Proceedings of the 39th IEEE Conference on Decision and Control
Sydney, Australia • December, 2000

Gaussian Filter for Nonlinear Filtering Problems

Kazufumi Ito *

Abstract

In this paper we develop and analyze real-time and accurate filters for nonlinear filtering problems based on the Gaussian distributions. We present the systematic formulation of Gaussian filters and develop efficient and accurate numerical integration of the proposed filter. We also discuss the mixed Gaussian filters in which the conditional probability density is approximated by the sum of Gaussian distributions. Our numerical testings demonstrate that new filters significantly improve the extended Kalman filter with no additional cost and the new Gaussian sum filter has a nearly optimal performance.

1 Introduction

In this paper we develop a class of Gaussian filters for nonlinear filtering based on the Kushner equation. The optimal nonlinear filter equation is described by the so-called Kushner equation, that is the conditional probability density function is governed by a nonlinear stochastic PDEs driven by a noisy observation. There has been increasing researches in developing robust and efficient numerical integrations of the Kushner equation as well as the Zakai equation. The Zakai equation is linear and mathematically equivalent to the Kushner equation by the change of probability measure. A relationship between the numerical integrations of Zakai and Kushner equations is discussed in [5]. Applications of such numerical methods have been successfully tested and demonstrated the superiority of the nonlinear filter comparing to, for example the extended Kalman-Bucy filter. However they are limited to a relatively lower dimensional signal process.

The objective of this paper is to develop and analyze a nonlinear filter based on sum of Gaussian distributions. An efficient implementation of of the proposed filter is developed based on the quadrature rule discussed in [6]. Our proposed filter greatly improves the performance of the extended Kalman-Bucy filter without increasing much of computational complexities. We discuss the nonlinear filtering problem for the continuous-time signal system for R^n-valued diffusion process $x(t)$. But our proposed method can be extended to the signal process with hidden Markov chains.

We will present an application of the proposed filter including the target detection and tracking using IR measurements and report numerical findings.

The following are an outline of the paper. In Section 2 we derive our proposed Gaussian filter. In Section 3 we analyze the stability and asymptotic behavior of the Gaussian filter. In Section 4 we discuss the mixed Gaussian filter. Finally, we discuss the implementation of the proposed filter in Section 5.

We formulate the optimal filter in a recursive form. Our proposed Gaussian filter is based on an approximation of the recursive form based on Gaussian distributions. We consider the continuous-time signal system for R^n-valued diffusion process $x(t)$:

$$dx(t) = f(x(t))\,dt + \sigma(x(t))\,dw_1(t), \quad x(0) = x_0 \quad (1.1)$$

and the observation process $y(t) \in R^p$ is of the form

$$y(t) = \int_0^t h(x(s))\,ds + w_2(t) \quad (1.2)$$

where w_1 and w_2 are Brownian motions. In addition, it is assumed that w_2 is independent of $x(t)$, and x_0 is a random variable with the density function $p_0 \in L^2(R^n)$. Throughout what follows $f : R^n \to R^n$, $\sigma : R^n \to R^{n\times n}$ and $h : R^n \to R^p$ are bounded continuous functions and f and σ are also Lipschitz continuous.

The diffusion processes $(x(t), y(t))$ are considered on a complete probability space $(\Omega, \mathcal{F}, P)$. Let us denote by $\mathcal{F}_t^y$ the P-completed σ-field generated by the observations $\{y(s), 0 \le s \le t\}$. It is a standard fact that for a bounded function ϕ, the best mean square estimator of $\phi(x(t))$ based on the observations $\{y(s), 0 \le s \le t\}$ is given by $\pi[\phi] := E[\phi(x(t))|\mathcal{F}_t^y]$. Moreover, a fundamental result of filtering theory (see e.g.[2], [8]) says that if $\phi \in C_b^2(R^n)$, then $\pi_t[\phi]$ satisfies the stochastic differential equation

$$\begin{aligned} d\pi_t[\phi] + \pi_t[A^*\phi]dt \\ = (\pi_t[h\phi] - \pi_t[\phi]\pi_t[h])(dy(t) - \pi_t[h]dt) \end{aligned} \quad (1.3)$$

