Asymptotic Properties of General Autoregressive Models and Strong Consistency of Least-Squares Estimates of Their Parameters

T. L. Lai* and C. Z. Wei†

Columbia University and University of Maryland

Communicated by P. R. Krishnaiah

This paper establishes several almost sure asymptotic properties of general autoregressive processes. By making use of these properties, we obtain a proof of the strong consistency of the least-squares estimates of the parameters of the process without any assumption on the roots of the characteristic polynomial.

1. Introduction

Consider the autoregressive AR(p) model

\[ y_n = \beta_1 y_{n-1} + \cdots + \beta_p y_{n-p} + \varepsilon_n, \]

(1.1)

where \( y_n \) is the observation and \( \varepsilon_n \) is the (unobservable) random disturbance (noise) at stage \( n \), and \( \beta_1, \ldots, \beta_p \) are the parameters of the model. Throughout the sequel, we shall assume that \( \{\varepsilon_n\} \) is a martingale difference sequence with respect to an increasing sequence of \( \sigma \)-fields \( \{\mathcal{F}_n\} \), i.e., \( \varepsilon_n \) and \( y_n \) are \( \mathcal{F}_n \)-measurable and \( E(\varepsilon_n | \mathcal{F}_{n-1}) = 0 \) a.s. for every \( n \). An important example is a sequence of independent random variables with zero means. We shall let

\[ \phi(z) = z^p - \beta_1 z^{p-1} - \cdots - \beta_p \]

(1.2)

denote the characteristic polynomial of the autoregressive model (1.1), and let

\[ B = \begin{pmatrix} \beta_1 & \cdots & \beta_{p-1} & \beta_p \\ I_{p-1} & 0 \end{pmatrix} \]

(1.3)

Received January 18, 1982.

AMS 1980 Subject Classifications: Primary 62J05, 62M10; secondary 60F15.

Key Words and Phrases: Autoregressive processes, characteristic polynomial, purely explosive and non-explosive models, least-squares estimates, strong consistency.

* Research supported by the National Science Foundation and the National Institutes of Health.

† Research supported by the National Science Foundation.

0047-259X/83 $3.00

Copyright © 1983 by Academic Press, Inc.
All rights of reproduction in any form reserved.
denote the associated companion matrix, where $I_k$ denotes the $k \times k$ identity matrix. Defining the $p$-dimensional vectors

$$Y_n = (y_n, y_{n-1}, \ldots, y_{n-p+1})', \quad e_n = (e_n, 0, \ldots, 0)',$$

the model (1.1) can be expressed in vector form as

$$Y_n = BY_{n-1} + e_n.$$  \hspace{1cm} (1.5)

For the AR($p$) model (1.1), a commonly used estimate of the parameter vector $\beta = (\beta_1, \ldots, \beta_p)'$ is the least-squares estimate

$$b_n = (X_n'X_n)^{-1} X_n' (y_{n-p+1}, \ldots, y_n)', \quad n > p,$$  \hspace{1cm} (1.6)

where

$$X_n = \begin{pmatrix} y_p & \cdots & y_1 \\ y_{p+1} & \cdots & y_2 \\ \vdots & \ddots & \vdots \\ y_{n-1} & \cdots & y_{n-p} \end{pmatrix}$$  \hspace{1cm} (1.7)

is the design matrix at stage $n$ and $(X_n'X_n)^{-1}$ is the Moore–Penrose generalized inverse of $X_n'X_n$. The weak consistency of $b_n$ was first obtained by Mann and Wald [12] under the assumptions that $\phi(z)$ has roots inside the unit circle and that the $e_n$ are i.i.d. with $Ee_1 = 0$, $Ee_1^2 > 0$ and $EE_1^4 < \infty$. The case where $\phi(z)$ has roots outside the unit circle was subsequently investigated by Rubin [16] and Anderson [1]. Rao [15] considered the case of two roots for $\phi(z)$ such that one root lies inside and the other outside the unit circle. The case where $\phi(z)$ has a root on the unit circle was first studied by White [19]. Muench [13] and Stigum [17] unified all these earlier efforts and obtained the weak consistency of $b_n$ for the general autoregressive model without any assumptions on the roots of $\phi(z)$.

These weak consistency results in the literature have been obtained by rather complicated computations of moments of certain linear and quadratic forms related to (1.6), and such an approach fails to show the stronger result that $b_n$ is in fact strongly consistent. Herein we establish the strong consistency in the general AR($p$) model by making use of certain almost sure asymptotic behavior of the sequence \{\{\{y_n\}\}\} and related quadratic forms. In Section 4, we prove the following:

**Theorem 1.** Suppose that in the AR($p$) model (1.1), $\{e_n\}$ is a martingale difference sequence such that

$$\liminf_{n \to \infty} E(e_n^2 | \mathcal{F}_{n-1}) > 0 \quad a.s.$$  \hspace{1cm} (1.8)
and

\[ \sup_n E(|\epsilon_n|^\alpha | \mathcal{F}_{n-1}) < \infty \text{ a.s. for some } \alpha > 2. \]

Define \( b_n, X_n \) as in (1.6) and (1.7). Then

\[ \lim inf_{n \to \infty} n^{-1} \lambda_{\text{min}}(X'_n X_n) > 0 \quad \text{a.s.} \quad (1.9) \]

and

\[ \lim_{n \to \infty} b_n = \beta \quad \text{a.s.} \quad (1.10) \]

In (1.9) and the sequel, we use the notations \( \lambda_{\text{min}}(X'_n X_n) \) and \( \lambda_{\text{max}}(X'_n X_n) \) to denote the minimum and the maximum eigenvalues of the matrix \( X'_n X_n \). When all roots of the characteristic polynomial \( \phi(z) \) lie on or inside the unit circle, we shall call the autoregressive model \textit{non-explosive}. We recently obtained in [8, Corollary 1] the strong consistency of \( b_n \) in non-explosive autoregressive models, generalizing an earlier result of Anderson and Taylor [2] for stationary autoregressive models in which all roots of \( \phi(z) \) lie inside the unit circle and \( E(\epsilon_{n-1}^2 | \mathcal{F}_{n-1}) = \sigma^2 > 0 \) a.s. for all \( n \). We shall call an AR\((p)\) model \textit{purely explosive} if all roots of its characteristic polynomial \( \phi(z) \) lie outside the unit circle. In Section 2, we obtain certain almost sure asymptotic properties of purely explosive autoregressive models. The almost sure asymptotic behavior of non-explosive models is studied in Section 3. Applying the results of Sections 2 and 3, we obtain a simple proof of Theorem 1 on the strong consistency of \( b_n \) in general AR\((p)\) models in Section 4, where we also combine the results of Sections 2 and 3 to analyze the asymptotic behavior of general AR\((p)\) models.

In the engineering literature, there has been considerable interest in the question of strong consistency of the least-squares estimate \( b_n \), especially in view of its commonly used recursive form for system identification and control (cf. [4, 5, 11]). In this area, a recent attempt to prove the strong consistency of \( b_n \) in the general AR\((p)\) model is due to Graupe [5]. Assuming first that the matrix \( B \) defined in (1.3) is diagonable so that there exists a non-singular matrix \( T \) such that \( TBT^{-1} = \text{diag}(\lambda_1, \ldots, \lambda_p) \), where the \( \lambda_i \) are the eigenvalues of \( B \), Graupe noted that it suffices to show that

\[ R_{n}^{-1} \sum_{k=p+1}^{n} z_k \epsilon_k = (R_{n}^{-1} D_n) \left( D_{n}^{-1} \sum_{k=p+1}^{n} z_k \epsilon_k \right) \to \Omega \quad \text{a.s.} \quad (1.11) \]

where

\[ z_k = (z_{k1}, \ldots, z_{kp})' = T(y_{k-i}, \ldots, y_{k-p})', \]

\[ R_n = \sum_{\rho+1}^{n} z_k z_k' = T(X'_n X_n) T', \quad D_n = \text{diag} \left( \sum_{\rho+1}^{n} z_{k1}^2, \ldots, \sum_{\rho+1}^{n} z_{kp}^2 \right). \]
Since
\[
D_n^{-1} \sum_{p+1}^{n} z_k e_k = \left( \sum_{p+1}^{n} z_{k1} e_k, \ldots, \sum_{p+1}^{n} z_{kp} e_k \right) \rightarrow 0 \quad \text{a.s.}
\]
if
\[
\sum_{k=p+1}^{\infty} z_{kj}^2 = \infty \quad \text{a.s. for } j = 1, \ldots, p,
\]
(1.13)
Graupe's approach was to show that (1.13) holds and that
\[
\left\{ R_n^{-1} D_n \right\} \text{ is uniformly bounded with probability 1.}
\]
(1.14)
However, the proof of (1.14) given in [5] has some errors, and it will be shown in Section 3 that (1.14) is in fact false. Therefore Graupe's approach fails to prove the strong consistency of \( b_n \) in the general AR(\( p \)) model. Our approach to this problem is entirely different and makes use of certain basic asymptotic properties of the AR(\( p \)) model in the non-explosive and the purely explosive cases. Moreover, our objective herein is not only to solve the strong consistency problem, but also to establish these and other almost sure asymptotic properties of general autoregressive models.

