Robust parameter estimation with noisy data

Didier Maquin[1], José Ragot[1], Fayçal Ben Hmida[2] and Moncef Gossa[2]

[1] Centre de Recherche en Automatique de Nancy
UMR 7039 - Nancy-Université, CNRS
2, Avenue de la forêt de Haye
54516 Vandœuvre-les-Nancy Cedex, France
{*Didier.Maquin, Jose.Ragot*}*@ensem.inpl-nancy.fr*
[2] Ecole Supérieure des Sciences et Techniques de Tunis (ESSTT)
5, Avenue Taha Hussein - BP 56, Bab Manara 1008 Tunis, Tunisie
U.R. Commande Surveillance et Sûreté de Fonctionnement des Systèmes (C3S)
{*Faycal.Benhmida, Moncef.Gossa*}*@esstt.rnu.tn*

Abstract. For process control improvement, coherency of information supplied by instrument lines and sensors must first be ensured; because of the presence of random and possibly gross errors, the model equations of the process are not generally satisfied. Moreover, the parameters of the considered model are not always exactly known. The problem of how to reconcile the measurements so that they satisfy the model constraints is considered in this article. The simultaneous presence of measurement errors in process input and output measurements coupled with the model parameter uncertainty poses serious problem in the rectification of data. In that paper, the problem is solved using a special filter to estimate both the parameters, the input and the output of a process represented by an autoregressive model.

1 Introduction

The estimation of the state of a process is a fundamental part of modeling, monitoring and control strategies. For example, in the field of diagnosis, the success of fault detection and isolation mainly depends on the estimation of the state of the process. Generally, for diagnosis purpose, estimation has to be performed through on-line recursive techniques. This has been extendedly studied and Kalman filter in the stochastic case [8] or Luenberger observer in the deterministic case are well known approaches. Some extensions have also been considered for processes with unknown parameters; in this case [13], general nonlinear estimation involving both data reconciliation and parameter estimation have been developed. However, in these approaches, the estimation problem is generally reduced to the state estimation, the input of the process being known.

Although our presentation is limited to linear model, the problem addressed in this article is more general than those mentioned in the previous works, since it is desired to simultaneously estimate the state, the input of the process and its parameters. In the field of process engineering, state estimation is generally seen through the classical concept of data reconciliation [9], [1]. Data reconciliation is mainly a physical problem: the variables of the process should obey the mass and energy conservation constraints. In the following, this concept of balance constraints will be considered. With a general point of view, the problem of state estimation of a process may be formulated under the following statement: knowing the measurements collected on the process, whose functioning is characterized by state variables, and knowing a model of the process involving unknown parameters, is it possible to give an estimate of the state of the process? Generally this problem is too complex and no analytical solution may be found. However, with some specifications on the measurement system and when considering particular descriptions of the model of the process, it is possible to establish the existence conditions of the solution and the solution itself [4], [5], [11]. For example, it is the case when the process and the measurement system are modeled by linear equations (in respect to the state):

$$\frac{dx(t)}{dt} = A(\theta)x(t) + Bu(t), \qquad y(t) = Cx(t)$$

for which, when the parameters θ are known, the observer theory [7] gives adequate solutions. The problem under consideration here is however more general while the parameters are unknown. Consequently, our aim is to estimate the state of the process and simultaneously the parameters of its model. On a general point of view, this problem may be addressed as a nonlinear estimation one. Cox [2] is probably one of the first being concerned with such difficulties and has proposed an iterative solution based of the maximization of the likelihood function of the measurement constrained by the model of the system. El Sherief [3] has also proposed an estimation method based on the Kalman filter. These methods are also known as "bootstrap methods" and Puthenpura [12] has given some refinement in order to increase the robustness of the estimation. Compared to the existing techniques, our contribution is directed in the following directions: first, one presents a complete analytic formulation of the estimation problem, second, one takes into account the presence on errors affecting the measurement of both input and output of the process and third, one gives to the user data which are representative of the system i.e. verifying all the state equations and with some good properties of smoothing.

