Katholieke Universiteit Leuven

Departement Elektrotechniek ESAT-SISTA/TR 06-156

Information, Covariance and Square-Root Filtering in the Presence of Unknown Inputs[1]

Steven Gillijns and Bart De Moor [2]

October 2006

Internal Report
Submitted for publication

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/gillijns/reports/TR-06-156.pdf*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SCD, Kasteelpark Arenberg 10, 3001 Leuven, Belgium, Tel. +32-16-32-17-09, Fax +32-16-32-19-70, WWW: *http://www.esat.kuleuven.be/scd*, E-mail: *steven.gillijns@esat.kuleuven.be*, Steven Gillijns is a research assistant and Bart De Moor is a full professor at the Katholieke Universiteit Leuven, Belgium. Research supported by Research Council KULeuven: GOA AMBioRICS, several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects, G.0407.02 (support vector machines), G.0197.02 (power islands), G.0141.03 (Identification and cryptography), G.0491.03 (control for intensive care glycemia), G.0120.03 (QIT), G.0452.04 (new quantum algoritheorems), G.0499.04 (Statistics), G.0211.05 (Nonlinear), research communities (ICCoS, ANMMM, MLDM); IWT: PhD Grants, GBOU (McKnow); Belgian Federal Science Policy Office: IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modelling', 2002-2006); PODO-II (CP/40: TMS and Sustainability); EU: FP5-Quprodis; ERNSI; Contract Research/agreements: ISMC/IPCOS, Data4s, TML, Elia, LMS, Mastercard

Abstract

The optimal filtering problem for linear systems with unknown inputs is addressed. Based on recursive least-squares estimation, information formulas for joint input and state estimation are derived. By establishing duality relations to the Kalman filter equations, covariance and square-root forms of the formulas follow almost instantaneously.

1 Introduction

Since the publication of Kalman's celebrated paper [10], the problem of state estimation for linear discrete-time systems has received considerable attention. In the decade immediately following the introduction of the Kalman filter, alternative implementations of the original formulas appeared. Most notable in the context of this paper are the information and square-root filters.

Information filters accentuate the recursive least-squares nature of the Kalman filtering problem [1, 3]. Instead of propagating the covariance matrix of the state estimate, these filters work with its inverse, which is called the information matrix. This approach is especially useful when no knowledge of the initial state is available.

To reduce numerical errors in a direct implementation of the Kalman filter equations, Potter and Stern [14] introduced the idea of expressing the equations in terms of the Cholesky factor of the state covariance matrix. Although computationally more expensive than the original formulation, these so-called square-root algorithms are numerically better conditioned than a direct implementation [16].

During the last decades, the problem of optimal filtering in the presence of unknown inputs has received growing attention due to its applications in environmental state estimation [11] and in fault detection and isolation problems [4]. Optimal state filters which assume that no prior knowledge of the input is available were for example developed by parameterizing the filter equations and then calculating the parameters which minimize the trace of the state covariance matrix under an unbiasedness condition [11, 5] or by transforming the system into a system which is decoupled from the unknown input and then deriving a minimum-variance unbiased state estimator based on this transformed system [7]. The problem of joint state and input estimation has also been intensively studied [8, 6]. In the latter reference, it is shown that the filter equations of [11] can be rewritten in a form which reveals optimal estimates of the input.

In this paper, we establish a relation between the filter of [6] and recursive least-squares estimation. We set up a least-squares problem to jointly estimate the state vector and the unknown input vector and derive information filter formulas by recursively solving this least-squares problem. We show that by converting the resulting formulas to covariance form, the filter of [6] is obtained. Finally, by establishing duality relations to the Kalman filter equations, a square-root implementation of the information filter follows almost instantaneously.

This paper is outlined as follows. In section 2, we formulate the filtering problem in more detail. Next, in section 3, we set up the least-squares problem and derive the filter equations by recursively solving the least-squares problem. In sections 4 and 5, we convert all filter equations into information form and covariance form and we discuss the relation to the results of [6]. Finally, in section 6, we develop a square-root implementation of the information filter.

We use the following notations. $\mathbb{E}[\cdot]$ denotes the expected value of a random variable, $^{\mathsf{T}}$

denotes matrix transposition, $\|a\|_B^2$ denotes the weighted norm $a^{\mathsf{T}}Ba$ and $\{a_i\}_{i=0}^k$ denotes the sequence $a_0, a_1, \ldots, a_k$. For a positive definite matrix A, $A^{1/2}$ denotes any matrix satisfying $A^{1/2}(A^{1/2})^{\mathsf{T}} = A$. We call $A^{1/2}$ a "square-root" of A. For conciseness of equations, we will also write $(A^{1/2})^{\mathsf{T}} = A^{\mathsf{T}/2}, (A^{1/2})^{-1} = A^{-1/2}$ and $(A^{-1/2})^{\mathsf{T}} = A^{-\mathsf{T}/2}$.

