Variance estimation in the particle filter

Nick Whiteley
University of Bristol

Joint work with Anthony Lee, University of Warwick

arXiv:1509.00394

Outline

Outline

Particle filters

Variance estimators

Applications of the estimators

HMM's

Ingredients

1. $(X_n, Y_n)_{n\geq 0}$, $(X_n)_{n\geq 0}$ unobserved, $(Y_n)_{n\geq 0}$ observed
2. $(X_n)_{n\geq 0}$ is Markov, $X_0 \sim M_0$, $X_n | X_{n-1} \sim M_n(X_{n-1}, \cdot)$
3. $(Y_n)_{n\geq 0}$ conditionally indep. given $(X_n)_{n\geq 0}$, $Y_n | (X_n)_{n\geq 0} \sim G_n(X_n, \cdot)$

Target measures

$$\gamma_n(A) := \mathbb{E}\left[\mathbb{I}_A(X_n) \prod_{p=0}^{n-1} G_p(X_p, y_p)\right]$$

$$\eta_n(A) := \quad \frac{\gamma_n(A)}{\gamma_n(\mathbb{X})} = \text{cond. probability of } X_n \in A \text{ given } y_0, \ldots, y_{n-1},$$

$$\gamma_n(\mathbb{X}) = \text{marg. likelihood of } y_0, \ldots, y_{n-1}.$$

Tempered targets

Ingredients

1. Unnormalized prob. densities $\bar{\pi}_0(x), \bar{\pi}_1(x)$

$$Z_i := \int \bar{\pi}_i(x)dx, \quad \pi_i(dx) := \frac{\bar{\pi}_i(x)dx}{Z_i}, \quad i \in \{0,1\}$$

   and a sequence of constants $0 = \beta_0 < \cdots < \beta_n = 1$.

2. $X_0 \sim \pi_0$ and for $p = 1, \ldots, n$, $X_p | X_{p-1} \sim M_p(X_{p-1}, \cdot)$, where M_p is invariant w.r.t. dist. $\propto \bar{\pi}_0^{1-\beta_p}(x)\bar{\pi}_1^{\beta_p}(x)$,
3. for $p = 0, \ldots, n-1$, $G_p(x) := \bar{\pi}_1(x)^{\beta_{p+1}-\beta_p} / \bar{\pi}_0(x)^{\beta_{p+1}-\beta_p}$

Target measures

$$\gamma_n(A) := \mathbb{E}\left[\mathbb{I}_A(X_n)\prod_{p=0}^{n-1} G_p(X_p)\right], \quad \eta_n(A) := \frac{\gamma_n(A)}{\gamma_n(\mathbb{X})}.$$

$$\eta_n(A) = \pi_n(A), \quad \gamma_n(\mathbb{X}) = \frac{Z_1}{Z_0}.$$

Generic formulation

- On a measurable space $(\mathbb{X}, \mathcal{X})$, define:
  - M_0 a prob. measure, $(M_n)_{n\geq 1}$ a seq. of Markov kernels.
  - $(G_n)_{n\geq 0}$ a seq. of strictly positive, upper bounded functions.
- Define measures $\gamma_0 := M_0$ and, recursively,

$$\gamma_n(A) := \int \gamma_{n-1}(\mathrm{d}x) G_{n-1}(x) M_n(x, A), \quad n \geq 1, \quad A \in \mathcal{X}.$$

- Probability measure counterparts

$$\eta_n(A) := \frac{\gamma_n(A)}{\gamma_n(1)} = \frac{\eta_{n-1}(G_{n-1} M_n(A))}{\eta_{n-1}(G_{n-1})}, \qquad A \in \mathcal{X}.$$

- Notation: $\mathcal{L}(\mathcal{X})$ the set of real-valued, bounded, $\mathcal{X}$-measurable functions, and for a measure μ on $\mathcal{X}$,

$$\mu(\varphi) := \int_{\mathsf{X}} \varphi(x) \mu(\mathrm{d}x), \qquad \varphi \in \mathcal{L}(\mathcal{X}).$$

A particle filter

At time 0, for each $i \in \{1, \ldots, N\}$, sample $\zeta_0^i \sim M_0(\cdot)$.

