数理解析研究所講究録 第2116巻 2019年 57-63

定常過程に対する MA ブートストラップ

広島大学・大学院理学研究科　藤本 智博 (Tomohiro Fujimoto)
Graduate School of Science, Hiroshima University

広島大学・大学院理学研究科　井上 昭彦 (Akihiko Inoue)
Graduate School of Science, Hiroshima University

オリックス生命保険株式会社　清水 亮 (Ryo Shimizu)
ORIX Life Insurance Corporation

1　設定

これは, 定常過程に対する MA ブートストラップの研究に関する我々の最近の結果の報告である. 証明等の詳細については，別の場所で発表予定である.

$d \in \mathbb{N}$ とする．$a \in \mathbb{R}^{d\times d}$ に対し，$\|a\| := \sup_{u\in\mathbb{R}^d, |u|=1} |au|$ をそのスペクトル・ノルムとする．平均 $\mu_X \in \mathbb{R}^d$ を持つ $\mathbb{R}^d$-値の定常過程 $\{X_t\}_{t\in\mathbb{Z}}$ はMA (∞) 表現

$$X_k - \mu_X = \sum_{j=-\infty}^{k} \psi_{k-j}\epsilon_j, \qquad k \in \mathbb{Z} \tag{1.1}$$

により記述されるとする．ここで $\{\psi_k\}_{k=0}^{\infty}$ は $\mathbb{R}^{d\times d}$-値の列で，次を満たすとする：

(A1) $\psi_0 = I_d$,

(A2) $\sum_{j=0}^{\infty} j\|\psi_j\| < \infty$,

(A3) $\Psi(z) := \sum_{j=0}^{\infty} z^j\psi_j$ は $\det\Psi(z) \neq 0$ $(z \in \mathbb{C},\ |z| \le 1)$ を満たす.

また，次も仮定する：

(A4) $\{\epsilon_k\}_{k\in\mathbb{Z}}$ は $\mathbb{R}^d$-値の i.i.d. 確率ベクトル列で，$E[\|\epsilon_0\|^4] < \infty$, $E[\epsilon_0] = 0$ および $E[\epsilon_0\epsilon_0^{\mathrm{T}}] > 0$ を満たす.

$X_1, \ldots, X_n$ を $\{X_t\}$ からの標本とする．経験自己共分散関数 $\{\widehat{\gamma}(j)\}$ を次の様に定義する：

$$\widehat{\gamma}(j) := \begin{cases} \dfrac{1}{n}\displaystyle\sum_{k=1}^{n-j}(X_{k+j} - \bar{X}_n)(X_k - \bar{X}_n)^{\mathrm{T}}, & j = 0, 1, \ldots, n-1, \\ \dfrac{1}{n}\displaystyle\sum_{k=-j+1}^{n}(X_{k+j} - \bar{X}_n)(X_k - \bar{X}_n)^{\mathrm{T}}, & j = -n+1, \ldots, -1. \end{cases}$$

ここで，$\bar{X}_n$ は次で定義される $X_1, \ldots, X_n$ の標本平均である：

$$\bar{X}_n := \frac{1}{n}\sum_{k=1}^{n} X_k.$$

2　推定 MA 係数

標本 $X_1, X_2, \ldots, X_n$ のサイズ n が増加するにつれて増加する MA 次数 $p(n)$ を考える．ここで，$\{p(n)\}$ は $\mathbb{N}$-値の数列で

$$p(n) \to \infty \quad (n \to \infty)$$

および

$$p(n) = o(n) \quad (n \to \infty)$$

を満たすとする．以下，簡単のため，$p(n)$ を p と書く．

$1 \ll n$ に対し, 推定 AR 係数 $(\widehat{\phi}_{1,n}, \widehat{\phi}_{2,n}, \ldots, \widehat{\phi}_{n,n})$ を次の経験 Yule–Walker 方程式の解として定義する：

$$\sum_{j=1}^{p} \widehat{\phi}_{j,n} \widehat{\gamma}(i-j) = \widehat{\gamma}(i), \qquad i = 1, 2, \ldots, p.$$

推定 MA 係数 $\widehat{\psi}_{k,n} \in \mathbb{R}^{d \times d}$ $(k = 0, 1, \ldots, p)$ を次により定義する：

$$\widehat{\psi}_{k,n} := \widehat{v}_{k,n} \widehat{v}_{0,n}^{-1}, \qquad k = 0, \ldots, p. \tag{2.1}$$

