Weak convergence of linear and quadratic forms and related statements on Lp-approximability

Mynbayev, Kairat and Darkenbayeva, Gulsim

Kazakh-British Technical University, Institute of Mathematics and Mathematical Modeling, al-Farabi Kazakh National University

May 2017

Weak convergence of linear and quadratic forms and related statements on L_p-approximability ☆

K.T. Mynbaev[a], G.S. Darkenbayeva[b]

[a] *Kazakh-British Technical University, Institute of Mathematics and Mathematical Modeling, Almaty, Kazakhstan*
[b] *Institute of Mathematics and Mathematical Modeling, al-Farabi Kazakh National University, IITU, Almaty, Kazakhstan*

Abstract

In this paper we obtain central limit theorems for quadratic forms of non-causal short memory linear processes with independent identically distributed innovations. Nabeya and Tanaka (1988) suggested the format, which links the asymptotic distribution to integral operators. In their approach, integral operators had to have continuous symmetric kernels. Mynbaev (2001) employed the theory of approximations to get rid of the continuity requirement. Here we go one step further by lifting the kernel symmetry condition. Also, we establish L_p-approximability of the special sequences which arise in the theory of regressions with slowly varying regressors.

Keywords:
central limit theorem, L_p-approximability, quadratic forms
2008 MSC: 47G10, 60F05, 62E20

1. Introduction

Convergence in distribution of sequences of random variables plays a central role in the theory of probabilities and statistics. Sequences of linear and quadratic forms are among the most important. Existence of a large number of different asymptotic statements is explained by the fact that different applications require different formats and conditions. We concentrate on weak convergence of linear and quadratic forms arising in regression analysis. The book by Tanaka [13] can serve as a comprehensive introduction to this area.

Central limit theorems (CLT's) deal with convergence in distribution of linear forms of type

$$\sum_{t=1}^{n} w_{nt} v_t \text{ as } n \to \infty, \tag{1.1}$$

☆This research is financially supported by a grant from the Ministry of Science and Education of the Republic of Kazakhstan (Grant No. 4084/GF4).

Email addresses: `kairat_mynbayev@yahoo.com, corresponding author` (K.T. Mynbaev), `g.spandiyar@gmail.com` (G.S. Darkenbayeva)

where $v = (v_1, ..., v_n)' \in \mathbf{R}^n$ is a random vector and $w_n = (w_{n1}, ..., w_{nn}) \in \mathbf{R}^n$ is a deterministic vector. One popular approach to modeling dependence of $\{v_t\}$ over time is to specify it as a linear process

$$v_t = \sum_{i \in \mathbf{Z}} c_i e_{t-i}, \tag{1.2}$$

where $\{e_i\}_{i \in \mathbf{Z}}$ are random variables and $\{c_i\}_{i \in \mathbf{Z}}$ are constants. If $c_i = 0$ for $i < 0$, the process (1.2) is called causal; otherwise it is called non-causal. The theory critically depends on whether $\alpha_c \equiv \sum |c_i| < \infty$ (short-memory process) or $\alpha_c = \infty$ but $\sum c_i^2 < \infty$ (long-memory process).

Our results hold for $\{e_i\}_{i \in \mathbf{Z}}$ martingale differences but, for simplicity, we assume that

Assumption 1. The innovations $e_t,\ t \in \mathbf{Z}$, are independent identically distributed (i.i.d.), satisfy $Ee_t = 0$, $\sigma_e^2 \equiv Ee_i^2 < \infty$, $Ee_i^4 < \infty$ for any t and the constants c_i satisfy $\alpha_c := \sum_{i \in Z} |c_i| < \infty$ (short-memory).

It follows that $v_t,\ t \in \mathbf{Z}$, are identically distributed. Our method is flexible in modeling $\{w_n\}$ but is limited to short-memory processes. Quadratic forms are of type

$$Q_n(k_n) = v' k_n v, \tag{1.3}$$

where k_n is a deterministic $n \times n$ matrix and the random vector v is the same as above.

Quadratic forms involving linear processes were considered by many authors. For example, Horvath and Shao [5] established approximations for quadratic forms of dependent random variables and obtained necessary and sufficient conditions for weak convergence of weighted functions of quadratic forms, Shao and Wu [14] considered asymptotic problems in spectral analysis of stationary causal processes, Bhansali et al. [1, 2] established central limit theorems for quadratic forms of causal linear processes with long-memory. Many authors, including Tanaka [13], Horvath and Shao [5], Phillips [11] employed properties of Brownian motion in their derivations.

All our results evolve around the L_p-approximability notion introduced in Mynbaev [6]. The general idea behind L_p-approximability is to represent converging sequences of deterministic vectors with functions of a continuous argument. It is realized as follows. Let $1 \le p < \infty$. The interpolation operator $\Delta_{np} : \mathbf{R}^n \to L_p(0, 1)$ is defined by

$$\left(\Delta_{np} w\right)(x) = n^{\frac{1}{p}} \sum_{t=1}^{n} w_t \mathbf{1}_{\left[\frac{t-1}{n}, \frac{t}{n}\right)}(x), \quad w \in \mathbf{R}^n. \tag{1.4}$$

If $w_n \in \mathbf{R}^n$ for each n and there exists a function $W \in L_p(0, 1)$ such that

$$\left\|\Delta_{np} w_n - W\right\|_{L_p(0,1)} \to 0, \quad n \to \infty,$$

then we say that $\{w_n\}$ is L_p-approximable and also that it is L_p-close to W.

Our results concerning quadratic forms (1.3) are cast in a different format and have a different area of applicability that those from the papers cited above. The format, which links the asymptotic distribution to integral operators, was suggested by Nabeya and Tanaka [10]. They required the integral operators to have continuous symmetric kernels and the $\{v_t\}$ to be independent. Using the L_p-approximability notion allowed Mynbaev [6] to get rid of the kernel continuity condition and replace independent $\{v_t\}$ by non-causal short-memory linear processes. Here we go one step further by lifting the kernel symmetry condition (see Section 2).

In case of CLT's for linear forms (1.1) the method developed in [7] has three advantages. Firstly, all sequences arising in the theory of regressions

$$y_t = \alpha + \beta L(t) + u_t \tag{1.5}$$

with a slowly varying regressor $L(t)$ turn out to be L_p-approximable. Secondly, as is shown in [7], using L_p-approximability allows one to bypass some difficulties arising in the Brownian motion method. Thirdly, as long as the linear process (1.2) is short-memory, to have convergence of (1.1) in distribution, it is enough to establish that $\{w_n\}$ is L_2-close to some $W \in L_2(0,1)$.

