Akademie věd České republiky
Ústav teorie informace a automatizace, v.v.i.

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

RESEARCH REPORT

E. Suzdaleva

Entry-wise Kalman Filtering and Prior Knowledge Processing

2207 November 2007

GAČR 201/06/P434

ÚTIA AVČR, P.O.Box 18, 182 08 Prague,
Czech Republic
Fax: (+420)286890378, http://www.utia.cz, E-mail: utia@utia.cas.cz

Contents

Chapter 1

Introduction

1.1 Addressed problem

The paper deals with the Kalman filtering [1], organized in an entry-wise manner. This version of Kalman filter enables to estimate the individual entries of the system state vector, which are also called *factors*. Due to this reason such the Kalman filtering is also mentioned as the factorized one.

The main application area of the research is the (urban) traffic control, where the unobservable system state is treated as a length of the car queue at the intersection. The factorized, or entry-wise, state estimation will contribute to the solution of estimation with the mixed-type traffic data.

1.2 State of the art and contribution of the paper

Speaking about the state of the art of the factorized filtering, it is necessary to note several research works in this field. The paper [2] proposed the recursive algorithm of the entry-wise organized filtering under Bayesian methodology [3], restricted by the reduced form of the state-space model, concluded in the triangular transition matrix. The papers [4, 5] 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 [6] offered the version of factorized Kalman filtering with Gaussian models, which was based on the $L'DL$ decomposition of the covariance matrices. The paper [7] expanded the line with $L'DL$-factorized covariance matrices and demonstrated the application of the solution to the traffic system state-space model.

The present work assumes the full factorization of the observation model and proposes the algorithm for Gaussian multiple-output state-space model. It takes into account the drawbacks of solutions, proposed in [5] and [6]. The paper presents the example of calculations of the state entry estimation for two-dimensional system state as well as illustrative experiments.

The separate part of the work is devoted to selection of initial conditions for the factorized filtering. The prior knowledge processing for Kalman filter was previously proposed in [8], which described the methodology of selection of the initial state distribution. The work [9] continued this research, expanded the methodology for multiple-output system, and described processing of one of the common type of the prior knowledge, namely, simulated data from the traffic control area. The present work proposes the methodology of specification of prior distributions for the individual state factors.

1.3 Paper outline

The outline of the paper includes the following parts. Chapter 2 provides preliminaries, describing the model, used in the work, and the basic facts about Bayesian prediction and filtering. Chapter 3 deals with the factorized version of the filtering. It describes Bayesian prediction and filtering as well as Kalman filtering in the factorized form. The chapter provides the algorithm and demonstrates the example of calculation of the state estimates for the two-dimensional state. The illustrative experiments are presented in Section 3.4. Chapter 4 is devoted to the factorized version of the prior knowledge processing for the initial state. The conclusion in Section 4.3 summarizes the paper and describes the future work.

Chapter 2

Preliminaries

The following specific notations are used throughout the paper. The notations, which are not shown here, are explained immediately in the text.

$\mathring{x}$ *denotes the number of members in the countable set* x^* or the *number of entries in the vector* x.

x_t^i *denotes an* i*-th entry of the vector* x *at time* t. *The subscript symbol is a time index (except the matrix element indices).*

$x^{i:k}$ *denotes the sequence* $x^i, \ldots, x^k$.

2.1 State-space model

The probabilistic description of the system, which state is to be estimated, is provided with the help of the following probability density functions (pdf).
The *model of observation*

$$f\left(y_t \middle| u_t, d^{1:t-1}, x_t\right), \ t \in t^*, \tag{2.1}$$

relates the system output y_t to the system input u_t, system state x_t and past data $d^{1:t-1}$ at discrete time moments $t \in t^* \equiv \{0, \ldots, \mathring{t}\}$, where $\mathring{t}$ is the cardinality of the set t^* and $\equiv$ means equivalence.
The *model of state evolution*

$$f\left(x_t \middle| u_t, d^{1:t-1}, x_{t-1}\right), \ t \in t^*, \tag{2.2}$$

describes the evolution of the state x_t.
The *model of control strategy*

$$f\left(u_t \middle| d^{1:t-1}\right), \ t \in t^*, \tag{2.3}$$

describes, generally randomized, generating of inputs u_t, which ignore the unobservable system states.

2.2 Prediction and filtering

Bayesian predictor of the output is given by the formula

$$f\left(y_t \middle| u_t, d^{1:t-1}\right) = \int f\left(y_t \middle| u_t, d^{1:t-1}, x_t\right) f\left(x_t \middle| u_t, d^{1:t-1}\right) \, dx_t. \tag{2.4}$$

Bayesian filtering, estimating the state x_t, includes the following coupled formulas.

<u>Data updating</u>

$$\begin{aligned} f\left(x_t \middle| d^{1:t}\right) &= \frac{f\left(y_t \middle| u_t, d^{1:t-1}, x_t\right) f\left(x_t \middle| u_t, d^{1:t-1}\right)}{f\left(y_t \middle| u_t, d^{1:t-1}\right)}, \\ &\propto f\left(y_t \middle| u_t, d^{1:t-1}, x_t\right) f\left(x_t \middle| u_t, d^{1:t-1}\right), \end{aligned} \tag{2.5}$$

($\propto$ means proportionality), that incorporates the experience contained in the data $d^{1:t}$.

<u>Time updating</u>

$$f\left(x_{t+1} \middle| u_{t+1}, d^{1:t}\right) = \int f\left(x_{t+1} \middle| u_{t+1}, d^{1:t}, x_t\right) f\left(x_t \middle| d^{1:t}\right) \, dx_t, \tag{2.6}$$

which fulfills the state prediction.

The filtering does not depend on the control strategy $\{f(u_t|d^{1:t-1})\}_{t\in t^*}$ but on the generated inputs only. The application to Gaussian state-space model with Gaussian prior on x_0 and Gaussian observations provides Kalman filter. The prior pdf $f(x_0)$, that expresses the subjective prior knowledge on the initial state x_0, starts the recursions.

