CONSISTENCY AND LIMIT DISTRIBUTIONS OF ESTIMATORS OF PARAMETERS IN EXPLOSIVE STOCHASTIC DIFFERENCE EQUATIONS[1]

By M. M. Rao[2]

Carnegie Institute of Technology

1. Introduction and Summary. Let $\{X_t, t \geq 1\}$ be a stochastic process which satisfies the following set of assumptions:

Assumption 1: For every t, X_t satisfies

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \cdots + \alpha_k X_{t-k} + u_t, \tag{1}$$

where $\alpha_1, \cdots, \alpha_k$ are k finite real numbers (unknown parameters) and u_t, t positive, are independent, identically distributed random variables with mean zero and a finite positive variance σ^2.

Assumption 2: The distribution[3] of u_t is continuous. (Actually $\Pr\{u_t = 0\} = 0$ suffices.)

Assumption 3: The roots $m_1, \cdots, m_k$ of the characteristic equation

$$m^k - \alpha_1 m^{k-1} - \alpha_2 m^{k-2} - \cdots - \alpha_k = 0, \tag{2}$$

of (1), are distinct.

Assumption 4: There is a unique root ρ of (2) such that $|\rho| > 1$, and $|\rho| > \max_{j=2,\cdots,k} |m_j|$. Here ρ is identified with m_1 for convenience.

Since complex roots enter in pairs, it follows from this assumption that ρ is real. Note that there can be m_j, $j > 1$, such that $|m_j| > 1$.

Assumption 5: For t non-positive, $u_t = 0$.

If Assumption 4 holds, the process $\{X_t, t \geqq 1\}$ is said to be (strongly) *explosive*, and the corresponding difference equation (1) is called an *explosive* (linear homogeneous) *stochastic difference equation*; this is the subject of the present paper.

Under the assumptions above, it follows (cf., C. Jordan [5], p. 564, Mann and Wald [8], p. 178, and also the footnote on p. 22 of [10]) that

$$X_t = \sum_{r=1}^{t} \sum_{q=1}^{k} \lambda_q m_q^{t-r} u_r,$$

Received December 14, 1959; revised May 5, 1960.

[1] This paper forms part of a dissertation submitted to the University of Minnesota, in partial fulfillment of the requirements for the Ph.D. degree. The work was done in part under Contract Nonr 25-82(00), Task NR 042-200 of the Office of Naval Research, at Minnesota. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2] Presented to the American Mathematical Society, September 2, 1959. (Cf. Abstract in *Notices*. Vol. 6 (1959), pp. 432–433.)

[3] Some authors use cumulative distribution for the same concept. Terminology here follows [3] and [4].

t positive, and that λ_q satisfy the relations

$$\delta_{1t} = \sum_{q=1}^{k} \lambda_q m_q^{t-1}, \qquad t = 1, 0, -1, \cdots, -(k-2), \tag{3}$$

where $\delta_{1t} = 1$ if $t = 1$ and 0 otherwise. (Note that $\sum_{q=1}^{k} \lambda_q = 1$.) For convenience, define the random variables

$$X_{i,t} = \sum_{r=1}^{t} m_i^{t-r} u_r, \qquad i = 1, 2, \cdots, k, \ (m_1 = \rho), \tag{4}$$

so that $X_{i,t} = 0$ for t non-positive. Thus one may write X_t as follows:

$$X_t = \lambda_1 X_{1,t} + \lambda_2 X_{2,t} + \cdots + \lambda_k X_{k,t}. \tag{5}$$

The first part of this paper is devoted to finding a consistent estimator of ρ and its limit distribution. Consequently, in Section 3 some lemmas will be proved for use in the consistency proof (Theorem I). Similarly, in Section 5, some lemmas leading to the proof of the limit distribution of the estimator (Theorem II) will be given.

In the second part, the consistency of the Least Squares (L.S.) or Maximum Likelihood (M.L.) estimators of the "structural parameters" α_i of (1) will be considered (Theorem III). The procedure becomes much more involved because the direct application of the usual limit theorems is not possible, since the process under consideration is explosive. It is noteworthy that Lemmas 9, 10, 14–16, and Theorem I are rather general, in that they hold under the only global Assumptions 1–5 above, and the further requirement $|m_j| < 1, j = 2, \cdots, k$, so essential for the rest of the analysis of this paper, is unnecessary for them.

The corresponding problem, in the case $|\rho| < 1$, has been completely solved by Mann and Wald [8]. If $k = 1$ in (1), the results of this paper reduce to those obtained by Rubin [13], White [14], and T. W. Anderson [1]. The vector case has also been treated by Anderson in [1], but a comparison of the results in this case with those of the present paper shows that they do not imply each other except in the first order. In the latter case, however, both reduce to Rubin's [13] result. The available results on stochastic difference equations are summarized in a table at the end of the paper. Some of the details and computations omitted in this paper may be found in [10].

In the following section, some known lemmas related to stochastic convergence are collected and stated in a convenient form, as they will be constantly referred to in both parts of the paper. (For proofs, see [2], [3], [4], [6] and [9].)

2. Lemmas Related to Stochastic Convergence. To avoid misunderstanding certain basic terms, often used in the paper, will first be defined. By a random variable (r.v.) is meant a finite real valued measurable function on the measure space in question. A random vector is one which has a finite number of r.v.'s as its components. Pr $\{S\}$ and $E_\theta(f(X))$ are the probability and expectation symbols (cf. [2]).

Let $\{X_n\}$ be a sequence of r.v.'s and X be a r.v. Then the stochastic convergence and convergence in distribution of X_n to X, as $n \to \infty$, will be written $X_n \xrightarrow{P} X$, and $X_n \xrightarrow{D} X$ respectively (cf. [6]). The stochastic equivalence of two sequences of r.v.'s, $\{X_n\}$ and $\{Y_n\}$, will be written $X_n \overset{P}{=} Y_n$ (i.e., $(X_n - Y_n \xrightarrow{P} 0)$.

A sequence of r.v.'s $\{X_n\}$ is *bounded in probability* if, for any $\epsilon > 0$, there exists a positive number M_ϵ (depending only on ϵ) such that

$$\limsup_{n\to\infty} \Pr\{|X_n| \geqq M_\epsilon\} \leqq \epsilon.$$

Note that a sequence of r.v.'s is always bounded in probability if their means and variances are bounded functions of n.

Unless stated otherwise, in what follows in this section, $\{X_n\}$ and $\{Y_n\}$ are two arbitrary sequences of r.v.'s. *All the limits are taken as $n \to \infty$, and the repetition of this phrase will be omitted.*

LEMMA 1: *If $X_n \xrightarrow{P} X$, and $Y_n \xrightarrow{P} Y$, then $X_nY_n \xrightarrow{P} XY$. If further* $\Pr\{Y = 0\} = 0$, *then $(X_n/Y_n) \xrightarrow{P} X/Y$.*

LEMMA 2: *If $X_n \xrightarrow{P} 0$, and $\{Y_n\}$ is bounded in probability, then $X_nY_n \xrightarrow{P} 0$.*

LEMMA 3: *If $X_n \xrightarrow{P} X$, then $X_n \xrightarrow{D} X$. (The lemma holds if $\{X_n\}$ is a vector sequence. Then the stochastic convergence is component-wise.)*

LEMMA 4: *Let $\{X_{n,0}, X_{n,1}, \cdots, X_{n,k}\}$ be a sequence of random vectors such that $(X_{n,0}, X_{n,1}, \cdots, X_{n,k}) \xrightarrow{D} (X_0, X_1, \cdots, X_k)$ and* $\Pr\{X_0 = 0\} = 0$. *Then $(\sum_{i=1}^k a_iX_{n,i}/X_{n,0}) \xrightarrow{D} (\sum_{i=1}^k a_iX_i/X_0)$, where a_i are some constants independent of n.* (e.g., *if $k = 1$, and X_0, X_1 are independent normal with zero means, then the latter limit distribution is Cauchy.*)

Obviously some generalizations hold.

LEMMA 5: *Let $\{X_n\}$ be a sequence of r.v.'s with $\{\mu_n\}$ and $\{\sigma_n^2\}$ as the mean and variance sequences respectively. If $\mu_n \to 0$, and $\sigma_n \to 0$, then $X_n \xrightarrow{P} 0$.*

LEMMA 6: *If $X_n \overset{P}{=} Y_n$, and $X_n \xrightarrow{D} X$, then $Y_n \xrightarrow{D} X$. (Sometimes $X_n \overset{P}{=} Y_n$ is also written as $X_n = Y_n + o_p(1)$.)*

LEMMA 7: *(Kolmogorov). Let $Y_1, Y_2, \cdots$ be a sequence of mutually independent r.v.'s with means zero and variances $\sigma_1^2, \sigma_2^2, \cdots$. Then, if $\sum_{n=1}^\infty \sigma_n^2 = \sigma^2 < \infty$, $\sum_{n=1}^\infty Y_n = X$ is convergent with probability one. Moreover, $E(X) = 0$, $E(X^2) = \sigma^2$, and if $X_n = \sum_{i=1}^n Y_i$, then for every $\epsilon > 0$,*

$$\Pr\{\operatorname{lub}_n |X_n| \geqq \epsilon\} \leqq \sigma^2/\epsilon^2.$$

PART I

3. Lemmas for Theorem I. Define a "normalizing factor" $s(n)$ as $s(n) = \lambda_1 |\rho|^n/(\rho^2 - 1)$. For convenience of writing, it is assumed that λ_1 is positive. Otherwise it will be replaced by $|\lambda_1|$. (Note that λ_1, being the coefficient of the term in ρ, i.e., $X_{1,t}$, is never zero.) Next define

$$V_n = (\rho^2 - 1)^{\frac{1}{2}} \sum_{r=1}^n \rho^{-r} u_r. \tag{6}$$

LEMMA 8: *There exists a r.v., V such that* (i) *$V_n \to V$ with probability one, and* (ii) $\Pr\{V = 0\} = 0$.

PROOF: (i) This is immediate from Lemma 7, on identifying the V_n here with X_n there, and V with X. Thus $E(V) = 0$, Var $V = E(V^2) = \sigma^2$. (ii) Since the distribution of u_r is continuous, the distribution of V is continuous. Hence $\Pr\{V = a\} = 0$, for any a, in particular $a = 0$.

LEMMA 9[4]: *If V is the r.v. defined in Lemma 8, then*

$$[s(n)]^{-2} \sum_{t=2}^{n} X_{t-1}^2 \xrightarrow{P} V^2, \quad \textit{and} \quad [s(n)]^{-2} \sum_{t=2}^{n} X_t X_{t-1} \xrightarrow{P} \rho V^2. \tag{7}$$

The proof of this lemma is omitted as it is similar to, and a special case of, a more general result to be given in Lemma 15 below.

4. A Consistent Estimator of ρ.

Using (4) and (5) and the relations $X_{i,t} = m_i X_{i,t-1} + u_t$, the r.v. X_t of (1) can be written as follows:

$$X_t - \rho X_{t-1} = \sum_{q=1}^{k} \lambda_q (X_{q,t} - \rho X_{q,t-1}). \tag{8}$$

Since $m_1 = \rho$, $\sum_{q=1}^{k} \lambda_q = 1$,

$$X_t - \rho X_{t-1} = u_t + \sum_{j=2}^{k} \lambda_j (m_j - \rho) X_{j,t-1} = v_t, \quad \text{say,} \tag{9}$$

so that

$$X_t = \rho X_{t-1} + v_t. \tag{10}$$

Note that the v_t, being dependent, can form an unstable sequence of r.v.'s.

