ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

UDC 519.218.7 **doi: 10.32620/reks.2022.1.17**

Oleg RUDENKO, Oleksandr BEZSONOV

Kharkiv National University of Radio Electronics, Ukraine

ADAPTIVE IDENTIFICATION UNDER THE MAXIMUM CORRENTROPY CRITERION WITH VARIABLE CENTER

The problem of identifying the parameters of a linear object in the presence of non-Gaussian noise is considered. The identification algorithm is a gradient procedure for maximizing the functional, which is a correntropy. This functionality allows you to get estimates that have robust properties. In contrast to the commonly used Gaussian kernels, the centers of which are at zero and effective for distributions with zero mean, this paper considers a modification of the criterion suitable for distributions with nonzero mean. The modification is to use correntropy with a variable center The use of Gaussian kernels with a variable center will allow us to estimate unknown parameters under Gaussian and non-Gaussian noises with zero and non-zero mean distributions and provide an opportunity to develop new technologies for data analysis and processing. It is important to develop a robust identification algorithm based on correntropy with a variable center. Their properties in the identification of stationary and non-stationary objects are ***the subject*** *of research. The* ***goal*** *is to develop a robust identification algorithm that maximizes the criterion of correntropy with a variable center using center configuration procedures and kernel width and to study its convergence in stationary and non-stationary cases under non-Gaussian noise. Expressions for the steady-state value of the estimation error are obtained, which depend on the type of noise distribution and the degree of non-stationarity of the estimated parameters The following* ***tasks*** *are solved: to investigate the convergence of the algorithm and determine the conditions for the stability of the established identification process.* ***Methods*** *of estimation theory (identification) and probability theory were used. The following* ***results*** *were obtained: 1) the developed algorithm provides robust estimates in the presence of noise having a distribution with zero and non-zero mean; 2) its convergence was studied in stationary and non-stationary cases under conditions of Gaussian and non-Gaussian noise; 3) simulation of the algorithm was carried out. 1) the developed algorithm consists of the development of a robust identification algorithm that maximizes the criterion of correntropy with a variable center; 2) its convergence in stationary and non-stationary cases in the conditions of Gaussian and non-Gaussian noise is investigated; 3) simulation of the algorithm is performed.* ***Conclusions****: The results of the current study will improve existing data processing technologies based on robust estimates and accelerate the development of new computing programs in real-time.*

Keywords: *correntropy; maximization; functional; gradient algorithm; asymptotic estimation; convergence; identification accuracy; steady state.*

Introduction

Many tasks of control, forecasting, pattern recognition, etc. associated with building a view model: ξ

$$y_{n+1} = c^{*T} x_{n+1} + \xi_{n+1}, \quad (1)$$

where y_{n+1} is the observed output signal; $x_{n+1} = (x_{1,n+1}, x_{2,n+1}, .. x_{N,n+1})^T$ is the vector of input signals $N \times 1$; $c^* = (c_1^*, c_2^*, .. c_N^*)^T$ is the vector of the required parameters $N \times 1$; ξ_{n+1} is the output -noise; n is the discrete time, and are reduced to minimizing some previously selected quality functional (identification criterion)

$$F[e_n] = \sum_{i=1}^{n} \rho(e_i), \quad (2)$$

where $e_i = y_i - \hat{y}_i$; $\hat{y}_i = c_{i-1}^T x_i$ is the model output; c is the vector estimate c^*;

$\rho(e_i)$ is some differentiable loss function satisfying the conditions:

1) $\rho(e_i) \ge 0$;
2) $\rho(0) = 0$;
3) $\rho(e_i) = \rho(-e_i)$;
4) $\rho(e_i) \ge \rho(e_j)$ for $|e_i| \ge |e_j|$.

The identification problem is to find an estimate $\hat{\theta}$ defined as a solution to the extremal minimum problem

$$F(\theta) = \min \quad (3)$$

or as a solution to the system of equations

$$\frac{\partial F(e)}{\partial \theta_j} = \sum_{i=1}^{n} \rho'(e_i) \frac{\partial e_i}{\partial \theta_j} = 0, \quad (4)$$

where $\rho'(e_i)) = \frac{\partial \rho(e_i)}{\partial e_i}$ is the function of influence.

If we introduce a weight function $\omega(e) = \rho'(e)/e$, then the system of equations (4) can be written as follows:

$$\sum_{i=1}^{n} \omega(e_i) e_i \frac{\partial e_i}{\partial \theta_j} = 0 \quad (5)$$

and minimization of functional (2) will be equivalent to minimization of the weighted quadratic functional, which is most often encountered in practice

$$\min \sum_{i=1}^{n} \omega(e_i) e_i^2 . \quad (6)$$

When choosing $\rho(e_i) = 0{,}5e_i^2$ the function of influence is $\rho'(e_i) = e_i$, i.e. grows linearly with increasing, which explains the instability of the OLS estimate to outliers and to noise, the distributions of which have large tails.

A stable M-estimate is also an estimate c defined as a solution to the extremal problem (3) or as a solution to the system of equations (4), but the loss function $\rho(e_i)$ is chosen to be different from the quadratic one.

There is a fairly large number of functionals that provide robust M-estimates; however, the most common are the combined functionals proposed by Huber [1] and Hempel [2] and consisting of a quadratic one, which ensures the optimality of estimates for a Gaussian distribution, and a modular one, which makes it possible to obtain more robust to distributions. with heavy tails (outliers) estimate. However, the efficiency of the obtained robust estimates substantially depends on the numerous parameters used in these criteria and selected on the basis of the researcher's experience.

The practical application of these functionals for solving the identification problem was considered in many works, in particular, in [3, 4]. Another approach to obtaining robust estimates, devoid of the indicated drawback, is the use of a combined criterion, using a combination of the quadratic criterion and the criterion of least modules [5, 6], the quadratic criterion and the fourth degree criterion [7], the fourth degree criterion and the least modulus criterion [8, 9]. It should be noted that the use of the combined criterion turned out to be very effective and much simpler in the implementation of the identification procedure.

Another approach that is currently widely used is the approach based on information characteristics of signals, entropy, in particular. The functional used in this case is an explicit functional of the probability density function (PDF) and includes all the higher-order statistical properties defined in PDF. Since entropy measures the mean uncertainty contained in a given PDF, minimizing it provides a reduction in error. Under minimum error entropy (MEE) criterion, several gradient-based adaptation algorithms, including the LMS-like algorithm, i.e., the stochastic information gradient algorithm, have been developed by the researchers [10]. In [11] was analyzed the structure of the MEE performance surface around the optimal solution, and derived the approximate upper bound for the step-size in ADALINE training. In [12] was developed an unified approach for mean-square convergence analysis for ADALINE training under MEE criterion. However, in these works, the case of non-Gaussian noise with a zero-symmetric distribution was considered.

In [13, 14], the concept of information theoretic learning (ITL) was introduced, using as a criterion the Rényi quadratic entropy, for which a nonparametric estimate based on Parzen windows with Gauss kernels is determined directly from data samples. In these works, it was proved that when using the Rényi entropy, as a result of training, the Rényi distance between the conditional probability of the density function of the desired and actual output signals for the given input signals is minimized.

The results of numerous studies [15,16] indicate that in the presence of non-Gaussian, in particular, impulse noise, in measurements, an approach based on information characteristics of signals is very effective, while a criterion that considers all statistics of a higher-order error signal turns out to be more appropriate.

Recently, when solving problems of identification, filtration, etc. robust algorithms obtained not on the basis of minimization (3), but on the basis of maximizing the correntropy criteria [17, 18] are gaining popularity. The maximum correntropy criterion (MCC) has recently triggered enormous research activities in engineering and machine learning communities since it is robust when faced with heavy-tailed noise or outliers in practice.

In [19] an indirect adaptive control system by reference model was proposed for modeling and controlling hydraulic pressure in water supply systems. The paper [20] is interested in distributed MCC algorithms, based on a divide-and-conquer strategy, which can deal with big data efficiently. In [21] a proportionate-type normalized maximum correntropy criterion with a correntropy induced metric zero attraction terms is presented, whose performance is also discussed for identifying sparse systems. In [22] the shortcomings of existing performance assessment methods and indicators are summarized firstly, and a novel evaluation method based on generalized correntropy criterion is proposed to evaluate the performance of non-Gaussian stochastic system.

