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

UDC 519.2

On manuscript rights

DARKENBAYEVA GULSIM SPANDIYAROVNA

Asymptotic theory of regressions with asymptotically collinear regressors

6D060100 – Mathematics

Dissertation is submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy (PhD)

Scientific advisers:

Mynbaev Kairat Turysbekovich,
Doctor of Physical and
Mathematical Sciences, Professor,
Institute of Mathematics and
Mathematical Modeling,
Kazakh British University and
Satbayev University

Carlos Brunet Martins-Filho,
PhD, Professor,
University of Colorado at
Boulder, CO, USA,
Senior Research Fellow,
International Food Policy Research
Institute IFPRI-MTID

Republic of Kazakhstan
Almaty, 2019

CONTENTS

NORMATIVE REFERENCES

In this dissertation thesis the following references for standards were used:

SOSE RK 5.04.034-2011. State obligatory standard of education. Postgraduate education. Doctoral studies.

State standard 7.32-2001 (changes dated 2006). Report on scientific-research work. Structure and rules of presentation.

State Standard 7.1-2003. Bibliographic record. Bibliographic description. General requirements and rules.

DEFINITIONS

In this section presented only the basic definitions and notions (there are taken from [1-2]). For convenience of narration we will recall some of them in the right places and other new ones will be introduced later.

A **norm** of an element x is denoted $\|x\|$ and defined by $\operatorname{dist}(x,0)$ in Euclidean space R^n with Euclidean distance

$$\operatorname{dist}(x,y)=\sqrt{\sum_i (x_i-y_i)^2}$$

between points x, $y\in R^n$. In an abstract situation, we can first axiomatically define the distance $\operatorname{dist}(x,0)$ from x to the origin and then the distance between any two points will be $\operatorname{dist}(x,y)=\operatorname{dist}(x-y,0)$.

A **norm** is a real-valued function $\|\cdot\|$ defined on linear space X if

1. $\|x\|\geq 0$ vector x (nonnegativety),
2. $\|ax\|=|a|\|x\|$ for all numbers a and vectors x (homogeneity) ,
3. $\|x+y\|\leq\|x\|+\|y\|$ for all vectors x and y (triangle inequality) and
4. $\|x\|=0$ implies $x=0$ (nondegeneracy).

A norm more general than $\operatorname{dist}(x,y)$ is obtained by replacing the index 2 by an arbitrary number $p\in[1,\infty)$. In other words, in R^n the function

$$\|x\|_p=\left(\sum_i |x_i|^p\right)^{1/p}$$

satisfies all axioms of a norm. For $p=\infty$, above norm is completed with

$$\|x\|_\infty=\sup_i |x_i|$$

because

$$\lim_{p\to\infty}\|x\|_p=\|x\|_\infty .$$

R^n provided with the norm $\|\cdot\|_p$ is denoted R^n_p ($1\leq p\leq\infty$).

The **space** l_p of infinite sequences of numbers $x=(x_1,x_2,...)$ that have a finite norm $\|x\|_p$ is the most immediate generalization of R^n_p (defined by norm in above definition of norm, where i runs over the set of naturals). More generally, the set of indices $I = \{i\}$ may depend on the context.

The j **th unit vector** in l_p is an infinite sequence $e_j = (0,...,0,1,0,...)$ with unity in the j th place and 0 in all others.

A σ**-field** $\Im$ is a nonempty family $\Im$ of subsets of some set Ω if

1. unions, intersections, differences and complements of any two elements of $\Im$ belongs to $\Im$,

2. the union of any sequence $\{A_n : n = 1, 2, ...\}$ of elements of $\Im$ belongs to $\Im$ and

3. Ω belongs to $\Im$.

In probabilities, σ-fields play the role of information sets.

A **measurable space** is a pair $(\Omega, \Im)$, where Ω is some set and $\Im$ is a σ-field of its subsets. A set function μ defined on elements of $\Im$ with values in the extended half-line $[0, \infty]$ is called *σ-additive measure* if for any disjoint sets A_1, A_2, $\ldots \in \Im$ one has

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j).$$

The space L_p. Let $(\Omega, \Im, \mu)$ be any space with a σ-additive measure μ and let $1 \le p < \infty$. The set of measurable functions $f : \Omega \to R$ provided with the norm

$$\| f \|_p = \left(\int_{\Omega} |f(x)|^p d\mu \right)^{1/p}, \ 1 \le p < \infty,$$

is denoted *space* $L_p = L_p(\Omega)$. In the case $p = \infty$ this definition is completed with

$$\| f \|_\infty = ess \sup_{x \in \Omega} |f(x)| = \inf_{\mu(A)=0} \sup_{x \in \Omega \setminus A} |f(x)|.$$

The term in the middle is, by definition, the quantity at the right and is called *essential supremum*. These definitions mean that the values taken by functions on sets of measure zero don't matter. An equality $f(t) = 0$ is accompanied by the caveat “almost everywhere” (a.e.) or “almost surely” (a.s.) in the probabilistic setup, meaning that there is a set of measure zero outside which $f(t) = 0$.

A sequence of random variables $\{e_t : t = 1, 2, \ldots\}$ is called **adapted to** $\{\Im_t\}$**,** where $\{\Im_t : t = 1, 2, \ldots\}$ is an increasing sequence of σ-fields contained in $\Im$: $\Im_1 \subseteq \cdots \subseteq \Im_n \subseteq \cdots \subseteq \Im$, if e_t is $\Im_t$-measurable for $t = 1, 2, \ldots$.

$\{e_t, \Im_t\}$ or, shorter, $\{e_t\}$ is a **martingale difference** (m.d.) **sequence**, if a sequence of integrable variables $\{e_t\}$ satisfies

1. $\{e_t\}$ is adapted to $\{\Im_t\}$

and

2. $E(e_t \mid \Im_{t-1}) = 0$ for $t = 1, 2, \ldots$, where $\Im_0 = \{\varnothing, \Omega\}$.

Martingale difference (m.d.) **array**. Consider a family $\{\{X_{nt}, \Im_{nt} : t = 1, ..., k_n\} : n = 1, 2, ...\}$, where $\{k_n\}$ is an increasing sequence of integers, $\{X_{nt}\}_{n \in N, t=1,...,k_n}$ are random variables and $\{\Im_{nt}\}_{n \in N, t=1,...,k_n}$ are nested sub-σ-fields of Ω, $\Im_{n,t-1} \subseteq \Im_{nt}$ for all n, t. Such a family is called *martingale difference (m.d.) array* if

1. X_{nt} is $\Im_{nt}$-measurable,
2. X_{nt} is integrable,
3. $E(X_{nt} \mid \Im_{n,t-1}) = 0$ for all n, t.

A family $\{X_\tau : \tau \in T\}$ of random variables is called **uniformly integrable** if

$$\lim_{m\to\infty} \sup_{\tau \in T} E|X_\tau| 1_{|X_\tau| \ge m} = 0.$$

Let $1 \le p < \infty$. For each natural n the set

$$\{t/n : t = 0, \ldots, n\}$$

is called a **uniform partition**. The intervals

$$i_t = [(t-1)/n, t/n)$$

form a disjoint covering of $[0,1)$ of equal length $1/n$. Denoting $[a]$ as the integer part of a real number a, we can see that the condition

$$x \in i_t$$

is equivalent to

$$t - 1 \le nx < t,$$

which, in turn, is equivalent to

$$t = [nx] + 1.$$

The function $[nx]+1$ can be called **a locator** because $x \in i_{[nx]+1}$ for all $x \in [0,1)$.

For each natural n, we can define a **discretization operator**

$$\delta_{np} : L_p \to R_p^n$$

by

$$\left(\delta_{np} F\right)_t = n^{1/q} \int_{i_t} F(x)\, dx,\ t = 1, \ldots, n,\ F \in L_p,$$

where q is the conjugate of p, i.e.

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Up to a scaling factor, the t -th component of $\delta_{np} F$ is the average of F over the interval i_t. For a given $F \in L_p$, the sequence $\{\delta_{np} F : n \in N\}$ is called L_p**-generated** by F.

The **interpolation operator**

$$\Delta_{np} : R^n \to L_p(0,1)$$

is defined by

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

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

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

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

A linear process. Let $\left\{\left\{e_{nt}, \Im_{nt} : t \in Z\right\} : n \in Z\right\}$ be a double-infinite m.d. array. Except that the set is wider, this satisfies the same requirements as a one-sided array $\left\{\left\{X_{nt}, \Im_{nt} : t = 1, \ldots, k_n\right\} : n = 1, 2, \ldots\right\}$. Fixing a summable sequence of numbers $\left\{\psi_j\right\}_{j \in Z}$, denote

$$v_{nt} = \sum_{j \in Z} e_{n,t-j} \psi_j \, , \; t \in Z \, .$$

The array $\{v_{nt} : t, n \in Z\}$ is called a *linear process* (with short-range dependence).

We suppose that for each $n \in N$ given a vector of weights $w_n \in R^n$. The **convolution operator** $T_n : R_p^n \to l_p(Z)$ defined by

$$(T_n w)_j = \sum_{t=1}^{n} w_t \psi_{t-j} \, , \; j \in Z \, ,$$

where $\{\psi_j\}_{j \in Z}$ is a summable sequence of real numbers.

Sometimes it is convenient to represent $T_n w$ as

$$T_n w = \begin{pmatrix} T_n^- w \\ T_n^0 w \\ T_n^+ w \end{pmatrix},$$

where $T_n^+ w : R_p^n \to l_p(j > n)$, $T_n^0 w : R_p^n \to R_p^n(1 \le j \le n)$, $T_n^- w : R_p^n \to l_p(j < 1)$ are defined by

$$\left(T_n^+ w\right)_j = (T_n w)_j , \; j > n \, ,$$
$$\left(T_n^0 w\right)_j = (T_n w)_j , \; 1 \le j \le n \, ,$$
$$\left(T_n^- w\right)_j = (T_n w)_j , \; j < 1 \, .$$

The above three operators are called **trinity** [1, P.55]. Naturally, T_n is called a T**-operator**.

T**-decomposition**. Consider the convolution operator $T_n : R_p^n \to l_p(Z)$ defined by

$$(T_n w)_j = \sum_{t=1}^{n} w_t \psi_{t-j} \, , \; j \in Z \, .$$

For any vector of weights $w_n \in R_2^n$ and linear process $\{v_{nt}\}_{n,t \in Z}$ the linear form $w_n' v_n$ defines T**-decomposition** as

$$w_n' v_n = \sum_{t=1}^{n} w_{nt} \sum_{j \in Z} e_{n,t-j} \psi_j = \sum_{i \in Z} e_{ni} \sum_{t=1}^{n} w_{nt} \psi_{t-j} = \sum_{i \in Z} e_{ni} \left(T_n w_n\right)_i \, .$$

The **discretization** and **interpolation** operators. Corresponding to uniform covering of $(0,1)$ we can define uniform covering of the square $Q=(0,1)^2$ consisting of small squares

$$q_{st}=i_s\times i_t,\ 1\le s,t\le n,$$

of area n^{-2}. For a given $F\in L_p\left((0,1)^2\right)$, $\delta_{np}F$ is defined by

$$\left(\delta_{np}F\right)_{st}=n^{2/q}\int\limits_{q_{st}}F(x)dx,\ 1\le s,t\le n,$$

here $x=(x_1,x_2)$ and dx is the Lebesgue measure on the plane. If f is a matrix of size $n\times n$, the step function $\Delta_{np}f$ is, by definition,

$$\Delta_{np}f=n^{2/p}\sum_{s,t=1}^{n}f_{st}1_{q_{st}}\ ,$$

here 1 is the indicator. δ_{np} and Δ_{np} are called *discretization* and *interpolation* operators, respectively.

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

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

A **Hilbert space** is a linear space that is endowed with a scalar product and is complete in the norm generated by that scalar product.

Let A be a compact linear operator in a Hilbert space with a scalar product $(\cdot,\cdot)$. The operator $H=\left(A^*A\right)^{\frac{1}{2}}$ is called the **modulus** of A, here A^* is the adjoint operator of A. If $A=A^*$, then we say that operator A is selfadjoint. The eigenvalues of H, denoted s_i, $i=1,2,\ldots$, and counted with their multiplicity, are called s-numbers of A.

The operator A is called **nuclear** if

$$\sum s_i<\infty\ (\sum|\lambda_i|<\infty \text{ when } A \text{ is selfadjoint}),$$

where $\{\lambda_i\}$ are eigenvalues of A.

NOTATIONS AND ABBREVIATIONS

$N=\{1,2,3,\ldots\}$	set of positive integers
$N_+=\{0,1,2,3,\ldots\}$	set of positive integers including 0 (zero)
$Z=\{\ldots,-3,-2,-1,0,1,2,3,\ldots\}$	set of integers
R	set of real numbers
R^n	set of n dimensional vectors with elements from R
R_p^n	set R^n provided with the norm $\lVert\cdot\rVert_p$, see definition of norm
l_p	space l_p of infinite sequences of numbers, see definition of the space l_p
L_p	space L_p, see definition of the space L_p
$\underset{n\to\infty}{\overset{P}{\to}}$	convergence in probability as $n\to\infty$
$\underset{n\to\infty}{\overset{d}{\to}}$	convergence in distribution $n\to\infty$
$a\in A$	a is an element of set (space) A
$a\notin A$	a is not an element of set (space) A
$A\subseteq B$	A subset (subspace) of B
$A\subset B$	A subset (subspace) of B but not equal
$\bigcup$	union of sets (spaces)
$\bigcap$	intersection of sets (spaces)
$\setminus$	relative complement of sets (spaces)
$[a]$	the integer part of a real number a
$\lVert f\rVert_p$	function f has a norm of space where it defined
1_A	indicator of set A
δ_{np}	discretization operator
Δ_{np}	interpolation operator
a.e.	almost everywhere
cdf	cumulative distribution function
pdf	probability density function
i.i.d.	independent identically distributed
m.d.	martingale difference
m.a.	martingale array
LIE	Law of iterated expectations
1-D	one dimensional
2-D	two dimensional
a.s.	almost surely

SV	slowly varying
i_t	interval $\left[(t-1)/n, t/n\right)$, where $1 \le t \le n$
$q_{st} = i_s \times i_t$	square, where $1 \le s,t \le n$,

INTRODUCTION

Topicality of the research theme. Consider a model

$$y_t = \alpha + \beta L(t) + u_t,\ t = 1,...,n,$$

where L is a positive, measurable on $[A,\infty)$, $A > 0$, and

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

function, or, shortly, L is a slowly varying function (SV). For the case when the errors $\{u_t\}$ are stationary, Phillips [3] obtained the asymptotic distribution of the OLS estimators $\hat{\alpha}$ and $\hat{\beta}$.

