Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT

Statistics & Probability Letters 65 (2003) 71–77

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

Spatial autoregression model: strong consistency

B.B. Bhattacharyya[a], J.-J. Ren[b], G.D. Richardson[b,*], J. Zhang[b]

[a]*Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203, USA*
[b]*Department of Mathematics, University of Central Florida, Orlando, FL 32816-1364, USA*

Received February 2003; received in revised form March 2003

Abstract

Let $(\hat{\alpha}_n, \hat{\beta}_n)$ denote the Gauss–Newton estimator of the parameter (α, β) in the autoregression model $Z_{ij} = \alpha Z_{i-1,j} + \beta Z_{i,j-1} - \alpha\beta Z_{i-1,j-1} + \varepsilon_{ij}$. It is shown in an earlier paper that when $\alpha = \beta = 1$, $\{n^{3/2}(\hat{\alpha}_n - \alpha, \hat{\beta}_n - \beta)\}$ converges in distribution to a bivariate normal random vector. A two-parameter strong martingale convergence theorem is employed here to prove that $n^r(\hat{\alpha}_n - \alpha, \hat{\beta}_n - \beta) \to \underline{0}$ almost surely when $r < \frac{3}{2}$.

MSC: 62F12; 60F05; 62M30

Keywords: Spatial autoregression; Unit roots; Two-parameter martingale

1. Introduction and preliminaries

Martin (1979) introduced the spatial autoregression model

$$Z_{ij} = \alpha Z_{i-1,j} + \beta Z_{i,j-1} - \alpha\beta Z_{i-1,j-1} + \varepsilon_{ij} \tag{1.1}$$

and indicated its applicability in a subsequent paper in 1990. The above model is used in the study of image processing by Jain (1981) and in the analysis of digital filtering and system theory by Tjostheim (1981). Basu and Reinsel (1994) illustrate the feasibility of (1.1) being nonstationary with a practical example. Moreover, Cullis and Gleeson (1991) include model (1.1) with $\alpha = \beta = 1$ in the class of models used to represent the error structure in linear regression to analyze field data. When $|\alpha| < 1$ and $|\beta| < 1$, various estimators of (α, β) are known to converge in distribution to a bivariate normal, where the normalizing term is n (Basu, 1990; Khalil, 1991). A similar conclusion is proved by Bhattacharyya (1995) when $\alpha = \beta = 1$ by using a Gauss–Newton estimator $(\hat{\alpha}_n, \hat{\beta}_n)$ and

* Corresponding author. Tel.: +1-407-823-2753; Fax: +1-407-823-6253.
E-mail addresses: bhattach@stat.ncsu.edu (B.B. Bhattacharyya), garyr@pegasus.cc.ucf.edu (G.D. Richardson).

0167-7152/$ - see front matter
doi:10.1016/j.spl.2003.07.004

a normalizing term $n^{3/2}$. It is proved here that $n^r(\hat{\alpha}_n - \alpha, \hat{\beta}_n - \beta) \to \underline{0}$ almost surely when $\alpha = \beta = 1$ and $r < \frac{3}{2}$.

It is assumed without further mention that $Z_{ij} = 0$ when $i \wedge j \leqslant 0$. The following axioms are listed below for future reference.

Assumptions.

(A.1) $\alpha = \beta = 1$
(A.2) ε_{ij}, $i, j \geqslant 1$, are i.i.d., each having mean zero, variance σ^2 and a finite fourth moment
(A.3) $E|\sum_{k=1}^{i} \varepsilon_{k1}|^{4+2\delta} = O(i^{2+\delta})$ for some $\delta > 0$, where ε_{ij} are i.i.d. and each has mean zero.
(A.4) $\bar{\alpha}_n$ and $\bar{\beta}_n$ are initial estimators obeying $n^r(\bar{\alpha}_n - \alpha) \to 0$ and $n^r(\bar{\beta}_n - \beta) \to 0$ almost surely when $r < \frac{3}{2}$.

