Препринт Санкт-Петербургского математического общества
Поступил 30.05.2020
Доступен на сайте http://www.mathsoc.spb.ru/preprint/

Online-centered Gaussian processes and applications

A.I. Nazarov* and Ya.Yu. Nikitin**

Abstract

We establish various properties of the online centering of Gaussian processes and discuss their application to goodness-of-fit testing.

Keywords: Gaussian processes, online centering, spectral equivalence

The operation of online centering was introduced in [8]. For a Gaussian process X on $[0,1]$, we consider the process

$$\widehat{X}(x) = X(x) - \frac{1}{x}\int\limits_0^x X(t)\,dt.$$

Proposition ([8, Example 4], [5, Proposition 6.3]). *The online centered Brownian motion is spectrally equivalent to the usual centered Brownian motion:*

$$\widehat{W}(x) \sim \overline{W}(x) := W(x) - \int\limits_0^1 W(t)\,dt. \tag{1}$$

*St.Petersburg Department of Steklov Mathematical Institute of Russian Academy of Science, Fontanka 27, St.Petersburg, 191023, Russia, and St.Petersburg State University, Universitetskaya emb. 7-9, St.Petersburg, 199034, Russia; E-mail: nazarov@pdmi.ras.ru

**St.Petersburg State University, Universitetskaya emb. 7-9, St.Petersburg, 199034, Russia, and National Research University – Higher School of Economics, Soyuza Pechatnikov 16, St.Petersburg, 190008, Russia; E-mail: y.nikitin@spbu.ru

In [8] this fact was proved by the Laplace transform while in [5] a direct calculation of the spectrum of $\widehat{W}$ was given.

In this paper we demonstrate the operator nature of the relation (1) and show that a similar identity holds for the entire class of online centered Gaussian processes.

Let A and B be compact operators in a Hilbert space H. We call A and B *spectrally equivalent* and write $A \sim B$ if their non-zero eigenvalues coincide (with the multiplicities). A typical example of such operators is given by operator products $AB \sim BA$, see, e.g., [1, Section 3.10].

Also we call two Gaussian random functions X and Y *spectrally equivalent* and write $X \sim Y$ if their covariance operators K_X and K_Y are spectrally equivalent. The notion of spectral equivalence was introduced in the recent paper [14] though examples of such functions, both in univariate and in multivariate case, were known much earlier, see the references in [14].

We define some operators in $L_2(0,1)$: operators of integration from the left and from the right

$$(Tu)(x) = \int_0^x u(t)\,dt, \qquad (T^*u)(x) = \int_x^1 u(t)\,dt,$$

the orthogonal projector onto the subspace of constants, the multiplication operator

$$(Pu)(x) = \int_0^1 u(t)\,dt, \qquad (Su)(x) = xu(x),$$

and the operator of online centering

$$(\widehat{\mathbb{T}}u)(x) = u(x) - \frac{1}{x}\int_0^x u(t)\,dt.$$

Proposition 1. The following identities hold:

$$\widehat{\mathbb{T}}^*\widehat{\mathbb{T}} = I - P; \qquad \widehat{\mathbb{T}}\,\widehat{\mathbb{T}}^* = I. \tag{2}$$

This statement shows that operators $\widehat{\mathbb{T}}$ and $\widehat{\mathbb{T}}^*$ form, respectively, the left and the right shifts in $L_2(0,1)$. This fact was proved in [2], see also [9, Theorem 1.1]. We give here an elementary proof for the reader's convenience. We have

$$\widehat{\mathbb{T}}^*\widehat{\mathbb{T}}u = (I - T^*S^{-1})(I - S^{-1}T)u,$$

i.e.

$$(\widehat{\mathbb{T}}^*\widehat{\mathbb{T}}u)(x) = u(x) - \frac{1}{x}\int_0^x u(t)\,dt - \int_x^1 \frac{u(t)}{t}\,dt + \int_x^1 \frac{1}{t^2}\int_0^t u(s)\,dsdt.$$

Since

$$\int_x^1 \Big(\frac{u(t)}{t} - \frac{1}{t^2}\int_0^t u(s)\,ds\Big)dt = \Big(\frac{1}{t}\int_0^t u(s)\,ds\Big)\Big|_x^1 = \int_0^1 u(t)\,dt - \frac{1}{x}\int_0^x u(t)\,dt,$$

we arrive at

$$(\widehat{\mathbb{T}}^*\widehat{\mathbb{T}}u)(x) = u(x) - \int_0^1 u(t)\,dt = ((I-P)u)(x).$$

On the other hand,

$$(\widehat{\mathbb{T}}\,\widehat{\mathbb{T}}^*u)(x) = u(x) - \frac{1}{x}\int_0^x u(t)\,dt - \int_x^1 \frac{u(t)}{t}\,dt + \frac{1}{x}\int_0^x\int_t^1 \frac{u(s)}{s}\,dsdt.$$

Integration by parts annihilates three last terms, and the statement follows. $\square$

Now we can formulate our first main result.

