

On New Design of Kalman Filter with Entry-wise Updating

EVGENIA SUZDALEVA

Department of Adaptive Systems

Institute of Information Theory and Automation of the ASCR

Pod vodárenskou věží 4, 18208 Prague

CZECH REPUBLIC

suzdalev@utia.cas.cz

Abstract: The paper deals with a factorized version of Kalman filter. Via factorization of a state-space model such the filter provides the state estimates of individual state vector entries in the factorized form and allows to update them entry-wise. The paper continues a series of research in the field of the factorized filtering and proposes the novel modified algorithm, including the simultaneous entry-wise organized fulfillment of data and time updating steps. A motivation of the research is a preparation of the universal algorithm for the joint filtering of variables of a mixed (continuous and discrete-valued) type. The illustrative example and comparison of computational complexity with other versions of Kalman filtering are presented.

Key-Words: factorized filtering, Kalman filter, state estimation.

1 Introduction

The paper deals with Kalman filter [1], organized in an entry-wise way. The entry-wise updating of the posterior state estimates is reached with the help of a special factorization of covariance matrices. To enable the proposed entry-wise updating, a modified algorithm of the factorized Kalman filtering has been evolved. This modification assumes the simultaneous fulfillment of the data and time updating steps for a system, whose state and output are described by the joint probability density function (pdf). The proposed filter provides the estimates of the state vector entries in the factorized form and allows to update them entry-wise.

A general motivation of the research in the field of the entry-wise filtering is a preparation of a technique of the joint estimation of the mixed-type (continuous and discrete-valued) data. The target application of the research is the traffic control systems, using state-space models. The state-space models with a car queue length as the main state variable have made a good showing. However, some of the state variables are of discrete-valued nature (signal lights, level of service, visibility, road surface, etc). Their involvement calls for the algorithm, which enables the joint filtering of the mixed-type data. The entry-wise Kalman filtering is a potential solution of this task. Due

to the entry-wise updating the involvement of the discrete-valued variables is seen optimistically.

The state of the art of the problem is as follows. Most works [2, 3] found in the field are devoted to the factorized implementations of Kalman filtering. The factorization of covariance matrices is also used in problems of systems classification, dealing with multivariate Gaussian random field [4]. However, the global aim of the mentioned works is, primarily, reduction of the computational complexity via a lesser rank of the covariance matrix. Exploitation of matrix factorization to approach the entry-wise filtering under Bayesian methodology [5] was proposed in [6] with a reduced form of the state-space model. The paper [7] removed this restriction and proposed the solution of factorized Bayesian prediction and filtering, based on applying the chain rule to the single output state-space model. The work [8] offered the version of factorized Kalman filtering with Gaussian models, which was based on the $L'DL$ decomposition of the covariance matrices. The paper [9] expanded the line with $L'DL$ -factorized covariance matrices and demonstrated the application of the solution to the traffic-control area. However, the mentioned works had problems with preserving of the distribution form of state entries and consequently with the entry-wise updating. The present paper proposes the novel algorithm

with LDL' -factorized covariance matrices and simultaneous data and time updating of the posterior state entry estimates. The entry-wise algorithm is verified and implemented without high computational complexity.

The layout of the paper is as follows. Section 2 is devoted to basic facts about Bayesian and Kalman filtering and used models. Section 3 provides a general derivation of Kalman filter with simultaneous data and time updating and presents the algorithm of its factorized version. The algorithm verification, examples and comparison of a level of computational complexity of different implementations of Kalman filter are shown in Section 4. Remarks and plans of future work in Section 5 close the paper.

2 Preliminaries

2.1 Models

The system is described by the joint probability density function (pdf)

$$f(x_t, y_t | x_{t-1}, u_t), \quad (1)$$

where x_t is the unobservable system state at discrete time moments $t \in t^* \equiv \{0, \dots, \hat{t}\}$, where \hat{t} is the cardinality of the set t^* , \equiv means equivalence, and y_t and u_t are the system output and input respectively. In the case of Gaussian state-space model, the pdf (1) includes the *state evolution model* (2) and *observation model* (3)

$$x_t = Ax_{t-1} + Bu_t + \omega_t, \quad (2)$$

$$y_t = Cx_{t-1} + Hu_t + v_t, \quad (3)$$

where $[A, C]$ and $[B, H]$ are known joint matrices of appropriate dimensions; ω_t is a process (Gaussian) noise with zero mean and known covariance matrix R_w ; v_t is a measurement (Gaussian) noise with zero mean and known covariance matrix R_v .

