

基于多维泰勒网的非线性时间序列预测方法及其应用

林屹^{1,2}, 严洪森¹, 周博¹

(1. 东南大学 a. 自动化学院, b. 复杂工程系统测量与控制教育部重点实验室, 南京 210096; 2. 南京信息工程大学 信息与控制学院, 南京 210044)

摘要: 针对非线性时间序列, 提出一种基于多维泰勒网的时间序列预测方法. 其特点在于利用非线性时间序列的观测数据, 通过多维泰勒网得到 n 元一阶多项式差分方程组, 在无需待预测系统的任何先验知识和机理的情况下获得动力学特性描述, 实现对非线性时间序列的预测. 最后分别采用 Lorenz 混沌时间序列, 以及某大型建筑在顶升施工安全性监测中的结构振动响应数据进行实证研究, 所得结果表明了该方法的有效性.

关键词: 时间序列; 多维泰勒网; 施工安全性监测; 预测

中图分类号: TP391

文献标志码: A

Nonlinear time series prediction method based on multi-dimensional Taylor network and its applications

LIN Yi^{1,2}, YAN Hong-sen¹, ZHOU Bo¹

(1a. School of Automation, 1b. Key Laboratory of Measurement and Control of CSE of Ministry of Education, Southeast University, Nanjing 210096, China; 2. College of Information and Control, Nanjing University of Information Science & Technology, Nanjing 210044, China. Correspondent: LIN Yi, E-mail: nuistly@163.com)

Abstract: For the nonlinear time series, a time series prediction method based on multi-dimensional Taylor network is proposed. The characteristics of this method is that based on the observed data of the nonlinear time series, n -variables first order polynomial difference equations of the system are obtained by the multi-dimensional Taylor network, and the dynamic characteristics can be described without prior knowledge and mechanism of the system, thus realizing the prediction of the nonlinear time series. The Lorenz chaotic time series and the structural vibration response data in lift-up construction safety monitoring are used to show the effectiveness of this method.

Key words: time series; multi-dimensional Taylor network; construction safety monitoring; prediction

0 引言

时间序列是根据时间顺序得到的等时间间隔的一系列观测数据. 对时间序列的研究是通过分析在过去数据序列中所包含的信息, 找到其中所蕴含的变化规律, 建立数学模型来进行时间序列的分析、预测和控制. 传统的时间序列预测方法主要有: 针对线性时间序列的标准的自回归滑动平均 (ARMA) 模型^[1]和针对非线性时间序列的非线性自回归滑动平均 (NARMA) 模型^[2]等. 这些传统的时间序列预测方法是利用观测的历史数据建立预测的统计模型, 虽然这些模型无需知道待预测系统的动力学过程和机理, 但这种统计模型一般只适用于因变量与自变量之间是线性关系或一些简单的非线性函数关系. 此外, 当样

本数据太少时, 这种模型不具有统计意义.

目前, 时间序列预测采用的智能预测方法主要有支持向量机方法和神经网络方法. 支持向量机方法可以通过核函数实现从样本空间到高维特征空间的非线性映射, 较好地解决了小样本、非线性、高维数、局部极小点等问题. 文献 [3] 采用支持向量机方法对时变控制参数条件下 Lorenz 系统产生的非平稳时间序列以及实际气候系统的非平稳时间序列进行预测, 结果表明, 支持向量机方法对于非平稳过程具有预测能力. 文献 [4] 采用支持向量机回归原理对 Mackey-Glass 混沌时间序列进行预测, 并对支持向量机参数进行自适应优化, 也取得了较好的预测效果. 然而, 文献 [3] 和文献 [4] 中所讨论的方法都是基于系统的精

收稿日期: 2013-02-25; 修回日期: 2013-09-24.

基金项目: 国家自然科学基金重点项目(60934008).

作者简介: 林屹(1977-), 女, 讲师, 博士生, 从事复杂系统建模、智能预测的研究; 严洪森(1957-), 男, 教授, 博士生导师, 从事生产计划与调度、预测、知识化制造等研究.

确数学模型已知这一前提展开的. 另外, 支持向量机方法尚存在核函数确定困难、可调参数的选取对性能影响大、优质有限小样本的选取困难等问题. 为解决支持向量机方法存在的这些问题, 文献[5]在建立基于最小二乘支持向量机的焦炉煤气柜位预测模型的基础上, 尝试通过构造梯度网格搜索算法来优选模型参数, 并采用大样本筛选方法选取训练样本等手段来提高预测精度.

非线性函数还可以通过神经网络进行逼近^[6], 基于神经网络的时间序列预测法也取得了一定的研究成果. 文献[7]通过混合学习算法获取知识, 确定模糊初始规则基, 再利用神经网络学习能力修改规则库中的模糊规则以及隶属函数和网络权值等参数以加快推理速度. 文献[8]采用输入及对应的连接权为时变函数的过程神经网络, 增加时间聚合运算算子, 构建预测模型及学习算法. 文献[9]利用遗传算法对神经网络的初始权值和阈值分布进行优化, 经过选择、交叉和变异得到BP神经网络的最优初始权值和阈值. 从上述文献的具体研究内容可以看出, 正是由于神经网络方法本身仍存在初始权值、阈值确定困难, 以及泛化能力差, 容易陷入局部极值等问题, 研究人员正力图采用一些新的、有效的方法来解决这些问题. 此外, 虽然基于神经网络的时间序列预测并不需要明确的机理, 但其实际上获得的是隐性的方程描述, 在这种隐性表述中, 历史观测数据对待预测数据的影响是无法具体明确表达的.

