Kybernetika

Jiří Michálek

The Rényi distances of Gaussian measures

Kybernetika, Vol. 35 (1999), No. 3, [333]--352

Persistent URL: http://dml.cz/dmlcz/135291

KYBERNETIKA — VOLUME 35 (1999), NUMBER 3, PAGES 333-352

THE RÉNYI DISTANCES OF GAUSSIAN MEASURES

JIŘÍ MICHÁLEK

The author in the paper evaluates the Rényi distances between two Gaussian measures using properties of nuclear operators and expresses the formula for the asymptotic rate of the Rényi distances of stationary Gaussian measures by the corresponding spectral density functions in a general case.

INTRODUCTION

This paper deals with the calculation of the Rényi distances of Gaussian measures defined by random processes with a continuous time. In the case of stationary measures the asymptotic rate of the Rényi distances is evaluated by the use of the corresponding spectral measures in a very general case.

The Rényi distances of probability measures are important information-theoretical measures of similarity between probabilities, often used in mathematical statistics. The Rényi distances are parametrized by a real parameter a. For us the case $a \in \langle 0,1\rangle$ will be of the main interest. The limit cases $a=0$ and $a=1$ are very closely connected with the Kullback–Leibler divergences. Generally speaking, the Rényi distances are derived from the Hellinger integrals. They are defined as follows

$$R_a(P|Q) = \frac{\ln H_a(P|Q)}{a(a-1)}$$

for $a \neq 0,\ a \neq 1$, where

$$H_a(P|Q) = \mathsf{E}_\mu\left\{\left(\frac{p}{q}\right)^a\right\}$$

is the Hellinger integral with $p = \frac{\mathrm{d}P}{\mathrm{d}\mu}$, $q = \frac{\mathrm{d}Q}{\mathrm{d}\mu}$ and μ is an arbitrary σ-finite dominating measure. For $a=1$ we put on the basis of continuity

$$R_1(P|Q) = \lim_{a\uparrow 1} R_a(P,Q) = I_1(P|Q),$$

for $a=0$

$$R_0(P|Q) = \lim_{a\downarrow 0} R_a(P,Q) = I_0(P|Q),$$

[1]This work is supported by the Grant Agency of the Czech Republic through Grant No. 201/96/0415 and the Grant Agency of the Academy of Sciences of the Czech Republic through Grant No. K 1075601.

where

$$I_1(P|Q) = \int p \ln \frac{p}{q} \, \mathrm{d}\mu, \qquad I_0(P|Q) = \int q \ln \frac{q}{p} \, \mathrm{d}\mu = I_1(Q|P).$$

Evidently,

$$I_1(P|Q) = \mathsf{E}_P \left\{ \ln \frac{p}{q} \right\}, \qquad I_0(P|Q) = \mathsf{E}_Q \left\{ \ln \frac{q}{p} \right\}.$$

There are many papers and books describing properties of the Rényi distances and other f-divergences. We will quote here the monograph due to Vajda [14].

Generally speaking, the Rényi distance is meaningful for each real a if we accept the value $+\infty$, too. The Rényi distances are very exceptionally metrics because they do not fulfill the triangular inequality, in general.

1. ABSOLUTE CONTINUITY OF GAUSSIAN MEASURES

A Gaussian probability measure is in a unique way determined by its mean and covariance function. A main advantage of Gaussian measures with respect to their absolute continuity is the fact that two particular cases can be analyzed separately. The first case is that both measures have different means but the same covariance, the other one is that means are zero and covariance functions are different. The case with different means is more transparent than the other one and some results concerning the Rényi distances in this situation can be found in Michálek [8]. In that paper the case with different covariances is solved too, but not in full generality. The existence of the Rényi distances of Gaussian measures was proved in Michálek [8] under the assumption of a strong equivalence of measures. This notion will be mentioned later.

Let two covariance functions $R(\cdot,\cdot)$ and $S(\cdot,\cdot)$ be given on $\langle 0,T\rangle$. Let P_T, Q_T be the corresponding Gaussian measures with vanishing means and covariances $R(\cdot,\cdot)$, $S(\cdot,\cdot)$ respectively, i.e.

$$\mathsf{E}_{P_T}\{x(s)\,x(t)\} = R(s,t), \quad \mathsf{E}_{Q_T}\{x(s)\,x(t)\} = S(s,t) \quad \text{respectively.}$$

The question of absolute continuity of Gaussian measures is in detail answered, e.g. in the monograph of Rozanov [13]. For a better understanding it is necessary to introduce the following basic notions.

Let L_T be a linear hull over all the observations $x(t)$, $t \in \langle 0,T\rangle$, i.e.

$$\xi \in L_T \Longleftrightarrow \xi = \sum_1^n c_i\, x(t_i), \quad t_i \in \langle 0,T\rangle.$$

As we have two Gaussian measures P_T, Q_T we must consider also two closures of L_T with respect to the convergence in the quadratic mean. In general, the closures $\overline{L}_T(P_T)$, $\overline{L}_T(Q_T)$ can be different although the linear hull L_T is everywhere dense in both topologies. It is easy to show that the case $\overline{L}(P_T) \neq \overline{L}_T(Q_T)$ leads to the orthogonality of P_T and Q_T, for details see Rozanov [13]. Then a necessary condition for the equivalence between P_T and Q_T sounds

$$\overline{L}_T = \overline{L}_T(P_T) = \overline{L}_T(Q_T).$$

This means that both the topologies generated by the convergences with respect the quadratic mean must be equivalent, i. e. that the relation

$$0 < C_1 \leq \frac{\|\xi\|_{P_T}}{\|\xi\|_{Q_T}} \leq C_2 < +\infty$$

must be valid on L_T. If we wish to study conditions for the equivalence between P_T and Q_T we must begin with two Hilbert spaces, namely

$$(\overline{L}_T,\, \|\cdot\|_{P_T}), \quad (\overline{L}_T,\, \|\cdot\|_{Q_T}).$$

At this moment we will define an operator A_T on $L_T \subset \overline{L}_T(Q_T)$, i. e. $\mathrm{Dom}(A_T) = L_T$ and $\mathrm{Range}(A_T) = L_T \subset \overline{L}_T(P_T)$. In other words, A_T is a linear mapping from $(\overline{L}_T,\, \|\cdot\|_{Q_T})$ into $(\overline{L}_T,\, \|\cdot\|_{P_T})$ defined by

$$A_T\, \xi = \xi$$

for every $\xi \in L_T$. A_T must have the adjoint operator A^* defined in the unique way by

$$\langle \eta,\, A_T\, \xi\rangle_{P_T} = \langle A_T^*\, \eta,\, \xi\rangle_{Q_T}.$$

Surely, $\mathrm{Dom}(A_T^*) \subset \overline{L}_T(P_T)$, $\mathrm{Range}(A_T^*) \subset \overline{L}_T(Q_T)$ and $\eta \in \mathrm{Dom}(A_T^*)$ if and only if $\langle \eta,\, A_T\, \xi\rangle_{P_T}$ is a bounded linear functional on $(\overline{L}_T,\, \|\cdot\|_{Q_T})$. For each couple $\xi,\, \eta \in L_T$ the operator $A_T^*\, A_T$ is then well defined because

$$\langle A_T^*\, A_T\, \eta,\, \xi\rangle_{Q_T} = \langle A_T\, \eta,\, A_T\, \xi\rangle_{P_T} = \langle \eta,\, \xi\rangle_{P_T}.$$

Hence this operator can be extended onto the whole space $\overline{L}_T$. Let us denote it by B_T.