We consider integrated errors

$$u_t = \rho u_{t-1} + v_t,\ t = 2,...,n,$$

where $\rho = 1$ under the null hypothesis and $\{v_t\}$ is a non-causal linear process

$$v_t = \sum_{i\in Z} c_i e_{t-i}\ .$$

Integrated errors and non-causal linear processes have many applications in statistics and econometrics. Results presented in this work can be used in derivation the limiting distribution of the unit root test statistic for our main regression model. Statement of this problem you can read in work of Uematsu [4]. This problem is open at present time.

Also, as one can see we restrict our attention to models with deterministic regressors. Models with such regressors have many applications [5-11].

Another application of this research is to address the problem of early detection of bubbles. This is a macroeconomic problem that has direct implications for monetary and fiscal policies. A school headed by P. Phillips has provided a decisive component of the statistical procedure [12-15].

The aims and objectives of the study. The work is devoted to studying:

1) central limit theorems for quadratic forms of linear processes $\{v_t\}$;

2) add a couple of sequences to the list of L_p-approximable sequences contained in Mynbaev [16];

3) prove Uematsu's result [4, P.10] on the asymptotic distribution of $\hat{\alpha}$ and $\hat{\beta}$ under less restrictive conditions.

The main provisions for the defense of the dissertation:

1) Obtained convergence of some quadratic forms used in regression analysis.

2) Obtained central limit theorems for linear and quadratic forms.

3) Added a couple of sequences to the list of L_p-approximable sequences contained in Mynbaev [16, P.321].

4) Proved Uematsu's result [4, P.10] on the asymptotic distribution of OLS estimations $\hat{\alpha}$ and $\hat{\beta}$ under less restrictive conditions.

5) Done Monte-Carlo simulations for the asymptotic distribution of OLS estimations $\hat{\alpha}$ and $\hat{\beta}$.

The objects of research regression with slowly varying regressors, regression with asymptotically collinear regressors, non-causal linear processes, quadratic forms, central limit theorems.

The research subjects L_p-approximable sequences, quadratic forms of linear processes, central limit theorems.

Research methods L_p-approximation method of Mynbaev (see [16, P.314]), central limit theorems.

Novelty of the dissertation research is that the main model with a slowly varying (SV) regressor in the presence of a unit root, also regression model has integrated errors $u_t = \rho u_{t-1} + v_t$, $t = 2,...,n$, and $\{v_t\}$ is a non-causal linear process.

Results in Section 3 generalize to the non-causal linear processes some statements from [17, P. 979; 18, P.172].

Result in Section 4 extends Mynbaev's theorems [16, P.323] on convergence of quadratic forms to the case of asymmetric kernels.

Section 5 considers a couple of new L_p-approximable sequences.

The study of the main model with integrated errors gave us the asymptotic distribution of $\hat{\alpha}$ and $\hat{\beta}$. Uematsu characterized convergence in distribution of $\hat{\alpha}$ and $\hat{\beta}$, and it turned out to be very different from Phillips [3, P.565] had with stable errors. In Section 6 we prove Uematsu's result [4, P.10] on the asymptotic distribution of $\hat{\alpha}$ and $\hat{\beta}$ under less restrictive conditions.

Theoretical and practical significance of the research. This research constitute step in solving problem about unit root test. Also, has attracted number of applications in Econometrics and Statistics (see [4, P.2; 5, P.1048; 6, P.19; 7, P.59; 8, P.501; 9, P.1; 10, P.1771; 11, P.1153]).

Connection of the dissertation thesis with the other scientific research works. The dissertation work was implemented within the scientific projects of the program of grant financing of fundamental researches in the areas of natural sciences of the Ministry of education and science of the Republic of Kazakhstan "Prediction of rare events and spatial effects in financial and commodity markets" (2015-2017 years, № 4084/GF4) and "Estimation of discontinuous densities and distribution functions in relation to applications in economics, finance and insurance" (2018-2020 years, AP05130154).

The work approbation. Results of the work were presented and discussed at the following conferences [19-21] and seminars: "Actual problems of pure and applied mathematics”, Almaty, 2015; Second International Conference on Statistical Distributions and Applications ICOSDA, Niagara Falls, Canada, 2016; XIII International scientific conference of students and young scientists “Lomonosov-2017”, Astana, 2017; the city scientific seminar "Differential operators and their applications", Almaty, 2017; scientific seminar of Institute of mathematics, physics and informatics, Almaty, 2017; the city scientific seminar "Differential operators and their applications", Almaty, 2019. Results of this dissertation were discussed with probability, statistics and econometrics specialist Carlos Brunet Martins-Filho during the scientific training in University of Colorado at Boulder, Colorado, USA, 2016.

Publications. Based on results of the dissertation 7 works were published: 5 journal articles (1 in Scopus indexed Journal [22] and 4 in journals recommended by the Committee for Control in Education and Science of the Ministry of Education and Science of the Republic of Kazakhstan [23-26]) and 2 in proceedings of international scientific conferences [19, P.24; 20, P.56; 21, P.9].

Volume and structure of the dissertation. The work includes the title page, contents, normative references, definitions, notations and abbreviations, introduction, 7 sections, conclusion and references. Total volume of dissertation is 72 pages, the work contains 5 illustrations, 1 table and 42 literature references.

Main content of the dissertation. The introduction includes actuality of the research theme, aims and objectives, the main provisions for the defense of the dissertation, the research object and subject, methods, novelty and theoretical and practical significance of the research, connection of the dissertation thesis with the other scientific research works, the work approbation, author’s publications, and volume, structure and content of the dissertation thesis.

The first section contains a more detailed introduction to the work.

The second section gives preliminary results like useful lemmas, theorems and assumptions which will be used in dissertation work.

The third section consists of research on convergence of some quadratic forms used in regression analysis.

The fourth section gives central limit theorems for linear and quadratic forms with proofs.

The fifth section is about slow variation and L_p-approximality. Here we add a couple of new sequences to the list of L_p-approximable sequences contained in Mynbaev with proofs.

The sixth section contains the proof of the asymptotic distribution of the OLS estimators $\hat{\alpha}$ and $\hat{\beta}$.

The seventh section contains the Monte-Carlo simulations for the OLS estimators $\hat{\alpha}$ and $\hat{\beta}$.

The conclusion lists and generalizes the main results obtained during the implementation of the dissertation thesis.

1 OVERVIEW AND BRIEFLY ABOUT RESULTS

Convergence in distribution of sequences of random variables plays a central role in the theory of probability 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. This dissertation work concentrates on weak convergence of linear and quadratic forms arising in regression analysis. The book by Tanaka [18] can serve as a comprehensive introduction to this area.

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

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

where

$$v = (v_1, \ldots, v_n)' \in R^n$$

is a random vector and

$$w_n = (w_{n1}, \ldots, w_{nn})' \in R^n$$

is a deterministic vector.

One popular approach to modeling dependence of $\{v_t\}$ over time is to specify it as a linear process defined by the convolution

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

where $\{e_i\}_{i \in Z}$ is a sequence of random variables and $\{c_i\}_{i \in Z}$ is a double-infinite sequence of numbers.

If $c_i = 0$ for $i < 0$, we have

$$v_t = \sum_{i=0}^{+\infty} c_i e_{t-i}, \ t \in Z, \tag{1.3}$$

the process (1.2) is called causal; otherwise it is called non-causal.

The theory critically depends on whether

$$\alpha_c \equiv \sum_{i \in Z} |c_i| < \infty \tag{1.4}$$

or

$$\alpha_c = \infty \text{ but } \sum_{i \in Z} c_i^2 < \infty . \tag{1.5}$$

If (1.4) holds we say that (1.2) is a short-memory process, if (1.5) holds we say that (1.2) is a long-memory process.

Square-integrable m.d. sequences are uncorrelated and have mean zero (see proof of this fact in Mynbaev [1, P.31]). The generality of the m.d. assumption is often reduced by the necessity to restrict the behavior of the second-order conditional moments by the condition

$$E\left(e_t^2 \mid \Im_{t-1}\right) = \sigma^2,\ t = 1, 2, \ldots.$$

Owing to the Law of Iterated Expectations (LIE) this condition implies

$$E\left(e_t^2\right) = \sigma^2,\ t = 1, 2, \ldots.$$

So, the results presented in this work hold for $\{e_t\}_{t \in Z}$ martingale differences but, for simplicity, we assume that

Assumption A. The innovations $\{e_t\}$, $t \in Z$, are independent identically distributed (i.i.d.), satisfy

$$Ee_t = 0,\ \sigma_e^2 \equiv Ee_t^2 < \infty,\ Ee_t^4 < \infty \text{ for any } t \in Z$$

and the constants $\{c_i\}$ satisfy

$$\alpha_c := \sum_{i \in Z} |c_i| < \infty \text{ (short-memory).}$$

It follows that $\{v_t\}_{t \in Z}$ are identically distributed. The method presented here is flexible in modeling $\{w_n\}$ but is limited to short-memory processes. We also consider quadratic forms of type

$$Q_n(k_n) = v' k_n v, \tag{1.6}$$

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 [27] established approximations for quadratic forms of dependent random variables and obtained necessary and sufficient conditions for weak convergence of weighted functions of quadratic forms, Wu and Shao [28] considered asymptotic problems in spectral analysis of stationary causal processes, Bhansali et al. [29], [30] established central limit theorems for quadratic forms of causal linear processes with long-memory. Many authors, including Tanaka, Horvath and Shao, Phillips employed properties of Brownian motion in their derivations.

All results presented in this work evolve around the L_p-approximability notion introduced in Mynbaev [16, P.308]. The general idea behind L_p-approximability is to represent sequences converging to deterministic vectors with functions of a continuous argument. It is realized as follows (for convenience here we recall the definitions). Let $1 \le p < \infty$. The interpolation operator

$$\Delta_{np} : R^n \to L_p\left(0,1\right)$$

is defined by

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

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

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

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

The Section 2 contains preliminary results like useful lemmas and theorems and assumptions which will be used in this dissertation.

In Section 3 we consider convergence in distribution of two quadratic forms arising in unit root tests for a regression with slowly varying regressor. The error term is a unit root process with linear processes as disturbances. We apply results from [31, P.348] and [1, P.122] for non-causal processes to characterize the asymptotic distribution of these quadratic forms. Obtained results generalize to the non-causal linear processes some statements from [17, P.979;18, P.172].

To present the results of Section 3 here, we introduce the following:

Assumption B. The process $\left\{u_t\right\}$ possesses a unit root under the null hypothesis $\rho = 1$ in

$$u_t = \rho u_{t-1} + v_t,$$

where $\{v_t\}$ is the same linear process as above.

Main results of Section 3 are the following:

Lemma 1 (Lemma 1) in [24, P.161]. Let

$$v_t = \sum_{j\in\mathbb{Z}} c_j e_{t-j} \ ,$$

here

$$\{e_i\} \text{ being i.i.d. with } Ee_i = 0, \ Ee_i^4 < \infty, \ \sigma^2 = Ee_i^2$$

and $\{c_j\}$ is satisfying

$$\alpha_c \equiv \sum_{i\in Z} |c_i| < \infty \ .$$

Suppose Assumption B holds. Denote

$$R_n = \frac{1}{n}\sum_{t=2}^{n}\left(\sum_{l=1}^{t-1} v_l\right) v_t \ .$$

Then

- in case $\beta_c \neq 0$ we have

$$R_n \underset{n\to\infty}{\overset{d}{\to}} \frac{\sigma^2}{2}\left(\beta_c^2 u^2 + \sum_{i\in\mathbb{Z}} c_i^2\right), \tag{1.7}$$

where u is standard normal random variable;

- in case $\beta_c = 0$ we have

$$R_n \underset{n\to\infty}{\overset{P}{\to}} \frac{\sigma^2}{2}\sum_{i\in\mathbb{Z}} c_i^2 \ . \tag{1.8}$$

Since convergence in distribution to a constant implies convergence in probability to the same constant, (1.8) is a part of (1.7).

For a related result, see Theorem 3.7 of [17, P.979].

Lemma 2 (Lemma 2) in [24, P.163]. Let

$$S_n = \frac{1}{n^2}\sum_{t=1}^{n}\left(\sum_{l=1}^{t} v_l\right)^2 ,$$

where $\{v_l\}$ is

$$v_t = \sum_{j\in\mathbb{Z}} c_j e_{t-j} \,,$$

here

$$\{e_i\} \text{ being i.i.d. with } Ee_i = 0\,,\ Ee_i^4 < \infty\,,\ \sigma^2 = Ee_i^2$$

and $\{c_j\}$ is satisfying

$$\alpha_c \equiv \sum_{i\in Z} |c_i| < \infty\,.$$

Denote

$$\beta_c = \sum_{j\in\mathbb{Z}} c_j \,.$$

Then
- in case $\beta_c \neq 0$ we have

$$S_n \xrightarrow[n\to\infty]{d} (\sigma\beta_c)^2 \sum_{k=1}^{\infty} \frac{1}{\left(\left(k-\frac{1}{2}\right)\pi\right)^2} u_k^2 \,,$$

where $\{u_i\}$ are independent standard normal random variables;
- in case $\beta_c = 0$ we have

$$S_n \xrightarrow[n\to\infty]{P} 0\,.$$

Comparing to Theorem 5.12 of [18, 172] we have the same result but under less stringent assumptions on linear processes.