2. Purely Explosive Autoregressive Models

In this section we study the almost sure asymptotic behavior of the AR(\( p \)) model (1.1) when all the roots of the characteristic polynomial (1.2) lie outside the unit circle. The main results are summarized in the following:

**Theorem 2.** Suppose that in the AR(\( p \)) model (1.1), \( \{e_n\} \) is a martingale difference sequence with respect to an increasing sequence of \( \sigma \)-fields \( \{\mathcal{F}_n\} \) such that (1.8) holds. Assume that the roots \( z_j \) of the characteristic polynomial \( p(z) \) as defined in (1.2) lie outside the unit circle, i.e., \( |z_j| > 1 \) for \( j = 1, \ldots, p \). Let \( M = \max_{1 \leq j \leq p} |z_{j}|, \) \( m = \min_{1 \leq j \leq p} |z_{j}|, \) and let \( B \) be the companion matrix defined in (1.3).

(i) Define \( Y_n, \varepsilon_n \) as in (1.4). Then
\[
B^{-n} Y_{n+p} \text{ converges a.s. to } Z = Y_p + \sum_{i=1}^{\infty} B^{-i} \varepsilon_{i+p}.
\]
Moreover,
\[
x'Z \text{ has a continuous distribution for all } x \in \mathbb{R}^p - \{0\}.
\]
(ii) Let $X_n$ be the design matrix defined in (1.7). Then

$$B^{-n}X_n'X_n(B^{-n})'$$ converges a.s. to \( F = \sum_{i=p+1}^{\infty} B^{-i}(ZZ')(B^{-i})' \). (2.3)

Moreover, \( F \) is positive definite with probability 1. Consequently,

$$\lim_{n \to \infty} n^{-1} \log \lambda_{\text{min}}(X_n'X_n) = 2 \log m \quad \text{a.s.,}$$

$$\lim_{n \to \infty} n^{-1} \log \lambda_{\text{max}}(X_n'X_n) = 2 \log M \quad \text{a.s.}$$ (2.4)

The proof of Theorem 2 makes use of certain ideas of [13] and the following lemma, whose proof is given in [9].

**Lemma 1.** Let \( \{\varepsilon_n\} \) be a martingale difference sequence with respect to an increasing sequence of \( \sigma \)-fields \( \{\mathcal{F}_n\} \) such that (1.8) holds. If \( y \) is \( \mathcal{F}_p \)-measurable and \( \{a_n\} \) is a sequence of real constants such that \( \sum_{n=p+1}^{\infty} a_n^2 < \infty \) and \( a_n \neq 0 \) for infinitely many \( n \), then \( y + \sum_{n=p+1}^{\infty} a_n \varepsilon_n \) has a continuous distribution.

In the sequel we shall let \( \|x\| \) denote the Euclidean norm of a \( p \)-dimensional vector \( x = (x_1, \ldots, x_p)' \), i.e., \( \|x\|^2 = x'x \). Moreover, viewing a \( p \times p \) matrix \( A \) as a linear operator, we define \( \|A\| = \sup_{\|x\|=1} \|Ax\| \). Thus, \( \|A\|^2 \) is equal to the maximum eigenvalue of \( A' A \) (cf. [14, p. 50]). In particular, for the companion matrix \( B \) of Theorem 2,

$$\log \|B^{-n}\| \sim \log \|(B')^{-n}\| \sim -n \log m,$$

$$\log \|B^n\| \sim \log \|(B')^n\| \sim n \log M,$$ (2.5)

(cf. [18, p. 65]). Moreover, if \( A \) is symmetric and non-negative definite, then \( \|A\| = \lambda_{\text{max}}(A) \).

**Proof of Theorem 2.** Let \( Z_n = Y_p + \sum_{i=1}^{n} B^{-i} \varepsilon_{i+p} \). By (1.5),

$$Y_{n+p} = B^n Y_p + \sum_{i=1}^{n} B^{n-i} \varepsilon_{i+p} = B^n Z_n.$$ (2.6)

The a.s. convergence of \( Z_n \) to \( Z \) follows from (2.5) and the martingale convergence theorem (cf. [10]). To prove (2.2), we apply Lemma 1 and it therefore suffices to show that for \( x \in \mathcal{R}^p - \{0\}, \)

$$x'B^{-i}(1, 0, \ldots, 0)' \neq 0 \quad \text{for infinitely many } i.$$ (2.7)

Let \( u = (1, 0, \ldots, 0)' \in \mathcal{R}^p \). The matrix \( \{u, Bu, \ldots, B^{n-1}u\} \) is upper triangular with diagonal elements all equal to 1 and is therefore non-singular. Since \( B \)
is non-singular, it then follows that for every \( k = 1, 2, \ldots \), the vectors \( B^{-k}u, B^{-k}Bu, \ldots, B^{-k}B^{p-1}u \) are linearly independent. Hence given \( x \in \mathbb{R}^p - \{0\} \), there exists \( j \in \{0, \ldots, p-1\} \) such that \( x'B^{-k+j}u \neq 0 \). This therefore proves (2.7).

To prove (2.3), we note that

\[
X'_n X_n = \sum_{i=p}^{n-1} Y'_i Y_i = \sum_{i=0}^{n-p-1} B^i Z_i(B^i)' \quad \text{by (2.6).} \tag{2.8}
\]

Let \( F_n = \sum_{i=p+1}^{n} B^{-i}Z_n Z'_n(B^{-i})' \). By (2.8),

\[
\|B^{-n}X'_nX_n(B^{-n})' - F_n\|
= \left\| \sum_{i=0}^{n-p-1} B^{-(n-i)}Z_i Z'_i(B^{-(n-i)})' - F_n \right\|
= \left\| \sum_{i=p+1}^{n} B^{-i}[Z_{n-i}Z'_{n-i} - Z_n Z'_n](B^{-i})' \right\|
\leq \sum_{i=p+1}^{n} \|B^{-i}\| \|(B^{-i})'\| (\|Z_{n-i}\| + \|Z_n\|) \|Z_n - Z_{n-i}\|
\to 0 \quad \text{a.s., by (2.5) and the a.s. convergence of } Z_n. \tag{2.9}
\]

Moreover, with probability 1,

\[
\|F_n - F\| \leq \sum_{i=n+1}^{\infty} \|B^{-i}ZZ'(B^{-i})'\| + \sum_{i=p+1}^{n} \|B^{-i}(Z_n Z'_n - ZZ')(B^{-i})'\|
\leq \|ZZ'\| \sum_{i=n+1}^{\infty} \|B^{-i}\| \|(B^{-i})'\|
+ \|Z_n Z'_n - ZZ'\| \sum_{i=p+1}^{\infty} \|B^{-i}\| \|(B^{-i})'\|
\to 0 \quad \text{a.s. by (2.5) and the a.s. convergence of } Z_n. \tag{2.10}
\]

From (2.9) and (2.10), (2.3) follows.