ここで，

$$\widehat{v}_{k,n} := \widehat{\gamma}(k) - \sum_{j=1}^{p} \widehat{\gamma}(k+j) \widehat{\phi}_{j,n}^{\mathrm{T}}, \qquad k = 0, \ldots, p. \tag{2.2}$$

さらに，簡単のため，次のようにおく：

$$\widehat{\psi}_{k,n} := 0, \quad k \geq p+1.$$

次の定理は [1, Theorem 3.2] の類似物である．

定理 2.1. (A1)–(A4) および $p(n) = O((n/\log n)^{1/4})$ $(n \to \infty)$ を仮定する. すると次を満たす確率変数 n_1 が存在する：

$$\sup_{n \geq n_1} \sum_{j=0}^{\infty} j \|\widehat{\psi}_{j,n}\| < \infty \quad \text{almost surely.}$$

次の定理は [1, Theorem 3.1] の類似物である．

定理 2.2. (A1)–(A4) および $p(n) = O((n/\log n)^{1/2})$ $(n \to \infty)$ を仮定する. すると，次が成り立つ：

$$\sup_{0 \leq j < \infty} \|\widehat{\psi}_{j,n} - \psi_j\| = o(1) \quad (n \to \infty) \quad \text{almost surely.}$$

3 MA ブートストラップ

n が増加するにつれて増加する $q(n)$ を考える．ここで，$\{q(n)\}$ は $\mathbb{N}$-値の数列で

$$q(n) \to \infty \quad (n \to \infty)$$

および

$$q(n) = o(n) \quad (n \to \infty)$$

を満たすとする．以下，簡単のため，$q(n)$ を q と書く．

定義 1. (1.1) の $\{\epsilon_k\}_{k=q+1}^n$ の推定値として，$\sigma(X_1, X_2, \ldots, X_n)$-可測な $n-q$ 個の確率変数 $\{\widehat{\epsilon}_{k,n}\}_{k=q+1}^n$ が得られているとする．これらを

$$\tilde{\epsilon}_{k,n} := \widehat{\epsilon}_{k,n} - \frac{1}{n-q} \sum_{j=q+1}^{n} \widehat{\epsilon}_{j,n}, \qquad k = q+1, \ldots, n$$

と中心化し，$\{\tilde{\epsilon}_{k,n}\}_{k=q+1}^n$ の経験分布

$$\frac{1}{n-q} \sum_{k=q+1}^{n} \delta_{\tilde{\epsilon}_{k,n}}$$

の分布関数を $\widehat{F}_{\epsilon,n}$ と表す．$\{\tilde{\epsilon}_{t,n}\}$ のリサンプリング $\{\epsilon_t^*\}_{t\in\mathbb{Z}}$ を

$\{\epsilon_t^*\}$ は i.i.d. でかつ各 ϵ_t^* の分布は $\widehat{F}_{\epsilon,n}$ に従う

ように取る．観測データ $X_1, \ldots, X_n$ のリサンプリング $\{X_t^*\}_{t\in\mathbb{Z}}$ を，次の**近似移動平均表現**に従い構成する：

$$X_t^* = \bar{X}_n + \sum_{j=0}^{p} \widehat{\psi}_{j,n} \epsilon_{t-j}^* \qquad (t \in \mathbb{Z}). \tag{3.1}$$

以上の構成によるブートストラップを **MA ブートストラップ**とよぶ．

MA ブートストラップは，標本 $X_1, \ldots, X_n$ による条件付確率 P^* を導く．P^* に関する量を * をつけて書く．

$\{\widehat{\epsilon}_{k,n}\}_{k=q+1}^n$ に対し，次の 2 つの性質を仮定する：

(B1) 任意の $\xi \in \mathbb{R}^d$ に対し，$E^*\left[(\xi^{\mathrm{T}}\epsilon_t^*)^2\right] = E\left[(\xi^{\mathrm{T}}\epsilon_t)^2\right] + o_P(1)$ $(n \to \infty)$.

(B2) $n \to \infty$ のとき，$\epsilon_t^* \xrightarrow{d^*} \epsilon_t$ $(n \to \infty)$ in probability.

