Non-central limit theorems for non-linear functionals of vector valued Gaussian stationary random sequences.

Major Péter

2018. május 16.

The problem we are interested in

Let $X(n) = (X_1(n), \dots, X_d(n))$, $n = \dots, -1, 0, 1, \dots$ be a stationary sequence of d-dimensional Gaussian random vectors. Let us have a function $H(x_1, \dots, x_d)$, define with their help the random variables

$$Y_n = H(X_1(n), \dots, X_d(n)), \quad n = \dots, -1, 0, 1, \dots$$

and their normalized partial sums

$$S_N = \frac{1}{A_N} \sum_{n=0}^{N-1} Y_n, \quad N = 1, 2, \dots,$$

with an appropriate norming constant A_N. Prove a new type of limit theorem for S_N (with a non-Gaussian limit) under appropriate conditions on the stationary sequence of d-dimensional random vectors and function $H(\cdot)$.

Precedents

With R. L. Dobrushin we have proved the following result.

Theorem. *Let us fix some integer $k \geq 1$. Let $X(n)$, $n = 0, \pm 1, \pm 2, \dots$, be a stationary Gaussian sequence, $EX(0) = 0$, $EX^2(0) = 1$, such that its covariance function $r(n) = EX(0)X(n) = n^{-\alpha}L(n)$, $n = 0, 1, 2, \dots$, with some $0 < \alpha < \frac{1}{k}$, $L(\cdot)$ is a slowly varying function at infinity. Let $H(x) = H_k(x)$ be the k-th Hermite polynomial, and take the normalized random sums*

$$S_N = \frac{1}{N^{1-k\alpha/2} L(N)^{k/2}} \sum_{n=0}^{N-1} H_k(X_n), \quad N = 1, 2, \dots.$$

The random variables S_N converge in distribution to a random variable S_0 which can be defined as a k-fold Wiener–Itô integral

$$S_0 = \int \frac{e^{i(x_1 + \cdots + x_k)} - 1}{i(x_1 + \cdots + x_k)} Z_G(\,dx_1) \dots Z_G(\,dx_k).$$

Here Z_G is a random spectral measure corresponding to the spectral measure $G(\,dx) = C|x|^{-(1-\alpha)}\,dx$ on R with an appropriate constant $C > 0$.
A similar result holds if $H(x) = \sum_{l=k}^{\infty} c_l H_l(x)$.

A. M. Arcones claimed to prove a similar result for vector vector valued stationary Gaussian sequences. But his proof is problematic.

The main problem with his proof.
Our proof was based on a thoroughly worked out theory of stationary Gaussian sequences. The theory of vector valued stationary sequences is not worked out. The discussion of Arcones is based on non-proved results and non-defined notions.

The basic step in proving the vector valued generalization of our result is to work out the theory of vector valued stationary sequences.

Let us have a stationary Gaussian random sequence $X(n)$, $n=0,\pm1,\pm2,\dots$, $EX(n)=0$. The covariance function $r(n)=EX(0)X(n)$, $n=0,1,2,\dots$ determines its distribution.
Problem in the scalar valued case: Characterize the sequences $r(n)$, $n=0,1,2,\dots$ which can be the covariance function of a stationary Gaussian sequence.
The multivariate version of this problem: Let $X(n)=(X_1(n),\dots,X_d(n))$, be a d-dimensional stationary Gaussian random sequence, $EX_j(n)=0$. Characterize the covariance function $r_{j,k}(n)=EX_j(0)X_k(n)$, $n=0,1,2,\dots$, $1\le j,k\le d$, of such a random sequence.

Answer in the scalar valued case: The function $r(n)$ can be written as

$$r(n)=\int e^{inx}G(\,dx) \quad \text{for all } n=0,\pm1,\pm2,\dots,$$

where the spectral measure $G(\cdot)$ has the following properties:

1) G is a measure on $[-\pi, \pi]$.
2) In particular, it is positive, i.e. $G(A) \geq 0$ for all sets A.
3.) G is finite, i.e. $G([-\pi, \pi]) < \infty$.
4.) G is even, i.e. $G(A) = G(-A)$ for all sets A.

The covariance function $r(n)$ determines the spectral measure G.