2 Objective of the method

When considering single-input – single-output system, let us note x the input and y the output, both being discretized at the same sampling constant rate; $\tilde{x}$ et $\tilde{y}$ will represent the corresponding measurements. It is supposed that the true data x and y are subject to an ARX constraint:

$$y_k = \sum_{i=1}^{n} a_i y_{k-i} + \sum_{i=1}^{n} b_i x_{k-i} \tag{1}$$

From available measurements $\tilde{x}$ et $\tilde{y}$, the aim is to estimate the parameters of the system model. Unfortunately, the measured data being subject to errors, they don't verify the model constraints; thus, one tries to simultaneously estimate the true data x and y. In the following, these estimates are noted $\hat{x}$ and $\hat{y}$. Both measurements $\tilde{x}$ and $\tilde{y}$ are subjected to additive errors:

$$\tilde{x}_k = x_k + \varepsilon_{x_k} \tag{2a}$$

$$\tilde{y}_k = y_k + \varepsilon_{y_k} \tag{2b}$$

where ε_{x_k} and ε_{y_k} are supposed to be realizations of independent random variables.

The principle that can be used for the extended estimation (estimation of the parameters and the variables) is the constrained maximization of the likelihood function of the measurement errors. Otherwise, without assumption concerning the distribution of the measurement errors, it seems reasonable to minimize the sum of square deviations between the measurements and the estimates:

$$\varPhi = \sum_{k=1}^{N} (\hat{y}_k - \tilde{y}_k)^2 + \sum_{k=1}^{N-1} (\hat{x}_k - \tilde{x}_k)^2 \tag{3}$$

N being the number of considered samples. Without restriction to generality, the deviations between estimates and measurements have not been weighted; however, it is easy to introduce weights, for example according to the precision of the measurements or to use more specific information about the probability density function of the errors of measurements. As previously said, the estimates have to satisfy the constraint:

$$\hat{y}(k) = \sum_{i=1}^{n} \hat{a}_i \hat{y}_{k-i} + \sum_{i=1}^{n} \hat{b}_i \hat{x}_{k-i} \tag{4}$$

The computation of the estimates of the variables x and y and the parameters a_i and b_i is achieved by minimizing the criterion (3) taking into account the constraint (4) which is applied at each sampling time. On a numerical point of view, the problem is reduced to the optimization of a quadratic criterion (in respect to model parameters and input-output variables) under nonlinear equality constraints (the nonlinearity resulting from the link between variables and parameters). Despite of a classical and well know formulation, because the dimension of the problem (number of variables and parameters in the dynamical model) and also because the noise affecting the measurements, conventional techniques for the resolution are not always powerful; consequently, we have developed an original way based on a robust hierarchical estimation.

To present the idea, let us consider a first order system $y_k = ay_{k-1} + bx_{k-1}$ observed on a finite time interval $k \in \{1, 2, 3, 4\}$. The problem to be solved is the estimation of the parameters of the model and the estimation of the true values of the input and the output of the system. The solution of the problem results in optimizing the following function:

$$L = \frac{1}{2}\sum_{j=1}^{3}(x_j - \tilde{x}_j)^2 + \frac{1}{2}\sum_{j=1}^{4}(y_j - \tilde{y}_j)^2 \tag{5}$$

under the constraints:

$$y_k = a y_{k-1} + b x_{k-1} \quad k = 2..4 \tag{6}$$

Optimizing this criterion gives both the estimates of the parameters and the estimates of the input and output. Classical Lagrange method can be used for that purpose. Let us note that the optimality equations, in respect to $\hat{x}$, $\hat{y}$, $\hat{a}$, $\hat{b}$ and λ (where λ is the Lagrange multiplier vector) are:

