Limit theory for quadratic forms of linear processes

L. Giraitis, Queen Mary, University of London
joint work with
R. J. Bhansali, University of Liverpool
P. S. Kokoszka, Utah State University

January 15, 2008

Problem: what simple verifiable condition for the CLT for general quadratic form of linear variables

<u>Sums</u> Assume that (X_t) is a stationary linear process

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j},$$

where (ε_t) is a sequence of i.i.d. (0,1) variables

$$E\varepsilon_t = 0, \quad E\varepsilon_t^2 = 1$$

$$\psi_j \text{ real} \qquad \sum_{j=0}^{\infty} \psi_j^2 < \infty.$$

Dependence: Variables X_t can be weakly and strongly dependent

Asymptotic theory for the sums: If

$$S_n = \sum_{j=1}^{n} X_j, \qquad and \qquad Var(S_n) \to \infty$$

then (Well known (Ibragimov, Linnik (1964)):

$$\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \to N(0, 1).$$

Note: no additional conditions needed,

Quadratic forms in i.i.d. variables

Question: can we have something like that for quadratic forms?

Answer: Yes, for i.i.d variables $X_t = \varepsilon_t$. Set

$$T_n = \sum_{t,s=1}^{n} a_n(t,s)\varepsilon_t\varepsilon_s$$

$A_n = (a_n(t,s))$ real symmetric matrix

Question: Does

$$Var(T_n) \to \infty$$

implies CLT for T_n?

Answer: Almost: we need slightly stronger condition:

Denote

$||A_n|| = (\sum_{t,k=1}^{n} a_n(t,s)^2)^{1/2}$ Euclidean norm

$||A_n||_{sp} = \max_{||x||=1} ||A_n x||$ spectral norm

CLT in zero diagonal case: $a_n(t,t) = 0$

Sufficient condition:

$$\frac{||A_n||_{sp}}{||A_n||} \to 0, \quad n \to \infty.$$

Then

$$\frac{T_n - E[T_n]}{\sqrt{Var(T_n)}} \to N(0,1).$$

Discussed by: Rotar (1973), Jong (1987), Guttorp and Lockhart (1988), Mikosch (1991)

Note: 1. Condition implies CLT if $||T_n|| \le C$,

$$Var(T_n) = 2||T_n||^2 \to \infty$$

2. Only second finite moment is needed: $E\varepsilon_t^2 < \infty$.

CLT in non-zero diagonal case: $a_n(t,t) \neq 0$. More subtle, we discuss it later

Asymptotic theory for quadratic forms

Assume now that X_t are dependent linear variables

Objectives: 1. asymptotic normality theory for quadratic form $Q_{n,X}$ in dependent linear random variables X_t

$$Q_{n,X} = \sum_{k,t=1}^{n} d_n(k-t) X_k X_t$$

2. for the use in kernel, and other estimators converging at a rate not necessarily

$n^{1/2}$.

3. suitable for all types of dependence (long, short and negative memory) of X_t

4. conditions should be easy to verify

Known results for dependent X_k

1. The case $d_n(t) \equiv d(t)$. Conditions for CLT with normalization $\sqrt{n}$ were derived in

Fox and Taqqu (1987), Avram (1988), Giraitis and Surgailis (1990) and others

4 finite moments needed, $EX_t^4 < \infty$, $E\varepsilon_t^4 < \infty$

Note Direct verification of CLT when $d_n(t)$ depends on n is difficult.

We wish to allow slow growth of $Var(Q_{n,X}) = o(n)$, to cover kernel estimation.

Method: we approximate $Q_{n,X}$ by a quadratic form

$$Q_{n,\varepsilon} = \sum_{k,t=1}^{n} e_n(k-t)\varepsilon_k \varepsilon_t,$$

in i.i.d. variables ε_t (innovations of $\{X_t\}$

Note: The existing research, based on this method,

Phillips and Solo (1992), Mikosch (1995), Kokoszka and Taqqu (1996) deals with the case

$$d_n(t) \equiv d(t), \text{ four moments}$$

It provides only the bound

$$Var(Q_{n,X} - Q_{n,\varepsilon}) = o(n)$$

and CLT with normalization $\sqrt{n}$. Not good enough, to cover the case $Var(Q_{n,X}) = o(n)$.