Remark 1.1. Assume model (1.1) obeys (A.1)–(A.3). According to (1.4), $X_{ij} = \alpha X_{i-1,j} + \varepsilon_{ij}$; define $\bar{\alpha}_n$ to be the least squares estimator of α. Since $\alpha = \beta = 1, X_{ij} = Z_{ij} - Z_{i,j-1}$ is observable and, moreover, when $r < \frac{3}{2}$, $n^r(\bar{\alpha}_n - \alpha) = n^{r-3} \sum_{i,j=1}^{n} X_{i-1,j}\varepsilon_{ij} / n^{-3} \sum_{i,j=1}^{n} X_{i-1,j}^2 \to 0$ a.s. by Theorem 2.2(i) and Theorem 2.4(i). The estimator $\bar{\beta}_n$ is defined similarly and hence this establishes the existence of initial estimators in axiom (A.4) whenever (A.1)–(A.3) are fulfilled.

Let $\theta_0' = (\alpha, \beta), \bar{\theta}_n' = (\bar{\alpha}_n, \bar{\beta}_n), f_{ij}(a, b) = aZ_{i-1,j} + bZ_{i,j-1} - abZ_{i-1,j-1}, F_{ij}'(a, b) = (\partial f_{ij}\,(a, b)/\partial a, \partial f_{ij}\,(a, b)/\partial b) = (Z_{i-1,j} - bZ_{i-1,j-1}, Z_{i,j-1} - aZ_{i-1,j-1})$ and $R_{ij}(a, b) = -(\alpha - a)(\beta - b)Z_{i-1,j-1}$. Denote $\hat{\delta}_n = [\sum_{i,j=1}^{n} F_{ij}(\bar{\theta}_n)\ F_{ij}'(\bar{\theta}_n)]^{-1} \sum_{i,j=1}^{n} F_{ij}(\bar{\theta}_n)(Z_{ij} - f_{ij}(\bar{\theta}_n))$; then $\hat{\theta}_n = \hat{\delta}_n + \bar{\theta}_n$ is said to be the "Gauss–Newton estimator" of θ_0. It is shown by Bhattacharyya (1995) that $\hat{\theta}_n$ obeys

$$\hat{\theta}_n - \theta_0 = \left[\sum_{i,j=1}^{n} F_{ij}(\bar{\theta}_n)F_{ij}'(\bar{\theta}_n)\right]^{-1} \sum_{i,j=1}^{n} F_{ij}(\bar{\theta}_n)(R_{ij}(\bar{\theta}_n) + \varepsilon_{ij}). \tag{1.2}$$

Define $X_{ij} = Z_{ij} - \beta Z_{i,j-1}, Y_{ij} = Z_{ij} - \alpha Z_{i-1,j}$ and observe that

$$F_{ij}'(\bar{\theta}_n) = (X_{i-1,j} + (\beta - \bar{\beta}_n)Z_{i-1,j-1}, Y_{i,j-1} + (\alpha - \bar{\alpha}_n)Z_{i-1,j-1}). \tag{1.3}$$

Moreover, using model (1.1),

$$X_{ij} = \alpha X_{i-1,j} + \varepsilon_{ij} \quad \text{and thus } X_{ij} = \sum_{k=1}^{i} \varepsilon_{kj} \text{ when } \alpha = \beta = 1. \tag{1.4}$$

Likewise,

$$Y_{ij} = \beta Y_{i,j-1} + \varepsilon_{ij} \quad \text{and} \quad Y_{ij} = \sum_{l=1}^{j} \varepsilon_{il} \text{ when } \alpha = \beta = 1. \tag{1.5}$$

The principal result of this work is stated below and proved in subsequent sections.

Theorem 1.2. *Assume that model* (1.1) *obeys axioms* (A.1)–(A.3); *then* $n^r(\hat{\theta}_n - \theta_0) \to \underline{0}$ *almost surely when* $r < \frac{3}{2}$.

Even though (A.3) implies the existence of fourth moments of ε_{ij}, condition (A.2) is listed separately since some results proved here are valid under (A.1)–(A.2). Note that (A.3) is satisfied when $\varepsilon_{ij},\ i,j \geqslant 1$, are i.i.d. and each is distributed as $N(0,\sigma^2)$, which leads to the corollary below. More generally, axiom (A.3) is fulfilled with $\delta=1$ when (A.2) is valid and $E|\varepsilon_{11}|^6 < \infty$.