$$\hat{x}_i - \tilde{x}_i - \lambda_i \hat{b} = 0, \quad i = 1, 2, 3 \tag{7a}$$
$$\hat{y}_1 - \tilde{y}_1 - \lambda_1 \hat{a} = 0 \tag{7b}$$
$$\hat{y}_2 - \tilde{y}_2 + \lambda_1 - \lambda_2 \hat{a} = 0 \tag{7c}$$
$$\hat{y}_3 - \tilde{y}_3 + \lambda_2 - \lambda_3 \hat{a} = 0 \tag{7d}$$
$$\hat{y}_4 - \tilde{y}_4 + \lambda_3 = 0 \tag{7e}$$
$$\lambda_1 \hat{x}_1 + \lambda_2 \hat{x}_2 + \lambda_3 \hat{x}_3 = 0 \tag{7f}$$
$$\lambda_1 \hat{y}_1 + \lambda_2 \hat{y}_2 + \lambda_3 \hat{y}_3 = 0 \tag{7g}$$
$$\hat{y}_i - \hat{a}\hat{y}_{i-1} - \hat{b}\hat{x}_{i-1} = 0, \quad i = 2, 3, 4 \tag{7h}$$

Despite of symmetrical aspect of these equations, no analytical solution may be exhibited. Therefore, numerical procedure must be established. For that purpose, one should notice that, using (7a) to (7e) $\hat{x}_i$ and $\hat{y}_i$ may be easily expressed in respect to λ_i, $\hat{a}$ and $\hat{b}$:

$$\hat{x}_i = \tilde{x}_i + \lambda_i \hat{b}, \quad i = 1, 2, 3 \tag{8a}$$
$$\hat{y}_1 = \tilde{y}_1 + \lambda_1 \hat{a} \tag{8b}$$
$$\hat{y}_2 = \tilde{y}_2 - \lambda_1 + \lambda_2 \hat{a} \tag{8c}$$
$$\hat{y}_3 = \tilde{y}_3 - \lambda_2 + \lambda_3 \hat{a} \tag{8d}$$
$$\hat{y}_4 = \tilde{y}_4 - \lambda_3 \tag{8e}$$

After some eliminations between (8) and (7f) to (7h), the following coupled systems may be obtained:

$$\begin{pmatrix} \lambda_1^2 + \lambda_2^2 + \lambda_3^2 & 0 \\ 0 & \lambda_1^2 + \lambda_2^2 + \lambda_3^2 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = -\begin{pmatrix} \lambda_1 \tilde{x}_1 + \lambda_2 \tilde{x}_2 + \lambda_3 \tilde{x}_3 - \lambda_1 \lambda_2 - \lambda_2 \lambda_3 \\ \lambda_1 \tilde{y}_1 + \lambda_2 \tilde{y}_2 + \lambda_3 \tilde{y}_3 \end{pmatrix}$$

$$\begin{pmatrix} 1 + \hat{a}^2 & -\hat{a} & 0 \\ -\hat{a} & 1 + \hat{a}^2 + \hat{b}^2 & -\hat{a} \\ 0 & -\hat{a} & 1 + \hat{a}^2 + \hat{b}^2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \tilde{y}_2 - \hat{a}\tilde{y}_1 - \hat{b}\tilde{x}_1 \\ \tilde{y}_3 - \hat{a}\tilde{y}_2 - \hat{b}\tilde{x}_2 \\ \tilde{y}_4 - \hat{a}\tilde{y}_3 - \hat{b}\tilde{x}_3 \end{pmatrix}$$

This system may be easily solved in respect to $\hat{a}$, $\hat{b}$ and λ_i by using a direct iteration procedure. In the following section, based on this principle, we suggest a general formulation of the estimation of the parameters and the state of the system.

