Transactions of the VŠB – Technical University of Ostrava, Mechanical Series
No. 2, 2013, vol. LIX
article No. 1971

Ján VACHÁLEK[*], Michal BARTKO[**]

ONLINE SYSTEM IDENTIFICATION METHOD USING MODIFIED REGULARIZED EXPONENTIAL FORGETTING

PRIEBEŽNÁ IDENTIFIKÁCIA SYTÉMOV S VYUŽITÍM MODIFIKOVANÉHO REGULARIZOVANÉHO EXPONENCIÁLNEHO ZABÚDANIA

Abstract

The paper deals with the use of regularized exponential forgetting (REF) in the process of online system identification. The deployment of this type of forgetting strategy is advantageous for very long runs with small changes in the identified input parameters (in the range of 100 000 steps). In these cases, the classical methods of forgetting, such as an exponential (EF) or directional forgetting (DF) lack the required quality and reach the limit of numerical stability of the calculations of system parameters, which may lead to the early termination of system identification procedure. To avoid this undesirable effect and maintain sufficient primary information about the identified system, a modified REF method is used that employs alternative covariance matrix (ACM) formulation to store the primary information of the identified system (REFACM) and prevents the numerical destabilization of the identification process. The quality of the modified REFACM forgetting method —along with its validation and comparison with REZ to verify its properties—is performed using standard tests.

Abstrakt

Príspevok sa zaoberá využitím regularizovaného exponenciálneho zabúdania (REZ) v procese priebežnej identifikácie systémov. Nasadenie tohto typu zabúdania je výhodné pre veľmi dlhé behy identifikovaného systému s málo sa meniacimi vstupnými parametrami (rádovo cez 100 000 krokov). V týchto prípadoch totiž klasické metódy zabúdania ako exponenciálne (EZ) alebo smerové zabúdanie (SZ) nedosahuje požadované kvality a dostáva sa na hranicu numerickej stability výpočtov parametrov identifikovaného systému, ktoré môže viesť až ku predčasnému ukončeniu celej identifikacie. Na zabránenie tohto nežiadúceho efektu a udržanie postačujúcej primárnej informácie o identifikovanom systéme použijeme nami navrhnutú modifikovanú metódu REZ, ktorá s využitím alternatívnej kovariančnej matice (AKM) primárne informácie o identifikovanom systéme uchováva (REZAKM) a zabraňuje tým numerickej destabilizácii identifikácie. Kvalitu modifikovanej metódy zabúdania REZAKM overíme a porovnáme s REZ na štandartných testoch, pre overenie jej vlastností.

Keywords

online identification, time varying parameters, covariance matrix, forgetting.

[*] Ing., Ph.D., Institute of Automation, Measurement and Applied Informatics, Faculty of Mechanical Engineering, Slovak University of Technology in Bratislava, Námestie Slobody 1, Bratislava, tel. (+421) 2 5249 7193, e-mail jan.vachalek@stuba.sk

[**] Ing., Institute of Automation, Measurement and Applied Informatics, Faculty of Mechanical Engineering, Slovak University of Technology in Bratislava, Námestie Slobody 1, Bratislava, tel. (+421) 2 5249 7193, e-mail michal.bartko@stuba.sk

1 INTRODUCTION

This paper is devoted to online identification methods and their practical application possibilities along with adaptive control. In this paper, we will focus on monitoring of long-run operation with time variant dynamic systems. Emphasis is given on long-run operation regimes and therefore the working mechanism with non-informative data. The process of algorithm realization is elaborated as well. Online identification methods are explored, where non-informative data that could possibly destabilize numerical computation of the identified system parameters is weighted by the chosen method to ensure „forgetting". The contribution of this paper lies in a novel algorithm based on the technique of utilizing an alternative covariance matrix. All algorithms are validated by simulations in the Matlab/Simulink software environment. Finally, the results obtained through the algorithms proposed in the article are compared to other commonly used algorithms. The observed parameters in simulations are the integral sum of the Euclidean norm of a deviation of the parameter estimates from their true values and a selected band prediction error count.