2 Problem formulation

Consider the linear discrete-time system

$$x_{k+1} = A_k x_k + G_k d_k + w_k, \tag{1a}$$
$$y_k = C_k x_k + v_k, \tag{1b}$$

where $x_k \in \mathbb{R}^n$ is the state vector, $d_k \in \mathbb{R}^m$ is an unknown input vector and $y_k \in \mathbb{R}^p$ is the measurement. The process noise w_k and the measurement noise $v_k \in \mathbb{R}^p$ are assumed to be mutually uncorrelated zero-mean white random signals with nonsingular covariance matrices $Q_k = \mathbb{E}[w_k w_k^{\mathsf{T}}]$ and $R_k = \mathbb{E}[v_k v_k^{\mathsf{T}}]$, respectively. We assume that an unbiased estimate $\hat{x}_0$ of the initial state x_0 is available with covariance matrix P_0. Also, we assume that rank $C_k G_{k-1} = m$ for all k. Finally, in the derivation of the information formulas, we also assume that A_k is invertible for all k.

In case d_k is known, is zero or is a zero-mean white random vector with known covariance matrix, the optimal filtering problem for the system (1) reduces to the Kalman filtering problem. On the other hand, if d_k is deterministic and its evolution in time is governed by a known linear system, optimal estimates of d_k and x_k can be obtained using an augmented state Kalman filter [1]. Like [11, 6], however, we consider the case where no prior knowledge about the time evolution or the statistics of d_k is available, that is, d_k is assumed to be completely unknown.

The derivation of the filters developed by [11, 6] is based on unbiased minimum-variance estimation. In this paper, however, we address the optimal filtering problem from the viewpoint of recursive least-squares estimation.

3 Recursive least-squares estimation

The relation between recursive least-squares estimation and the Kalman filter is well established. Let $d_k = 0$ and consider the least-squares problem

$$\begin{aligned} &\min_{\{x_i\}_{i=0}^k} \ J_k \\ &\text{subject to (1a) and (1b),} \end{aligned} \tag{2}$$

where the performance index J_k is given by

$$J_k = \|x_0 - \hat{x}_0\|^2_{P_0^{-1}} + \sum_{i=0}^{k} \|v_i\|^2_{R_i^{-1}} + \sum_{i=0}^{k-1} \|w_i\|^2_{Q_i^{-1}}.$$

Let $\{x^\star_{i|k}\}_{i=0}^k$ denote the solution to the minimization problem (2), then based on the recursion

$$J_k = J_{k-1} + \|v_k\|^2_{R_k^{-1}} + \|w_{k-1}\|^2_{Q_{k-1}^{-1}},$$

it can be shown that $x^\star_{k|k}$ can be computed recursively. More precisely, $x^\star_{k|k}$ can be derived from $x^\star_{k-1|k-1}$ and y_k by solving the minimization problem

$$\min_{x_k} \|y_k - C_k x_k\|^2_{R_k^{-1}} + \|x_k - A_{k-1} x^\star_{k-1|k-1}\|^2_{\bar{P}_k^{-1}}, \tag{3}$$

where $P^\star_k$ is given by

$$P^\star_k = \mathbb{E}[(x_k - A_{k-1} x^\star_{k-1|k-1})(x_k - A_{k-1} x^\star_{k-1|k-1})^\mathsf{T}].$$

In particular, it can be proved that $x^\star_{k|k} = \hat{x}^{\text{KF}}_k$ and $P^\star_k = P^{\text{KF}}_k$ where $\hat{x}^{\text{KF}}_k$ and P^{KF}_k denote the estimate of x_k and its covariance matrix obtained with the Kalman filter, respectively [15, 9, 1].

