

一类非参数的 ARMA 模型

王少锋, 王海斌, 蔡俊娟

(厦门大学数学科学学院, 福建 厦门 361005)

摘要: 用任意的一元函数代替常数作为线性自回归滑动平均(ARMA)模型中自回归项的系数, 提出并研究一类新的非参数 ARMA 模型. 首先研究该模型的概率性质, 获得了该模型的平稳性条件. 分别用局域线性回归和全局的最小二乘法估计模型中的函数系数和参数, 在函数系数的局域线性估计中, 推广了一个 GCV 准则以选择最优的窗宽. 为了检验特殊的参数化模型是否已经足够描述实际数据的动态结构, 提出了一个 Bootstrap 检验方法. 随机仿真的例子表明本文的估计和检验方法是正确的和可行性的. 进一步, 用该模型成功地分析了一个实际数据集.

关键词: 时间序列; ARMA 模型; 局域线性估计; GCV; Bootstrap

中图分类号: O 211.61

文献标识码: A

文章编号: 0438-0479(2006)05-0628-06

在时间序列分析中, Box 和 Jenkins 提出的线性自回归滑动平均(ARMA)模型^[1]已经形成了相当完善的理论. 然而, 实际中出现的一些非线性现象却不能由该模型得到合理解释. 为了得到更完美的结果, 一些参数化的非线性模型继而被提出, 比较典型的有门限自回归(TAR)模型^[2,3]、指数自回归(EXPAR)模型^[4]和双线性(BL)模型^[5]三种. 问题在于, 与线性函数相对应的非线性函数有无穷多种形式, 要全部研究是不可能的, 事实上也是不必要的. 随着计算能力的提高, 人们更多地采用非参数模型, 即并不假定函数的具体形式, 而由数据自动去寻找. 最近这方面的工作主要有, 函数系数自回归(FAR)模型^[6,7]、自适应函数系数自回归(AAR)模型^[8]等. 当然, 上述模型, 无论是参数的, 还是非参数的, 均是在最经典的 ARMA 模型, 特别是自回归模型的基础上推广得到的.

本文建议一类新的非参数时间序列模型

$$X_t = \sum_{m=1}^p f_m(X_{t-d})X_{t-m} + \varepsilon_t + \sum_{m=1}^q b_m \varepsilon_{t-m} \quad (1)$$

其中 $\{\varepsilon_t\}$ 是均值为0、方差有限的独立同分布序列, 正整数 d 称为延迟参数, $f_1(\cdot), \dots, f_p(\cdot)$ 是任意的一元函数. 把模型(1)称为函数系数自回归滑动平均(FARMA)模型, 记为 FARMA(p, q, d)模型. 事实上, 若 $f_1(\cdot), \dots, f_p(\cdot)$ 均为常数, 则就是 ARMA 模型, 因此, FARMA 模型是 ARMA 模型的直接推广.

FARMA(p, q, d)模型包含了许多已经被研究过的非线性和非参数模型. 若 $q=0$, 则就是函数系数自回归模型, 进一步取特殊的函数形式, 分段常数或指数

函数, 则分别对应于门限自回归模型和指数自回归模型.

1 概率性质

首先讨论 FARMA 模型的概率性质. 对于模型的平稳性条件, 有以下定理.

定理1 对于 FARMA 模型(1), 假定 $|f_j(\cdot)| \leq c_j, j=1, \dots, p, \varepsilon$ 绝对连续且其密度函数在实轴上处处为正. 如果特征函数 $\lambda^p - c_1 \lambda^{p-1} - \dots - c_p = 0$ 所有的根都在单位圆内, 则 FARMA 模型(1)是平稳的.

证明 令 $X_t = (X_t, \dots, X_{t-p+1})^T, \varepsilon = (\varepsilon_t, \dots, \varepsilon_{t-q})^T$,

$$A(X) =$$

$$\begin{bmatrix} f_1(X) & f_2(X) & \dots & f_{p-1}(X) & f_p(X) \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}_{p \times p},$$

$$B = \begin{bmatrix} 1 & b_1 & \dots & b_q \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}_{p \times q},$$

则模型(1)可以表示为

$$X_t = A(X_{t-d})X_{t-1} + B\varepsilon \quad (2)$$

下面只需证明模型(1)的序列 $\{X_t\}$ 是几何遍历. 首先, 由于 ε 绝对连续且其密度函数在实轴上处处为正, $|f_j(\cdot)| \leq c_j, j=1, \dots, p$, 所以 $\{X_t\}$ 是非周期、不可约的. 设矩阵

$$C = \begin{bmatrix} c_1 & c_2 & \cdots & c_{p-1} & c_p \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}_{p \times p}.$$

