International Journal of Systems Signal Control and Engineering Application

MEDWELL PUBLICATIONS SINCE 2000

Models of Stohastic Processes and Their use in Optimal Linear Inteprolation and Forecasting

Vladimir Alekseevich Golovkov

Institute of Optoelectronic Instrumentation Engineering, Leningrad Region, 188540 Sosnovy Bor, Russia

Key words: Stochastic process, differentiability, wiener-Hopf filtering, forecasting, interpolation

Corresponding Author:
Vladimir Alekseevich Golovkov
Institute of Optoelectronic Instrumentation Engineering Leningrad Region, 188540 Sosnovy Bor, Russia

Page No.: 113-116
Volume: 10, Issue 4, 2017
ISSN: 1997-5422
International Journal of Systems Signal Control and Engineering Application

Abstract: In this study, I consider models of stochastic process correlation functions and, by way of numerical calculation, prove that the efficiency of optimal linear interpolation and forecasting is determined by the existing highest derivative of a stochastic process. I also set out the results of numerical calculations pertaining to efficiency assessment of interpolation and forecasting of finitely differentiable stochastic processes with correlation functions commonly used in practice for Wiener-Hopf filtering.

INTRODUCTION

Needless to state once again how important it is to develop methods for optimal linear forecasting and interpolation of stochastic processes and assessment of their efficiency, by which we should understand forecast/interpolation result dispersion for a stochastic process. When using the Wiener-Hopf filter (Widrow and Stearns, 1985) that shapes output signal ξ as a linear regression of a sample $||Z|| = ||\xi_1, \xi_2, ..., \xi_k||$ or:

$$\xi \sum_{i=1}^{k} w_i \xi_k \qquad (1)$$

where vector $||W|| = ||w_1, w_2, ..., w_k||$ are coefficients; the optimal vector of these coefficients is defined as:

$$\|W_{opt}\| = \|R\|^{-1} \cdot \|P\|^{T} \qquad (2)$$

where, $||R||$ is the matrix of cross-correlation of sample components $||Z|| = ||\xi_1, \xi_2, ..., \xi_k||$, $||P||^T$ is the column vector of cross-correlation between signal ξ and sample components $||Z|| = ||\xi_1, \xi_2, ..., \xi_k||$. Minimal variance of ξ is obtained as:

$$\sigma^2(\xi/\|Z\|) = \sigma^2 - \|P\| \cdot \|R\|^{-1} \|P\|^{T} \qquad (3)$$

where σ^2 is stochastic process variance. In real contexts, it is inconceivable to assume the finiteness of spectral density of a stochastic process under scrutiny; at the same time, sample dimension for forecasting and interpolation purposes is significantly limited by equipment functionality. Whence it follows that we must admit residual variance of the results of forecasting or interpolation of a stochastic process. To know probability's multivariate density of a stochastic process is also impossible in many a case; however, for optimal linear forecasting or interpolation of stochastic processes the knowledge of their spectral-correlation properties is sufficient. For convenience of description (Tikhonov and Khimenko, 1987), we shall designate the stochastic process $\eta(t)$ and its implementation in the similar way as