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

The algorithms that maximize correntropy, are simple to implement and efficient.

The paper structure is the following. Section 1 discusses the correntropy existing of algorithms that maximize correntropy. In section 2 are given correntropy learning algorithms with variable center. Sections 3 and 4 describe the study of the issues of convergence of the algorithm in a stationary and non-stationary cases. Numerical experiments are discusses in section 5. The last section concludes research and describes future steps.

1. Formulation of the problem

Correntropy, defined as a localized measure of similarity, has proven to be very effective for obtaining robust estimates due to the fact that it is less sensitive to outliers [15, 16].

For two random variables X and Y, the correntropy is defined as

$$V(X,Y) = M\{k_\sigma(X,Y)\}, \qquad (7)$$

where $k_\sigma(\bullet)$ is the rotation invariant Mercer kernels; σ is the kernel width. The most widely used in calculating the correntropy are Gaussian, defined by the formula

$$k_\sigma(x,y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-y)^2}{2\sigma^2}\right). \qquad (8)$$

When calculating the correntropy, it is necessary to know the joint distribution of random variables X and Y, which, as a rule, is not known. In practice, there are usually a finite number of samples $\{x_i, y_i\}, i = 1,2,\ldots,N.$ Therefore, the most simple estimate of the correntropy is calculated as follows:

$$\hat{V}(X,Y) = \frac{1}{N}\sum_{i=1}^{N} k_\sigma(x_i - y_i). \qquad (9)$$

In tasks of identification, filtering, etc. as a functional, the correntropy between the required output signal d_i and the output signal of the model (real) y_i is used. When using Gaussian kernels, the optimized functional takes the form

$$J_{corr}(n) = \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{N}\sum_{i=n-N+1}^{N}\exp\left(-\frac{e_i^2}{2\sigma^2}\right), \qquad (10)$$

where $e_i = d_i - y_i$ is the identification error.

It is easy to see that the choice $\rho(e_i) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{e_i^2}{2\sigma^2}\right)$ satisfies the above requirements for $\rho(e_i)$.

The gradient optimization algorithm (10) at $N = 1$ will have the form [15, 16]

$$w_{n+1} = w_n + \gamma\exp\left(-\frac{e_{n+1}^2}{2\sigma^2}\right)e_{n+1}x_{n+1}, \qquad (11)$$

where γ is the parameter that affects the rate of convergence.

This algorithm has a significant drawback - a low convergence rate, which significantly limits the possibility of its use in research of non-stationary objects. It should be noted that finding the optimal value of the parameter γ that provides the maximum convergence rate of the algorithm equal, as it is easy to show,

$$\gamma_{n+1} = \left(\psi_{n+1}\|x_{n+1}\|^2\right)^{-1}, \qquad (12)$$

where $\psi_{n+1} = \exp\left(-\frac{e_{n+1}^2}{2\sigma^2}\right)$.

In [17], to combat impulse noise, a recurrent weighted least squares method (RLS) was proposed, which minimizes the criterion

$$\psi_{n+1} = \exp\left(-\frac{e_{n+1}^2}{2\sigma^2}\right) \qquad (13)$$

and having the form

$$c_{n+1} = c_n + \\ + \frac{\psi_{n+1}P_n x_{n+1}}{\lambda + \psi_{n+1}x_{n+1}^T P_n x_{n+1}}(y_{n+1} - c_n^T x_{n+1}); \qquad (14)$$

$$P_{n+1} = \lambda^{-1}\left(P_n - \frac{\psi_{n+1}P_n x_{n+1}x_{n+1}^T P_n}{\lambda + \psi_{n+1}x_{n+1}^T P_n x_{n+1}}\right). \qquad (15)$$

Here $0 \le \lambda < 1$ is the weighing factor.

Thus, when deriving the formula for calculating P_{n+1} (15), we used the approximation

$$P_{n+1} = \lambda P_n + \psi_{n+1}x_{n+1}x_{n+1}^T. \qquad (16)$$

As you know, introducing a parameter λ into an algorithm is advisable when identifying non-stationary parameters. Another approach used to estimate non-stationary parameters is the use of a limited number of measurements in RLS, which leads to the algorithm of the current regression analysis method.

The papers [18] considered the issue of configuring all network parameters (weights, radii and window width). In [23], the LSM was used for this purpose. It is known that this method is inconvenient when constructing a model in real time and when estimating non-stationary parameters.

When using the criterion of maximum correlation, the optimal model M^* is obtained as follows

$$M^* = \underset{M \in M}{\operatorname{argmax}} V_\sigma(T,Y) = M\{G_\sigma(e)\}, \qquad (17)$$

where M is the space of models $V_\sigma(T,Y)=M\{G_\sigma(e)\}$ is the correntropy between target T and input signals Y; $e=T-Y$ is the error; is the $G_\sigma(e)$ is the Gaussian kernel

$$G_\sigma(x,y)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{e^2}{2\sigma^2}\right). \qquad (18)$$

Since a function $G_\sigma(e)$ is a local function of error e, correntropy can be used as an indicator of error in information processing and machine learning problems. It can be seen from (18) that the center of the Gaussian nucleus is at zero. This circumstance can lead to the fact that if the distribution of errors (noise) has a nonzero mean, function (18) will not correspond to this distribution. Therefore, the problem arises of choosing such a correntropy function that would be effective for noises having a nonzero mean.

One of the approaches to solving this problem is the use of correntropy with variable center [24, 25]

$$V_{\sigma,c}(T,Y)=M\{G_{\sigma,c}(e)\}G_\sigma(x,y)=$$
$$=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\{e-c\}^2}{2\sigma^2}\right), \qquad (19)$$

where $c\in R$ is the center.

Questions of practical application of such algorithms are considered in [26, 27].

The main task of this paper is to develop algorithm for robust identification which maximizes the criterion of correntropy with variable center and investigation of its convergence performance in the stationary and non-stationary cases in conditions of non-Gaussian noises. We construct verification rules and illustrate results of checking as well.

2. Correntropy learning algorithms

Using (19) we have

$$V_{\sigma,c}(T,Y)=$$
$$=\frac{1}{\sqrt{2\pi}\sigma}\sum_{n=0}^{\infty}\frac{(-1)^n}{2^n n!}M\left(\frac{(e-c)^{2n}}{\sigma^{2n}}\right). \qquad (20)$$

When increasing σ, the moments of higher orders relative to the center will decrease faster, therefore, the moment of the second order will prevail in the value $V_{\sigma,c}(T,Y)$. In particular, for $c=M\{e\}$ and $\sigma\to\infty$, maximizing the center c correntropy is equivalent to minimizing the error variance.

Minimizing functional (19) with respect to the parameters of the model, we obtain

$$\frac{\partial E_{n+1}}{\partial w}=-\exp\left(\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)\frac{(e_{n+1}-c)}{2\sigma^2}x_{n+1}; \qquad (21)$$

$$\frac{\partial E_{n+1}}{\partial c}=w\exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)\frac{(e_{n+1}-c)}{\sigma^2}; \qquad (22)$$

$$\frac{\partial E_{n+1}}{\partial \sigma^2}=$$
$$=-w\exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)\frac{(e_{n+1}-c)^2}{\sigma^3}. \qquad (23)$$

Taking these expressions into account, the algorithms for correcting the network parameters will have the form

$$w_{n+1}=w_n+$$
$$+\gamma_w\exp\left(-\frac{(e_{n+1}-c_{n+1})^2}{2\sigma_{n+1}^2}\right)(e_{n+1}-c_{n+1})x_{n+1}, \qquad (24)$$

$$c_{n+1}=c_n+$$
$$+\gamma_c\exp\left(-\frac{(e_{n+1}-c_{n+1})^2}{2\sigma_{n+1}^2}\right)(e_{n+1}-c_n); \qquad (25)$$

$$\sigma_{n+1}^2=\sigma_n^2-$$
$$-\gamma_\sigma w_{n+1}\exp\left(-\frac{(e_{n+1}-c_{n+1})^2}{2\sigma_n^2}\right)\frac{(e_{n+1}-c_{n+1})^2}{\sigma_n^3}, \qquad (26)$$

where $\gamma_w,\gamma_c,\gamma_\sigma$ are the algorithm parameters that regulate the step size and affect the rate of its convergence.

