

文章编号: 1005—0523(2007)01—0145—04

基于最小最大概率回归的混沌时间序列全局预测

刘遵雄, 刘建辉

(华东交通大学 信息工程学院, 江西 南昌 330013)

摘要:提出了使用最大化最小概率机器回归 MPMR 方法进行混沌时间序列的全局预测, 在分析最大化最小概率机器 MPM 及用于回归的原理的基础上, MPMR 方法被用于 Mackey—Glass 混沌时间序列多步预测. 在该序列进行 1 至 20 步的预测中, MPMR 方法较局步加权线性回归方法具有更好的预测效果. 核函数形状参数及回归管道宽度 ϵ 的选取将直接影响 MPMR 回归技术的系统性能, 试验中我们使用了交叉验证的方法确定这两个参数值. 试验证明, 核函数形状参数和回归管道宽度取值都为 2 和 3 的情况下, 预测效果更为理想.

关 键 词:混沌序列; 全局预测; 最大最小概率; 回归方法

中图分类号: TM 715; TP 18

1 引言

混沌时间序列预测广泛地应用在自然科学和社会科学的各个领域, 如预测太阳黑子的数目、水文数据、股票行情和电力负荷等. 给定混沌时间序列 $\{x_i\}_{i=1}^N$, 预测问题就是建立时间序列动力学关系 $x_{t+\Delta} = F(X_t)$, 其中 $X_t = (x_t, x_{t-\Delta}, \dots, x_{t-(m-1)\Delta})$ 是重构相空间(嵌入维为 m 延时为 Δ)的相点^[1]. 根据求解映射函数使用时间序列重构得相空间相点的不同, 将混沌时间序列预测方法分为全局预测和局部预测. 所谓全局预测, 是指利用全部已知相点建立函数 F , 包括指数模型、神经网络模型以及非参数核回归方法等^[2,3,4]; 而局部预测基于相邻的相点具有相似的发展趋势这一假设, 选择与预测相点近邻的相点求解 F ^[5].

最小最大概率回归 (Minimax Probability Regression, MPMR) 是将最大化最小概率分类 (Minimax Probability Machine Classification, MPMC)^[6] 推广用于解决回归和函数逼近问题的一种崭新的回归方法^[7]. MPMC 基于线性分类, 实现了对未知样本正确分类的最小概率最大化, 并且可以引入核函数获得

非线性分类效果. MPMC 不需要指定类条件密度的分布, 利用凸优化技术及矩问题方法获得与分布无关的结果, 能够直接求得对未知样本错误分类的概率最坏情况下的界. MPMR 实现了预测值落在实际回归函数的 ϵ 值管道内的最小概率最大化, 同样可以直接求出该最小概率的界.

本文使用最小最大概率回归方法进行混沌时间序列的全局多步预测, 获得很好的预测精度, 并就预测性能和相关条件同其它方法进行了分析比较.

2 最小最大概率分类与回归

建立回归模型的目的就是, 根据给定训练集获得近似回归函数 $\hat{f}$, 于是给定输入 x , 有输出估计值 $\hat{y} = \hat{f}(x)$; 同时实现 $\hat{f}(x)$ 位于真实值 y 的 ϵ 值管道 ($\epsilon > 0$) 的最小概率最大.

2.1 最小最大概率分类

最小最大概率分类器^[6]的建立首先是用来分开由均值和协方差矩阵为 $\bar{u}, \bar{V}$ 和 Σ_u, Σ_v 的分布产生的两类线性可分样本 $\{u_i\}_{i=1}^m$ 和 $\{V_i\}_{i=1}^m$. 需要求解能分开这两类样本的超平面 $a^T z = b$ ($a, z \in R^n, b \in R$)

满足对未知样本正确分类的概率下确界最大,即:

$$\max_{a, b} \text{约束条件: } \begin{cases} 1 - \alpha \geq \sup Pr\{a^T u \leq b\} \\ 1 - \alpha \geq \sup Pr\{a^T v \geq b\} \end{cases} \quad (1)$$

通过应用 Bertsimas 和 Sethuraman 扩展的定理^[8], $\sup Pr\{a^T v \geq b\} = 1/(1+d^2)$, $d^2 = \inf_{a^T v \geq b} (v - \bar{v})^T \Sigma_v^{-1} (v - \bar{v})$, 将以上问题可以通过求解优化问题: $\min_a$

$\sqrt{a^T \Sigma_u a} + \sqrt{a^T \Sigma_v a}$, 约束条件 $a^T (\bar{u} - \bar{v}) = 1$, 求得 a^* ; $b^* = a^{*T} \bar{u} - k^* \sqrt{a^{*T} \Sigma_u a} = a^{*T} \bar{v} - k^* \sqrt{a^{*T} \Sigma_v a}$, $k^* (\sqrt{a^{*T} \Sigma_u a} + \sqrt{a^{*T} \Sigma_v a}) = 1$.

