Conditional averaging models of exponentially correlated data

Abstract. The article investigates a real model of exponential correlation which is obtained by signals with a limited band with features of white noise passing through physical inertial systems. In the investigated models, the subsequent conditionally averaged implementations of the signal exceeding the threshold \( x_p \) with a derivative with a random sign are significantly correlated, while the implementations of the signal exceeding the threshold \( x_p \) with a derivative with a given sign can be practically considered as uncorrelated. Such a method of conditional averaging can be recommended in practical applications.

Streszczenie. W artykule badano realny model skorelowania wykładniczego, który uzyskują sygnały z ograniczonym pasmem o cechach szumu białego, przechodzące przez fizyczne układy inercyjne. W badanych modelach kolejne warunkowo uśrednione realizacje sygnału przekraczające próg \( x_p \) z pochodną o dowolnym znaku są istotnie skorelowane a realizacje sygnału przekraczające próg \( x_p \) z pochodną jednego znaku można praktycznie uznać za nieskorelowane. Taki sposób uśredniania może być zalecany w zastosowaniach praktycznych. (Modele warunkowego uśredniania danych skorelowanych wykładniczo).

Keywords: random signal, exponential correlation function, conditional averaging.

Introduction

The correlation of signals and measurement data impedes the estimation of their characteristic statistics. In the majority of existing metrological documents and recommendations intended for use the influence of data correlation is most often not taken into account. Many authors point out in their specialist publications (e.g. [1-3]) to the necessity of accounting for the influence of autocorrelation in the evaluation of measurement uncertainty.

The exponential shape of the autocorrelation function is a relatively prevalent model in the processing and description of analogue stochastic signals. In practice, the model of exponential correlation is obtained by signals with a limited bandwidth with the features of white noise passing through physical inertial systems. Due to the central limit theorem and the inertia of typical processing systems, distributions of physical signals are usually normal or quasi-normal.

In measurement tasks, in order to evaluate data one can make use of conditional probability characteristics, and in particular conditional expected values and conditional variances [4].

In algorithms of conditional averaging using the maximum number of conditions \( x(t) = x_p \) starting the averaging, the degree of correlation of subsequently averaged portions of the signal becomes problematic. Further on in this article, selected test results of correlating the implementations of subsequent instances of exceeding the level \( x_p \), initiating conditional averaging for a normal distribution and exponential correlation of the signal \( x(t) \).

Averaging models

In the basic applications of conditional averaging of Gaussian random signals, characteristics of a linear regression are used. For a stationary signal \( x(t) \) with the distribution \( N(0, \sigma_x) \) and a normalised autocorrelation function \( \rho_x(\tau) \) the conditional expected value and the conditional variance are represented by relations:

\[
E(x_2|x_1) = \rho_x(\tau)x_1,
\]

\[
Var(x_2|x_1) = \sigma^2_x \left( 1 - \rho^2_x(\tau) \right),
\]

where: \( x_1 \) and \( x_2 \) - signal values \( x(t) \) in the moments \( t_1 \) and \( t_2 \) respectively; \( \tau = t_2 - t_1 \).

An example of implementation of a random signal with the selected parameters in the categories of time and value marked is provided in Figure 1a. In practice, the model of exponential correlation is obtained by signals with the features of white noise with a limited bandwidth of frequency \( B \) passing through physical inertial systems with the time constant \( T \), while \( BT \geq 2.5 \). This model of exponential correlation is illustrated in Figure 1b and described by the autocorrelation function (3). For \( \tau = 0 \) the autocorrelation function (3) has a finite value of the derivative [4].

\[
R_x(\tau) = \frac{\sigma^2_x}{2\pi B} \int_{-\pi}^{\pi} \cos(\omega\tau) \frac{1}{1 + (\omega T)^2} d\omega.
\]

Fig.1. Implementation of random signal (a) and an experimental exponential correlation characteristic (b)

In conditional averaging of the signal \( x(t) \) in real time, in order to obtain a full mapping of the normalised...
exponential autocorrelation function \( \rho_x(\tau) \) one ought to assume the averaging time \( T_r \) equal to the maximum correlation interval \( \tau_{kw} \), which for the exponential correlation model equals \( 3\tau_k \), where \( \tau_k \) is the correlation interval. The subsequently averaged portions with duration time \( T_r + \tau_{di} \) are uncorrelated.

For a stochastic input signal \( x(t) \) in the form of white noise limited by a low-pass RC filter with the exponential autocorrelation function \( \rho_x(\tau) \) and the distribution \( N(0, \sigma_x^2) \), the average time between the subsequent instances of passing the given level \( x_p \) with a derivative with a given sign equals [4]:

\[
\tau(x_p) = \frac{2\pi}{\omega_{1x}}e^{\frac{x_p^2}{2\sigma_x^2}}
\]

(4)

\[
= \frac{1}{(RC)^2} \left( \frac{2\pi BRC}{\arctg(2\pi BRC)} - 1 \right)
\]

Example 1

For low-pass white noise in the band \( B = 25 \text{kHz} \) filtered with an RC system with the time constant \( T = 10^{-4} \text{s} \), Table 1 presents the calculated and rounded values of the relation \( \frac{\tau_{pm}}{\tau_{km}} \) for several values \( \nu = x_p/\sigma_x \). The subsequent averaged portions of the signal for \( \nu > 1 \) are uncorrelated.

