

文章编号: 1000-5641(2014)04-0018-08

一类由布朗运动驱动的滑动平均的参数矩估计

何伟平¹, 张世斌²

(1. 内蒙古大学 数学系, 呼和浩特 010021; 2. 上海海事大学 数学系, 上海 201306)

摘要: 研究一类自相关结构具有周期性和递减性共存性质的由布朗运动驱动的滑动平均的参数矩估计. 通过研究模型参数与自协方差函数间的联系, 构造了参数的矩估计量. 借助滑动平均离散抽样过程谱密度的研究, 分析了其强混合系数的特点, 进而证明了该矩估计量的相合性和渐近正态性. 模拟显示估计量在小样本场合下也呈现良好的估计效果. 实例分析表明该模型可用于刻画船体应力随时间的变化情况.

关键词: 滑动平均; 相合性; 渐近正态性; 谱密度; α -混合

中图分类号: O212; O211.64 **文献标识码:** A

DOI: 10.3969/j.issn.1000-5641.2014.04.003

Moment estimation for a class of moving averages driven by Brownian motions

HE Wei-ping¹, ZHANG Shi-bin²

(1. *Department of Mathematics, Inner Mongolia University, Hohhot 010021, China;*
2. *Department of Mathematics, Shanghai Maritime University, Shanghai 201306, China*)

Abstract: This paper was concerned with moment estimation for a class of moving averages driven by Brownian motions, whose autocorrelation functions take on both periodic and regressive properties. Based on investigation of the relationship between the parameter and auto-covariance function, the moment estimator of the parameter was constructed. By analyzing the spectral density of the discretely sampled trajectory, the properties of the mixing coefficient were captured. Further, the consistency and asymptotic normality of the estimator were proved. A simulation study showed that the estimator achieves good performance even in the small-sample circumstances. The real data analysis evidences that the model can be used to depict the ship hull stress data.

Key words: moving average; consistency; asymptotic normality; spectral density; α -mixing

收稿日期: 2013-09

基金项目: 上海市自然科学基金(13ZR1419100); 上海市教委科研创新项目(14YZ115)

第一作者: 何伟平, 女, 硕士研究生, 研究方向为概率论与数理统计. E-mail: 524540664@qq.com.

通信作者: 张世斌, 男, 副教授, 硕士生导师, 研究方向为概率论与数理统计.

E-mail: sbzhang@shmtu.edu.cn.

0 引言

设 $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, P)$ 是一概率空间, $\{B_1(s), s \geq 0\}$, $\{B_2(s), s \geq 0\}$ 是两个定义在该概率空间上的相互独立的标准布朗运动. 通过 $B(s) = B_1(s)\mathbf{1}_{\{s \geq 0\}} + B_2(-s)\mathbf{1}_{\{s < 0\}}$ 定义双边布朗运动 $\{B(s), s \in \mathbb{R}\}$.

本文讨论在 \mathbb{R} 上定义的一类由布朗运动驱动的滑动平均

$$X_t = \sigma \int_{-\infty}^t \cos(\beta(t-s))e^{-\alpha(t-s)} dB(s), \quad t \in \mathbb{R}, \quad (0.1)$$

其中参数 $\alpha > 0, \beta > 0$ 且 $\sigma \neq 0$. 从文献 [1,2] 可知, 对任意的正整数 n 及 $t_1 < t_2 < \cdots < t_n \in \mathbb{R}$, $(X_{t_1}, X_{t_2}, \cdots, X_{t_n}) \sim N(\mathbf{0}, \Sigma)$, 其中 $\Sigma = (\Sigma_{kj})_{n \times n}$, 且

$$\begin{aligned} \Sigma_{kj} &= \text{Cov}[X_{t_k}, X_{t_j}] \\ &= \sigma^2 \left(\frac{2\alpha^2 + \beta^2}{4\alpha(\alpha^2 + \beta^2)} \cos(\beta|t_k - t_j|) - \frac{\beta}{4(\alpha^2 + \beta^2)} \sin(\beta|t_k - t_j|) \right) e^{-\alpha|t_k - t_j|}. \end{aligned}$$

进而, X_t 的自相关函数为

$$\text{Corr}[X_t, X_{t+h}] = \rho_h = \left(\cos(\beta|h|) - \frac{\beta\alpha}{2\alpha^2 + \beta^2} \sin(\beta|h|) \right) e^{-\alpha|h|}, \quad t, h \in \mathbb{R}.$$