2.2 Factorization by the chain rule

The joint pdf (1) can be decomposed, according to the chain rule [10], into the following factorized form.

$$\prod_{i=1}^{\hat{x}} f(x_{i;t} | x_{i+1:\hat{x};t}, u_t, x_{t-1}, y_t), \quad (4)$$

$$\times \prod_{j=1}^{\hat{y}} f(y_{j;t} | y_{j+1:\hat{y};t}, u_t, x_{t-1}),$$

where \hat{x} and \hat{y} are numbers of entries of respective (column) vectors, $i = \{1, \dots, \hat{x}\}$, $j = \{1, \dots, \hat{y}\}$,

and such a notation as $x_{i+1:\hat{x};t}$ denotes a sequence $\{x_{i+1;t}, x_{i+2;t}, \dots, x_{\hat{x};t}\}$.

2.3 Bayesian filtering

Bayesian filtering [10] includes two coupled formulae.

$$f(x_t | u_t, d^{t-1}) = \int f(x_t | u_t, x_{t-1}), \quad (5)$$

$$\times f(x_{t-1} | d^{t-1}) dx_{t-1},$$

$$f(x_t | d^t) \propto f(y_t | u_t, x_t) f(x_t | u_t, d^{t-1}), \quad (6)$$

where \propto means proportionality, and data are defined as $d^t = (d_0, \dots, d_{\hat{t}})$, $d_t \equiv (y_t, u_t)$. Relation (5) represents the time updating of the state estimate, while (6) – the data updating. The application of (5) and (6) to Gaussian state-space model with Gaussian prior distribution and Gaussian observations provides Kalman filter.

3 Factorized Kalman Filtering with Simultaneous Data and Time Updating

3.1 Simultaneous Data and Time Updating

In general, the data updating (6) is obtained from Bayes rule [10]. Let the pdfs on the right side of (6) be written as a single joint pdf $f(y_t, x_t | u_t, d^{t-1})$. Using the operation of marginalization and the model (1), one obtains the expression of filtering with relations (5) and (6) to be fulfilled simultaneously (proof is available in [11]).

$$f(x_t | d^t) \propto \int f(x_t, y_t | u_t, x_{t-1}) f(x_{t-1} | d^{t-1}) dx_{t-1}. \quad (7)$$

Substitution of (4) in (7) and decomposition of the initial state according to the chain rule provide the following factorized form of (7).

$$f(x_t | d^t) \propto \int \prod_{i=1}^{\hat{x}} f(x_{i;t} | x_{i+1:\hat{x};t}, u_t, x_{t-1}, y_t), \quad (8)$$

$$\times \prod_{j=1}^{\hat{y}} f(y_{j;t} | y_{j+1:\hat{y};t}, u_t, x_{t-1}),$$

$$\times \prod_{i=1}^{\hat{x}} f(x_{i;t-1} | x_{i+1:\hat{x};t-1}, d^{t-1}) dx_{t-1}.$$

A presence of vector x_{t-1} in all pdfs in (8) means involvement of all entries of the respective vector in integration. Such a notation is used only for shortening of the equation.

3.2 Algorithm of Factorized Kalman Filtering

The entry-wise updating assumed in (8) and preserving of the factorized form (4) of the posterior state estimate $f(x_t|d^t) = \prod_{i=1}^{\hat{x}} f(x_{i;t}|x_{i+1;\hat{x};t}, d^t)$ can be reached via LDL' decomposition [5] of the precision (i.e. inverse covariance) matrices. Such the decomposition supposes L to be a lower triangular matrix with unit diagonal and D to be a diagonal one. This kind of matrix decomposition is used throughout the paper.

The key moments of the entry-wise organized Kalman filter (8), applied to Gaussian models (2-3) are as follows.

