定义如下

$$L_\epsilon(e_n) = \begin{cases} |e_n| - \epsilon, & \text{若 } |e_n| \geq \epsilon \\ 0, & \text{若 } |e_n| < \epsilon \end{cases} \quad (2)$$

同时, 余项的大小可以用本模型参数的范数 L_2 调整。由于 $\epsilon \neq 0$, SVM 允许估计分类或回归函数仅取决于观测样本所减少的数目, 这些样本称为支持向量。SVM 可扩展到非线性分类器和回归器。我们使用非灵敏- ϵ 函数和 L_2 调整, 包括 AR 和 MA 模型系数, 来建立 SVM-ARMA 模型。它对应于下列的非约束极值

$$L_p(a_i, b_j, e_n) = \frac{1}{2} \left(\sum_{i=1}^P a_i^2 + \sum_{j=1}^Q b_j^2 \right) + C \sum_{n=k_0}^N L_\epsilon(e_n) \quad (3)$$

其中参数 C 是固定系数, 以控制余项的价值和调整项之间的折中。

以上处理的等价公式是最小化问题

$$L_p(a_i, b_j, \xi_n, \xi_n^*) = \frac{1}{2} \left(\sum_{i=1}^P a_i^2 + \sum_{j=1}^Q b_j^2 \right) + C \sum_{n=k_0}^N (\xi_n + \xi_n^*) \quad (4)$$

这里 ξ_n 和 ξ_n^* 是余项变量, 或损失约束项。

* 收稿日期: 2004-10-11

基金项目: 广东省自然科学基金资助项目 (04105503); 佛山市科技发展专项资金资助项目 (04010052)

作者简介: 胡学骏 (1965年生), 男, 副教授; E-mail: fodajun@163.net

$$y_n - \sum_{i=1}^P a_i y_{n-i} - \sum_{j=1}^Q b_j x_{n-j+1} \leq \epsilon + \xi_n \quad (5)$$

$$-y_n - \sum_{i=1}^P a_i y_{n-i} - \sum_{j=1}^Q b_j x_{n-j+1} \leq \epsilon + \xi_n^* \quad (6)$$

$$\xi_n^{(*)} \geq 0 \quad (7)$$

$n = k_0, \dots, N$, 这里 $\xi_n^{(*)}$ 代表了 ξ_n 和 ξ_n^* 。为了解决问题, 对于每个约束条件, 我们在拉格朗日乘法器中引入一个非负系数, 有

$$\begin{aligned} L_{PD}(a_i, b_j, \xi_n, \xi_n^*, \alpha_n, \alpha_n^*, \beta_n, \beta_n^*) = & C \sum_{n=k_0}^N (\xi_n + \xi_n^*) + \\ & \frac{1}{2} \left(\sum_{i=1}^P a_i^2 + \sum_{j=1}^Q b_j^2 \right) - \sum_{n=k_0}^N (\beta_n \xi_n + \beta_n^* \xi_n^*) + \\ & \sum_{n=k_0}^N \alpha_n \left(y_n - \sum_{i=1}^P a_i y_{n-i} - \sum_{j=1}^Q b_j x_{n-j+1} - \epsilon - \xi_n \right) + \\ & \sum_{n=k_0}^N \alpha_n^* \left(-y_n + \sum_{i=1}^P a_i y_{n-i} + \sum_{j=1}^Q b_j x_{n-j+1} - \epsilon - \xi_n^* \right) \end{aligned} \quad (8)$$

这里乘法器被约束为 $\alpha_n^{(*)} \geq 0$ 和 $\beta_n^{(*)} \geq 0$ 。公式 (8) 最小化必须考虑初始变量 a_i, b_j 和 $\xi_n^{(*)}$, 其最大化要根据拉格朗日乘法器变量因子。从

$$\frac{\partial L_{PD}}{\partial a_i} = 0; \frac{\partial L_{PD}}{\partial b_j} = 0; \frac{\partial L_{PD}}{\partial \xi_n^{(*)}} = 0 \quad (9)$$

