The central limit theorem regarding cubic forms of random noise in spatial process

Ming-Yu Deng[1], Levent Kutlu[2], and Mingxi Wang[1]

[1]School of International Trade and Economics, University of International Business and Economics, Beijing 100029, China

[2]Department of Economics, University of Texas Rio Grande Valley, TX, USA

This draft: June 15, 2023

Let $\{X_{i,n}, 1 \leq i \leq k_n, n \geq 1\}$ ($k_n \to \infty$ as $n \to \infty$) be an array of random variables defined on a probability space $(\Omega, \mathfrak{F}; P)$ with $\mathrm{E}|X_{i,n}| < \infty$. Let $\{\mathfrak{F}_{i,n}, 1 \leq i \leq k_n, n \geq 1\}$ be an array of sub-sigma fields with $\mathfrak{F}_{i-1,n} \subseteq \mathfrak{F}_{i,n}$ ($\mathfrak{F}_{0,n} = \{\varnothing, \Omega\}$). We then call $\{X_{i,n}, \mathfrak{F}_{i,n}, 1 \leq i \leq k_n, n \geq 1\}$ a martingale difference array if $X_{i,n}$ is $\mathfrak{F}_{i,n}$-measurable and $\mathrm{E}[X_{i,n}|\mathfrak{F}_{i,n}] = 0$.

Theorem 1 (Deng and Wang (2023)). *Define* $F_n = \sum_{i=1}^{n} c_{ni}\varepsilon_{ni}^3 + \varepsilon_n' A_n \varepsilon_n + \mathbf{b}_n' \varepsilon_n$, *where* $\varepsilon_n = [\varepsilon_{n1}, \cdots, \varepsilon_{nn}]^\top$ *is a random vector,* c_{ni}*'s are scalars for* $i = 1, \cdots, n$, A_n *is a* $n \times n$ *matrix, and* $\mathbf{b}_n = [b_{n1}, \cdots, b_{nn}]^\top$ *is a* n*-dimensional column vector. Suppose that*

(i) ε_{ni}*'s are totally i.i.d., and* $\mathrm{E}[\varepsilon_{ni}] = 0$ *for each* i. *Moreover,* $\sup_n \mathrm{E}[|\varepsilon_{ni}|^{6+\eta_1}] < \infty$ *for some* $\eta_1 > 0$.

(ii) $\sup_{j,n} \sum_{i=1}^n |a_{ij,n}| < \infty$.

(iii) $\sup_n \frac{1}{n} \sum_{i=1}^n |b_{ni}|^{2+\eta_2} < \infty$ *for some* $\eta_2 > 0$, *and* $\sup_n \frac{1}{n} \sum_{i=1}^n |c_{ni}|^{2+\eta_3} < \infty$ *for some* $\eta_3 > 0$.

Then, $(F_n - \mu_{F,n})/\sigma_{F,n} \xrightarrow{d} N(0,1)$ *provided that* $\frac{1}{n}\sigma_{F,n}^2 \geq \nu$ *for some* $\nu > 0$, *where* $\mu_{F,n} = \sum_{i=1}^n (c_{ni}\mu_{\varepsilon,3} + a_{ii,n}\sigma_\varepsilon^2)$, *and*

$$\sigma_{F,n}^2 = \sum_{i=1}^{n} \Bigg\{ c_{ni}^2(\mu_{\varepsilon,6} - \mu_{\varepsilon,3}^2) + a_{ii,n}^2(\mu_{\varepsilon,4} - \sigma_\varepsilon^4) + 4\sigma_\varepsilon^4 \sum_{j=1}^{i-1} a_{ij,n}^2 + b_{ni}^2\sigma_\varepsilon^2 + 2b_{ni}c_{ni}\mu_{\varepsilon,4} + 2a_{ii,n}c_{ni}(\mu_{\varepsilon,5} - \mu_{\varepsilon,3}\sigma_\varepsilon^2) + 2a_{ii,n}b_{ni}\mu_{\varepsilon,3} \Bigg\}$$

with $\sigma_\varepsilon^2 = \mathrm{E}[\varepsilon_{ni}^2]$ *and* $\mu_{\varepsilon,s} = \mathrm{E}[\varepsilon_{ni}^s]$ *for* $s = 3, 4, 6$.