The "first order least squares" estimator $\tilde{\rho}_n$ of ρ is given by

$$\tilde{\rho}_n = \sum_{t=1}^{n} X_t X_{t-1} \Big/ \sum_{t=1}^{n} X_{t-1}^2 = \rho + \sum_{t=1}^{n} v_t X_{t-1} \Big/ \sum_{t=1}^{n} X_{t-1}^2. \tag{11}$$

THEOREM I: *Let $\{X_t, t \geqq 1\}$ satisfy Assumptions 1 to 5 of Section 1. Then $\tilde{\rho}_n \xrightarrow{P} \rho$, or $\lim_{n\to\infty} \Pr[|\tilde{\rho}_n - \rho| > \epsilon] = 0$ for any positive ϵ,* (i.e., *$\tilde{\rho}_n$ is a consistent estimator of ρ) where $\tilde{\rho}_n$ is defined by* (11).

PROOF: The proof of this theorem is an immediate consequence of Lemmas 1, 8 and 9.

A separate proof of Theorem I and Lemma 9 can be found in [10], without recourse to Lemma 15. A similar remark applies to Lemma 10 also.

5. Lemmas for the Limit Distribution of $\tilde{\rho}_n$.

In this section some lemmas, useful in the derivation of the limit distribution of the estimator $\tilde{\rho}_n$ of ρ, will be proved. Define the r.v.,

$$U_{n,i} = (\rho^2 - 1)^{\frac{1}{2}} \sum_{t=i+1}^{n} \rho^{-(n-t+1)} u_t. \tag{12}$$

---

[4] Lemmas 9 and 10 have been proved earlier by T. W. Anderson, [1], for the case $k = 1$.

LEMMA 10: *The r.v.* $s^{-1}(n) \sum_{t=2}^{n} u_t X_{t-1} \overset{P}{=} a^n U_{n,1} V_n$, *where* $U_{n,1}$, V_n *and* $s(n)$ *are defined respectively in* (12), (6), *and* $a = \rho/|\rho|$.

This is a special case of a result to be given in Lemma 14 below.

LEMMA 11: *The r.v.* V_n *is asymptotically uncorrelated with* $U_{n,1}$ *and* $X_{j,n}$, $j = 2, \cdots, k$, *where* $X_{j,n} = \sum_{r=1}^{n} m_j^{n-r} u_r$ *and* $|m_j|$ *is less than unity.*

PROOF: From definitions $E(U_{n,1}) = 0 = E(V_n) = E(X_{j,n})$, all n,

$$\text{Cov}\,(U_{n,1}, V_n) = \sigma^2 \rho^{-n-1}(\rho^2 - 1)\,(n - 1) \to 0, \tag{13}$$

since $|\rho| > 1$. Further,

$$\text{Cov}(V_n, X_{j,n}) \begin{cases} = \sigma^2(\rho^2 - 1)^{\frac{1}{2}} m_j^n n & \text{if} \quad \rho m_j = 1, \\ = \sigma^2(\rho^2 - 1)^{\frac{1}{2}} m_j^n \dfrac{[(\rho m_j)^{-1} - (m_j \rho)^{-n-1}]}{[1 - (m_j \rho)^{-1}]} & \text{if} \quad m_j \rho \neq 1, \end{cases} \tag{14}$$

which also $\to 0$, since $|\rho| > 1 > |m_j|$.

LEMMA 12: *The r.v.'s* $U_{n,1}$ *and* $X_{j,n}$ *are correlated even asymptotically. If* $m_j \rho = 1$ *then the asymptotic correlation is* ± 1.

The statements are verified analogously.

The next lemma plays an important role in the limit distribution of $\tilde{\rho}_n$.

LEMMA 13: *The random vectors* $(V_n, U_{n,1}, \cdots, X_{k,n})$ *converge in distribution to a random vector which will be denoted as* $(V, U, W_2, \cdots, W_k)$. *Moreover,* V *and* $(U, W_2, \cdots, W_k)$ *are independently distributed. Here the* V_n, $U_{n,1}$, *and* $X_{i,n}$ *have the same meanings as in Lemmas* 11 *and* 12.

(The factor $(\rho^2 - 1)^{\frac{1}{2}}$, in $U_{n,1}$ and V_n, which is irrelevant here, will be omitted in the proof for convenience and symmetry, slightly abusing the notation. Note that the W_i's are defined as the limit in distribution of $X_{i,n}$, the existence of such a limit being part of the conclusion of the lemma.)

PROOF. Let $m_1 = 1/\rho$ and $m_2, \cdots, m_k$ be as before. Then $|m_i| < 1$, $i = 1, \cdots, k$. Since $V_n \to V$, with probability one by Lemma 8, so $V_n \xrightarrow{D} V$. Let $X^{(n)} = (U_{n,1}, X_{2,n}, \cdots, X_{k,n})'$ (prime for transpose). To prove the lemma, it suffices to show that (i) $X^{(n)} \xrightarrow{D} X = (U, W_2, \cdots, W_k)'$ and (ii) V and X are independent. Since $X^{(n)} \xrightarrow{D} X$ if, and only if, for any (real) vector $a = (a_1, \cdots, a_k)$, $aX^{(n)} = (a_1 U_{n,1} + \cdots + a_k X_{k,n}) \xrightarrow{D} aX = (a_1 U + \cdots + a_k W_k)$ (see e.g., [7], Proposition 7.1), consider $aX^{(n)}$, and let $\psi(t) = E(e^{itu_r})$.

$$\begin{aligned} \varphi_n(t) &= E\left(\exp\left[it \sum_{r=1}^{n} u_r(a_1 m_1^{(n-r)} + \cdots + a_k m_k^{(n-r)})\right]\right) \\ &= \prod_{r=1}^{n} \psi(t[a_1 m_r^{n-r} + \cdots + a_k m_k^{n-r}]) = \prod_{r=0}^{n-1} \psi(t[a_1 m_1^r + \cdots + a_k m_k^r]) \\ &= E\left(\exp\left[it \sum_{r=0}^{n-1} u_{r+1}(a_1 m_1^r + \cdots + a_k m_k^r)\right)\right] \\ &= E(\exp\,[it(a_1 \bar{V}_{n,1} + \cdots + a_k \bar{V}_{n,k})]), \quad \text{say}. \end{aligned} \tag{15}$$

Since $|m_i| < 1$, and u_r are independent with means zero etc., it follows by Lemma

7 that $\bar{V}_{n,i} \to \bar{V}_i$ with probability one, and consequently, using Lemma 3, one has that $(\bar{V}_{n,1}, \cdots, \bar{V}_{n,k}) \xrightarrow{D} (\bar{V}_1, \cdots, \bar{V}_k) = (U, W_2, \cdots, W_k)$. Hence, $\varphi_n(t) = E(e^{itaX^{(n)}}) = E(e^{ita\bar{V}^{(n)}}) \to \varphi(t) = E(e^{ita\bar{V}}) = E(e^{itaX})$. This proves (i). Next (ii) will be proved by slightly extending an argument of J. R. Blum ([1], p. 679). Let [n/2] = integral part of $n/2$. Define $V_n^* = \sum_{r=1}^{[n/2]} m_1^r u_r$, and $\tilde{V}_n = \sum_{r=[n/2]+1}^{n} m_1^r u_r$. Clearly $V_n^* \stackrel{P}{=} V_n$ since $\tilde{V}_n \xrightarrow{P} 0$, and hence, by Lemma 6, they have the same limit distribution. Similarly, let

$$aX^{(n)*} = \sum_{r=[n/2]+1}^{n} u_r(a_1 m_1^{n-r} + \cdots + a_k m_k^{n-r}),$$

and

$$a\tilde{X}^{(n)} = \sum_{r=1}^{[n/2]} u_r(a_1 m_1^{n-r} + \cdots + a_k m_k^{n-r}), \qquad (\xrightarrow{P} 0)$$

so that $aX^{(n)} \stackrel{P}{=} aX^{(n)*}$, and since $X^{(n)}$ has a limit distribution by part (i) of this lemma, $X^{(n)*}$ has the same limit distribution. But from definition, for every n, V_n^* and $X^{(n)*}$ are independent, and hence they are also independent as $n \to \infty$, proving (ii), q.e.d.

COROLLARY TO LEMMA 13: *Let $X_{i,n}$ be as in Lemma 13 and $\bar{X}_{j,n} = \sum_{r=1}^{n-1} (n-r) m_j^{n-r-1} u_r$ for some $j(|m_j|) < 1)$. Then $(X_{i,n}, \bar{X}_{j,n})$ has a joint limit distribution, and if any $X_{j,n}$ in Lemma 13 is replaced by $\bar{X}_{j,n}$ then the conclusions of the lemma remain valid.*

PROOF: The proof runs on the same lines as in the lemma. In fact (taking $i = 1$)

$$\tag{16} \begin{aligned} \varphi_n(t) = E(\exp[it(a_1 X_{1,n} + a_2 X_{j,n})]) &= \prod_{r=1}^{n-1} \psi(t[a_1 m_1^{n-r} + a_2(n - r) m_j^{n-r-1}]) \\ &= \prod_{r=1}^{n-1} \psi(t[a_1 m_1^r + a_2 r m_j^{r-1}]) = E\{\exp[it(a_1 \bar{V}_{n,1} + a_2 \bar{\bar{V}}_{n,j})]\}, \quad \text{say.} \end{aligned}$$

It was noted that $\bar{V}_{n,1} \to \bar{V}_1$ with probability one. But it also follows from Lemma 7, that $\bar{\bar{V}}_{n,j} = \sum_{r=1}^{n} r m_j^{r-1} u_r \to \bar{\bar{V}}_j$ (say), with probability one, since the r.v. $\bar{\bar{V}}_{n,j}$ has also a bounded variance. Thus $\varphi_n(t) \to \varphi(t)$, as in the lemma itself. The proof of the last statement is identical to that in the lemma, q.e.d.

6. Limit Distribution of $\tilde{\rho}_n$. A complete (i.e., self-contained) statement of the theorem on the limit distribution of $\tilde{\rho}_n$ will be given here, even if it involves some repetition. (The W_j's below are the same as in Lemma 13.)

THEOREM II: *Let $\{X_t, t$ positive integer$\}$ be a stochastic process satisfying the following conditions:*

CONDITION 1: *For each t, $X_t = \alpha_1 X_{t-1} + \cdots + \alpha_k X_{t-k} + u_t$, k finite, where $\alpha_1, \cdots, \alpha_k$ are finite real numbers (unknown parameters), and u_t, t positive, are independent, identically distributed r.v.'s with mean zero and a finite positive variance σ^2.*

CONDITION 2: *The distribution of u_t is continuous.*

CONDITION 3: *The roots $m_1, \cdots, m_k$ of the characteristic equation $m^k - \alpha_1 m^{k-1} -$*

$\cdots - \alpha_k = 0$ *are simple (i.e.,* $m_i \neq m_j$ *if* $i \neq j$*), with one root, say* $\rho = m_1$, *and the other roots* m_j, $j = 2, \cdots, k$, *satisfy the inequalities* $|\rho| > 1 > |m_j|$.

CONDITION 4: *For t non-positive,* $u_t = 0$.

Then, it follows that $s(n)\ (\hat{\rho}_n - \rho)$ *has a limit distribution which is that of the r.v.* $(GU + W + Hu)/V$, *where* $W = \sum_{j=2}^{k} C_j W_j$, $s(n) = \lambda_1 |\rho|^n/(\rho^2 - 1)$ (λ_1 *is positive) and* $\hat{\rho}_n = \sum_{t=1}^{n} X_t X_{t-1} / \sum_{t=1}^{n} X_{t-1}^2$, *the* λ_1, G, H, *and* C_j *are some constants that depend only on the roots* $(\rho, m_2, \cdots, m_k)$.

