ISSN 8756-6990, Optoelectronics, Instrumentation and Data Processing, 2016, Vol. 52, No. 3, pp. 280–285. © Allerton Press, Inc., 2016.
Original Russian Text © E.A. Kochegurova, E.S. Gorokhova, 2016, published in Avtometriya, 2016, Vol. 52, No. 3, pp. 79–85.

AUTOMATION SYSTEMS
IN SCIENTIFIC RESEARCH AND INDUSTRY

Current Estimation of the Derivative of a Nonstationary Process Based on a Recurrent Smoothing Spline

E. A. Kochegurova and E. S. Gorokhova

Tomsk Polytechnical University, pr. Lenina 30, Tomsk, 634050 Russia
E-mail: kocheg@mail.ru

Received December 2, 2015

Abstract—A method of real-time reconstruction of the useful signal and its lower derivatives on the basis of a recurrent smoothing spline is presented. A calculation technique for a spline with the number of measurements at each segment greater than the number of nodes is given, and the spline coefficients are found by the variational approach.

Keywords: reconstruction of the function and its derivatives, recurrent algorithm, variational smoothing spline.

DOI: 10.3103/S8756699016030109

INTRODUCTION AND FORMULATION OF THE PROBLEM

One of the forms of presentation of initial data in monitoring, control, and decision making systems is time series. The main information about the observed object is often formed by the object itself and is registered by measurement tools within a certain time interval. Because of noise distorting the object parameters and nonstationary behavior of the object, the observed process becomes nonstationary.

Processing of nonstationary random processes involves a number of problems; one of them is indirect estimation of derivatives on the basis of measured information. This problem is particularly important in such applied fields as object control, monitoring, decision making, and prediction of indicators and behavior of objects and systems.

Of particular interest are methods of real-time reconstruction of the useful signal and its lower derivatives.

In this paper, we consider a class of algorithms used in measurement instruments with a discrete input and analog output, where a sequence of samples $y(t_i)$ with random noise $\xi(t_i)$ is recorded at time instants t_i, $i = 1, 2, \ldots$:

$$\begin{gathered} y(t_i) = f(t_i) + \xi(t_i); \\ M\{\xi(t)\} = 0; \quad M\{\xi^2(t)\} = \sigma_\xi^2; \quad M\{f(t), \xi(t)\} = 0. \end{gathered} \tag{1}$$

Such devices are based on a converter approximating input samples into analog output signals.

APPROACHES TO SOLVING THE PROBLEM

Smoothing of time series is one of the most powerful tools of studying them. The smoothed process can be considered as an ideal variant of its behavior. Approximation of a smoothed series in the form of an analytical function offers a number of prospects for estimation of derivatives. The resultant analytical function reflecting the basic properties of the observed process allows one to estimate derivatives of a given order. The most popular methods of smoothing are the regression analysis, orthogonal polynomials, splines, wavelet functions, Fourier functions, and other methods of the time-frequency analysis. The main classes of the methods for approximation of experimental data were analyzed in [1].

Methods based on using spline functions [2, 3] combined with spectral or genetic methods are often used for noisy signal reconstruction.

The best-developed apparatus of interpolation splines involves solving systems of algebraic equations, which is computationally expensive. Moreover, interpolation splines cannot be used for smoothing if the measured results are rather noisy.

It is known that the solution of the problem of differentiation by direct methods is unstable because the rate of change of the function drastically enhances sharp fluctuations of the original function registered with certain errors. For discrete data, this is manifested as an increase in the input data error. A possible approach providing stable results is parametrization of solutions in the class of smoothing splines.

The use of a spline providing the minimum norm of the derivative allows obtaining the sought smooth solution [4–7]. The smoothing spline $S(t)$ is based on optimization of a special type of the functional

$$J(S) = \lambda \int_a^b [S''(t)]^2 dt + \sum_{i=0}^{n} [S(t_i) - y(t_i)]^2, \qquad (2)$$

where the first term together with the smoothing parameter λ defines the penalty of the curvature, whereas the second term minimizes the residual sums of the squares and corresponds to the classical least squares technique.