We now prove that \( F \) is positive definite with probability 1. Let \( \varphi_k(z) = \varphi(z)/(z - z_k) \) if \( z_k \) is real and \( \varphi_k(z) = \varphi(z)/((z - z_k)(z - z'_{-k})) \) otherwise, noting that the roots of the characteristic polynomial \( \varphi(z) \) exist in conjugate pairs. Since \( \varphi(z) \) is the minimal polynomial of the companion matrix \( B \), it is well known that the following implication holds for all \( w \in \mathbb{R}^p - \{0\} \):

\[ w, Bw, \ldots, B^{p-1}w \text{ are linearly dependent} \]

\[ \Rightarrow w \text{ belongs to the null space of } \varphi_k(B) \text{ for some } k = 1, \ldots, p \tag{2.11} \]
Let $N_k$ denote the null space of $\varphi_k(B)$. Then $\dim N_k < p$, so there exists $x \in \mathbb{R}^p - \{0\}$ such that $x'y = 0$ for all $y \in N_k$. Since $P[x'Z = 0] = 0$ by (2.2), it then follows that $P[Z \in N_k] = 0$ for all $k = 1, \ldots, p$. This and (2.11) in turn establish that

$$V \overset{\text{def}}{=} (Z, BZ, \ldots, B^{p-1}Z)$$

is nonsingular a.s. \hfill (2.12)

Noting that

$$V V' = \sum_{i=0}^{p-1} B^i Z Z'(B^i)' = B^2 P \sum_{i=p+1}^{2p} B^{-i} Z Z'(B^{-i})' (B')^{2p},$$

we obtain from (2.12) the a.s. positive definiteness of $VV'$ and therefore also of $F = \sum_{i=p+1}^{\infty} B^{-i} Z Z'(B^{-i})'$.

Let $U_n = B^{-n} X_n' X_n (B^{-n})'$. By (2.3), with probability 1, the symmetric matrix $U_n$ converges to a positive definite matrix, and it therefore follows that

$$\lim_{n \to \infty} \lambda_{\max}(U_n) < \infty \quad \text{a.s.}, \quad \lim_{n \to \infty} \lambda_{\min}(U_n) > 0 \quad \text{a.s.} \hfill (2.13)$$

Since $X_n' X_n = B^n U_n (B^n)'$ and $(X_n' X_n)^{-1} = (B^{-n})' U_n^{-1} B^{-n}$, the desired conclusion (2.4) follows from (2.5), (2.13), and Lemma 2 below, noting that $\lambda_{\min}(X_n' X_n) = 1/\lambda_{\max}((X_n' X_n)^{-1})$.

**Lemma 2.** Let $A, C$ be $p \times p$ matrices such that $C$ is symmetric and non-negative definite. Then

$$\lambda_{\max}(C) \lambda_{\max}(AA') \geq \lambda_{\max}(ACA') \geq \lambda_{\min}(C) \lambda_{\max}(AA'). \hfill (2.14)$$

**Proof.** To prove (2.14), we note that $\lambda_{\max}(ACA')$ is equal to

$$\sup_{\|x\| = 1} x' ACA' x \leq \lambda_{\max}(C) \sup_{\|x\| = 1} \|A' x\|^2 = \lambda_{\max}(C) \lambda_{\max}(AA').$$

On the other hand,

$$\sup_{\|x\| = 1} x' ACA' x \geq \sup_{\|x\| = 1} \{ \lambda_{\min}(C) \|A' x\|^2 \} = \lambda_{\min}(C) \lambda_{\max}(AA'). \hfill \square$$

The following corollary of Theorem 2 will be used in the sequel.

**Corollary 1.** Under the same assumptions and notations as in Theorem 2.

(i) \hfill \lim_{n \to \infty} \sum_{i=p}^{n-1} \|B^{-n} Y_i\| = \sum_{i=p+1}^{\infty} \|B^{-i} Z\| < \infty \quad \text{a.s.,}
(ii) \[ \lim_{n \to \infty} Y'_{n-k} \left( \sum_{i=0}^{n} Y_i Y_i' \right)^{-1} Y_{n-k} = (B^{-p-k-1}Z)' F^{-1}(B^{-p-k-1}Z) > 0 \text{ a.s. for } k = 0, \pm 1, \ldots, \]

(iii) \[ \lim_{n \to \infty} B^{-n}(Y_{n+p}, Y_{n+p+1}, \ldots, Y_{n+2p-1}) = (Z, BZ, \ldots, B^{p-1}Z) \text{ a.s.,} \]

(iv) \[ \lim_{n \to \infty} n^{-1} \log \| Y_n \| = \log M \text{ a.s.} \]

**Proof.** Since \( \sum_{i=0}^{n-1} B^{-n}Y_i \| = \sum_{i=0}^{n-1} B^{-n-i}B^{-i}Y_{i+p} \| \) (i) follows from (2.1) and (2.5). Obviously (iii) also follows from (2.1). To prove (ii), we note that

\[ Y'_{n-k} \left( \sum_{i=0}^{n} Y_i Y_i' \right)^{-1} Y_{n-k} = Y'_{n-k}(X'_{n+1}X_{n+1})^{-1} Y_{n-k} = (B^{-n}Y_{n-k})' \{ (B^{n+1})'(X'_{n+1}X_{n+1})^{-1} B^{n+1} \} B^{-n-1}Y_{n-k}, \]

and apply (2.1), (2.2), (2.3), and the a.s. positive definiteness of \( F \).

We now prove (iv). By (iii) and (2.12), with probability 1,

\[ B^{-n}(Y_{n+p}, \ldots, Y_{n+2p-1})(Y_{n+p}, \ldots, Y_{n+2p-1})' (B^{-n})' \text{ converges to a symmetric positive definite matrix.} \] (2.15)

Hence, by (2.5) and Lemma 2,

\[ \log \lambda_{\max}\{ (Y_{n+p}, \ldots, Y_{n+2p-1})(Y_{n+p}, \ldots, Y_{n+2p-1})' \} \sim 2n \log M \text{ a.s.} \] (2.16)

From (2.16) and the inequality

\[ \lambda_{\max}(C) \leq \text{tr}(C) \leq p\lambda_{\max}(C) \] (2.17)

for every symmetric nonnegative definite matrix \( C \), it follows that

\[ \lim_{n \to \infty} n^{-1} \log \left( \sum_{i=0}^{p-1} ||Y_{n+p+i}||^2 \right) = 2 \log M \text{ a.s.} \] (2.18)

By (1.5), \( ||Y_{n+p+i}|| \leq ||B|| ||Y_{n+p}|| + ||\epsilon_{n+p+1}|| \), and more generally for \( i = 1, \ldots, p-1, \)

\[ ||Y_{n+p+i}|| \leq ||B||^i ||Y_{n+p}|| + \sum_{j=1}^{i} ||B||^j |\epsilon_{n+p+j}|. \] (2.19)

Since \( \epsilon_n = o(n^{1/2}) \) a.s. (cf. [7]), (iv) follows from (2.18) and (2.19).
3. Non-Explosive Autoregressive Models

In this section we study the asymptotic behavior of non-explosive AR\(p\) models, i.e., the roots of the characteristic polynomial (1.2) lie on or inside the unit circle. Unlike the exponential rate of growth for \(\| Y_n \|, \lambda_{\text{max}}(X'_n X_n), \lambda_{\text{min}}(X'_n X_n)\) in Theorem 2 and Corollary 1 for the purely explosive case, the following theorem shows that these quantities grow at most algebraically fast in non-explosive models.

**Theorem 3.** Suppose that in the AR\(p\) model (1.1), \(\{ \epsilon_n \}\) is a martingale difference sequence with respect to an increasing sequence of \(\sigma\)-fields \(\mathcal{F}_n\) such that (1.8) holds. Assume that the roots \(z_j\) of the characteristic polynomial \(\varphi(z)\) as defined in (1.2) lie on or inside the unit circle, i.e., \(|z_j| < 1\) for \(j = 1, \ldots, p\).