COROLLARY II_1 : *If the process satisfies the conditions* 1, 3 *and* 4, *and* u_t *are Gaussian, then the limit distribution of* $s(n)\ (\hat{\rho}_n - \rho)$ *is a Cauchy distribution.*

COROLLARY II_2 : *Under the hypothesis of Theorem* II, *it follows that the limit distribution of* $(\sum_{t=1}^{n} X_{t-1}^2)^{\frac{1}{2}} \cdot (\hat{\rho}_n - \rho)$ *is the same as that of the r.v.* $GU + W + Hu$, *and under the hypothesis of Corollary* II_1, *this limit distribution, i.e., of* $(\sum_{t=1}^{n} X_{t-1}^2)^{\frac{1}{2}} \cdot (\hat{\rho}_n - \rho)$, *is Gaussian with mean zero, and a finite variance depending on* $(\rho, m_2, \cdots, m_k)$.

The proof of the theorem and the corollaries will be given in succession.

PROOF OF THEOREM II: The general idea of the proof is this: From (11)

$$s(n)(\hat{\rho}_n - \rho) = \left(\sum_{t=1}^{n} \frac{v_t X_{t-1}}{s(n)}\right) \Big/ \left(\sum_{t=1}^{n} \frac{X_{t-1}^2}{s^2(n)}\right) = R_n/Q_n, \tag{17}$$

where R_n and Q_n stand respectively for the expressions in the numerator and denominator. First R_n and Q_n will be expressed in terms of stochastically equivalent quantities, call them $\bar{R}_n$ and $\bar{Q}_n$ (they are given precisely in steps 13 and 14 below), i.e.,

$$R_n = \bar{R}_n + o_p(1) \quad \text{and} \quad Q_n = \bar{Q}_n + o_p(1). \tag{18}$$

Then, by Lemma 6, R_n and $\bar{R}_n$, and Q_n and $\bar{Q}_n$ converge in distribution to the same r.v.'s. Then, by Lemma 4, R_n/Q_n and $\bar{R}_n/\bar{Q}_n$ converge in distribution to the same limit if $\bar{Q}_n$ does not converge to a degenerate distribution at the origin. Therefore, the main task here is to obtain $\bar{R}_n$ and $\bar{Q}_n$ which are simpler to deal with, and to show that the latter is non-degenerate at the origin. It will be seen that the r.v.'s $\bar{R}_n$ and $\bar{Q}_n$ are "nice" functions of $U_{n,1}$, V_n, and $X_{j,n}$. Then using Lemma 13 the limit distribution will be obtained.

This plan is carried through in several steps as follows:

1. $Q_n = \sum_{t=1}^{n} X_{t-1}^2/s^2(n) = V_{n-1}^2 + o_p(1)$, by Lemma 9.

In steps 2–12, R_n will be simplified.

2. $R_n = \sum_{t=1}^{n} \frac{u_t X_{t-1}}{s(n)} + \sum_{j=2}^{k} \sum_{i=1}^{k} \lambda_i \lambda_j (m_j - \rho) \sum_{t=1}^{n} \frac{X_{i,t-1} X_{j,t-1}}{s(n)}$, using (10), $=$ $A + \bar{B}$, say.

3. $A = \sum_{t=1}^{n} u_t X_{t-1}/s(n) = a^n U_{n,1} V_n + o_p(1)$, by Lemma 10.

4. $\bar{B} = \frac{\lambda_1}{s(n)} \sum_{j=2}^{k} (m_j - \rho)\lambda_j \sum_{t=1}^{n} X_{1,t-1} X_{j,t-1} + \sum_{j=2}^{k} \sum_{q=2}^{k} \lambda_j \lambda_q (m_j - \rho) \sum_{t=1}^{n}$ $\frac{X_{q,t-1} X_{j,t-1}}{s(n)} = A_1 + \sum_{j=2}^{k} \sum_{q=2}^{k} \lambda_j \lambda_q (m_j - \rho) \frac{1}{s(n)} B_{qj}$, say.

5. $s(n)^{-1}B_{qj} \xrightarrow{P} 0$, all $q, j > 1$. For, by Schwarz' inequality (for the first time $|m_q| < 1$ will be used)

$$\left|\frac{1}{s(n)} B_{qj}\right| = \left|\frac{1}{s(n)} \sum_{t=1}^{n} X_{q,t-1} X_{j,t-1}\right| \leqq \left(\sum_{t=1}^{n} \frac{X_{q,t-1}^2}{s(n)}\right)^{\frac{1}{2}} \left(\sum_{t=1}^{n} \frac{X_{j,t-1}^2}{s(n)}\right)^{\frac{1}{2}}, \tag{19}$$

$$\begin{aligned} E\left(\sum_{t=1}^{n} \frac{X_{q,t-1}^2}{s(n)}\right) &= \frac{1}{s(n)} \sum_{t=2}^{n} E\left(\sum_{r=2}^{t-1} m_q^{t-1-r} u_r\right)^2 \\ &= \sigma^2 \frac{(\rho^2 - 1)}{\lambda_1 |\rho|^n} \left[\frac{n}{1 - m_q^2} - \frac{m_q^2 - m_q^{2n}}{(1 - m_q^2)^2}\right] \to 0, \end{aligned} \tag{20}$$

since $|m_q| < 1 < |\rho|$.

By Markov's inequality, since the r.v.'s are non-negative and $q, j > 1$ implies $|m_q| < 1$ and $|m_j| < 1$, it follows that the right side of (19) converges in probability to zero, so that step 5 is proved. Consequently, the rest of the analysis is concerned with A_1.

6. $A_1 = \sum_{j=2}^{k} (m_j - \rho)\lambda_j \cdot \frac{\lambda_1}{s(n)} \sum_{t=1}^{n} X_{1,t-1} X_{j,t-1} = \sum_{j=2}^{k} \lambda_j(m_j - \rho)A_{1j}$, say.

7. It will be shown that A_{1j} is stochastically equivalent to a r.v. in terms of $U_{n,i}$, V_{n-i} and $X_{i,n}$.

8. $$\begin{aligned} A_{1j} &= \frac{\rho^2 - 1}{|\rho|^n} \sum_{t=2}^{n} X_{j,t-1} \left(\sum_{r=1}^{t-1} \rho^{t-1-r} u_r\right) \\ &= a^n \left[\left((\rho^2 - 1)^{\frac{1}{2}} \sum_{t=2}^{n} \rho^{-(n-t+1)} X_{j,t-1}\right)\left((\rho^2 - 1)^{\frac{1}{2}} \sum_{r=1}^{n} \rho^{-r} u_r\right)\right] \\ &\quad - a^n \left[(\rho^2 - 1) \sum_{t=2}^{n} \rho^{-(n-t+1)} X_{j,t-1} \sum_{r=t}^{n} \rho^{-r} u_r\right] \\ &= a^n A^* V_n - a^n B^*, \quad \text{say.} \end{aligned}$$

The r.v. B^* was added and subtracted. Thus

9. $A_{1j} = a^n(A^* V_n - B^*)$.

It will now be shown that

10. $B^* \xrightarrow{P} 0$. Since $X_{j,t-1}$ and $\sum_{r=t}^{n} \rho^{-r+1} u_r$ are independent and each has mean zero, $E(B^*) = 0$. However, different terms are *not* independent. Consequently, an upper bound for the variance of B^* is obtained, using the elementary inequality Var $(X + Y) \leqq [\text{S.D.}(X) + \text{S.D.}(Y)]^2$, and that is shown to converge to zero.

Regrouping the terms in B^*, one gets the following:

$$\begin{aligned} B^* = \frac{\rho^2 - 1}{\rho} \Bigg[\rho^{-n} \sum_{r=1}^{n-1} X_{j,r} u_{r+1} + \rho^{-n-1} \sum_{r=1}^{n-2} X_{j,r} u_{r+2} \\ + \cdots + \rho^{-(2n-2)} X_{j,1} u_n\Bigg]. \end{aligned} \tag{21}$$

The variance of the general term in B^* is given by

$$\text{Var}\left[\rho^{-(n+i)}\sum_{r=1}^{n-i}X_{j,r}u_{r+i}\right] = \sigma^2\rho^{-2(n+i)}\sum_{r=1}^{n-i}\text{Var}\,X_{j,r} \tag{22}$$
$$\leqq [(n-i)/(1-m_j^2)]\sigma^2\rho^{-2n}, \quad \text{since} \quad |m_j| < 1.$$

$$\text{S.D.}\ (\rho^{-(n+i)}\sum_{r=1}^{n-i}X_{j,r}u_{r+i}) \leqq |\rho|^{-n}(n-i)(1-m_0^2)^{-\frac{1}{2}}$$
$$\text{where} \quad |m_0| = \max_j |m_j| < 1.$$

Consequently

$$\begin{aligned}\text{Var}\,B^* &\leqq (\rho^2 - 1/\rho)^2[\rho^{-2n}/(1-m_0^2)][(n-1)+(n-2)+\cdots+1]^2 \\ &= \text{constant}\cdot(\rho^2-1/\rho)^2[\rho^{-2n}/(1-m_0^2)]n^4 \to 0,\end{aligned} \tag{23}$$

so that $B^* \xrightarrow{P} 0$.

11. Now A^* will be simplified.

$$\begin{aligned}A^* &= \left(\frac{\rho^2-1}{\rho^2}\right)^{\frac{1}{2}}\sum_{t=2}^{n}\rho^{-(n-t)}X_{j,t-1} \\ &= \left(\frac{\rho^2-1}{\rho^2}\right)^{\frac{1}{2}}[\rho^{-(n-2)}u_1 + \rho^{-(n-3)}(m_j u_1 + u_2) \\ &\qquad + \cdots + 1\cdot(m_j^{n-2}u_1 + \cdots + u_{n-1})],\end{aligned} \tag{24}$$

since $X_{j,t} = \sum_{r=1}^{t} m_j^{t-r}u_r$. Another regrouping gives

$$\begin{aligned}A^* = \left(\frac{\rho^2-1}{\rho^2}\right)^{\frac{1}{2}}\Bigg[&u_1 m_j^{n-2}\sum_{r=2}^{n}(\rho m_j)^{-(n-r)} \\ &+ u_2 m_j^{n-3}\sum_{r=3}^{n}(\rho m_j)^{-(n-r)} + \cdots + u_{n-1}\Bigg].\end{aligned} \tag{25}$$

Case 1: $m_j\rho \neq 1$. Then (excluding the trivial case $m_j = 0$)

$$\begin{aligned}A^* &= \left(\frac{\rho^2-1}{\rho^2}\right)^{\frac{1}{2}}\cdot(1-\rho m_j)^{-1}\left[\sum_{r=1}^{n-1}\rho^{-(n-1-r)}u_r - \rho m_j\sum_{r=1}^{n-1}m_j^{n-1-r}u_r\right] \\ &= (1-\rho m_j)^{-1}[U_{n-1,0} - m_j(\rho^2-1)^{\frac{1}{2}}X_{j,n-1}].\end{aligned} \tag{26}$$

Case 2: $\rho m_j = 1$. Then (25) becomes

$$A^* = \left(\frac{\rho^2-1}{\rho^2}\right)^{\frac{1}{2}}\sum_{r=1}^{n-1}m_j^{n-r-1}u_r = \bar{X}_{j,n}, \quad \text{say.} \tag{27}$$

In this case, the corollary to Lemma 13 will be used to conclude that $(V_n, \bar{X}_{2n}, \cdots, \bar{X}_{k,n})$ has a joint limit distribution which is continuous in the V component, and the rest of the analysis is unchanged. Therefore, in what follows, only the relatively harder Case 1 will be considered.

