

基于RBF神经网络的非线性时间序列在线预测

张冬青, 宁宣熙, 刘雪妮

(南京航空航天大学 经济与管理学院, 江苏 南京 210016)

摘要: 针对非线性非高斯时间序列, 提出观测噪声服从隐马尔可夫模型(HMM)的径向基函数(RBF)神经网络(RBF-HMM)预测模型, 其特点在于模型输入包含误差反馈项、RBF网络隐含层节点数的可变性和观测噪声的隐马尔可夫性; 并采用序列蒙特卡罗(SMC)方法实现基于RBF-HMM模型的时间序列在线预测. 最后采用太阳黑子数平滑月均值数据和CRU国际钢材价格指数月数据进行实证研究, 结果表明该模型的有效性.

关键词: 预测; 径向基函数神经网络; 隐马尔可夫模型; 序列蒙特卡罗方法

中图分类号: O211.6

文献标识码: A

On-line prediction of nonlinear time series using RBF neural networks

ZHANG Dong-qing, NING Xuan-xi, LIU Xue-ni

(College of Economics & Management, Nanjing University of Aeronautics and Astronautics, Nanjing Jiangsu 210016, China)

Abstract: For the nonlinear and non-Gaussian time series, a novel predictive model—the RBF-HMM model is proposed based on the radial basis function(RBF) neural network with measurement noise being assumed to be of a hidden Markov model(HMM). The characteristics of this model include: 1) the predictive errors of RBF neural network are associated with the input of RBF-HMM model; 2) the number of hidden neurons varies with time; 3) the measurement noise is assumed to be HMM distributed. Sequential Monte Carlo(SMC) method is then applied to the on-line prediction for time series in RBF-HMM model. Finally, the smoothed data of the monthly mean sunspot numbers and CRU(the Britain Commodity Research University) steel price index are analyzed. The experimental results indicate that the RBF-HMM model is effective.

Key words: prediction; radial basis function neural networks; hidden Markov model; sequential Monte Carlo method

1 引言(Introduction)

目前, 大多数时间序列均是采用标准的自回归滑动平均(ARMA)模型^[1]来进行建模, 即假设被研究的时间序列服从2阶平稳且噪声驱动机制为高斯白噪声的线性模型. 然而, 现实生活中绝大多数时间序列却呈现非线性, 为此Billings 等人提出非线性自回归滑动平均(NARMA)模型^[2], 它具有形式简洁、高度概括等特点. 然而对于NARMA模型而言, 除了少数情况外, 其非线性函数都难以用精确的数学模型描述, 由于神经网络能以任意精度逼近任意的非线性函数^[3], 因此本文采用径向基函数(RBF)神经网络进行非线性时间序列建模. 另一方面, 已有关于神经网络用于时间序列分析的文献, 都是假设观测噪声服从高斯分布^[4~6], 事实上, 这个假设过于简单, 而且在有些场合并不适合, 因此本文尝试用隐马尔可夫模型(HMM)来对噪声进行非高斯建模.

本文提出观测噪声服从隐马尔可夫模型的径向

基函数神经网络(RBF-HMM)预测模型, 并用序列蒙特卡罗(SMC)方法实现其在线预测.

2 问题描述(Problem statement)

2.1 NARMA模型(NARMA model)

Billings等^[2]于1989年提出如下NARMA模型:

$$\hat{y}_t = f(y_{t-1}, \dots, y_{t-p}, e_{t-1}, \dots, e_{t-q}) + e_t \quad (1)$$

式中: $\{y_{t-1}, \dots, y_{t-p}\}$ 为观测序列, $\hat{y}_t$ 为预测值, $\{e_t, e_{t-1}, \dots, e_{t-q}\}$ 为观测噪声序列, 由于很难得到, 本文用相应的预测误差来代替, 即 $e_{t-1} = y_{t-1} - \hat{y}_{t-1}, \dots, e_{t-q} = y_{t-q} - \hat{y}_{t-q}$; $f(y_{t-1}, \dots, y_{t-p}, e_{t-1}, \dots, e_{t-q})$ 为非线性函数.

