Density-based Nonlinear Estimation

Sum of Gaussians Filter

Dynamic Stochastic Filtering, Prediction and Smoothing

PhD Program in Electrical Engineering and Computers

Spring Semester 2009/2010

Author:
Bruno Gil 54222/D

Description of the system:

$$x_{k+1} = f_{k+1}(x_k, u_k) + w_k \quad , \quad k = 0,1,2,...$$
$$z_k = h_k(x_k) + v_k \quad , \quad k = 0,1,2,...$$

Filtering:
The aim is to find the probability density function (*pdf*) of the state x_k conditioned by the measurements $z^k = \left[z_0, z_1, ..., z_k \right]$:

$$p(x_k \mid z^k) = ?$$

One-step prediction:
The aim is to find the *pdf* of the state at the next time instant *k+1* conditioned by the measurements $z^k = \left[z_0, z_1, ..., z_k \right]$ and the evolution of the plant dynamics:

$$p(x_{k+1} \mid z^k) = ?$$

The solution of filtering/one-step prediction problem is given by the bayesian recursive relations:

$$p(x_k \mid z^k) = \frac{p(x_k \mid z^{k-1})\, p(z_k \mid x_k)}{\int p(x_k \mid z^{k-1})\, p(z_k \mid x_k)\, dx_k}$$

$$p(x_{k+1} \mid z^k) = \int p(x_k \mid z^k)\, p(x_{k+1} \mid x_k)\, dx_k$$

where $p(x_0 \mid z^{-1}) = p(x_0)$

(proof)

$$\boxed{p(x_k \mid z^k) = \frac{p(z_k \mid x_k)\, p(x_k \mid z^{k-1})}{p(z_k \mid z^{k-1})}}$$

Bayes' Rule

$$\boxed{p(x_k \mid z^{k-1}, z^{k-2}, \ldots, z^{-1}) = p(x_k \mid z^{k-1})}$$

Markov Property

(proof)

$$p(x_k \mid z^k) = \frac{p(z_k \mid x_k)p(x_k)}{p(z^k)} = \frac{p(z_k, z^{k-1} \mid x_k)p(x_k)}{p(z_k \mid z^{k-1})} = \frac{p(z_k \mid z^{k-1}, x_k)p(z^{k-1} \mid x_k)p(x_k)}{p(z_k \mid z^{k-1})p(z^{k-1})} =$$

$$= \frac{p(z_k \mid x_k)p(x_k \mid z^{k-1})p(z^{k-1})p(x_k)}{p(z_k \mid z^{k-1})p(z^{k-1})p(x_k)} = \frac{p(z_k \mid x_k)p(x_k \mid z^{k-1})}{p(z_k \mid z^{k-1})}$$

1) $p(z_k, x_k \mid z^{k-1}) = p(z_k \mid x_k, z^{k-1})p(x_k \mid z^{k-1}) = p(z_k \mid x_k)p(x_k \mid z^{k-1})$

$$\boxed{p(z_k \mid z^{k-1}) = \int p(z_k \mid x_k)p(x_k \mid z^{k-1})dx_k}$$

2) time update: $p(x_{k+1}, x_k \mid z^k) = p(x_{k+1} \mid x_k, z^k)p(x_k \mid z^k) = p(x_{k+1} \mid x_k)p(x_k \mid z^k)$

$$\boxed{p(x_{k+1} \mid z^k) = \int p(x_{k+1} \mid x_k)p(x_k \mid z^k)dx_k}$$

Facts:

. $p(x_k \mid z^k)$ gives the most complete description possible of x_k, i.e., no more complete description is possible.

. The introduction of an estimation criterion reduces the state information to a finite collection of numbers, yielding an incomplete/inadequate description of x_k.

. The EXCEPTION occurs when the system is linear and gaussian.