先给定 C^n 一个上界, 设 I_p 为 p 阶单位阵, 考虑 $X - C$ 的行列式. 注意到

$$|X - C| = |X - c_1 X^{-1} - \cdots - c_p|,$$

因此, 特征函数的根即矩阵 C 的特征根. 设 $\lambda_{\max}$ 为 C 的最大特征根. 由 $\|C^m\|^{1/n} \rightarrow |\lambda_{\max}| < 1$, 知存在一正常数 $\delta < 1$ 和一整数 m , 使得 $\|C^m\| < \delta$.

对于序列 $\{X_t\}$ 的子序列 $\{X_{mt}, t = 1, 2, \dots\}$, 利用递推公式(2), 有

$$X_{m(t+1)} = \prod_{i=0}^{m-1} A(X_{m(t+1)-i-d}) X_{mt} + \sum_{i=1}^m \left[\prod_{j=i}^{m-1} A(X_{m(t+1)-j-1-d}) \right] B \varepsilon_{mt+i} \quad (3)$$

注意到, 对任意向量 $k = (k_1, \dots, k_p)^T$, $(d_1, \dots, d_p)^T = A(X)k$, 均有

$$|d_1| = |f_1(X)k_1 + \cdots + f_p(X)k_p| \leq c_1 |k_1| + \cdots + c_p |k_p|,$$

且 $|d_i| = |k_{i-1}|$, $i = 2, \dots, p$. 因此,

$$\|A(X)k\| \leq \|C\| \|k\|,$$

其中 $\|k\| = (|k_1|, \dots, |k_p|)^T$. 重复对式(3) 应用上述方法, 可以得到

$$\|X_{m(t+1)}\| \leq \|C^m\| \|X_{mt}\| + \left\| \sum_{i=1}^m C^{m-i} B \varepsilon_{mt+i} \right\|,$$

由于

$$\|C^m\| \|X_{mt}\| \leq \delta \|X_{mt}\|,$$

因此

$$E\{\|X_{m(t+1)}\| | X_{mt} = x\} \leq$$

$$\delta \|x\| + E\left\|\sum_{i=1}^m C^{m-i} B \varepsilon_{mt+i}\right\|.$$

注意到第2项是有界, 并且不依赖于 x , 设其界为 D . 则

$$E\{\|X_{m(t+1)}\| | X_{mt} = x\} \leq \delta \|x\| + D,$$

令 $\rho \in (\delta, 1)$ 且设 $M = D(\rho - \delta)^{-1}$. 则对于所有的 $\|x\| \geq M$,

$$E\{\|X_{m(t+1)}\| | X_{mt} = x\} \leq \rho \|x\|.$$

因此, 根据 Tweedie 准则, 对于 $K = \{x: \|x\| \leq M\}$, 序列 $\{X_{mt}\}$ 是几何遍历的, 进而由 Tj stheim 准则, 模型(2) 的序列 $\{X_t\}$ 是几何遍历的. 从而 FARMA 模型(2) 是平稳的. 证毕.

容易看出, FARMA 模型的可逆性条件与 MA 模型的可逆性条件是一样的, 即 $X + b_1 X^{-1} + \cdots + b_q = 0$ 的根都在单位圆内.

2 估计

2.1 参数和非参数的估计

对于 FARMA 模型(1), 需要估计的不仅有参数 $\theta = (b_1, \dots, b_q)^T$ 和 σ^2 , 还有系数函数 $\{f_m(\cdot)\}_{m=1}^p$. 估计通常通过最小化残差平方和来获得. 设 $\{x_t, 1 \leq t \leq N\}$ 为模型(1) 的一系列观察值, 则残差平方和为

$$\sum_{t=p+1}^N \varepsilon_t^2 = \sum_{t=p+1}^N [x_t - \sum_{m=1}^p f_m(x_{t-d}) x_{t-m} - \sum_{m=1}^q b_m \varepsilon_{t-m}]^2 \quad (4)$$

式中 $p' = \max(p, q, d)$, 等号右边的 $\{\varepsilon\}$ 由

$$\varepsilon_t = x_t - \sum_{m=1}^p f_m(x_{t-d}) x_{t-m} - \sum_{m=1}^q b_m \varepsilon_{t-m} \quad (5)$$

递推得到, 初值设为 $x_t = \varepsilon_t = 0, t \leq 0$. 这是一个典型的非线性优化问题. 为叙述方便, 假定(4) 中等号右边的 $\{\varepsilon_{t-m}\}$ 已知. 采用条件估计的方法, 即固定 $\{f_m(\cdot)\}_{m=1}^p$, 全局估计 $\theta = (b_1, \dots, b_q)^T$; 固定 $\theta = (b_1, \dots, b_q)^T$, 局域估计 $\{f_m(\cdot)\}_{m=1}^p$.