At each time $n \geq 1$, for each $i \in \{1, \ldots, N\}$, sample

1. $A_{n-1}^i \sim \text{Categorical}\left(\frac{G_{n-1}(\zeta_{n-1}^1)}{\sum_{j=1}^N G_{n-1}(\zeta_{n-1}^j)}, \ldots, \frac{G_{n-1}(\zeta_{n-1}^N)}{\sum_{j=1}^N G_{n-1}(\zeta_{n-1}^j)}\right)$,
2. $\zeta_n^i \sim M_n(\zeta_{n-1}^{A_{n-1}^i}, \cdot)$.

$$\eta_n^N(A) := \frac{1}{N} \sum_{i \in [N]} \delta_{\zeta_n^i}(A), \quad \gamma_n^N(A) := \eta_n^N(A) \prod_{p=0}^{n-1} \eta_p^N(G_p).$$

A particle filter

At time 0, for each $i \in \{1, \ldots, N\}$, sample $\zeta_0^i \sim M_0(\cdot)$.

At each time $n \geq 1$, for each $i \in \{1, \ldots, N\}$, sample

1. $A_{n-1}^i \sim \mathrm{Categorical}\left(\frac{G_{n-1}(\zeta_{n-1}^1)}{\sum_{j=1}^N G_{n-1}(\zeta_{n-1}^j)}, \ldots, \frac{G_{n-1}(\zeta_{n-1}^N)}{\sum_{j=1}^N G_{n-1}(\zeta_{n-1}^j)}\right)$,
2. $\zeta_n^i \sim M_n(\zeta_{n-1}^{A_{n-1}^i}, \cdot)$.

$$\eta_n^N(A) := \frac{1}{N} \sum_{i \in [N]} \delta_{\zeta_n^i}(A), \quad \gamma_n^N(A) := \eta_n^N(A) \prod_{p=0}^{n-1} \eta_p^N(G_p).$$

integrating out the A_{n-1}^i's gives

$$\begin{aligned} \zeta_n^i \quad \sim \quad & \frac{\sum_{j=1}^N G_{n-1}(\zeta_{n-1}^j) M_n(\zeta_{n-1}^j, \cdot)}{\sum_{j=1}^N G_{n-1}(\zeta_{n-1}^j)} = \frac{\eta_{n-1}^N(G_{n-1} M_n(A))}{\eta_{n-1}^N(G_{n-1})} \\ & \xrightarrow{N \to \infty} \frac{\eta_{n-1}(G_{n-1} M_n(A))}{\eta_{n-1}(G_{n-1})} = \eta_n(A). \end{aligned}$$

Quality of the approximations

Extensive literature on these approximations includes:

1. Central limit theorems: Del Moral and Guionnet [1999], Chopin [2004], Künsch [2005], Douc and Moulines [2008].
2. Nonasymptotic results: Del Moral and Miclo [2001], Cérou et al. [2011].
3. Many others!

- These theoretical results validate the methodology.
- They do not allow us to extract information from a realization of a *single* particle filter in order to report numerical measures of Monte Carlo error.
- Chan and Lai [2013] provide an estimate of the asymptotic variance of an "updated" variant of $\eta_n^N(\varphi)$.

A particle filter

At time 0, for each $i \in \{1, \ldots, N\}$, sample $\zeta_0^i \sim M_0(\cdot)$.