本文提出一种新型的非线性时间序列预测方法, 即基于多维泰勒网的非线性时间序列预测方法. 该方法有别于传统统计方法和上述智能预测方法之处在于, 它在不需系统任何先验知识的条件下, 仅利用非线性时间序列的观测数据和多维泰勒网对时间序列进行学习预测, 泛化能力强. 同时, 多维泰勒网可以表示一般意义下的动力学特性, 可以任意逼近一般的非线性系统, 从而获得该非线性系统的状态方程描述. 本文通过对Lorenz混沌时间序列, 以及大型建筑顶升施工过程中结构振动响应数据进行实证研究, 算例实验表明, 这种基于多维泰勒网的非线性时间序列预测方法思想简单, 实施方便, 逼近精度高, 是一种有效的非线性时间序列预测方法.

1 多维泰勒网(MTN)

若观测到的时间序列为 $\{x(t), t = 1, 2, \dots, L\}$, 则在 n 维状态空间中重构的一点状态矢量可表示为 $\mathbf{x}(t) = [x(t), x(t-\tau), \dots, x(t-(n-1)\tau)]^T$. 其中: n 为嵌入维数, τ 为延迟时间. 由Takens定理, 有

$$\mathbf{x}(t+1) = \mathbf{f}(\mathbf{x}(t)). \quad (1)$$

令 $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$, 其中

$$x_1(t) = x(t),$$

$$x_2(t) = x_1(t) - x_1(t-1) = x(t) - x(t-1),$$

$$x_3(t) = x_2(t) - x_2(t-1) =$$

$$x(t) - x(t-1) - (x(t-1) - x(t-2)),$$

⋮

$$x_n(t) = x_{n-1}(t) - x_{n-1}(t-1),$$

则有

$$\mathbf{x}(t+1) = \mathbf{f}(\mathbf{x}(t)) = \mathbf{f}(x_1(t), x_2(t), \dots, x_n(t)). \quad (2)$$

为了便于表述, 下文中 $x_i(t)$ ($i = 0, 1, \dots, n$) 均用 x_i ($i = 0, 1, \dots, n$) 表示.

引理 1 任何定义于一个闭区间的连续函数可以用多项式函数任意准确地逼近^[10].

证明 设 $[a, b]$ 是有界闭区间, f 在 $[a, b]$ 上连续, 若 $[a, b] = [0, 1]$, 则由Bernstein定理可知, 对于任意 $\varepsilon > 0$, 存在多项式 P 使对于所有 $x \in [0, 1]$, $|f(x) - P(x)| < \varepsilon$, f 由多项式 P 一致逼近. 若 $[a, b] \neq [0, 1]$, 则考虑 y 的函数 $f(a+y(b-a))$ 在 $[0, 1]$ 上有定义且连续, 故存在多项式 $Q(y)$ 使对于所有 $y \in [0, 1]$ 有 $|f(a+y(b-a)) - Q(y)| < \varepsilon$. 当 $x \in [a, b]$ 时, $\frac{x-a}{b-a} \in [0, 1]$, 于是对于任意 $\varepsilon > 0$, 有

$$\left| f(x) - Q\left(\frac{x-a}{b-a}\right) \right| < \varepsilon, \quad P(x) = Q\left(\frac{x-a}{b-a}\right),$$

故存在多项式 P 使对于所有 $x \in [a, b]$, $|f(x) - P(x)| < \varepsilon$, 即 f 可由多项式 P 一致逼近. \square

定理 1 对于定义于一个闭区间的连续函数 $f(x_1, x_2, \dots, x_n)$, 可以用 $\sum_{q=1}^w \lambda_q \prod_{i=1}^n x_i^{\sigma_{q,i}}$ 逼近. 其中: w 为逼近展开式中乘积项的总项数, λ_q 为逼近展开式中第 q 个乘积项之前的权值, $\sigma_{q,i}$ 为展开式中第 q 个乘积项中变量 x_i 的幂次.

证明 由引理1可知, 任何定义于一个闭区间的连续函数均可用多项式函数任意准确地逼近. 因此, $f(x_1, x_2, \dots, x_n)$ 可以用下式表示:

$$\begin{aligned} f(x_1, x_2, \dots, x_n) \approx & \alpha + (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n) + \\ & (\gamma_{11} x_1^2 + \gamma_{12} x_1 x_2 + \dots + \\ & \gamma_{22} x_2^2 + \gamma_{23} x_2 x_3 + \dots + \gamma_{nn} x_n^2) + \dots \end{aligned} \quad (3)$$

在该逼近展开式中, 各幂次的乘积项单元分别为: 常数项单元1; 一次项单元 x_1, x_2, \dots, x_n ; 二次项单元 $x_1^2, x_1 x_2, \dots, x_n^2, \dots$; n 次项单元 $x_1^n, x_1 x_2^{n-1}, \dots, x_n^n$. 展开式中所有乘积项的组成和排列如图1所示.