Derivatives can be obtained on the basis of observation results by means of interpretation of direct [8, 9] or indirect [10] measurements. In the latter case, the sought derivative is the solution of the integral equation of the first kind and is related to using Tikhonov's regularization technique.

Interpretation of direct measurements is based on differentiation of the approximating function. A high-power polynomial basis ensures calculation of the kth derivative. As the spline has a certain group of parameters (power index, defect, and nodes), it is possible to vary the accuracy and time characteristics of the spline. At the same time, the form of presentation and the method of obtaining the smoothing spline largely determine possible regimes of estimating the derivative. It is reasonable to form recurrent computational schemes for real-time estimation of the observed process parameters. The method of sequential spline smoothing [11, 12] does not require solving systems of equations, but data processing is performed a posteriori, and the number of sequential smoothing events is a regularization parameter.

In this work, we provide a computational scheme of a recurrent smoothing spline (RSS) with the number of measurements of each segment being greater than the number of nodes. The recurrent smoothing spline $S(t)$ is found by means of optimization of the extreme functional, which was modified for the proposed spline and has the form [13]

$$J(S) = (1-\rho)(h\Delta t)^2 \int_{t_0^i}^{t_h^i} [S''(t)]^2 dt + \rho \sum_{j=0}^{h} [S(t_j^i) - y(t_j^i)]^2, \qquad (3)$$

where ρ is the weight coefficient establishing the balance between the smoothing and interpolating properties of the spline $S(t)$, $\rho \in [0, 1]$, Δt is the interval of sampling of the observed process, t_0^i and t_h^i are the initial and final abscissa values of the ith segment of the spline, respectively, $h = (t_h^i - t_0^i)/\Delta t$ is the number of measurements inside the ith segment of the spline (in what follows, $h = \text{const}$ for all spline segments on the process observation interval), and $y(t_i)$ are the measurements defined by Eq. (1).

For the ith segment, the analyzed cubic spline $S_i(\tau) = a_0^i + a_1^i\tau + a_2^i\tau^2 + a_3^i\tau^3$ can be calculated at any internal point of the segment $\tau \in [1, h]$.

The conditions of conjugation of the spline segments define the formulas for its continuous coefficients. The discontinuous spline coefficients are found from the condition of minimization of the extreme functional (3) as a function of the spline defect d. The discontinuous coefficients are a_2^i and a_3^i for $d = 2$ and a_3^i for $d = 1$.

Several modes of RSS operation are identified depending on the relationship between the moments of conjugation $q \in [0, h-1]$ and calculation $\tau \in [1, h]$ of the ith spline (Fig. 1), for example, the current mode of estimation, where the spline segment is calculated after the next measurement becomes available, with the use of the previously obtained $h - 1$ measurements. Conjugation of the continuous derivatives in this regime is performed at the time instant $t_{q+1}^{i-1} = t_q^i$.

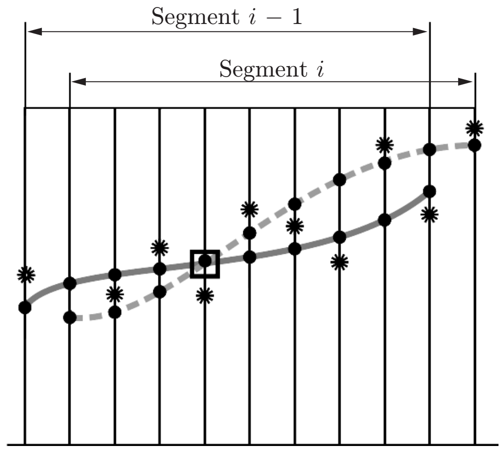

Fig. 1. Current mode of RSS operation: observed process (∗), calculated values of τ (•), and instant of conjugation of the RSS segments q (□).