It is this last fact that lets us concentrate on establishing L_p-approximability of certain sequences, which we do in Section 3. Here is an overview of the related results.

Consider a polynomial trend $f_n = \left(1^{k-1}, ..., n^{k-1}\right)$ or a logarithmic trend $f_n = \left(\ln^k 1, ..., \ln^k n\right)$ and normalize it to get $w_n = f_n / \left(\sum_{j=1}^n \left|f_{nj}\right|^p\right)^{1/p}$. Then $\{w_n\}$ is L_p-approximable for $1 \le p < \infty$ [8, Theorem 2.7.1]. Here and below $k \ge 0$ is an integer.

Using the fact that certain spatial matrices are L_2-approximable, Mynbaev [9] gave the first derivation of the asymptotic distribution of the OLS estimator for spatial models that does not rely on high level conditions.

A real-valued, positive, measurable function L on $[A, \infty)$ is slowly varying (SV) if

$$\lim_{x\to\infty} \frac{L(rx)}{L(x)} = 1 \text{ for any } r > 0. \tag{1.6}$$

Denote

$$\varepsilon(x) = \frac{xL'(x)}{L(x)}, \quad G(t,n) = \frac{L(t) - L(n)}{L(n)\varepsilon(n)}, \quad w_{nt} = n^{-\frac{1}{p}} G^k(t,n), \quad t = 1, ..., n. \tag{1.7}$$

Phillips [11] pointed out the importance of function $G(t,n)$ for regression (1.5) with stable errors and established a series of its properties, among them the fact that

$$G(rn,n) = \log r[1 + o(1)] \text{ uniformly in } r \in [a,b] \text{ for any } 0 < a < b < \infty. \tag{1.8}$$

Then under some conditions $\{w_n\}$ is L_p-close to $\log^k x$ [8, Theorem 4.4.1].

Denote

$$\eta(x) = \frac{x\varepsilon'(x)}{\varepsilon(x)}, \quad \mu(x) = \frac{1}{2}\left[\varepsilon(x) + \eta(x)\right], \quad H(t,n) = \frac{G(t,n) - \log \frac{t}{n}}{\mu(n)},$$

$$w_{nt} = n^{-1/p} H(t,n), \quad t = 1, ..., n. \tag{1.9}$$

Then $\{w_n\}$ is L_p-close to $\log^2 x$ [8, Theorem 4.4.8].

Sequences (1.7) and (1.9) appear in the theory of regression (1.5) with stationary errors $\{u_t\}$. In case of nonstationary errors, we need three more sequences:

$$F(t,n) = \frac{1}{nL(n)} \sum_{j=t}^n L(j), \quad w_{nt} = n^{-\frac{1}{p}} F^k(t,n), \quad t = 1, ..., n, \tag{1.10}$$

$$I(t,n) = \frac{1}{n}\sum_{j=t}^{n} G(j,n), \quad w_{nt} = n^{-\frac{1}{p}} I^k(t,n), \quad t = 1,...,n, \tag{1.11}$$

$$J(t,n) = \frac{1}{n}\sum_{j=t}^{n} \left(L(j) - \bar{L}\right), \quad \text{where } \bar{L} = \frac{1}{n}\sum_{k=1}^{n} L(k), \quad w_{nt} = n^{-\frac{1}{p}} J^k(t,n), \quad t = 1,...,n. \tag{1.12}$$

(1.10) is L_p-close to $(1-x)^k$ (Theorem 3.1 below), (1.11) is L_p-close to $\left(x \log \frac{1}{x} - 1 + x\right)^k$ (Theorem 3.2 below) and (1.12) is L_p-close to $\left(t \log \frac{1}{t}\right)^k$ (Theorem 3.3 below).

As one can see from this list, by looking at a sequence it is difficult to guess its L_p-limit.

2. Central limit theorems for linear and quadratic forms

The main subject of this section is convergence in distribution of quadratic forms (1.3). For this we need some facts from the theory of operators in Hilbert spaces (all of them can be found in [4]). Let A be a compact linear operator in a Hilbert space with a scalar product $(\cdot,\cdot)$. The operator $H = (A^*A)^{\frac{1}{2}}$ is called the modulus of A, here A^* is the adjoint operator of A. The eigenvalues of H, denoted s_i, $i = 1, 2, ...$ and counted with their multiplicity, are called s-numbers of A. U denotes a partially isometric operator that isometrically maps the range $R(A^*)$ onto the range $R(A)$. Then we have the polar representation $A = UH$. Denote $r(A)$ the dimension of the range $R(A)$ ($r(A) \le \infty$).

Let $\left\{\phi_j\right\}$ be an orthonormal system of eigenvectors of H which is complete in $R(H)$. Then, we have the representation

$$Ax = \sum_{i=1}^{r(A)} s_i (x, \phi_i) U\phi_i$$

or, denoting $\psi_i = U\phi_i$,

$$Ax = \sum_{i=1}^{r(A)} s_i (x, \phi_i) \psi_i, \tag{2.1}$$

where $\{\phi_i\}$ and $\{\psi_i\}$ are orthonormal systems, $H\phi_i = s_i\phi_i$, $\lim_{i\to\infty} s_i = 0$. In particular, when A is selfadjoint, ϕ_i are eigenvectors of A and $s_i = |\lambda_i|$, where λ_i are eigenvalues of A.

Let us apply (2.1) to an integral operator

$$(\mathcal{K} f)(s) = \int_0^1 K(s,t) f(t) dt, \quad f \in L_2(0,1),$$

with a square-integrable kernel $K \in L_2\left((0,1)^2\right)$. From

$$\int K(s,t) f(t) dt = \sum_i s_i \int f(t)\phi_i(t) dt \psi_i(s)$$

we get

$$\int \left[K(s,t) - \sum_i s_i \psi_i(s)\phi_i(t)\right] f(t) dt = 0 \ \ a.e.$$

Because f is arbitrary, we have the decomposition

$$K(s,t) = \sum_{i=1}^{r(A)} s_i \psi_i(s)\phi_i(t), \tag{2.2}$$

where s_i and ϕ_i are, respectively, the eigenvalues and eigenvectors of $(\mathcal{K}^*\mathcal{K})^{\frac{1}{2}}$ and $\psi_j = U\phi_j$.