Corollary 1.3. *Suppose that $\alpha=\beta=1$ in model (1.1) and $\varepsilon_{ij},\ i,j \geqslant 1$, are i.i.d. and each is distributed as $N(0,\sigma^2)$. Then $n^r(\hat{\theta}_n - \theta_0) \to \underline{0}$ almost surely when $r < \frac{3}{2}$.*

2. Strong martingale

Almost sure convergence results are established in this section by appealing to a martingale convergence theorem for doubly indexed processes due to Walsh (1979, 1986). Let N^2 denote the set of all ordered pairs of positive integers and define $\underline{s}=(s_1,s_2) \leqslant (t_1,t_2)=\underline{t}$ iff $s_i \leqslant t_i$ and $\underline{s} < \underline{t}$ provided $s_i < t_i,\ i=1,2$. Suppose that $(\Omega,\mathfrak{F},P)$ denotes the underlying probability space and $\mathfrak{F}_{\underline{t}},\ \underline{t}\in N^2$, is an increasing sequence of sub-σ-fields of $\mathfrak{F}$; that is, $\mathfrak{F}_{\underline{s}} \subseteq \mathfrak{F}_{\underline{t}} \subseteq \mathfrak{F}$ when $\underline{s} \leqslant \underline{t}$. Moreover, assume that $V_{\underline{t}},\ \underline{t}\in N^2$, is an $\mathfrak{F}_t$-measurable random variable. Then $\{V_{\underline{t}},\mathfrak{F}_{\underline{t}},\ \underline{t}\in N^2\}$ is called a *strong martingale* if it obeys the following conditions:

(a) $E[V_{\underline{t}}|\mathfrak{F}_{\underline{s}}]=V_{\underline{s}}$ when $\underline{s} \leqslant \underline{t}$ (martingale)
(b) $E[V(\underline{s},\underline{t}]|\mathfrak{F}^*_{\underline{s}}]=0$ provided $\underline{s} < \underline{t}$,

where $V(\underline{s},\underline{t}]=V_{\underline{t}} - V_{s_1t_2} - V_{t_1s_2} + V_{\underline{s}}$ and $\mathfrak{F}^*_{\underline{s}}$ denotes the smallest σ-field containing each $\mathfrak{F}_{ij}$ with either $i \leqslant s_1$ or $j \leqslant s_2$. Walsh (1979, 1986) proved that an L^1-bounded strong martingale converges almost surely.

An extension of Kronecker's lemma to two indices is given below.

Lemma 2.1. *Assume that $\{x_{ij} : i,j \geqslant 1\}$ and $\{c_{ij} : i,j \geqslant 0\}$ are sequences of real numbers, where $c_{ij}=0$ when $i \wedge j=0$, satisfying the following conditions:*

(a) $\nabla c_{ij}=c_{ij}-c_{i-1,j}-c_{i,j-1}+c_{i-1,j-1}>0$ *when* $i \wedge j \geqslant 1$
(b) $c_{Mn}/c_{nn} \to 0$ *and* $c_{nM}/c_{nn} \to 0$ *as* $n \to \infty$, *for each fixed* $M \geqslant 1$
(c) $y_{mn}=\sum_{k,l=1}^{m,n} x_{kl}/c_{kl} \to y$ *as* $m \wedge n \to \infty$ *and* $\{y_{mn} : m,n \geqslant 1\}$ *is bounded.*

Then $(1/c_{nn})\sum_{i,j=1}^{n} x_{ij} \to 0$ *as* $n \to \infty$.

