Improving the precision of the selection of piecewise linear useful signal component under conditions of a priori uncertainty

I S Shrayfel, V I Marchuk, A B Mihailov, I M Malcev, A V Fisunov
Don State Technical University, Laboratory of electrochemical and hydrogen energy, Town of Shakhty, Rostov Region, Russia

Abstract. The paper is devoted to the analytical substantiation of a new approach to the processing of signals representing a collection of some piecewise linear signals under conditions of a priori uncertainty about its parameters. The results of computer simulation fully confirm the main theoretical results and allow us to conclude that the new method is highly efficient in processing piecewise linear signals under conditions of a priori uncertainty about the useful signal function and the statistical characteristics of the additive noise component. The developed method makes it possible to reduce the dispersion of the additive noise component up to 10 times, however, as the dispersion of the noise increases, the efficiency decreases.

1 Problem statement
Currently, when conducting research, a priori information about the function of the measured process and the statistical characteristics of the additive noise component is usually absent, or extremely small. With this formulation of the problem, the number of methods and the algorithm for their implementation are directly dependent on the amount of a priori information, the less it is, the fewer methods for its effective processing [1-5]. In this regard, the development of new methods that make it possible to implement the processing of measurement results with a limited amount of a priori information about the useful signal function and statistical characteristics of the additive noise component is a very urgent problem, which has a high both theoretical and practical value. [6-20].

The purpose of this work is to develop a new method that allows to reduce the error of extracting a useful signal, which is a piecewise linear function with a limited amount of information about its parameters, based on the use of the least squares method. Let, as a result of measurements, the received signal can be represented as the sum $y(t) = u(t) + \eta(t)$, where the values $y_1, y_2, ..., y_n$ are obtained at the moments of time $t_1 < t_2 < ... < t_n$. We will consider the deterministic signal $u(t)$ is a piecewise linear function.

2 Problem solution
Let $u_1, u_2, ..., u_n$ be a sequence of useful signal values unknown to us: $u_i = u(t_i), i = 1, 2, ..., n$. It is obvious that there is a need to select the approximation of the useful signal among the sequences $x_1, x_2, ..., x_m$, for which there is an integer $1 \leq m \leq n$. A sequence of integers $(n_k)_{k=0}^m$ and two sequences of real numbers $(a_k)_{k=1}^m, (b_k)_{k=1}^m$ satisfy the indicated below properties (1), (2).

$$n_0 = 0 < n_1 < n_2 < ... < n_m = n;$$

(1)
\[ x_i = a_k t_i + b_k, i = n_{k-1} + 1, n_{k-1} + 2, \ldots, n_k; k = 1, 2, \ldots, m. \]  

(2)

In addition, there is no doubt that for each \(1 \leq k \leq m\) as the values \(a_k, b_k\) in (2) coefficients should be taken \(a_k^*, b_k^*\) OLS-line equations, built from the points \((t_i, y_i)\), \(n_{k-1} + 1 \leq i \leq n_k\). In this regard, condition (1) is strengthened by supplementing it with the requirement

\[ n_k \geq n_{k-1} + l, k = 1, 2, \ldots, m, \]  

(3)

where \(l\) is a predetermined integer from the segment \([2; n]\).