The fundamental idea of Nabeya and Tanaka [10] was to postulate that the matrices k_n in (1.3) approach in some sense a function K on $(0,1)^2$ and express the limit properties of $Q_n(k_n)$ in terms of the properties of the associated integral operator $\mathcal{K}$. The operator $\mathcal{K}$ is called nuclear if $\sum s_i < \infty$ ($\sum |\lambda_i| < \infty$ when $\mathcal{K}$ is selfadjoint). Nabeya and Tanaka required K to be continuous and symmetric and $\mathcal{K}$ to be nuclear. Mynbaev [6] used L_p-approximability to relax the continuity assumption and replace i.i.d. $\{v_t\}_{t\in\mathbf{Z}}$ with linear processes. Here we develop his approach further by lifting the symmetry condition.

Let $K \in L_2\left((0,1)^2\right)$. For each natural n and $1 \le p < \infty$, we define an $n \times n$ matrix

$$\left(\delta_{np}K\right)_{ij} = n^{2(1-\frac{1}{p})} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \int_{\frac{j-1}{n}}^{\frac{j}{n}} K(s,t)\,dsdt, \quad 1 \le i, \quad j \le n. \tag{2.3}$$

We say that the sequence $\{k_n\}$ is L_2-close to K if

$$\left(\sum_{i,j} (k_n - \delta_{n2}K)_{ij}^2\right)^{\frac{1}{2}} = \|k_n - \delta_{n2}K\|_2 \to 0.$$

Unlike the one-dimensional case, where L_2-approximability of $\{w_n\}$ is enough to have convergence in distribution, in the two-dimensional case one has to impose a stronger condition on the rate of approximation. Mynbaev [8] proposed two such conditions. In Theorem 3.9.1 the conditions on the innovations are weaker (e_t^2 must be uniformly integrable) and the requirement on the rate of approximation

$$\|k_n - \delta_{n2}K\|_2 = o\left(\frac{1}{n}\right) \tag{2.4}$$

is stronger than in Theorem 3.9.7, where the fourth moments Ee_t^4 must exist but the rate of approximation

$$\|k_n - \delta_{n2}K\|_2 = o\left(\frac{1}{\sqrt{n}}\right) \tag{2.5}$$

is less restrictive. For simplicity, we adhere to Assumption 1, which allows us to use (2.5), remembering that in cases (2.4) and (2.5) Mynbaev's conditions on $\{v_t\}_{t\in\mathbf{Z}}$ from Theorems 3.9.1 and 3.9.7 can be repeated word for word. So, one of the main results of this section is the following:

Theorem 2.1. *Let $\{v_t\}_{t\in\mathbf{Z}}$ satisfy Assumption 1 and let (2.5) hold. If $\mathcal{K}$ is nuclear, then*

$$Q_n(k_n) \xrightarrow{d} \left(\sigma_e \sum_i c_i\right)^2 \sum_{i\ge 1} s_i u_i^{(1)} u_i^{(2)},$$

where $\left\{u_i^{(1)}\right\}$, $\left\{u_i^{(2)}\right\}$ *are systems of independent (within a system) standard normals,* s_i *are s numbers of* $\mathcal{K}$ *and*

$$cov(u_i^{(1)}, u_j^{(2)}) = (\psi_i, \phi_j) \quad \textit{for all} \quad i, j.$$

If K is symmetric, then $u_i^{(1)} = u_i^{(2)}$ *for all i.*

Proof. We can exclude symmetric K covered in [8]. The proof is similar to that of Theorem 3.9.7 (all references are to [8]), so we indicate only the modifications. (2.2) above is analogous to equation (3.38), which holds in the symmetric case. Hence, the initial segment of (2.2) is $K_L(s,t) = \sum_{i=1}^{L} s_i\psi_i(s)\phi_i(t)$. Subtracting from (2.2) its initial segment and applying Lemma 3.6.1(iii) we get

$$\left(\delta_{n2}^2 K - \delta_{n2}^2 K_L\right)_{s,t} = \sum_{i>L} s_i \left(\delta_{n2}^1\psi_i\right)_s \left(\delta_{n2}^1\phi_i\right)_t, \tag{2.6}$$

where $\delta_{n2}^2 = \delta_{n2}$ is the two-dimensional discretization operator defined in (2.3) and δ_{n1}^1 is its one-dimensional version defined by

$$\left(\delta_{np}F\right)_i = n^{1-\frac{1}{p}} \int_{\frac{i-1}{n}}^{\frac{i}{n}} F(x)dx, \quad i = 1, ..., n.$$

Combining (1.3) and (2.6) we have

$$\begin{aligned} Q_n\left(\delta_{n2}^2 K\right) - Q_n\left(\delta_{n2}^2 K_L\right) &= \sum_{i>L} s_i \sum_{s,t=1}^{n} \left(\delta_{n2}^1\psi_i\right)_s v_s \left(\delta_{n2}^1\phi_i\right)_t v_t \\ &= \sum_{i>L} s_i \left[\left(\delta_{n2}^1\psi_i\right)' v\right]\left[\left(\delta_{n2}^1\phi_i\right)' v\right]'. \end{aligned} \tag{2.7}$$

By Section 3.5.5 about the T-decomposition for means of quadratic forms

$$\begin{aligned} \left|E\left(\left[\left(\delta_{n2}^1\psi_i\right)' v\right]\left[\left(\delta_{n2}^1\phi_i\right)' v\right]'\right)\right| &= \sigma_e^2 \left|\left(T_n^0\delta_{n2}^1\psi_i, T_n^0\delta_{n2}^1\phi_i\right)\right. \\ &\quad \left. + \left(T_n^-\delta_{n2}^1\psi_i, T_n^-\delta_{n2}^1\phi_i\right) + \left(T_n^+\delta_{n2}^1\psi_i, T_n^+\delta_{n2}^1\phi_i\right)\right| \end{aligned}$$

(applying the Cauchy-Schwarz inequality)

$$\begin{aligned} &\le \sigma_e^2[\left\|T_n^0\delta_{n2}^1\psi_i\right\|_2 \left\|T_n^0\delta_{n2}^1\phi_i\right\|_2 + \left\|T_n^-\delta_{n2}^1\psi_i\right\|_2 \left\|T_n^-\delta_{n2}^1\phi_i\right\|_2 \\ &\quad + \left\|T_n^+\delta_{n2}^1\psi_i\right\|_2 \left\|T_n^+\delta_{n2}^1\phi_i\right\|_2] \end{aligned}$$

(using boundedness of the operators T_n^0, T_n^-, T_n^+, see Section 2.3.2)

$$\le 3\left(\sigma_e\alpha_c\right)^2 \left\|\delta_{n2}^1\phi_i\right\|_2 \left\|\delta_{n2}^1\psi_i\right\|_2$$