对于线性不可分的两类样本 $\{u_i\}_{i=1}^{nu}$ 和 $\{v_i\}_{i=1}^m$, 可以通过引入核方法解决. 将线性不可分的 m 维输入空间通过函数 φ 映射到维数为 h 的高维空间, 有: $u \rightarrow \varphi(u)$, 对应均值和协方差阵 $(\overline{\varphi(u)}, \Sigma_{\varphi(u)})$; 有: $v \rightarrow \varphi(v)$, 对应均值和协方差阵 $(\overline{\varphi(v)}, \Sigma_{\varphi(v)})$. 二元分类器形式如下(对于样本属于第一类, $c = -1$; 属第二类, $c = +1$):

$$c = \text{sign} \left[\sum_{i=1}^{nu+m} \gamma_i K^c(z_i, z) + b_c \right] \quad (2)$$

其中 $K^c(z_i, z) = \varphi(z_i)^T \varphi(z)$, $z_i = u_i$, $(i = 1, \dots, nu)$, $z_i = v_i$, $(i = nu + 1, \dots, nu + m)$, $\gamma = (\gamma_1, \dots, \gamma_{nu+m})$, γ 通过求解以下最优化问题得出:

$$\min_{\gamma} \left\{ \left\| \frac{\overline{K_u}}{\sqrt{nu}} \gamma \right\|_2 = \left\| \frac{\overline{K_v}}{\sqrt{m}} \gamma \right\|_2 \right\} \quad s. t. \quad \gamma^T (\overline{k_u} - \overline{k_v}) = 1 \quad (3)$$

其中 $\overline{K_u} = K_u - 1_{nu} \overline{k_u}$, $\overline{K_v} = K_v - 1_m \overline{k_v}$, $\overline{k_u}$ 和 $\overline{k_v}$ 维数是 $nu + m$, 定义为: $[\overline{k_u}]_i = \frac{1}{nu} \sum_{j=1}^{nu} k^c(u_j, z_i)$,

$[\overline{k_v}]_i = \frac{1}{m} \sum_{j=1}^m k^c(v_j, z_i)$, 1_k 是元素全为 1 的列向量. K_u 包含 Gram 矩阵 K 的前 nu 行元素, K_v 包含 Gram 矩阵 K 的最后 m 行元素. K 是元素为 $K_{ij} = k^c(z_i, z_j)$ 的方阵. b_c 通过下式求得:

$$b_c = \gamma^T \overline{k_u} - k \sqrt{\frac{1}{Nu} \gamma^T \overline{K_u}^T \overline{K_u} \gamma} = \gamma^T \overline{k_v} + k \sqrt{\frac{1}{Nv} \gamma^T \overline{K_v}^T \overline{K_v} \gamma} \quad (4)$$

其中 $k = (\sqrt{\frac{1}{nu} \gamma^T \overline{K_u}^T \overline{K_u} \gamma} + \sqrt{\frac{1}{m} \gamma^T \overline{K_v}^T \overline{K_v} \gamma})^{-1}$, MPMC 框架的显著优点是, 给定 $\bar{u}$, Σ_u , $\bar{v}$ 和 Σ_v 信息, 不正确分类的最大概率受 $1 - \alpha$ 界定, α 可以通过 k 直接求出: $\alpha = \frac{k^2}{1 + k^2}$, 该结果将被用来求近似回归函数位于真实回归函数 ϵ 值管道内的概率的下界.