If the object under study has several outputs, then the output signal will be a vector signal and the error will also be a vector value.

$$w_{n+1}=w_n+\gamma\exp\left(-\|e_{n+1}-c\|_{R^{-1}}^2\right)e_{n+1}x_{n+1}, \qquad (27)$$

where $\|e_{n+1}-c\|_{R^{-1}}^2=(e_{n+1}-c)^T R^{-1}(e_{n+1}-c)$; R^{-1} is the covariance matrix.

$$R_{n+1}^{-1}=R_n^{-1}-$$
$$-\gamma_R w_{n+1}\exp\left(-\|e_{n+1}-c_{n+1}\|_{R_n^{-1}}^2\right) \qquad (28)$$
$$\times(e_{n+1}-c_{n+1})(e_{n+1}-c_{n+1})^T.$$

Comparing (24)-(26) with (11), we see that the implementation of the proposed algorithm with parameters settings c and σ (or R^{-1}) is somewhat more complicated than the implementation of a simple gradient algorithm. It is clear that such a complication requires additional memory costs for storing these parameters.

Consider the issues of convergence of the proposed algorithm.

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

3. Study of the issues of convergence of the algorithm in a stationary case

Consider the estimation error

$$\Theta_{n+1} = c_{n+1} - c^*. \quad (29)$$

Then

$$e_{n+1} = \Theta_{n+1}^T x_{n+1} + \xi_{n+1} = e_{n+1}^a + \xi_{n+1}, \quad (30)$$

where $e_{n+1}^a = \Theta_{n+1}^T x_{n+1}$ is the aprior error.

In this case, the estimation algorithm can be written as

$$w_{n+1} = w_n + \gamma f(e_{n+1}) x_{n+1}, \quad (31)$$

where $f(e_{n+1}) = \exp\left(\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)(e_{n+1}-c)$.

Writing down algorithm (31) with respect to estimation errors, we have

$$\theta_{n+1} = \theta_n - \gamma f(e_{n+1}) x_{n+1}.$$

Multiplying both sides of this expression on the left θ_{n+1}^T by we get

$$\|\theta_{n+1}\|^2 = \|\theta_n\|^2 - 2\gamma f(e_{n+1}) e_{n+1}^a + \gamma^2 f^2(e_{n+1}) \|x_{n+1}\|^2. \quad (32)$$

Averaging both sides of (32), i.e.

$$M\left\{\|\theta_{n+1}\|^2\right\} = M\left\{\|\theta_n\|^2\right\} - 2\gamma M\left\{f(e_{n+1}) e_{n+1}^a\right\} + \gamma^2 M\left\{f^2(e_{n+1}) \|x_{n+1}\|^2\right\}, \quad (33)$$

we obtain the condition for the convergence of algorithm (31) in the mean square

$$0 < \gamma \le \frac{2M\left\{f(e_{n+1}) e_{n+1}^a\right\}}{M\left\{f^2(e_{n+1}) \|x_{n+1}\|^2\right\}}. \quad (34)$$

Consider a steady state. Since in steady state

$$\lim_{n\to\infty} M\left\{\|\theta_{n+1}\|^2\right\} = \lim_{n\to\infty} M\left\{\|\theta_n\|^2\right\},$$

it follows from (33) that

$$2\lim_{n\to\infty} M\left\{e_{n+1}^a f(e_{n+1})\right\} = \gamma \operatorname{tr} R_x \lim_{n\to\infty} M\left\{f^2(e_{n+1})\right\}, \quad (35)$$

where $R_x = M\left\{x_{n+1} x_{n+1}^T\right\}$ is the covariance matrix of input signals.

To calculate the steady-state value of the estimation error, we define $M\left\{f^2(e_{n+1})\|x_{n+1}\|^2\right\}$ и $M\left\{f^2(e_{n+1})\right\}$.

Consider the case of Gaussian noise ($\xi \sim N\left(0, \sigma_\xi^2\right)$). Using Price's theorem [28], we obtain

$$\lim_{n\to\infty} M\left\{e_{n+1}^a f(e_{n+1})\right\} = \lim_{n\to\infty} M\left\{e_{n+1}^a f(e_{n+1}^a + \xi_{n+1})\right\} =$$
$$= \lim_{n\to\infty} M\left\{\left(e_{n+1}^a\right)^2\right\} M\left\{f'(e_{n+1})\right\} =$$
$$= \lim_{n\to\infty} SM\left\{\exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)\left(1-\frac{(e_{n+1}-c)^2}{\sigma^2}\right)\right\} = \quad (36)$$
$$= \frac{S}{\sqrt{2\pi}\sigma_e} \lim_{n\to\infty} \int_{-\infty}^{+\infty} \exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)\left(1-\frac{(e_{n+1}-c)^2}{\sigma^2}\right) * \exp\left(-\frac{(e_{n+1}-c_e)^2}{2\sigma_e^2}\right) de_{n+1},$$

where

$$\sigma_e^2 = M\left\{\left(e_{n+1}^a\right)^2 + \sigma_\xi^2\right\}; \quad S = \lim_{n\to\infty} M\left\{\left(a_{n+1}^a\right)^2\right\};$$

c_e is the the center of the Gaussian error e_{n+1}.

Similarly, we define

$$M\left\{f^2(e_{n+1})\right\} = \lim_{n\to\infty} M\left\{\exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)(e_{n+1}-c)^2\right\} =$$
$$= \frac{1}{\sqrt{2\pi}\sigma_e} \lim_{n\to\infty} \int_{-\infty}^{+\infty} \exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)(e_{n+1}-c)^2 \times \exp\left(-\frac{(e_{n+1}-c_e)^2}{2\sigma_e^2}\right) de_{n+1}. \quad (37)$$

Substitution of (36) and (37) into (33) gives the expression for the steady-state error $\lim_{n\to\infty} M\left\{\left(e_{n+1}^a\right)^2\right\}$

$$\lim_{n\to\infty} M\left\{\left(e_{n+1}^a\right)^2\right\} = \frac{A}{2B}, \quad (38)$$

where

$$A = \gamma \operatorname{tr} R_x \lim_{n\to\infty} \int_{-\infty}^{+\infty} \exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)(e_{n+1}-c)^2 \times \exp\left(-\frac{(e_{n+1}-c_e)^2}{2\sigma_e^2}\right) de_{n+1};$$

$$B = \lim_{n\to\infty} \int_{-\infty}^{+\infty} \exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)\left(1-\frac{(e_{n+1}-c)^2}{\sigma^2}\right) \times \exp\left(-\frac{(e_{n+1}-c_e)^2}{2\sigma_e^2}\right) de_{n+1}.$$

or

$$\lim_{n\to\infty} M\left\{\left(e_{n+1}^a\right)^2\right\} = \frac{A}{B}, \quad (39)$$

where

$$A=\gamma \mathrm{tr} R_x \lim_{n\to\infty}\int_{-\infty}^{+\infty}\exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)(e_{n+1}-c)^2\times$$
$$\times\exp\left(-\frac{(e_{n+1}-c_e)^2}{2\sigma_e^2}\right)de_{n+1};$$
$$B=\lim_{n\to\infty}\int_{-\infty}^{+\infty}\exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)\left(1-\frac{(e_{n+1}-c)^2}{\sigma^2}\right)\times$$
$$\times\exp\left(-\frac{(e_{n+1}-c_e)^2}{2\sigma_e^2}\right)de_{n+1}.$$