既然过程 (0.1) 的任意有限维分布是高斯分布, 那么它应属于高斯滑动平均过程, 并且其自相关函数呈现周期性和递减性并存的特点 (如图 1). 图 1 给出了当 $\alpha = 0.2$, $\beta = 1$ 和 $\sigma = 1$ 时, 模拟的一条长度为 2000 的样本轨道 (左图) 及其该样本的经验自相关函数 (实线) 的拟合图. 自相关函数具有该特点的随机过程可用于刻画船体应力随时间的变化情况 (见文献 [3]). 近来, 高斯滑动平均无论在理论还是应用领域都广受关注. 如文献 [4,5] 关心它们的半鞅性; 具有分形性质的高斯滑动平均的研究综述可参见文献 [6]; 关于高斯滑动平均及其推广过程幂变差的理论和应用的研究参见文献 [7] 等.

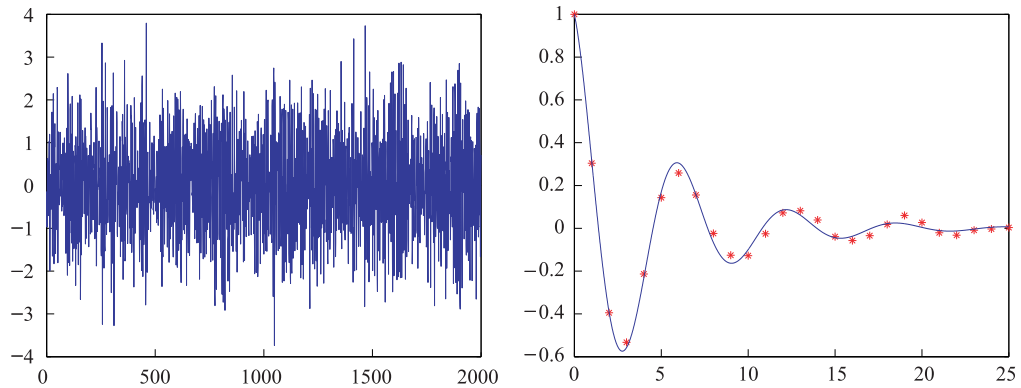


图 1 $\alpha=0.2, \beta=1$ 和 $\sigma=1$ 时, 长度为 2000 的样本轨道 (左图) 及理论和经验自相关函数的拟合图 (右图)

Fig. 1 A discretely sampled trajectory of length 2000 (left figure), fitting the theoretical auto-correlation function to the empirical auto-correlation function (right figure); $\alpha=0.2$, $\beta=1$ and $\sigma=1$

既然过程 (0.1) 的任意有限维分布是高斯分布, 在理论上对其进行最大似然估计是可行的. 但事实上, 过程 (0.1) 不是 Markov 过程, 其离散抽样的样本的似然函数就是它的有限维分布密度函数. 在最大化似然函数时, 需要涉及到有限维分布协方差阵的求逆和求行列式计算. 但是其计算复杂度极高, 几乎不可能实现. 故本文通过研究模型参数与自协方差函数间的联系, 用样本自协方差矩近似代替模型自协方差矩, 从而实现对模型参数的估计. 参数估计量的渐近性质则通过借助滑动平均离散抽样过程的谱密度, 分析其强混合系数的特点, 进而利用相依序列的遍历定理和中心极限定理对其加以证明.

本文以下部分的结构如下. 第 1 节给出了模型参数的估计方法并证明了参数估计量的相合性和渐近正态性. 第 2 节介绍了模型的模拟算法并借助模拟算法研究了参数估计量的小样本性质. 第 3 节是利用估计方法对船体应力数据做了实证研究.