由式(1)可知, NARMA模型预测的关键是确定非线性函数具体表达式, 然而这并非易事, 因此许多学者借助神经网络来逼近它. 相对于其他神经网络而言, RBF网络具有学习速度快、逼近能力强等特点而备受瞩目, 所以本文选取节点数可变的RBF网络来

描述NARMA模型中的非线性函数。

关于噪声项 e_t 的分布,本文假设它服从HMM过程。这个假设有其深刻的哲学含义:因为在实际经济社会系统或工程系统中,任何时间序列的产生都有其内在驱动机制,只是有些时间序列的内在驱动机制不太明显,只能通过相应的观测序列进行推断,HMM正好能够从更深层次刻画它。本文预测模型结构如图1所示。

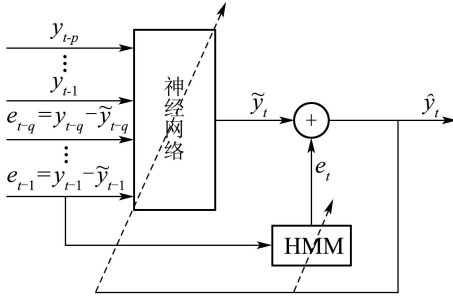


图1 基于神经网络预测模型的结构框图

Fig. 1 Framework of neural networks prediction model

2.2 可变节点数的RBF神经网络(Variable hidden neurons RBF model)

本节讨论图1中神经网络的建模,采用形如Holmes等提出由 k_t 个隐节点的RBF网络和线性回归项组成的混合RBF网络模型^[7]:

$$\begin{cases} \tilde{y}_t = b_t + \beta_t^T \mathbf{x}_t, & k_t = 0, \\ \tilde{y}_t = \sum_{j=1}^{k_t} a_{t,j} \phi(\|\mathbf{x}_t - \mathbf{c}_{t,j}\|) + b_t + \beta_t^T \mathbf{x}_t, & k_t \geq 1. \end{cases} \quad (2)$$

式中: $\mathbf{x}_t = (y_{t-1}, \dots, y_{t-p}, e_{t-1}, \dots, e_{t-q})^T$, $\tilde{y}_t$ 分别为RBF网络的输入和输出; $\|\cdot\|$ 为距离度量; $\phi(\cdot)$ 为径向基函数; $\mathbf{c}_{t,j}$, $a_{t,j}$ 分别为 t 时刻RBF网络第 j 个隐节点的中心和输出权重; b_t 和 $\beta_t = (\beta_{t,1}, \dots, \beta_{t,p+q})$ 为回归系数。简单起见,令 $D_t = (y_{t-1}, \dots, y_{t-p}, e_{t-1}, \dots, e_{t-q}, \phi(\mathbf{x}_t, \mathbf{c}_{t,1}), \dots, \phi(\mathbf{x}_t, \mathbf{c}_{t,k_t}))$, $\alpha_t = (b_t, \beta_{t,1}, \dots, \beta_{t,d}, a_{t,1}, \dots, a_{t,k_t})^T$, 其中 $d = p + q$, 式(2)可表示为如下向量形式:

$$\tilde{y}_t = D_t \alpha_t. \quad (3)$$

至于径向基函数选择,借鉴文献[8],本文选取形如 $\phi(x) = e^{-\lambda x^2}$ 高斯径向基函数, λ 为扩展系数。

2.3 HMM噪声模型(HMM noise model)

图1中观测噪声 e_t 可以用HMM进行建模。HMM最早出现于1957年,由于其具有坚实的数学理论基础,近年来得到广泛应用。HMM是一个双内嵌式随机过程,即HMM是由两个随机过程组成:一个是隐含的状态序列;另一个是与隐状态有关的观测序列。

其中隐状态序列不可观测,只能通过观测序列进行推断^[9]。

$$\text{HMM的状态空间模型为} \begin{cases} z_t \sim p(z_t | z_{t-1}, A_t), \\ e_t \sim f(e_t | z_t, B_t), \end{cases}$$

