On some LS-method modification for parameter estimating of discrete dynamic models

Andrey Nikitenko

Dept. of Applied Math, National Technical University "KPI", Kyiv, Ukraine

feeasul@gmail.com

Abstract. *The simple and widely used method of identification of discrete dynamic models is based on approximation of dynamic equation at each discrete time point. In case of linear by parameters model this leads to well-known LS-method. In contrary to regression problems MLS estimation of parameters leads to displaced values. This work suggests two modifications of LS-method that decrease the displacement in parameter estimations. Suggested methods give significant improvements in parameter estimations, so that they can be used not only as initial values for sophisticated methods, but by themselves (in cases, that do not need extremely high precise). One of this methods is based on integrating of difference equations on "sliding interval" therefore the question of influence of "sliding interval" length on method accuracy is investigated.*

Keywords

Discrete dynamic model, identification, parameter estimation, integrating on sliding interval.

Introduction

The issue of parameter estimating for discrete dynamic systems is all-important since displacement in parameter values leads to considerable changes in model behavior, especially for long time-series. For models, that are near to stability bounds such parameters displacement can even change the inner stability properties, switching to unstable model dynamic. As simple as possible schemes of parameter estimation, yet sufficiently précised are very important for GMDH algorithms [1], since such schemes are intended for multiple usages in multilevel model structure identification. Parameter estimation for dynamic models also important in cases, when they can be used as initial values for gradient methods of model more precise definition, e.g. based on variations calculus.

Basic LS-method and its modifications

Here we deal with a discrete dynamic model

$$x_{k+1} = \sum_{i=1}^{m} a_i \cdot f_i(x_k, x_{k-1}, ..., x_{k-s}; u_k, u_{k-1}, ..., u_{k-p}), \quad k = 0,...N \tag{1}$$

that is linear by parameters a_i. Parameter values should be estimated using observed noise-corrupted data $x_k^* = x_k + \xi_k, \quad k = -s,...,N$ and measured values of control variables $u_k^*, \quad k = -p,...,N$. We will denote vector $\mathbf{Y} = \{x_1^*, x_2^*, ..., x_N^*\}$; also we will denote matrix

$$\mathbf{X}=\begin{bmatrix} f_1\left(x_0^*,x_{-1}^*,...x_{-s}^*;u_0^*,u_{-1}^*,...,u_{-p}^*\right) & ... & f_n\left(x_0^*,x_{-1}^*,...x_{-s}^*;u_0^*,u_{-1}^*,...,u_{-p}^*\right) \\ f_1\left(x_1^*,x_{-0}^*,...x_{-s+1}^*;u_1^*,u_0^*,...,u_{-p+1}^*\right) & ... & f_n\left(x_1^*,x_{-0}^*,...x_{-s+1}^*;u_1^*,u_0^*,...,u_{-p+1}^*\right) \\ ... & ... & ... \\ f_1\left(x_{N-1}^*,x_{N-2}^*,...x_{N-s}^*;u_{N-1}^*,u_{N-2}^*,...,u_{N-p}^*\right) & ... & f_n\left(x_{N-1}^*,x_{N-2}^*,...x_{N-s}^*;u_{N-1}^*,u_{N-2}^*,...,u_{N-p}^*\right) \end{bmatrix}$$

MLS estimations

The simplest idea of discrete dynamic model identification is the next: for real values x_k equations (1) for each k are strict, but for observed values x_k^* there will be some difference between left and right parts of (1), so

$$\Delta_{k+1}=x_{k+1}^*-\sum_{i=1}^{m}a_i\cdot f_i(x_k^*,x_{k-1}^*,...,x_{k-s}^*;u_k^*,u_{k-1}^*,...,u_{k-p}^*)$$

Minimizing $\sum_{k=1}^{N}\Delta_k^2$ one can estimate parameters $\overline{a}_i$, $i=1,...m$. Denoting vector $\mathbf{a}=\{a_1,a_2,...,a_m\}$ and using matrix $\mathbf{X}$ and vector $\mathbf{Y}$, we may rewrite MLS-estimated parameters as

$$\overline{\mathbf{a}}=\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\cdot\mathbf{X}^T\mathbf{Y} \tag{2}$$

Theoretically [2] parameter estimations $\overline{\mathbf{a}}$ have nonzero asymptotical displacement even fore simplest autoregression models, and therefore are not accurate.