At each time $n \geq 1$, for each $i \in \{1, \ldots, N\}$, sample

1. $A_{n-1}^i \sim \mathrm{Categorical}\left(\frac{G_{n-1}(\zeta_{n-1}^1)}{\sum_{j=1}^N G_{n-1}(\zeta_{n-1}^j)}, \ldots, \frac{G_{n-1}(\zeta_{n-1}^N)}{\sum_{j=1}^N G_{n-1}(\zeta_{n-1}^j)} \right)$,
2. $\zeta_n^i \sim M_n(\zeta_{n-1}^{A_{n-1}^i}, \cdot)$.

$$\eta_n^N(A) := \frac{1}{N} \sum_{i \in [N]} \delta_{\zeta_n^i}(A), \quad \gamma_n^N(A) := \eta_n^N(A) \prod_{p=0}^{n-1} \eta_p^N(G_p).$$

Outline

A particle filter + "Eve" variables

At time 0, for each $i \in \{1, \ldots, N\}$, sample $\zeta_0^i \sim M_0(\cdot)$ and set $E_0^i = i$.

At each time $n \geq 1$, for each $i \in \{1, \ldots, N\}$, sample

1. $A_{n-1}^i \sim \mathrm{Categorical}\left(\frac{G_{n-1}(\zeta_{n-1}^1)}{\sum_{j=1}^N G_{n-1}(\zeta_{n-1}^j)}, \ldots, \frac{G_{n-1}(\zeta_{n-1}^N)}{\sum_{j=1}^N G_{n-1}(\zeta_{n-1}^j)}\right)$,
2. $\zeta_n^i \sim M_n(\zeta_{n-1}^{A_{n-1}^i}, \cdot)$ and set $E_n^i = E_{n-1}^{A_{n-1}^i}$.

Variance estimation

Define

$$V_n^N(\varphi) := \eta_n^N(\varphi)^2 - \left(\frac{N}{N-1}\right)^{n-1} \frac{1}{N(N-1)} \sum_{i,j: E_n^i \neq E_n^j} \varphi(\zeta_n^i)\varphi(\zeta_n^j)$$

Thm. (Lee & W.) If for all $0 \leq p < n$, $G_p(x) > 0$, $\forall x$, and $\sup_x G_p(x) < \infty$, then for any $\varphi \in \mathcal{L}(\mathcal{X})$,

$$\mathbb{E}\left[\gamma_n^N(1)^2 V_n^N(\varphi)\right] = \mathrm{var}\left[\gamma_n^N(\varphi)\right]$$

$$N\gamma_n^N(1)^2 V_n^N(\varphi) \xrightarrow{p} \sigma^2_{\gamma_n^N} := \lim_{N\to\infty} N \,\mathrm{var}\left[\gamma_n^N(\varphi)\right],$$

$$NV_n^N(\varphi - \eta_n^N(\varphi)) \xrightarrow{p} \sigma^2_{\eta_n^N} := \lim_{N\to\infty} N\mathbb{E}\left[\left|\eta_n^N(\varphi) - \eta_n(\varphi)\right|^2\right].$$

Variance estimation - interpretation

$$V_n^N(\varphi) := \eta_n^N(\varphi)^2 - \left(\frac{N}{N-1}\right)^{n-1} \frac{1}{N(N-1)} \sum_{i,j:E_n^i \neq E_n^j} \varphi(\zeta_n^i)\varphi(\zeta_n^j)$$

think: $X^1, \ldots, X^N$ with sample mean $\bar{X}$ and sample variance:

$$\bar{X}^2 - \frac{1}{N(N-1)} \sum_{i \neq j} X^i X^j = \frac{1}{N(N-1)} \sum_i (X^i - \bar{X})^2$$

Variance estimation - interpretation

$$V_n^N(\varphi) := \eta_n^N(\varphi)^2 - \left(\frac{N}{N-1}\right)^{n-1} \frac{1}{N(N-1)} \sum_{i,j:E_n^i \neq E_n^j} \varphi(\zeta_n^i)\varphi(\zeta_n^j)$$

think: $X^1, \ldots, X^N$ with sample mean $\bar{X}$ and sample variance:

$$\bar{X}^2 - \frac{1}{N(N-1)} \sum_{i \neq j} X^i X^j = \frac{1}{N(N-1)} \sum_i (X^i - \bar{X})^2$$

$$\#_n^i := \mathrm{card}\{j : E_n^j = i\}, \qquad \Delta_n^i := \frac{1}{\#_n^i} \sum_{j:E_n^j = 1} \varphi(\zeta_n^i) - \eta_n^N(\varphi)$$