3 General formulation of the problem

The model of the system is taken under the ARX representation:

$$y_k = a_1 y_{k-1} + \cdots + a_p y_{k-p} + b_1 x_{k-1} + \cdots + b_p x_{k-p} \tag{9}$$

where, without loss of generality, the orders of the AR and the X parts have been chosen equal to p. Let us define the following true value and measurement vectors:

$$z = \begin{pmatrix} x_1 \ldots x_{p+N-1} \; y_1 \ldots y_{p+N} \end{pmatrix}^T, \quad z \in \mathbb{R}^q, q = 2N-1 \tag{10a}$$

$$\tilde{z} = \begin{pmatrix} \tilde{x}_1 \ldots \tilde{x}_{p+N-1} \; \tilde{y}_1 \ldots \tilde{y}_{p+N} \end{pmatrix}^T, \quad \tilde{z} \in \mathbb{R}^q \tag{10b}$$

where:

$$\tilde{y}_i = y_i + \varepsilon_{y_i} \qquad i = 1..p+N \tag{11a}$$

$$\tilde{x}_i = y_i + \varepsilon_{x_i} \qquad i = 1..p+N-1 \tag{11b}$$

The model (9) written on the given time interval $[1,\ p+N-1]$, can be expressed by:

$$M\theta - N = 0 \tag{12}$$

with the definitions:

$$M = \begin{pmatrix} y_p & \ldots & y_1 & x_p & \ldots & x_1 \\ y_{p+1} & \ldots & y_2 & x_{p+1} & \ldots & x_2 \\ \vdots & & & & & \\ y_{p+N-1} & \ldots & y_N & x_{p+N-1} & \ldots & x_N \end{pmatrix}, \quad M \in \mathbb{R}^{N\times 2p} \tag{13}$$

$$\theta = \begin{pmatrix} a_1 \ldots a_p \; b_1 \ldots b_p \end{pmatrix}^T, \quad \theta \in \mathbb{R}^{2p} \tag{14}$$

$$N = \begin{pmatrix} y_{p+1} \; y_{p+2} \ldots y_{p+N} \end{pmatrix}^T, \quad N \in \mathbb{R}^N \tag{15}$$

In order to emphasize the linearity of the model (12) in respect to the input and output, a more convenient equivalent form of (12) consists in expanding the matrices M and N in respect to z:

$$M = \sum_{j=1}^{q} z_j M_j \quad M_j = \frac{\partial M}{\partial z_j}, \quad M_j \in \mathbb{R}^{N\times 2p} \tag{16a}$$

$$N = \sum_{j=1}^{q} z_j N_j \quad N_j = \frac{\partial N}{\partial z_j}, \quad N_j \in \mathbb{R}^{N} \tag{16b}$$

where the matrices M_j (resp. N_j) are built from the matrices M (resp. N) as follows. If the (l, m) element of M (resp. N) is equal to z_j, then the (l, m) element of M_j(resp. N_j) is equal to 1; else it is equal to 0. Thus the matrices M_j and N_j only describe the occurrence of the variable z_j.

Thus, using (16), model (12) may be alternatively written:

$$\sum_{j=1}^{q} z_j \left(M_j\theta - N_j\right) = 0 \tag{17a}$$

Thus, the simultaneous parameter and state estimation problem may be formulated as the optimization of the Lagrangian:

$$L = \frac{1}{2}\sum_{j=1}^{q} w_j^{-1} \left(z_j - \tilde{z}_j\right)^2 + \lambda^T \sum_{j=1}^{q} z_j \left(M_j\theta - N_j\right) \tag{18}$$

where w_j is a weighting factor associated to each measurement and where $\lambda \in \mathbb{R}^N$. If the measurements are only subject to random errors with zero mean, the optimal weight corresponds to the variance v_i of the measurement. On the other hand, if the measurements are corrupted by outliers, the weight can be adapted to remove their influence on the estimate. For that purpose, the following weight can be proposed [6], [10]:

$$w_j^{-1} = v_j^{-1} \frac{1}{\left(1 + \left(\frac{\hat{z}_j - z_j}{r}\right)^2\right)^2 r^2} \tag{19}$$