Usage of equivalent integral form for dynamic equation

The dynamical equation (1) may be rewritten in difference form

$$\Delta x_k=\sum_{i=1}^{m}a_i\cdot f_i(x_k,x_{k-1},...,x_{k-s};u_k,u_{k-1},...,u_{k-p})-x_k$$

then x_k can be expressed as

$$x_k=x_0+\sum_{j=0}^{k-1}\Delta x_j=x_0+\sum_{j=0}^{k-1}\left(\sum_{i=1}^{m}a_i\cdot f_i(x_j,x_{j-1},...,x_{j-s};u_j,u_{j-1},...,u_{j-p})-x_j\right). \tag{3}$$

Last equation expresses model output in equivalent integral form of differential equation for the discrete case. Changing the order of some operations in right part of last expression we obtain the final expression which will define the model output x_k as a result of integrating differences Δx_j starting from x_0 for arbitrary k

$$x_k=x_0+\sum_{i=1}^{m}a_i\sum_{j=0}^{k-1}f_i(x_j,x_{j-1},...,x_{j-s};u_j,u_{j-1},...,u_{j-p})-\sum_{j=0}^{k-1}x_j$$

To prepare this last expression for estimating parameters $\mathbf{a}$ we may transpose last summ and x_0 from right part to the left obtaining

$$\sum_{j=1}^{k} x_j = \sum_{i=1}^{m} a_i \sum_{j=0}^{k-1} f_i(x_j, x_{j-1}, \dots, x_{j-s}; u_j, u_{j-1}, \dots, u_{j-p}).$$

Analogously to LS-method matrix of normal equations can be built. Introducing $N \times N$ matrix

$$\mathbf{D}(N) = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

and denoting $\mathbf{U}(N) = \mathbf{D}(N) \cdot \mathbf{D}^T(N)$ we will obtain

$$\overline{\mathbf{a}} = \left(\mathbf{X}^T \cdot \mathbf{U}(N) \cdot \mathbf{X}\right)^{-1} \cdot \mathbf{X}^T \cdot \mathbf{U}(N) \cdot \mathbf{Y}.$$

Usage of integrating dynamic equation on “sliding interval”

To gain model output at moment k we may use integral form of model dynamic equation with one difference: we may integrate not from initial state x_0 but from other moment proceeding x_k by fixed L discrete intervals [3]. So we may define

$$x_k = x_{k-L} + \sum_{j=k-L}^{k-1} \left(\sum_{i=1}^{m} a_i \cdot f_i(x_j, x_{j-1}, \dots, x_{j-s}; u_j, u_{j-1}, \dots, u_{j-p}) - x_j \right)$$

This change will lead to alteration in matrix $\mathbf{D}$: its dimensions change to $N \times N - L$, each column has L ones placed sequentially (other elements have zero value):

$$\mathbf{D}(N, L) = \begin{bmatrix} 1 & 0 & 0 & \dots & \dots & 0 & 0 \\ 1 & 1 & 0 & \dots & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & .. \\ 1 & 1 & 1 & \dots & \dots & 0 & 0 \\ 0 & 1 & 1 & \dots & \dots & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \dots & 0 & 1 \end{bmatrix}$$

Notice that both modifications of estimating method introduces in comparison with (2) some weighting matrix $\mathbf{U}(N)$ or $\mathbf{U}(N, L) = \mathbf{D}(N, L) \cdot \mathbf{D}^T(N, L)$. All three methods are very close, having different weighting matrix. Also previous two methods can be considered as partial cases of “sliding interval” method: MLS has “sliding interval” length $L = 1$ ($\mathbf{U}(N, L)$ converts to unique weight matrix); for full integral form $L = N$.

Numerical experiments

All numerical experiments were carried out for sample model $x_{k+1} = a_1 x_k + a_2 x_{k-1} + bu_k$ with different values of a_1, a_2. Observed values x_k were corrupted with different level of independent uniformly distributed noise. Lengths of the time-series and of "sliding interval" were varied.

Fig.1 illustrates that modeling error (while integrating sample model from initial state $\overset{*}{x}_0, \overset{*}{x}_{-1}$) is accumulated to the end of time-series and even not so large parameter displacement for basic MLS leads to improper behavior on modeling interval. At the same time both modifications preserve adequate model behavior.