Our approach:

1. Write

$$Q_{n,X} = Q_{n,\varepsilon} + [Q_{n,X} - Q_{n,\varepsilon}]$$

2. Use CLT for $Q_{n,\varepsilon}$ with non-vanishing diagonal

3. Main technical problem: show that

$$Q_{n,\varepsilon} \text{ dominates } [Q_{n,X} - Q_{n,\varepsilon}]$$

we need very sharp upper bound for

$$[Q_{n,X} - Q_{n,\varepsilon}]$$

Assumptions on X_t (linear process)

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}.$$

Property: There exists $d \in (-1/2, 1/2)$ such that ψ_j satisfy

$$\psi_j = O(j^{-1+d}), \quad |\psi_j - \psi_{j+1}| = O(j^{-2+d}), \qquad \text{if } d \neq 0$$

and, $\sum_{j=0}^{\infty} \psi_j = 0$, if $d < 0$.

If $d = 0$, then there exists $\alpha > 1$ such that

$$\sum_{j=n}^{\infty} |\psi_j| = O(n^{-\alpha}).$$

Example: the above property holds if $\{X_t\}$ is a defined by

$$(1 - \mathsf{L})^d A(L) X_t = \varepsilon_t, \qquad A(L) = \sum_{j=0}^{\infty} a_j L^j$$

where $-1/2 < d < 1/2$ is memory parameter, L is the lag operator, and AR coefficients decay fast:

$$a_j = O(r^j), \quad \text{for some } 0 < r < 1.$$

For example, ARFIMA(p,d,q) models. Then $\{X_t\}$ has spectral density

$$f(\lambda) = (2\pi)^{-1} |\sum_{j=0}^{\infty} \psi_j e^{-i\lambda j}|^2$$

with property

$$f(\lambda) = |\lambda|^{-2d}(b_0 + O(|\lambda|^2)), \quad \lambda \to 0.$$

Question: how to construct the approximating quadratic form $Q_{n,Z\varepsilon}$?

Denote by

$$I_n(\lambda) = \frac{1}{2\pi n}\left|\sum_{j=1}^{n} X_j e^{i\lambda j}\right|^2, \quad I_{n,\varepsilon}(\lambda) = \frac{1}{2\pi n}\left|\sum_{j=1}^{n} \varepsilon_j e^{i\lambda j}\right|^2$$

periodograms of $\{X_t\}$ and $\{\varepsilon_t\}$.

Under

Assumption on $d_n(t)$: there exist a real even function $\eta_n(\lambda)$ such that

$$d_n(t) = \int_{-\pi}^{\pi} \eta_n(\lambda) e^{i\lambda t} d\lambda$$

we can write

$$Q_{n,X} = \sum_{k,t=1}^{n} d_n(k-t)X_kX_t \equiv (2\pi n)\int_{-\pi}^{\pi} \eta_n(\lambda)I_n(\lambda)d\lambda$$

<u>Bartlett decomposition</u>

$$I_n(\lambda) = 2\pi f(\lambda)I_{n,\varepsilon}(\lambda) + \text{"small term"}$$

suggests that we can write

$$Q_{n,X} = Q_{n,\varepsilon} + \text{"small term"}$$

where

$$Q_{n,\varepsilon} = \sum_{k,t=1}^{n} e_n(k-t)\varepsilon_k\varepsilon_t \equiv (2\pi n)\int_{-\pi}^{\pi} \eta_n(\lambda)2\pi f(\lambda)I_{n,\varepsilon}(\lambda)d\lambda$$

here

$$e_n(t) = \int_{-\pi}^{\pi} (2\pi\eta_n(\lambda)f(\lambda))e^{i\lambda t}d\lambda.$$

Approximation of $Q_{n,X}$:

Objective: to derive sharp upper bounds of

$$\mathrm{Var}(Q_{n,X} - Q_{n,\varepsilon}) \text{ and } E|Q_{n,X} - Q_{n,\varepsilon}|$$

Assumption on $\eta_n(\lambda)$: There exists $-1 < \beta < 1$ and $k_n \geq 0$,

$$|\eta_n(\lambda)| \leq k_n |\lambda|^{-\beta}, \quad \lambda \in [-\pi, \pi], \quad n \geq 1.$$

Note that the weight function in approximating form $Q_{n,\varepsilon}$, has the bound

$$|\eta_n(\lambda)| f(\lambda) \leq C k_n |\lambda|^{-(2d+\beta)}, \quad \lambda \to 0$$