Chapter 3

Factorized filtering

3.1 Factorized Bayesian prediction and filtering

Bayesian predictor $f\left(y_t \middle| u_t, d^{1:t-1}\right)$ is the last pdf of the sequence of the predictors, indexed by $j = 1, \ldots, \mathring{x}$.

$$f\left(y_t^j \middle| u_t, d^{1:t-1}, y_t^{j+1:\mathring{y}}, x_t^{i+1:\mathring{x}}\right) = \int f\left(y_t^j \middle| u_t, d^{1:t-1}, y_t^{j+1:\mathring{y}}, x_t^{i:\mathring{x}}\right) f\left(x_t^i \middle| u_t, d^{1:t-1}, x_t^{i+1:\mathring{x}}\right) dx_t^i. \tag{3.1}$$

The pdfs $f\left(x_t^i \middle| u_t, d^{1:t-1}, x_t^{i+1:\mathring{x}}\right)$, $i = 1, \ldots, \mathring{x}$, determining the state estimate through the chain rule $f\left(x_t \middle| u_t, d^{1:t-1}\right) = \prod_{i=1}^{\mathring{x}} f\left(x_t^i \middle| u_t, d^{1:t-1}, x_t^{i+1:\mathring{x}}\right)$ evolve according the following coupled formulas.

Data updating

$$\begin{aligned} f\left(x_t^i \middle| d^{1:t}, x_t^{i+1:\mathring{x}}\right) &= \frac{f\left(y_t^j \middle| u_t, d^{1:t-1}, y_t^{j+1:\mathring{y}}, x_t^{i:\mathring{x}}\right) f\left(x_t^i \middle| u_t, d^{1:t-1}, x_t^{i+1:\mathring{x}}\right)}{f\left(y_t^j \middle| u_t, d^{1:t-1}, y_t^{j+1:\mathring{y}}, x_t^{i+1:\mathring{x}}\right)} \\ &\propto f\left(y_t^j \middle| u_t, d^{1:t-1}, y_t^{j+1:\mathring{y}}, x_t^{i:\mathring{x}}\right) f\left(x_t^i \middle| u_t, d^{1:t-1}, x_t^{i+1:\mathring{x}}\right), \end{aligned} \tag{3.2}$$

which incorporates the experience contained in the data $d^{1:t}$, consisting of the output y_t and the input u_t.

Time updating

The pdf $f\left(x_{t+1} \middle| u_{t+1}, d^{1:t}\right)$ is the last member of the sequence of the partially conditioned state estimates, indexed by $l = 1, \ldots, \mathring{x}$

$$\begin{aligned} & f\left(x_{t+1} \middle| u_{t+1}, d^{1:t}, x_t^{l+1:\mathring{x}}\right) \\ = & \int f\left(x_{t+1} \middle| u_{t+1}, d^{1:t}, x_t^{l:\mathring{x}}\right) f\left(x_t^l \middle| d^{1:t}, x_t^{l+1:\mathring{x}}\right) dx_t^l. \end{aligned} \tag{3.3}$$

The factorized version of the time-updated state estimate is obtained by straightforward application of the chain rule for pdfs. The proofs are available in [4].

3.2 Factorized Kalman filtering

The linear Gaussian models (2.1-2.2), used for demonstrating of the proposed methodology, are given by

$$\textit{observation model} \qquad y_t = Cx_t + Du_t + v_t, \tag{3.4}$$

$$\textit{state evolution model} \qquad x_{t+1} = Ax_t + Bu_{t+1} + \omega_{t+1}, \tag{3.5}$$

where x_t, y_t and u_t are column vectors with dimensions $\mathring{x}$, $\mathring{y}$ and $\mathring{u}$ respectively; v_t is a measurement (Gaussian) noise with zero mean and covariance $(R_v R_v')^{-1}$, where R_v is a lower triangular matrix; ω_t is a process (Gaussian) noise with zero mean and covariance $(R_\omega R_\omega')^{-1}$, where R_ω is a lower triangular matrix; A, B, C and D are known matrices of appropriate dimensions.

The algorithm of the factorized filtering with linear Gaussian state-space model includes the following steps. The initial state estimate of the i-th factor is to be set in the following form of normal distribution

$$f\left(x_t^i \middle| u_t, d^{1:t-1}, x_t^{i+1:\mathring{x}}\right) = \mathcal{N}_{x_t^i}\left(\hat{\mu}_{t|t-1}^i + \sum_{k=i+1}^{\mathring{x}} g_{t|t-1}^{ik} x_t^k, p_{t|t-1}^i\right), \tag{3.6}$$

where $\hat{\mu}^i_{t|t-1}$ is the term, depending on the data $u_t, d^{1:t-1}$ only, $g^{ik}_{t|t-1}$ are coefficients, which are data and state independent, similarly as the variance $p^i_{t|t-1} > 0$.

Factorized output prediction

To simplify presentation, dimensions of the state and the output are chosen to be identical, and matrix $D = 0$. Fully conditioned Gaussian observation model (2.1), cf (3.4), takes the following form.

$$f\left(y^i_t \middle| u_t, d^{1:t-1}, y^{i+1:\mathring{y}}_t, x^{i:\mathring{x}}_t\right) = \mathcal{N}_{y^i_t}\left(\underbrace{\sum_{k=i+1}^{\mathring{y}} h^{ki} y^k_t}_{\rho^i_t} + \sum_{k=i}^{\mathring{x}} \underbrace{(C_{ik} - \sum_{l=i+1}^{\mathring{x}} C_{lk} h^{li})}_{\tilde{c}^i} x^k_t, \underbrace{((R_{vii})^2)^{-1}}_{r^i_t}\right), \tag{3.7}$$

where $h^{ki} = \frac{R_{vki}}{R_{vii}}$, $k = i+1, \dots, \mathring{x}$, and the state independent offsets ρ^i_t, coefficients $\tilde{c}^i$ evolve according to the following recursions, for $i = 1, \dots, \mathring{x}$

$$\begin{aligned} \rho^{i+1}_t &= \rho^i_t + \tilde{c}^i \hat{\mu}^i_{t|t-1}, \\ \tilde{c}^{i+1} &= \tilde{c}^i g^{ik}_{t|t-1}, \text{ for } k > i, \\ r^{i+1}_t &= r^i_t + p^i_{t|t-1} (\tilde{c}^i)^2. \end{aligned} \tag{3.8}$$

The recursions are obtained by integration of Gaussian pdf in (3.1) and square completion.