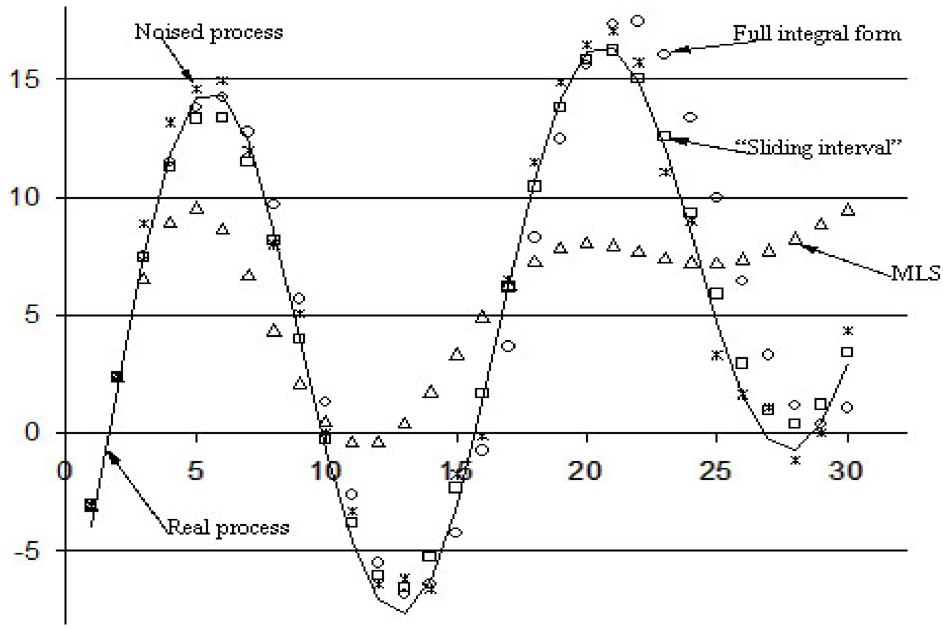

Fig. 1. Model dynamic built by different MLS-modifications.

Accuracy of identification was determined by average square error in differences of real and estimated model parameters.

$$\Delta S = \frac{\sum_{n=1}^{S} \sqrt{(a_1 - \overline{a}_1)^2 + (a_2 - \overline{a}_2)^2 + (b - \overline{b})^2}}{S}$$

Average square error ΔS was calculated while S growth. To gain a a steady ΔS value the number of noise realizations S raised up to 1000.

For the same sample model Fig.2 presents how model accuracy ΔS is changed in average for different noise realizations. Better accuracy of both MLS-modifications is stable and not accidental.

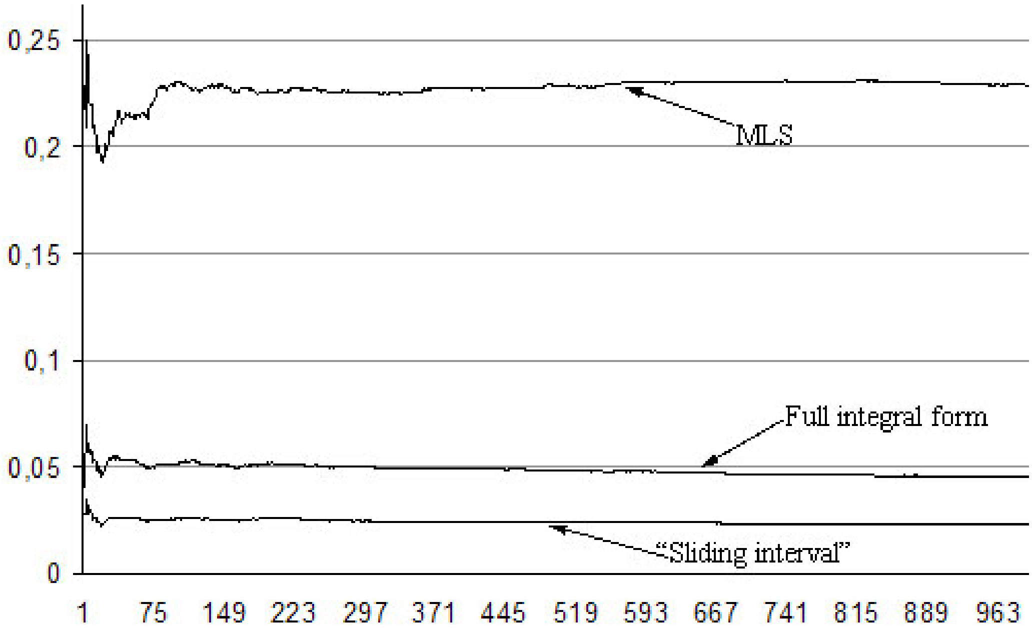

Fig. 2. Comparison of average square error ("sliding interval" length $L = 6$).

Experiments showed that integrating on "sliding interval" has some advantages; however the question of finding appropriate "sliding interval" length is raised. Estimating with different "sliding intervals" proved the hypothesis of existence of the stable "sliding interval" length that provides the best accuracy.