Thus, using the value of A^* in A_{1j} of step 9, and A_{1j} in A_1 of step 6, and A_1 in $\bar{B}$ of 4, and finally A from 3 and $\bar{B}$ from 4 in 2, one obtains the following expression for R_n :

12. $R_n = \Big[a^n U_{n,1} + \sum_{j=2}^{k} \frac{(m_j - \rho)\lambda_j}{1 - \rho m_j} U_{n-1,1} - (\rho^2 - 1)^{\frac{1}{2}} \sum_{j=2}^{k} \frac{(m_j - \rho)\lambda_j m_j}{1 - \rho m_j} X_{j,n-1}\Big] V_{n-1} + o_p(1)$,

because $V_{n-1} \overset{P}{=} V_n$ and $U_{n,0} \overset{P}{=} U_{n,1}$. But $U_{n,1} = \rho^{-1}U_{n-1,1} + \rho^{-1}(\rho^2 - 1)^{\frac{1}{2}} u_n$, so that

$$R_n = \Big[a^n\rho^{-1}U_{n-1,1} + a^n\rho^{-1}(\rho^2 - 1)^{\frac{1}{2}}u_n + \sum_{j=2}^{k} \frac{(m_j - \rho)\lambda_j}{1 - \rho m_j} U_{n-1,1}$$

$$- (\rho^2 - 1)^{\frac{1}{2}} \sum_{j=2}^{k} \frac{(m_j - \rho)\lambda_j m_j}{1 - \rho m_j} X_{j,n-1}\Big] V_{n-1} + o_p(1) \tag{28}$$

$$= \Big[GU_{n-1,1} + Hu_n + \sum_{j=2}^{k} C_j X_{j,n-1}\Big] V_{n-1} + o_p(1),$$

where G, H and C_j are constants that depend only on the roots $(\rho, m_2, \cdots, m_k)$. Note that $(|a| = 1)a^n$ has no influence on any statements since $U_{n,i}$ as well as u_n are r.v.'s with zero means, the same variances as before the multiplication of a^n, and their distributions are still continuous.

13. $R_n = \bar{R}_n + o_p(1)$, where $\bar{R}_n = [GU_{n-1,1} + Hu_n + \sum_{j=2}^{k} C_jX_{j,n-1}] V_{n-1}$. Clearly u_n is independent of $U_{n-1,1}$, $X_{j,n-1}$ and V_{n-1}.

14. $Q_n = \bar{Q}_n + o_p(1)$, where $\bar{Q}_n = V_{n-1}^2$. (Cf. step 1.)

By Lemma 13, $(u_n, U_{n-1,1}, V_{n-1}, X_{2,n-1}, \cdots, X_{k,n-1})$ has a limit distribution which is continuous in the V component, and the limit distribution is that of $(u, U, V, W_2, \cdots, W_k)$, where, in fact, V is independent of the other r.v.'s. Hence it follows, by Lemma 4, that the limit distribution of $s(n)(\hat{\rho}_n - \rho) = R_n/Q_n$ is the same as that of $\bar{R}_n/\bar{Q}_n$, i.e., of the r.v. $[(GU + Hu + \sum_{j=2}^{k} C_jW_j) V]/V^2$, since $\Pr\{V = 0\} = 0, = [GU + Hu + W]/V$, where G, H, C's are constants depending on the characteristic roots $(\rho, m_2, \cdots, m_k)$, q.e.d.

Proof of Corollary II_1 : It was seen that u and V are independent, and V is independent with U and W by Lemma 13 above. If u_t are $N(0, \sigma^2)$, then it follows that U, u, W all have Gaussian distributions (U and W being linear combinations of u_t) with zero means and finite variances. The same is true of V. Consequently $[GU + Hu + W]/V$ has a Cauchy distribution.

The continuity of the distribution of u_t with two moments, which is condition 2, is clearly satisfied, q.e.d.

Proof of Corollary II_2 : It suffices to observe that

$$\Big(\sum_{t=1}^{n} X_{t-1}^2\Big)^{\frac{1}{2}} (\hat{\rho}_n - \rho) = \sum_{t=1}^{n} v_t X_{t-1} \Big/ \Big(\sum_{t=1}^{n} X_{t-1}^2\Big)^{\frac{1}{2}}$$

$$= \left(\frac{s^2(n)}{\sum_{t=1}^{n} X_{t-1}^2}\right) \frac{\sum_{t=1}^{n} v_t X_{t-1}}{s(n)} \xrightarrow{D} (V^2)^{-\frac{1}{2}}(GU + Hu + W)V$$

$$= GU + Hu + W,$$

which, when u_t are Gaussian mean zero, (and with finite positive variance) is a Gaussian r.v. with mean zero and a finite variance depending on the roots $(\rho, m_2, \cdots, m_k)$, q.e.d.

Some Remarks: If some $|m_j| \geqq 1$ for $2 \leqq j \leqq k$, then the conclusions of the above Theorem II and the corollaries need not be valid since, as seen in the proof, the fact that $|m_j| < 1$ is used in an essential way. If $k = 1$, the Corollary II_1 was proved by White [14] and the theorem and its corollaries by T. W. Anderson [1]. The result of Theorem I, also for $k = 1$, was proved by Rubin [13].

Sometimes it may be of interest to find a lower bound for the variance of $\tilde{\rho}_n$. However, because of Corollary II_1, variance need not be a good risk function to consider, and some more general risk functions as in [11], may be more appropriate. This problem will not be considered here.

PART II

7. Introduction. Let $\{X_t, t \geqq 1\}$ be a stochastic process satisfying the Assumptions 1 to 5 of Section 1. The problem considered in this part is the consistency of the L.S. (or M.L.) estimators of the "structural parameters", or the regression coefficients, of

$$X_t = \alpha_1 X_{t-1} + \cdots + \alpha_k X_{t-k} + u_t . \tag{29}$$

The ordinary "normal equations" for the estimators $\hat{\alpha}_i$ of α_i are

$$\sum_{i=1}^{k} \hat{\alpha}_i \sum_{t=[i,j]+1}^{n} X_{t-i}X_{t-j} = \sum_{t=j+1}^{n} X_t X_{t-j}, \quad j = 1, 2, \cdots, k, \tag{30}$$

where $[i, j] = \max(i, j)$. Introducing the notation

$$C_{ij}^n = \sum_{t=[i,j]+1}^{n} X_{t-i}X_{t-j}, \quad \text{and} \quad A_i^n = \sum_{t=i+1}^{n} u_t X_{t-i}, \tag{31}$$

the equations (30) can be written as

$$(C_{ij}^n)(\hat{\alpha}_n - \alpha)' = A^{n'}, \tag{32}$$

where $\hat{\alpha}_n = (\hat{\alpha}_1, \cdots, \hat{\alpha}_k)$ and $A^n = (A_1^n, \cdots, A_k^n)$ etc. Since for every fixed n, the inverse of (C_{ij}^n) exists, (32) may be written as

$$(\hat{\alpha}_n - \alpha)' = (C_{ij}^n)^{-1} A^{n'}. \tag{33}$$

The question considered in this part reduces to this: Is the following equation true?

$$\text{plim } (\hat{\alpha}_n - \alpha)' = \text{plim } (C_{ij}^n)^{-1} A^{n'} = 0. \tag{34}$$

To carry out the work, several auxiliary results are required, including the generalized versions of Lemmas 9 and 10. These are proved in Section 8, and the main problem is considered in Section 9.

8. Lemmas for Theorem III. The first two lemmas were stated for $i = j = 1$, $(k \geqq 1)$ in Part I. Employing the same notation, their general form is given here.

LEMMA 14: *The r.v.* $s(n)^{-1} A_i^n \overset{P}{=} a^n \rho^{-(i-1)} U_{n,i} V_n$.

PROOF: By equation (5),

$$s^{-1}(n) A_i^n = \frac{\lambda_1}{s(n)} \sum_{t=i+1}^{n} u_t X_{1,t-i} + \sum_{j=2}^{k} \lambda_j \sum_{t=i+1}^{n} \frac{u_t X_{j,t-i}}{s(n)} = A_i + B_i, \text{ say} \tag{35}$$

It will be shown that

$$B_i = \sum_{j=2}^{k} \lambda_j B_{ij} \xrightarrow{P} 0. \tag{36}$$

Since u_t and $X_{j,t-i}$ are independent, for all j, and $i \geqq 1$,

$$E(B_{ij}) = 0, \quad \text{all } i, j \geqq 1, \text{ and} \quad \operatorname{Var} B_{ij} = \frac{\sigma^2}{s^2(n)} \sum_{t=i+1}^{n} E(X_{j,t-i}^2).$$

But

$$E(X_{j,t-i}^2) = \sigma^2 \sum_{r=1}^{t-i} m_j^{2(t-i-r)} \leqq t\sigma^2 m^{2t} \times \text{constant}.$$

Hence, $\operatorname{Var} B_{ij} \leqq$ constant $\times\ \rho^{-2n} n^2 m_j^{2n} \to 0$, since $|m_j/\rho| < 1$. It follows that $B_{ij} \xrightarrow{P} 0$. Hence, $B_i \xrightarrow{P} 0$. Consider,

$$\begin{aligned} A_i &= \frac{\lambda_1}{s(n)} \sum_{t=i+1}^{n} u_t X_{1,t-i} \\ &= a^n \rho^{-(i-1)} \Big[\Big((\rho^2 - 1)^{\frac{1}{2}} \sum_{t=i+1}^{n} \rho^{-(n-t+1)} u_t \Big) \Big((\rho^2 - 1)^{\frac{1}{2}} \sum_{r=1}^{n} \rho^{-r} u_r \Big) \\ &\qquad - (\rho^2 - 1) \sum_{t=i+1}^{n} \rho^{-(n-t+1)} u_t \sum_{r=t-i+1}^{n} \rho^{-r} u_r \Big] \\ &= \rho^{-(i-1)} [a^n U_{n,i} V_n - B_i^*], \quad \text{say}. \end{aligned} \tag{37}$$

The lemma would follow if it is shown that

$$B_i^* \xrightarrow{P} 0, \; i = 1, \cdots, k. \tag{38}$$

Rewrite B_i^* as,

$$B_i^* = a^n (\rho^2 - 1) \rho^{-(n+1)} \Big[\sum_{t=i+1}^{n} u_{t-i+1} u_t + \rho^{-1} \sum_{t=i+1}^{n-1} u_{t-i+2} u_t + \cdots + \rho^{-(n-2)} u_n u_{i+1} \Big].$$

In the above each term has mean zero (except the one involving $\rho^{-(n+i)} \sum_{t=i+1}^{n-i} u_t^2$, which by Markov's inequality $\xrightarrow{P} 0$) and variance $\leqq n\rho^{-2n}$, and each component is independent of the other (excluding the squares term). Hence their variance is not greater than $n^2 \rho^{-2n} \to 0$. It follows that $B_i^* \xrightarrow{P} 0$. Hence $A_i \overset{P}{=} a^n \rho^{-(i-1)} U_{n,i} V_n$.

REMARK: The exact expression is

$$s^{-1}(n) A_i^n = a^n \rho^{-(i-1)} U_{n,i} V_n - \rho^{-(i-1)} B_i^* + B_i, \tag{39}$$

where $B_i^* \xrightarrow{P} 0$, $B_i \xrightarrow{P} 0$, and $|a| = 1$.

