UDC 519.22 (045)

[1]I. G. Prokopenko,
[2]I. A. Martynchuk

NONPARAMETRIC DETECTION ALGORITHM OF THE SIGNALS THAT USING THE KERNEL ESTIMATION OF DISTRIBUTION FUNCTION OF A NOISE AND BETA-DISTRIBUTION OF THE MODIFIED RANKS

Aeronavigation institute, National Aviation University, Kiev, Ukraine
E-mails: [1]prokop-igor@yandex.ru, [2]igor9852@hotmail.com

Abstract. *The nonparametric detection algorithms of noise-type radar signals is considered, which will be applied as in cases of change of scale parameter of distribution of signal sample in the presence of the signal, and shift parameter and shows a big efficiency in comparison with classical ranks algorithms.*

Keywords: kernel function; learning sample; nonparametric detection algorithm; rank; probability density distribution.

Introduction

To construct the algorithm for detecting signals in the proposed approach, using kernel estimate the cumulative distribution function of noise received from learning sample of noise. Kernel approximation of the cumulative distribution function of noise is used for the nonlinear transformation of the signal sample. In the absence of a signal converted sample has a uniform distribution on the interval [0, 1], in the presence, the distribution of the transformed sample differs from uniform.

Analysis of researches and publications

Recently, the problems of signal detection are studied and applied nonparametric methods. Such methods in pure form, although not imply knowledge least some information about the functional form of the distribution of noise and composition of signal and noise (e.g. Wilcoxon rank algorithm). In developing nonparametric detection algorithm is necessary to ensure only one condition: the distribution of decision statistic of algorithm should not depend on the noise distribution. This implies that at a constant distribution function of decision statistic remains fixed false alarm probability. As known, the probability of false alarm determines the quality of detector, so its stabilization, at changing the noise environment, is a very important task.

Among the non-parametric methods, ranking detection criteria occupy the first place, due to their ease of implementation. They are widely used in a variety of information-measuring systems. Widely known criteria are such as Wilcoxon, Kolmogorov, and others. Processing algorithms, using these statistics, provide stability of probability of the first kind errors, that is a false alarm, but a number of problems of their low efficiency. Their performance can improve by using a priori or a posteriori information about the probability distribution of noise and composition of signal and noise.

Problem

The task of constructing and studying nonparametric algorithm for detecting noise-like signals, which would use a priori information about the class of noise distributions. In the term “non-parametric” weakened the wording adopted in the signal detection theory [4], because it uses information about the class of distributions and only one smoothing parameter of kernel estimation of distribution function of noise.

Will be considered only two classes of noise distributions: the first class includes unimodal functions that have a convex shape, the second class includes functions that have a convex shape, great length on the right, one mode and are defined only on the positive axis. This selection is associated with the fact that in this article algorithm is tested for the detector that is intended for a Gaussian noise-like signal in which the amplitude demodulator is used.

Further research is to check. Is whether the algorithm to ensure the stability of a false alarm if the noise will change in a given class?

Statistic synthesis of detection algorithm

To construct the detection statistic that is free from distribution uses two signals: learning $y(t)$ (which contains only a noise) and working $x(t)$ (consisting of known signal and noise composition). After their passage through the ADC respectively obtained two samples: y – noise sample of size Ny and x – signal sample of size Nx. On the noise sample kernel estimation of a probability distribution function (CDF) is constructed with the optimal scaling parameter. Gaussian kernel is chosen because it provides the smoothness and continuity of the resulting estimates of the distribution function. Optimal scale parameter is chosen according to the sample size and the supposed distribution function of noise that belongs to a given class.

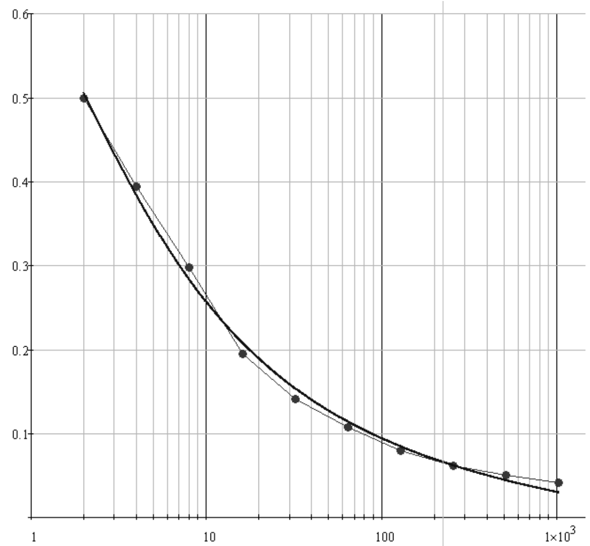
Let us assume that the supposed distribution function of noise before amplitude demodulator will be Gaussian, and after amplitude demodulator will be Rayleigh. For Gaussian distribution dependence of optimal scale parameter on the sample size shall be equal to:

$$h_{\text{opt}}(Ny) = k_{\text{opt}}(Ny)\Delta = \left[\frac{0.389}{\lg(Ny + 2.401)} - 0.1\right]\Delta.$$

