SMOOTHING INTERVAL ESTIMATION FOR A RANDOM PROCESS IN REAL TIME

Mark Polyak

Saint-Petersburg State University of Aerospace Instrumentation
Saint-Petersburg, Russia

Abstract

This paper presents a mathematical algorithm of smoothing interval estimation for telemetry data processing in soft real time system. Smoothing interval depends on the properties of a random process describing the noise component of the telemetry signal.

I. INTRODUCTION

The task of filtering data often arises in expert systems, control systems, monitoring systems and other systems that estimate state of a complex unit or vehicle, such as carrier rocket or nuclear reactor, in real time. Although processing power of computers has grown rapidly in the last few decades, modern computers (not counting mainframes) are not capable to process hundreds or even thousands of telemetry parameters with frequencies varying from 50 to 200 Hz in real time. The data being transferred is often redundant as sensor's frequency is intentionally overrated to provide enough information for reconstructing an emergency situation if such situation occurs. Unfortunately, this redundancy leads to additional computational expenses when the data is processed by real time expert control system. Transmission channel adds random noise to telemetry data making it more difficult for data processing system to correctly interpret received values. Eventually, time delays and control errors caused by incorrect estimation of current state of the monitored complex unit may lead to irreversible consequences: loss of control, different faults, accidents with various consequences and even global catastrophes.

This is where the problem of filtering data arises: noisy (corrupted) data should be removed [1] and redundancy should be reduced. Smoothing, which means substituting a set of data points with a single point (their average), allows reducing redundancy to some predefined extent. Time interval on which smoothing is performed is the tuning parameter used to control the extent of smoothing. So, to successfully filter telemetry data a criterion of

accuracy of the estimated smoothing interval must be defined first.

II. CRITERION OF ACCURACY

About 90% of all possible noise found in telemetry data (noise from transmission channel, noise from sensor vibrations, etc.) could be described just with two types of stationary random processes $X_1(t)$ and $X_2(t)$ with corresponding correlation functions:

$$R_{x_1}(\tau) = \sigma_{x_1}^2 e^{-\gamma|\tau|}, \qquad (1)$$

$$R_{x_2}(\tau) = \sigma_{x_2}^2 \left(1 + \gamma|\tau|\right) e^{-\gamma|\tau|}. \qquad (2)$$

Expectation function $M_z(t)$ of both processes has the following form:

$$M_z(t) = A_c \cos\left(2\pi F_c t + \varphi\right), \qquad (3)$$

Let's take minimum of variance of estimation of expectation function $M_z^*(t)$ as criterion of accuracy. Variance of $M_z^*(t)$ is (4):

$$D[M*_Z(t)] = 2\int_0^\infty \left[|\Phi_n(F_c)|^2 \left|S_x(F_c) + |\Phi_n(F_c) - 1|^2 S_{M_Z}(F_c)\right]dF_c\right.$$

$$\qquad (4)$$

where: $S_X(F_c)$, $S_{M_Z}(F_c)$ – are spectral densities of $X_i(t)$ and $M_z(t)$ accordingly; $\Phi_n(F_c) = \left(\dfrac{\sin \pi F_c \Delta T}{\pi F_c \Delta T}\right)^n$ – filter of current average; and ΔT – smoothing interval.

Thus, criterion of accuracy is $\dfrac{d}{d\Delta T} D[M*_Z(t)] \to 0$.

III. FIRST NOISE MODEL

Let's estimate the smoothing interval (ΔT) for the first type of random process. This process has correlation function of the form (1) and following spectral densities of $X_1(t)$ and $M_z(t)$ [2]:

$$S_X(F_c) = \frac{2\sigma_x^2 \gamma}{\gamma^2 + \left(2\pi F_c\right)^2}, \qquad (5)$$

$$S_{M_z}(F_c) = \frac{\pi A_c^2}{2}\left[\delta(\omega - \omega_c) + \delta(\omega + \omega_c)\right], (6)$$

where: $\omega_c = 2\pi F_c$

After inserting (5) and (6) into (4) we obtain expressions for variance of estimation of expectation function. Then, differentiating those expressions with respect to ΔT and setting the result equal to zero, we obtain the formula for optimal value of ΔT (7):

$$\alpha^2 \frac{1}{\Delta T^2 \gamma^2}\left[\Delta T\gamma - 2 + e^{-\Delta T\gamma}(\Delta T\gamma + 2)\right] +$$

$$+ \left(1 - \frac{\sin \pi F_c \Delta T}{\pi F_c \Delta T}\right)\left(\cos \pi F_c \Delta T - \frac{\sin \pi F_c \Delta T}{\pi F_c \Delta T}\right) = 0 \quad (7)$$

where: $\alpha = \dfrac{\sigma_x \sqrt{2}}{A_c}$ – ratio of standard deviations

(coefficient of variation) of processes $X_1(t)$ and $M_z(t)$; ΔT – smoothing interval; γ – sampling frequency; F_c – natural frequency of random process.