In the case of d-dimensional stationary random sequences there exists for all $1 \leq j, k \leq d$ a complex number valued measure $G_{j,k}$ on $[\pi, \pi]$ such that

$$r_{j,k}(n) = EX_j(0)X_k(n) = \int e^{inx} G_{j,k}(\,dx), \quad n = 0, 1, 2, \ldots.$$

Take the matrix valued measure $G = (G_{j,k})$, $1 \leq j, k \leq d$, and let us characterize it.

1.) All elements $G_{j,k}$ of G are complex number valued measures.
2.) All complex number valued measures $(G_{j,k})$ have finite total variation.
3.) The matrix $G(A) = (G_{j,k}(A))$, $1 \le j, k \le d$, is positive semidefinite for all (measurable) sets A.
4.) The matrix valued measure G is even i.e. $G(-A) = \overline{G(A)}$ for all measurable sets A.
Here overline denotes complex conjugate.

The covariance functions $r_{j,k}(\cdot)$ determine the matrix valued measure G. We call it matrix valued spectral measure.

The covariance function of a stationary sequence can be calculated as the Fourier transform of the spectral measure. Our next step is to construct such random spectral measures whose „random Fourier transforms" express the elements of the stationary random sequence.

Random spectral measures

Let G be a spectral measure. There exists a random spectral measure Z_G with the following properties:

(i) The random variables $Z_G(A)$ are complex valued, jointly Gaussian random variables. (The random variables $\mathrm{Re}\, Z_G(A)$ and $\mathrm{Im}\, Z_G(A)$ with possibly different sets A are jointly Gaussian.)
(ii) $EZ_G(A) = 0$.
(iii) $EZ_G(A)\overline{Z_G(B)} = G(A \cap B)$ for all sets A and B.
(iv) $\sum\limits_{j=1}^{n} Z_G(A_j) = Z_G\left(\bigcup\limits_{j=1}^{n} A_j\right)$ if $A_1, \dots, A_n$ are disjoint sets.
(v) $Z_G(A) = \overline{Z_G(-A)}$ for all sets A.

A natural random integral can be defined with respect to a random spectral measure, and

$$X(n) = \int e^{inx} Z_G(\,dx), \quad n = 0, 1, 2, \dots,$$

is a stationary Gaussian random sequence with spectral measure G.

A similar result holds for vector valued stationary random sequences.

Let $(G_{j,k})$, $1 \le j, k \le d$, be a matrix valued spectral measure. There exists a vector valued random spectral measure $Z_G = (Z_{G_1}, \dots, Z_{G_d})$ with the following properties.

(i) The random variables $Z_{G_j}(A)$ are complex valued, and their real and imaginary parts are jointly Gaussian, i.e. for any positive integer N and sets A_s, $1 \le s \le N$, the random variables $\operatorname{Re} Z_{G_j}(A_s)$, $\operatorname{Im} Z_{G_j}(A_s)$, $1 \le s \le N$, $1 \le j \le d$, are jointly Gaussian.
(ii) $EZ_{G_j}(A) = 0$ for all $1 \le j \le d$ and A.
(iii) $EZ_{G_j}(A)\overline{Z_{G_k}(B)} = G_{j,k}(A \cap B)$ for all $1 \le j, k \le d$ and sets A, B.
(iv) $\sum_{s=1}^{n} Z_{G_j}(A_s) = Z_{G_j}\left(\bigcup_{s=1}^{n} A_s\right)$ if $A_1, \dots, A_n$ are disjoint sets, $1 \le j \le d$.
(v) $Z_{G_j}(A) = \overline{Z_{G_j}(-A)}$ for all $1 \le j \le d$ and sets A.

A random integral by Z_{G_j} can be defined for all $1 \le j \le d$, and

$$X_j(n) = \int e^{inx} Z_{G_j}(\,dx), \quad n = 0, \pm 1, \pm 2, \dots, \;\; 1 \le j \le d,$$

is a vector valued stationary Gaussian sequence with the matrix valued spectral measure G.

Continuous time stationary random sequences can be handled similarly. But we omit this subject. We discuss instead so-called generalized stationary sequences. The reason for it is that they have useful spectral and random spectral measures.