Clearly, the further away the estimate is from measurement, the less the corresponding weight is important, r being a tuning parameter The first order optimality conditions of the Lagrangian (18) in respect to z, θ and λ are expressed:

$$w_j^{-1}(\hat{z}_j - \tilde{z}_j) + \lambda^T(M_j\hat{\theta} - N_j) = 0 \tag{20a}$$

$$\lambda^T \sum_{j=1}^{q} \hat{z}_j M_j = 0 \tag{20b}$$

$$\sum_{j=1}^{q} \hat{z}_j(M_j\hat{\theta} - N_j) = 0 \tag{20c}$$

Solving system (20) in respect to $\hat{z}_j$, θ and λ is not easy in the general case. Approximate solution can be obtained through iterative algorithm as proposed in the following section.

4 Solution

We propose to solve the nonlinear system (20) with a direct iterative algorithm. First, from (20a), we derive:

$$\hat{z}_j = \tilde{z}_j - w_j \lambda^T (M_j \hat{\theta} - N_j) \tag{21}$$

More compactly:

$$\hat{z} = \tilde{z} - W R^T(\hat{\theta}) \lambda \tag{22a}$$
$$W = \operatorname{diag}\left(w_1 \; \dots \; w_q\right) \tag{22b}$$
$$R(\hat{\theta}) = \left(R_1(\hat{\theta}) \; \dots \; R_q(\hat{\theta})\right) \tag{22c}$$
$$R_j(\hat{\theta}) = M_j \hat{\theta} - N_j \tag{22d}$$

Then using (20c) and (21) with the definition:

$$\hat{P} = \sum_{j=1}^{q} w_j (M_j \hat{\theta} - N_j)(M_j \hat{\theta} - N_j)^T \tag{23}$$
$$= R(\hat{\theta}) W R^T(\hat{\theta}) \tag{24}$$

we deduce:

$$\lambda = \hat{P}^{-1} \sum_{j=1}^{q} \tilde{z}_j (M_j \hat{\theta} - N_j) \tag{25}$$

$$\lambda = \hat{P}^{-1} R(\hat{\theta}) \tilde{z} \tag{26}$$

By analogy with (16), let us introduce the following matrices:

$$\tilde{M} = \sum_{j=1}^{q} \tilde{z}_j M_j, \qquad \tilde{N} = \sum_{j=1}^{q} \tilde{z}_j N_j, \qquad \hat{M} = \sum_{j=1}^{q} \hat{z}_j M_j \tag{27}$$

Taking into account the definitions (22c), (22d) and (16), equation (25) can be written as:

$$\lambda = \hat{P}^{-1} (\tilde{M} \hat{\theta} - \tilde{N}) \tag{28}$$

The input-output estimate is then obtained using (22a) and (28):

$$\hat{z} = \left(I - W R^T(\hat{\theta}) \left(R(\hat{\theta}) W R^T(\hat{\theta}) \right)^{-1} R(\hat{\theta}) \right) \tilde{z} \tag{29}$$

Reporting (28) into (20b) gives:

$$\sum_{j=1}^{q} \hat{z}_j M_j^T \hat{P}^{-1} (\tilde{M} \hat{\theta} - \tilde{N}) = 0 \tag{30}$$

or equivalently using (16a):

$$\hat{M}^T \hat{P}^{-1}(\tilde{M}\hat{\theta} - \tilde{N}) = 0 \tag{31}$$

from which the parameter may be deduced:

$$\hat{\theta} = (\hat{M}^T \hat{P}^{-1} \tilde{M})^{-1} \hat{M}^T \hat{P}^{-1} \tilde{N} \tag{32}$$