Note that in fact the value of \(l\) should be greater than 2. This is explained by the fact that the equality \(n_k = n_{k-1} + 2\) implies the equalities \(x_{n_{k-1} + 1} = y_{n_{k-1} + 1}, x_{n_k} = y_{n_k}\), which reduce the quality of the approximation. Let us give a recommendation for choosing the value of \(l\). We will consider the moments \(t_i\) equidistant on the time axis \(t\) (which is typical for discretization of a continuous signal): \(t_i = (i - 1)h, i = 1, 2, \ldots, n\), where \(h(>0)\) is a given constant. Note that the approximation of a pure signal by piecewise linear functions makes sense only if the step \(h\) is sufficiently small in comparison with the lengths \(L_k\) \(I_k\) linearity intervals \(I_k\) of the function \(u(t)\). Let \(\Delta = \min L_k\). Let us show that if \(\Delta\) exceeds the value \(lh\), each of the intervals \(I_k\) contains at least, \(l\) points \(t_i\) even in the case when the ends of the interval do not belong to it. First, we note that at least one of these points will appear in \(I_k\) (otherwise it would entail an impossible inequality: \(h \geq L_k \geq \Delta > lh \Rightarrow l < 1\)). Let \(j_k h / I_k\) be the smallest (largest) integer for which \(j_k h \in I_k\) (respectively, \(J_k h \in I_k\)). Then \(lh < \Delta \leq L_k \leq (J_k + 1)h - (j_k - 1)h = (j_k - j_k + 2)h\), whence \(l < j_k - j_k + 2 \Leftrightarrow l \leq j_k - j_k + 1\), as required to prove \((j_k - j_k + 1\) is the number of points \(t_i\) from the interval \(I_k\).

Let us now assume that the value of \(\Delta_0\), is known, practically reliably less than \(\Delta\). Then the choice of the step \(h \leq \Delta_0/l(\leq \Delta/l)\) with the same reliability will allow us to assume that constraint (3), in fact, is not such. This follows from the just proved statement: since each linearity interval of a pure signal contains at least \(l\) points \(t_i\), it is natural to demand the same from its approximation. The above-mentioned requirement \(l > 2 \Leftrightarrow l \geq 3\) is updated in the case of choosing the step \(h \leq \Delta_0/3\). After setting the step \(h\), the best allowable value \(l\) will be \(l_0 = [\Delta_0/h]\) (the largest integer \(l\), satisfying the inequality \(h \leq \Delta_0/l\). Here \([x]\) – is the integer part of the number \(x\). If it is impossible to set \(\Delta_0\) due to the lack of a priori information about the lengths of the intervals \(I_k\), \(l_0 = 3\) should be taken.

Due to the above, filtering the signal \(y(t)\) is reduced to minimizing the sum \(\sum_{k=1}^{m} \sum_{i=n_{k-1}+1}^{n_k} (a_k^* h(i-1) + b_k^* - y_i)^2\) over the set of all sequences \((n_k)_{k=0}^{m}\), satisfying conditions (1) and inequalities \(n_k \geq n_{k-1} + l_0, k = 1, 2, \ldots, m\). In turn, this problem can be solved by plotting the shortest route on a weighted digraph with vertices \(0, 1, 2, \ldots, n\) and arcs \((p, q)\), where \(0 \leq p < q \leq n\).

The weight of the arc \((p, q)\) is given by the formula \(d_{pq} = \sum_{i=p+1}^{q} (a_{pq} h(i-1) + b_{pq} - y_i)^2\) provided that \(q \geq p + l_0\), \(a_{pq} t + b_{pq}\) is the equation of the OLS-line built from the points \((t_i, y_i)\), \(i = p + 1, p + 2, \ldots, q\). The weights of all other arcs of the digraph are set equal \(+\infty\).

In conclusion, we derive formulas for calculating the coefficients \(a_{pq}, b_{pq}\), which are necessary for specifying the finite weights \(d_{pq}\) of the arcs of the digraph \(G\). The ordered pair \((a_{pq}, b_{pq})\) is the only solution of the system of linear equations

\[
\begin{align*}
\alpha_{pq} h a_{pq} + (q-p) b_{pq} &= \sum_{i=p+1}^{q} y_i, \\
\beta_{pq} h^2 a_{pq} + \alpha_{pq} h b_{pq} &= \sum_{i=p+1}^{q} y_i h(i-1),
\end{align*}
\]  

(4)

