Automatica 78 (2017) 153–162

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

An estimation approach for linear stochastic systems based on characteristic functions☆

Moshe Idan [a], Jason L. Speyer [b]

[a] Aerospace Engineering, Technion - Israel Institute of Technology, Haifa 3200003, Israel

[b] Mechanical and Aerospace Engineering, University of California, Los Angeles, CA 90095, USA

ARTICLE INFO

Article history:
Received 25 May 2016
Received in revised form
21 September 2016
Accepted 15 November 2016

Keywords:
Estimation
Conditional probability density function
Filtering for non-Gaussian noise
Characteristic functions

ABSTRACT

This paper presents an alternative, characteristic function based approach for the Bayesian design of estimators for dynamic linear systems and linear detection problems. For a measurement update, the *a posteriori* characteristic function of the unnormalized conditional probability density function (ucpdf) of the state given the measurement history is obtained as a convolution of the *a priori* characteristic function of the ucpdf with the characteristic function of the measurement noise. It is shown that this convolution holds for a general measurement structure. Time propagation involves the product of the updated characteristic function of the ucpdf and the characteristic function of the process noise. Some estimation problems are found to be naturally tractable using only characteristic functions, such as the multivariable linear system with additive Cauchy measurement and process noise. It is shown that even the derivation of the Kalman filter algorithm has advantages when formulated using the characteristic function approach. Finally, in some instances the estimation problem can only be formulated in terms of characteristic functions. This is illustrated by a one-update scalar example for symmetric-α-stable distributions.

1. Introduction

Normally, in designing dynamic estimators, the problem is formulated in terms of probability density functions (pdf's) of the associated initial conditions and the measurement and process noises. The estimation problem is solved by determining the update and propagation of the conditional pdf of the state given the measurement history. For linear systems with additive Gaussian noise, this is the standard approach (Speyer & Chung, 2008). For scalar-state linear dynamic systems with additive Cauchy noise, the conditional pdf was also generated (Idan & Speyer, 2010). Sometimes using characteristic functions presents an attractive alternative (Idan & Speyer, 2012). In fact, when deriving the multivariate estimator for linear systems with Cauchy noises *only* the characteristic function approach led to a recursive, closed-form analytical solution (Idan & Speyer, 2014).

For a measurement update, the *a posteriori* characteristic function of the unnormalized conditional probability density function (ucpdf) of the state given the measurement history is obtained as a convolution of the *a priori* characteristic function of the ucpdf with the characteristic function of the measurement noise. In Idan and Speyer (2014), only the scalar measurement was considered. It is shown here that this convolution holds for a general measurement structure. Propagation involves the product of the updated characteristic function of the ucpdf and the characteristic function of the process noise and was derived in Idan and Speyer (2014). The characteristic function approach is advocated in this paper and may provide solutions for cases in which determining the associated conditional pdfs can be intractable (Idan & Speyer, 2014).

The general problem of state estimation for the multi-input-multi-output linear system, where the initial conditions and additive uncertainties have known characteristic functions, is presented in Section 2. In Section 3 we derive the update and propagation equations for the characteristic function of the unnormalized conditional probability density function of the state given the measurement history, followed by the expressions used to

☆ This work was supported by the United States-Israel Binational Science Foundation, Grant 2012122. The material in this paper was partially presented at the 2016 American Control Conference, July 6–8, 2016, Boston, MA, USA. This paper was recommended for publication in revised form by Associate Editor Gianluigi Pillonetto under the direction of Editor Torsten Söderström.

E-mail addresses: moshe.idan@technion.ac.il (M. Idan), speyer@g.ucla.edu (J.L. Speyer).

http://dx.doi.org/10.1016/j.automatica.2016.12.038
0005-1098/

determine the minimum conditional variance estimate and the associated estimation error variance. In particular, while the scalar linear measurement case was addressed in Idan and Speyer (2014), this result is extended to a state coefficient matrix in the measurement of any dimension and rank. In Section 4 two examples are given. In Section 4.1 the conditional mean and conditional variance are determined for a scalar state given a scalar measurement where the state and the measurement noise are in the class of symmetric α-stable ($S\alpha$-S) distributions (Samorodnitsky & Taqqu, 1994). Then, the derivation of the Kalman filter algorithm using the characteristic function approach is presented in Section 4.2. This approach allows for singularities in certain covariance matrices, normally assumed full rank in the standard derivation. Some concluding remarks are drawn Section 5.

2. Problem formulation

We consider the multi-input-multi-output linear system

$$x_{k+1} = \Phi x_k + \Gamma w_k, \qquad z_k = Hx_k + v_k, \tag{1}$$

with state vector $x_k \in \mathbb{R}^n$, vector measurement $z_k \in \mathbb{R}^m$, and known matrices $\Phi \in \mathbb{R}^{n\times n}$, $\Gamma \in \mathbb{R}^{n\times p}$, and $H \in \mathbb{R}^{m\times n}$. The independent vector-valued process noise sequence $w_k \in \mathbb{R}^p$ is assumed to have a known characteristic function $\phi_W(\nu_w)$ where $\nu_w \in \mathbb{R}^p$. Similarly, the independent measurement noise sequence $v_k \in \mathbb{R}^m$ is specified by its characteristic function $\phi_V(\nu_v)$ with $\nu_v \in \mathbb{R}^m$. The initial conditions at $k = 1$ are also assumed to be random variables with a known characteristic function $\phi_{X_1}(\nu)$ where $\nu \in \mathbb{R}^n$. In this work, the random variables w_k, v_k and x_1 are assumed to be independent.

The goal is to compute the minimum conditional variance estimate of x_k given the measurement history or $y_k = \{z_1 \quad z_2 \quad \cdots \quad z_k\}$.