1 参数矩估计及其渐近性质

假设从 $t = 0$ 开始抽样, 间隔为 1, 即观测样本为 X_0, X_1, \dots, X_{n+2} , 记 $\mathbf{X} = (X_0, X_1, \dots, X_{n+2})$. 根据 (0.2) 式可知, 滞后阶数不同过程的自协方差函数值的形式相同. 对于固定的样本观测, 一般来讲滞后阶数较低的自协方差函数值的信息更容易从样本中提取. 故在进行矩估计时, 优先选择滞后阶数较低的自协方差函数值. 因 X_t 的期望函数为 0, 故可选择样本方差, 滞后 1 阶和滞后 2 阶的自协方差函数值. 记 $\boldsymbol{\theta} = (\alpha, \beta, \sigma)^T$ 表示待估参数, $\boldsymbol{\theta}_0 = (\alpha_0, \beta_0, \sigma_0)^T$ 代表其真值. 设 $\mathbf{g}(\boldsymbol{\theta}) = (g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}), g_3(\boldsymbol{\theta}))^T$, 其中

$$\begin{aligned} g_1(\boldsymbol{\theta}) &= E[X_0^2] = \frac{\sigma^2(2\alpha^2 + \beta^2)}{4\alpha(\alpha^2 + \beta^2)}, \\ g_2(\boldsymbol{\theta}) &= E[X_0X_1] = \sigma^2 \left(\frac{2\alpha^2 + \beta^2}{4\alpha(\alpha^2 + \beta^2)} \cos \beta - \frac{\beta}{4(\alpha^2 + \beta^2)} \sin \beta \right) e^{-\alpha}, \\ g_3(\boldsymbol{\theta}) &= E[X_0X_2] = \sigma^2 \left(\frac{2\alpha^2 + \beta^2}{4\alpha(\alpha^2 + \beta^2)} \cos(2\beta) - \frac{\beta}{4(\alpha^2 + \beta^2)} \sin(2\beta) \right) e^{-2\alpha}. \end{aligned}$$

用样本矩代替总体矩, 得到方程组

$$\begin{cases} g_1(\boldsymbol{\theta}) = \frac{1}{n} \sum_{i=1}^n X_i^2, \\ g_2(\boldsymbol{\theta}) = \frac{1}{n} \sum_{i=1}^n X_i X_{i+1}, \\ g_3(\boldsymbol{\theta}) = \frac{1}{n} \sum_{i=1}^n X_i X_{i+2}. \end{cases} \quad (1.1)$$

参数 $\boldsymbol{\theta}$ 的矩估计量可由方程组 (1.1) 解出, 记为 $\hat{\boldsymbol{\theta}}_n = \hat{\boldsymbol{\theta}}_n(X_0, X_2, \dots, X_{n+2})$.

注 1.1 由于方程组 (1.1) 是一非线性方程组, 故对其用 Newton-Raphson 法进行求解 (参见文献 [8]). 因 Newton-Raphson 法是一种迭代算法, 需确定 (α, β) 迭代初值为

$$(\tilde{\alpha}, \tilde{\beta}) = \arg \min_{(\alpha, \beta)} \sum_{k=1}^M (\rho_k - \hat{\rho}_k)^2, \quad (1.2)$$

其中 M 为某一正整数且 $M < n$, $\hat{\rho}_k$ 表示滞后 k 阶的经验自相关函数值. 再将 $(\tilde{\alpha}, \tilde{\beta})$ 代入式 (1.1) 中的第一式, 得到 $\tilde{\alpha}$; 然后选取 $(\tilde{\alpha}, \tilde{\beta}, \tilde{\alpha})$ 作为迭代初值.

下面引入强混合 (参见文献 [9]) 和谱密度 (参见文献 [10]) 的定义.

定义 1.1 平稳过程 $\{X_t, t \in \mathbb{R}\}$ 称为强混合的, 如果

$$\alpha(k) = \sup_{S \in \mathcal{F}_{-\infty}^0, T \in \mathcal{F}_k^\infty} |P(S \cap T) - P(S)P(T)| \rightarrow 0, \quad k \rightarrow \infty;$$

其中, $\mathcal{F}_{-\infty}^0$ 和 \mathcal{F}_k^∞ 分别表示由 $\{X_t, t \leq 0\}$ 和 $\{X_t, t \geq k\}$ 生成的 σ -代数.

定义 1.2 若 $\sum_n |E[X_0 X_k]| < \infty$, 则平稳序列 $\{X_k, k \in \mathbb{Z}\}$ 的谱密度定义为

$$f(\lambda) = \sum_{k=-\infty}^{\infty} E[X_0 X_k] e^{ik\lambda}, \quad \lambda \in [-\pi, \pi].$$

为了得到矩估计量的渐近结论, 需对 X_t 的参数附加如下条件.