Main approximation result

Theorem 2.1 Assume that

$$\delta := 2d + \beta < 1/2.$$

(i) If $E\varepsilon_t^4 < \infty$, then

$$[\mathrm{Var}(Q_{n,X} - Q_{n,\varepsilon})]^{1/2} = O(r_n)$$

where

$$r_n = \begin{cases} k_n, & \text{if } d = 0, \\ k_n n^{\max(\delta,0)} \log n, & \text{if } d \neq 0. \end{cases}$$

(ii) If $E\varepsilon_t^2 < \infty$, then

$$E|Q_{n,X} - Q_{n,\varepsilon}| = O(\bar{r}_n)$$

where

$$\bar{r}_n = \begin{cases} k_n, & \text{if } d = 0, \\ k_n n^{\max(\delta,d,0)} \log n, & \text{if } d \neq 0 \end{cases}$$

(Note that $r_n \leq \bar{r}_n$.)

Comment: 1. Approximation rate depends on $\delta = 2d + \beta$. It allows memory compensation (d positive, β negative)

2. k_n plays a secondary role

3. Approximation precision is very high when δ is small or negative

4. In some case existing results are improved by $n^{-1/2}$. For example, if $d = 0$, $\eta_n(\lambda) \equiv \eta(\lambda)$ and $\beta = 0$, then Brockwell and Davis (1991) approximation is

$$[\mathrm{Var}(Q_{n,X} - Q_{n,\varepsilon})]^{1/2} = o(n^{1/2})$$

Derived approximation is

$$[\mathrm{Var}(Q_{n,X} - Q_{n,\varepsilon})]^{1/2} = O(1)$$

Central Limit theorem for Quadratic forms

To derive the CLT for $Q_{n,X}$ we have to

1. assume that the main term $Q_{n,\varepsilon}$ dominates the approximation error.

2, Use CLT for $Q_{n,\varepsilon}$

<u>Notations</u>: denote E_n the matrix $E_n = (e_n(t-k))_{t,k=1,\ldots,n}$

$||E_n|| = (\sum_{t,k=1}^n e_n^2(t-k))^{1/2}$ Euclidean norm.

Recall

$$|\eta_n(\lambda)| \leq k_n |\lambda|^{-\beta}$$

Theorem 2.2 Assume that $\beta + 2d < 1/2$ and $E\varepsilon_t^4 < \infty$. Suppose that

$$\frac{k_n}{||E_n||} \to 0, \quad when \quad d = 0$$

$$\frac{k_n n^{\max(\beta+2d,0)} \log n}{||E_n||} \to 0, \quad when \quad d \neq 0$$

then,

$$\text{Var}(Q_{n,X})/\text{Var}(Q_{n,\varepsilon}) \to 1, \quad \text{Var}(Q_{n,X}) \asymp ||E_n||^2$$

and

$$(\text{Var}(Q_{n,X}))^{-1/2}(Q_{n,X} - EQ_{n,X}) \xrightarrow{d} N(0,1).$$

Note: conditions of the CLT are comparable to those of the classical CLT in case of i.i.d. variables

Statistical applications involve the integrated periodograms

$$T_{n,X} = \int_{-\pi}^{\pi} \eta_n(\lambda) I_n(\lambda) d\lambda$$

Theorem 2.3 gives conditions when centering $ET_{n,X}$ can be replaced by explicit constant

$$ET_{n,\varepsilon} = \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda.$$

Recall that $\delta = \beta + 2d$,

$$\bar{r}_n = \begin{cases} k_n, & \text{if } d = 0, \\ k_n n^{\max(\delta, d, 0)} \log n, & \text{if } d \neq 0 \end{cases}$$

Theorem 2.3. Assume that $\delta < 1/2$ and

$$\frac{\bar{r}_n}{||E_n||} \to 0$$

(i) If $E\varepsilon_t^4 < \infty$, then

$$[\mathrm{Var}(T_{n,X})]^{-1/2}(T_{n,X} - \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda) \xrightarrow{d} N(0,1).$$

(ii) If

$$E\varepsilon_t^{2+\delta} < \infty \text{ and } \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda = o(n^{-1/2}||E_n||)$$

then

$$\mathrm{Var}(T_{n,X}) = \frac{||E_n||^2}{2(\pi n)^2}(1 + o(1))$$

and

$$\frac{\sqrt{2}\pi n}{||E_n||}(T_{n,X} - \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda) \xrightarrow{d} N(0,1).$$

