Ústav teorie informace a automatizace AV ČR, v.v.i.

Institute of Information Theory and Automation of the ASCR

RESEARCH REPORT

E. Suzdaleva

Filtering with discrete-valued states

2229 August 2008

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 Bayesian filtering [1], applied to discrete-valued data. The report presents a preparation part of research work in the field of filtering with mixed-type (continuous and discrete-valued) states. The mixed-type filtering is planned to be implemented as the entry-wise Kalman filtering [2] with an involved discrete-valued state.

1.2 Preliminaries

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

$$f\left(x_t, y_t \middle| x_{t-1}, u_t\right), \tag{1.1}$$

where x_t is the unobservable system state at discrete time moments $t \in t^* \equiv \{0, \ldots, \mathring{t}\}$, $\mathring{t}$ is the cardinality of the set t^*, $\equiv$ means equivalence, y_t is the system output and u_t is the system input.

1.2.1 Bayesian filtering

The state-space model in the form of joint pdf (1.1) allows to rewrite the Bayesian filtering [1] so that the posterior state estimate is updated by data and predicted in time simultaneously. According to Bayes rule [1], it holds

$$\begin{aligned} f\left(y_t, x_t \middle| u_t, d^{1:t-1}\right) &= f\left(y_t \middle| u_t, d^{1:t-1}, x_t\right) f\left(x_t \middle| u_t, d^{1:t-1}\right), & (1.2) \\ &= f\left(x_t \middle| d^{1:t}\right) f\left(y_t \middle| u_t, d^{1:t-1}\right). & (1.3) \end{aligned}$$

From last expression it can be seen, that data updating takes a form

$$f\left(x_t \middle| d^{1:t}\right) = \frac{f\left(y_t, x_t \middle| u_t, d^{1:t-1}\right)}{f\left(y_t \middle| u_t, d^{1:t-1}\right)}, \tag{1.4}$$

with Bayesian predictor [1] as a denominator. With the help of operation of marginalization and according to the used model, one can obtain the simultaneous data & time updating as follows.

$$\begin{aligned} f\left(x_t \middle| d^{1:t}\right) &= \frac{f\left(y_t, x_t \middle| u_t, d^{1:t-1}\right)}{f\left(y_t \middle| u_t, d^{1:t-1}\right)}, & (1.5) \\ &= \frac{\int f\left(x_t, y_t, x_{t-1} \middle| u_t, d^{1:t-1}\right) \, dx_{t-1}}{f\left(y_t \middle| u_t, d^{1:t-1}\right)}, & (1.6) \\ &= \frac{\int f\left(x_t, y_t \middle| u_t, x_{t-1}\right) f\left(x_{t-1} \middle| d^{1:t-1}\right) \, dx_{t-1}}{f\left(y_t \middle| u_t, d^{1:t-1}\right)}, & (1.7) \\ &\propto \int f\left(x_t, y_t \middle| u_t, x_{t-1}\right) f\left(x_{t-1} \middle| d^{1:t-1}\right) \, dx_{t-1}, & (1.8) \end{aligned}$$

where $\propto$ means proportionality. This version of Bayesian filtering[1] is assumed to be used throughout the paper.

[1] Application of Bayesian filtering to Gaussian state-space model with Gaussian observations and Gaussian prior distribution provides Kalman filter [3], implemented for the used model with simultaneous data and time updating.

Chapter 2

Filtering with discrete-valued states

2.1 Filtering with scalar state

Let's apply the filtering (1.8) to the discrete-valued data. According to the chain rule [1], the joint pdf inside the integral can be decomposed in the following way.

$$f\left(x_t\middle|\, d^{1:t}\right) \;\propto\; \int f\left(x_t|\, y_t, u_t, x_{t-1}\right) f\left(y_t|\, u_t, x_{t-1}\right) f\left(x_{t-1}\middle|\, d^{1:t-1}\right)\, dx_{t-1}. \tag{2.1}$$

Let's assume that the state and output are scalars such that $x_t \in \{0, 1\}$ and $y_t \in \{0, 1\}$. The system input u_t is the known constant and is out of interest for the considered problem.

