文章编号：1000−8152(2006)05−0671−08

非线性时间序列预报的隐多分辨ARMA模型

高　伟[1], 田　铮[1,2]

(1. 西北工业大学 应用数学系, 陕西 西安 710072; 2. 中国科学院 自动化研究所 模式识别国家重点实验室, 北京 100080)

摘要: 研究一类用于非线性时间序列预报的隐多分辨自回归滑动平均(ARMA)模型, 该模型以ARMA模型为初始细水平模型(即隐多分辨模型的基本块). 证明了模型的建模精度由水平间的方差决定. 研究了新模型的自相关函数结构, 给出了参数估计的Bayes方法和Metropolis-Hasting算法. 进一步提出了一种可以直接用于不同基本块的隐多分辨模型的非线性时间序列预报方法, 证明了其比其他的线性预报方法和隐多分辨模型预报方法降低了预报误差. 最后通过数值模拟和实例验证了模型和预报方法, 并和其他模型进行比较, 结果表明新提出模型和预报方法能够更好地描述数据的特征, 提高预报的精度.

关键词: 非线性时间序列预报; 隐多分辨自回归滑动平均模型; 自相关函数

中图分类号: O211　　**文献标识码**: A

Hidden multi-resolution ARMA models for nonlinear time series forecast

GAO Wei[1], TIAN Zheng[1,2]

(1. Department of Applied Mathematics, Northwest Polytechnical University, Xi'an Shaanxi, 710072, China;
2. National Key Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China)

Abstract: A class of hidden multi-resolution autoregressive moving average (ARMA) model is studied for forecasting nonlinear time series. The model has ARMA model as the original fine level model, that is, the building blocks. The precision of the model for approximating the true one is determined by the variance among the levels. The autocorrelation functions (ACF) structure of the new model is then studied. The estimation of parameters is easily performed via Bayes method and Metropolis-Hasting algorithm. Furthermore, a new method for nonlinear time series forecast is proposed. The method can be directly applied to hidden multi-resolution model with different building blocks, and reduce the forecasting error compared with other linear method and hidden multi-resolution model forecast method. Finally, the model and approaches are illustrated through the use of both simulated and real series. The new model and forecasting method appear to capture features of the data better and provide more precise forecasting than other competing models do.

Key words: nonlinear time series forecast; hidden multi-resolution autoregressive moving average model; autocorrelation functions

1　引言(Introduction)

非线性时间序列的最佳预报与最佳控制的理论和方法是时间序列分析在工程中应用的一个重要方面. 而最佳控制是在最佳预报的基础上进行的, 因此对于工程应用而言, 获得非线性时间序列的最佳预报就具有重要意义[1]. 线性时间序列预报方法(如Durbin-Levinson递归预报、新息方法和递归预报方法等)本质上是考虑观测数据间的低阶相关, 即基于观测数据的样本自相关函数和样本偏自相关函数, 这些方法难以反映动态数据更复杂的概率结构, 如多峰、拖尾等等. 文献[2]提出了一种隐分辨模型(Hidden Resolution Model), 综合了小波多分辨分析和Kalman滤波的优点, 可以处理具有高阶相关特征的数据. 但是文献[2]没有从理论上对利用隐多分辨模型进行建模和预报的精度进行定量分析, 还应注意到用比AR(1)更复杂的模型(如ARMA模型)作为基本块建立的隐多分辨模型, 其对复杂概率结构动态数据建模和预报能力可能要优于隐多分辨AR(1)模型.

综上所述, 本文分析了隐多分辨模型的建模精度, 证明了所建立模型的自相关函数(autocorrelation functions, 简记为ACF)结构与真实ACF结构的误差由水平间的方差决定. 研究了以ARMA模型作为基本块的隐分辨模型, 讨论了其更复杂的ACF结构. 提出了一种可以直接用于各种不同基本块的隐多分辨模型的预报方法. 预报的初始均值向量和协方

收稿日期: 2005−06−01; 收修改稿日期: 2005−10−31.
基金项目: 国家自然科学基金资助项目 (60375003); 国家航空基础项目 (03153059).

差矩阵来自于所构造的隐多分辨模型, 而不是原来方法中假设的初始过程, 由于隐多分辨模型比假设的初始过程能更准确的描述数据, 因此新的预报方法建立在更精确的初始信息之上. 证明了在合适的模型假定下, 新的预报方法在预报误差方差上优于原有的方法. 仿真及实例分析的结果都表明隐多分辨ARMA模型是一类有着广泛应用前景和研究价值的非线性时间序列模型.

2 隐多分辨ARMA模型及建模精度分析(Hidden multi-resolution ARMA models and modeling accuracy analysis)

用$X = \{X_t\}(t = 1, 2, \cdots)$表示要研究的时间序列, 其到时刻$n_x$的随机列向量记为$X_{1:n_x} = (X_1, \cdots, X_{n_x})'$. $p(\cdot)$和$q(\cdot)$为分布密度函数. 将X视为细水平过程, 对确定的正整数m(粗化窗宽), 构造粗水平过程Y为细水平的加噪平均, $Y_s = m^{-1}\sum\limits_{i=1}^{m} X_{(s-1)m+i} + u_s$, $s = 1, 2, \cdots$, 其中$u_s \sim N(0, \tau)$, 是i.i.d.噪声项, τ为水平间的方差.

假设 1 设细水平是某一确定模型, 本文用平稳可逆的ARMA(1,1)模型

$$X_t = \phi_x X_{t-1} + \varepsilon_t + \theta_x \varepsilon_{t-1}, \tag{1}$$

其中$\varepsilon_t \sim N(0, \sigma_x^2)$, 是i.i.d.噪声序列, 则

$$X_{1:n_x} \sim N(0, V_x). \tag{2}$$

假设 2 设得到关于粗水平的附加信息, 如$Y_{1:n_y}$是ARMA(1,1)模型

$$Y_s = \phi_y Y_{s-1} + \eta_s + \theta_y \eta_{s-1}, \tag{3}$$

其中$\eta_s \sim N(0, \sigma_y^2)$, 则$Y_{1:n_y} \sim N(0, Q_y)$.