(using boundedness of the operators δ_{n2}^1, see Section 2.1.3(ii))

$$\le 3\left(\sigma_e\alpha_c\right)^2 \|\phi_i\|_2 \|\psi_i\|_2 = 3\left(\sigma\alpha_c\right)^2. \tag{2.8}$$

By nuclearity of $\mathcal{K}$ from (2.7)-(2.8) we have

$$\left|E\left(Q_n\left(\delta_{n2}^2 K\right)\right) - Q_n\left(\delta_{n2}^2 K_L\right)\right| \le 3\left(\sigma_e \alpha_c\right)^2 \sum_{i>L} s_i \to 0, \quad L \to \infty.$$

The conclusion is the same as in Section 3.9.3:

$$\operatorname{plim}_{L\to\infty}\left[Q_n\left(\delta_{n2}K\right) - Q_n\left(\delta_{n2}K_L\right)\right] = 0 \quad \text{uniformly in} \quad n.$$

Turning to the analog of Section 3.9.4, note that by selecting

$$w_n^l = \delta_{n2}\psi_l, \ \ l = 1, ..., L; \ \ w_n^l = \delta_{n2}\phi_l, \ \ l = L+1, ..., 2L,$$

we satisfy condition (ii) of Theorem 3.5.2 with

$$F_l = \psi_l, l = 1, ..., L; \ \ F_l = \phi_l, \ \ l = L+1, ..., 2L.$$

With $W_n = \left(w_n^1, ..., w_n^{2L}\right)$ by Theorem 3.5.2 we have

$$W_n' v \xrightarrow{d} N\left(0, \left(\sigma_e \sum_i c_i\right)^2 G\right), \tag{2.9}$$

where

$$G = \begin{pmatrix} (\psi_1,\psi_1) & \dots & (\psi_1,\psi_L) & (\psi_1,\phi_1) & \dots & (\psi_1,\phi_L) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (\psi_L,\psi_1) & \dots & (\psi_L,\psi_L) & (\psi_L,\phi_1) & \dots & (\psi_L,\phi_L) \\ (\phi_1,\psi_1) & \dots & (\phi_1,\psi_L) & (\phi_1,\phi_1) & \dots & (\phi_1,\phi_L) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (\phi_L,\psi_1) & \dots & (\phi_L,\psi_L) & (\phi_L,\phi_1) & \dots & (\phi_L,\phi_L) \end{pmatrix}.$$

Since both systems $\{\phi_i\}$, $\{\psi_i\}$ are orthonormal, this can be written as $G = \begin{pmatrix} I & H \\ H' & I \end{pmatrix}$, where the identities are of size $L \times L$ and H has elements $\left(\phi_i, \psi_j\right)$. It follows that (2.9) is equivalent to

$$W_n' v \xrightarrow{d} \left|\sigma_e \sum_i c_i\right| \begin{pmatrix} u^{(1)} \\ u^{(2)} \end{pmatrix}, \tag{2.10}$$

where $u^{(1)}$, $u^{(2)}$ are standard normal vectors and $cov\left(u^{(1)}, u^{(2)}\right) = H$.

Similarly to equation (2.7),

$$Q_n\left(\delta_{n2}^2 K_L\right) = \sum_{i=1}^{L} s_i \left(\delta_{n2}^1 \psi_i\right)' v \left(\delta_{n2}^1 \phi_i\right)' v.$$

This is a continuous function of the vector at the left of (2.10). By the continuous mapping theorem then

$$Q_n\left(\delta_{n2}^2 K_L\right) \xrightarrow{d} \left(\sigma_e \sum c_i\right)^2 \sum_{i=1}^{L} s_i u_i^{(1)} u_i^{(2)}, \ \ n \to \infty.$$

Establishing the analog of 3.9.4 is complete.

3.9.6 goes through with obvious changes. 3.9.10 is not impacted by the fact that $\mathcal{K}$ is not symmetric. The proof of the generalization of Theorem 3.9.7 is complete. □

Recall the discussion about rates of approximation (2.4), (2.5). An interesting question is: under what conditions on matrices k_n and the kernel K just $\|k_n - \delta_{n2}K\|_2 = o(1)$ would be enough for the CLT to hold? The answer contained in the next theorem means that this is true when essentially the two-dimensional case can be reduced to the one-dimensional.

Theorem 2.2. *Let Assumption 1 hold and suppose that f_n is L_2-close to F and g_n is L_2-close to G:*

$$\|f_n - \delta_{n2}F\|_2 \to 0, \ \ \|g_n - \delta_{n2}G\|_2 \to 0. \tag{2.11}$$

Here f_n, $g_n \in R^n$ for each n, F, $G \in L_2(0,1)$. Put $k_n = f_n g_n'$, $K(s,t) = F(s)G(t)$. The integral operator $\mathcal{K}$ with this kernel is not symmetric but it is nuclear (it is degenerate). Denote $F_0 = F/\|F\|_2$, $G_0 = G/\|G\|_2$. Then

$$Q_n(k_n) = v'k_n v_n \xrightarrow{d} \left(\sigma \sum_{i\in\mathbf{Z}} c_i\right)^2 \|F\|_2 \, \|G\|_2 \, u_1 u_2, \tag{2.12}$$

where u_1, u_2 are standard normal and $cov\,(u_1, u_2) = \int_0^1 F_0(t)G_0(t)dt$

Proof. In the proof of Theorem 2.1 we showed how to deal with the fact that K is not symmetric. Here we show how to lift the restriction (2.5). By Lemma 3.6.1(iii) $\left(\delta_{n2}^2 K\right)_{st} = \left(\delta_{n2}^1 F\right)_s \left(\delta_{n2}^1 G\right)_t$. For an $n \times n$ matrix A denote

$$g(A) = \left[E(v_n' A v_n)^2\right]^{1/2}.$$

Since $g(A)$ is a seminorm, we have

$$g\left(k_n - \delta_{n2}^2 K\right) = g\left(f_n g_n' - \left(\delta_{n2}^1 F\right)\left(\delta_{n2}^1 G\right)'\right)$$

$$\leq g\left(\left(f_n - \delta_{n2}^1 F\right) g_n'\right) + g\left(\left(\delta_{n2}^1 F\right)\left(g_n - \delta_{n2}^1 G\right)'\right). \tag{2.13}$$