(5)

where \(\alpha_{pq} = \frac{1}{h} \sum_{i=p+1}^{q} t_i = \frac{\sum_{i=p+1}^{q} (i-1)}{2} = \frac{(p+q-1)(q-p)}{2} \) ;
\[ \beta_{pq} = \frac{1}{h^2} \sum_{i=p+1}^{q} t_i^2 = \sum_{i=p+1}^{q} (i-1)^2 = \sum_{i=1}^{q-1} i^2 - \sum_{i=1}^{p-1} i^2 = \frac{(q-1)q(2q-1)-(p-1)p(2p-1)}{6}; \quad \text{here we used the well-known formula } \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}, n = -1, 0, 1, 2, \ldots \ (\text{for } j > s \text{ the sum of the form } \sum_{i=j}^{s} \text{ is considered equal to zero}). \]

We divide both sides of Eq. (5) by \( h \) and introduce the notation \( Y_j = \sum_{i=1}^{j} y_i, \bar{Y}_j = \sum_{i=1}^{j} y_i(i-1), j = 0, 1, 2, ..., n. \) Then the right side of equation (4) (equation (5)) can be given the form \( Y_q - Y_p \) (respectively, the form \( \bar{Y}_q - \bar{Y}_p \)). Now, according to Cramer's rule, we obtain
\[
\begin{align*}
\Delta_{pq} &= h(\alpha_{pq} - (q-p)\beta_{pq}), A_{pq} = \alpha_{pq}(Y_q - Y_p) - (q - p)(\bar{Y}_q - \bar{Y}_p), B_{pq} = h(\alpha_{pq}(\bar{Y}_q - \bar{Y}_p) - \beta_{pq}(Y_q - Y_p)).
\end{align*}
\]

The obtained processing results were checked using machine modeling on a piecewise linear function model in the presence of additive normal noise with zero mathematical expectation, which are shown in Figure 1.

![Figure 1](image-url)

**Figure 1.** Useful signal model \( U(t) \) and noisy model of the measured process \( U(t) + \eta(t) \)

The simulation was carried out using the model of a piecewise linear function presented in Figure 1 and an additive noise component with variance taking values from 1 to 5 with a step equal to 1. Figure 2 shows the results of processing the measured process \( U(t) + \eta(t) \) when the noise variance is equal to 3. When analyzing the research results, it should be taken into account that there was no a priori information about the initial and final coordinates of each of the steps. There was also no information about the slope of the piecewise linear function on each of the intervals. In the figure presented, the solid line shows the model of the useful component of the initial implementation of the measurement results, in the dotted line the estimate of the useful signal obtained as a result of using the proposed method of processing the measurement results.
Figure 2. Model processing results $U(t) + \eta(t)$ where $\sigma^2=3$

Comparative analysis of the simulation results allows us to conclude that the evaluation of the useful signal with almost a single probability determined the initial and final coordinates of each of the steps. The error lies in determining the slope of the piecewise linear function, which is explained by the influence of the noise structure.

In conclusion, let us consider the dependence of the error on the variance of the additive noise component, which is shown in Figure 3.

Figure 3. Dependence of the coefficient of attenuation of the variance of additive noise on the value of its variance

The results presented in Figure 3 allow us to conclude that the efficiency of the proposed method decreases with increasing variance. For example, when $\sigma^2=1$, there is a decrease in variance by more
than 9 times, with $\sigma^2=3$ only 3 times, and with $\sigma^2=5$ only by 15 percent, which is explained by the fact that the probability of correct estimation of the slope decreases piecewise linear function.

**Discussion**

A new method of filtering a useful signal presented as a piecewise linear function is proposed. Optimization of this method using the least squares method is analytically proved.

**Conclusions**

1. The results of computer modeling confirmed the theoretical results of processing piecewise linear functions under conditions of a priori uncertainty.
2. The developed method makes it possible to reduce the dispersion of the additive noise component up to 10 times, however, as the dispersion of the noise increases, the efficiency decreases.

**References**

[1] Marchuk V I *Estimation of the Error of Approximation of the Useful Component When Dividing the Implementation of Measurement Results Into Intervals.* Modern information technologies. 2014. № 19, pp. 153-159

[2] Marchuk V, Makov S, Timofeev D, Pismenskova M, Fisunov A *A Method of Signal Estimation Error Reduction in a Priori Indeterminacy.* В сборнике: 2015 23rd Telecommunications Forum, TELFOR 2015. 23. 2015. pp. 400-403.