This expression shows that $\lim_{n\to\infty} M\left\{\left(e_{n+1}^a\right)^2\right\}=0$ when choosing $\gamma\to 0$. Consider the case of non-Gaussian interference. In this case, we use the Taylor series expansion. In the steady state, the estimated parameters change (are corrected) insignificantly. Therefore, we can rewrite (33) as follows

$$2M\{e^a f(e)\}=\gamma \mathrm{tr} R_x M\{f^2(e)\} \quad (40)$$

We expand the function $f(e)$ in a Taylor series, limiting ourselves to terms of the second order of smallness

$$f(e)=f(e^a+\xi)=f(\xi)+$$
$$+f'(\xi)e^a+0.5f''(\xi)\left(e^a\right)^2+o\left(\left(e^a\right)^2\right), \quad (41)$$

where

$$f'(\xi)=\exp\left(-\frac{(\xi-c)^2}{2\sigma^2}\right)\left(1-\frac{(\xi-c)^2}{\sigma^2}\right); \quad (42)$$

$$f''(\xi)=$$
$$=\exp\left(-\frac{(\xi-c)^2}{2\sigma^2}\right)\left(\frac{(\xi-c)^3}{\sigma^4}-\frac{3(\xi-c)}{\sigma^2}\right). \quad (43)$$

Assuming that the interference does not correlate with the signals and the aprior error e^a, we can write

$$M\{e^a f(e)\}=$$
$$=M\left\{e^a f(\xi)+f'(\xi)\left(e^a\right)^2+o\left(e^a\right)^2\right\}\approx \quad (44)$$
$$\approx SM\{f'(\xi)\};$$

$$M\{f^2(e)\}\approx M\{f^2(\xi)\}+$$
$$+SM\left\{f(\xi)f''(\xi)+\left|f'(\xi)\right|^2\right\} \quad (45)$$

Substituting (44) and (45) into (41), we have

$$S=\frac{\gamma \mathrm{tr} R_x M\{f^2(\xi-c)\}}{A}, \quad (46)$$

where

$$A=2M\{f'(\xi-c)\}-$$
$$-\gamma \mathrm{tr} R_x M\left\{f(\xi-c)f''(\xi-c)+\left|f'(\xi-c)\right|^2\right\}.$$

Substitution of (42), (43) into (46) gives

$$S=\frac{\gamma \mathrm{tr} R_x M\left\{K(\xi-c)^2\right\}}{B}, \quad (47)$$

where

$$B=2M\left\{K'\left(1-\frac{(\xi-c)^2}{\sigma^2}\right)\right\}-$$
$$-\gamma \mathrm{tr} R_x M\left\{K\left(1+\frac{2(\xi-c)^4}{\sigma^4}-\frac{5(\xi-c)^2}{\sigma^2}\right)\right\};$$
$$K=\exp\left(-\frac{(\xi-c)^2}{\sigma^2}\right);\quad K'=\exp\left(-\frac{(\xi-c)^2}{2\sigma^2}\right).$$

The obtained conditions for the convergence of the proposed algorithm in the stationary case and the steady-state value of the estimation error depend on the type of noise distribution.

4. Study of the algorithm convergence in a non-stationary case

Let us assume that the estimated parameters are non-stationary, i.e.

$$c_{n+1}^*=c_n^*+\Delta c^*, \quad (48)$$

where $\Delta c^*=(\Delta c_1^*,\Delta c_2^*,\ldots,\Delta c_N^*)^T$ is the vector of a random sequence $N\times 1$ whose components have zero mathematical expectation, the correlation matrix of which is equal to $R_c=M\{c^* c^{*T}\}$.

Consider the error vector $\theta_{n+1}=c_{n+1}-c_{n+1}^*$

$$\theta_{n+1}=\theta_n-c_{n+1}^*+\gamma f(e_{n+1})x_{n+1}=$$
$$=\theta_n-\Delta c^*+\gamma f(e_{n+1})x_{n+1}, \quad (49)$$

Multiplying both sides of (49) on the left by θ_{n+1}^T and calculating the mathematical expectation, we obtain

$$M\left\{\|\theta_{n+1}\|^2\right\}=M\left\{\|\theta_n\|^2\right\}-2\gamma M\left\{x_{n+1}^T\theta_n f(e_{n+1})\right\}+$$
$$+\gamma^2 M\left\{f^2(e_{n+1})\|x_{n+1}\|^2\right\}+$$
$$+M\left\{\left\|\Delta c^*\right\|^2\right\}+M\left\{x_{n+1}^T\Delta c^*\right\}+M\left\{\Delta c^{*T}x_{n+1}\right\}-$$
$$-2\gamma M\left\{x_{n+1}^T\Delta c^* f(e_{n+1})\right\},$$

Taking into account the statistical properties of signals and noise, we have

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

$$M\left\{\left\|\theta_{n+1}\right\|^2\right\}=M\left\{\left\|\theta_n\right\|^2\right\}-\\-2\gamma M\left\{e^a_{n+1}f(e_{n+1})\right\}+\\+\gamma^2M\left\{f^2(e_{n+1})\left\|x_{n+1}\right\|^2\right\}+M\left\{\left\|\Delta c^*\right\|^2\right\}, \quad (50)$$

For Gaussian interference, using Price's theorem gives

$$\lim_{n\to\infty}M\left\{e^a_{n+1}f(e_{n+1})\right\}=\lim_{n\to\infty}M\left\{e^a_{n+1}f(e^a_{n+1}+\xi_{n+1})\right\}=\\=\lim_{n\to\infty}M\left\{\left(e^a_{n+1}\right)^2\right\}M\{f'(e_{n+1})\}=\\=\lim_{n\to\infty}SM\left\{\exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)\left(1-\frac{(e_{n+1}-c)^2}{\sigma^2}\right)\right\}=\\=\frac{S}{\sqrt{2\pi}\sigma_e}\lim_{n\to\infty}\int_{-\infty}^{+\infty}\exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)\left(1-\frac{(e_{n+1}-c)^2}{\sigma^2}\right)\times\\\times\exp\left(-\frac{(e_{n+1}-c_e)^2}{2\sigma_e^2}\right)de_{n+1}=\frac{S\sigma^3}{\left(\sigma^2+\sigma_\xi^2+S\right)^{\frac{3}{2}}} \quad (51)$$

$$M\left\{f^2(e_{n+1})\right\}=\\=\lim_{n\to\infty}M\left\{\exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)(e_{n+1}-c)^2\right\}=\\=\frac{1}{\sqrt{2\pi}\sigma_e}\lim_{n\to\infty}\int_{-\infty}^{+\infty}\exp\left(-\frac{(e_{n+1}-c)^2}{2\sigma^2}\right)(e_{n+1}-c)^2\times\\\times\exp\left(-\frac{(e_{n+1}-c_e)^2}{2\sigma_e^2}\right)de_{n+1}=\\=\frac{\sigma^3\left(S+\sigma_\xi^2\right)}{\left(2\sigma_\xi^2+\sigma^2+2S\right)^{\frac{3}{2}}}. \quad (52)$$

Considering that $M\left\{\left\|\Delta c^*\right\|^2\right\}=M\left\{\Delta c^*\Delta c^{*T}\right\}=trR_c$, for steady state when $\lim_{n\to\infty}M\left\{\left\|\theta_{n+1}\right\|^2\right\}=\lim_{n\to\infty}M\left\{\left\|\theta_n\right\|^2\right\}$ from expression (49) we obtain

$$\frac{2S}{\left(\sigma^2+\sigma_\xi^2+S\right)^{\frac{3}{2}}}=\\=\frac{\gamma trR_x\left(\sigma_\xi^2+S\right)}{\left(\sigma^2+2\sigma_\xi^2+2S\right)^{\frac{3}{2}}}+\frac{trR_c}{\gamma\sigma^3}. \quad (53)$$

From this ratio, we can determine the value S

$$S=\frac{\gamma trR_x\left(\sigma_\xi^2+S\right)\left(\sigma^2+\sigma_\xi^2+S\right)^{\frac{3}{2}}}{\left(\sigma^2+2\sigma_\xi^2+2S\right)^{\frac{3}{2}}}+\\+\frac{trR_c\left(\sigma^2+\sigma_\xi^2+S\right)^{\frac{3}{2}}}{2\gamma\sigma^3}. \quad (54)$$