Theorem 1. *For any zero mean-value Gaussian process $\mathcal{X}(x)$ on $[0,1]$, the online centered process $\widehat{\mathcal{X}}(x)$ is spectrally equivalent to the usual centered process:*

$$\widehat{\mathcal{X}}(x) \sim \overline{\mathcal{X}}(x) := \mathcal{X}(x) - \int_0^1 \mathcal{X}(t)\,dt. \tag{3}$$

Proof. It is easy to see that the covariance operators of the online centered and the usual centered process admit the representation

$$K_{\widehat{\mathcal{X}}} = \widehat{\mathbb{T}}\,K_{\mathcal{X}}\widehat{\mathbb{T}}^*; \qquad K_{\overline{\mathcal{X}}} = (I-P)K_{\mathcal{X}}(I-P).$$

Therefore, we can write down the following chain:

$$\widehat{\mathbb{T}}\cdot\big[K_{\mathcal{X}}\widehat{\mathbb{T}}^*\big] \sim \big[K_{\mathcal{X}}\widehat{\mathbb{T}}^*\big]\cdot\widehat{\mathbb{T}} \overset{\bullet}{=} K_{\mathcal{X}}(I-P) = \big[K_{\mathcal{X}}(I-P)\big]\cdot(I-P) \sim (I-P)\cdot\big[K_{\mathcal{X}}(I-P)\big],$$

(the equality ($\bullet$) follows from Proposition 1), and the statement follows. $\square$

Remark 1. Notice that the online centered Brownian motion is a *Green Gaussian process*, i.e. its covariance function is the Green function of a boundary value problem for an ordinary differential operator, see [5, Proposition 6.3]. Since the spectral theory of ODOs is well developed, this fact always helps a lot in search of the spectrum, see, e.g., [13], [10]. In contrast, for many Green Gaussian processes, the corresponding online centered process is NOT a Green Gaussian process. For instance, this is the case for the Brownian bridge. However, the equivalence (3) associates the spectrum of $\widehat{B}$ with that of $\overline{B}$ which is known long ago, see [18].

Another interesting example is related to the fractional Brownian motion (FBM) W^H that is a zero mean-value Gaussian process with covariance function

$$G_{W^H}(x,y) = \frac{1}{2}\big(x^{2H} + y^{2H} - |x-y|^{2H}\big), \qquad x, y \in [0,1],$$

(here $H \in (0,1)$ is the so-called Hurst index, the case $H = \frac{1}{2}$ corresponds to the standard Brownian motion). Using Theorem 1 we obtain

$$\widehat{W^H}(x) \sim \overline{W^H}(x).$$

Notice that all fractional processes are not Green Gaussian processes, and their spectrum is not known exactly. However, recently the sharp spectral asymptotics for $\overline{W^H}$ were obtained in [11] using a breakthrough approach of [3].

Remark 1 generates a natural question, for which Green Gaussian processes corresponding online centered process is again a Green Gaussian process. A partial answer is given by the following theorem.

Theorem 2. *Let $X(x)$ be a zero mean-value Green Gaussian process on $[0,1]$. Denote by $\mathcal{X}(x)$ the left-integrated process*

$$\mathcal{X}(x) = \int\limits_0^x X(t)\,dt, \qquad x \in [0,1].$$

Then the online centered process $\widehat{\mathcal{X}}(x)$ is also a Green Gaussian process.

Before giving the proof we recall that the assumption of theorem means that the covariance function G_X satisfies

$$LG_X(\cdot, y) = \delta(\cdot - y); \qquad G_X(\cdot, y) \in Dom(L), \tag{4}$$

where L is a self-adjoint ordinary differential operator of order 2ℓ,

$$L \equiv (-1)^\ell D^\ell \left(p_\ell(x) D^\ell\right) + D^{\ell-1}\left(p_{\ell-1}(x) D^{\ell-1}\right) + \cdots + p_0(x)$$

(here D stands for the differentiation operator, and $p_\ell(x) > 0$) and the domain $Dom(L)$ is defined by 2ℓ boundary conditions. In operator terms, (4) can be written as $LK_X = I$.