(i) Define \(Y_n\) as in (1.4). Let \(\rho = 0\) if all roots of \(\varphi(z)\) lie inside the unit circle, otherwise let \(\rho\) be the largest multiplicity of all the distinct roots on the unit circle. Then

\[
\| Y_n \| = o(n^{1/2}) \quad \text{a.s. if } \rho = 0, \quad (3.1)
\]

\[
= O(n^{\rho - 1/2} (\log \log n)^{1/2}) \quad \text{a.s. if } \rho \geq 1.
\]

(ii) Let \(X_n\) be the design matrix defined in (1.7). Then

\[
\liminf_{n \to \infty} n^{-1} \lambda_{\text{min}}(X'_n X_n) > 0 \quad \text{a.s.} \quad (3.2)
\]

Moreover,

\[
\lambda_{\text{max}}(X'_n X_n) = O(n) \quad \text{a.s. if } \rho = 0,
\]

\[
= O(n^{2\rho} (\log \log n)) \quad \text{a.s. if } \rho \geq 1. \quad (3.3)
\]

Proof. The proofs of (3.1) and (3.2) are given in [10, Theorem 2] and [8, Corollary 1], respectively, in a somewhat more general context. Since \(\lambda_{\text{max}}(X'_n X_n) \leq \text{tr}(\sum_{i=1}^{n-1} Y_i Y'_i) = \sum_{i=1}^{n-1} \| Y_i \|^2\), (3.3) follows from (3.1) in the case \(\rho \geq 1\). For the case \(\rho = 0\), (3.3) follows from Corollary 2 of [8].

The following example illustrates some of the results of Theorem 3 and shows why the approach of Graupe [5] outlined in Section 1 fails to prove the strong consistency of the least-squares estimate \(b_n\).

**Example 1.** Consider the AR(2) model

\[
y_n = \beta_1 y_{n-1} + \beta_2 y_{n-2} + \epsilon_n, \quad n \geq 1 \quad (y_0 = y_{-1} = 0), \quad (3.4)
\]

where \(\epsilon_1, \epsilon_7, \ldots\) are i.i.d. random variables with mean 0 and variance \(\sigma^2 > 0\).
The least-squares estimate $b_n$ of $(\beta_1, \beta_2)'$ is given by (1.6) with

$$
X_n'X_n = \begin{pmatrix}
\sum_{i=2}^{n-1} y_i^2 & \sum_{i=2}^{n-1} y_i y_{i-1} \\
\sum_{i=2}^{n-1} y_i y_{i-1} & \sum_{i=2}^{n-1} y_{i-1}^2
\end{pmatrix}.
$$

(3.5)

Assume that $\beta_1 = 1$ and $\beta_2 = 0$. Then $y_n = S_n$, where $S_n = \sum_1^n e_i$, and the characteristic polynomial $\phi(z) = z^2 - z$ has 1 and 0 as its roots. Moreover, by the law of the iterated logarithm,

$$
\limsup_{n \to \infty} |y_n|/(2n \log \log n)^{1/2} = \sigma \quad \text{a.s.},
$$

(3.6)

providing an example of the log log behavior in (3.1) with $\rho = 1$. As shown in Example 4 of [8],

$$
\lambda_{\min}(X_n'X_n) \sim \frac{1}{2} \sum_{i=1}^{n} e_i^2 \sim \frac{1}{2} n\sigma^2 \quad \text{a.s.}, \quad \lambda_{\max}(X_n'X_n) \sim 2 \sum_{i=1}^{n} S_i^2 \quad \text{a.s.},
$$

(3.7)

$$
\limsup_{n \to \infty} \left( \sum_{i=1}^{n} S_i^2 \right) \left( n^2 \log \log n \right) = 8\sigma^2/\pi^2 \quad \text{a.s.},
$$

$$
\liminf_{n \to \infty} \left( \sum_{i=1}^{n} S_i^2 \right) \left( n^2/\log \log n \right) - \sigma^2/4 \quad \text{a.s.}
$$

Since $\beta_1 = 1$ and $\beta_2 = 0$, the companion matrix $B$ in (1.3) reduces to

$$
B = \begin{pmatrix}
1 & 0 \\
1 & 0
\end{pmatrix}
$$

(3.8)

and is diagonalizable, i.e., $TB_{B^{-1}} = \text{diag}(1, 0)$, where

$$
T = \begin{pmatrix}
1 & 0 \\
1 & -1
\end{pmatrix}.
$$

(3.9)

Letting $z_k = (z_{k1}, z_{k2})' = T(y_{k-1}, y_{k-2})'$, it follows from (3.9) that

$$
z_k = (y_{k-1}, y_{k-1} - y_{k-2})' = (S_{k-1}, e_{k-1})'.
$$

(3.10)

Hence $D_n = \text{def} \ \text{diag}(\sum_{k=1}^{n} z_{k1}^2, \sum_{k=1}^{n} z_{k2}^2) = \text{diag}(\sum_{i=2}^{n-1} S_i^2, \sum_{i=3}^{n-1} e_i^2)$, and the diagonal elements of $D_n$ diverge to \infty a.s., i.e., (1.13) holds for the present example.

As described in Section 1, the approach of Graupe [5] to establish the strong consistency of $b_n$ consists of proving (1.13) and (1.14). We now show
that for the present example (1.14) fails to hold. Letting $R_n = T(X'_n X_n) T'$, it follows from (3.7) and (3.9) that

$$\det R_n = \det(X'_n X_n) = \lambda_{\max}(X'_n X_n) \lambda_{\min}(X'_n X_n) \sim \left(\sum_{i=1}^{n} \varepsilon_i^2\right) \left(\sum_{i=1}^{n} S_i^2\right) \text{ a.s.} \quad (3.11)$$

We note that

$$R_n^{-1} D_n = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \sum_{i=2}^{n-1} S_i^2 & \sum_{i=1}^{n-1} S_i S_{i-1} \\ \sum_{i=2}^{n-1} S_i S_{i-1} & \sum_{i=2}^{n-1} S_{i-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} D_n$$

$$= \left(\sum_{i=2}^{n-1} S_i^2 \sum_{i=1}^{n-1} S_i \varepsilon_i - \sum_{i=2}^{n-1} \varepsilon_i^2 S_i \varepsilon_i \right) D_n$$

$$= \left(\sum_{i=2}^{n-1} S_i^2 \sum_{i=1}^{n-1} S_i \varepsilon_i - \sum_{i=2}^{n-1} \varepsilon_i^2 S_i \varepsilon_i \right) \begin{pmatrix} \sum_{i=2}^{n-1} S_i^2 / \det R_n & 0 \\ 0 & \sum_{i=2}^{n-1} \varepsilon_i^2 / \det R_n \end{pmatrix}.$$ (3.12)

By (3.11), with probability 1,

$$\left(\sum_{i=2}^{n-1} S_i \varepsilon_i \right) \left(\sum_{i=2}^{n-1} S_i^2 / \det R_n \right) \sim \left(\sum_{i=1}^{n-1} S_i \varepsilon_i \right) \left(\sum_{i=1}^{n} \varepsilon_i^2 \right) \sim \frac{1}{2} \left\{ \sum_{i=2}^{n-1} (S_i^2 - S_{i-1}^2 + \varepsilon_i^2) \right\} \left(\sum_{i=1}^{n} \varepsilon_i^2 \right) \sim \frac{1}{2} + \frac{1}{2} S_{n-1}^2 / n \sigma^2,$$ by (3.7). (3.13)

Since $\lim \sup_{n \to \infty} S_n^2 / n = \infty$ a.s. by the law of the iterated logarithm, (3.12) and (3.13) show that $P[\{R_n^{-1} D_n\} \text{ is uniformly bounded}] = 0$, and therefore (1.14) fails to hold for the present example.

While Corollary 1(ii) shows that for fixed $k = 0, 1, \ldots, Y_{n-k} \left(\sum_{i=p}^{n} Y_i Y_i'\right)^{-1} Y_{n-k}$ converges a.s. to a positive limit in the purely explosive case, the following theorem says that $Y_{n-k} \left(\sum_{i=p}^{n} Y_i Y_i'\right)^{-1} Y_{n-k}$ converges a.s. to 0 in non-explosive models. This result plays an important
role in the proof of Theorem 1 on the strong consistency of $b_n$ in general AR($p$) models.