条件 1.1 $\{X_t, t \in \mathbb{R}\}$ 的参数 $\varepsilon \leq \alpha \leq M_1, \beta \leq M_2, \delta \leq |\sigma| \leq M_3$, 其中, $\varepsilon, \delta, M_1, M_2, M_3$ 均为正数, 且 $\varepsilon \leq M_1, \delta \leq M_3$.

引理 1.1 若条件 1.1 成立, 则序列 $\{X_k, k \in \mathbb{Z}\}$ 的谱密度满足

$$f(\lambda) \geq a, \quad \lambda \in [-\pi, \pi],$$

其中 $a > 0$ 为一常数.

证 明 根据文献 [11] 中第 51 页的三角序列的结论, 经计算可得

$$\begin{aligned} \sum_{k=1}^{\infty} \sin(k\beta) \cos(k\lambda) e^{-\alpha k} &= \frac{1}{4} \left(\frac{\sin(\beta + \lambda)}{\cosh \alpha - \cos(\beta + \lambda)} + \frac{\sin(\beta - \lambda)}{\cosh \alpha - \cos(\beta - \lambda)} \right), \\ \sum_{k=1}^{\infty} \cos(k\beta) \cos(k\lambda) e^{-\alpha k} &= \frac{1}{4} \left(\frac{\sinh \alpha}{\cosh \alpha - \cos(\beta + \lambda)} + \frac{\sinh \alpha}{\cosh \alpha - \cos(\beta - \lambda)} - 2 \right). \end{aligned}$$

因为

$$(\cosh \alpha - \cos \beta \cos \lambda) - (\cosh \alpha \cos \lambda - \cos \beta) = (\cosh \alpha + \cos \beta)(1 - \cos \lambda) \geq 0,$$

$$(\cosh \alpha - \cos \beta \cos \lambda) + (\cosh \alpha \cos \lambda - \cos \beta) = (\cosh \alpha - \cos \beta)(1 + \cos \lambda) \geq 0,$$

故 $|\cosh \alpha \cos \lambda - \cos \beta| \leq \cosh \alpha - \cos \beta \cos \lambda$. 从而谱密度

$$\begin{aligned} f(\lambda) &= \sum_{k=-\infty}^{\infty} E[X_0 X_k] e^{ik\lambda} = \sum_{k=-\infty}^{\infty} \sigma^2 \left(\frac{2\alpha^2 + \beta^2}{4\alpha(\alpha^2 + \beta^2)} \cos(\beta|k|) - \frac{\beta}{4(\alpha^2 + \beta^2)} \sin(\beta|k|) \right) e^{-\alpha|k|} e^{ik\lambda} \\ &= \sigma^2 \frac{(2\alpha^2 + \beta^2) \sinh \alpha (\cosh \alpha - \cos \beta \cos \lambda) - \alpha \beta \sin \beta (\cosh \alpha \cos \lambda - \cos \beta)}{4\alpha(\alpha^2 + \beta^2) (\cosh \alpha - \cos(\beta + \lambda)) (\cosh \alpha - \cos(\beta - \lambda))} \\ &\geq \sigma^2 \frac{((2\alpha^2 + \beta^2) \sinh \alpha - |\alpha \beta \sin \beta|) (\cosh \alpha - \cos \beta \cos \lambda)}{4\alpha(\alpha^2 + \beta^2) (\cosh \alpha - \cos(\beta + \lambda)) (\cosh \alpha - \cos(\beta - \lambda))} \\ &\geq \frac{2\sigma^2 \alpha^3 (\cosh \alpha - 1)}{4\alpha(\alpha^2 + \beta^2) (\cosh \alpha + 1)^2} \geq \frac{2\delta^2 \varepsilon^3 (\cosh \varepsilon - 1)}{4M_1(M_1^2 + M_2^2) (\cosh M_1 + 1)^2} =: a. \end{aligned}$$

证毕.

引理 1.2 若条件 1.1 成立, 则序列 $\{X_k, k \in \mathbb{Z}\}$ 是强混合的, 且强混合系数 $\alpha(k) \leq b e^{-\varepsilon k}$. 其中 $b > 0$ 为一常数.