Thus, (29) and (32) are implicit form of the solution $\hat{z}$ and $\hat{\theta}$. A natural way for solving this system consists in using an iterative algorithm expressed by:

$$\hat{P}_{(k)} = R(\hat{\theta}_{(k)}) W R(\hat{\theta}_{(k)})^T \tag{33a}$$

$$\hat{\theta}_{(k+1)} = \left(\hat{M}^T_{(k)} P^{-1}_{(k)} \tilde{M}\right)^{-1} \hat{M}^T_{(k)} P^{-1}_{(k)} \tilde{N} \tag{33b}$$

$$\hat{z}_{(k+1)} = \left(I - W R^T(\hat{\theta}_{(k)}) \left(R(\hat{\theta}_{(k)}) W R(\hat{\theta}_{(k)})\right)^{-1} R(\hat{\theta}_{(k)})\right) \tilde{z} \tag{33c}$$

It should be noticed that (33) looks like a least square estimator, by minimizing the residual $\tilde{M}\hat{\theta} - \tilde{N}$ with a weighting matrix $\hat{P}$ and an instrumental matrix $\hat{M}$. Summarizing, the estimation procedure is expressed by:

- 0 - initialize the parameter estimation: $\hat{\theta}_{(0)}$ (e.g. $\hat{\theta}_{(0)} = 0$)
  and the input and output estimate $z_{(0)}$, (e.g. $z_{(0)} = \tilde{z}$), $k = 0$
- 1 - compute the matrices M_j and N_j (13), (15), (16)
- 2 - compute the matrix $\hat{P}_{(k)}$ (33a)
- 3 - update the parameters $\hat{\theta}_{(k+1)}$ (33b)
- 4 - compute the input and output estimates $\hat{z}_{(k+1)}$ (33c)
- 5 - test the convergence
  if the test don't fire, $k = k + 1$ and go to step 2
  else $\hat{\theta} = \hat{\theta}_{(k+1)}$, $\hat{z} = \hat{z}_{(k+1)}$ and stop

5 Example

Data have been obtained from the simulation of a first order system with parameter values $a = 0.800$ and $b = 0.2000$. The output has been corrupted with a centered random noise and with outliers of magnitude 0.5 and -0.5. The standard least square estimation gives the following results: $a = 0.6840$ and $b = 0.16$ which are strongly affected by the outlier values. With the proposed approach, the estimated parameters are: $a = 0.800$ and $b = 0.2030$; they have been obtained in 6 iterations of the algorithm initialized with the standard least square solution.

The figure 1 shows from top to down: the input, the output, the estimated output, the output error estimation, the weights. In order to ease the representation, the weights have been normalized between 0 and 1. The estimated input has not been represented as, for that example, the estimate is close to the measurement. The dashed vertical lines indicate the time instants of the outlier occurrences. Considering the output error, the outliers have been "detected" and "removed" from the measured output; this is due to the output weights which have been automatically adjusted in order to aside the corrupted measurements.

