Systems & Control Letters 108 (2017) 8–15

Contents lists available at ScienceDirect

Systems & Control Letters

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

Square-root algorithms for maximum correntropy estimation of linear discrete-time systems in presence of non-Gaussian noise

M.V. Kulikova

CEMAT (Center for Computational and Stochastic Mathematics), Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal

ARTICLE INFO

Article history:
Received 28 December 2016
Received in revised form 25 June 2017
Accepted 27 July 2017

Keywords:
Maximum correntropy criterion
Kalman filter
Square-root filtering
Robust estimation

ABSTRACT

Recent developments in the realm of state estimation of stochastic dynamic systems in the presence of non-Gaussian noise have induced a new methodology called the maximum correntropy filtering. The filters designed under the maximum correntropy criterion (MCC) utilize a similarity measure (or *correntropy*) between two random variables as a cost function. They are shown to improve the estimators' robustness against outliers or impulsive noises. In this paper we explore the numerical stability of *linear* filtering technique proposed recently under the MCC approach. The resulted estimator is called the maximum correntropy criterion Kalman filter (MCC-KF). The purpose of this study is two-fold. First, the previously derived MCC-KF equations are revised and the related Kalman-like equality conditions are proved. Based on this theoretical finding, we improve the MCC-KF technique in the sense that the new method possesses a better estimation quality with the reduced computational cost compared with the previously proposed MCC-KF variant. Second, we devise some square-root implementations for the newly-designed improved estimator. The square-root algorithms are well known to be inherently more stable than the *conventional* Kalman-like implementations, which process the full error covariance matrix in each iteration step of the filter. Additionally, following the latest achievements in the KF community, all square-root algorithms are formulated here in the so-called *array* form. It implies the use of orthogonal transformations for recursive update of the required filtering quantities and, thereby, no loss of accuracy is incurred. Apart from the numerical stability benefits, the array form also makes the modern Kalman-like filters better suited to parallel implementation and to very large scale integration (VLSI) implementation. All the MCC-KF variants developed in this paper are demonstrated to outperform the previously proposed MCC-KF version in two numerical examples.

1. Introduction

In the past few years, the study of filtering techniques under the maximum correntropy criterion (MCC) has become an important aspect of a hidden state estimation of stochastic dynamic systems in the presence of non-Gaussian noise [1–4]. The MCC methodology implies that a statistical metric of a *similarity* between two random variables (or *correntropy*) is used as a cost function (or performance index) for designing the corresponding estimation method. The resulted MCC filters have become the methods of choice in signal processing and machine learning due to its robustness against outliers or impulsive noises compared to the classical Kalman filtering (KF); e.g., see the discussion in [5–9] and many others.

Being a *linear* estimator, the KF is an attractive and simple technique that requires only the computation of mean and covariance for constructing the optimal estimate of unknown dynamic state under the minimum mean square (MMS) criterion. For Gaussian systems, this estimate is optimal, i.e. the KF reduces to an MMS estimate rather than a *linear* MMS estimate. It is clear that in non-Gaussian setting, the classical KF exhibits sub-optimal behavior only. Due to this fact, there was a need for a new estimator that improves the KF robustness against outliers or impulsive noises.

For linear non-Gaussian state-space models, the robust maximum correntropy Kalman filter (MCKF) and the maximum correntropy criterion Kalman filter (MCC-KF) have been recently developed in [10,11] and [12], respectively. As all Kalman-like filtering algorithms, they compute the first two moments (i.e. the mean and the covariance) for constructing the optimal estimate. However, in contrast to the classical KF, these recent developments utilize the robust MCC as the optimality criterion, instead of using the MMS cost function. As a result, the new filters are shown to outperform the classical KF and several nonlinear Kalman-like filtering techniques in the presence of non-Gaussian uncertainties in the state-space models. Nevertheless, little attention

E-mail address: maria.kulikova@ist.utl.pt.

http://dx.doi.org/10.1016/j.sysconle.2017.07.016
0167-6911/

is paid to numerical stability of the Kalman-like filters developed under the MCC strategy, although the classical KF is widely known to suffer from the influence of roundoff errors, severely; see [13,14]. Our research has tended to focus on the MCC-KF technique and the design of its numerically stable square-root implementations.

