Commentationes Mathematicae Universitatis Carolinae

Tomáš Cipra

Note on approximate non-Gaussian filtering with nonlinear observation relation

Commentationes Mathematicae Universitatis Carolinae, Vol. 31 (1990), No. 3, 601--605

Persistent URL: http://dml.cz/dmlcz/106893

Terms of use:

© Charles University in Prague, Faculty of Mathematics and Physics, 1990

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://project.dml.cz

Note on approximate non-Gaussian filtering with nonlinear observation relation

TOMÁŠ CIPRA

Abstract. Masreliez's theorem [3] on approximate non-Gaussian filtering with linear state and observation relations is generalized for a simple case of nonlinear observation relations.

Keywords: Approximate conditional-mean filter, non-Gaussian nonlinear filter, robust Kalman filter

Classification: 60G35, 62M10

1. Introduction.

The well-known Kalman filter (see e.g. [1], [2]) deals with recursive estimation of an n-dimensional observations $\{z_1, z_2, \dots, z_k\}$ in a discrete linear system

$$x_{k+1} = F_k x_k + w_k, \tag{1.1}$$

$$z_k = H_k x_k + v_k, \tag{1.2}$$

where F_k is a known state transition matrix of the type $n \times n$ in the state relation (1.1), H_k is a known observation matrix of the type $r \times n$ in the observation relation (1.2) and w_k and v_k are disturbance vectors of the type $n \times 1$ and $r \times 1$, respectively, which form zero mean uncorrelated and mutually uncorrelated sequences with known covariance matrices. Moreover, some initial conditions are given.

Under the assumption of normality the Kalman filter gives the optimal minimum variance state estimator which is equal to the conditional expectation

$$\hat{x}_k^k = E(x_k \mid Z^k), \tag{1.3}$$

where the denotation $Z^k = \{z_1, z_2, \dots, z_k\}$ is used. However, in many practical situations one must face highly non-Gaussian densities or the ones which have Gaussian shape in the middle but differ from normality in the tails so that robustification of the Kalman filter is necessary to protect the state estimator against outliers. Since non-Gaussian minimum variance filters are usually difficult to implement one looks for various approximations which retain the numerically simple recursive structure of the Kalman filter also in non-Gaussian cases.

One of successful approaches to non-Gaussian filtering in the linear system (1.1) and (1.2) consists in the assumption that the predicted state density $p(x_k \mid Z^{k-1})$ is approximately Gaussian although the observation disturbances v_k may be non-Gaussian (see [3]):

Let us denote

$$\hat{x}_k^{k-1} = E(x_k \mid Z^{k-1}), \tag{1.4}$$

$$P_k^k = E\{(x_k - \hat{x}_k^k)(x_k - \hat{x}_k^k)' \mid Z^k\}, \tag{1.5}$$

$$P_k^{k-1} = E\{(x_k - \hat{x}_k^{k-1})(x_k - \hat{x}_k^{k-1})' \mid Z^{k-1}\}. \tag{1.6}$$

Further let $p(z_k \mid Z^{k-1})$ be the predicted observation density which is assumed to exist twice differentiable.

Theorem 1. (Masreliez [3]) *Assume that the predicted state density $p(x_k \mid Z^{k-1})$ is Gaussian with mean $\hat{x}_k^{k-1}$ and covariance matrix P_k^{k-1}. Then it holds*

$$\hat{x}_k^k = \hat{x}_k^{k-1} + P_k^{k-1} H_k' g_k(z_k), \tag{1.7}$$

$$P_k^k = P_k^{k-1} - P_k^{k-1} H_k' G_k(z_k) H_k P_k^{k-1}, \tag{1.8}$$

where

$$g_k(z_k) = -[p(z_k \mid Z^{k-1})]^{-1} \frac{\partial p(z_k \mid Z^{k-1})}{\partial z_k}, \tag{1.9}$$

$$G_k(z_k) = \frac{\partial g_k(z_k)}{\partial z_k}. \tag{1.10}$$

Theorem 1 has important applications. For example, it presents a theoretical background for so called ACM (Approximate Conditional-Mean) filters which are useful instruments in robust time series analysis (see [4], [5]).

The aim of this note is to generalize the Masreliez's theorem for the discrete systems with nonlinear observation relations. This generalization can be looked upon as a theoretical result for possible non-Gaussian nonlinear filtering.

2. Nonlinear observation relation.

We consider the simple scalar case

$$z_k = h(x_k) + v_k \tag{2.1}$$

with a function $h : R^1 \to R^1$ which can be approximated sufficiently by the second--order Taylor expansion so that

$$z_k = h(\hat{x}_k^{k-1}) + h_x(\hat{x}_k^{k-1})(x_k - \hat{x}_k^{k-1}) + \frac{1}{2} h_{xx}(\hat{x}_k^{k-1})(x_k - \hat{x}_k^{k-1})^2 + v_k, \tag{2.2}$$

where

$$h_x(\hat{x}_k^{k-1}) = \frac{\partial h(\hat{x}_k^{k-1})}{\partial x}, \quad h_{xx}(\hat{x}_k^{k-1}) = \frac{\partial^2 h(\hat{x}_k^{k-1})}{\partial x^2}.$$