In this section, filter equations for the case d_k unknown will be derived based on recursive least-squares estimation. Let $\hat{x}_{k-1}$ denote the estimate of x_{k-1}, then in accordance to (3), $\hat{x}_k$ is obtained by solving the minimization problem

$$\min_{x_k, d_{k-1}} \|y_k - C_k x_k\|^2_{R_k^{-1}} + \|x_k - \hat{\bar{x}}_k - G_{k-1} d_{k-1}\|^2_{\bar{P}_k^{-1}}, \tag{4}$$

where

$$\hat{\bar{x}}_k := A_{k-1} \hat{x}_{k-1} \tag{5}$$

and where

$$\bar{P}_k := \mathbb{E}[(\bar{x}_k - \hat{\bar{x}}_k)(\bar{x}_k - \hat{\bar{x}}_k)^\mathsf{T}],$$

with

$$\bar{x}_k := A_{k-1} x_{k-1} + w_{k-1}. \tag{6}$$

We now derive explicit update formula by solving the minimization problem (4). First, note that (4) is equivalent to the least-squares problem

$$\min_{\mathcal{X}_k} \|\mathcal{Y}_k - \mathcal{A}_k \mathcal{X}_k\|^2_{\mathcal{W}_k}, \tag{7}$$

where

$$\mathcal{A}_k := \begin{bmatrix} C_k & 0 \\ I & -G_{k-1} \end{bmatrix}, \tag{8}$$

$$\mathcal{Y}_k := \begin{bmatrix} y_k \\ \hat{\bar{x}}_k \end{bmatrix}, \qquad \mathcal{X}_k := \begin{bmatrix} x_k \\ d_{k-1} \end{bmatrix},$$

and $\mathcal{W}_k := \mathrm{diag}(R_k^{-1}, \bar{P}_k^{-1})$. In order for (7) to have a unique solution, $\mathcal{A}_k$ must have full column rank, that is, G_{k-1} must have full column rank. The solution can then be written as

$$\hat{\mathcal{X}}_k = (\mathcal{A}_k^\mathsf{T} \mathcal{W}_k \mathcal{A}_k)^{-1} \mathcal{A}_k^\mathsf{T} \mathcal{W}_k \mathcal{Y}_k. \tag{9}$$

This solution has covariance matrix $(\mathcal{A}_k^\mathsf{T} \mathcal{W}_k \mathcal{A}_k)^{-1}$. Using (8), it follows that

$$\mathcal{A}_k^\mathsf{T} \mathcal{W}_k \mathcal{A}_k = \begin{bmatrix} \breve{P}_k^{-1} & -\bar{P}_k^{-1} G_{k-1} \\ -G_{k-1}^\mathsf{T} \bar{P}_k^{-1} & \breve{D}_{k-1}^{-1} \end{bmatrix},$$

where $\breve{P}_k^{-1}$ and $\breve{D}_{k-1}^{-1}$ are given by

$$\breve{P}_k^{-1} = \bar{P}_k^{-1} + C_k^\mathsf{T} R_k^{-1} C_k, \tag{10}$$

$$\breve{D}_{k-1}^{-1} = G_{k-1}^\mathsf{T} \bar{P}_k^{-1} G_{k-1}. \tag{11}$$

Furthermore, using [2, Prop. 2.8.7] it follows that the covariance matrix of $\hat{\mathcal{X}}_k$ can be written as

$$\begin{aligned} (\mathcal{A}_k^\mathsf{T} \mathcal{W}_k \mathcal{A}_k)^{-1} &= \begin{bmatrix} \breve{P}_k^{-1} & -\bar{P}_k^{-1} G_{k-1} \\ -G_{k-1}^\mathsf{T} \bar{P}_k^{-1} & \breve{D}_{k-1}^{-1} \end{bmatrix}^{-1}, \\ &= \begin{bmatrix} P_k & P_k \bar{P}_k^{-1} G_{k-1} \breve{D}_{k-1} \\ D_{k-1} G_{k-1}^\mathsf{T} \bar{P}_k^{-1} \breve{P}_k & D_{k-1} \end{bmatrix}, \end{aligned} \tag{12}$$

where the inverses of P_k and D_{k-1} are given by

$$D_{k-1}^{-1} = \breve{D}_{k-1}^{-1} - G_{k-1}^\mathsf{T} \bar{P}_k^{-1} \breve{P}_k \bar{P}_k^{-1} G_{k-1}, \tag{13}$$

$$P_k^{-1} = \breve{P}_k^{-1} - \bar{P}_k^{-1} G_{k-1} \breve{D}_{k-1} G_{k-1}^\mathsf{T} \bar{P}_k^{-1}. \tag{14}$$

Note that P_k and D_{k-1} can be identified as the covariance matrices of $\hat{x}_k$ and $\hat{d}_{k-1}$, that is,

$$\begin{aligned} P_k &= \mathbb{E}[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^\mathsf{T}], \\ D_{k-1} &= \mathbb{E}[(d_{k-1} - \hat{d}_{k-1})(d_{k-1} - \hat{d}_{k-1})^\mathsf{T}]. \end{aligned}$$