The purpose of this paper is two-fold. First, we revise the previously derived MCC-KF equations and prove the related Kalman-like equality conditions. Based on this theoretical finding, we improve the previously proposed MCC-KF algorithm in the sense that the new filter (abbreviated as IMCC-KF) possesses a better estimation quality with the reduced computational cost. Second, we devise some square-root IMCC-KF implementations grounded in numerically robust orthogonal transformations. The square-root strategy is the most popular approach used for enhancing the filter numerical robustness; see [15–18] *etc.* It implies the Cholesky decomposition of error covariance matrix and, then, recursive re-calculation of its Cholesky factors instead of using full matrix. Following the latest achievements in the KF community, all square-root algorithms are formulated here in the so-called *array* form. This means that numerically stable orthogonal transformations are used as far as possible for updating the Cholesky factors in each iteration step. This provides a more reliable estimation procedure as explained in [19, Chapter 12]. Apart from numerical advantages, array Kalman-like algorithms are easier to implement than the explicit filter equations, because all required quantities are simply read off from the corresponding filter post-arrays. As mentioned in [18], this makes the modern KF-like algorithms better suited to parallel implementation and to very large scale integration (VLSI) implementation. Finally, all algorithms developed in this paper are demonstrated to outperform the previously proposed MCC-KF technique in two numerical examples.

2. Maximum correntropy criterion Kalman filter

Consider the state-space equations

$$x_k = F_{k-1}x_{k-1} + G_{k-1}w_{k-1}, \quad k \ge 1, \tag{1}$$

$$z_k = H_k x_k + v_k \tag{2}$$

where $x_k \in \mathbb{R}^n$ and $z_k \in \mathbb{R}^m$ are the unknown dynamic state and the observable measurement vector, respectively. The processes $\{w_k\}$ and $\{v_k\}$ are zero-mean, white, uncorrelated, and have known covariance matrices Q_k and R_k, respectively. They are also uncorrelated with the initial state x_0, which has the mean $\bar{x}_0$ and the covariance matrix Π_0.

The KF associated with state-space model (1), (2) yields the linear MMS estimate, $\hat{x}_{k|k}$, of the unknown dynamic state, given the available measurements $\{z_1, \dots, z_k\}$. To improve the filter estimation quality in the presence of non-Gaussian noise, the MCC optimality criterion can be used instead of the MMS cost function for deriving the corresponding filtering equations. The performance index to be optimized under the MCC (with Gaussian kernel) approach is given as follows [4,12]:

$$J_m(x_k) = G_\sigma\left(\|z_k - H_k x_k\|\right) + G_\sigma\left(\|x_k - F_{k-1}x_{k-1}\|\right)$$

where $G_\sigma(\|x_k - y_k\|) = \exp\left\{-\|x_k - y_k\|^2/(2\sigma^2)\right\}$, and $\sigma > 0$ is the kernel size or bandwidth.

Minimization of the objective function J_m with respect to x_k implies $\partial J_m/\partial x_k = 0$ and yields the equation [4]:

$$(x_k - F_{k-1}x_{k-1}) = \frac{G_\sigma\left(\|z_k - H_k x_k\|\right)}{G_\sigma\left(\|x_k - F_{k-1}x_{k-1}\|\right)} H_k^T(z_k - H_k x_k). \tag{3}$$

We note that the best estimate for state vector x_{k-1} at time point $k-1$ is *a posteriori* estimate $\hat{x}_{k-1|k-1}$. Hence, from (3) one obtains the following nonlinear equation, which needs to be solved with respect to x_k:

$$x_k = F_{k-1}\hat{x}_{k-1|k-1} + \frac{G_\sigma\left(\|z_k - H_k x_k\|\right)}{G_\sigma\left(\|x_k - F_{k-1}\hat{x}_{k-1|k-1}\|\right)} H_k^T(z_k - H_k x_k). \tag{4}$$

