Journal of Process Control 24 (2014) 81–87

Contents lists available at ScienceDirect

Journal of Process Control

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

Kalman filter with both adaptivity and robustness

Guobin Chang *

Tianjin Institute of Hydrographic Surveying and Charting, Tianjin, China

ARTICLE INFO

Article history:
Received 27 August 2013
Received in revised form
28 December 2013
Accepted 30 December 2013
Available online 28 January 2014

Keywords:
Kalman filter
Adaptive
Robust
Fading memory
Enhancing memory

ABSTRACT

Adaptive and robust methods are two opposite strategies to be adopted in the Kalman filter when the difference between the predictive observation and the actual observation, i.e. the innovation vector is abnormally large. The actual observation is more weighted in the former one, and is less weighted in the later one. This article addresses the subject of making a choice between the adaptive and robust methods when abnormal innovation occurs. An adaptive method with fading memory and a robust method with enhancing memory is proposed in the Kalman filter based on the chi-square distribution of the square of the Mahalanobis distance of the innovation. A heuristic method of recursively choosing among the adaptive, the robust, and the standard Kalman filter approaches in the occurrence of abnormal innovations is proposed through incorporating the observations at the next instance. The proposed method is both adaptive and robust, i.e. having the ability of strongly tracking the variation of the state and being insensitive to gross errors in observation. Numerical simulations of a simple illustrating example validate the efficacy of the proposed method.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Kalman filter (KF) has been applied in many areas such as signal processing, GPS positioning, integrated navigation, target tracking, and so on. Although there was no assumptions about Gaussian distributed process and measurement noises, and linear state and measurement functions in the seminal paper [1], it can be proven that KF is optimal in the sense of being unbiased, consistent, and asymptotic efficient only when the above assumptions hold [2]. Under the Gaussian assumption, the innovation, i.e. the difference between the predictive observation and the actual observation at an arbitrary instance, say k, should be a zero-mean Gaussian distributed stochastic vector, and the square of the Mahalanobis distance of the innovation, defined as the judging index in this contribution, should be χ^2 distributed with the dimension of the observation as its degree of freedom. According to the hypothesis testing theory, for this χ^2 distribution, given a significance level, say α, which is predetermined as a small value, the judging index should be smaller than the α-quantile with the probability $1-\alpha$ which is a high probability as α is assumed a small value. If it is not the case, it could be concluded with the same high probability $(1-\alpha)$ that there are some kinds of violation to the a priori assumptions.

Generally speaking, those violations may be due to modeling errors and/or gross errors. The model errors fall into two categories, i.e. functional and stochastic model errors. The functional model errors may be due to some sudden variation of parameters including the discontinuous changes of the states and/or the system parameters. The stochastic model errors are due to the incorrect knowledge of the statistics of the process and/or measurement noises. Both the functional and stochastic model errors can be dealt with using some kind of adaptive method, in which the varying parameters or incorrect statistics are recursively tuned to accord with the actual incoming observations. The gross errors, which show themselves as outliers, may occur in the state prediction and/or update stages. It is noted that outliers may not necessarily be gross errors implying some outlying data can still be good ones. However, redundant information is always unavailable to detect gross errors, while outliers can be distinguished relatively easily, so outliers, other than gross errors are addressed in most practical problems. In this distribution, it is our basic assumption that an outlier is due to gross errors. Those gross errors are often due to random hardware failures or man-made mistakes, and are often dealt with using robust method, in which the influence of those outliers are reduced. In this study, only the sudden variation of the states and the gross error in the measurement are involved to represent the above mentioned two kinds of violations.

There are many kinds of adaptive and robust methods in Kalman filtering in the literature. Fading memory filters [3], covariance matching filters which matches process noise covariance, or measurement noise covariance, or both [4], Gauss sum filters [5], interactive multiple model filters [6], are some of the familiar ones. Robust methods include median estimates [7], H∞ filters [8], federated filters [9], measurement-noise-inflating filters [10],

* Tel.: +86 18630876058.
E-mail address: guobinchang@hotmail.com

0959-1524/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jprocont.2013.12.017

filter-gain-rescaling filters [11], M-estimation based filters [12,13], among others. In this paper two simple but effective methods are adopted, i.e. fading memory filter to make the filter adaptive and measurement-noise-inflating filter to make the filter robust, and a heuristic method of making a choice recursively among the adaptive, the robust, and the standard KF is proposed.

The remaining of the paper is organized as follows, in Section 2, the problem is formulated and the available adaptive and/or robust method is reviewed. In Section 3, an adaptive method with fading memory and a robust method with enhancing memory both based on the χ^2 distributed judging index are proposed, and a new scheme to recursively select among adaptive, robust, and standard strategies in KF is proposed. In Section 4, single-axis INS/GPS integration is simulated to justify the method proposed. Conclusions are made in Section 5.

2. Problem formulation

Below, the symbols "$\hat{}$" and "$\sim$" above a variable represent an estimate and a measurement; the superscripts "−" and "+" represent the a priori and a posteriori estimates respectively; $N(x, \mu, P)$ denotes that the variable x obeys Gaussian distribution with mean μ and covariance P.