3.2.1 Factorization of State Evolution Model

Gaussian state evolution model (2) is factorized in the following way. The process noise covariance matrix R_w is inverted into a precision matrix and decomposed so that

$$R_w^{-1} = L_w D_w L_w'. \quad (9)$$

The factorized form of the model (2) becomes now

$$[L_w' x_t - z_t - \Xi x_{t-1}]' D_w [L_w' x_t - z_t - \Xi x_{t-1}], \quad (10)$$

where

$$z_t = L_w' B u_t, \quad (11)$$

$$\Xi = L_w' A. \quad (12)$$

Gaussian distribution of an individual state entry has the following form.

$$\mathcal{N}_{x_{i;t}}(z_i - \sum_{k=i+1}^{\hat{x}} L_{w;ki} x_{k;t} + \sum_{l=1}^{\hat{x}} \Xi_{il} x_{l;t-1}, \frac{1}{D_{w;ii}}). \quad (13)$$

where $L_{w;ki}$, Ξ_{il} and $D_{w;ii}$ are the elements of matrices L_w , Ξ and D_w respectively.

3.2.2 Factorization of Observation Model

Factorization of Gaussian observation model (3) is fulfilled similarly. The measurement noise covariance matrix R_v is inverted into the precision matrix and decomposed so that

$$R_v^{-1} = L_v D_v L_v', \quad (14)$$

The factorized form of the model (3) is as follows.

$$[L_v' y_t - \rho_t - \mathcal{A} x_{t-1}]' D_v [L_v' y_t - \rho_t - \mathcal{A} x_{t-1}], \quad (15)$$

where

$$\rho_t = L_v' H u_t, \quad (16)$$

$$\mathcal{A} = L_v' C. \quad (17)$$

Gaussian distribution of an individual output entry takes the form

$$\mathcal{N}_{y_{j;t}}(\rho_j - \sum_{k=j+1}^{\hat{y}} L_{v;kj} y_{k;t} + \sum_{l=1}^{\hat{x}} \mathcal{A}_{jl} x_{l;t-1}, \frac{1}{D_{v;jj}}). \quad (18)$$

where $L_{v;kj}$, \mathcal{A}_{jl} and $D_{v;jj}$ are the elements of matrices L_v , \mathcal{A} and D_v respectively.

3.2.3 Initial State Factorization

The initial state Gaussian distribution $f(x_{t-1}|d^{t-1}) \sim \mathcal{N}(\mu_{t-1}, P_{t-1})$, where μ_{t-1} is a known vector of mean values and P_{t-1} is a known covariance matrix, is also factorized in the similar way. Decomposition of the precision matrix is done so that

$$P_{t-1}^{-1} = L_{p|t-1} D_{p|t-1} L_{p|t-1}'. \quad (19)$$

The factorized form of the initial state looks like

$$[L_{p|t-1}' x_{t-1} - \mu_{t-1}^f]' D_{p|t-1} [L_{p|t-1}' x_{t-1} - \mu_{t-1}^f], \quad (20)$$

with

$$\mu_{t-1}^f = L_{p|t-1}' \mu_{t-1}. \quad (21)$$

Via (19) the initial state entries have the following factorized form of Gaussian distribution.

$$\mathcal{N}_{x_{i;t-1}}(\mu_{i;t-1}^f - \sum_{k=i+1}^{\hat{x}} L_{p|t-1;ki} x_{k;t-1}, \frac{1}{D_{p|t-1;ii}}). \quad (22)$$

where $L_{p|t-1;ki}$ and $D_{p|t-1;ii}$ are the elements of matrices $L_{p|t-1}$ and $D_{p|t-1}$ respectively.

3.2.4 Factorized Simultaneous Data & Time Updating

The simultaneous data and time updating in the factorized form is proposed as follows. After integration in (8) and completion of squares the following quadratic form for the factorized state is obtained.

$$[L_w' x_t - \mu_t^*]' \tilde{D}_t [L_w' x_t - \mu_t^*], \quad (23)$$

where

$$\mu_t^* = z_t + \tilde{D}^{-1}, \quad (24)$$

$$\times (D_w \Xi \tilde{A}_t^{-1} (\mathcal{A}' D_v (L_v' y_t - \rho_t) + L_{p|t-1} D_{p|t-1} \mu_{t-1}^f)),$$

$$\Omega_t = \text{diag} [D_w, D_v, D_{p|t-1}], \quad (25)$$

$$\tilde{A}_t = [\Xi; \mathcal{A}; L_{p|t-1}]' \Omega_t [\Xi; \mathcal{A}; L_{p|t-1}'], \quad (26)$$

$$\tilde{D}_t = D_w - D_w \Xi \tilde{A}_t^{-1} \Xi' D_w. \quad (27)$$

The matrix \tilde{D}_t , obtained in (27) is decomposed so that

$$\tilde{D}_t = L_{u|t} D_{u|t} L_{u|t}'. \quad (28)$$

After decomposition and factorization of the quadratic form (23), the updated factorized form is obtained.