*Dept. of Mathematics, North Carolina State University, Raleigh, NC 27695–8205, USA, kito@eos.ncsu.edu,Research partially supported by ONR grant N00014-96-1-1-0265, DARPA grant F30602-99-2-0548

0-7803-6638-7/00$10.00

where the generator A^* is the formal adjoint operator to A

$$-A\phi(x) = \frac{\partial}{\partial x_i}\left(a_{i,j}(x)\frac{\partial}{\partial x_j}\phi(x)\right) - \frac{\partial}{\partial x_i}\left(a_i(x)\phi(x)\right).$$

with

$$a_{i,j} = (\frac{1}{2}\sigma\sigma^t)_{i,j} \quad \text{and} \quad a_i = f_i - \frac{\partial}{\partial x_j}a_{i,j}.$$

If the measure $\pi_t(dx) = E[\xi(x(t) \in dx)|\mathcal{F}_t^y]$ admits a smooth density $\pi(t,x)$ with respect to the Lebesgue measure, it is easy checked that $\pi(t,x)$ satisfies the Kushner equation

$$d\pi(t) + A\pi(t)\,dt = (h - \pi_t[h])\pi(t)\,(dy(t) - \pi_t[h]\,dt) \tag{1.4}$$

where $\pi(0,x) = p_0(x)$ and $\pi_t[h] = \int_{R^n} h(x)\pi(t,x)\,dx$.

From (1.3) and (1.4) the optimal state estimate $\hat{x}(t) = E[x(t)|\mathcal{F}_t^y$ satisfies

$$d\hat{x}(t) = \widehat{f(x)}(t)\,dt + L(t)(dy(t) - \widehat{h(x)}(t)\,dt) \tag{1.5}$$

where

$$\widehat{f(x)}(t) = \int_{R^n} f(x)\pi(t,x)\,dx,\ \widehat{h(x)}(t) = \int_{R^n} h(x)\pi(t,x)\,dx$$

and

$$L(t) = \int_{R^n}(x - \hat{x}(t))(h(x) - \widehat{h(x)}(t))^t\pi(t,x)\,dx.$$

2 Gaussian Filter

We consider a Gaussian approximation of the optimal nonlinear filter by approximating π by the normal distribution $N(\hat{x}(t), P(t))$

$$\pi(t,x) \simeq N(t,x) = N(\hat{x}(t), P(t))(x)$$

$$= \frac{1}{((2\pi)^n\, det P(t))^{1/2}}\, e^{-\frac{1}{2}(x-\hat{x}(t))^t P(t)^{-1}(x-\hat{x}(t))}.$$

Then we obtain the filter equation

$$d\hat{x}^{(1)}(t) = \widehat{f(x)}^{(1)}(t)dt + L^{(1)}(t)(dy(t) - \widehat{h(x)}^{(1)}(t)dt) \tag{2.1}$$

with

$$\widehat{f(x)}^{(1)}(t) = \int_{R^n} f(x)N(\hat{x}^{(1)}(t), P(t))\,dx,$$
$$\widehat{h(x)}^{(1)}(t) = \int_{R^n} h(x)N(\hat{x}^{(1)}(t), P(t))\,dx \tag{2.2}$$

and

$$L^{(1)}(t) = \int_{R^n}(x - \hat{x}^{(1)}(t))(h(x) - \widehat{h(x)}^{(1)}(t))^t)N(t,x)\,dx. \tag{2.3}$$

In order to obtain the update of $P(t)$ we observe

$$d(x(t) - \hat{x}(t) = (f(x(t)) - \widehat{f(x)}(t)))\,dt$$
$$-L^{(1)}(t)(h(x(t)) - \widehat{h(x)}(t))\,dt$$
$$+\sigma(x(t))\,dw_1(t) - L(t)\,dw_2(t).$$

By Ito's rule

$$d(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))^t$$
$$= d(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))^t$$
$$+(x(t) - \hat{x}(t))d(x(t) - \hat{x}(t))^t$$
$$+(\sigma(x(t))\sigma(x(t))^t + L(t)L(t)^t)\,dt.$$