Consider the discrete-time linear stochastic state space model with Gaussian distributed process and measurement noise,

$$x_k = F_{k-1}x_{k-1} + w_{k-1} \tag{1}$$

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

where x_k is the n-dimensional state vector at time instance k and is to be estimated. The symbol y_k is the m-dimensional observation vector which can be measured at time instance k. The symbols F_k, H_k are the $n \times n$ and $m \times n$ dimensional propagation matrix and observational matrix respectively. And w_k, v_k are process and measurement noises respectively, both of which are zero-mean uncorrelated Gaussian white noise satisfying $E[w_k w_j^T] = Q_k\delta_{kj}$, $E[v_k v_j^T] = R_k\delta_{kj}$, $E[w_k v_j^T] = 0$, where Q_k and R_k are the corresponding covariances, and δ_{kj} is the Kronecker delta function. The initial state estimation is assumed to be Gaussian with mean $\hat{x}_0^+$ and covariance P_0^+, and uncorrelated with any process and measurement noises.

Systems (1) and (2) are solved by KF to give the a posteriori estimate of x_k given the observations $\tilde{y}_1, \tilde{y}_2, \cdots, \tilde{y}_k$. The KF equations are as follows:

$$\hat{x}_k^- = F_{k-1}\hat{x}_{k-1}^+ \tag{3}$$

$$P_k^- = F_{k-1}P_{k-1}^+F_{k-1}^T + Q_{k-1} \tag{4}$$

$$\eta_k = \tilde{y} - H_k\hat{x}_k^- \tag{5}$$

$$P_{\eta_k} = H_k P_k^- H_k^T + R_k \tag{6}$$

$$K_k = P_k^- H_k^T (P_{\eta_k})^{-1} \tag{7}$$

$$\hat{x}_k^+ = \hat{x}_k^- + K_k\eta_k \tag{8}$$

$$P_k^+ = P_k^- - K_k H_k P_k^- \tag{9}$$

where K_k is the KF gain at instance k, and η_k is the innovation vector stated above.

If the assumptions about (1) and (2) hold, the innovation η_k should be zero-mean Gaussian-distributed with covariance $H_k P_k^- H_k^T + R_k$, so the square of the Mahalanobis distance of the innovation should be χ^2 distributed, i.e.

$$M^2 = \eta_k^T (P_{\eta_k})^{-1} \eta_k \sim \chi_m^2 \tag{10}$$

According the hypothesis testing theory, for a given significance level, say α, we have

$$Pr(M^2 < \chi_{m,\alpha}^2) = 1 - \alpha \tag{11}$$

where $\Pr(\cdot)$ represents the probability of a random event and $\chi_{m,\alpha}^2$ is the α-quantile of the distribution χ_m^2.

If (11) does not hold, it can be concluded with high probability $(1-\alpha)$ that there are some violations to the assumptions about the system, e.g. a discontinuous change of the state x_k or some gross errors in the observation $\tilde{y}_k$. In this case, we call the innovation η_k abnormally large, or abnormal for brevity. Attention should be paid that the so called "abnormal innovation" here is just an expression without any statistical meaning, or in other word, it does not necessarily mean that this innovation is not a normally, or Gaussian distributed random variable.

Fading KFs work on the prerequisite that the actual observations are correctly obtained. The abnormal innovation implies that there are certain modeling errors which violate the normal operation of KF, in this paper only the discontinuous variation of the state is assumed to represent such modeling errors. Some kind of fading factors could be used to rescale the weights of the predictive state estimation so as to adapt the filter to the actual observation, i.e. by enlarging P_k^-, $\hat{x}_k^-$ is less weighted in the update stage, as a result the actual observation is more weighted. Fading factors can be introduced into the prior covariance matrix through three different ways, i.e. rescaling P_{k-1}^+, see e.g. [3]; rescaling Q_{k-1}, see e.g. [14], or directly rescaling P_k^-, see e.g. [11]. Fading factors can be either single ones [14] or multiple ones in the form of diagonal matrix [3]. In [15] a special structure to rescale P_k^- is proposed. This structure is different from most of the ordinary fading filters and is derived rigorously through linear matrix inequality. In [11], the stability analysis of the fading filter is carried out. Fading memory strategies have been introduced into the newly developed derivative-free nonlinear filters such as unscented KF [14]. Generally the performance of different fading filters may be slightly different in certain applications, but it is the basic common point that all of them have the ability of adaptation. It is noted that (11) is not the only criterion to detect the abnormal operation of the filter, many other criterion are used in the above papers.

Contrary to the adaptive, or specifically the fading filters, some robust filters assume that the process model is correctly constructed, so P_k^- and $\hat{x}_k^-$ are correctly obtained. The abnormally large innovation vector implies that there are some errors in the observation, so the observation is less weighted in the update stage of the KF through modifying the observation noise covariance, as a result, $\hat{x}_k^-$ is more weighted. This can be achieved by directly enlarging R_k [16]. An enlarged R_k results in a reduced filter gain K_k, so robustness can also be achieved by directly rescaling K_k[11]. M-estimation based robust filters have been widely studied in recent years. Bayesian estimator, as a generalization of KF was robustified in [17]. As a direct application of [17], the M-estimation based KF for rank deficient observation model was proposed in [18]. Then the M-estimator based KF is extended to the derivative-free nonlinear filters through the statistical linear regression [19]. But these nonlinear robust KFs achieve robustness at the cost of reducing the accuracy of the nonlinear transformation itself, so they are modified through exploiting the intuitive meaning of the M-estimator, i.e. that the process of M-estimator is equivalent to constructing pseudo observations or inflating the covariances [12]. This modified robust filter is further studied in [13] and extended to divided difference filter in [20].

It is apparent from the above analysis that adaptive and robust method are two opposite strategies, i.e. actual observation is assumed correct in the adaptive ones while predictive state estimation is assumed correct in the robust ones. Both the adaptive and

Download English Version:

https://daneshyari.com/en/article/689072

Download Persian Version:

https://daneshyari.com/article/689072

Daneshyari.com