Substituting (12) in (9) then yields

$$\hat{\mathcal{X}}_k = \begin{bmatrix} P_k & P_k \bar{P}_k^{-1} G_{k-1} \breve{D}_{k-1} \\ D_{k-1} G_{k-1}^\mathsf{T} \bar{P}_k^{-1} \breve{P}_k & D_{k-1} \end{bmatrix} \begin{bmatrix} C_k^\mathsf{T} R_k^{-1} & \bar{P}_k^{-1} \\ 0 & -G_{k-1}^\mathsf{T} \bar{P}_k^{-1} \end{bmatrix} \mathcal{Y}_k,$$

from which it follows that

$$P_k^{-1} \hat{x}_k = \bar{P}_k^{-1} \hat{\bar{x}}_k + C_k^\mathsf{T} R_k^{-1} y_k - \bar{P}_k^{-1} G_{k-1} \breve{D}_{k-1} G_{k-1}^\mathsf{T} \bar{P}_k^{-1} \hat{\bar{x}}_k, \tag{15}$$

and

$$D_{k-1}^{-1}\hat{d}_{k-1} = -G_{k-1}^{\mathsf{T}}\bar{P}_k^{-1}\hat{\bar{x}}_k + G_{k-1}^{\mathsf{T}}\bar{P}_k^{-1}\breve{P}_k(\bar{P}_k^{-1}\hat{\bar{x}}_k + C_k^{\mathsf{T}}R_k^{-1}y_k). \tag{16}$$

Finally, we derive a closed form expression for $\bar{P}_k$. It follows from (6) and (5) that

$$\bar{x}_k - \hat{\bar{x}}_k = A_{k-1}(x_{k-1} - \hat{x}_{k-1}) + w_{k-1}.$$

Consequently, $\bar{P}_k$ is given by

$$\bar{P}_k = A_{k-1}P_{k-1}A_{k-1}^{\mathsf{T}} + Q_{k-1}. \tag{17}$$

By defining $\hat{\bar{x}}_k$, we did actually split the recursive update of the state estimate into two steps. The first step, which we call the *time update*, is given by (5). The second step, which we call the *measurement update*, is given by (15). Note that the time update is given in covariance form, whereas the measurement update is given in information form. In section 4, we will convert all equations in information form, in section 5 into covariance form.

Based on the least-squares formulation of the measurement update, (7), it is straightforward to derive a single least-squares problem for the combination of the time and measurement update. The resulting least-squares problem can be written as

$$\min_{\bar{\mathcal{X}}_k} \left\|\bar{\mathcal{Y}}_k - \bar{\mathcal{A}}_k\bar{\mathcal{X}}_k\right\|^2_{\bar{\mathcal{W}}_k}, \tag{18}$$

where

$$\bar{\mathcal{A}}_k := \begin{bmatrix} C_k & 0 & 0 \\ I & -G_{k-1} & 0 \\ -A_k & 0 & I \end{bmatrix},$$

$$\bar{\mathcal{Y}}_k := \begin{bmatrix} y_k \\ \hat{\bar{x}}_k \\ 0 \end{bmatrix}, \qquad \bar{\mathcal{X}}_k := \begin{bmatrix} x_k \\ d_{k-1} \\ \bar{x}_{k+1} \end{bmatrix},$$

and $\bar{\mathcal{W}}_k := \text{diag}(R_k^{-1}, \bar{P}_k^{-1}, Q_k^{-1})$. The least-squares problem (18) yields a method to recursively calculate $\hat{x}_k$. Indeed, let $\hat{\bar{x}}_k$ and $\bar{P}_k^{-1}$ be known, then the least-squares problem can be used to obtain the estimates $\hat{x}_k, \hat{d}_{k-1}$ and $\hat{\bar{x}}_{k+1}$ together with their covariance matrices. Once the measurement y_{k+1} is available, it can be used together with $\hat{\bar{x}}_{k+1}$ and $\bar{P}_{k+1}^{-1}$ as input data of a new least-squares problem of the form (18).

4 Information filtering

In this section, we convert the time update into information form, we derive more convenient formula for the measurement update and the estimation of the unknown input and we establish duality relations to the Kalman filter. The resulting equations are especially useful when no knowledge of the initial state is available ($P_0^{-1} = 0$) since in that case the covariance formulas of e.g. [6, 11] can not be used.