注意 3.1. (B2) をもっと明示的に書くと次の通りである: 任意の $\xi \in \mathbb{R}^d$ と $x \mapsto P(\xi^{\mathrm{T}}\epsilon_t \le x)$ の任意の連続点 x に対し，

$$P^*(\xi^{\mathrm{T}}\epsilon_t^* \le x) = P(\xi^{\mathrm{T}}\epsilon_t \le x) + o_P(1) \quad (n \to \infty).$$

定義 2. 推定値 $\{\widehat{\epsilon}_{k,n}\}_{k=q+1}^{n}$ が次の近似 AR 方程式により与えられる場合を考える：

$$\widehat{\epsilon}_{k,n} = \sum_{j=0}^{q} \widehat{\phi}_{j,n}(X_{t-j} - \bar{X}_n). \tag{3.2}$$

ただし，$q = p$ とする．この場合の MA ブートストラップを**部分 MA ブートストラップ**とよぶ．

次は [2, Lemmas 5.3 and 5.4] の多次元への拡張であり，証明も同様である．

定理 3.2. (A1)–(A4) および $p(n) = q(n) = o((n/\log n)^{1/2})$ を仮定する．すると，(3.2) で決まる $\{\widehat{\epsilon}_{k,n}\}_{k=q+1}^{n}$ は，(B1) と (B2) を満たす．

定理 2.1 と定理 2.2 および仮定 (B1)，(B2) を用いると，AR ブートストラップに対する結果 [2, Lemma 5.5]（の多次元版）に対して，次の MA 類似を証明することができる．

定理 3.3. (A1)–(A4), (B1)，(B2) および $p(n) = O((n/\log n)^{1/2})$ $(n \to \infty)$ を仮定する．すると次が成り立つ：

$$X_k^* \xrightarrow{d^*} X_k \quad (n \to \infty) \quad \text{in probability.}$$

定義 3. MA(p) 表現の変形

$$\varepsilon_k = X_k - \bar{X}_n - \sum_{l=1}^{p} \widehat{\psi}_{l,n}\varepsilon_{k-l}$$

と初期値 $\varepsilon_k = 0$ $(k = 0, \ldots, p-1)$ を用いて，ε_k $(k = p, \ldots, n)$ を求めることにより，推定値

$$\widehat{\varepsilon}_{k,n} = \varepsilon_k, \qquad k = p+1, \ldots, n$$

を定める．ただし，$q = p$ とする．この $\{\widehat{\epsilon}_{k,n}\}_{k=p+1}^{n}$ による MA ブートストラップを**完全 MA ブートストラップ**とよぶ．

完全 MA ブートストラップは，次節のシミュレーションの結果から分かるように，ブートストラップとしてよい性質を持つ．しかし，完全 MA ブートストラップの $\{\widehat{\epsilon}_{k,n}\}_{k=p+1}^{n}$ に対しては，まだ (B1) と (B2) の性質は証明はされておらず，open problem である．

4　シミュレーション

ここでは，[1] による AR ブートストラップ，前節の部分ブートストラップおよび完全 MA ブートストラップに対するシミュレーションの結果を比較する．

以下のシミュレーションでは，$d = 1$ とし，$\sigma_n^2 = n\mathrm{var}(T_n)$ の推定を行う．ただし，

$$T_n = \mathrm{median}\{X_1, \ldots, X_n\}$$

であり，サンプルサイズ n は 512 とする．次のモデルについてシミュレーションを行う：

(M1) AR(15), $X_t = \varepsilon_t + \sum_{j=1}^{15} \phi_j X_{t-j}$, $\phi_j = (-1)^{j+1} 7.5/(j+1)^3$ $(j = 1, \ldots, 15)$.

(M2) MA(15), $X_t = \varepsilon_t + \sum_{j=1}^{15} \sigma_j \varepsilon_{t-j}$, $\sigma_j = (-1)^{j+1} 1.5/(j+1)^3$ $(j = 1, \ldots, 15)$.

(M3) ARMA(2, 15), $X_t = \varepsilon_t + \sum_{j=1}^{2} \phi_j X_{t-j} + \sum_{j=1}^{15} \sigma_j \varepsilon_{t-j}$,
$\phi_j = (-1)^{j+1} 7.5/(j+1)^3$, $(j = 1, 2), \sigma_j = (-1)^{j+1} 1.5/(j+1)^3$ $(j = 1, \ldots, 15)$.