由粗水平构造过程, 到时刻n_x构造的粗水平随机列向量$Y_{1:n_y} = (Y_1, \cdots, Y_{n_y})'$, 其中$n_y = n_x/m$, 水平间的变换系数矩阵$A$为满足$Y_{1:n_y} = AX_{1:n_x}$的$n_y \times n_x$矩阵. 水平间的关联方程即已知$X_{1:n_x}$条件下$Y_{1:n_y}$ 的条件密度函数

$$\begin{aligned}(Y_{1:n_y}|X_{1:n_x}) \sim \prod_{s=1}^{n_y} N(m^{-1}\sum_{i=1}^{m} X_{(s-1)m+i}, \tau) = \\ N(AX_{1:n_x}, U).\end{aligned} \tag{4}$$

其中U为向量$u_{1:n_y}$的协方差阵.

为便于参数估计, 令$\tau = \lambda(AV_xA')_{11}$, 其中$(AV_xA')_{11}$表示矩阵$AV_xA'$的第1行第1列元素, 即将$\tau$参数化为$AV_xA'$的函数.

由Bayes公式[3]和线性回归的方法, 已知$Y_{1:n_y}$条件下$X_{1:n_x}$的条件密度函数

$$(X_{1:n_x}|Y_{1:n_y}) \sim N(BY_{1:n_y}, V_x - BWB'). \tag{5}$$

其中: $B = V_xA'W^{-1}, W = AV_xA' + U$.

应用Bayes定理和Jeffrey条件准则[2], 修正的细水平变量的联合分布为

$$\begin{aligned}q(X_{1:n_x}) = \int p(X_{1:n_x}|Y_{1:n_y})q(Y_{1:n_y})dY_{1:n_y} \propto \\ \int N(BY_{1:n_y}, V_x - BWB')N(0, Q_y)\mathrm{d}Y_{1:n_y}.\end{aligned} \tag{6}$$

所得到的分布$q(X_{1:n_x})$(即要建立的隐多分辨模型) 是多维正态分布: $X_{1:n_x} \sim N(0, Q_x)$. 其中: $Q_x = V_x - B(W - Q_y)B'$, $W = AV_xA' + U$, $B = V_xA'W^{-1}$.

由上述过程构造的隐多分辨模型用粗水平信息修正初始细水平假设, 模型对数据的建模精度由下面定理给出.

定理 1 隐多分辨ARMA模型用粗水平的信息修正了初始细水平假设, 所构造的新模型的协方差阵与真实的细水平模型协方差阵之间的误差由水平间的协方差阵U 决定.

证 真实模型的协方差阵表示为V_x^{T}, 由式(4)

$$\begin{aligned}Q_y =& AV_x^{\mathrm{T}}A' + U,\\ Q_x =& V_x - B(W - Q_y)B' = \\ & V_x - BA(V_x - V_x^{\mathrm{T}})A'B'.\end{aligned} \tag{7}$$

由于Q_x的复杂形式, Q_x对V_x^{T}的估计精度很难直接获得, 考虑到计算问题, 将式(7)转化到粗水平

$$\begin{aligned}AQ_xA' + U =& AV_xA' + U - AB(W - Q_y)B'A' = \\ & W - AB(W - Q_y)B'A'.\end{aligned} \tag{8}$$

由于A为常数矩阵, 要证明Q_x对V_x^{T}的误差比V_x对V_x^{T} 的误差小, 只需计算式(8)中的矩阵对Q_y的估计精度. 即计算$AB(W - Q_y)B'A'$对$(W - Q_y)$的估计精度, 因此矩阵AB与单位阵的误差就可作为精度的度量.

由式(4)和参数化$\tau = \lambda(AV_xA')_{11}$的方法

$$AB = AV_xA'W^{-1} = (W - U)W^{-1} = I - UW^{-1},$$

矩阵AB与单位阵的误差$(I - AB)W = U$. 定理证毕.

通过第4节介绍的隐多分辨模型的预报过程, 定理1不仅给出了模型的建模精度, 还给出了模型对未来值进行预报的预报精度.

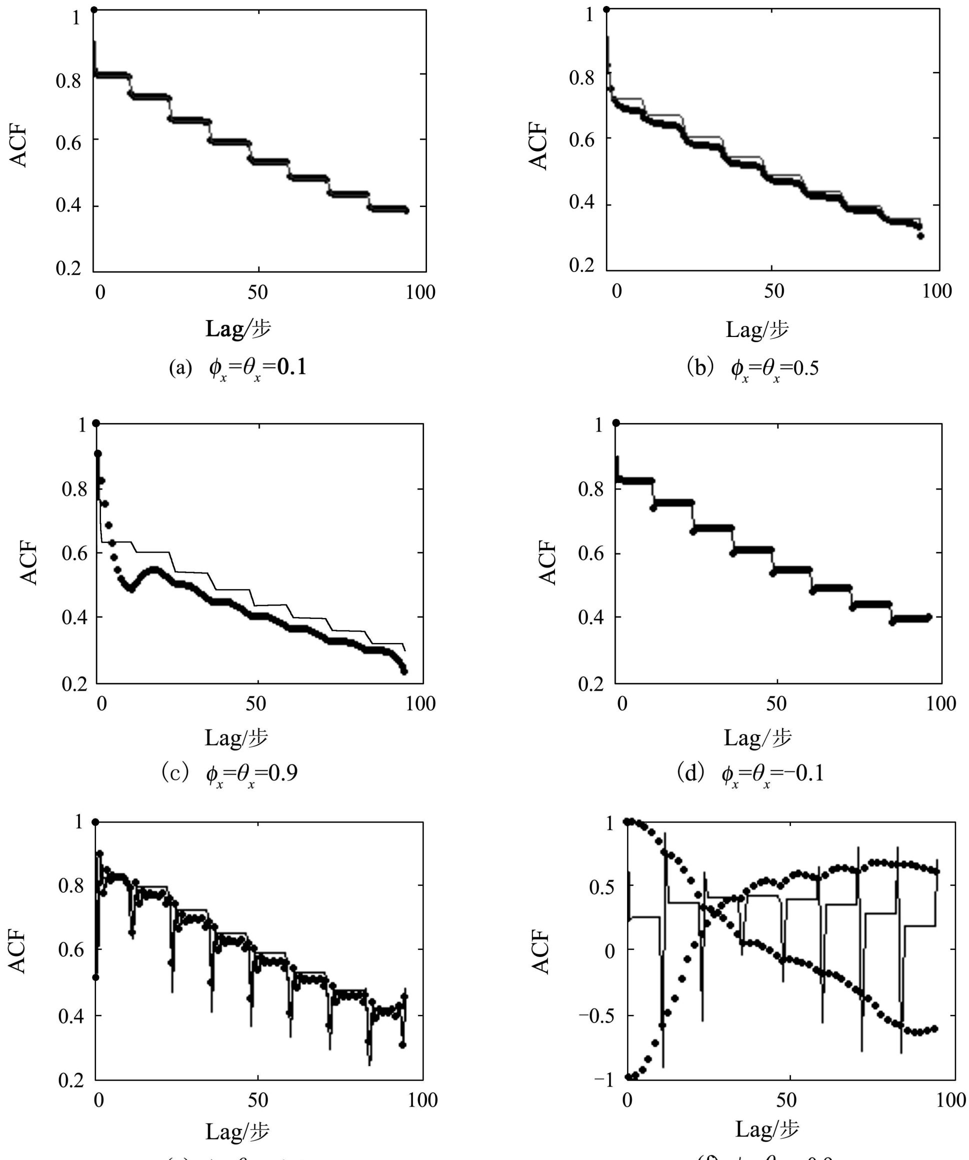
3 不同基本块的隐多分辨模型的自相关函数结构比较(Comparison of the ACF structure in hidden multi-resolution models with different building blocks)