Proof. We follow the framework in Kelejian and Prucha (2001). That is, checking the conditions in Lemma 1 (in Appendix). Notice that the condition (0.5) follows from

$$\sum_{i=1}^{k_n} \mathrm{E}\left\{ \mathrm{E}\left[X_{i,n}^2 \mid \mathfrak{F}_{i-1,n} \right] \right\} \xrightarrow{p} 0, \tag{0.1}$$

so it only remains to check (0.1) and (0.6) (see ibid, p. 240 for details).

Without loss of generality, we can reasonably assume that A_n is symmetric since $\varepsilon_n^\top A_n \varepsilon_n = (1/2)\varepsilon_n^\top (A_n + A_n')\varepsilon_n$. Because ε_{ni}'s are independent and $\mathrm{E}[\varepsilon_{ni}] = 0$, $\mu_{F,n} = \sum_{i=1}^n (c_{ni}\mu_{\varepsilon,3} + a_{ii,n}\sigma_\varepsilon^2)$. Let

$$F_n - \mu_{F,n} = \sum_{i=1}^n Y_{ni} = \sum_{i=1}^n \left[c_{ni}(\varepsilon_{ni}^3 - \mu_{\varepsilon,3}) + a_{ii,n}(\varepsilon_{ni}^2 - \sigma_\varepsilon^2) + 2\varepsilon_{ni}\sum_{j=1}^{i-1} a_{ij,n}\varepsilon_{nj} + b_{ni}\varepsilon_{ni} \right].$$

Consider the σ-fields $\mathfrak{F}_{0,n} = \{\varnothing, \Omega\}$, $\mathfrak{F}_{i,n} = \sigma(\varepsilon_{n1}, \cdots, \varepsilon_{ni-1})$, $1 \le i \le n$. Obviously, $\mathfrak{F}_{i-1,n} \subseteq \mathfrak{F}_{i,n}$, Y_{ni} is $\mathfrak{F}_{i,n}$-measurable, and $\mathrm{E}[Y_{ni}|\mathfrak{F}_{i,n}] = 0$. Therefore, $\{Y_{i,n}, \mathfrak{F}_{i,n}, 1 \le i \le n, n \ge 1\}$ forms a martingale difference array. Consequently $\sigma_{F,n}^2 = \sum_{i=1}^n \mathrm{E}[Y_{i,n}^2]$, where

$$\begin{aligned} Y_{i,n}^2 =& c_{ni}^2(\varepsilon_{ni}^3 - \mu_{\varepsilon,3})^2 + a_{ii,n}^2(\varepsilon_{ni}^2 - \sigma_\varepsilon^2)^2 + 4\varepsilon_{ni}^2\Big(\sum_{j=1}^{i-1} a_{ij,n}\varepsilon_{nj}\Big)^2 + b_{ni}^2\varepsilon_{ni}^2 \\ &+ 4c_{ni}\varepsilon_{ni}(\varepsilon_{ni}^3 - \mu_{\varepsilon,3})\sum_{j=1}^{i-1} a_{ij,n}\varepsilon_{nj} + 4b_{ni}\varepsilon_{ni}^2\sum_{j=1}^{i-1} a_{ij,n}\varepsilon_{nj} + 2b_{ni}c_{ni}\varepsilon_{ni}(\varepsilon_{ni}^3 - \mu_{\varepsilon,3}) \\ &+ 2a_{ii,n}c_{ni}(\varepsilon_{ni}^3 - \mu_{\varepsilon,3})(\varepsilon_{ni}^2 - \sigma_\varepsilon^2) + 4a_{ii,n}(\varepsilon_{ni}^2 - \sigma_\varepsilon^2)\varepsilon_{ni}\sum_{j=1}^{i-1} a_{ij,n}\varepsilon_{nj} + 2a_{ii,n}b_{ni}\varepsilon_{ni}(\varepsilon_{ni}^2 - \sigma_\varepsilon^2). \end{aligned} \tag{0.2}$$

