Exponential decay rate of partial autocorrelation coefficients of ARMA and short-memory processes

Akimichi Takemura*

January, 2016

Abstract

We present a short proof of the fact that the exponential decay rate of partial autocorrelation coefficients of a short-memory process, in particular an ARMA process, is equal to the exponential decay rate of the coefficients of its infinite autoregressive representation.

1 Introduction

The autocorrelation coefficients and the partial autocorrelation coefficients are basic tools for model selection in time series analysis based on ARMA models. For AR models, by the Yule-Walker equation, the autocorrelation coefficients satisfy a linear difference equation with constant coefficients and hence the autocorrelation coefficients decay to zero exponentially with the rate of the reciprocal of the smallest absolute value of the roots of the characteristic polynomial of the AR model. This also holds for ARMA models, because their autocorrelations satisfy the same difference equation defined by their AR part, except for some initial values.

On the other hand, it seems that no clear statement and proof is given in standard textbooks on time series analysis concerning the exponential decay rate of the partial autocorrelation coefficients for MA models and ARMA models. For example, in Section 3.4.2 of [2] the following is stated without clear indication of the decay rate.

Hence, the partial autocorrelation function of a mixed process is infinite in extent. It behaves eventually like the partial autocorrelation function of a pure moving average process, being dominated by a mixture of damped exponentials and/or damped sine waves, depending on the order of the moving average and the values of the parameters it contains.

In Section 3.4 of [3] the following is stated on MA(q) processes again without clear indication of the decay rate.

In contrast with the partial autocorrelation function of an AR(p) process, that of an MA(q) process does not vanish for large lags. It is however bounded in absolute value by a geometrically decreasing function.

*Graduate School of Information Science and Technology, University of Tokyo
The purpose of this paper is to give a clear statement on the decay rate and its short proof. Because of the duality between AR models and MA models, it is intuitively obvious that the partial autocorrelation coefficients of an ARMA model decay to zero at the rate of the reciprocal of the smallest absolute value of the roots of the characteristic polynomial of the MA part of the model. Note that this rate is also the decay rate of the coefficients of the AR(∞) representation.

In literature the sharpest results on asymptotic behavior of partial autocorrelation functions have been given by Akihiko Inoue and his collaborators (e.g. [4], [6], [5], [1]). They give detailed and deep results on the polynomial decay rate of the partial autocorrelation coefficients for the case of long-memory processes. Concerning ARMA processes, the most clear result seems to have been given by Inoue in Section 7 of [5]. However his result is one-sided, giving an upper bound on the exponential rate, whereas Theorem 2.1 in this paper gives an equality.

2 Main result and its proof

We consider a zero-mean causal and invertible weakly stationary process \(\{X_t\}_{t \in \mathbb{Z}}\) having an AR(∞) representation and an MA(∞) representation given by

\[
X_t = \pi_1 X_{t-1} + \pi_2 X_{t-2} + \cdots + \epsilon_t, \quad \pi(B)X_t = \epsilon_t, \quad \sum_{i=1}^\infty |\pi_i| < \infty, \quad (1)
\]

\[
X_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \cdots = \psi(B)\epsilon_t, \quad \sum_{i=1}^\infty |\psi_i| < \infty. \quad (2)
\]

For an ARMA(p, q) process

\[
\phi(B)X_t = \theta(B)\epsilon_t,
\]

\(\pi_1, \pi_2, \ldots\), decay exponentially with the rate of the reciprocal of the smallest absolute value of the roots of \(\theta(B) = 0\) and similarly \(\psi_1, \psi_2, \ldots\), decay, with \(\theta(B)\) replaced by \(\phi(B)\). The autocovariance function of \(\{X_t\}\) is

\[
E(X_tX_{t+k}) = \gamma_k = \gamma_{-k} = \sigma^2_\epsilon (\psi_k + \sum_{i=1}^\infty \psi_{k+i}\psi_i), \quad k \geq 0,
\]

where \(\sigma^2_\epsilon = E(\epsilon^2_t)\). Let \(H\) denote the Hilbert space spanned by \(\{X_t\}\) and for a subset \(I \subset \mathbb{Z}\) of integers, let \(P_I\) denote the orthogonal projector onto the subspace \(H_I\) spanned by \(\{X_t\}_{t \in I}\). The \(k\)-th partial autocorrelation is defined by \(\phi_{kk}\) in

\[
P_{[t-k,t-1]}X_t = \phi_{k1}X_{t-1} + \cdots + \phi_{kk}X_{t-k}.
\]

We state our theorem, which shows that the radius of convergence is common for the infinite series with coefficients \(\{\pi_n\}\) and coefficients \(\{\phi_{nn}\}\).