They behave similarly to the usual spectral and random spectral measures. But they do not have to be finite. We need such more general random spectral measures, because they appear in the limit theorems we are interested in.

Generalized stationary random sequences

The idea of the definition: Take a (continuous time) stationary random sequence $X(t)$ and a nice class of test functions $\mathcal{F}$. Define the random integrals $X(\varphi) = \int X(t)\varphi(t)\,dt$ for all $\varphi \in \mathcal{F}$. If the class of functions $\mathcal{F}$ is sufficiently large, then the random field $X(\varphi)$, $\varphi \in \mathcal{F}$, characterizes the starting random sequence. A useful choice for the class of test functions is the Schwartz space $\mathcal{S}$ which consists of those functions which tend to zero at $\pm\infty$ together with all of their derivatives faster than any polynomial degree. We shall define the properties of the generalized random fields in such a form which reflects the most important properties of the above random integrals. We give this definition in the case of vector valued fields. A d-dimensional generalized random field is a random field whose elements are d-dimensional random vectors

$$(X_1(\varphi), \dots, X_d(\varphi)) = (X_1(\varphi, \omega), \dots, X_d(\varphi, \omega))$$

defined for all functions $\varphi \in \mathcal{S}$, where $\mathcal{S}$ is the Schwartz space.

Definition of vector valued generalized random fields. *We say that the set of random vectors* $(X_1(\varphi), \dots, X_d(\varphi))$, $\varphi \in \mathcal{S}$, *is a* d*-dimensional vector valued generalized random field over the Schwartz space* $\mathcal{S}$ *of rapidly decreasing, smooth functions if:*

(a) $X_j(a_1\varphi + a_2\psi) = a_1 X_j(\varphi_j) + a_2 X_j(\psi)$ *with probability 1 for the* j*-th coordinate of the random vectors* $(X_1(\varphi), \dots, X_d(\varphi))$ *and* $(X_1(\psi), \dots, X_d(\psi))$. *This relation holds for each coordinate* $1 \le j \le d$, *all real numbers* a_1 *and* a_2, *and pair of functions* φ, ψ *from the Schwartz space* $\mathcal{S}$. *(The exceptional set of probability 0 where this identity does not hold may depend on* a_1, a_2, φ *and* ψ.*)*

(b) $X_j(\varphi_n) \Rightarrow X_j(\varphi)$ *stochastically for any* $1 \le j \le d$ *if* $\varphi_n \to \varphi$ *in the topology of* $\mathcal{S}$.

We also introduce the following definition. Here $\stackrel{\Delta}{=}$ denotes equality in distribution.

Definition of stationarity and Gaussian property for a vector valued generalized random field. *The d-dimensional vector valued generalized random field $X = \{(X_1(\varphi)\dots, X_d(\varphi)),\ \varphi \in \mathcal{S}\}$ is stationary if*

$$(X_1(\varphi)\dots, X_d(\varphi)) \stackrel{\Delta}{=} (X_1(T_t\varphi)\dots, X_d(T_t\varphi))$$

for all $\varphi \in \mathcal{S}$ and $t \in R$, where $T_t\varphi(x) = \varphi(x-t)$. It is Gaussian if $(X_1(\varphi), \dots, X_d(\varphi))$ is a Gaussian random vector for all $\varphi \in \mathcal{S}$. We call a vector valued generalized random field a vector valued generalized random field with zero expectation if $EX_j(\varphi) = 0$ for all $\varphi \in \mathcal{S}$ and coordinates $1 \le j \le d$.

The covariance function $r_{j,k}(\varphi, \psi) = EX_j(\varphi)X_k(\psi)$, $\varphi, \psi \in \mathcal{S}$, determines the distribution of a vector valued stationary random field, and we want to give a „spectral measure" type representation for it. The following result gives such a representation.

Theorem about the covariance function of vector valued Gaussian stationary generalized random fields with zero expectation.

Let $(X_1(\varphi),\dots,X_d(\varphi))$ be a vector valued, Gaussian stationary generalized random field with expectation zero. Its covariance function $r_{j,k}(\varphi,\psi)=EX_j(\varphi)X_k(\psi)$ with $\varphi\in\mathcal{S}$ and $\psi\in\mathcal{S}$ is given by the formula

$$r_{j,k}(\varphi,\psi)=EX_j(\varphi)X_k(\psi)=\int\tilde\varphi(x)\bar{\tilde\psi}(x)G_{j,k}(\,dx)$$

for all $\varphi,\psi\in\mathcal{S}$ where $\tilde{}$ denotes Fourier transform, and $\bar{}$ is complex conjugate, and we integrate with respect to complex number valued measures $G_{j,k}$ which have the following properties.