2.2 最小最大概率回归^[7]

应用 MPMC 方法形成 MPMR 框架, 需要根据每个样本数据 (x_i, y_i) 产生相应的两个 $1 + d$ 维向量 u_i 和 v_i ($i = 1, \dots, N$), $u_i = (y_i + \epsilon, x_{i1}, x_{i2}, \dots, x_{id})$, $v_i = (y_i - \epsilon, x_{i1}, x_{i2}, \dots, x_{id})$, 从而形成两类点集合, 通过最小化式(4)求得 γ . 这样可以获得两类点 u_i 和 v_i 的 MPM 分类面:

$$\sum_{i=1}^{2N} \gamma_i K^c(z_i, z) + b_c = 0 \quad (6)$$

可以认为是回归面, 因为它分开了训练样本集产生的两类点. 给定输入 $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$, 计算回归模型输出 $\hat{y}$, 相当于求取满足式(6)的 $\hat{y}$, 其中 $z = (y, x_1, x_2, \dots, x_d)$. 根据上面的讨论, 有 $z_i = u_i$ ($i = 1, \dots, N$), $z_i = v_{i-N}$ ($i = N + 1, \dots, 2N$). 如果 $K^c(z_i, z)$ 非线性, 求解 $\hat{y}$ 问题变成了一般的非线性单变量优化问题.

为了免于使用非线性优化技术从式(6)求解 $\hat{y}$, 限定核函数 $K^c(z_i, z)$ 满足条件: $K^c(z_1, z) = y'_1 \hat{y} + K(x_i, x)$, $K(x_i, x) = \varphi(x_i)^T \varphi(x)$ 是满足 Mercer 条件的核, $z = (y, x_1, x_2, \dots, x_d)$, $z_i = u_i$, $y'_i = y_i + \epsilon$ ($i = 1, \dots, N$), $z_i = v_{i-N}$, $y'_i = y_i - \epsilon$ ($i = N + 1, \dots, 2N$). 于是式(5)具有解析解: $\beta_i = -2\epsilon(\gamma_i + \gamma_{i+N})$, $b = -2\epsilon b_c$. 式(4)求解最优 γ 值问题转化成以下线性最小二乘问题:

$$\min_t \|\overline{K_u}(\gamma_0 + Ft)\|_2 \quad (7)$$

其中 $\gamma = \gamma_0 + Ft$, $\gamma_0 = (\overline{k_u} - \overline{k_v}) / \|\overline{k_u} - \overline{k_v}\|_2$, t 是 $2N - 1$ 维向量, $F \in R^{2N \times (2N - 1)}$, 是正交阵, 其列张成正交于 $\overline{k_u} - \overline{k_v}$ 的向量子空间.

3 混沌时序预测实例

考察 Mackey-Glass 延时微分动力学方程

$$\dot{x}(t) = \frac{0.2x(t - \tau)}{1 + x^{10}(t - \tau)} - 0.1x(t)$$

产生的混沌时间序列的预测问题, 这是一个广泛地使用的混沌时间序列的建模预测的用例. 我们使用 4 阶 Runge-Kutta 算法求解以上微分方程生成整数点时间序列值 $\{x_t\}_{t=1}^N$, 取 $x_0 = 1.2$, $\tau = 17$; 并且当 $t < 0$ 时, $t_t = 0$, 作为本文建模预测试验的数据, 原混沌时间序列数据如图 1.

根据通常的 MG 时间序列预测的惯例有, 该序列重构相空间的嵌入维 $m = 4$, 延时 $\Delta = 6$, 所以已知 t 时时间序列值 x_t , 构造的相点形式为 $X_i = (x_t, x_{t-6}, x_{t-12}, x_{t-18})$, 成为系统建模方法的输入量, 对

应的系统输出 x_{t+6} 就是系统在 t 时其轨迹的预测值. 通过对训练数据集中输入输出值对 (X_t, x_{t+6}) 的学习, 找出对应的函数映射关系 $x_{t+6} = F(X_t)$, 这里采用最小最大概率回归方法来解决.

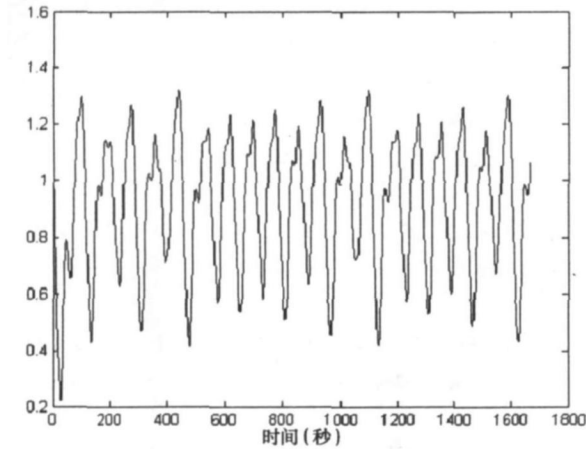


图 1 Mackey-Glass 混沌时间序列

构造 1 000 点输入输出值对 $\{x_{t+6}, x_t, x_{t-6}, x_{t-12}, x_{t-18}\}_{t=1}^{1117}$, 将这 1000 点前 500 点作为训练数据, 后 500 点作为测试数据. 使用最小最大概率回归方法对该系统数据进行建模学习, 这里核函数选用最为普遍的径向基函数 RBF, RBF 函数具有性质: $\varphi(\chi_i)^T(\chi_j) = e^{-\theta \|x_i - x_j\|^2}$. 对于核函数形状参数和回归管道宽度, 我们使用交叉验证, 将选取样本集的