The operator B_T can be defined by a simpler way, namely using the relation connecting the scalar products on $\overline{L}_T$. Thanks to the equivalence between the norms $\|\cdot\|_{Q_T}$, $\|\cdot\|_{P_T}$ the scalar product $\langle \eta,\, \xi\rangle_{P_T}$ must be a linear bounded functional on the space $(\overline{L}_T,\, \|\cdot\|_{Q_T})$, i. e.

$$\langle B_T\, \eta,\, \xi\rangle_{Q_T} = \langle \eta,\, \xi\rangle_{P_T}.$$

We immediately see that B_T is positive because

$$\langle B_T\, \eta,\, \eta\rangle_{Q_T} = \langle \eta,\, \eta\rangle_{P_T} \geq 0.$$

It is well known (see Rozanov [13]) that the measures P_T, Q_T are mutually equivalent if and only if the operator $I - B_T$, where I is the identity on $\overline{L}_T$, is of the Hilbert–Schmidt type and B_T^{-1} exists and is bounded. This condition of existence and boundedness of B_T^{-1} is equivalent to that of equivalence between $\langle\cdot,\cdot\rangle_{P_T}$ and $\langle\cdot,\cdot\rangle_{Q_T}$, which implies the coincidence of $\overline{L}_T(P_T)$ and $\overline{L}_T(Q_T)$. In other words, the operator $I - B_T$ must not have 1 as its eigenvalue.

As known from the theory of Hilbert–Schmidt operators such an operator is totally continuous and it possesses at most a countable number of proper values $\{\alpha_i\}_{i=1}^{\infty}$ that satisfy

$$\sum_{1}^{\infty} \alpha_i^2 < +\infty.$$

In our case it means for the operator $I - B_T$ that

$$\sum_{1}^{\infty}(1-\alpha_i)^2 < \infty, \tag{1}$$

where $\{\alpha_i\}_{i=1}^{\infty}$ are the proper values of the operator B_T. The convergence of the series (1) immediately gives, as $\alpha_i \to 1$ when $i \to \infty$, that the series

$$\sum_{i=1}^{\infty}\left(1-\frac{1}{\alpha_i}\right)^2$$

is convergent, too. But this fact says that the operator $I - B_T^{-1}$ is of the Hilbert–Schmidt type, too and this property proves the mutual equivalence of the measures P_T and Q_T. If a stronger condition about the convergence of series (1) is valid, namely if

$$\sum_{1}^{\infty}|1-\alpha_i| < \infty,$$

then in accordance with Hájek [4] we speak about the strong equivalence of Gaussian measures. Then the series

$$\sum_{1}^{\infty}\left|1-\frac{1}{\alpha_i}\right|$$

is also convergent and hence this type of equivalence of probability measures is mutual, too. It is important to notice that all the proper values of B_T must be strictly positive so that their reciprocals $1/\alpha_i$ are proper values belonging to the inverse operator B_T^{-1}. The condition of strong equivalence says that the operators $I - B_T$ and $I - B_T^{-1}$ are nuclear because their traces are finite.

Let $\{\xi_i\}_{i=1}^{\infty}$ be the proper vectors of B_T in the space $\overline{L}_T$. Then the corresponding Radon–Nikodym derivative of P_T with respect to Q_T equals

$$\frac{\mathrm{d}P_T}{\mathrm{d}Q_T}(\omega) = \lim_{n\to\infty}\left\{\left(\prod_{1}^{n}\alpha_i\right)^{-1/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\frac{1-\alpha_i}{\alpha_i}\,\xi_i^2(\omega)\right\}\right\}$$

in the sense a.s. $[Q_T]$, for details see Hájek [4]. If the measures P_T, Q_T are strongly equivalent then the infinite product $\prod_1^{\infty}\alpha_i$ is convergent because it is nothing but the Fredholm determinant of the operator $I - B_T$ and the corresponding Radon–Nikodym derivative can be expressed in a more closed form, namely

$$\frac{\mathrm{d}P_T}{\mathrm{d}Q_T}(\omega) = \left(\prod_{1}^{\infty}\alpha_i\right)^{-1/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^{\infty}\frac{1-\alpha_i}{\alpha_i}\,\xi_i^2(\omega)\right\}.$$

Some conditions ensuring the strong equivalence of Gaussian measures can be found in Hájek [4]. But in general, the Fredholm determinant need not exist and we can use the properties following only from the convergence of the series

$$\sum_{1}^{\infty}(1-\alpha_i)^2 < +\infty.$$

As was said earlier this convergence ensures the finiteness of the series

$$\sum_1^\infty \frac{(1-\alpha_i)^2}{\alpha_i},$$

because $\alpha_i \to 1$ as $i \to \infty$. One easily sees that for each $n \in \mathcal{N}$

$$\begin{aligned}\sum_{i=1}^n \frac{(1-\alpha_i)^2}{\alpha_i} &= \sum_{i=1}^n \left(\frac{1}{\alpha_i} + \alpha_i - 2\right) = \sum_{i=1}^n \left(\frac{1}{\alpha_i} - 2 + \alpha_i - \ln\alpha_i - \ln\frac{1}{\alpha_i}\right) \\ &= \sum_{i=1}^n (\alpha_i - \ln\alpha_i - 1) + \sum_{i=1}^n \left(\frac{1}{\alpha_i} - \ln\frac{1}{\alpha_i} - 1\right).\end{aligned}$$

As $\ln x < x - 1$ for every positive $x \neq 1$, the last series have positive terms.Since the original series having positive numbers too was split into two positive series, they must be convergent too. Hence the proper values of B_T must satisfy

$$\sum_{i=1}^\infty \alpha_i - \ln\alpha_i - 1 < +\infty, \qquad \sum_{i=1}^\infty \frac{1}{\alpha_i} - \ln\frac{1}{\alpha_i} - 1 < +\infty.$$

Using these facts, the Radon–Nikodym derivative can be rewritten into the form

$$\frac{\mathrm{d}P_T}{\mathrm{d}Q_T}(\omega) = \lim_{n\to\infty}\left\{\left(\prod_1^n \alpha_i\, e^{1-\alpha_i}\right)^{-1/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^n \frac{1-\alpha_i}{\alpha_i}\,\xi_i^2(\omega) - (1-\alpha_i)\right\}\right\}.$$

Since

$$\ln\prod_{i=1}^n \alpha_i\, e^{1-\alpha_i} = \sum_{i=1}^n \ln(\alpha_i\, e^{1-\alpha_i}) = -\sum_{i=1}^n \alpha_i - \ln\alpha_i - 1,$$

it follows the convergence of the infinite product

$$\prod_1^\infty \alpha_i\, e^{1-\alpha_i}.$$

Due to the properties of martingales the other part of the Radon–Nikodym derivative must be also convergent and hence we finally reach the following expression

$$\frac{\mathrm{d}P_T}{\mathrm{d}Q_T}(\omega) = D_1^{-1/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^\infty \frac{1-\alpha_i}{\alpha_i}\,\xi_i^2(\omega) - (1-\alpha_i)\right\}.$$

The quantity $D_1 = \prod_{i=1}^\infty (1-(1-\alpha_i))\, e^{1-\alpha_i}$ is called the regularized Fredholm determinant of the operator $I - B_T$, see, e. g. Gohberg and Krein [1]. This form of a determinant substitutes the Fredholm determinant for operators of the Hilbert–Schmidt type.

Now we will present several notes about the proper vectors $\{\xi_i\}_{i=1}^\infty$ of the operator B_T. These vectors belong to the space $\overline{L}_T$ and are Gaussian with respect to both the probability measures P_T and Q_T. Thanks to the orthogonality with respect to both the scalar products, they are independent and satisfy

$$\xi_i \underset{P_T}{\sim} N(0, \alpha_i), \qquad \xi_i \underset{Q_T}{\sim} N(0, 1).$$

These facts follow from the spectral decomposition of the operator $I - B_T$ in $(\overline{L}_T, \|\cdot\|_{Q_T})$

$$(I - B_T)\,\eta = \sum_{i=1}^{\infty}(1-\alpha_i)\,\langle \eta, \xi_i\rangle\,\xi_i.$$

The construction of the proper vectors ξ_i, $i = 1, 2, \dots$ is enabled by the simultaneous "diagonalization" of covariance functions or covariance operators generated by them, which is a generalization of a famous fact of matrix calculation, for details, see Kadota [6].