The formulas for estimating the coefficients in this mode for $q = 0$ and $d = 2$ have the simplest form:

$$\begin{aligned} a_0^i &= a_0^{i-1} + a_1^{i-1} + a_2^{i-1} + a_3^{i-1}; \\ a_1^i &= a_1^{i-1} + 2a_2^{i-1} + 3a_3^{i-1}; \\ a_2^i &= \frac{\rho(F_1^i C - F_2^i A)}{BC - A^2}; \\ a_3^i &= \frac{\rho(F_2^i B - F_1^i A)}{BC - A^2}; \end{aligned} \qquad \begin{aligned} A &= 6(1-\rho)h^4 + \rho H_5; \\ B &= 4(1-\rho)h^3 + \rho H_4; \\ C &= 12(1-\rho)h^5 + \rho H_6; \\ F_1^i &= \sum_{k=0}^{h} y(t_k^i)k^2 - a_0^i H_2 - a_1^i H_3; \\ F_2^i &= \sum_{k=0}^{h} y(t_k^i)k^3 - a_0^i H_3 - a_1^i H_4; \\ H_n &= \sum_{k=0}^{h} k^n. \end{aligned} \tag{4}$$

The coefficients a_0^i, a_1^i are found from the condition of conjugation of the segments $S^{(k)}(t_{q+1}^{i-1})_+ = S^{(k)}(t_q^i)_-$, $k = 0, 1$, and the coefficients a_2^i and a_3^i are obtained from the conditions of minimization of functional (3) $\frac{\partial J(S)}{\partial a_2^i} = 0$ and $\frac{\partial J(S)}{\partial a_3^i} = 0$.

The advantage of this approach is the use of explicit formulas for calculating the spline coefficients, which prevents using numerical algorithms and reduces the computation time of real-time estimation of the derivatives.

RESULTS OF THE COMPUTATIONAL EXPERIMENT

The quality of estimation of derivatives is largely determined by the RSS setting parameters. Such parameters of the spline are the balance factor ρ and the number of measurements in the spline segment h. The quality of reconstruction of the function and its derivatives in the presence of noise with a prescribed dispersion σ_ξ^2 is estimated by using the mean squared error (MSE) criterion [14]. To provide a better illustration, we express this criterion in the form of the mean squared percentage error (MSPE) as

$$\begin{gathered} \mathrm{MSE}(\widehat{x}_j(t)) = \sqrt{\frac{1}{K}\sum_{j=1}^{K}\frac{1}{n}\sum_{i=1}^{n}(x(t_i) - \widehat{x}_j(t_i))^2}; \\ \mathrm{MSPE} = \frac{\mathrm{MSE}}{|x_{\max} - x_{\min}|} \cdot 100\ \%, \end{gathered} \tag{5}$$