Factorized data updating

The state estimate of the i-th factor preserves its functional form (3.6) in the data updating. The data-updated state estimate evolves in the following way.

$$f\left(x^i_t \middle| d^{1:t}, x^{i+1:\mathring{x}}_t\right) = \mathcal{N}_{x^i_t}\left(\hat{\mu}^i_{t|t} + \sum_{k=i+1}^{\mathring{x}} g^{ik}_{t|t} x^k_t, p^i_{t|t}\right), \tag{3.9}$$

with

$$\begin{aligned} K^i_{t|t} &\equiv \frac{\tilde{c}^i_t p^i_{t|t-1}}{r^{i+1}_t}, \\ \hat{\mu}^i_{t|t} &= \hat{\mu}^i_{t|t-1} + K^i_{t|t}(y^i_t - \rho^i_t - \tilde{c}^i \hat{\mu}^i_{t|t-1}), \\ g^{ik}_{t|t} &= g^{ik}_{t|t-1} - K^i_{t|t}(g^{ik}_{t|t-1}\tilde{c}^i + \tilde{c}^k) \text{ for } k > i, \\ p^i_{t|t} &= \frac{r^i_t}{r^{i+1}_t} p^i_{t|t-1}. \end{aligned} \tag{3.10}$$

The recursions are obtained by square completion in (3.2) and obtaining of the necessary Gaussian quadratic form.

Factorized time updating

The fully conditioned state evolution model (2.2), cf (3.5) can be given by the form

$$\begin{aligned} f\left(x_{t+1} \middle| u_{t+1}, d^{1:t}, x_t\right) &= \prod_{i=1}^{\mathring{x}} f\left(x^i_{t+1} \middle| x^{i+1:\mathring{x}}_{t+1}, u_{t+1}, d^{1:t}, x_t\right), \\ &= \prod_{i=1}^{\mathring{x}} \mathcal{N}_{x^i_{t+1}}\left(\zeta^i_{t+1} + \sum_{k=i+1}^{\mathring{x}} \alpha^{ik} x^k_{t+1} + \sum_{k=1}^{\mathring{x}} \beta^{ik} x^k_t, R^i_{t+1}\right), \end{aligned} \tag{3.11}$$

where for all $i \in \{1, \dots, \mathring{x}\}$ the offset ζ^i_{t+1}, coefficients α^{ik}, $k = i+1, \dots, \mathring{x}$, β^{ik}, $k = 1, \dots, \mathring{x}$ and variances R^i_{t+1} are assumed to be known functions of $u_{t+1}, d^{1:t}$, namely

$$\begin{aligned} \alpha^{ik} &= \frac{R_{\omega ki}}{R_{\omega ii}}, \\ \beta^{ik} &= A_{ik} + \sum_{l=i+1}^{\mathring{x}} \alpha^{ik} A_{li}, \\ \zeta^i_{t+1} &= \sum_{l=1}^{\mathring{u}} u^l_t (B_{il} + \sum_{k=i+1}^{\mathring{x}} B_{kl} \alpha^{ik}), \\ R^i_{t+1} &= ((R_{\omega ii})^2)^{-1}. \end{aligned} \tag{3.12}$$

Time updating preserves the functional form of the state estimate and gives

$$f\left(x_{t+1}^{i}\middle| x_{t+1}^{i+1:\mathring{x}}, u_{t+1}, d^{1:t}, x_{t}^{j+1:\mathring{x}}\right) = \mathcal{N}_{x_{t+1}^{i}}\left(\hat{\mu}_{t+1|t}^{i} + \sum_{k=i+1}^{\mathring{x}} g_{t+1|t}^{ik} x_{t+1}^{k} + \sum_{k=j+1}^{\mathring{x}} \tilde{\beta}^{ik} x_{t}^{k}, p_{t+1|t}^{i}\right), \quad (3.13)$$

where for $i \in \{1, \ldots, \mathring{x}\}$ offsets $\hat{\mu}_{t+1|t}^{i}$, coefficients $g_{t+1|t}^{ik}$, $k = i+1, \ldots, \mathring{x}$, $\tilde{beta}^{ik}$, $k = j+1, \ldots, \mathring{x}$ and variances $p_{t+1|t}^{i} > 0$ are the state independent. The following recursions over $j = 1, \ldots, \mathring{x}$ hold

$$\begin{aligned}
\hat{\mu}_{t+1|t}^{i} &= \zeta_{t+1}^{i} + \sum_{l=i+1}^{\mathring{x}} U_{li}\zeta_{t+1}^{l} - U_{(\mathring{x}+1)i}\hat{\mu}_{t|t;j}, \quad (3.14)\\
g_{t+1|t}^{ik} &= -\alpha^{ik} - \sum_{l=k-1}^{\mathring{x}} U_{li}\alpha^{lk} + U_{ki},\\
\tilde{\beta}^{ik} &= \beta^{ik} + \sum_{l=i+1}^{\mathring{x}} U_{li}\beta^{lk} - U_{(\mathring{x}+1)i} g_{t|t}^{jk},\ k > j.
\end{aligned}$$

The elements of the upper triangular $(\mathring{x}+1, \mathring{x}+1)$ matrix U with the unit diagonal as well as the positive scalars $p_{t+1|t}^{i}$ are obtained via the $U'DU$ decomposition

$$\begin{aligned}
\omega - \frac{\omega\gamma\gamma'\omega}{\gamma'\omega\gamma} &\equiv U'\mathrm{diag}\left[\frac{1}{p_{t+1|t}^{1}}, \ldots, \frac{1}{p_{t+1|t}^{\mathring{x}}}\right] U \text{ with} \quad (3.15)\\
\gamma' &\equiv [\beta_{1j}, \ldots, \beta_{\mathring{x}j}, 1],\\
\omega &\equiv \mathrm{diag}\left[\frac{1}{R_{t+1}^{1}}, \ldots, \frac{1}{R_{t+1}^{\mathring{x}}}, \frac{1}{p_{t|t}^{j}}\right].
\end{aligned}$$

The results are obtained by integration of Gaussian pdf in (3.3) and square completion. The proof is available in [4].

3.3 Example with single-output two-state system

The linear Gaussian models (3.4-3.5) with dimension $\mathring{x} = 2$ and scalars y_t and u_t are used to illustrate the factorized version of Kalman filter. The measurement (Gaussian) noise v_t in (3.4) has a zero mean and variance r_v; ω_t is a process (Gaussian) noise with zero mean and covariance $(R_\omega R_\omega')^{-1}$, where $R_\omega = [r_{\omega 11} \quad 0;\ r_{\omega 21} \quad r_{\omega 22}]$; A, B, C and D are known matrices of appropriate dimensions.