图1中, 由各变量组成的每一个乘积项均按照各

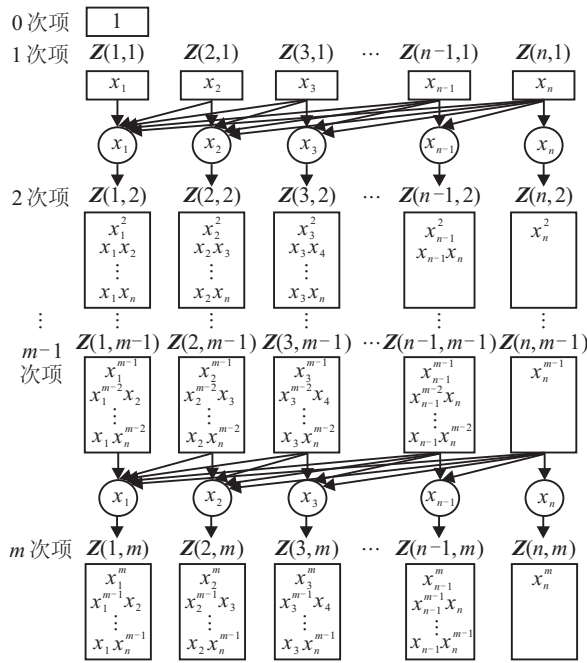


图 1 逼近展开式中乘积项组成和排列示意

变量 x_i 的下标 i 的升序从左至右排列. 例如: 由变量 x_1 和 x_2 构成的幂次之和为 2 的乘积项写为 x_1x_2 ; 由变量 x_1 的一次方、 x_2 的一次方和 x_3 的二次方构成的幂次之和为 4 的乘积项写为 $x_1x_2x_3^2, \dots$, 以此类推. 所有这些乘积项根据各变量的幂次之和的不同, 按幂次之和的升序划分为 0 次项 (即常数项) 集合, 1 次项 (如 x_1, x_2, \dots, x_n) 集合, 2 次项 (如 $x_1^2, x_1x_2, \dots, x_n^2$) 集合, \dots , n 次项 (如 $x_1^n, x_1x_2^{n-1}, \dots, x_n^n$) 集合. 图 1 中, 常数项集合用 $Z(n, 0)$ 表示, 其余各集合分别用 $Z(i, j) (i = 1, 2, \dots, n, j = 1, 2, \dots, m)$ 表示. 其中: i 表示变量, j 表示次项.

如图 1 所示, 除常数项集合 $Z(n, 0)$ 以外, 任意一个子集 $Z(i, j)$ 均可由相应子集 $Z(i, j - 1), Z(i + 1, j - 1), \dots, Z(n, j - 1)$ 的并集中的所有乘积项依次与 x_i 相乘后产生的乘积项构成. 例如, 子集 $Z(1, 2)$ 中的所有乘积项 $x_1^2, x_1x_2, \dots, x_1x_n$ 为子集 $Z(1, 1)$ 的所有乘积项 x_1 、子集 $Z(2, 1)$ 的所有乘积项 x_2 、 \dots 、子集 $Z(n, 1)$ 的所有乘积项 x_n 分别与 x_1 相乘产生的. 再如, 子集 $Z(2, 3)$ 中的所有乘积项 $x_2^3, x_3x_4, \dots, x_3x_n$ 为子集 $Z(3, 1)$ 的所有乘积项 x_3 、子集 $Z(4, 1)$ 的所有乘积项 x_4 、 \dots 、子集 $Z(n, 1)$ 的所有元素 x_n 分别与 x_3 相乘产生的.

令 Z_i^j 为展开式中任意一个子集 $Z(i, j)$ 中所包含的所有乘积项的个数, W_i^j 为从 0 次项集合 (常数项集合) 开始到子集 $Z(i, j)$ 截止时展开式中所有乘积项的个数.

如上所述, 由于图 1 中除常数项集合 $Z(n, 0)$ 以外, 任意一个子集 $Z(i, j)$ 均可由相应子集 $Z(i, j - 1)$,

$Z(i + 1, j - 1), \dots, Z(n, j - 1)$ 的并集中所有乘积项依次与 x_i 相乘后产生的乘积项构成, 可以得到

$$Z_i^j = \sum_{k=1}^n Z_k^{j-1}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m; \quad (4)$$

$$\begin{aligned} W_i^j &= W_n^{j-1} + Z_1^j + Z_2^j + \dots + Z_i^j = \\ &= \left(1 + \sum_{k=1}^n Z_k^1 + \sum_{k=1}^n Z_k^2 + \dots + \sum_{k=1}^n Z_k^{j-1}\right) + \\ &= Z_1^j + Z_2^j + \dots + Z_i^j = \\ &= 1 + \sum_{l=1}^{j-1} \sum_{k=1}^n Z_k^l + \sum_{p=1}^i Z_p^j, \end{aligned} \quad (5)$$

$i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m.$

当式 (5) 中 $i = n$ 且 $j = m$ 时, $W_i^j = W_n^m$, 即为全部乘积项的项数. 若以 λ_q 表示各幂次的乘积项单元相应的权值, 以 $\sigma_{q,i}$ 表示展开式中第 q 个乘积项中变量 x_i 的幂次, 其中 $q (q = 1, 2, \dots, W_n^m)$ 为按图 1 中规律排列的幂次乘积项单元的项数, 则由式 (3) 可得

$$\begin{aligned} f(x_1, x_2, \dots, x_n) &\approx \\ &= \alpha + (\beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n) + \\ &= (\gamma_{11}x_1^2 + \gamma_{12}x_1x_2 + \dots + \gamma_{22}x_2^2 + \\ &= \gamma_{23}x_2x_3 + \dots + \gamma_{nn}x_n^2) + \dots = \\ &= \lambda_1 \cdot 1 + \lambda_2x_1 + \lambda_3x_2 + \dots + \\ &= \lambda_{n+1}x_n + \lambda_{n+2}x_1^2 + \dots = \sum_{q=1}^{W_n^m} \lambda_q \prod_{i=1}^n x_i^{\sigma_{q,i}}. \end{aligned} \quad (6)$$

用 w 表示逼近展开式中乘积项的总项数 W_n^m , 可得

$$f(x_1, x_2, \dots, x_n) \approx \sum_{q=1}^w \lambda_q \prod_{i=1}^n x_i^{\sigma_{q,i}}. \quad (7)$$