The main subject of Section 4 is convergence in distribution of quadratic forms (1.6). Results concerning quadratic forms (1.6) are cast in a different format and have a different area of applicability than those from the papers cited above.

The format, which links the asymptotic distribution to integral operators, was suggested by Nabeya and Tanaka [32, P.129]. 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 [16, P.308] 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.

Main results of Section 4 are the following:

Theorem 1 (Theorem 2.1) in [22, P.1309]. Let $\{v_t\}_{t\in Z}$ satisfy Assumption A, and let

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

hold, where $K\in L_2\left((0,1)^2\right)$. If K is nuclear, then

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

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

$$\operatorname{cov}\left(u_i^{(1)},u_j^{(2)}\right)=\left(\psi_i,\phi_j\right) \text{ for all } i,\ j,$$

here functions $\{\psi_i\}$ and $\{\phi_j\}$ are from the representation of the operator K. If $\boldsymbol{K}$ is symmetric, then $u_i^{(1)}=u_i^{(2)}$ for all i.

In [26, P.34] shown analyzing variance of above CLT.

Theorem 2 (Theorem 2.2) in [22, P.1312]. Let Assumption A hold and suppose that f_n is L_2-close to F and g_n is L_2-close to G:

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

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 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

$$Q_n(k_n) = v' k_n v \xrightarrow{d} \left(\sigma_e \sum_i c_i\right)^2 \|F\|_2 \|G\|_2 u_1 u_2,$$

where u_1, u_2 are standard normal and

$$\operatorname{cov}(u_1, u_2) = \int_0^1 F_0(t) G_0(t) dt.$$

In case of CLT's for linear forms (1.1) the method developed in [33, P.748] has three advantages. Firstly, all sequences arising in the theory of regressions

$$y_t = \alpha + \beta L(t) + u_t \tag{1.9}$$

with slowly varying regressor $L(t)$ turn out to be L_p-approximable. Secondly, as is shown in [33, P.748], 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) in distribution, it is enough to establish that $\{w_n\}$ is L_p-close to some $W \in L_p(0,1)$.

It is this last fact that lets us concentrate on establishing L_p-approximability of certain sequences, which we do in Section 5. Here is an overview of the related results.

Consider a polynomial trend

$$f_n = \left(1^{k-1}, \ldots, n^{k-1}\right)$$

or a logarithmic trend

$$f_n = \left(\ln^k 1, \ldots, \ln^k n\right)$$

and normalize it to get

$$w_n = f_n \bigg/ \left(\sum_{j=1}^n \left|f_{nj}\right|^p\right)^{1/p}.$$

Then $\{w_n\}$ is L_p-approximable to

$$F(x) = \left((k-1)p + 1\right)^{1/p} x^{k-1},\ k \in N_+,$$

and

$$F \equiv 1,\ k \in N_+,$$

respectively, for $1 \le p < \infty$ (Theorem 2.7.1) in [1, P.80]. Here and below $k \ge 0$ is an integer.

Using the fact that certain spatial matrices are L_2-approximable, Mynbaev [34] 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

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

Denote

$$\varepsilon(x) = \frac{xL'(x)}{L(x)},\ G(t,n) = \frac{L(t)-L(n)}{L(n)\varepsilon(n)},\ w_{nt} = n^{-\frac{1}{p}}G^k(t,n),\ t = 1,\ldots,n. \tag{1.10}$$

Phillips pointed out the importance of function $G(t,n)$ for regression (1.9) with stable errors and established a series of its properties, among them the fact that

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

Then under some conditions $\{w_n\}$ is L_p-close to $\log^k x$ (Theorem 4.4.1) in [1, P.149].

Denote

$$\eta(x) = \frac{x\varepsilon'(x)}{\varepsilon(x)},\ \mu(x) = \frac{1}{2}\left[\varepsilon(x)+\eta(x)\right],\ H(t,n) = \frac{G(t,n)-\log\frac{t}{n}}{\mu(n)}, \tag{1.11}$$
$$w_{nt} = n^{-\frac{1}{p}}H(t,n),\ t = 1,\ldots,n.$$

Then $\{w_n\}$ is L_p-close to $\log^2 x$ (Theorem 4.4.8) in [1, P.157].

Sequences (1.10) and (1.11) appear in the theory of regression (1.9) with stationary errors $\{u_t\}$. In case of nonstationary errors, we need three more sequences:

$$F(t,n)=\frac{1}{nL(n)}\sum_{j=t}^{n}L(j),\ w_{nt}=n^{-\frac{1}{p}}F^{k}(t,n),\ t=1,...,n, \tag{1.12}$$

$$I(t,n)=\frac{1}{n}\sum_{j=t}^{n}G(j,n),\ w_{nt}=n^{-\frac{1}{p}}I^{k}(t,n),\ t=1,...,n, \tag{1.13}$$

$$J(t,n)=\frac{1}{n}\sum_{j=t}^{n}\left(L(j)-\bar{L}\right), \tag{1.14}$$

where

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

Section 5 consists of proofs (Section 3) in [22, P.1313] that (1.12) is L_p-close to

$$(1-t)^k,$$

(1.13) is L_p-close to

$$\left(t\log\frac{1}{t}-1+t\right)^k$$

and (1.14) is L_p-close to

$$\left(t\log\frac{1}{t}\right)^k.$$

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

As we said earlier, we restrict out attention to models with deterministic regressors (1.9). Most papers concentrated on quickly growing regressors, like polynomials [35-37]. Phillips [3, P.558] was the first to consider SV regressors, of the type $\log t$, $\log(\log t)$, $\frac{1}{\log t}$, etc. Models with such regressors have many applications, see [5, P.1048; 6, P.19; 7, P.1; 8, P. 501; 9, P. 1, 10, P.1772; 11, P.1156].

Consider the model (9). For the case when the errors $\{u_t\}$ are stationary, Phillips [3, P.565] obtained the asymptotic distribution of the OLS estimators $\hat{\alpha}$, $\hat{\beta}$.

We consider integrated errors (Assumption B)

$$u_t = \rho u_{t-1} + v_t,\ t = 2,...,n, \tag{1.15}$$

where $\rho = 1$ under the null hypothesis and $\{v_t\}_{t\in Z}$ is a non-causal linear process.

In Section 6 we have to find the asymptotic distribution of $\hat{\alpha}$ and $\hat{\beta}$. Uematsu [4, P.10] characterized convergence in distribution of $\hat{\alpha}$ and $\hat{\beta}$, and it turned out to be very different from what Phillips [3, P.565] had with stable errors. Section 6 consists of a proof Uematsu's result ([4, P.10]) on the asymptotic distribution of $\hat{\alpha}$ and $\hat{\beta}$ under less restrictive conditions. He relied on the Brownian motion method suggested by Phillips [3, P.565]. As illustrated in [33, P.748], the L_p-approximability approach is more powerful than the Brownian motion method in case of problems involving SV functions.

The importance of L_p-approximability is explained by the next result Mynbaev [33, P.307] which will be applied in this dissertation work multiple times:

Theorem A. Suppose $\{v_t\}$ satisfy Assumption A. If $\{w_n\}$ is L_2-close to W, then for the sums $S_n = \sum_{t=1}^{n} w_{nt} v_t$, we have

$$S_n \underset{n\to\infty}{\overset{d}{\to}} N\left(0, \sigma^2 \int_0^1 W^2(t)\,dt\right),$$

where $\sigma^2 = \sigma_e^2 \left(\sum_i c_i\right)^2$, $\sigma_e^2 = Ee_1^2$.

Main results of Section 6 are the following:

Lemma 3 (Lemma 6) in [23, P.94]. Denote by

$$\sigma^2 = \left(\sigma_e \sum_{j\in Z} c_j\right)^2.$$

Under Assumptions A, B and some assumptions on function L (Assumption 5.1 in Section 5) the following statements are true:

i) If $0 < \theta < 1$, then

$$\frac{1}{n\sqrt{n}L(n)}\sum_{t=1}^{n}L(t)u_t \xrightarrow[n\to\infty]{d} N\left(0,\frac{1}{3}\sigma^2\right).$$

ii) If $0<\theta<1$, then

$$\frac{1}{n\sqrt{n}L(n)\varepsilon(n)}\sum_{t=1}^{n}\left(L(t)-\overline{L}\right)u_t \xrightarrow[n\to\infty]{d} N\left(0,\frac{2}{27}\sigma^2\right).$$

Theorem 3 (Theorem 3) in [23, P.95]. If L satisfies Assumption A, $0<\theta<1$ and u_t satisfies Assumption B, then

$$\begin{pmatrix} \dfrac{\varepsilon(n)}{\sqrt{n}}(\hat{\alpha}-\alpha) \\ \dfrac{L(n)\varepsilon(n)}{\sqrt{n}}\left(\hat{\beta}-\beta\right) \end{pmatrix} \xrightarrow[n\to\infty]{d} N\left(0,\frac{2}{27}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\right).$$

Section 7 contains Monte-Carlo simulations [25, P.87] for OLS estimators $\hat{\alpha}$ and $\hat{\beta}$ of the above regression model and compare this study with theoretical results obtained in Section 6. Simulations were done in MatLab software packages.

2 PRELIMINARY RESULTS AND ASSUMPTIONS

In this section we introduce useful lemmas, theorems and assumptions as well as the main tools used in other sections to prove results.

2.1 Useful lemmas and theorems

Theorem 2.1. (Theorem 3.9.1) in [1, P.122]: Suppose that

(i) e_{nt}^2 are uniformly integrable and $E\left(e_{nt}^2 \mid \Im_{n,t-1}\right)=\sigma^2$ for all t and n;

(ii) the sequence $\{\psi_j\}_{j\in Z}$ is summable, i.e. $\alpha_\psi<\infty$;

(iii) the sequence $\{k_n\}_{n\in N}$ is L_2-close to some symmetric function $K\in L_2\left((0,1)^2\right)$ with the next rate of approximation

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

(iv) the integral operator K with the kernel K is nuclear.

Then we can assert that

- If $\beta_\psi\neq 0$, then the quadratic form

$$Q_n\left(k_n\right)=v_n'k_nv_n,$$

converges in distribution to

$$\left(\sigma\beta_\psi\right)^2\sum_i \lambda_i u_i^2,$$

where $\{u_i\}$ are independent standard normal and $\{\lambda_i\}$ are the eigenvalues of K.

- If $\beta_\psi=0$, then

$$Q_n\left(k_n\right)\underset{n\to\infty}{\overset{P}{\rightarrow}}0.$$

Theorem 2.2 (Theorem 1) in [31, P.348]: Assume that $H^{(r)}(x)$ exists and is continuous for all x, $E\varepsilon_1^4<\infty$. If $\left\{K,\{a_j\}_{j=-\infty}^{\infty},\{\varepsilon_j\}_{j=-\infty}^{\infty}\right\}$ satisfies $C(2,\lambda)$ for some $\lambda>0$, then

$$N^{-1/2}\sum_{n=1}^{N}\left(K\left(X_n^l\right)-EK\left(X_n^l\right)\right)\xrightarrow{d}N\left(0,\sigma^2\right),$$

where K is defined in [31] and

$$\sigma^2=\lim_{n\to\infty}N^{-1}Var\left(\sum_{n=1}^{N}\left(K\left(X_n^l\right)-EK\left(X_n^l\right)\right)\right),\ X_n^l=\sum_{j=-\infty}^{\infty}c_j^l\varepsilon_{n-j}$$

Lemma 2.1 (Lemma 3.6.1 (iii)) in [1, P.106]: To distinguish the 2-D and 1-D cases, denote δ_{np}^2 as the operator defined in 2-D and δ_{np}^1 its 1-D cousin. If

$$F\left(x,y\right)=G\left(x\right)H\left(y\right)\in L_p\left(\left(0,1\right)^2\right),$$

then

$$\left(\delta_{np}^2F\right)_{st}=\left(\delta_{np}^1G\right)_s\left(\delta_{np}^1H\right)_t \text{ for all } s,t.$$

Lemma 2.2 (Lemma 2.3.2) in [1, P.55]: Let $\left\{\psi_j\right\}_{j\in Z}$ be a summable sequence of real numbers, T_n is a T-operator. If $\alpha_\psi<\infty$ and $1\le p\le\infty$, then

$$\sup_n\left\|T_n\right\|\le\alpha_\psi \text{ and } \sup_n\max\left\{\left\|T_n^+\right\|,\left\|T_n^-\right\|,\left\|T_n^0\right\|\right\}\le\alpha_\psi.$$

Lemma 2.3 (Lemma 2.1.3 (ii)) in [1, P.46]: If $F\in L_p$, $1\le p\le\infty$, then

$$\left\|\delta_{np}F\right\|_p\le\left\|F\right\|_p.$$

Theorem 2.3 [16, P.322]: Let $\{e_{nt},\Im_{nt}\}$ be a m.d. array and let W_n be a sequence of $n\times L$ matrices with columns $w_n^1,...,w_n^L$. Suppose that

i) $\left\{e_{nt}^2\right\}$ are uniformly integrable and $E\left(e_{nt}^2\mid\Im_{n,t-1}\right)=\sigma^2$ for all t,n.

ii) the sequence $\left\{w_n^l:n\in N\right\}$ is L_2-close to $F_l\in L_2$, $l=1,...,L$.

Then

$$\lim_{n\to\infty}V\left(W_n'e_n\right)=\sigma^2G$$

and

$$W_n' e_n \underset{n\to\infty}{\overset{d}{\to}} N\left(0,\sigma^2 G\right),$$

where G is the Gram matrix of $F_1,...,F_L$.

Theorem 2.4 (Theorem 3.5.2) in [1, P.103]): Let $\{e_{nt},\Im_{nt}\}$ be a double infinite m.d. array and let W_n be a sequence of $n\times L$ matrices with columns $w_n^1,...,w_n^L$. Suppose that

i) $\left\{e_{nt}^2\right\}$ are uniformly integrable and $E\left(e_{nt}^2 \mid \Im_{n,t-1}\right)=\sigma^2$ for all t,n.

ii) the sequence $\left\{w_n^l : n\in N\right\}$ is L_2-close to $F_l\in L_2$, $l=1,...,L$, and

iii) $\alpha_\psi<\infty$.