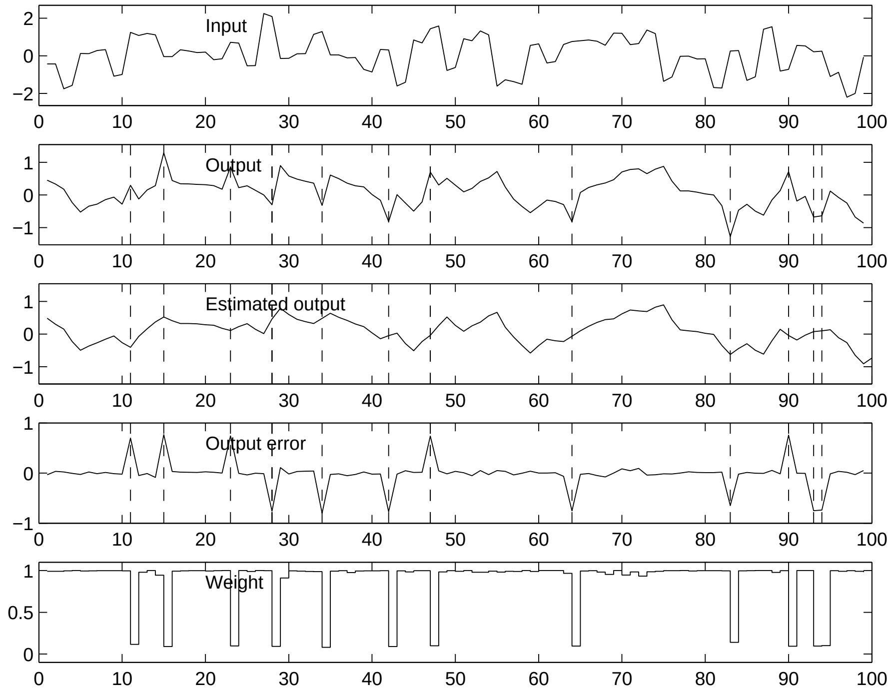

Fig. 1. Results

6 Discussion and conclusion

The problem of estimating simultaneously the input, the output and the model parameters of a system described by an ARX model has been investigated in this paper. Initially, the proposed method has been developed for signals corrupted by additive centered noises (Gaussian noises for example). Next, thanks to the introduction of adaptive weights (based on Cauchy's function or Lorentzian function) in the optimization criterion, the method has been made robust to the presence of outliers in the measured signals. The result analysis of numerous simulations shows that the parameters are always correctly estimated and the estimated output signal is filtered in a satisfactory way. Moreover, as shown by the proposed example, the occurrence of outliers can easily be taken into account and their influence minimized. On the other hand, the filtering of the input signal is rather weak and it will be necessary to study this effect and to propose some modifications of the algorithm by incorporating a filtering effect of the reconstruction of the input. A possible improvement also relates to the development of a recursive algorithm (actually, the proposed algorithm operates on a given sliding observation window).

References

1. Bousghiri, S., Kratz, F., Ragot, J. Comparison of the data reconciliation and the finite memory observer at the inverted pendulum. IFAC/IMACS Symposium on Fault Detection, Supervision and Safety for Technical Processes, SAFEPROCESS'94, Espoo, Finland, June 13-15, 1994.
2. Cox, H. On the estimation of state variables and parameters for noisy dynamic systems. IEEE Transactions on Automatic Control, AC-9, 5-12, 1964.
3. El Sherief, H., Sinha, N. Bootstrap estimation of parameters and states of linear multivariable systems. IEEE Transactions on Automatic Control, 24 (2), 340-343, 1979.
4. Gee, D.A., Ramirez, W.F. On-line state estimation and parameter identification for batch fermentation. Biotechnology Progress, 12 (1), 132-140, 1996.
5. Gove, J.H., Hollinger, D.Y. Application of a dual unscented Kalman filter for simultaneous state and parameter estimation in problems of surface-atmosphere exchange Journal of Geophysical Research. 111, 21 p., 2006.
6. Huber, P.J. Robust statistics. John Wiley, New-York, 1981.
7. Kailath, T. Linear systems theory. Englewood Cliffs, NJ, Prentice-Hall, 1980.
8. Karjala, T.W., Himmelblau, D.M.. Dynamic rectification of data via recurrent neural nets and the extended Kalman filter. AIChE Journal, 42 (8), 2225-2239, 1996.
9. Liebman, M.J., Edgar, T.F., Lasdon, L.S. Efficient data reconciliation and estimation for dynamic process using non linear programming techniques. Computers and Chemical Engineering, 16 (10/11), 963-986, 1992.
10. Özyurt, D.B., Pike, R.W. Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. Computers and Chemical Engineering, 28 (3), 381-402, 2004.
11. Liu, C.S., Peng, H. A state and parameter identification scheme for linearly parametrized systems. ASME Journal of Dynamic Systems, Measurement and Control, 120 (4), 524-528, 1998.
12. Puthenpura, S., Sinha, N.K. Robust bootstrap method for joint estimation of states and parameters of linear systems. Journal of Dynamic Systems, Measurement and Control, 108, 255-263, 1986.
13. Robertson, D.G., Lee, J.H., Rawlings, J.B. A moving horizon-based approach for least-squares estimation. AIChE Journal, 42 (8), 2209-2224, 1996.