**Theorem 4.** Under the same assumptions and notations as in Theorem 3,

$$\lim_{n \to \infty} \max_{1 \leq j \leq n} Y_j' \left( \sum_{i=1}^{n} Y_i Y_i' \right)^{-1} Y_j = 0 \quad a.s. \quad (3.14)$$

We preface the proof of Theorem 4 by the following five lemmas, some of which will also be used in the proof of Theorem 1.

**Lemma 3.** Let $\{a_n\}$ be a sequence of nonnegative numbers such that

$$\sum_{i=1}^{n} a_i = o(n^\delta) \quad \text{for all } \delta > 0, \quad (3.15)$$

and there exist $C > 0$ and $\gamma > 0$ such that

$$a_{n+1} \leq a_n + Cn^{-\gamma} \quad \text{for all large } n. \quad (3.16)$$

Then $\lim_{n \to \infty} a_n = 0$.

**Proof.** Condition (3.16) implies that for every $0 < \rho < 1$,

$$\min_{n > j > n} a_j \geq a_n - 2Cn^{-\gamma} \quad \text{for all large } n. \quad (3.17)$$

In particular, choosing $0 < \rho < \min(1, \gamma/2)$, we obtain from (3.17) that

$$\sum_{j=1}^{n} a_j \geq n^\rho(a_n - 2Cn^{-\gamma}) \geq n^\rho a_n - 2C \quad \text{for all large } n. \quad (3.18)$$

From (3.15) and (3.18), it follows that $\lim_{n \to \infty} a_n = 0$. □

**Lemma 4.** Let $g_1, \ldots, g_r, h_1, \ldots, h_s$ be real numbers and let $p = r + s$. Define the $s \times p$, $r \times p$ and $p \times p$ matrices $M_1$, $M_2$, $M$ by

$$M_1 = \begin{pmatrix} 1 & g_1 & \cdots & g_r & 0 & \cdots & 0 \\ 0 & 1 & g_1 & \cdots & g_r & 0 & \cdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & g_1 & \cdots & g_r \end{pmatrix},$$

$$M_2 = \begin{pmatrix} 1 & h_1 & \cdots & h_s & 0 & \cdots & 0 \\ 0 & 1 & h_1 & \cdots & h_s & 0 & \cdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & h_1 & \cdots & h_s \end{pmatrix}, \quad M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}. \quad (3.19)$$
Define the polynomials

\[ P_1(z) = z^r + g_1 z^{r-1} + \cdots + g_r, \quad P_2(z) = z^s + h_1 z^{s-1} + \cdots + h_s. \quad (3.20) \]

(i) If \( P_1, P_2 \) are relatively prime (over the real field), then \( M \) is nonsingular.

(ii) Let \( \phi(z) = P_1(z) P_2(z) = z^n - \beta_1 z^{n-1} - \cdots - \beta_p \). For a given sequence of real numbers \( \{e_n\} \) and initial values \( y_0, \ldots, y_{p-1} \), define \( y_n = \beta_1 y_{n-1} + \cdots + \beta_p y_{n-p} + e_n \), \( n \geq 1 \). Moreover, define

\[ u_n = y_n + g_1 y_{n-1} + \cdots + g_r y_{n-r}, \quad v_n = y_n + h_1 y_{n-1} + \cdots + h_s y_{n-s}. \quad (3.21) \]

Then for \( n \geq 1 \),

\[ u_n + h_1 u_{n-1} + \cdots + h_s u_{n-s} = e_n = v_n + g_1 v_{n-1} + \cdots + g_r v_{n-r}. \quad (3.22) \]

**Proof.** (i) follows from the well-known theory of Sylvester’s determinants (cf. [3]). To prove (ii), note that \( L^p \phi(L^{-1}) y_n = e_n \), where \( L \) denotes the unit delay operator (i.e., \( Ly_n = y_{n-1} \)). Hence \( e_n = L^p P_1(L^{-1}) \{L^s P_2(L^{-1}) y_n\} = L^p P_1(L^{-1}) u_n \) by \( (3.21) \). Likewise \( e_n = L^s P_2(L^{-1}) \{L^s P_1(L^{-1}) y_n\} = L^s P_2(L^{-1}) u_n \).

**Lemma 5.** Let \( A \) be a \( p \times p \) symmetric positive definite matrix.

(i) If \( A^{-1} = I_p + V + W \), where \( V, W \) are symmetric \( p \times p \) matrices such that \( V \) is nonnegative definite and \( \|W\| < 1 \), then

\[ \|A\| \leq 1/(1 - \|W\|). \quad (3.23) \]

(ii) If \( A \) is partitioned as

\[ A = \begin{pmatrix} P & H \\ H' & Q \end{pmatrix}, \]

where \( P, Q \) are respectively \( r \times r \) and \( s \times s \) matrices such that \( p = r + s \), then for \( u \in \mathbb{R}^r \),

\[ \left( \begin{array}{c} u \\ 0 \end{array} \right)^T A^{-1} \left( \begin{array}{c} u \\ 0 \end{array} \right) \leq u^T P^{-1} u (1 + \|A^{-1}\| \text{tr}(Q)). \quad (3.24) \]

**Proof.** To prove (i), note that

\[ \lambda_{\min}(A^{-1}) = \inf_{\|x\|=1} x^T A^{-1} x = \inf_{\|x\|=1} x^T x - \sup_{\|x\|=1} |x^T W x| \geq 1 - \|W\|. \]

Since \( \|A\| = \lambda_{\max}(A) = 1/\lambda_{\min}(A^{-1}) \), \( (3.23) \) follows.
To prove (ii), note that
\[
A^{-1} = \begin{pmatrix} P^{-1} + H'P^{-1}H & -P^{-1}H \\ -H'P^{-1} & 1 \end{pmatrix},
\]
where
\[
\Gamma^{-1} = Q - H'P^{-1}H \text{ is positive definite}
\]
(cf. [14, p. 29]). By (3.25), \(\|\Gamma\| \leq \|A^{-1}\|\), and by (3.26), \(\text{tr}(H'P^{-1}H) \leq \text{tr} Q\).
Therefore
\[
u'P^{-1/2}HH'P^{-1/2}u \leq \|P^{-1/2}u\|^2 \|P^{-1/2}HH'P^{-1/2}x\|
\leq \|P^{-1/2}u\|^2 \|\Gamma\| \sup_{\|x\|=1} x'P^{-1/2}HH'P^{-1/2}x
\leq (u'P^{-1}u) \|\Gamma\| \text{tr}(H'P^{-1}H)
\leq (u'P^{-1}u) \|A^{-1}\| \|A^{-1}\| \|\Gamma\|
\]
(3.27)
From (3.25) and (3.27), (3.24) follows.

Lemma 6. With the same notations and assumptions as in Theorem 3, let \(C_n = \sum_{i=p}^n Y_i Y_i' = X_{n+p} X_{n+p}'\). Let \(N = \inf\{n \geq p: C_n \text{ in nonsingular}\}\) (inf \(\phi = \infty\)). Then

(i) \(N < \infty\) a.s. and \(\|C_n^{-1/2}\| = O(n^{-1/2})\) a.s.,

(ii) \(Y_i' C_n^{-1} Y_i \leq 1\) for \(n \geq N\) and
\[
\sum_{i=N}^n Y_i' C_n^{-1} Y_i = O((\log n)^{1/2}) \text{ a.s.,}
\]
(3.28)

(iii) \(\|C_n^{-1/2} \sum_{i=p}^n Y_i \epsilon_{i+1}\| = O((\log n)^{1/2})\) a.s.

Proof. By (3.2), \(\lim inf_{n \to \infty} n^{-1} \lambda_{\min}(C_n) > 0\) a.s. Since \(\|C_n^{-1/2}\|^2 = \lambda_{\max}(C_n^{-1}) = 1/\lambda_{\min}(C_n)\), (i) follows. By Lemma 2 of [7], \(Y_i' C_n^{-1} Y_i \leq 1\) for \(n \geq N\) and \(\sum_{i=N}^n Y_i' C_n^{-1} Y_i = O(\log \lambda_{\max}(C_n))\) a.s. Since \(\lambda_{\max}(C_n) = O(\log n)\) a.s. by (3.3), (3.28) follows. By Lemma 1 of [7],
\[
\|C_n^{-1/2} \sum_{i=p}^n Y_i \epsilon_{i+1}\|^2 = \left( \sum_{i=p}^n Y_i \epsilon_{i+1} \right)' C_n^{-1} \left( \sum_{i=p}^n Y_i \epsilon_{i+1} \right)
= O(\log \lambda_{\max}(C_n)) \text{ a.s.,}
\]
and therefore (iii) follows.