It follows that

$$\begin{aligned} \mathrm{E}[Y_{i,n}^2] =& c_{ni}^2(\mu_{\varepsilon,6} - \mu_{\varepsilon,3}^2) + a_{ii,n}^2(\mu_{\varepsilon,4} - \sigma_\varepsilon^4) + 4\sigma_\varepsilon^4\sum_{j=1}^{i-1} a_{ij,n}^2 + b_{ni}^2\sigma_\varepsilon^2 \\ &+ 2b_{ni}c_{ni}\mu_{\varepsilon,4} + 2a_{ii,n}c_{ni}(\mu_{\varepsilon,5} - \mu_{\varepsilon,3}\sigma_\varepsilon^2) + 2a_{ii,n}b_{ni}\mu_{\varepsilon,3}. \end{aligned} \tag{0.3}$$

Let $X_{i,n} = Y_{i,n}/\sigma_{F,n}$, then $\{X_{i,n}, \mathfrak{F}_{i,n}, 1 \le i \le n, n \ge 1\}$ also forms a martingale difference array. We now prove that

$$\frac{F_n - \mu_{F,n}}{\sigma_{F,n}} = \sum_{i=1}^n X_{i,n} \xrightarrow{d} \mathcal{N}(0,1)$$

by checking that $X_{i,n}$'s satisfy conditions (0.1) and (0.6) under our assumptions.

We consider the case where $0 < \eta \leqslant \frac{1}{2}\min(\eta_1, \eta_2, \eta_3)$, which make Conditions (i)–(iii) simultaneously holds. Under Condition (i) on the ε_{ni}, there exists then some finite constant K_e such that $\mathrm{E}\,|\varepsilon_{ni}|^s \leqslant K_e$ for $s = 1, \cdots, 6$, and $\mathrm{E}\,|\varepsilon_{ni}|^t\,\mathrm{E}\,|\varepsilon_{ni}|^t \leqslant K_e$, $\mathrm{E}\left|\varepsilon_{ni}^2 - \sigma_\varepsilon^2\right|^t \leqslant K_e$, and $\mathrm{E}\left|\varepsilon_{ni}^3 - \mu_{\varepsilon,3}\right|^t \leqslant K_e$ for $t \leqslant 2 + \eta$. Similarly, under Conditions (ii)–(iii) on the $(a_{ij,n}, b_{ni}, c_{ni})$, there also exists some finite constant K_p such that $\sum_{j=1}^n |a_{ij,n}| \leqslant K_p$, $\frac{1}{n}\sum_{j=1}^n |b_{ni}|^t \leqslant K_p$ and $\frac{1}{n}\sum_{j=1}^n |c_{ni}|^t \leqslant K_p$ for $t \leqslant 2 + \eta$. Observe that $\sum_{j=1}^n |a_{ij,n}|^r \leqslant K_p^r$ for $r \geqslant 1$ and therefore, $\sum_{k=1}^n |a_{ik,n}| \cdot |a_{jk,n}| \leqslant [\sum_{k=1}^n a_{ik,n}^2]^{1/2}[\sum_{k=1}^n a_{jk,n}^2]^{1/2} \leqslant K_p^2$ by the Cauchy-Schwarz inequality.

In what follows, let $q = 2 + \eta$ and there must exist $p > 1$ such that $1/q + 1/p = 1$. By applying the inequalities $|a+b+c|^q \le 3^q(|a|^q + |b|^q + |c|^q)$, $|a+b|^q \le 2^q(|a|^q + |b|^q)$, and Hölder's inequality, we have

