Automatica 82 (2017) 158–164

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Feedback quadratic filtering☆

Filippo Cacace [a], Francesco Conte [b], Alfredo Germani [c], Giovanni Palombo [d]

[a] Università Campus Bio-Medico di Roma, Via Álvaro del Portillo, 21, 00128 Roma, Italy
[b] DITEN, Università degli studi di Genova, Via all'Opera Pia 11A, 16145, Genova, Italy
[c] DISIM, Università dell'Aquila, Via Vetoio, 67100 L'Aquila, Italy
[d] IASI CNR, Roma, Italy

ARTICLE INFO

Article history:
Received 17 March 2016
Received in revised form
11 January 2017
Accepted 4 April 2017

Keywords:
Filtering theory
Non-Gaussian processes
Kalman filters

ABSTRACT

This paper concerns the state estimation problem for linear discrete-time non-Gaussian systems. It is known that filters based on quadratic functions of the measurements processes (Quadratic Filter) improve the estimation accuracy of the optimal linear filter. In order to enlarge the class of systems, which can be processed by a Quadratic Filter, we rewrite the system model by introducing an output injection term. The resulting filter, named the Feedback Quadratic Filter, can be applied also to non asymptotically stable systems. We prove that the performance of the Feedback Quadratic Filter depends on the gain parameter of the output term, which can be chosen so that the estimation error is always less than or equal to the Quadratic Filter.

1. Introduction

In this paper, we study the state estimation problem for linear discrete-time non-Gaussian systems. In many applications, the widely used Gaussian assumption must be removed (see Spall, 1985, 2003 and Wu & Chen, 1993). In these cases, the conditional expectation, which gives the optimal minimum variance estimation, is the solution of an infinite dimensional problem (Zakai, 1969). Methods to approximate the state conditional probability density function include Monte Carlo methods (Arulampalam, Maskell, Gordon, & Clapp, 2002), sums of Gaussian densities (Arasaratnam, Haykin, & Elliott, 2007) and weighted sigma points (Julier & Uhlmann, 2004) among others. These general solutions can cope with nonlinearities and/or with the presence of noise outliers (Stojanovic & Nedic, 2015) or unknown parameters (Stojanovic & Nedic, 2016), and they generally have high computational cost. In the context of linear non-Gaussian systems, many research works aim at filtering algorithms that are easily computable (see Afshar, Yang, & Wang, 2012; Bilik & Tabrikian, 2010; Carravetta, Germani, & Raimondi, 1996; Gordon, Salmond, & Smith, 1993; Kassam & Thomas, 1976; Maryak, Spall, & Heydon, 2004; Picinbono & Devaut, 1988; Spall, 1995; Zhang, Kuai, Ren, Luo, & Lin, 2016 and the references therein). In the minimum variance framework a natural development is to use quadratic or polynomial functions of the observations to improve the estimation accuracy while preserving easy computability and recursion (Carravetta et al., 1996; De Santis, Germani, & Raimondi, 1995; Verriest, 1985). The suboptimal polynomial estimate is obtained by applying the KF to a system augmented with the powers of state and observations. A drawback of this approach is that the resulting augmented system is bilinear and the noise variance depends on the state variance of the original system. Thus, if the variance of the state grows unboundedly, so does the equivalent noise. As a consequence the stability of the resulting Quadratic Filter (QF) is guaranteed only for asymptotically stable systems. In this paper, we propose to use an output injection term to overcome this problem and obtain an internally stable QF. Furthermore, we show that any recursively implementable filter based on the use of powers of measurements has an error that depends on the choice of the gain of this output injection term, in contrast with the linear case. Thus, the gain can be chosen to achieve a smaller estimation error than the QF. A preliminary version of this work has been published in Cacace, Conte, Germani, and Palombo (2014), where the theoretical analysis was missing.

☆ The material in this paper was presented at the 53rd IEEE Conference on Decision and Control, December 15–17, 2014, Los Angeles, CA, USA. This paper was recommended for publication in revised form by Associate Editor Wei Xing Zheng under the direction of Editor Torsten Söderström.

E-mail addresses: f.cacace@unicampus.it (F. Cacace), fr.conte@unige.it (F. Conte), alfredo.germani@univaq.it (A. Germani), giovanni.palombo@univaq.it (G. Palombo).

http://dx.doi.org/10.1016/j.automatica.2017.04.046
0005-1098/

2. Discussion on quadratic filtering

Consider the problem of state estimation for a discrete-time linear system with non-Gaussian noise in the form

$$x(k+1) = Ax(k) + Bu(k) + f_k, \qquad x(0) = x_0 \tag{1}$$

$$y(k) = Cx(k) + g_k \tag{2}$$

where $x(k) \in \mathbb{R}^n$, $u(k) \in \mathbb{R}^p$, $y(k) \in \mathbb{R}^q$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$ and $C \in \mathbb{R}^{q\times n}$. $\{f_k\}$ and $\{g_k\}$ are sequences of non-Gaussian random variables with values in $\mathbb{R}^n$ and $\mathbb{R}^q$, respectively. The system is assumed fully observable, i.e. rank $\mathcal{O}(A, C) = n$, where $\mathcal{O}(A, C)$ is the observability matrix of the pair (A, C).