Cor. (Lee & W.)

$$NV_n^N(1) = \frac{1}{N} \sum_i (\#_n^i - 1)^2 - n + O_p(1/N)$$

$$NV_n^N(\varphi - \eta_n^N(\varphi)) = \frac{1}{N} \sum_i (\#_n^i \Delta_n^i)^2 + O_p(1/N)$$

Proof ideas

- objective is to obtain a numerical estimator of var $\left[\gamma_n^N(\varphi)\right]$, from which estimators of $\sigma^2_{\gamma_n^N}$ and $\sigma^2_{\eta_n^N}$ may then be derived

Proof ideas

- objective is to obtain a numerical estimator of $\mathrm{var}\left[\gamma_n^N(\varphi)\right]$, from which estimators of $\sigma^2_{\gamma_n^N}$ and $\sigma^2_{\eta_n^N}$ may then be derived

- $\mathrm{var}\left[\gamma_n^N(\varphi)\right]$ can be decomposed into terms by order in N^{-1}

Proof ideas

- objective is to obtain a numerical estimator of $\mathrm{var}\left[\gamma_n^N(\varphi)\right]$, from which estimators of $\sigma^2_{\gamma_n^N}$ and $\sigma^2_{\eta_n^N}$ may then be derived
- $\mathrm{var}\left[\gamma_n^N(\varphi)\right]$ can be decomposed into terms by order in N^{-1}
- "information" relevant to estimation of these terms is carried by the genealogical structure of the particle system

Proof ideas

- objective is to obtain a numerical estimator of $\text{var}\left[\gamma_n^N(\varphi)\right]$, from which estimators of $\sigma^2_{\gamma_n^N}$ and $\sigma^2_{\eta_n^N}$ may then be derived
- $\text{var}\left[\gamma_n^N(\varphi)\right]$ can be decomposed into terms by order in N^{-1}
- "information" relevant to estimation of these terms is carried by the genealogical structure of the particle system
- main insight is to find a suitable parameterization of this structure – the Eve variables enter here

Proof ideas

- objective is to obtain a numerical estimator of var $\left[\gamma_n^N(\varphi)\right]$, from which estimators of $\sigma^2_{\gamma_n^N}$ and $\sigma^2_{\eta_n^N}$ may then be derived
- var $\left[\gamma_n^N(\varphi)\right]$ can be decomposed into terms by order in N^{-1}
- "information" relevant to estimation of these terms is carried by the genealogical structure of the particle system
- main insight is to find a suitable parameterization of this structure – the Eve variables enter here
- as $N \to \infty$, one has "enough" pairs of particles to consistently estimate the zero'th-order terms in $N\text{var}\left[\gamma_n^N(\varphi)\right]$, hence $\sigma^2_{\gamma_n^N}$ and $\sigma^2_{\eta_n^N}$

Second moment of $\gamma_n^N(\varphi)$

Cérou et al. [2011]: for certain measures $\{\mu_{\mathbf{b}} : \mathbf{b} \in \{0,1\}^{n+1}\}$ on $\mathcal{X}^{\otimes 2}$,

$$\mathbb{E}\left[\gamma_n^N(\varphi)^2\right] = \sum_{\mathbf{b}\in\{0,1\}^{n+1}} \left[\prod_{p=0}^{n} \left(\frac{1}{N}\right)^{b_p} \left(1-\frac{1}{N}\right)^{1-b_p}\right] \mu_{\mathbf{b}}(\varphi^{\otimes 2}).$$