Lemma 7. With the same notations and assumptions as in Theorem 3,
assume furthermore that $B$ is non-singular. Define $C_n$ and $N$ as in Lemma 6. Then

(i) $\| C_n^{1/2} B' C_n^{-1} B C_n^{-1/2} \| \leq 1 + O(n^{-1/2} \log n)^{1/2})$ a.s.

(ii) Let $\rho > 1/\alpha$, where $\alpha$ is as given by (1.8). Then

$$\limsup_{n \to \infty} n^{1/2-\rho} (Y_{n+1} C_n^{-1} Y_{n+1} - Y_n C_n^{-1} Y_n) \leq 0 \quad \text{a.s.} \quad (3.29)$$

Proof. To prove (i), we apply Lemma 5(i) to $A_n = C_n^{1/2} B' C_n^{-1} B C_n^{1/2}$. Defining $e_n$ as in (1.4), we obtain by (1.5) that

$$C_{n+1} = \sum_{i=p-1}^{n} Y_{i+1} Y_{i+1}' = \sum_{i=p-1}^{n} (B Y_i + e_{i+1})(B Y_i + e_{i+1})'$$

$$= B(C_n + Y_{p-1} Y_{p-1}') B' + B \sum_{p-1}^{n} Y_i e_i' + \sum_{p-1}^{n} e_i Y_i B' + \sum_{p-1}^{n} e_i e_i'.$$

Therefore

$$A_n^{-1} = C_n^{-1/2} B^{-1} C_n^{1/2} (B')^{-1} C_n^{-1/2} = I_p + V_n + \tilde{W}_n + \tilde{\tilde{W}}_n, \quad (3.30)$$

where

$$V_n = C_n^{-1/2} \left( Y_{p-1} Y_{p-1}' + B^{-1} \left( \sum_{i=p}^{n+1} e_i e_i' \right)(B')^{-1} \right) C_n^{-1/2}$$

is nonnegative definite,

$$\tilde{W}_n = C_n^{-1/2} \left( \sum_{p-1}^{n} Y_i e_i' \right)(B')^{-1} C_n^{-1/2},$$

$$\tilde{\tilde{W}}_n = C_n^{-1/2} B^{-1} \left( \sum_{p-1}^{n} e_i Y_i' \right) C_n^{-1/2}.$$

Since $\sum_{i=p-1}^{n} Y_i e_i' = \sum_{i=p-1}^{n} Y_i (e_i', 0, ..., 0)$,

$$\| \tilde{\tilde{W}}_n \| \leq \| (B')^{-1} \| \| C_n^{-1/2} \| \| C_n^{-1/2} \sum_{i=p-1}^{n} Y_i e_i' \|$$

$$= O(n^{-1/2} \log n)^{1/2}) \quad \text{a.s.,} \quad (3.31)$$

by Lemma 6(i,iii). Likewise

$$\| \tilde{W}_n \| = O(n^{-1/2} \log n)^{1/2}) \quad \text{a.s.} \quad (3.32)$$
In view of (3.30), (3.31), (3.32), and the fact that $\tilde{W}_n + \tilde{W}_n$ is symmetric, we can apply Lemma 5(i) to obtain that $\|A_n\| \leq 1 + O(n^{-1/2}(\log n)^{1/2})$ a.s.

To prove (ii), we obtain by (1.5) that for $n \geq N$

$$Y_n' C_{n+1}^{-1} Y_{n+1}$$

$$= (B Y_n + e_n+1)' C_{n+1}^{-1} (B Y_n + e_n+1)$$

$$= Y_n' B' C_{n+1}^{-1} B Y_n + Y_n' C_{n+1}^{-1} e_n+1 + e_n+1 C_{n+1}^{-1} Y_{n+1} - e_n+1 C_{n+1}^{-1} e_n+1$$

$$\leq (C_{n+1}^{-1/2} Y_n)' C_{n+1}^{1/2} B' C_{n+1}^{-1} B C_{n+1}^{1/2} (C_{n+1}^{-1/2} Y_n) \leq 2 (C_{n+1}^{-1/2} Y_n)' (C_{n+1}^{-1/2} e_n+1)$$

$$\leq \left( 1 + O(n^{-1/2}(\log n)^{1/2}) \right) \|C_{n+1}^{-1/2} Y_n\|^2$$

$$+ 2 \|C_{n+1}^{-1/2} Y_{n+1}\| \|C_{n+1}^{-1/2}\| \|e_{n+1}\| \quad \text{a.s.}$$  \hspace{1cm} (3.33)

by (i). Since $\rho > 1/\alpha$, $\|e_{n+1}\| = \|e_{n+1}\| = O(n^\alpha)$ a.s. (cf. [10]). Moreover, $\|C_{n+1}^{-1/2}\| = O(n^{-1/2})$ a.s. and $\|C_{n+1}^{-1/2} Y_n\|^2 = Y_n' C_{n+1}^{-1} Y_n \leq 1$ by Lemma 6. Therefore (3.33) implies that

$$Y_n' C_{n+1}^{-1} Y_{n+1} \leq Y_n' C_{n+1}^{-1} Y_n + O(n^{-1/2+\rho}) \quad \text{a.s.}$$  \hspace{1cm} (3.34)

or, equivalently, that (3.29) holds.

**Proof of Theorem 4.** Define $C_n$ and $N$ as in Lemma 6. For $j \geq N$, $C_j^{-1} - C_{j+1}^{-1}$ is non-negative definite (cf. (1.4b) of [7]), and therefore

$$Y_j' C_j^{-1} Y_j \geq Y_j' C_n^{-1} Y_j \quad \text{for} \quad n \geq j \geq N.$$  \hspace{1cm} (3.35)

Since $\lim_{n \to \infty} Y_j' C_n^{-1} Y_j = 0$ a.s. for every fixed $j$ by Lemma 6(i) and since (3.35) holds, it suffices for the proof of (3.14) to show that

$$\lim_{j \to \infty} Y_j' C_j^{-1} Y_j = 0 \quad \text{a.s.}$$  \hspace{1cm} (3.36)

First consider the case where $B$ is non-singular. Then by Lemma 7(ii), (3.34) holds with $\rho < \frac{1}{2}$. Moreover, by Lemma 6, (3.28) also holds. Hence we can apply Lemma 3 to obtain (3.36).

Now consider the case where $B$ is singular. Then 0 is a root of the characteristic polynomial $\varphi(z)$ in (1.2), so

$$\varphi(z) = z^r (z^s - \beta_1 z^{s-1} - \cdots - \beta_s), \quad \beta_s \neq 0, \quad \beta_{s+1} = \cdots = \beta_p = 0,$$  \hspace{1cm} (3.37)

where $r(\geq 1)$ is the multiplicity of the root 0 and $s = p - r$. First assume that $r < p$. Letting $g_1 = \cdots = g_r = 0$ and $h_1 = -\beta_1, \ldots, h_s = -\beta_s$ in Lemma 4 and defining $M$ as in (3.19), we obtain from (3.37) and Lemma 4(i) that $M$ is nonsingular. Moreover, by Lemma 4(ii),

$$MY_n = \begin{pmatrix} U_n \\ V_n \end{pmatrix}.$$  \hspace{1cm} (3.38)
where \( V_n = (e_n, \ldots, e_{n-r+1})' \), and \( U_n = (y_n, \ldots, y_{n-s+1})' \) satisfies

\[
U_n = B_1 U_{n-1} + (e_n, 0, \ldots, 0)'.
\] (3.39)

The matrix \( B_1 \) in (3.39) is given by

\[
B_1 = \begin{pmatrix}
\beta_1 & \cdots & \beta_s \\
I_{s-1} & 0 
\end{pmatrix}
\]

and is non-singular. Therefore, as shown in the preceding paragraph,

\[
U_n \left( \sum_{i=p}^n U_i U_i' \right)^{-1} U_n \to 0 \quad \text{a.s.} \] (3.40)