展开式的乘积项中各变量的幂次 $\sigma_{q,i}$ 可以通过递归法得到. 用 V_r^h 表示各变量的幂次之和为 h 的第 r 组乘积项集合的起始序号, 则有

$$\begin{aligned} V_r^h &= W_n^h - Z_r^{h+1}, \\ r &= 1, 2, \dots, n, \quad h = 1, 2, \dots, m. \end{aligned} \quad (8)$$

由图 1 所示的展开式中乘积项组成和排列示意关系, 可以得到

$$\sigma_{V_r^{h+s}, i} = \begin{cases} \sigma_{V_r^{h-1+s}, i} + 1, & i = r; \\ \sigma_{V_r^{h-1+s}, i}, & i \neq r. \end{cases} \quad (9)$$

其中: $\{r = 1, 2, \dots, n\}, \{h = 1, 2, \dots, m\}, \{i = 1, 2, \dots, n\}, \{s = 1, 2, \dots, Z_r^h\}$. 初始值为

$$\sigma_{V_n^0+s, i} = \begin{cases} 1, & i = s; \\ 0, & i \neq s. \end{cases} \quad (10)$$

其中: $i = 1, 2, \dots, n, \quad s = 1, 2, \dots, n, \quad V_n^0 = 1.$

至此定理 1 得证. \square

对于观测到的时间序列 $\{x(t)\}$, 将式 (2) 写成标

量形式为

$$\begin{cases} x_1(t+1) = f_1(x_1(t), x_2(t), \dots, x_n(t)), \\ x_2(t+1) = f_2(x_1(t), x_2(t), \dots, x_n(t)), \\ \vdots \\ x_n(t+1) = f_n(x_1(t), x_2(t), \dots, x_n(t)). \end{cases} \quad (11)$$

由定理 1 可知, $x_1(t+1), x_2(t+1), \dots, x_n(t+1)$ 可分别表示为

$$\begin{cases} x_1(t+1) = \sum_{q=1}^w \lambda_{1,q} \prod_{i=1}^n x_i^{\sigma_{q,i}}, \\ x_2(t+1) = \sum_{q=1}^w \lambda_{2,q} \prod_{i=1}^n x_i^{\sigma_{q,i}}, \\ \vdots \\ x_n(t+1) = \sum_{q=1}^w \lambda_{n,q} \prod_{i=1}^n x_i^{\sigma_{q,i}}. \end{cases} \quad (12)$$

其中: w 为展开式中乘积项的总项数; $\lambda_{1,q}, \lambda_{2,q}, \dots, \lambda_{n,q}$ 为各展开式中第 q 个乘积项相应的权值; $\sigma_{q,i}$ 为展开式中第 q 个乘积项中变量 x_i 的幂次. 由此, 可以构造多维泰勒网模型, 如图 2 所示. 多维泰勒网为一个 3 层前向网络, 输入为 $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$, 用于完成 n 个变量向网络的输入; 中间层为网络处理层, 由各幂次乘积项单元和相应连接权值组成, 用于完成各输入变量在各幂次乘积项单元的加权聚合, 其中 $\lambda_l = \{\lambda_{l,1}, \lambda_{l,2}, \dots, \lambda_{l,w}\} (l = 1, 2, \dots, n)$ 为连接线上的权值集合; 输出层有 n 个节点, $\mathbf{x}(t+1) = [x_1(t+1), x_2(t+1), \dots, x_n(t+1)]^T$, 用于完成相应输入变量的网络输出.

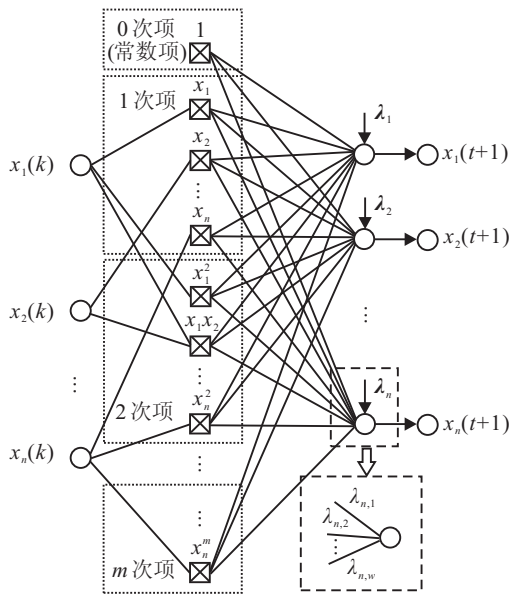


图 2 多维泰勒网模型

2 基于多维泰勒网的时间序列预测模型

时间序列预测就是根据时间序列的历史观测数据对其进行估计. 利用人工神经网络以及支持向量机方法进行预测时, 本质上获得的是一元 n 阶隐非线性

线性差分方程, 可描述为 $x(t+1) = f[x(t), x(t-1), \dots, x(t-n)]$. 在这种隐性表述中, 无法具体明确表达 $x(t), x(t-1), \dots, x(t-n)$ 对 $x(t+1)$ 的影响. 而本文提出的基于多维泰勒网的预测方法, 最终可以获得 n 元一阶显性多项式差分方程组, 多维泰勒网的模型参数可以通过样本学习得到.