证明 记 $\Delta_k(f) = \inf\{\|f - P\|_\infty\}$, 其中 $\|\cdot\|_\infty$ 表示上确界范数, $f(\lambda)$ 是 X_k 的谱密度, $P(\lambda) = \sum_{|t|<k} c_t e^{i\lambda t}$. 即 $\Delta_k(f)$ 表示 $k-1$ 阶三角多项式对 f 的最优一致近似. 取 $c_t = E[X_0 X_t]$, 可得

$$\Delta_k(f) \leq \left| \sum_{|t| \geq k} E[X_0 X_t] e^{it\lambda} \right| \leq \sigma^2 \frac{2\alpha^2 + \beta^2 + \alpha\beta}{2\alpha(\alpha^2 + \beta^2)} \frac{e^{-\alpha k}}{1 - e^{-\alpha}} \leq \frac{3\sigma^2 e^{-\alpha k}}{2\alpha(1 - e^{-\alpha})} \leq \frac{3M_3^2 e^{-\varepsilon k}}{2\varepsilon(1 - e^{-\varepsilon})}.$$

令 $\eta = \frac{3M_3^2}{2\varepsilon(1 - e^{-\varepsilon})}$, 再根据引理 1.1 及文 [10] 第 58 页中推论 1, 强混合系数

$$\alpha(k) \leq \frac{1}{a} \Delta_k(f) \leq \frac{\eta}{a} e^{-\varepsilon k} =: b e^{-\varepsilon k}.$$

显然, 当 $k \rightarrow \infty$ 时, $\alpha(k) \rightarrow 0$.

定理 1.1 若条件 1.1 成立, 且方程组 (1.1) 在 θ_0 的一邻域 $U(\theta_0)$ 内存在唯一的解 $\hat{\theta}_n$. 则 $n \rightarrow \infty$ 时, $\hat{\theta}_n$ 依概率收敛于 θ_0 , 即 $\hat{\theta}_n \xrightarrow{P} \theta_0$.

证明 由引理 1.2 知, $\{X_k, k \in \mathbb{Z}\}$ 是强混合的. 再根据文献 [12], 该序列具有遍历性. 故当 $n \rightarrow \infty$ 时,

$$g_1(\hat{\theta}_n) \xrightarrow{P} g_1(\theta), \quad g_2(\hat{\theta}_n) \xrightarrow{P} g_2(\theta), \quad g_3(\hat{\theta}_n) \xrightarrow{P} g_3(\theta).$$

由 $g_1(\theta), g_2(\theta)$ 及 $g_3(\theta)$ 的连续性, 若 $U(\theta_0)$ 内存在唯一的 (1.1) 的解 $\hat{\theta}_n$, 则 $\hat{\theta}_n \xrightarrow{P} \theta_0$.

定理 1.1 若定理 1.1 的条件满足, 则当 $n \rightarrow \infty$ 时,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{L} N_3(\mathbf{0}, D^{-1} B (D^T)^{-1}).$$

其中 \xrightarrow{L} 表示依分布收敛, $B = E[\mathbf{Y}_0 \mathbf{Y}_0^T] + \sum_{k=1}^{\infty} (E[\mathbf{Y}_0 \mathbf{Y}_k^T] + E[\mathbf{Y}_k \mathbf{Y}_0^T])$, $\mathbf{Y}_k = (X_k^2 - E[X_k^2], X_k X_{k+1} - E[X_0 X_1], X_k X_{k+2} - E[X_0 X_2])^T$,

$$D = \begin{pmatrix} \frac{\partial g_1}{\partial \alpha}(\theta_0) & \frac{\partial g_1}{\partial \beta}(\theta_0) & \frac{\partial g_1}{\partial \sigma}(\theta_0) \\ \frac{\partial g_2}{\partial \alpha}(\theta_0) & \frac{\partial g_2}{\partial \beta}(\theta_0) & \frac{\partial g_2}{\partial \sigma}(\theta_0) \\ \frac{\partial g_3}{\partial \alpha}(\theta_0) & \frac{\partial g_3}{\partial \beta}(\theta_0) & \frac{\partial g_3}{\partial \sigma}(\theta_0) \end{pmatrix}.$$

证明 由引理 1.2 知, 强混合系数 $\alpha(k) \leq b e^{-\varepsilon k}$, 故对某一正数 ξ , $\sum_{k=1}^{\infty} \alpha(k)^{\xi/(2+\xi)} < \infty$ 成立. 再根据文献 [9] 中定理 18.5.3 及文献 [13] 中引理 1, 当 $n \rightarrow \infty$ 时, $\sqrt{n} \cdot \frac{1}{n} \sum_{k=1}^n \mathbf{Y}_k \xrightarrow{L} N_3(\mathbf{0}, B)$, 即

$$\sqrt{n}(g(\hat{\theta}_n) - g(\theta_0)) \xrightarrow{L} N_3(\mathbf{0}, B), \quad n \rightarrow \infty.$$

最后, 再利用文献 [14] 中定理 4.3, 结论成立.