Thus

$$\Phi(t) = E[(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))^t|\mathcal{F}_t^y]$$
$$= \int_{R^n}(x - \hat{x}(t))(x - \hat{x}(t))^t\pi(t,x)\,dx$$

satisfies

$$\Phi(t) = \Phi(0)$$
$$+ \int_0^t [\int_{R^n}\{(f(x) - \hat{f}(t)) - L(t)(h(x) - \widehat{h(x)}(t))\}(x - \hat{x}(t))^t$$
$$+(x - \hat{x}(t))\{(f(x) - \hat{f}(t)) - L(t)(h(x) - \hat{h}(t))\}^t\}\pi(t,x)\,dx$$
$$+ \int_{R^n}(\sigma(x) - \hat{\sigma}(t))(\sigma(x) - \hat{\sigma}(t))^t\pi(t,x)\,dx$$
$$+\hat{\sigma}(t)\hat{\sigma}(t)^t + L^{(1)}(t)L^{(1)}(t)^t]\,dt. \tag{2.4}$$

For the simplicity of our discussion we assume $\sigma(x) = \sigma$ and set $\Sigma = \sigma\sigma^t$. Now we take the Gaussian normalization by replacing $\pi(t,x)$ by the Gaussian approximation $N(\hat{x}^{(1)}(t), P(t)),\ P(t) = P^{(1)}(t)$ and obtain

$$\frac{d}{dt}P^{(1)}(t) = F^{(1)}(\hat{x}^{(1)}(t), P^{(1)}(t)) - L^{(1)}(t)L^{(1)}(t)^t + \Sigma \tag{2.5}$$

where

$$F^{(1)}(\hat{x}^{(1)}, P) = \int_{R^n}[(f(x) - \widehat{f(x)}^{(1)})(x - \hat{x}^{(1)})^t$$
$$+(x - \hat{x}^{(1)}(t))(f(x) - \widehat{f(x)}^{(1)})^t]N(\hat{x}^{(1)}, P)\,dx \tag{2.6}$$

Note that if $f(x) = Ax$, $h(x) = Hx$ and $p_0 = N(x_0, P_0)$ then (2.5) reduces to the Riccati equation

$$\frac{d}{dt}P(t) = AP(t) + P(t)A^t - P(t)H^tHP(t) + \Sigma$$

i.e.,

$$F^{(1)}(\hat{x}^{(1)}, P) = AP + PA^t \quad \text{and} \quad L^{(1)}(t) = P(t)H^t$$

Note that $\widehat{f(x)}^{(1)}(t)$ and $\widehat{h(x)}^{(1)}(t)$ depend on $P(t)$ and thus (2.1) and (2.5) are coupled. In order to the filter equation (2.1) separated from the covariance update $P(t)$ we propose the following variant.

$$d\hat{x}^{(2)} = f(x^{(2)}(t))\,dt + L^{(2)}(t)(dy(t) - h(\hat{x}^{(2)}(t))\,dt)$$

$$\frac{d}{dt}P^{(2)}(t) = F^{(2)}(\hat{x}^{(2)}(t), P^{(2)}(t)) - L^{(2)}(t)L^{(2)}(t)^t + \Sigma \tag{2.7}$$

where

$$L^{(2)}(t) = \int_{R^n} (x - \hat{x}^{(2)}(t))(h(x) - h(\hat{x}^{(2)}(t)))^t N\,dx$$

$$F^{(2)}(\hat{x}^{(2)}, P) = \int_{R^n} [(f(x) - f(\hat{x}^{(2)}))(x - \hat{x}^{(2)})^t$$

$$+(x - \hat{x}^{(2)})(f(x) - f(\hat{x}^{(2)}))^t] N\,dx \tag{2.8}$$

with $N = N(\hat{x}^{(2)}(t), P^{(2)}(t))(x)$.

3 Stability and Performance Bound

In this section we analyze the stability of the Gaussian filters (2.1)-(2.6) and (2.7)–(2.8). We assume that

$$(f(x_1) - f(x_2), x_1 - x_2) \le \omega\,|x_1 - x_2|^2 \tag{3.1}$$

for $x_1, x_2 \in R^n$. Then we have the following theorem.

