

MAREK BEŚKA (Gdańsk)

## NOTE ON THE VARIANCE OF THE SUM OF GAUSSIAN FUNCTIONALS

*Abstract.* Let  $(X_i, i = 1, 2, \dots)$  be a Gaussian sequence with  $X_i \in N(0, 1)$  for each  $i$  and suppose its correlation matrix  $R = (\rho_{ij})_{i,j \geq 1}$  is the matrix of some linear operator  $R: l_2 \rightarrow l_2$ . Then for  $f_i \in L^2(\mu)$ ,  $i = 1, 2, \dots$ , where  $\mu$  is the standard normal distribution, we estimate the variation of the sum of the Gaussian functionals  $f_i(X_i)$ ,  $i = 1, 2, \dots$ .

**1. Introduction.** Let  $(X, Y)$  be a Gaussian random vector such that  $X, Y \in N(0, 1)$  and  $E(XY) = \rho$ ,  $(|\rho| < 1)$ . We denote by  $\mu$  the normalized one-dimensional Gaussian measure, i.e.

$$\mu(dx) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx.$$

In  $L^2(\mu)$  we have the scalar product

$$(f, g)_\mu = \int_{\mathbb{R}} f(x)g(x) \mu(dx).$$

Introducing a random variable  $Z \in N(0, 1)$  such that  $Z, Y$  are independent, we find that the Gaussian vectors  $(X, Y)$  and  $(U, Y)$  with  $U = \rho Y + \sqrt{1 - \rho^2} Z$  have the same joint distribution. Thus, for  $f, g \in L^2(\mu)$  we have

$$(1.1) \quad E(f(X)g(Y)) = E(f(U)g(Y)) = E(P_\rho(Y)g(Y)),$$

where

$$P_\rho f(y) = E(f(U) | Y = y) = \int_{\mathbb{R}} f(\rho y + \sqrt{1 - \rho^2} z) d\mu(z), \quad y \in \mathbb{R},$$

is called the *Ornstein–Uhlenbeck operator*. The Ornstein–Uhlenbeck operator has a representation in terms of orthonormal Hermite polynomials

---

2010 *Mathematics Subject Classification*: Primary 60E15; Secondary 60F15.

*Key words and phrases*: Gebelein inequality, Ornstein–Uhlenbeck operator, Hermite polynomials.

$\{h_n\}_{n \geq 0} \subset L^2(\mu)$ , namely

$$(1.2) \quad P_\rho f = \sum_{n=0}^{\infty} \rho^n (f, h_n)_\mu h_n, \quad f \in L^2(\mu).$$

In particular,

$$P_\rho h_n = \rho^n h_n, \quad n \geq 0.$$

From (1.2) we obtain Gebelein’s inequality (see [G] and [DK]):

PROPOSITION 1.1. *If  $f \in L^2$  and  $(f, 1)_\mu = 0$ , then*

$$(1.3) \quad \|P_\rho f\|_2 \leq |\rho| \cdot \|f\|_2,$$

*or equivalently for any  $g \in L^2$  and  $f$  as above,*

$$|(P_\rho f, g)_\mu| \leq |\rho| \cdot \|f\|_2 \cdot \|g\|_2.$$

*In both inequalities we have equality if and only if  $f(x) = cx$ .*

Consider a Gaussian sequence  $(X_i, i = 1, 2, \dots)$  of random variables with  $X_i \in N(0, 1)$  for each  $i$ . It is assumed that the correlation matrix  $R = (\rho_{ij})_{i,j \geq 1}$ , where  $\rho_{ij} = E(X_i X_j)$ ,  $i, j = 1, 2, \dots$ , satisfies

$$(1.4) \quad C = \sup_{i \geq 1} \sum_{j \geq 1} |\rho_{ij}| < \infty.$$

It is evident that  $C \geq 1$ . The Frobenius Theorem (see [HLP]) implies that  $R$  is the matrix (in the standard basis) of a continuous linear operator (which we denote by the same letter)  $R : l_p \rightarrow l_p$  for  $1 \leq p \leq \infty$  with  $\|R\| \leq C$ . Hence, it is easily seen that for  $C < 2$  the linear operator  $R$  is invertible. Using Gebelein’s inequality (1.3), one can prove (see [BC1], [BC2], [V])

LEMMA 1.1. *Let the Gaussian sequence  $(X_i, i = 1, 2, \dots)$  with  $X_i \in N(0, 1)$  for each  $i$  satisfy the hypothesis (1.4) and let  $(f_i, i = 1, 2, \dots) \subset L^2(\mu)$ . Then for each  $n \geq 1$  we have*

$$(1.5) \quad (2 - C) \sum_{i=1}^n \text{Var}(f_i(X_i)) \leq \text{Var}\left(\sum_{i=1}^n f_i(X_i)\right) \leq C \sum_{i=1}^n \text{Var}(f_i(X_i)).$$

For  $C \geq 2$  the left inequality in (1.5) holds trivially. In fact, we can say more: an inequality of the form

$$(1.6) \quad M \sum_{i=1}^n \text{Var}(f_i(X_i)) \leq \text{Var}\left(\sum_{i=1}^n f_i(X_i)\right),$$

where  $M$  is a positive constant, is not satisfied in general when  $C \geq 2$ .