3. Measurement update and time propagation equations

In the proposed design method, the sequential estimator is derived by propagating the characteristic function of the un-normalized conditional pdf of the state given the measurement history, that is the joint pdf of the state and observations history. In this section we derive those update and propagation equations, followed by the expressions used to determine the minimum conditional variance estimate and the associated estimation error variance.

3.1. Measurement update

Consider $\bar{f}_{X_k|Y_{k-1}}(x_k|y_{k-1})$, the un-normalized conditional pdf of the state at time k given the measurement history up to time step $k-1$, defined as

$$\bar{f}_{X_k|Y_{k-1}}(x_k|y_{k-1}) = f_{X_k|Y_{k-1}}(x_k|y_{k-1})\, f_{Y_{k-1}}(y_{k-1}). \tag{2}$$

Here, $f_{Y_{k-1}}(y_{k-1})$ is the pdf of the past measurement sequence, which, for a given measurement history, is a constant computed by integrating $\bar{f}_{X_k|Y_{k-1}}(x_k|y_{k-1})$ with respect to x_k. The conditional pdf is not normalized for computation efficiency and simplicity by avoiding the above mentioned integration. We assume that the characteristic function of the un-normalized pdf, i.e.

$$\bar{\phi}_{X_k|Y_{k-1}}(\nu) = \int_{x_k} \bar{f}_{X_k|Y_{k-1}}(x_k|y_{k-1})\, e^{j\nu^T x_k} dx_k, \tag{3}$$

is known. In (3), the notation $\int_{x_k}$ implies n integrals with respect to the components of the vector x_k over the range $(-\infty, \infty)$. A similar notation will be used throughout the paper, where the number of integrals is implied by the dimension of the indicated integration vector. The normalization factor in (2) can be easily computed from the characteristic function of the un-normalized pdf, i.e.

$$f_{Y_{k-1}}(y_{k-1}) = \bar{\phi}_{X_k|Y_{k-1}}(0). \tag{4}$$

It should be stressed that in the proposed approach only the characteristic function of the un-normalized pdf is determined, without specifically computing the pdfs of the variables.

Assuming that $\bar{\phi}_{X_k|Y_{k-1}}(\nu)$ is known, the measurement update implies processing the next measurement

$$z_k = Hx_k + v_k \tag{5}$$

to determine $\bar{\phi}_{X_k|Y_{k-1},Z_k}(\nu) = \bar{\phi}_{X_k|Y_k}(\nu)$. Using Bayes' theorem it can be easily verified that

$$\begin{aligned}\bar{f}_{X_k|Y_k}(x_k|y_k) &= f_{X_k|Y_k}(x_k|y_k)\, f_{Y_k}(y_k)\\ &= \bar{f}_{X_k|Y_{k-1}}(x_k|y_{k-1})\, f_{Z_k|X_k,Y_{k-1}}(z_k|x_k, y_{k-1})\\ &= \bar{f}_{X_k|Y_{k-1}}(x_k|y_{k-1})\, f_{Z_k|X_k}(z_k|x_k)\\ &= \bar{f}_{X_k|Y_{k-1}}(x_k|y_{k-1})\, f_V(z_k - Hx_k),\end{aligned} \tag{6}$$

where the last two equalities result from the fact that v_k is an independent sequence uncorrelated with w_k and x_{k-1}. The characteristic function of $\bar{f}_{X_k|Y_k}(x_k|y_k)$ is given by

$$\bar{\phi}_{X_k|Y_k}(\nu) = \int_{x_k} \bar{f}_{X_k|Y_{k-1}}(x_k|y_{k-1})\, f_V(z_k - Hx_k)\, e^{j\nu^T x_k} dx_k. \tag{7}$$

The above resembles a Fourier transform of a product of the two functions, i.e., of $\bar{f}_{X_k|Y_{k-1}}(x_k|y_{k-1})$ and $f_V(z_k - Hx_k)$. Using the dual convolution property, this integral can be solved by a convolution in the ν domain between the characteristic function of $\bar{f}_{X_k|Y_{k-1}}(x_k|y_{k-1})$, i.e., the currently known $\bar{\phi}_{X_k|Y_{k-1}}(\nu)$ of (3), and the characteristic function of $f_V(z_k - Hx_k)$, which we denote by $\hat{\phi}_{V_k}(\nu)$. It is given by

$$\hat{\phi}_{V_k}(\nu) = \int_{x_k} f_V(z_k - Hx_k)\, e^{j\nu^T x_k} dx_k. \tag{8}$$

Hence, $\bar{\phi}_{X|Z}(\nu)$ can be computed by the convolution integral

$$\bar{\phi}_{X_k|Y_k}(\nu) = \frac{1}{(2\pi)^n} \int_{\eta} \bar{\phi}_{X_k|Y_{k-1}}(\nu - \eta)\, \hat{\phi}_V(\eta)\, d\eta, \tag{9}$$

where $\eta \in \mathbb{R}^n$. $\hat{\phi}_V(\nu)$ is now determined for several cases. The scalar case, i.e., rank$H = m = 1$, was addressed in Idan and Speyer (2014). Here we extend this result to H of any dimension and rank while addressing four different cases.

3.1.1. Full-rank H cases

First we address the cases that H is full rank, i.e., rank$H = \min(n, m)$, with different number of measurements m relative to the number of states n.

Case 1: $m = n$

This is the simplest case, where the number of measurements equals the number of states and $H \in \mathbb{R}^{m\times m}$ is full rank, i.e., invertible. To address the integral (8), we introduce the following change of variables

$$\mu = z_k - Hx_k, \tag{10}$$

that yield the following relations:

$$x_k = H^{-1}z_k - H^{-1}\mu \quad \rightarrow \quad dx_k = \frac{d\mu}{|H|}. \tag{11}$$