Theorem 3.1 There exists a unique path-wise non-negative symmetric solution $P^{(2)}(t)$ to equation

$$\begin{aligned} \tfrac{d}{dt}P^{(2)}(t) &= \Phi(x^{(2)}, P^{(2)}) \\ &= F^{(2)}(\hat{x}^{(2)}(t), P^{(2)}(t)) - L^{(2)}(t)L^{(2)}(t)^t + \Sigma \end{aligned} \tag{3.2}$$

with $P^{(2)}(0) = P_0 \ge 0$.

Proof: Let $S = \{P \in R^{n\times n} : P = P^t \text{ and } P \ge 0\}$. First note that since h is Lipschitz continuous,

$$|L^{(2)}| \le M_1 \int_{R^n} |s|^2 N(0, P)\,ds \le M_1\,|P^{(2)}|$$

for $P \in S$ and some $M_1 > 0$ independent of $x^{(2)}$. Also, it can be proved that Φ is locally Lipschitz continuous uniformly on S in $x^{(2)}$. Assume that $P_0 \ge \epsilon\, I$ and $\Sigma \ge \epsilon\, I$ for some $\epsilon > 0$. Thus there exits a locally defined unique solution P in S. We show that there exists $\delta = \delta(\epsilon)$ such that $P(t) \ge \delta\, I$. Note that there exists a constant $\delta > 0$ such that $\Phi(x^{(2)}, P) \ge 0$ for all $x^{(2)}$ and $P \in S$ satisfying $|P| \le \delta$, since $\Phi(x, 0) = \Sigma > 0$. Thus $\frac{d}{dt}P \ge 0$ on $S \cap \{|P| \le \delta\}$ and hence $P(t) \ge \delta > 0$. Since $\epsilon > 0$ is arbitrary, it follows that $P(t) \in S$.

Next we establish a priori bound of $P(t)$ of solutions to (3.2). Since

$$tr\, F^{(2)}(x^{(2)}, P) = 2\int_{R^n} (f(x) - f(\hat{x}^{(2)}), x - \hat{x}^{(2)}) N(\hat{x}^{(2)}, P)\,dx$$

$$\le 2\omega \int_{R^n} |x - \hat{x}^{(2)}|^2 N(\hat{x}^{(2)}, P)\,dx = 2\omega\, tr\, P,$$

it follows from (2.8) that

$$\frac{d}{dt} tr\, P(t) \le 2\omega\, tr\, P(t) + tr\, \Sigma.$$

By Gronwall's inequality

$$tr\, P(t) \le e^{2\omega t} tr\, P(0) + \int_0^t e^{2\omega\,(t-s)}\, tr\, \Sigma\, ds. \square$$

Suppose there exists a bounded continuous function $L : \hat{x} \in R^n \to L(\hat{x}) \in R^{n\times p}$ such that

$$(f(x) - f(\hat{x}) - L(\hat{x})(h(x) - h(\hat{x})), x - \hat{x}) \le -\omega_0\, |x - \hat{x}|^2 \tag{3.3}$$

for $\omega_0 > 0$ and x, $\hat{x} \in R^n$. Then we have the following bound.

Theorem 3.2 Assume that assumption (3.3) holds. Then the solution $P^{(2)}(t)$ satisfies the bounded

$$tr\, P(t) \le \frac{tr\, \Sigma + \|L\|^2}{2\omega_0}. \tag{3.4}$$

Proof: Note that

$$\Phi(\hat{x}, P) = \int_{R^n} [(f(x) - f(\hat{x}) + L(\hat{x})(h(x) - h(\hat{x})))(x - \hat{x})^t$$

$$+(x - \hat{x})(f(x) - f(\hat{x}) - L(\hat{x})(h(x) - h(\hat{x})))^t] N(\hat{x}, P)\,dx$$

$$-MM^t + L(\hat{x})L(\hat{x})^t + \Sigma$$

where

$$M = L(\hat{x}) - \int_{R^n} (x - \hat{x})(h(x) - h(\hat{x}))^t N(\hat{x}, P)\,dx.$$

Using the same arguments as in the proof of Theorem 3.1 we obtain the desired bound. □

We can show the stability result of the Gaussian filter (2.1)–(2.6) as follows.

