Particle Filters with Random Resampling Times

Dan Crisan

Imperial College London

Recent Advances in Sequential Monte Carlo
University of Warwick
19-21 September, 2012

- The Continuous Time Filtering Problem
- Particle Filters with Random Resampling Times
- Convergence Results
- The Central Limit Theorem
- Final remarks

Joint work with O. Obanubi (Imperial College London)

Research partially by the EPSRC Grant No: EP/H0005500/1 and an EPSRC Doctoral Training Award.

D. C. O. Obanubi, Particle Filters with Random Resampling Times
SPA, Vol 122, Issue 4, 2012.

$(\Omega, \mathcal{F}, P)$ probability space $\quad Z = (X, Y) = \{(X_t, Y_t), t \geq 0\}$

- X the signal process - "hidden component"
- Y the observation process - "the data" - $Y_t = f(X, \text{"}noise\text{"})$.

The filtering problem: Find the conditional distribution of the *signal* X_t given $\mathcal{Y}_t = \sigma(Y_s, s \in [0, t])$, i.e.,

$$\pi_t(\varphi) = \mathbb{E}[\varphi(X_t)|\mathcal{Y}_t], \qquad t \geq 0, \qquad \varphi \in \mathcal{B}(\mathbb{R}^d).$$

The model

$V = (V_t^i)_{i=1}^p, t \geq 0\}$, $W = \{\left(W_t^i\right)_{i=1}^m, t \geq 0\}$ independent Brownian motions

$$\begin{aligned} X_t &= X_0 + \int_0^t f(X_s)\, ds + \int_0^t \sigma(X_s)\, dV_s \\ Y_t &= \int_0^t h(X_s)\, ds + W_t, \end{aligned}$$

The process Y becomes a Brownian motion via a change of measure (Girsanov's theorem)

$$\left.\frac{d\tilde{P}}{dP}\right|_{\mathcal{F}_t} = Z_t \triangleq \exp\left(-\int_0^t \sum_{k=1}^m \gamma_k\left(X_s\right) dU_s^k - \frac{1}{2}\int_0^t \sum_{k=1}^m \gamma_k\left(X_s\right)^2 ds\right), t \geq 0.$$

Under $\tilde{P}$, Y becomes a Brownian motion independent of X.
The law of X remains unchanged.

The Kallianpur-Striebel formula

$$\pi_t\left(\varphi\right) = \frac{\rho_t\left(\varphi\right)}{\rho_t(\mathbf{1})},$$

where

$$\rho_t\left(\varphi\right) = \tilde{\mathbb{E}}\left[\varphi(X_t)\exp\left(\int_0^t \sum_{k=1}^m \gamma_k\left(X_s\right) dY_s^k - \frac{1}{2}\int_0^t \sum_{k=1}^m \gamma_k\left(X_s\right)^2 ds\right)\middle| \mathcal{Y}_t\right] \qquad (1)$$

Discrete Time Framework

- Pierre Del Moral, Arnaud Doucet and Ajay Jasra, Bernoulli Volume 18, Number 1 (2012), 252-278

  The result relies on a ingenious coupling argument. The particle filter with random resampling times $(T_k^n)_{k\geq 0}$ is coupled with one with resampling times $(\bar{T}_k)_{k\geq 0}$ where $(\bar{T}_k)_{k\geq 0}$ can be any deterministic times or times that could depend on the observation process only (and not on the current state of the particle filter). The authors show that, as n increases, T_k^n are exponentially close to $\bar{T}_k$. Since time runs discreetly, they must be equal with high probability and the convergence result follows by analyzing the particle filter with observation dependent resampling times. This argument cannot be applied in a continuous-time framework. In continuous time, the corresponding equivalent of T_k^n and $\bar{T}_k$ can be different no matter how close they are.

- Douc, R. and Moulines, E. (2008). Limit theorems for weighted samples with applications to sequential Monte Carlo methods. Ann. Statist.

The approximation will take the form

$$\pi_t^n = \sum_{j=1}^n \bar{a}_j^n(t)\,\delta_{v_j^n(t)}. \qquad \rho_t^n = \sum_{j=1}^n a_j^n(t)\,\delta_{v_j^n(t)}.$$

Time 0.

- choose $v_j^n(0)$ i.i.d., $\mathcal{L}\left(v_j^n(0)\right) = \pi_0$,

$$\pi_0^n = \sum_{j=1}^n \frac{1}{n}\delta_{v_j^n(0)} \qquad \rho_0^n = \sum_{j=1}^n \frac{\rho_0(\mathbf{1})}{n}\delta_{v_j^n(0)}.$$

$\{T_k^n\}_{k\in\mathbb{N}}$ be a strictly increasing sequence of predictable stopping times.

Time interval $\left[T_k^n, T_{k+1}^n\right]$.

- The particles move with the same law as X

$$v_j^n(T) = v_j^n(T_k) + \int_{T_k}^T f(v_j^n(s))\,\mathrm{d}s + \int_{T_k}^T \sigma(v_j^n(s))\,\mathrm{d}V_s^{(j)}, \quad j = 1, \ldots, n, \quad (2)$$

$(V^{(j)})_{j=1}^n$ are independent $\mathcal{F}_t$-adapted p-dimensional Bm independent of Y and independent of all other random variables in the system.