With the same W_n and G as in Theorem 2.3, the following statements are true:

a) If $\beta_\psi\neq 0$, then

$$W_n' v_n \underset{n\to\infty}{\overset{d}{\to}} N\left(0,\left(\sigma\beta_\psi\right)^2 G\right).$$

b) If $\beta_\psi=0$, then

$$W_n' v_n \underset{n\to\infty}{\overset{P}{\to}} 0.$$

In both cases

$$\lim_{n\to\infty} V\left(W_n' v_n\right)=\left(\sigma\beta_\psi\right)^2 G.$$

Lemma 2.4 (Lemma 2.5.2 (i)) in [1, P.69]: Let $\{f_n\}$ be L_p-approximable. Then

$$\sup_n \| f_n\|_p<\infty.$$

Theorem 2.5 [38]: Let L be defined on $[A,\infty)$, $A>0$. Then L is SV if and only if there exist a number $B\geq A$ and functions μ,ε on $[B,\infty)$ with properties:

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

ii) μ is bounded, measurable and the limit $c=\lim_{x\to\infty}\mu(x)$ exists ($c\in R$);

iii) ε is continuous on $[B,\infty)$ and $\lim_{x\to\infty}\varepsilon(x)=0$.

Corollary 2.1 (Corollary 4.4.3) in [1, P.152]: If $L=K(\varepsilon,\phi_\varepsilon)$ and $\theta<1$, then

$$\frac{1}{nL^k(n)}\sum_{t=1}^{n}L^k(t)=1-k\varepsilon(n)\left[1+o(1)\right].$$

Corollary 2.2 (Corollary 4.4.2) in [1, P.151]: If $L=K(\varepsilon,\phi_\varepsilon)$ and $\theta k<1$, then

$$\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}G^k(t,n)=(-1)^k k!.$$

Lemma 2.5 (Lemma 2.7.2) in [1, P.80]: Let $p\le\infty$. Let F be continuous on $[0,1]$ and suppose that a sequence $\{p_n\}$, with $p_n\in R^n$ for all n, satisfies

$$\max_{1\le t\le n}\left|p_{nt}-F\left(\frac{t}{n}\right)\right|\to 0,\ n\to\infty.$$

Denote $f_n=n^{-1/p}p_n$. Then $\{f_n\}$ is L_p-close to F.

Lemma 2.6 (Lemma 7.3) in [3, P.587]: If $L=K(\varepsilon)$, $\varepsilon=K(\eta)$, $\eta=K(\mu)$, and $\eta(n)=o(\varepsilon(n))$, then

$$\frac{1}{n}\sum_{t=1}^{n}\left(L(t)-\bar{L}\right)^2=L^2(n)\varepsilon^2(n)\left(1+o(1)\right).$$

Lemma 2.7 (Lemma 4.6.4) in [1, P.178]: If $\{v_n\}$ is L_p-close to V, $\{w_n\}$ is L_p-close to W and numerical sequences $\{a_n\}$ and $\{b_n\}$ converge to a and b, respectively, then $\{a_nv_n+b_nw_n\}$ is L_p-close to $aV+bW$.

2.2 Assumptions

Assumption 1. Let $\{e_t,F_t:t\in Z\}$ be a double-infinite m.d. array. Fixing a summable sequence of numbers $\{c_j:j\in Z\}$, denote

$$v_t=\sum_{j\in\mathbb{Z}}c_je_{t-j},\ t\in Z.$$

The array $\{v_t:t\in Z\}$ is called a linear process.

Assumption 2. $\{e_t^2\}_{t\in Z}$ are uniformly integrable and

$$E\left(e_t^2\,|F_{t-1}\right)=\sigma_e^2 \text{ for all } t.$$

Assumption 3. The sequence $\{c_j,\ j\in Z\}$ is summable, i.e. $\alpha_c<\infty$.

Assumption 4. The process $\{u_t\}$ possesses a unit root under the null hypothesis $\rho=1$ in

$$u_t=\rho u_{t-1}+v_t,\ t=1,...,n,$$

where $\{v_t\}$ is the same linear process as in Assumption 1.

Assumption 5 (this assumption implies Assumptions 1 and 2). The innovations $\{e_t\}_{t\in Z}$, are independent identically distributed (i.i.d.), satisfy

$$Ee_t=0,\ \sigma_e^2\equiv Ee_t^2<\infty,\ Ee_t^4<\infty \text{ for any } t\in Z.$$

3 CONVERGENCE OF SOME QUADRATIC FORMS USED IN REGRESSION ANALYSIS

The goal of this section is find an asymptotic distribution of two quadratic forms that arise in statistical applications.

Consider the linear processes $\{v_n\}$ satisfying Assumptions 1, 2 and $\{u_n\}$ satisfying Assumption 4. Denote

$$R_n = \frac{1}{n}\sum_{t=2}^{n} u_{t-1}v_t = \frac{1}{n}\sum_{t=2}^{n}\left(\sum_{l=1}^{t-1} v_l\right)v_t \tag{3.1}$$

and

$$S_n = \frac{1}{n^2}\sum_{t=1}^{n} u_t^2 = \frac{1}{n^2}\sum_{t=1}^{n}\left(\sum_{l=1}^{t} v_l\right)^2,\ n\in N \tag{3.2}$$

with

$$e_t' = \left(\underbrace{1,\ldots,1,}_{t}\ \ \underbrace{0,\ldots,0}_{n-t}\right),\ d_t' = \left(\underbrace{0,\ldots,0}_{t-1},1,\underbrace{0,\ldots,0}_{n-t}\right),\ v' = (v_1,\ldots,v_n)$$

we can write

$$\sum_{l=1}^{s} v_l = e_s'v,\ v_s = d_s'v \text{ for } s = 1,\ldots,n. \tag{3.3}$$

So for (3.1) by (3.3) we obtain

$$R_n = \frac{1}{n}\sum_{t=2}^{n} e_{t-1}'vd_{t-1}'v = v'\frac{1}{n}\sum_{t=2}^{n} e_{t-1}d_t'v = v'a_nv,$$

where a_n is an upper triangular matrix of size $n\times n$: $a_n = \frac{1}{n}\begin{pmatrix} 0 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 0 & 1 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$.

One can show that a_n is L_2-close to

$$F(s,t)=\begin{cases}1, & s<t\\ 0, & s\ge t\end{cases},$$

but the degree of approximation and the integral operator with F as a kernel are not good enough to apply Theorem 2.1 for convergence in distribution of R_n. This is because a_n is not a symmetric matrix. We note that

$$R_n=\left(v'a_nv\right)'=v'a_n'v,$$

so we can write

$$R_n=\frac{1}{2}v'\left(a_n+a_n'\right)v=\frac{1}{2}v'\left(a_n+a_n'+\frac{I}{n}\right)v-\frac{1}{2n}v'v,$$

here I is the identity matrix of size $n\times n$. Denoting

$$k_n=a_n+a_n'+\frac{I}{n}$$

we have

$$R_n=\frac{1}{2}v'k_nv-\frac{1}{2n}v'v. \qquad (3.4)$$

Let $K(s,t)\equiv 1$ on $(0,1)^2$. Then

$$\left(k_n\right)_{ij}-\left(\delta_{n2}K\right)_{ij}=\frac{1}{n}-n\int_{q_{ij}}dx=\frac{1}{n}-\frac{1}{n}=0 \text{ for all } 1\le i,j\le n,$$

here δ_{n2} is a discretization operator. Thus,

$$\left\|k_n-\delta_{n2}K\right\|_2=0.$$

The integral operator $\mathcal{K}$ is associated with $\boldsymbol{K}$

$$(\mathcal{K}f)(x)=\int_0^1 f(t)dt$$

has only one non-zero eigenvalue $\lambda=1$; the corresponding eigenspace consists of constants and is one-dimensional, all other eigenvalues are zero. By Theorem 2.1 we can assert that

1. If $\beta_c \neq 0$, then

$$v'k_n v \underset{n\to\infty}{\overset{d}{\rightarrow}} (\sigma\beta_c)^2 u^2, \tag{3.5}$$

where u is a standard normal, or

2. If $\beta_c = 0$, then

$$v'k_n v \underset{n\to\infty}{\overset{P}{\rightarrow}} 0. \tag{3.6}$$

To deal with the second term in (3.4), we apply a CLT from [31, P.348]. Suppose a function H has r-th derivative and let

$$H_\lambda^{(r)}(x) = \sup_{|y|\le\lambda} \left| H^{(r)}(x+y) \right|, \ \lambda > 0.$$

Definition 3.1. The triplet $\{H,\{c_j\},\{e_j\}\}$ is said to satisfy condition $C(r,\lambda)$ if there exists a number $\lambda \in (0,\infty)$ such that

1. $H^{(r)}(x)$ exists and is continuous for all $x \in R$,
2. For all $x \in R$

$$\sup_{I\subset\mathbb{Z}} E\left[H_\lambda^{(r)}\left(x + \sum_{i\in I} c_i e_i \right) \right]^4 < \infty \tag{3.7}$$

where supremum is taken over all subsets $I \subset Z$.

Let $\{e_i\}$ be independent identically distributed (i.i.d.) and satisfies Assumption 5. Put $H(x) = x^2$. Then (3.7) is trivially satisfied with $r=2$. By Theorem 1

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} (v_t^2 - Ev_t^2) \underset{n\to\infty}{\overset{d}{\rightarrow}} N\left(0,\sigma_1^2\right), \tag{3.8}$$

where

$$\sigma_1^2 = \lim_{n\to\infty} \frac{1}{n} Var\left[\sum_{t=1}^{n} v_t^2 \right] = \sigma^2 \sum_{i\in\mathbb{Z}} c_i^4 .$$

From (3.8) we have

$$\frac{1}{n}v'v=\frac{1}{n}\sum_{t=1}^{n}v_t^2=\frac{1}{n}\sum_{t=1}^{n}\left(v_t^2-Ev_t^2\right)+\frac{1}{n}\sum_{t=1}^{n}Ev_t^2$$

$$=\frac{1}{\sqrt{n}}\left[\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(v_t^2-Ev_t^2\right)\right]+\frac{1}{n}\sum_{t=1}^{n}Ev_t^2=o_P(1)+\frac{1}{n}\sum_{t=1}^{n}Ev_t^2 . \tag{3.9}$$

Since

$$Ev_t^2=E\sum_{i,j\in\mathbb{Z}}c_ic_je_{t-i}e_{t-j}=\sigma_e^2\sum_{i\in\mathbb{Z}}c_i^2 ,$$

(3.9) implies

$$\lim_{n\to\infty}\frac{1}{n}v'v\xrightarrow[n\to\infty]{P}\sigma^2\sum_{i\in\mathbb{Z}}c_i^2 . \tag{3.10}$$

From (3.4), (3.5), (3.6) and (3.10) we obtain the following statement.

Lemma 3.1. Let

$$v_t=\sum_{j\in\mathbb{Z}}c_je_{t-j} ,$$

here $\{e_i\}_{i\in Z}$ being i.i.d. with $Ee_i=0$, $Ee_i^4<\infty$, $\sigma_e^2=Ee_i^2$ and $\{c_j\}_{j\in Z}$ is summable sequence of numbers. Suppose Assumption 4 holds. Denote

$$R_n=\frac{1}{n}\sum_{t=2}^{n}u_{t-1}v_t .$$

Then

- in case $\beta_c\neq 0$ we have

$$R_n\xrightarrow[n\to\infty]{d}\frac{\sigma^2}{2}\left(\beta_c^2u^2+\sum_{i\in\mathbb{Z}}c_i^2\right), \tag{3.11}$$

where u is standard normal random variable;

- in case $\beta_c=0$ we have

$$R_n \xrightarrow[n\to\infty]{P} \frac{\sigma^2}{2}\sum_{i\in\mathbb{Z}} c_i^2 . \tag{3.12}$$

Since convergence in distribution to a constant implies convergence in probability to the same constant, (3.12) is a part of (3.11).

For a related result, see Theorem 3.7 of [17, P.979].

