Simultaneous State and Parameter Estimation of Nonlinear Models by Evolution Strategies Based Particle Filters*

Katsuji Uosaki Toshiharu Hatanaka
Department of Information and Physical Sciences
Graduate School of Information Science and Technology, Osaka University
Suita, Osaka 565-0871, Japan

Abstract

Recently, particle filters have attracted attentions for nonlinear state estimation. In this approaches, a posterior probability distribution of the state variable is evaluated based on observations in simulation using so-called importance sampling. We proposed a new filter, Evolution Strategies based particle (ESP) filter to circumvent degeneracy phenomena in the importance weights, which deteriorates the filter performance, and apply it to simultaneous state and parameter estimation of nonlinear state space models. Results of numerical simulation studies illustrate the applicability of this approach.

Keywords. Nonlinear filtering, particle filters, Bayesian approach, evolution strategies, importance sampling, selection.

1 Introduction

The problem of state estimation of dynamic systems using a sequence of their noisy observations has been an active research area in control system sciences for many years. We focus here on Bayesian estimation approaches, that is, inference on the unknown state can be performed according to the posteriori probability density function (pdf), which is obtained by combining a prior pdf for the unknown state with a likelihood function relating to the observations. When observations come sequentially in time, recursive state estimation is often interested, where the evolving posterior pdf is evaluated recursively in time. However, in many realistic problems, state space models include nonlinear and non-Gaussian elements that preclude a closed form of expression for the posteriori pdf, and hence many approximations have been proposed such as the extended Kalman filter (EKF) and Gaussian sum filter [6]. Recent progress of computing ability allowed to the rebirth of Monte Carlo integration and its application of Bayesian filtering, or Monte Carlo filters. A class of Monte Carlo filters, known as "particle filters" [4, 2] is discussed here. In this approach, the integrals in Bayes' rule is approximated by a weighted sum based on the discrete grids with associated weights sequentially chosen by the importance sampling. A common problem in the particle filter is the degeneracy phenomenon, where almost all importance weights tend to zero after some iteration and a large computational effort is wasted to updating the particles with negligible weights. In order to resolve this difficulty, several modifications have been proposed such as resampling particle filter (SIR) [5] that introduces a resampling steps. Applying the concept of Evolution Strategies [7], we also developed the Evolution Strategies based prticle (ESP) flter [8]. In this paper, the ESP filter is applied to simultaneous state and parameter estimation of nonlinear state space models. Numerical simulation studies have been conducted to exemplify the applicability of this approach.

*This work is partially supported by the Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (C)(2)14550447.

2 Particle Filters

Consider the following nonlinear state space model.

$$x_{k+1} = f(x_k, v_k) \quad (1)$$

$$y_k = g(x_k, w_k) \quad (2)$$

where x_k and y_k are the state variable and observation, respectively, f and g are known possibly nonlinear functions, v_k and w_k are independently identically distributed (i.i.d.) system noise and observation noise sequences, respectively. We assume v_k and w_k are mutually independent. The main objective here is to find the best estimate of the state variable x_k in some sense based on the all available data of observations $y_{1:k} = \{y_1, y_2, \dots, y_k\}$. We can solve the problem by calculating the posteriori pdf of the state variable x_k of time instant k based on all the available data of observation sequence $y_{1:k}$.

The posteriori pdf $p(x_k|y_{1:k})$ of x_k based on the observation sequence $y_{1:k}$ satisfies the following recursion:

$$p(x_k|y_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|y_{1:k-1})dx_{k-1} \quad (3)$$

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})} \quad (4)$$

with a prior pdf $p(x_0|y_0) \equiv p(x_0)$ of the initial state variable x_0. Here normalizing constant

$$p(y_k|y_{1:k-1}) = \int p(y_k|x_k)p(x_k|y_{1:k-1})dx_k$$

depends on the likelihood $p(y_k|x_k)$, which is determined by the observation equation (2).