可以得出几个结论。首先, 变量 $\alpha_n^{(*)}$ 和 $\beta_n^{(*)}$ 约束条件的上下范围是

$$0 \leq \alpha_n^{(*)}; \beta_n^{(*)} \leq C \quad (10)$$

第二, 可以观察

$$a_i = \sum_{n=k_0}^N (\alpha_n - \alpha_n^*) y_{n-i} \quad (11)$$

$$b_j = \sum_{n=k_0}^N (\alpha_n - \alpha_n^*) x_{n-j+1} \quad (12)$$

说明了模型系数和变量参数的分析联系。

第三, 这些条件可以被引入到 (8) 式的拉格朗日函数来移动初始变量。这样, 提出的鲁棒 SVM ARMA 算法可以用余项价值函数和调整项来描述, 其表示式如下

$$\begin{aligned} L_P^{SVM}(a_i, b_j, \xi_n, \xi_n^*) = & \frac{1}{2} \left(\sum_{i=1}^P a_i^2 + \sum_{j=1}^Q b_j^2 \right) + \\ & \frac{1}{2\gamma} \sum_{n \in I_1} (\xi_n^2 + \xi_n^{*2}) + C + \sum_{n \in I_2} (\xi_n + \xi_n^*) \end{aligned} \quad (13)$$

这里 I_1 是 $\epsilon \leq \xi_n^* \leq \epsilon_C$ 的样本设置, I_2 是 $\xi_n^{(*)} > \epsilon_C$ 的样本设置, 且 $\epsilon_C = \epsilon + \gamma C$, γ 是一个调整小系数, P 和 Q 分别表示了局部时间输入/输出 DTP 自相关函数值的 P 和 Q 个样本估计。

2 统计描述

在 SVM ARMA 算法 (13) 中, 研究余项和拉格朗日乘法器的解析关系。设 $\alpha_n^0, \alpha_n^{*0}$ 是拉格朗日系数, 那么余项 e_n 和拉格朗日系数的关系可以表示为

$$\alpha_n^0 - \alpha_n^{*0} = f(e_n) = \begin{cases} -C, & e_n \geq \epsilon_C \\ \frac{1}{\gamma}(e_n + \epsilon), & -\epsilon_C \leq e_n \leq -\epsilon \\ 0, & -\epsilon \leq e_n \leq \epsilon \\ \frac{1}{\gamma}(e_n - \epsilon), & \epsilon \leq e_n \leq \epsilon_C \\ C, & e_C \leq e_n \end{cases} \quad (14)$$

这种分析特性的证明见文献 [6]。这种余项和拉格朗日系数在基于向量 ARMA 模型的可分析关系可以使我们的算法进行比较。这里, 我们给出 ARMA 模型进行预测误差方法 (PEMs) 时的比较, 其中 PEMs 应用了我们熟知的最小二乘法 (LS)。

2.1 SVM ARMA 和 PEM 比较

如果 $\epsilon = 0, \epsilon_C = +\infty$, 那么拉格朗日系数乘法器只和余项成比例。简便起见, 考虑数据 $\{x, y\}$ 子空间上的向量 $v \in R^3$ 投影, v_{LS} 产生一个非位移空间正交平均误差 d ; 而 v_{SVM} 有一个规则的近似均差 e 如图 1。按照 IS 准则, 此投影可表示为

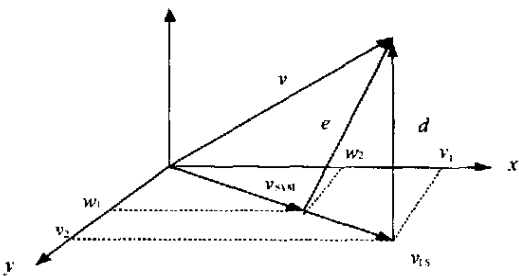


图 1 误差的几何描述

Fig 1 Geometrical interpretation of error bias

$v_{LS} = v_1 x + v_2 y$, 而 SVM 投影是 $v_{SVM} = w_1 x + w_2 y$ 。SVM 投影可以通过下式的极小值结果得到

$$L = \frac{1}{2}(w_1^2 + w_2^2) + \frac{1}{2\gamma} \|e\|^2 \quad (15)$$

约束条件是 $e = v - v_{SVM}$ 。从

$$\frac{\partial L}{\partial w_1} = 0, \frac{\partial L}{\partial w_2} = 0 \quad (16)$$