For (3.2) by using (3.3) we have

$$S_n = \frac{1}{n^2}\sum_{t=1}^{n} e_t' v e_t' v = v' \frac{1}{n^2}\sum_{t=1}^{n} e_t e_t' v = v' b_n v , \tag{3.13}$$

where b_n is a symmetric matrix of size $n\times n$:

$$b_n = \frac{1}{n}\begin{pmatrix} 1 & 1-\frac{1}{n} & 1-\frac{2}{n} & \cdots & 1-\frac{n-2}{n} & 1-\frac{n-1}{n} \\ 1-\frac{1}{n} & 1-\frac{1}{n} & 1-\frac{2}{n} & \cdots & 1-\frac{n-2}{n} & 1-\frac{n-1}{n} \\ 1-\frac{2}{n} & 1-\frac{2}{n} & 1-\frac{2}{n} & \cdots & 1-\frac{n-2}{n} & 1-\frac{n-1}{n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1-\frac{n-2}{n} & 1-\frac{n-2}{n} & 1-\frac{n-2}{n} & \cdots & 1-\frac{n-2}{n} & 1-\frac{n-1}{n} \\ 1-\frac{n-1}{n} & 1-\frac{n-1}{n} & 1-\frac{n-1}{n} & \cdots & 1-\frac{n-1}{n} & 1-\frac{n-1}{n} \end{pmatrix}.$$

In other words,

$$b_n = \left(\frac{1}{n}\left(1-\max\left\{\frac{i-1}{n},\frac{j-1}{n}\right\}\right)\right)_{i,j=1}^{n} .$$

For given $s,t\in(0,1)^2$ choose $1\le i, j\le n$ such that

$$\frac{i-1}{n} < s < \frac{i}{n},\ \frac{j-1}{n} < t < \frac{j}{n} .$$

It is easy to see that b_n is L_2-close to a symmetric function

$$B(s,t) = 1-\max\{s,t\} .$$

We need to know the rate of approximation. For this consider

$$\left|b_{n_{ij}}-(\delta_{n2}B)_{ij}\right|=\left|\frac{1}{n}\left(1-\max\left\{\frac{i-1}{n},\frac{j-1}{n}\right\}\right)-n\int_{q_{ij}}\left(1-\max\{s,t\}\right)dsdt\right|$$

$$=\left|\int_{q_{ij}}\left[n^2\cdot\frac{1}{n}\left(1-\max\left\{\frac{i-1}{n},\frac{j-1}{n}\right\}\right)-n\left(1-\max\{s,t\}\right)\right]dsdt\right|$$

$$\le\int_{q_{ij}}\left|n\left(\max\{s,t\}-\max\left\{\frac{i-1}{n},\frac{j-1}{n}\right\}\right)\right|dsdt\le\frac{1}{n^2}.$$

Thus,

$$\left\|b_n-\delta_{n2}B\right\|_2^2=\sum_{i,j=1}^{n}\left|b_{n_{i,j}}-\left(\delta_{n2}B\right)_{ij}\right|^2=o\left(\frac{1}{n^2}\right).$$

The integral operator K is associated with B in the following way

$$(Kf)(t)=\int_0^1 B(x,t)\,f(x)dx$$

and let us consider the integral equation for λ and $f(t)$

$$f(t)=\lambda\int_0^1 B(x,t)\,f(x)dx,$$

here λ is the eigenvalue of the kernel $B(s,t)$; the corresponding solution $f(t)$ is an eigenfunction for the eigenvalue λ. According to [18, P139] eigenvalues of kernel $B(s,t)$ are

$$\lambda_k=\frac{1}{\left(\left(k-\frac{1}{2}\right)\pi\right)^2},\ k\in N,$$

with multiplicity 1 and corresponding eigenvectors are

$$f_k(t) = c\cos\sqrt{\lambda_k t} \text{ with } c \neq 0.$$

By Theorem 2.1 we can assert that
1. If $\beta_c \neq 0$, then

$$v' b_n v \xrightarrow[n\to\infty]{d} (\sigma\beta_c)^2 \sum_{i=1}^{\infty} \lambda_i u_i^2, \tag{3.14}$$

where $\{u_i\}_{i\in N_+}$ are independent standard normal random variables, or
2. If $\beta_c = 0$, then

$$v' b_n v \xrightarrow[n\to\infty]{P} 0. \tag{3.15}$$

From (3.13), (3.14) and (3.15) we obtain the next statement.

Lemma 3.2. Let

$$S_n = \frac{1}{n^2} \sum_{t=1}^{n} \left(\sum_{l=1}^{t} v_l \right)^2,$$

where $\{v_l\}_{l\in Z}$ as in Assumption 1. And let Assumption 2 hold. Denote

$$\beta_c = \sum_{j\in\mathbb{Z}} c_j .$$

Then
- in case $\beta_c \neq 0$ we have

$$S_n \xrightarrow[n\to\infty]{d} (\sigma\beta_c)^2 \sum_{k=1}^{\infty} \frac{1}{\left(\left(k - \frac{1}{2}\right)\pi\right)^2} u_k^2,$$

where $\{u_i\}$ are independent standard normal random variables;
- in case $\beta_c = 0$ we have

$$S_n \xrightarrow[n\to\infty]{P} 0.$$

Comparing to Theorem 5.12 of [18, P.172] we have the same result but under less stringent assumptions on linear processes.

4 CENTRAL LIMIT THEOREMS FOR LINEAR AND QUADRATIC FORMS

The main subject of this section is convergence in distribution of quadratic forms (1.6). For this let us recall some definitions from Definitions section and we need some facts from the theory of operators in Hilbert spaces (all of them can be found in [2]). Let A be a compact linear operator in a Hilbert space with a scalar product $(\cdot,\cdot)$. The operator $H=\left(A^*A\right)^{\frac{1}{2}}$ is called the modulus of A, here A^* is the adjoint operator of A. If $A=A^*$, then we say that operator A is selfadjoint. The eigenvalues of H, denoted s_i, $i=1,2,\ldots$, and counted with their multiplicity, are called s-numbers of A. U denotes a partially isometric operator that isometrically maps the range $R\left(A^*\right)$ 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 $\{\phi_j\}$ be an orthonormal system of eigenvectors of H which is complete in $R(H)$. Then, we have the representation

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

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

$$Ax=\sum_{i=1}^{r(A)} s_i\left(x,\phi_i\right)\psi_i, \tag{4.1}$$

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, $\{\phi_i\}$ are eigenvectors of A and

$$s_i=|\lambda_i|,$$

where $\{\lambda_i\}$ are eigenvalues of A.

Let us apply (4.1) to an integral operator

$$\left(Kf\right)(s)=\int_0^1 K\left(s,t\right)f\left(t\right)dt,\ f\in L_2(0,1),$$

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

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

Because f is arbitrary, we have the decomposition

$$K(s,t) = \sum_{i=1}^{r(A)} s_i\psi_i(s)\phi_i(t), \tag{4.2}$$

where s_i and ϕ_i are, respectively, the eigenvalues and eigenvectors of $\left(K^*K\right)^{\frac{1}{2}}$ and

$$\psi_j = U\phi_j.$$

The fundamental idea of Nabeya and Tanaka [32] was to postulate that the matrices k_n in (1.6) approach in some sense a function $\boldsymbol{K}$ on $(0,1)^2$ and express the limit properties of $Q_n(k_n)$ in terms of the properties of the associated integral operator K.

Let us recall definition of nuclear operator.

Definition 4.1. The operator K is called **nuclear** if

$$\sum s_i < \infty \; (\sum|\lambda_i| < \infty \text{ when } K \text{ is selfadjoint}).$$

Nabeya and Tanaka [32, P.219] required $\boldsymbol{K}$ to be continuous and symmetric and K to be nuclear. Mynbaev [16, P.307] used L_p-approximability, which allowed him to relax the continuity assumption and replace i.i.d. $\{v_t\}_{t\in Z}$ with linear processes. Here we develop his approach further by lifting the symmetry condition.

Definition 4.2. 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

$$\left(\delta_{np}K\right)_{ij}=n^{2\left(1-\frac{1}{p}\right)}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\int_{\frac{j-1}{n}}^{\frac{j}{n}}K(s,t)\,dsdt,\ 1\le i,j\le n. \tag{4.3}$$

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

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

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 impose a stronger condition on the rate of approximation. Mynbaev [1, P.122] 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

$$\left\|k_n-\delta_{n2}K\right\|_2=o\left(\frac{1}{n}\right) \tag{4.4}$$

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

$$\left\|k_n-\delta_{n2}K\right\|_2=o\left(\frac{1}{\sqrt{n}}\right) \tag{4.5}$$

is less restrictive. For simplicity, we adhere to Assumptions 3 and 5, which allows us to use (4.5), remembering that in cases (3.4) and (3.5) Mynbaev's conditions on $\{v_t\}_{t\in Z}$ from Theorems 3.9.1 and 3.9.7 can be repeated word for word.

Theorem 4.1. Let $\{v_t\}_{t\in Z}$ satisfy Assumption 5 and let Assumption 3 and (4.5) hold. If 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 $\{u_i^{(1)}\}$, $\{u_i^{(2)}\}$ are systems of independent (within a system) standard normals, $\{s_i\}$ are s -numbers of K and

$$\operatorname{cov}\left(u_i^{(1)},u_j^{(2)}\right)=\left(\psi_i,\phi_j\right)\text{ for all } i,\ j.$$

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

Proof. We can exclude symmetric K covered in Mynbaev [1, P.126]. The proof is similar to that of Theorem 3.9.7 (all references are to Mynbaev [1]), so we indicate only the modifications. (4.2) above is analogous to equation (3.38), which holds in the symmetric case. Hence, the initial segment of (4.2) is

$$K_L(s,t) = \sum_{i=1}^{L} s_i \psi_i(s) \varphi_i(t).$$

Subtracting from (4.2) its initial segment and applying Lemma 2.1 we get

$$\left(\delta_{n2}^2 K - \delta_{n2}^2 K_L\right)_{s,t} = \sum_{i>L} s_i \left(\delta_{n2}^1 \psi_i\right)_s \left(\delta_{n2}^1 \phi_i\right)_t, \tag{4.6}$$

where $\delta_{n2}^2 = \delta_{n2}$ is the two-dimensional discretization operator defined in (4.3) and δ_{np}^1 is its one-dimensional version defined by

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

Combining (1.6) and (4.6) we have

$$Q_n\left(\delta_{n2}^2 K\right) - Q_n\left(\delta_{n2}^2 K_L\right) = \sum_{i>L} s_i \sum_{s,t=1}^{n} \left(\delta_{n2}^1 \psi_i\right)_s v_s \left(\delta_{n2}^1 \phi_i\right)_s v_t$$

$$= \sum_{i>L} s_i \left[\left(\delta_{n2}^1 \psi_i\right)' v\right]\left[\left(\delta_{n2}^1 \phi_i\right)' v\right]'. \tag{4.7}$$

By Section 3.3.5 about the *T*-decomposition for means of quadratic forms

$$\left| E\left(\left[\left(\delta_{n2}^1 \psi_i\right)' v\right]\left[\left(\delta_{n2}^1 \phi_i\right)' v\right]'\right)\right| =$$

$$= \sigma_e^2 \left| \left(T_n^0 \delta_{n2}^1 \psi_i, T_n^0 \delta_{n2}^1 \phi_i\right) + \left(T_n^- \delta_{n2}^1 \psi_i, T_n^- \delta_{n2}^1 \phi_i\right) + \left(T_n^+ \delta_{n2}^1 \psi_i, T_n^+ \delta_{n2}^1 \phi_i\right) \right|$$

(applying the Cauchy-Schwarz inequality)

$$\le \sigma_e^2\left[\left\|T_n^0\delta_{n2}^1\psi_i\right\|_2\left\|T_n^0\delta_{n2}^1\phi_i\right\|_2+\left\|T_n^-\delta_{n2}^1\psi_i\right\|_2\left\|T_n^-\delta_{n2}^1\phi_i\right\|_2\right.$$

$$\left.+\left\|T_n^+\delta_{n2}^1\psi_i\right\|_2\left\|T_n^+\delta_{n2}^1\phi_i\right\|_2\right]$$

(using boundedness of the operators T_n^0, T_n^-, T_n^+, see Lemma 2.2)

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

(using boundedness of the operators δ_{n2}^1, see Lemma 2.3)

$$\le 3\left(\sigma_e\alpha_c\right)^2\left\|\phi_i\right\|_2\left\|\psi_i\right\|_2=3\left(\sigma\alpha_c\right)^2. \tag{4.8}$$

By nuclearity of K from (4.7) - (4.8) we have

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

The conclusion is the same as in Section 3.9.3:

$$\underset{L\to\infty}{p\lim}\left[Q_n\left(\delta_{n2}^2K\right)-Q_n\left(\delta_{n2}^2K_L\right)\right]=0 \text{ uniformly in } n.$$

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 2.4 we have

$$W_n'v\xrightarrow{d}N\left(0,\left(\sigma_e\sum_i c_i\right)^2G_L\right),\ n\to\infty, \tag{4.9}$$

where

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

Since both systems $\{\phi_i\}$, $\{\psi_i\}$ are orthonormal, this can be written as

$$G_L = \begin{pmatrix} I & H_L \\ H_L' & I \end{pmatrix},$$

where the identities are of size $L \times L$ and H_L has elements $\left(\phi_i, \psi_j\right)$. It follows that (4.9) is equivalent to

$$W_n'v \xrightarrow[n\to\infty]{d} \left|\sigma_e \sum_i c_i\right| \begin{pmatrix} u^{(1)} \\ u^{(2)} \end{pmatrix}, \tag{4.10}$$

where $u^{(1)}$, $u^{(2)}$ are standard normal vectors and

$$\operatorname{cov}\left(u^{(1)}, u^{(2)}\right) = H_L.$$

Similarly to equation (4.7),

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

This is a continuous function of the vector at the left of (4.10). By the continuous mapping theorem then

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

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 K is not symmetric. The proof of the generalization of Theorem 3.9.7 is complete. #

Remark. We use this opportunity to comment on the relationship between Theorem 2.4 used in the above proof and a classical result [39]. The conditions on the stochastic and deterministic parts of the process in this work are more general than in Theorem 3.5.2. However, there is no analysis of the limit of the normalizing sequence $\{\sigma_n\}$. In our results, it is the main selling point, without which the statement on the covariance structure in Theorem 4.1 would be impossible. On a related note, in [40] there is analysis of the normalizing sequence but the focus is different: the variance of the limit distribution is tied to the spectral density.

Recall the discussion about rates of approximation (4.4), (4.5). 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 hold? The answer contained in the next theorem means that it is true when essentially the two-dimensional case can be reduced to the one-dimensional.