According to (3.6), the initial state estimate is assumed in the form of the following pdfs.

$$f\left(x_t^1\middle| u_t, d^{1:t-1}, x_t^2\right) f\left(x_t^2\middle| u_t, d^{1:t-1}\right) = \mathcal{N}_{x_t^1}(\hat{\mu}_{t|t-1}^1 + g_{t|t-1}^{12} x_t^2, \hat{p}_{t|t-1}^1)\mathcal{N}_{x_t^2}(\hat{\mu}_{t|t-1}^2, \hat{p}_{t|t-1}^2), \quad (3.16)$$

where

$$\begin{aligned}
\hat{\mu}_{t|t-1}^1 &= \mu^1 + \frac{p_{21}}{p_{11}}\mu^2,\\
\hat{\mu}_{t|t-1}^2 &= \mu^2,\\
g_{t|t-1}^{12} &= -\frac{p_{21}}{p_{11}},\\
\hat{p}_{t|t-1}^1 &= ((p_{11})^2)^{-1},\\
\hat{p}_{t|t-1}^2 &= ((p_{22})^2)^{-1},
\end{aligned}$$

and μ^1 and μ^2 are the initial mean values of unconditional factors x_t^1 and x_t^2 respectively, and $(PP')^{-1}$ is a covariance matrix of the joint distribution $f\left(x_t^1, x_t^2\middle| u_t, d^{1:t-1}\right)$, $P = [p_{11} \quad 0;\ p_{21} \quad p_{22}]$.

Output prediction

According to factorization (3.7), the single-output observation model (2.1), cf. (3.4), with nonzero matrix D, takes the following form.

$$f\left(y_t\middle| u_t, d^{1:t-1}, x_t^1, x_t^2\right) = \mathcal{N}_{y_t}(\underbrace{d_{11}u_t}_{\rho_t^1} + c_{11}x_t^1 + c_{12}x_t^2, \underbrace{r_v^1}_{r^1}). \quad (3.17)$$

The factorized form $f\left(y_t \middle| u_t, d^{1:t-1}, x_t^2\right)$ of Bayesian prediction (3.1) evolves in the following way.

$$\begin{aligned}\rho_t^2 &= \rho_t^1 + c_{11}\hat{\mu}_{t|t-1}^1, \\ \tilde{c}^2 &= c_{12} + c_{11} g_{t|t-1}^{12}, \\ r_v^2 &= r_v^1 + (c_{11})^2 \hat{p}_{t|t-1}^1.\end{aligned} \tag{3.18}$$

Data updating

According to (3.2), the data updating takes the following form.

$$\begin{aligned}f\left(x_t^1 \middle| d^{1:t}, x_t^2\right) &= \frac{f\left(y_t \middle| u_t, d^{1:t-1}, x_t^1, x_t^2\right) f\left(x_t^1 \middle| u_t, d^{1:t-1}, x_t^2\right)}{f\left(y_t \middle| u_t, d^{1:t-1}, x_t^2\right)} \\ &\propto f\left(y_t \middle| u_t, d^{1:t-1}, x_t^1, x_t^2\right) f\left(x_t^1 \middle| u_t, d^{1:t-1}, x_t^2\right)\end{aligned} \tag{3.19}$$

As it is shown in (3.9), data updating preserves the form of state estimate (3.16) and results in the following Gaussian distribution.

$$f\left(x_t^1 \middle| d^{1:t}, x_t^2\right) = \mathcal{N}_{x_t^1}(\hat{\mu}_{t|t}^1 + g_{t|t}^{12} x_t^2, \hat{p}_{t|t}^1), \tag{3.20}$$

where

$$\begin{aligned}K_{t|t}^1 &= \frac{c_{11}\hat{p}_{t|t-1}^1}{r_v^2}, \\ \hat{\mu}_{t|t}^1 &= \hat{\mu}_{t|t-1}^1 + K_{t|t}^1(y_t - \rho_t^1 - c_{11}\hat{\mu}_{t|t-1}^1), \\ g_{t|t}^{12} &= g_{t|t-1}^{12} - K_{t|t}^1(c_{12} + c_{11} g_{t|t-1}^{12}), \\ \hat{p}_{t|t}^1 &= \frac{r_v^1}{r_v^2}\hat{p}_{t|t-1}^1.\end{aligned}$$

The resulted mean and variance are calculated with the help of square completion for x_t^1 in (3.19) and obtaining of the necessary Gaussian form. Additional notations

$$\hat{\mu}_{t|t-1}^1 \frac{r_v^1}{r_v^2} = \hat{\mu}_{t|t-1}^1(1 - K_{t|t}^1 c_{11}), \tag{3.21}$$

$$g_{t|t-1}^{12} \frac{r_v^1}{r_v^2} = g_{t|t-1}^{12}(1 - K_{t|t}^1 c_{11}), \tag{3.22}$$

are used in the obtained Gaussian form to obtain the recursions. The state estimate of the factor x_t^2 is obtained in the similar way.

$$f\left(x_t^2 \middle| d^{1:t}\right) = \mathcal{N}_{x_t^2}(\hat{\mu}_{t|t}^2, \hat{p}_{t|t}^2), \tag{3.23}$$

where

$$\begin{aligned}r_v^3 &= r_v^2 + (\tilde{c}^2)^2 \hat{p}_{t|t-1}^2, \\ K_{t|t}^2 &= \frac{\tilde{c}^2 \hat{p}_{t|t-1}^2}{r_v^3}, \\ \hat{\mu}_{t|t}^2 &= \hat{\mu}_{t|t-1}^2 + K_{t|t}^2(y_t - \rho_t^2 - \tilde{c}^2\hat{\mu}_{t|t-1}^2), \\ \hat{p}_{t|t}^2 &= \frac{r_v^2}{r_v^3}\hat{p}_{t|t-1}^2.\end{aligned}$$

Time updating

Time updating (3.3) comes to the following integration of the product of Gaussian pdfs.

$$\begin{aligned}f\left(x_{t+1}^1, x_{t+1}^2 \middle| u_{t+1}, d^{1:t}, x_t^2\right) &= \int f\left(x_{t+1}^1 \middle| x_{t+1}^2, u_{t+1}, d^{1:t}, x_t^1, x_t^2\right) f\left(x_{t+1}^2 \middle| u_{t+1}, d^{1:t}, x_t^1, x_t^2\right), \\ &\times \; f\left(x_t^1 \middle| d^{1:t}, x_t^2\right) \, dx_t^1.\end{aligned} \tag{3.24}$$

The fully conditioned factorized version of the state evolution model (2.2), cf. (3.5), is as follows.