本节以粗水平为AR(1), 细水平分别为AR(1), MA(1)和ARMA(1,1)的隐多分辨模型为例, 研究模

型的自相关函数结构.

图1和图2描述了粗水平相同, 细水平模型和参数变化时隐多分辨模型的ACF变化. 各参数值为:

$\phi_y = 0.9, \sigma_y^2 = 1, \sigma_x^2 = 1,$
$\lambda = 0.1, n_x = 96,$
$m = 12.$



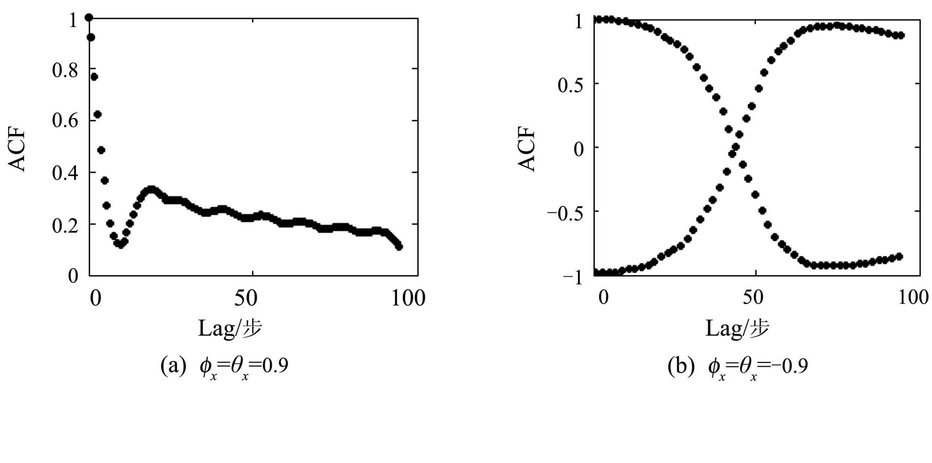
(a) $\phi_x=\theta_x=0.1$　(b) $\phi_x=\theta_x=0.5$

(c) $\phi_x=\theta_x=0.9$　(d) $\phi_x=\theta_x=-0.1$

(e) $\phi_x=\theta_x=-0.5$　(f) $\phi_x=\theta_x=-0.9$

图 1 粗水平为AR(1), 细水平分别为AR(1)(点线)和MA(1)(实线)的隐多分辨模型的自相关函数

Fig. 1 ACF of hidden multi-resolution models with AR(1) as coarse level and AR(1) (dot line) and MA(1)(real line) as fine level respectively



(a) $\phi_x=\theta_x=0.9$　(b) $\phi_x=\theta_x=-0.9$

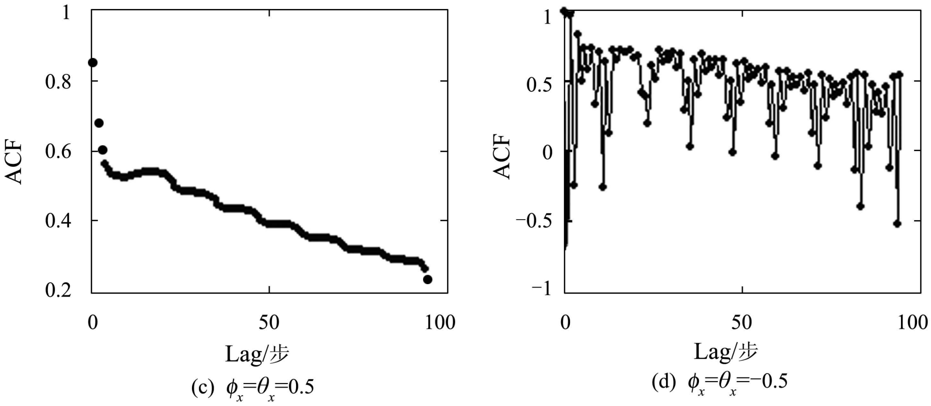


图 2 粗水平为AR(1), 细水平为ARMA(1,1)的隐多分辨模型的自相关函数

Fig. 2 ACF of hidden multi-resolution models with AR(1) as coarse level and ARMA(1,1) as fine level

通过比较可发现隐多分辨ARMA模型(图2)表示的ACF结构是隐多分辨AR和隐多分辨MA模型(图1)所不具有的. 由于ARMA模型包含AR模型和MA模型, 从这个意义上隐多分辨ARMA模型更值得研究.