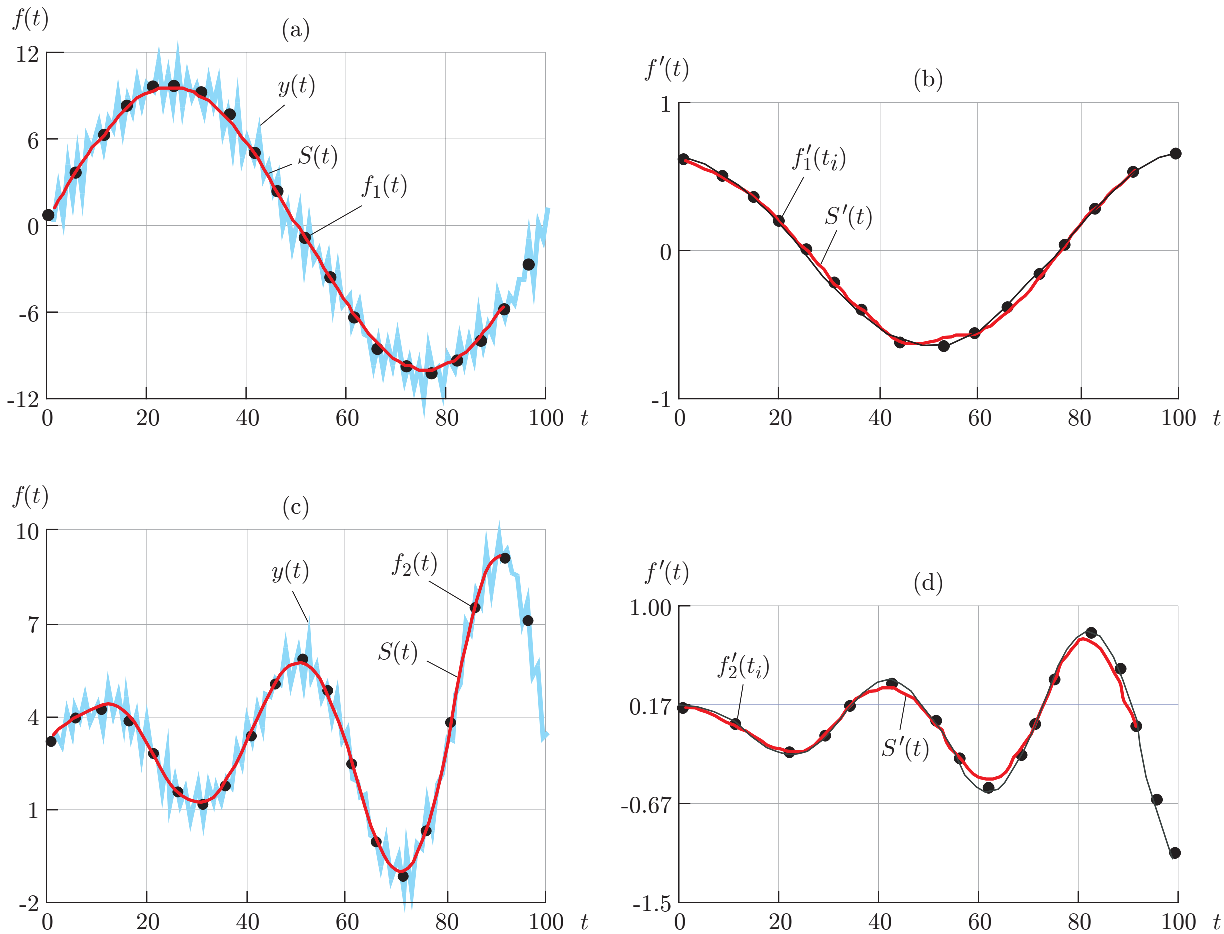

Fig. 2. Test functions and reconstructed derivatives.

where $\widehat{x}_j(t_i)$ is the estimate of the function (derivative) at the time instant t_i in the jth experiment, $x(t_i)$ is the true value of the function (derivative) at the time t_i, K is the number of experiments, n is the number of the function values, and $|x_{\max} - x_{\min}|$ is the range of variation of the test function (derivative). For comparisons with the results of derivative estimation [15], we use similar parameters of the computational experiment: $K = 200$ and $n = 100$.

The optimal values of the RSS setting parameters are found through the computational experiment with the use of the test functions

$$f_1(t) = 10\sin\left(\frac{2\pi i}{100}\right), \qquad f_2(t) = \sin\left(\frac{\pi i}{20}\right)\mathrm{e}^{002t} + 3,$$

which are plotted in Figs. 2a and 2c, respectively.

The figures show the useful signal (points), the reconstructed functions (red curves, see Figs. 2a and 2c), and the reconstructed derivatives (red curves, Figs. 2b and 2d). The results are obtained for a particular case with $h = 10$ and $\sigma_\xi = 10$ % of the maximum value of the useful signal. It is visually seen from the graphs that the derivative is adequately reconstructed. However, despite a similar error of reconstruction (smaller than 2%), the derivative in Fig. 2b is smoother than that in Fig. 2d. The derivative reconstruction smoothness is controlled by the parameter ρ. Excessive enhancement of smoothness gives rise to a systematic deviation, which is another undesirable effect of reconstruction.

