



28-33

基于径向基函数的非线性预测模型^①

王明进¹, 程乾生²

0211.61

(1. 北京大学光华管理学院, 北京 100871; 2. 北京大学数学科学学院, 北京 100871)

摘要: 针对混沌时间序列的预测问题提出一种兼有全局和局部特点的新的预测模型, 通过数值实验探讨这一模型的性能和特点, 并给出对太阳黑子数据较以往文献更好的预测结果。

关键词: 混沌时间序列; 径向基函数; 太阳黑子数据

非线性预测

分类号: N94 **文献标识码:** A **文章编号:** 1007-9807(1999)04-0028-06

0 引言

一个混沌时间序列是指对某一确定的混沌动力系统观测采样而得到的单变量(也可以是多变量)的时间序列, 以 $\{s_t\}$ 记之。混沌序列研究的中心问题在于能否以及如何根据 $\{s_t\}$ 的变化来推知整个动力系统的性质。这是一个具有重要理论及应用价值的课题。这一研究的理论基础在于重构相空间理论, 即可以根据 $\{s_t\}$ 在 R^m 中构造一条轨道 $x(t)$,

$$x(t) = (s_t, s_{t-\tau}, s_{t+2\tau}, \dots, s_{t+(m-1)\tau}) \quad (1)$$

其中 m 称为嵌入维数, τ 为时滞。按照 Takens^[1] 的结论, 一般地如果 $m \geq 2d + 1$ (d 表示原动力系统的相空间的维数), 则 $x(t)$ 是原动力系统相应的一条轨道到 R^m 中的嵌入。由此即可得到 R^m 上的一个离散的动力系统 $F: R^m \rightarrow R^m$, 使得

$$x(t+1) = F(x(t)) \quad (2)$$

或者得到一个函数 $F: R^m \rightarrow R$, 满足

$$s_{t+(m-1)\tau+1} = f(s_t, s_{t+\tau}, \dots, s_{t+(m-1)\tau}) \quad (3)$$

通常所说的非线性预测问题即根据 $\{s_t\}$ 如何确定 f 或构造 f 的一个近似形式 \hat{f} (也可以是对 F)。当然由于系统的混沌性质, \hat{f} 或 \hat{F} 首先应该是非线性的。自从 80 年代后期以来, 这方面的研究已得到了大量的成果。比如, 局部线性模型^[2]、神经网络预测模型 (BP 网络模型^[3] 和小波网络模型^[4])、径向基函数预测模型^[5,6]、基于不变测度

的函数重构模型^[7] 等, 主要分成局部预测模型和全局预测模型两类^[8], 其中在整个重构吸引子上寻求 \hat{f} 的统一形式即全局预测模型, 相反, 在待预测点的某一邻域内建立的预测模型属于局部模型。

在文[9]中提出的如下的一类预测模型也属于全局模型:

$$\hat{f}(x) = \sum_{j=1}^K \lambda_j \Phi(\|x - c_j\|) + \mu x + \mu_0 \quad (4)$$

其中 $\Phi(r): R^+ \rightarrow R$ 是径向基函数, c_j 称为中心向量。先采用 Kohonen 自组织网络选择 $\{c_j\}$, $j = 1, 2, \dots, K$ 。再用最小二乘法来确定参数 $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_K)$, μ, μ_0 。在下文中将其记成模型 I。

全局预测模型一旦由学习集合确定下来, 就得到了一个固定的模型, 任何一个待预测的点都可以由这一模型求出来; 而做为局部模型, 为了预测一个点后面的行为, 首先要确定这一点在学习集中的邻域, 一旦这一邻域被确定下来, 给出的模型就能很好地反映出吸引子的局部性质, 并且, 一个明显的特征是随着训练集合所含样本个数的增加, 吸引子会逐渐呈显示更加精细的结构, 此时局部模型给出的预测效果也会更好。这一点对于全局预测模型是不成立的。但是, 随着样本向量的逐渐增加, 要对每一个待预测的点都到样本集中搜寻它的一个邻域所花费的时间就会越来越长, 而且, 局部模型显然不是一个光滑

① 收稿日期: 1998-09-28; 修订日期: 1999-06-28。

作者简介: 王明进 (1970-), 男 (汉族), 山东惠民县人, 理学博士, 北京大学光华管理学院讲师。

的函数.

为了尽量兼顾全局和局部两种模型上面所述的各自的优点并弥补其不足,做为一种尝试,本文在模型 I 的基础上提出了一种新的预测模型,与以往单纯的全局或局部的模型不同,不妨称其为介于全局与局部的预测模型.