2. THE RÉNYI DISTANCES

Now we are ready to evaluate the Rényi distances in the case of two Gaussian measures P_T, Q_T on $\langle 0, T\rangle$ with different covariance functions. We start with the expression of the Radon–Nikodym derivative using the proper values of $I - B_T$. We have

$$\ln \frac{\mathrm{d}P_T}{\mathrm{d}Q_T} = -\frac{1}{2}\ln D_1 - \frac{1}{2}\sum_{i=1}^{\infty}\left\{\frac{1-\alpha_i}{\alpha_i}\,\xi_i^2 - (1-\alpha_i)\right\}.$$

Hence

$$R_1(P_T, Q_T) = \mathsf{E}_{P_T}\left\{\ln\frac{\mathrm{d}P_T}{\mathrm{d}Q_T}\right\} = -\frac{1}{2}\ln D_1 = \frac{1}{2}\sum_{i=1}^{\infty}(\alpha_i - \ln\alpha_i - 1),$$

because $\mathsf{E}_{P_T}\{\xi_i^2\} = \alpha_i$ and the series in the expression of $\ln\frac{\mathrm{d}P_T}{\mathrm{d}Q_T}$ is convergent in the quadratic mean, too.

In a quite analogous way we can obtain

$$R_0(P_T|Q_T) = R_1(Q_T|P_T) = \frac{1}{2}\sum_{i=1}^{\infty}\left(\frac{1}{\alpha_i} - \ln\frac{1}{\alpha_i} - 1\right)$$

which is the regularized Fredholm determinant of the operator $I - B_T^{-1}$.

Now we will express $R_a(P_T|Q_T)$ for $a \in (0,1)$. We must calculate the expected value with respect to the measure Q_T

$$\begin{aligned}
\mathsf{E}_{Q_T}\left\{\left(\frac{\mathrm{d}P_T}{\mathrm{d}Q_T}\right)^a\right\} &= D_1^{-a/2}\,\mathsf{E}_{Q_T}\left\{\exp\left\{-\frac{a}{2}\sum_{j=1}^{\infty}\left(\frac{1-\alpha_i}{\alpha_i}\,\xi_i^2 - 1 + \alpha_i\right)\right\}\right\}\\
&= D_1^{-a/2}\mathsf{E}_{Q_T}\left\{\prod_{i=1}^{\infty}\left(\exp\left\{-\frac{a}{2}\,\frac{1-\alpha_i}{\alpha_i}\,\xi_i^2\right\}\exp\left\{\frac{a}{2}(1-\alpha_i)\right\}\right)\right\}\\
&= D_1^{-a/2}\prod_{i=1}^{\infty}\exp\left\{\frac{a}{2}(1-\alpha_i)\right\}\mathsf{E}_{Q_T}\left\{\exp\left\{-\frac{a}{2}\,\frac{1-\alpha_i}{\alpha_i}\,\xi_i^2\right\}\right\}\\
&= \prod_{i=1}^{\infty}\left(\alpha_i\, e^{1-\alpha_i}\right)^{-a/2}\prod_{i=1}^{\infty} e^{-\frac{a}{2}(\alpha_i-1)}\sqrt{\frac{\alpha_i}{\alpha_i + (1-\alpha_i)\,a}}
\end{aligned}$$

because

$$\mathsf{E}_{Q_T}\left\{e^{-\frac{a(1-\alpha_i)\xi_i^2}{2\alpha_i}}\right\} = \sqrt{\frac{\alpha_i}{\alpha_i + (1-\alpha_i)\,a}}.$$

Due to the convergence of the regularized Fredholm determinants of $I - B_T$ and of $I - B_T^{-1}$ we finally obtain

$$\mathsf{E}_{Q_T}\left\{\left(\frac{\mathrm{d}P_T}{\mathrm{d}Q_T}\right)^a\right\} = \exp\left\{-\frac{1}{2}\sum_{i=1}^{\infty}\ln\left(a\,\alpha_i^{a-1} + (1-a)\,\alpha_i^a\right)\right\}.$$

Using the last formula expressing the corresponding Hellinger integral the Rényi distance between P_T and Q_T equals

$$R_a(P_T|Q_T) = \frac{1}{2(1-a)\,a}\sum_{i=1}^{\infty}\ln\left(a\,\alpha_i^{a-1} + (1-a)\,\alpha_i^a\right),$$

where $\{\alpha_i\}_{i=1}^{\infty}$ are proper values of the operator B_T.

One immediately sees that for each $a \in (0,1)$

$$\ln\left(a\,x^{a-1} + (1-a)\,x^a\right) > 0$$

for each $x > 0,\ x \neq 1$, as

$$\ln\left(a\,x^{a-1} + (1-a)\,x^a\right) > a\ln x^{a-1} + (1-a)\ln x^a = 0.$$

Hence the infinite series standing in the formula for $R_a(P_T|Q_T)$ consists of positive members and its sum is always meaningful if we accept the value $+\infty$, too. Let us prove that in the case of absolute continuity of the pair P_T, Q_T the Rényi distance is finite for each $a \in \langle 0,1\rangle$.

Theorem 1. Let two covariance functions $R(\cdot,\cdot)$ and $S(\cdot,\cdot)$ be given on $\langle 0,T\rangle$. Let P_T, Q_T be the corresponding Gaussian measures with vanishing means and covariances $R(\cdot,\cdot)$, $S(\cdot,\cdot)$, respectively. The measures P_T, Q_T are mutually absolute continuous when there exists at least one $a_0 \in \langle 0,1\rangle$ with

$$R_{a_0}(P_T|Q_T) < +\infty.$$

Then for others $a \in \langle 0,1\rangle$ the corresponding $R_a(P_T|Q_T)$ exists and is finite, too. The converse is also true, i.e. P_T, Q_T are absolutely continuous then $R_a(P_T|Q_T) < \infty$ for all $a \in \langle 0,1\rangle$.

Proof. Let us denote by

$$\varphi_a(x) = \frac{\ln(a\,x^{a-1} + (1-a)\,x^a)}{a(1-a)}$$

for $a \in (0,1)$. We know that $\varphi_a(x) > 0$ for $x > 0,\ x \neq 1$, $\varphi_a(1) = 0$. Let us calculate $\varphi_a'(1)$, $\varphi_a''(1)$. After it we immediately see that in a neighbourhood of the point 1

$$\varphi_a(x) = \frac{1}{2}(x-1)^2 + o(|x-1|^2).$$

This fact gives the following conclusion: when the series $\sum_{i=1}^{\infty}(1-\lambda_i)^2$ is convergent then the series $\sum_{i=1}^{\infty}\varphi_a(\lambda_i)$ is also convergent for each $a \in (0,1)$ and for $a = 0$, $a = 1$ as was proved above. Now we prove the opposite conclusion. If there exists $a_0 \in \langle 0,1\rangle$ such that the series

$$\sum_{i=1}^{\infty}\varphi_{a_0}(\lambda_i) < +\infty$$

then $\sum_{i=1}^{\infty}(1-\lambda_i)^2$ is convergent, too and the operator $I - B_T$ is of the Hilbert–Schmidt type. □

Remark. The value of $R_a(P_T|Q_T)$ can be expressed in a somewhat different form, namely using directly the operator B_T. The operator B_T is bounded with the spectral decomposition

$$B_T\,\eta = \sum_{i=1}^{\infty}\alpha_i\langle\eta,\,\xi_i\rangle\,\xi_i,$$

where $\sup_{i\in\mathcal{N}}\alpha_i < +\infty$ because $\lim_{i\to\infty}\alpha_i = 1$. Hence we can consider the operator

$$\varphi_a(B_T)\,\eta = \sum_{i=1}^{\infty}\varphi_a(\alpha_i)\,\langle\eta,\,\xi_i\rangle\,\xi_i.$$