The present study allowed us to estimate the MSPE values of reconstruction of the function σ_0 and its derivative σ_1 as functions of the noise level. Figure 3 shows the reconstruction errors in the range of noise variation $\sigma_\xi \in$ 0–20%. The derivative estimation error (blue curves) is slightly higher than the function reconstruction error (red curves); the error values completely depend on the form and properties of the original function.

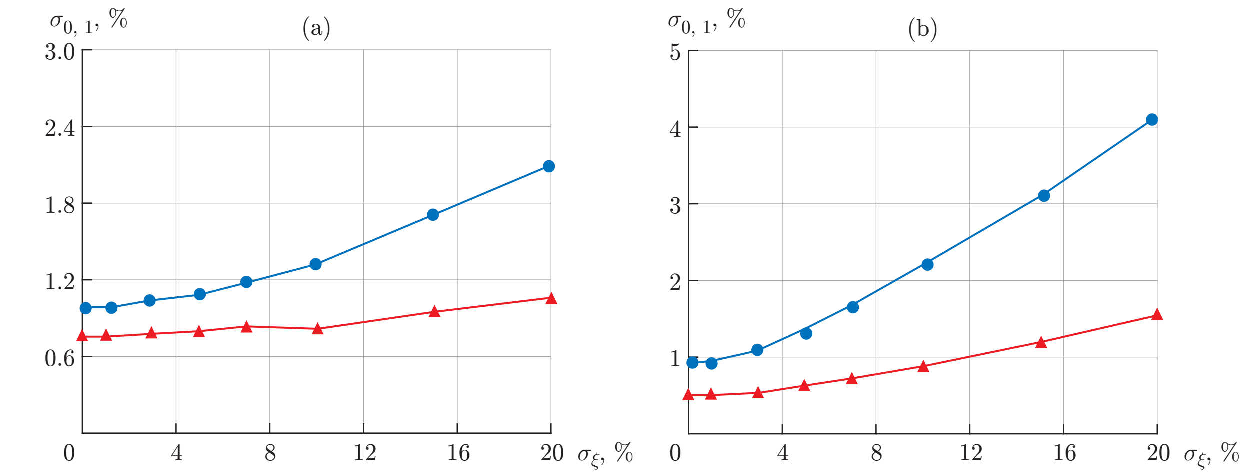

Fig. 3. Error of reconstruction of the function σ_0 and its derivative σ_1 versus the noise error σ_ξ: (a) $f_1(t)$; (b) $f_2(t)$.

Method	$\sigma_\xi = 3$ %	$\sigma_\xi = 7$ %	$\sigma_\xi = 10$ %	Comment
P-spline	0.093 0.442 %	0.204 0.972 %	0.273 1.300 %	MSE [15] MSPE
RSS	1.09 %	1.688 %	2.218 %	$f_1(t)$, h, $\rho_{\rm opt}$, $\tau = 1$
RSS	1.042 %	1.176 %	1.334 %	$f_2(t)$, h, $\rho_{\rm opt}$, $\tau = 1$

The error values are given in the figure for optimal values of the setting parameters h and ρ. The error of the initial data is seen to decrease noticeably in the entire estimated interval: the MSPE reduction factor varies from 3 (for $\sigma_\xi < 10$ %) to 10 (for $\sigma_\xi > 15$ %).

As the noise error σ_ξ increases, the optimal length h of the spline also increases: from $h_{\rm opt}$ = 8–10 measurements for $\sigma_\xi = 1$ % to $h_{\rm opt}$ = 14–19 measurements for $\sigma_\xi = 20$ %; in this case, $\rho_{\rm opt} \in [0, 1]$.

To conclude, let us compare the MSPE of derivative reconstruction by the proposed method (RSS) with the results of using a P-spline (Parametric Penalized Spline Smoothing) reported in [15].