LEMMA 15: *The r.v.* $s^{-2}(n)C_{ij}^n \overset{P}{=} \rho^{-(i+j-2)}V_{n-\max(i,j)}^2$.

PROOF: Since

$$X_t = \sum_{i=1}^{k} \lambda_i X_{i,t} \tag{40}$$

it follows that

$$\begin{aligned} s^{-2}(n)C_{ij}^n &= s^{-2}(n) \sum_{t=[i,j]+1}^{n} X_{t-i}\, X_{t-j}\,, \\ &= \frac{\lambda_1^2}{s^2(n)} \sum_{t=[i,j]+1}^{n} X_{1,t-i}\, X_{1,t-j} + \frac{1}{s^2(n)} \sum_{t=[i,j]+1}^{n} \sum_{q,q'=1}^{k}{}^{*} \lambda_q \lambda_{q'}\, X_{q,t-i}\, X_{q',t-j}\,, \end{aligned}$$

where $[i, j] = \max(i, j)$, and where $\sum^*$ above stands for the sum in which $q = q' = 1$ is omitted; or

$$s^{-2}(n)C_{ij}^n = A_{ij} + \bar{Q}_{ij}\,, \quad \text{say.} \tag{41}$$

Note that the special case $i = j = 1$ has been stated as Lemma 9. The general case is proved here. Also some of the computations (e.g., of A_{ii}) given here will be needed later.

First consider A_{ii}.

$$\begin{aligned} A_{ii} &= \frac{(\rho^2 - 1)^2}{\rho^{2n}} \sum_{t=i+1}^{n} X_{1,t-i}^2 \\ &= \frac{(\rho^2 - 1)^2}{\rho^{2n}} \sum_{t=i+1}^{n} \left[\sum_{r=1}^{t-i} \rho^{2(t-r-i)} u_r^2 + \sum_{r \neq r'=1}^{t-i} \rho^{2(t-i)-r-r'} u_r\, u_{r'} \right] \\ &= A_i^* + B_i^{**}\,, \quad \text{say.} \end{aligned}$$

But A_i^* may be simplified as follows (more details can be found in [10]):

$$\begin{aligned} A_i^* &= \frac{\rho^2 - 1}{\rho^{2n}} \left[u_1^2 \rho^{2(n-i)} + u_2^2 \rho^{2(n-i-1)} + \cdots + \rho^{2(n-i-(n-i-1))} u_{n-i}^2 - \sum_{r=1}^{n-i} u_r^2 \right] \\ &= \frac{\rho^2 - 1}{\rho^{2i}} \sum_{r=1}^{n-i} \rho^{-2(r-1)} u_r^2 - \frac{\rho^2 - 1}{\rho^{2n}} \sum_{r=1}^{n-i} u_r^2 . \end{aligned} \tag{42}$$

Similarly B_i^{**} can be rewritten as,

$$\begin{aligned} B_i^{**} &= 2\, \frac{\rho^2 - 1}{\rho^{2n}} [\rho u_1\, u_2(\rho^{2(n-i-1)} - 1) \\ &\qquad + \rho^2 u_1 u_3(\rho^{2(n-i-2)} - 1) + \cdots + \rho u_{n-i}\, u_{n-i-1}(\rho^2 - 1)] \\ &= 2\, \frac{\rho^2 - 1}{\rho^{2i}} \sum_{\substack{r<r' \\ r,r'=1}}^{n-i} \rho^{-(r+r'-2)} u_r\, u_{r'} - 2\, \frac{\rho^2 - 1}{\rho^{2n}} \left[\sum_{r=1}^{n-i-1} u_r\, u_{r+1} \right. \\ &\qquad \left. + \rho^2 \sum_{r=1}^{n-i-2} u_r\, u_{r+2} + \cdots + \rho^{n-i-1} u_1\, u_{n-i} \right]. \end{aligned} \tag{43}$$

Substituting (42) and (43) in A_{ii},

$$A_{ii} = \frac{\rho^2 - 1}{\rho^{2i}} \left[\sum_{r=1}^{n-i} \rho^{-2(r-1)} u_r^2 + 2 \sum_{\substack{r<r' \\ r,r'=1}}^{n-i} \rho^{-(r+r'-2)} u_r u_{r'} \right] -$$

$$(44) \qquad \frac{\rho^2 - 1}{\rho^{2n}} \left[\sum_{r=1}^{n-i} u_r^2 + 2 \left(\rho \sum_{r=1}^{n-i-1} u_r u_{r+1} + \cdots + \rho^{n-i-1} u_1 u_{n-i} \right) \right]$$

$$= \frac{\rho^2 - 1}{\rho^{2i}} \left(\sum_{r=1}^{n-i} \rho^{-r+1} u_r \right)^2 - \bar{R}_i, \text{ say.}$$

$$(45) \qquad A_{ii} = \rho^{-2(i-1)} V_{n-i}^2 - \bar{R}_i,$$

where $\bar{R}_i$ is defined by

$$(46) \quad \bar{R}_i = \frac{\rho^2 - 1}{\rho^{2n}} \left[\sum_{r=1}^{n-i} u_r^2 + 2 \left(\rho \sum_{r=1}^{n-i-1} u_r u_{r+1} + \cdots + \rho^{n-i-1} u_1 u_{n-i} \right) \right].$$

It is not difficult to see that

$$(47) \qquad \bar{R}_i \xrightarrow{P} 0, \qquad i = 1, 2, \cdots, k.$$

Hence,

$$(48) \qquad A_{ii} \overset{P}{=} \rho^{-2(i-1)} V_{n-i}^2.$$

Using this result it will be shown that (cf. (41), for definition of $\bar{Q}_{ij}$)

$$(49) \qquad \bar{Q}_{ij} \xrightarrow{P} 0, \qquad i, j = 1, \cdots, k.$$

By Schwarz' inequality,

$$(50) \quad \left| \frac{1}{s^2(n)} \sum_{t=[i,j]+1}^{n} X_{q,t-i} X_{q',t-j} \right| \leqq \left(\frac{1}{s^2(n)} \sum_{t=i+1}^{n} X_{q,t-i}^2 \right)^{\frac{1}{2}} \left(\frac{1}{s^2(n)} \sum_{t=j+1}^{n} X_{q',t-j}^2 \right)^{\frac{1}{2}}.$$

For $q > 1$,

$$E\left(s^{-2}(n) \sum_{t=i+1}^{n} X_{q,t-i}^2 \right) = \frac{(\rho^2 - 1)^2}{\lambda_1^2 \rho^{2n}} \sum_{t=i+1}^{n} E\left(\sum_{r=1}^{t-i} m_q^{t-i-r} u_r \right)^2$$

$$\leqq n^2 m_q^{2n} \rho^{-2n} \times \text{constant} \to 0,$$

since $|m_q/\rho| < 1$. Hence, the r.v.'s being non-negative, it follows that,

$$(51) \qquad s^{-2}(n) \sum_{t=i+1}^{n} X_{q,t-i}^2 \xrightarrow{P} 0, \qquad i = 1, \cdots, k,\ q > 1.$$

If both q, $q' > 1$, then the right side of (50) $\xrightarrow{P} 0$, so that $\bar{Q}_{ij} \xrightarrow{P} 0$. If $q = 1$ (so that $q' > 1$), then the right side of (50) has a factor

$$s^{-2}(n) \sum_{t=i+1}^{n} X_{1,t-i}^2 = \lambda_1^{-2} A_{ii} \overset{P}{=} \rho^{-2(i-1)} V_{n-i}^2, \qquad \xrightarrow{P} \rho^{-2(i-1)} V^2,$$

by (48) and Lemma 8. Hence by Lemma 1 (for products), it follows that

$\bar{Q}_{ij} \xrightarrow{P} 0$. To complete the proof of this lemma, it remains to consider A_{ij} for $i \neq j$. Since $A_{ij} = A_{ji}$, let $i > j$. Then, from the definition of A_{ij} (cf. (41)),

$$\begin{aligned} A_{ij} &= (\rho^2 - 1)^2 \rho^{-2n} \sum_{t=i+1}^{n} \left(\sum_{r=1}^{t-i} \rho^{t-i-r} u_r \right)\left(\sum_{r=1}^{t-j} \rho^{t-j-r} u_r \right) \\ &= \frac{(\rho^2 - 1)^2}{\rho^{2n}} \rho^{i-j} \sum_{t=i+1}^{n} X_{1,t-i}^2 \\ &\qquad + \frac{(\rho^2 - 1)^2}{\rho^{2n}} \sum_{t=i+1}^{n} \left(\sum_{r=1}^{t-i} \rho^{t-i-r} u_r \right)\left(\sum_{r=1}^{i-j} \rho^{i-j-r} u_{t-i-r} \right) \\ &= \rho^{i-j} A_{ii} + \bar{S}_{ij}, \end{aligned}$$

where A_{ii} is the same as in (41), and $\bar{S}_{ij}$ is the second term above, and for symmetry let $\bar{S}_{ii} = 0$. Hence using (45), A_{ij} can be written as

$$\begin{aligned} A_{ij} &= \rho^{i-j}[\rho^{-2(i-1)} V_{n-i}^2 - \bar{R}_i] + \bar{S}_{ij} \\ &= \rho^{-(i+j-2)} V_{n-i}^2 - \rho^{i-j}\bar{R}_i + \bar{S}_{ij}. \end{aligned} \tag{52}$$

Since it is shown that $\bar{R}_i \xrightarrow{P} 0$, $\bar{Q}_{ij} \xrightarrow{P} 0$, the lemma would follow if it is shown that

$$\bar{S}_{ij} \xrightarrow{P} 0. \tag{53}$$

After a slight rearrangement one obtains

$$\begin{aligned} \bar{S}_{ij} = \frac{(\rho^2 - 1)^2}{\rho^{2n}} [&(u_1 u_{i-j+1} + u_2 u_{i-j+2} + \cdots + u_{n-i} u_{n-j}) \\ &+ \rho(u_1 u_{i-j} + u_1 u_{i-j+2} + \cdots + u_{n-i} u_{n-j-1}) \\ &+ \cdots + \rho^{i-j+1}(u_1 u_2 + \cdots + u_{n-i} u_{n-j+1}) + \cdots + \rho^{n-j-2} u_1 u_{n-i+1}]. \end{aligned} \tag{54}$$

Now all terms in the square bracket on the right of (54) have means zero, and the variance of the whole expression is bounded by the quantity

$$\begin{aligned} (\rho^2 - 1)^4 \rho^{-4n}[(n - i) + \rho^2(n - i - 1) + \cdots + \rho^{2(n-j-2)} \cdot 1] \\ \leqq n^2 \rho^{-2n} \times \text{constant} \to 0. \end{aligned}$$

Hence it follows that $\bar{S}_{ij} \xrightarrow{P} 0$, $i, j = 1, \cdots, k$. Therefore, from (41), using (52), one obtains

$$s^{-2}(n) C_{ij}^n = A_{ij} + \bar{Q}_{ij} \stackrel{P}{=} \rho^{-(i+j-2)} V_{n-[i,j]}^2. \tag{55}$$

REMARK: The exact expression is the following: $(i \geqq j)$

$$s^{-2}(n) C_{ij}^n = \rho^{-(i+j-2)} V_{n-i}^2 - \rho^{i-j}\bar{R}_i + \bar{S}_{ij} + \bar{Q}_{ij} \tag{56}$$

where $\bar{R}_i \xrightarrow{P} 0$, $\bar{S}_{ij} \xrightarrow{P} 0$, and $\bar{Q}_{ij} \xrightarrow{P} 0$.