2.1 模型结构

根据时间序列 $\{x(t)\}$ 的历史观测数据 $\{x(t), x(t-1), \dots, x(t-n)\}$ 对 $x(t+1)$ 进行预测时, 本文所提出的基于多维泰勒网的时间序列预测模型采用的是一种四元多维泰勒网络, 其输入为 $\mathbf{x}(t) = [x_1(t), x_2(t), x_3(t), x_4(t)]^T$, 其中

$$\begin{cases} x_1(t) = x(t), \\ x_2(t) = x_1(t) - x_1(t-1), \\ x_3(t) = x_2(t) - x_2(t-1), \\ x_4(t) = x_3(t) - x_3(t-1). \end{cases}$$

中间处理层由各乘积项单元和相应连接权组成, 多维泰勒网输出为 $\mathbf{x}(t+1) = [x_1(t+1), x_2(t+1), x_3(t+1), x_4(t+1)]^T$. 其中

$$\begin{cases} x_1(t+1) = \sum_{q=1}^{W_4^4} \lambda_{1,q} \prod_{i=1}^4 x_i^{\sigma_{q,i}}, \\ x_2(t+1) = \sum_{q=1}^{W_4^4} \lambda_{2,q} \prod_{i=1}^4 x_i^{\sigma_{q,i}}, \\ x_3(t+1) = \sum_{q=1}^{W_4^4} \lambda_{3,q} \prod_{i=1}^4 x_i^{\sigma_{q,i}}, \\ x_4(t+1) = \sum_{q=1}^{W_4^4} \lambda_{4,q} \prod_{i=1}^4 x_i^{\sigma_{q,i}}. \end{cases} \quad (13)$$

预测值为

$$\hat{y}(t+1) = x_1(t+1) = \sum_{q=1}^{W_4^4} \lambda_{1,q} \prod_{i=1}^4 x_i^{\sigma_{q,i}}. \quad (14)$$

其中: W_4^4 为展开式中乘积项的总项数, $\lambda_{1,q}$ 为第 q 个乘积项之前的权值, $\sigma_{q,i}$ 为第 q 个乘积项中变量 x_i 的幂次.

2.2 MTN 学习算法

对于多维泰勒网的学习是指对多维泰勒网参数, 即展开式中各乘积项的权值的学习.

根据时间序列 $\{x(t)\}$ 的历史观测数据 $\{x(t), x(t-1), \dots, x(t-n)\}$ 预测 $x(t+1)$, 进行网络学习时, 设样本集为 $\{(\mathbf{x}(t), y(t+1))\}_{t=1}^d$. 其中: $\mathbf{x}(t) = [x_1(t), x_2(t), x_3(t), x_4(t)]^T, y(t+1) = x_1(t+1)$. 相应地, 模型的预测值表示为 $\hat{y}(t+1)$. 令

$$\mathbf{b}_t = \left(\prod_{i=1}^4 x_i^{\sigma_{1,i}}(t), \prod_{i=1}^4 x_i^{\sigma_{2,i}}(t), \dots, \prod_{i=1}^4 x_i^{\sigma_{W_4^4,i}}(t) \right)^T,$$

$$\begin{aligned} \mathbf{B} &= (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_d), \\ \boldsymbol{\lambda} &= (\lambda_{1,1}, \lambda_{1,2}, \dots, \lambda_{1,W_4})^T, \\ \mathbf{y} &= (y(2), y(3), \dots, y(d+1))^T, \end{aligned}$$

则式(14)可写为

$$\hat{y}(t+1) = \boldsymbol{\lambda}^T \mathbf{b}_t. \quad (15)$$

网络参数寻优需要最小化误差平方和

$$E = \frac{1}{2} \sum_{t=1}^d (y(t+1) - \hat{y}(t+1))^2. \quad (16)$$

将式(15)代入(16), 得到目标函数

$$E = \frac{1}{2} \sum_{t=1}^d (y(t+1) - \boldsymbol{\lambda}^T \mathbf{b}_t)^2. \quad (17)$$

这是二次函数的无约束最优化问题, 可采用共轭梯度法求解. 对目标函数求偏导, 有

$$\begin{aligned} \nabla E(\boldsymbol{\lambda}) &= \sum_{t=1}^d (\mathbf{b}_t \mathbf{b}_t^T) \cdot \boldsymbol{\lambda} - \sum_{t=1}^d y(t+1) \cdot \mathbf{b}_t = \\ & \mathbf{B} \mathbf{B}^T \boldsymbol{\lambda} - \mathbf{B} \mathbf{y}. \end{aligned} \quad (18)$$

令 \mathbf{u} 为共轭向量, 假设初始的共轭向量 \mathbf{u}_0 取为随机初始点 $\boldsymbol{\lambda}_0$ 处的负梯度 $-\nabla E(\boldsymbol{\lambda}_0)$, 而以下各共轭方向 \mathbf{u}_p 由第 p 迭代点 $\boldsymbol{\lambda}_p$ 处的负梯度 $-\nabla E(\boldsymbol{\lambda}_p)$ 与已得到的共轭向量 \mathbf{u}_{p-1} 的线性组合来确定, 从 $\boldsymbol{\lambda}_p$ 出发按相应共轭方向进行搜索, 得到第 $p+1$ 个迭代点 $\boldsymbol{\lambda}_{p+1}$, 其迭代公式为

$$\boldsymbol{\lambda}_{p+1} = \boldsymbol{\lambda}_p + \delta_p \mathbf{u}_p. \quad (19)$$

其中: δ_p 为步长, \mathbf{u}_p 为共轭方向.