Since a closed form solution of recursions (3) and (4) is not admitted except in very restrictive cases including linear Gaussian state space models, where the Kalman filter [1] can be applied, some approximations should be introduced such as the extended Kalman filter (EKF)[6] and particle filters [4, 2]. Since EKF uses a linearization technique based on a first order Taylor expansions of the nonlinear system and observation equations about the current estimate and approximates the posteriori pdf to be Gaussian, it can never describe the true non-Gaussian density well. Particle filters approximate the true posteriori pdf $p(x_k|y_{1:k})$ by a large set of $n \gg 1$ particles $\{x_k^{(i)},\ (i = 1, \ldots, n)\}$, where each particle has an assigned relative weight, $\{w_k^{(i)}, (i = 1, \ldots, n)\}$, $w_k^{(i)} > 0$, $\sum_{i=1}^{n} w_k^{(i)} = 1$ as follows:

$$p(x_k|y_{1:k}) \approx \sum_{i=1}^{n} w_k^{(i)} \delta(x_k - x_k^{(i)}) \tag{5}$$

where $\delta(\cdot)$ is Dirac's delta function ($\delta(x) = 1$ for $x = 0$ and $\delta(x) = 0$ otherwise).

Here, the particles are generated and associated weights are chosen using the principle of "importance sampling". If the samples $x_k^{(i)}$ in (5) were drawn from an importance density $q(x_k^{(i)}|y_{1:k})$, then the associated normalized weights are defined by

$$w_k^{(i)} \propto \frac{p(x_k^{(i)}|y_{1:k})}{q(x_k^{(i)}|y_{1:k})}. \tag{6}$$

When the importance density $q(x_k|y_{1:k-1})$ is chosen to factorize such that

$$q(x_k|y_{1:k}) = q(x_k|x_{k-1}, y_{1:k})q(x_{k-1}|y_{1:k-1}), \tag{7}$$

we can obtain samples $x_k^{(i)}$ by augmenting each of the existing samples $x_{k-1}^{(i)}$ sampled from the importance density $q(x_{k-1}|y_{1:k-1})$ with the new state sampled from $q(x_k|x_{k-1}^{(i)}, y_{1:k})$.

Noting that the posteriori pdf can be rewritten using Bayes' rule as

$$\begin{aligned} p(x_k|y_{1:k}) &= \frac{p(y_k|x_k, y_{1:k-1})p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})} \\ &\propto p(y_k|x_k)p(x_k|x_{k-1})p(x_{k-1}|y_{1:k-1}) \end{aligned} \tag{8}$$

and inserting (7) and (8) into (6), the weights are recursively updated as

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(y_k|x_k^{(i)})p(x_k^{(i)}|x_{k-1}^{(i)})}{q(x_k^{(i)}|x_{k-1}^{(i)}, y_{1:k})}. \tag{9}$$

The particle filter with these steps is called "Sequential Importance Sampling Particle Filter" (SIS).

It is known that the SIS filter suffers from the degeneracy phenomenon, where all but one of the normalized importance weights are very close to zero after a few iterations. By this degeneracy, a large computational effort is wasted to updating trajectories whose contribution to the final estimate is almost zero. In order to prevent this phenomenon, several modifications have been introduced. Among them, resampling process, which eliminates trajectories whose normalized importance weights are small, is a common approach. It involves generating new grid points $x_k^{*(i)}$ $(i = 1, \ldots, n)$ by resampling from the grid approximation (5) randomly with probability

$$\Pr(x_k^{*(i)} = x_k^{(j)}) = w_k^{(j)} \tag{10}$$

and the weights are reset to $w_k^{*(i)} = 1/n$, when

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{n}(w_k^{(i)})^2} \tag{11}$$

with the associated normalized weight $w_k^{(i)}$ is less than a predefined threshold $N_{thres} < 1$. Particle filter with this resampling process is called "Sampling Importance Resampling Particle Filter" (SIR).