It is seen from the table that the P-spline ensures a smaller error, but it is the RSS that ensures real-time reconstruction of the function and its derivative.

CONCLUSIONS

The paper describes a regression model of reconstruction of the function and its derivative, which is based on the variational approach. The main requirement to the smoothed function is the possibility of using it in the real time. A recurrent computational scheme containing explicit formulas was developed for this purpose. A comparative analysis of the proposed RSS and P-spline shows that the reconstruction error ensured by the RSS is slightly higher (by a factor of 1.5 on the average); nevertheless, it is fairly reasonable for engineering applications. It should be noted that the RSS allows real-time implementation and the estimates provided by such procedures only asymptotically tend to the estimates of a posteriori algorithms, which are represented here by the P-spline.

REFERENCES

1. A. N. Golubinskii, "Methods of Approximating Experimental Data and Constructing Models," Vestn. Voronezh. Inst. MVD Possii, No. 1, 138–143 (2007).
2. Yu. S. Zav'yalov, B. I. Kvasov, and V. L. Miroshnichenko, *Methods of Spline Functions* (Nauka, Moscow, 1980).
3. V. V. Vershinin, Yu. S. Zav'yalov, and N. N. Pavlov, *Extreme Properties of Splines and Smoothing Problems* (Nauka, Novosibirsk, 1988).

4. A. I. Rozhenko, "Theory and Algorithms of Variational Spline Approximation," Doctor's Dissertation, Inst. Comp. Math. Math. Geophys. Novosibirsk, 2003.
5. Yu. E. Voskoboinikov and A. B. Kolker, "Approximation of Images by Smoothing Splines," Avtometriya **39** (4), 3–12 (2003).
6. V. I. Denisov and A. V. Faddeenkov, "Spline Regression with Variable Penalty Coefficients," Avtometriya **51** (3), 3–10 (2015) [Optoelectron., Instrum. Data Process. 51 (3), 213–219 (2015)].
7. I. N. Khaimovich and L. S. Klentak, "Improvement of Methods of Smoothing Complex Surfaces with the use of Interpolation Splines," Fund. Issled., No. 10 (part 12), 2634–2638 (2013).
8. K. de Brabanter, J. de Brabanter, and B. de Moor, "Derivative Estimation with Local Polynomial Fitting," J. Mach. Learning Res. **14** (1), 281–301 (2013).
9. E. G. Zhilyakov, A. A. Chernomorets, and N. S. Pabolkova, "Differentiation of Signals over Discrete Samples on the Basis of the Frequency Concept," Nauch. Ved. Belgor. Gos. Univ., No. 19, 88–92 (2010).
10. V. I. Dmitriev and Zh. G. Ingtem, "Solving an Integral Equation of the First Kind by Spline Approximation," Computat. Math. Modeling **15** (2), 99–104 (2004).
11. A. I. Grebennikov, "Method of Sequential Smoothing and Spline Algorithms," in: Numerical Analysis: Methods, Algorithms, Applications (Izd. Mosk. Gos. Univ., Moscow, 1985), pp. 82–95.
12. J. W. Schmidt and I. Scholz, "A Dual Algorithm for Convex-Concave Data Smoothing by Cubic C^2-Splines," Numer. Math. **57**, 330–350 (1990).
13. E. A. Kochegurova and E. V. Shebeko, "Application of a Variational Smoothing Spline in the Problem of Short-Term Prediction," Izv. TPU **309** (7), 36–39 (2006).
14. A. I. Nizamitdinov, "Application of a Penalty Spline (P-Spline) and a Cubic Smoothing Spline for Time Series Predictions," Vestn. Tadzhik. Gos. Univ. Prava, Biznesa i Politiki, Ser. Gum. Nauki, No. 4, 124–131 (2012).
15. J. Cao, J. Cai, and L. Wang, "Estimating Curves and Derivatives with Parametric Penalized Spline Smoothing," Stat. Comput. **22** (5), 1059–1067 (2012).