For $\sigma^2\to\infty$, we have the value of S for the least squares

$$\lim_{\sigma\to\infty}S=\frac{\gamma trR_x\sigma_\xi^2+\gamma^{-1}trR_c}{2-\gamma trR_x}.$$

In the case of non-Gaussian noise, we have

$$M\left\{e^a_{n+1}f(e_{n+1})\right\}\approx\\\approx M\left\{e^a_{n+1}f(\xi_{n+1})+e^a_{n+1}f'(\xi_{n+1})\right\}\approx\\\approx SM\{f'(\xi_{n+1})\}. \quad (55)$$

$$M\left\{f^2(e_{n+1})\right\}\approx\\\approx M\left\{\left(f(\xi_{n+1})+e^a_{n+1}f'(\xi_{n+1})+0{,}5f''(\xi_{n+1})e^{a2}_{n+1}\right)^2\right\}\approx\\\approx M\left\{f^2(\xi_{n+1})\right\}+SM\left\{f(\xi_{n+1})f''(\xi_{n+1})+\left(f'(\xi_{n+1})\right)^2\right\}, \quad (56)$$

where

$$f'(\xi_{n+1})=\exp\left(-\frac{(\xi-c)^2}{2\sigma^2}\right)\left(1-\frac{(\xi-c)^2}{\sigma^2}\right);$$

$$f''(\xi_{n+1})=\exp\left(-\frac{(\xi_{n+1}-c)^2}{2\sigma^2}\right)\left(\frac{\xi^3_{n+1}}{\xi^4_{n+1}}-\frac{3\xi_{n+1}}{\sigma^2}\right).$$

Substituting (55) and (56) into (50), after simple transformations we obtain

$$S=\frac{\gamma A+\gamma^{-1}B}{C-\gamma D}, \quad (57)$$

where

$$A=trR_xM\left\{(\xi_{n+1}-c)^2\exp\left(-\frac{(\xi_{n+1}-c)^2}{\sigma^2}\right)\right\};$$

$$B=trR_c;$$

$$C=2M\left\{\left(1-\frac{(\xi_{n+1}-c)^2}{2\sigma^2}\right)\exp\left(-\frac{(\xi_{n+1}-c)^2}{\sigma^2}\right)\right\};$$

$$D=trR_xM\{F\},$$

$$F=\left(1+\frac{2(\xi_{n+1}-c)^4}{\sigma^4}-\frac{5(\xi_{n+1}-c)^2}{\sigma^2}\right)\times\\\times\exp\left(-\frac{(\xi_{n+1}-c)^2}{\sigma^2}\right).$$

This expression shows that S is a monotonically non-increasing function of the parameter γ .

From the condition $\partial S/\partial\gamma = 0$, an equation can be obtained to determine the optimal value of the parameter that provides the minimum value of S

$$AC\gamma^2 + BD\gamma - BC = 0.$$

As can be seen from the results presented in this section, the steady-state value of the estimation error depends not only on the type of noise distribution, but also on the degree of non-stationarity (Δc^*) of the estimated parameters.

5. Numerical experiments

Experiment 1. The problem of identification of a stationary linear object, which is described by equation (1) with the following parameters

$$\theta^* = (-128; -96; -85; -64; -57; -48; -31;$$
$$-21; 0; 2; 20; 32; 62; 97; 108; 127)^T$$

was considered.

Sequences of normally distributed quantities $x(k) \sim N(0;1)$ were chosen as the input signal $x(k)$.

When testing the robustness of the algorithms, an independent noise distributed according to the Rayleigh law with σ = 1 was added to the output signal of the object.

The histogram of such noise is shown in fig. 1. The simulation results for various values of the parameter are shown in fig. 2.

This figure shows the graphs of changes in the error when choosing the LSM algorithm and algorithm (31) respectively, here

$$RMSE = \sqrt{\frac{1}{2}\left\|c_n - c^*\right\|^2},$$

where c_n and c^* denote estimated and target parameters vectors respectively.

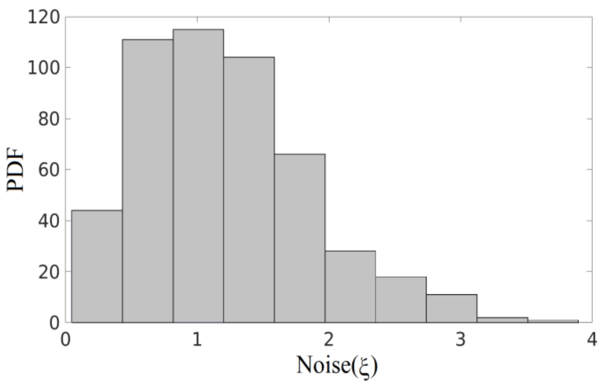

Fig. 1. Noise distribution

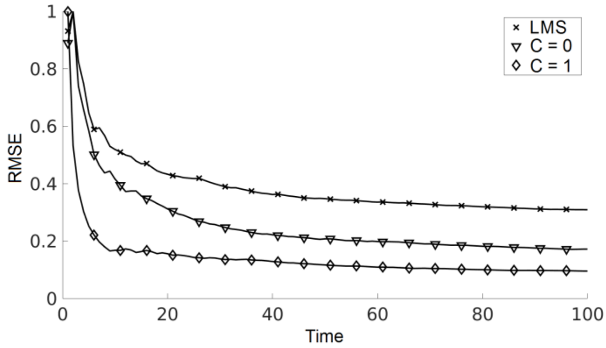

Fig. 2. Different algorithms results

Experiment 2. The comparison analysis was performed to compare robust properties of the LSM algorithm and algorithm (31) with different parameters. Parameters of the linear object was taken the same as in the Experiment 1. The results of the experiment are shown in the Table 1. In this table the values of the normalised error after 100 iteratios of each algorithm simulation are presented.

Table 1

The Experiment 2 results

Algorithm/Noise	**Rayleigh,** σ = 1	**Laplas,** σ = 3	**Normal,** σ = 3
LMS	0.36	0.29	0.21
(31) with c = 0	0.12	0.08	0.054
(31) with c = 1	0.015	0.009	0.006

As it can be seen from the Table 1 the (31) algorithm with c = 1 performs well for all considered types of noise.

6. Discussion of the results

As the research results have shown, the use of the correntropy functional for teaching ADALINE objects in conditions of non-Gaussian noise makes it possible to obtain robust estimates. The conditions for the convergence of the gradient algorithm for learning a non-stationary object in the presence of Gaussian and non-Gaussian measurement noises are obtained, which are determined by expression (34). In addition, a relation was derived for the maximum achievable (asymptotic) a priori error $\lim_{n\to\infty} M\left\{\left(e^a_{n+1}\right)^2\right\}$ in the presence of Gaussian noise in the form of formula (38). This expression shows that $\lim_{n\to\infty} M\left\{\left(e^a_{n+1}\right)^2\right\}$ when choosing $\gamma \to 0$.

The use of the Taylor series expansion made it possible to obtain relations (46), (47) characterizing the asymptotic learning error in the presence of non-Gaussian measurement noises.

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

Similar estimates are obtained for the non-stationary case.

In addition, an equation is given, the solution of which provides the minimum value of S.

The estimates obtained are quite general and depend both on the degree of non-stationarity of the object $M\left\{\left\|\Delta c^*\right\|^2\right\}$ and on the statistical characteristics of useful signals R_x and noises σ_ξ^2. Since these parameters are usually not known, for the practical application of the obtained relations, one should use the estimates of the indicated parameters. And so the training takes place in the on-line mode, you can apply any recurrent procedure for evaluating these parameters and use the resulting estimates for step-by-step refinement of the parameters included in the algorithms.

It should be noted that the estimates obtained in this work depend on the parameter used in the algorithm σ (kernel width) and γ (convergence rate coefficient), the problem of choosing values of which remains open.

However, these estimates allow the researcher, when solving practical problems, to preliminarily estimate the limiting capabilities of this algorithm and the effectiveness of its application.

The results of online learning simulations have shown the superiority and reliability of the new method.

Conclusions

In this paper, we have developed an adaptive robust identification algorithm under the maximum correntropy criterion with variable center.

The properties of its convergence in the stationary and non-stationary cases in conditions of non-Gaussian noises are investigated.