In rewriting the estimation of the unknown input and the measurement update, we will use the following equation, which follows by applying the matrix inversion lemma to (10),

$$\breve{P}_k = \bar{P}_k - \bar{P}_k C_k^{\mathsf{T}} \tilde{R}_k^{-1} C_k \bar{P}_k, \tag{19}$$

where

$$\tilde{R}_k := C_k \bar{P}_k C_k^{\mathsf{T}} + R_k. \tag{20}$$

4.1 Input estimation

A more convenient expression for (13) will now be derived. It follows from (11) and (13) that

$$D_{k-1}^{-1} = G_{k-1}^{\mathsf{T}} (\bar{P}_k^{-1} - \bar{P}_k^{-1} \breve{P}_k \bar{P}_k^{-1}) G_{k-1}. \tag{21}$$

Substituting (19) in (21), yields

$$\begin{aligned} D_{k-1}^{-1} &= F_k^{\mathsf{T}} \tilde{R}_k^{-1} F_k, & (22)\\ &= F_k^{\mathsf{T}} R_k^{-1} F_k - F_k^{\mathsf{T}} R_k^{-1} C_k (C_k^{\mathsf{T}} R_k^{-1} C_k + \bar{P}_k^{-1})^{-1} C_k^{\mathsf{T}} R_k^{-1} F_k, & (23) \end{aligned}$$

where $F_k := C_k G_{k-1}$ and where the last step follows by applying the matrix inversion lemma to (20).

A more convenient expression for (16) is obtained as follows. First, note that (16) can be rewritten as

$$D_{k-1}^{-1} \hat{d}_{k-1} = -G_{k-1}^{\mathsf{T}} (I - \bar{P}_k^{-1} \breve{P}_k) \bar{P}_k^{-1} \hat{\bar{x}}_k + G_{k-1}^{\mathsf{T}} \bar{P}_k^{-1} \breve{P}_k C_k^{\mathsf{T}} R_k^{-1} y_k. \tag{24}$$

Substituting (19) in (24), then yields

$$D_{k-1}^{-1} \hat{d}_{k-1} = F_k^{\mathsf{T}} R_k^{-1} y_k - F_k^{\mathsf{T}} R_k^{-1} C_k (C_k^{\mathsf{T}} R_k^{-1} C_k + \bar{P}_k^{-1})^{-1} (C_k^{\mathsf{T}} R_k^{-1} y_k + \bar{P}_k^{-1} \hat{\bar{x}}_k). \tag{25}$$

4.2 Measurement update

Now, we consider the measurement update. It follows from (10) and (14) that the information matrix P_k^{-1} can be written as

$$P_k^{-1} = \bar{P}_k^{-1} + C_k^{\mathsf{T}} R_k^{-1} C_k - \bar{P}_k^{-1} G_{k-1} (G_{k-1}^{\mathsf{T}} \bar{P}_k^{-1} G_{k-1})^{-1} G_{k-1}^{\mathsf{T}} \bar{P}_k^{-1}. \tag{26}$$

An expression for $\hat{x}_k$ in information form has already been derived,

$$P_k^{-1} \hat{x}_k = \bar{P}_k^{-1} \hat{\bar{x}}_k + C_k^{\mathsf{T}} R_k^{-1} y_k - \bar{P}_k^{-1} G_{k-1} (G_{k-1}^{\mathsf{T}} \bar{P}_k^{-1} G_{k-1})^{-1} G_{k-1}^{\mathsf{T}} \bar{P}_k^{-1} \hat{\bar{x}}_k, \tag{27}$$

see (15).

4.3 Time update

Since (5) and (17) take the form of the time update of the Kalman filter, information formulas follow almost immediately,

$$\begin{aligned}\bar{P}_k^{-1}\hat{\bar{x}}_k &= (I - L_{k-1})A_{k-1}^{-\mathsf{T}}P_{k-1}^{-1}\hat{x}_{k-1},\\ \bar{P}_k^{-1} &= (I - L_{k-1})H_{k-1},\end{aligned}$$

where

$$\begin{aligned}H_{k-1} &= A_{k-1}^{-\mathsf{T}}P_{k-1}^{-1}A_{k-1}^{-1},\\ \tilde{Q}_{k-1} &= (H_{k-1} + Q_{k-1}^{-1})^{-1},\\ L_{k-1} &= H_{k-1}\tilde{Q}_{k-1},\end{aligned}$$

see e.g. [1].