The operator $\varphi_a(B_T)$ is bounded if and only if

$$\sup_{i\in\mathcal{N}}\varphi_a(\alpha_i) < +\infty.$$

But, we will prove even more using the fact that the series $\sum_{i=1}^{\infty}\varphi_a(\lambda_i)$ is convergent if P_T, Q_T are equivalent. Then we can assert that the operator $\varphi_a(B_T)$ is even nuclear and

$$R_a(P_T,\,Q_T) = \frac{1}{2(1-a)\,a}\,\mathrm{tr}(\varphi_a(B_T)),$$

where tr means the trace of an operator. Further, we can also prove that there exists a limit in the nuclear norm

$$\lim_{a\nearrow 1}\frac{I - a\,B_T^{a-1} - (1-a)\,B_T^{a}}{a(a-1)} = B_T - \ln B_T - I$$

and symmetrically

$$\lim_{a\searrow 0}\frac{I - a\,B_T^{a-1} - (1-a)\,B_T^{a}}{a(a-1)} = B_T^{-1} - \ln B_T^{-1} - I.$$

These facts are based on the following

Theorem 2. For each $a \in \langle 0,1\rangle$ the operator

$$\psi_a(B_T) = a\, B_T^{a-1} + (1-a)\, B_T^a - I$$

is positive and nuclear.

Proof. If $\{\alpha_i\}_{i=1}^{\infty}$ are proper values of B_T then the operator $\psi_a(B_T)$ has its proper values of the form

$$a\,\alpha_i^{a-1} + (1-a)\,\alpha_i^a.$$

Using the inequality $x - 1 \geq \ln x$ one can state that with $\alpha_i > 0$, $i = 1, 2, \ldots$ and $\alpha_i \neq 1$, $i = 1, 2, \ldots$

$$a\,\alpha_i^{a-1} + (1-a)\,\alpha_i^a > 0,$$

which proves the pcsitiveness of $\psi_a(B_T)$. To prove the property of a finite trace we must study the infinite series

$$\sum_{i=1}^{\infty} \left[a\,\alpha_i^{a-1} + (1-a)\,\alpha_i^a - 1\right].$$

The convergence of this series immediately follows from the local behaviour of the function $\psi_a(\cdot)$ in a neighbourhood of 1 because $\alpha_i \to 1$ as $i \to \infty$. By a simple calculation we obtain the relation

$$\psi_a(x) = \frac{a(1-a)}{2}\,(x-1)^2 + o(|x-1|^2).$$

Due to the absolute continuity of P_T and Q_T the series

$$\sum_{1}^{\infty}(1-\alpha_i)^2$$

is convergent and hence the series

$$\sum_{i=1}^{\infty} \psi_a(\alpha_i)$$

is convergent, too. This fact shows that the operator $\psi_a(B_T)$ is nuclear for each $a \in \langle 0,1\rangle$. □

Till now we have considered evaluating $R_a(P_T|Q_T)$ for $a \in \langle 0,1\rangle$ only. Now we will make some remarks about $a \notin \langle 0,1\rangle$. First, $a < 0$. In order to be able to speak about the finiteness of $R_a(P_T|Q_T)$ it is necessary to ensure the existence of the mean value

$$\mathsf{E}_{Q_T}\left\{e^{-\frac{a}{2}\frac{(1-\alpha_i)}{\alpha_i}\xi_i^2}\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{a(1-\alpha_i)}{\alpha_i}+1\right)x^2}\,\mathrm{d}x$$

figuring in the formula for $\mathsf{E}_{Q_T}\left\{\left(\frac{\mathrm{d}P_T}{\mathrm{d}Q_T}\right)^a\right\} = H_a(P_T|Q_T)$.

The integral on the right hand side is convergent if and only if

$$\frac{a(1-\alpha_i)}{\alpha_i} + 1 > 0.$$

As $\alpha_i > 0$, then this inequality is equivalent to another one, namely

$$a(1-\alpha_i) + \alpha_i > 0$$

for each $i = 1, 2, 3, \ldots$ But, it means that the inequality

$$\alpha_i > -\frac{a}{1-a}$$

must be valid simultaneously for each $i = 1, 2, \ldots$. When there exists $\alpha_i < 1$ and as the sequence $\{\alpha_i\}$ has the only limit point equal to one, there exists an index $i_0 \in \mathcal{N}$ such that

$$0 < \alpha_{i_0} = \min_{i \in \mathcal{N}} \alpha_i.$$

As $\alpha_{i_0} > 0$, one can find $a_0 < 0$ such that

$$\alpha_i > \frac{-a_0}{1-a_0}$$

for each $i = 1, 2, \ldots$. In the case of $\alpha_i > 1$ for every $i = 1, 2, \ldots$ the inequality $\alpha_i > -\frac{a}{1-a}$ holds for every $a < 0$. This choice is, of course, dependent on the character of the operator $I - B_T$. Therefore we can assert that in the case of absolute continuity between P_T and Q_T there exists a left neighbourhood of 0, let us say $(a_0, 0)$ where the Rényi distance is still finite.

A similar situation occurs in the other case with a right hand neighbourhood of 1. For finiteness of $R_a(P_T, Q_T)$ we must demand that for each $i = 1, 2, \ldots$

$$\alpha_i < \frac{a}{a-1}.$$

As the sequence $\{\alpha_i\}_{i=1}^{\infty}$ is convergent to 1 and this is the only limit point of this sequence there exists $\alpha_{i_1} = \max_{i \in \mathcal{N}} \alpha_i$ under the assumption that at least one $\alpha_i > 1$. But then there exists such an $a_1 > 1$ that

$$\alpha_i < \frac{a_1}{a_1 - 1}$$

for each $i = 1, 2, \ldots$, because $\lim_{a \searrow 1} \frac{a}{a-1} = +\infty$. When $\alpha_i < 1$ for every $i = 1, 2, \ldots$ the previous inequality is valid for each $a_1 > 1$.

We have just proved that for every absolutely continuous pair of Gaussian measures P_T, Q_T there exists a right neighbourhood $(1, a_1)$ of the point 1 where the Rényi distance is still finite. This result will be summarized in the following

Theorem 3. Let two covariance functions $R(\cdot,\cdot)$ and $S(\cdot,\cdot)$ be given on $\langle 0,T\rangle$. Let P_T, Q_T be the corresponding Gaussian measures with vanishing means and covariances $R(\cdot,\cdot)$, $S(\cdot,\cdot)$, respectively. If $P_T \sim Q_T$ there exist two positive real numbers a_0, a_1 depending on the pair $(P_T,\, Q_T)$ such that the corresponding Rényi distance $R_a(P_T,\, Q_T)$ is finite in an open interval $(-a_0,\, 1+a_1)$.

Proof. See above. □

The interval $(-a_0,\, 1+a_1)$ described in the Theorem 2 is maximal in the following sence:
$$\lim_{a\searrow -a_0} R_a(P_T,\, Q_T) = +\infty \quad \text{and} \quad \lim_{a\nearrow 1+a_1} R_a(P_T,\, Q_T) = +\infty.$$

The next theorem uses the behaviour of the Hellinger integral $H_a(P_T,\, Q_T)$ in $\langle 0,1\rangle$ and leads to an interesting property of the proper values of the operator $I - B_T$.

Theorem 4. Let the Gaussian measures P_T, Q_T considered in Theorem 3 be strongly equivalent. Then $\|B_T\| > 1$ and also $\|B_T^{-1}\| > 1$.