$$\left[\begin{array}{c} L_{u|t}' L_w' x_t - L_{u|t}' \mu_t^* \\ L_{p|t}' \end{array} \right]' \begin{array}{c} D_{u|t} \\ D_{p|t} \end{array} \left[\begin{array}{c} L_{p|t}' x_t - \mu_t^f \\ \end{array} \right]. \quad (29)$$

It means, that the updating of the decomposed matrices and the factorized mean value is as follows.

$$D_{p|t} = D_{u|t}, \quad (30)$$

$$L_{p|t} = L_{u|t} L_w, \quad (31)$$

$$\mu_t^f = L_{u|t}' \mu_t^*, \quad (32)$$

which allows to preserve the form (20) of the prior state

$$\left[L_{p|t}' x_t - \mu_t^f \right]' D_{p|t} \left[L_{p|t}' x_t - \mu_t^f \right]. \quad (33)$$

Hence, Gaussian distribution of the i -th state entry also keeps its factorized form (22), i.e. finally one obtains

$$\mathcal{N}_{x_{i:t}}(\mu_{i:t}^f - \sum_{k=i+1}^{\hat{x}} L_{p|t;ki} x_{k;t}, \frac{1}{D_{p|t;ii}}). \quad (34)$$

The obtained results are proved by direct calculation of the integral (8).

Remark 1 *Exploitation of the joint pdf (1) as a system model and, therefore, presence of the state x_{t-1} in both state evolution and observation models (2-3) enables a full factorization of the observation model (3). It means, that in practice the proposed algorithm is not restricted by a single-output model, as its previous versions [7].*

Remark 2 *The proposed algorithm can be sensitive to preserving of positive-definiteness of matrix \tilde{D}_t , used in calculation of final variances of the state entries. For more stability a QR factorization, where Q is an orthogonal matrix and R is an upper triangular one, can be used. However, the present paper is focused on the proposed LDL' factorization due to a lower computational complexity.*

4 Verification and Examples

Correct performance of the proposed entry-wise updating is verified by return of the obtained results (33-34) into non-factorized form and their comparison with solution of the non-factorized integral (7), which gives the following mean value of the state x_t

$$\mu_t = B u_t + \tilde{K}_t^{-1}, \quad (35)$$

$$\times (R_w^{-1} A \tilde{S}_t^{-1} (C' R_v^{-1} (y_t - H u_t) + P_{t-1}^{-1} \mu_{t-1})), \text{ where}$$

$$\tilde{K}_t = R_w^{-1} - R_w^{-1} A \tilde{S}_t^{-1} A' R_w^{-1}, \quad (36)$$

$$\tilde{S}_t = A' R_w^{-1} A + C' R_v^{-1} C + P_{t-1}^{-1}, \quad (37)$$

and the following covariance matrix

$$P_t = (R_w^{-1} - R_w^{-1} A \tilde{S}_t^{-1} A' R_w^{-1})^{-1}. \quad (38)$$

It should be noted, that the non-factorized filtering (7) operates primarily with the precision matrices. Such a number of inversions in (35-38) are shown only for transparency of the results.

The transformation of the results (33) into non-factorized form $\mathcal{N}(\mu_t, P_t)$ is fulfilled as follows.

$$\mu_t = (L_{p|t}')^{-1} \mu_t^f, \quad (39)$$

$$P_t = (L_{p|t} D_{p|t} L_{p|t}')^{-1}, \quad (40)$$

Moreover, the correct performance can be double verified by the LDL' factorization of the state estimate (35-38) as

$$P_t^{-1} = LDL', \quad (41)$$

$$\mu_t^f = L' \mu_t. \quad (42)$$

4.1 Experiments

A simple example of realizations of both versions of the algorithm demonstrates the verification of the proposed factorization. A number of entries of the state \hat{x} and of the output \hat{y} respectively

is equal to 2, the system input $u_t = 0.5$. The following simulated data are used in the models (2-3).