4.4 Duality to the Kalman filter

There is a duality between the recursion formula for the covariance matrix in the Kalman filter and equations (23) and (26). Consider the system

$$\begin{aligned}x_{k+1} &= A_k x_k + E_k w_k,\\ y_k &= C_k x_k + v_k,\end{aligned}$$

and let $\hat{x}_{k+1|k}^{\mathrm{KF}}$ denote the estimate of x_{k+1} given measurements up to time instant k obtained with the Kalman filter. The covariance matrix $P_{k+1|k}^{\mathrm{KF}}$ of $\hat{x}_{k+1|k}^{\mathrm{KF}}$ then obeys the recursion

$$P_{k+1|k}^{\mathrm{KF}} = A_k P_{k|k-1}^{\mathrm{KF}} A_k^{\mathsf{T}} + E_k Q_k E_k^{\mathsf{T}} - A_k P_{k|k-1}^{\mathrm{KF}} C_k^{\mathsf{T}} (C_k P_{k|k-1}^{\mathrm{KF}} C_k^{\mathsf{T}} + R_k)^{-1} C_k P_{k|k-1}^{\mathrm{KF}} A_k^{\mathsf{T}}. \quad (29)$$

The duality between (29) and (23), (26) is summarized in Table 1 and will be used in section 6 to derive square-root information algorithms for the measurement update and the estimation of the unknown input.

It follows from Table 1 that the dual of deriving a square-root covariance algorithm for the measurement update is deriving a square-root information algorithm for the Kalman filter equations of a system with perfect measurements. The latter problem is unsolvable. Therefore, we will not consider square-root covariance filtering for systems with unknown inputs.

5 Covariance filtering

In this section, we derive covariance formulas for the time update, the measurement update and the estimation of the unknown input. Also, we establish relations to the filters of [6] and [11].

Table 1: Duality between the recursion for $P^{\text{KF}}_{k|k-1}$ in the Kalman filter (29), the measurement update (26) and the estimation of the input (23).

Kalman filter, Eq. (29)	Eq. (26)	Eq. (23)
$P^{\text{KF}}_{k\|k-1}$	$\bar{P}_k^{-1}$	R_k^{-1}
A_k	I	F_k^{T}
R_k	0	$\bar{P}_k^{-1}$
C_k	G_{k-1}^{T}	C_k^{T}
E_k	C_k^{T}	0
Q_k	R_k^{-1}	0

5.1 Input estimation

First, we consider the estimation of the unknown input. An expression for the covariance matrix D_{k-1} is obtained by inverting (22), which yields

$$D_{k-1} = (F_k^{\mathsf{T}} \tilde{R}_k^{-1} F_k)^{-1}. \tag{30}$$

The expression for $\hat{d}_{k-1}$ then follows by premultiplying left and right hand side of (25) by (30), which yields

$$\hat{d}_{k-1} = (F_k^{\mathsf{T}} \tilde{R}_k^{-1} F_k)^{-1} F_k^{\mathsf{T}} \tilde{R}_k^{-1} \tilde{y}_k, \tag{31}$$

where $\tilde{y}_k := y_k - C_k \hat{\bar{x}}_k$. Note that $\hat{d}_{k-1}$ equals the solution to the least-squares problem

$$\min_{d_{k-1}} \|d_{k-1} - F_k \tilde{y}_k\|^2_{\tilde{R}_k^{-1}}.$$

Finally, note that (31) exists if and only if rank $F_k =$ rank $C_k G_{k-1} = m$.

5.2 Measurement update

Now, we consider the measurement update. By noting that (10) takes the form of the measurement update of the Kalman filter, it immediately follows that

$$\breve{P}_k = (I - K_k^x C_k) \bar{P}_k,$$

where K_k^x is given by

$$K_k^x = \bar{P}_k C_k^{\mathsf{T}} \tilde{R}_k^{-1}.$$

An expression for the covariance matrix P_k is obtained by applying the matrix inversion lemma to (14), which yields after some calculation

$$P_k = \breve{P}_k + (I - K_k^x C_k) G_{k-1} D_{k-1} G_{k-1}^{\mathsf{T}} (I - K_k^x C_k)^{\mathsf{T}}. \tag{32}$$

By premultiplying left and right hand side of (15) by P_k, we obtain the following expression for $\hat{x}_k$,

$$\hat{x}_k = \hat{\bar{x}}_k + K_k^x \tilde{y}_k + (I - K_k^x C_k) G_{k-1} \hat{d}_{k-1}. \tag{33}$$

5.3 Time update

Equations for the time update have already been derived,

$$\begin{aligned} \hat{\bar{x}}_k &= A_{k-1} \hat{x}_{k-1}, \\ \bar{P}_k &= A_{k-1} P_{k-1} A_{k-1}^{\mathsf{T}} + Q_{k-1}. \end{aligned} \tag{34}$$

see (5) and (17).