1 模型的描述

新的预测模型是在模型 I^[9] 的基础上来完成的. 径向基函数插值是一种“具有很好的局部性质的全局插值”^[5], 这意味着如果将径向基函数的中心选在某一个点的小邻域内, 则可以得到对这个点附近动力系统性质的局部刻画. 因此, 首先利用 Kohonen 自组织神经网络由给定的学习样本集合输出 K 个中心, 这样按照各自到哪一个中心的距离最近相应地也将吸引子上的点分归于 K 个不同的区域, 接着针对每个不同区域, 分别构造不同的全局模型 $\tilde{f}_1, \tilde{f}_2, \dots, \tilde{f}_K$, 每一个 \tilde{f}_i ($i=1, 2, \dots, K$) 既是一种对系统整体的刻画(它是整个吸引子

区域上的光滑函数), 又能够反映出吸引子在第 i ($1 \leq i \leq K$) 个区域里的一些比较细致的结构, 尤其是在各自中心附近的位置. 这样, 当要预测某一个点的后继行为时, 只需首先判断一下这一点属于哪一类, 并由此选择相应的一个 \tilde{f}_i 做为预测所用的模型, 这样避免了局部模型中所要求的要针对待预测的每一个点, 都要在训练集中搜索其一个局部邻域的复杂且费时的过程, 特别是在该集合容量很大的时候.

定义 \tilde{f}_i ($i=1, 2, \dots, K$) 的方式可以有許多, 在这里不妨仍选择径向基函数插值的形式. 设学习样本向量经由 Kohonen 神经网络所产生的 K 个中心分别为 c_i , $i=1, 2, \dots, K$. 这样, 整个学习样本集合被划分成 K 个子集, 记成 $\bigcup_{i=1}^K A_i$, 其中向量属于 A_i 当且仅当该向量到 c_i 的距离比到其它中心的距离要小. 设集合 A_i 中向量的个数为 N_i , 对于某一个 i , 在 A_i 中 c_i 的附近按照一定方式选取 n_i 个样本点 $y_{i1}, y_{i2}, \dots, y_{in_i}$, 将 \tilde{f}_i 写成如下的形式

$$\tilde{f}_i(x) = \sum_{l=1}^{N_i} \lambda_l \Phi_1(\|x - c_i\|) + \sum_{l=1}^{n_i} \mu_l \Phi_2(\|x - y_{il}\|) + \nu \cdot x + \nu_0 \quad (5)$$

其中 $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_{N_i})'$, $\mu = (\mu_1, \mu_2, \dots, \mu_{n_i})'$, ν, ν_0 是待定的参数, $\Phi_1(r)$ 和 $\Phi_2(r)$ 是径向基函数, 在本文中, 均采用 Gaussian 型的径向基函数

$$\Phi_p(r) = \exp(-r^2/\sigma_p^2), \quad p=1, 2 \quad (6)$$

考虑到在文[9]中所运用的预测模型, 也可以将式(5)写成如下的形式:

$$\begin{aligned} \tilde{f}_i(x) &= \mu_i g_0(x) + g_i(x) \\ &= \mu_i g_0(x) + \sum_{l=1}^{n_i} \mu_{il} \Phi_2(\|x - y_{il}\|) \end{aligned} \quad (7)$$

其中 $g_0(x)$ 即是按照文[9]的方法得出的形如式(4)的预测模型, 它的系数是经过对全体样本向量的学习而得到的; 为了进一步得到 $\mu = (\mu_i, \mu_{i1}, \dots, \mu_{in_i})'$, 采用类似于逐次逼近的方式 (successive approximation method)^[6], 再次利用最小二乘法, 此时的学习样本为 A_i 中的向量, 设为 x_{i1}, \dots, x_{in_i} , 即求解如下的方程:

$$d_i = \tilde{A}_i \cdot \mu_i, \quad (8)$$

其中

$$d_i = (s_{i1+(m-1)\nu_i+1}, s_{i2-(m-1)\nu_i+1}, \dots, s_{in_i+(m-1)\nu_i-1}) \quad (9)$$

$$\tilde{a}_i(k, 1) = g_0(x_{ik}) \quad (10)$$

$$\begin{aligned} \tilde{a}_i(k, (l+1)) &= \Phi_2(\|x_{ik} - y_{il}\|); \\ 1 \leq k \leq N_i, 1 \leq l \leq n_i, \end{aligned} \quad (11)$$

(8) 的最小平解为

$$\mu_i = \tilde{A}_i^+ \cdot d_i, \quad (12)$$

由此即可得到一种新的预测模型(7). 容易看出这样得到的 \tilde{f}_i 仍然是吸引子整体范围的一个光滑的函数.