The estimates obtained are quite general and depend both on the degree of non-stationarity of the object and on the statistical characteristics of useful signals and interference.

The results of the current study presented in Fig. 1 and in Table 1 are expected to improve existing robust estimation-based data processing technologies and accelerate the development of new real-time computing applications.

In connection with the fact that the question of the optimal choice of parameter values σ and γ remains open, it seems important and expedient to conduct research in the direction of

1) studying the effectiveness of the developed approach in teaching in the non-stationary case, when a model other than the first-order Markov model is used to describe non-stationarity;

2) establishing the dependence of the speed of the learning algorithm on the degree of non-stationarity of the object under study, i.e.;

3) development of recommendations for choosing the optimal values of the parameters σ and γ or the rules for their correction.

Future research and development steps can be dedicated to computation of parameters of discrete atomic transform, which is a core of the algorithm DAC, as well as construction of verification rules.

In addition, it seems appropriate to apply the approach using the maximum correntropy criterion to the problems of image recognition and processing in the presence of non-Gaussian noise, considered in [29, 30].

Contribution of authors. Authors together formulated of the problem, developed main concept, models and algorithms, carried on and analysed results of numerical experiments. All authors have read and agreed to the published version of the manuscript.

References (GOST 7.1:2006)

1. *Huber, P. J. Robust Statistics [Text] / P. J. Huber. – N.Y. : John Wiley and Sons, 1981. – 317 p.*

2. *Robust Statistics. The Approach Based on Influence Functions. [Text] / F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, W. A. Stahel. – N.Y. : John Wiley and Sons, 1986. – 526 p.*

3. *Руденко, О. Г. Робастное обучение вейвлет-нейросетей [Text] / О. Г. Руденко, А. А. Бессонов // Проблемы управления и информатики. – 2010. – № 5. – С. 66–79.*

4. *Rudenko, O. Function approximation using robust radial basis function networks [Text] / O. Rudenko, O. Bezzonov // J. of Intelligent Learning Systems and Applications. – 2011. – Vol. 3, No 1. – P. 17–25. DOI: 10.4236/jilsa.2011.31003.*

5. *Chambers, J. A Robust Mixed-Norm Adaptive Filter Algorithm [Text] / J. Chambers, A. Avlonitis // IEEE Signal Processing Letters. – 1997. – Vol. 4, No 2. – P. 46–48. DOI: 10.1109/97.554469.*

6. *Papoulis, E. V. A Normalized Robust Mixed-Norm Adaptive Algorithm for System Identification [Text] / E. V. Papoulis, T. Stathaki // IEEE Signal Processing Letters. – 2004. – Vol. 11, No 1. – P. 56–59. DOI: 10.1109/LSP.2003.819353.*

7. Zerguine, A. A variable-parameter normalized mixed-norm (VPNMN) adaptive algorithm [Text] / A. Zerguine // EURASIP Journal on Advances in Signal Processing. – 2012. – Article no. 55. DOI: 10.1186/1687-6180-2012-55.

8. Робастна ідентифікація об'єктів за наявностю негаусівських завад [Text] / О. Г. Руденко, О. О. Безсонов, Н. М. Сердюк, К. О. Олійник, *О. С. Романюк // Бионика интеллекта. –2019. – №2 (93). – С. 7–12.*

9. Робастна ідентифікація об'єктів на основі мінімізації комбінованого функціоналу [Text] / О .Г. Руденко, О. О. Безсонов, Н. М. Сердюк, К. О. Олійник, О. С. Романюк // Системи обробки інформації. – 2020. – №. 1 (160). – С. 80–88.

10. Erdogmus, D. Generalized information potential criterion for adaptive system training [Text] / D. Erdogmus, J. C. Principe // IEEE Trans. Neural Netw. – 2002. – Vol. 13, No 5. – P. 1035–1044.

11. Erdogmus, D. Convergence properties and data efficiency of the minimum error entropy criterion in ADALINE training. [Text] / D. Erdogmus, J. C. Principe // IEEE Trans. Signal Process. – 2003. – Vol. 51, No 7. – P. 1966–1978.

12. Chen, B. Mean-Square Convergence Analysis of ADALINE Training With Minimum Error Entropy Criterion [Text] / B. Chen, Y. Zhu, J. Hu // IEEE Transactions on neural networks. – 2010. – Vol. 21, No 7. – P. 1168–1179.

13. Santamaría, I. Generalized Correlation Function: Definition, Properties, and Application to Blind Equalization [Text] / I. Santamaría, P. P. Pokharel, J. C. Principe // IEEE Trans. on Signal Processing. – 2006. – Vol. 54, No 6. – P. 2187–2197. DOI: 10.1109/TSP.2006.872524.

14. Liu, W. Correntropy: Properties and Applications in Non-Gaussian Signal Processing [Text] / W. Liu, P. P. Pokharel, J. C. Principe // IEEE Trans. on Signal Processing. – 2007. – Vol. 55, No 11. – P. 5286–5298. DOI: 10.1109/TSP.2007.896065.

15. Wang, W. An adaptive kernel width update method of correntropy for channel estimation [Text] / W. Wang, J. Zhao, H. Qu, B. Chen, J. C. Principe // IEEE International Conference on Digital Signal Processing (DSP). – 2015. – P. 916–920. DOI: 10.1109/ICDSP.2015.7252010.

16. Gunduz, A. Correntropy as a novel measure for nonlinearity tests [Text] / A. Gunduz, J. C. Principe // Signal Processing. – 2009. – Vol. 89. – P. 14–23. DOI: 10.1016/j.sigpro.2008.07.005.

17. Lu, L. Active impulsive noise control using maximum correntropy with adaptive kernel size [Text] / L. Lu, H. Zhao // Mechanical Systems and Signal Processing. – 2017. – Vol. 87, Part A. – P. 180–191. DOI: 10.1016/j.ymssp.2016.10.020.

18. Huang, F. Adaptive filtering under a variable kernel width maximum correntropy criterion [Text] / F. Huang, J. Zhang, S. Zhang // IEEE Transactions on Circuits and Systems II: Express Briefs. – 2017. – Vol. 64, No 10. – P. 1247–1251. DOI: 10.1109/TCSII. 2017.2671339.

19. Adaptive Pressure Control System Based on the Maximum Correntropy Criterion [Text] / T. K. S. Flores, J. M. M. Villanueva, H. P. Gomes, S. Y. C. Catunda // Sensors. – 2021. – Vol. 21, no. 15. – Article No. 5156. DOI: 10.3390/s21155156.

20. Maximum Correntropy Criterion with Distributed Method. [Text] / F. Xie, T. Hu, S. Wang, D. Wang // Mathematics. – 2022. – Vol. 10. – Article No. 304. DOI: 10.3390/math10030304.

21. Li, Y. A Proportionate Normalized Maximum Correntropy Criterion Algorithm with Correntropy Induced Metric Constraint for Identifying Sparse Systems. [Text] / Y. Li, Y. Wang, L. Sun // Symmetry. – 2018. – Vol. 10. – Article No. 683. DOI: 10.3390/sym10120683.

22. Zhang, J. Generalized Correntropy Criterion-Based Performance Assessment for Non-Gaussian Stochastic Systems. [Text] / J. Zhang, G. Huang, G. L. Zhan // Entropy. – 2021. – Vol. 23. 764. https://doi.org/10.3390/e23060764

23. Wang, X. Affine Projection Algorithm by Employing Maximum Correntropy Criterion for System Identification of Mixed Noise [Text] / X. Wang, J. Han // IEEE Access. – 2019. – Vol. 7. – P. 182515–182526.

24. Maximum Correntropy Criterion with Variable Center [Text] / B. Chen, X. Wang, Y. Li, J. C. Principe // IEEE Signal Process. Letter. – 2019. – Vol. 26, No 5. – P. 1212–1216. arXiv: 1904.06501v1 [stat.ML] 13 Apr 2019. DOI: 10.1109/LSP.2019.2925692.

25. Adaptive Filtering Under the Maximum Correntropy Criterion With Variable Center [Text] / L. Zhu, C. Song, L. Pan, J. Li // IEEE Access. – 2019. – P. 105902–105908. DOI: 10.1109/ACCESS.2019. 2932201.