(iii) If

$$E\varepsilon_t^2 < \infty \text{ and } \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda \equiv 0$$

then

$$\frac{\sqrt{2}\pi n}{||E_n||}(T_{n,X} - \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda) \xrightarrow{d} N(0,1).$$

Discussion of the results

Assumption on weights ψ_j can be replaced by a stronger condition

Assumption. There exists $d \in (-1/2, 1/2)$ and a constant $c \neq 0$ such that

$$\psi_j = \begin{cases} cj^{-1+d}(1+O(j^{-1})), & \text{if } d \in (0, 1/2), \\ cj^{-1+d}(1+O(j^{-1})) \text{ and } \sum_{j=0}^{\infty} \psi_j = 0, & \text{if } d \in (-1/2, 0). \end{cases}$$

Comment: 1. Assumption motivated by common time series models

2. It implies that the spectral density

$$f(\lambda) = c|\lambda|^{-2d}(1 + o(1)), \quad \text{as} \quad \lambda \to 0,$$

3. Assumption $|\eta_n(\lambda)| \leq k_n|\lambda|^{-\beta}$ is weak, easy to check

4. In case $k_n \equiv K$, $2d + \beta \leq 0$ (for example, $d = 0$, $\beta = 0$), the bound

$$E|Q_{n,X} - Q_{n,\varepsilon}| = O(\log n),$$

is a much sharper bound than

$$E|Q_{n,X} - Q_{n,\varepsilon}| = O(n^{1/2}),$$

in Brockwell and Davis (1991), Kokoszka and Taqqu (1996,97)

5. In case $\eta_n(\lambda) \equiv \eta(\lambda)$, Taqqu, Fox (1986), Giraitis and Surgailis (1991) used the same condition

$$2d + \beta < 1/2$$

If $2d + \beta > 1/2$ CLT might not hold.

6. Applications show that CLT in kernel estimation, instead of 4, requires only $2{+}\delta$ moments

$$E\varepsilon_t^{2+\delta} < \infty$$

7. In applications, the main term $Q_{n,\varepsilon}$ dominates the reminder $Q_{n,X} - Q_{n,\varepsilon}$, and allows

$$Var(Q_{n,X}) = o(n).$$

CLT for quadratic forms of i.i.d. random variables

Problem: asymptotic normality result is based on approximation result, and normality of $Q_{n,\varepsilon}$.

Consider the general quadratic form

$$T_n = \sum_{t,k=1}^{n} a_n(t,k)\varepsilon_t\varepsilon_k$$

where $A_n = (a_n(t,k))_{t,k=1,\dots,n}$ a real symmetric matrix.

Comment: The case of zero diagonal $a_n(t,t) = 0$ is well investigated. Next theorem allows non-zero diagonal

Theorem 4.1 Assume that

$$\frac{||A_n||_{sp}}{||A_n||} \to 0.$$

(i)[**Non-zero diagonal**]. If $E\varepsilon_t^4 < \infty$, then

$$(Var(T_n))^{-1/2}(T_n - ET_n) \xrightarrow{d} N(0,1).$$

(ii) [**Vanishing Diagonal**] If

$$E\varepsilon_t^{2+\delta} < \infty \quad \text{(for some } \delta > 0\text{) and} \quad \sum_{t=1}^{n} a_n^2(t,t) = o(||A_n||^2),$$

then

$$\frac{1}{\sqrt{2}||A_n||}(T_n - ET_n) \xrightarrow{d} N(0,1).$$

(iii) [**Zero diagonal**] If

$$E\varepsilon_t^2 < \infty \quad \text{and} \quad a_n(t,t) = 0, \quad t = 1, ..., n,$$

then CLT(ii) is valid

Special case: A_n is a Toeplitz matrix with entries

$$a_n(t,k) = \int_{-\pi}^{\pi} e^{i(t-k)x} g_n(x)dx, \quad t,k = 1,...,n,$$

where $g_n(x)$, $|x| \leq \pi$ is an even real function.

Theorem 4.2 Let A_n be a Toeplitz matrix and for some $0 \leq \alpha < 1$

$$|g_n(\lambda)| \leq k_n |\lambda|^{-\alpha}, \quad n \geq 1.$$

(i) Then

$$||A_n||_{sp} \leq C k_n n^{\alpha} \quad n \geq 1.$$

(ii) If

$$\frac{k_n n^{\alpha}}{||A_n||} \to 0$$

then

$$\frac{||A_n||_{sp}}{||A_n||} \to 0.$$

Comment: 1. If $|g_n(\lambda)| \leq C$, and $Var(T_n) \to \infty$, then T_n satisfies CLT.

