濟大論文集(人文·社會),35,309-316,1992
Cheju Univ. Jour. (Humanities ·
Social Sci.), 35, 309-316, 1992

Behavior of the AR and MA Parameters of Batch-Means Processes

*Young-Soo Chun**

Contents

Ⅰ. Introduction

It would be nice not to have to worry about the AR and MA orders of the model to be fitted. It also is reasonable to conjecture that the autocorrelation function of a batch-means process becomes simpler as the batch size grows, until it can be well approximated by an ARMA(p, q) model with low values of p and q. The present may depend on the far past, but only through the effect of the recent past, which in turn depends on its recent past. If we increase the batch size sufficiently, the data on the past having direct influence on the present ($X_t(m)$) may be entirely contained in the previous batch mean ($X_{t-1}(m)$). In this study we will prove that this actually occurs in some cases by showing that the AR and MA parameters of a batch-means process derived from an ARMA(p, q) process tends to dominate the lower-order ones as the batch size grows.

* 경상대학 경영학과(Dept. of Business Administration, Cheju Univ., Cheju-do, 690-756, Korea)

Ⅱ. Derivation

From Kang and Schmeiser, the batch means of the stationary ARMA(*p, q*) process is the stationary ARMA(*p, q'*) process with

$$q' = p - \lfloor (p-q)/m \rfloor \cdots\cdots\cdots\cdots\cdots\cdots\cdots (1)$$

The AR parameters, $\phi_1(m), \cdots, \phi_p(m)$, of the batch-means process for the batch size *m* are the coefficients of $x^1, \cdots, x^p$ of $\Pi_{j=0}^{p}(1-a_j^m x)$, respectively, where $\{a_1, \cdots, a_p\}$ are the roots, complex or real, of the polynomial equation

$$\Pi_{j=0}^{p}\phi_j a^{p-j} = 0$$

(Note the function $\Pi_{j=0}^{p}\phi_j z^j$ in z is called the *characteristic function.*)
That is, there exists a set of values $\{a_1, \cdots, a_p\}$ such that

$$\phi_1 = a_1 + a_2 + \cdots + a_p,$$
$$\phi_2 = -(a_1a_2 + a_1a_3 + \cdots + a_1a_p + a_2a_3 + \cdots + a_{p-1}a_p),$$
$$\cdot$$
$$\cdot$$
$$\phi_p = (-1)^{p-1}a_1a \cdots a_p.$$

Also,

$$\phi_h(m) = (-1)^{h-1}\Sigma\delta(i_1, i_2 \cdots, i_h)(a_{i1}a_{i2} \cdots a_{ih})^m,\ h = 1, 2, \cdots, p, \cdots\cdots\cdots\cdots (2)$$

where

$$\delta(i_1, i_2, \cdots, i_h) = \begin{cases} 1 & \text{if } i_1, i_2, \cdots, i_h \text{ are all different} \\ 0 & \text{otherwise.} \end{cases}$$

That is,

$$\phi_1(m) = a_1^m + a_2^m + \cdots + a_p^m,$$
$$\phi_2(m) = -\{(a_1a_2)^m + (a_1a_3)^m + \cdots + (a_1a_p)^m + (a_2a_3)^m + \cdots + (a_{p-1}a_p)^m\}$$
$$\cdots\cdots$$
$$\phi_p(m) = (-1)^{p-1}(a_1a_2 \cdots a_p)^m.$$

Notice that $a_j < 1$ for $1 \le j \le p$ by stationarity of the original process. Therefore $|\phi_h(m_2)| < |\phi_h(m_1)|$ for $m_2 > m_1$ if the a_j's are all real and m_2 and m_1 are even. That is, $\phi_h(m)$ dies out with increasing m. Also, we can see that $\phi_1(m)$ tends to dominate the other AR parameters, $\phi_2(m), \cdots, \phi_p(m)$, as m grows because higer-order AR parameters would die out faster than lower-order AR parameters unless more than one of the a_h's are close to 1.