一般地, 常常选择 $\sigma_1^2 \geq \sigma_2^2$, 为的是突出式(5)中第一、二两项是对吸引子不同尺度的插值. 当然还可以将 $g_i(x)$ 选择成其它的形式, 比如多项式函数等. 为了叙述的方便, 在下文中称(5)为模型 II.

在给出数值实验的结果之前, 值得一提的是,

将上面所提出的预测模型同汤家豪(H. Tong)提出的阈值自回归模型(Threshold Autoregression Model,简称 TAR^[2])做一下对比是很有启发意义的.一个典型的也是最为简单的阈值自回归模型是如下的一阶模型

$$x_t = a_{t-1} + a_t \cdot x_{t-1} + \varepsilon_t, \gamma_{j-1} \leq x_{t-1} < \gamma_j \quad (13)$$

其中 $-\infty = \gamma_0 < \gamma_1 < \dots < \gamma_{K-1} < \gamma_K = +\infty$ 定义了实轴上的一个划分.整个模型是由 K 个一阶自回归线性模型组成的,为了预测 x_t 的值,首先必须要确定 x_{t-1} 所属的区间,由此决定使用哪一个自回归模型.这与本节算法中先要确定预测点的所属类别,再以此选择所要采用的非线性预测模型的道理是一样的.

2 几个例子

下面给出几个混沌数据的例子来分析一下上一节提出的预测模型的一些性质.

首先选用具有典型意义的 Henon 数据和 Lorenz 数据,在文[9]数值试验的基础上继续本文所提出的算法的检验.在数值实验当中,选择了 $\sigma_1^2 =$ 点到各自中心距离平方的平均值, $\sigma_2^2 = 25 \times \sigma_1^2$;为了简单起见,还令 $n_i = 2m + 1, i = 1, 2, \dots, K$, 并且在 A_i 中的 n_i 个 $\Phi_{i,c}(r)$ 的中心位置都直接地选为距离 c_i 最近的 n_i 个点.这里的选择都是主观设定的(其它的选择很有可能会得到更好的预测结果,在此不做细究).两种数据实验的部分结果分别列在表 1 和表 2 中.做为对比,同时列出了应用模型 I^[9] 的预测结果.表 1、2 中的 σ_2^2 表示预测的正规化均方误差, N_L 表示用于建模的重构向量集合即学习集中元素的个数^[9].

表 1 Henon 数据的部分预测结果

m	r	N_L	K	n_i	模型 I $\log(\sigma_2^2)$	模型 II $\log(\sigma_2^2)$
2	1	100	4	5	-3.825 2	-5.369 9
2	1	300	4	5	-4.158 3	-5.412 7
2	1	500	4	5	-3.985 2	-5.413 8
2	1	700	4	5	-4.415 8	-5.540 6
2	1	100	8	5	-10.337 4	-11.325 8
2	1	300	8	5	-12.220 2	-13.659 9
2	1	500	8	5	-12.355 6	-13.976 3
2	1	700	8	5	-12.225 5	-13.734 5

表 2 Lorenz 数据的部分预测结果

m	r	N_L	K	n_i	模型 I $\log(\sigma_2^2)$	模型 II $\log(\sigma_2^2)$
3	1	200	4	7	-6.375 4	-6.600 0
3	1	400	4	7	-5.763 3	-6.880 1
3	1	600	4	7	-5.837 9	-6.964 6
3	1	800	4	7	-6.079 9	-7.359 2
3	1	200	8	7	-9.604 4	-10.051 4
3	1	400	8	7	-8.886 7	-9.683 0
3	1	600	8	7	-9.255 3	-10.432 6
3	1	800	8	7	-9.693 7	-10.876 7

分析表 1 和表 2 可以看出利用本文所提出的模型 II 的预测效果,一般都会比单纯地使用全局模型(3)(模型 I)的预测效果有不同程度的改进;另外,本章算法的预测误差基本上呈现出了随着学习样本个数的增加而减少的趋势,这种趋势性反映了样本增多时,各类中心 $c_i, i = 1, \dots, K$ 附近吸引子的越来越精细结构得到了体现,从而

为中心近旁的一些点的预测提高了精度.这一点符合局部预测模型的特点.因此,它是一种介于全局和局部之间的模型,是一种折中的形式.

