

基于小波网络的非线性经济时序预测模型

方先明¹ ,唐德善²

(1. 河海大学商学院 ,江苏 南京 210098 ; 2. 河海大学水利水电工程学院 ,江苏 南京 210098)

摘要 :为对经济时序准确预测 ,必须先对其数据结构进行分析 ,相空间重构技术为之提供了理论基础 .通过关联维数的计算 ,区分确定性系统和随机系统 .在此基础上确定最佳嵌入维数、最佳采样时间间隔及小波元的个数 ,并通过带有偏差单元的递归小波网络的学习 ,进行模型参数的辨识 .实验研究表明 ,模型对非线性经济时序具有良好的逼近能力 ,因此该模型用于非线性经济时序预测具有可行性 .

关键词 经济时序 ;关联维 ;小波网络 ;预测

中图分类号 :F224.9 文献标识码 :A 文章编号 :1000-198X(2004)03- 0344-05

经济系统是一个服从耗散结构的、非线性、开放的动态系统^[1,2] ,需要把非线性科学理论与方法应用于经济系统的研究 ,以分析系统的“ 不稳定性 ”、“ 结构变化 ”、“ 时空尺度 ”等^[3~7] .经济时序预测作为非线性经济研究的内容之一 ,主旨在于吸引子的刻画 ,因此其可以分为两种不同的形式 :一是建立直接的非线性模型 ,如逻辑斯蒂模型、洛伦兹模型、埃龙映射等 ;二是从已有的混沌时间序列出发 ,通过相空间重构技术重构吸引子 ,用逆处理非线性预测方法对系统的未来行为进行预测 .在逆处理非线性预测方法中 ,利用小波变换良好的时频局域化性质 ,结合神经网络的自学习能力 ,可建立基于小波网络的非线性经济时序预测模型 .通过网络的训练进行参数辨识 ,从而为非线性经济时序预测提供一种方法 .

1 数据分析

1.1 关联维数

设 ϵ 为一个正数 , m 为嵌入维数 ,考虑状态空间的 N 个点在 m 维嵌入空间中有多少“ 点对 ”之间的距离小于 ϵ ,把这些“ 点对 ”数在所有“ 点对 ”中所占的比例记为关联函数^[8] ,即

$$C(m, N, \epsilon, \tau) = \frac{1}{N^2} \sum_{i,j} H(\epsilon - |x_i^m - x_j^m|) \tag{1}$$

$$\hat{N} = N - (m - 1)\tau \qquad H(z) = \begin{cases} 0 & z < 0 \\ 1 & z \geq 0 \end{cases}$$

式中 : τ ——延滞参数 ; z ——函数自变量 ; $|\cdot|$ ——最大模 .

在实际应用中 ,一般根据所取的 ϵ 值 ,作 $\ln C(m, N, \epsilon, \tau) \sim \ln \epsilon$ 曲线 ,其直线部分的斜率就是关联维数 $d_2(m)$ 的估计值 ;不断提高嵌入维数 m ,重复上述步骤 ,直到 m 达到某一值 m_c 时 ,相应的 $d_2(m)$ 不再随 m 的增大而发生有意义的变化为止 .此时的嵌入维数 m_c 称为饱和嵌入维数 ,在 $\ln C(m, N, \epsilon, \tau) \sim \ln \epsilon$ 中一系列曲线的直线部分的斜率不再随 m 的变化而变化 ,此时 $d_2 = d_2(m_c)$ 就是吸引子的关联维数 .

饱和嵌入维数 m_c 的存在与否 ,决定了被分析的时间序列的性质 .如果 $d_2(m)$ 的估计值随 m 的增大趋向无穷 ,则被研究的系统是一个随机系统 ;如果 m_c 存在 ,则所研究的系统是一个有限维的确定系统 ,此时 m_c 的大小即表征了系统自由度数目的多少 . m_c 与 d_2 分别给出了系统所包含的基本变量数目的上限和下限 .

1.2 m 的选取

m 值关系相空间动力学行为的最佳表示 , m 太小不足以展示复杂行为的细致结构 , m 太大则费时费力 .Takens 证明 ,只有当 m 足够大时 ,重构的动力系统与原系统才是几何等价的 .一般地 ,当 $m \geq \text{in}(2d_2 + 1)$,则

得到的就是原来动力系统相应的一条轨道到 R^m 中的嵌入.

1.3 最佳采样时间间隔 τ 的选取

文献 [8] 中提到, Theiler 建议可以通过计算时序的自相关函数 $\Psi(\tau)$ 来求得最佳的时序采样间隔 τ .

$$\Psi(\tau) = \frac{1}{N} \sum_{t=1}^{N-\tau} (x_{t+\tau} - \bar{x})(x_t - \bar{x}) \approx \frac{1}{e}$$

(2)

式中 $\bar{x}$ 为时间序列 x_t 的均值.