(i) 固定 $\{f_m(\cdot)\}_{m=1}^p$, 对 $\theta = (b_1, \dots, b_q)^T$ 进行全局估计.

给定 $\{f_m(\cdot)\}_{m=1}^p$, 残差平方和式(4) 简化为

$$\sum_{t=p+1}^N [y_t^{(0)} - \sum_{m=1}^q b_m \varepsilon_{t-m}]^2,$$

这里 $y_t^{(0)} = x_t - \sum_{m=1}^p f_m(x_{t-d}) x_{t-m}$. 由最小二乘原理, $\theta = (b_1, \dots, b_q)^T$ 的估计为

$$\hat{\theta} = (X_0^T X_0)^{-1} X_0^T Y_0 \quad (6)$$

其中 $Y_0 = [y_{p+1}^{(0)}, \dots, y_N^{(0)}]^T$,

$$X_0 = \begin{bmatrix} \varepsilon_{p+1}' & \cdots & \varepsilon_{p+1-q}' \\ \vdots & \ddots & \vdots \\ \varepsilon_{N-1}' & \cdots & \varepsilon_{N-q}' \end{bmatrix}_{(N-p') \times q}.$$

(ii) 固定 $\theta = (b_1, \dots, b_q)^T, \{f_m(\cdot)\}_{m=1}^p$ 的局域线性估计.

给定 $\theta = (b_1, \dots, b_q)^T$, 采用局域线性回归方法来估计 $\{f_m(\cdot)\}_{m=1}^p$. 假定每一个 $f_m(\cdot)$ 均具有连续的二阶导数, 则对于任意给定的 z_0 , 设 z 为 z_0 邻域内一点, 由 Taylor 展开, 有

$$f_m(z) \approx f_m(z_0) + f_m'(z_0)(z - z_0) \equiv c_m + d_m(z - z_0),$$

据此, 为估计 $f_1(z_0), \dots, f_p(z_0)$, 极小化

$$\sum_{t=p+1}^N [y_t - \sum_{m=1}^p (c_m + d_m(x_{t-d} - z_0)) x_{t-m}]^2 \cdot K_h(x_{t-d} - z_0),$$

式中 $y_t = x_t - \sum_{m=1}^q b_m \varepsilon_{t-m}, K_h(\cdot) = K(\cdot/h)/h$ 为给定

的核函数, $h > 0$ 为给定的窗宽. 设 $\Phi = (c_1, \dots, c_p, d_1, \dots, d_p)^T$, 则由广义最小二乘理论, 得到

$$\Phi(z_0) =$$

$$X(z_0) = \begin{bmatrix} x_{p'} & \dots & x_{p'+1-p} & (x_{p'+1-d} - z_0)x_{p'} & \dots & (x_{p'+1-d} - z_0)x_{p'+1-p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N-1} & \dots & x_{N-p} & (x_{N-d} - z_0)x_{N-1} & \dots & (x_{N-d} - z_0)x_{N-p} \end{bmatrix}_{(N-p') \times 2p},$$

取 $\Phi(z_0)$ 的前 p 个元素 $\hat{c}_1, \dots, \hat{c}_p$ 分别作为 $f_1(\cdot), \dots, f_p(\cdot)$ 在 z_0 的估计值, $z_0 = x_1, \dots, x_{N-p}$.

(iii) 具体算法

由于 $\{\varepsilon_t\}$ 是未知的, 一旦得到 θ 或 $\{f_m(\cdot)\}_{m=1}^p$ 的新的估计值, 均需要重新由式(5)来递推计算, 因此估计模型中的参数和非参数的过程事实上是一个复杂的迭代过程. 具体步骤如下:

1) 设定 θ 和 $\{f_m(\cdot)\}_{m=1}^p$ 的初值 $\theta^{(0)}, \{f_m^{(0)}(\cdot)\}_{m=1}^p$.