Second moment of $\gamma_n^N(\varphi)$

Cérou et al. [2011]: for certain measures $\{\mu_{\mathbf{b}} : \mathbf{b} \in \{0,1\}^{n+1}\}$ on $\mathcal{X}^{\otimes 2}$,

$$\mathbb{E}\left[\gamma_n^N(\varphi)^2\right] = \sum_{\mathbf{b}\in\{0,1\}^{n+1}} \left[\prod_{p=0}^{n} \left(\frac{1}{N}\right)^{b_p} \left(1 - \frac{1}{N}\right)^{1-b_p}\right] \mu_{\mathbf{b}}(\varphi^{\otimes 2}).$$

Note:

$$\begin{aligned} \gamma_n^N(\varphi)^2 &= \eta_n^N(\varphi)^2 \prod_{p=0}^{n-1} \eta_p^N(G_p)^2 \\ &= N^{-2} \left[\sum_i \varphi(\zeta_n^i)^2 + \sum_{i\neq j} \varphi(\zeta_n^i)\varphi(\zeta_n^j)\right] \\ &\quad \times \prod_{p=0}^{n-1} N^{-2} \left[\sum_i G_p(\zeta_p^i)^2 + \sum_{i\neq j} G_p(\zeta_p^i) G_p(\zeta_p^j)\right] \end{aligned}$$

Measures $\mu_{\mathbf{b}}$ in the second moment formula

- For $\mathbf{b} \in \{0,1\}^{n+1}$, $\mu_{\mathbf{b}}(\varphi) = \mathsf{E}_{\mathbf{b}}\left[\varphi(X_n, X'_n)\prod_{p=0}^{n} G_p(X_p)G_p(X'_p)\right]$, with $(X_p, X'_p) \sim \tilde{M}_p^{b_p}(X_{p-1}, X'_{p-1}, \cdot)$,
- where:
  - when $b_p = 0$,

$$\tilde{M}_p^0(x, x'; \mathrm{d}z, \mathrm{d}z') := M_p(x, \mathrm{d}z)M_p(x', \mathrm{d}z'),$$

  - when $b_p = 1$,

$$\tilde{M}_p^1(x, x'; \mathrm{d}z, \mathrm{d}z') := M_p(x, \mathrm{d}z)\delta_z(\mathrm{d}z').$$

- case $\mathbf{b} = \mathbf{0}$, we obtain $\mu_{\mathbf{0}}(\varphi^{\otimes 2}) = \gamma_n(\varphi)^2$.

Genealogical tracing variables

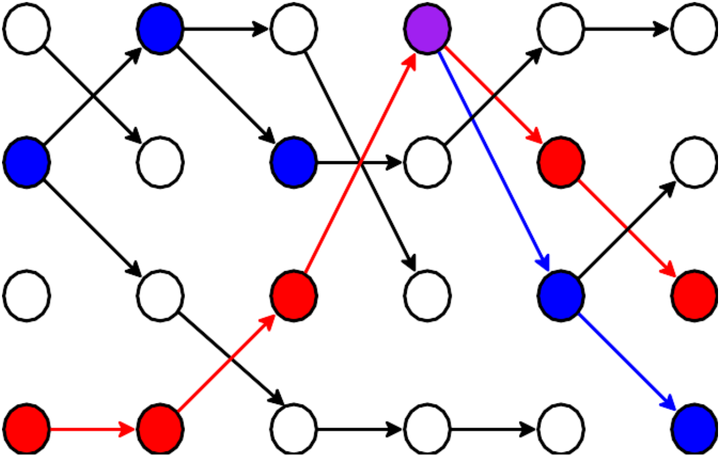
- Consider the particle system simulated up to time n.