2 模拟实验

下面用文献 [15] 中的模拟方法来产生 $\{X_k, k \in \mathbb{Z}\}$ 的样本轨道.

假设要产生随机向量 $\mathbf{X} = (X_0, X_1, \dots, X_{n-1})^T$, 可由以下两步来完成.

步骤1: 将 \mathbf{X} 的自协方差矩阵 \mathbf{R} 嵌入到一个 m 阶循环矩阵 \mathbf{C} , 其中 $m = 2^j$ (j 为某一整数), 且 $m \geq 2(n-1)$.

步骤2: 利用快速傅里叶变换, 产生一个随机向量 $\mathbf{W} = (W_0, W_1, \dots, W_{m-1})^T \sim N_m(\mathbf{0}, \mathbf{C})$, 那么 $(W_0, W_1, \dots, W_{n-1})^T \sim N_n(\mathbf{0}, \mathbf{R})$.

步骤1中的自协方差矩阵 \mathbf{R} 和循环矩阵 \mathbf{C} 分别为

$$\mathbf{R} = \begin{pmatrix} r_0 & r_1 & \cdots & r_{n-1} \\ r_1 & r_0 & \cdots & r_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n-1} & r_{n-2} & \cdots & r_0 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} c_0 & c_1 & \cdots & c_{m-1} \\ c_{m-1} & c_0 & \cdots & c_{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{pmatrix}.$$

其中 $r_k = \text{Cov}[X_0, X_k]$ ($k = 0, 1, \dots, n-1$), $c_j = \begin{cases} r_j, & 0 \leq j \leq \frac{m}{2}, \\ r_{m-j}, & \frac{m}{2} < j \leq m-1. \end{cases}$ 可以看出, 矩阵 \mathbf{C} 的左上角的 n 阶矩阵是 \mathbf{X} 的自协方差矩阵.

假设 \mathbf{C} 是非负定的, 则步骤2的具体操作过程如下. 循环矩阵 \mathbf{C} 可以被分解成 $\mathbf{C} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^*$ 的形式, 其中 $\mathbf{\Lambda}$ 是由 \mathbf{C} 的特征值构成的对角阵, 酉矩阵 \mathbf{Q} 的形式为

$$(\mathbf{Q})_{jk} = \frac{1}{\sqrt{m}} \exp\left(-2\pi i \frac{jk}{m}\right), \quad j, k = 0, 1, \dots, m-1.$$

\mathbf{Q}^* 表示 \mathbf{Q} 的转置的复共轭矩阵. \mathbf{C} 的特征值可表示为

$$\lambda_k = \sum_{j=0}^{m-1} c_j \exp\left(-2\pi i \frac{jk}{m}\right), \quad k = 0, 1, \dots, m-1.$$

令 $\mathbf{W} = \mathbf{Q}\mathbf{\Lambda}^{1/2}\mathbf{Q}^*\mathbf{V}$, 其中 $\mathbf{\Lambda}^{1/2} = \text{diag}\{\lambda_0^{1/2}, \lambda_1^{1/2}, \dots, \lambda_{m-1}^{1/2}\}$, \mathbf{V} 是由 m 个相互独立的 $N(0, 1)$ 随机变量构成的向量. 因 $\mathbf{Q}\mathbf{Q}^* = \mathbf{I}$, 故 $\mathbf{W} \sim N_m(\mathbf{0}, \mathbf{C})$. 其中 \mathbf{I} 表示单位阵. 从而 $\mathbf{W} = (W_0, W_1, \dots, W_{n-1})^T$ 和 \mathbf{X} 有相同的分布.

注2.1 循环矩阵 \mathbf{C} 不一定是非负定矩阵, 在条件1.1成立时, 可通过增加 m 以实现 \mathbf{C} 的正定化 (参见文献[15]命题2).