Table 1: Normalized correlation functions and spectral densities

Correlation functions			
1	2	3	4
$\rho(\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}S(\omega)e^{j\omega\tau}d\omega$	$-\rho''(0)$	$S(\omega)=\int_{-\infty}^{\infty}\rho(\tau)e^{-j\omega\tau}d\tau$	Δf,
$e^{-\alpha\lvert\tau\rvert}$	-	$\frac{2\alpha}{\alpha^2+\omega^2}$	$\frac{\alpha}{2}$
$(1+\alpha)e^{-\alpha[\tau]}$	2α	$\frac{4\alpha^3}{(\alpha^2+\omega^2)^2}$	$\frac{1}{4}\alpha$
$\left[1+\alpha\lfloor\tau\rfloor+\frac{1}{3}(\alpha\tau)^2\right]e^{-\alpha\lvert\tau\rvert}$	$\frac{1}{3}\alpha^2$	$\frac{16\alpha^5}{3(\alpha^2+\omega^2)^3}$	$\frac{3}{16}\alpha$
$\left[1+\alpha\lvert\tau\rvert+\frac{2}{5}(\alpha\tau)^2+\frac{1}{15}(\alpha\lvert\tau\rvert)^3\right]e^{-\alpha\lvert\tau\rvert}$	$\frac{3}{5}\alpha^2$	$\frac{32\alpha^7}{5(\alpha^2+\omega^2)^4}$	$\frac{5}{32}\alpha$
$\cos(\omega_0\tau)$	ω_0^2	$\frac{1}{2\pi}\left[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)\right]$	-
$e^{-\alpha\tau^2}$	2α	$\sqrt{\frac{\pi}{\alpha}}\exp\left(-\frac{\omega^2}{4\alpha}\right)$	$\sqrt{\frac{\alpha}{\pi}}$
$\frac{\sin\left(\frac{\Delta\omega\tau}{2}\right)}{\left(\frac{\Delta\omega\tau}{2}\right)}$	$\frac{1}{2}(\Delta\omega)^2$	$\begin{cases}\frac{2\pi}{\Delta\omega} & \lvert\omega\rvert\le\frac{\Delta\omega}{2}\\ 0 & \lvert\omega\rvert\ge\frac{\Delta\omega}{2}\end{cases}$	$\frac{1}{2\pi}\Delta\omega$

$\xi(t)$. Assuming that $\xi(t)$is stationary with dispersion σ^2 and correlation function $R(\tau) = \sigma^2\rho(\tau)$ where, $\rho(\tau)$ is the normalized correlation function, we shall consider various spectral correlation properties of stochastic processes in the form of a (Table 1).

In Table 1, we can see normalized correlation functions $\rho(\tau)$ and spectral densities $S(\omega)$ of stochastic processes most commonly used in practice. Besides this, in Table 1 there is $-\rho''(0)$, i.e., the second spectral moment of a stochastic process (Tikhonov and Khimenko, 1987) and $\Delta f_Э$-effective width of spectral density of a random process.

Stochastic process No. 1 with normalized correlation function $\rho(\tau) = e^{-\alpha|\tau|}$ is not differentiable albeit continuous. Most commonly such correlation functions are applied to describing simple markovian stochastic processes. Random process No. 2 with normalized correlation function is d ifferentiable only once whereas that of No. 3-twice, process No. 4 can be differentiated three times (Tikhonov and Khimenko, 1987). Implementation of process No. 5 with correlation function $\rho(\tau) = \cos(\omega_0\tau)$ is described as $\xi(t) = A \sin(\omega_0 t+\varphi)$ where A-amplitude, ω_0-constant frequency, φ-random value with probability coefficient $p(\varphi) = 1/2\pi$ at $-\pi\le\varphi\le\pi$. Such a process is differentiable arbitrarily many times. Processes with correlation functions Eq. 6 and 7 are also differentiable arbitrarily many times. However, the process with normalized correlation function $\rho(\tau) = \exp(-\alpha\tau^2)$ exhibits spectral density equal to zero only at $\omega\to 0$ whereas the process with normalized correlation function $\rho(\tau) = \sin(\omega\tau/2)/(\omega\tau/2)$ has ω-axis finite spectral density. One can easily recognize that the question of whether infinitely differentiable stochastic processes do exist is more than doubtful as it lets calculation of their implementation for infinite time both ways of the axis, thus leading us to the notion of linear singularity (Rozanov, 1990) or random process degeneration.