Proof. On the basis of the previous result we know that in the case of equivalence between P_T and Q_T there exists finite $R_a(P_T|Q_T)$ in $\langle 0,1\rangle$ and hence the Hellinger integral exists also because the sum
$$\sum_{i=1}^{\infty} \ln\left(a\,\alpha_i^{a-1} + (1-a)\,\alpha_i^a\right)$$
is convergent. At the first sight $H_0(P_T|Q_T) = H_1(P_T|Q_T) = 1$. At this moment, we can use a result given in Vajda [14] about the existence of the derivative of any order. Hence the function $H_a(P_T,\, Q_T)$ in the parameter a is differentiable in $(0,1)$ and we can calculate its derivative. We will immediately see why the strong equivalence for the calculation of this derivative must be assumed. Namely,
$$\frac{\partial}{\partial a} H_a(P_T|Q_T) = H_a(P_T|Q_T)\cdot \frac{1}{2}\sum_{i=1}^{\infty} \frac{[1-\alpha_i]\,[\alpha_i + a(a-1)]}{\alpha_i[a + (1-a)\,\alpha_i]}.$$
From this formula we see why we demand the strong equivalence between P_T, Q_T. The convergence
$$\sum_{i=1}^{\infty} |1-\alpha_i| < +\infty$$
ensures the existence of an absolutely convergent majorant and thus we can interchange the infinite summation and differentiation. Now, let us suppose that $\sup_{i\in\mathcal{N}} \alpha_i \le 1$. Then $\alpha_i < 1$ for each $i \in \mathcal{N}$ because 1 is not a proper value of B_T and $\|B_T\| \le 1$ would be. But in such a case $1-\alpha_i > 0$ and the derivative of $H_a(P_T|Q_T)$ would be positive, i. e. the function $H_a(P_T,\, Q_T)$ would be in $(0,1)$ strictly increasing. This of course contradicts the fact $H_1(P_T,\, Q_T) = 1 = H_0(P_T,\, Q_T)$. Similarly we can continue under the assumption $\|B_T^{-1}\| \le 1$. As B_T^{-1} has the proper values $\frac{1}{\alpha_i}$, we would reach the $\alpha_i > 1$ for each $i \in \mathcal{N}$ and the derivative $H_a'(P_T|Q_T)$ would be negative in $(0,1)$. We would come to a contradiction again. □

Remark. The previous theorem says that the proper values of $I - B_T$ by the strong equivalence between P_T and Q_T must be situated on both sides of the limit point 0.

In the next part of the paper we will study the asymptotic behaviour of $R_a(P_T, Q_T)$ when $T \to \infty$ under the stationarity of P_T and Q_T. For this reason we must mention in detail some notions and relations, which concern the asymptotic behaviour of proper values of Toeplitz operators. Here we use mainly results given in Grenander and Szegő [3].

3. TOEPLITZ OPERATORS

The Toeplitz matrices play a very important role in the theory of weakly stationary random sequences because they present covariance matrices. The Toeplitz operators are in a certain sense a generalization of the notion of Toeplitz matrices in the case of a continuous time.

Let $(X, \mathcal{X}, \mu)$ and (S, σ, ν) be two spaces with measures. We do not exclude the case $X = S$ and $\mu = \nu$. Let a kernel $\varphi(x,s)$ be such that the integral operator given on $X \times S$

$$(T\,u)\,(x) = \int_S \varphi(x,s)\,u(s)\,\mathrm{d}\nu(s)$$

is defined for each $u(\cdot) \in L_2(S, \sigma, \nu)$. We demand $T u(\cdot) \in L_2(X, \mathcal{X}, \mu)$ and T preserves a norm, i.e.

$$\|T\,u\|_X = \|u\|_S.$$

If the range of the operator T is everywhere dense in $L_2(X)$, then the operator T can be enlarged into a unitary mapping $T : L_2(S) \to L_2(X)$. Under fulfilling these conditions the kernel $\varphi(\cdot,\cdot)$ is called an orthogonal kernel.

Let $f(\cdot)$ be an integrable function on $(X, \mathcal{X})$, let λ be a real number and let us define

$$e_\lambda = \{x : f(x) \le \lambda\} = f^{-1}\{(-\infty, \lambda\rangle\})$$

$e_\lambda \in \mathcal{X}$ for each $\lambda \in R_1$ and for each $u(\cdot) \in L_2(S)$ let us define a decomposition

$$(T\,u)\,(x) = (T\,u)\,(x)\,\psi_{e_\lambda}(x) + (T\,u)\,(x)\,(1 - \psi_{e_\lambda}(x)),$$

where $\psi_{e_\lambda}(x) = 1$ for each $x \in e_\lambda$ and $\psi_{e_\lambda}(x) = 0$ otherwise. As T is unitary, there exists T^{-1} and let us put

$$\mathsf{E}_\lambda\,u(\cdot) = T^{-1}((T\,u)\,\psi_{e_\lambda}) = T^{-1}\,\psi_{e_\lambda} T,$$

where the transformation ψ_{e_λ} means a linear mapping presented by multiplying by the function $\psi_{e_\lambda}(\cdot)$ on the space $L_2(X, \mathcal{X}, \mu)$. One can prove that the system of operators $\{\mathsf{E}_\lambda\}$ on the space $L_2(S, \sigma, \nu)$ forms a resolution of identity, i.e. $\{\mathsf{E}_\lambda\}$ is a collection of projectors

$$\mathsf{E}_\lambda\,\mathsf{E}_\mu = \mathsf{E}_\mu\,\mathsf{E}_\lambda = \mathsf{E}_{\min(\lambda,\mu)}$$

and

$$\lim_{\lambda\to+\infty} \mathsf{E}_\lambda = I, \qquad \lim_{\lambda\to-\infty} \mathsf{E}_\lambda = 0.$$

Using this resolution we can define a new operator K, namely

$$K = \int_{-\infty}^{+\infty} \lambda \, \mathrm{d}\mathsf{E}_\lambda.$$

If we prove that the definition domain of K will be dense everywhere in $L_2(S)$ then the operator K will be well defined. In general, K need not be bounded. If the generating function $f(\cdot)$ is bounded then K is also bounded. The definition domain $\mathrm{Dom}(K)$ is determined by all $u \in L_2(S)$ for which the following integral exists

$$\begin{aligned} \|K\,u\|_S^2 &= \int_{-\infty}^{+\infty} \lambda^2 \, \mathrm{d}\langle u,\, \mathsf{E}_\lambda u\rangle = \int_{-\infty}^{+\infty} \lambda^2 \, \mathrm{d} \int_{e_\lambda} |(T\,u)\,(x)|^2 \, \mathrm{d}\mu(x) \\ &= \int_{-\infty}^{+\infty} |(T\,u)\,(x)|^2 \, f^2(x) \, \mathrm{d}\mu(x). \end{aligned}$$

We will consider the case $f(\cdot) \geq 0$ only. Then obviously, the operator K is positive and its definition domain contains all the functions $u(\cdot) \in L_2(S)$ for which $(T\,u)\,(\cdot)$ are bounded on X a.e. $[\mu]$ and $\mathcal{X}$-measurable. Then the operator K will be self-adjoint, but unbounded in general. The bilinear form $\langle Ku,\, v\rangle_S$ on $L_2(S)$ can be expressed as

$$\begin{aligned} \langle Ku,\, v\rangle_S &= \int_{-\infty}^{+\infty} \lambda \, \mathrm{d}\langle \mathsf{E}_\lambda U,\, v\rangle = \int_{-\infty}^{+\infty} \lambda \, \mathrm{d} \int_{e_\lambda} T\,u \overline{T\,u} \, \mathrm{d}\mu(x) \\ &= \int_X T\,u \overline{T\,v} \, f(x) \, \mathrm{d}\mu(x). \end{aligned}$$

As

$$T\,u = \int_S \varphi(x,s)\,u(s)\,\mathrm{d}\nu(s), \qquad T\,v = \int_S \varphi(x,s)\,v(s)\,\mathrm{d}\nu(s),$$

then

$$\begin{aligned} \langle Ku,\, v\rangle_S &= \int_X \left(\int_S \varphi(x,s)\,u(s)\,\mathrm{d}\nu(s) \int_S \overline{\varphi(x,t)}\,\overline{v(t)}\,\mathrm{d}\nu(t) \right) f(x)\,\mathrm{d}\mu(x) \\ &= \int_S \int_S \left(\int_X \varphi(x,s)\,\overline{\varphi(x,t)}\,f(x)\,\mathrm{d}\mu(x) \right) u(s)\,\overline{v(t)}\,\mathrm{d}\nu(s)\,\mathrm{d}\nu(t). \end{aligned}$$

This relation provides a formula for the operator K, namely

$$\langle Ku,\, v\rangle = \int_S \int_S K(s,t)\,u(s)\,\overline{v(t)}\,\mathrm{d}\nu(s)\,\mathrm{d}\nu(t)$$

where

$$(K\,u)\,(t) = \int_S K(s,t)\,u(s)\,\mathrm{d}\nu(s).$$

Any operator defined in this way will be called a Toeplitz operator.