Ⅲ. Dominance of Low-Order Parameters

Even though it is not always the case, assume for the moment that the a_h's are all real and m is even. We can sort $\{a_1, \cdots, a_p\}$ into $\{a_{(1)}, \cdots, a_{(p)}\}$ in descending order of absolute value. i.e. $1 > |a_{(1)}| \ge |a_{(2)}| \ge \cdots \ge |a_{(p)}| \ge 0$. Here we can assume $|a_{(1)}| > 0$. (If $|a_{(1)}| = 0$, it is implied that $a_1 = a_2 = \cdots = a_p = 0$, and thus $\phi_1 = \phi_2 = \cdots = \phi_p = 0$. Therefore, the process does not have the autoregressive property, making it very simple to estimate the autocorrelations.) Then we have from (2)

$$\phi_1(m) = a_1^m + \cdots + a_p^m \ge a_{(1)}^m > 0,$$

and

$$0 \le |\phi_h(m)| \le {}_pC_h(a_{(1)}a_{(2)} \cdots a_{(h)})^m.$$

Therefore,

$$0 \le |\phi_h(m)/\phi_1(m)| \, {}_pC_h(a_{(1)}a_{(2)} \cdots a_{(h)})^m/a_{(1)}$$
$$= {}_pC_h(a_{(2)} \cdots {}_{(h)})^m \to 0 \text{ as } m \to \infty, \; 2 \le h \le p$$

because ${}_pC_h$ is a constant and $0 \le (a_{(2)} \cdots a_{(h)})^2 < 1$. That is, $\phi_1(m)$ tends to dominate higher-order AR parameters as m grows. The speed of the convergence is governed by $|(a_{(2)} \cdots a_{(h)})^2|$, and would be fast unless $a_{(2)}, \cdots, a_{(p)}$ are all close to 1.

IV. Empirical study

It need not be true that the a_h's are all real, since they may include some pairs of conjugate complex numbers. In such cases, $\phi_1(m)$ shows damped fluctuation[1] with m possibly becoming close to 0, causing $\phi_h(m)/\phi_1(m)$ to explode even with large m. That is, the ratio $\phi_h(m)/\phi_1(m)$ may not converge to zero when some a_j's have an imaginary component.

To check how often this failure of convergence happens, we executed an experiment with 200 randomly chosen AR(3) models. In this experiment, we generated a candidate set of parameter values of an AR(3) model by use of a random-number generator such that $|\phi_h| < {}_pC_h$, which obviously should be the case from (2). The generated model went through a stationarity check discussed by Pandit and Wu. We kept it if it passed the stationarity check, and otherwise threw it away to generate and test another candidate set of parameters. In this way we generated many sets of parameters until we got 200 stationary sets. It should be noted that a stationary set is chosen even if the values of the roots (a_j's) of the associated characteristic equation may be complex.

Taking each set as the parameters ϕ_1, ϕ_2, ϕ_3 of an original stationary AR(3) process, we calculated $\phi_1(m)$, $\phi_2(m)$, $\phi_2(m)$, $\theta_1(m)$, $\theta_2(m)$, and $\theta_3(m)$, the AR and MA parameters of a batch-means process with batch size m for a various values of m. The results are summarized in Table 1.

In Table 1 we see that $\phi_2(m)$ tends to die out with increasing m much faster than does $\phi_1(m)$. For the original process, $\phi_2(m)/\phi_1(m) < 0.10$ in just 5% of the cases. However, this frequency increases with the batch size m to become 83.5% for $m = 32$. The dominance of $\phi_1(m)$ over $\phi_3(m)$ is even greater. In 98% of the cases, $|\phi_1(m)| > 10|\phi_3(m)|$ for the batch size $m = 16$. In summary, we can conclude that we may neglect all the AR