For a random process whose normalized correlation function $\rho(\tau) = \cos(\omega\tau)$ is indicated by Eq. 5 in Table 1, when using sample $\|Y\| = \xi(t),\xi(t)\|$ to forecast implementation $\xi(t+\theta)$ at calculating regression coefficients relying on (Tikhonov and Khimenko, 1987) we'll construct matrix $\|R\|$ and vector $\|P\|$:

$$\|R\| = \sigma^2\begin{Vmatrix}1 & 0\\ 0 & -\rho''(0)\end{Vmatrix} = \sigma^2\begin{Vmatrix}1 & 0\\ 0 & -\omega^2\end{Vmatrix} \tag{4}$$

$$\|P\| = \sigma^2\left\|\cos(\omega\theta) \quad \omega\sin(\omega\theta)\right\| \tag{5}$$

By inverting matrix Eq. 4 we have regression coefficients as follows:

$$w_1(\theta) = \cos(\omega\theta) \tag{6}$$

$$w_2(\theta) = \frac{\sin(\omega\theta)}{\omega} \tag{7}$$

Therefore, if we use sample $\|Y\| = \xi(t), \xi(t)\|$ for forecasting and if the normalized correlation function is $\rho(\tau) = \cos(\omega\tau)$ the forecasting algorithm can be represented as:

$$\xi(t+\theta/t) = \cos(\omega\theta)\cdot[\xi(t)-\mu] + \frac{\sin(\omega\theta)}{\omega}\xi(t) \qquad (8)$$

We may observe that for concrete implementation this algorithm lets to unambiguously calculate future implementation of a stochastic process for any forecast time θ. Thus, for example, given a random phase $\varphi = 0$ and mathematical expectation $\mu = 0$, implementation of the stochastic process will look like $\xi(t) = A \sin(\omega)$ while implementation forecast $-\xi(t+\theta/t) = A \sin[\omega(t+\theta)]$. If the second derivative implementation of a stochastic process ξ(t) is used, matrix $\|R\|$ will become degenerated, its determinant equal to zero and the forecast algorithm unobtainable.

The whole matter looks more complicated if for the interference model we choose a random process that is finitely differentiable. As numerical studies show, the maximum dimension of the forecasting filter is determined by the existing highest derivative of a stochastic process. So, if a random process is totally non-differentiable, then all information on its past and future is defined by only one of its values. An example of it is a process with normalized correlation function $\rho(\tau) = e^{-\alpha|\tau|}$ (a simple Markovian process). If a random process in differentiable once then the maximum sample dimension is equal to 2 process samples; if process differentiability takes the value of 2 then the maximum sample dimension is equal to 3 process samples and so forth. Subsequent expansion of sample size brings about no reduction in forecast dispersion. Consider the example of a triply differentiable random process with a normalized correlation function:

$$\rho(\tau) = \left[1+\alpha|\tau| + \frac{2}{5}(\alpha\tau)^2 + \frac{1}{15}(\alpha|\alpha|\tau\|)^3\right]e^{-\alpha|\tau|} \qquad (9)$$

As correlation functions of finitely differentiable random processes are not explicitly differentiable, then it seems reasonable to consider its value sample $\|X\| = \|\xi(t), \xi(t-\Delta t), ..., \xi(t-n\Delta t), ..., \|$. To obtain vector (2) of optimal factors $\|W_{opt}\|$ from the sample vector $\|X\|$ and the forecast value $\xi(t+\theta)$ is not a problem. Now we'll plot graphs of the dependence of normalized forecast variance of a stochastic process on forecast time interval θ in the following form:

$$l(\theta) = \frac{\sigma^2[\xi(t+\theta)/\xi(t-\Delta t), ..., \xi(t-n\Delta t)]}{\sigma^2} \qquad (10)$$

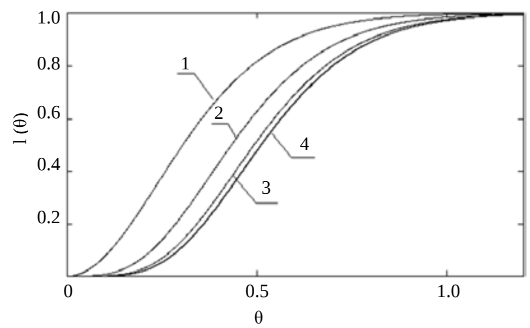

Fig. 1: Forecast variance at $\rho(\tau) = [1+\alpha\lfloor\tau\rfloor + 2/5(\alpha\tau)^2 + 1/15(\alpha|\tau|)^3]e^{-\alpha|\tau|}$, (1) For sample ξ(t), (2) For sample ξ(t), ξ(t-Δt), (3) For sample ξ(t), ξ(t-Δt), ξ(t-2Δt) and (4) For sample ξ(t), ξ(t-Δt), ξ (t-2Δt), ξ (t-3Δt)