At this moment for a better understanding we will present some examples, which will be useful in the sequel.

Let $(X, \mathcal{X}, \mu) = (S, \boldsymbol{\sigma}, \nu) = (\mathcal{R}_1, \mathcal{B}_1, \mathrm{Leb})$ and let

$$\varphi(x,s) = \frac{1}{\sqrt{2\pi}}\, e^{ixs}.$$

Then T is nothing but the well known Fourier–Plancherel transformation defined on $L_2(\mathcal{R}_1, \mathrm{Leb})$

$$(T\,u)\,(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ixs}\,u(s)\,\mathrm{d}s.$$

Then the kernel $K(\cdot,\cdot)$ expresses a stationary covariance function as follows from the Bochner theorem

$$K(s,t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i(s-t)x} f(x)\,\mathrm{d}x.$$

In the other example we will consider a kernel $\varphi(\cdot,\cdot)$ on $X \times S$ satisfying the condition

$$\int_X \varphi(x,s)\,\overline{\varphi(x,t)}\,\mathrm{d}\mu(x) = \delta_{st},$$

where δ_{st} is the Kronecker symbol. Then one can easily prove that the kernel $\varphi(\cdot,\cdot)$ is orthogonal, because the mapping T defined by

$$(T\,u)\,(x) = \int_S \varphi(x,s)\,\mathrm{d}\nu(s)$$

is unitary and the mapping K is surely a Toeplitz operator defining a covariance function

$$K_f(s,t) = \int_X \varphi(x,s)\,\overline{\varphi(x,t)}\,f(x)\,\mathrm{d}\mu(x)$$

belonging into the Karhunen class of covariance functions. In fact, we obtained a very interesting subclass of the Karhunen class. Some properties of random sequences having covariance functions of this type, i.e. $S = Z$ (integers) can can be found in Michálek and Rüschendorf [9]. Another information about Toeplitz operators one can find in Grenander [2] or in Kac [5].

Using the operators of the Toeplitz type we can say something about the asymptotic behaviour of their proper values. Let us formulate now the problem of asymptotic behaviour of proper values distribution of a Toeplitz operator.

Let $K_f(\cdot,\cdot)$ be a kernel of a Toeplitz operator on $L_2(S,\boldsymbol{\sigma},\nu)$ generated by a function $f(\cdot)$ defined on $(X,\mathcal{X},\mu)$ as described earlier. For simplicity, let $|f(\cdot)| \le$ const on X, then the corresponding operator $K(f)$ is bounded. Let in $(S,\boldsymbol{\sigma})$ be given a nondecreasing family of measurable subsets S_α, $\alpha \in \Lambda$ such that $S_\alpha \nearrow S$ and let the orthogonal kernel $\varphi(\cdot,\cdot)$ defining a unitary mapping T be bounded on S_α, i.e.

$$|\varphi(x,s)| \le C_\alpha$$

on $X \times S_\alpha$. Let us denote by $\mathcal{F}_\alpha$ the set of $\boldsymbol{\sigma}$-measurable functions bounded on S_α and vanishing outside of S_α in S. It is clear that $\bigcup_\alpha \mathcal{F}_\alpha$ is everywhere dense in $L_2(S)$. Further, we require that

$$\left\{T\,u : u(\cdot) \in \bigcup_\alpha \mathcal{F}_\alpha\right\} \text{ is dense everywhere in } L_2(X).$$

Now, for each $\alpha \in \Lambda$ (index set) we can define an operator K_α on $L_2(S_\alpha)$ by the relation

$$(K_\alpha(f)\, u)\ (t) = \int_{S_\alpha} K_f(s,t)\, u(s)\, \mathrm{d}\nu(s).$$

It is clear that $K_\alpha(f)$ is an operator of the Hilbert–Schmidt type when $\nu(S_\alpha) < +\infty$.

Let $\{\lambda_j^{(\alpha)}\}_{j=1}^\infty$ be proper values belonging to $K_\alpha(f)$ and the problem is to find an asymptotic distribution of $\{\lambda_j^{(\alpha)}\}_{j=1}^\infty$ on the real line if $S_\alpha \nearrow S$. As we see that $K_\alpha(f) \to K(f)$ in some sense it is suitable to use the behaviour of proper values $\{\lambda_j^{(\alpha)}\}_{j=1}^\infty$ for describing proper values $\{\lambda_j\}_{j=1}^\infty$ of the operator $K(f)$ that cannot be attainable in a usual way, e. g. due to computation problems. Using this approach we immediately see how we could utilize asymptotic properties of proper values belonging to a Toeplitz operator: proper values are included in the formula for Rényi distances $R_a(P_T,\, Q_T)$.

In the rest of paper we will be interested in Toeplitz operators with orthogonal Fourier kernels and their suitable approximation by operators of the Hilbert–Schmidt type.

Let $\varphi(x,s) = \frac{1}{\sqrt{2\pi}}\, e^{ixs}$ and $f(\cdot)$ be a real nonnegative function, $f(\cdot) \le M$ and integrable in R_1. Then the covariance function

$$K_f(s,t) = K_f(s-t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i(s-t)x}\, f(x)\, \mathrm{d}x$$

can be called as a Fourier kernel. In our case $S = \mathcal{R}_1^+$ and we put $S_T = \langle 0, T\rangle$ (it is possible, of course, to consider $S = \mathcal{R}_1$ and $S_T = \langle -\frac{T}{2},\, \frac{T}{2}\rangle$). We will study asymptotic behaviour of proper values of an integral operator $K(f)$ generated by the integral equation

$$\int_0^T K_f(s-t)\, \varphi(t) = \lambda\, \varphi(s)$$

for $T \nearrow +\infty$, i. e. for large T.

The kernel $K_f(s-t)$ will be approximated by a periodical kernel that generates a Hilbert–Schmidt operator whose proper values and functions can be found in an easy way and their asymptotic behaviour is very close to the asymptotic distribution of proper values of $K(f)$. For this purpose, let $A > 0$ and define a function

$$f_A(x) = \frac{1}{2\pi A} \int_{-\infty}^{+\infty} \left(\frac{\sin \frac{1}{2A(x-u)}}{(x-u)} \right)^2 f(x)\, \mathrm{d}x.$$

This function is bounded, $f_A(x) \le M$, nonnegative and integrable. Using this function let us define an approximating covariance function

$$K_A(s-t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i(s-t)x}\, f_A(x)\, \mathrm{d}x.$$

As well known

$$\begin{aligned} K_A(\tau) &= \left(1 - \frac{|\tau|}{A}\right) K(\tau) \quad \text{for} \quad |\tau| \le A \\ &= 0 \qquad\qquad\qquad\quad \text{for} \quad |\tau| > A. \end{aligned}$$

The similarity between the operators $K(f)$ and $K_A(f)$ can be measured by the Hilbert–Schmidt norm

$$\frac{1}{T}\|K(f) - K_A(f)\|^2 \le \frac{2A^2}{T^2}\int_0^A |K(s)|^2 \, \mathrm{d}s + 2\int_A^\infty |K(s)|^2 \, \mathrm{d}s.$$

Now we define a periodic kernel $L_{A,T}(s)$,

$$L_{A,T}(s) = K_A(s) \quad \text{for } -\frac{1}{2}T < s < \frac{1}{2}T,$$

with the period T. Then, similarly as above,

$$\frac{1}{T}\|K_A(f) - L_{A,T}\|^2 \le \frac{2A}{T}\int_0^\infty |K(s)|^2 \, \mathrm{d}s.$$

Combining the last two inequalities under $A < \frac{1}{2}T$ we can assert that

$$\frac{1}{T}\|K(f) - L_{A,T}\|^2 \le [O(T^{-2}) + 2\varepsilon]^{1/2} + O(T^{-1/2}),$$

where ε is chosen such that $\int_A^\infty |K(s)|^2 \, \mathrm{d}s < \varepsilon$.