- Define auxiliary random variables $\mathbf{K}^1 = (K_0^1, \ldots, K_n^1)$ and $\mathbf{K}^2 = (K_0^2, \ldots, K_n^2)$, with the following sampling *interpretation*:

1. $\mathbf{K}^1$ is an ancestral lineage: sample K_n^1 uniformly from $\{1, \ldots, N\}$, then for $p = n, \ldots, 1$ set $K_{p-1}^1 = A_{p-1}^{K_p^1}$.
2. $\mathbf{K}^2$ consists of possibly "broken" ancestral lineages: sample K_n^2 uniformly from $\{1, \ldots, N\}$, and trace back an ancestral lineage as above, but when a "collision" $K_p^2 = K_p^1$ occurs, sample K_{p-1}^2 with probability proportional to $G_{p-1}(\zeta_{p-1}^{k_{p-1}^2})$.

- Let $\mathbf{C}(\mathbf{A}, \zeta; \mathbf{k}^{1:2})$ be the conditional p.m.f. of $\mathbf{K}^1, \mathbf{K}^2$ given all ancestor indices $\mathbf{A}$ and particles $\boldsymbol{\zeta}$ up to time n

A realization of $(\mathbf{K}^1, \mathbf{K}^2)$ (red, blue)



$\mathbf{k}^1 = (4, 4, 3, 1, 2, 3)$, $\mathbf{k}^2 = (2, 1, 2, 1, 3, 4)$.

Particle approximations of $\mu_{\mathbf{b}}$

Define, for $\mathbf{b} \in \{0,1\}^{n+1}$, and with $N \geq 2$,

$$\mu_{\mathbf{b}}^N := \left[\prod_{p=0}^{n} N^{b_p} \left(\frac{N}{N-1}\right)^{1-b_p}\right] \gamma_n^N(1)^2 \sum_{\mathbf{k}^{1:2} \in \mathcal{I}(\mathbf{b})} \mathbf{C}(\mathbf{A}, \zeta; \mathbf{k}^{1:2}) \delta_{(\zeta_n^{k_n^1}, \zeta_n^{k_n^2})},$$

where $\mathcal{I}(\mathbf{b}) := \{\mathbf{k}^{1:2} \in \{1,\ldots,N\}^2 : k_p^1 = k_p^2 \iff b_p = 1\}$.

Thm. (Lee & W.) For any $\mathbf{b} \in \{0,1\}^{n+1}$ and $\varphi \in \mathcal{L}(\mathcal{X}^{\otimes 2})$,

$$\mathbb{E}\left[\mu_{\mathbf{b}}^N(\varphi)\right] = \mu_{\mathbf{b}}(\varphi),$$

and

$$\sup_{N \geq 1} \sqrt{N} \mathbb{E}\left[\left(\mu_{\mathbf{b}}^N(\varphi) - \mu_{\mathbf{b}}(\varphi)\right)^2\right]^{\frac{1}{2}} < +\infty.$$

Variance estimators: consistency

Define

$$V_n^N(\varphi) := \left[\gamma_n^N(\varphi)^2 - \mu_{\mathbf{0}}^N(\varphi^{\otimes 2})\right] / \gamma_n^N(1)^2$$

and

$$v_{p,n}^N(\varphi) := \left[\mu_{\mathbf{b}_p}^N(\varphi^{\otimes 2}) - \mu_{\mathbf{0}}^N(\varphi^{\otimes 2})\right] / \gamma_n^N(1)^2.$$

Thm. (Lee and W.) For any $\varphi \in \mathcal{L}(\mathcal{X})$, and as $N \to \infty$.

1. $NV_n^N(\varphi) \xrightarrow{p} \bar{\sigma}^2_{\gamma_n^N}(\varphi)$ and $NV_n^N(\varphi - \eta_n^N(\varphi)) \xrightarrow{p} \sigma^2_{\eta_n^N}(\varphi)$,
2. $v_{p,n}^N(\varphi) \xrightarrow{p} v_{p,n}(\varphi)$ and $v_{p,n}^N(\varphi - \eta_n^N(\varphi)) \xrightarrow{p} v_{p,n}(\varphi - \eta_n(\varphi))$,

where

$$\bar{\sigma}^2_{\gamma_n^N}(\varphi) = \frac{\sigma^2_{\gamma_n^N}(\varphi)}{\gamma_n(1)^2} = \sum_{p=0}^{n} v_{p,n}(\varphi), \quad \sigma^2_{\eta_n^N}(\varphi) = \bar{\sigma}^2_{\gamma_n^N}(\varphi - \eta_n(\varphi)).$$