where, $\sigma^2[\xi(t+\theta)/\xi(t), \xi(t-\Delta t), ..., \xi(t-n\Delta t)]$ is forecast ξ(t) dispersion for time point t+θ of time point t, Δt-time interval between sample units, n-integer. As other studies have demonstrated, time interval Δt must be minimal for ensuring minimal forecast dispersion. Let us now choose parameter $\alpha = 32/5$ for this correlation function, then the effective width of random process spectrum is $\Delta f_э = 1$. Figure 1 shown dependence of normalized forecast dispersion l(θ) for a random process with a normalized correlation function Eq. 9. The calculations were done for $\Delta t = 0.01$. As can be seen from Fig. 1, no increase in the efficiency of the forecast filter is observed when se employ samples of higher than the fourth order.

In this way, random processes that are finitely differentiable can be predicted by transversal filters while the order of such a transversal filter is determined by the order of the existing highest derivative of the random process.

Forecasting filter sample dimension for infinitely differentiable stochastic processes (correlation functions No. 6 and 7 in Table 1) is limited due to other considerations. Thus, when choosing a normalized correlation function of gaussoide type $\rho(\tau) = \exp(-\alpha\tau^2)$, the dimension of sample in forecasting should not be in excess of 3-4 as further increase in its size does not result in any considerable gain in forecasting efficiency. If the normalized correlation function $\rho(\tau) = \sin(\alpha\tau)/(\alpha\tau)$ is selected, the dimension of the sample is limited by quantizing noise generated by analog to digital conversion (Bodrenok, 1997).

The same sort of considerations hold for interpolating of random processes. It is known (Khurgin, 1971) that if both the random process value and its derivatives are used for interpolation, then the interval between process samples may be increased in proportion to the order of the involved derivatives. For practical engineering tasks, as a rule, what is the focus of interest is interpolation of a stochastic process within a chosen time interval. In this

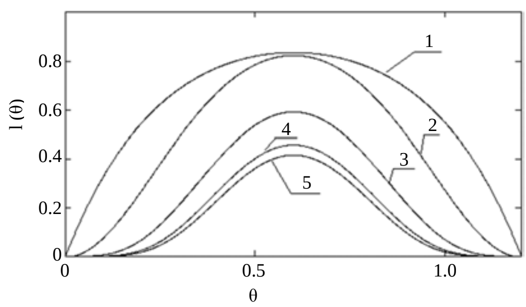

Fig. 2: Interpolation variance for $\rho(\tau) = e^{-\alpha|\tau|}$ и$\rho(\tau)[1+\alpha\lfloor\tau\rfloor + 2/5(\alpha\tau)^2 + 1/15(\alpha|\tau|)^3]e^{-\alpha|\tau|}$ (1) For $\rho(\tau) = e^{-\alpha|\tau|}$ and sample $\|X\| = \|..., \xi(t_0-\Delta t), \xi(t_0), \xi(t_0+T), \xi(t_0+T+\Delta t), ...\|$, (2) For $\rho(\tau) = [1+\alpha\lfloor\tau\rfloor +2/5(\alpha\tau)^2 + 1/15(\alpha|\tau|)^3 e^{-\alpha|\tau|}$ and sample $\|X\| = \|\xi(t_0), \xi(t_0+T)\|$, (3) For sample $\|X\| = \|\xi(t)$, (4) For $\|X\| = \|\xi(t_0-2\Delta t), \xi(t_0-\Delta t), \xi(t_0), \xi(t_0+T), \xi(t_0+T+\Delta t), \xi(t_0+T+2\Delta t)$ and (5) For sample $\|X\| = \|\xi(t-3\Delta t), \xi(t_0-2\Delta t), \xi(t_0-\Delta t), \xi(t_0), \xi(t_0+T), \xi(t_0+T+\Delta t), \xi(t_0+T+2\Delta t), \xi(t+T+3\Delta t)\|$