In the sequel a second order stochastic difference equation will be considered in detail. In that connection, it would be of interest to know more about $\bar{R}_i$ and $\bar{S}_{ij}$. More precisely, LEMMA 16: *The r.v.'s* $s(n)\bar{R}_i$ *and* $s(n)\bar{S}_{ij}$ *are bounded in probability.*

PROOF: Since $s(n) = \lambda_1 |\rho|^n/(\rho^2 - 1)$, from (46) above, one notes that the first term on the right of $s(n)\bar{R}_i \xrightarrow{P} 0$, by Markov's inequality. The second term, given in (57), will be shown to be bounded in probability.

$$(57) \qquad \lambda_1 |\rho|^{-n} \left[\rho \sum_{t=1}^{n-i-1} u_t u_{t+1} + \rho^2 \sum_{t=1}^{n-i-2} u_t u_{t+2} + \cdots + \rho^{n-i-1} u_1 u_{n-i} \right]$$

is composed of terms, each with mean zero, which are independent of each other. The variance of (57) is

$$(58) \quad \lambda_1^2 \rho^{-2n} \sigma^4 [\rho^2 (n - i - 1) + \rho^4 (n - i - 2) + \cdots + \rho^{2(n-i-1)}] \leqq M < \infty,$$

where M is independent of n. It follows that (57), and hence $s(n)\bar{R}_i$, is bounded in probability.

For $\bar{S}_{ij}$, the case $k = 2$ will be considered. It is the only case that is required. Then from definition,

$$(59) \quad s(n)\bar{S}_{21} = \frac{\lambda_1}{\rho} \frac{\rho^2 - 1}{|\rho|^n} \left[\rho \sum_{r=1}^{n-2} u_r\, u_{r+1} + \rho^2 \sum_{r=1}^{n-3} u_r\, u_{r+2} + \cdots + \rho^{n-2} u_1\, u_{n-1} \right].$$

But this is a special case of (57) (for $i = 1$), so that $s(n)\bar{S}_{ij}$ is bounded in probability, q.e.d.

LEMMA 17: *If* $|m_q| < 1$, *the r.v.* $\lambda_1 s^{-1}(n) \sum_{t=[i,j]+1}^{n} X_{1,t-i} X_{q,t-j}$ *is bounded in probability* $(q > 1)$.

PROOF: First consider the case $i \geqq j$,

$$(60) \quad \begin{aligned} &\frac{\lambda_1}{s(n)} \sum_{t=i+1}^{n} X_{1,t-i}\, X_{q,t-j} \\ &\qquad = a^n \rho^{-i+1} (\rho^2 - 1)^{\frac{1}{2}} \sum_{t=i+1}^{n} \rho^{-(n-t+1)} X_{q,t-j} (\rho^2 - 1)^{\frac{1}{2}} \sum_{r=1}^{t-i} \rho^{-r} u_r \\ &\qquad = a^n \rho^{-(i-1)} [\bar{A}_i\, V_{n-i} - T_{ij}^n], \end{aligned}$$

where V_{n-i} is defined in (6), $\bar{A}_i = (\rho^2 - 1)^{\frac{1}{2}} \sum_{t=i+1}^{n} \rho^{-(n-t+)} X_{q,t-j}$, and

$$T_{ij}^n = (\rho^2 - 1) \sum_{t=i+1}^{n} \rho^{-(n-t+1)} X_{q,t-j} \sum_{r=t-i+1}^{n-i} \rho^{-r} u_r\,.$$

Note that $E(\bar{A}_i) = 0$, and since $|m_q| < 1$ and $|\rho| > 1$,

$$\operatorname{Var} \bar{A}_i \leqq \frac{\rho^2 - 1}{\rho^2} \left[\sum_{t=i+1}^{n} |\rho|^{-(n-t)} \text{ S.D.}(X_{q,t-j}) \right]^2 \leqq M < \infty.$$

Hence $\bar{A}_i$ is bounded in probability.

It is known from Lemma 8 that V_{n-i} is a r.v. which is bounded in probability for $i = 1, \cdots, k$.

Next, as in (38), it is seen that $|E(T_{ij}^n)| \leqq (\rho^2 - 1) n^2 |\rho|^{-n-2} \to 0$, since $|m_q| < 1 < |\rho|$, and

$$\operatorname{Var} T_{ij}^n \leqq \rho^{-2n} [n + (n - 1) + \cdots + 1]^2 \leqq \text{constant} \times n^4 \rho^{-2n} \to 0.$$

Hence $T_{ij}^n \xrightarrow{P} 0$, $i, j = 1, \cdots, k$, by Lemma 5.

It follows that the r.v. in (60) is bounded in probability. To complete the proof of the lemma, it remains to consider the case $i < j$. Then, a rearrangement shows,

$$(61)\quad \begin{aligned} &\frac{\lambda_1}{s(n)} \sum_{t=j+1}^{n} X_{1,t-i} X_{q,t-j} \\ &= \rho^{j-i} \frac{\lambda_1}{s(n)} \sum_{t=j+1}^{n} X_{q,t-j} X_{1,t-j} + \frac{\lambda_1}{s(n)} \sum_{t=j+1}^{n} X_{q,t-j} \sum_{r=1}^{j-i} \rho^{j-i-r} u_{t-j+r} . \end{aligned}$$

The first term on the right of (61) is the same as that in (60) with $i = j$ and hence is bounded in probability. The second term, because $|m_q| < 1$, $\xrightarrow{P} 0$, since it has mean zero and variance $\to 0$. (Compare with $s(n)\bar{S}_{21}$ of (59).)

LEMMA 18: *If $|m_q| < 1$, then $s(n)\bar{Q}_{ij}$ is a r.v., bounded in probability, where*

$$\bar{Q}_{ij} = \frac{1}{s^2(n)} \sum_{t=[i,j]+1}^{n} \sum_{\substack{q,q'=1 \\ (q=q'=1 \text{ not allowed})}}^{k} \lambda_q \lambda_{q'} X_{q,t-i} X_{q',t-j} .$$

PROOF: Only the case $k = 2$ is needed below, and it will be considered.

$$(62)\quad \begin{aligned} s(n)\bar{Q}_{ij} &= \frac{\lambda_1 \lambda_2}{s(n)} \sum_{t=[i,j]+1}^{n} (X_{1,t-i} X_{2,t-j} + X_{2,t-i} X_{1,t-j}) \\ &\quad + \frac{\lambda_2^2}{s(n)} \sum_{t=[i,j]+1}^{n} X_{2,t-i} X_{2,t-j} . \end{aligned}$$

Since $|m_q| < 1$, the first term on the right of (62) is bounded in probability by Lemma 17. The second term is easily seen to converge stochastically to zero, q.e.d.

In proving the consistency of the estimators of the regression coefficients a simplification of the following expression, given as the final lemma, is *all*-important.

LEMMA 19: *If $|m_2| < 1 < |\rho|$, then $s^2(n)\ (\bar{Q}_{11} - 2\rho\bar{Q}_{21} + \rho^2\bar{Q}_{22})/n$ converges in probability to* $[2\lambda_1\lambda_2\sigma^2 + \lambda_2^2\sigma^2(1 - 2\rho m_2 + \rho^2)/(1 - m_2^2)]$.

PROOF: From the definition of $\bar{Q}_{ij}$ (cf. (41)), for $k = 2$, it is seen that

$$(63)\quad \begin{aligned} \bar{Q}_{ij} = s^{-2}(n) \Bigg[&\lambda_1 \lambda_2 \sum_{t=[i,j]+1}^{n} (X_{1,t-i} X_{2,t-j} + X_{2,t-i} X_{1,t-j}) \\ &+ \lambda_2^2 \sum_{t=[i,j]+1}^{n} X_{2,t-i} X_{2,t-j} \Bigg] = a_{ij} + b_{ij} , \quad \text{say.} \end{aligned}$$

Thus, from (63),

$$(\bar{Q}_{11} + \rho^2\bar{Q}_{22} - 2\rho\bar{Q}_{21}) = (a_{11} + \rho^2 a_{22} - 2\rho a_{21}) + (b_{11} + \rho^2 b_{22} - 2\rho b_{21}).$$

All the simplifications depend on the fact that $|m_2| < 1 < |\rho|$. If u_t were assumed to have four moments, then the results of Mann and Wald ([8], p. 182) imply that $s^2(n)b_{ij}/n \xrightarrow{P} \lim_{n\to\infty} E(s^2(n)b_{ij}/n) = M < \infty$, where M is given in (64) below.

$$s^2(n)b_{ij}/n = (\lambda_2^2/n) \sum_{t=i+1}^{n} X_{2,t-i}X_{2,t-j}, \quad \text{for } i \geqq j, \qquad \text{since } b_{ij} = b_{ji},$$

$$= (\lambda_2^2/n) \sum_{t=i+1}^{n} \left(\sum_{r=1}^{t-i} m_2^{t-i-r} u_r\right)\left(\sum_{r=1}^{t-j} m_2^{t-j-r} u_r\right).$$

Hence $E(s^2(n)b_{ij}/n) = (\lambda_2^2/n) m_2^{i-j} \sigma^2 \sum_{t=i+1}^{n} \sum_{r=1}^{t-i} m_2^{2(t-i-r)}$, so that

$$(64) \qquad \lim_{n\to\infty} E(s^2(n)b_{ij}/n) = \lambda_2^2 m_2^{i-j}\sigma^2/(1 - m_2^2) = M.$$

However, this result can be obtained with the assumption of only two moments for the u_t, as assumed in this paper. This will be proved here ($i \geqq j$, $i, j = 1, 2$).

Consider

$$\frac{\lambda_2^2}{n} \sum_{t=i+1}^{n} X_{2,t-i} X_{2,t-j}$$

$$(65) \qquad = \frac{\lambda_2^2}{n} \sum_{t=i+1}^{n} X_{2,t-i} \left[m_2^{i-j} \sum_{r=1}^{t-i} m_2^{t-i-r} u_r + \sum_{r=t-i+1}^{t-j} m_2^{t-j-r} u_r \right]$$

$$= \frac{\lambda_2^2}{n} m_2^{i-j} \sum_{t=i+1}^{n} X_{2,t-i}^2 + \frac{\lambda_2^2}{n} \sum_{t=i+1}^{n} X_{2,t-i} \left(\sum_{r=1}^{i-j} m_2^{i-j-r} u_{t-i+r} \right).$$

The last term on the right drops out for $i = j$, and if $i = 2$, $j = 1$, it becomes, $(1/n) \sum_{t=i+1}^{n} u_{t-1}X_{2,t-i} \xrightarrow{P} 0$, since it has mean zero, and variance of the order $1/n$, as $|m_2| < 1$.