$p(x_k \mid z^k)$ is also gaussian, so x_k is completely defined by 2 parameters: $\begin{cases} E\left[x_k \mid z^k\right] \\ E\left[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T \mid z^k\right] \end{cases}$

. Except when the system dynamics is linear and the *a piori* distributions are gaussian, it is generally impossible to determine $p(x_k \mid z^k)$ in a closed form.

Suboptimal State Estimation

Gaussian Sum Approximations

. Originally proposed by Sorenson and Alspach [1971].

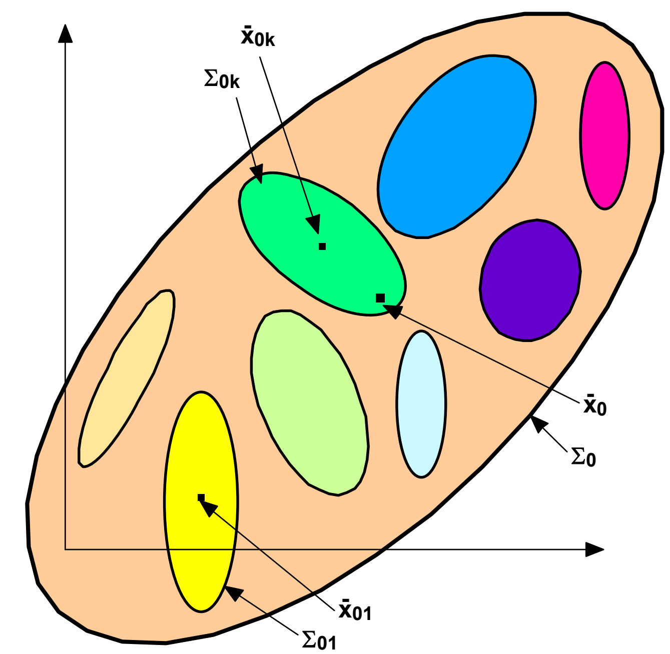
. The goal was to overcome the performance shortcomings and possible divergence of the Extended Kalman Filter.

Fact: EKF performs well if the initial state-vector covariance is small so the true-state is "near" its mean.

Approach: GS approximate the "large" uncertainty of the system with a set of EKFs, each characterised by a *pdf* having a different mean and a "small" covariance.

↓

One expects that one of the EKFs with its mean close to the true state should "work".



This approximation makes $p(x_k \mid z^k)$ more tractable since it is impossible to obtain exact representations of the *a posteriori* density for nonlinear systems.

Gaussian Sum Approximations

The gaussian sum p_A of a density function p associated with a vector-valued random variable x is defined as:

$$p_A(x) = \sum_{i=1}^{L} \alpha_i N(x - a_i, P_i)$$

with $\sum_{i=1}^{L} \alpha_i = 1 \quad , \quad \alpha_i \geq 0 \qquad \forall i$

Properties of GS:

- $\lim\limits_{\substack{l \to \infty \\ P_i \to 0}} p_A$ converges uniformly to any density function of practical concern;
- $\bar{x} = E\{x\} = \int_{-\infty}^{+\infty} x p(x) dx \approx \int_{-\infty}^{+\infty} x \sum_{i=1}^{L} \alpha_i N(x - a_i, P_i) dx = \sum_{i=1}^{L} \alpha_i \int_{-\infty}^{+\infty} x N(x - a_i, P_i) dx = \sum_{i=1}^{L} \alpha_i a_i$
- $\Sigma = \sum_{i=1}^{L} \alpha_i \left[P_i + (\bar{x} - a_i)(\bar{x} - a_i)' \right]$

Gaussian Sum Filter for Nonlinear Systems

Basic Idea: to use the GS representation of the *a posteriori* density function in conjunction with the linearization procedure that has proven so effective in Kalman filter applications.

$$p(x_k \mid z^{k-1}) = \sum_{i=1}^{\varepsilon_k^{'}} \alpha_{ki}^{'} N(x_k - a_{ki}, P_{ki}^{'})$$