2. Condition on g is precise. For example, if

$$g_n(x) = |x|^{-\alpha}, \quad 0 \leq \alpha < 1$$

then

$$||A_n|| \sim n^{\max(1/2,\alpha)}.$$

Hence, a) for $0 \leq \alpha < 1/2$,

$$\frac{||A_n||_{sp}}{||A_n||} \leq Cn^{\alpha-1/2} \to 0$$

and CLT holds.

b) if $1/2 < \alpha < 1$, then non-CLT holds (Giraitis, Taqqu, Terrin (1998))

Applications

CLT for quadratic forms is one of the main tools in inference of time series

A number of estimators/tests can be written as a quadratic form $Q_{n,X}$ (integrated periodogram with kernel $\eta_n(\lambda)$)

Important applications:

1. spectral estimation
2. kernel estimation
3. Whittle estimation
4. goodness-of-fit test

Example of application in kernel estimation

Illustration. Assume that $\{X_t\}$ is a linear short memory sequence with $d = 0$.

We estimate $f(0)$ using kernel estimator

$$\widehat{f}(0) = \int_{-\pi}^{\pi} \eta_n(\lambda) I_n(\lambda) d\lambda$$

where

$$\eta_n(\lambda) = (2\pi q)^{-1} |\sum_{j=1}^{q} e^{ij\lambda}|^2$$

is the Fejér kernel.

The estimator $\widehat{f}(0)$ uses the Bartlett window. The bandwidth

$$q \to \infty, \quad q = o(n), \quad \text{as} \quad n \to \infty.$$

Existing results: Anderson (1994), Theorem 9.4.1, shows that

$$(n/q)^{1/2}(\widehat{f}(0) - E\widehat{f}(0)) \to N(0, V^2)$$

It assumes finite fourth moment

$$E\varepsilon_t^4 < \infty.$$

Note: 1. Centering by $E[\widehat{f}(0)]$ is not convenient,

2. analysis of the bias $f(0) - E[\widehat{f}(0)]$ difficult.

Our results: 1. imply asymptotic normality of $\widehat{f}(0)$:

a) under $2 + \delta$ moments, $E\varepsilon_t^{2+\delta} < \infty$.

b) allows simple deterministic centering

Assumptions: Assume that f is continuous and

$$f(\lambda) = f(0) + O(\lambda^2), \text{ as } \lambda \to 0.$$

Since

$$|\eta_n(\lambda)| \le Cq$$

then

$$|\eta_n(\lambda)| \le k_n |\lambda|^{-\beta}$$

So,

$$k_n = Cq, \quad \beta = 0, \quad and \quad \delta = \beta + 2d = 0.$$

It is straightforward to check that

$$||E_n||^2 \;=\; \sum_{t,k=1}^{n} e_n^2(t-k) = \int_{-\pi}^{\pi} |\sum_{j=1}^{n} e^{it(x+y)}|^2 \eta_n(x) f(x) \eta_n(y) f(y) dx dy$$

$$\sim \quad qn\,(8/3)\pi^2 f(0)^2,$$

Then

$$\frac{\bar{r}_n}{||E_n||} = \frac{k_n}{||E_n||} \sim C\frac{q}{\sqrt{qn}} = C\frac{\sqrt{q}}{\sqrt{n}} \to 0.$$

Since f is continuous, then

$$\int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda = (2\pi q)^{-1} \int_{-\pi}^{\pi} |\sum_{j=1}^{q} e^{ij\lambda}|^2 f(\lambda) d\lambda \to f(0) = o(n^{-1/2}||E_n||)$$

since $n^{-1/2}||E_n|| \sim cq^{1/2} \to \infty$. Hence

$$(n/q)^{1/2}(\widehat{f}(0) - \int_{-\pi}^{\pi} \eta_n(x) f(x) dx) \xrightarrow{d} N(0, \frac{4}{3} f^2(0)).$$

Note: this convergence does not follow from any existing CLT's for quadratic forms of linear processes because

a) it involves rate of convergence different than $\sqrt{n}$

b) function η_n depends on n.

c) condition $f(\lambda) = f(0) + O(\lambda^2)$ allows to obtain the upper bound of the bias:

$$\int_{-\pi}^{\pi} \eta_n(x) f(x) dx - f(0) = O(q^{-1}).$$