Consider $\bar{A}_{ii} = \frac{1}{n} m_2^{i-j} \sum_{t=i+1}^{n} X_{2,t-i}^2$. Comparing $\bar{A}_{ii}$ with A_{ii} of (44), one obtains, on identifying m_2 with ρ,

$$(66) \qquad \bar{A}_{ii} = \frac{1}{n} m_2^{i-j} \left[\frac{m_2^2}{m_2^2 - 1} \left(\sum_{r=1}^{n-i} m_2^{n-i-r} u_r \right)^2 - \frac{1}{m_2^2 - 1} \bar{\bar{R}}_i \right],$$

where $\bar{\bar{R}}_i$ is given in (66′). But the first term in (66), on the right, is non-negative and has a mean that tends to zero since $|m_2| < 1$. Hence, it converges in probability to zero. Consequently, $\bar{A}_{ii} \stackrel{P}{=} m_2^{i-j}(1 - m_2^2)^{-1} n^{-1} \bar{\bar{R}}_i$, and

$$(66') \qquad \bar{\bar{R}}_i = \sum_{r=1}^{n-i} u_r^2 + 2 \left[m_2 \sum_{r=1}^{n-i-1} u_r u_{r+1} + \cdots + m_2^{n-i-1} u_1 u_{n-i} \right].$$

Notice that the terms in square brackets of $\bar{\bar{R}}_i$ have means zero, and the variance of the r.v. in [ ] is bounded by a constant times n since $|m_2| < 1$. Hence

$$(67) \qquad n^{-1} \left[m_2 \sum_{r=1}^{n-i-1} u_r u_{r+1} + \cdots + m_2^{n-i-1} u_1 u_{n-i} \right] \xrightarrow{P} 0.$$

On the other hand, by the strong law of large numbers,

$$(68) \qquad n^{-1} \sum_{r=1}^{n-1} u_r^2 \to E(u_r^2) = \sigma^2 > 0, \quad \text{with probability one.}$$

Therefore, (65) simplifies to (64). From this it follows that

(69) $(s^2(n)/n)(b_{11} + \rho^2 b_{22} - 2\rho b_{21}) \xrightarrow{P} [\lambda_2^2\sigma^2/(1 - m_2^2)][1 + \rho^2 - 2\rho m_2]$.

The following algebraic identity may be verified. (Details can be found in [10], p. 89.)

(70)
$$\frac{s^2(n)}{n}(a_{11} + \rho^2 a_{22} - 2\rho a_{21}) = 2\lambda_1\lambda_2\left[n^{-1}\sum_{r=1}^{n-1} u_r^2 + \frac{(m_2 - \rho)}{n}\sum_{t=3}^{n} u_{t-1}X_{2,t-2}\right].$$

The second term on the right has mean zero, variance of the order (n^{-1}), so that this r.v. $\xrightarrow{P} 0$. On the other hand, $n^{-1}\sum_{r=1}^{n-1} u_r^2 \to \sigma^2$ with probability one. Hence,

(71) $$(s^2(n)/n)(a_{11} + \rho^2 a_{22} - 2\rho a_{21}) \xrightarrow{P} 2\lambda_1\lambda_2\sigma^2.$$

Combining (71) with (69), one obtains

$$\frac{s^2(n)}{n}(\bar{Q}_{11} + \rho^2\bar{Q}_{22} - 2\rho\bar{Q}_{21}) \xrightarrow{P} [\lambda_2^2\sigma^2/(1 - m_2^2)][1 + \rho^2 - 2\rho m_2] + 2\lambda_1\lambda_2\sigma^2.$$

9. Consistency of the Estimators $\hat{\alpha}_i$. The previous lemmas enable the presentation of the main problem of this part. The complete statement of the theorem is given for convenience. The details are given for the second order stochastic difference equation and hence the statement is given only for that case.

THEOREM III: *Let a process $\{X_t, t \geqq 1\}$ satisfy the following conditions:*

CONDITION 1: *For each t, $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + u_t$, where α_1, α_2 are finite real constants to be estimated, and the u_t (t positive) are independent, identically distributed (with a continuous distribution) having mean zero, and a finite positive variance, σ^2.*

CONDITION 2: *The roots m_1, m_2 of the characteristic equation $m^2 - \alpha_1 m - \alpha_2 = 0$ are simple* (i.e., $m_1 \neq m_2$), *with one root,* say $\rho = m_1$, *satisfying* $|\rho| > 1 > |m_2|$.

CONDITION 3: *For $t \leqq 0$, $u_t = 0$.*

Then, it follows that the L.S. estimators $\hat{\alpha}_i$ of α_i (see eq. (32)) *are consistent, i.e.,* plim $(\hat{\alpha}_i - \alpha_i) = 0$.

PROOF: For convenience C_{ij} and A_i will be written instead of C_{ij}^n, and A_i^n. Then the "normal equations" for $\hat{\alpha}_i$ given by (33), can be written explicitly as

(72) $$(\hat{\alpha}_1 - \alpha_1) = \frac{C_{22}A_1 - C_{12}A_2}{C_{11}C_{22} - C_{12}^2},$$

(73) $$(\hat{\alpha}_2 - \alpha_2) = \frac{C_{11}A_2 - C_{12}A_1}{C_{11}C_{22} - C_{12}^2}.$$

It suffices to consider one of these equations, say (72). Then (72) may be written as

(74) $$(\hat{\alpha}_1 - \alpha_1) = s^{-1}(n)\left[\frac{C_{22}}{s^2(n)}\frac{A_1}{s(n)} - \frac{C_{12}}{s^2(n)}\frac{A_2}{s(n)}\right] \Big/ \left[\frac{C_{11}}{s^2(n)}\frac{C_{22}}{s^2(n)} - \left(\frac{C_{12}}{s^2(n)}\right)^2\right],$$

where, as usual, the normalizing factor $s(n) = \lambda_1|\rho|^n/(\rho^2 - 1)$. From equation (39),

$$A_i/s(n) = a^n\rho^{-(i-1)}U_{n,i}V_n - \rho^{-(i-1)}B_i^* + B_i\,,\ a = \rho/|\rho|, \tag{75}$$

where $U_{n,i}$, V_n, B_i^* and B_i are defined in, respectively, Lemma 14, equations (38) and (36), and it was shown that $B_i^* \xrightarrow{P} 0$ and $B_i \xrightarrow{P} 0$.

Similarly from (56), for $i \geqq j$,

$$s^{-2}(n)C_{ij} = \rho^{-(i+j-2)}V_{n-i}^2 - \rho^{(i-j)}\bar{R}_i + \bar{S}_{ij} + \bar{Q}_{ij}\,, \tag{76}$$

the quantities $\bar{R}_i$, $\bar{S}_{ij}$, and $\bar{Q}_{ij}$ are defined in (46), (53) and (41).

The numerator and denominator of (74) will be considered separately. Call them N_n and D_n. It will be shown in the following that $(s(n)/n)N_n \xrightarrow{P} 0$, and $(s(n)/n)D_n \xrightarrow{P}$ to a positive r.v. The theorem follows by an application of Lemma 1 (for ratios). The detailed steps follow:

I. Noting that $C_{12} = C_{21}$

$$N_n = \frac{C_{22}}{s^2(n)}\frac{A_1}{s(n)} - \frac{C_{21}}{s^2(n)}\frac{A_2}{s(n)}\,,$$

$$\begin{aligned} &= a^n\rho^{-2}V_{n-2}^2V_n(U_{n,1} - U_{n,2}) + \rho^{-2}V_{n-2}^2(B_1 - \rho B_2 - B_1^* + B_2^*) \\ &\quad - (\bar{R}_2 - \bar{Q}_{22})(a^nU_{n,1}V_n - B_1^* + B_2^*) \\ &\quad + (\rho\bar{R}_2 - \bar{S}_{21} - \bar{Q}_{21})(a^n\rho^{-1}U_{n,1}V_n - \rho^{-1}B_2^* + B_2). \end{aligned} \tag{77}$$

II. From the definition of $U_{n,i}$,

$$U_{n,1} = (\rho^2 - 1)^{\frac{1}{2}}\rho^{-(n-1)}u_2 + U_{n,2}\,. \tag{78}$$

Hence, from (77),

$$a^n(U_{n,1} - U_{n,2})V_nV_{n-2}^2 = a^nV_{n-2}^2V_n(\rho^2 - 1)^{\frac{1}{2}}\rho^{-(n-1)}u_2\,. \tag{79}$$

From (79), it follows that

$$(s(n)/n)a^n(U_{n,1} - U_{n,2})V_nV_{n-2}^2 \xrightarrow{P} 0. \tag{80}$$

III. Consider $s(n)(\bar{R}_2 - \bar{Q}_{22})(a^nU_{n,1}V_n - B_1^* + B_2)$. Here the fact that $|m_2| < 1$ will be used. Thus from Lemmas 15 and 18 it follows that $s(n)(\bar{R}_2 - \bar{Q}_{22})$ is bounded in probability, and $U_{n,1}V_n$ is also bounded in probability (from the definition), while $B_1^* \xrightarrow{P} 0$ and $B_1 \xrightarrow{P} 0$ (cf. eq. (75)). Consequently, by Lemma 1,

$$(s(n)/n)(\bar{R}_2 - \bar{Q}_{22})(a^nU_{n,1}V_n - B_1^* + B_1) \xrightarrow{P} 0. \tag{81}$$

IV. Consider $s(n)(\rho\bar{R}_2 - \bar{S}_{21} - \bar{Q}_{21})(a^n\rho^{-1}U_{n,1}V_n - \rho^{-1}B_2^* + B_2)$. Since $|m_2| < 1$, as in step III, by Lemmas 15 and 18, it follows that (because the r.v. in IV is bounded in probability)

$$(s(n)/n)(\rho\bar{R}_2 - \bar{Q}_{21} - \bar{S}_{21})(a^n\rho^{-1}U_{n,1}V_n - \rho^{-1}B_2^* + B_2) \xrightarrow{P} 0. \tag{82}$$

V. Finally consider $s(n)\rho^{-2}V_{n-2}^2(B_1 - \rho B_2 - B_1^* + B_2^*)$. The following algebraic identity may be verified. (See [10], Appendix 1, for details.)

(83) $$B_2^* - B_1^* = a^n \frac{\rho^2 - 1}{\rho^2} \rho^{-(n-2)} \left[\sum_{r=1}^{n-1} u_r u_{r+1} - \left(\frac{\rho^2}{\rho^2 - 1} \right)^{\frac{1}{2}} u_2 V_n \right].$$

Notice that $s(n)(B_2^* - B_1^*)$ is not necessarily bounded in probability. But,

$$n^{-1} s(n)(B_2^* - B_1^*) = \lambda_1 \left[n^{-1} \sum_{r=1}^{n-1} u_r u_{r+1} - n^{-1} u_2 V_n (\rho^2 - 1)^{\frac{1}{2}} \rho^{-1} \right].$$

Clearly $n^{-1} u_2 V_n \xrightarrow{P} 0$, and $u_r u_{r+1}$, $r = 1, 2, \cdots$, are independent identically distributed r.v.'s with means zero. Thus, $n^{-1} \sum_{r=1}^{n-1} u_r u_{r+1} \to$ its mean $(= 0)$ with probability one. Hence

(84) $$(s(n)/n)(B_2^* - B_1^*) \xrightarrow{P} 0.$$

Also notice that $s(n)B_i = \lambda_2 \sum_{t=i+1}^{n} u_t X_{2,t-i}$, $E(s(n)B_i) = 0$, and since $|m_2| < 1$, $E(s(n)B_i/n)^2 = 0(n^{-1})$, $i = 1, 2$, so that $(s(n)B_i/n) \xrightarrow{P} 0$, implying

(85) $$n^{-1} s(n)(B_1 - \rho B_2) \xrightarrow{P} 0.$$

From (84) and (85), it is seen that

(86) $$(s(n)/n)\rho^{-2} V_{n-2}^2 (B_1 - \rho B_2 - B_1^* + B_2^*) \xrightarrow{P} 0.$$

Consequently, from (77), (80)–(82) and (86), $(s(n)N_n/n) \xrightarrow{P} 0$, if $|m_2| < 1 < |\rho|$.