Here the matrices $A_1 = f_n - \delta_{n2}^1 F$, $A_2 = \delta_{n2}^1 F$ are just columns and the matrices $B_1 = g_n'$, $B_2 = \left(g_n - \delta_{n2}^1 G\right)'$ are just rows. Applying the last inequality of Section 3.9.9, we have

$$E\left(v_n' A_i B_i v_n\right)^2 \leq c\,\|A_i\|_2^2 \, \|B_i\|_2^2, \ \ i = 1, 2,$$

which is just another way of writing

$$g\left(\left(f_n - \delta_{n2}^1 F\right) g_n'\right) \leq c \left\|f_n - \delta_{n2}^1 F\right\|_2 \|g_n\|_2,$$

$$g\left(\left(\delta_{n2}^1 F\right)\left(g_n - \delta_{n2}^1 G\right)'\right) \leq c \left\|\delta_{n2}^1 F\right\|_2 \left\|g_n - \delta_{n2}^1 G\right\|_2. \tag{2.14}$$

By Lemmas 2.1.3(ii) and 2.5.2(i) $\sup_n \|g_n\| < \infty$ and $\sup_n \left\|\delta^1_{n2} F\right\|_2 < \infty$, so (2.11), (2.13), (2.14) imply $g\left(k_n - \delta^2_{n2} K\right) \to 0$. This gives [8, (3.50)]. The rest of the proof of convergence in distribution is the same.

We need to justify the format of the limit distribution. The operators $\mathcal{K}$ and $\mathcal{K}^*$ are given by

$$(\mathcal{K} f)(s) = \int K(s,t) f(t) dt = F(s) \int G(t) f(t) dt = F(s) (G, f), \tag{2.15}$$

and

$$(\mathcal{K}^* g)(u) = \int K(s,u) g(s) ds = G(u) \int F(s) g(s) = G(u) (F, g). \tag{2.16}$$

Hence,

$$(\mathcal{K}^*\mathcal{K} f)(u) = G(u)(G, f) \|F\|_2^2 .$$

If f is an eigenvector of $\mathcal{K}^*\mathcal{K}$, it should be proportional to $G : f = cG$, and from the above $\mathcal{K}^*\mathcal{K} f = \lambda f$ implies

$$G(u) c \|G\|_2^2 \|F\|_2^2 = \lambda c G(u).$$

This gives $\lambda = \|G\|_2^2 \|F\|_2^2$ and $s_1 = \|G\|_2 \|F\|_2$. The corresponding eigenvector is G_0. The subspace H_1 of functions proportional to G is one-dimensional. Let $f \perp H_1$, that is, $(G, f) = 0$. (2.15)-(2.16) show that $\mathcal{K}^*\mathcal{K} f = 0$ on all such functions. Hence, $s_j = 0$ for $j > 1$. From (2.15)-(2.16) we see that the range $R(\mathcal{K}^*)$ is spanned by $G_0 = G/\|G\|_2$ and the range $R(\mathcal{K})$ is spanned by $F_0 = F/\|F\|_2$. The required partially isometric operator obtains by setting $UG_0 = F_0$. Thus, (2.12) follows from Theorem 2.1, where

$$s_1 = \|F\|_2 \|G\|_2, \quad s_j = 0 \quad \text{for} \quad j > 1, \quad u^{(1)}, \quad u^{(2)} \quad \text{are standard normal and}$$
$$cov\left(u^{(1)}, u^{(2)}\right) = (\phi_1, \psi_1) = (F_0, G_0). \ \square$$

3. Slow variation and L_p-approximability

The seemingly innocuous condition (1.6) in fact entails many strong properties. We shall be using, often without explicitly mentioning, the following standard properties of SV functions (see [12]):

a) If L is SV, then L^a is SV for any $a \in R$.

b) If L and M are SV, then $L + M$ and LM are SV.

c) If L is SV, then (1.6) is actually uniform in $r \in [a, b]$, for any $0 < a < b < \infty$ (uniform convergence theorem).

d) If L is SV, then $x^\gamma L(x) \to \infty$, $x^{-\gamma} L(x) \to 0$ for any $\gamma > 0$.

If L is SV, then by Karamata's theorem there exist a number $B \geq A > 0$ and functions μ, ε on $[B, \infty)$ such that

$$L(x) = \exp\left(\mu(x) + \int_B^x \varepsilon(t) \frac{dt}{t}\right), \tag{3.1}$$

here μ is bounded, measurable, $\lim\limits_{x\to\infty} \mu(x)$ exists and is finite, ε is continuous on $[B, \infty)]$ and $\lim\limits_{x\to\infty} \varepsilon(x) = 0$.

Following Phillips [11], we make a simplifying assumption that $\mu = const$. Phillips argues that asymptotically this does not affect regression estimation. To this justification we can add that if μ is good in the sense that μ is continuously differentiable and $\lim_{x\to\infty} x\mu'(x) = 0$, then the Phillips assumption is satisfied because (3.1) can be equivalently written as

$$L(x) = c_L \exp\left(\int_B^x \varepsilon(t)\frac{dt}{t}\right) \tag{3.2}$$

with a new continuous function ε on $[B, \infty)$ such that $\lim_{x\to\infty} \varepsilon(x) = 0$ (see [8, p.133]). When (3.2) holds, we write $L = K(\varepsilon)$. Further, it is convenient to assume that L is continuous and does not vanish, which can be achieved by properly extending the function ε to $[0, B)$.

Expressions arising in regression statistics involve values $L(t)$ for $1 \le t \le n$. For a fixed $\delta \in (0, 1)$, the values $L(t)$ with $\delta n \le t \le n$ can be handled using the uniform convergence theorem. The values $L(t)$ with $1 \le t \le c$, for any $c > 0$, asymptotically do not present a problem because of continuity of L. To cover the remaining values $L(t)$ with $c \le t \le \delta n$, we need one more condition. Let us call a remainder a positive function ϕ on $[0, \infty)$ with properties:

i) ϕ is nondecreasing and $\lim_{x\to\infty} \phi(x) = \infty$,

ii) there exist positive numbers θ, X such that $x^{-\theta}\phi(x)$ is nonincreasing on $[X, \infty)$.