Theorem 3.3 There exists a unique path-wise non-negative symmetric solution $P^{(1)}(t)$ to equation

$$\frac{d}{dt}P^{(1)}(t) = F^{(1)}(\hat{x}^{(1)}(t), P^{(1)}(t)) - L^{(1)}(t)L^{(1)}(t)^t + \Sigma \tag{3.5}$$

with $P^{(1)}(0) = P_0 \geq 0$. Assume that for $L \in R^{n\times p}$

$$(f(x)-f(s)-L(h(x)-h(s)), x-s) \leq -\omega_0\,|x-s|^2. \quad (3.6)$$

Then the solution $P^{(1)}(t)$ to (3.5) satisfies (3.4).

Proof: The proof of the theorem is exactly same as those for Theorems 3.1–3.2 observing

$$\int_{R^n} (f(x) - \widehat{f(x)}^{(1)})^t(x - \hat{x}^{(1)})N(\hat{x}^{(1)}, P)(x)\,dx$$

$$= \frac{1}{2}\int_{R^n\times R^n} (f(x) - f(s))(x-s)^t N(x)N(s)\,dxds. \square$$

Similarly we have the stability result for the optimal filter (1.5).

Corollary 3.4 If assumption (3.6) holds, then

$$E[|x(t) - \hat{x}(t)|^2|\mathcal{F}_t^y] \leq \frac{tr\,\Sigma + \sup\,tr\,LL^t}{2\omega_0}.$$

Proof: Using the same arguments that lead to (2.4), we have $\Psi(t) = E[|x(t) - \hat{x}(t)|^2|\mathcal{F}_t^y]$ satisfies

$$\Psi(t) \leq \Psi(0) + tr\,(LL^t + \Sigma)$$

$$+\frac{1}{2}\int_{R^n\times R^n} [(f(x) - f(s)) - L(h(x) - h(s)]^t(x-s)$$
$$\times\pi(t,x)\pi(t,s)\,dxds.$$

The remaining of the proof is similar to those for Theorem 3.2. $\square$

Finally, we discuss the performance of the (Gaussian) filter of the form (2.7).

Theorem 3.5 Consider the filter of the form

$$d\hat{x}(t) = f(\hat{x}(t))\,dt + L(t,\hat{x})(dy(t) - h(\hat{x}(t))\,dt)$$

where $L : R^+ \times R^n \to R^{p\times n}$ is a bounded Lipschitz continuous function. Assume that

$$(f(x) - f(\hat{x}) - L(t,\hat{x})(h(x) - h(\hat{x})), x - \hat{x}) \leq \omega(t)\,|x - \hat{x}|^2 \quad (3.7)$$

Then we have the estimate

$$\begin{aligned} E[|x(t) - \hat{x}(t)|^2|\mathcal{F}_t^y] &\leq e^{\int_0^t 2\omega(s)\,ds}\,tr\,P(0) \\ &+ \int_0^t e^{\int_s^t 2\omega(\sigma)\,d\sigma}\,(|L(t,\hat{x}(t))|^2 + tr\,\Sigma)\,ds. \end{aligned} \quad (3.8)$$

Proof: Note that

$$E[|x(t) - \hat{x}(t)|^2|\mathcal{F}_t^y]$$

$$= 2\int_{R^n} (f(x) - f(\hat{x}(t)) + L(t,\hat{x}(t))(h(x) - h(\hat{x}(t)),$$
$$x - \hat{x}(t))\,\pi(t,x)\,dx$$

$$-|L(t,\hat{x}(t)) - \tilde{L}(t)|^2 + |L(t,\hat{x}(t))|^2 + tr\,\Sigma$$

where

$$\tilde{L}(t) = \int_{R^n} (x - \hat{x}(t))(h(x) - h(\hat{x}(t)))^t\pi(t,x)\,dx.$$

Thus (3.8) follows from the same arguments as in the proof of Theorem 3.1. $\square$

4 Mixed Gaussian Filter

In this section we discuss the mixed Gaussian filter. The mixed Gaussian filter assumes the form

$$\psi(t,x) = \sum_{k=1}^{m} \alpha_i(t)\,N(\hat{x}_i, P_i(t)).$$

The i-th Gaussian distribution $N(\hat{x}_i, P_i(t))$ is computed by the Gaussian filter described in Section 2 in a parallel manner. The weights $\alpha_i(t)$, $1 \leq i \leq m$ determines the likelihood of each Gaussian distribution. We propose the following update formula.