理论上产生 $\{X_t, t \in \mathbb{R}\}$ 离散抽样轨道的这种方法是精确的, 并且该方法的主要优点是其计算速度. 即样本容量为 n 时, 其计算复杂度为 $O(n \log n)$ (见文献[16]). 图1给出了当 $\alpha = 0.2$, $\beta = 1$ 和 $\sigma = 1$ 时, 用该方法模拟的一条长度为2000的样本轨道 (左图), 及其该样本的经验自相关函数和理论自相关函数的拟合图 (右图). 从图1可见, 经验自相关函数与理论自相关函数吻合的较好.

表1给出了当样本容量 n 从100增加到2000时, 400个矩估计值的均值和均方误差 (括号中的值). 表1显示参数矩估计量在小样本情形下效果良好. 对于固定的 α , β 和 σ , 随着样本容量的增大, 估计的精确度越来越高, 且越来越稳定. 这与估计值的渐近性质相吻合.

3 实证研究

下面讨论滑动平均过程 $\{X_t, t \in \mathbb{R}\}$ 在实际数据中的应用. 分析数据来自于安装在一艘名为“育锋”轮的集装箱船上的船体强度检测系统. 这艘船, 载重量为10,000t, 总长132m, 常年航行于中国东南海域. 采集的应力数据 (单位: MPa) 是2009年12月13日下

午 15:00 到 18:00 的右舷船舳处的一测点的连续 5 401 个应力值, 抽样时间间隔为 2s. 在进行参数估计时, 首先采用将原始数据 \mathbf{X} 同时减去其平均值 39.903 3 的方式进行预处理得 \mathbf{X}' . 这里所分析的样本数据及其预处理方式同文献 [3].

表 1 样本容量 n 从 100 到 2000 时, 对不同真值的参数估计

Tab. 1 Parameter estimation of true values, sample size n from 100 to 2000				
真值	$n=100$	$n=500$	$n=1\ 000$	$n=2\ 000$
$\alpha=0.2$	0.213 1(0.047 7)	0.202 7(0.023 0)	0.200 3(0.015 7)	0.200 1(0.011 2)
$\beta=1.0$	1.000 5(0.070 4)	1.003 0(0.033 9)	0.999 8(0.020 6)	0.999 8(0.032 0)
$\sigma=1.0$	1.011 9(0.072 5)	1.003 7(0.081 2)	1.002 1(0.065 7)	0.999 4(0.014 1)
$\alpha=0.2$	0.210 8(0.043 9)	0.201 0(0.021 7)	0.200 9(0.014 6)	0.200 2(0.009 5)
$\beta=2.0$	1.999 7(0.054 7)	1.997 8(0.026 0)	1.999 7(0.018 4)	1.999 9(0.007 4)
$\sigma=1.0$	1.020 9(0.084 1)	0.985 8(0.059 5)	1.006 1(0.031 1)	1.001 2(0.012 5)
$\alpha=0.5$	0.521 7(0.118 6)	0.503 3(0.052 8)	0.501 7(0.035 2)	0.500 1(0.030 1)
$\beta=1.0$	0.996 1(0.120 8)	1.001 9(0.065 6)	1.001 4(0.044 0)	1.000 0(0.039 6)
$\sigma=1.0$	1.024 1(0.042 0)	1.004 1(0.036 7)	1.008 4(0.028 3)	1.000 3(0.018 0)
$\alpha=0.5$	0.517 0(0.088 5)	0.501 2(0.043 0)	0.500 9(0.029 2)	0.499 9(0.029 4)
$\beta=2.0$	2.012 9(0.125 5)	1.999 8(0.059 6)	2.000 7(0.040 5)	2.000 2(0.017 0)
$\sigma=1.0$	0.990 0(0.135 5)	1.033 3(0.053 8)	0.998 6(0.032 1)	0.999 1(0.023 3)

取式 (1.2) 中的 $M = 10$, 用最小二乘法先求得 $(\tilde{\alpha}, \tilde{\beta})$, 再将 $(\tilde{\alpha}, \tilde{\beta})$ 代入式 (1.1) 中的第一式得到 $\tilde{\sigma}$. 选取 $(\tilde{\alpha}, \tilde{\beta}, \tilde{\sigma})$ 作为迭代初值, 用本文的矩估计法对数据 \mathbf{X}' 进行分析, 可求得模型 (0.1) 的参数估计值 $\hat{\alpha} = 0.556\ 3$, $\hat{\beta} = 1.651\ 2$ 和 $\hat{\gamma} = 3.327\ 3$. 图 2 给出了 $\alpha = 0.556\ 3$, $\beta = 1.651\ 2$ 和 $\sigma = 3.327\ 3$ 时, $\{X_t, t \in \mathbb{R}\}$ 的自相关函数(实线)及船体应力数据的经验自相关函数的拟合图, 由图可见, 拟合效果良好. 因样本数据的平稳分布近似为正态分布, 故该实例分析表明该模型可用于刻画船体应力随时间的变化情况.