5.4 Relation to existing results

The covariance formulas derived in this section equal the filter equations of [6]. Furthermore, as shown in the latter reference, the state updates (34) and (33) are algebraically equivalent to the updates of [11].

6 Square-root information filtering

Square-root implementations of the Kalman filter exhibit improved numerical properties over the conventional algorithms. They recursively propagate Cholesky factors or "square-roots" of the error covariance matrix or the information matrix using numerically accurate orthogonal transformations. Square-root formulas in information form have been derived directly from the information formulas or based on duality considerations.

In this section, we use the duality relations established in Table 1 to derive a square-root implementation for the information formulas derived in the previous section. Like the square-root implementations for the Kalman filter, the algorithm applies orthogonal transformations to triangularize a pre-array, which contains the prior estimates, forming a post-array which contains the updated estimates.

6.1 Time update

First, we consider the time update. The duality to the time update of the Kalman filter yields (see e.g. [13])

$$\left[\begin{array}{cc} Q_{k-1}^{-\mathsf{T}/2} & -A_{k-1}^{-\mathsf{T}} P_{k-1}^{-\mathsf{T}/2} \\ 0 & A_{k-1}^{-\mathsf{T}} P_{k-1}^{-\mathsf{T}/2} \\ \hline 0 & \hat{x}_{k-1}^{\mathsf{T}} P_{k-1}^{-\mathsf{T}/2} \end{array}\right] \Theta_{1,k} = \left[\begin{array}{cc} \tilde{Q}_{k-1}^{-\mathsf{T}/2} & 0 \\ -L_{k-1} \tilde{Q}_{k-1}^{-\mathsf{T}/2} & \bar{P}_k^{-\mathsf{T}/2} \\ \hline \star & \hat{\bar{x}}_k^{\mathsf{T}} \bar{P}_k^{-\mathsf{T}/2} \end{array}\right],$$

where the "$\star$" in the post-array denotes a row vector which is not important for our discussion. The orthogonal transformation matrix $\Theta_{1,k}$ which lower-triangularizes the pre-array, may be implemented as a sequence of numerically accurate Givens rotations or Householder reflections [1].

6.2 Measurement update

Now, we consider the measurement update. A square-root information algorithm for the measurement update can be derived based on the duality to the Kalman filter. Using Table 1 and the square-root covariance algorithm for the Kalman filter developed in [12], yields the following update,

$$\left[\begin{array}{ccc} 0 & G_{k-1}^{\mathsf{T}}\bar{P}_k^{-\mathsf{T}/2} & 0 \\ 0 & \bar{P}_k^{-\mathsf{T}/2} & C_k^{\mathsf{T}}R_k^{-\mathsf{T}/2} \\ \hline 0 & \hat{\bar{x}}_k^{\mathsf{T}}\bar{P}_k^{-\mathsf{T}/2} & y_k^{\mathsf{T}}R_k^{-\mathsf{T}/2} \end{array}\right]\Theta_{2,k} = \left[\begin{array}{ccc} \breve{D}_{k-1}^{-\mathsf{T}/2} & 0 & 0 \\ \star & P_k^{-\mathsf{T}/2} & 0 \\ \hline \star & \hat{x}_k^{\mathsf{T}}P_k^{-\mathsf{T}/2} & \star \end{array}\right]. \tag{35}$$

The algebraic equivalence of this algorithm to equations (26) and (27) can be verified by equating inner products of corresponding block rows of the post- and pre-array.

A standard approach to convert between square-root covariance and square-root information implementations of the Kalman filter is to augment the post- and pre-array such that they become nonsingular and then invert both of them [12]. However, adding a row to the pre-array in (35) such that the augmented array becomes invertible and the inverse contains square-roots of covariance matrices, is not possible (due to the zero-matrix in the upper-left entry). This again shows that developing a square-root covariance algorithm is not possible.