[3] Marchuk V I *Estimation of the Error of Approximation of the Useful Component When Dividing the Implementation of Measurement Results Into Intervals.* Telecommunications. 2010. № 8. pp. 12-16

[4] Marchuk V, Chernyshov D, Sadrtdinov I, Miniev A *Research of the Probability of the "flip" of Approximating Function During the Processing of Measurement Results.* EDP Sciences, 2019. – vol. 104.

[5] Marchuk V, Shrayfel I, Malcev I *Solving the Problem of Mirroring the Signal Function with Respect to a Straight Line.* AIP Conf. Proceedings: Fundamental Methods of System Analysis, Modeling and Optimization of Dynamic Systems, 2019, XV International Scientific-Technical Conference “Dynamics of Technical Systems” (DTS-2019), Rostov-on-Don, Russian Federation, September 11-13, 2019. https://doi.org/10.1063/1.5138435

[6] Chakrabarty S, Levkowitz H *Denoising and stability using independent component analysis in high dimensions—visual inspection still required // 2019 23rd International Conference Information Visualisation (IV).* – IEEE, 2019. – P. 181-185.

[7] Frigo G et al. *Statistical model of measurement noise in real-world PMU-based acquisitions // 2019 International Conference on Smart Grid Synchronized Measurements and Analytics (SGSMA).* – IEEE, 2019. – P. 1-8.

[8] Guo F, Zhang C *Edge preserving mixed noise removal // Multimedia Tools and Applications.* – 2019. – T. 78. – №. 12. – P. 16601-16613.

[9] Yang Y et al. *Distributed fusion for nonlinear uncertain systems with multiplicative parameters and random delay // Signal Processing.* – 2019. – T. 157. – P. 198-212.

[10] Liu X, Liu Y, Dong Z *An ENOB Evaluation Method for an Acquisition Channel // Journal of Circuits, Systems and Computers.* – 2019. – T. 28. – №. 11. – P. 1950185.

[11] Saulig N et al. *Extraction of useful information content from noisy signals based on structural affinity of clustered TFDs’ coefficients // IEEE Transactions on Signal Processing.* – 2019. – T. 67. – №. 12. – P. 3154-3167.

[12] Jung H M, Kim B H, Kim M Y *Residual Forward-Subtracted U-Shaped Network for Dynamic and Static Image Restoration // IEEE Access.* – 2020. – T. 8. – P. 145401-145412.

[13] Liu X et al. *An Efficient Collaborative Filtering Method for Image Noise and Artifact Removal // IEEE Access.* – 2020. – T. 8. – P. 124158-124171.
[14] Jin Z et al. *Noise Characterization for Time Interleaved Photonic Analog to Digital Converters* // Journal of Lightwave Technology. – 2019. – T. 38. – №. 6. – P. 1230-1242.

[15] Zhou Y et al. *Spatial Information and Angular Resolution of Sensor Array* // Signal Processing. – 2020. – P. 107635.

[16] Jiang P, Wang Q, Wu J *Efficient Noise Level Estimation Based on Principal Image Texture* // IEEE Transactions on Circuits and Systems for Video Technology. – 2019.

[17] Chen Y et al. *Five-dimensional Seismic Data Reconstruction Using the Optimally Damped Rank-reduction Method* // Geophysical Journal International. – 2019. – T. 218. – №. 1. – P. 224-246.

[18] Dao T et al. *Error Estimation of Polynomial Chaos Approximations in Transient Structural Dynamics* // International Journal of Computational Methods – 2019.

[19] Orović I et al. *An Analog Hardware Solution for Compressive Sensing Reconstruction Using Gradient-based Method* // EURASIP Journal on Advances in Signal Processing. – 2019. – T. 2019. – №. 1. – P. 1-11.

[20] Gao Z et al. *Fundamental Limits of Exact Support Recovery in High Dimensions* // Bernoulli. – 2020. – T. 26. – №. 4. – P. 2605-2638.