由模型的构造过程, 细水平模型假设主要用来参数化初始协方差阵V_x. 不同的细水平模型假设通过协方差阵影响隐多分辨模型. 将AR(1)和MA(1)的协方差阵分别表示为$V_x^{\rm AR}$和$A_x^{\rm MA}$. 当ϕ_x的绝对值较大时, AR(1)的ACF衰减缓慢, 当θ_x的绝对值较大时, ACF明显表现出一阶截尾性质, $V_x^{\rm AR}$和$A_x^{\rm MA}$的差别就很明显. 当参数ϕ_x和θ_x的绝对值较小, 且$\phi_x=\theta_x$时, 有$V_x^{\rm AR}\approx V_x^{\rm MA}$. 因此给定相同的粗水平模型的条件下, 细水平模型的假设AR(1)和MA(1)$\phi_x=\theta_x$ 导致相同的隐多分辨模型ACF结构. 图1表明当参数的绝对值足够大时(图1(c)和(f)), 隐多分辨AR(1)和隐多分辨MA(1)模型的ACF具有明显区别, 因此在对细水平模型进行假设时, 应结合样本的ACF来考虑.

4 隐多分辨ARMA模型的参数估计和预报(Parameter estimation and forecast of hidden multi-resolution ARMA models)

4.1 隐多分辨ARMA模型的参数估计(Parameter estimation of hidden multi-resolution ARMA models)

对$X_{1:n_x}=(X_1,\cdots,X_{n_x})$, 用小写符号表示其对应的观测值, $x_{1:n_x}=(x_1,\cdots,x_{n_x})$, 构造的粗水平值. $y_{1:n_y}=(y_1,\cdots,y_{n_y})$. 隐多分辨ARMA(1,1)模型的参数包含粗水平参数$\varGamma=(\phi_y,\theta_y,\sigma_y^2)$和细水平及关联方程参数$\varDelta=(\phi_x,\theta_x,\sigma_x^2,\lambda)$两部分. 模型建立过程的式(5)和式(6)表明对应于不同水平的参数可以独立进行估计.

设细水平和关联方程参数的先验分布分别为: $\phi_x\sim{\rm Tr}N_{(-1,1)}(m_{\phi_x},S_{\phi_x})$, $\theta_x\sim{\rm Tr}N_{(-1,1)}(m_{\theta_x},S_{\theta_x})$, $\sigma_x^2\sim{\rm I}Ga(v_{\sigma_x}/2,v_{\sigma_x}s_{\sigma_x}/2)$, $\lambda\sim{\rm TrI}Ga_{(0,10)}(v_\lambda/2,v_\lambda s_\lambda/2)$. 其中: I$G$表示逆伽马分布, 即$X\sim{\rm I}Ga(a,b)\Leftrightarrow 1/X\sim Ga(a,b)$; ${\rm Tr}N_{(a,b)}$和${\rm TrI}Ga_{(a,b)}$分别表示在区间(a,b)上截断的正态分布和逆伽马分布.

参数$\varDelta=(\phi_x,\theta_x,\sigma_x^2,\lambda)$用Metropolis-Hasting方法进行估计. 对隐多分辨ARMA(1,1)模型有下式成立:

$$\begin{aligned}&p(\phi_x,\theta_x,\sigma_x^2,\lambda|X_{1:n_x},\lambda|Y_{1:n_y})\propto\\&p(\phi_x,\theta_x,\sigma_x^2)p(X_{1:n_x}|\phi_x,\theta_x,\sigma_x^2)\cdot\\&p(\lambda)\frac{p(Y_{1:n_y}|X_{1:n_x},\phi_x,\theta_x,\sigma_x^2,\lambda)}{p(Y_{1:n_y}|\phi_x,\theta_x,\sigma_x^2,\lambda)}=\\&p(\phi_x,\theta_x,\sigma_x^2)N(0,V_x)p(\lambda)\cdot\\&\frac{\prod\limits_i^{n_y}N(A_iX_{1:n_x},\lambda(AV_xA')_{11})}{N(0,AVA'+U)}.\end{aligned}\tag{9}$$

其中A_i表示矩阵A的第i个行向量. 式(9)用于简化和加速参数$(\phi_x,\theta_x,\sigma_x^2,\lambda)$的模拟和计算每一步的接受概率.

4.2 基于隐多分辨ARMA模型的预报(Forecast based on hidden multi-resolution ARMA models)

本节讨论基于隐多分辨ARMA模型的预报问题. 现有的方法以AR(1)为基本块建立隐多分辨模型, 令$r=x_{n_x-m}(\phi_x,\phi_x^2,\cdots,\phi_x^m)'$, $R_{ij}=\sigma_x^2\phi_x^{|i-j|}(1-\phi_x^{2\min(i,j)})/(1-\phi_x^2)$分别为用初始AR(1)模型构造的的细水平$m$-步预报均值向量和误差协方差阵. Y_s在给定Y直到时刻$s-1$的值下的初始分布, $(Y_s|Y_{1:s-1})\sim N(p_s,P_s)$可用Kalman滤

波计算.

已有的方法如下得到预报分布:

a) $(Y_{n_y}|Y_{1:(n_y-1)}, X_{1:(n_x-m)}) \sim N(f_y, F_y)$, (10)

其中

$$F_y = [\sigma_y^{-2} + (m^{-2}\overrightarrow{1}'R\overrightarrow{1} + \tau)^{-1} + P_{n_y}^{-1}]^{-1}, \quad (11)$$

$$f_y = F_y[\sigma_y^{-2}\phi_y y_{n_y-1} + m^{-1}(m^{-2}\overrightarrow{1}'R\overrightarrow{1} + \tau)^{-1}\overrightarrow{1}'r + P_{n_y}^{-1}p_{n_y}], \quad (12)$$

其中$\overrightarrow{1}$表示元素为1的m维列向量.

b) $(X_{(n_x-m+1):n_x}|Y_{n_y}, X_{1:(n_x-m)}) \sim N(f_x, F_x)$, (13)

其中

$$F_x = R - m^{-2}R\overrightarrow{1}(m^{-2}\overrightarrow{1}'R\overrightarrow{1} + \tau)^{-1}\overrightarrow{1}'R,$$

$$f_x = r + m^{-1}R\overrightarrow{1}(m^{-2}\overrightarrow{1}'R\overrightarrow{1} + \tau)^{-1}(\hat{y}_{n_y} - m^{-1}\overrightarrow{1}'r).$$

$\hat{y}_{n_y}$为上面a)步所得到的预报值.