The pdfs in (2.1) are supposed to have alternative distribution. The table with probabilities of taking possible values for the prior pdf $f\left(x_{t-1}\middle|\, d^{1:t-1}\right)$ is as follows.

Table 2.1: Alternative distribution for prior pdf

$x_{t-1} = 0$	$x_{t-1} = 1$
$p_{0;t-1} = p_{t-1}$	$p_{1;t-1} = (1 - p_{t-1})$

In the product form the prior alternative distribution takes the form

$$f\left(x_{t-1}\middle|\, d^{1:t-1}\right) = p_{t-1}^{\delta_{x_{t-1},0}} (1 - p_{t-1})^{\delta_{x_{t-1},1}} = \prod_{i=0}^{1} p_{i;t-1}^{\delta_{x_{t-1},i}}, \quad \text{where } \sum p_{i;t-1} = 1,\ p_{i;t-1} > 0\ \forall\ i. \tag{2.2}$$

Similarly, the table with alternative distribution for the pdf $f\left(x_t|\, y_t, u_t, x_{t-1}\right)$ is the following one with a

Table 2.2: Alternative distribution for pdf $f\left(x_t|\, y_t, u_t, x_{t-1}\right)$

	$x_t = 0$	$x_t = 1$
$y_t = 0, x_{t-1} = 0$	$p_{0\|00}$	$p_{1\|00}$
$y_t = 1, x_{t-1} = 0$	$p_{0\|10}$	$p_{1\|10}$
$y_t = 0, x_{t-1} = 1$	$p_{0\|01}$	$p_{1\|01}$
$y_t = 1, x_{t-1} = 1$	$p_{0\|11}$	$p_{1\|11}$

corresponding pdf in the product form

$$f\left(x_t|\, y_t, u_t, x_{t-1}\right) = \prod_{x_t|y_t,x_{t-1}} p_{x_t|y_t,x_{t-1}}^{\delta_{x_t|y_t,x_{t-1},\hat{x}_t|\hat{y}_t,\hat{x}_{t-1}}}, \tag{2.3}$$

where $\hat{x}_t$, $\hat{y}_t$ and $\hat{x}_{t-1}$ denote possible values from the table.

Analogously, Table (2.3) represents the alternative distribution for the pdf $f\left(y_t|\, u_t, x_{t-1}\right)$, which can be

Table 2.3: Alternative distribution for pdf $f\left(y_t|\, u_t, x_{t-1}\right)$

	$y_t = 0$	$y_t = 1$
$x_{t-1} = 0$	$p_{0\|0}$	$p_{1\|0}$
$x_{t-1} = 1$	$p_{0\|1}$	$p_{1\|1}$

expressed in the following product form.

$$f\left(y_t|\, u_t, x_{t-1}\right) = \prod_{y_t|x_{t-1}} p_{y_t|x_{t-1}}^{\delta_{y_t|x_{t-1},\hat{y}_t|\hat{x}_{t-1}}}. \tag{2.4}$$

The integral (2.1) is transformed into regular summation in the case of discrete-valued data. It takes the following form.

$$f\left(x_t\middle|\, d^{1:t}\right) = \sum_{x_{t-1}\in\{0,1\}} \prod_{x_t|y_t,x_{t-1}} p_{x_t|y_t,x_{t-1}}^{\delta_{x_t|y_t,x_{t-1},\hat{x}_t|\hat{y}_t,\hat{x}_{t-1}}} \prod_{y_t|x_{t-1}} p_{y_t|x_{t-1}}^{\delta_{y_t|x_{t-1},\hat{y}_t|\hat{x}_{t-1}}} \prod_{i=0}^{1} p_{i;t-1}^{\delta_{x_{t-1},i}}. \tag{2.5}$$