1) Suppose the first s pairs of $\{a_j, 1 \le j \le p\}$ are conjugate complex numbers and that they can be represented as $r_1 e^{i\omega^1}$, $r_1 e^{-i\omega^2}$, $\cdots$, and $r_s e^{i\omega^s m}$ $r_s e^{-i\omega^s m}$ where r_j's and ω_j's are real numbers, $|r_j| < 1$, and i = sqrt(-1). Then

$$\begin{aligned}\phi_1(m) &= a_1^m + \cdots + a_p^m \\ &= r_1 e^{i\omega^1} + r_1 e^{-i\omega^1} + \cdots r_s e^{i\omega^s} + r_s e^{-i\omega^s} + a_{2s+1}^m + \cdots + a_p^m \\ &= r_1^m \cos(m\omega_1) + \cdots + r_s^m/\cos(m\omega_s) + a_{2s+1}^m + \cdots + a_p^m\end{aligned}$$

Table 1. Relative frequencies that $|\phi_1(m)/\phi_h(m)| > c$ and $|\theta_1(m)/\theta_h(m)| > c$
Original process : AR(3)
Number of replications : 200 (Unit : %)

	Batch size	c=10	c=20
$\lvert\phi_1(m)/\phi_2(m)\rvert$	1	5.0	2.5
	2	12.5	5.5
	4	22.5	12.5
	8	37.5	26.5
	16	64.0	46.5
	32	83.5	72.5
	64	90.5	87.5
$\lvert\phi_1(m)/\phi_3(m)\rvert$	1	6.5	3.0
	2	33.0	21.0
	4	63.0	51.0
	8	91.0	83.0
	16	98.0	92.5
	32	100.0	99.0
	64	100.0	100.0
$\lvert\theta_1(m)/\theta_2(m)\rvert$	1	N.A.	N.A.
	2	20.5	7.5
	4	33.0	18.0
	8	46.5	30.5
	16	72.0	55.5
	32	86.5	76.5
	64	94.5	88.0
$\lvert\theta_1(m)/\theta_3(m)\rvert$	1	N.A.	N.A.
	2	N.A.	N.A.
	4	80.5	64.5
	8	93.0	91.0
	16	99.5	97.5
	32	100.0	99.5
	64	100.0	99.5

parameters except the first-order one when we fit an ARMA model to a batch-means process if the batch size is large enough.

The dominance relationship observed in AR parameters of a batch-means process is also expected in MA parameters. From (1), $\dot{q} = p$ if $p > q$ and m is sufficiently large. For this value of m, the batch-means process can be well approximated by an ARMA(1, p)

model if the dominance of $\phi_1(m)$ has already been realized. Now consider an even larger batch size, say, ms, a batch size s times as large as m. The MA order q' for this batch size must be $q' = 1 - \lfloor (1-p)/s \rfloor$ again by (1). Therefore, q' would be at most 1 for the batch size ms if s is also large enough.

In the experiment mentioned already in this section, we calculated MA parameters of batch-means processes for the 200 randomly chosen AR(3) processes. As shown in Table 1, we can conclude that $\theta_1(m)$ tends to be prevalent among MA parameters as the batch size grows.

Actually, we carried out the same type of experiments for the ARMA(2, 2), ARMA(3, 3), and ARMA(4, 4) processes. And we got the similar observations.

V. Conclusions

In summary, it appears that the batch-means data can be well approximated by an ARMA(1, 1) model if the batch size is sufficiently large. Thus, we have found an alternative to the model-selection c.i. procedure introduced by Schriber and Andrews. Instead of trying to find p and q, the appropriate AR and MA orders of the model to be fitted to the original data, we can try to fit, say, an ARMA(1, 1) model to a batch-means sequence if we have a rule to determine the batch size appropriate for ARMA(1, 1) fitting.

References

Amemiya T. and R. Y. Wu. The effect of Aggregation on Prediction in the Autoregressive Model. *Journal of the American Statistical Association.* 67, 1972, pp. 628~632.