Using expression (7) it is possible to plot a nomogram for estimation of smoothing interval. An example of such a nomogram is shown on fig. 1.

Nomogram may be used to check an answer obtained from an exact calculation method or for approximate graphical computations. More precise values are usually stored in a table.

For smoothing interval estimation an algorithm of linear search for a root was used. This algorithm finds the first (smallest) intersection of function (7) with X-axis (zero level). Total number of solutions of equation (7) is infinite because the function is periodic (see fig. 2). Distinctive spikes clearly seen on this nomogram appear because the value of first local minimum of function (7) (this minimum gives the root of the equation in most cases) goes up when the coefficient of variation α is increased.

Measure of inaccuracy for the values found with this algorithm is $\pm 5 \times 10^{-5}$ s.
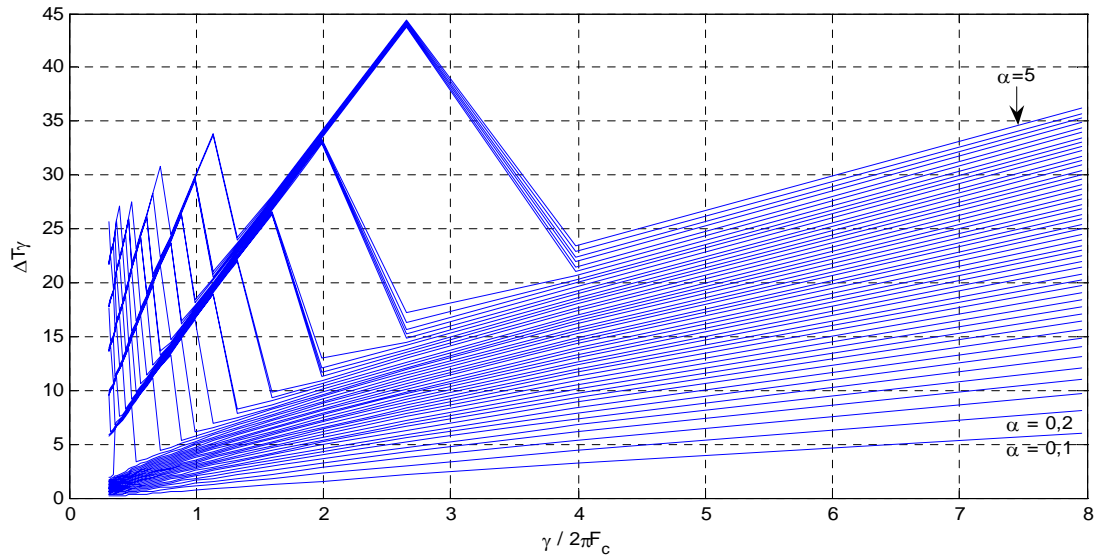

Fig. 1. Nomogram for smoothing interval estimation for a random process $X_1(t)$ with $\gamma = 50$ Hz
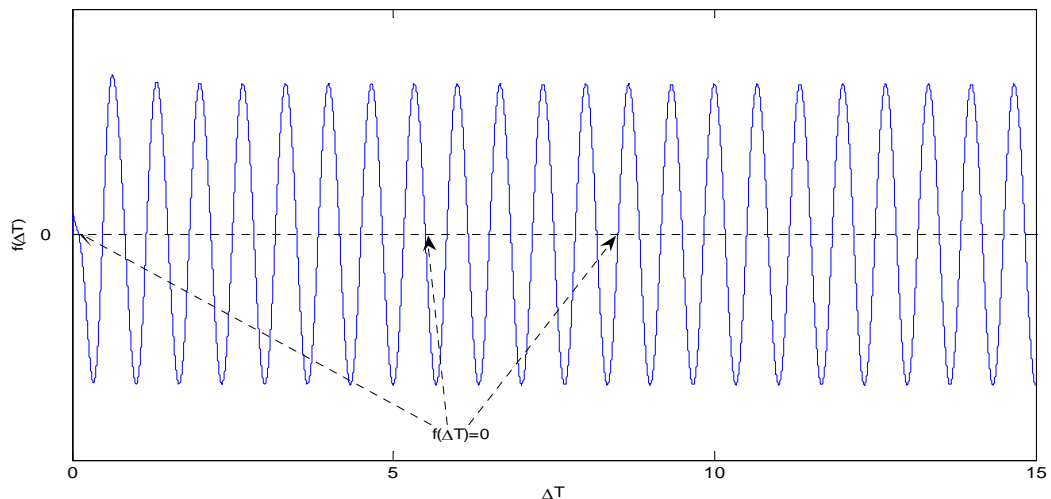
Fig. 2. General view of function (7) depending on ΔT with $\alpha = 1, \gamma = 200\,\text{Hz}, F_c = 3\,\text{Hz}$