Throughout the paper we use the following notations. $x^{[i]}$ is the ith Kronecker power of a vector x. $st_q^{-1}(\cdot)$ is the inverse of the stack function, which transforms a vector in $\mathbb{R}^{q\cdot l}$ into a $q \times l$ matrix (see Carravetta et al., 1996; De Santis et al., 1995). Given a random vector $x \in \mathbb{R}^i$, $\psi_x^{(i)} = E[(x - E[x])^{[i]}]$, that is, its centered ith moment.

The random sequences $\{f_k\}$ and $\{g_k\}$ and x_0 satisfy the following conditions for $k \geq 0$:

(1) $x_0 \sim \mathcal{N}(\bar{x}_0, \Psi_{x_0})$;
(2) $\{f_k\}$ and $\{g_k\}$ are sequences of zero mean temporally independent random vectors;
(3) $\{f_k\}$, $\{g_k\}$ and x_0 are statistically independent;
(4) x_0, f_k and g_k have finite fourth moments;
(5) $\psi_{x_0}^{(i)}$, $\psi_f^{(i)}$ and $\psi_g^{(i)}$, $i = 2, 3, 4$, are known vectors;
(6) $[C\ \Psi_g]$, $\Psi_g = st_q^{-1}(\psi_g^{(2)})$, is full row rank (FRR).

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{G}$ be a given sub σ-algebra of $\mathcal{F}$ and $L^2(\mathcal{G}, n)$ be the Hilbert space of the n-dimensional, $\mathcal{G}$-measurable random variables with finite second moment. We write $L^2(X, n)$ to denote $L^2(\sigma(X), n)$, where $\sigma(X)$ is the σ-algebra generated by X. $\Pi\,[\cdot|\,\mathcal{M}]$ is the orthogonal projection onto a given Hilbert space $\mathcal{M}$. Given system (1)–(2), the output sequence $Y_k = \mathrm{col}(y(0), \ldots, y(k))$ and the auxiliary vector $Y_k' = \mathrm{col}(1, Y_k) \in \mathbb{R}^{l+1}$, $l = (k+1)q$, the minimum variance estimate of $x(k)$ is the orthogonal projection of $x(k)$ onto the Hilbert space $L^2(Y_k', n)$,

$$\hat{x}(k) = E\,[x(k)|\,\sigma(Y_k)] = \Pi\left[x(k)|\,L^2(Y_k', n)\right]. \tag{3}$$

If the sequences $\{x(k)\}$ and $\{y(k)\}$ are jointly Gaussian, this projection is equivalent to the projection on the closed subspace $\mathcal{L}_y^k \subset L^2(Y_k', n)$ of all affine functions of Y_k,

$$\mathcal{L}_y^k = \left\{z : \Omega \to \mathbb{R}^n : \exists T \in \mathbb{R}^{n\times l+1} : z = TY_k'\right\}. \tag{4}$$

The KF recursively computes the projection $\Pi[x(k)|\mathcal{L}_y^k]$, the best affine estimate of $x(k)$ in the minimum variance sense. This coincides with $E\,[x(k)|\,\sigma(Y_k)]$ only in the Gaussian case. When $\{x(k)\}$ and/or $\{y(k)\}$ are non-Gaussian, the computation of (3) is challenging.

Since the best affine estimate is obtained by projecting onto $\mathcal{L}_y^k$, better suboptimal estimates can be obtained by projecting the state $x(k)$ onto larger sub-spaces. For example, we may consider the space of second-order polynomial (quadratic) transformations of Y_k, denoted by $\mathcal{Q}_y^k$.

Let $Y_k^{(2)} = \mathrm{col}(Y_k',\ Y_k^{[2]}) \in \mathbb{R}^{\tilde{l}}$, $\tilde{l} = 1 + l + l^2$, then

$$\mathcal{Q}_y^k = \left\{z : \Omega \to \mathbb{R}^n : \exists T \in \mathbb{R}^{n\times\tilde{l}} : z = TY_k^{(2)}\right\}. \tag{5}$$

Since $\mathcal{L}_y^k \subset \mathcal{Q}_y^k \subset L^2(Y_k', n)$, projecting the state onto $\mathcal{Q}_y^k$ will return an estimate, having an error variance equal to or smaller than that of the affine estimate.