可以得到两种结果的系数关系

$$w_1 = \frac{1}{1 + \gamma} v_1, w_2 = \frac{1}{1 + \gamma} v_2 \quad (17)$$

$$\langle e, x \rangle = v_1 - w_1 = (1 + \gamma)w_1 - w_1 = \gamma w_1$$
$$\langle e, y \rangle = v_2 - w_2 = (1 + \gamma)w_2 - w_2 = \gamma w_2 \quad (18)$$

按照(11)和(12)式所给出的 SVM-ARMA 系数,本结果在子空间投影最小范数能够通过输入信号、输出信号和它们的延迟形式获得。其模型系数计算由非线性误差和子空间向量基之间的点积完成。

如果噪声是高斯噪声,所观测的数据实际上由 ARMA 模型产生,例如, $v = u_1x + u_2y$, 那么 v_{IS} 是该模型系数的一个非位移估计,且第 i 个系数 ($i = 1, 2$) 有

$$E\{v_i\} = u_i \quad (19)$$

这里, $E\{\cdot\}$ 表示了一个随机变量的统计期望。由于(20)式,而 v_{SVM} 是一个位移估计

$$E\{w_i\} = E\left\{\frac{1}{1 + \gamma}v_i\right\} = \frac{1}{1 + \gamma}u_i \quad (20)$$

另外, LS 估计的方差是

$$\sigma_{v_i}^2 = E\{(v_i - u_i)^2\} = E\{v_i^2\} - u_i^2 \quad (21)$$

而 SVM 估计的方差是

$$\sigma_{w_i}^2 = E\{(w_i - u_i)^2\} = \frac{1}{1 + \lambda}\sigma_{v_i}^2 + \frac{1}{1 + \lambda}u_i^2 \quad (22)$$

所以,在 SVM ARMA 的方差低于 LS 方差的情况下,SVM ARMA 会提供更好的估计。

3 应用实例分析

这里把 SVM-ARMA 算法应用于反馈训练过程,并和传统几个系统识别方法比较,观测其主要特性和性能指标。

传统的反馈训练过程的一个实际例子是通过二进制电压管理的一个可变加热控制器设备。我们的目标是对输入(二进制电压)和输出(出口空气温度)关系 LTIS 进行识别。

首先,输入/输出 DTP 的 $N = 80$ 个样本作为训练序列,接着设置 800 个有效样值。SVM ARMA 的参数固定过程如下:初始化,40 个样本用于建立模型自参数的设置,40 个样值用于估算判断预测误差的测试。研究的分类是 $\gamma \in (10^{-10}, 1)$, $\epsilon_c \in (10^{-5}, \max_n |y_n|)$, $\epsilon \in (0, 0.10\max_n |y_n|)$ 每个参数观测使用 20 个点。一旦自由参数已选择,全部的 80 个样本用于获得模型的系数。和 SVM ARMA 比较的几个方法是:最小二乘法 LS 算法 ARX 模型;输出误差模型 OE。图 2 显示样本有效设置的预测误差。可以看到,OE 模型最差,而 ARX 和 SVM ARMA 呈现较好的性能,且在图 2(a)中 $Q = 6$ 时,SVM-ARMA 比 ARX 模型改善了约 1.2 dB 误差增益。