IV. SECOND NOISE MODEL

Now we will estimate the smoothing interval for the second type of random process. This process has correlation function of the form (2) and following spectral density of $X_2(t)$ [3]:

$$S_X(F_c) = \frac{4\sigma_x^2\gamma}{\left(\gamma^2 + \left(2\pi F_c\right)^2\right)^2}. \qquad (8)$$

Spectral density of $M_z(t)$ stays unaltered, it has the form (6).

After inserting (8) and (6) into (4), differentiating the resulting expressions with respect to ΔT and setting the final result to equal zero, we obtain the formula for optimal value of ΔT for the second type of random process:

$$\alpha^2 \frac{2}{\Delta T^2\gamma^2}\left(\frac{1}{\gamma^2 + (2\pi F_c)^2} + \frac{1}{\gamma^2}\right)\left[\Delta T\gamma - 2 + e^{-\Delta T\gamma}(\Delta T\gamma + 2)\right] +$$

$$+ \left(1 - \frac{\sin\pi F_c\Delta T}{\pi F_c\Delta T}\right)\left(\cos\pi F_c\Delta T - \frac{\sin\pi F_c\Delta T}{\pi F_c\Delta T}\right) = 0$$

$$(9)$$

where: $\alpha = \frac{\sigma_x\sqrt{2}}{A_c}$ – ratio of standard deviations (coefficient of variation) of processes $X_2(t)$ and $M_z(t)$; ΔT, γ and F_c – the same as in formula (7).

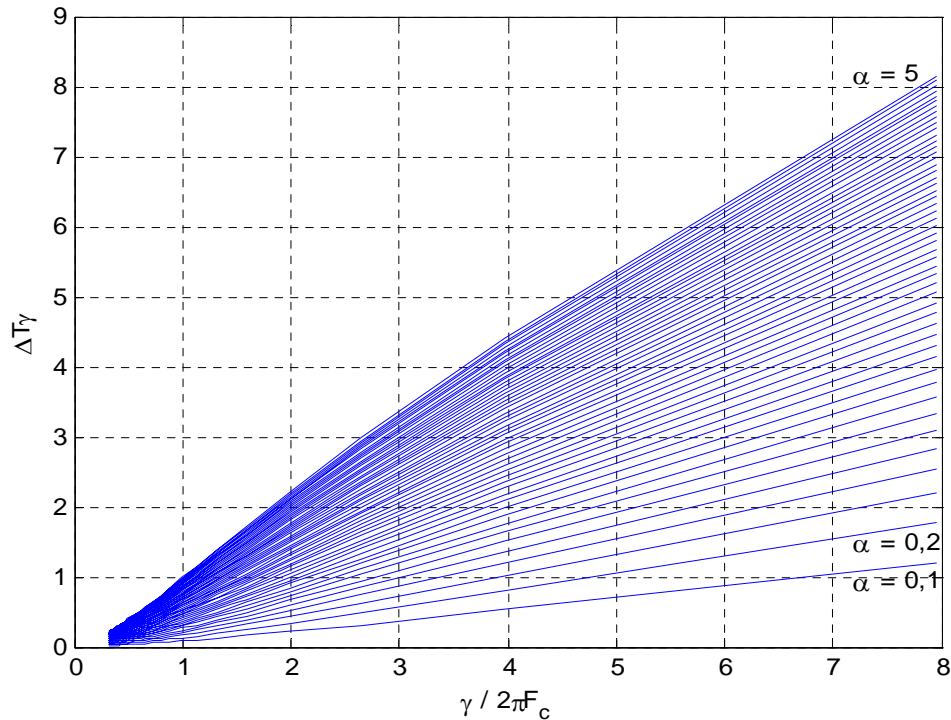A nomogram which was build using formula (9) is shown on fig. 3.

Fig. 3. Nomogram for smoothing interval estimation for a random process $X_2(t)$ with $\gamma = 50$ Hz

V. REAL TIME PROCESSING

With presented nomograms it is possible to find correct smoothing interval for the two examined models of noise in telemetry data. Although this is good enough for theoretical research, for the real time signal processing system the time interval on which smoothing can be performed is of little importance. Instead such systems should be supplied with the *number* of sample points that could be replaced with their average without loss of significant value. The number of sample points on the smoothing interval is defined by the formula

$$n = \lfloor \Delta T \times \gamma \rfloor, \qquad (10)$$

where: ΔT – is the smoothing interval; γ – sampling frequency.