$$A = \begin{bmatrix} 0.1 & -0.9 \\ 0.9 & 0.01 \end{bmatrix}, B = \begin{bmatrix} -0.4 \\ -0.4 \end{bmatrix}, \quad (43)$$

$$C = \begin{bmatrix} -0.1 & 1 \\ 0.1 & -0.5 \end{bmatrix}, H = \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \quad (44)$$

The noise covariances R_w and R_v are computed as a mean of squares of differences between the state (or output respectively) value and its conditional mean. The mean is substituted by the samples of a periodic course of the state (or output), which is constructed as a spline approximation of several last periodic courses. The resulted covariance matrices are as follows.

$$R_w = \begin{bmatrix} 0.3974 & -0.1060 \\ -0.1060 & 0.4011 \end{bmatrix}, \quad (45)$$

$$R_v = \begin{bmatrix} 0.4844 & 0.3598 \\ 0.3598 & 0.9887 \end{bmatrix}. \quad (46)$$

The chosen initial state mean vector and covariance matrix are

$$\mu_t = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, P_t = \begin{bmatrix} 2.029 & -2.784 \\ -2.784 & 5.432 \end{bmatrix}. \quad (47)$$

The data have been used in *MATLAB*[®] [12] implementation both of the filtering versions (7) and (8), or more precisely, proposed in Subsection 3.2.4 and in (35-38) respectively. The results have been transformed to the same form, i.e. the non-factorized state estimate has been factorized, and vice versa. Table 1 shows an example of the identical factorized results of both implementations for the first three steps. To save the place and due to the considered dimension of the state, only the element $L_{p|t;21}$ is shown in the Table 1 (i.e. $L_{p|t;11} = L_{p|t;22} = 1, L_{p|t;12} = 0$).

Table 1: Factorized state estimates

t	$L_{p t;21}$	$D_{p t}$		μ_t^f
1	0.0238	1.7222	0	-2.4860
		0	1.1401	-1.3068
2	0.1114	1.8573	0	1.3908
		0	1.1543	-2.3459
3	0.1016	1.8496	0	1.5480
		0	1.2089	1.1411

The identical results have been also obtained for the rest of the filtering steps $t = 200$. Table 2

demonstrates the non-factorized results of both implementations for the first three steps. The rest of the steps also give the results, identical for both implementations. The obtained identical results

Table 2: Non-factorized state estimates

t	P_t		μ_t
1	0.5811	-0.0209	-2.4549
	-0.0209	0.8772	-1.3068
2	0.5492	-0.0965	1.6521
	-0.0965	0.8663	-2.3459
3	0.5492	-0.0841	1.4320
	-0.0841	0.8272	1.1411

prove correctness of the proposed factorized algorithm. Fig. 1 shows the estimation of the state entries. Good correspondence between simulated and estimated values verifies the adequate performance of the proposed version of Kalman filter.

The proposed algorithm is expected to have an increased level of the computational complexity. Denoting a number of computational operations during implementation by n^{num} , one can compare a level of complexity of different implementations of Kalman filter. For "classic" Kalman filter, implemented according to [1], $n^{num} = 5$; for Kalman filter with simultaneous data and time updating, realized according to (35-38) $n^{num} = 4$; and, finally, a number of operations $n^{num} = 7$ for the factorized algorithm, proposed in Subsection 3.2.4. It can be seen, that the number of computational operations is increased rather insignificantly. Moreover, the proposed algorithm does not contain numerically dangerous operations.

5 Conclusion

The paper presents a modified algorithm of the entry-wise Kalman filtering. The modification proposes the simultaneous data and time updating of the posterior state entry estimates. To summarize the paper, one can note, that the obtained results look optimistically. Advantages, provided by the entry-wise updating, are planned to be used for construction of technique of the joint filtering with data of a mixed (both continuous and discrete-valued) type. The future work includes involvement of discrete-valued states and experiments with the data from a traffic-control area, which is the main target application.