Proof. Let $y_{ij}=\sum_{k=1}^{i}\sum_{l=1}^{j} x_{kl}/c_{kl}$ when $i\wedge j \geqslant 1$ and $y_{ij}=0$ elsewhere. Denote $\nabla y_{ij}=y_{ij}-y_{i-1,j}-y_{i,j-1}+y_{i-1,j-1}$ and thus $\nabla y_{ij}=x_{ij}/c_{ij}$ when $i \wedge j \geqslant 1$. Then $\sum_{i,j=1}^{n} x_{ij}=\sum_{i,j=0}^{n} c_{ij}y_{ij}-\sum_{i,j=0}^{n} c_{ij}y_{i-1,j}-\sum_{i,j=0}^{n} c_{ij}y_{i,j-1}+\sum_{i,j=0}^{n} c_{ij}y_{i-1,j-1}=\sum_{i,j=0}^{n} c_{ij}y_{ij}-\sum_{i=0}^{n-1}\sum_{j=0}^{n} c_{i+1,j}y_{ij}-\sum_{i=0}^{n}\sum_{j=0}^{n-1} c_{i,j+1}y_{ij}+\sum_{i,j=0}^{n-1} c_{i+1,j+1}y_{ij}=\sum_{i,j=0}^{n-1} y_{ij}\nabla c_{i+1,j+1}-\sum_{i=0}^{n-1}(c_{i+1,n}-c_{in})y_{in}-\sum_{j=0}^{n-1}(c_{n,j+1}-c_{nj})y_{nj}+c_{nn}y_{nn}$. Define $U_n=(1/c_{nn})\sum_{i,j=0}^{n-1} y_{ij}\nabla c_{i+1,j+1}$, $V_n=(1/c_{nn})\sum_{i=0}^{n-1}(c_{i+1,n}-c_{in})y_{in}$ and $W_n=(1/c_{nn})\sum_{j=0}^{n-1}(c_{n,j+1}-c_{nj})y_{nj}$.

Observe that since $\nabla c_{ij} > 0$ and $c_{ij}=0$ when $i\wedge j=0$, it follows that $c_{ij} > 0$ provided $i\wedge j > 0$ and $c_{ij} < c_{st}$ if $i \leqslant s, j \leqslant t$ and (i,j) is distinct from (s,t). First, it is shown that $U_n \to y$ as $n \to \infty$. Since $\sum_{i,j=0}^{n-1} \nabla c_{i+1,j+1}=c_{nn}$, $|U_n-y|=(1/c_{nn})|\sum_{i,j=0}^{n-1}(y_{ij}-y)\nabla c_{i+1,j+1}| \leqslant (1/c_{nn})\sum_{i,j=0}^{N-1}|y_{ij}-y|\nabla c_{i+1,j+1}+(1/c_{nn})\sum_{i=0}^{N-1}\sum_{j=N}^{n-1}|y_{ij}-y|\nabla c_{i+1,j+1}+(1/c_{nn})\sum_{i=N}^{n-1}\sum_{j=0}^{N-1}|y_{ij}-y|\nabla c_{i+1,j+1}+(1/c_{nn})\sum_{i,j=N}^{n-1}|y_{ij}-y|\nabla c_{i+1,j+1}$. Assumption (c) implies that for $\varepsilon > 0, |y_{ij}-y| < \varepsilon$ when $i \wedge j \geqslant N$ and also since $\{y_{ij}\}$ is bounded, $|y_{ij}-y| \leqslant M$ for all $i \wedge j \geqslant 1$. Then $|U_n - y| \leqslant (M/c_{nn})c_{NN} + (M/c_{nn})c_{Nn} + (M/c_{nn})c_{nN} + (\varepsilon/c_{nn})(c_{nn} - c_{Nn} - c_{nN} + c_{NN})$ and thus by assumption (b), $U_n \to y$ as $n \to \infty$. Likewise, $|V_n - y| = (1/c_{nn})|\sum_{i=0}^{n-1}(c_{i+1,n} - c_{in})(y_{in} - y)| \leqslant (1/c_{nn})|\sum_{i=0}^{N-1}(c_{i+1,n} - c_{in})(y_{in} - y)| + (1/c_{nn})|\sum_{i=N}^{n-1}(c_{i+1,n} - c_{in})(y_{in} - y)| \leqslant (M/c_{nn})c_{Nn} + (\varepsilon/c_{nn})(c_{nn} - c_{Nn})$ and thus $V_n \to y$ as $n \to \infty$. Similarly, $W_n \to y$ as $n \to \infty$ and thus by assumption (c), $(1/c_{nn})\sum_{i,j=1}^{n} x_{ij} \to 0$ as $n \to \infty$. □

Recall from (1.4) and (1.5), when $\alpha = \beta = 1$ in model (1.1), $X_{ij} = \sum_{k=1}^{i} \varepsilon_{kj}, Y_{ij} = \sum_{l=1}^{j} \varepsilon_{il}$ and, moreover, $Z_{ij} = \sum_{k,l=1}^{i,j} \varepsilon_{kl}$.