2) 假设得到了 θ 和 $\{f_m(\cdot)\}_{m=1}^p$ 的第 i 步迭代值 $\theta^{(i)}, \{f_m^{(i)}(\cdot)\}_{m=1}^p$.

a) 递推计算

$$\varepsilon_t = x_t - \sum_{m=1}^p f_m^{(i)}(x_{t-d})x_{t-m} - \sum_{m=1}^q b_m^{(i)}\varepsilon_{t-m},$$

$$t = 1, \dots, N,$$

初值取 $x_t = \varepsilon_t = 0, t < 1$, 由式(6)计算 $\theta^{(i+1)}$;

b) 重新计算

$$\varepsilon_t = x_t - \sum_{m=1}^p f_m^{(i)}(x_{t-d})x_{t-m} - \sum_{m=1}^q b_m^{(i+1)}\varepsilon_{t-m},$$

$$t = 1, \dots, N,$$

由式(7)得到 $\{f_m(\cdot)\}_{m=1}^p$ 的第 $i+1$ 步估计值 $\{f_m^{(i+1)}(\cdot)\}_{m=1}^p$.

3) 重复第 2) 步直到收敛.

最后, 用

$$\sigma^2 = \sum_{t=p+1}^N \varepsilon_t^2 / (N - p')$$

作为 σ^2 的估计.

2.2 窗宽的选择

在估计系数函数 $f_1(\cdot), \dots, f_p(\cdot)$ 时, 选择一个合适的窗宽 h 是至关重要的, 因为小的窗宽虽然能够使被估的函数值有较小的偏差, 但相应的方差通常比较大; 反之, 大的窗宽虽然有效地减少了被估的函数值的方差, 却造成大的偏差. 用 GCV 准则选择最优的窗宽 h . 具体地, 在由式(7)得到 $f_1(x_{t-d}), \dots, f_p(x_{t-d})$ 后, 用

$$\hat{y}_t = \sum_{m=1}^p f_m(x_{t-d})x_{t-m} =$$

$$(x_{t-1}, \dots, x_{t-p}, 0, \dots, 0)\Phi(x_{t-d})$$

$$[X^T(z_0)w(z_0)X(z_0)]^{-1}X^T(z_0)w(z_0)Y \quad (7)$$

式中 $Y = (y_{p'+1}, \dots, y_N)^T$, $w(z_0) = \text{diag}(K_h(x_{p'+1-d} - z_0), \dots, K_h(x_{N-d} - z_0))$,

来预测 $y_t, t = p' + 1, \dots, N$, 写成矩阵形式, 即 $(\hat{y}_{p'+1}, \dots, \hat{y}_N)^T = H(h)Y$, 这里 $H(h)$ 为一个 $(N - p') \times (N - p')$ 矩阵. 选择使

$$\text{GCV}(h) = \sigma^2 / \{1 - (N - p')^{-1} \text{trace}(H(h))\}^2$$

达到最小值的 h 作为最优的窗宽.

2.3 延迟参数的选择

在实际问题中, 延迟参数 d 通常也是未知的. 当然可以通过给定数据的实际物理背景来选择. 但是, 如果没有任何信息可以利用, 最直接的方法就是利用数据本身获得. 通过比较不同的 d 相对应的最小 GCV 值, 再从中选择最小者, 所对应的 d 作为延迟参数.

3 检验

非参数模型的重要作用之一是用来检验一个参数模型是否已经足够描述给定数据的结构. 相对应地, 当拟合好一个 FARMA 模型后, 自然会问同线性 ARMA 模型或其它特殊的参数非线性模型是否已经足够了昵?这导致下面的假设检验.

$$H_0: f_m(z) = \beta_m(z, \gamma_m), \quad m = 1, \dots, p \quad (8)$$

这里 $\beta_m(z, \gamma_m)$ 为给定的特殊形式的函数, γ_m 为其中的参数向量.