We consider the discrete-time measurement case. That is, we have observation

$$y_k = h(x(t_k)) + v_k \quad (4.1)$$

where $t_k = k\,\Delta t$ and v_k are white noises with covariances R, and independent of $x(t)$. Let $Y_k = \{y_i,\ 1 \leq i \leq k\}$. The probability density function $p_{k|k}$ of the conditional expectation $E[x(t_k)|Y_k]$ is given by Bayes' formula

$$p_{k|k}(x) = c\,e^{-\frac{1}{2}(y-h(x))^t R^{-1}(y-h(x))}p_{k|k-1}(x) \quad (4.2)$$

where the one-step prediction $p_{k|k-1} = p(t_k)$ is given by the Fokker-Planck

$$\frac{d}{dt}p(t) + Ap(t) = 0 \text{ with } p(t_{k-1}) = p_{k-1|k-1}. \quad (4.3)$$

We apply the Gaussian filter to (4.2)–(4.3) and obtain

Predictor Step

$$\begin{aligned} \frac{d}{dt}\hat{x}^{(1)}(t) &= \widehat{f(x)}^{(1)}(t) \\ \frac{d}{dt}P^{(1)}(t) &= F^{(1)}(\hat{x}^{(1)}(t), P^{(1)}(t)) + \Sigma \end{aligned} \quad (4.4)$$

with $\hat{x}((t_{k-1})^+) = x_{k-1|k-1}$ and $P^{(1)}((t_{k-1})^+) = P_{k-1|k-1}$. Set $x_{k|k-1} = \hat{x}^{(1)}((t_k)^-)$ and $P_{k|k-1} = P^{(1)}((t_k)^-)$.

Corrector Step

$$\begin{aligned} x_{k|k} &= x_{k|k-1} + L_k(y_k - \widehat{h(x)}^{(1)}(x_{k|k-1})) \\ P_{k|k} &= P_{k|k-1} - L_k P_{xz}^t \end{aligned} \quad (4.5)$$

where the filter gain L_k is defined by

$$L_k = P_{xy}(R + P_{yy})^{-1} \tag{4.6}$$

and the covariance matrices P_{xy}, P_{yy} is defined by

$$P_{yy} = \int_{R^n} (h(x) - \hat{h})(h(x) - \hat{h})^t N(x_{k|k-1}, P_{k|k-1})\, dx$$

$$P_{xy} = \int_{R^n} (x - x_{k|k-1})(h(x) - \hat{h})^t N(x_{k|k-1}, P_{k|k-1})\, dx.$$

We apply the Gaussian filter (4.2)–(4.6) to each Gaussian distribution $N(x^{(i)}_{k-1|k-1}, P^{(i)}_{k-1|k-1})$ and obtain the update $N(x^{(i)}_{k|k}, P^{(i)}_{k|k})$. Each update is independent from the others and can be performed in a parallel manner. Next, we update the weights $\alpha^{(i)}_k$ for the new update $p_{k|k}(x)$ at the end of corrector step.

We determine the weights $\alpha^{(i)}_k$ by the minimization of the sum of collocation errors:

$$\text{(4.7)}\quad \sum_{i=1}^{m} \left| p_{k|k}(x^{(i)}_{k|k}) - \sum_{j=1}^{m} \alpha^{(j)}_k N(x^{(j)}_{k|k}, P^{(j)}_{k|k})(x^{(i)}_{k|k}) \right|^2 .$$

over $\alpha \in R^m$ satisfying $\alpha \geq \alpha_0 > 0$, where $p_{k|k}$ is defined as in the corrector step (4.2) with

$$p_{k|k-1}(x) = \sum_{j=1}^{m} \alpha^{(j)}_{k-1} N(x^{(j)}_{k|k-1}, P^{(j)}_{k|k-1})(x).$$