隐多分辨模型的预报过程和建模过程类似: 用各种观测信息和预报信息进行修正得到最终关于细水平的预报结果. 细水平的预报分布协方差阵由(5)而来, 均值向量用粗水平的估计值作修正.

本文提出的预报方法按如下步骤得到预报分布:

1) 计算初始值$r^{(0)}$和$R^{(0)}$.

设对观测数据已建立隐多分辨模型

$$X_{1:n_x} \sim N(0, Q_x),$$

$$Q_x = V_x - B(W - Q_y)B'.$$

初始m-步向前预报均值向量和协方差阵为n_x维正态分布$N(0, Q_x)$的条件均值和条件协方差矩阵: $r^{(0)} = S_{12}S_{22}^{-1}(x_1, \cdots, x_{n_x-m})'$和$R^{(0)} = S_{11} - S_{12}S_{22}^{-1}S_{21}$, 其中$S_{11}$是一$m \times m$矩阵, 其元素为$Q_x$的$n_x - m + 1$到$n_x$行和列; S_{22}是由Q_x的第1到$n_x - m$行和第1到$n_x - m$列组成的$(n_x - m) \times (n_x - m)$矩阵; $S_{21} = S_{12}'$, 具体如下:

$$S_{11} = \begin{bmatrix} Q_{n_x-m+1,n_x-m+1} & \cdots & Q_{n_x-m+1,n_x} \\ \vdots & \vdots & \vdots \\ Q_{n_x,n_x-m+1} & \cdots & Q_{n_x,n_x} \end{bmatrix},$$

$$S_{22} = \begin{bmatrix} Q_{1,1} & \cdots & Q_{1,n_x-m} \\ \vdots & \vdots & \vdots \\ Q_{n_x-m,1} & \cdots & Q_{n_x-m,n_x-m} \end{bmatrix},$$

$$S_{12} = \begin{bmatrix} Q_{n_x-m+1,1} & \cdots & Q_{n_x-m+1,n_x-m} \\ \vdots & \vdots & \vdots \\ Q_{n_x,1} & \cdots & Q_{n_x,n_x-m} \end{bmatrix}.$$

2) 获得m-步向前预报的均值向量和协方差阵$f_x^{(1)}$和$F_x^{(1)}$.

a) $(Y_{n_y}|Y_{1:(n_y-1)}, X_{1:(n_x-m)}) \sim N(f_y^{(1)}, F_y^{(1)})$,

其中:

$$F_y^{(1)} = [\sigma_y^{-2} + (m^{-2}\overrightarrow{1}'R^{(0)}\overrightarrow{1} + \tau)^{-1} + (P_{n_y})^{-1}]^{-1}, \quad (14)$$

$$f_y^{(1)} = F_y^{(1)}[\sigma_y^{-2}\phi_y y_{n_y-1} + m^{-1}(m^{-2}\overrightarrow{1}'R^{(0)}\overrightarrow{1} + \tau)^{-1}\overrightarrow{1}'r^{(0)} + (P_{n_y})^{-1}p_{n_y}]; \quad (15)$$

b) $(X_{(n_x-m+1):n_x}|Y_{n_y, X_{1:(n_x-m)}}) \sim N(f_x^{(1)}, F_x^{(1)})$,

其中:

$$F_x^{(1)} = R^{(0)} - m^{-2}R^{(0)}\overrightarrow{1}(m^{-2}\overrightarrow{1}'R^{(0)}\overrightarrow{1} + \tau)^{-1}\overrightarrow{1}'R^{(0)},$$

$$f_x^{(1)} = r^{(0)} + m^{-1}R^{(0)}\overrightarrow{1}(m^{-2}\overrightarrow{1}'R^{(0)}\overrightarrow{1} + \tau)^{-1}(\hat{y}_{n_y} - m^{-1}\overrightarrow{1}'r^{(0)}).$$

3) 设已得到第$l \times m$-步的预报结果, 计算第$(l+1) \times m$步的初始预报信息$r^{(l)}$和$R^{(l)}$.

设已得到第$l \times m$-步向前预报值, 预报均值向量和协方差阵分别为

粗水平的预报信息:

$$\hat{y}_{n_y}, \cdots, \hat{y}_{n_y+l-1}, f_y^{(l)}, F_y^{(l)};$$

细水平的预报信息:

$$\hat{x}_{n_x-m+1}, \cdots, \hat{x}_{n_x+(l-1)m},$$

$$f_x^{(l)} = (f_{x1}^{(l)}, \cdots, f_{xm}^{(l)}), F_x^{(l)}.$$

令$F_y^{(0)}$, $F_x^{(0)}$为零矩阵. 则预报的初始均值向量和协方差阵分别为

$$r^{(l)} = S_{12}(S_{22}^l)^{-1}(x_{1+lm}, \cdots, x_{n_x-m}, \hat{x}_{n_x-m+1} - f_{x1}^{(1)}, \cdots, \hat{x}_{n_x} - f_{xm}^{(1)}, \cdots, \hat{x}_{n_x+(l-2)m+1} - f_{x1}^{(l)}, \cdots, \hat{x}_{n_x+(l-1)m} - f_{xm}^{(l)})',$$

$$R^{(l)} = S_{11} - S_{12}(S_{22}^{(l)})^{-1}S_{21}.$$

其中$S_{22}^{(l)} = \begin{bmatrix} \cdots & \cdots \\ \cdots & F_x^{(l)} \end{bmatrix}$, 即$S_{22}$右下角的$m \times m$矩阵被$F_x^{(l)}$代替, 其他值不变所得到的矩阵.