其中: z_t 是时刻 t 系统所处的隐状态,且 $z_t \in S$, $S = \{S_1, \dots, S_{n_s}\}$ 为隐状态集合, n_s 为隐状态数,不失一般性,本文选取HMM的隐状态数为3(即 $n_s = 3$),表示神经网络对 y_t 的预测性能,即预测值偏高、偏低和基本符合实际3种情况; $z_t \sim p(z_t | z_{t-1}, A_t)$ 为状态转移概率分布,由状态转移矩阵 $A_t = \{a_{ij}\}$ 决定,其中: $a_{ij} = P\{z_t = S_j | z_{t-1} = S_i\}$; $e_t \sim f(e_t | z_t, B_t)$ 是 t 时刻在隐状态 z_t 下测量值的概率分布,由条件概率 $B_t = \{g_i(e_t), 1 \leq i \leq n_s\}$ 刻画,其中 $g_i(e_t) = p\{e_t | z_t = S_i\}$, 可以用混合高斯模型来逼近^[10]。假设每个隐状态下混合高斯模型的个数为 n_m , 即有

$$g_i(e_t) = \sum_{j=1}^{n_m} v_{ij} N(e_t; \mu_{ij}, \sigma_{ij}^2), \quad i = 1, \dots, n_s. \quad (4)$$

式中: $N(e_t; \mu_{ij}, \sigma_{ij}^2)$ 是 t 时刻 S_i 状态下第 j 个高斯分布; $\mu_{ij}, \sigma_{ij}^2, v_{ij}$ 分别为该高斯模型的均值、方差和权重,其中权重应满足 $\sum_{j=1}^{n_m} v_{ij} = 1, v_{ij} \geq 0, i = 1, \dots, n_s$ 。关于HMM详细介绍请参见文献[9]。

综上所述,得到如下两个结论: 1) 当HMM参数 $n_s = 1, n_m = 1$ 时, HMM就退化为高斯模型,所以观测噪声服从高斯分布只是服从HMM分布的特例,而观测噪声服从HMM分布则是高斯分布的推广; 2) 当 $f(y_{t-1}, \dots, y_{t-p}, e_{t-1}, \dots, e_{t-q})$ 为线性函数,且HMM参数 $n_s = 1, n_m = 1$ 时,则本文提出噪声服从HMM分布的NARMA模型就退化为高斯白噪声的ARMA模型。

3 基于SMC方法的RBF-HMM模型在线预测(On-line prediction based on RBF-HMM model with SMC method)

本节旨在研究: 已知至 $t-1$ 时刻所有测量值 $y_{1:t-1} = (y_1, \dots, y_{t-1})$, 如何求出 t 时刻的预测值 $\hat{y}_t$? 采用SMC方法来实现基于RBF-HMM模型的时间序列在线预测。SMC方法源于20世纪60年代,因其能够解决非线性、非高斯等难题而倍受瞩目。SMC方法的基本思想是利用重要性采样方法,用离散随机粒子逼近概率分布,以实现概率分布的递推计算,当随机粒子数量足够大时,它能够无限逼近任意的概率密度函数。有关SMC方法的详细描述和其他相关问题参见文献[11]。

3.1 观测方程与状态方程(Observation equation and state equation)

为了成功应用SMC方法, 必须给出RBF-HMM模型的观测方程和状态转移方程, 由第2节分析可知观测方程可写成

$$\hat{y}_t = \tilde{y}_t + e_t = D_t \alpha_t + e_t. \quad (5)$$

由于RBF-HMM模型参数未知, 将其作为状态向量的一部分进行演化滤波。目前多数文献都采用离散或连续形式的随机游走模型来进行演化, 以此实现模型参数的动态调整和增加粒子的多样性^[5,11,12]。本文亦采用同样方法, 对于连续参数 $\{\alpha_t, c_t, \mu_t, \sigma_t^2, A_t, v_t\}$, 采用添加零均值的高斯随机噪声进行参数演化, 其中 μ_t, σ_t^2, v_t 分别为 t 时刻HMM所有隐状态、所有混合高斯模型的均值、方差和权重组成向量。状态转移方程具体如下:

$$\begin{cases} k_t \sim P(k_t|k_{t-1}), c_t = c_{t-1} + \varepsilon_c, \\ \alpha_t = \alpha_{t-1} + \varepsilon_\alpha, A_t = A_{t-1} + \varepsilon_A, \\ \mu_t = \mu_{t-1} + \varepsilon_\mu, \ln(\sigma_t^2) = \ln(\sigma_{t-1}^2) + \varepsilon_\sigma, \\ v_t = v_{t-1} + \varepsilon_v, z_t \sim P(z_t|z_{t-1}, A_t), \\ I_t \sim P(I_t|z_t, v_t). \end{cases} \quad (6)$$