Computational complexity

- The definition of $\mu_{\mathbf{b}}^N$ does not itself suggest an efficient algorithm to compute $\mu_{\mathbf{b}}^N(\varphi^{\otimes 2})$.
- Efficient algorithms for computing $V_n^N(\varphi)$ and $v_{p,n}^N(\varphi)$ satisfy

Estimate	Time complexity	Space complexity
$\gamma_n^N(\varphi)$ or $\eta_n^N(\varphi)$	$\mathcal{O}(Nn)$	$\mathcal{O}(N)$
$V_n^N(\varphi)$	$\mathcal{O}(Nn)$	$\mathcal{O}(N)$
$v_{p,n}^N(\varphi)$	$\mathcal{O}(Nn)$	$\mathcal{O}(Nn)$

- Calculating $V_n^N(\varphi)$ is $\mathcal{O}(N)$ after computing $\gamma_n^N(\varphi)$.
- Calculating $v_{p,n}^N(\varphi)$ requires some recursive computations and storage of the genealogies $A_0, \ldots, A_{n-1}$.
- Evaluation only of φ and the potentials $(G_p)_{p\geq 0}$ is required, using the output of a single particle filter.

Updated models

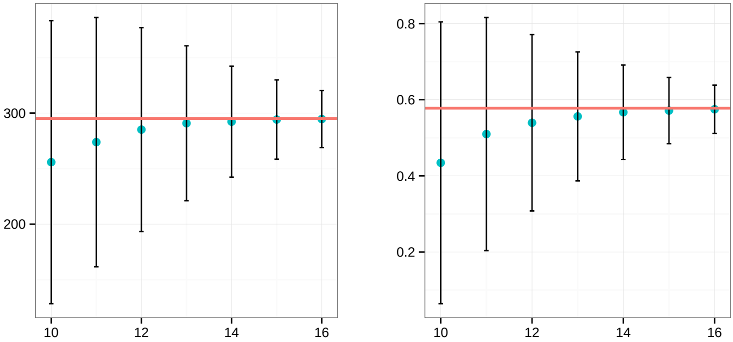
- In many applications, particularly HMMs, there is interest in the "updated" sequence of measures

$$\hat{\gamma}_n(A) := \int_A G_n(x)\gamma_n(\mathrm{d}x), \qquad \hat{\eta}_n(A) := \frac{\hat{\gamma}_n(A)}{\hat{\gamma}_n(1)}, \quad A \in \mathcal{X}.$$

- Corresponding estimators of the variance for their particle approximations: $\hat{V}_n^N(\varphi)$ and $\hat{v}_{p,n}^N(\varphi)$.

- The estimator of $\sigma^2_{\hat{\eta}_n^N}$, $N\hat{V}_n^N(\varphi - \hat{\eta}_n^N(\varphi))$, is exactly $(\frac{N}{N-1})^{n+1}$ times the estimator of Chan and Lai [2013].

Linear Gaussian example: $N\hat{V}_n^N(\varphi)$, estimating $\bar{\sigma}^2_{\hat{\gamma}_n^N}(\varphi)$



(a) $\varphi \equiv 1$ (b) $\varphi = Id - \hat{\eta}_n^N(Id)$

Figure: Estimated asymptotic variances $N\hat{V}_n^N(\varphi)$ (blue dots and error bars for the mean $\pm$ one standard deviation) against $\log_2 N$. The red lines correspond to the true asymptotic variances.