VI. Next consider the denominator D_n,

$$D_n = s(n) \left[\frac{C_{11}}{s^2(n)} \frac{C_{22}}{s^2(n)} - \left(\frac{C_{12}}{s^2(n)} \right)^2 \right] = s(n) \bar{D}_n, \quad \text{say.}$$

It will be shown that $(s^2(n)/n)\bar{D}_n \xrightarrow{P}$ positive r.v. Consider $(C_{21} = C_{12})$,

$$\bar{D}_n = \frac{C_{11}}{s^2(n)} \frac{C_{22}}{s^2(n)} - \left(\frac{C_{21}}{s^2(n)} \right)^2 = (V_{n-1}^2 - \bar{R}_1 + \bar{Q}_{11})(\rho^{-2} V_{n-2}^2 - \bar{R}_2 + \bar{Q}_{22})$$
$$- (\rho^{-1} V_{n-2}^2 - \rho \bar{R}_2 + \bar{S}_{21} + \bar{Q}_{21})^2,$$

from (75). Now adding and subtracting $V_{n-2}^2(\bar{R}_2 - \bar{Q}_{22})$ suitably, one gets, after rearrangement,

(87) $$\begin{aligned} \bar{D}_n = \rho^{-2} V_{n-2}^2 [((V_{n-1}^2 - V_{n-2}^2) - (\bar{R}_1 - \rho^2 \bar{R}_2 + 2\rho \bar{S}_{21})) \\ + (\bar{Q}_{11} - 2\rho \bar{Q}_{21} + \rho^2 \bar{Q}_{22})] - (V_{n-1}^2 - V_{n-2}^2)(\bar{R}_2 - \bar{Q}_{22}) \\ + (\bar{R}_1 - \bar{Q}_{11})(\bar{R}_2 - \bar{Q}_{22}) - (\rho \bar{R}_2 - \bar{S}_{21} - \bar{Q}_{21})^2. \end{aligned}$$

VII. It is not difficult to see that the following identity obtains (see [10], Appendix II, for details):

(88) $$\frac{s^2(n)}{n} [(V_{n-1}^2 - V_{n-2}^2) - (\bar{R}_1 - \rho^2 \bar{R}_2 + 2\rho \bar{S}_{21})] = \lambda_1^2 \sum_{r=1}^{n-1} u_r^2/n.$$

Since the u_r^2 are independent identically distributed with means σ^2, it follows that

$$(89)\qquad \lambda_1^2 \sum_{r=1}^{n-1} u_r^2/n \to \lambda_1^2\sigma^2 > 0, \quad \text{with probability one.}$$

VIII. It was shown in Lemma 19 that

$$(90)\qquad \frac{s^2(n)}{n}\,[\bar{Q}_{11} - 2\rho\bar{Q}_{21} + \rho^2\bar{Q}_{22}] \xrightarrow{P} 2\lambda_1\lambda_2\sigma^2 + \frac{\lambda_2^2\sigma^2}{1-m_2^2}(1+\rho^2-2m\rho).$$

Thus the term in square brackets on the right side of (87), multiplied by $s^2(n)/n$, converges in probability to the following limit.

$$(91)\qquad \lambda_1^2\sigma^2 + 2\lambda_1\lambda_2\sigma^2 + \frac{\lambda_2^2\sigma^2}{1-m_2^2}(1+\rho^2-2m\rho) = \sigma^2 + \frac{\lambda_2^2\sigma^2}{1-m_2^2}(\rho-m_2)^2,$$

since $\lambda_1 + \lambda_2 = 1$. Note that this constant is strictly positive.

IX. The following statements are immediate consequences of Lemmas 16 and 18.

$$(92)\qquad n^{-1}[s^2(n)(V_{n-1}^2 - V_{n-2}^2)(\bar{R}_2 - \bar{Q}_{22})] \xrightarrow{P} 0,$$

$$(93)\qquad (s^2(n)/n)(\bar{R}_1 - \bar{Q}_{11})(\bar{R}_2 - \bar{Q}_{22}) \xrightarrow{P} 0,$$

$$(94)\qquad (s^2(n)/n)(\bar{R}_2 - \bar{S}_{21} - \bar{Q}_{21})^2 \xrightarrow{P} 0.$$

Summarizing the work in steps VI–IX it can be inferred that (on noting $\rho^{-2}V_{n-2}^2 \to \rho^{-2}V^2$ ($\neq 0$) with probability one, cf., Lemma 8),

$$(95)\qquad \frac{s^2(n)}{n}\bar{D}_n \xrightarrow{P} \rho^{-2}V^2\left[\sigma^2 + \frac{\lambda_2^2\sigma^2}{1-m_2^2}(\rho-m_2)^2\right],$$

which is a positive r.v.

Hence, by Lemma 1 (for ratios) it follows that $\hat{\alpha}_1 - \alpha_1 = N_n/D_n \xrightarrow{P} 0$, if $|m_2| < 1 < |\rho|$, q.e.d.

10. Remarks on Theorem III.

1. The assumption that the other root $|m_2| < 1$ is used to get the required bounds in probability for the r.v.'s B_i and $\bar{Q}_{ij}$ when multiplied by $s(n)/n$ and $s^2(n)/n$. This is only a sufficient condition. It is probably true that the consistency of $\hat{\alpha}_i$ holds without any restriction on the other roots (i.e., other than the maximum), but due to the computational difficulties these relaxations were not attempted.

2. The long route followed in the proof was necessitated by the fact that the numerator as well as the denominator $\xrightarrow{P} 0$. *The usual assumption, that the matrix $(s^{-2}(n)C_{ij})$ is non-singular in the limit, is not tenable here, and the classical procedure is not applicable.*

Proof of 2: Let $M_n = (C_{ij}^n)$ and $s(n) = \lambda_1|\rho|^n/(\rho^2-1)$, as usual.

$$s^{-2}(n)M_n = (s^{-2}(n)C_{ij}^n) \overset{P}{=} (\rho^{-(i+j-2)}V_{n-\max(i,j)}^2), \qquad \text{by Lemma 15,}$$

$$\xrightarrow{P} (\rho^{-(i+j-2)})V^2, \qquad \text{by Lemma 8.}$$

But the right side is a singular matrix of rank one. Note that only $|\rho| > |m_j|$ and $|\rho| > 1$ are used here.

3. From this theorem the following important conclusion obtains. The nonsingularity of the matrix $(s^{-2}(n)C_{ij})$, for all n, is not a necessary condition for the consistency of the estimators $\hat{\alpha}_i$ of the regression coefficients α_i.

4. If $|\rho|, |m| > 1$ (i.e., all roots are greater than one in absolute value), then T. W. Anderson's results [1] imply the consistency of $\hat{\alpha}_i$. Theorem III and Anderson's results [1], taken together, exhaust all cases for the second order difference equations if the roots have distinct moduli. But the question is still open in higher order cases ($k \geqq 3$).

5. It is apparent that the assumption $|\rho| > 1 > |m_2|$ implies that the terms involving m_2 are bounded in probability, while those involving ρ "disturb the stability" of the process. For the same reason, it is clear that the condition $|\rho| > 1 > |m_j|, j = 2, \cdots, k$, with an arbitrary but finite (known) k, is sufficient to state the results of Theorem III for higher order difference equations. In other words, $(\hat{\alpha}_i - \alpha_i) \xrightarrow{P} 0$, $i = 1, \cdots, k$, if $|\rho| > 1 > |m_j|, j = 2, \cdots, k$.

6. If the difference equation is of second or higher order, k, and the maximal root ρ is greater than one in absolute value, and $|m_j| = 1$ for any $j = 2, \cdots, k$, then no result is available on the consistency of $\hat{\alpha}_i$. Also nothing is known about the estimators of the α_i of an explosive linear stochastic difference equation if there is a constant term. The following table summarizes the available results.

Table of the Available Results

kth order	Roots $(\rho, m_2, \cdots, m_k)$	Results [(i) = consistency, (ii) = efficiency and (iii) = limit distribution]
$k = 1$	$\|\rho\| >, =, < 1$	(i) Rubin [13], Mann-Wald [8] (i), (ii) Theorem 4:A of [10] (iii) White [14], T. W. Anderson [1], and Mann-Wald [8]
$k = 2$	$\|\rho\| > 1 > \|m_2\|$	(i) Theorems I and III for $\hat{\rho}_n$ and $\hat{\alpha}_i$ (iii) Theorem II for $\hat{\rho}_n$
	$\|\rho\|, \|m_2\| > 1$	(i), (iii) T. W. Anderson [1]
	$\|\rho\| < 1$	(i), (iii) Mann-Wald [8]
	$\|\rho\| = 1$ or $\|m_2\| = 1$	no result available
$k \geqq 3$	$\|\rho\| > 1 > \|m_j\|$	(iii) Theorem II for $\hat{\rho}_n$ (i) Theorem III and remark 5, for $\hat{\alpha}_i$
	$\|\rho\| > \|m_j\|, \|\rho\| > 1$	(i) Theorem I for $\hat{\rho}_n$, no result for $\hat{\alpha}_i$
	$\|\rho\|, \|m_j\| > 1$	(i), (iii) T. W. Anderson [1] for $\hat{\alpha}_i$
	$\|\rho\| < 1$	(i), (iii) Mann-Wald [8]
	$\|\rho\| < 1$, u_t Gaussian	(ii) Rubin [12] (i), (ii) Theorem 3:C of [10]

Table continues on next page

$\lvert\rho\rvert = 1$ or $\lvert m_j\rvert = 1$ or $\lvert\rho\rvert > \lvert m_2\rvert > 1 > \lvert m_j\rvert$	no result available.

11. Acknowledgments. I wish to express my gratitude to Professors L. Hurwicz and I. R. Savage, for the direction and encouragement throughout the course of writing this paper. I am indebted to Professors M. D. Donsker and B. R. Gelbaum for many helpful comments and encouragement in the preparation of the paper. Also I am thankful to the editor, W. Kruskal, for helpful remarks on stylistic problems which have improved the presentation, involving revisions of the manuscript, of the paper.

REFERENCES

[1] T. W. Anderson, "On asymptotic distribution of estimates of parameters of stochastic difference equations," *Ann. Math. Stat.*, Vol. 30 (1959), pp. 676–687.

[2] Herman Chernoff, "Large-sample theory: parametric case," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 1–22.

[3] Harald Cramér, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1946.

[4] J. L. Doob, *Stochastic Processes*, John Wiley and Sons, New York, 1953.

[5] Charles Jordan, *Calculus of Finite Differences*, Röttig and Romwalter, Budapest, 1939.

[6] Michel Loève, *Probability Theory*, D. Van Nostrand, New York, 1955.

[7] Z. A. Lomnicki and S. K. Zaremba, "On some moments and distributions occurring in the theory of linear stochastic processes-I," *Monatshefte Math.*, Vol. 61 (1957), pp. 318–358.

[8] H. B. Mann and A. Wald, "On the statistical treatment of linear stochastic difference equations," *Econometrica*, Vol. 11 (1943), pp. 173–220.

[9] H. B. Mann and A. Wald, "On stochastic limit and order relationships," *Ann. Math. Stat.*, Vol. 14 (1943), pp. 217–226.

[10] M. M. Rao, "Properties of maximum likelihood estimators in nonstable stochastic difference equations," Technical Report No. 7, Department of Statistics, University of Minnesota, July 1959.

[11] M. M. Rao, "Lower bounds for risk functions in estimation," *Proc. Nat. Acad. Sci.*, Vol. 45 (1959), pp. 1168–1171.

[12] Herman Rubin, "Systems of linear stochastic equations," unpublished dissertation, University of Chicago, 1948.

[13] Herman Rubin, "Consistency of maximum likelihood estimates in the explosive case," in *Statistical Inference in Dynamic Economic Models*, Ed. T. C. Koopmans, Cowles Commission Monograph No. 10, John Wiley and Sons, New York (1950), pp. 356–364.

[14] John S. White, "The limiting distribution of the serial correlation coefficient in the explosive case I, II," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 1188–1197, Vol. 30 (1959), pp. 831–834.