$$
\begin{aligned}
&|Y_{i,n}|^q = |c_{ni}(\varepsilon_{ni}^3 - \mu_{\varepsilon,3}) + a_{ii,n}(\varepsilon_{ni}^2 - \sigma_\varepsilon^2) + 2\varepsilon_{ni}\sum_{j=1}^{i-1} a_{ij,n}\varepsilon_{nj} + b_{ni}\varepsilon_{ni}|^q \\
\le& 3^q \Bigg[|c_{ni}|^q \cdot |\varepsilon_{ni}^3 - \mu_{\varepsilon,3}|^q + |b_{ni}|^q \cdot |\varepsilon_{ni}|^q \\
&+ \Big(|a_{ii,n}|^{1/p} \cdot |a_{ii,n}|^{1/q} \cdot |\varepsilon_{ni}^2 - \sigma_\varepsilon^2| + 2\sum_{j=1}^{i-1} |a_{ij,n}|^{1/p} \cdot |a_{ij,n}|^{1/q} \cdot |\varepsilon_{ni}| \cdot |\varepsilon_{nj}| \Big)^q \Bigg] \\
\le& 3^q \left[|c_{ni}|^q \cdot |\varepsilon_{ni}^3 - \mu_{\varepsilon,3}|^q + |b_{ni}|^q \cdot |\varepsilon_{ni}|^q \right] \\
&+ 3^q 2^q \Big(\sum_{j=1}^{i-1} |a_{ij,n}|^{p/p} \Big)^{q/p} \Big[|a_{ii,n}|^{q/q} \cdot |\varepsilon_{ni}^2 - \sigma_\varepsilon^2|^q + 2^q \sum_{j=1}^{i-1} |a_{ij,n}|^{q/q} \cdot |\varepsilon_{ni}|^q \cdot |\varepsilon_{nj}|^q \Big] \\
\le& 3^q \left[|c_{ni}|^q \cdot |\varepsilon_{ni}^3 - \mu_{\varepsilon,3}|^q + |b_{ni}|^q \cdot |\varepsilon_{ni}|^q \right] + 12^q \left(K_p\right)^{q/p} \Big[|a_{ii,n}| \cdot |\varepsilon_{ni}^2 - \sigma_\varepsilon^2|^q + \sum_{j=1}^{i-1} |a_{ij,n}| \cdot |\varepsilon_{ni}|^q \cdot |\varepsilon_{nj}|^q \Big].
\end{aligned}
$$

It yields that

$$
\begin{aligned}
&\sum_{i=1}^{n} \mathrm{E}\left\{ \mathrm{E}\left[|Y_{i,n}|^q \mid \mathfrak{F}_{i,n} \right] \right\} \\
\le& 3^q \sum_{i=1}^{n} \left[|c_{ni}|^q \cdot \mathrm{E}|\varepsilon_{ni}^3 - \mu_{\varepsilon,3}|^q + |b_{ni}|^q \cdot \mathrm{E}|\varepsilon_{ni}|^q \right] \\
&+ 12^q \left(K_p\right)^{q/p} \sum_{i=1}^{n} \left[|a_{ii,n}| \cdot \mathrm{E}|\varepsilon_{ni}^2 - \sigma_\varepsilon^2|^q + \sum_{j=1}^{i-1} |a_{ij,n}| \cdot \mathrm{E}|\varepsilon_{ni}|^q \cdot \mathrm{E}|\varepsilon_{nj}|^q \right] \\
\le& n \cdot 2 \cdot 3^q K_e K_p + 12^q K_e K_p^{q/p} \sum_{i=1}^{n} \sum_{j=1}^{i} |a_{ij,n}| \le n \cdot \left(2 \cdot 3^q K_e K_p + 12^q K_e K_p^{1+q/p} \right).
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
0 \le \sum_{i=1}^{n} \mathrm{E}\left\{ \mathrm{E}[|X_{i,n}|^{2+\eta} \mid \mathfrak{F}_{i,n}] \right\} &= \frac{1}{(\sigma_{F,n}^2)^{1+\eta/2}} \sum_{i=1}^{n} \mathrm{E}\left\{ \mathrm{E}[|Y_{i,n}|^q \mid \mathfrak{F}_{i,n}] \right\} \\
&\le \frac{1}{n^{\eta/2}} \cdot \frac{1}{(\frac{1}{n}\sigma_{F,n}^2)^{1+\eta/2}} \left(2 \cdot 3^q K_e K_p + 12^q K_e K_p^{1+q/p} \right) \\
&\le \frac{1}{n^{\eta/2}} \cdot O(1) \to 0 \quad (\text{as } n \to \infty),
\end{aligned}
$$

where the last inequality follows from $0 < \nu \le \frac{1}{n}\sigma_{F,n}^2$. This proves that condition (0.1) holds.