3 Evolution Strategies Based Particle Filter

Evolution Strategies (ES) is one of the Evolutionary Computation approaches, computational models simulating natural evolutionary processes to design and implement computer-based problem solving systems (see the extensive surveys, for examples[3]. It has been applied to continuous function optimization in real-valued n-dimensional space via selection and perturbation processes called mutation. Mutation process is realized by the additive process.

$$\begin{aligned} \sigma_j' &= \sigma_j \exp(\tau' N(0,1) + \tau N_j(0,1)) \\ x_j' &= x_j + \sigma_j' N_j(0,1) \end{aligned} \tag{12}$$

where $N(0,1)$ and $N_j(0,1)$ denote a realization of normal random variable and normal random variables sampled anew for counter j with zero mean and unit variance, respectively, and σ_j denotes the mean step size. The factors τ and τ' are chosen dependent on the population size. The μ individuals of higher fitness are chosen deterministically out of the union of μ parents and λ offspring ($(\mu + \lambda)$-selection) or λ offspring only ((μ, λ)-selection) to form the parents of the next generation in order to evolve towards better search region. It can be seen that SIR and ES have similarities; both the importance sampling process in SIR filter and mutation process in ES give perturbation to the parent individuals $x_{k-1}^{(i)}$ with extrapolation by $f(x_{k-1}^{(i)})$, and both resampling process in SIR filter and selection process in ES select offspring among the perturbed individuals. However, there is a difference between them, i.e., resampling in SIR is carried out randomly and the weights are reset as $1/n$, while the selection in ES is deterministic and the fitness function is never reset. Hence, by replacing the resampling process in SIR by the selection process in ES, we can derive a new particle filter as follows.

Based on the particles $\{x_{k-1}^{(i)},\ (i=1,\ldots,n)\}$ sampled from the importance density $q(x_{k-1}|\,y_{1:k-1})$, we generates ℓ samples $\{x_k^{(i,j)},\ (j=1,\ldots,\ell)\}$ according to the importance density function $q(x_k|\,x_{k-1}^{(i)}, y_{1:k})$. Corresponding weights $w_k^{(i,j)}$ are evaluated by

$$w_k^{(i,j)} \propto w_{k-1}^{(i)} \frac{p(y_k|\,x_k^{(i,j)})p(x_{k-1}^{(i,j)}|\,x_{k-1}^{(i)})}{q(x_k^{(i,j)}|\,x_{k-1}^{(i)}, y_{1:k})} \tag{13}$$

From the set of $n\ell$ particles and weights $\{x_k^{(i,j)}, w_k^{(i,j)}, (i=1,\ldots,n,\ j=1,\ldots,\ell)\}$, we choose n sets with the larger weights, and set as $x_k^{(i)}, w_k^{(i)} (i=1,\ldots,n)$. This process corresponds to $(n, n\ell)$-selection in ES. Hence, we call this particle filter using $(n, n\ell)$-selection in ES as Evolution Strategies based particle filter Comma (ESP(,)). When we add the particles $x_k^{(i,0)} = f(x_{k-1}^{(i)}),\ (i=1,\ldots,n)$ in addition to $n\ell$ $x_k^{(i,j)},\ (i=1,\ldots,n,\ j=1,\ldots,\ell)$ sampled from the importance density function $q(x_k|x_{k-1}^{(i)}, y_{1:k})$ as above and evaluate the weights $\{w_k^{(i,j)}, (i=1,\ldots,n, j=0,\ldots,\ell)\}$ by (13), and then choose n sets of $(x_k^{(i)}, w_k^{(i,j)})$ with larger weights from the ordered set of $n(\ell+1)$ particles $\{x_k^{(i,j)}, w_k^{(i,j)}, (i=1,\ldots,n,\ j=0,\ldots,\ell)\}$, we can obtain another ESP filter. Since this ESP filter uses the selection corresponding to $(n+n\ell)$-selection in ES, we can call this filter as Evolution Strategies based particle filter Plus (ESP(+)).