$$f\left(x_{t+1}^1 \middle| x_{t+1}^2, u_{t+1}, d^{1:t}, x_t^1, x_t^2\right) f\left(x_{t+1}^2 \middle| u_{t+1}, d^{1:t}, x_t^1, x_t^2\right), \tag{3.25}$$

or, as Gaussian distributions, according to (3.11)

$$\mathcal{N}_{x^1_{t+1}}(\zeta^1_{t+1} + \alpha^{12}x^2_{t+1} + \beta^{11}x^1_t + \beta^{12}x^2_t, R^1_{t+1})\mathcal{N}_{x^2_{t+1}}(\zeta^2_{t+1} + \beta^{21}x^1_t + \beta^{22}x^2_t, R^2_{t+1}), \tag{3.26}$$

where

$$\begin{aligned}
\zeta^1_{t+1} &= b_{11}u_t + b_{21}u_t\frac{r_{\omega 21}}{r_{\omega 11}},\\
\alpha^{12} &= -\frac{r_{\omega 21}}{r_{\omega 11}},\\
\beta^{11} &= a_{11} + a_{21}\frac{r_{\omega 21}}{r_{\omega 11}},\\
\beta^{12} &= a_{12} + a_{22}\frac{r_{\omega 21}}{r_{\omega 11}},\\
R^1_{t+1} &= ((r_{\omega 11})^2)^{-1},\\
\zeta^2_{t+1} &= b_{21}u_t,\\
\beta^{21} &= a_{21},\\
\beta^{22} &= a_{22},\\
R^2_{t+1} &= ((r_{\omega 22})^2)^{-1}.
\end{aligned}$$

The time-updated state estimates of x^1_{t+1} and x^2_{t+1} preserve the functional form (3.16) and are calculated in the following way, see (3.14).

$$\mathcal{N}_{x^1_{t+1}}(\hat{\mu}^1_{t+1|t} + g^{12}_{t+1|t}x^2_{t+1} + \tilde{\beta}^{12}x^2_t, \hat{p}^1_{t+1|t}), \tag{3.27}$$

where

$$\begin{aligned}
\hat{\mu}^1_{t+1|t} &= \zeta^1_{t+1} + u_{21}\zeta^2_{t+1} - u_{31}\hat{\mu}^1_{t|t},\\
g^{12}_{t+1|t} &= \alpha^{12} - u_{21},\\
\tilde{\beta}^{12} &= \beta^{12} + u_{21}\beta^{22} - u_{31}g^{12}_{t|t},\\
\hat{p}^1_{t+1|t} &= p_1,
\end{aligned}$$

and respectively for x^2_{t+1}

$$\mathcal{N}_{x^2_{t+1}}(\hat{\mu}^2_{t+1|t} + \tilde{\beta}^{22}x^2_t, \hat{p}^2_{t+1|t}), \tag{3.28}$$

where

$$\begin{aligned}
\hat{\mu}^2_{t+1|t} &= \zeta^2_{t+1} - u_{32}\hat{\mu}^1_{t|t},\\
\tilde{\beta}^{22} &= \beta^{22} - u_{32}g^{12}_{t|t},\\
\hat{p}^2_{t+1|t} &= p_2.
\end{aligned}$$

The elements of the upper triangular (3×3) matrix U with unit diagonal as well as variances $\hat{p}^1_{t+1|t}$ and $\hat{p}^2_{t+1|t}$ are obtained via decomposition, proposed in (3.15).

$$U'\mathrm{diag}\left[p_1^{-1} \;\; p_2^{-1} \;\; p_3^{-1}\right]U = \omega - \frac{\omega\gamma\gamma'\omega}{\gamma'\omega\gamma}, \tag{3.29}$$

where

$$\begin{aligned}
\gamma' &= [\beta^{11} \;\; \beta^{21} \;\; 1],\\
\omega &= \mathrm{diag}[(R^1_{t+1})^{-1} \;\; (R^2_{t+1})^{-1} \;\; (\hat{p}^1_{t|t})^{-1}].
\end{aligned}$$

3.4 Experiments

The experiments of comparison of the estimation results between the proposed factorized filtering and Kalman filter have been made. The experiments used 250 data, simulated by the state-space model (3.4-3.5) with single input and output and two-dimensional state. The known model parameters, used for simulation, are as follows.

$$
\begin{aligned}
A &= \begin{bmatrix} 0.08975 & 0.998 \\ 0.026 & 0.02107 \end{bmatrix}, \\
B &= \begin{bmatrix} -0.0439 \\ -0.12 \end{bmatrix}, \\
C &= \begin{bmatrix} -0.0705 & 0.2372 \end{bmatrix}, \\
D &= \begin{bmatrix} 0.5329 \end{bmatrix}, \\
r_v &= 0.3549, \\
R_\omega &= \begin{bmatrix} 0.0354 & -0.0202 \\ -0.0202 & 0.0451 \end{bmatrix}.
\end{aligned} \tag{3.30}
$$

The results of the state estimation of the factor x_t^1 by the proposed factorized filtering are compared with Kalman filter at Fig. 3.1, which demonstrates the adequate correspondence of the tested state estimates. The

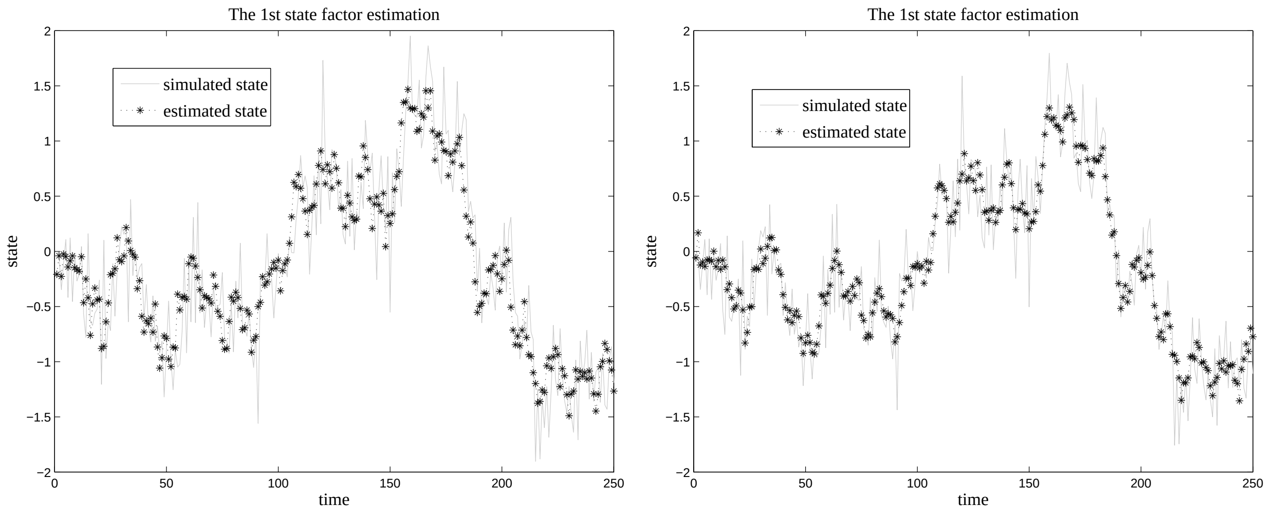

Figure 3.1: State estimation of x_t^1 by factorized filter (left) and by Kalman filter (right)