Theorem 4.2. Let Assumptions 1 and 3 hold and suppose that f_n is L_2-close to F and g_n is L_2-close to G:

$$\|f_n - \delta_{n2}F\|_2 \to 0,\ \|g_n - \delta_{n2}G\|_2 \to 0. \tag{4.11}$$

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 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

$$Q_n(k_n) = v'k_n v \xrightarrow{d} \left(\sigma_e \sum_i c_i\right)^2 \|F\|_2 \|G\|_2 u_1 u_2, \tag{4.12}$$

where u_1, u_2 are standard normal and

$$\operatorname{cov}(u_1, u_2) = \int_0^1 F_0(t) G_0(t)\, dt.$$

Proof. In the proof of Theorem 4.1 we showed how to deal with the fact that K is not symmetric. Here we show how to lift the restriction (4.5). By Lemma 2.1

$$\left(\delta_{n2}^{2}K\right)_{st}=\left(\delta_{n2}^{1}F\right)_{s}\left(\delta_{n2}^{1}G\right)_{t}.$$

For an $n\times n$ matrix A denote

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

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

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

$$\leq g\left(\left(f_n-\delta_{n2}^{1}F\right)g_n'\right)+g\left(\left(\delta_{n2}^{1}F\right)\left(g_n-\delta_{n2}^{1}G\right)'\right). \tag{4.13}$$

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

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

which is just another way of writing

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

$$g\left(\left(\delta_{n2}^{1}F\right)\left(g_n-\delta_{n2}^{1}G\right)'\right)\leq c\left\|\delta_{n2}^{1}F\right\|_2\left\|g_n-\delta_{n2}^{1}G\right\|_2. \tag{4.14}$$

By Lemma 4.3 and Lemma 2.4

$$\sup_n\|g_n\|<\infty\ \text{and}\ \sup_n\left\|\delta_{n2}^{1}F\right\|<\infty,$$

so (4.11), (4.13), (4.14) imply

$$g\left(k_n - \delta_{n2}^2 K\right) \to 0 .$$

This gives Equation (3.50) in [1, P.123]. The rest of the proof of convergence in distribution is the same.

We need to justify the format of the limit distribution. The operators K and K^* are given by

$$(K f)(s) = \int K(s,t)\, f(t)\,dt = F(s)\int G(t)\, f(t)\,dt = F(s)(G, f) \tag{4.15}$$

and

$$\left(K^* g\right)(u) = \int K(s,u)\, g(s)\,ds = G(u)\int F(s)\, g(s)\,ds = G(u)(F,g). \tag{4.16}$$

Hence,

$$\left(K^* K f\right)(u) = G(u)(G, f)\|F\|_2^2 .$$

If f is an eigenvector of K^*K, it should be proportional to G, i.e.

$$f = cG,$$

and from the above

$$K^* K f = \lambda f$$

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,

$$(G, f) = 0.$$

(4.15) - (4.16) show that

$$K^* K f = 0$$

on all such functions. Hence, $s_j = 0$ for $j > 1$. From (4.15) - (4.16) we see that the range $R\left(K^*\right)$ is spanned by

$$G_0 = G \big/ \|G\|_2$$

and the range $R(K)$ is spanned by

$$F_0 = F \big/ \|F\|_2 .$$

The required partially isometric operator obtains by setting

$$UG_0 = F_0 .$$

Thus, (4.12) follows from Theorem 4.1, where

$$s_1 = \|G\|_2 \|F\|_2 , \; s_j = 0 \text{ for } j > 1,$$

$$u^{(1)}, \; u^{(2)} \text{ are standard normal}$$

and

$$\operatorname{cov}\left(u^{(1)}, u^{(2)}\right) = \left(\phi_1, \psi_1\right) = \left(F_0, G_0\right).$$

#

5 SLOW VARIATION AND L_p-APPROXIMABILITY

The main goal of this section is prove L_p-approximations of some new sequences introduced in Overview.

First, let us recall the definition of SV functions and consider examples.

Definition 5.1. A positive measurable function on $[A,\infty)$, $A>0$, is slowly varying (SV) if

$$\lim_{x\to\infty}\frac{L(rx)}{L(x)}=1 \text{ for any } r>0. \tag{5.1}$$

Examples of such functions are $L_1(x)=\log x$ and $L_2(x)=\log(\log x)$ because

$$\frac{L_1(rx)}{L_1(x)}=\frac{\log r+\log x}{\log x}\underset{x\to\infty}{\to}1,$$

$$\frac{L_2(rx)}{L_2(x)}=\frac{\log(\log r+\log x)}{\log(\log x)}=\frac{\log(\log x)+\log(1+(\log r/\log x))}{\log(\log x)}\underset{x\to\infty}{\to}1.$$

Similarly, $L_3(x)=1/\log x$ and $L_4(x)=1/\log(\log x)$ are SV. The function $L_5(x)=x^a$, $a\neq 0$, is not SV because

$$(rx)^a/x^a=r^a$$

does not tend to 1 unless $r=1$.

The seemingly innocuous condition (5.1) in fact entails many strong properties. We shall be using, often without explicitly mentioning, the following standard properties of SV functions [38]:

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 (5.1) as 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 the above Karamata theorem there exist a number $B\geq A>0$ and functions μ, ε on $[B,\infty)$ such that

$$L(x)=\exp\left(\mu(x)+\int_B^x\varepsilon(t)\frac{dt}{t}\right), \tag{5.2}$$

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

Following Phillips [3, P.560], we make a simplifying assumption that

$$\mu = const.$$

Phillips argues that asymptotically this does not affect regression estimation because the asymptotic behavior of representation in Theorem 2.5 (i) is equivalent to that with $\mu = const$. 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 if μ is continuously differentiable, then

$$L(x) = \exp\left(\mu(x) + \int_B^x \varepsilon(t)\frac{dt}{t}\right) = \exp\left(\int_B^x \mu'(t)dt + \mu(B) + \int_B^x \varepsilon(t)\frac{dt}{t}\right)$$

$$= \exp\left(\mu(B) + \int_B^x \left(t\mu'(t) + \varepsilon(t)\right)\frac{dt}{t}\right).$$

So, if additionally

$$\lim_{x\to\infty} x\mu'(x) = 0,$$

we have a new representation of the same function L with a constant μ. Thus, (5.2) can be equivalently written as

$$L(x) = c_L \exp\left(\int_B^x \varepsilon(t)\frac{dt}{t}\right) \tag{5.3}$$

with a new continuous function ε on $[B,\infty)$ such that $\lim_{x\to\infty}\varepsilon(x)=0$ [1, P.133]. When (5.3) holds, we write $L = K(\varepsilon)$ in this case, omitting the constant c_L from the notation. The function in this representation is called an ε-function of L.

Further, ε can be extended to the segment $[0,B]$ in such a way that the integral $\int_0^B \frac{\varepsilon(t)}{t}dt$ exists (e.g., one can set ε equal to 0 in the neighborhood of 0 and interpolate continuously between that neighborhood and $[B,\infty)$). This will amount to redefining L on $[0,B]$, which does not matter asymptotically because only a finite number of regression equations are affected. In any case, L can be considered continuous and positive on $[0,\infty)$. From now on we assume that such adjustments are made.

For some expansions we need to assume that $|\varepsilon(x)|$ is also SV and ε has Karamata representation

$$\varepsilon(x) = c_\varepsilon \exp\left(\int_B^x \eta(t)\frac{dt}{t}\right) \text{ for } x \geq B$$

for some (possibly negative) constant c_ε, where η is continuous and $\eta(x) \to 0$ as $x \to \infty$. In such cases we also write $\varepsilon = K(\eta)$, remembering that ε can be negative. The number of different conditions in this theory may be daunting. To reduce it, in some cases we assume a little more than is required by a puristic approach.

So, we understand that it is convenient to assume that L is continuous and does not vanish on $[0,\infty)$, which can be achieved by properly extending the function ε on $[0,B)$.

For $L = K(\varepsilon)$,

$$\varepsilon(x) = \frac{xL'(x)}{L(x)} \to 0 \text{ as } x \to \infty$$

Using this formula we calculate and collect in Table 1 below expressions for ε and η in the sequence $L = K(\varepsilon), \varepsilon = K(\eta)$, (the role of the function

$$\mu(x) = \frac{1}{2}(\varepsilon(x) + \eta(x))).$$

is disclosed in Mynbaev's book (Section 4.2.7) in [1, P.142]. In Table 5.1 we denote $l_1(x) = \log(x)$, $l_2(x) = \log(\log(x))$ and assume $\gamma > 0$ it is the (Table 4.1) in [1, P.134]. The table contains the functions of most practical interest against which the plausibility of new assumptions should be checked.

Expressions arising in regression statistics involve values $L(t)$ for $1 \leq t \leq n$. For a fixed $\delta \in [0,1)$, the values $L(t)$ with $\delta n \leq t \leq n$ 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 non-decreasing and $\lim_{x\to\infty}\phi(x) = \infty$,

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

L is called SV with remainder ϕ if for any $r > 0$ instead of (5.1) one has

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

Table 5.1 – Basic SV functions

L	ε	η	μ	$L\varepsilon\mu$
$L_1 = l_1^\gamma$	$\frac{\gamma}{l_1}$	$-\frac{1}{l_1}$	$\frac{1}{2}\frac{\gamma-1}{l_1}$	$\frac{\gamma(\gamma-1)}{2}l_1^{\gamma-2}$
$L_2 = l_2$	$\frac{1}{l_1 l_2}$	$-\frac{1+l_2}{l_1 l_2}$	$-\frac{1}{2l_1}$	$-\frac{1}{2l_1^2}$
$L_3 = \frac{1}{l_1}$	$-\frac{1}{l_1}$	$-\frac{1}{l_1}$	$-\frac{1}{l_1}$	$\frac{1}{l_1^3}$
$L_4 = \frac{1}{l_2}$	$-\frac{1}{l_1 l_2}$	$-\frac{1+l_2}{l_1 l_2}$	$-\frac{2+l_2}{2l_1 l_2}$	$\frac{2+l_2}{2l_1^3 l_2^3}$

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

Lemma 5.1 [38, P.102]: If L is SV with remainder ϕ, then for any $b > \theta$ there exist constants $M_b > 0$ and $B_b > B$ such that

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

Assumption 5.1 (on SV function L). a) $L = K(\varepsilon)$, that is, (5.2) holds, with ε described after (5.3).

b) ε is SV in the general sense (5.1).

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

$$\frac{1}{c\phi_\varepsilon(x)} \le |\varepsilon(x)| \le \frac{1}{\phi_\varepsilon(x)} \text{ for all } x \ge c. \tag{5.4}$$

We write $L = K(\varepsilon, \phi_\varepsilon)$ to mean that L satisfies Assumption 5.1. Note that all practically important SV functions from Table 5.1 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 arbitrary close to zero. For our final results on L_p-approximability, on top of Assumption 5.1 we shall have to impose more conditions, and all of them hold for functions from Table 5.1.

Let us consider the function F defined in (1.12). Let $[a]$ denote the integer part of $a \in R$. Now we can proceed with our new results contained in the next lemmas and theorems:

Lemma 5.2. 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

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

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

$$\frac{\delta}{2} \le \delta - \frac{\alpha+1}{n} \le \frac{[nr]-1}{n} \le \frac{1}{\delta} - \frac{\alpha+1}{n} \le \frac{1}{\delta}. \tag{5.5}$$

By Corollary 2.1

$$\frac{1}{nL(n)}\sum_{t=1}^{n}L(t)=1-\varepsilon(n)\left[1+o(1)\right], \tag{5.6}$$

so

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

$$=\left(1-\varepsilon(n)\left[1+o(1)\right]\right)-\frac{\left(\left[rn\right]-1\right)L\left(\left[rn\right]-1\right)}{nL(n)}\left(1-\varepsilon\left(\left[rn\right]-1\right)\left[1+o(1)\right]\right). \tag{5.7}$$

According to the definition of ε we can continue (5.7) and have

$$F\left(\left[rn\right],n\right)=\left(1+o(1)\right)-\frac{\left(\left[rn\right]-1\right)L\left(\left[rn\right]-1\right)}{nL(n)}\left(1+o(1)\right). \tag{5.8}$$

The $o(1)$ here is uniform in r because by (5.5) $\left[rn\right]-1\geq n\frac{\delta}{2}$. By the uniform convergence theorem (5.7) also implies

$$L\left(\frac{\left[rn\right]-1}{n}\cdot n\right)\Big/L(n)=1+o(1).$$

Hence, continuing (5.8)

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

uniformly in r.

To prove b) consider two cases.

Case 1. $\left(B_b+1\right)/n\leq r\leq\delta$, where B_b is the constant from Lemma 5.1. Obviously,

$$\left|F\left(\left[rn\right],n\right)\right|\leq\left|\sum_{t=1}^{n}\frac{L(t)}{nL(n)}\right|+\left|\sum_{t=1}^{B_b}\frac{L(t)}{nL(n)}\right|+\left|\sum_{t=B_b+1}^{[rn]-1}\frac{L(t)}{nL(n)}\right|. \tag{5.9}$$

By (5.6), the first term at the right is $1+o(1)$. L is continuous and bounded on $[0,B_b]$, so

$$\left|\sum_{t=1}^{B_b}\frac{L(t)}{nL(n)}\right|\le\frac{cB_b}{nL(n)}\to 0,\ n\to\infty. \tag{5.10}$$

The first 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 5.1

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

$$\le\frac{M_b}{n\phi(n)}\sum_{t=B_b+1}^{[rn]-1}\left(\frac{t}{n}\right)^{-b}+\frac{1}{n}([rn]-B_b-1). \tag{5.11}$$

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 $0<a<N$

$$\sum_{t=a-1}^{N}t^{-b}\le\int_a^N t^{-b}dt\le\int_0^N t^{-b}dt. \tag{5.12}$$

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 (5.11) and get

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

$$= \frac{c_1}{\phi(n)}\left(\frac{[rn]}{n} - \frac{1}{n}\right)^{1-b} + 1 \le c_1 \frac{r^{1-b}}{\phi(n)} + 1 \le c_2. \tag{5.13}$$

(5.9), (5.10) and (5.13) 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 (5.9) is empty; the rest of the proof does not change. #

Theorem 5.1. For $p \in [1,\infty)$ and integer $k \ge 0$ define a vector $w_n \in R^n$ by (1.12). 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 [1, P.149]: definition of interpolation operator is equivalent to

$$(\Delta_{np} w)(u) = n^{1/p} w_{[nu+1]},\ 0 \le u < 1. \tag{5.14}$$

Therefore in our case by using (5.13) we obtain

$$(\Delta_{np} w)(u) = F^k([nu+1], n),\ 0 \le u < 1.$$

Let $0 < \delta \le \dfrac{1}{2}$, $\delta \le u < 1$. 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,

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

This and Lemma 5.2 (a) imply

$$F([nu+1],n) = F([rn],n) = 1 - u + o(1),\ n \to \infty, \tag{5.15}$$

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 5.2 (b)

$$\left|F([nu+1],n)\right| = \left|F([rn],n)\right| \le c \text{ for } u \in (0,\delta). \tag{5.16}$$

Obviously,

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

$$+\left(\int_0^\delta \left|(1-u)^k\right|^p du\right)^{1/p} + \left(\int_0^\delta \left|F^k([nu+1],n)\right|^p du\right)^{1/p}.$$

By (5.15) - (5.16) 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 (1.13).