Acknowledgements: The research was sup-

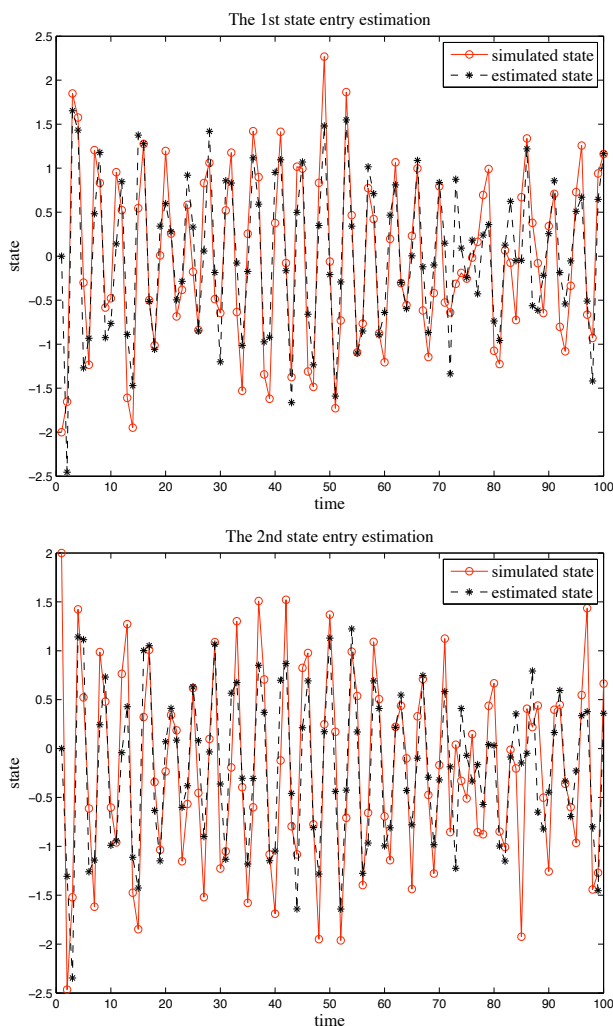


Figure 1: Estimation of state entries

ported by GA ČR grant No. 201/06/P434.

References:

[1] Greg Welch and Gary Bishop, “An Introduction to the Kalman Filter”, Tech. Rep. 95-041, UNC-CH Computer Science, 1995.

[2] G. Dimitriu, “Implementation issues related to data assimilation using Kalman filtering”, in *Proceedings for 3th World Conference on Computational Statistics & Data Analysis*, Limassol, Cyprus, October 28-31 2005, International Association for Statistical Computing.

[3] I. Patras and M. Pantic, “Particle filtering with factorized likelihoods for tracking facial features”, in *Proceedings of the 6th IEEE International Conference on Automatic Face*

and Gesture Recognition (FGR04), Seoul, Korea, May 17-19 2004.

[4] J. Saltyte-Benth and K. Ducinskas, “Linear discriminant analysis of multivariate spatial-temporal regressions”, Tech. Rep. 2, Department of Mathematics, University of Oslo, July 2003, ISBN 82-553-1395-8.

[5] M. Kárný, J. Böhm, T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma, and L. Tesar, *Optimized Bayesian Dynamic Advising: Theory and Algorithms*, Springer, London, 2005.

[6] E. Suzdaleva, “On entry-wise organized filtering”, in *Proceedings of 15th International Conference on Process Control'05*, High Tatras, Slovakia, June 7-10 2005, pp. 1-6, Department of Information Engineering and Process Control, FCFT STU, ISBN 80-227-2235-9.

[7] E. Suzdaleva, “Towards state estimation in the factorized form”, in *Proceedings of the 2nd International Conference on From Scientific Computing to Computational Engineering*, D. Tsahalis, Ed., Athens, July 2006, pp. 1-6, University of Patras.

[8] E. Suzdaleva, “Factorized Kalman filtering”, in *Proceedings of the 7th International PhD Workshop on Interplay of societal and technical decision-making*, J. Přikryl, J. Andryšek, and V. Šmidl, Eds., Hrubá Skála, Czech Republic, September 25-30 2006, pp. 226-233, ÚTIA AV ČR.

[9] E. Suzdaleva, “Early experiments with traffic system model and factorized Kalman filtering”, in *Proceedings of the IEEE Fourth International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications IDAACS'2007*, Dortmund, Germany, September 6-8 2007, IEEE, ISBN 1-4244-1348-6.

[10] V. Peterka, “Bayesian system identification”, in *Trends and Progress in System Identification*, P. Eykhoff, Ed., pp. 239-304. Pergamon Press, Oxford, 1981.

[11] E. Suzdaleva, “Factorized filtering with simultaneous data and time updating”, Tech. Rep. 2221, ÚTIA AV ČR, Praha, May 2008.

[12] The Matlab Inc., *Matlab documentation*, vol. MA 01760-2098, Natick, 2000.