6.3 Input estimation

Finally, we consider the estimation of the unknown input. Using the duality given in Table 1, we obtain the following array algorithm,

$$\left[\begin{array}{cc} \bar{P}_k^{-\mathsf{T}/2} & C_k^{\mathsf{T}}R_k^{-\mathsf{T}/2} \\ 0 & F_k^{\mathsf{T}}R_k^{-\mathsf{T}/2} \\ \hline \hat{\bar{x}}_k^{\mathsf{T}}\bar{P}_k^{-\mathsf{T}/2} & y_k^{\mathsf{T}}R_k^{-\mathsf{T}/2} \end{array}\right]\Theta_{3,k} = \left[\begin{array}{ccc} \breve{P}_k^{-\mathsf{T}/2} & 0 & 0 \\ \star & D_{k-1}^{-\mathsf{T}/2} & 0 \\ \hline \star & \hat{d}_{k-1}^{\mathsf{T}}D_{k-1}^{-\mathsf{T}/2} & \star \end{array}\right].$$

The algebraic equivalence of this algorithm to equations (23) and (25) can be verified by equating inner products of corresponding block rows of the post- and pre-array.

7 Conclusion

The problem of recursive least-squares estimation for systems with unknown inputs has been considered in this paper. It was shown that the solution to this least-squares problem yields information formulas for the filters of [11, 6]. By establishing duality relations to the Kalman filter equations, a square-root information implementation was developed almost instantaneously. Finally, it was shown that square-root covariance filtering for systems with unknown inputs is not possible.

Acknowledgements

Steven Gillijns and Niels Haverbeke are research assistants and Bart De Moor is a full professor at the Katholieke Universiteit Leuven, Belgium. This work is supported by Research Council KUL: GOA AMBioRICS, CoE EF/05/006 Optimization in Engineering, several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects, G.0407.02 (support vector machines), G.0197.02 (power islands), G.0141.03 (Identification and cryptography), G.0491.03 (control for intensive care glycemia), G.0120.03 (QIT), G.0452.04 (new quantum algorithms), G.0499.04 (Statistics), G.0211.05 (Nonlinear), G.0226.06 (cooperative systems and optimization), G.0321.06 (Tensors), G.0302.07 (SVM/Kernel, research communities (ICCoS, ANMMM, MLDM); IWT: PhD Grants, McKnow-E, Eureka-Flite2; Belgian Federal Science Policy Office: IUAP P6/04 (Dynamical systems, control and optimization, 2007-2011); EU: ERNSI.

References

[1] B.D.O. Anderson and J.B. Moore. *Optimal filtering.* Prentice-Hall, 1979.

[2] D.S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas with Application to Linear Systems Theory.* Princeton University Press, Princeton, New Jersey, 2005.

[3] G.J. Bierman. *Factorization Methods for Discrete Sequential Estimation.* Academic Press, New York, 1977.

[4] J. Chen and R.J. Patton. Optimal filtering and robust fault diagnosis of stochastic systems with unknown disturbances. *IEEE Proc. Contr. Theory Appl.*, 143:31–36, 1996.

[5] M. Darouach, M. Zasadzinski, and M. Boutayeb. Extension of minimum variance estimation for systems with unknown inputs. *Automatica*, 39:867–876, 2003.

[6] S. Gillijns and B. De Moor. Unbiased minimum-variance input and state estimation for linear discrete-time systems. *Automatica*, 43:111–116, 2007.

[7] M. Hou and P.C. Müller. Disturbance decoupled observer design: A unified viewpoint. *IEEE Trans. Autom. Control*, 39(6):1338–1341, 1994.

[8] C.S. Hsieh. Robust two-stage Kalman filters for systems with unknown inputs. *IEEE Trans. Autom. Control*, 45(12):2374–2378, 2000.

[9] A.H. Jazwinski. *Stochastic processes and filtering theory.* Academic Press, New York, 1970.

[10] R. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME – J. Basic Engr.*, 83:35–45, 1960.

[11] P.K. Kitanidis. Unbiased minimum-variance linear state estimation. *Automatica*, 23(6):775–778, 1987.

[12] M. Morf and T. Kailath. Square-root algorithms for least-squares estimation. *IEEE Trans. Autom. Control*, 20:487–497, 1975.

[13] P. Park and T. Kailath. New square-root algorithms for Kalman filtering. *IEEE Trans. Autom. Control*, 40(5):895–899, 1995.

[14] J.E. Potter and R.G. Stern. Statistical filtering of space navigation measurements. *Proc. of AIAA Guidance and Control Conf.*, 1963.

[15] H.W. Sorenson. Least-squares estimation: From Gauss to Kalman. *IEEE Spectrum*, 7:63–68, 1970.

[16] M. Verhaegen and P. Van Dooren. Numerical aspects of different Kalman filter implementations. *IEEE Trans. Autom. Control*, 31(10):907–917, 1986.