(from BRRs)

(from measurement equation)

$$p(x_k \mid z^k) = c_k \left[\sum_{i=1}^{\varepsilon_k^{'}} \alpha_{ki}^{'} N(x_k - a_{ki}, P_{ki}^{'}) p\left(\boxed{z_k - h_k(x_k)}\right) \right]$$

As in the EKF, one linearizes h_k relative to a_{ki}, so that $p(z_k - h_k(x_k))$ can be approximated by a gaussian in the region around a_{ki} :

$$p(x_k \mid z^k) = c_k \left[\sum_{i=1}^{\varepsilon_k^{'}} \alpha_{ki}^{'} N(x_k - a_{ki}, P_{ki}^{'}) N(\xi_i, \boxed{R_k}) \right]$$

(from sensor disturbances ~ N(0,Rk))

where $\xi_i = z_k - h_k(a_{ki}) - H_{ki}\left(x_k - a_{ki}\right)$

$$H_{ki} = \left.\frac{\partial h_k}{\partial x_k}\right|_{a_{ki}}$$

Gaussian Sum Filter for Nonlinear Systems

Last expression reduces to:

$$p(x_k \mid z^k) = \sum_{i=1}^{\varepsilon_k} \alpha_{ki} N(x_k - y_{ki}, P_{ki})$$

with
$$y_{ki} = a_{ki} + K_{ki}\left(z_k - h_k(a_{ki})\right)$$

$$P_{ki} = P'_{ki} - K_{ki} H_{ki} P'_{ki}$$

$$K_{ki} = P'_{ki} H_{ki}^T \left[H_{ki} P'_{ki} H_{ki}^T + R_k \right]^{-1}$$ (filter gains)

$$\alpha_{ki} = \frac{\alpha'_{ki} \beta_{ki}}{\sum_{j=1}^{\varepsilon'_k} \alpha'_{kj} \beta_{kj}}$$

$$\beta_{kj} = N(z_k - h_k(a_{ki}), H_{ki} P'_{ki} H_{ki}^T + R_k)$$

The parameters above are obtaining using the Extended Kalman Filter equations.

This way, GS representation is formed as the convex combination of the output of several Kalman filters operating in parallel.

... and now, how to obtain $p(x_{k+1} \mid z^k)$ for GS filter?

$$p(x_{k+1} \mid z^k) = \int \boxed{p(x_k \mid z^k)} \boxed{p(x_{k+1} \mid x_k)} dx_k$$

$$x_{k+1} = f_{k+1}(x_k, u_k) + w_k$$

<u>Two possible scenarios:</u>

1) when there is little or no plant noise:

$$w_k \sim N(0, Q_k) \quad , \quad Q_k \leq P_{ki}$$

2) when there is a significant amount of plant noise:

$$p(w_k) = \sum_{l=1}^{q_k} \gamma_k N(w_k - \omega_{kl}, Q_{kl})$$

(non-normal disturbances included)

1st scenario:

. Covariance of the plant noise comparable to P_{ki}

. So, linearize the plant equation f_{k+1} relative to the mean values of the gaussian sum y_{ki}:

$$p(x_{k+1} \mid x_k) = \sum_{i=1}^{\varepsilon_k} N(x_{k+1} - f_{k+1}(y_{ki}) - F_{(k+1)i}(x_k - y_{ki}), Q_k)$$

$$\Big\downarrow$$

$$p(x_{k+1} \mid z^k) = \sum_{i=1}^{\varepsilon'_{(k+1)}} \alpha'_{(k+1)i} N(x_{k+1} - a_{(k+1)i}, P_{(k+1)i})$$

with $a_{(k+1)i} = f_{k+1}(y_{ki})$

$$P'_{(k+1)i} = F_{(k+1)i} P_{ki} F^T_{(k+1)i} + Q_k$$

$$\alpha'_{(k+1)i} = \alpha_{ki}$$

$$\varepsilon'_{(k+1)} = \varepsilon_k$$

$$F'_{(k+1)} = \left. \frac{\partial f_{k+1}}{\partial x_k} \right|_{y_{ki}}$$

. Important remark: the number of terms in the sum do not increase.