90% 样本作为训练样本, 其余 10% 的样本作为测试样本, 验证了取 $\theta = 0.3 * d$, d 为模型输入量的维数 [2], 这里 θ 等于 2, 管道宽度 ϵ 等于 3.

以上通过最小最大概率回归方法建立的模型是一个单步预测模型, 为了实现多步预测, 我们应该把前面预测的结果作为新息并用于预测下一步值. 例如, 求得 $t+6$ 点预测值 $\overline{x}_{t+6}$ 后为了预测 $t+12$ 点值 $\overline{x}_{t+12}$, 必须使用 $\overline{x}_{t+6}$ 生成对应的输入量, 即 $\overline{x}_{t+12} = F(\overline{x}_{t+6}, x_t, x_{t-6}, x_{t-12})$. 如此反复从而实现多步预测, 这里我们利用最小最大概率回归方法进行 20 步预测. 使用两种指标评价系统的预测性能, 其一为平均绝对误差 MAPE, $MAPE = \frac{\sum_{t=1}^n |\frac{x_i - \overline{x}_t}{x_t}|}{n}$; 其二是均方根误差 RMSE, $RMSE = \sqrt{\frac{\sum_{t=1}^n (x_t - \overline{x}_t)^2}{n}}$. 其中 n 为进行多步预测的预测点数, 此处 $n=500$. 图 2 是在 500 点的测试数据集上进行 18 步预测的预测曲线, 可见系统预测的精度较好, 此时求得的 MAPE 和 RMSE 值分别为 0.045 28 和 0.051 747. 对测试数据集进行 1 至 20 步预测, 得到各种情况下的指标 MAPE 和 RMSE 见图 3 (图中标示为 MPMR), 表明随着预测步数的增加, 系统的预测性能相应下降, 也印证了混沌时间序列短期预测的可行性.

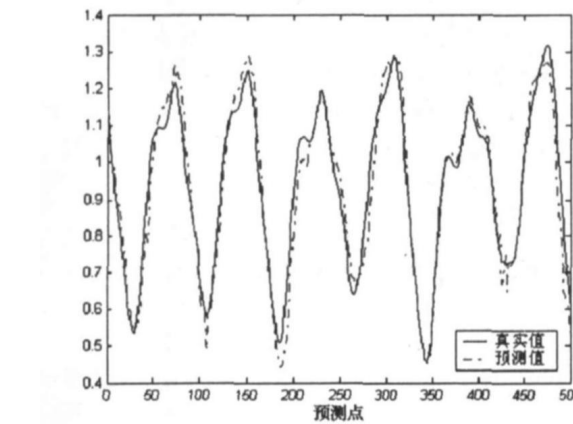


图 2 测试数据集 18 步预测曲线

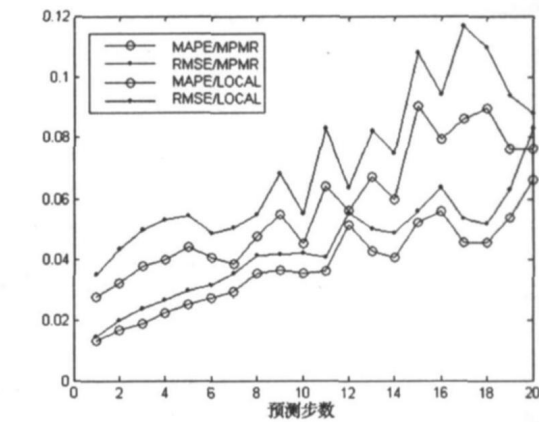


图 3 测试数据集 1 至 20 步预测标 MAPE 及 RMSE 比较

表 1 MPMR 方法和局域线性加权预测方法 (LOCAL) 预测指标数据

预测步数		1	5	10	15	20
MAPE $\times 10^{-2}$	MPMR	1.300 1	2.521	3.539 6	5.212 1	6.616 2
	LOCAL	2.754 4	4.403 3	4.510 5	9.030 2	7.637 5
RMSE $\times 10^{-2}$	MPMR	1.458 9	2.958 5	4.187 3	5.608	8.299
	LOCAL	3.495 9	5.467 3	5.518 4	10.804	8.787 1