**Theorem 2.1.** Let \(\{X_t\}_{t \in \mathbb{Z}}\) be a zero-mean causal and invertible weakly stationary process with its AR(∞) representation given by (1) and let \(\phi_{nn}\) be the \(n\)-th partial autocorrelation coefficient. Then

\[
\limsup_{n \to \infty} |\phi_{nn}|^{1/n} = \limsup_{n \to \infty} |\pi_n|^{1/n}. \quad (3)
\]
By our assumptions, both \( \{\pi_n\} \) and \( \{\phi_{mn}\} \) are bounded and hence we have \( \limsup_{n \to \infty} |\pi_n|^{1/n} \leq 1 \), \( \limsup_{n \to \infty} |\phi_{mn}|^{1/n} \leq 1 \). Note that (3) only gives the exponential decay rates of \( \pi_n \) and \( \phi_{mn} \) and does not distinguish polynomial rates since \( (n^k)^{1/n} \to 1 \) as \( n \to \infty \) for any power of \( n \). Akihiko Inoue and his collaborators provided detailed analyses of the polynomial decay rate of \( \phi_{mn} \) for the case of long-memory processes (e.g. [4], [6], [5], [1]).

For proving Theorem 2.1 we present two lemmas.

**Lemma 2.2.** Suppose \( \limsup_{n \to \infty} |\pi_n|^{1/n} < 1 \). Then \( \limsup_{n \to \infty} |\phi_{mn}|^{1/n} \leq \limsup_{n \to \infty} |\pi_n|^{1/n} \).

**Proof.** Let \( \limsup_{n \to \infty} |\pi_n|^{1/n} = c_0 < 1 \). Then for every \( c \in (c_0, 1) \), there exist \( n_0 \) such that 
\[
|\pi_n| < c^n, \quad \forall n \geq n_0.
\]

We denote the \( h \)-period \((h \geq 1)\) ahead prediction by 
\[
P_{[t-k,t-1]}X_{t+h-1} = \phi_{k_1}^{(h)}X_{t-1} + \cdots + \phi_{k_k}^{(h)}X_{t-k} \quad (\phi_{k}^{(1)} = \phi_{k}).
\]

Here \( \phi_{k_1}^{(h)} \) is the partial regression coefficient of \( X_{t-1} \) in regressing \( X_{t+h-1} \) to \( X_{t-1}, \ldots, X_{t-k} \). Hence it is written as
\[
\phi_{k_1}^{(h)} = \frac{\text{Cov}(P_{[t-k,t-2]}X_{t+h-1}, P_{[t-k,t-2]}X_{t-1})}{\text{Var}(P_{[t-k,t-2]}X_{t-1})},
\]
where \( P_{[t-k,t-2]} \) is the projector onto the orthogonal complement of \( H_{[t-k,t-2]} \). Then \( |\phi_{k_1}^{(h)}| \) is uniformly bounded from above as
\[
|\phi_{k_1}^{(h)}| \leq \sqrt{\frac{\text{Var}(P_{[t-k,t-2]}X_{t+h-1})\text{Var}(P_{[t-k,t-2]}X_{t-1})}{\text{Var}(P_{[t-k,t-2]}X_{t-1})}} = \sqrt{\frac{\text{Var}(X_{t+h-1})}{\text{Var}(X_{t-1})}} = \sqrt{\frac{\sigma^2}{\varphi^2}} \tag{4}
\]
\[
\text{In (1)} \text{ we apply } P_{[t-k,t-1]} \text{ to } X_t. \text{ Then }
\]
\[
\phi_{k_1}X_{t-1} + \cdots + \phi_{k_k}X_{t-k} = P_{[t-k,t-1]}X_t = P_{[t-k,t-1]}P_{(-\infty,t-1]}X_t = P_{[t-k,t-1]} \left( \sum_{l=1}^{\infty} \pi_lX_{t-l} \right) = \pi_1X_{t-1} + \cdots + \pi_kX_{t-k} + \sum_{l=k+1}^{\infty} \pi_lP_{[t-k,t-1]}X_{t-l}. \tag{5}
\]

Now by time reversibility of the covariance structure of weakly stationary processes we have
\[
P_{[t-k,t-1]}X_{t-k-h} = \phi_{k_1}^{(h)}X_{t-k} + \cdots + \phi_{k_k}^{(h)}X_{t-1}.
\]
By substituting this into (5) and considering the coefficient of $X_{r-k}$ we have

$$\phi_{kk} = \pi_k + \sum_{h=1}^{\infty} \pi_{k+h} \phi_{k1}^{(h)},$$

where the right-hand side converges absolutely under our assumptions. Then

$$|\phi_{kk}| \leq |\pi_k| + \sum_{h=1}^{\infty} |\pi_{k+h}| |\phi_{k1}^{(h)}|.$$

For $k \geq n_0$, in view of (4), the right-hand side is bounded as

$$|\phi_{kk}| \leq c^k (1 + \sum_{h=1}^{\infty} c^h \sqrt{\gamma_0/\sigma_z^2}) = c^k (1 + \frac{c \sqrt{\gamma_0/\sigma_z^2}}{1 - c}).$$

Then

$$\limsup_{k \to \infty} |\phi_{kk}|^{1/k} \leq c.$$

Since $c > c_0$ was arbitrary, we let $c \downarrow c_0$ and obtain

$$\limsup_{n \to \infty} |\phi_{nn}|^{1/n} \leq c_0 = \limsup_{n \to \infty} |\pi_n|^{1/n}.$$

\[ \Box \]

**Lemma 2.3.** Suppose $\limsup_{n \to \infty} |\phi_{nn}|^{1/n} < 1$. Then $\limsup_{n \to \infty} |\pi_n|^{1/n} \leq \limsup_{n \to \infty} |\phi_{nn}|^{1/n}$.