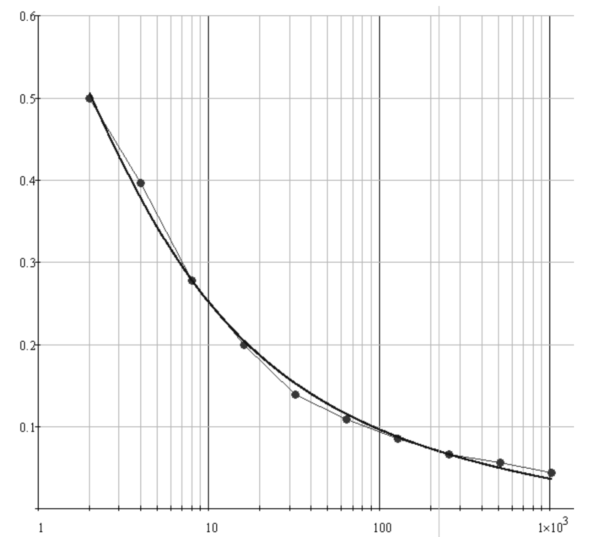
For Rayleigh distribution this dependence is accepted the following:

$$h_{\text{opt}}(Ny) = k_{\text{opt}}(Ny)\Delta = \left[\frac{0.366}{\lg(Ny + 2.147)} - 0.086\right]\Delta.$$

where Δ is sample span.



a



b

Fig. 1. Dependence of $k_{\text{opt}}(Ny)$: experimental dependence is points connected by lines; regression dependence is solid line: (a) for the Gaussian distribution; (b) for the Rayleigh distribution

Experimental and regression dependence of $k_{\text{opt}}(Ny)$ for these distributions is shown in Fig. 1a and b respectively. As can be seen from the formulas and figures, dependences of $k_{\text{opt}}(Ny)$ for Gaussian and Rayleigh distributions differ little, so in practice it is natural to replace their on some averaged dependence

$$h_{\text{opt}}(Ny) = k_{\text{opt}}(Ny)\Delta = \left[\frac{0.38}{\lg(Ny + 2.3)} - 0.09\right]\Delta.$$

Choosing a model

$$h_{\text{opt}}(N) = \frac{a}{\lg(N + c)} + b$$

is taken from [5] about kernel probability density estimation with the optimal scale parameter for different sample sizes. The only thing that should be clarified, since this expansion $h_{\text{opt}}(Ny) = k_{\text{opt}}(Ny)\Delta$. This linear dependence is completely valid for Gaussian distribution, but for Rayleigh distribution, it can take only approximately. Multiplication by sample span eliminates the need to know the formula for $h_{\text{opt}}(N)$, if the distribution will change the scale. The disadvantage of this approach is to decrease the accuracy.

We mentioned above two classes of distributions with which we work. For them, the optimal scale parameter is assumed equal to the expression (1). It is clear that in the case of non-Gaussian noise, the scale parameter will not be optimum, which will affect the accuracy of the estimate of the distribution function of noise. But it's tolerable risk in the absence of a priori information about the true distribution of noise.

We denote the kernel estimation of CDF of noise as $Ky^*(\bullet)$. By nonlinear transformation $rx_n = Ky^*(x_n)$, $n \in 0 \ldots Nx - 1$ formed the sample rx, which we call sample of "modified" ranks. Due to the fact that the values of rx would have sense of ranks of the signal sample regarding the learning sample, divided by Ny, if $Ky^*(\bullet)$ was a stepwise empirical CDF.

Under the hypothesis H_0 there is no signal, the distribution of statistic of "modified" rank will be described uniform law.

Under the hypothesis H_1 the signal is present, the distribution of statistic of "modified" rank will be described by the law, other than the uniform.