检验统计量取广义似然比

$$T_{N,p'} = \frac{N-p'}{2} \log \left(\frac{\sigma_0^2}{\sigma_1^2} \right) \quad (9)$$

其中 σ_0^2 和 σ_1^2 分别为零假设和 FARMA 模型下 σ^2 的估计, 即

$$\sigma_0^2 = \sum_{t=p+1}^N [x_t - \sum_{m=1}^p \beta_m(x_{t-d}, \gamma_m)x_{t-m} - \sum_{m=1}^q b_m \varepsilon_{t-m}]^2 / (N - p'),$$

$$\sigma_1^2 = \sum_{t=p+1}^N [x_t - \sum_{m=1}^p f_m(x_{t-d})x_{t-m} - \sum_{m=1}^q b_m \varepsilon_{t-m}]^2 / (N - p'),$$

这里 γ_m 为 γ_m 的估计. 若 $T_{N,p'}$ 值比较大时, 则拒绝零假设.

下面用非参数 Bootstrap 方法来获得检验的 p 值:

(i) 从 FARMA 模型(1) 得到的残差 $\{\varepsilon_t\}_{t=1}^N$ 中抽取 Bootstrap 残差 $\{\varepsilon_t^*\}_{t=1}^N$, 由初值 $x_t^* = \varepsilon_t^* = 0, t < 1$ 递推

$$x_t^* = \sum_{m=1}^p \beta_m(x_{t-d}^*, y_m)x_{t-m}^* + \varepsilon_t^* + \sum_{m=1}^q \theta_m \varepsilon_{t-m}^*, \quad t = 1, 2, \dots, N;$$

获得 Bootstrap 样本 $\{x_t^*\}_{t=1}^N$;

(ii) 利用 Bootstrap 样本 $\{x_t^*\}_{t=1}^N$, 分别重新拟合两个模型并计算出各自的残差方差估计 $\hat{\sigma}^{2(*)}$ 和 $\hat{\sigma}^{2(*)'}$. 令

$$T_{N,p}^* = \frac{N-p'}{2} \log \left(\hat{\sigma}^{2(*)} / \hat{\sigma}^{2(*)'} \right).$$

(iii) 重复上面两步 B 次, 用 $\{T_{N,p}^*\}$ 的经验分布作为广义似然比检验统计量(9) $T_{N,p}'$ 的近似分布函数;

(iv) 取 B 次中 $\{T_{N,p}^*\}$ 大于 $T_{N,p}'$ 的频率作为检验的 P 值的估计.

为简单起见, 在估计 $\{T_{N,p}^*\}$ 时, 使用和估计 $T_{N,p}'$ 时一样的窗宽^[9,10].

4 数值演示

在估计过程中, 使用 Biweight 核

$$K(x) = \frac{15}{16}(1-x^2)^2 I(|x| \leq 1).$$

最优窗宽在 $h_j = 1.1^{-j} \times (\text{样本极差的一半}), j = 0, 1, 2, \dots, 20$ 上选. θ 和 $\{f_m(\cdot), 1 \leq m \leq p\}$ 的初始值均取 0. 当残差平方和(RSS) 的相对误差小于 0.5×10^{-4} 或 RSS 在最近的 10 次迭代中都没减小时终止迭代.

4.1 仿 真

首先, 用两个仿真的例子来验证本文所提出方法的有效性和可行性. 用平方误差的平方根(RASE)

$$\text{RASE}_m = \left\{ \sum_{k=1}^n [f_m(z_k) - f_m(z_k)]^2 / n \right\}^{1/2}$$

$$\text{RASE} = \left\{ \sum_{m=1}^p \text{RASE}_m^2 \right\}^{1/2} \quad (10)$$

来衡量 $\{f_m(\cdot)\}$ 的效果, 这里 $\{z_k, k = 1, \dots, n\}$ 为所有离均值的长度小于 1.5 个样本标准差的观察值.

例 1 考虑 FARMA(2, 1) 模型

$$X_t = f_1(X_{t-1})X_{t-1} + f_2(X_{t-1})X_{t-2} + \varepsilon_t - 0.5\varepsilon_{t-1},$$

其中 $\{\varepsilon_t\}$ 独立同服从 $N(0, 0.15^2)$, 两个系数函数分别为 $f_1(u) = 0.138 + (0.316 + 0.982u)e^{-3.89u^2}$ 和 $f_2(u) = -0.437 - (0.659 + 1.260u)e^{-3.89u^2}$.