式中: $P(k_t|k_{t-1})$ 是RBF网络隐含层节点数随时间变化的概率分布; $\varepsilon_c, \varepsilon_\alpha, \varepsilon_A, \varepsilon_\mu, \varepsilon_\sigma, \varepsilon_v$ 维数分别和 $c_{t-1}, \alpha_{t-1}, A_{t-1}, \mu_{t-1}, \sigma_{t-1}^2, v_{t-1}$ 相同, 且它们中每个元素分别服从均值为0、方差为 $\delta_c^2, \delta_\alpha^2, \delta_A^2, \delta_\mu^2, \delta_\sigma^2, \delta_v^2$ 的高斯分布。 $P(z_t|z_{t-1}, A_t)$ 为隐状态随时间变化的概率分布; $P(I_t|z_t, v_t)$ 为HMM中某状态下混合高斯模型中子模型指标 I_t 的条件概率分布, 它由 z_t 和 v_t 决定。由于参数 α_t 和 c_t 的维数依赖于 k_t , 所以当 k_t 发生变化时, α_t 和 c_t 也随之变动。简单起见, 本文假设 k_t 只有如下3种变化: 1) 增加1个节点; 2) 减少1个节点; 3) 节点数不变。当 k_t 增加时, 则新增加的节点中心为此时刻的网络输入加上随机高斯扰动得到; 当 k_t 减少时, 则根据输出权重的大小决定删除某个节点, 即删除输出权重最小的那个隐含层节点。至于参数 α_t 随 k_t 的变化情况参见文献[12]。上述各参数初始值最好根据具体问题的先验知识来确定, 如果没有相应先验知识, 则可以随机生成。图2给出开始4步RBF-HMM模型各参数动态演化的示意图。

3.2 Rao-Blackwellised 粒子滤波法(Rao-Blackwellised particle filter)

考虑RBF-HMM模型, 首先要计算 $p(\theta_t|y_{1:t})$, 其中 $\theta_t = \{k_t, \alpha_t, c_t, \mu_t, \sigma_t^2, A_t, v_t, z_t, I_t\}$ 。如果是线性高斯模型, 则可以用卡尔曼滤波等精确方法来进行计算。但在多数情况下, 此概率密度的精确求解并不容易, 可以采用SMC方法来近似, 如果此时模型参数很多, 采用SMC方法求解则需要大量粒子,

从而影响计算速度, 因此可采用Rao-Blackwellised粒子滤波(RBPF)方法^[13], 其基本思想就是把高维状态空间划分成不同子空间, 其中部分子空间的概率密度函数用粒子抽样实现, 而剩余子空间的条件概率分布可以用诸如卡尔曼滤波等精确方法进行求解, 这样就能有效降低SMC方法采样空间的维数。文献[13]证明要达到相同估计精度, RBPF需要的粒子数远远少于标准SMC方法所需要的粒子数; 而文献[14]证明与标准SMC方法相比, RBPF方法能够有效降低估计方差。

对于本文RBF-HMM模型, 显然给定参数 $\{k_t, c_t, \mu_t, \sigma_t^2, A_t, v_t, z_t, I_t\}$ 条件下, 参数 α_t 的条件概率 $p(\alpha_t|r_t, y_{1:t})$ 服从高斯分布, 可以用卡尔曼滤波精确计算, 所以本文把参数 θ_t 分成2个子集合, 即 $\theta_t = \{r_t, s_t\}$, 其中: $s_t = \{\alpha_t\}$, $r_t = \{k_t, c_t, \mu_t, \sigma_t^2, A_t, v_t, z_t, I_t\}$ 。根据乘法定理, 则有

$$p(r_t, s_t|y_{1:t}) = p(r_t|y_{1:t})p(s_t|r_t, y_{1:t}). \quad (7)$$

其中用SMC方法近似计算 $p(r_t|y_{1:t})$, 而用卡尔曼滤波来求解 $p(s_t|r_t, y_{1:t})$, 这样就可以有效地降低粒子采样空间的维数, 提高计算效率。