Detection statistic synthesized according to the Neyman-Pearson lemma. If we assume that the "modified" ranks are statistically independent, then

$$T(rx_0, rx_1, \dots, rx_{Nx-1}; b) = \prod_{n=0}^{Nx-1} \frac{fH1^*(rx_n; b)}{fH0^*(rx_n)} = \prod_{n=0}^{Nx-1} \frac{fH1^*(rx_n; b)}{1} = \prod_{n=0}^{Nx-1} fH1^*(rx_n; b) \quad (2)$$

where $fH1^*(\bullet;\bullet)$ is the empirical probability density function (PDF) constructed from rx under the hypothesis H_1; $fH0^*(\bullet)$ is empirical PDF constructed from rx under the hypothesis H_0; b is signal parameter (it may be a signal-to-noise ratio).

Algorithm for finding empirical PDF $fH1^*$

Algorithm for finding empirical PDF consists of the following.

1. It is computed dotty empirical CDF on the sorted sample rx of the "modified" ranks, which points are expressed as

$$\left\{ rx_n; \frac{n+1}{Nx+1} \right\}, \; n \in 0 \dots Nx-1.$$

2. As a model of the cumulative probability distribution function of the "modified" ranks accepted beta distribution

$$\text{Beta}(x, s0, s1) = \frac{\Gamma(s0+s1)}{\Gamma(s0)\Gamma(s1)} \int_0^x t^{s0-1} (1-t)^{s1-1} \, dt,$$

where $\Gamma(\bullet)$ is a gamma function

$$\Gamma(z) = \int_0^\infty t^{z-1} \cdot \exp(-t) dt,$$

$s0$, $s1$ are shape parameters, $x \in (0;1)$ is an independent variable. This model is flexible enough to change its shape, which is determined by the parameters $s0 \in (0;1)$, $s1 \in (0;1)$.

3. Bounded intervals for $s0$ and $s1$ allows you to organize search such $s0_{\min}$ and $s1_{\min}$, which minimize the error estimation of CDF (the sum of squared differences)

$$\text{error}(s0, s1) = \sum_{n=0}^{Nx-1} \left(\text{Beta}(rx_n, s0, s1) - \frac{n+1}{Nx+1} \right)^2. \quad (3)$$

Minimization is carried out by direct search of values $\text{error}(s0_n, s1_m)$. Here

$$s0_n = \Delta 0 + n \cdot \Delta 0, n = 0 \dots N_{s0} - 1, \Delta 0 = \frac{1}{N_{s0}+1};$$

$$s1_m = \Delta 1 + m \cdot \Delta 1, m = 0 \dots N_{s1} - 1, \Delta 1 = \frac{1}{N_{s1}+1};$$

N_{s0} and N_{s1} are the number of nodes in the grid along the axes $s0$ and $s1$ respectively, which are chosen for reasons of accuracy and computational speed.

4. Probability density function is the derivative of the cumulative probability distribution function, so

$$\text{Beta}(x, s0, s1)' = \frac{\Gamma(s0_{\min} + s1_{\min})}{\Gamma(s0_{\min})\Gamma(s1_{\min})} x^{s0_{\min}-1} (1-x)^{s1_{\min}-1},$$

Then

$$fH1^*(x) = \frac{\Gamma(s0_{\min} + s1_{\min})}{\Gamma(s0_{\min})\Gamma(s1_{\min})} x^{s0_{\min}-1} (1-x)^{s1_{\min}-1}.$$

Comparison of detection performances

We compare the detection performances of the algorithm (2) with the Kolmogorov rank algorithm and optimal parametric with known parameters in the following tasks:

1. Detecting a noise-like signal before demodulation (before non-linear part of the receiver) in Gaussian noise (Fig. 2). In the presence of a signal, variance of received realization increases.

2. Detecting a noise-like signal after amplitude demodulator, when before demodulator Gaussian noise (Fig. 3). In the presence of a signal, the mean and variance of the realization at the demodulator output increases.

For the comparative analysis of efficiency Kolmogorov's algorithm is chosen

$$T(R_0, R_1, \dots, R_{Nx-1}) = \max_{n \in 0 \dots Nx-1} \left| F\left(\frac{R_n}{Ny}\right) - F^*\left(\frac{R_n}{Ny}\right) \right|$$

as more sensitive to the situation where the presence of the signal leads to an increase in the scale parameter of the distribution of composition. Here $F(\bullet)$ is uniform CDF with parameters $a = 0, b = 1$; $F^*(\bullet)$ is step empirical CDF which is built on sample $r_n = \frac{R_n}{Ny}$, $n \in 0 \dots Nx-1$; R is a vector of ranks of signal sample concerning noise sample.