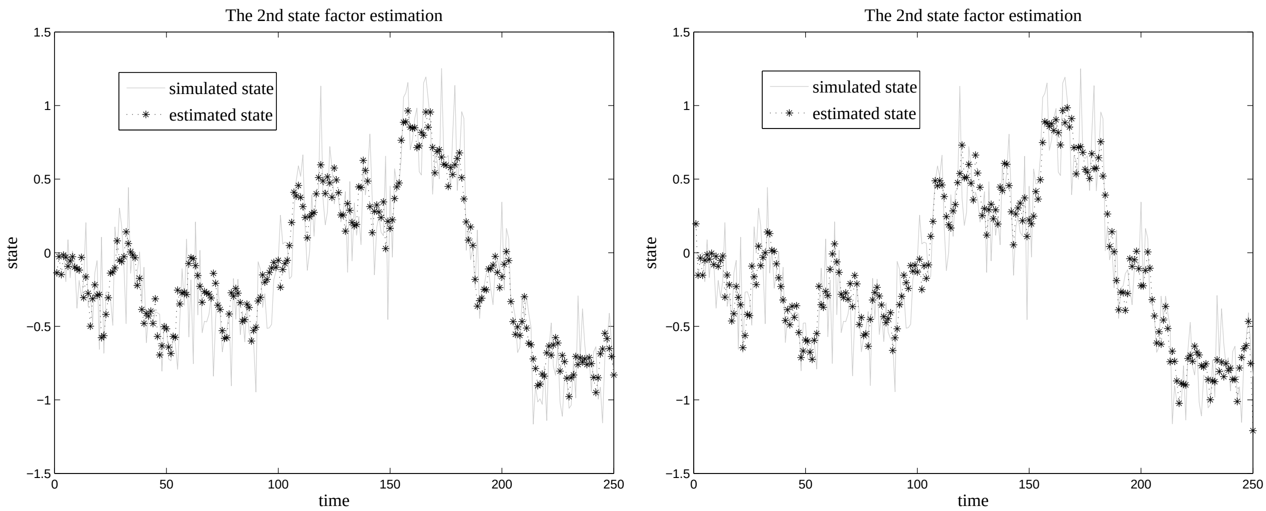

Figure 3.2: State estimation of x_t^2 by factorized filter (left) and by Kalman filter (right)

analogous comparison has been made for the state factor x_t^2. The results of this comparison are demonstrated at Fig. 3.2, which shows the good correspondence between state estimates.

The estimation error, defined as the difference between the estimated and simulated state factors, can be seen in Table 3.1. Table 3.1 shows, that the difference between the simulated and estimated state factor x_t^1 is

Table 3.1: Comparison of estimation error

State factor	Factorized Kalman filter	Kalman filter
x_t^1	5.54867	4.37693
x_t^2	3.85893	4.62868

less in the case of the Kalman filtering. However, the estimation of the state factor x_t^2 demonstrates the less error in the case of the factorized filter.

The experiments with different random realizations of the system input and noises have been made. The results of some of them are shown at Fig. 3.3. The state estimates, made by the factorized filter and Kalman one are in good correspondence. The estimation errors for experiments with different random realizations are shown in Table 3.2. It can be seen, that although the state estimate of x_t^1 is worse in the case of the factorized filtering, nevertheless the factorized estimate of the state factor x_t^2 is better for most experiments. Moreover, the factorized filter has demonstrated more stability in comparison with the "classic" Kalman filtering. It can be seen with the help of differences between the state estimates of factors x_t^1 and x_t^2, made by both filters with different realizations.

Table 3.2: Estimation error with different random realizations

State factor	Factorized Kalman filter	Kalman filter
x_t^1	5.80583	3.86407
x_t^2	4.04565	4.86154
x_t^1	5.60503	4.14361
x_t^2	3.89044	7.20127