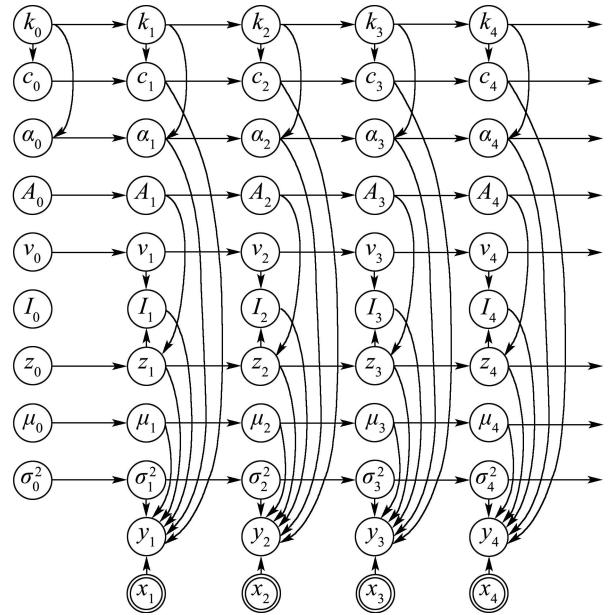


图2 RBF-HMM预测模型参数动态演化示意图

Fig. 2 Dynamic evolution graph of RBF-HMM model

3.3 在线预测算法(Algorithm of on-line prediction)

假设 $t-1$ 时刻概率密度函数 $p(s_{t-1}, r_{t-1}|y_{1:t-1})$ 可由一组已知带权粒子 $\{(r_{t-1}^{(i)}, s_{t-1}^{(i)}), w_{t-1}^{(i)}\}_{i=1}^N$ 来逼近, 即

$$p(r_{t-1}, s_{t-1}|y_{1:t-1}) \approx \sum_{i=1}^N w_{t-1}^{(i)} \delta((r_{t-1}, s_{t-1}) - (r_{t-1}^{(i)}, s_{t-1}^{(i)})),$$

其中: $\delta(\cdot)$ 是Kronecker Delta函数, N 为粒子数.

下面分别递推计算 $p(\hat{y}_t|y_{1:t-1})$ 和 $p(\mathbf{r}_t, \mathbf{s}_t|y_{1:t})$, 以此实现时间序列的在线预测, 具体步骤如下:

Step 1 粒子 $\mathbf{r}_{t-1}$ 更新.

根据RBF-HMM模型的转移方程更新粒子 $\mathbf{r}_{t-1}$, 即抽样

$$\begin{aligned} k_t^{(i)} &\sim P(k_t|k_{t-1}^{(i)}), \mathbf{c}_t^{(i)} \sim \mathbf{c}_{t-1}^{(i)} + \boldsymbol{\varepsilon}_c, \\ A_t^{(i)} &\sim A_{t-1}^{(i)} + \boldsymbol{\varepsilon}_A, \boldsymbol{\mu}_t^{(i)} \sim \boldsymbol{\mu}_{t-1}^{(i)} + \boldsymbol{\varepsilon}_\mu, \\ \ln(\boldsymbol{\sigma}_t^{2(i)}) &\sim \ln(\boldsymbol{\sigma}_{t-1}^{2(i)}) + \boldsymbol{\varepsilon}_\sigma, \mathbf{v}_t^{(i)} \sim \mathbf{v}_{t-1}^{(i)} + \boldsymbol{\varepsilon}_v, \\ z_t^{(i)} &\sim P(z_t|z_{t-1}^{(i)}, A_t^{(i)}), I_t^{(i)} \sim P(I_t|z_t^{(i)}, \mathbf{v}_t^{(i)}), \end{aligned}$$

得新粒子

$$\mathbf{r}_t^{(i)} = \{k_t^{(i)}, \mathbf{c}_t^{(i)}, A_t^{(i)}, \mathbf{v}_t^{(i)}, \boldsymbol{\mu}_t^{(i)}, \boldsymbol{\sigma}_t^{2(i)}, z_t^{(i)}, I_t^{(i)}\},$$

则

$$p(\mathbf{r}_t|y_{1:t-1}) \approx \sum_{i=1}^N w_{t-1}^{(i)} (\mathbf{r}_t - \mathbf{r}_t^{(i)}). \quad (8)$$

Step 2 测量值预测.