2 PROBLEM STATEMENT

Let us consider a stochastic system on which observations are made at discrete time instants k = 1, 2... . A directly manipulated input u_k and an indirectly affected output y_k (booth possibly multivariate) can be distinguished in the data pair $d_k = (u_k, y_k)$. The collection of all data observed on the system up to time t is denoted by $D_t = (d_1, d_2, ... d_t)$. The dependence of a new pair of data (u_k, y_k) on previous observations D_{k-1} can be described by a conditional probability density function p.d.f. with the following structure

$$p(y_k, u_k | D_{k-1}, \theta_k) = p(y_k | u_k, D_{k-1}, \theta_k) p(u_k | D_{k-1}) \tag{1}$$

Incomplete knowledge of the system behavior is expressed through a vector of unknown, time varying parameters $\theta_k \in \theta$. Note that the input generator described by the second term does not depend on these parameters directly, instead it is expected to utilize only prior information and information contained in observed data.

2.1 Bayes parametric inference

When the unknown parameter θ is interpreted as a random variable, the uncertainity of θ given the observed data D_t is naturally described by the posterior p.d.f. $p(\theta/D_t)$ conditional on D_t. This is generally determined by the Bayes theorem. Provided the input generator employs no other information about $p(u_k/D_{k-1}, \theta_k) = p(u_k/D_{k-1})$ k = 1,....t, the Bayes rule simplifies to the formula:

$$p(\theta | D_t) \propto p(\theta) \prod_{k=1}^{t} p(y_k | u_k, D_{k-1}, \theta) \tag{2}$$

An important special case of the first member of Eq. (1) occurs when the output y_k depends on previous data u_k, D_{k-1} over a known finite-dimensional vector function $\phi(u_k, D_{k-1}) = \phi_k$ in the following special case:

$$y_k = \theta' \Phi_k + e_k, e_k \approx N(0, \sigma^2) \tag{3}$$

Let us consider, that apriori p.d.f. is normal $N(\hat{\theta}_0, P_0)$. Then we can easily derive the shape of the posterior p.d.f. $p(\theta / D_t)$ formed by Bayesian rule (2), provided $N(\hat{\theta}_t, P_t)$ with modified recursive statistics:

$$\hat{\theta}_k = \hat{\theta}_{k-1} + \frac{P_{k-1}\Phi_k}{\sigma^2 + \Phi'_k P_{k-1} \Phi_k} \hat{e}_k, P_k^{-1} = P_{k-1}^{-1} + \frac{\Phi_k \Phi'_k}{\sigma^2} \tag{4}$$

where $\hat{e}_u = y_k - \hat{\theta}'_{k-1}\Phi_k$ is the error prediction. Recursive equation (4) is identical to the well-known algorithm of recursive least squares method [6], [5].

3 REF TECHNIQUES

Suppose that no explicit model of parameter changes is known. Yet, we can quantify our prior information (and possibly information taken from data already available) by introducing an alternative probability density function $p^*(\theta_{k+1}/D_k)$. The problem is then to construct a p.d.f. $p(\theta_{k+1}/D_k)$ based on two hypotheses described by the p.d.f. $p(\theta_k/D_k)$ (the case of no parameter changes) and the alternative p.d.f. $p^*(\theta_{k+1}/D_k)$ (the case of worst expected changes). In order to simplify notation in this section, we use $p_0(\theta)$, $p_1(\theta)$ a $p^*(\theta)$ for the posterior, alternative and resulting p.d.f.'s, respectively. The task of choosing p^* given p_0 and p_1 was formulated as a Bayesian decision making problem by Kulhavý and Kraus [3]. Next, we will make a short review of their solutions: let

$$p_{\hat{\theta},P} = \frac{1}{\sqrt{2\Pi}} |P|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(\theta - \hat{\theta}\right)' P^{-1}\left(\theta - \hat{\theta}\right)\right) \tag{5}$$

where $\hat{\theta}$ and P denote the mean and covariance of a particular p.d.f., then the following solutions were derived:

EF:

$$\hat{\theta}^* = \hat{\theta}_0, P^{*-1} = \lambda P_0^{-1} + (1-\lambda)P_1^{-1} \tag{6}$$

LF:

$$\hat{\theta}^* = \hat{\theta}_0, P^* = \lambda P_0 + (1-\lambda)P_1 \tag{7}$$

Let us consider the model of a system with time varying parameters θ_k [3]. In order to be able to track parameter variations, we complement the standard recursive last square (RLS) algorithm by exponential or linear forgetting according to (6) or (7) respectively. In addition the alternative mean is set equal to posterior mean $\hat{\theta}^{alt}_{k+1|k} = \hat{\theta}_{k|k}$, and for simplicity the alternative covariance is set equal to the prior covariance $P^{alt}_{k+1|k} = P_{1,0} = Q$. With this choice, we can use a general forgetting algorithm with the following choice of forgetting operator

$$F\left\{P_{k|k}, Q\right\} = \left[\lambda \qquad P_{k|k}^{-1} + (1-\lambda)Q^{-1}\right]^{-1} \tag{8}$$

which construct harmonic mean for recursive exponential forgetting (REF) and the prior covariance matrix Q is not forgotten and is repetitively taken into account in every step k [1], [8].