每次随机产生 400 个观察值, 利用不同的随机种子, 重复 200 次试验. 图 1 概括了模拟结果, 从图 1 中可

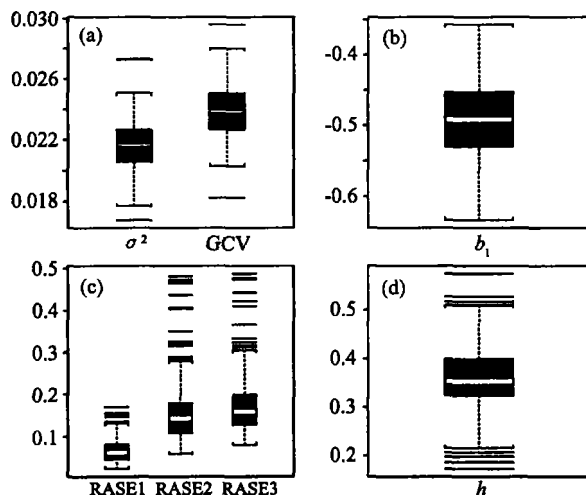


图 1 例 1 的模拟结果

(a) 参数 σ^2 的估计和最小 GCV 的箱形图; (b) 参数 b_1 估计的箱形图; (c) RASE 的箱形图; (d) 所选窗宽 h 的箱形图.

Fig.1 Simulation results for example 1

以发现, 模型的两个参数 b_1 和 σ^2 的估计值与真实值非常接近, 建议的窗宽选取方法运行的也相当稳定.

为进一步说明系数函数的估计效果, 图 2 刻画了真实的曲线和利用一个典型样本得到的估计曲线 $f_1(u), f_2(u)$. 该样本取自 200 次模拟中估计的 RASE 在最中间的一个, 相应的窗宽为 $h_{opt} = 0.323$. 从图 2 中可以看出, 估计的曲线完全描述了真实曲线的非线性特性. 总的来说, 建议的估计方法性能优良.

为了检验第 4 节描述的 Bootstrap 检验方法的功效, 用上面的典型样本来检验假设

$$H_0: \{x_t\} \text{ 来自 } \text{ARMA}(2, 1) \text{ 模型.}$$

利用 ARMA 模型的参数估计方法, 得到:

$$X_t = 0.21X_{t-1} - 0.885X_{t-2} + \varepsilon_t - 0.241\varepsilon_{t-1},$$

相应的 $\hat{\sigma}^2 = 0.04$. 由此得到 $T_{N,p}' = 124.26$. 在 1 000

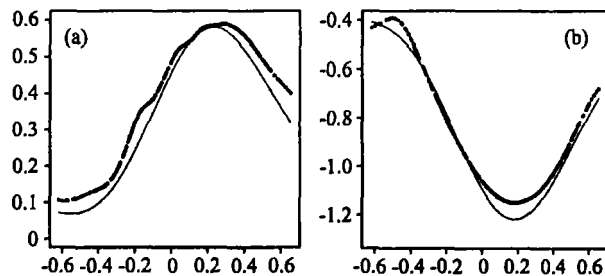


图 2 例 1 中估计的函数(点线) 与实际函数(实线)

$$(a) f_1(u) = 0.138 + (0.316 + 0.982u)e^{-3.89u^2};$$

$$(b) f_2(u) = -0.437 - (0.659 + 1.260u)e^{-3.89u^2}$$

Fig.2 The plots of the local linear estimators (dot curve) and the real functional coefficients(solid curve) for example 1

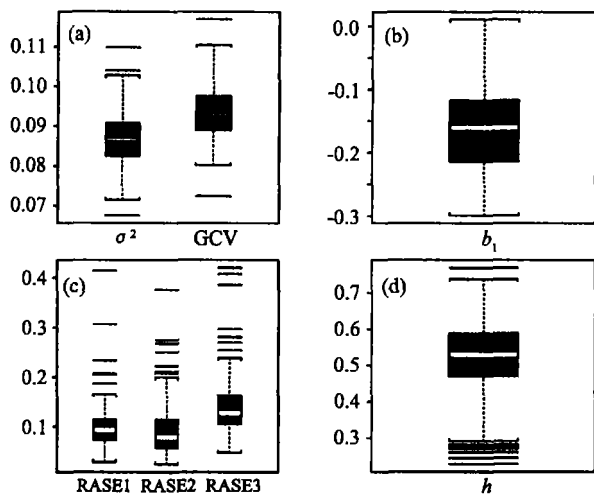


图3 例2的模拟结果

(a) 参数 σ^2 的估计和 200 次模拟中相应的最小 GCV 的箱形图; (b) 参数 b_1 估计的箱形图; (c) RASE 的箱形图; (d) 所选窗宽 h 的箱形图

Fig.3 Simulation results for example 2

个 Bootstrap 样本下, p 值的估计为 0, 图 5 的左图给出了 $\{T_{N,p}^*\}$ 的经验分布. 这提供了强烈的证据, 说明 ARMA(2, 1) 模型是极不合适的, 表明该样本包含了非常强的非线性动态特性.