当 $\mathbf{r}_t^{(i)} = \{k_t^{(i)}, \mathbf{c}_t^{(i)}, A_t^{(i)}, \mathbf{v}_t^{(i)}, \boldsymbol{\mu}_t^{(i)}, \boldsymbol{\sigma}_t^{2(i)}, z_t^{(i)}, I_t^{(i)}\}$ 已知时, 根据定义可得 $D_t^{(i)}$. 又由于 e_t 服从HMM过程, 在特定隐状态、特定混合高斯子模型下它为高斯分布, 即 $p(e_t^{(i)}|z_t^{(i)}, I_t^{(i)}) = N(e_t^{(i)}; \boldsymbol{\mu}_t^{(i)}(z_t^{(i)}, I_t^{(i)}), \boldsymbol{\sigma}_t^{2(i)}(z_t^{(i)}, I_t^{(i)}))$, 所以可取 $e_t^{(i)} = \boldsymbol{\mu}_t^{(i)}(z_t^{(i)}, I_t^{(i)})$. 根据式(5)可得 $\hat{y}_t^{(i)} = D_t^{(i)} \boldsymbol{\alpha}_{t-1}^{(i)} + \boldsymbol{\mu}_t^{(i)}(z_t^{(i)}, I_t^{(i)})$. 则

$$p(\hat{y}_t|y_{1:t-1}) \approx \sum_{i=1}^N w_{t-1}^{(i)} (\hat{y}_t - \hat{y}_t^{(i)}), \quad (9)$$

所以 $\hat{y}_t = \sum_{i=1}^N w_{t-1}^{(i)} (D_t^{(i)} \boldsymbol{\alpha}_{t-1}^{(i)} + \boldsymbol{\mu}_t^{(i)}(z_t^{(i)}, I_t^{(i)}))$.

Step 3 粒子选择及权重更新.

当获得 t 时刻的真实测量值 y_t 时, 计算 $p(y_t|\mathbf{r}_t^{(i)}, y_{1:t-1})$ 得到每个新粒子 $\mathbf{r}_t^{(i)}$ 的权重. 如果多数粒子的权重很小(接近于0), 则会引起退化现象, 为了防止这种情况出现, 本文采用残差重采样方法^[15]和马尔可夫链蒙特卡罗(MCMC)方法^[16]对粒子进行重采样. 经过上述操作后, 得到 N 个权重相等的新粒子, 即 $w_t^{(i)} = 1/N$, 简洁起见, 这里将这些新粒子仍然记为 $\mathbf{r}_t^{(i)}$, 则

$$p(\mathbf{r}_t|y_{1:t-1}) \approx \sum_{i=1}^N w_t^{(i)} (\mathbf{r}_t - \mathbf{r}_t^{(i)}). \quad (10)$$

Step 4 粒子 $\mathbf{s}_{t-1}$ 更新.

由于条件概率 $p(\mathbf{s}_t|\mathbf{r}_t, y_{1:t})$ 在特定隐状态、特定混合高斯模型下服从高斯分布, 因此可以用卡尔曼滤波来计算 $p(\mathbf{s}_t|\mathbf{r}_t^{(i)}, y_{1:t})$ 得到 $\mathbf{s}_t^{(i)}$, 详见文献[17]. 即得新粒子 $\{\mathbf{s}_t^{(i)}\}_{i=1}^N$, 其中 $\mathbf{s}_t^{(i)} = \boldsymbol{\alpha}_t^{(i)}$.

通过以上几节的讨论, 可得 $p(\hat{y}_t|y_{1:t-1})$, $p(\mathbf{r}_t|y_{1:t})$ 和 $p(\mathbf{s}_t|\mathbf{r}_t, y_{1:t})$. 根据式(7), 则可计算得 $p(\mathbf{r}_t, \mathbf{s}_t|y_{1:t})$, 同时得到 t 时刻的预测值 $\hat{y}_t$.

4 实验研究(Experimental research)

4.1 太阳黑子数平滑月均值序列(Time series of smoothed monthly mean of sunspot numbers)

本节选取太阳黑子数平滑月均值序列, 样本区间为1964年1月至2005年10月(共502个数据), 数据来源于http://sidc.oma.be/. 为了提高预测精度, 首先对原始数据进行1阶差分以去除序列低频成分, 然后用RBF-HMM模型进行单步在线预测, 最后还原成太阳黑子数月均值预测值. 并取1997年1月至2005年10月共106个样本来检验模型预测精度.