作为比较,我们使用局域线性加权预测方法对该混沌数据序列进行了建模预测,在 500 点的测试数据集上进行了 1 至 20 步预测(此处近邻点数选为 8),预测性能指标见图 3(图中标示为 LOCAL)。可以看出,MPMR 方法的两种指标都优于局域线性加权预测方法,验证了最小最大概率回归方法可以很好地用来对混沌时间进行全局预测。表一显示了 MPMR 方法和局域线性加权预测方法对测试数据集进行 1, 5, 10, 15, 20 步预测的指标 MAPE 和 RMSE 的具体数据。

最小最大概率回归的优化算法的计算代价与训练样本的大小成正比,在试验中选取了 500 点训练数据集的前 200 点训练模型,在测试数据集上也获得了很好的预测精度。而使用人工神经网络试验时要求的训练样本多,模型的学习时间相对长,所以在对混沌时间序列进行全局预测方面 MPMR 方法的性能时间明显高于人工神经网络。

4 结 论

MPMR 方法利用凸优化技术及矩问题方法获得与分布无关的结果,能够直接求得预测值落在实际回归函数的 ϵ 值管道内的最小概率的界。使用 MPMR 方法对 Mackey-Glass 混沌时间序列进行了全局多步预测,获得了不错的性能时间比和预测精度。同局域加权线性回归预测方法比较,我们认为 MPMR 方法可以很好地用于混沌时间序列的全局预测。

今后,我们将 MPMR 方法推广到实际的工程应用中,并结合 MPMR 方法到混沌时间序列的局域预

测方法中。

参考文献:

- [1] Takens, F., Detecting strange attractors in fluid turbulence[J]. Springer lecture notes in mathematics (Springer, Berlin), 1981, 898, 366—381.
- [2] P. F. Verdes, P. M. Granitto, H. D. Navone and H. A. Cecatto, "Forecasting Chaotic Time Series: Global vs. Local Methods"[J]. in Novel Intelligent Automation and Control Systems, 1998, 1, 129—145.
- [3] K. A. de Oliveira, A. Vannucci, E. C. da Silva, Using artificial neural network to forecast chaotic time series[J]. Physica A, 2000, 284, 393.
- [4] Tronci S, Giona M, Baratti R, Reconstruction of chaotic time series by neural models: a case study[J]. Neurocomputing, 2003, 55, 581—591.
- [5] 丁涛, 周惠成. 混沌时间序列局域预测方法[J]. 系统工程与电子技术, 2004, 26(3), 338—340.
- [6] G. R. G. Lanckriet, L. E. Ghaoui, C. Bhattacharyya, and M. I. Jordan. Minimax probability machine[M]. In T. G. Dietterich, S. Becker, and Z. Ghahramani, editors, Advances in Neural Information Processing Systems 14, Cambridge, MA, 2002. MIT Press.
- [7] T. Strohmann and G. Z. Grudic. A Formulation for Minimax Probability Machine Regression[C]. Accepted for Publication at NIPS * 2002, Neural Information Processing Systems: Natural and Synthetic, Monday, Dec. 3 — Saturday, Dec. 8, 2001 Vancouver, British Columbia, Canada.
- [8] I. Popescu and D. Bertsimas. Optimal inequalities in probability theory: A convex optimization approach[R]. Technical Report TM62, INSEAD, Dept. Math. O.R., Cambridge, Mass., 2001.

Chaotic Time Series Prediction Based on Minimax Probability Machine Regression

LIU Zun-xiong, LIU Jian-hui

(School of Electronics and Electric Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: Minimax Probability Machine Regression (MPMR) is proposed for chaotic time series global prediction in this paper. In MPMR, regression function maximizes the minimum probability that future predication will be within an ϵ to the true regression function. Multi-step predictions up to 20 steps were done on Mackey-Glass chaotic time series with MPMR and Local Weighted Linear Regression (LWLR). The results demonstrated that MPMP had satisfactory prediction efficiency, compared with LWLR. Kernel function shape parameter and regression tube value will influence the MPMR-based system performance. In experiments, we use cross validation to choose the two parameters.

Key words: chaotic time series; global prediction; minimax probability; regression