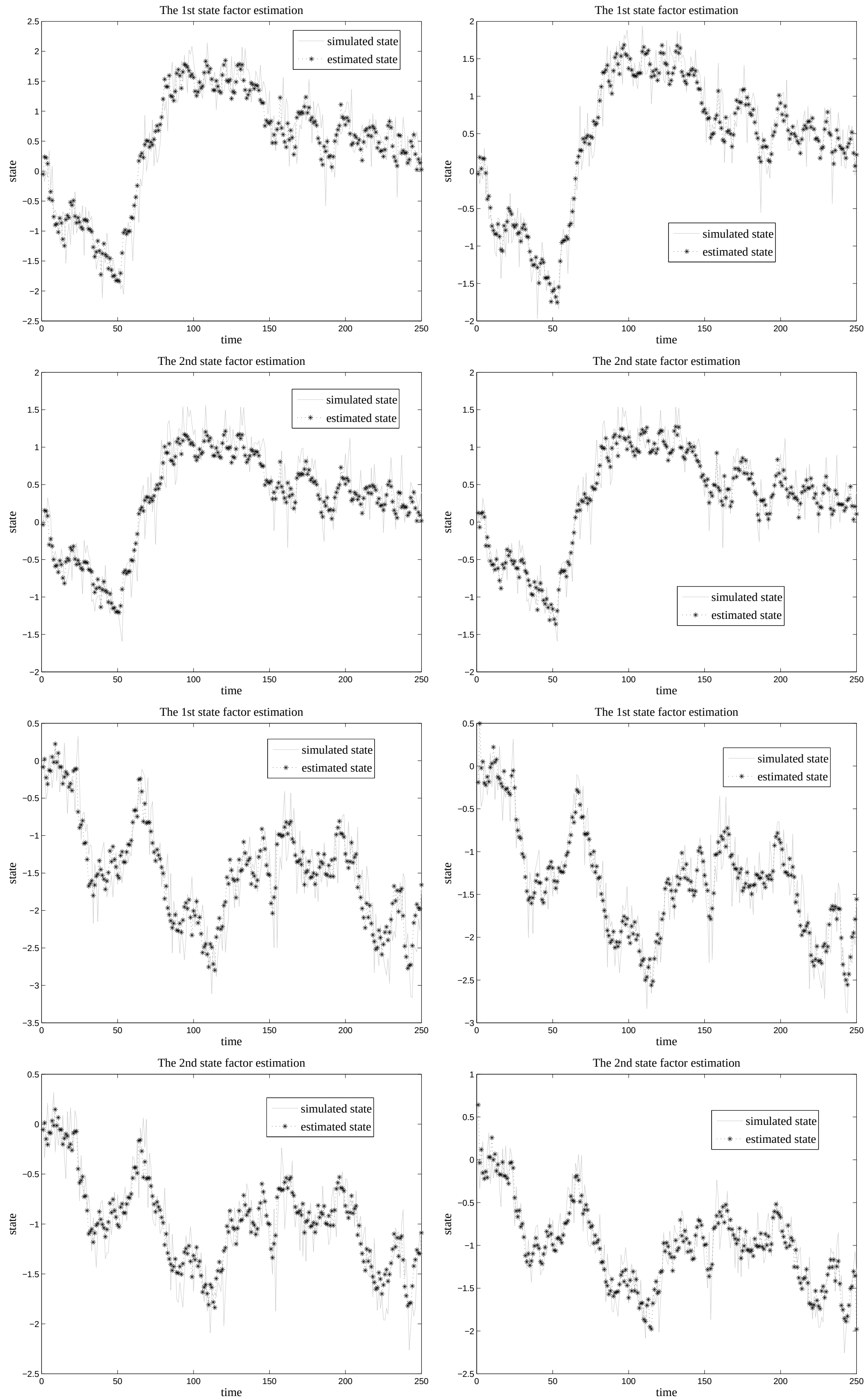

Figure 3.3: Comparison with different random realizations: factorized filter (left) and Kalman filter (right)

Chapter 4

Prior knowledge elicitation for factorized Kalman filter

One of the problems, occurred when using the Kalman filter, is a choice of the prior distribution, i.e. the distribution of the initial state. In practice its distribution is primarily not being chosen systematically. The papers [8, 9] proposed the methodology of specification of the prior distribution for Kalman filter. Before to go to the choice of the factorized form of the initial state, it is worth describing briefly the methodology of selection of the initial state for Kalman filter.

4.1 Prior knowledge processing for initial state of Kalman filter

Let the prior knowledge, provided by experts, be described by the pdfs $f_{\tau^*} \equiv \{f(d^{\,\tau})\}_{\tau\in\tau^*}$, where τ^* is a finite set of time moments. The index τ emphasizes, that the quantities, denoted by it, are related to the prior (not current) knowledge. The observed values are denoted by index t.

The distribution of the initial state x_0 should be chosen, taking into account the provided prior knowledge f_{τ^*}. According to [8, 9], the prior (flat) pdf $f(x_0)$ is proposed to be transformed into

$$f(x_0|f_{\tau^*}) \quad \propto \quad f(x_0)\exp\left[\frac{1}{\mathring{\tau}}\sum_{\tau\in\tau^*}\int f(d^{\,\tau})\ln[\mathcal{Z}(d^{\,\tau}|x_0)]\,dd^{\,\tau}\right], \tag{4.1}$$

where $\mathcal{Z}(d^{\,\tau}|x_0)$ relates to the local state-space model, expressing relation of available system inputs and outputs and the initial state. It means, that the form of $\mathcal{Z}(d^{\,\tau}|x_0)$ depends on the cardinality of the set τ^*, denoted by $\mathring{\tau}$, and is as follows.

$$\begin{aligned}\mathcal{Z}(d^{\,\tau}|x_0) \quad = \quad & \int f\left(y_\tau|\,x_\tau,u_\tau\right) f\left(x_\tau|\,u_\tau,x_{\tau-1}\right) f\left(x_{\tau-1}|\,u_{\tau-1},x_{\tau-2}\right)\ldots \\ \times \quad & f\left(x_{\tau-i}|\,u_{\tau-i},x_0\right)\,dx_\tau\ldots\,dx_{\tau-i}\mathcal{Z}(d^{\,\tau-1}|x_0),\end{aligned} \tag{4.2}$$

where $i=\tau-1$, $\tau=1,\ldots,\mathring{\tau}$. The transformed prior pdf (4.1) takes a simple form, when the model (4.2) belongs to the exponential family. Gaussian model, which Kalman filter deals with, does belong to the exponential family. In this case the model (4.2) can be expressed as

$$\mathcal{Z}(d^{\,\tau}|x_0) = \mathcal{A}(x_0)\exp\left\langle \mathcal{B}(d^{\,\tau}),\mathcal{C}(x_0)\right\rangle, \tag{4.3}$$

where $\mathcal{A}(x_0)$ is a non-negative scalar function; $\mathcal{B}(d^{\,\tau})$ and $\mathcal{C}(x_0)$ are multivariate functions of compatible and finite dimensions; the functional $\langle\cdot,\cdot\rangle$ is linear in the first argument. In this case pdf (4.1) will get the following form

$$f(x_0|f_{\tau^*}) \quad \propto \quad f(x_0)\mathcal{A}(x_0)\exp\left\langle \mathring{\tau}V,\mathcal{C}(x_0)\right\rangle, \tag{4.4}$$

where the array V is

$$V \quad \equiv \quad \frac{1}{\mathring{\tau}}\sum_{t\in t^*}\int f(d^{\,\tau})\mathcal{B}(d^{\,\tau})\,dd^{\,\tau}. \tag{4.5}$$

The conjugate Gaussian prior $f(x_0)$ in (4.4) can be chosen as

$$f(x_0) \quad = \quad \frac{\mathcal{A}(x_0)\exp\left\langle \bar{V},\mathcal{C}(x_0)\right\rangle}{\int \mathcal{A}(x_0)\exp\left\langle \bar{V},\mathcal{C}(x_0)\right\rangle dx_0}. \tag{4.6}$$

With such a conjugate prior the transformed pdf $f(x_0|\tau^*)$ in (4.4) keeps the exponential family form with recursively calculated array

$$V_\tau = V_{\tau-1} + \mathcal{B}(d^{\,\tau}),\ V_0 \equiv \bar{V} + \mathring{\tau}V. \tag{4.7}$$

The mean and covariance matrix of the initial state are obtained via partition of the array (matrix) $V = [V_1\ V_2';\ V_2\ V_3]$, where V_1 and V_3 are square matrices. The partition is obtained naturally by the straightforward calculation of (4.7). The subsequent completion of squares for x_0 [10] gives the following Gaussian distribution for $f(x_0|\tau^*)$

$$\textit{covariance matrix} \qquad \hat{P}_0 = V_3^{-1}, \tag{4.8}$$
$$\textit{mean} \qquad \hat{x}_0 = -V_3^{-1}V_2. \tag{4.9}$$

It means, that in general, the initial state distribution is obtained in the following quadratic form

$$f(x_0|f_{\tau^*}) \propto \mathcal{A}(x_0)\exp\left\{-\frac{1}{2}tr[1\ \ x_0']V[1\ \ x_0']'\right\}, \tag{4.10}$$
$$\propto \mathcal{A}(x_0)\exp\left\{-\frac{1}{2}tr[x_0-\hat{x}_0]'\hat{P}_0[x_0-\hat{x}_0] - \frac{1}{2}tr\lambda\right\}, \tag{4.11}$$

where tr is a trace of matrix, $\mathcal{A}(x_0)$ is only a notion of the scalar function, which is as well as the reminder λ out of interest.