case, at both ends of the interval it is reasonable to take the sample out of the value of the process and its derivatives, or out of several values of the random process. For random processes with normalized correlation functions marked 6 and 7 in Table 1 this issue has been extensively studied. Suppose, interpolation is being carried out on $t_0 \le t \le t_0+T$ time interval, consequently interpolation takes place on interval $t_0+\theta$ where θ denotes universal time.

Let us consider the sample $\|X\| = \|, ..., \xi(t_0-2\Delta t), \xi(t_0-\Delta t), \xi(t_0), \xi(t_0+T), \xi(t_0+T+\Delta t), \xi(t_0+T+2\Delta t), ..., \|$. In is not a hard task to construct matrix $\|R\|$ using sample vector $\|X\|$ in order to obtain vector $\|P\|$ as that of mutual correlation $\xi(t_0+\theta)$ and components of vector $\|X\|$ and adducing them here is not worthwhile. From expression Eq. 2 we can find weight vector $\|W_{opt}\|$. Using expression Eq. 3 we can find dispersion of interpolated values $\sigma^2\ [\xi(t_0+\theta)/\|X\|]$. The normalized efficiency of interpolation can be represented as:

$$e(\theta) = \frac{\sigma^2\left[\xi(t_0+\theta)/\|X\|\right]}{\sigma^2} \qquad (11)$$

Now, we consider a stochastic process with normalized correlation function No. 1 in Table 1. We select parameter α = 2for this correlation function, then the efficient spectrum width for the random process is $\Delta f_Э$. In addition to this, we shall consider a random proces with a normalized correlation function as per expression Eq. 9 and the efficient spectrum width $\Delta f_Э = 1$, the same as for the case analyzed above when we estimated forecast variance. Let us no calculate the transversal filter at $\Delta t = 0.01$. In Fig. 2 you can see graphs e(θ) for these cases of interpolating a random process on interval T = 1, 2.

As can be seen from Fig. 2, interpolation efficiency for random processes is determined by the highest derivative of a random process. Studies have shown that if a random process is not differentiable, then for its interpolation at the ends of the time interval it will suffice to know its values on these ends. If a stochastic process is differentiable only once, then the sample at the ends of the interpolation interval is appropriate to be taken already for two values; if the process is differentiable twice, then the sample at the ends of the time interval is to be taken for three values of the stochastic process, and so forth. Subsequent increase in the sample size does not result in better interpolation efficiency.

CONCLUSION

Summing up the aforesaid, we can state that the efficiency of optimal linear forecasting and interpolation of random processes using Wiener-Hopf filtering is determined by the existing highest derivative of these. Application of various correlation function models is motivated by the investigative tasks set by the researcher. Thus, we can use finitely differentiable random processes for eliminating "paradoxical" effects assuming the form of stochastic process singularity during optimal linear interpolation or interpolation. Linear singularity of random processes that are differentiable without restriction can be eliminated by adding quantizing noises during analogue-digital conversion or by adding white noise.

REFERENCES

Bodrenok, Z.E., 1997. [On the issue of choosing filter parameters for random process forecasting (In Russian)]. Radio Eng., 1: 32-33.

Khurgin, Y.I., 1971. Finite Functions in Physics and Engineering. Nauka Publisher, Moscow, Russia, Pages: 408.

Rozanov, Y.A., 1990. [Stationary Random Processes]. Nauka Publishing, Moscow, Russia, (In Russian).

Tikhonov, V.I. and V.I. Khimenko, 1987. [Stochastic Process Trajectory Overshoots]. Nauka Publishing, Moscow, Russia, (In Russian).

Widrow, B.W. and S.D. Stearns, 1985. Adaptive Signal Processing. Prentice-Hall, New Jersey, USA..