In this way we have constructed a suitable approximation of the original Toeplitz operator $K(f)$ by a sequence of Hilbert–Schmidt operators with periodic kernels.

Let us consider the integral equation with the periodic kernel $L_{A,T}(\cdot)$:

$$\int_0^T L_{A,T}(s-t)\,\varphi(t)\,\mathrm{d}t = \lambda\,\varphi(s).$$

One can prove that under the choice $\varphi_{j,T}(t) = \exp\{2\pi tij/T\}$, $j \in Z$ we obtain proper functions of this equation with proper values

$$\lambda_{j,t} = \int_0^T L_{A,T}(t)\,e^{-2\pi ijt/T}\,\mathrm{d}t.$$

As well known, the system of functions $\{\varphi_{j,T}(t)\}$ is complete in $L_2\langle 0,T\rangle$ and hence no other proper functions and values exist. Using properties of Fourier transform we easily find that

$$\lambda_{j,T} = f_A\left(\frac{2\pi j}{T}\right), \quad j = 0, \pm 1, \pm 2, \dots$$

Let $p \in \mathcal{N}$. It is possible to prove (using the Poisson summation formula) that

$$\frac{1}{T}\sum_{j=-\infty}^{+\infty} (\lambda_{j,T})^p \underset{\substack{T\to\infty \\ A\to\infty}}{\longrightarrow} \frac{1}{2\pi}\int_{-\infty}^{+\infty} f^p(x)\,\mathrm{d}x.$$

This convergence proves the asymptotic behaviour of the moments belonging to the distribution of $\{\lambda_{j,T}\}$ for $T \to \infty$. The obtained result, very important in the sequel, can be summarized as follows.

Theorem 5. (Grenander and Szegő [3]) Let f be a bounded real function integrable over $\mathcal{R}_1$, let $K(\cdot,\cdot)$ be a Toeplitz kernel (covariance function) defined by Fourier kernel

$$K_f(s-t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i(s-t)x} f(x)\,\mathrm{d}x.$$

Then the proper values of the integral equation

$$\int_0^T K(s-t)\,\varphi(t)\,\mathrm{d}t = \lambda\,\varphi(s), \quad 0 \le x \le T$$

have an asymptotic distribution equal to that of a random variable $f(X)$, where the random variable X has an improper probability density function equal to $\frac{1}{2\pi}$ all over the real line.

This theorem will be extensively used in the next chapter in evaluating the asymptotic rates of the Rényi distances between two Gaussian and stationary probability measures.

4. ASYMPTOTIC RATE OF THE RÉNYI DISTANCES

Let two covariance functions $R(\cdot,\cdot)$, $S(\cdot,\cdot)$ be given on $\langle 0,T\rangle$. Let P_T, Q_T be the corresponding Gaussian measures with vanishing means.

In Chapter 2 we introduced necessary and sufficient conditions under which P_T and Q_T are mutually absolutely continuous and we evaluated the corresponding the Rényi distances using proper values of the operator B_T. This operator is defined on a somewhat abstract space constructed from the observations $x(t)$, $t \in \langle 0,T\rangle$. It is possible, of course, to construct a similar Hilbert–Schmidt operator connected with two covariance functions $R(\cdot,\cdot)$ and $S(\cdot,\cdot)$ defining the measures P_T and Q_T. This situation describes the following

Theorem 6. (Pitcher [12]) Let two continuous covariance functions $R(\cdot,\cdot)$ and $S(\cdot,\cdot)$ be given on $\langle 0,T\rangle$. Let $R(\cdot,\cdot)$ be strictly positive and both belong to $L_2\langle 0,T\rangle$. Let $\hat{R}$ and $\hat{S}$ be the corresponding integral operators defined on $L_2\langle 0,T\rangle$. Then the Gaussian measures generated by $R(\cdot,\cdot)$ and $S(\cdot,\cdot)$ are equivalent if and only if the operator

$$I - \hat{R}^{-1/2}\,\hat{S}\,\hat{R}^{-1/2}$$

is of Hilbert–Schmidt type.

We obviously see that the role of the operator $I - B_T$ is played here by $I - \hat{R}^{-1/2}\,\hat{S}\,\hat{R}^{-1/2}$, which is defined on $L_2\langle 0,T\rangle$. We have

$$(\hat{R}\,f)\,(t) = \int_0^T R(s,t)\,f(s)\,\mathrm{d}s, \qquad (\hat{S}\,f)\,(t) = \int_0^T S(s,t)\,f(s)\,\mathrm{d}s.$$

As we assume positive definiteness of $R(\cdot,\cdot)$, the corresponding operrator $\hat{R}$ must be positive, too, and it is possible to consider its inverse $\hat{R}^{-1}$ which however need not be bounded, in general. The operator $\hat{R}^{-1}$ is also positive and hence its square root is well defined. P_T and Q_T are mutually absolute continuous and this fact implies that $\hat{S}^{-1}$ must exist, too. The nonexistence of an inverse operator to $\hat{S}$ would lead in such a situation to perpedicularity of Gaussian measures.

Let us start with the formula for $R_a(P_T|Q_T)$ and properties of the operator $a\,B_T^{a-1} + (1-a)\,B_T^a$ where $B_T = \hat{R}^{-1/2}\,\hat{S}\,\hat{R}^{-1/2}$ on $L_2\langle 0, T\rangle$. The Rényi distance can be written in the form

$$R_a(P_T|Q_T) = \frac{1}{2(1-a)\,a}\,\mathrm{tr}\ln B_a,$$

where $B_a = a\,B_T^{a-1} + (1-a)\,B_T^a$. As the operator $B_T - I$ is nuclear according to Theorem 2, the distance $R_a(P_T|Q_T)$ can be also expressed as

$$R_a(P_T|Q_T) = \frac{1}{2(1-a)\,a}\ln\det B_a,$$

where $\det B_a = \prod_1^\infty \left(a\,\alpha_i^{a-1} + (1-a)\,\alpha_i^a\right)$ is convergent.