2nd scenario:

The large plant noise will increase the variance of each term in the GS, creating conditions for the overlap of the individual terms.

<u>Consequences:</u>

. linearizations are no longer valid;

. the next measurement will cause the various terms to have nearly the same mean.

<u>Solution:</u>

. $p(w_k)$ is approximated by a gaussian sum;

. recalculate the term $p(x_{k+1} \mid z^k)$

$$p(w_k) = \sum_{l=1}^{q_k} \gamma_k N(w_k - \omega_{kl}, Q_{kl})$$

$$p(x_{k+1} \mid x_k) = \sum_{i=1}^{q_k} N(x_{k+1} - f_{k+1}(y_{ki}) - F_{(k+1)i}(x_k - y_{ki}) - \omega_{kl}, Q_{kl})$$

$$p(x_{k+1} \mid z^k) = \sum_{i=1}^{\varepsilon'_{(k+1)}} \alpha'_{(k+1)i} N(x_{k+1} - a_{(k+1)i}, P_{(k+1)i})$$

with $a_{(k+1)i} = f_{k+1}(y_{ki}) + w_{kl}$

$\varepsilon'_{(k+1)} = \boxed{\varepsilon_k q_k}$ the number of terms in the sum has increased

$\alpha'_{(k+1)i} = \alpha_{ki}\gamma_{kl}$

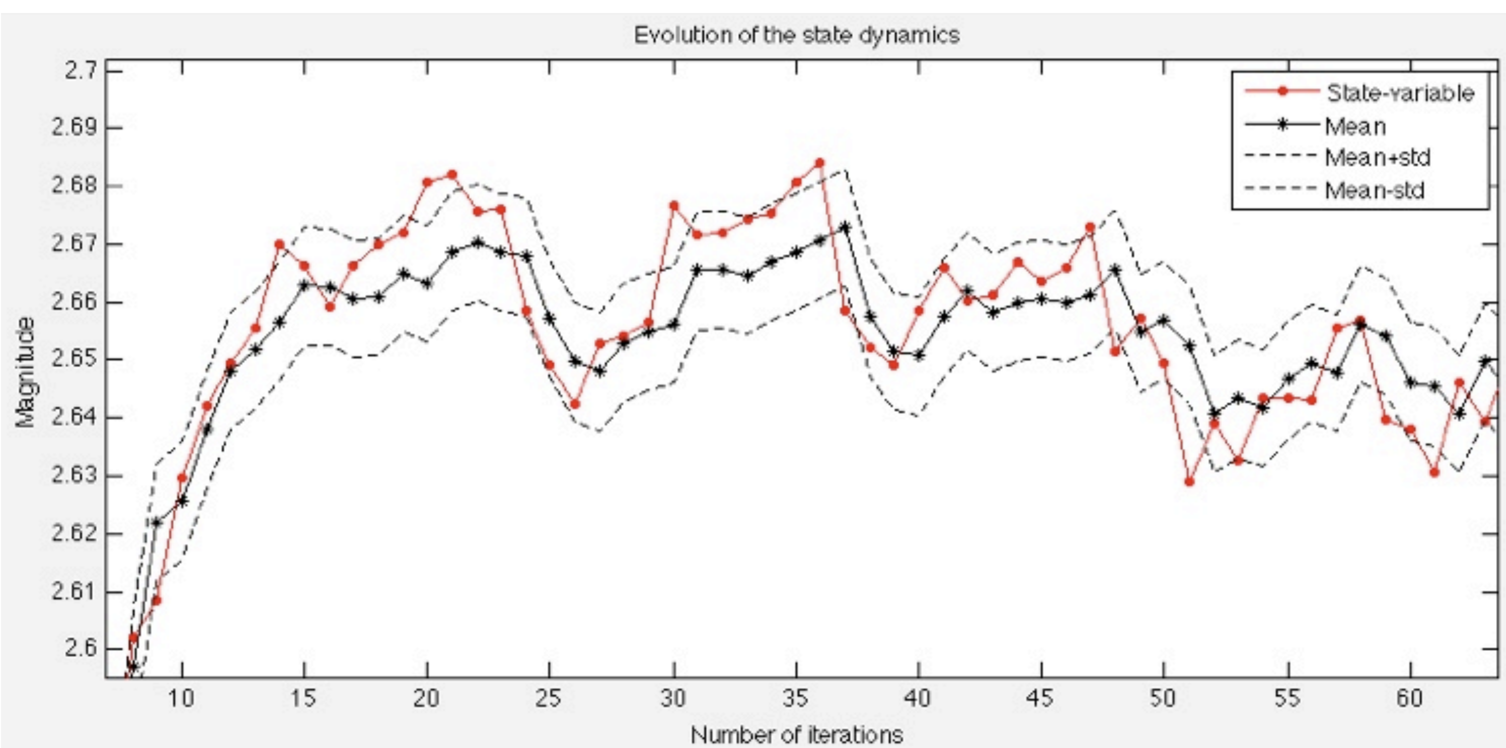
$P'_{(k+1)i} = F_{(k+1)i} P_{ki} F^T_{(k+1)i} + Q_{kl}$