2 基于小波网络的预测模型

小波分析源于信号分析,是泛函分析、Fourier 分析、样条分析、调和分析、数值分析的完美结晶.小波网络是以小波函数为基底的一种新型函数连接神经网络.它以小波空间作为模式识别的特征空间,通过将小波基与信号向量的内积进行加权来实现信号的特征提取.结合了小波变换良好的时频局域化性质及传统神经网络的自学习能力,综合了神经网络和小波分解的优点,因而具有良好的逼近与容错能力.

2.1 模型的提出

设混沌时序 $x_1, x_2, x_3, \dots, x_N$ 最佳嵌入维数为 m , 采样时间间隔为 $\tau_i, i = 1, 2, 3, \dots, N, \Delta\tau_i = \tau_{i+1} - \tau_i$ 满足最佳采样时间间隔, 则可以在相空间中重构 m 维相量 $\mathbf{v}(n)$,

$$\mathbf{v}(n) = (x_n, x_{n+1 \cdot \tau}, x_{n+2 \cdot \tau}, \dots, x_{n+(m-1) \cdot \tau})$$

(3)

由此得

$$\mathbf{v}(1), \mathbf{v}(2), \mathbf{v}(3), \dots, \mathbf{v}(N - m + 1)$$

(4)

预测问题即利用式 (4) 求 $\mathbf{v}(N - m + 2), \mathbf{v}(N - m + 3), \dots$. Takens 证明了在 $\ln(d_2 + 1) \leq m \leq \ln(2d_2 + 1)$ 条件下, 存在 R^m 上的函数 G 和参数 α 满足式 (5)

$$\mathbf{v}(n + 1) = G(\mathbf{v}(n), \alpha)$$

(5)

如果已经得到了映射 $G(\mathbf{v}, \alpha)$ 和已知 R^m 上的向量 $\mathbf{v}$, 则

$$\mathbf{v} \rightarrow \mathbf{v}(1) = G(\mathbf{v}, \alpha) \dots \mathbf{v}(n - 1) \rightarrow \mathbf{v}(n) = G(\mathbf{v}(n - 1), \alpha) = G^n(\mathbf{v}, \alpha)$$

(6)

可见预测问题就是构造映射函数 $G(\mathbf{v}, \alpha)$, 并确定参数 α . 对于非线性确定系统产生的时间序列而言, 这种函数是确定的, 因而预测问题与函数逼近问题等价, 即对于有 n 个自变量的函数, 在 $n + 1$ 维空间内对超曲面进行拟合. 于是可利用小波级数的有限项来逼近非线性映射函数 $G(\mathbf{v}, \alpha)$.

令

$$\hat{G}(X) = \exp\left\{-\frac{1}{2}\left[\exp\left(-\frac{h((UX - b)/a) \cdot \omega_1 + X \cdot \omega_r + z_1 \cdot \omega_{\text{bias}}}{2}\right)\omega_2 + z_2 \cdot \omega_{\text{obias}}\right]^2\right\}$$

(7)

$$h(a, b, t) = \frac{1}{\sqrt{|a|}} h_{\text{basic}}\left(\frac{t - b}{a}\right)$$

式中: ω_1, ω_2 ——权值; z_1, z_2 ——偏差值; $\omega_{\text{bias}}, \omega_{\text{obias}}$ ——偏差单元的权值; ω_r ——回归单元权值; $h(a, b, t)$ ——小波; h_{basic} ——基本小波; $\frac{1}{\sqrt{|a|}}$ ——归一化系数; a, b ——小波基的伸缩因子、平移因子.

2.2 参数辨识

模型 (7) 的目标函数为

$$J(E) = \frac{1}{2} \sum_{t=1}^{N-m+1} (\hat{G}(x_t) - x_t)^2$$

(8)

$$E = [\omega_1 \ \omega_2 \ \omega_{\text{bias}} \ \omega_{\text{obias}} \ a \ b \ U]^T$$

于是参数辨识问题变为

$$\min_E J(E)$$

具体算法可借鉴 LMS 迭代算法, 为避免网络陷入局部极小点, 采用带有偏差单元的递归网络, 通过网络的学习确定模型参数.

2.3 确定小波元、小波元维数及小波元个数

小波元的选择目前尚无理论上的指导原则, 只能根据不同的研究背景选择不同的小波元. 一般 Morlet 小波多用于分类、图像识别和特征提取, 高斯一阶导数多用于函数估计, 墨西哥草帽小波多用于系统辨识. 文章

采用余弦调制的高斯波,即令 $h(t)=\cos(rt)\exp(-\frac{t^2}{2})$.