Consider the following simple example: Let  $(Y_i, i = 1, 2, \dots) \subset N(0, 1)$  be a sequence of independent Gaussian random variables. Let  $a, b \in \mathbb{R}$  be such that  $a^2 + b^2 = 1$  and define

$$X_{3k-2} = -Y_{2k}, \quad X_{3k-1} = a Y_{2k-1} - b Y_{2k}, \quad X_{3k} = a Y_{2k-1} + b Y_{2k}, \quad k \geq 1.$$

Moreover, we put

$$f_{3k-2}(t) = 2bt, \quad f_{3k-1}(t) = -t, \quad f_{3k}(t) = t, \quad t \in \mathbb{R}, k \geq 1.$$

It is easy to check that

$$C = \sup_{i \geq 1} \sum_{j \geq 1} |\rho_{ij}| = 1 + |b| + \max\{|b|, |1 - 2b^2|\} \geq 2$$

and

$$\text{Var}\left(\sum_{i=1}^{3n} f_i(X_i)\right) = 0 \quad \text{and} \quad \sum_{i=1}^{3n} \text{Var}(f_i(X_i)) > 0, \quad n \geq 1.$$

**2. Main result.** In this section we are going to prove the inequality (1.5) under a slightly weaker condition than (1.4). First let us introduce some notations. For a given correlation matrix  $R = (\rho_{ij})_{i,j \geq 1}$ , we put

$$R_n^{(m)} = (\rho_{ij}^{(m)})_{1 \leq i,j \leq n}, \quad m, n \geq 1,$$

and let  $\lambda_{n,1}^{(m)}$  and  $\lambda_{n,n}^{(m)}$  denote the least and the greatest of the eigenvalues of the matrix  $R_n^{(m)}$ . By the Schur lemma (see [B]) the matrix  $R_n^{(m)}$  is nonnegative definite. Hence, the eigenvalues  $\lambda_{n,1}^{(m)}$  are nonnegative. For the matrix  $R_n = R_n^{(1)}$  we use the well known decomposition

$$R_n = U_n D_n U_n^T,$$

where

$$D_n = \begin{pmatrix} \lambda_{n,1}^{(1)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_{n,n}^{(1)} \end{pmatrix}$$

is a diagonal matrix with eigenvalues  $\lambda_{n,i}^{(1)}$ ,  $i = 1, \dots, n$ , of  $R_n$  on the main diagonal. The matrix  $U_n = (u_{n,ij})_{1 \leq i,j \leq n}$  is an orthogonal matrix and  $T$  stands for transposition. It follows that

$$(2.1) \quad \rho_{ij} = \sum_{k=1}^n \lambda_{n,k}^{(1)} u_{n,ik} u_{n,jk}, \quad 1 \leq i, j \leq n.$$

Now we can state the following simple result.

LEMMA 2.1. *Fix  $n \geq 1$ . Then the least and the greatest eigenvalues of the matrix  $R_n^{(m)}$  are monotonic with respect to  $m$ , i.e.*

$$(2.2) \quad \lambda_{n,1}^{(m+1)} \geq \lambda_{n,1}^{(m)} \quad \text{and} \quad \lambda_{n,n}^{(m+1)} \leq \lambda_{n,n}^{(m)}, \quad \text{for } m = 1, 2, \dots$$



*Proof.* Since the matrix  $R_n^{(m+1)}$  is symmetric, we have

$$(2.3) \quad \lambda_{n,1}^{(m+1)} = \inf_{\|c\|=1} (R_n^{(m+1)}c, c) = \inf_{\|c\|=1} \sum_{i,j=1}^n \rho_{ij}^{m+1} c_i c_j,$$

where  $c = (c_1, \dots, c_n) \in l_2^n$  and  $l_2^n$  is the  $n$ -dimensional Euclidean space with the scalar product denoted here by  $(\cdot, \cdot)$ . From (2.3) and (2.1) we conclude that for every  $c = (c_1, \dots, c_n) \in l_2^n$  with  $\|c\| = 1$  we have