The fixed point correntropy filter developed in [4] and the MCC-KF method proposed in [12] suggest to use a fixed point rule for solving the mentioned nonlinear equation with initial approximation $x_k^{(0)} = \hat{x}_{k|k-1}$ at the right-hand side of (4). Besides, both techniques imply only one iteration of the fixed point rule and, hence, by substituting $x_k \approx \hat{x}_{k|k-1}$ into the right-hand side of formula (4) we obtain the following recursion

$$\hat{x}_{k|k} = F_{k-1}\hat{x}_{k-1|k-1} + \frac{G_\sigma\left(\|z_k - H_k\hat{x}_{k|k-1}\|\right)}{G_\sigma\left(\|\hat{x}_{k|k-1} - F_{k-1}\hat{x}_{k-1|k-1}\|\right)} \times H_k^T(z_k - H_k\hat{x}_{k|k-1}).$$

Next, the MCC-KF method designed in [12] integrates the KF minimum-variance estimation with the maximum correntropy filtering. In particular, the cited paper utilizes the norm $\|\cdot\|_{R_k^{-1}}$ induced by the inverse measurement covariance matrix R_k^{-1} in the numerator and the norm $\|\cdot\|_{P_{k|k-1}^{-1}}$ induced by the inverse predicted process covariance matrix $P_{k|k-1}^{-1}$ in the denominator of the recursion above. Thus, the MCC-KF is given as follows; see Algorithm 2 in [12]:

Initialization:

$$\hat{x}_{0|0} = \mathbf{E}\{x_0\}, \qquad P_{0|0} = \mathbf{E}\left\{(x_0 - \hat{x}_{0|0})(x_0 - \hat{x}_{0|0})^T\right\}. \tag{5}$$

Prior estimation:

$$\hat{x}_{k|k-1} = F_{k-1}\hat{x}_{k-1|k-1}, \tag{6}$$

$$P_{k|k-1} = F_{k-1}P_{k-1|k-1}F_{k-1}^T + G_{k-1}Q_{k-1}G_{k-1}^T. \tag{7}$$

Posterior estimation:

$$L_k = \frac{G_\sigma\left(\|z_k - H_k\hat{x}_{k|k-1}\|_{R_k^{-1}}\right)}{G_\sigma\left(\|\hat{x}_{k|k-1} - F_{k-1}\hat{x}_{k-1|k-1}\|_{P_{k|k-1}^{-1}}\right)}, \tag{8}$$

$$K_k^L = (P_{k|k-1}^{-1} + L_k H_k^T R_k^{-1} H_k)^{-1} L_k H_k^T R_k^{-1}, \tag{9}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k^L(z_k - H_k\hat{x}_{k|k-1}), \tag{10}$$

$$P_{k|k} = (I - K_k^L H_k)P_{k|k-1}(I - K_k^L H_k)^T + K_k^L R_k (K_k^L)^T. \tag{11}$$

In the equations above, we use the new notation K_k^L for the gain matrix $(P_{k|k-1}^{-1} + L_k H_k^T R_k^{-1} H_k)^{-1} L_k H_k^T R_k^{-1}$ appeared in Algorithm 2 in [12], emphasizing the dependence of this quantity on the scalar L_k. This also helps us to distinguish this matrix from the classical KF feedback gain in the rest of our paper.

The readers are referred to [12] for a detailed derivation and properties of the MCC-KF estimator under consideration. In the cited paper, the MCC-KF is shown to outperform the classical KF, the fixed point correntropy filter from [4] and several nonlinear filtering techniques when the non-Gaussian uncertainties arise in stochastic system (1), (2).

It is worth noting here that because of utilizing only one iteration of a fixed point rule for solving the underlying nonlinear equation (4), we have $G_\sigma\left(\|\hat{x}_{k|k-1} - F_{k-1}\hat{x}_{k-1|k-1}\|\right) = G_\sigma\left(\|0\|\right) = 1$ since $\hat{x}_{k|k-1} = F_{k-1}\hat{x}_{k-1|k-1}$. Hence, both methods in [4,12] can be simplified since the denominator in (8) is equal to 1. For further iterates, this is not the case and the difference might be considerable. For this reason, the general form of (8) is used in this paper.