例 2 对于模型

$$X_t = \frac{\cos(1.5X_{t-2})}{(1 + X_{t-2}^2)} X_{t-1} - \exp(-\frac{1}{2}X_{t-2}^2) X_{t-2} +$$

$$\varepsilon_t - 0.2\varepsilon_{t-1}, \quad \{\varepsilon_t\} \text{ i.i.d} \sim N(0, 0.3^2).$$

这是一个 FARMA(2, 1, 2), 相应的

$$f_1(u) = \frac{\cos(1.5u)}{1 + u^2}, f_2(u) = -\exp(-\frac{u^2}{2}).$$

与例 1 一样, 模拟 200 次, 每次取 400 个样本. 模拟的结果如图 3. 图 4 同样刻画了一个由典型样本得到的 $f_1(u), f_2(u)$ 的估计曲线, 相应的 RASE 在 200 个值中处于最中间, 最优窗宽 $h_{opt} = 0.705$.

对于零假设

$$H_0: \{x_t\} \text{ 来自 ARMA}(2, 1) \text{ 模型},$$

相应的广义似然比统计值 $T_{n,p}^* = 86.34$. 在 1 000 个 Bootstrap 样本下, p 值的估计也为 0. 同样说明了利用线性模型建模是不够的.

4.2 实际数据

南方涛动(The Southern Oscillation) 现象是气候学中一个令人关注的现象, 一些学者希望通过对它的研究来揭示最近几十年来气候变化的规律. 考虑南太平洋的塔希提岛(Tahiti) 和澳大利亚北部海滨城市达尔文(Darwin) 两地从 1950 年到 1995 年每月海平面平均压力之差数据集, 共 540 个观察值.

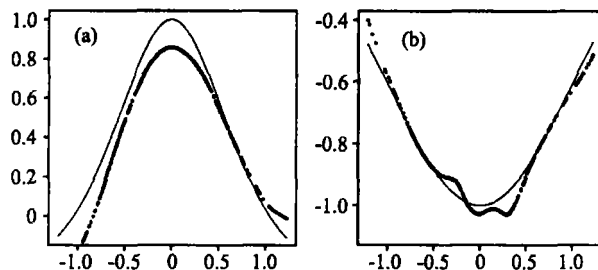


图4 例2中估计的曲线(点线)与实际曲线(实线)

$$(a) f_1(u) = \frac{\cos(1.5u)}{1 + u^2};$$

$$(b) f_2(u) = -\exp(-\frac{1}{2}u^2)$$

Fig.4 The plots of the local linear estimators (dot curve) and the real functional coefficients (solid curve) for example 2

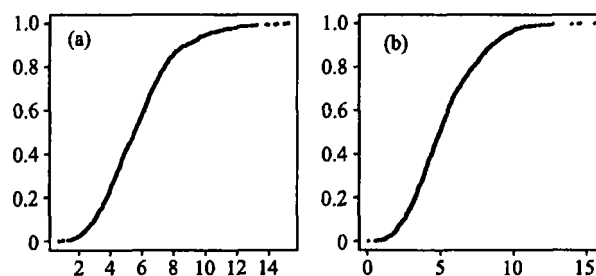


图5 广义似然比统计量 $\{T_{N,p}^*\}$ 的经验分布

(a) 例1; (b) 例2

Fig.5 Empirical distribution function of $\{T_{N,p}^*\}$

表 1 延迟参数 d 为 1 到 5 对应的最小 GCV 的值

Tab.1 Minimum GCV correspond to parameter d from 1 to 5

d	1	2	3	4	5
GCV	1.656	1.627	1.611	1.623	1.644
窗宽	1.338	2.869	2.869	2.371	2.371

利用标准的 ARMA 模型建模技术, 在 Bayes 信息准则下, ARMA(1, 1) 模型

$$X_t = 0.878X_{t-1} + \varepsilon_t - 0.44\varepsilon_{t-1} \quad (11)$$

是最优的, 相应的方差为 1.651. 进一步用 FARMA(1, 1, d) 模型

$$X_t = f_1(X_{t-d})X_{t-1} + \varepsilon_t + b_1\varepsilon_{t-1}$$

来对该数据集进行建模. 限定 $1 \leq d \leq 5$, 从表 1 可以看到, 在 GCV 准则下, 最优的延迟参数应该是 $d = 3$, 由此得到模型

$$X_t = f_1(X_{t-3})X_{t-1} + \varepsilon_t - 0.487\varepsilon_{t-1} \quad (12)$$

其中的系数函数 $f_1(\cdot)$ 见图 6 左半部分, 相应的残差方差 $\sigma^2 = 1.577$, 比 ARMA(1, 1) 模型(11) 的方差减少了大约 4.5%.