Proof. It is easy to see that the covariance operators of $\mathcal{X}$ and $\widehat{\mathcal{X}}$ admit the representation

$$K_{\mathcal{X}} = TK_XT^*; \qquad K_{\widehat{\mathcal{X}}} = \widehat{\mathbb{T}}TK_XT^*\widehat{\mathbb{T}}^*.$$

We begin with the identity

$$\widehat{\mathbb{T}}T = S^{-1}TS, \tag{5}$$

which follows from a simple integration by parts formula

$$\int_0^x tf(t)\,dt = x\int_0^x f(t)\,dt - \int_0^x\int_0^t f(s)\,dsdt.$$

So, we obtain

$$K_{\widehat{\mathcal{X}}} = S^{-1}TSK_XST^*S^{-1}.$$

We invert the operator factors consequently and arrive at $\widehat{\mathcal{L}}K_{\widehat{\mathcal{X}}} = I$, where $\widehat{\mathcal{L}}$ is an ODO of order $2\ell+2$ given by

$$\widehat{\mathcal{L}} \equiv xDx^{-1}Lx^{-1}Dx.$$

Since for every $u \in L_2(0,1)$ we have

$$(ST^*S^{-1}u)(1) = 0, \qquad K_XST^*S^{-1}u \in Dom(L), \qquad (K_{\widehat{\mathcal{X}}}u)(0) = 0,$$

the domain $Dom(\widehat{\mathcal{L}})$ is defined by $2\ell+2$ boundary conditions

$$u(0) = 0; \qquad x^{-1}Dxu \in Dom(L); \qquad (Lx^{-1}Dxu)(1) = 0.$$

Thus, the covariance function $G_{\widehat{\mathcal{X}}}$ satisfies

$$\widehat{\mathcal{L}}G_{\widehat{\mathcal{X}}}(\cdot,y) = \delta(\cdot - y); \qquad G_{\widehat{\mathcal{X}}}(\cdot,y) \in Dom(\widehat{\mathcal{L}}),$$

and the statement follows. □

Remark 2. In the case where L is an operator with constant coefficients, the fundamental system of solutions to the equation $\widehat{\mathcal{L}}u = \mu u$ can be written in terms of elementary functions, see [12]. We stress that this family of explicitly solvable ODEs is not included into classical handbooks [4, 16].

Next, by virtue of the Karhunen–Loève expansion, spectrally equivalent Gaussian functions have equally distributed L_2-norms. By Theorem 1, the following identity in law holds for any zero mean-value Gaussian process $\mathcal{X}(x)$ on $[0,1]$:

$$\|\widehat{\mathcal{X}}\|^2_{L_2(0,1)} \stackrel{d}{=} \|\overline{\mathcal{X}}\|^2_{L_2(0,1)}.$$

However, we provide a much stronger statement.

Theorem 3. *Let $X(x)$ be arbitrary random process on $[0,1]$, square integrable a.s. Then*

$$\|\widehat{\mathcal{X}}\|^2_{L_2(0,1)} = \|\overline{\mathcal{X}}\|^2_{L_2(0,1)} \qquad \text{a.s.}$$

Proof. Using Proposition 1 we derive

$$\int_0^1 \widehat{\mathcal{X}}^2(t)\,dt = \int_0^1 (\widehat{\mathbb{T}}\mathcal{X})^2(t)\,dt = \int_0^1 (\widehat{\mathbb{T}}^*\widehat{\mathbb{T}}\mathcal{X})(t)\cdot\mathcal{X}(t)\,dt$$
$$= \int_0^1 ((I-P)\mathcal{X})(t)\cdot\mathcal{X}(t)\,dt = \int_0^1 ((I-P)\mathcal{X})^2(t)\,dt = \int_0^1 \overline{\mathcal{X}}^2(t)\,dt,$$

and the statement follows. □

The results obtained above may have an unexpected application to nonparametrric statistics, namely to goodness-of-fit testing. Consider the classical empirical process built on a uniform sample on $[0,1]$

$$\xi_n(t) = \sqrt{n}(F_n(t) - t),\ 0 \le t \le 1,$$

where $F_n(t)$ is the empirical distribution function. The functionals of the empirical process are the famous nonparametric statistics such as Kolmogorov, Cramér–von Mises, Watson, Anderson–Darling statistics, and many others.

All them are used for goodness-of-fit testing. It is well known, see, e.g., the classical monograph [17] that the empirical process converges weakly in the Skorokhod space $D[0,1]$ to the Brownian bridge, and therefore the limiting distributions of statistics listed above coincide with these of Brownian bridge, and therefore are well studied.