Theorem 2.2. *Suppose that model* (1.1) *and axioms* (A.1)–(A.2) *are fulfilled. Then*

(i) $\frac{1}{n^r}\sum_{i,j=1}^{n} X_{i-1,j}\varepsilon_{ij} \to 0$ *a.s. when* $r > \frac{3}{2}$
(ii) $\frac{1}{n^r}\sum_{i,j=1}^{n} Y_{i,j-1}\varepsilon_{ij} \to 0$ *a.s. when* $r > \frac{3}{2}$
(iii) $\frac{1}{n^r}\sum_{i,j=1}^{n} Z_{i-1,j-1}\varepsilon_{ij} \to 0$ *a.s. when* $r > 2$.

Proof. (i) Define, for each $\underline{t}=(t_1,t_2)\in N^2, V_{\underline{t}}=\sum_{i,j=1}^{t_1,t_2} X_{i-1,j}\varepsilon_{ij}$ and let $\mathfrak{F}_{\underline{t}}$ be the smallest σ-field for which each ε_{ij} is measurable, $i \leqslant t_1, j \leqslant t_2$. First, it is shown that $\{V_{\underline{t}},\mathfrak{F}_{\underline{t}},\ \underline{t}\in N^2\}$ is a martingale. Fix $\underline{s}=(s_1,s_2) \leqslant (t_1,t_2)=\underline{t}$; employing standard properties of conditional expectation,

$$E[V_{\underline{t}}\,|\,\mathfrak{F}_{\underline{s}}] = V_{\underline{s}} + \sum_{i=1,j=s_2+1}^{t_1,t_2} E(X_{i-1,j})E(\varepsilon_{ij}) + \sum_{i=s_1+1,j=1}^{t_1,s_2} E[X_{i-1,j}\varepsilon_{ij}\,|\,\mathfrak{F}_{\underline{s}}]$$

$$= V_{\underline{s}} + \sum_{i=s_1+1,j=1}^{t_1,s_2}\left[\sum_{k=1}^{s_1}\varepsilon_{kj}E(\varepsilon_{ij}\,|\,\mathfrak{F}_{\underline{s}}) + \sum_{k=s_1+1}^{i-1}E(\varepsilon_{kj}\varepsilon_{ij})\right] = V_{\underline{s}}$$

by (A.2). Hence $\{V_{\underline{t}},\mathfrak{F}_{\underline{t}},\ \underline{t}\in N^2\}$ is a martingale.

It remains to show that the above is a strong martingale. Assume that $\underline{s}=(s_1,s_2) < (t_1,t_2)=\underline{t}$; then $V(\underline{s},\underline{t})=\sum_{i=s_1+1,j=s_2+1}^{t_1,t_2} X_{i-1,j}\varepsilon_{ij}$. Hence $E[V(\underline{s},\underline{t})\,|\,\mathfrak{F}_{\underline{s}}^*]=\sum_{i=s_1+1,j=s_2+1}^{t_1,t_2} E[X_{i-1,j}\varepsilon_{ij}\,|\,\mathfrak{F}_{\underline{s}}^*]=\sum_{i=s_1+1,j=s_2+1}^{t_1,t_2}\left[\sum_{k=1}^{s_1}\varepsilon_{kj}E(\varepsilon_{ij})+\sum_{k=s_1+1}^{i-1}E(\varepsilon_{kj})E(\varepsilon_{ij})\right]=0$, and thus $\{V_{\underline{t}},\mathfrak{F}_{\underline{t}},\ \underline{t}\in N^2\}$ is a strong martingale. Define $W_{\underline{t}}=\sum_{i,j=1}^{t_1,t_2} X_{i-1,j}\varepsilon_{ij}/i^p j^q$, where $p > 1$ and $q > \frac{1}{2}$. Then $\{W_{\underline{t}},\mathfrak{F}_{\underline{t}},\ \underline{t}\in N^2\}$ is also a strong martingale and, moreover, $E(W_{\underline{t}}^2) = \sum_{i,j=1}^{t_1,t_2} E(X_{i-1,j}^2)\sigma^2/i^{2p}j^{2q} = \mathrm{O}(1)$. According to Walsh (1986, Corollary 2.8), $W_{\underline{t}} \to W$ almost surely as $t_1 \wedge t_2 \to \infty$, for some W. Employing Lemma 2.1 with $c_{ij} = i^p j^q$, it follows that $(1/n^r)\sum_{i,j=1}^{n} X_{i-1,j}\varepsilon_{ij} \to 0$ almost surely when $r > \frac{3}{2}$. Parts (ii)–(iii) are verified using a similar argument. □