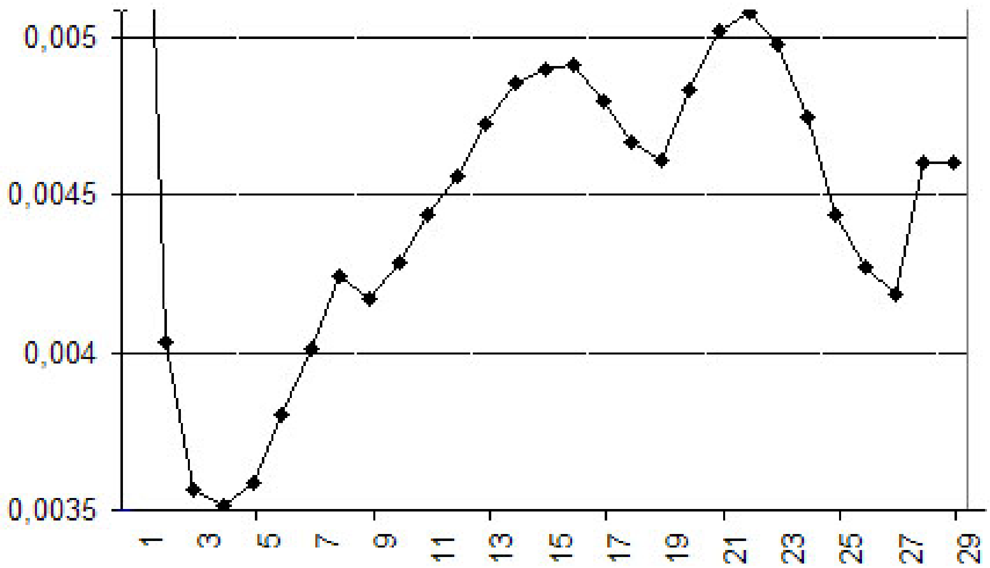

Fig. 3. Typical dependency of model average error on "sliding interval" length.

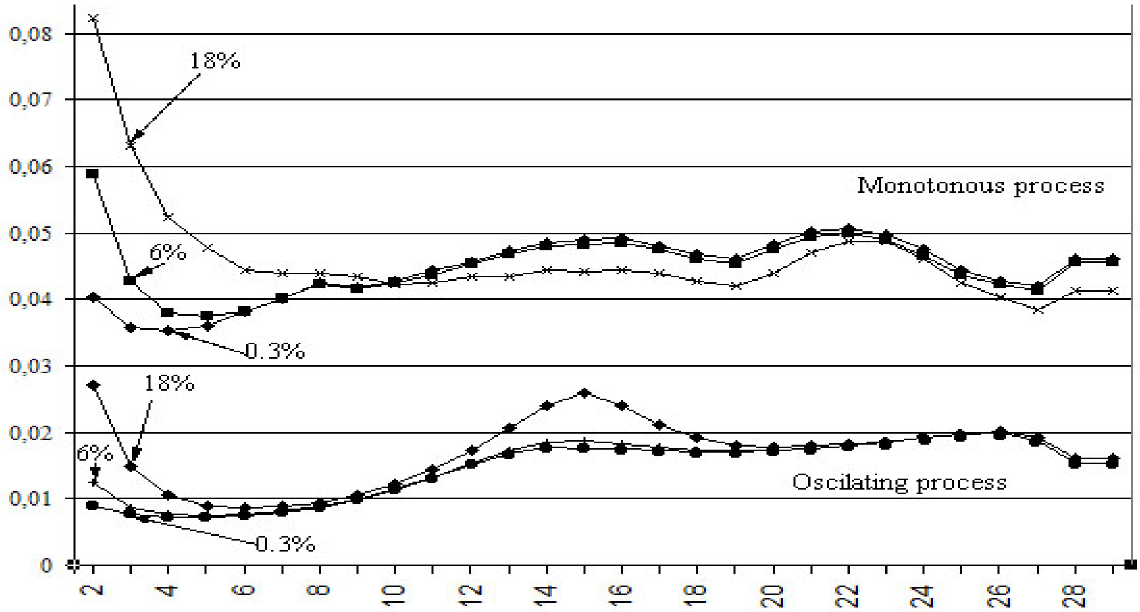

Fig. 4. Typical dependency of average error on "sliding interval" length for different noise levels and process type.

On the last figure the ΔS values are scaled by noice level to be conveniently presented on one drawing.

Conclusion

Comparing with ordinary MLS its modifications based on integral forms of dynamic equation allow gaining more precise parameter estimations for dynamic model. At that:

- MLS-modification with "sliding interval" has stable advantage over other estimating schemes;
- there always exists the best reasonable "sliding interval" length, which is probably equals doubled model dimension;
- the length of "sliding interval" practically does not depend on noise level (while it's not too high) and noise realization.

References

[1] Madala H.R., Ivakhnenko A.G.: Inductive Learning Algorithms for Complex Systems Modeling, - CRC Press, 1994 - 368p.

[2] Eykhoff P.: System Identification: Parameter and State Estimation, - Chichister, England: Wiley, 1974 – 555p.

[3] Voronova L.I., Krementulo Yu.V.: A New Method of Defining of complex dynamic systems characteristics. Complex control systems, - Kyiv, Naukova Dumka, 1968 - 286p (in Russian).