In nonparametric statistics, the researchers are very interested in transformations of the empirical process and in functionals from such processes in the hope of finding new, more powerful or efficient tests for fit. The examples of such transformations are the extracting of the martingale part and more general constructions due to Khmaladze [6], [7] or the so-called Deheuvels empirical process [15].

The online centered empirical process has never been considered in this context. It has the form

$$\widehat{\xi}_n(t) = \xi_n(t) - \frac{1}{t}\int_0^t \xi_n(s)ds, \qquad 0 \le t \le 1.$$

This process converges in Skorokhod space to the process $\widehat{B}$ which is spectrally equivalent to the usual centered Brownian bridge $\overline{B}$ by Theorem 1.

The corresponding ω^2-type statistic $\int_0^1 \widehat{\xi}_n^2(t)dt$ not only has the same limiting distribution as $\|\overline{B}\|^2_{L_2(0,1)}$ but equals the Watson statistics $\int_0^1 \overline{\xi}_n^2(t)dt$ almost surely by Theorem 3. However, it would be interesting to calculate and compare the local Bahadur asymptotic efficiency of other statistics based on $\widehat{\xi}_n$ against standard alternatives.

Acknowledgement

This work was supported by the Russian Foundation of Basic Research Grant 20-51-12004.

References

[1] Birman, M.S., Solomyak, M.Z. *Spectral theory of self-adjoint operators in Hilbert space*, 2nd ed., revised and extended. Lan', St.Petersburg, 2010 (Russian); English transl. of the 1st ed.: Mathematics and Its Applications. Soviet Series, **5**, Kluwer, Dordrecht etc. 1987.

[2] Brown, A., Halmos, P.R., Shields, A.L. Cesàro operators, Acta Sci. Math. (Szeged) **26** (1965), N1-2, 125–137.

[3] Chigansky, P., Kleptsyna, M. Exact asymptotics in eigenproblems for fractional Brownian covariance operators. Stoch. Proc. Appl., **128** (2018), N6, 2007–2059.

[4] Kamke, E. *Differentialgleichungen Losungsmethoden und Losungen*, vol. II. Chelsea Publishing Company, New York, 1947.

[5] Karol', A., Nazarov, A., Nikitin, Ya. Small ball probabilities for Gaussian random fields and tensor products of compact operators, Trans. AMS, **360** (2008), N3, 1443–1474.

[6] Khmaladze, E.V., Martingale approach in the theory of goodness-of-fit tests. Theory of Probability & Its Applications. **26** (1982), N2, 240–257.

[7] Khmaladze, E. Unitary transformations, empirical processes and distribution free testing. Bernoulli, **22** (2016), N1, 563–588.

[8] Kleptsyna, M.L., Le Breton, A. A Cameron–Martin type formula for general Gaussian processes – a filtering approach, Stochastics, **72** (2002), N3, 229–250.

[9] Kruglyak, N., Maligranda, L., Persson, L.-E., Structure of the Hardy operator related to Laguerre polynomials and the Euler differential equation, Revista Matemática Complutense, **19** (2006), N2, 467–476.

[10] Nazarov, A.I. Exact L_2-small ball asymptotics of Gaussian processes and the spectrum of boundary-value problems, J. Theor. Probab., **22** (2009), N3, 640–665.

[11] Nazarov, A.I., Spectral asymptotics for a class of integro-differential equations arising in the theory of fractional Gaussian processes, Preprint available at https://arxiv.org/abs/1908.10299. 30p. To appear in Comm. Contemp. Math.

[12] Nazarov, A.I., On a family of ordinary differential equations integrable in elementary functions, 2020, to appear in Math. Notes.

[13] Nazarov, A.I., Nikitin, Ya.Yu. Exact small ball behavior of integrated Gaussian processes under L_2-norm and spectral asymptotics of boundary value problems. Probab. Theory Rel. Fields, **129** (2004), N4, 469–494.

[14] Nazarov, A.I., Nikitin, Ya.Yu., Spectral equivalence of Gaussian random functions: operator approach, Preprint available at https://arxiv.org/abs/2005.05771. 14p.

[15] Nikitin Y., Sporysheva P. On asymptotic efficiency of tests of fit based on the Deheuvels empirical process. J. Math. Sci. **159** (2009), 317–323.

[16] Polyanin, A.D., Zaitsev, V.F., *Handbook of Exact Solutions for Ordinary Differential Equations*, CRC Press, Boca Raton, 2003

[17] Shorack, G.R., Wellner, J.A. *Empirical Processes with Applications to Statistics*, SIAM, 2009.

[18] Watson, G.S. Goodness-of-fit tests on a circle. Biometrika, **48** (1961), N 1/2, 109–114.