4 Simultaneous State and Parameter Estimation by Evolution Strategies Based Particle Filter

The proposed ESP filter is applied here to simultaneous state and parameter estimation of nonlinear systems. Consider the nonlinear state space model (1) with unknown parameter θ and (2), where a posteriori pdf $p(x_k, \theta|y_{1:k})$ should be approximated to estimate state and parameter simultaneously, Application of Bayes' rule (4) provides

$$\begin{aligned} p(x_{k+1}, \theta|y_{1:k+1}) \quad \propto \quad & p(y_{k+1}|x_{k+1}, \theta)p(x_{k+1}|\theta, y_{1:k+1}) \\ & \times p(\theta|y_{1:k+1}) \end{aligned}$$

Since the form of the theoretical pdf $p(\theta|y_{1:k})$ is not known for unknown parameter case, we replace θ by θ_k at time k, and simply include θ_k in an augmented state vector $\boldsymbol{x}_k = (x_k, \theta_k)^T$, where θ_k evolves as

$$\theta_{k+1} = \theta_k + \eta_k \tag{14}$$

and η_k is a normal random disturbance with zero-mean and very small variance. Then approximation of the true posteriori pdf is given by

$$p(\boldsymbol{x}_k|y_{1:k}) \approx \sum_{i=1}^{n} w_k^{(i)} \delta(\boldsymbol{x}_k - \boldsymbol{x}_k^{(i)}) \tag{15}$$

If particles $\boldsymbol{x}_k^{(i)}$ in (15) were drawn from an importance density

$$\begin{aligned} q(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{k-1}^{(i)}, y_{1:k}) \quad = \quad & q_x(x_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k}) \\ & \times q_\theta(\theta_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k}) \end{aligned} \tag{16}$$

with importance densities for x_k and θ_k, $q_x(x_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k})$ and $q_\theta(\theta_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k})$, and the associated normalized weights are evaluated by

$$\begin{aligned} w_k^{(i)} \quad \propto \quad & w_{k-1}^{(i)} \frac{p(y_k|x_k^{(i)}, \theta_k^{(i)})}{q_x(x_k^{(i)}, \theta_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k})} \\ & \times \frac{p(x_k^{(i)}, \theta_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)})}{q_\theta(\theta_k^{(i)}|x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k})}. \end{aligned} \tag{17}$$

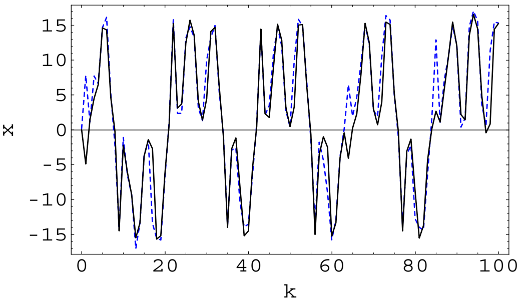
Then, the SIS, SIR and ESP filters are defined as above.

4.1 Numerical Examples

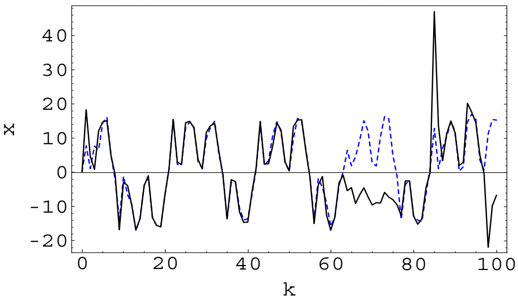
Numerical simulation are carried out to exemplify the applicability of the proposed ESP filter. First, we consider the following nonlinear state space model

$$\begin{aligned} x_k &= \frac{x_{k-1}}{2} + \frac{\theta x_{k-1}}{1+x_{k-1}^2} + 8\cos(1.2k) + v_k \\ &= f(x_{k-1}, \theta) + v_k \qquad (18) \\ y_k &= \frac{x_k^2}{20} + w_k \qquad (19) \end{aligned}$$

where v_k and w_k are i.i.d. zero-mean normal random variables with variance 10 and 1, respectively, and value of the parameter θ is known to be 25. The normal distribution with mean $f(x_{k-1}^{(i)})$ and variance 10 is chosen as the importance density $q(x_k|x_{k-1}^{(i)}, y_{1:k})$. Sample paths of the estimates by the proposed ESP(,) $(n=100, \ell=4)$ and EKF as well for comparison are given in Fig.1. Proposed ESP filter works well in nonlinear state estimation, while the estimate by EKF cannot follow the true state.