Each particle is assigned a normalized weights $\bar{a}_j^n(T)$, $j = 1, \ldots, n$, for arbitrary stopping time $T \in [T_k, T_{k+1})$ given by

$$\bar{a}_j^n(T) := \frac{a_j^n(T)}{\sum_{k=1}^n a_k^n(T)}$$

where

$$a_j^n(T) = \exp\left(\int_{T_k}^T h(v_j^n(s))^\top \, \mathrm{d}Y_s - \frac{1}{2}\int_{T_k}^T \|h(v_j^n(s))\|^2 \, \mathrm{d}s\right). \tag{3}$$

For $T \in [T_k, T_{k+1})$, define

$$\pi_T^n = \sum_{j=1}^n \bar{a}_j^n(T)\delta_{v_j^n(T)}.$$

At the end of the (random) interval $[T_k, T_{k+1})$, the correction procedure (e.g. sample n-times from $\pi^n_{T_{k+1}-}$) is implemented, the particles are re-indexed and their weights reinitialized to 1.The aim of the correction procedure is to avoid the sample degeneracy.

Examples:

- The effective sample size of the system. Define

$$\hat{n}_{eff} = \frac{1}{\sum_{j=1}^{n} (\bar{a}_j^n(T))^2}, \tag{4}$$

  - $1 \leq \hat{n}_{eff} \leq n$.
  - Define $T_k := \inf\{t \geq T_{k-1} : n_{eff} \leq \lambda_{thres} n\}$, $\lambda_{thres} \in (0, 1)$.

- The coefficient of variation CV, where

$$CV := \left(\frac{1}{n} \sum_{j=1}^{n} \left(n\bar{a}_j^n(t)^2 - 1 \right)^2 \right)^{\frac{1}{2}} \tag{5}$$

  - $0 \leq CV \leq \sqrt{n-1}$.
  - $CV = (n/n_{eff} - 1)^{1/2}$.
  - Define $T_k := \inf\{t \geq T_{k-1} : CV > \alpha\}$, $\alpha \in \left(0, \sqrt{n-1}\right)$.

- The (soft) maximum of the unnormalized weights $a_j^n(T) = \exp(w_j^n(T))$ where

$$w_j^n(T) := \int_{T_k}^{T} h(v_j^n(s))^\top \,\mathrm{d}Y_s - \frac{1}{2}\int_{T_k}^{T} \|h(v_j^n(s))\|^2 \,\mathrm{d}s, \quad j = 1, \ldots, n.$$

Then

$$\max_{1\le j\le n} a_j^n(T) = \exp(\max_{1\le j\le n} w_j^n(T)). \tag{6}$$

  - $\max_{1\le j\le n} w_j^n(T) = \lim_{r\to\infty} \frac{\log\sum_{j=1}^n \exp(r\, w_j^n(T))}{r} \equiv SM(r)$.
  - Define $T_k := \inf\{t \ge T_{k-1} : SM(r) \ge \alpha\}$, $\alpha > 1$.

- The entropy E_t, where

$$E_t = -\sum_{j=i}^{n} \bar{a}_j^n(t) \log \bar{a}_j^n(t) \tag{7}$$

  - $E_t = -\sum_{j=i}^n \bar{a}_j^n(t) \log \bar{a}_j^n(t)$.
  - Define $T_k := \inf\{t \ge T_{k-1} : E_t \le \beta\}$, $\beta \in (0, \log n)$.

Introduce the measure-valued process $\rho^n = \{\rho_t^n : t \geq 0\}$ to be defined by

$$\rho_t^n := \xi_t^n \pi_t^n, \quad t \geq 0,$$

where $\xi^n = \{\xi_t^n : t \geq 0\}$ is the process defined as $\xi_t^n := \prod_{i=1}^{\infty} \frac{1}{n} \sum_{j=1}^{n} a_j^{n,i}(t)$ with

$$a_j^{n,i}(t) := \exp\left(\int_{T_{i-1}\wedge t}^{T_i \wedge t} h(v_j^n(s))^\top \, \mathrm{d}Y_s - \frac{1}{2} \int_{T_{i-1}\wedge t}^{T_i \wedge t} \|h(v_j^n(s))\|^2 \, \mathrm{d}s \right). \tag{8}$$

Define

- N_t^n the number of resampling instances that occur before time t
- For μ finite signed measure define

$$||\mu||_p = \sup_{\{\varphi \in C_b^1(\mathbb{R}^d), \|\varphi\|_{1,\infty} \leq 1\}} \left(\mathbb{E}\left[(\mu(\varphi))^p \right] \right)^{1/p}$$

Theorem (D.C, O. Obanubi, 2012)

Assume that there exists $p > 1$ such that for all $t > 0$

$$\sup_{n>0} \mathbb{E}[(N_t^n)^p] < \infty. \tag{9}$$

Then for any $T \geq 0$ and for any $r < p$, there exists a constant $\alpha = \alpha(T)$, independent of n such that

$$\sup_{t\in[0,T]} ||\rho_t^n - \rho_t||_{2r} \leq \frac{\alpha}{\sqrt{n}} \qquad \sup_{t\in[0,T]} ||\pi_t^n - \pi_t||_r \leq \frac{\alpha}{\sqrt{n}} \tag{10}$$

Remarks:

- We do not assume that the resampling times converge as typically they will depend on π_n and therefore their convergence cannot be a priori assumed.
- Condition (9) implies that $\lim_{k\to\infty} T_k^n = \infty$. In particular, that there are only a finite number of resampling times in any finite interval.
- Condition (9) is satisfied for any sequence of *deterministic* times that converge to ∞ and for random times determined by the ESS, CV and the soft maximum weights criteria.