Graph of number of sample points n as a function of natural frequency of the random process is shown on fig. 4.

To reduce computations in real time all values of n are stored in static tables. Thus, instead of long computations a real time system performs a table lookup operation which is considerably faster. In other words, a table lookup means instant memory access with time complexity *O(1)*. To achieve such results the mentioned tables must be organized as follows:

- each cell contains a single value of n;
- F_c defines a row and α defines a column (or vice versa);
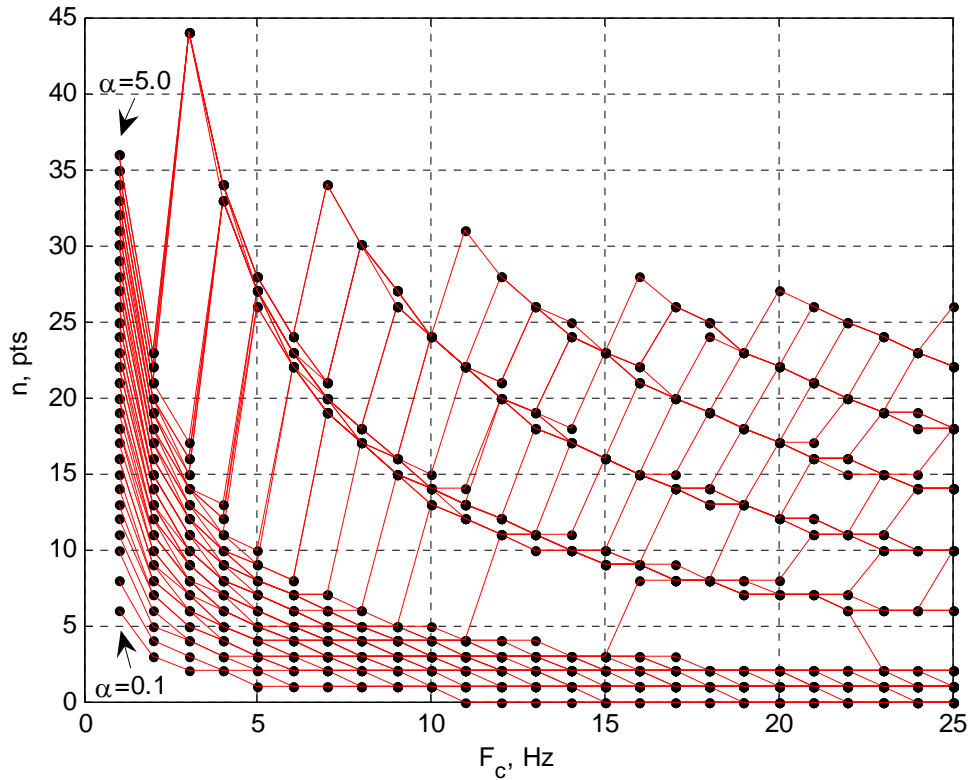- there should be different table for each sampling frequency (e.g. 50, 100 and 200 Hz – three tables).

Fig. 4. Graph of n as a function of F_c with $\gamma=50$ Hz for the first type of random process

VI. CONCLUSION

Fig. 4 clearly shows that as the natural frequency of a random process goes up the number of points for smoothing drops rapidly. So the best smoothing results could be achieved on random processes with low natural frequency, while smoothing of a random process with relatively high natural frequency is likely to give no sample size shrinking at all.

Nevertheless, smoothing method discussed in this paper allows reducing redundancy of telemetry data without loosing significant values. The final outputs of examined algorithm are optimized tables of values for fast data access which is compliant for real time data processing.

REFERENCES

[1] Каргин В.А., Николаев Д.А. Самойлов Е.Б. Обнаружение и отбраковка аномальных результатов измерений для формирования исходной измерительной информации по ракете-носителю типа "Союз", Информация и Космос 2008, выпуск 4, С-Петербург, стр.83-86.

[2] Волгин В.В., Каримов Р.Н. Оценка корреляционных функций в промышленных системах управления. — М.: Энергия, 1979. — 80 с., ил. — (Библиотека по автоматике; Вып. 600.)

[3] Отбраковка аномальных результатов измерений / А.Ф. Фомин, О.Н. Новоселов, А.В. Плющев. — М.: Энергоатомиздат, 1985. — 200 с., ил.