Letting \( A_n = MC_n M' = \sum_{i=p}^n M Y_i Y_i' M' \), we obtain from (3.38) that

\[
A_n = \begin{pmatrix}
\sum_{i=p}^n U_i U_i' & \sum_{i=p}^n U_i V_i' \\
\sum_{i=p}^n V_i U_i' & \sum_{i=p}^n V_i V_i'
\end{pmatrix}
\] (3.41)

From (3.38), it also follows that

\[
Y_n C_n^{-1} Y_n = (M Y_n)' A_n^{-1} (M Y_n)
\]

\[
= \left( \begin{array}{cc}
U_n' & 0 \\
0 & V_n'
\end{array} \right) A_n^{-1} \left( \begin{array}{c}
0 \\
V_n'
\end{array} \right) + 2 \left( \begin{array}{c}
U_n' \\
V_n'
\end{array} \right) A_n^{-1} \left( \begin{array}{c}
0 \\
V_n'
\end{array} \right)
\]

\[
= I_{n1} + I_{n2} + 2I_{n3}, \quad \text{say.} \] (3.42)

Since \( M \) is non-singular and \( \|C_n^{-1}\| = O(n^{-1}) \) a.s. by Lemma 6(i),

\[
\|A_n^{-1}\| \leq \|M^{-1}\| \|C_n^{-1}\| \|(M')^{-1}\| = O(n^{-1}) \quad \text{a.s.} \] (3.43)

Since \( V_n = (e_n, \ldots, e_{n-r+1})' \) and (1.8) holds,

\[
\|V_n\| = o(n^{1/2}) \quad \text{a.s.,} \quad \text{tr} \left( \sum_{i=p}^n V_i V_i' \right) = \sum_{i=p}^n \|V_i\|^2 = O(n) \quad \text{a.s.}
\] (3.44)

(cf. [7]). From (3.43) and (3.44),

\[
0 \leq I_{n2} \leq \|A_n^{-1}\| \|V_n\|^2 = O(n^{-1}) o(n) = o(1) \quad \text{a.s.} \] (3.45)
In view of (3.41), we can apply Lemma 5(ii) to obtain that

\[ 0 \leq I_{n1} \leq U_n' \left( \sum_{i=p}^{n} U_i U_i' \right)^{-1} U_n \left\{ 1 + \| A_n^{-1} \| \text{tr} \left( \sum_{i=p}^{n} V_i V_i' \right) \right\} \rightarrow 0 \quad \text{a.s. by (3.40), (3.43), and (3.44).} \]  

(3.46)

By the Schwarz inequality,

\[ |I_{n3}| \leq I_{n1}^{1/2} I_{n2}^{1/2}. \]  

(3.47)

From (3.42), (3.45), (3.46), and (3.47), the desired conclusion (3.36) follows.

For the case \( r = p \) in (3.37), \( Y_n = V_n = (\epsilon_n, \ldots, \epsilon_{n-p+1})' \) and therefore as shown in (3.45),

\[ Y_n' C_n^{-1} Y_n \leq \| C_n^{-1} \| \| V_n \|^2 = O(n^{-1}) o(n) = o(1) \quad \text{a.s.} \]

4. General AR(p) Models and Proof of Theorem 1

For the AR(p) model \( \{ y_n \} \) defined in (1.1) we have considered in Sections 2 and 3 the purely explosive case in which all roots of the characteristic polynomial \( \phi(z) \) lie outside the unit circle, and the non-explosive case in which all roots of \( \phi(z) \) lie on or inside the unit circle. The only remaining case is therefore the mixed model in which \( \phi(z) \) can be factorized as

\[ \phi(z) = P_1(z) P_2(z), \]  

(4.1)

where

\[ P_1(z) = z^r + g_1 z^{r-1} + \cdots + g_r, \quad P_2(z) = z^s + h_1 z^{s-1} + \cdots + h_s, \]

where \( r \geq 1, s \geq 1 \), and all roots of \( P_1(z) \) lie on or inside the unit circle, while all roots of \( P_2(z) \) lie outside the unit circle. Given the coefficients \( g_1, \ldots, g_r, h_1, \ldots, h_s \) of (4.1), define the matrix \( M \) by (3.19). By Lemma 4(i), \( M \) is non-singular. Define \( Y_n \) as in (1.5). By Lemma 4(ii),

\[ MY_n = \begin{pmatrix} U_n \\ V_n \end{pmatrix}, \]  

(4.2)

where \( U_n = (u_n, \ldots, u_{n-s+1})' \), \( V_n = (v_n, \ldots, v_{n-r+1})' \), and

\[ L^s P_2(L^{-1}) u_n = \epsilon_n, \quad L^r P_1(L^{-1}) v_n = \epsilon_n, \]  

(4.3)

\( L \) being the unit delay operator (i.e., \( Lu_n = u_{n-1} \)). Thus, the linear transfor-
mation (4.2) decomposes \( \{y_n\} \) into two autoregressive processes \( \{u_n\} \) and \( \{v_n\} \) such that \( \{u_n\} \) is a purely explosive AR(s) process and \( \{v_n\} \) is a non-explosive AR(r) process in view of (4.3). The results of Sections 2 and 3 can then be applied to \( \{u_n\} \) and \( \{v_n\} \) and therefore in turn also imply certain asymptotic properties of the mixed model \( \{y_n\} \). We first use this technique in the following proof of Theorem 1 on the strong consistency of the least-squares estimate \( b_n \) in general AR(p) models.

**Proof of Theorem 1.** Defining \( Y_n \) as in (1.5), we note that

\[
 b_n - \beta = (X_n'X_n)^{-1} X_n'(e_{p+1}, \ldots, e_n)' = \left( \sum_{i=p}^{n-1} Y_i Y_i' \right)^{-1} \sum_{i=p}^{n-1} Y_i e_{i+1}. \tag{4.4}
\]

First consider the case of a mixed model and introduce the linear transformation (4.2). Then

\[
 M \sum_{i=p}^{n-1} Y_i e_{i+1} = \begin{bmatrix}
 \sum_{i=p}^{n-1} U_i e_{i+1} \\
 \sum_{i=p}^{n-1} V_i e_{i+1}
\end{bmatrix}
\]

\[
 M \left( \sum_{i=p}^{n-1} Y_i Y_i' \right) M' = \begin{bmatrix}
 \sum_{i=p}^{n-1} U_i U_i' & \sum_{i=p}^{n-1} U_i V_i' \\
 \sum_{i=p}^{n-1} V_i U_i' & \sum_{i=p}^{n-1} V_i V_i'
\end{bmatrix}. \tag{4.5}
\]

From (4.3), it follows that

\[
 U_n = B_1 U_{n-1} + (e_n, 0, \ldots, 0)', \tag{4.6}
\]

where

\[
 B_1 = \begin{pmatrix}
 -h_1 & \cdots & -h_s \\
 I_{s-1} & 0
\end{pmatrix}.
\]

Since \( \{u_n\} \) is purely explosive, we obtain by Theorem 2(ii) that

\[
 \lim_{n \to \infty} B_1^{-n} \left( \sum_{i=p}^{n-1} U_i U_i' \right) (B_1^{-n})' = F \quad \text{is positive definite a.s.} \tag{4.7}
\]

Moreover, by Corollary 1(i),

\[
 \lim_{n \to \infty} \sum_{i=p}^{n-1} \|B_1^{-n} U_i\| < \infty \quad \text{a.s.} \tag{4.8}
\]
Let $C_{n-1} = \sum_{i=p}^{n-1} V_i V_i'$. Since $\{v_n\}$ is non-explosive, we obtain by Theorem 4 that

$$\lim_{n \to \infty} \max_{p \leq i \leq n-1} V_i' C_{n-1}^{-1} V_i = 0 \quad \text{a.s.} \quad (4.9)$$