26. Sparsity Constrained Recursive Generalized Maximum Correntropy Criterion With Variable Center Algorithm [Text] / Q. Sun, H. Zhang, X. Wang, W. Ma, B. Chen // IEEE Transactions on Circuits and Systems II: Express Briefs. – 2020. – Vol. 67, No 12. – P. 3517–3521. DOI: 10.1109/TCSII.2020.2986053.

27. Yang, J. Robust Maximum Mixture Correntropy Criterion-Based Semi-Supervised ELM With Variable Center [Text] / J. Yang, J. Cao, A. Xue // IEEE Transactions on Circuits and Systems II: Express Briefs. – 2020. – Vol. 67, No 12. – P. 3572–3576. DOI: 10.1109/TCSII.2020.2995419.

28. Price, R. A useful theorem for nonlinear devices having Gaussian inputs. [Text] / R. Price // IREN Trans. Inform Theory. – 1958. – No 4. – P. 69–72. DOI: 10.1109/TIT.1958.1057444.

29. Модель і метод навчання класифікатора контекстів спостереження на зображеннях відеоінспекції стічних труб [Text] / В. В. Москаленко, М. О. Зарецький, Я. Ю. Ковальський, С. С. Мартиненко // Радіоелектронні і комп'ютерні системи. – 2020. – № 3. – С. 59-66. DOI: 10.32620/reks.2020.3.06.

30. Naumenko, I. Information-extreme machine training of on-board recognition system with optimization of RGB-component digital images[Text] / I. Naumenko, M. Myronenko, T. Savchenko // Радіоелектронні і комп'ютерні системи. – 2021. – № 4. – С. 59-70. DOI: 10.32620/reks.2021.4.05.

References (BSI)

1. Huber, P. J. *Robust Statistics*. N.Y., John Wiley and Sons Publ., 1981. 317 p.

2. Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., Stahel, W. A. *Robust Statistics. The Approach Based on Influence Functions.* N.Y., John Wiley and Sons Publ., 1986. 526 p.

3. Rudenko, O. G., Bezsonov, O. O. Robastnoe obuchenie veivlet-neirosetei [Robust learning of wavelet neural networks]. *Problems of control and informatics*, 2010, no. 5, pp. 66–79.

4. Rudenko, O. G., Bezsonov, O. O. Function approximation using robust radial basis function networks. *J. of Intelligent Learning Systems and Applications*, 2011, vol. 3, no. 1, pp. 17–25. DOI: 10.4236/jilsa.2011.31003.

5. Chambers, J., Avlonitis, A. A Robust Mixed-Norm Adaptive Filter Algorithm. *IEEE Signal Processing Letters*, 1997, vol. 4, iss. 2, pp. 46–48. DOI: 10.1109/97.554469.

6. Papoulis, E. V., Stathaki, T. A Normalized Robust Mixed-Norm Adaptive Algorithm for System Identification. *IEEE Signal Processing Letters,*2004, vol. 11, no. 1, pp. 56–59. DOI: 10.1109/LSP.2003.819353.

7. Zerguine, A. A variable-parameter normalized mixed-norm (VPNMN) adaptive algorithm. *EURASIP Journal on Advances in Signal Processing*, 2012, article no. 55. DOI: 10.1186/1687-6180-2012-55.

8. Rudenko, O. G., Bezsonov, O. O., Serdiuk, N. N., Oliinyk, K. O., Romanyk, O. S. Robastna identyfikatsiya ob"yektiv za nayavnostyu nehausivs'kykh zavad [Robust identification of objects in the presence of non-Gaussian interference]. *Bionics of the intellect,* 2019, no. 2 (93), pp. 7–12.

9. Rudenko, O. G., Bezsonov, O. O., Serdiuk, N. N., Oliinyk, K. O., Romanyk, O. S. Robastna identyfikatsiya ob"yektiv na osnovi minimizatsiyi kombinovanoho funktsionalu [Robust object identification based on minimization of combined functional]. *Information processing systems*, 2020, no. 1 (160), pp. 80–88.

10. Erdogmus, D., Principe, J. C. Generalized information potential criterion for adaptive system training. *IEEE Trans. Neural Netw.*, 2002, vol. 13, iss. 5, pp. 1035–1044.

11. Erdogmus, D., Principe, J. C. Convergence properties and data efficiency of the minimum error entropy criterion in ADALINE training. *IEEE Trans. Signal Process.*, 2003, vol. 51, iss. 7, pp. 1966–1978.

12. Chen, B., Zhu, Y., Hu, J. Mean-Square Convergence Analysis of ADALINE Training with Minimum Error Entropy Criterion. *IEEE Transactions on Neural Networks*, 2010, vol. 21, iss. 7, pp. 1168–1179.

13. Santamaría, I., Pokharel, P. P., Principe, J. C. Generalized Correlation Function: Definition, Properties, and Application to Blind Equalization. *IEEE Trans. on Signal Processing,* 2006. vol. 54, iss. 6, pp. 2187–2197. DOI: 10.1109/TSP.2006.872524.

14. Liu, W., Pokharel, P. P., Principe, J. C. Correntropy: Properties and Applications in Non-Gaussian Signal Processing. *IEEE Trans. on Signal Processing,* 2007, vol. 55, no. 11, pp. 5286–5298. DOI: 10.1109/TSP.2007.896065.

15. Wang, W., Zhao, J., Qu, H., Chen, B., Principe J. C. An adaptive kernel width update method of correntropy for channel estimation. *IEEE International Conference on Digital Signal Processing (DSP),* 2015, pp. 916–920. DOI: 10.1109/ICDSP.2015.7252010.

16. Gunduz, A., Principe, J. C. Correntropy as a novel measure for nonlinearity tests. *Signal Processing,* 2009, vol. 89, pp. 14–23. DOI: 10.1016/j.sigpro.2008.07.005.

17. Lu, L., Zhao, H. Active impulsive noise control using maximum correntropy with adaptive kernel size. *Mechanical Systems and Signal Processing,* 2017, vol. 87, Part A, pp. 180–191. DOI: 10.1016/j.ymssp.2016.10.020.

18. Huang, F., Zhang, J., Zhang, S. Adaptive filtering under a variable kernel width maximum correntropy criterion. *IEEE Transactions on Circuits and Systems II: Express Briefs,* 2017, vol. 64, no. 10, pp. 1247–1251. DOI: 10.1109/TCSII.2017.2671339.

19. Flores, T. K. S., Villanueva, J. M. M., Gomes, H. P., Catunda, S. Y. C. Adaptive Pressure Control System Based on the Maximum Correntropy Criterion. *Sensors,* 2021, vol. 21, no. 15, article no. 5156. DOI: 10.3390/s21155156.

20. Xie, F., Hu, T., Wang, S., Wang, D. Maximum Correntropy Criterion with Distributed Method. *Mathematics*, 2022, vol. 10, article no. 304. DOI: 10.3390/math10030304.

21. Li, Y., Wang, Y., Sun, L. A Proportionate Normalized Maximum Correntropy Criterion Algorithm with Correntropy Induced Metric Constraint for Identifying Sparse Systems. *Symmetry,* 2018, vol. 10, article no. 683. DOI: 10.3390/sym10120683.

22. Zhang, J., Huang, G., Zhan, L. Generalized Correntropy Criterion-Based Performance Assessment for Non-Gaussian Stochastic Systems. *Entropy,* 2021, vol. 23, article no. 764. DOI: 10.3390/e23060764.

23. Wang, X., Han, J. Affine Projection Algorithm by Employing Maximum Correntropy Criterion for

System Identification of Mixed Noise. *IEEE Access*, 2019, vol. 7, pp. 182515–182526.

24. Chen, B., Wang, X., Li, Y., Principe, J. C. Maximum Correntropy Criterion with Variable Center. *IEEE Signal Process. Letter*, 2019, vol. 26, no. 5, pp. 1212-1216 arXiv:1904.06501v1 [stat.ML] 13 Apr 2019. DOI: 10.1109/LSP.2019.2925692.