Table 1. Values of the relation \( \frac{\tau_{pm}}{\tau_{km}} \)

<table>
<thead>
<tr>
<th>( \nu )</th>
<th>0</th>
<th>1</th>
<th>( \sqrt{2} )</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{\tau_{pm}}{\tau_{km}} )</td>
<td>0.70</td>
<td>1.15</td>
<td>1.90</td>
<td>5.16</td>
</tr>
</tbody>
</table>

In a simplified model of averaging, with \( M \) uncorrelated and averaged (after exceeding the threshold \( x_1(t) = x_p \)) fragments of the implementation of the signal \( x(t) \), the evaluation of the relative standard uncertainty of the conditional value of the arithmetic mean (CVAM) equals:

\[
\varepsilon(\tau) = \frac{\sqrt{\text{Var}(x_2 | x_1)}}{E(x_2 | x_1)} = \frac{\sigma_x}{\sqrt{Mx_p}} \sqrt{1 - \rho_x^2(\tau)}
\]

(5)

It results from the formula (4) that in order to decrease the uncertainty of evaluation, the value \( x_p \) should be as high as possible. However, its increase will at the same time cause a decrease of the number \( M \) of averaged implementations (for a high value of \( \nu \) exceeding the threshold level \( x_p \), initiating averaging happens less frequently), which in turn will lead to an increase of uncertainty. Both given conditions are mutually opposed and they require determining a compromise value \( \nu \) with the selected practical measurement method.

In the measurement and analysis system the registration of the signal \( x(t) \) with the length \( T_0 \) occurs. Next, portions of the signal \( x(t) \) of the length \( T_r \) are analysed from the moments of exceeding the level \( x_p \) by the signal \( x(t) \). In the time of observation \( T_0 \) the number of analysed portions of the signal \( x(t) \) will occur on average \( M \) times:

\[
M = \frac{T_0}{T_r + \tau(x_p)} = \frac{T_0}{T_r + \frac{2\pi}{\omega_{1x}}e^{\frac{x_p^2}{2\sigma_x^2}}}
\]

(6)

After accounting for the relations (4) and (6) and introducing the designations:

\[
a = \frac{T_r}{T_0}, \quad b = \frac{2\pi}{(RC)^2}, \quad c = \frac{1}{2\sigma_x^2 - 1}
\]

(7)

one obtains an evaluation of the relative standard uncertainty of CVAM:

\[
\varepsilon(\tau) = \frac{1}{x_p} \frac{1}{\rho_{xy}(\tau)} - 1.
\]

(8)

The expression thus arrived at allows one to determine the conditions at which \( \varepsilon(\tau) \) assumes minimum values. In order to do that, one needs to calculate the derivative of \( \frac{d\varepsilon(\tau)}{d\nu} \) and set it to zero:

\[
\frac{d\varepsilon(\tau)}{d\nu} = -\frac{c}{v^2} \sqrt{a + be^2} + \frac{c}{2} \frac{v^2}{\sqrt{a + be^2}} = 0.
\]

(9)

When taking into account the data from Example 1, one obtains a condition for determining the optimal value \( \nu_{opt} \):

\[
b = \frac{v_{opt}^2}{2} = \frac{1}{(RC)^2} \left( \frac{2\pi BRC}{\arctg(2\pi BRC)} - 1 \right).
\]

(10)
Assuming the parameters of the signal $x(t)$ allows one to calculate the relation of $b / a$, and then to solve the equation (10) and determine the optimal value of factor $\nu_{opt}$.

For the data from Example 1: $T_r = \tau_{km} = 3RC$, $b / a = 2.1$ and $\nu_{opt} = 1.517$.

For other models of low-pass noises the value of the threshold $x_p$ initiating the averaging with uncorrelated samples and ensuring a small variance of the CVAM characteristic should be included in the range $\sigma_x / \sqrt{2} \leq x_p \leq 2\sigma_x$ [4].

In conditional averaging within the post-processing of a digitally registered long implementation of the signal $x(t)$ and using all subsequent conditions of exceeding the level $x_p$ with the derivative with a random sign, the subsequent implementations are correlated [4]. While the value of $x_p$ increases, the variance of the conditional arithmetic mean $\bar{x}_i | x_p$ depends largely on the value of $\bar{T}_{pz}$ of the average time the signal $x(t)$ remains over the level $x_p$ and the correlation value $\rho_1(\bar{T}_{pz}) = \rho_1$ (Fig. 2).