L is called SV with remainder ϕ if for any $r > 0$ instead of (1.6) one has

$$\frac{L(rx)}{L(x)} = 1 + O\left(\frac{1}{\phi(x)}\right), \quad x \to \infty.$$

The following result allows us to handle the values $L(t)$ with $c \le t \le \delta n$:

Lemma 3.1 (Seneta, 1985, p.102)**.** *If L is SV with remainder ϕ, then for any $b > \theta$ there exist constants $M_b > 0$ and $B_b \ge B$ such that*

$$\left|\frac{L(rx)}{L(x)} - 1\right| \le M_b r^{-b}/\phi(x) \quad \textit{for} \quad x \ge B_b, \quad \frac{B_b}{x} \le r \le 1.$$

Assumption 2 (on SV function L). a) $L = K(\varepsilon)$, that is, (3.2) holds, with ε described after (3.2).

b) ε is SV in the general sense (1.6).

c) There exists a remainder ϕ_ε with properties i), ii) above such that for some $c > 0$

$$\frac{1}{c\phi_\varepsilon(x)} \le |\varepsilon(x)| \le \frac{c}{\phi_\varepsilon(x)} \quad \text{for all} \quad x \ge c. \tag{3.3}$$

We write $L = K(\varepsilon, \phi_\varepsilon)$ to mean that L satisfies Assumption 2. Note that all practically important SV functions from [8] (Table 4.1) satisfy this assumption with $\varepsilon(x) = \frac{xL'(x)}{L(x)}$, $\phi(x) = \frac{1}{|\varepsilon(x)|}$ and number $\theta > 0$ which can be chosen arbitrarily close to zero. For our final results on L_p-approximability, on top of Assumption 2 we shall have to impose more conditions, and all of them hold for functions from Table 4.1.

Let us consider the function F defined in (1.10). Let $[a]$ denote the integer part of $a \in \mathbf{R}$. Now we can proceed with our new results contained in the next lemmas and theorems:

Lemma 3.2. *If $L = K(\varepsilon, \phi_\varepsilon)$, $\theta < 1$, then*

a) $F([rn], n) = 1 - r + o(1)$, $n \to \infty$, uniformly in $r \in \left[\delta, \frac{1}{\delta}\right]$ for any $\delta \in (0, 1)$.

b) For all large n we have

$$|F([rn], n)| \le c \quad \textit{uniformly in} \quad r \in (0, \delta]$$

with a constant c independent of $\delta \in (0, 1/2]$.

Proof. a) $r \in \left[\delta, \frac{1}{\delta}\right]$ implies $n\delta \le rn \le \frac{n}{\delta}$. Since $nr = [nr] + \alpha$ with $0 \le \alpha < 1$, we have for all large n

$$\frac{\delta}{2} \le \delta - \frac{\alpha + 1}{n} \le \frac{[nr] - 1}{n} \le \frac{1}{\delta} - \frac{\alpha + 1}{n} \le \frac{1}{\delta}. \tag{3.4}$$

By [8], Corollary 4.4.3

$$\frac{1}{nL(n)} \sum_{t=1}^{n} L(t) = 1 - \varepsilon(n)[1 + o(1)], \tag{3.5}$$

so

$$F([rn], n) = \frac{1}{nL(n)} \left(\sum_{t=1}^{n} L(t) - \sum_{t=1}^{[rn]-1} L(t) \right)$$

$$= (1 - \varepsilon(n)[1 + o(1)]) - \frac{([rn] - 1)L([rn] - 1)}{nL(n)} (1 - \varepsilon([rn] - 1)[1 + o(1)]) \,. \tag{3.6}$$

According to the definition of ε we can continue (3.6) and have

$$F([rn], n) = (1 + o(1)) - \frac{([rn] - 1)L([rn] - 1)}{nL(n)}(1 + o(1)). \tag{3.7}$$

The $o(1)$ here is uniform in r because by (3.4) $[rn] - 1 \ge n\frac{\delta}{2}$. By the uniform convergence theorem (3.5) also implies $L\left(\frac{[rn]-1}{n} \cdot n\right)/L(n) = 1 + o(1)$. Hence, continuing (3.7)

$$F([rn], n) = 1 + o(1) - \frac{[rn] - 1}{n} \cdot \frac{L([rn] - 1)}{L(n)} \cdot (1 + o(1)) = 1 - r + o(1)$$

uniformly in r.

To prove b), consider two cases.

Case 1. $(B_b + 1)/n \le r \le \delta$, where B_b is the constant from Lemma 3.1. Obviously,

$$|F([rn], n)| \le \left| \sum_{t=1}^{n} \frac{L(t)}{nL(n)} \right| + \left| \sum_{t=1}^{B_b} \frac{L(t)}{nL(n)} \right| + \left| \sum_{t=B_b+1}^{[rn]-1} \frac{L(t)}{nL(n)} \right|. \tag{3.8}$$

By (3.5), the first term at the right is $1 + o(1)$. L is continuous and bounded on $[0, B_b]$, so

$$\left| \sum_{t=1}^{B_b} \frac{L(t)}{nL(n)} \right| \le \frac{cB_b}{nL(n)} \to 0, \quad n \to \infty. \tag{3.9}$$

The third term is the most difficult to bound. From $B_b + 1 \le t \le [rn] - 1$ and $r \le \delta \le \frac{1}{2}$ we have $B_b/n < t/n \le ([rn]-1)/n \le r \le 1$, so by Lemma 3.1

$$\left|\sum_{t=B_b+1}^{[rn]-1} \frac{L(t)}{nL(n)}\right| \le \frac{1}{n}\sum_{t=B_b+1}^{[rn]-1}\left|\frac{L\left(\frac{t}{n}\cdot n\right)}{L(n)} - 1\right| + \frac{1}{n}\sum_{t=B_b+1}^{[rn]-1} 1$$

$$\le \frac{M_b}{n\phi(n)}\sum_{t=B_b+1}^{[rn]-1}\left(\frac{t}{n}\right)^{-b} + \frac{1}{n}\left([rn] - B_b - 1\right). \tag{3.10}$$

Recall that $0 < \theta < 1$ and the number $b > \theta$ is arbitrarily close to θ, so we can choose $0 < b < 1$. Geometrically it is obvious that for any integer $0 < a < N$

$$\sum_{t=a-1}^{N} t^{-b} \le \int_a^N t^{-b}dt \le \int_0^N t^{-b}dt \tag{3.11}$$

and therefore

$$\sum_{t=B_b+1}^{[rn]-1} t^{-b} \le \int_0^{[rn]-1} t^{-b}dt = \frac{([rn]-1)^{1-b}}{1-b}.$$

Using this we can continue (3.10) and get

$$\left|\sum_{t=B_b+1}^{[rn]-1} \frac{L(t)}{nL(n)}\right| \le M_b\frac{n^{b-1}}{\phi(n)}\frac{([rn]-1)^{1-b}}{1-b} + 1$$

$$= \frac{c_1}{\phi(n)}\left(\frac{[rn]}{n} - \frac{1}{n}\right)^{1-b} + 1 \le c_1\frac{r^{1-b}}{\phi(n)} + 1 \le c_2. \tag{3.12}$$

(3.8), (3.9) and (3.12) prove boundedness in Case 1.