Anderson O. D., *Time Series Analysis and Forecasting: The Box Jenkins Approach.* Butterworths, London, 1979.

Anderson, T. W. *The statistical Analysis of Time Series.* Wiley, New York. 1970.

Box, G. E. P. and G. M. Jenkins. *Time Series Analysis: Forecasting and Control.* Holden-Day, California, 1976.

Brillinger, D. R. Estimation of the Mean of a Stationary Time Series by Sampling. *Journal of Applied Statistics.* 10, 1973, pp. 419~431.

Cryer, J. D. *Time Series Analysis.* Duxbury Press, Boston, 1986.

Fishman, G. S. *Concepts and Methods in Discrete Event Digital Similation.* Wiley, New York, 1978.

Fishman, G. S., *Principles of Discrete Event Simulation,* Wiley, New York, 1978.

Kang K. and B. Schmeiser. Properties of Batch Means from Stationary ARMA Time Series. *Operations Research Letters,* 1987.

Law, A. M., Confidence Intervals in Discrete event simulation : A Comparison of Replication and Batch Means. *Naval Research Logistics Quarterly.* 24, 1977, pp. 667~678.

Law, A. M. and W. D. Kelton, *Simulation Modeling and Analysis.* McGraw-Hill, New York, 1982a.

Pandit, S. M. and S. Wu. *Time Series and System Analysis with Application.* Wiley, New York, 1983.

Schrage, L. A More Portable Fortran Random Number Generator, *ACM. Trans. Math. Software.* 5, 1979, pp. 132~138.

〈국문초록〉

구간평균 프로세스의 AR 및 MA모수의 변화특성

천 영 수

본 연구는 구간평균 프로세스의 모수들이 구간규모의 증가에 따라 어떻게 변화하는지 분석하는 것을 목적으로 수행되었다. 본 연구에서 규명된 바는 구간규모가 증가함에 따라 모든 모수들의 값이 0에 수렴하지만 특히 고차의 모수들이 더욱 빠른 속도로 0에 수렴하기 때문에, 어느 정도의 구간규모에 대한 구간평균 프로세스는 간단한 ARMA, 예컨대 ARMA(1, 1), 모델로 잘 대표될 수 있다는 점이다. 이 발견은 프로세스의 정상상태 평균에 대하여 신뢰구간을 설정하는 작업을 수월하게 한다는 점에서 각별한 의의를 갖는다.

어느 프로세스의 정상상태 평균에 대하여 추정하기 위하여 그 프로세스의 관측치인 시계열 자료로부터 표본평균의 분산을 추정한다. 이때, 시계열자료에 내재하는 자동상관은 표본평균의 분산에 대한 추정량에 편기를 유발한다. 이 편기를 제거하기 위한 수단으로서 원 자료를 가급적 커다란 구간으로 묶어서 일련의 구간평균을 얻는다. 그리고는 구간평균 프로세스를 i. i. d. 프로세스로 간주하여 신뢰구간을 구한다. 또 다른 방법은 원래의 프로세스를 잘 표현할 수 있는 ARMA 모델을 추정하여 이 모델이 가지고 있는 자동상관구조에 대한 정보를 활용하여 편기없는 분산추정량을 구하려고 시도하는 것이다. 그러나 구간평균 프로세스가 i. i. d. 프로세스에 충분히 수렴하도록 구간의 크기를 증가시키면 필연적으로 정보를 상실하게 된다. 또한 원래의 시계열 자료를 잘 표현할 수 있는 ARMA 모델을 찾아내는 것은 쉬운 일이 아니다.

그러나 본연구에서 발견된 결과에 따르면 적절한 구간규모에 상응하는 구간평균 프로세스를 ARMA(1, 1) 모델로 대표하여 모수를 추정하고 이 추정값으로부터 분산추정량을 용이하게 계산할 수 있게 된다.