4) 计算第$(l+1) \times m$步的预报.

a) $(Y_{n_y+l}|Y_{1:(n_y+l-1)}, X_{1:(n_x+(l-1)m)}) \sim$

$N(f_y^{(l+1)}, F_y^{(l+1)})$, 其中:

$$F_y^{(l+1)}=[(\sigma_y^2+F_y^{(l)}\phi_y^2)^{-1}+(m^{-2}\overrightarrow{1}'R^{(l)}\overrightarrow{1}+\tau)^{-1}+P_{n_y}^{-1}]^{-1},$$

$$f_y^{(l+1)}=F_y^{(l+1)}[(\sigma_y^2+F_y^{(l)}\phi_y^2)^{-1}\phi_y\hat{y}_{n_y+l-1}+m^{-1}(m^{-2}\overrightarrow{1}'R^{(l)}\overrightarrow{1}+\tau)^{-1}\overrightarrow{1}'r^{(l)}+P_{n_y}^{-1}p_{n_y}];$$

b) $(X_{(n_x+(l-1)m+1):n_x+lm}|Y_{n_y+l}, X_{1:(n_x+(l-1)m)}) \sim N(f_x^{(l+1)}, F_x^{(l+1)})$, 其中:

$$F_x^{(l+1)}=R^{(l)}-m^{-2}R^{(l)}\overrightarrow{1}(m^{-2}\overrightarrow{1}'R^{(l)}\overrightarrow{1}+\tau)^{-1}\overrightarrow{1}'R^{(l)},$$

$$f_x^{(l+1)}=r^{(l)}+m^{-1}R^{(l)}\overrightarrow{1}(m^{-2}\overrightarrow{1}'R^{(l)}\overrightarrow{1}+\tau)^{-1}\cdot(\hat{y}_{n_y+l}-m^{-1}\overrightarrow{1}'r^{(l)}),\ l=0,1,2,\cdots.$$

本文提出的上述预报方法与原有的方法比较,降低了粗水平的预报误差,由下面定理给出.

定理 2 在水平间方差τ充分小的条件下,式(13)的预报方差$F_y^{(1)}$优于式(11)的预报方差F_y,而F_y又优于初始细水平模型AR(1)的预报方差σ_y^2.

证

$$F_y=[\sigma_y^{-2}+(m^{-2}\overrightarrow{1}'R\overrightarrow{1}+\tau)^{-1}+P_{n_y}^{-1}]^{-1}=[\sigma_y^{-2}+a^{-1}+b^{-1}]^{-1}=\sigma_y^2ab/(ab+a\sigma_y^2+b\sigma_y^2)<\sigma_y^2,$$

$$F_y^{(1)}=[\sigma_y^{-2}+(m^{-2}\overrightarrow{1}'R^{(0)}\overrightarrow{1}+\tau)^{-1}+P_{n_y}^{-1}]^{-1}=[\sigma_y^{-2}+(a^*)^{-1}+b^{-1}]^{-1}.$$

根据定理1的结果,方差τ充分小,隐多分辨模型的预报协方差阵$R^{(0)}$优于AR(1)的预报协方差阵R,在对粗水平上,就转化为$a^*<a$. 考虑

$$ab/(ab+a\sigma_y^2+b\sigma_y^2)-a^*b/(a^*b+a^*\sigma_y^2+b\sigma_y^2),$$

$$ab(a^*b+a^*\sigma_y^2+b\sigma_y^2)-a^*b(ab+a\sigma_y^2+b\sigma_y^2)=b^2(a-a^*)\sigma_y^2>0.$$

故$F_y^{(1)}<F_y$. 而$F_y<\sigma_y^2$可由式(11)推出.

证毕.

根据细水平预报的构造过程和整个预报的逐步迭代过程,新方法中细水平预报的精确度也优于用式(13)的预报结果为基础得到的预报.

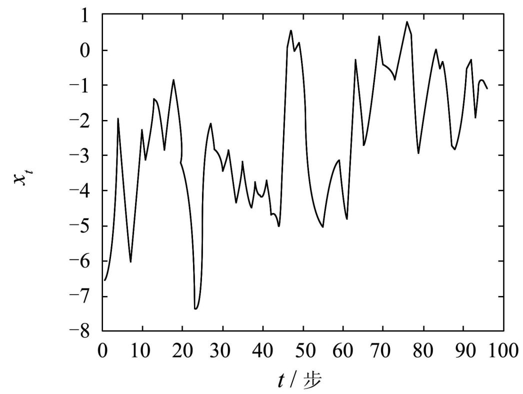
本文提出的预报方法有两个个优点:其一是用比单纯的细水平模型能更好的拟合数据的隐多分辨模型来得到初始预报信息(r, R与$r^{(0)}, R^{(0)}$);其二,由步骤1)和3),新方法不需要初始的细水平信息,可直接推广到细水平模型的各种情形.

5 数值模拟与实例分析(Numerical simulation and example analysis)

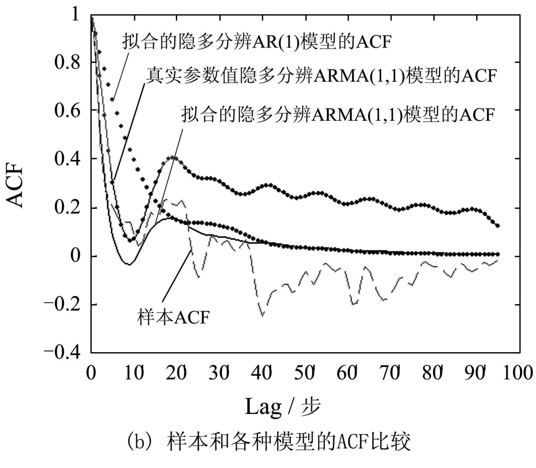
本节首先用模拟产生的数据验证了隐多分辨ARMA模型的建模精度优于隐多分辨AR模型. 对美国季失业率数据分别用隐多分辨ARMA(1,1)模型和隐多分辨AR(1)模型进行建模和预报,进一步验证了研究复杂模型作为基本块的隐多分辨模型的实际意义. 最后以对Huron湖水平面高度数据的隐多分辨AR(1)模型为例,检验了本文提出的预报方法比原有的预报方法优越.