对于由时间延迟序列构造的系统吸引子,文献 [10,11] 中给出了相关数学证明,如果 f 是光滑的,则存在一个最小的上限,如果吸引子的维数是 d_2 ,则嵌入维 $m\geqslant 2d_2+1$. 由此构造网络的输入矢量 X , $X=\{x_i\}_{i=1}^m$,小波元的维数随之确定^[9].

有关小波元个数的确定问题,相关研究较多,但未见有突破性的进展,具体可参见文献 [10,11] 中的论述.

2.4 预测长度

预测长度 k 通过 Kolmogorov 熵求得,该数值是一个描述时间序列混沌特征的重要参数^[12]. 当 k 等于某一有限值时系统是混沌系统,且混沌程度随 k 的加大而加大. 此时只能对系统作短期预测, $1/k$ 为系统的平均可预报时间尺度. 运用 G-P 方法:

$$k=\frac{1}{\varepsilon}\ln\frac{\alpha(m,N,\varepsilon,\tau)}{\alpha(m+1,N,\varepsilon,\tau)}$$

(9)

在实际应用中,通常将 k 随 $(m+1)$ 变化的稳定值作为 k 的估计值.

2.5 模型检验

模型检验选用评价预测模型效果的常用指标:

平方和误差

$$S=\sum_{t=1}^n(x_t-\hat{x}_t)^2$$

均方误差

$$M=\frac{1}{n}\sqrt{\sum_{t=1}^n(x_t-\hat{x}_t)^2}$$

3 实验研究

3.1 数据分析

利用江苏省 1996~2000 年工业生产总值资料^①,对上述理论模型进行实验研究.

实测数据只有 50 个,采用了 3 次样条插值得 1471 个实验数据. 由图 1 可以看出关联函数 $\ln C(m,N,\varepsilon,\tau)$ 对度量尺度 $\ln\varepsilon$ 存在一个线性依赖区. 这些直线段的斜率就是关联维数的估计值. 同时,图 1 显示随着 m 的增加,图中直线部分趋于平行,这说明关联维数的估计值随着 m 的增加趋于一个定值,这个定值即为要求的关联维数. 根据图 1 及表 1 可得饱和嵌入维 $m_c=7$,关联维数的收敛值 $d_2=1.76$. 由式 (2)(9) 得最佳采样时间间隔 $\tau=6\cdot\Delta t$,平均预测长度为 10.

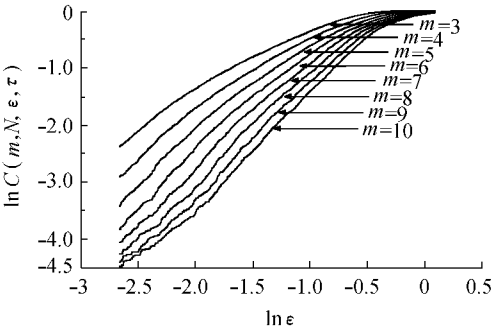


图 1 $\ln C(m,N,\varepsilon,\tau)\sim\ln\varepsilon$ 曲线

Fig.1 Relation between $\ln C(m,N,\varepsilon,\tau)$ and $\ln\varepsilon$

3.2 网络训练

取小波元个数为 8,最佳嵌入维数为 4,利用原始数据(1466 个数据用于训练,后 5 个数据用于模型检验)

① 江苏统计信息网 <http://www.jssb.gov.cn/sjzl/tjnj.htm>.

对网络进行训练 ,训练结果见图 2 3.

表 1 嵌入维 m 和关联分维 D 的关系

Table 1 Relation between embedded dimension m and correlation fractal dimension D								
m	3	4	5	6	7	8	9	10
D	1.112	1.167	1.31	1.39	1.71	1.76	1.80	1.82

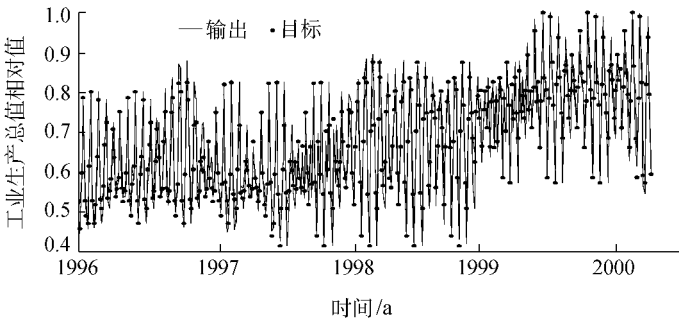


图 2 网络训练结果
Fig.2 Network training results

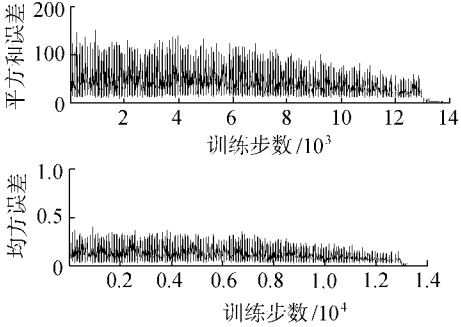


图 3 误差曲线
Fig.3 Variation of error

3.3 预测

利用训练好的网络对江苏省工业生产总值进行预测 ,结果见表 2.