Theorem 2. *Under the assumption of Theorem 1 it holds*

$$\hat{x}_k^k = \begin{cases} \hat{x}_k^{k-1} + B_k g_k(z_k) & \text{for } h_{xx}(\hat{x}_k^{k-1}) = 0, \\ = \hat{x}_k^{k-1} + \frac{h_x(\hat{x}_k^{k-1})}{h_{xx}(\hat{x}_k^{k-1})}\left[B_k^{-1}\frac{\int_{-\infty}^{z_k}\exp(B_k^{-1}u)p(u|Z^{k-1})\,du}{\exp(B_k^{-1}z_k)p(z_k|Z^{k-1})} - 1\right] & \\ & \text{for } h_{xx}(\hat{x}_k^{k-1}) > 0, \\ = \hat{x}_k^{k-1} + \frac{h_x(\hat{x}_k^{k-1})}{h_{xx}(\hat{x}_k^{k-1})}\left[-B_k^{-1}\frac{\int_{z_k}^{-\infty}\exp(B_k^{-1}u)p(u|Z^{k-1})\,du}{\exp(B_k^{-1}z_k)p(z_k|Z^{k-1})} - 1\right] & \\ & \text{for } h_{xx}(\hat{x}_k^{k-1}) < 0, \end{cases} \tag{2.3}$$

where $p(u \mid Z^{k-1})$ denotes the predicted observation density in the integrand and

$$B_k = P_k^{k-1} h_{xx}(\hat{x}_k^{k-1}). \tag{2.4}$$

PROOF : One can write according to Bayes law

$$\hat{x}_k^k - \hat{x}_k^{k-1} = \int_{-\infty}^{\infty} (x_k - \hat{x}_k^{k-1}) p(x_k \mid Z^k)\, dx_k =$$
$$= P_k^{k-1}[p(z_k \mid Z^{k-1})]^{-1} \int_{-\infty}^{\infty} (P_k^{k-1})^{-1}(x_k - \hat{x}_k^{k-1}) p(x_k \mid Z^{k-1}) p(z_k \mid x_k)\, dx_k.$$

Since the density $p(x_k \mid Z^{k-1})$ is Gaussian with mean $\hat{x}_k^{k-1}$ and variance P_k^{k-1}, one has

$$\frac{\partial p(x_k \mid Z^{k-1})}{\partial x_k} = -(P_k^{k-1})^{-1}(x_k - \hat{x}_k^{k-1}) p(x_k \mid Z^{k-1})$$

so that

$$\hat{x}_k^k - \hat{x}_k^{k-1} = -P_k^{k-1}[p(z_k \mid Z^{k-1})]^{-1} \int_{-\infty}^{\infty} \frac{\partial p(x_k \mid Z^{k-1})}{\partial x_k} p(z_k \mid x_k)\, dx_k.$$

Integrating by parts one obtains

$$\hat{x}_k^k - \hat{x}_k^{k-1} = P_k^{k-1}[p(z_k \mid Z^{k-1})]^{-1} \int_{-\infty}^{\infty} \frac{\partial p(z_k \mid x_k)}{\partial x_k} p(x_k \mid Z^{k-1})\, dx_k.$$

Further one can write according to (2.2)

$$\hat{x}_k^k - \hat{x}_k^{k-1} =$$
$$= P_k^{k-1}[p(z_k \mid Z^{k-1})]^{-1} \int_{-\infty}^{\infty} [-h_x(\hat{x}_k^{k-1}) - h_{xx}(\hat{x}_k^{k-1})(x_k - \hat{x}_k^{k-1})] \frac{\partial p(z_k \mid x_k)}{\partial z_k} p(x_k \mid Z^{k-1})\, dx_k$$
$$= -P_k^{k-1} h_x(\hat{x}_k^{k-1})[p(z_k \mid Z^{k-1})]^{-1} \times$$
$$\times \frac{\partial}{\partial z_k}\{p(z_k \mid Z^{k-1}) \int_{-\infty}^{\infty} p(z_k \mid x_k) p(x_k \mid Z^{k-1})[p(z_k \mid Z^{k-1})]^{-1}\, dx_k\} -$$
$$- P_k^{k-1} h_{xx}(\hat{x}_k^{k-1})[p(z_k \mid Z^{k-1})]^{-1} \times$$
$$\times \frac{\partial}{\partial z_k}\{p(z_k \mid Z^{k-1}) \int_{-\infty}^{\infty} (x_k - \hat{x}_k^{k-1}) p(z_k \mid x_k) p(x_k \mid Z^{k-1})[p(z_k \mid Z^{k-1})]^{-1}\, dx_k\} =$$
$$= -P_k^{k-1} h_x(\hat{x}_k^{k-1})[p(z_k \mid Z^{k-1})]^{-1} \frac{\partial p(z_k \mid Z^{k-1})}{\partial z_k} -$$
$$- P_k^{k-1} h_{xx}(\hat{x}_k^{k-1})[p(z_k \mid Z^{k-1})]^{-1} \frac{\partial}{\partial z_k}[p(z_k \mid Z^{k-1})(\hat{x}_k^k - \hat{x}_k^{k-1})]$$

since

$$\int_{-\infty}^{\infty} p(x_k \mid Z^k)\, dx_k = 1, \quad \int_{-\infty}^{\infty} (x_k - \hat{x}_k^{k-1}) p(x_k \mid Z^k)\, dx_k = \hat{x}_k^k - \hat{x}_k^{k-1}.$$