For the boundary points $a = 0$ and $a = 1$ the operators

$$B_0 = B_T^{-1} - \ln B_T^{-1} - I, \qquad B_1 = B_T - \ln B_T - I$$

have also finite traces. Let $\{u_i\}_{i=1}^\infty$ be an orthonormal system of proper functions corresponding to proper values $\{\alpha_i\}_{i=1}^\infty$ of the operator B_T. Then using a result of Hájek [4] we can approximate the Fredholm determinant of B_T by a sequence of matrix determinants, namely

$$\det B_T = \lim_{n\to\infty}\det\{\langle B_T\,u_i,\,u_j\rangle\}_{i,j=1}^n.$$

Since

$$\langle B_T\,u_i,\,u_j\rangle = \langle\hat{S}\,u_i,\,\hat{R}^{-1}uj\rangle = \langle\hat{R}^{-1}\,\hat{S}\,u_i,\,u_j\rangle,$$

we obtain

$$\det\{\langle B_T\,u_i,\,u_j\rangle\} = \det\{\langle\hat{R}^{-1}u_i,\,u_j\rangle\}\,\det\{\langle S\,u_i,\,u_j\rangle\}.$$

This easily yields

$$\ln\det\{\langle B_T\,u_i,\,u_j\rangle\}_{i,j=1}^n = \sum_{i=1}^n \ln(\mu_i) - \ln(\lambda_i),$$

where $\{\mu_i\}_{i=1}^\infty$ are proper values of $\hat{S}$ and $\{\lambda_i\}_{i=1}^\infty$ are proper values of $\hat{R}$. In this way we established the possibility to express $R_a(P_T,\,Q_T)$ by means of proper values of $\hat{S}$ and $\hat{R}$,

$$R_a(P_T|Q_T) = \frac{1}{2(1-a)\,a}\sum_{i=1}^\infty \ln\left(a\left(\frac{\mu_i}{\lambda_i}\right)^{a-1} + (1-a)\left(\frac{\mu_i}{\lambda_i}\right)^a\right).$$

Now it is natural to apply Theorem 5 about asymptotic behaviour of proper values belonging to Toeplitz operators with Fourier kernels.

Theorem 7. Let Gaussian stationary measures P_T, Q_T be absolutely continuous for each $T > 0$. Let f_P, f_Q be their corresponding spectral density functions. We need f_P, f_Q to be bounded. Then there exists the asymptotic rate of the Rényi distance in the form

$$\lim_{T\to\infty} \frac{1}{T} R_a(P_T|Q_T) = \frac{1}{4\pi(1-a)\,a} \int_{-\infty}^{+\infty} \ln\left(a\left(\frac{f_Q(x)}{f_P(x)}\right)^{a-1} + (1-a)\left(\frac{f_Q(x)}{f_P(x)}\right)^{a}\right) \mathrm{d}x$$

but we must accept the value $+\infty$, too.

Proof. We have shown that det B_T can be approximated by determinants of covariance functions $R(\cdot,\cdot)$ and $S(\cdot,\cdot)$. Therefore

$$\ln \det B_A = \lim_{n\to\infty} \left(a \ln \det {}_nB_T + \ln \det(a\,{}_nB_T^{-1} + (1-a)\,E_n)\right),$$

where ${}_nB_T$ is an n-dimensional approximation of the operator B_T, which can be expressed by using a matrix form. In this way we get to the formula

$$R_a(P_T, Q_T) = \frac{1}{2(1-a)\,a} \sum_{i=1}^{\infty} \ln\left(a\left(\frac{\mu_{i,T}}{\lambda_{i,T}}\right)^{a-1} + (1-a)\left(\frac{\mu_{i,T}}{\lambda_{i,T}}\right)^{a}\right),$$

where $\{\lambda_{i,T}\}$ and $\{\mu_{i,T}\}$ are proper values of $\hat{R}$ and $\hat{S}$ operating only on $L_2\langle 0, T\rangle$. At this place we need the following properties of Toeplitz operators. Let f_1, f_2 be two real functions generating Toeplitz operators $K(f_1)$, $K(f_2)$. Then

$$K(a\,f_1 + b\,f_1) = a\,K(f_2) + b\,K(f_2), \qquad K(f_1\,f_2) = K(f_1)\,K(f_2).$$

On the basis on this fact we can easily prove that if $\{\lambda_{i,T}\}$ and $\{\mu_{i,T}\}$ are proper values of the operators $K(f_1)$, $K(f_2)$ operating on $L_2\langle 0, T\rangle$ then for any polynomial $f_N(x,y)$ of two variables

$$\frac{1}{T} \sum_{i=1}^{\infty} f_N(\lambda_{i,T}, \mu_{i,T}) \underset{T\to\infty}{\longrightarrow} \frac{1}{2\pi} \int_{-\infty}^{+\infty} f_N(f_1(x), f_2(x))\,\mathrm{d}x.$$

As every continuous function $g(x,y)$ can be approximated by a sequence of polynomials, we can conclude that

$$\frac{1}{T} \sum_{i=1}^{\infty} g(\lambda_{i,T}, \mu_{i,T}) \underset{T\to\infty}{\longrightarrow} \frac{1}{2\pi} \int_{-\infty}^{+\infty} g(f_1(x), f_2(x))\,\mathrm{d}x$$

provided the infinite sums and integral are convergent. Hence we get the formula for the asymptotic rate of the Rényi distance

$$\overline{R_a}(P, Q) = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \ln\left(a\left(\frac{f_Q(x)}{f_P(x)}\right)^{a-1} + (1-a)\left(\frac{f_Q(x)}{f_P(x)}\right)^{a}\right) \mathrm{d}x. \qquad \square$$

Remark. Unfortunately, the absolute continuity of two stationary Gaussian measures P_T and Q_T for each $T > 0$ does not ensure the convergence of the integral figuring above. There is an example, even with rational spectral density functions where $P_T \sim Q_T$ for each $T > 0$ but

$$\overline{R_1}(P, Q) < +\infty \quad \text{and} \quad \overline{R_1}(Q|P) = +\infty,$$

see Pisarenko [11].

In Pinsker. [10] are given formulae for the asymptotic rate of the Rényi distances only for the boundary cases $a = 0$ and $a = 1$, claimed to be valid even for multidimensional Gaussian processes. Unfortunately these results were stated without proof. Analogous result concerning the one dimensional case was given later in Kullback, Keegal and Kullback [7]. But the proof presented there is also incomplete.

(Received December 11, 1997.)

REFERENCES

[1] I. Gohberg and M. Krein: An Introduction to the Theory of Linear Nonselfadjoint Operators in Hilbert Space. Nauka, Moscow 1965. In Russian.

[2] V. Grenander: A contribution to the theory of Toeplitz matrices. Trans. Amer. Math. Soc. *79* (1955), 124–140.

[3] V. Grenander and G. Szegő: Toeplitz Forms and their Application. Innostrannaja literatura, Moscow 1961. In Russian.

[4] J. Hájek: On linear statistical problems. Czechoslovak Math. J. *12 (87)* (1962), 3, 404–444.

[5] M. Kac: Toeplitz matrices, translation kernels and related problems in probability theory. Duke Math. J. *21* (1954), 501–509.

[6] T. Kadota: Simultaneous diagonalization of two covariance kernels and application to second order stochastic process. SIAM J. Appl. Math. *15* (1967), 1470–1480.

[7] S. Kullback, J. C. Keegel and J. H. Kullback: Topics in Statistical Information Theory (Lecture Notes in Statistics 42). Springer–Verlag, Berlin 1987.

[8] J. Michálek: Asymptotic Rényi's rate of Gaussian processes. Problem Control Inform. Theory *18* (1990), 8, 209–227.

[9] J. Michálek and L. Rüschendorf: Karhunen class processes forming a basis. In: Trans. 12th Prague Conference on Inform. Theory, Statist. Decis. Functions, Random Process., Prague 1992, Academy of Sciences of the Czech Republic, pp. 158–160.

[10] M. S. Pinsker: Information and Information Stability of Random Variables and Processes. Izv. AN SSSR, Moscow 1960. In Russian.

[11] B. Pisarenko: On the problem of detection of random signal in noise. Radiotekhn. i Elektron. *6* (1961), 4, 514–528 In Russian.

[12] T. Pitcher: An integral expression for the log likelihood ratio of two Gaussian processes. SIAM J. Appl. Math. *14* (1966), 228–233.

[13] J. Rozanov: Infinite–Dimensional Gaussian Probability Distributions. Nauka, Moscow 1968. In Russian.

[14] I. Vajda: Theory of Statistical Inference and Information. Kluwer, Dordrecht 1989.

RNDr. Jiří Michálek, CSc., Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 182 08 Praha 8. Czech Republic.

e-mail: michalek@utia.cas.cz