Moreover, by Lemma 6(i,iii),

$$\|C_{n-1}^{-1/2}\| = O(n^{-1/2}) \quad \text{a.s.,} \quad (4.10)$$

and therefore

$$\|C_{n-1}^{-1}\| = O(n^{-1}) \quad \text{a.s.,} \quad (4.11)$$

and

$$\left\| C_{n-1}^{-1/2} \sum_{i=p}^{n-1} V_i \epsilon_{i+1} \right\| = O((\log n)^{1/2}) \quad \text{a.s.} \quad (4.12)$$

We now show that

$$D_n M \left( \sum_{i=p}^{n-1} Y_i Y_i' \right) M' D_n' \to \begin{pmatrix} F & 0 \\ 0 & I_r \end{pmatrix} \quad \text{a.s.,} \quad (4.13)$$

where $F$ is as defined in (4.7) and

$$D_n = \begin{pmatrix} B_1^{-n} & 0 \\ 0 & C_{n-1}^{-1/2} \end{pmatrix}. \quad (4.14)$$

From (4.5) and (4.13), we obtain

$$D_n M \left( \sum_{i=p}^{n-1} Y_i Y_i' \right) M' D_n' = \begin{pmatrix} B_1^{-n} \left( \sum_{i=p}^{n-1} U_i U_i' \right) (B_1^{-n})' & B_1^{-n} \left( \sum_{i=p}^{n-1} U_i V_i \right) C_{n-1}^{-1/2} \\ C_{n-1}^{-1/2} \left( \sum_{i=p}^{n-1} V_i U_i' \right) (B_1^{-n})' & I_r \end{pmatrix}. \quad (4.15)$$

Hence to prove (4.12), it suffices to show that

$$B_1^{-n} \left( \sum_{i=p}^{n-1} U_i V_i' \right) C_{n-1}^{-1/2} = \sum_{i=p}^{n-1} (B_1^{-n} U_i)(C_{n-1}^{-1/2} V_i)' \to 0 \quad \text{a.s.} \quad (4.14)$$

By (4.9),

$$\max_{p \leq i \leq n-1} \|C_{n-1}^{-1/2} V_i\|^2 = \max_{p \leq i \leq n-1} V_i' C_{n-1}^{-1} V_i \to 0 \quad \text{a.s.} \quad (4.15)$$

From (4.8) and (4.15), (4.14) follows.
Since $M$ and $D_n$ are non-singular, we obtain from (4.4) that

$$b_n - \beta = (n^{1/2}M'D_n^t) \left\{ D_nM \left( \sum_{i=p}^{n-1} Y_iY_i' \right) M'D_n^t \right\}^{-1} \times \left( n^{-1/2}D_nM \sum_{i=p}^{n-1} Y_i \varepsilon_{i+1} \right).$$

(4.16)

By (4.5) and (4.13),

$$n^{-1/2}D_nM \sum_{i=p}^{n-1} Y_i \varepsilon_{i+1} = \begin{pmatrix} n^{-1/2}B_1^{-n} \sum_{i=p}^{n-1} U_i \varepsilon_{i+1} \\ n^{-1/2}C_{n-1}^{-1/2} \sum_{i=p}^{n-1} V_i \varepsilon_{i+1} \end{pmatrix}.$$  

(4.17)

Since $\varepsilon_n = o(n^{1/2})$ a.s. (cf. [7]), we obtain from (4.8) that

$$\left\| n^{-1/2}B_1^{-n} \sum_{i=p}^{n-1} U_i \varepsilon_{i+1} \right\| \leq (n^{-1/2} \max_{p+1 \leq i \leq n} \| \varepsilon_i \|) \sum_{i=p}^{n-1} \| B_1^{-n}U_i \| \to 0 \text{ a.s.}$$

(4.18)

Moreover, by (4.11),

$$\left\| n^{-1/2}C_{n-1}^{-1/2} \sum_{i=p}^{n-1} V_i \varepsilon_{i+1} \right\| = O(n^{-1/2}(\log n)^{1/2}) \text{ a.s.}$$

(4.19)

By (2.5), (4.10), and (4.13),

$$\| n^{1/2}M'D_n^t \| = n^{1/2} \| M' \| \| (B_1^{-n})^{-1} \| + \| C_{n-1}^{-1/2} \| = O(1) \text{ a.s.}$$

(4.20)

From (4.12), (4.16), (4.17), (4.18), (4.19), and (4.20), it then follows that $b_n \to \beta$ a.s.

If all roots of $\phi(z)$ lie on or inside the unit circle, then $Y_n = V_n$ and the strong consistency of $b_n$ follows from (4.4), (4.10), and (4.11). Finally, if all roots of $\phi(z)$ lie outside the unit circle, then $B_1 = B$ and $Y_n = U_n$; moreover, since

$$b_n - \beta = (B^{-n})' \left\{ B^{-n} \left( \sum_{i=p}^{n-1} Y_iY_i' \right) (B^{-n})' \right\}^{-1} B^{-n} \sum_{i=p}^{n-1} Y_i \varepsilon_{i+1},$$

it follows from (2.5), (4.7), and (4.18) that $b_n \to \beta$ a.s.

We now show that (1.9) holds. In the non-explosive case, (1.9) follows
from (3.2), and in the purely explosive case, (1.9) follows from (2.4). Now consider the case of a mixed model. By (4.12),

\[(X_n^t X_n)^{-1} = M_n' D_n^t \Sigma_n D_n M_n,\]  

(4.21)

where \(\lim_{n \to \infty} \Sigma_n\) is a.s. positive definite. Hence by Lemma 2,

\[\lambda_{\text{max}}((X_n^t X_n)^{-1}) = O(\lambda_{\text{max}}(M_n' D_n^t D_n M_n)) = O(n^{-1}) \quad \text{a.s. by (4.20).} \]  

(4.22)

From (4.22), (1.9) follows. □

The following theorem shows that in a mixed AR(p) model, \(\|Y_n\|\) and \(\lambda_{\text{max}}(X_n^t X_n)\) grow exponentially fast as in the purely explosive case, while \(\lambda_{\text{min}}(X_n^t X_n)\) can only have an algebraic rate of growth.

**Theorem 5.** With the same notations and assumptions as in Theorem 1, assume that the characteristic polynomial \(\varphi(x)\) has roots outside the unit circle and also has roots on or inside the unit circle. Let \(p = \max_{1 \leq j \leq p} |z_j|\), where \(z_1, \ldots, z_p\) are the roots of \(\varphi(x)\). Define \(Y_n\) as in (1.5). Then

\[\lim_{n \to \infty} n^{-1} \log \|Y_n\| = \log p \quad \text{a.s.,} \]  

(4.23)

\[\lim_{n \to \infty} n^{-1} \log \lambda_{\text{max}}(X_n^t X_n) = 2 \log p \quad \text{a.s.,} \]  

(4.24)

\[\lambda_{\text{min}}(X_n^t X_n) = O(n^a) \quad \text{a.s. for some } a > 0. \]  

(4.25)

**Proof.** We use the same notations as in the preceding proof of Theorem 1. By (4.2), \(\|M_n Y_n\|^2 = \|U_n\|^2 + \|V_n\|^2\). Since \(\{u_n\}\) is purely explosive, \(n^{-1} \log \|U_n\| \to \log p\) a.s. by Corollary 1(iv). Since \(\{v_n\}\) is non-explosive, \(n^{-1} \log \|V_n\| \to 0\) a.s. by Theorem 3(i). This proves (4.23).

To prove (4.25), we note from (4.21) and Lemma 2 that

\[\lambda_{\text{max}}((M')^{-1} M^{-1}) \lambda_{\text{max}}((X_n^t X_n)^{-1}) \geq \lambda_{\text{max}}((M')^{-1}(X_n^t X_n)^{-1} M^{-1}) \]

\[\geq \lambda_{\text{min}}(\Sigma_n) \lambda_{\text{max}}(D_n^t D_n),\]

and therefore

\[\lambda_{\text{min}}(X_n^t X_n) = O(\lambda_{\text{min}}((D_n^t D_n)^{-1})) \quad \text{a.s.} \]  

(4.26)

Since \(\lambda_{\text{min}}((D_n^t D_n)^{-1}) \leq \lambda_{\text{min}}(C_{n-1})\) by (4.13), we obtain (4.25) from (4.26) and Theorem 3(ii). A similar argument also proves (4.24). □
REFERENCES