Example I

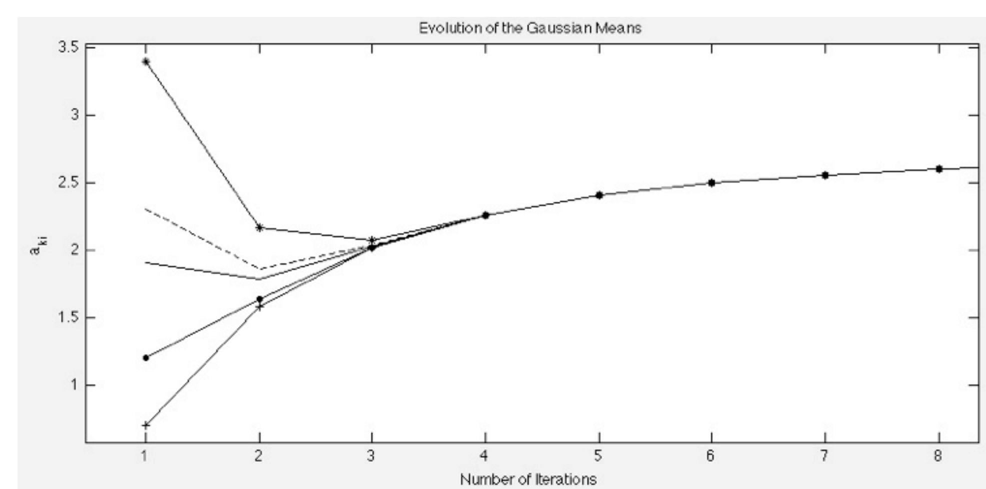
Consider the following nonlinear system:

$$\begin{cases} x_{k+1} = 0{,}5x_k + 1 + \sin(0{,}04\pi x_k) + w_k & , \quad w \sim N(0;10^{-3}) \\ z_k = x_k^2 + v_k & , \quad v \sim N(0;10^{-3}) \end{cases}$$

$$p(x_0 \mid z^{-1}) = 0{,}2N(1{,}2;0{,}02) + 0{,}3N(0{,}7;0{,}01) + 0{,}1N(3{,}4;0{,}02) + 0{,}1N(2{,}3;0{,}04) + 0{,}3N(1{,}9;0{,}01)$$