Lemma 5.3. 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)

$$\left|I([rn],n)\right|\le c$$

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

$$1\ge t\ge[rn]/n\ge[\delta n]/n\ge\delta-1/n\ge\delta/2$$

for a large n. By (1.10)

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

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

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

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

By Stirling's formula [41, 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,

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

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

$$=(1+o(1))\left[\left(\frac{s}{n}-\frac{1}{2n}\right)\log\left(\frac{s-1}{n}\right)^{-1}+\frac{s-1}{n}-1+\frac{\theta(n)}{12n^2}-\frac{\theta(s-1)}{12n^2\,\frac{s-1}{n}}\right]. \tag{5.17}$$

Since

$$rn-1\le s=[rn]\le rn,$$

we have

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

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

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

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

This and (5.17) prove part a).

b) First consider the case

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

and write

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

From

$$rn-1 \le \left[rn\right] \le t \le n$$

we have

$$\frac{B_b}{n} \le r-\frac{1}{n} \le \frac{t}{n} \le 1,$$

so by Lemma 5.1 and (5.12)

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

$$= \frac{M_b n^{b-1}}{\left|\varepsilon(n)\right|\phi(n)}\sum_{t=[rn]}^{n} t^{-b} \le \frac{M_b}{1-b}\cdot\frac{1}{\left|\varepsilon(n)\right|\phi(n)} \le C. \tag{5.18}$$

The last bound holds by (5.4).

Now let

$$0 < r < \frac{B_b+1}{n}.$$

Then for

$$t \le \left[rn\right]-1$$

we have

$$t \le rn \le B_b+1$$

and $\left|L(t)\right| \le c_1$. By Corollary 2.2, we have

$$\frac{1}{n}\sum_{t=1}^{n} G(t,n) \to 1,$$

so

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

$$\le c_2 + \frac{1}{n\left|\varepsilon(n)\right|}\sum_{t=1}^{[rn]-1}\left(\frac{\left|L(t)\right|}{L(n)}+1\right) \le c_2 + \frac{[rn]-1}{n\left|\varepsilon(n)\right|}\left(\frac{c_1}{L(n)}+1\right) \le c_3. \qquad (5.19)$$

This is because $\left|\varepsilon(n)\right|$, $\left|\varepsilon(n)L(n)\right|$ are SV and

$$n\left|\varepsilon(n)\right| \to \infty,\ n\left|\varepsilon(n)L(n)\right| \to \infty.$$

(5.18) and (5.19) prove b). #

Theorem 5.2. For $p \in [1,\infty)$ and integer $k \ge 0$ define a vector $w_n \in R^n$ by (1.13). 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 5.1, just replace Lemma 5.2 with Lemma 5.3. #

Now consider the function J defined in (1.14).

Lemma 5.4. 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)

$$\left|J([rn],n)\right| \le c \text{ for } r \in (0,\delta],$$

where c does not depend on δ.
Proof. Obviously,

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

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

Use here (5.6) 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 5.3a))

$$=(1+o(1))\left(r\log\frac{1}{r}-1+r\right)+(1+o(1))(1-r+o(1))$$

$$=(1+o(1))r\log\frac{1}{r}.$$

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 5.3 instead of part a). #

Theorem 5.3. For $p\in[1,\infty)$ and integer $k\ge 0$ define a vector $w_n\in R^n$ by (1.14). 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 5.2 in the proof of Theorem 5.1 with Lemma 5.4.

6 ASYMPTOTIC DISTRIBUTION OF OLS ESTIMATORS

The goal of this section is to prove asymptotic distribution of OLS estimators in (1.9).

For the simple regression model

$$y_t = \alpha + \beta L(t) + u_t,\ t = 1,...,n,$$

one version of the formulas for the OLS estimates $\hat{\alpha}$ and $\hat{\beta}$ is (Section 3.1) [42]

$$\hat{\beta} - \beta = \sum_{t=1}^{n} \left(L(t) - \bar{L}\right) u_t \left[\sum_{t=1}^{n} \left(L(t) - \bar{L}\right)^2\right]^{-1} \tag{6.1}$$

and

$$\hat{\alpha} - \alpha = \bar{u} - \bar{L}\left(\hat{\beta} - \beta\right), \tag{6.2}$$

where

$$\bar{u} = \frac{1}{n}\sum_{t=1}^{n} u_t,\ \bar{L} = \frac{1}{n}\sum_{t=1}^{n} L(t).$$

are the averages.

The main result of this section is:

Theorem 6.1. If Assumptions 1, 5 and 5.1 hold and, additionally, $\varepsilon = K(\eta)$, $\eta = K(\mu)$, $\eta(n) = o\left(\varepsilon(n)\right)$, then

$$\begin{pmatrix} \dfrac{\varepsilon(n)}{\sqrt{n}}(\hat{\alpha} - \alpha) \\ \dfrac{L(n)\varepsilon(n)}{\sqrt{n}}\left(\hat{\beta} - \beta\right) \end{pmatrix} \xrightarrow[n \to \infty]{d} N\left(0, \frac{2\sigma^2}{27}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\right).$$

We will prove this theorem later. First we start studying of convergence in distribution of

$$z_{1n} = \frac{\bar{u}}{n},\ z_{2n} = \frac{L(n)\varepsilon(n)}{\sqrt{n}}\left(\hat{\beta} - \beta\right)$$

which are part of $\hat{\alpha}$, $\hat{\beta}$.

Lemma 6.1. a) $z_{1n} \underset{n\to\infty}{\overset{d}{\to}} N\left(0, \frac{\sigma^2}{3}\right)$.

b) If $\varepsilon = K(\eta)$, $\eta = K(\mu)$ and $\eta(n) = o(\varepsilon(n))$, then z_{2n} admits the representation

$$z_{2n} = \frac{1+o(1)}{n\sqrt{n}L(n)\varepsilon(n)} \sum_{t=1}^{n} \left(L(t) - \bar{L}\right)u_t$$
$$= \frac{1}{n\sqrt{n}} \sum_{t=1}^{n} \left(G(t,n)+1\right)u_t + o_P(1) \tag{6.3}$$

and

$$z_{2n} \underset{n\to\infty}{\overset{d}{\to}} N\left(0, \frac{2\sigma^2}{27}\right). \tag{6.4}$$

Proof. a) We have

$$z_{1n} = \frac{1}{n\sqrt{n}} \sum_{t=1}^{n} u_t = \frac{1}{n\sqrt{n}} \sum_{t=1}^{n} \sum_{l=1}^{t} v_l = \frac{1}{n\sqrt{n}} \sum_{l=1}^{n} v_l \sum_{t=l}^{n} 1$$
$$= \frac{1}{n\sqrt{n}} \sum_{t=1}^{n} v_l (n-l+1) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(1 - \frac{t-1}{n}\right) v_t .$$

By Lemma 2.5 the sequence

$$w_n = \frac{1}{\sqrt{n}}\left(1, 1-\frac{1}{n}, \dots, 1-\frac{n-1}{n}\right) \text{ is } L_2\text{-close to } F(x) = 1-x. \tag{6.5}$$

Since $\int_0^1 F^2(x)dx = \frac{1}{3}$, statement a) follows from Theorem A.

b) By Lemma 2.6

$$\frac{1}{n} \sum_{t=1}^{n} \left(L(t) - \bar{L}\right)^2 = L^2(n)\varepsilon^2(n)\left(1+o(1)\right).$$

Combining this with (6.1), we get

$$z_{2n} = \frac{1+o(1)}{n\sqrt{n}L(n)\varepsilon(n)}\sum_{t=1}^{n}\left(L(t)-\bar{L}\right)u_t .$$

Adding and subtracting $L(n)$ and using Corollary 2.3, we obtain

$$z_{2n} = \frac{1+o(1)}{n\sqrt{n}}\sum_{t=1}^{n}\frac{L(t)-L(n)}{L(n)\varepsilon(n)}u_t + \frac{1+o(1)}{n\sqrt{n}}\sum_{t=1}^{n}\frac{L(t)-\bar{L}}{L(n)\varepsilon(n)}u_t$$

$$= \frac{1+o(1)}{n\sqrt{n}}\sum_{t=1}^{n}G(t,n)u_t + \frac{1+o(1)}{n\sqrt{n}}\sum_{t=1}^{n}u_t$$

$$= \frac{1}{n\sqrt{n}}\sum_{t=1}^{n}\left(G(t,n)+1\right)u_t + \frac{o(1)}{n\sqrt{n}}\sum_{t=1}^{n}\left(G(t,n)+1\right)u_t . \tag{6.6}$$

Changing the order of summation,

$$\frac{1}{n\sqrt{n}}\sum_{t=1}^{n}\left(G(t,n)+1\right)u_t = \frac{1}{n\sqrt{n}}\sum_{l=1}^{n}v_l\sum_{t=1}^{n}\left(G(t,n)+1\right)$$

$$= \sum_{l=1}^{n}v_l\left[\frac{1}{\sqrt{n}}\left(I(l,n)+1-\frac{l-1}{n}\right)\right]. \tag{6.7}$$

The sequence $\frac{1}{\sqrt{n}}\left(I(1,n),...,I(n,n)\right)$ is L_2-close to $t\log\frac{1}{t}-1+t$ (Theorem 5.2), so the sum of this sequence and the one in (6.5) by Lemma 2.7 L_2-close to $W(t)=t\log\frac{1}{t}$. Since $\int_0^1 W^2(t)dt = \frac{2}{27}$, by Theorem A the variable in (6.7) converges in distribution to $N\left(0,\frac{2\sigma^2}{27}\right)$. Now (6.3) and (6.4) follow from (6.6). #

Proof of Theorem 6.1. We know from Lemma 6.1 (b) that

$$\frac{L(n)\varepsilon(n)}{\sqrt{n}}\left(\hat{\beta}-\beta\right)\xrightarrow[n\to\infty]{d} N\left(0,\frac{2\sigma^2}{27}\right).$$

By (6.2) and Corollary 2.3

$$\frac{\varepsilon(n)}{\sqrt{n}}(\hat{\alpha}-\alpha)=\frac{\varepsilon(n)}{\sqrt{n}}\bar{u}-\frac{\varepsilon(n)\bar{L}}{\sqrt{n}}(\hat{\beta}-\beta)$$

$$=\varepsilon(n)z_{1n}-(1+o(1))\frac{\varepsilon(n)\bar{L}}{\sqrt{n}}(\hat{\beta}-\beta)$$

$$=\varepsilon(n)z_{1n}-(1+o(1))z_{2n}.$$

By Lemma 6.1 this equals to $-z_{2n}+o_P(1)$. This proves the theorem. #

7 MONTE-CARLO STUDY FOR OLS ESTIMATORS FOR REGRESSION WITH SLOWLY VARYING REGRESSOR

In this section we will present some simulations in MatLab.

We implement $\frac{\varepsilon(n)}{\sqrt{n}}(\hat{\alpha}_n-\alpha)$, $\frac{L(n)\varepsilon(n)}{\sqrt{n}}(\hat{\beta}_n-\beta)$ for four types of slowly varying functions $L(t)=\log(t)$, $L(t)=\log(\log(t))$, $L(t)=\frac{1}{\log(t)}$, $L(t)=\frac{1}{\log(\log(t))}$ and two types of coefficients $c_i=\frac{1}{i^2}$ and $c_i=\frac{1}{(i+i^2)^2}$. The fit quickly improves as the number of iterations and the number of observations increases. These are shown in Figures 1-4 for $a(n)=\frac{\varepsilon(n)}{\sqrt{n}}(\hat{\alpha}-\alpha)$, $b(n)=\frac{L(n)\varepsilon(n)}{\sqrt{n}}(\hat{\beta}-\beta)$, where $L(t)=\log(\log(t))$, numbers of iterations are equal to 1000 and 10000, the number of observations is equal to 300. Based on Theorem 5.1, we expect this to work for all types of slowly varying functions and any absolutely convergent coefficients. As you can see in these figures, $a(n)$, $b(n)$ data come from normal distribution, against the alternative hypothesis that the cumulative distribution function of the data is not from the normal distribution, MatLab program returned value of $h=0$ indicates that test fails to reject the null hypothesis at significance level 5%. The cdf of $b(n)$ is shown in Figure 5.