Now two probabilities, i.e. for $x_t = 0$ and for $x_t = 1$ can be calculated from (2.5).

$$\begin{aligned} f\left(x_t = 0\middle|\, d^{1:t}\right) &= p_t, \qquad (2.6)\\ &= \sum_{x_{t-1}\in\{0,1\}} \prod_{x_t|y_t,x_{t-1}} p_{x_t|y_t,x_{t-1}}^{\delta_{x_t|y_t,x_{t-1},0|\hat{y}_t,\hat{x}_{t-1}}} \prod_{y_t|x_{t-1}} p_{y_t|x_{t-1}}^{\delta_{y_t|x_{t-1},\hat{y}_t|\hat{x}_{t-1}}} \prod_{i=0}^{1} p_{i;t-1}^{\delta_{x_{t-1},i}},\\ &= \underbrace{p_{0|00}p_{0|10}p_{0|0}p_{1|0}p_{t-1}}_{sum\ for\ x_{t-1}=0} + \underbrace{p_{0|01}p_{0|11}p_{0|1}p_{1|1}(1-p_{t-1})}_{sum\ for\ x_{t-1}=1}. \end{aligned} \tag{2.7}$$

The result is obtained by substitution of probabilities from Table 2.2, corresponding to the state $x_t = 0$ and all probabilities from Tables 2.1-2.3 and subsequent summation over respective values of the state x_{t-1}. The probability for $x_t = 1$ is a complement for (2.6), it means, that it is enough to calculate

$$f\left(x_t = 1\middle|\, d^{1:t}\right) = 1 - p_t. \tag{2.8}$$

In such a way, the pdf $f\left(x_t\middle|\, d^{1:t}\right)$ preserves the form (2.2)

$$f\left(x_t\middle|\, d^{1:t}\right) = p_t^{\delta_{x_t,0}}(1-p_t)^{\delta_{x_t,1}} = \prod_{i=0}^{1} p_{i;t}^{\delta_{x_t,i}}, \quad \text{where } \sum p_{i;t} = 1,\ p_{i;t} > 0\ \forall\ i. \tag{2.9}$$

The updated probabilities, written in the table, are as follows. For $x_t \in x^*$, where x^* is a set of discrete values,

Table 2.4: Updated probabilities

$x_t = 0$	$x_t = 1$
$p_{0;t} = p_t$	$p_{1;t} = (1 - p_t)$

$x^* \equiv \{0, 1, \ldots, n\}$ with finite number n, the pdfs (2.2), and, respectively, (2.9), take the following form

$$f\left(x_t\middle|\, d^{1:t}\right) = \prod_{i=0}^{n} p_{i;t}^{\delta_{x_t,i}}, \quad \text{where } \sum p_{i;t} = 1,\ p_{i;t} > 0\ \forall\ i. \tag{2.10}$$

The filtering (2.5) in this case exploits the similar form of the prior distribution (2.2) with the models in the product forms (2.3) and (2.4) with summation over $x_{t-1} \in x^*$.

2.2 Filtering with state vector

The discrete-valued filtering with a state vector adds up to the scalar case through reduction of a state dimension via respective denoting. It means, that for state $x_t \equiv [x_{1;t}, \ldots, x_{k;t}]'$ with finite number k and $x_{i;t} \in x^*$, each state entry with all possible values is treated as an individual possible value of some scalar state $\tilde{x}_t$. For example, for $x_t \equiv [x_{1;t}, x_{2;t}]'$, $x_{1;t} \in \{0, 1\}$ and $x_{2;t} \in \{0, 1\}$ the possible values $\{(0, 0), (1, 0), (0, 1), (1, 1)\}$ are denoted as the new set $\tilde{x}^* \equiv \{1, 2, 3, 4\}$ for the scalar state $\tilde{x}_t$.

Bibliography

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

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

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