With a limited time $T_0$ of observing the signal $x(t)$ in order to increase the number $M$ of averages of conditional implementations it may be necessary to use all the conditions of exceeding the threshold $\nu$ by the signal. In practical analyses the number of averaged implementations should not be below $M = 50$. In calculating the approximate consequences of correlation one may use the arithmetic mean of values of average times of positive and negative instances of exceeding the level $\nu$ by the signal $x(t)$:

\[
\bar{T}_{pzu} = \frac{\bar{T}_{pzu}}{2} = \frac{\pi}{\alpha x} e^{2\tau}. 
\]

As an example, for the data provided in Table 1, one arrives at the value $\bar{T}_{pzu} = 0.95\tau_{km}$ with $\nu = \sqrt{2}$ and $\rho_1 = \rho_1(\bar{T}_{pzu}) = 0.06$.

In the last model of conditional averaging, with an appropriately large number $M$ of implementations, the variance of the CVAM characteristic in point $\bar{x}_i$ can be represented by the relation [4]:

\[
\text{Var}[\bar{x}_i | x_p] \approx \frac{\sigma_x^2}{M} \left[ 1 + \rho_1(\bar{T}_{pzu}) \right] = 1.13 \frac{\sigma_x^2}{M}, 
\]

where $\sigma_x^2$ is a variance of the set of averaged implementations of the signal $x(t)$ at the i-th cross-section.

In another statistical model, in the overall number $2M$ of arithmetically conditionally averaged implementations, at first one averages $M$ implementations with positive values of the derivative with the threshold condition $x_p$, followed by averaging of $M$ implementations with a negative derivative, and at the end both conditionally obtained means as random variables are also correlated arithmetically.

For averaging the uncorrelated implementations $x_a$ and $x_b$ exceeding the threshold $x_p$ with a positive and negative derivative, the conditional variances of the implementations equal:

\[
\sigma_{\bar{x}_a}^2(\tau) = \sigma_{\bar{x}_b}^2(\tau) = \sigma_x^2 \left[ 1 + \rho_1(\bar{T}_{pzu}) \right]. 
\]

With the assumption averaging implementations the beginnings of which are on average $\bar{T}_{pzu}$ apart in the signal, the conditional variance of the sum of averaged and correlated variables $\bar{x}_a$ and $\bar{x}_b$ for $\tau = 0$ will equal:

\[
\sigma_{\bar{x}_a}^2 = 2\sigma_x^2 \left[ 1 + \rho_1(\bar{T}_{pzu}) \right]. 
\]

For times of delay $\tau > 0$ the conditional variance of the sum of variables $\bar{x}_a$ and $\bar{x}_b$ described by the relation:

\[
\sigma_{\bar{x}_a}^2(\tau) = 2\sigma_x^2 \left[ 1 + \rho_1(\bar{T}_{pzu}) \right]. 
\]

The occurrence of correlation results in an increase of the variance of the arithmetical mean. For example for the data:

$\nu = \sqrt{2}$ ; \hspace{1cm} $\bar{T}_{pzu} = 0.15\tau_{km}$ ; \hspace{1cm} $\rho_1(\bar{T}_{pzu}) = 0.64$ ; 

$\sigma_{\bar{x}_a} = \sqrt{2}\sigma_x \left( 1 + 0.64 \right) = 1.81\sigma_x$ 

the general formula expressing the relative standard uncertainty of CVAM in the investigated model of correlation equals:
Examples of conditional implementations

In Figures 3 and 4 observed conditional implementations were presented. What is visible is a larger variance of implementation while averaging with derivatives of both signs (4a) when compared with averaging with the derivative with one sign (4b).

In Figure 5 an example of implementations initiated by all subsequent positive and negative instances of exceeding the threshold \( x_p = 1 \) \( V \) are presented. The distribution of implementations for \( \tau = 10 \) \( \mu s \) is comparable with Figure 4a and it is an example of calculation based on the expression (15).

\[
\sigma(\tau) = \frac{\sigma_x \sqrt{\left[ 1 + \rho_x(\tau) \right]^2 - 1 + \rho_x^2(\tau)}}{\sqrt{M - x_p \cdot \rho_x(\tau)}}.
\]

**Summary**

1. In practice, the real model of exponential correlation occurs frequently and it is obtained by signals with a limited bandwidth with the features of white noise passing through physical inertial systems.
2. In measuring applications for the assessment of data correlation with normal and quasi-normal distribution, operations of conditional data averaging can be used. Determining statistical links with the use of conditional averaging of signals is particularly beneficial in the conditions of strong data correlation.
3. In the model of exponential correlation, subsequent averaged implementations of the signal exceeding the threshold with a derivative of any sign are significantly correlated, while the correlation causes the conditional variance of the average value to increase.
4. In models of exponential correlation the averaged implementations of the signal exceeding the threshold \( x_p \) with a derivative of one sign can practically be considered uncorrelated. Such a way of conditional averaging can be recommended for practical applications.

**Authors**: Assoc. Prof. Adam Kowalczyk, Msc Eng Rafał Chorzępa, Rzeszow University of Technology, Department of Metrology and Diagnostic Systems, W. Pola 2, 35-959 Rzeszów, E-mail: kowadam@prz.edu.pl, rchorz@prz.edu.pl.

**REFERENCES**

[1] Zhang N.F., Calculation of the uncertainty of the mean of autocorrelated measurements, Metrology, 43 (2006), 276-281