$$(2.4) \quad \begin{aligned} \sum_{i,j=1}^n \rho_{ij}^{m+1} c_i c_j &= \sum_{i,j=1}^n \rho_{ij}^m \rho_{ij} c_i c_j \\ &= \sum_{i,j=1}^n \rho_{ij}^m \sum_{k=1}^n \lambda_{n,k}^{(1)} u_{n,ik} u_{n,jk} c_i c_j = \sum_{k=1}^n \lambda_{n,k}^{(1)} \left( \sum_{i,j=1}^n \rho_{ij}^m c_i u_{n,ik} c_j u_{n,jk} \right) \\ &\geq \sum_{k=1}^n \lambda_{n,k}^{(1)} \sum_{i=1}^n c_i^2 u_{n,ik}^2 \inf_{\|b\|=1} (R_n^{(m)}b, b) = \lambda_{n,1}^{(m)}, \end{aligned}$$

since

$$\sum_{k=1}^n \lambda_{n,k}^{(1)} \sum_{i=1}^n c_i^2 u_{n,ik}^2 = \sum_{i=1}^n c_i^2 \sum_{k=1}^n \lambda_{n,k}^{(1)} u_{n,ik}^2 = 1$$

by (2.1). Taking the infimum in (2.4) over all  $c = (c_1, \dots, c_n) \in l_2^n$  with  $\|c\| = 1$  we obtain the first inequality of (2.2). The proof of the second one runs similarly. ■

We can now formulate our main result.

**THEOREM 2.1.** *Let  $(X_i, i = 1, 2, \dots)$  be a Gaussian sequence with  $X_i \in N(0, 1)$  for each  $i$  and suppose its correlation matrix  $R = (\rho_{ij})_{i,j \geq 1}$  is the matrix of some operator  $R : l_2 \rightarrow l_2$ . Then for  $f_i \in L^2(\mu)$ ,  $i = 1, 2, \dots$ , and for every  $n \geq 1$  we have*

$$(2.5) \quad \lambda_{\min} \sum_{i=1}^n \text{Var}(f_i(X_i)) \leq \text{Var}\left(\sum_{i=1}^n f_i(X_i)\right) \leq \lambda_{\max} \sum_{i=1}^n \text{Var}(f_i(X_i)),$$

where

$$\lambda_{\min} = \inf_{\|x\|=1} (Rx, x), \quad \lambda_{\max} = \sup_{\|x\|=1} (Rx, x).$$

**REMARK.** Let us point out that the assumption concerning the correlation matrix  $R = (\rho_{ij})_{i,j \geq 1}$  of the sequence  $(X_i, i = 1, 2, \dots)$  is slightly weaker than the hypothesis (1.4). To see this, consider the following example: Let  $(Y_i, i = 1, 2, \dots) \subset N(0, 1)$  be a sequence of independent Gaussian random

variables and define

$$X_1 = aY_1 + \sum_{j=2}^{\infty} Y_j/j, \quad \text{where } a = \sqrt{2 - \pi^2/6},$$

$$X_i = Y_i \quad \text{for } i \geq 2.$$

It follows immediately that the correlation matrix  $R = (\rho_{ij})_{i,j \geq 1}$  of the sequence  $(X_i, i = 1, 2, \dots)$  is the matrix of some linear operator  $R : l_2 \rightarrow l_2$  and the hypothesis (1.4) is not satisfied.

*Proof of Theorem 1.1.* First we prove the left inequality of (2.5). Without loss of generality we assume that  $E(f_i(X_i)) = 0, i = 1, 2, \dots$ . If  $\lambda_{\min} = 0$  then the inequality holds trivially. Assume that  $\lambda_{\min} \neq 0$ . Expanding each  $f_i, i \geq 1$ , with respect to the Hermite basis in  $L^2(\mu)$  we obtain

$$(2.6) \quad f_i = \sum_{k=1}^{\infty} c_{ik}h_k, \quad \|f_i\|_{\mu}^2 = \sum_{k=1}^{\infty} c_{ik}^2, \quad i = 1, 2, \dots$$

From (1.1) and (1.2) and the orthonormality of Hermite polynomials  $\{h_n\}_{n \geq 1} \subset L^2(\mu)$  it follows that

$$(2.7) \quad E[h_n(X_i)h_m(X_j)] = \rho_{ij}^n \delta_m^n, \quad n, m, i, j = 1, 2, \dots,$$

where  $\delta_m^n$  is the Kronecker delta. Combining (2.6) with (2.7) and using Lemma 2.1 we get