**Proof.** This follows from the Durbin-Levinson algorithm. Consider $j = n$ in

$$\phi_{n+1,n} = \phi_{n,n} - \phi_{n+1,n+1} \phi_{n,n-1}, \quad j = 1, 2, \ldots, n. \quad (6)$$

The initial value is

$$\phi_{n+1,n} = \phi_{n,n} - \phi_{n+1,n+1} \phi_{n,1}.$$

Using (6) for $n$ replaced by $n + 1$, $j = n$, and substituting the initial value, we obtain

$$\phi_{n+2,n} = \phi_{n+1,n} - \phi_{n+2,n+2} \phi_{n+1,2}$$

$$= \phi_{n,n} - \phi_{n+1,n+1} \phi_{n,1} - \phi_{n+2,n+1} \phi_{n+1,2}.$$

Repeating the substitution, we have

$$\phi_{n+h,n} = \phi_{n,n} - \phi_{n+1,n+1} \phi_{n,1} - \cdots - \phi_{n+h,n+h} \phi_{n+h-1,h}.$$

As $h \to \infty$, the left-hand side converges to $\pi_n$ (cf. Theorem 7.14 of [8]). Hence

$$\pi_n = \phi_{n,n} - \sum_{h=1}^{\infty} \phi_{n+h,n+h} \phi_{n+h-1,h}$$

and

$$|\pi_n| \leq |\phi_{n,n}| + \sum_{h=1}^{\infty} |\phi_{n+h,n+h}| |\phi_{n+h-1,h}|$$
Now arguing as in (4), we see that $|\phi_{n+h-1,h}|$ is uniformly bounded as

$$|\phi_{n+h-1,h}| \leq \sqrt{\frac{\gamma_0}{\text{Var}(P_{(-\infty,-1]} \cup [1,\infty)) X_t)}\frac{\gamma_0}{\text{Var}(P_{(-\infty,-1]} \cup [1,\infty)) X_0)}.$$  

Here the denominator is positive, because under our assumptions $\{X_t\}$ is “minimal” (cf. Theorem 8.11 of [8]. [7, Section 2]). The rest of the proof is the same as in the proof of Lemma 2.2

By the above two lemmas, Theorem 2.1 is proved as follows.

Proof of Theorem 2.1 As noted after Theorem 2.1 both $\limsup_{n \to \infty} |\phi_n|^{1/n}$ and $\limsup_{n \to \infty} |\pi_n|^{1/n}$ are less than or equal to 1. If

$$\limsup_{n \to \infty} |\phi_n|^{1/n} < 1 \quad \text{or} \quad \limsup_{n \to \infty} |\pi_n|^{1/n} < 1,$$

then by the above two lemmas both of them have to be less than one and they have to be equal. The only remaining case is that they are equal to 1.

Acknowledgment This research is supported by JSPS Grant-in-Aid for Scientific Research No. 25220001. The author is grateful for useful comments by Tomonari Sei.

References

[1] N. H. Bingham, A. Inoue, and Y. Kasahara. An explicit representation of Verblunsky coefficients. Statist. Probab. Lett., 82(2):403–410, 2012.

[2] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. Time Series Analysis: Forecasting and Control. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, fifth edition, 2015.

[3] P. J. Brockwell and R. A. Davis. Time Series: Theory and Methods. Springer Series in Statistics. Springer-Verlag, New York, second edition, 1991.

[4] A. Inoue. Asymptotic behavior for partial autocorrelation functions of fractional ARIMA processes. Ann. Appl. Probab., 12(4):1471–1491, 2002.

[5] A. Inoue. AR and MA representation of partial autocorrelation functions, with applications. Probab. Theory Related Fields, 140(3-4):523–551, 2008.

[6] A. Inoue and Y. Kasahara. Explicit representation of finite predictor coefficients and its applications. Ann. Statist., 34(2):973–993, 2006.

[7] N. P. Jewell and P. Bloomfield. Canonical correlations of past and future for time series: definitions and theory. Ann. Statist., 11(3):837–847, 1983.

[8] M. Pourahmadi. Foundations of Time Series Analysis and Prediction Theory. Wiley Series in Probability and Statistics: Applied Probability and Statistics. Wiley-Interscience, New York, 2001.