We now check (0.6). Noticing that the ε_{ni}'s are independent with zero mean, it follows from (0.2) that

$$
\begin{aligned}
&\mathrm{E}\left[|Y_{i,n}|^2 \mid \mathfrak{F}_{i-1,n} \right] = \mathrm{E}\left[|Y_{i,n}|^2 \mid \varepsilon_{n1}, \cdots, \varepsilon_{ni-1} \right] \\
=& c_{ni}^2(\mu_{\varepsilon,6} - \mu_{\varepsilon,3}^2) + a_{ii,n}^2(\mu_{\varepsilon,4} - \sigma_\varepsilon^4) + 4\sigma_\varepsilon^2 \sum_{j=1}^{i-1}\sum_{k=1}^{i-1} a_{ij,n}a_{ik,n}\varepsilon_{nj}\varepsilon_{nk} + 4(c_{ni}\mu_{\varepsilon,4} + b_{ni}\sigma_\varepsilon^2)\sum_{j=1}^{i-1} a_{ij,n}\varepsilon_{nj} + b_{ni}\sigma_\varepsilon^2 \\
&+ 2b_{ni}c_{ni}\mu_{\varepsilon,4} + 2a_{ii,n}c_{ni}(\mu_{\varepsilon,5} - \mu_{\varepsilon,3}\sigma_\varepsilon^2) + 4a_{ii,n}\mu_{\varepsilon,3}\sum_{j=1}^{i-1} a_{ij,n}\varepsilon_{nj} + 2a_{ii,n}b_{ni}\mu_{\varepsilon,3}.
\end{aligned}
$$

Recalling that $\sigma_{F,n}^2 = \sum_{i=1}^n \mathrm{E}[Y_{i,n}^2]$ and working with (0.2), it yields that

$$
\begin{aligned}
&\sum_{i=1}^{n} \mathrm{E}\left[X_{i,n}^2 \mid \mathfrak{F}_{i-1,n}\right] - 1 = \frac{1}{\sigma_{F,n}^2} \sum_{i=1}^{n} \left\{\mathrm{E}\left[Y_{i,n}^2 \mid \mathfrak{F}_{i-1,n}\right] - \mathrm{E}[Y_{ni}^2]\right\} \\
=&\frac{1}{\sigma_{F,n}^2} \sum_{i=1}^{n} \left\{ 4\sigma_\varepsilon^2 \sum_{j=1}^{i-1} \sum_{k=1,k\neq j}^{i-1} a_{ij,n} a_{ik,n} \varepsilon_{nj}\varepsilon_{nk} + 4\sigma_\varepsilon^2 \sum_{j=1}^{i-1} a_{ij,n}^2 (\varepsilon_{nj}^2 - \sigma_\varepsilon^2) \right. \\
&\left. +4(a_{ii,n}\mu_{\varepsilon,3} + b_{ni}\sigma_\varepsilon^2) \sum_{j=1}^{i-1} a_{ij,n}\varepsilon_{nj} + 4c_{ni}\mu_{\varepsilon,4} \sum_{j=1}^{i-1} a_{ij,n}\varepsilon_{nj} \right\} \\
\triangleq & \frac{1}{\frac{1}{n}\sigma_{F,n}^2} \left[4H_{1,n} + 4H_{2,n} + 4H_{3,n} + 4H_{4,n}\right].
\end{aligned}
$$