A positive constant α_0 is chosen so that the likelihood of each Gaussian distribution is nonzero (e.g., $\alpha_0 = 0.001$). Problem (4.7) is formulated as the quadratic programming
(4.8)

$$\min \quad \frac{1}{2}\alpha^t A^t A\alpha - \alpha^t A^t b + \frac{\delta}{2}|\alpha|^2 \quad \text{subject to } \alpha \geq \alpha_0,$$

where $\delta > 0$ is chosen so that the singularity of the matrix $A^t A$ is avoided and the matrices (A, b) are defined by

$$A_{i,j} = N(x^{(j)}_{k|k}, P^{(j)}_{k|k})(x^{(i)}_{k|k})$$

$$b_i = \sum_{j=1}^{m} N(h(x^{(i)}_{k|k}, R)(y_k) N(x^{(j)}_{k|k-1}, P^{(j)}_{k|k-1})(x^{(i)}_{k|k})$$

Thus, we solve (4.8) to obtain the weights $\alpha^{(j)}_k$ at each corrector step by using the existing numerical optimization method The theoretical foundation of the Gaussian sum approximation as above is that any probability density function can be approximated as closely as desired by a Gaussian sum.

5 Quadrature Rules

In this section we discuss an implementation of the proposed filters. In [6] we develop the approximation methods for the integral of the form

$$I = \int_{R^n} F(t)\, \tfrac{1}{((2\pi)^n \det \Sigma)^{1/2}}\, e^{-\frac{1}{2}(t-\bar{x})^t \Sigma^{-1}(t-\bar{x})}\, dt. \tag{5.1}$$

If we assume $\Sigma = S^t S$ and change the coordinate of integration by $t = S^t s + \bar{x}$, then

$$I = \int_{R^n} \tilde{F}(s) \tfrac{1}{(2\pi)^{n/2}}\, e^{-\frac{1}{2}|s|^2}\, dt. \tag{5.2}$$

with $\tilde{F}(s) = F(S^t s + \bar{x})$. We apply the Gauss-Hermite quadrature rule. The Gauss quadrature rule is given by

$$\int_{-\infty}^{\infty} g(x)\, \tfrac{1}{(2\pi)^{1/2}}\, e^{-x^2}\, dx = \sum_{i=1}^{m} w_i g(x_i)$$

where the equality holds for all polynomials of degree up to $2m-1$ and the quadrature points x_i and the weights are determined (e.g., see [3]) as follows. Let J be the symmetric tri-diagonal matrix with zero diagonals and $J_{i,i+1} = \sqrt{i/2}$, $1 \leq i \leq m-1$. Then $\{x_i\}$ are the eigenvalues of J and w_i equal to $|(v_i)_1|^2$ where $(v_i)_1$ is the first element of the i-th normalized eigenvector of J. Thus, I is approximate by

$$I_m = \sum_{i_1=1}^{m} \cdots \sum_{i_n=1}^{m} \tilde{F}(q_{i_1}, q_{i_2}, \ldots, q_{i_n})\, w_{i_1} w_{i_2} \ldots w_{i_n} \tag{5.3}$$

where $q_i = \frac{x_i}{\sqrt{2}}$, $1 \leq i \leq m$ and I_m is exact for all polynomials of the form $s^{i_1} s^{i_2} \ldots s^{i_n}$ with $1 \leq i_k \leq 2m-1$. In order to evaluate I_m we need m^n-point function evaluations. For example $m = 3$ we have

$$q_1 = -\sqrt{3},\ q_2 = 0,\ q_3 = \sqrt{3} \quad \text{and} \quad w_1 = w_3 = \frac{1}{6},\ w_2 = \frac{2}{3}.$$

and I_3 requires 3^n-point function evaluations.

Next, we consider be the polynomial approximation P_1 of $\tilde{F}$, defined by

$$P_1(s) = \tilde{F}(0) + \textstyle\sum_{i=1}^{n} \dfrac{\tilde{F}(he_i) - \tilde{F}(-he_i)}{2h}\, s_i + \textstyle\sum_{i=1}^{n} \dfrac{1}{2} H_{i,i} s_i^2$$

where

$$H_{i,i} = \frac{\tilde{F}(he_i) - 2\tilde{F}(0) + \tilde{F}(-he_i)}{h^2},$$

with $h = \sqrt{3}$, which is only based on the values $\tilde{F}(\pm he_i) = P_1(\pm he_i)$, $1 \leq i \leq n$ and $\tilde{F}(0) = P_1(0)$, and uses the the diagonal second order correction $\sum_{i=1}^{n} \frac{1}{2} H_{ii} s_i^2$ of the central difference approximation of $\tilde{F}$. Then the integrals I is approximated by

$$I_c = \int_{R^n} P_1(s)\, \tfrac{1}{(2\pi)^{n/2}}\, e^{-\frac{1}{2}|s|^2}\, ds = \tilde{F}(0) + \textstyle\sum_{i=1}^{n} \dfrac{1}{2} H_{i,i} \tag{5.4}$$