3 太阳黑子数据的预测

做为两种预测模型(即模型 I 和模型 II)的

应用,下面考虑太阳黑子数据(Sunspot Data^[11])的预测问题。(当然,这并不意味着太阳黑子数据是来自一个混沌动力系统的混沌时间序列,这是另外的一个问题)。利用从 1770 年到 1979 年的一

段观测数据,该段数据在近几年的文献中曾被利用基于神经网络的预测模型^[12]、径向基函数逐次逼近模型^[6]等比较新的预测方法加以研究,并取得了较为理想的预测效果。

表 3 Sunspot 数据的部分预测结果

m	K	n_i	模型 I			模型 II		
			Var_f	$MSE_{1921-1955}$	$MSE_{1956-1979}$	Var_f	$MSE_{1921-1955}$	$MSE_{1956-1979}$
3	4	4	202.6	229.8	659.2	182.9	193.9	530.5
3	6	7	179.1	226.9	704.6	139.2	180.2	581.6
3	8	7	175.7	223.2	634.7	123.6	185.9	453.9
3	10	7	164.4	208.5	750.9	116.5	167.5	580.9
3	12	7	159.3	172.5	947.9	108.8	220.7	840.5
3	14	4	154.3	147.8	571.9	120.6	171.1	577.4
3	15	5	150.8	140.3	562.8	113.8	164.7	532.7
3	16	4	147.0	136.7	560.6	114.2	160.7	563.1
3	18	5	146.5	141.9	486.7	109.4	188.6	485.0
3	20	5	143.8	140.2	422.9	102.0	182.6	415.3
4	18	5	162.6	288.4	865.5	101.5	249.7	815.7
6	12	5	151.0	447.2	333.2	129.1	408.8	292.9

为了以往的文献中结果进行对比,采用和他们相同的学习集合与检验集合,即利用从 1700 年到 1920 年的数据建立模型,此时 $N_L = 218$,用 1921—1955 的一段数据和 1956—1979 的一段数据分别做为模型的检验时用。选取时滞参数 $\tau = 1$,嵌入到不同维数的欧氏空间中去,并且选择 $\sigma_1^2 = 5000$, σ_2^2 同前面例子,表 3 中给出了两种模型对太阳黑子数据的部分预测结果。其中, $Var_f = \sigma_1^2 \cdot Var(s)$, $MSE = \sigma_1^2 \cdot Var(s)$ (见文[9])。

从表 3 中所列出的结果看到,模型 II 基本上保证了在拟合效果上比模型 I 更加精确,至于在某些情况下预测的效果反而比模型 I 更差些,一

方面可以理解为检验集合比较短,尚不能反映出预测误差的变化趋势;另一方面如果检验集合中的点距离各自中心比较远时也会出现类似情况。我们看到嵌入 $m = 3$ 维空间选择 $K = 16$, $n_i = 4$ 时由模型 I 给出了 1921—1955 年段中最为理想的预测结果(当然是在所进行的数值实验范围内),而 $m = 6$, $K = 12$, $n_i = 5$ 时由模型 II 给出了 1956—1979 年段中最好的预测结果,综合来看,选择 $m = 3$, $K = 20$, $n_i = 5$ 所得到的预测效果是比较令人满意的。

做为对比,在下面列出在其他文献中使用别的方法所得到的预测结果:

表 4 Sunspot 数据预测结果的比较

	Var_f	$MSE_{1921-1955}$	$MSE_{1956-1979}$
阈值自回归模型(TAR)(转引自 ^[12])	148.9	148.9	429.8
Weigend et al. 文[12]中神经网络模型	125.9	132.0	537.25
He et al. 文[6]中径向基函数逐次逼近模型	142.3	140.8	
径向基函数严格插值模型(转引自文[6])	10^{-3}	10^5	

结合表3及表4,可以看到,当选择 $m=3$, $K=14\sim 20$ 利用模型I给出的对1921—1955年段预测的结果都好于阈值自回归模型的结果,且大多也都好于He. et al.文[6]按照径向基函数逐次逼近模型得出的预测结果.尤其是 $K=16$

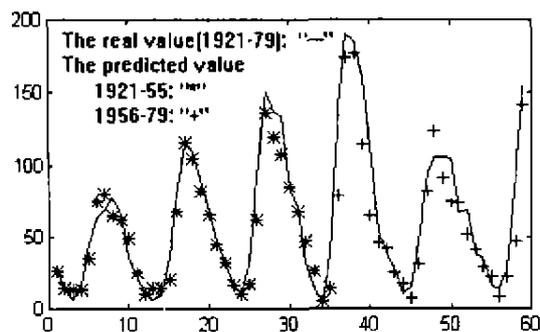


图1 Sunspot Data 的预测值与真值的比较

时的结果比较接近于文[12].如果单纯考虑对1956—1979年段内的预测的话,显然 $m=6$, $K=12$, $n=5$ 时两种模型得到的结果相比上述的方法都好的多.综合比较可以看出,选择 $m=3$, $N=20$, $n=5$ 时得到的对两个年段内的预测效果应该是更为可取的.