步长 δ_p 由精确线性搜索得到, 即

$$\delta_p = \frac{\nabla E(\boldsymbol{\lambda}_p)^T \nabla E(\boldsymbol{\lambda}_p)}{\mathbf{u}_p^T (\mathbf{B} \mathbf{B}^T) \mathbf{u}_p}. \quad (20)$$

共轭方向 \mathbf{u}_p 可由下式得到:

$$\mathbf{u}_p = \begin{cases} -\nabla E(\boldsymbol{\lambda}_p), & p=0; \\ -\nabla E(\boldsymbol{\lambda}_p) + \rho_{p-1} \mathbf{u}_{p-1}, & p \geq 1. \end{cases} \quad (21)$$

其中

$$\rho_{p-1} = \frac{\nabla E(\boldsymbol{\lambda}_p)^T (\mathbf{B} \mathbf{B}^T) \mathbf{u}_{p-1}}{\mathbf{u}_{p-1}^T (\mathbf{B} \mathbf{B}^T) \mathbf{u}_{p-1}}.$$

具体算法如下.

Step 1: 获取训练数据, 随机选取初始 MTN 网络参数 $\boldsymbol{\lambda}_0$, 设定允许误差 ε , 最大迭代次数 p_{\max} .

Step 2: 构造初始搜索方向, 计算 $\mathbf{u}_0 = -\nabla E(\boldsymbol{\lambda}_0)$, 置迭代计数 $p=0$.

Step 3: 对目标函数 $E(\boldsymbol{\lambda})$ 进行直线搜索, 按式(20)和(19)求出 δ_p 和 $\boldsymbol{\lambda}_{p+1}$, 使得

$$E(\boldsymbol{\lambda}_p + \delta_p \mathbf{u}_p) = \min_{\delta \geq 0} E(\boldsymbol{\lambda}_p + \delta \mathbf{u}_p).$$

Step 4: 检查是否满足终止准则, 计算 $\nabla E(\boldsymbol{\lambda}_{p+1})$. 若 $\|\nabla E(\boldsymbol{\lambda}_{p+1})\| < \varepsilon$, 则迭代终止, $\boldsymbol{\lambda}_{p+1}$ 为最优参数, 存为 $\boldsymbol{\lambda}_{\text{opt}}$; 否则, 转 Step 5.

Step 5: 计算

$$\begin{aligned} \rho_p &= \frac{\nabla E(\boldsymbol{\lambda}_{p+1})^T (\mathbf{B} \mathbf{B}^T) \mathbf{u}_p}{\mathbf{u}_p^T (\mathbf{B} \mathbf{B}^T) \mathbf{u}_p}, \\ \mathbf{u}_{p+1} &= -\nabla E(\boldsymbol{\lambda}_{p+1}) + \rho_p \mathbf{u}_p. \end{aligned}$$

Step 6: 令 $p = p + 1$.

Step 7: 检查迭代次数, 若 p 大于最大迭代次数 p_{\max} , 则停机, 输出寻优得到的多维泰勒网最优参数 $\boldsymbol{\lambda}_{\text{opt}}$; 否则, 转 Step 3.

2.3 预测过程

基于多维泰勒网的时间序列预测过程如下:

1) 对时间序列的原始数据进行数据重构. 根据已知的历史观测数据 $\{x(t), x(t-1), \dots, x(t-n)\}$ 生成多维泰勒网的输入输出向量, 即多维泰勒网预测所需的输入向量 $\mathbf{x}(t)$ ($n \times 1$ 维) 和输出向量 $\mathbf{x}(t+1)$ ($n \times 1$ 维), 预测值为 $\hat{y}(t+1) = x_1(t+1)$, 从而得到样本集 $\{(\mathbf{x}(t), y(t+1))\}_{t=1}^d$.

2) 将重构的时间序列分成两部分: 训练集和测试集. 训练集用于模型参数的寻优, 测试集用于检测所建模型的性能.

3) 进行训练集的多维泰勒网参数学习, 即展开式中各乘积项的权值的学习. 采用共轭梯度法进行参数寻优, 得到最优的 MTN 网络参数, 从而建立 MTN 预测模型. 与此同时, 通过得到的最优多维泰勒网参数, 可以表示一般意义下该时间序列的动力学特性.

4) 对时间序列预测集进行预测, 输出预测结果.

3 应用实例

例1 首先采用 Lorenz 系统这一典型混沌系统的时间序列, 对上述基于多维泰勒网的非线性时间序列预测方法进行预测仿真实验. Lorenz 系统方程为

$$\begin{cases} \dot{x} = -\xi x + yz, \\ \dot{y} = -\theta(y - z), \\ \dot{z} = -xy + \eta y - z. \end{cases}$$

取 $\xi = 8/3, \theta = 10, \eta = 28$, 初始值为 $[0 \ 0 \ 0]$. 选择 Lorenz 系统 x 分量的 1310 个数据, 其中前 1010 个数据作为训练样本, 后 300 个数据作为测试样本. 首先对数据归一化处理; 然后建立 MTN 模型, 进行输入数据的重构. 令模型的输入为 $x_1(t), x_2(t), x_3(t), x_4(t)$. 其中: $x_1(t)$ 为 Lorenz 系统的 x 分量数据, $x_2(t) = x_1(t) - x_1(t-1), x_3(t) = x_2(t) - x_2(t-1), x_4(t) = x_3(t) - x_3(t-1)$. 输出为 $t+1$ 时刻的 x 分量数据 $x_1(t+1)$, 即预测值 $\hat{y}(t+1)$. 预测结果与实际值之间的比较如图 3 所示.