4 AUGMENTING REF WITH ACM

The involved REF augmentation considers addition and keeping the initial information in the alternative covariance matrix (ACM) form [2]. The augmentation is based on the modified Dyadic reduction algorithm, where instead of adding a-priori covariance matrix Q, ACM is computed at each step. ACM is stabilizing the evolution of matrix P(0) after the recursive update. It is necessary for the REF algorithms to be augmented by the stabilization component in the ACM form. The aforementioned stabilization component prevents the destabilization of the original algorithms at long running employments, when slow time changes are to be expected in the observed parameters in relation to the sampling period. The modified algorithm REF augmented with ACM shall be named as REFACM.

5 SIMULATIONAL ALGORITHM VERIFICATION METHODOLOGY

Two different models were created for the verification of the properties of the introduced algorithms for the observation of time variant parameters of dynamic systems. These two models (model no. 1. and no. 2.) have a different approach to input excitation (input signal generator A and B). All algorithms were subject to the same test with identical length using the two featured models [9] with Matlab Simulink representation in Fig. 1.

All results were graphically evaluated and analysed where algorithm quality has been shown numerically through the parameters known as the integral sum (IS) of the Euclidian norm of parameter error and prediction error (PE).

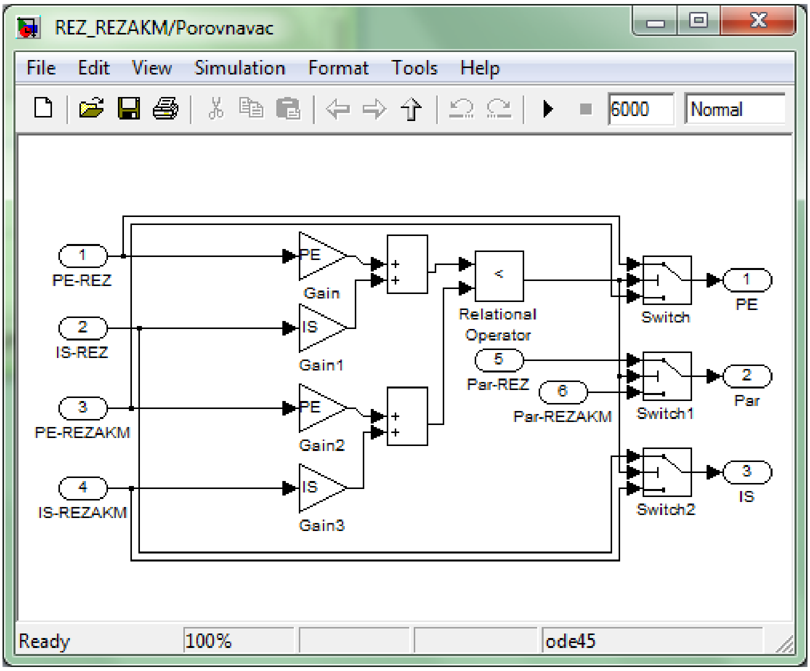

Fig 1. Matlab Simulink model for REF (REZ) and REFACM (REZAKM) verification methology

5.1 Description of model no. 1 and no. 2

In the case of model 1., a second order model is considered with external disturbance $v_{(t)}$ according to:

$$y_k = \sum_{i=1}^{2} a_i y_{k-i} + \sum_{i=0}^{2} b_i u_{k-i} + \sum_{i=0}^{2} d_i v_{k-i} + e_k, e_k \approx N\left(0, \sigma^2\right) \tag{9}$$

The values of constant parameters are given by: a_2 = -0.9, b_0 = 0.5, b_1 = -0.25, b_2 = 0.1, d_1 = 0.8, d_2 = 0.2 and σ = 0.1.