5.1 Examples

First we consider the one dimensional equation

$$dx(t) = 5(x - x^3)\, dt + .5\, dw_1(t)$$

$$dy = (x(t) - .5)^2\, dt + +R^{\frac{1}{2}}\, dw_2(t)$$

Then the filter equation (2.1)–(2.6) is given by

$$\begin{aligned} d\hat{x}(t) &= 5(\hat{x} - (3p\hat{x} + \hat{x}^3))\, dt \\ &\quad + pR^{-1}2(\hat{x} - .5)\,(dy(t) - ((\hat{x} - .5)^2 + p)\, dt) \end{aligned}$$

$$\frac{d}{dt}p(t) = 10(-3\hat{x}^2 + 1)p - (40 + 4R^{-1}(\hat{x} - .5)^2)p^2 + Q$$

where $R = .01$ and $Q = .25$.

Next we consider the Lorenz's system

$$dx(t) = \begin{pmatrix} 10(x_2 - x_1) \\ 28x_1 - x_2 - x_1x_3 \\ x_1x_2 - \frac{8}{3}x_3 \end{pmatrix} dt + \sigma \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} dw_1(t)$$

$$dy(t) = h(x(t))\, dt + R^{\frac{1}{2}}\, dw_2(t)$$

with $h(x) = (x_1 - 5)^2 + x_2^2 + x_3^2$. Since the system is quadratic

$$F^{(2)}(\hat{x}, P) = f_x(\hat{x})P + Pf_x(\hat{x})^t$$

$$L^{(2)} = Ph_x(\hat{x})^t R^{-1}$$

an thus the filter equation (2.1)–(2.6) is given by

$$d\hat{x}(t) = [f(\hat{x}) + (0, -P_{13}, P_{12})^t]\, dt + Ph_x(\hat{x})^t R^{-1}\,(dy(t) - (h(\hat{x}) + tr\, P)\, dt)$$

$$\frac{d}{dt}P(t) = f_x(\hat{x})P + Pf_x(\hat{x})^t - Ph_x(\hat{x})^t R^{-1} h_x(\hat{x})P + Q$$

where $Q = \sigma^2\, bb^t$.

References

[1] A.Bensoussan, Nonlinear filtering theory, Progress in Automation and Information Systems, Recent Advances in Stochastic Calculus, ed. J.S.Baras and V.Mirelli.

[2] M.Fujisaki, G.Kallianpur and H.Kunita, Stochastic differential equations for the nonlinear filtering problem, Osaka J. Math., 9 (1972), 19-44.

[3] G. H. Golub, Some modified matrix eigenvalue problems, SIAM Rev., vol. 15 (1973), 318-334.

[4] K.Ito, Approximation of Zakai equation for nonlinear filtering, SIAM J. Control & Optim., 34 (1996), 620-634.

[5] K.Ito and B.L.Rozovskii, Approximation of the Kushner equation for nonlinear filtering, SIAM J. Control & Optim., (2000), to appear.

[6] K.Ito and K.Xiong, New Gaussian filters for nonlinear filtering problems, IEEE Trans. AC, (2000), to appear.

[7] H. J. Kushner, Approximations to optimal nonlinear filters, IEEE Trans. Autom. Control, 12 (1967), 546-556.

[8] R.Sh.Lipster and A.N.Shiryayev, *The Statistics of Random Process*, I,II, Springer-Verlag, Berlin (1977).

[9] E.Pardoux, Equations du filtrage non linéaire, de la prédiction wt du lissage, Stochastics, 3 (1979), 127-168.

[10] B.L.Rozovskii, *Stochastic Evolution Systems, Linear Theory and Application to Nonlinear Filtering*, MIA, Kluwer Academic Publishers, (1991).