Due to (1.9) and (2.4) we have finally

$$(2.5)\quad \hat{x}_k^k - \hat{x}_k^{k-1} = P_k^{k-1} h_x(\hat{x}_k^{k-1}) g_k(z_k) - B_k[\frac{\partial}{\partial z_k}(\hat{x}_k^k - \hat{x}_k^{k-1}) - g_k(z_k)(\hat{x}_k^k - \hat{x}_k^{k-1})].$$

The relation (2.5) can be looked upon as the differential equation of the first order for the unknown function $\hat{x}_k^k - \hat{x}_k^{k-1}$ of the argument z_k. By solving this differential equation of the type $y' + u(x)y = v(x)$ one obtains (2.3) as its particular solution fulfilling the initial condition

$$(2.6)\qquad E(\hat{x}_k^k - \hat{x}_k^{k-1} \mid Z^{k-1}) = \int_{-\infty}^{\infty} (\hat{x}_k^k - \hat{x}_k^{k-1}) p(z_k \mid Z^{k-1})\, dz_k = 0.$$

For example, if $h_{xx}(\hat{x}_k^{k-1}) > 0$, i.e.$B_k > 0$, one can verify (2.6) in the following way

$$\int_{-\infty}^{\infty} (\hat{x}_k^k - \hat{x}_k^{k-1}) p(z_k \mid Z^{k-1})\, dz_k =$$
$$= \frac{h_x(\hat{x}_k^{k-1})}{h_{xx}(\hat{x}_k^{k-1})}\left[B_k^{-1} \int_{-\infty}^{\infty} \exp(-B_k^{-1} z_k) \int_{-\infty}^{z_k} \exp(B_k^{-1} u) p(u \mid Z^{k-1})\, du - 1\right] =$$
$$= \frac{h_x(\hat{x}_k^{k-1})}{h_{xx}(\hat{x}_k^{k-1})}\left[\int_0^{\infty} \exp(-B_k^{-1} y) p(z_k - y \mid Z^{k-1})\, dy\right]_{z_k=-\infty}^{\infty} = 0,$$

where we have used the integration by parts and the substitution $y \approx z_k - u$. ■

Remark 1. If $h_{xx}(\hat{x}_k^{k-1}) = 0$ then (2.3) reduces to (1.7) from Theorem 1.

Remark 2. Proceeding similarly as in the proof of Theorem 1 one obtains after some tedious manipulations the following differential equation of the second order for P_k^k

$$(2.7)\qquad P_k^k - B_k^2\{\frac{\partial^2}{\partial z_k^2}P_k^k - g_k(z_k)\frac{\partial}{\partial z_k}P_k^k - [G_k(z_k) - (g_k(z_k))^2]P_k^k\} =$$
$$= P_k^{k-1} - (P_k^{k-1})^2\{(h_x(\hat{x}_k^{k-1}))^2[G_k(z_k) - (g_k(z_k))^2] - h_{xx}(\hat{x}_k^{k-1})g_k(z_k) -$$
$$-2h_x(\hat{x}_k^{k-1})h_{xx}(\hat{x}_k^{k-1})[p(z_k \mid Z^{k-1})]^{-1}\frac{\partial^2}{\partial z_k^2}[p(z_k \mid Z^{k-1})(\hat{x}_k^k - \hat{x}_k^{k-1})] -$$
$$-(h_x(\hat{x}_k^{k-1}))^2[p(z_k \mid Z^{k-1})]^{-1}\frac{\partial^2}{\partial z_k^2}[p(z_k \mid Z^{k-1})(\hat{x}_k^k - \hat{x}_k^{k-1})^2]\} - (\hat{x}_k^k - \hat{x}_k^{k-1})^2.$$

The explicit formula for P_k^k is very complicated. For simplicity one can accept the formula (1.8) from Theorem 1 which is the special case of (2.7) for $h_{xx}(\hat{x}_k^{k-1}) = 0$.

References

[1] Anderson B.D.O, Moore J.B., *Optimal Filtering*, Prentice–Hall, Englewood Cliffs, New Jersey, 1979.

[2] Jazwinski A.H., *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[3] Masreliez C.J., *Approximate non-Gaussian filtering with linear state and observation relations*, IEEE Transactions on Automatic Control **AC–20** (1975), 107–110.

[4] Martin R.D., *Robust estimation of time series autoregressions*, in: Robustness in Statistics (R.L. Launer and G. Wilkinson, eds.), Academic Press, New York, 1979.

[5] Stockinger N., Dutter R., *Robust time series analysis: A survey*, Supplement to the journal Kybernetika **23** (1987).

Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 00 Praha 8, Czechoslovakia

(*Received* May 17, 1990)