$G_{j,k}$, $1 \le j, k \le d$, are complex number valued measures on R with locally finite variation such that the matrix $G(A) = (G_{j,k}(A))$, $1 \le j, k \le d$, is positive semidefinite for all bounded sets A. Besides, they satisfy, instead of boundedness, the following weaker condition.

$$\int (1+|x|)^{-r} G_{j,j}(\,dx) < \infty \quad \text{for all } 1 \le j \le d$$

with some number $r > 0$.

There can be defined a vector valued random spectral measure $(Z_{G_1}, \dots, Z_{G_d})$ corresponding to a matrix valued spectral measure $(G_{j,k})$, $1 \le j, k \le d$, and random integral with respect to it such that

$$X_j(\varphi) = \int \tilde{\varphi}(x)\, Z_{G_j}(\,dx), \quad \varphi \in \mathcal{S}, \ \ 1 \le j \le d,$$

is a generalized random field with spectral measure $(G_{j,k})$, $1 \le j, k \le d$. (There exists e.g. a random spectral measure corresponding $G(\,dx) = |x|^{-\alpha}\, dx, \ \ \alpha < 1$.)

The goal for defining multiple Wiener–itô integrals.

Let us have a Gaussian stationary random sequence $X(n)$, $n = 0, \pm 1, \dots$, or $X(n) = (X_1(n), \dots, X_d(n))$, $n = 0, \pm 1, \dots$, and let us consider the (real) Hilbert space $\mathcal{H}$ consisting of the square integrable random variables measurable with respect to the σ-algebra generated by the elements of this sequence of random variables.
We want to give a good representation of the elements of this Hilbert space.

First step: Let $\mathcal{H}_1$ denote the subspace of $\mathcal{H}$ consisting of the closure of the finite linear combinations $\sum c_k X(n_k)$ (scalar valued case) or $\sum_{j=1}^{d} \sum c_{j,k} X_j(n_k)$ (vector valued case). Give a good representation of the elements of $\mathcal{H}_1$.

If $X(n) = \int e^{inx} Z_G(\,dx)$ (scalar valued case) or $X_j(n) = \int e^{inx} Z_{G_j}(\,dx)$, $1 \le j \le d$, (vector valued case), $n = 0, \pm 1, \dots$,then the elements of $\mathcal{H}_1$ can be written as one-fold random integrals with respect to the random spectral measure Z_G or Z_{G_p}, as $\int \varphi(x) Z_G(\,dx)$ or $\sum_{j=1}^{d} \int \varphi_j(x) Z_{G_j}(\,dx)$ with kernel functions such that $\int |\varphi^2(x)| G(\,dx) < \infty$, $\int |\varphi_j^2(x)| G_{j.j}(\,dx) < \infty$, and $\varphi(-x) = \overline{\varphi(x)}$, $\varphi_j(-x) = \overline{\varphi_j(x)}$ for all $x \in R$.

We can define multiple Wiener–Itô integrals

$$\int f(x_1, \dots, x_k) Z_G(\,dx_1) \dots Z_G(\,dx_k)$$

with random spectral measure Z_G (scalar valued case),

$$\int f(x_1, \dots, x_k) Z_{G_{j_1}}(\,dx_1) \dots Z_{G_{j_k}}(\,dx_k)$$

$1 \le j_p \le d$ for all $1 \le p \le k$ with the vector valued random spectral measure $(Z_{G_1}, \dots, Z_{G_d})$ (vector valued case).

The integral

$$\int f(x_1, \ldots, x_k) Z_{G_{j_1}}(\,dx_1) \ldots Z_{G_{j_k}}(\,dx_k)$$

is defined for kernel functions $f(\cdot)$ such that

$$\int |f(x_1, \ldots, x_k)|^2 G_{j_1,j_1}(\,dx_1) \ldots G_{j_k,j_k}(\,dx_k) < \infty,$$

and

$$f(-x_1, \ldots, -x_k) = \overline{f(x_1, \ldots, x_k)}$$

for all coordinates $(x_1, \ldots, x_k)$.