通过对1964年1月至1996年12月的396个样本进行离线训练, 得到RBF网络输入参数为 $p = 10$, $q = 5$, 隐含层节点数 $k = 10$, 其中 $\lambda = 1$. 本例根据试值法确定HMM的混合高斯模型个数为2. 将离线训练得到的RBF-HMM模型参数作为在线预测的初始参数, 然后根据第3节的算法实现在线预测. 关于粒子数目选择, 从理论上讲, 当粒子数目趋于无穷时, 用一些带权的随机粒子能无限逼近任意概率分布. 但由于计算机的运算和储存能力限制, 粒子数目不可能取无穷大, 在实际运用中, 既要考虑现实环境的限制, 又要兼顾估计准确性, 借鉴文献[18]中建议, 本例中粒子数目取1000. 此外, 本文还采用前期法和观测噪声服从零均值高斯分布的RBF网络(RBF-GAUSS)模型对该序列进行预测. 前期法就是把前一时刻实际值作为当前时刻预测值, 即 $\hat{y}_t = y_{t-1}$. 并给出文献[19]中方法的预测结果, 各种预测方法的平均相对误差(MAPE)见表1.

表1 太阳黑子数平滑月均值的平均相对预测误差
Table 1 MAPE of smoothed mean sunspot numbers

预测方法	MAPE / %
RBF-HMM模型	1.24
RBF-GAUSS模型	1.42
文献[19]中方法	2.30
前期法	3.98

从表1可看出, RBF-HMM模型和RBF-GAUSS模型预测结果优于文献[19]中方法, 其原因在于RBF-HMM模型和RBF-GAUSS模型的相关参数及网络隐节点数随着时间可动态调整, 而文献[19]中模型结构和参数经训练后即保持不变; RBF-HMM模型预测结果优于RBF-GAUSS模型, 这是因为任何观测序列都有其根本驱动力, HMM正是从本质上对此进行描述, 即根据噪声内在驱动机制的不同状态对神经网络预测结果(即神经网络输出)进行在线补偿校正, 而RBF-GAUSS模型仅仅假设噪声服从零均值高斯分布, 没有对神经网络的输出进行补偿校正; 前期法只是把第 $t-1$ 期的实际值作为第 t 期的预测值, 事实

上第 t 期通常不会是第 $t-1$ 期的简单重复, 所以其预测效果最差。

4.2 CRU钢材价格指数序列(Time series of CRU steel price index)

本小节选取英国商品研究所发布的CRU国际钢材价格指数月数据序列进行研究, 样本区间为1994年4月至2008年3月(共168个数据), 数据来源于<http://cruonline.crugroup.com>。用2007年1月至2008年3月的15个样本来检验模型预测精度。对1994年4月至2006年12月的153个样本进行离线训练, 通过试值法确定RBF网络输入参数为 $p=5$, $q=4$, 隐含层节点数为 $k=7$, 其余参数确定方法同4.1节。此外还采用前期法和RBF-GAUSS模型对该序列进行预测, 具体结果见表2。由此可看出, 基于RBF-GAUSS模型的预测效果优于前期法, 而RBF-HMM模型的预测效果又优于RBF-GAUSS模型, 进一步验证本文模型的有效性。

表2 CRU钢材价格指数的平均相对预测误差
Table 2 MAPE of CRU steel price index

预测方法	MAPE / %
RBF-HMM模型	1.68
RBF-GAUSS模型	2.56
前期法	3.01

5 结论(Conclusion)

针对观测噪声为非高斯分布的非线性时间序列, 提出噪声服从HMM的RBF-HMM在线预测模型。该模型具有以下3个显著特点: 1) 模型输入包含误差反馈项; 2) 采用节点数可变的RBF网络对NARMA模型中非线性项进行建模; 3) 用HMM对观测噪声项进行建模。并采用SMC方法实现基于RBF-HMM模型的时间序列在线预测。最后采用太阳黑子数平滑月均值数据和CRU国际钢材价格指数进行实验研究, 并和前期法、RBF-GAUSS模型的预测结果进行对比研究, 结果表明本文模型的有效性。

参考文献(References):