The time variant parameter has been chosen as $a_{(1)}$ = 0.98, which has been kept constant half of the n simulation steps, then at time t = n / 2 changed its value to $a_{(1)}$ = − 0.98. The outside disturbance has been simulated as a square signal periodically changing its value from +1 to −1 each hundred simulation steps. The identification has been made difficult mainly by the rarely occurring disturbances, which contained minimal information about the parameter $d_{(i)}$.

For the needs of the simulation, two input signal generators have been assumed:

- Input signal generator A: discrete white noise generator

- Input signal generator B: the input signal has been generated using the following equation: $u_k^* = 0.8u_{k-1}^* + 0.2u_k$, where u^*_k is normally distributed white noise and u^*_{k-1} is the previous input value. For model no. 2. only one change has been realized in comparison to model no. 1. This has been carried out by altering the time variant parameter $a_{(1,k)}$ = 0.98 cos(2π k/250). In this case, two different input generators were considered as well:

- Input signal generator A: discrete white noise generator

• Input signal generator B: the input signal has been generated similarly to model no 1., where $u_{(k)}$ has been only chosen from the interval $u_{(k)} \sim (0.5, 1)$.

6 VERIFICATION IN MATLAB SIMULINK ENVIROMENT

For the simulation verification procedure, a set of S-Function libraries has been created along with a common universal user interface. This interface allows the user to select input data, simulated model and the observed algorithm. The output of these simulations is a graphical representation of the observed parameters, along with a data file containing the results for the following analysis. Integral sum of the Euclidian norm of parameter error and prediction error has been shown, which is the amount exceeded by the interval $\pm 3\sigma 2$. The simulation experiments will be marked by the character pair XY, where X is the number of the utilized model (no. 1 or no. 2) and Y represents the generator utilized (A, respectively B).

7 EVALUATION OF SIMULATION RESULTS

This section introduces the results of the validation tests in a tabular form. As for the detailed description of algorithm behaviour during the simulations with different lengths, simulations lasting n = 6000, 12 000 and 120 000 have been evaluated as featured in Tab. 1 to 3.

It is clear from Tab. 2 that using simulation length n = 12 000 steps, the artefacts characteristic of long lasting runs are already appearing. The result is the confirmation of REFACM algorithm quality in comparison to REF, which in the case 1A achieved better results than pure REF.

Tab. 1 Simulation length 6 000 steps

6 000 steps	1A	1B	2A	2B
REF	IS=118,3	IS=120,3	IS=740,9	IS=1336,7
	PE=18	PE=12	PE=174	PE=103
REFACM	IS=127,4	IS=177,8	IS=963,8	IS=1985,1
	PE=27	PE=34	PE=239	PE=166

Tab. 2 Simulation length 12 000 steps

12 000 steps	1A	1B	2A	2B
REF	IS=168,9	IS=253,9	IS=1457,2	IS=2525,3
	PE=13	PE=14	PE=351	PE=188
REFACM	IS=162,9	IS=296,3	IS=1844,9	IS=4507,3
	PE=21	PE=31	PE=654	PE=439

The data featured in Tab. 3 fully confirm the previous considerations of the REFACM algorithm quality. It is clear that using ACM is well to reduce model parameter trending, which also implies the improvement of IS parameters in comparison of the results achieved by REF. The convergence of the REF covariance matrix is faster and finite in contrast to REFACM, where the convergence is slower and also the addition of excited ACM cannot be finite.

The achieved simulation results and REFACM algorithm behaviour at 6 000 and 120 000 simulation steps, show that as the running length increases the quality improves in contrast to REF see in Fig. 2. In Tab. 4 we have shown influence of weighting factor λ to quality of REFACM algorithms. The best simulation results were observed using a setting of λ=0.8.

Tab. 3 Simulation length 120 000 steps only for best performing algorithms REF and REFACM

120 000 steps	1A	1B
REF	IS=1035,4	IS=3419,5
	PE=17	PE=12
REFACM	IS=880,8	IS=2963,7
	PE=51	PE=29

Tab. 4 Influence of weighing factor λ, for REFACM algorithm on 6 000 steps length

REFACM	λ = 0.8	λ = 0.5	λ = 0.2
1A	IS=127,4	IS=164,2	IS=182,8
	PE=27	PE=34	PE=48
1B	IS=177,8	IS=298,9	IS=496,9
	PE=34	PE=50	PE=94
2A	IS=963,6	IS=870,1	IS=1031,6
	PE=239	PE=289	PE=327
2B	IS=1985,1	IS=1520,0	IS=1732,8
	PE=166	PE=145	PE=167