Let N^φ be an $(\mathcal{F}_t \vee \mathcal{Y})$-adapted square-integrable martingale given by

$$
\begin{aligned}
N_t^\varphi &= \int_0^t \int_{\mathbb{R}^d} \sqrt{\tilde{\rho}_s\left((\nabla\varphi)^\top \sigma\sigma^\top (\nabla\varphi)\right)}\, B(dx, ds) \\
&\quad + \sum_{k=1}^\infty \mathbf{1}_{[0,t]}(T_k)\rho_{T_k}(\mathbf{1})\sqrt{\pi_{T_k}(\varphi^2) - \pi_{T_k-}(\varphi)^2}\,\Upsilon_k
\end{aligned}
$$

where $B(dx, ds)$ is a Brownian sheet or space-time white noise, $\{\Upsilon_k\}_{k\in\mathbb{N}}$ is a sequence of i.i.d standard normal r.v. mutually independent given $\mathcal{Y}$.

Theorem

If $U := \{U_t : t \geq 0\}$ is a $D_{M_F(\mathbb{R}^d)}[0,\infty)$-valued process such that for $\varphi \in C_0^2(\mathbb{R}^d)$

$$
U_t(\varphi) = U_0(\varphi) + \int_0^t U_s(A\varphi)\, ds + N_t^\varphi + \sum_{k=1}^m \int_0^t U_s(h^k\varphi)\, dY_s^k \tag{11}
$$

then U is pathwise unique.

Theorem (D.C, O. Obanubi, 2012)

Assume that for any $k \geq 0$, $\lim_{n\to\infty} T_k^n = T_k$, *where* $(T_k)_{k\geq 0}$ *is a strictly increasing sequence of predictable stopping times such that*

$$\sup_{n>0} \mathbb{E}[(N_t^n)^p] < \infty. \quad (12)$$

$$\lim_{\delta\to 0} \sup_{n>0} [\sup_{s\in[0,t]} \mathbb{E}[(N_{s+\delta}^n - N_s^n)^p | \mathcal{F}_s]] = 0. \quad (13)$$

Then $\{\sqrt{n}(\rho_t^n - \rho_t)\}_n$ *converges in distribution to* U *satisfying*

$$U_t(\varphi) = U_0(\varphi) + \int_0^t U_s(A\varphi)\,ds + N_t^\varphi + \sum_{k=1}^m \int_0^t U_s(h^k\varphi)\,dY_s^k, \quad \varphi \in C_0^2(\mathbb{R}^d), \quad (14)$$

where N^φ *is a martingale with quadratic variation*

$$\langle N_\cdot^\varphi \rangle_t = \int_0^t \tilde{\rho}_s\left((\nabla\varphi)^\top \sigma\sigma^\top (\nabla\varphi)\right) ds + \sum_{k=1}^\infty \left[\mathbf{1}_{[0,\,t]}(T_k)\rho_{T_k}(\mathbf{1})^2 \left[\pi_{T_k}(\varphi^2) - \pi_{T_k-}(\varphi)^2 | T_{k-}\right]\right].$$

Corollary

Let $\bar{U}^n : \{\bar{U}^n_t : t \geq 0\}$ be the process defined as $\bar{U}^n_t := \sqrt{n}(\pi^n_t - \pi_t), \quad t \geq 0$. Under the same condtions as above, the process $\{\bar{U}^n\}_n$ converges in distribution to the measure-valued process $\bar{U} : \{\bar{U}_t : t \geq 0\}$ defined as

$$\bar{U}_t = \frac{1}{\rho_t(\mathbf{1})}\left(U_t - U_t(\mathbf{1})\pi_t\right), \quad t \geq 0, \tag{15}$$

where U satisfies (14).

- We assume now that the resampling times converge.
- Condition (13) is satisfied for random times determined by the ESS, CV and the soft maximum weights criteria.
- Condition (13) is *not* satisfied for any sequence of *deterministic* times that converge. For this a variation of (13) needs to be imposed.

- We analyse (continuous time) particle filters where resampling takes place at times that form a sequence of (predictable) stopping times
- We prove that, under very general conditions imposed on the sequence of resampling times, the corresponding particle filters converge.
- The conditions are verified when the resampling times are chosen in accordance to effective sample size of the system of particles, the coefficient of variation of the particles' weights and, respectively, the (soft) maximum/minimum of the particles' weights.
- We also deduce central-limit theorem type results for the approximating particle system with random resampling times.