Since $0 < \nu \leq \frac{1}{n}\sigma_{F,n}^2$ condition (0.6) holds if $H_{i,n} = o_p(1)$ for $i = 1, 2, 3, 4$. Here $4H_{1,n}$, $4H_{2,n}$, and $4H_{3,n}$ are fully identical to those in Kelejian and Prucha (2001, p. 245-246). Hence, we omit analysis for triple, and the proof will be accomplished only by proving the last $H_{4,n} = o_p(1)$. Observe that

$$
H_{4,n} = \sum_{i=1}^{n} \varphi_{i,n}\varepsilon_{ni}, \quad \varphi_{i,n} = \frac{1}{n} \sum_{j=i+1}^{n} a_{ji,n} c_{nj} \mu_{\varepsilon,4},
$$

where $\varphi_{i,n}\varepsilon_{ni}/|\varphi_{i,n}|$'s are then independent with zero-mean and uniformly integrable under our moment assumptions. Furthermore, we have

$$
\begin{aligned}
\limsup_{n\to\infty} \sum_{i=1}^{n-1} |\varphi_{i,n}| &\leq \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} |c_{nj}| \mu_{\varepsilon,4} \cdot |a_{ji,n}| \\
&\leq K_e \limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} |c_{nj}| \sum_{i=1}^{n} |a_{ji,n}| \leq K_e K_p^2 < \infty
\end{aligned}
$$

and

$$
\lim_{n\to\infty} \sum_{i=1}^{n-1} \varphi_{i,n}^2 \leq \lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n-1} \left[\sum_{j=i+1}^{n} \mu_{\varepsilon,4} |c_{nj}| \cdot |a_{ji,n}| \right]^2 \leq K_e^2 \lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n-1} \left[\sum_{j=i+1}^{n} |c_{nj}| \cdot |a_{ji,n}| \right]^2. \tag{0.4}
$$

Observing that

$$
\sum_{j=1}^{n} |c_{nj}| \cdot |a_{ji,n}| \leq n^{1/q} \left(\frac{1}{n} \sum_{j=1}^{n} |c_{nj}|^q \right)^{1/q} \cdot \left(\sum_{j=1}^{n} |a_{ji,n}|^p \right)^{1/p} \leq n^{1/q} K_p^{1+1/q}
$$

by the Hölder's inequality, (0.4) simplifies to

$$
\begin{aligned}
\lim_{n\to\infty} \sum_{i=1}^{n-1} \varphi_{i,n}^2 &\leq K_e^2 \lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} \left(n^{1/q} K_p^{1+1/q} \right)^2 \leq K_e^2 K_p^{2(1+1/q)} \lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} n^{2/q} \\
&= K_e^2 K_p^{2(1+1/q)} \lim_{n\to\infty} n^{2/q-1} = 0,
\end{aligned}
$$

where the final "=" follows from $q > 2$. By the weak law of large numbers (for martingale difference arrays) in Davidson (1994, p. 299), $H_{4,n} = o_p(1)$ holds. This demonstrates that also condition

(0.6) holds. The proof is accomplished. $\square$

Appendix useful lemma

Lemma 1. *(Kelejian and Prucha, 2001, Lemma A.1) Let* $\{X_{i,n}, \mathfrak{F}_{i,n}, 1 \leq i \leq k_n, n \geq 1\}$ *be a square integrable martingale difference array. Suppose that for all* $\delta > 0$*,*

$$\sum_{i=1}^{k_n} \mathrm{E}\left[X_{i,n}^2 I_{\{|X_{i,n}|>\delta\}} \mid \mathfrak{F}_{i-1,n}\right] \xrightarrow{p} 0 \tag{0.5}$$

and

$$\sum_{i=1}^{k_n} \mathrm{E}\left[X_{i,n}^2 \mid \mathfrak{F}_{i-1,n}\right] \xrightarrow{p} 1. \tag{0.6}$$

Then, $\sum_{i=1}^{k_n} X_{i,n} \xrightarrow{d} \mathcal{N}(0,1)$.

References

Davidson, J., 1994. Stochastic Limit Theory. Oxford University Press, Oxford.

Deng, M., Wang, M., 2023. GMM inference in spatial autoregressive stochastic frontier analysis. Manuscript .

Kelejian, H., Prucha, I.R., 2001. On the asymptotic distribution of the Moran I test statistic with applications. Journal of Econometrics 104, 219–257.