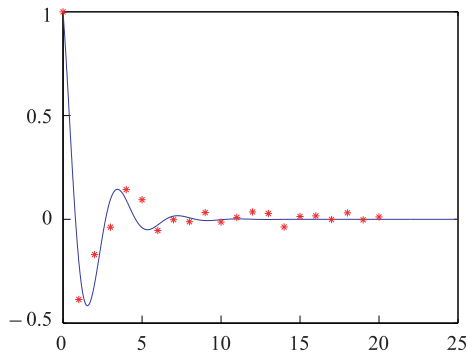


图 2 $\alpha=0.556\ 3$, $\beta=1.651\ 2$ 和 $\sigma = 3.327\ 3$ 时, X_t 的自相关函数及应力数据的经验自相关函数的拟合图

Fig. 2 Fitting the empirical auto-correlation function of the stress data to the auto-correlation function of X_t ; $\alpha=0.556\ 3$, $\beta=1.651\ 2$ and $\sigma = 3.327\ 3$

[参 考 文 献]

- [1] RAJPUT B S, ROSINSKI J. Spectral representations of infinitely divisible processes [J]. Probability Theory and Related Fields, 1989, 82: 451-487.
- [2] ZHANG S. On some dependence structures for multidimensional lévy driven moving averages [J]. Journal of the Korean Statistical Society, 2012, 41: 555-562.

- [3] ZHANG S, LIN Z, ZHANG X. A least squares estimator for Lévy-driven moving averages based on discrete time observations [J]. Communications in Statistics-Theory and Methods, 2013, DOI:10.1080/03610926.2012.763093. (published online)
- [4] BASSE A. Gaussian moving averages and semimartingales [J]. Electronic Journal of Probability, 2008, 13: 1140-1165.
- [5] BASSE A. Spectral representation of Gaussian semimartingales [J]. Journal of Theoretical Probability, 2009, 22: 811-826.
- [6] TAQQU M S. Benoît Mandelbrot and fractional Brownian motion [J]. Statistical Science, 2013, 28: 131-134.
- [7] BARNDORFF-NIELSEN O E, CORCUERA J M, PODOLSKIJ M. Multipower variation for Brownian semistationary processes [J]. Bernoulli, 2011, 17: 1159-1194.
- [8] PRESS W H, TEUKOLSKY S A, VETTERLING W T, et al. Numerical Recipes [M]//Fortran 77, the Art of Scientific Computing. 2nd ed. Cambridge: Cambridge University Press, 1992.
- [9] IBRAGIMOV I A, LINNIK Y V. Independent and Stationary Sequences of Random Variables [M]. Groningen: Wolters-Noordhoff, 1971.
- [10] DOUKHAN P. Mixing: Properties and Examples [M]. New York: Springer, 1994.
- [11] GRADSHTEYN I S, RYZHIK I M. Table of Integrals, Series, and Products [M]. New York: Academic Press, 2007.
- [12] FRISTEDT B, GRAY L. A Modern Approach to Probability Theory [M]. Boston: Birkhäuser, 1996.
- [13] POLITIS D N, ROMANO J P. A general resampling scheme for triangular arrays of α -mixing random variables with application to the problem of spectral density estimation [J]. Annals of Statistics, 1992, 20: 1985-2007.
- [14] 陈希孺. 高等数理统计学 [M]. 合肥: 中国科学技术大学出版社, 1999.
- [15] WOOD A T A, CHAN G. Simulation of stationary Gaussian processes in $[0, 1]^d$ [J]. Journal of Computational and Graphical Statistics, 1994, 3: 409-432.
- [16] DIEKER A B, MANDJES M. On spectral simulation of fractional Brownian motion [J]. Probability in the Engineering and Informational Sciences, 2003, 17: 417-434.

(责任编辑 王善平)