The next result is due to Hu et al. (1989).

Theorem 2.3. *Assume that* $\{V_{nj} : 1 \leqslant j \leqslant n, n \geqslant 1\}$ *is an array of random variables satisfying*:

(i) $E(V_{nj}) = 0$
(ii) $V_{n1}, V_{n2}, \ldots, V_{nn}$ *are independent, for each fixed* $n \geqslant 1$
(iii) $\sup_{n,j} E|V_{nj}|^{2p+\delta} < \infty$, *for some* $1 \leqslant p < 2$ *and* $\delta > 0$.

Then $(1/n^{1/p}) \sum_{j=1}^{n} V_{nj} \to 0$ *almost surely.*

Theorem 2.4. *Suppose that model* (1.1) *obeys axioms* (A.1)–(A.3). *Then*

(i) $\frac{1}{n^3} \sum_{i,j=1}^{n} X_{i-1,j}^2 \to \frac{\sigma^2}{2}$ *a.s.*

(ii) $\frac{1}{n^3} \sum_{i,j=1}^{n} Y_{i,j-1}^2 \to \frac{\sigma^2}{2}$ *a.s.*

Proof. (i) Denote $V_{nj}=(1/n^2) \sum_{i=1}^{n} [X_{i-1,j}^2 - E(X_{i-1,j}^2)]$; Theorem 2.3 is used to show that $(1/n) \sum_{j=1}^{n} V_{nj} \to 0$ a.s. Since $X_{ij}=\sum_{k=1}^{i} \varepsilon_{kj}$, axiom (A.2) implies that $V_{n1}, V_{n2}, \ldots, V_{nn}$ are i.i.d. random variables and thus Theorem 2.3(ii) is satisfied. It follows from axiom (A.3) that $E|V_{nj}|^{2+\delta} \leqslant (C_1/n^{2(2+\delta)})n^{1+\delta} \sum_{i=1}^{n} E|X_{i-1,j}|^{4+2\delta} \leqslant (C_2/n^{3+\delta}) \sum_{i=1}^{n} i^{2+\delta} = \mathrm{O}(1)$. Employing Theorem 2.3 with $p=1$, $(1/n^3) \sum_{i,j=1}^{n} [X_{i-1,j}^2 - E(X_{i-1,j}^2)] \to 0$ a.s. Since $E(X_{i-1,j}^2) = (i-1)\sigma^2$, $(1/n^3) \sum_{i,j=1}^{n} E(X_{i-1,j}^2) \to \sigma^2/2$ and thus $(1/n^3) \sum_{i,j=1}^{n} X_{i-j}^2 \to \sigma^2/2$ a.s. Part (ii) is proved in a similar manner. □

3. Proof of Theorem 1.2

Bhattacharyya (1995) proves that $\{n^{3/2}(\hat{\theta}_n - \theta_0)\}$ converges in distribution to a bivariate normal random vector. Under assumptions (A.1)–(A.3), it is shown below that $n^r(\hat{\theta}_n - \theta_0) \to \underline{0}$ almost surely when $r < \frac{3}{2}$.