To determine the upper limit of effectiveness was built detection performance of optimal parametric algorithm

$$T(x_0, x_1, \dots, x_{Nx-1}; b) = \prod_{n=0}^{Nx-1} \frac{fH1(x_n; b)}{fH0(x_n)}.$$

Simulation parameters, at which were built the detection performances, summarized in the table.

Simulation parameters

$\sigma = 1$	Standard deviation of noise
$\mu = 0$	Mean of noise
$Ny = 64$	Size of learning sample
$Nx = 64$	Size of signal (working) sample
$ND = 10000$	Number of statistical tests to simulate the dependence of the probability of correct detection on the signal/noise ratio
$\alpha = 0.01$	Probability of a false alarm
$N_{s0} = 100$	The number of nodes in the mesh along the axis $s0$, which determine the accuracy of the minimization of a function (3)
$N_{s1} = 100$	The number of nodes in the mesh along the axis $s1$, which determine the accuracy of the minimization of a function (3)
$b_{op} = 1$	Signal parameter, which is set optimal parametric algorithm

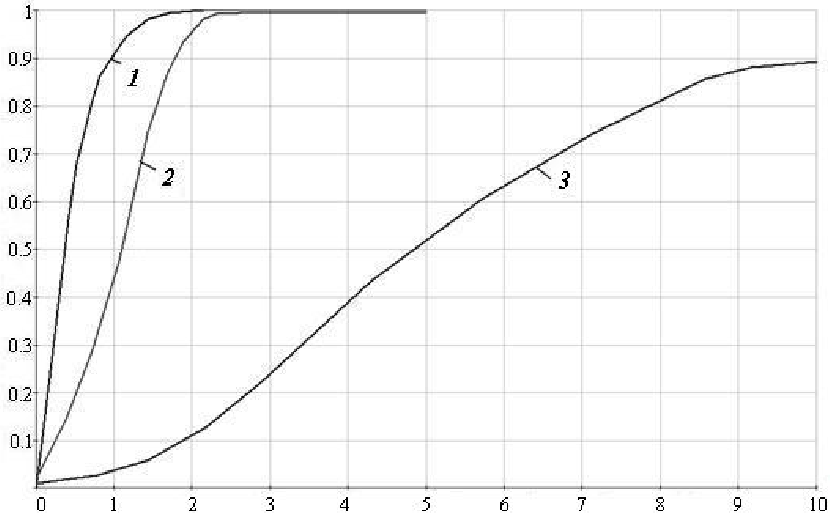


Fig. 2. Detection performances for task of processing before demodulator (on the x-axis signal parameter b): 1 is the optimal parametric algorithm; 2 is the algorithm (2); 3 is the Kolmogorov's algorithm

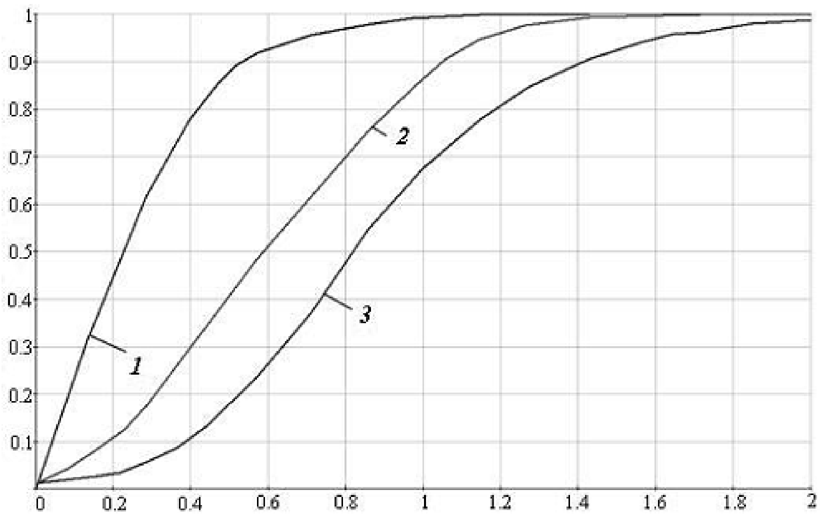


Fig. 3. Detection performances for task of processing after demodulator (on the x-axis signal parameter b): *1* is the optimal parametric algorithm; *2* is the algorithm (2); *3* is the Kolmogorov's algorithm

Conclusions

Synthesized algorithm is effective in detecting problems when changing the scale parameter, and the problem of detecting when changing the shift parameter of distribution of composition signal and noise.

Gain in the signal/noise ratio compared with the Kolmogorov's algorithm for the level $D = 0.9$ in task of processing before the demodulator is 7.4 dB, and in task of processing after the demodulator is 1.2 dB.