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^n f_i(X_i)\right) &= E\left(\sum_{i=1}^n f_i(X_i)\right)^2 \\ &= \lim_{N \rightarrow \infty} E\left(\sum_{i=1}^n \sum_{k=1}^N c_{ik}h_k(X_i)\right)^2 = \lim_{N \rightarrow \infty} E\left(\sum_{k=1}^N \sum_{i=1}^n c_{ik}h_k(X_i)\right)^2 \\ &= \lim_{N \rightarrow \infty} \sum_{k,l=1}^N E\left[\left(\sum_{i=1}^n c_{ik}h_k(X_i)\right)\left(\sum_{j=1}^n c_{jl}h_l(X_j)\right)\right] \\ &= \lim_{N \rightarrow \infty} \sum_{k=1}^N E\left[\sum_{i=1}^n c_{ik}h_k(X_i)\right]^2 = \lim_{N \rightarrow \infty} \sum_{k=1}^N \sum_{i,j=1}^n \rho_{ij}^k c_{ik}c_{jk} \\ &\geq \lim_{N \rightarrow \infty} \sum_{k=1}^N \lambda_{n,1}^{(k)} \sum_{i=1}^n c_{ik}^2 \geq \lim_{N \rightarrow \infty} \sum_{k=1}^N \lambda_{n,1}^{(1)} \sum_{i=1}^n c_{ik}^2 \\ &\geq \lambda_{\min} \sum_{i=1}^n \sum_{k=1}^{\infty} c_{ik}^2 = \lambda_{\min} \sum_{i=1}^n E[f_i(X_i)]^2 = \lambda_{\min} \sum_{i=1}^n \text{Var}(f_i(X_i)). \end{aligned}$$

This proves the left inequality of (2.5). The proof of the right one is similar. ■

REMARK. Let us point out that under the assumptions of Theorem 2.1 the inequality (1.6) holds for all  $f_i \in L^2(\mu)$ ,  $i = 1, 2, \dots$ , with a positive constant  $M$  if and only if the operator  $R : l_2 \rightarrow l_2$  is invertible. ■

Adapting now the methods from [BC1] and [BC2] we can write the following two statements:

LEMMA 2.2 (Borel–Cantelli Lemma). *Let  $(X_i, i = 1, 2, \dots)$  be a Gaussian sequence with  $X_i \in N(0, 1)$  for  $i \geq 1$  and suppose its correlation matrix  $R = (\rho_{ij})_{i,j \geq 1}$  is the matrix of some linear operator  $R : l_2 \rightarrow l_2$ . Then for every sequence of Borel sets  $(A_i, i = 1, 2, \dots)$  such that  $\sum_{i=1}^{\infty} P\{X_i \in A_i\} = \infty$  we have  $P\{X_i \in A_i \text{ i.o.}\} = 1$ . ■*

THEOREM 2.2 (Strong Law of Large Numbers). *Let  $(X_i, i = 1, 2, \dots)$  be a Gaussian sequence with  $X_i \in N(0, 1)$  for  $i \geq 1$  and suppose its correlation matrix  $R = (\rho_{ij})_{i,j \geq 1}$  is the matrix of some linear operator  $R : l_2 \rightarrow l_2$ . Then for  $f \in L^1(\mu)$  we have*

$$\frac{1}{n} \sum_{i=1}^n f(X_i) \xrightarrow[n \rightarrow \infty]{} Ef(X_1) \quad \text{a.s.} \quad \blacksquare$$

### References

- [B] R. Bellman, *Introduction to Matrix Analysis*, McGraw-Hill, New York, 1960.
- [BC1] M. Beška and Z. Ciesielski, *Ergodic type theorems for Gaussian systems*, Proc. Steklov Inst. Math. 248 (2005), 35–40.
- [BC2] —, —, *Gebelein's inequality and its consequences*, in Banach Center Publ. 72, Inst. Math., Polish Acad. Sci., 2006, 11–23.
- [DK] H. Dym and H. P. McKean, *Gaussian Processes, Function Theory and Inverse Spectral Problem*, Academic Press, 1976.
- [G] H. Gebelein, *Das statistische Problem der Korrelation als Variations und Eigenwertproblem und sein Zusammenhang mit der Ausgleichsrechnung*, Z. Angew. Math. Mech. 21 (1941), 364–379.
- [HLP] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge Univ. Press, Cambridge, 1967.
- [V] M. Veraar, *Correlation inequalities and applications to vector-valued Gaussian random variables and fractional Brownian motion*, Potential Anal. 30 (2009), 341–370.

Marek Beška  
 Faculty of Applied Mathematics  
 Gdańsk University of Technology  
 Narutowicza 11/12  
 80-233 Gdańsk, Poland  
 E-mail: beska@mif.pg.gda.pl