The elements of $\mathcal{H}$ can be expressed as a sum of multiple Wiener–Itô integrals. Moreover, by a result called Itô's formula we can do this in an explicit way. In particular: If $X = \int g(x) Z_G(\,dx)$, $EX^2 = 1$, then

$$H_k(X) = \int g(x_1) \cdots g(x_k)\, Z_G(\,dx_1) \ldots Z_G(\,dx_k).$$

This result has a multivariate generalization.

Why is it useful to work with multiple Wiener–Itô integrals?

Shift transformations

Given a (vector valued) stationary random sequence $(X(n) = X_1(n), \dots, X_d(n))$, $n = 0, \pm 1, \pm 2, \dots$ we can define the shift transformation T_k for all $k = 0, \pm 1, \pm 2, \dots$ by the formula $T_k X_j(n) = X_j(n+k)$ for all $n = 0, \pm 1, \pm 2, \dots$, $1 \le p \le d$, or more generally
$T_k g(X_{j_1}(n_1), \dots, X_{j_s}(n_s)) = g(X_{j_1}(n_1+k), \dots, X_{j_s}(n_s+k))$ for a general function g and indices $j_1, \dots, j_s$, $n_1, \dots, n_s$. In such a way the shift transformation can be extended to $\mathcal{H}$.
For a random variable given in the form of a multiple integral there is a simple and useful formula to calculate its shift transformation. If

$$Y = \int h(x_1, \dots, x_n) Z_{G_{j_1}}(\,dx_1) \dots Z_{G_{j_n}}(\,dx_n),$$

then

$$T_k Y = \int e^{ik(x_1 + \cdots + x_n)} h(x_1, \dots, x_n) Z_{G_{j_1}}(\,dx_1) \dots Z_{G_{j_n}}(\,dx_n).$$

An application

Let $X(n)$, $n = 0, \pm 1, \pm 2, \dots$, be a Gaussian stationary random sequence, $EX(0) = 0$, and $EX^2(0) = 1$, such that $X(n) = \int e^{inx} Z_G(\,dx)$ with a random spectral measure Z_G. $H_k(x)$ is the k-th Hermite polynomial (with leading coefficient 1).
By Itô's formula

$$H_k(X(0)) = \int Z_G(\,dx_1) \dots Z_G(dx_k),$$

hence

$$H_k(X(n)) = T_n H_k(X(0)) = \int e^{in(x_1+\cdots+x_k)} Z_G(\,dx_1) \dots Z_G(\,dx_k)$$

and

$$\sum_{n=0}^{N-1} H_k(X(n)) = \int \frac{e^{iN(x_1+\cdots+x_k)} - 1}{e^{i(x_1+\cdots+x_k)} - 1} Z_G(\,dx_1) \dots Z_G(\,dx_k).$$

It is useful to apply the change of variables $y_s = Nx_s$, $1 \le s \le k$. Some calculation shows that for

$$S_N = \frac{1}{N^{1-k\alpha/2} L(N)^{k/2}} \sum_{n=0}^{N-1} H_k(X_n)$$

$$S_N = \int \frac{e^{i(y_1+\cdots+y_k)} - 1}{N\left(e^{i\frac{1}{N}(y_1+\cdots+y_k)} - 1\right)} Z_{G_N}(\,dy_1)\ldots Z_{G_N}(dy_k)$$

with $Z_{G_N}(A) = \frac{N^{\alpha/2}}{L(N)^{1/2}} Z_G(\frac{A}{N})$ for all measurable sets $A \subset [-N\pi, N\pi]$. It is a random spectral measure corresponding to the spectral measure G_N, $G_N(A) = \frac{N^\alpha}{L(N)} G(\frac{A}{N})$ on $[-N\pi, N\pi]$. Clearly,

$$\frac{e^{i(y_1+\cdots+y_k)} - 1}{N\left(e^{i\frac{1}{N}(y_1+\cdots+y_k)} - 1\right)} \to \frac{e^{i(y_1+\cdots+y_k)} - 1}{i(y_1+\cdots+y_k)} \quad \text{as } N \to \infty.$$

It can be proved that if $r(n) = n^{-\alpha} L(n)$, $n \geq 0$, then the sequence of the spectral measures G_N, $G_N(A) = \frac{N^\alpha}{L(N)} G(\frac{A}{N})$ tend (vaguely) to the spectral measure G_0 (of a generalized stationary random sequence) with density function $C|x|^{\alpha-1}$ with some $C > 0$.