5.1 数值模拟(Numerical simulation)

图3(a)是模拟产生的一组隐多分辨ARMA(1,1)模型数据,各参数真值分别为: $\phi_y=0.9, \sigma_y^2=1, \phi_x=0.8, \theta_x=0.9, \sigma_x^2=1, \lambda=0.01, n_x=96, m=12$.



(a) 模拟序列



(b) 样本和各种模型的ACF比较

图 3 模拟数据

Fig. 3 Simulated data

表1分别为用隐多分辨ARMA(1,1)模型和隐多分辨AR(1)模型拟合数据的参数后验均值和均方误差. 可以看到隐多分辨ARMA(1,1)模型的参数值与真值很接近;对隐多分辨AR(1)模型,虽然只有5个要估计的参数(隐多分辨ARMA(1,1)模型有6个),但参数$\phi_y, \sigma_y^2, \sigma_x^2$ 的均方误差大于隐多分

辨ARMA(1,1)模型中相应参数的均方误差, 尤其是ϕ_y, 分别为61.3391和0.2096, 这表明用隐多分辨AR(1)模型对数据建模是不合适的.

表 1 不同基本块隐多分辨模型的参数估计结果

Table 1 Estimated parameters of hidden multi-resolution model with different building blocks

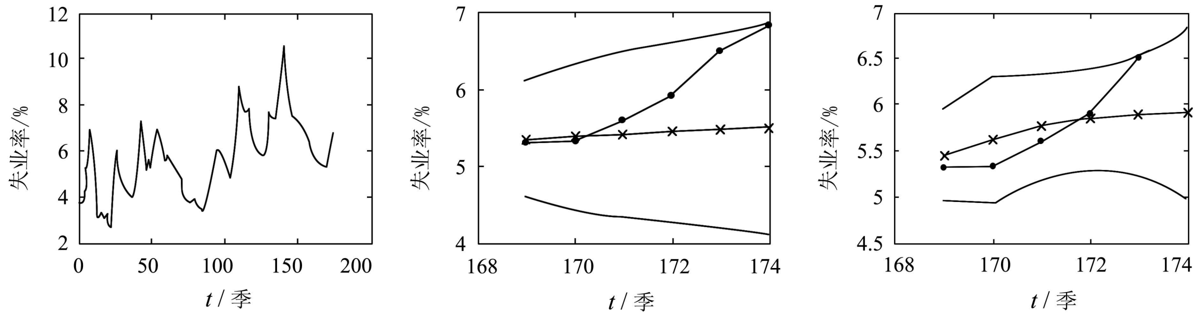
	隐多分辨ARMA(1,1)模型		隐多分辨AR(1)模型	
	均值	均方误差	均值	均方误差
ϕ_y	0.6170	0.0105	0.4836	0.0785
σ_y^2	0.9821	0.2099	4.5348	61.3391
ϕ_x	0.6801	0.0363	0.9250	0.0052
θ_x	0.7978	0.0078	–	–
σ_x^2	0.8188	0.0199	1.3364	0.0396
λ	0.0112	5.6718×10^{-6}	0.0111	5.2887×10^{-6}

图3(b)为两个模型对ACF的拟合结果. 与真实模型的ACF相比, 隐多分辨ARMA (1,1)模型要优于隐多分辨AR(1)模型, 均方误差分别为0.0450和0.0482. 与样本的ACF相比隐多分辨ARMA(1,1)和隐多分辨AR(1)模型的均方误差分别为0.0138和0.0299. 另外, 由定理1, 两个模型所计算的的误差分别为$\hat{\tau}^{\rm AR}=0.0772$和$\hat{\tau}^{\rm ARMA}=4.2877\times10^{-4}$, 真实参数的误差为$\tau=4.5984\times10^{-4}$. 矩阵$(AB)^{\rm AR}$和矩阵$(AB)^{\rm ARMA}$与单位矩阵之差的所有元素绝对值之和分别为0.3689和0.1531, 验证了定理1的结果. 这个例子也表明研究带有更复杂基本块的隐多分辨模型是很必要的.

5.2 实例分析(Example analysis)

5.2.1 美国季失业率数据建模和预报(Modeling and forecast for the quarterly U.S. civilian unemployment rate data)

美国季失业率数据(从1948年到1993年, 共174个观测值[4], 经过去季节项修正)曾被许多研究者用多种不同的模型分析过. 原始数据序列见图4(a), 取粗化窗宽为m=6, 分别用隐多分辨ARMA(1,1)和隐多分辨AR(1)模型对序列前168个数据进行建模, 参数估计结果见表2.



(a) 原始数据序列 (b) 隐多分辨AR(1)模型的预报均值和95%预报区间 (c) 多隐分辨ARMA(1,1)模型的预报均值95%预报区间

——:95%预报区间; —•—:真值; —×—:预报均值

图 4 美国季失业率数据

Fig. 4 Quarterly U.S. civilian unemployment rate data

表 2 不同基本块隐多分辨模型的参数估计结果

Table 2 Estimated parameters of hidden multi-resolution model with different building blocks

	隐多分辨ARMA(1,1)模型		隐多分辨AR(1)模型	
	均值	均方误差	均值	均方误差
ϕ_y	0.6488	2.0934×10^{-6}	0.5454	0.0112
σ_y^2	1.3687	0.0037	2.0078	0.6467
ϕ_x	0.8701	0.0043	0.9044	0.0014
θ_x	0.5515	0.0065	–	–
σ_x^2	0.1017	1.5690×10^{-4}	0.1746	5.1483×10^{-4}
λ	0.0108	4.8043×10^{-6}	0.4664	0.0150

图4(b)和(c)分别给出两个模型用本文提出的预报方法对后6个数据的预报均值和95%预报区间, 预报的均方误差(MSE)和每步的预报方差(即细水平预报分布中正态分布协方差阵的主对角元素)见表3, 其中$\mathrm{MSE}=\frac{1}{m}\sum\limits_{i=1}^{l}(\hat{x}_{T+i}-x_{T+i})^2$, x_{T+i}和$\hat{x}_{T+i}$分别为变量X_{T+i}的真实观测值和预报均值.