4.2 Prior knowledge processing for the factorized Kalman filter

The pdfs $f\left(x_0^i\middle|x_0^{i+1:\mathring{x}}, f_{\tau^*}\right)$, $i = 1,\dots,\mathring{x}$, determining the distribution of the initial state (4.11) through the chain rule $f(x_0) = \prod_{i=1}^{\mathring{x}} f\left(x_0^i\middle|x_0^{i+1:\mathring{x}}\right)$ can be obtained in the following way.

$$f\left(x_0^i\middle|x_0^{i+1:\mathring{x}}, f_{\tau^*}\right) = \mathcal{N}_{x_0^i}\left(\hat{\mu}_0^i - \sum_{k=i+1}^{\mathring{x}} l_{ki}x_0^k, p_0^i\right), \tag{4.12}$$

where

$$\hat{\mu}_0^i = \hat{x}_0^i + \sum_{k=i+1}^{\mathring{x}} l_{ki}\hat{x}_0^k, \tag{4.13}$$

where $\hat{x}_0^i$, $i = 1,\dots,\mathring{x}$, are the entries of the mean vector of the initial state in (4.11); l_{ki} are the elements of lower triangular $(\mathring{x},\mathring{x})$ matrix L with unit diagonal, obtained via decomposition of covariance matrix in (4.11) $\hat{P}_0 = LDL'$; variances $p_0^i = \frac{1}{d_{ii}}$ are obtained with the help of the mentioned decomposition by means of

$$\hat{P}_0 = L\text{diag}\left[\frac{1}{p_0^1}\cdots\frac{1}{p_0^{\mathring{x}}}\right]L'. \tag{4.14}$$

4.2.1 Example with three-dimensional state

According to (4.11-4.14), for the dimension $\mathring{x} = 3$, the necessary Gaussian form of the transformed prior pdf$f(x_0|f_{\tau^*})$, defining the initial state estimate (3.6), evolves in the following way.

$$[x_0-\hat{x}_0]'\hat{P}_0[x_0-\hat{x}_0] = \frac{(x_0^1-\hat{x}_0^1+(x_0^2-\hat{x}_0^2)l_{21}+(x_0^3-\hat{x}_0^3)l_{31})^2}{p_0^1}, \tag{4.15}$$
$$+ \frac{(x_0^2-\hat{x}_0^2+(x_0^3-\hat{x}_0^3)l_{32})^2}{p_0^2} + \frac{(x_0^3-\hat{x}_0^3)^2}{p_0^3},$$

which helps to specify the prior distribution for the state factor x_0^1

$$f\left(x_0^1\middle|x_0^2,x_0^3,f_{\tau^*}\right) = \mathcal{N}_{x_0^1}\left(\hat{\mu}_0^1 - l_{21}x_0^2 - l_{31}x_0^3, p_0^1\right), \tag{4.16}$$

where

$$\hat{\mu}_0^1 = \hat{x}_0^1 + l_{21}\hat{x}_0^2 + l_{31}\hat{x}_0^3, \tag{4.17}$$

and for the second state factor x_0^2

$$f\left(x_0^2 \middle| x_0^3, f_{\tau^*}\right) = \mathcal{N}_{x_0^2}\left(\hat{\mu}_0^2 - l_{32}x_0^3, p_0^2\right), \quad (4.18)$$

where

$$\hat{\mu}_0^2 = \hat{x}_0^2 + l_{32}\hat{x}_0^3, \quad (4.19)$$

and respectively for the last state factor x_0^3

$$f\left(x_0^3 \middle| f_{\tau^*}\right) = \mathcal{N}_{x_0^3}\left(\hat{x}_0^3, p_0^3\right). \quad (4.20)$$

4.3 Conclusion

The paper described the algorithm of the factorized Kalman filtering for Gaussian multiple-output state-space model. The example of calculations of the state entry estimate for two-dimensional system state as well as illustrative experiments were presented. The experiments with the simulated state-space model demonstrated optimistic results, compared with the state estimates, obtained by Kalman filter.

The separate part of the paper provided the methodology of selection of initial conditions for the factorized Kalman filtering. The paper proposed the factorized version of prior knowledge processing, resulting in the specified initial state factor distribution.

The immediate future work will be primarily concerned with the experiments. The planned experiments with the factorized Kalman filter and processing of the prior knowledge for the factorized initial state will be based on data, obtained from the realistic traffic microsimulator.

Bibliography

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

[2] 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.

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

[4] E. Suzdaleva and M. Kárný, "Factorized Filtering", Tech. Rep. 2161, ÚTIA AV ČR, Praha, 2006.

[5] 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.

[6] E. Suzdaleva, "Factorized Kalman filtering", in *Proceedings of the 7th International PhD Workshop on Interplay of societal and technical decision-making*, J. Přikryl, J. Andrýsek, and V. Šmidl, Eds., Hruba Skala, Czech Republic, September 25-30 2006, pp. 226–233, ÚTIA AV ČR.

[7] 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, ISBN 1-4244-1348-6.

[8] E. Suzdaleva, "Initial conditions for Kalman filtering: Prior knowledge specification", in *Proceedings of the 7th WSEAS International Conference on Systems Theory and Scientific Computation (ISTASC'07)*, M.H. Le, M. Demiralp, V. Mladenov, and Z. Bojkovic, Eds., Athens, Greece, August 24-26 2007, Electrical and Computer Engineering, pp. 45–49, WSEAS Press.

[9] E. Suzdaleva, "Prior knowledge processing for initial state of Kalman filter", *International Journal of Adaptive Control and Signal Processing,* , 2007, submitted.

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