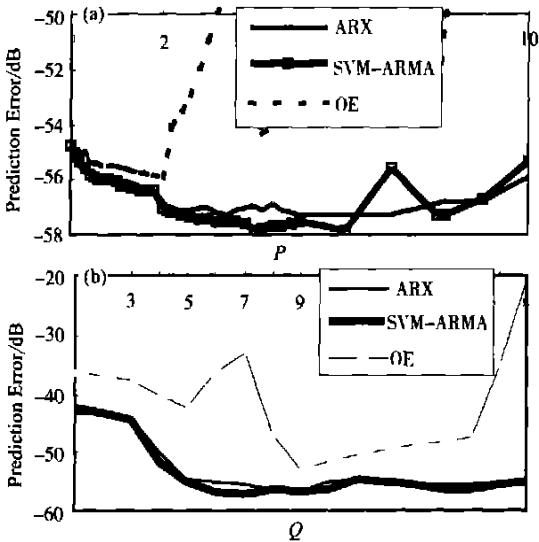


图 2 反馈训练过程结果

Fig. 2 Results for feedback's process trainer

(a) $Q = 6$ 且 $P \in (1, 10)$ 预测误差;
(b) $P = 5$ 且 $Q \in (1, 15)$ 预测误差

4 结 论

本文提出了一种基于向量 SVM 建立 ARMA 模型的有效方法,并给出了相应的算法公式和 SVM-ARMA 的统计描述。理论上,SVM ARMA 算法允许减少余项和数据的相关性。应用结果说明,和其它传统算法比较,SVM ARMA 算法通过对自由参数的近似选择,其系统辨识性能很好并不受分离向量的影响。这种方案可以扩展成自适应 SVM 训练方法以用于其它信号处理。使用非线性函数核(即允许获得输入/输出状态的传输函数),本系统将会成为一个鲁棒的非线性辨识过程。

参考文献:

[1] SIDIROPOULOS N D, DIMIC G Z. Blind multiuser detection in W-CDMA systems with large delay spread[J]. IEEE Signal Processing Lett, 2003, 8: 87—89.
[2] VAPNIK V. The nature of statistical learning theory[M]. New York: Springer-Verlag, 1995.
[3] SCHOLKOPF B, SUNG K. Comparing support vector machines with gaussian kernels to radial basis function classifiers[J]. IEEE Trans Signal Processing, 1997, 45: 2758—2765.
[4] PONTIL M, VERRI A. Support vector machines for 3D object recognition[J]. IEEE Trans Pattern Anal Machine Intell, 1998, 20: 637—646.
[5] BURGESS C J C. A tutorial on support vector machines for pattern recognition[J]. Data Mining Knowl Discovery, 1998, 22: 1—32.
[6] ROJO-ALVAREZ J L, de PRADO-CUMPLIDO M, ANIBAL R. Support vector method for robust system identification[J]. IEEE Trans on Signal Processing, 2004, 52: 155—163.

(下转第 48 页)

Solution of QoS Multicast Routing Problems Based on HGA

XIE Li ming¹, YU Feng ren², QIU Hai ming²

(1. Guangdong Industry Technical College, Guangzhou 510300, China;

2. Department of Electronics and Communication Engineering, Sun Yat sen University, Guangzhou 510275, China)

Abstract: A hybrid genetic algorithm (HGA) was presented to solve the delay and delay jitter constrained least cost multicast routing problem. Genetic algorithm (GA) is a kind of parallel optimization algorithm that simulates the evolution process of a creature, and is suitable for finding optimal solution in a large and complicated search space. GA does not need continuous solution to the problem, and can be easily implemented in parallel distributed process, so that it is effective in solving QoS multicast routing.

Key words: hybrid genetic algorithm (HGA); multicast; QoS; simulated annealing algorithm (SA); instructional preliminary colony building; heuristic crossover operation

(上接第 41 页)

A Robust ARMA System Identification Based on the Vector Method

HU Xue jun, ZHAO Hui min, CHEN Guang yi

(Foshan University, Guangdong Foshan 528000, China)

Abstract: To present a new approach to autoregressive and moving average (ARMA) modeling based on the support vector method, a statistical analysis of the characteristics of the proposed method is carried out. The results show, compared SVM-ARMA with LS, precisions of validation prediction error of the SVM-ARMA improved 1.2 dB than LS in some conditions. Besides, the effect of outliers can be cancelled. With using nonlinear kernels, the system will lead to robust, nonlinear system identification procedures.

Key words: ARMA modeling; support vector method; system identification; times series