Lemma 3.1. *Suppose that model* (1.1) *and axioms* (A.1)–(A.2) *are satisfied. Then*

(i) $\operatorname{var}\left(\sum_{i,j=1}^{n} X_{i-1,j}^2\right) = \mathrm{O}(n^5)$,
(ii) $\operatorname{var}\left(\sum_{i,j=1}^{n} Y_{i,j-1}^2\right) = \mathrm{O}(n^5)$,
(iii) $\operatorname{var}\left(\sum_{i,j=1}^{n} Z_{i-1,j-1}^2\right) = \mathrm{O}(n^8)$,
(iv) $\operatorname{var}\left(\sum_{i,j=1}^{n} X_{i-1,j} Y_{i,j-1}\right) = \mathrm{O}(n^4)$,
(v) $\operatorname{var}\left(\sum_{i,j=1}^{n} X_{i-1,j} Z_{i-1,j-1}\right) = \mathrm{O}(n^6)$,
(vi) $\operatorname{var}\left(\sum_{i,j=1}^{n} Y_{i,j-1} Z_{i-1,j-1}\right) = \mathrm{O}(n^6)$.

Proof. Only verification of (v) is given here since the other arguments are similar. Since $X_{i-1,j} = \sum_{k=1}^{i-1} \varepsilon_{kj}$ and $Z_{i-1,j-1} = \sum_{k=1}^{i-1} \sum_{l=1}^{j-1} \varepsilon_{kl}$ are independent, mean zero random variables,

$\mathrm{cov}(X_{i-1,j}Z_{i-1,j-1}, X_{i'-1,j'}Z_{i'-1,j'-1}) = 0$ unless $j = j'$. Moreover, $\mathrm{cov}(X_{i-1,j}Z_{i-1,j-1}, X_{i'-1,j}Z_{i'-1,j-1}) = E(X_{i-1,j}X_{i'-1,j})E(Z_{i-1,j-1}Z_{i'-1,j-1}) = \mathrm{O}((i \wedge i')^2 j)$ and when $i \leqslant i'$, $\mathrm{var}\left(\sum_{i,j=1}^{n} X_{i-1,j}Z_{i-1,j-1}\right) = \sum_{i'=1}^{n} \sum_{i=1}^{i'} \sum_{j=1}^{n} \mathrm{O}(i^2 j) = \mathrm{O}(n^6)$. Therefore $\mathrm{var}\left(\sum_{i,j=1}^{n} X_{i-1,j}Z_{i-1,j-1}\right) = \mathrm{O}(n^6)$. □

Let $\bar{\theta}_n = (\bar{\alpha}_n, \bar{\beta}_n)$ denote the initial estimator defined in Remark 3.2, $A_n = \mathrm{diag}(n^{-3/2}, n^{-3/2})$, $A = \mathrm{diag}(\sigma^2/2, \sigma^2/2)$, $G_n = \sum_{i,j=1}^{n} F_{ij}(\bar{\theta}_n)F'_{ij}(\bar{\theta}_n)$ and $R_{ij}(\bar{\theta}_n) = -(\alpha - \bar{\alpha}_n)(\beta - \bar{\beta}_n)Z_{i-1,j-1}$, where $F'_{ij}(\bar{\theta}_n)$ is given in (1.3).

Lemma 3.3. *Suppose that model* (1.1) *and axioms* (A.1)–(A.3) *are fulfilled. Then*

(i) $A_nG_nA_n \to A$ *a.s.*
(ii) $A_n \sum_{i,j=1}^{n} F_{ij}(\bar{\theta}_n)R_{ij}(\bar{\theta}_n) \to \underline{0}$ *a.s.*
(iii) $n^{-\delta}A_n \sum_{i,j=1}^{n} F_{ij}(\bar{\theta}_n)\varepsilon_{ij} \to \underline{0}$ *a.s. when* $\delta > 0$.