- [1] BOX G E P, JENKINS G M, REINSEL G C. *Time Series Analysis: Forecasting and Control* [M]. 3rd edition. Beijing: Posts & Telecom Press, 2005: 19–180.
- [2] CHEN S, BILLINGS S A. Representation of non-linear systems: the NARMAX model[J]. *International Journal of Control*, 1989, 49(3): 1013–1042.
- [3] HORNIK K, STINCHCOMBE M, WHITE H. Multi-layer feed-forward networks are universal approximators[J]. *Neural Networks*, 1989, 2(3): 359–366.
- [4] 张高煜, 江水, 梁继民, 等. 采用序贯滤波的红外/雷达机动目标跟踪[J]. *控制理论与应用*, 2007, 24(5): 811–814.
(ZHANG Gaoyu, JIANG Shui, LIANG Jimin, et al. Infrared/radar maneuvering target tracking via sequence filtering[J]. *Control Theory & Applications*, 2007, 24(5): 811–814.)
- [5] DE FREITAS N, ANDRIEU C, HOJEN-SORENSEN P, et al. *Sequential Monte Carlo methods for neural networks*[M]//*Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag, 2001.
- [6] XU X, LI B. Adaptive Rao-Blackwellised particle filter and its evaluation for tracking in surveillance[J]. *IEEE Transactions on Image Processing*, 2007, 16(3): 838–849.
- [7] HOLMES C, MALLICK B K. Bayesian radial basis function of variable dimension[J]. *Neural Computation*, 1998, 10(5): 1217–1233.
- [8] POGGIO T, GIROSI F. Networks for approximation and learning[J]. *Proceedings of the IEEE*, 1990, 78(9): 1481–1497.
- [9] RABINER L R. A tutorial on hidden Markov Models and selected applications in speech recognition[J]. *Proceedings of the IEEE*, 1989, 77(2): 257–286.
- [10] LIPORACE L A. Maximum likelihood estimation for multivariate observations of Markov sources[J]. *IEEE Transactions on Information Theory*, 1982, IT-28(5): 729–734.
- [11] AVRULAMPALAM S, MASKELL S, GORDON N, CLAPP T. A tutorial on particle filters for on-line non-linear/ non-Gaussian Bayesian tracking[J]. *IEEE Transactions on Signal Processing*, 2002, 50(2): 174–188.
- [12] ANDRIEU C, DE FREITAS J F G, DOUCET A. Sequential Bayesian estimation and model selection applied to neural networks, Technical Report CUED/F-INFENG/TR 341[R]. Cambridge: Department of Engineering, Cambridge University, 1999.
- [13] DOUCET A, DE FREITAS J F G, MURPHY K, RUSSELL S. Rao-Blackwellised particle filtering for dynamic Bayesian networks[C]//*Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*. Stanford: Morgan Kaufmann Publishers Inc, 2000: 176–183.
- [14] DOUCET A, GORDON N J, KRISHNAMURTHY V. *Particle filters for state estimation of jump Markov linear systems*(Technical Report CUED/F-INFENG/TR 359)[R]. Cambridge: Department of Engineering, Cambridge University, 1999.
- [15] LIU J S, CHEN R. Sequential Monte Carlo methods for dynamic systems[J]. *Journal of the American Statistical Association*, 1998, 93(443): 1022–1031.
- [16] GILKS W R, BERZUINI C. Following a moving target-monte carlo inference for dynamic Bayesian models[J]. *Journal of the Royal Statistical Society B*, 2001, 63(1): 127–146.
- [17] DE FREITAS J F G. *Bayesian Methods for Neural Networks*[D]. Cambridge: University of Cambridge, 1999.
- [18] Goldenstein S K. A gentle introduction to predictive filters[J]. *Revista de Informática Teórica e Aplicada*, 2004, 6(1): 235–297.
- [19] 丁刚, 钟诗胜. 基于时变阈值过程神经网络的太阳黑子数预测[J]. *物理学报*, 2007, 56(2): 1224–1230.
(DING Gang, ZHONG Shisheng. Sunspot number prediction based on process neural network with time-varying threshold functions[J]. *Acta Physica Sinica*, 2007, 56(2): 1224–1230.)

作者简介:

张冬青 (1971—), 女, 博士研究生, 目前研究方向为经济信号处理、时间序列预测等, E-mail: hollycase@sina.com;

宁宣熙 (1938—), 男, 教授, 博士生导师, 目前研究方向为管理科学与工程、系统工程等, E-mail: nhning@nuaa.edu.cn;

刘雪妮 (1978—), 女, 博士研究生, 目前研究方向为区域经济等, E-mail: applelxl@hotmail.com.