Theorem 1. *Suppose system (1)–(2) satisfies conditions (1)–(6). Let $\hat{x}^{\mathcal{L}}(k) = \Pi[x(k)|\mathcal{L}_y^k]$, $\hat{x}^{\mathcal{Q}}(k) = \Pi[x(k)|\mathcal{Q}_y^k]$, with errors $e^{\mathcal{L}}(k) = x(k) - \hat{x}^{\mathcal{L}}(k)$, $e^{\mathcal{Q}}(k) = x(k) - \hat{x}^{\mathcal{Q}}(k)$. Then, $E[e^{\mathcal{L}}(k)^T e^{\mathcal{L}}(k)] \geq E[e^{\mathcal{Q}}(k)^T e^{\mathcal{Q}}(k)]$.*

Proof. In virtue of the Hilbert projection theorem, $\hat{x}^{\mathcal{Q}}(k)$ has the minimum distance from $x(k)$ among all the elements of $\mathcal{Q}_y^k$. Therefore, since $\mathcal{L}_y^k \subset \mathcal{Q}_y^k$,

$$\left\|x(k) - \hat{x}^{\mathcal{L}}(k)\right\|^2_{L^2(X,n)} \geq \left\|x(k) - \hat{x}^{\mathcal{Q}}(k)\right\|^2_{L^2(X,n)}, \tag{6}$$

and the thesis follows from the definition of the norm in $L^2(X, n)$, $\|v\|^2_{L^2(X,n)} = \int_\Omega v^T v dP = E[v^T v]$. $\square$

To compute the optimal quadratic estimate the idea is to derive an augmented version of (1)–(2) with vectors $\mathcal{X}(k) = \mathrm{col}(x(k), x^{[2]}(k))$, $\mathcal{Y}(k) = \mathrm{col}(y(k), y^{[2]}(k))$ and use a recursive linear filter. To this aim we have to consider the following issues:

(1) $\Pi[x(k)|\mathcal{Q}_y^k]$ must be recursively computable;
(2) the augmented system must be detectable;
(3) the noise sequences of the augmented system must be second-order asymptotically stationary processes (as defined in Carravetta et al., 1996).

As for the first point, the computation of $\Pi[x(k)|\mathcal{Q}_y^k]$ would require a growing filter size, due to the presence in $\mathcal{Q}_y^k$ of terms of the kind $y^i(k_1)y^j(k_2)$, with $i, j \leq 2$. A possible solution (De Santis et al., 1995) is to replace $Y_k^{(2)}$ with

$$\overline{Y}_k^{(2)} = \mathrm{col}\left(Y_k',\ y(0)^{[2]}, \ldots, y(k)^{[2]}\right) \in \mathbb{R}^{\bar{l}}, \tag{7}$$

$\bar{l} = 1 + l + (k+1)q^2$, that is, a vector containing only the observations and their Kronecker squares from time 0 to k. We obtain the projection subspace

$$\overline{\mathcal{Q}}_y^k = \left\{z : \Omega \to \mathbb{R}^n : \exists T \in \mathbb{R}^{n\times\bar{l}} : z = T\overline{Y}_k^{(2)}\right\}, \tag{8}$$

with $\mathcal{L}_y^k \subset \overline{\mathcal{Q}}_y^k \subset \mathcal{Q}_y^k$. Thus, the corresponding *recursively computable quadratic estimate* will not be the optimal quadratic estimate, but it will still have an error variance not larger than the best linear one.

Detectability (issue 2) may not be satisfied even when the original pair (A, C) is fully observable, that is, the quadratic part of $\mathcal{X}(k)$ may not be completely observable. To solve these issues, we use output injection to rewrite (1)–(2) as an equivalent system with an asymptotically stable and hence detectable stochastic part.

This solves issue 3 as well. The problem stems from the fact that the state noise of the augmented system depends on $x(k)$, see Theorem 3.3.4 in Carravetta et al. (1996) or Theorem 1 in De Santis et al. (1995). By rewriting the system as above, the augmented system becomes asymptotically stationary.

3. The proposed approach

The proposed method, named Feedback Quadratic Filter (FQF), consists of the following steps.

(a) System (1)–(2) is rewritten as a system with a feedback given by an output injection term.
(b) The modified system is decomposed in the sum of a deterministic and a stochastic component.
(c) The augmented quadratic system is derived for the asymptotically stable stochastic component.
(d) The KF for mutually correlated state and output noises is applied to the augmented quadratic system. The final estimate is the sum of the deterministic component and the linear part of the augmented estimate.