Gaussian	Mean	Covariance	Alpha
1.0000	2.6522	0.0101	0.8933
2.0000	2.6522	0.0101	0.1067
3.0000	2.6522	0.0101	0.0000
4.0000	2.6522	0.0101	0.0000
5.0000	2.6522	0.0101	0.0000



Example II

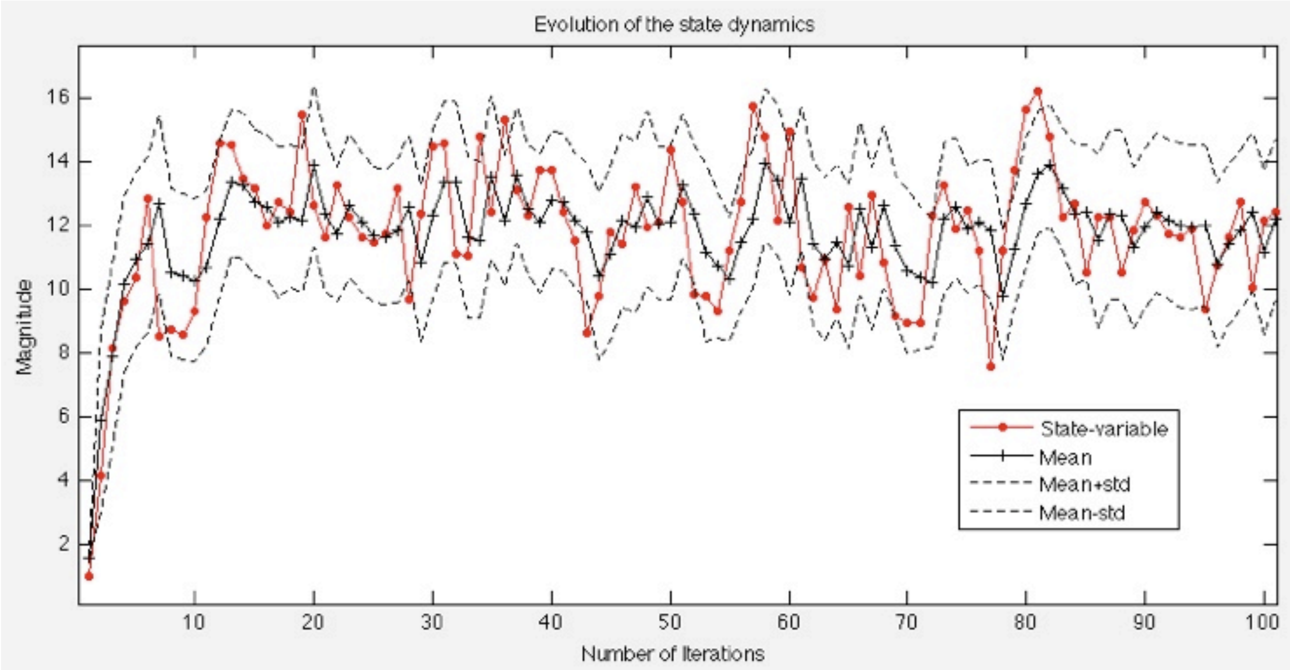
The same system with a Gaussian sum for the noise plant:

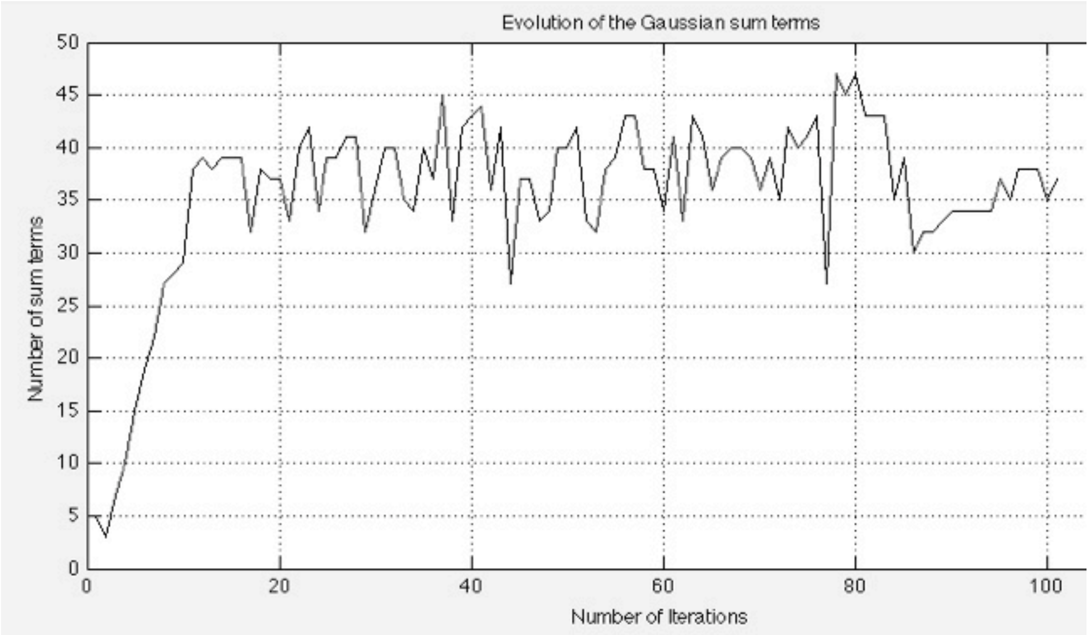
$$\begin{cases} x_{k+1} = 0,5x_k + 1 + \sin(0,04\pi x_k) + w_k \\ z_k = x_k^2 + v_k \end{cases}$$

$$p(w_k) = 0,29N(2,14;0,72) + 0,18N(7,45;8,05) + 0,53N(4,31;2,29)$$

$$p(v_k) = N(0;10^{-5})$$

$$p(x_0 \mid z^{-1}) = 0,2N(1,2;0,02) + 0,3N(0,7;0,01) + 0,1N(3,4;0,02) + 0,1N(2,3;0,04) + 0,3N(1,9;0,01)$$





Conclusions

. GSF results in the parallel operation of several EKF: there are as many individual filters as there are terms in the Gaussian sum.

. The computational overhead of the GSF can be significantly greater than those of the EKF.

. Nonetheless, the filtering estimates of the GSF are still less biased than those of the EKF.

. There are some theoretical bounds on the errors for the approximations made in $p(x_k \mid z^k)$ and $p(x_{k+1} \mid z^k)$. However, they are hard to apply to a specific case.

$\downarrow$

One needs to rely on some *ad hoc* rules to keep the measurement residual consistent with its theoretical properties.

. When inconsistency occurs, one proceeds by reinitialising the filter parameters ($\downarrow P_{ki}^{'}$ or extra terms in the sum)

References

[1] H. Tanizaki, *Nonlinear Filters - Estimation and Applications*, Springer Verlag, 1996.

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

[3] B. D. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall Inc., 1979.

[4] H. W. Sorenson and D. L. Alspach, *Recursive Bayesian Estimation Using Gaussian Sums*, Automatica, Vol. 7, July 1971, pp. 465-479.

[5] D. L. Alspach and H. W. Sorenson, *Nonlinear Bayesian Estimation Using Gaussian Sum Approximations*, IEEE Trans. Auto. Control, Vol. AC-17, August 1972, pp. 439-448.

[6] T. Schön, *On Computational Methods for Nonlinear Estimation*, Linköping Studies in Science and Technology, Thesis No. 1047, 2003.

[7] N. Bergman, *Recursive Bayesian Estimation - Navigation and Tracking Applications*, Linköping Studies in Science and Technology, Dissertations No. 579, 1999.