Proof. (i) Denote

$$A_nG_nA_n = \begin{bmatrix} b_n & c_n \\ c_n & d_n \end{bmatrix}.$$

Then $b_n = (1/n^3)\sum_{i,j=1}^{n} [X_{i-1,j} + (\beta - \bar{\beta}_n)Z_{i-1,j-1}]^2 = (1/n^3)\sum_{i,j=1}^{n} X_{i-1,j}^2 + (2/n^3)(\beta - \bar{\beta}_n)\sum_{i,j=1}^{n} X_{i-1,j} Z_{i-1,j} + ((\beta - \bar{\beta}_n)^2/n^3)\sum_{i,j=1}^{n} Z_{i-1,j-1}^2 := V_{n1} + V_{n2} + V_{n3}$. According to Theorem 2.4(i), $V_{n1} \to \sigma^2/2$ a.s. It follows from Lemma 3.1(v) and the Borel–Cantelli Lemma that $(1/n^{7/2+\delta})\sum_{i,j=1}^{n} X_{i-1,j}Z_{i-1,j-1} \to 0$ a.s. when $\delta > 0$, and hence by Remark 1.1, $n^{3/2-\delta}(\beta - \bar{\beta}_n)(1/n^{7/2+\delta})\sum_{i,j=1}^{n} X_{i-1,j}Z_{i-1,j-1} \to 0$ a.s. In particular, $V_{n2} \to 0$ a.s. A similar argument using Lemma 3.1(iii) and the Borel–Cantelli Lemma shows that $V_{n3} \to 0$ a.s. Hence $b_n \to \sigma^2/2$ a.s.; likewise, $c_n \to 0$ a.s. and $d_n \to \sigma^2/2$ a.s.

(ii) Verification is similar to part (i).
(iii) Appealing to (1.3), $n^{-\delta}A_n \sum_{i,j=1}^{n} F_{ij}(\bar{\theta}_n)\varepsilon_{ij} = n^{-3/2-\delta}\left(\sum_{i,j=1}^{n} X_{i-1,j}\varepsilon_{ij}, \sum_{i,j=1}^{n} Y_{i,j-1}\varepsilon_{ij}\right)' + n^{-3/2-\delta} \sum_{i,j=1}^{n} Z_{i-1,j-1}\varepsilon_{ij}(\beta - \bar{\beta}_n, \alpha - \bar{\alpha}_n)' \to \underline{0}$ a.s. by Theorem 2.2 and Remark 1.1. □

Verification of Theorem 1.2 now follows immediately from Lemma 3.3. Indeed, employing (1.2), $n^{-\delta}A_n^{-1}(\hat{\theta}_n - \theta_0) = (A_nG_nA_n)^{-1}n^{-\delta}A_n\sum_{i,j=1}^{n} F_{ij}(\bar{\theta}_n)(R_{ij}(\bar{\theta}_n) + \varepsilon_{ij}) \to A^{-1}\underline{0} = \underline{0}$ a.s.

4. Further reading

Martin (1990) may also be of interest to the reader.

Acknowledgements

The authors are grateful to M. Amezziane for bringing Hu et al.'s work to our attention.

References

Basu, S., 1990. Analysis of first-order spatial bilateral ARMA models. Ph.D. Dissertation, University of Wisconsin, Madison, WI.

Basu, S., Reinsel, G.C., 1994. Regression models with spatially correlated errors. J. Amer. Statist. Assoc. 89, 88–99.

Bhattacharyya, B.B., Khalil, T.M., Richardson, G.D., 1995. Gauss–Newton estimation of parameters for a spatial autoregression model. Statist. Probab. Lett. 28, 173–179.

Cullis, B.R., Gleeson, A.C., 1991. Spatial analysis of field experiments—an extension to two dimensions. Biometrics 47, 1449–1460.

Hu, T.-C., Móricz, F., Taylor, R.L., 1989. Strong laws of large numbers for arrays of rowwise independent random variables. Acta Math. Hungar. 54, 153–162.

Jain, A.K., 1981. Advances in mathematical models for image processing. Proc. IEEE 69, 502–528.

Khalil, T.M., 1991. A study of the doubly geometric process, stationary cases and a nonstationary case. Ph.D. Dissertation, North Carolina State University, Raleigh, NC.

Martin, R.J., 1979. A subclass of lattice processes applied to a problem in planar sampling. Biometrika 66, 209–217.

Martin, R.J., 1990. The use of time-series models and methods in the analysis of agricultural field trials. Comm. Statist. Theory Methods 19, 55–81.

Tjostheim, D., 1981. Autoregressive modeling and spectral analysis of array data in the plane. IEEE Trans. Geosci. Remote Sensing 19, 15–24.

Walsh, J.B., 1979. Convergence and regularity of multiparameter strong martingales. Z. Wahrsch. Verw. Gebiete 46, 177–192.

Walsh, J.B., 1986. Martingales with a multidimensional parameter and stochastic integrals. Lecture Notes in Mathematics, Vol. 1215. Springer, New York, pp. 329–491.