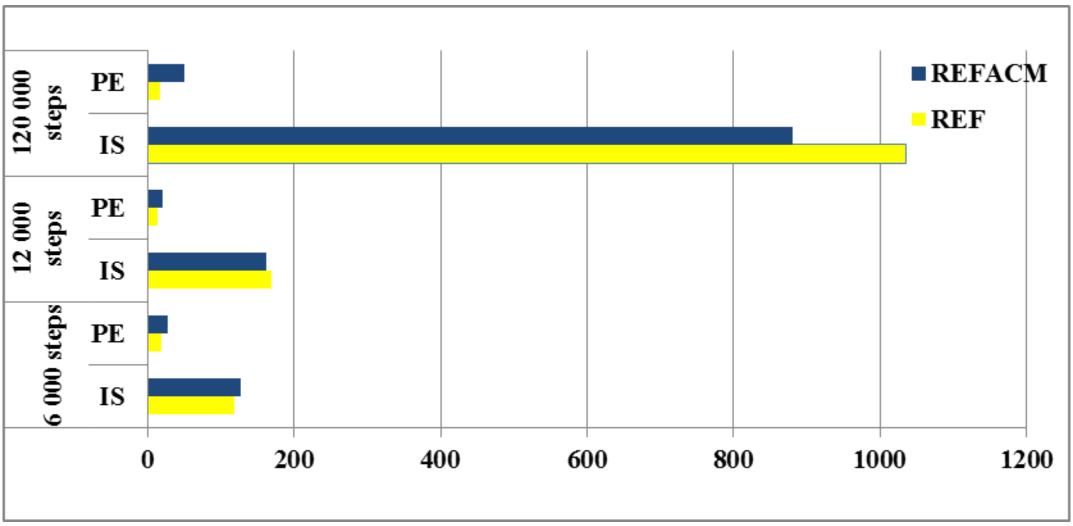

Fig 2. Simulation results for REF and REFACM algorithms (lower value of PE and IS parameter are better, focus of simulation is on long-run experiments 120 000 steps)

8 CONCLUSIONS

The simulation verification tests evaluated in the Matlab Simulink environment featured in the previous section confirm that the quality of the tested algorithms is diverse and dependant on the type of emulated experimental scenario. From the viewpoint of our interest—that is the long running simulations—the best results are achieved by the REFACM algorithm, which is the main contribution of this paper. The quality of REFACM in comparison with the other algorithms confirms the advantages of using ACM, given the specific conditions featured in this work.

The author would like to thank the financial Slovak Grant Agency APVV, project ID: APVV-0090-10 and APVV-0131-10. This support is very gratefully acknowledged.

REFERENCES

[1] SCHMITZ, U., HABER R., BARS R. (2003). A predictive On-Off controller for nonlinear processes. 14th. Int. Conference Process Control 2003, June 8-11, 2003, Štrbské Pleso, Slovakia

[2] VACHÁLEK, J. (2004). Priebežná identifikácia laboratórneho modelu s využitím dátového úložiska pre množinu linearizovaných modelov. Proceedings the 6th international scientific-technical conference Process control 2004, ŘÍP 2004, 8-11 June 2004, Kouty nad Desnou, Czech republic

[3] KULHAVÝ, R., KRAUS, F.J. (1996). On Duality of regularized Exponencial and Linear Forgetting. Automatica, 1996, vol. 32, No 10, pp. 1403-1415.

[4] PRALY, L. (1993). Robustness of model reference adaptive control. In 3th Yale WorkShop on application of adaptive system theory, 1993, Yale University, USA.

[5] PARKUM, J.E., POULSEN, N.K., HOLST, J. (1991). Analysis of Forgetting Algorithms. 9th IFAC/IFORS Symposium, Budapest, Hungary, 1991, pp. 134-139

[6] MILEK, J.J., KRAUS, F.J. (1991). Stabilized least squares estimators for time variant processes., Proceedings of the IFAC Symposium on Design Methods Of Control Systems., Vol.1, Zurich, Switzerland, 1991, pp. 430-435

[8] PETERKA, V. (1986). Control of uncertain processes: Applied theory and algorithms. Supplement to the Journal Kybernetika vol 22, 1986

[9] NIKOUKHAH, R., DELEBECQUE, F., CAMPBELL, S.L., HORTON, K. (2000). Multi-model identification an the separability index., INRIA, Rocquencourt,BP 105, Le Chesnay cedex, France, 2000.