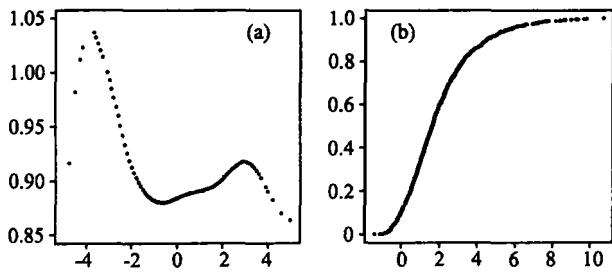


图6 南方涛动现象的建模研究

(a) 估计的系数函数 $f_1(\bullet)$; (b) 广义似然比统计量 $\{T_{N,p}^*\}$ 的经验分布

Fig.6 The research of The Southern Oscillation

不仅要问, ARMA(1, 1) 模型(11) 是否足以描述了南方涛动现象? 由广义似然比统计量(9) 得到 $T_{n,p}' = 12.411$. 在 1 000 个 Bootstrap 样本下, p 值为 0, $\{T_{N,p}^*\}$ 的经验分布见图6右半部分. 因此, 几乎可以肯定地说, ARMA(1, 1) 模型(11) 是不可行的, 并没有完全捕捉到南方涛动现象的非线性特性, 相比之下, FARMA(1, 1, 3) 模型(12) 更合理.

参考文献:

- [1] Box G E P, Jenkins G M. Time Series Analysis, Forecasting and Control[M]. San Francisco: Holden-Day, 1976.
- [2] Tong H. Threshold Models in Nonlinear Time Series Analysis[M]. New York: Springer-Verlag, 1983.

- [3] Tong H. Nonlinear Time Series: A Dynamical Systems Approach[M]. Oxford: Oxford University Press, 1990.
- [4] Haggan V, Ozaki T. Modeling nonlinear vibrations using an amplitude-dependent autoregressive time series model[J]. Biometrika, 1981, 68: 189–196.
- [5] Subba Rao T. On the theory of bilinear time series models[J]. Journal of the Royal Statistical Society, Series B, 1981, 43: 244–255.
- [6] Chen R, Tsay R S. Functional coefficient autoregressive models[J]. Journal of the American Statistical Association, 1993, 88: 298–308.
- [7] Cai Z, Fan J, Yao Q. Functional coefficient regression models for nonlinear time series[J]. Journal of the American Statistical Association, 2000, 95: 941–956.
- [8] Fan J, Yao Q, Cai Z. Adaptive varying coefficient linear model[J]. Journal of the Royal Statistical Society, Series B, 2003, 65: 57–80.
- [9] Wahba G. A survey of some smoothing problems and the method of generalizing cross-validation for solving them[M]// Application of Statistics. Amsterdam: North Holland, 1977: 507–523.
- [10] Craven W S, Wahba G. Smoothing noisy data with spline function: estimating the correct degree of smoothing by the method of generalized cross-validation[J]. Numerische Mathematik, 1979, 31: 377–403.

A Class of Nonparametric ARMA Models

WANG Shao-feng, WANG Hai-bin, CAI Jun-juan

(School of Mathematical Science, Xiamen University, Xiamen 361005, China)

Abstract: A new class of nonparametric autoregressive moving average models, in which arbitrary univariate functions act as the coefficients of autoregressive terms instead of constants, is proposed and discussed. The probabilistic property of the models is investigated and a stationarity condition is derived. It provides global estimate for the parameters and local linear smoother for the functional coefficients in the models respectively. In particular, the optimal bandwidth is selected via a modified generalized cross-validation criterion. A bootstrap test is applied to test whether the functional coefficients are some specified parametric forms. The feasibility and validity of the proposed methods are justified by simulated examples. A real data set is analyzed by the proposed models.

Key words: time series; ARMA model; Local linear estimate; Generalized Cross-Validation; Bootstrap