Case 2. $0 < r < (B_b + 1)/n$. In this case $[rn] - 1 \le rn - 1 < B_b + 1$. The third sum in (3.8) is empty; the rest of the proof does not change. □

Theorem 3.1. *For $p \in [1,\infty)$ and integer $k \ge 0$ define a vector $w_n \in R^n$ by*

$$w_{nt} = n^{-\frac{1}{p}}F^k(t,n), \quad t = 1,...,n.$$

If $L = K(\varepsilon,\phi_\varepsilon)$, $\theta < 1$, then w_n is L_p-close to $f_k(t) = (1-t)^k$.

Proof. We need the following fact (see [8], pp.149-150): definition (1.4) is equivalent to

$$\left(\Delta_{np}w\right)(u) = n^{1/p}w_{[nu+1]}, \quad 0 \le u < 1. \tag{3.13}$$

Therefore in our case by using (3.13) we obtain $\left(\Delta_{np}w\right)(u) = F^k([nu+1],n)$, $0 \le u < 1$.

Let $0 < \delta \le \frac{1}{2}$, $\delta \le u < 1$. Define $r = \frac{[nu+1]}{n}$. From the inequality $nu < [nu+1] \le nu + 1$ we have

$$\delta \le u < r \le u + \frac{1}{n} < \frac{1}{\delta},$$

if n is sufficiently large. Hence,

$$r = u + o(1), \quad n \to \infty; \quad r \in \left[\delta, \frac{1}{\delta}\right].$$

This and Lemma 3.2 (a) imply

$$F([nu+1], n) = F([rn], n) = 1 - u + o(1), \quad n \to \infty, \tag{3.14}$$

uniformly in $u \in [\delta, 1)$.

Now let $0 < u < \delta$. Then

$$0 < \frac{[nu+1]}{n} = r \le \frac{nu+1}{n} < \delta + \frac{1}{n} \le 2\delta \le 1,$$

if n is large enough. By Lemma 3.2 (b)

$$|F([nu+1], n)| = |F([rn], n)| \le c \quad \text{for} \quad u \in (0, \delta). \tag{3.15}$$

Obviously,

$$\left\|\Delta_{np} - f_k\right\|_{L_p(0,1)} \le \left(\int_\delta^1 \left|F^k([nu+1], n) - (1-u)^k\right|^p du\right)^{1/p}$$
$$+ \left(\int_0^\delta \left|(1-u)^k\right|^p dt\right)^{1/p} + \left(\int_0^\delta \left|F^k([nu+1], n)\right|^p du\right)^{1/p}.$$

By (3.14)-(3.15) this can be made as small as desired, by selecting first a small δ and then a large n. □

Next consider the function I defined in (1.11).

Lemma 3.3. *If $L = K(\varepsilon, \phi_\varepsilon)$, $\theta < 1$, then for each $\delta \in (0,1)$*

a) $I([rn], n) = (1 + o(1))\left(r \log \frac{1}{r} - 1 + r\right)$, $n \to \infty$, uniformly in $r \in \left[\delta, \frac{1}{\delta}\right]$ (the $o(1)$ depends on δ),

b) $|I([rn], n)| \le c$ for $r \in (0, \delta]$, where c does not depend on δ.

Proof. a) If $r \ge \delta$ and $n \ge t \ge [rn]$, then $1 \ge t/n \ge [rn]/n \ge [\delta n]/n \ge \delta - 1/n \ge \delta/2$ for large n. By (1.8)

$$G(t, n) = G\left(\frac{t}{n} n, n\right) = (1 + o(1)) \log \frac{t}{n} \quad \text{for} \quad [rn] \le t \le n,$$

where $o(1)$ does not depend on t. Hence, denoting $s = [rn]$,

$$I([rn], n) = (1 + o(1)) \frac{1}{n} \sum_{t=s}^{n} \log \frac{t}{n} = (1 + o(1)) \frac{1}{n} \log \frac{s(s+1)...n}{n^{n-s+1}}$$

$$= (1 + o(1)) \frac{1}{n} \log \frac{n!}{(s-1)! n^{n-s+1}}.$$

By Stirling's formula [3, p.371], for each natural n there exists $\theta = \theta(n) \in (0, 1)$ such that

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{\theta}{12n}}.$$

So

$$I([rn], n) = (1 + o(1))\frac{1}{n}\log\frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{\theta(n)}{12n}}}{\sqrt{2\pi(s-1)}\left(\frac{s-1}{e}\right)^{s-1} e^{\frac{\theta(s-1)}{12(s-1)}} n^{n-s+1}}$$

$$= (1 + o(1))\frac{1}{n}\log\left[\left(\frac{n}{s-1}\right)^{s-1/2} e^{s-1-n+\frac{\theta(n)}{12n}-\frac{\theta(s-1)}{12(s-1)}}\right]$$

$$= (1 + o(1))\left[\left(\frac{s}{n} - \frac{1}{2n}\right)\log\left(\frac{s-1}{n}\right)^{-1} + \frac{s-1}{n} - 1 + \frac{\theta(n)}{12n^2} - \frac{\theta(s-1)}{12n^2\frac{s-1}{n}}\right]. \tag{3.16}$$

Since $rn - 1 \le s = [rn] \le rn$, we have

$$\frac{s}{n} = r + o(1), \frac{s-1}{n} = r + o(1) \quad \text{uniformly in} \quad r \in \left(\delta, \frac{1}{\delta}\right).$$

Using the fact that $\frac{s-1}{n}$ is bounded and bounded away from zero, $\frac{1}{\delta} \ge r - \frac{1}{n} \ge \frac{s-1}{n} \ge r - \frac{2}{n} \ge \frac{\delta}{2}$ for large n, we have the bounds

$$\left|\frac{1}{n}\log\left(\frac{s-1}{n}\right)^{-1}\right| \le \frac{1}{n}C(\delta), \quad \left|\frac{1}{n^2(s-1)/n}\right| \le \frac{2}{n^2\delta}.$$

This and (3.16) prove part a).

b) First consider the case $\frac{B_b+1}{n} \le r < \delta$ and write

$$|I([rn], n)| \le \frac{1}{n|\varepsilon(n)|}\sum_{t=[rn]}^{n}\left|\frac{L(t)}{L(n)} - 1\right|.$$

From $rn - 1 \le [rn] \le t \le n$ we have $\frac{B_b}{n} \le r - \frac{1}{n} \le \frac{t}{n} \le 1$, so by Lemma 3.1 and (3.11)

$$|I([rn], n)| \le \frac{M_b}{n|\varepsilon(n)|\phi(n)}\sum_{t=[rn]}^{n}\left(\frac{t}{n}\right)^{-b} =$$

$$= \frac{M_b n^{b-1}}{|\varepsilon(n)|\phi(n)}\sum_{t=[rn]}^{n} t^{-b} \le \frac{M_b}{1-b}\cdot\frac{1}{|\varepsilon(n)|\phi(n)} \le C. \tag{3.17}$$

The last bound by (3.3).