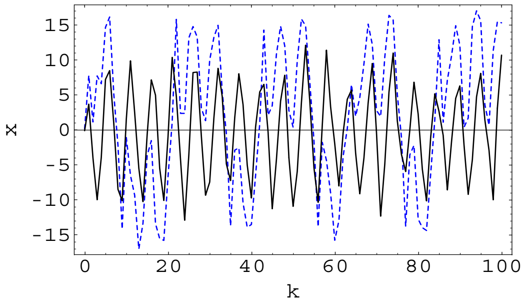


(a) Estimate by ESP(,)

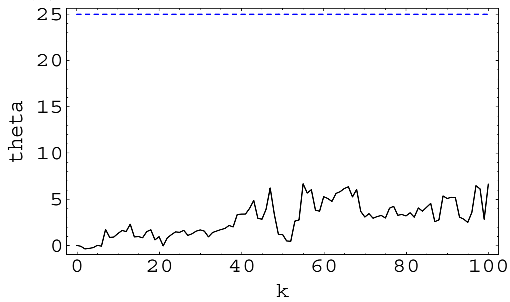


(b) Estimate by EKF

Figure 1: Sample paths of state estimates (dashed line: estimate, solid line: true state)



(a) A sample path of the state estimate



(b) A sample path of the parameter estimate

Figure 2: Simulation results in simultaneous state and parameter estimation by ESP(,)

Next, we consider the unknown parameter case where the true value of $\theta = 25$ in (18) is not known. Here, only the results by ESP(,) with the importance densities $q_x(x_k^{(i)} | x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k}) \sim \mathcal{N}(f(x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, 10)$ and $q_\theta(\theta_k^{(i)} | x_{k-1}^{(i)}, \theta_{k-1}^{(i)}, y_{1:k}) \sim \mathcal{N}(\theta_{k-1}^{(i)}, 0.01)$ are shown in Fig.2 since the EKF does not work as before. Though the estimate approach to the true ones, the convergence speed is slow and the filter leaves much for improvement. For examples, better choice of design parameters n, N_{eff} and ℓ and choice of evolution operations should be pursued since the estimation performance, of course, depends on the choice of them.

5 Conclusions

The novel particle filter, which is developed by recognizing the similarity and the difference between the importance sampling and resampling processes in the SIR filter and mutation and selection processes in ES and substituting (μ, λ)-selection in ES into resampling process in SIR, is applied to simultaneous state and parameter estimation of nonlinear state space models. It works stably and provides small mean square errors compared to EKF filter. Application of other evolution operations such as crossover and modification of mutation will have the potential to create much higher performance particle filters.

References

[1] B. D. O. Anderson and J. B. Moore: *Optimal Filtering*, Prentice-Hall, 1979.

[2] S. Arulampalam, S. Maskell, N. Gordon and T. Clapp: A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking, *IEEE Trans. on Signal Processing*, SP-50, No.2, pp.174–188, 2002.

[3] T. Bäck, *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. Oxford University Press, 1996.

[4] A. Doucet, N. de Freitas and N. Gordon (Eds.): *Sequential Monte Carlo Methods in Practice*, Springer-Verlag, 2001.

[5] N. Gordon, D. Salmond and A. Smith: A novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proc. on Rader and Signal Processing*, 140, pp.107-113, 1993.

[6] A. H. Jazwinski: *Stochastic Process and Filtering Theory*, Academic Press, 1970.

[7] H.-P. Schwefel: *Evolution and Optimum Seeking*, J. Wiley, 1995.

[8] K. Uosaki, Y. Kimura and T. Hatanaka: State estimation of nonlinear stochastic systems by evolution strategies based particle filter. *Proc. 9th Int. Symposium on Artificial Life and Robotics*, pp.406–409, 2004.