25. Zhu, L., Song, C., Pan, L., Li, J. Adaptive Filtering Under the Maximum Correntropy Criterion With Variable Center. *IEEE Access*, 2019, pp. 105902–105908. DOI: 10.1109/ACCESS.2019.2932201.

26. Sun, Q., Zhang, H., Wang, X., Ma, W., Chen, B. Sparsity Constrained Recursive Generalized Maximum Correntropy Criterion With Variable Center Algorithm. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2020, vol. 67, no. 12, pp. 3517–3521. DOI: 10.1109/TCSII.2020.2986053.

27. Yang, J., Cao, J., Xue, A. Robust Maximum Mixture Correntropy Criterion-Based Semi-Supervised ELM With Variable Center. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2020, vol. 67, no. 12, pp. 3572–3576. DOI: 10.1109/TCSII.2020. 2995419.

28. Price, R. A useful theorem for nonlinear devices having Gaussian inputs. *IREN Trans. Inform Theory*, 1958, no. 4, pp. 69–72. DOI: 10.1109/TIT.1958.1057444

29. Moskalenko, V. V., Zarets'kyy, M. O. Koval's'kyy, Ya. Yu., Martynenko, S. S. Model' i metod navchannya klasyfikatora kontekstiv sposterezhennya na zobrazhennyakh videoinspektsiyi stichnykh trub [Model and method of training the classifier of observation context on video inspection images of sewer pipes]. *Radioelektronni i komp'uterni sistemi – Radioelectronic and computer systems*, 2020, no.3, pp. P.59 –66. doi: 10.32620/reks.2020.3.06

30. Naumenko, I., Myronenko, M., Savchenko, T. Information-extreme machine training of on-board recognition system with optimization of RGB-component digital images. *Radioelektronni i komp'uterni sistemi – Radioelectronic and computer systems*, 2021, no. 4, pp. 59-70. DOI: 10.32620/reks.2021.4.05.

Надійшла до редакції 4.12.2021, розглянута на редколегії 16.02.2022

АДАПТИВНА ІДЕНТИФІКАЦІЯ ЗА КРИТЕРІЄМ МАКСИМАЛЬНОЇ КОРЕНТРОПІЇ ЗІ ЗМІННИМ ЦЕНТРОМ

О. Г. Руденко, О. О. Безсонов

Розглядається задача ідентифікації параметрів лінійного об'єкта за наявністю негаусівських завад. Алгоритм ідентифікації є градієнтною процедурою максимізації функціоналу, що являє собою корентропію. Такий функціонал дозволяє отримати оцінки, які володіють робастними властивостями. На відміну від гаусівських ядер, що зазвичай застосовуються, центри яких знаходяться в нулі і ефективних для розподілів з нульовим середнім, у роботі розглядається модифікація критерію, придатна для розподілів з ненульовим середнім. Модифікація полягає у використанні корентропії зі змінним центром. Використання гаусівських ядер зі змінним центром дозволить оцінювати невідомі параметри за умов гаусівських та негаусівських завад, що мають розподіли з нульовим та ненульовим середнім і забезпечити можливість розробки нових технологій аналізу та обробки даних. Важливим є розроблення алгоритму робастної ідентифікації на основі корентропії зі змінним центром. Їх властивості при ідентифікації стаціонарних та нестаціонарних об'єктів є **предметом** дослідження. **Метою** дослідження є розроблення алгоритму робастної ідентифікації, що максимізує критерій корентропії зі змінним центром з використанням процедур налаштування центрів та ширини ядра та дослідження його збіжності в стаціонарному та нестаціонарному випадках за умов негаусівських завад. Отримано вирази для значень сталого стану помилки оцінювання, які залежать від виду розподілу завад та степені нестаціонарності параметрів, що оцінюються. Розв'язуються такі **завдання**: дослідити збіжність алгоритму та визначити умови сталості усталеного процесу ідентифікації. Використовуються **методи** теорії оцінювання (ідентифікації) та теорії ймовірності. Отримано такі **результати:** 1) розроблений алгоритм забезпечує отримання робастних оцінок за наявності завад, що мають розподіл з нульовим та ненульовим середнім; 2) досліджено його збіжність у стаціонарному та нестаціонарному випадках в умовах гаусівських та негаусівських завад; 3) проведено імітаційне моделювання роботи алгоритму. **Висновки:** результати дослідження дозволять покращити існуючі технології обробки даних на основі робастних оцінок та прискорять розробку нових обчислювальних програм у реальному часі.

Ключові слова: корентропія; максимізація; функціонал; градієнтний алгоритм; асимптотична оцінка; збіжність; точність ідентифікації; усталений режим.

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

АДАПТИВНАЯ ИДЕНТИФИКАЦИЯ ПО КРИТЕРИЮ МАКСИМАЛЬНОЙ КОРЕНТРОПИИ С ПЕРЕМЕННЫМ ЦЕНТРОМ

О. Г. Руденко, А. А. Бессонов

Рассматривается задача идентификации параметров линейного объекта при наличии негауссовских помех. Алгоритм идентификации является градиентной процедурой максимизации функционала, представляющего собой коррэнтропию. Такой функционал позволяет получить оценки, обладающие робастными свойствами. В отличие от обычно применяемых гауссовских ядер, центры которых находятся в нуле и эффективных для распределений с нулевым средним, в работе рассматривается модификация критерия, пригодная для распределений с ненулевым средним. Модификация заключается в использовании коррэнтропии с переменным центром Использование гауссовских ядер с переменным центром позволит оценивать неизвестные параметры в условиях гауссовских и негауссовских помех, имеющих распределение с нулевым и ненулевым средним, и обеспечить возможность разработки новых технологий анализа и обработки данных. Это объясняет важность разработки алгоритмов робастной идентификации на основе коррэнтропии с переменным центром Получены выражения для значений величин устойчивого состояния ошибок оценивания, которые зависят от вида распределения помех и степени нестационарности оцениваемых параметров. Их свойства при идентификации стационарных и нестационарных объектов являются **предметом** исследования. **Целью** исследования является разработка алгоритма робастной идентификации, максимизирующего критерий коррэнтропии с переменным центром с применением процедур настройки центров и ширины ядра и исследование его сходимости в стационарном и нестационарном случаях в условиях негауссовских помех. Получены выражения для установившегося значения ошибки оценивания, которые зависят от вида распределения помех и степени нестационарности оцениваемых параметров. Решаются следующие **задачи:** исследовать сходимость алгоритма и определить условия постоянства установившегося процесса идентификации. Используются **методы** теории оценивания (идентификации) и теории вероятности. Получены следующие **результаты:** 1) разработанный алгоритм обеспечивает получение робастных оценок при наличии помех, имеющих распределение с нулевым и ненулевым средним; 2) исследована его сходимость в стационарном и нестационарном случаях в условиях гауссовских и негауссовских помех; 3) проведено имитационное моделирование работы алгоритма. **Выводы:** результаты исследования позволят улучшить существующие технологии обработки данных на основе робастных оценок и ускорят разработку новых вычислительных программ в реальном времени.

Ключевые слова: коррэнтропия; максимизация; функционал; градиентный алгоритм; асимптотическая оценка; сходимость; точность идентификации; установившийся режим.

Руденко Олег Григорійович – д-р техн. наук, проф., зав. каф. комп'ютерних інтелектуальних технологій та систем, Харківський національний університет радіоелектроніки, Харків, Україна.

Безсонов Олександр Олександрович – д-р техн. наук, проф., проф. каф. комп'ютерних інтелектуальних технологій та систем, Харківський національний університет радіоелектроніки, Харків, Україна.

Oleg Rudenko – Doctor of Technical Science, Professor, Head of Department Computer Intelligent Technologies and Systems, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine,
e-mail: oleh.rudenko@nure.ua, Scopus Author ID: 7005837246, ORCID: 0000-0003-0859-2015.

Oleksandr Bezsonov – Doctor of Technical Science, Professor, Professor of Department Computer Intelligent Technologies and Systems, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine,
e-mail: oleksandr.bezsonov@nure.ua, Scopus Author ID: 57221227175, ORCID: 0000-0001-6104-4275.