模型的预测精度采用均方根误差 (RMSE) 和相对误差 (perr) 加以评价, 各误差指标定义如下:

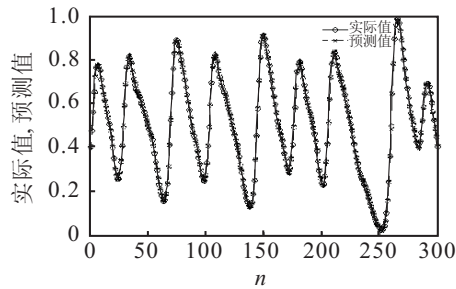


图3 预测结果对比

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum (y(t+1) - \hat{y}(t+1))^2},$$

$$\text{perr} = \frac{\sum (y(t+1) - \hat{y}(t+1))^2}{\sum y(t+1)^2},$$

其中 $y(t+1)$ 和 $\hat{y}(t+1)$ 分别为实际值和预测值。

对 Lorenz 系统 x 分量的 300 个数据预测的具体误差指标如表 1 所示。

表1 预测误差指标对比

方法	RMSE	perr
MTN 方法	5.047e-03	8.138e-05
文献[11]方法	7.176e-03	1.319e-04
文献[12]方法	1.147e-02	4.755e-04

由表 1 可以看出, 本文提出的基于多维泰勒网的非线性时间序列预测方法预测效果优于文献[11]和文献[12]方法, 具有更好的预测精度。

例 2 在大型建筑平移或顶升施工过程中, 需要实时监控建筑结构响应, 以保障在施工过程中建筑结构的安全。传统的建筑平移或顶升施工中采取的监控方法是通过多处的传感器实时监测当前结构的加速度响应, 根据监测到的数据来判断结构的实时安全性。然而, 结构状态的测量、信号的传输都要耗费一定时间, 而且传统监测中一旦监测到异常响应实际上为时已晚, 此时建筑结构已经受到了损伤, 对安全性产生了影响。所以如果能在施工过程中提前准确预测到下一时刻的结构响应, 并做出相应的提前反应, 则将大大减少施工过程中可能出现的结构安全性损伤。为此, 要求寻找一种合适而有效的方法, 根据结构前几个时刻的结构响应预测未来时刻的结构响应。

针对大型建筑顶升施工过程中监测采样的结构响应数据这一非线性时间序列, 本文尝试用文中提出的基于多维泰勒网的非线性时间序列预测方法来实现根据结构前几个时刻的结构响应预测未来时刻的结构响应。选取某 3 层结构的大型建筑顶升施工数据, 应用本文方法建立多维泰勒网进行加速度响应的预测。采用该大型建筑顶升施工过程中位于第 3 层结构的传感器测得的该层加速度响应中前 200 个数据作为训练数据, 以获得 MTN 模型的系数, 后 30 个数据作为测试数据, 以检验预测性能。建立结构响应 MTN 模型, 进行输入数据的重构。令模型的输入为 $x_1(t)$,

$x_2(t), x_3(t), x_4(t)$ 。其中: $x_1(t)$ 为 t 时刻的加速度响应, $x_2(t) = x_1(t) - x_1(t-1)$, $x_3(t) = x_2(t) - x_2(t-1)$, $x_4(t) = x_3(t) - x_3(t-1)$ 。输出为 $t+1$ 时刻的加速度响应 $x_1(t+1)$, 即预测值 $\hat{y}(t+1)$ 。第 3 层结构的传感器加速度响应预测结果如图 4 所示。鉴于神经网络也是一种常用的预测方法, 为了说明本文预测方法的有效性, 图 4 中还给出了在结构响应预测中采用神经网络方法(文献[13]方法)的预测结果, 该方法将第 2 层的前两个时刻加速度响应和第 3 层的前两个时刻加速度响应作为输入。

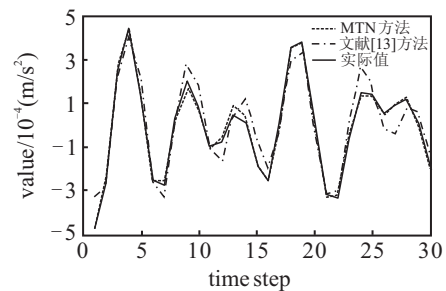


图4 加速度响应预测结果对比

由图 4 可以看出, 与文献[13]方法相比较, 本文所提出的多维泰勒网方法的预测曲线与实际曲线的吻合度更好, 预测准确率更高。

采用本文方法进行预测时的 30 个测试数据的实际值和预测值见表 2。

表2 MTN 方法预测值与实际值比较 $\times 10^{-4}$

序号	实际值	预测值	序号	实际值	预测值
1	-4.839	-4.719	16	-2.550	-2.530
2	-2.630	-2.316	17	0.084	-0.001
3	2.485	2.799	18	3.547	3.604
4	4.451	4.424	19	3.868	3.787
5	1.416	1.447	20	0.367	0.144
6	-2.538	-2.563	21	-3.256	-3.448
7	-2.720	-2.544	22	-3.358	-3.196
8	0.237	0.288	23	-0.590	-0.440
9	2.066	1.717	24	1.499	1.338
10	0.881	0.616	25	1.358	1.241
11	-0.947	-1.121	26	0.610	0.502
12	-0.782	-0.524	27	0.847	0.970
13	0.477	0.934	28	1.184	1.335
14	0.122	0.453	29	-0.012	-0.120
15	-1.890	-1.829	30	-2.046	-2.410

本文在对图 4 中的两种预测方法进行误差分析的同时, 为了更好地说明本文所提出 MTN 方法的有效性, 还进行了输入量重构后神经网络方法的预测误差分析, 即仍采用神经网络结构, 但不采用在结构响应预测中通常采用的第 2 层前两个时刻的加速度响应和第 3 层前两个时刻的加速度响应作为输入量, 而是采用与 MTN 方法相同的重构后的信号 $x_1(t), x_2(t), x_3(t), x_4(t)$ 作为输入量进行预测。上述 3 种方法的具体误差指标如表 3 所示。

表3 加速度响应预测误差指标对比

方法	RMSE	perr
MTN方法	2.039e-05	8.325e-03
文献[13]方法	7.341e-05	1.079e-01
ANN方法(输入重构)	2.451e-05	1.203e-02

由表3可以看出, 本文提出的基于多维泰勒网(MTN方法)的非线性时间序列预测方法预测效果优于文献[13]方法. 另外, 将人工神经网络方法(ANN方法)的输入进行重构, 即采用与多维泰勒网完全相同的经过多阶差分处理后的数据作为输入量, 其预测误差指标虽然已较文献[13]方法有了明显改进, 但仍然明显不如本文方法的预测误差指标, 从而进一步验证了本文方法的有效性.