Loss to optimal parametric detector does not exceed 2 dB.

References

[1]. Prokopenko, I. G. "Statistical signal processing": teaching book. Kyiv, NAU. 2011. 220 p. (in Ukrainian).

[2]. Levin, B. R. "Theoretical bases of statistical radio engineering." In three books. The second book. Moscow. *Soviet radio*. 1975. 392 p. (in Russian).

[3]. Levin, B. R. "Theoretical bases of statistical radio engineering." In three books. The third book. Moscow. *Soviet radio*. 1976. 288 p. (in Russian).

[4]. Akimov, P. S.; Bakut, P. A.; Bogdanovich, V. A.; etc. "Theory of signals detection." Moscow. *Radio and communication*. 1984. 440 p. (in Russian).

[5]. Prokopenko, I. G.; Martynchuk, I. A. "Optimization of smoothing parameter by kernel estimation of probability density distribution." *Herald of NAU*. 2014. Kyiv, NAU. (in press).

Received 22 November 2013.

Prokopenko Igor Grigorevich. Dr. Sci. Tech. Professor.
Aviation radio-electronic complexes department, National Aviation University, Kyiv, Ukraine.
Education: Kyiv institute of engineers of civil aviation, Ukraine (1972).
Research interests: signal and data processing.
Publications: more then 200.
E-mail: prokop-igor@yandex.ru

Martynchuk Igor Andreevich. Post-graduate student. Engineer.
Aviation radio-electronic complexes department, National Aviation University, Kyiv, Ukraine.
Education: National Aviation University, Kyiv, Ukraine (2012).
Research interests: signal and data processing.
Publications: 8.
E-mail: igor9852@hotmail.com

І. Г. Прокопенко, І. А. Мартинчук. Непараметричний алгоритм виявлення сигналів, який використовує ядерну оцінку функції розподілу завади і бета-розподіл модифікованих рангів.

Розглянуто непараметричний алгоритм виявлення радіолокаційних шумоподібних сигналів, який може використовуватись як у випадках зміни параметра масштабу розподілу сигнальної вибірки при наявності сигналу, так і зсуву і показує більшу ефективність в порівнянні із класичними ранговими алгоритмами.

Ключові слова: навчальна вибірка; непараметричний алгоритм виявлення; ранг; щільність розподілу ймовірності; ядерна функція.

Прокопенко Ігор Григорович. Доктор технічних наук. Професор.
Кафедра авіаційних радіоелектронних комплексів, Національний авіаційний університет, Київ, Україна.
Освіта: Київський інститут інженерів цивільної авіації. Київ, Україна. (1972).
Напрям наукової діяльності: обробка сигналів і даних.
Кількість публікацій: більше 200.
E-mail: prokop-igor@yandex.ru

Мартинчук Ігор Андрійович. Аспірант. Інженер.
Кафедра авіаційних радіоелектронних комплексів, Національний авіаційний університет, Київ, Україна.
Освіта: Національний авіаційний університет, Київ, Україна. (2012).
Напрям наукової діяльності: обробка сигналів і даних.
Кількість публікацій: 8.
E-mail: igor9852@hotmail.com

И. Г. Прокопенко, И. А. Мартынчук. Непараметрический алгоритм обнаружения сигналов, использующий ядерную оценку функции распределения помехи и бета-распределение модифицированных рангов

Рассмотрен непараметрический алгоритм обнаружения радиолокационных шумоподобных сигналов, который может применятся как в случаях изменения параметра масштаба распределения сигнальной выборки при наличии сигнала, так и сдвига и показывает большую эффективность по сравнению с классическими ранговыми алгоритмами.

Ключевые слова: непараметрический алгоритм обнаружения; обучающая выборка; плотность распределения вероятности; ранг; ядерная функция.

Прокопенко Игорь Григорьевич. Доктор техничных наук. Профессор.
Кафедра авиационных радиоэлектронных комплексов, Национальный авиационный университет, Киев, Украина.
Обучение: Киевский институт инженеров гражданской авиации. Киев, Украина. (1972).
Направление научной деятельности: обработка сигналов и данных.
Количество публикаций: больше 200.
E-mail: prokop-igor@yandex.ru

Мартынчук Игорь Андреевич. Аспирант. Инженер.
Кафедра авиационных радиоэлектронных комплексов, Национальный авиационный университет, Киев, Украина.
Обучение: Национальный авиационный университет, Киев, Украина (2012).
Направление научной деятельности: обработка сигналов и данных.
Количество публикаций: 8.
E-mail: igor9852@hotmail.com