表 2 外推预测结果及误差

Table 2 Results of prediction and error					
预测点位置	模型输出值/ 10^8 元	相对误差/%	预测点位置	模型输出值/ 10^8 元	相对误差/%
1467	2 055.5	- 4.477 6	1472	2 300.5	—
1468	3 077.4	0.048 4	1473	2 226.7	—
1469	3 502.1	- 0.545 1	1474	2 301.5	—
1470	2 988.2	0.217 5	1475	2 225.7	—
1471	2 315.2	3.731 9	1476	2 302.6	—

由表 2 可知 ,模型对原始数据具有良好的逼近能力 ,相对误差较小 ,误差有波动 ,这是由于原始数据有波动 ,加之网络有一个自学习过程而引起的 ,预测误差随预测步数的增加有增大的趋势 ,这验证了混沌时序只能作短期预测的论断.

4 结 论

- a. 构造的映射 $\hat{G}(v, \alpha)$ 也许和真正的映射 $G(v, \alpha)$ 不完全一样 ,但只要通过检验 ,满足短期预测的要求 ,就可以接受.
- b. 用带有偏差单元的递归小波网络进行预测 ,虽增加了模型的复杂程度及参数辨识的难度 ,但其对非线性时序具有良好的逼近 ,预测误差较小.
- c. 关于小波网络的广泛应用 ,仍有一些关键问题有待解决 ,如母小波的选取、小波元的个数、适合于小波网络的学习算法等.

参考文献 :

[1] SCHENK K R. Random dynamical systems in economics[J]. Stochastics and Dynamics 2001 ,1(1) :63—68.
[2] 范如国 ,黄本笑.复杂经济系统的非均衡分形特征研究[J]. 经济学动态 2002 (6) :18—21.
[3] MARTIN C. Nonlinear prediction of chaotic time series[J]. Phys D ,1989 ,35(7) :335—356.
[4] 张永安.混沌经济系统非线性预测模型的分析[J]. 西北轻工业学院学报 ,2000 ,18(4) :74—79.
[5] BAUMOL W J. BENHABIB J. Chaos : significance ,mechanism ,and economic application[J]. J of Economic Perspartices ,1989 (1) : 77—105.

- [6] PINGEL D ,SCHMELCHER P. Theory and application of the systematic detection of unstable periodic orbits in dynamical systems[J]. Phys Rev E ,2000 ,62(8) :2119—2134.
- [7] KUGIUNTZIS D ,LINGJNRDE O C ,CHRISTOPBERSEN N. Regularized local linear prediction of chaotic time series[J]. Phys D ,1998 ,112(2) :344—360.
- [8] 陈国华 ,马军海 ,陈春旺 ,等. 经济时序动力系统最佳嵌入维数的选取及应用[J]. 天津大学学报 ,2001 ,34(2) :158—162.
- [9] 顾成奎 ,王正欧. 基于神经网络动态非线性非平稳经济系统预测[J]. 系统工程学报 ,2002 ,17(2) :12—15.
- [10] 杨一文 ,刘贵忠 ,张宗华 ,等. 基于小波网络的非线性时间预测及其在股市中的应用[J]. 模式识别与人工智能 ,2001 ,14(2) :243—247.
- [11] CAO Liang-yue ,HONG Yi-guang ,FANG Hai-ping ,et al. Predicting chaotic time series with wavelet-networks[J]. Physica D ,1995 ,85(1) :225—238.
- [12] HANSEN J V ,NELSON R D. Neural network and traditional time series methods : a synergistic combination in state economic forecasts [J]. IEEE Transaction on Neural Networks ,1997 ,8(4) :863—873.

Wavelet network-based non-linear economic time series prediction model

FANG Xian-ming¹ , TANG De-shan²

(1. *Business School , Hohai Univ. , Nanjing 210098 , China ;*

2. College of Water Conservancy and Hydropower Engineering , Hohai Univ. , Nanjing 210098 , China)

Abstract : For accurate prediction of economic time series , the data structure analysis should be made , and the phase space technique provides the theoretical basis for the analysis. Through correlation dimension calculation , the determinate system and random system are classified , and the optimum embedded dimension , the optimum sampling time interval , and the number of wavelet units are further determined. Then by training of recursive wavelet networks with deviation units , the parameters of the model are identified. Experimental study shows that the model is of good behavior in approximation of nonlinear economic time series. The good agreement between experimental results and theoretical analysis validates the feasibility of the model applied to prediction of non-linear economic time series.

Key words : economic time series ; correlation dimension ; wavelet network ; prediction