The proof of our result with Dobrushin consists of a justification of a limiting procedure suggested by the above limit relations. The k-fold Wiener–Itô integral defining the limit must exist. To satisfy this we need the condition $k\alpha < 1$.

The above result can be generalized to vector valued Gaussian stationary random sequences. One has to prove a multivariate version of Itô's formula, and of the result which describes the asymptotic behaviour of the spectral measure of a vector valued stationary random sequence $X(n)$, $n = 0, \pm 1, \pm 2, \dots$, whose covariance functions behave at infinity as

$$r_{j,k}(n) = EX_j(0)X_k(n) \sim C_{j,k}|n|^{-\alpha} L(n).$$

The following result can be proved.

Let us fix some integer $k \geq 1$, and let $X(n) = (X_1(n), \dots, X_d(n))$, $n = 0, \pm 1, \dots$ be a vector valued Gaussian stationary random sequence with covariance functions $r_{j,k}(\cdot)$ such that $r_{j,k}(n) = EX_j(0)X_k(n) \sim C_{j,k} n^{-\alpha} L(n)$, with a slowly varying function $L(n)$ as $n \to \infty$, with some $0 < \alpha < \frac{\nu}{k}$, and $EX_j(0)X_k(0) = \delta_{j,k}$, $1 \leq j, k \leq d$,
Put

$$H(x_1, \dots, x_d) = \sum_{\substack{(k_1,\dots,k_d),\ k_j \geq 0,\ 1 \leq j \leq d, \\ k_1 + \cdots + k_d = k}} c_{k_1,\dots,k_d} H_{k_1}(x_1) \cdots H_{k_d}(x_d)$$

where $H_k(x)$ denotes the k-th Hermite polynomial.
Define the normalized partial sums

$$S_N = \frac{1}{N^{\nu - k\alpha/2} L(N)^{k/2}} \sum_{n=0}^{N-1} H(X_1(n), \dots, X_d(n))$$

for all all $N = 1, 2, \dots$. Then the sequence S_N converges in distribution to the following random variable S_0 as $N \to \infty$.

Let $(G_{j,k})$, $1 \le j, k \le d$, denote the spectral measure of the random sequence $X(n) = (X_1(n), \dots, X_d(n))$, $n = 0, \pm 1, \dots$, and define their normalizations $G_{j,k}^{(N)}(A) = \frac{N^\alpha}{L(N)} G_{j,k}(\frac{A}{N})$. Then the sequence $(G_{j,k}^{(N)})$ converges to the spectral measure $(G_{j,k}^{(0)})$ of a vector valued stationary generalized random field as $N \to \infty$. Let $Z_{G^{(0)}} = (Z_{G_1^{(0)}}, \dots, Z_{G_d^{(0)}})$ be a vector valued random spectral measure which corresponds to the matrix valued spectral measure $(G_{j,k}^{(0)})$, $1 \le j, k \le d$. Then the random sums S_N converge to the sum of multiple Wiener–Itô integrals

$$S_0 = \sum_{\substack{(k_1,\dots,k_d),\ k_j \ge 0,\ 1 \le j \le d, \\ k_1 + \cdots + k_d = k}} c_{k_1,\dots,k_d} \int \frac{e^{i(x_1+\cdots+x_k)} - 1}{i(x_1 + \cdots + x_k)} Z_{G^{(0)}_{j(1|k_1,\dots,k_d)}}(\,dx_1) \dots Z_{G^{(0)}_{j(k|k_1,\dots,k_d)}}(\,dx_k)$$

as $N \to \infty$. Here we use the notation $j(s|k_1, \dots, k_d) = r$ if $\sum_{u=1}^{s-1} k_u < r \le \sum_{u=1}^{s} k_u$, $1 \le s \le k$.