(M4) ARMA(10, 10), $X_t = \varepsilon_t + \sum_{j=1}^{10} \phi_j X_{t-j} + \sum_{j=1}^{10} \sigma_j \varepsilon_{t-j}$,
$\phi_j = (-1)^{j+1} 5.5/(j+1)^5$, $(j = 1, \ldots, 10), \sigma_j = (-1)^{j+1} 1.5/(j+1)^2$ $(j = 1, \ldots, 10)$.

(M5) MA(6), $X_t = \varepsilon_t + 0.1\varepsilon_{t-2} - 0.3\varepsilon_{t-6}$.

ただし，(M1) から (M4) に対しては

$$\varepsilon_t \text{ i.i.d. } \sim N(0, 1)$$

とし，(M5) に対しては

$$\varepsilon_t \text{ i.i.d. } \sim 0.95N(0, 1) + 0.05N(0, 100)$$

とする．また，次のモデルについてもシミュレーションを行う：

(M6) ARFIMA$(0, -0.25, 0)$.

分散 σ_n^2 は 1000 回のシミュレーションから求める．ブートストラップによる推定は次のように行う：

(a) 標本を発生させ，有限近似次数 p を AIC を最小にする $0 \le p \le 10 \log_{10} n$ から選ぶ．ただし，AR および 部分 MA ブートストラップの場合には AR モデルに対する AIC を用い，完全 MA ブートストラップの場合には MA モデルに対する AIC を用いる．

(b) リサンプリング $X_1^*, \ldots, X_n^*$ を発生させ $T_n^* = \text{median}\{X_1^*, \ldots, X_n^*\}$ を計算する．

(c) (b) を 300 回行い，300 個の T_n^* から $(\sigma_n^2)^* = n\text{var}^*(T_n^*)$ を計算する．

(d) (a) から (c) を 100 回行い $\text{E}[(\sigma_n^2)^*]$ と $\text{SD}((\sigma_n^2)^*)$ を求める．

シミュレーションの結果は表 1 から 3 のようになった．これらから次のようなことが見て取れる．

(1) 一般に，MA ブートストラップは，$\text{SD}((\sigma_n^2)^*)$ の値が AR ブートストラップよりも小さな値をとる傾向にあるという利点を持つ．

(2) (M3) の ARMA(2, 15) モデルに対しては，MA ブートストラップの方が AR ブートストラップよりも良い結果を与えている．

(3) 一方で (M1) の AR モデルに対しては，部分 MA ブートストラップは AR ブートストラップと比較して，大きく異なる値を推定してしまっているという点で劣ってしまう．

(4) モデル (M2), (M5) の MA モデルに対する結果を見ると，ノイズが正規か否かにかかわらず，いずれも同程度の良い推定値を与えている．

(5) 完全 MA ブートストラップにおいては，(M3) のようなモデルでは他の 2 つよりも良い結果が得られ，他の場合でも同程度の結果が得られている．

	σ_n^2	$E[(\sigma_n^2)^*]$	$\mathrm{SD}((\sigma_n^2)^*)$
$n = 512$			
(M1)	16.4	17.2	4.3
(M2)	1.7	1.9	0.4
(M3)	13.4	14.0	3.5
(M4)	3.1	3.1	0.9
(M5)	1.5	1.4	0.3
(M6)	0.7	1.0	0.2

表 1: AR ブートストラップ

	σ_n^2	$E[(\sigma_n^2)^*]$	$\mathrm{SD}((\sigma_n^2)^*)$
$n = 512$			
(M1)	16.4	10.6	3.1
(M2)	1.7	1.8	0.3
(M3)	13.4	13.1	3.0
(M4)	3.1	2.8	0.7
(M5)	1.5	1.6	0.4
(M6)	0.7	0.9	0.2

表 2: 部分 MA ブートストラップ

参考文献

[1] Bühlmann, P. (1995). Moving-average representation of autoregressive approximations. *Stochastic Process. Appl.* **60** 331–342.

[2] Bühlmann, P. (1997). Sieve bootstrap for time series. *Bernoulli* **3** 123–148.

	σ_n^2	$E[(\sigma_n^2)^*]$	$\mathrm{SD}((\sigma_n^2)^*)$
$n = 512$			
(M1)	16.4	14.9	3.7
(M2)	1.7	1.9	0.4
(M3)	13.4	13.7	2.8
(M4)	3.1	3.0	0.6
(M5)	1.5	1.6	0.4
(M6)	0.7	0.9	0.2

表 3: 完全 MA ブートストラップ