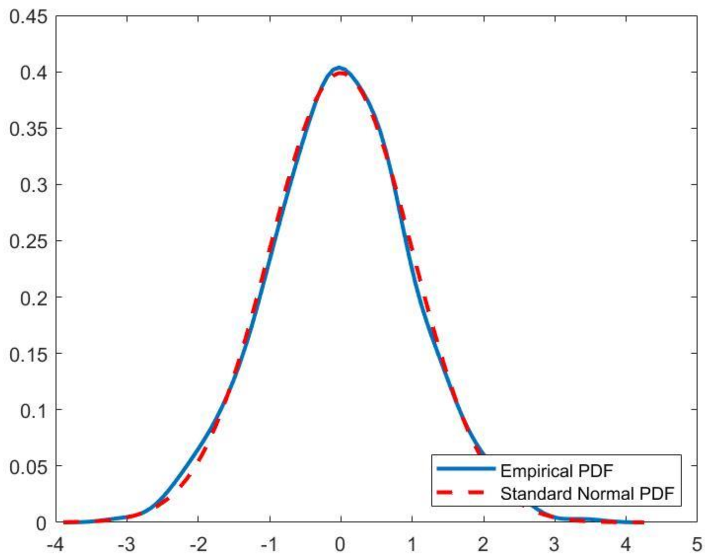

Figure 1 – Empirical pdf of $a(300)$ with 1000 number of iterations

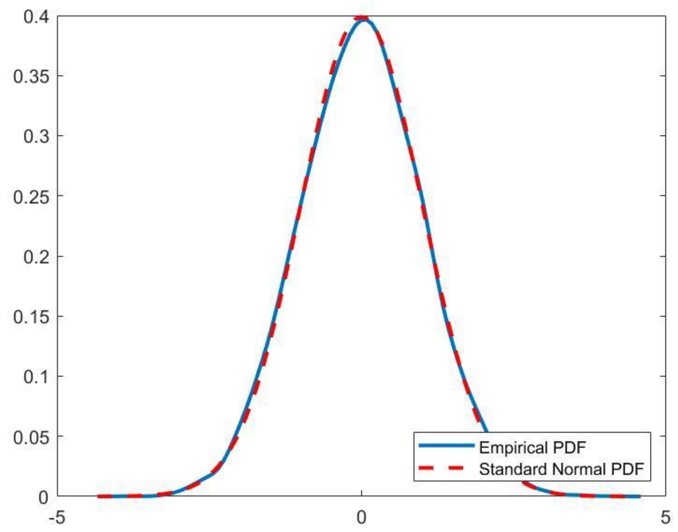

Figure 2 – Empirical pdf of $a(300)$ with 10000 number of iterations

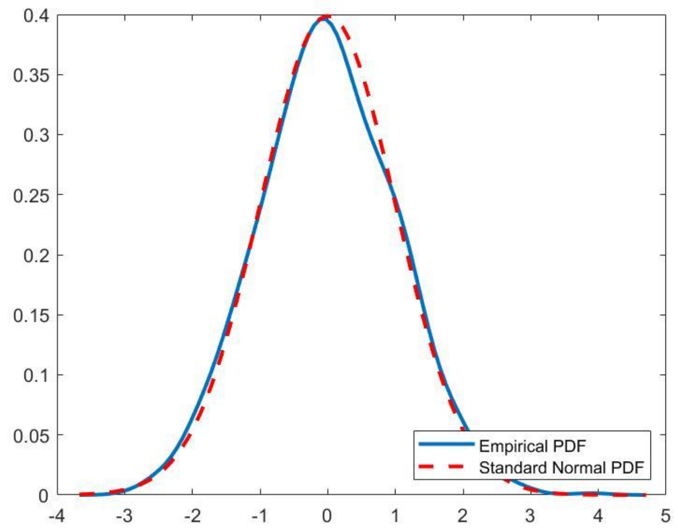

Figure 3 – Empirical pdf of $b(300)$ with 1000 number of iterations

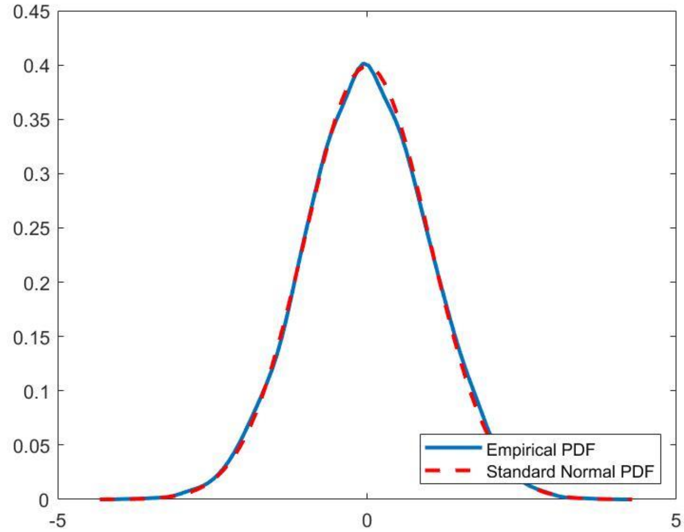

Figure 4 – Empirical pdf of $b(300)$ with 10000 number of iterations

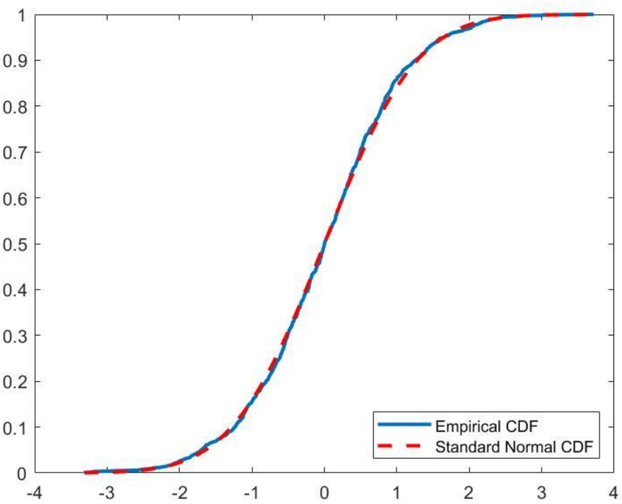

Figure 5 – Empirical pdf of $b(300)$ with 10000 number of iterations

CONCLUSION

The dissertation considers:

1) central limit theorems for quadratic forms of linear processes;

2) a couple of new L_p-approximable sequences;

3) a model

$$y_t = \alpha + \beta L(t) + u_t,\ t = 1,...,n,$$

with a slowly varying (SV) regressor, integrated errors ($u_t = \rho u_{t-1} + v_t,\ t = 2,...,n$) under the unit root and $\{v_t\}$ is a non-causal linear process, i.e. $v_t = \sum_{i \in Z} c_i e_{t-i}$.

Assessment of the completeness of the aims of the work. All results are new and based on our own methods and tools. In this work we have:

1) obtained convergence of some quadratic forms used in regression analysis.

2) obtained central limit theorems for linear and quadratic forms.

3) added a couple of sequences to the list of L_p-approximable sequences contained in Mynbaev [16, P.322].

4) proved Uematsu's result [4, P.10] on the asymptotic distribution of OLS estimations $\hat{\alpha}$ and $\hat{\beta}$ under less restrictive conditions.

5) done Monte-Carlo simulations for the asymptotic distribution of OLS estimations $\hat{\alpha}$ and $\hat{\beta}$.

Therefore, the work objectives were completed.

Suggestions on applications of the obtained results. The results obtained in this area can be used during the study of unit root test statistics, the problem of early detection of bubbles, and other statistical and econometrical problems.

Assessment of scientific level of the work in comparison with the achievements in the scientific direction. The results obtained are on par with the best achievements of foreign colleagues and contribute to the study of statistics and econometrics.

REFERENCES

1 Mynbaev K. Short-Memory Linear Processes and Econometric Applications. – Hoboken, New Jersey: Wiley and Sons, 2011. – 429 p.

2 Gohberg I., Krein M. Introduction to the theory of Linear Nonselfadjoint Operators in Hilbert Space: Translations of Mathematical Monographs. – Providence, Rhode Island: American Mathematical Society, 1969. – Vol. 18. – 378 p.

3 Phillips P.C.B. Regression with slowly varying regressors and nonlinear trends // Econometric Theory. – 2007. – Vol. 23 (4). – P. 557–614.

4 Uematsu Y. Regression with a Slowly Varying Regressor in the Presence of a Unit Root // Global COE Hi-Stat Discussion Paper Series gd11-209, Institute of Economic Research, Hitotsubashi University. – 2011. – 29 p.

5 Robinson P.M. Log-periodogram regression of time series with long range dependence // Annals of Statistics. – 1995. – Vol. 23 – P. 1048-1072.

6 Hurvich C.M., Deo R., Brodsky J. The mean squared error of Geweke and Porter-Hudakís estimator of the memory parameter of a long-memory time series // Journal of Time Series Analysis. – 1998. – Vol.19 – P.19-46.

7 Phillips P.C.B. Discrete Fourier Transforms of Fractional Processes. – Yale University, mimeographed. – 1999. – 59 p.

8 Wu C.F. Asymptotic theory of nonlinear least squares estimation // Annals of Statistics. – 1981. – Vol. 9 – P. 501-513.

9 Barro R., Sala-i-Martin X. Economic Growth. – Cambridge, Massachusetts: The MIT Press. – 2004. – 160 p.

10 Phillips P.C.B., Sul D. Transition modeling and econometric convergence tests // Econometrica. – 2007. -Vol. 75(6) – P. 1771-1855.

11 Phillips P.C.B., Sul D. Economic transition and growth // Journal of Applied Econometrics – 2009. – Vol. 27 – P. 1153-1185.

12 Phillips P.C.B., Shi Shu-Ping, Yu J. Testing for Multiple Bubbles: Limit Theory of Real Time Detectors // International Economic Review. – 2015. – Vol. 56 – P. 1079-1134.

13 Phillips Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P 500 // International Economic Review. – 2015. – Vol. 56 – P. 1043-1078.

14 Phillips P.C.B. Exploring the Mysteries of Trends and Bubbles. // Australia’s Economy and its International Context: The Joseph Fisher Lectures. - Adelaide: University of Adelaide Press, 2012. – Vol. II (1956–2012), Ch. 54. – P. 599–616.

15 Phillips P.C.B., Yu J. Dating the Timeline of Financial Bubbles during the Subprime Crisis // Quantitative Economics. – 2010. – Vol. 1 – P. 455-491.

16 Mynbaev K. L_p-approximable sequences of vectors and limit distribution of quadratic forms of random variables // Advances in Applied Mathematics – 2001. – Vol. 26 (4). – P. 302–329.

17 Phillips P.C.B., Solo V. Asymptotics for linear processes // The Annals of Statistics. – 1992. – Vol. 20(2) – P. 971-1001.

18 Tanaka K. Time Series Analysis: Nonstationary and Noninvertible Distribution Theory. 1st ed. – Wiley and Sons, 1996. -623 p.

19 Darkenbayeva G. Регрессий с медленно меняющимися регрессорами // Тезисы докладов международной конференции «Актуальные проблемы математики и математического моделирования», посвященный 50-летию создания Института математики и механики. – Алматы, 2015. – С.24.

20 Darkenbayeva G. Convergence of some quadratic forms used in regression analysis // Second International Conference on Statistical Distributions and Applications ICOSDA. – Niagara Falls, Canada, 2016. – P. 56.

21 Darkenbayeva G. Сходимость некоторых видов квадратичных форм, возникающих в регрессионном анализе // Тезисы докладов XIII Международной конференции студентов, магистрантов и молодых ученых «Ломоносов - 2017». – Астана, 2017. – С. 9.

22 Mynbaev K., Darkenbayeva G. Weak convergence of linear and quadratic forms and related statements on L_p-approximability // Journal of Mathematical Analysis and Applications. – 2019. – Vol. 473 – P.1305-1319.

23 Mynbaev K., Darkenbayeva G. The asymptotic distribution of the OLS estimator of the regression with slowly varying regressor // Математический журнал. – 2015. - Т. 15, №2. – С. 80-98.

24 Mynbaev K.,Darkenbayeva G. Convergence of some quadratic forms used in regression analysis // Математический журнал. – 2016. - Т. 16, №3. – С. 156-165.

25 Mynbaev K., Darkenbayeva G. Monte-Carlo study for OLS estimators for regression with slowly varying regressor // Математический журнал. – 2017. - Т. 17, №3. – С. 84-90.

26 Mynbaev K., Darkenbayeva G. Analyzing variance in central limit theorems // Kazakh Mathematical Journal. – 2019. - V. 19, №3. – P. 30-39.

27 Horvath L., Shao Q.-M. Limit theorems for quadratic forms with applications to Whittle’s estimate // Ann. Appl. Probab. – 1999. – Vol. 9 (1) – P. 146-187.

28 Wu W., Shao X. Asymptotic spectral theory for nonlinear time series // Ann. Statist. – 2007. – Vol. 35 (4) – P. 1773–1801.

29 Bhansali R., Giraitis L., Kokoszka P. Convergence of quadratic forms with nonvanishing diagonal // Statist. Probab. Lett. – 2007. – Vol. 77(7) – P. 726–734.

30 Bhansali R., Giraitis L., Kokoszka P. Approximations and limit theory for quadratic forms of linear processes // Stochastic Process. Appl. – 2007. – Vol. 117 (1) – P. 71–95.

31 Cheng T.-L., Ho H.-Ch. Asymptotic normality for non-linear functional of non-causal linear processes with summable weights // Journal of Theoretical Prob. – 2005. – Vol 18(2) – P. 345-358.

32 Nabeya S., Tanaka K. Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative // Ann. Statist. – 1988. – Vol. 16 (1) – P. 218–235.

33 Mynbaev K. Central limit theorems for weighted sums of linear processes: -approximability versus Brownian motion // Econometric Theory. – 2009. – Vol 25(3) – P.748-763.

34 Mynbaev K., Ullah A. Asymptotic distribution of the OLS estimator for a purely autoregressive spatial model // J. Multivariate Anal. – 2008. – Vol. 99(2) – P. 245-277.

35 Canjels E., Watson M.W. Estimating Deterministic Trends In The Presence Of Serially Correlated Errors // The Review of Economics and Statistics, MIT Press. – 1997. – Vol. 79 – P. 184-200.

36 Phillips P.C.B. and Perron P. Testing for a Unit Root in Time Series Regression // Biometrica. – 1988. - Vol. 75 – P.335-346.

37 Vogelsang T. Trend Function Hypothesis Testing in the Presence of Serial Correlation // Econometrica. – 1998. – Vol. 66 – P. 123-148.

38 Seneta E. Regularly Varying Functions. – Moscow: Nauka, 1985.

39 Peligrad M., Utev S. Central limit theorem for linear processes // Ann Probab. – 1997. – Vol. 25 (2) – P. 443-456.

40 Peligrad M., Utev S. Central Limit theorem for stationary linear processes // Ann. Probab. – 2006. – Vol. 34 (4) – P. 1608-1622.

41 Fihtenholz G. Differential and Integral Calculus 2. – Moscow: Nauka, 1968.

42 Theil H. Principles of Econometrics. – New York: Wiley, 1971. – Vol. 1. – 736 p.