Now let $0 < r < \frac{B_b+1}{n}$. Then for $t \le [rn] - 1$ we have $t \le rn \le B_b + 1$ and $|L(t)| \le c_1$. By [8], Corollary 4.4.2, we have $\frac{1}{n}\sum_{t=1}^{n} G(t, n) \to 1$, so

$$|I([rn], n)| \le \left|\frac{1}{n}\sum_{t=1}^{n} G(t, n)\right| + \frac{1}{n}\left|\sum_{t=1}^{[rn]-1} G(t, n)\right|$$

$$\le c_2 + \frac{1}{n|\varepsilon(n)|}\sum_{t=1}^{[rn]-1}\left(\frac{|L(t)|}{L(n)}+1\right) \le c_2 + \frac{[rn]-1}{n\,|\varepsilon(n)|}\left(\frac{c_1}{L(n)}+1\right) \le c_3. \tag{3.18}$$

This is because $|\varepsilon(n)|$, $|\varepsilon(n)L(n)|$ are SV and $n|\varepsilon(n)| \to \infty$, $n\,|\varepsilon(n)L(n)| \to \infty$. (3.17) and (3.18) prove b). □

Theorem 3.2. *For $p \in [1,\infty)$ and integer $k \ge 0$ define a vector $w_n \in R^n$ by*

$$w_{nt} = n^{-\frac{1}{p}} I^k(t,n), \quad t = 1,...,n.$$

If $L = K(\varepsilon,\phi_\varepsilon)$, $\theta < 1$, then w_n is L_p-close to $f_k(t) = \left(t\log\frac{1}{t} - 1 + t\right)^k$.

Proof. The proof is similar to that of Theorem 3.1, just replace Lemma 3.2 with Lemma 3.3. □

Now consider the function J defined in (1.12).

Lemma 3.4. *Suppose $L = K(\varepsilon,\phi_\varepsilon)$, $\theta < 1$. Then for each $\delta \in (0,1)$*
a) $J([rn],n) = (1+o(1))r\log\frac{1}{r}$, $n \to \infty$, uniformly in $r \in \left[\delta,\frac{1}{\delta}\right]$.
b) $|J([rn],n)| \le c$ for $0 < r \le \delta$, where c does not depend on δ.

Proof. Obviously,

$$J([rn],n) = \frac{1}{n}\sum_{t=[rn]}^{n}\frac{L(t)-L(n)}{L(n)\varepsilon(n)} + \frac{1}{n}\sum_{t=[rn]}^{n}\frac{L(n)-\bar{L}}{L(n)\varepsilon(n)}$$
$$= I([rn],n) + \frac{1}{n}\sum_{[rn]}^{n}\frac{L(n)-\bar{L}}{L(n)\varepsilon(n)}.$$

Use here (3.5) to get

$$J([rn],n) = I([rn],n) + (1+o(1))\frac{1}{n}\sum_{t=[rn]}^{n}1$$
$$= I([rn],n) + (1+o(1))\frac{n-[rn]+1}{n}$$

(applying Lemma 3.3a))

$$= (1+o(1))(r\log(1/r) - 1 + r) + (1+o(1))(1 - r + o(1)) = (1+o(1))r\log(1/r).$$

In all of the above the $o(1)$ does not depend on r.
b) In case $0 < r \le \delta$ just use part b) of Lemma 3.3 instead of part a). □

Theorem 3.3. *For $p \in [1,\infty)$ and integer $k \ge 0$ define a vector $w_n \in R^n$ by*

$$w_{nt} = n^{-\frac{1}{p}} J^k(t,n), \quad t = 1,...,n.$$

If $L = K(\varepsilon,\phi_\varepsilon)$, $\theta < 1$, then w_n is L_p-close to $f_k(t) = \left(t\log\frac{1}{t}\right)^k$.

Proof. Just replace Lemma 3.2 in the proof of Theorem 3.1 with Lemma 3.4.□

References

[1] R. Bhansali, L. Giraitis, P. Kokoszka, Convergence of quadratic forms with nonvanishing diagonal, Statistics and Probability Letters 77 (7) (2007) 726–734.

[2] R. Bhansali, L. Giraitis, P. Kokoszka, Approximations and limit theory for quadratic forms of linear processes, Stochastic Processes and their Applications 117 (1) (2007) 71–95.

[3] G. Fichtenholz, Differential and Integral Calculus II, Nauka, 1968.

[4] I. Gohberg, M. Krein, Introduction to the Theory of Linear Nonselfadjoint Operators, Translations of Mathematical Monographs, 1969.

[5] L. Harvath, Q.-M. Shao, Limit theorems for quadratic forms with applications to Whittle's estimate, The Annals of Applied Probability 9 (1) (1999) 146–187.

[6] K. Mynbaev, L_p-approximable sequences of vectors and limit distribution of quadratic forms of random variables, Advances in Applied Mathematics 26 (4) (2001) 302–329.

[7] K. Mynbaev, Central Limit Theorems for weighted sums of linear processes: L_p-approximability versus Brownian mothion, Econometric Theory 25 (3) (2009) 748–763.

[8] K. Mynbaev, Short-Memory Linear Processes and Econometric Applications, Wiley and Sons, 2011.

[9] K. Mynbaev, A. Ullah, Asymptotic distribution of the OLS estimator for a purely autoregressive spatial model, Journal of Multivariate Analysis 99 (2) (2008) 245–277.

[10] S. Nabeya, K. Tanaka, Asymptotic Theory of a Test for the Constancy of Regression Coefficients Against the Random Walk Alternative, The Annals of Statistics 16 (1) (1988) 218–235.

[11] P. Phillips, Regression with slowly varying regressors and nonlinear trends, Economertic Theory 23 (4) (2007) 557–614.

[12] E. Seneta, Regularly Varying Functions, Nauka, 1985.

[13] K. Tanaka, Time Series Analysis: Nonstationary and Noninvertible Distribution Theory, Wiley and Sons, 1996.

[14] W. Wu, X. Shao, Asymptotic spectral theory for nonlinear time series, The Annals of Statistics 35 (4) (2007) 1773–1801.