参考文献:

- [1] Tankens F. Detecting strange attractors in turbulence[M]. In Rand D, Young L S. ed. Dynamical System and Turbulence. Springer, 1981, 366
- [2] Farmer J D, Sidorowich J J. Predicting chaotic time series[J]. Phys. Rev. Lett., 1987, (24): 845~848
- [3] Gencay R. Nonlinear prediction of noisy time series with feedforward networks[J]. Phys. Lett. A, 1994, (187): 397~403
- [4] Cao L, Hong Y, Fang H, He G. Predicting chaotic time series with wavelet networks[J]. Physica D, 1995, (85): 225~238
- [5] Casdagli M. Nonlinear prediction of chaotic time series[J]. Physica D, 1989, (35): 335~356
- [6] He X, Lapedes A. Nonlinear modeling and prediction by successive approximation using radial basis functions[J]. Physica D, 1994, (70): 289~301
- [7] Giona M, Lentini F, Cimagalli V. Functional reconstruction and local prediction of chaotic time series[J]. Phys. Rev. A, 1991, (44): 3496~3502
- [8] Weigend A S, Gershenfeld N A. ed. Time series prediction; forecasting the future and understanding the past[J]. SFT Studies in the Sciences of Complexity, Vol. XV, Addison-Wesley, 1993
- [9] 王明进, 程乾生. Kohonen 自组织网络在混沌时间序列预测中的应用[J]. 系统工程理论与实践, 1997, 17(7): 12~17
- [10] Tong H. Non-linear time series; a dynamical system approach[M]. Oxford Univ. Press, 1990
- [11] Yule G. On a method of investigating periodicity in disturbed series with special reference to Wolfer's Sunspot numbers[J]. Phil. Tran. Roy. Soc. London A, 1927, (226): 267~298

在图1中给出了选择 $m=3$, $K=16$ 时由模型I得到的预测结果.图中仅列出了对检验集合的一步预测情况,其中,'*'表示1921—1955年段的预测值,'+'表示1956—1979年段的预测值.

4 结束语

本文可以看作是文[9]的继续.在Kohonen自组织网络对吸引子进行聚类的基础上,进一步提出了对混沌时间序列的一种新的预测模型,它实际上是产生几个类就包括几个模型,每个模型既保留了函数全局的光滑性又兼顾了对在相应类中心处吸引子的局部性质的刻画.因此它兼有全局和局部模型的一些特点,比如在聚成同类数的情况下,预测效果将会随着样本个数的增加而变好.这种模型体现了一些同阈值自回归模型[10]相似的思路.进一步研究了两种模型在实际数据中的预测效果,尤其是对太阳黑子数据得到了十分理想的结果.

- [12] Weigend A S, Huberman B A, Rumelhart D E. Predicting sunspots and exchange rates with connectionist networks [J]. in: *Nonlinear Modeling and Forecasting, SFI Studies in the Sciences of Complexity, Proc. Vol. XII*, Eds. M. Casdagli & S. Eubank, Addison-Wesley, 1992, 395~432

Nonlinear prediction model based on radial basis functions

WANG Ming-jin¹, CHENG Qian-sheng²

1. Guanghua School of Management, Peking University, Beijing 100871;

2. School of Mathematical Science, Peking University, Beijing 100871

Abstract: In this paper, a new nonlinear prediction model of chaotic time series which has both the global and local properties has been given and explained in detail. Firstly, the Kohonen Self Organization Network is used to produce K centers and to cluster the points of the reconstructed attractor using time-delay coordinate into K classes; Secondly, a nonlinear model based on radial basis functions is fitted to every class. So the prediction model is composed of K functions which are all defined on the whole attractor and also have the good properties near their correspondent centers, such as the predicting ability can be improved by increasing size of the learning set. Compared with the local prediction model, this new model has an explicit form (only K functions) and avoids the long time searching for the local points as in the previous one.

The predicting power and the specific properties are confirmed by the numeric experiments in the paper. Henon data and Lorenz data are used as the examples of chaotic time series and different learning sets and centers are chosen. Another model which has been developed in our previous work is also used at the same time to compare the results. At last, we apply these models to predict the sunspot data and obtain the lower mean square errors than using the connectionist network and some other models in the previous references.

Keywords: chaotic time series; radial basis function; sunspot data