表 3 不同模型的预报能力

Table 3 Forecast power of different models

步数	1	2	3	4	5	6	MSE
AR(1)	0.1519	0.2457	0.3101	0.3655	0.4255	0.4970	0.5241
ARMA(1,1)	0.0660	0.1191	0.0952	0.0763	0.1115	0.2216	0.2323

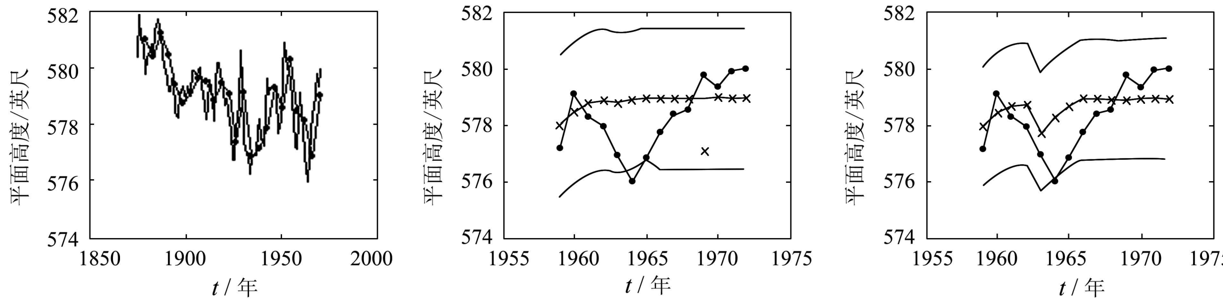
隐多分辨ARMA(1,1)模型整体上预报效果优于隐多分辨AR(1)模型. 与其他非线性模型(门限自回归和Markov转移模型[4])比较, 隐多分辨ARMA(1,1)模型用较少的参数就可以很好的

反映数据的统计特征并进行精确预报.

5.2.2 Huron湖水平面高度数据建模和预报(Modeling and forecast for the level height data of Huron Lake)

用隐多分辨AR(1)模型对Huron湖水平面高度数据(从1875年到1972年, 共98个观测值[5])进行建模和分析. 原数据序列见图5(a). 设细水平为年时间尺度, 同时存在一个粗化窗宽为$m=4$的不可观测的粗水平.

下面用从1875到1958年的84个数据来拟合模型(表4), 分别用原有的方法和本文的方法计算从1959到1972年的14个预报均值和95%的预报区间(图5(c)和(d)), 从图5可以发现, (d)的预报结果要明显优于(c)的. 用线性时间序列对数据建模, 分别用AR(2)模型$x_t=2.0993+1.0159x_{t-1}-0.2462x_{t-2}+\varepsilon_t, \sigma_\varepsilon^2=0.4559$和初始的AR(1)模型$x_t=2.9388+0.6776x_{t-1}+\varepsilon_t$对后14个数据进行预报. 预报结果和比较见表5:



(a) 原始数据序列(实线)和m=4的平均值(虚线)　(b) [2]的方法得到的预报均值和95%预报区间　(c) 本文方法得到的预报均值和95%预报区间

——:95%预报区间; —•—•—:真值; —×—×—:预报均值

图5　Huron湖水平面高度

Fig. 5　Level height data of Huron Lake

表4　隐多分辨AR(1)模型的参数估计

Table 4　Estimated parameters of hidden multi-resolution AR(1) model

	均值	均方误差
ϕ_y	0.4620	–
σ_y^2	0.9010	–
ϕ_x	0.6776	0.0115
σ_x^2	0.3437	0.0038
τ	0.0559	0.0023
λ	0.1042	0.0011

表5　不同预报的均方误差

Table 5　MSE of the forecast based different models

	隐多分辨AR(1) (本文方法)	隐多分辨AR(1) (本文方法)	AR(2)	AR(1)
MSE	1.2866	1.6449	1.6555	1.6991

6　结论(Conclusion)

本文研究一类用于非线性时间序列预报的隐多分辨模型, 对其建模和预报精度进行了定量分析, 证明了所建立模型的ACF与真实模型的ACF之间的误差由水平间的方差决定. 针对隐多分辨ARMA模型, 分析了其ACF结构, 给出了参数估计方法. 提出了一种新的预报方法, 不依赖于初始细水平假定, 而是利用所建立的隐多分辨模型来获得更为合理的初始预报信息, 从而提高了预报的精度, 并证明了新的预报方法比原来方法减少了预报误差. 实例也表明本文的方法可以很好的改进预报精度.

参考文献(References):

[1] 杨叔子, 吴雅. 时间序列分析的工程应用[M]. 武汉: 华中理工大学出版社, 1994.
(YANG Suzi, WU Ya. *Time Series Analysis in Engineering Application*[M]. Wuhan: Huazhong University of Science and Technology Press, 1994.)

[2] FERREIRA M A R. *Bayesian multi-scale modelling*[D]. Duke: Duke University, 2002.

[3] 茆诗松, 王静龙, 濮晓龙. 高等数理统计[M]. 北京: 高等教育出版社; 柏林: Springer-Verlag, 1998: 346 – 347.
(MAO Shisong, WANG Jinglong, PU Xiaolong. *Advanced Mathematical Statistics*[M]. Beijing: China Higher Education Press; Heidelberg Berlin, Springer-Verlag , 1998: 346 – 347.)

[4] TSAY R S. *Analysis of Financial Time Series*[M]. New York: John Wiley & Sons, 2002: 164 – 167.

[5] BROCKWELL P J, DAVIS R A. 时间序列的理论与方法[M]. 田铮, 译. 北京: 高等教育出版社, 2001: 430.
(BROCKWELL P J, DAVIS R A. *Time series: Theory and Methods*[M]. TIAN Zheng, Translation. Beijing: China Higher Education Press, 2001: 430.)

作者简介:

高　伟　(1978—), 女, 博士研究生, 主要研究方向为非线性时间序列分析的理论与应用, E-mail: gaoww525@tom.com;

田　铮　(1948—), 女, 教授, 博士生导师, 主要从事非线性时间序列与信息处理、多尺度随机模型与图像处理等方面的研究, E-mail: Zhtian@nwpu.edcu.cn.