此外, 虽然神经网络能逼近任意的非线性函数, 并且基于神经网络的时间序列预测虽然也不需要明确的机理, 但该方法无法具体明确表达历史观测数据对预测数据的影响, 无法以显性方程形式加以描述. 而本文所提出的基于多维泰勒网的时间序列预测, 在训练集参数学习结束后, 可以确定最优多维泰勒网络参数 λ_{opt} , 进而可以表示出如式(11)的系统一般意义下的状态动力学特性. 这种显性表述正是本文方法的又一优势.

4 结 论

本文提出了基于多维泰勒网的非线性时间序列预测方法, 该方法具有以下显著优点: 1) 不需要系统的任何先验知识, 仅利用非线性系统的输入输出数据; 2) 利用多维泰勒网可以表示系统一般意义下的状态动力学特性; 3) 算法实现方便、预测精度高. 分别采用Lorenz混沌时间序列, 以及某3层结构的大型建筑顶升施工过程中的数据进行应用研究, 并将预测结果与其他方法进行对比, 结果表明了本文所提出方法的有效性.

参考文献(References)

- [1] Box G E, Jenkins G M, Reinsel G C. Time series analysis forecasting and control[M]. Beijing: Posts & Telecom Press, 2005: 19-180.
- [2] Chen S, Billings S A. Representation of non-linear systems: The NARMAX model[J]. Int J of Control, 1989, 49(3): 1013-1042.
- [3] 王革丽, 杨培才, 毛宇清. 基于支持向量机方法对非平稳时间序列的预测[J]. 物理学报, 2008, 57(2): 714-719.
(Wang G L, Yang P C, Mao Y Q. On the application of non-stationary time series prediction based on the SVM method[J]. Acta Physica Sinica, 2008, 57(2): 714-719.)
- [4] 刘涵, 刘丁, 李琦. 基于支持向量机的混沌时间序列非线性预测[J]. 系统工程理论与实践, 2005, 25(9): 94-99.

- (Liu H, Liu D, Li Q. Chaotic time series nonlinear prediction based on support vector machines[J]. Systems Engineering-Theory & Practice, 2005, 25(9): 94-99.)
- [5] 张晓平, 赵珺, 王伟, 等. 基于最小二乘支持向量机的焦炉煤气柜位预测模型及应用[J]. 控制与决策, 2010, 25(8): 1178-1183.
(Zhang X P, Zhao J, Wang W, et al. COG holder level prediction model based on least square support vector machine and its application[J]. Control and Decision, 2010, 25(8): 1178-1183.)
- [6] Hornik K, Stinchcombe M, White H. Multi-layer feedforward networks are universal approximators[J]. Neural Networks, 1989, 2(3): 359-366.
- [7] 谭文, 王耀南, 周少武, 等. 混沌时间序列的模糊神经网络预测[J]. 物理学报, 2003, 52(4): 795-801.
(Tan W, Wang Y N, Zhou S W, et al. Prediction of the chaotic time series using neuro-fuzzy networks[J]. Acta Physica Sinica, 2003, 52(4): 795-801.)
- [8] 丁刚, 钟诗胜. 基于过程神经网络的时间序列预测及其应用研究[J]. 控制与决策, 2006, 21(9): 1037-1041.
(Ding G, Zhong S S. Time series prediction based on process neural networks and its applications[J]. Control and Decision, 2006, 21(9): 1037-1041.)
- [9] 李松, 刘力军, 解永乐. 遗传算法优化BP神经网络的短时交通流混沌预测[J]. 控制与决策, 2011, 26(10): 1581-1585.
(Li S, Liu L J, Xie Y L. Chaotic prediction for short-term traffic flow of optimized BP neural network based on genetic algorithm[J]. Control and Decision, 2011, 26(10): 1581-1585.)
- [10] Klambauer G. Mathematical analysis[M]. New York: Marcel Dekker INC, 1975: 236-237.
- [11] 孟庆芳, 张强, 牟文英. 混沌时间序列多步自适应预测方法[J]. 物理学报, 2006, 55(4): 1666-1671.
(Meng Q F, Zhang Q, Mu W Y. A novel multi-step adaptive prediction method for chaotic time series[J]. Acta Physica Sinica, 2006, 55(4): 1666-1671.)
- [12] 闫华, 魏平, 肖先赐. 基于Bernstein多项式的自适应混沌时间序列预测算法[J]. 物理学报, 2007, 56(9): 5112-5116.
(Yan H, Wei P, Xiao X C. An adaptive approach based on Bernstein polynomial to predict chaotic time series[J]. Acta Physica Sinica, 2007, 56(9): 5112-5116.)
- [13] 徐赵东, 沈亚鹏, 郭迎庆. 神经网络对结构地震响应的预测及试验研究[J]. 振动与冲击, 2003, 22(2): 8-11.
(Xu Z D, Shen Y P, Guo Y Q. Neural network prediction for seismic response of structure and experimental study[J]. J of Vibration and Shock, 2003, 22(2): 8-11.)

(责任编辑: 李君玲)