Linear Gaussian example: $\hat{v}^N_{p,n}(1)$, $N = 10^5$

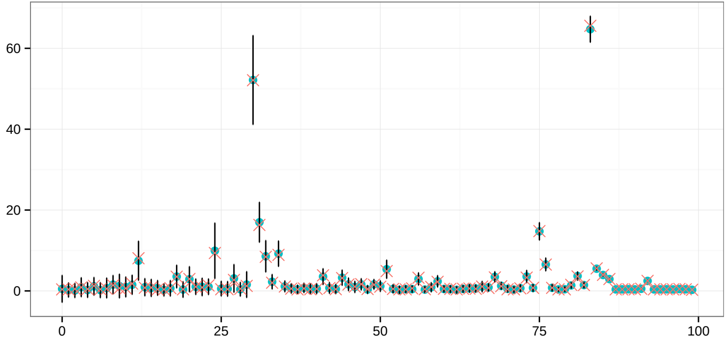


Figure: Plot of $\hat{v}^N_{p,n}(\varphi)$ (blue dots and error bars for the mean $\pm$ one standard deviation) and $\hat{v}_{p,n}(\varphi)$ (red crosses) at each $p \in \{0, \ldots, n\}$ with $\varphi \equiv 1$.

Linear Gaussian example: $\hat{v}^N_{p,n}(Id - \hat{\eta}^N_n(Id))$, $N = 10^5$

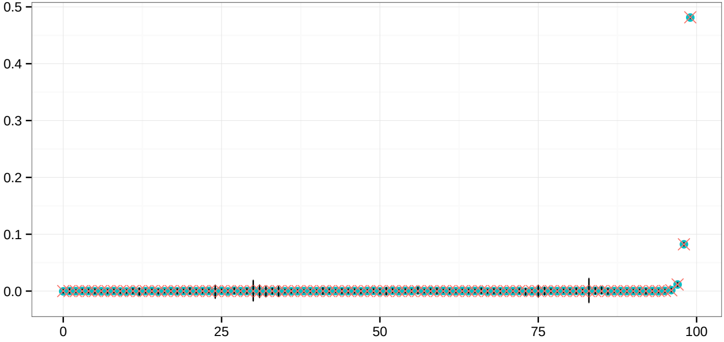
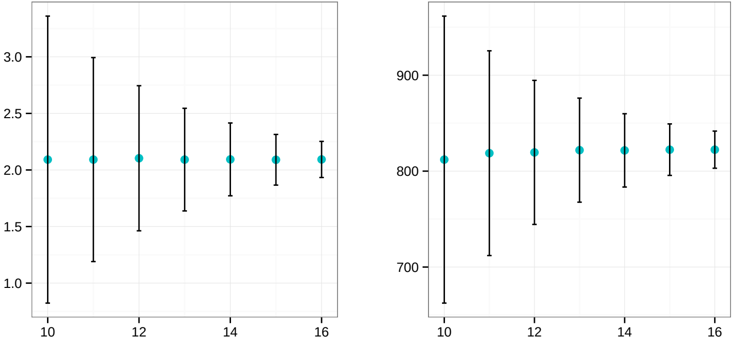


Figure: Plot of $\hat{v}^N_{p,n}(\varphi)$ (blue dots and error bars for the mean $\pm$ one standard deviation) and $\hat{v}_{p,n}(\varphi)$ (red crosses) at each $p \in \{0, \ldots, n\}$ with $\varphi = Id - \hat{\eta}^N_n(Id)$.

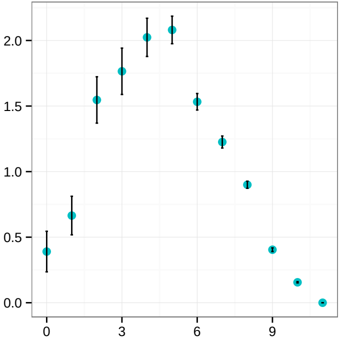
SMC sampler example: $NV_n^N(\varphi)$, estimating $\bar{\sigma}^2_{\gamma_n^N}(\varphi)$



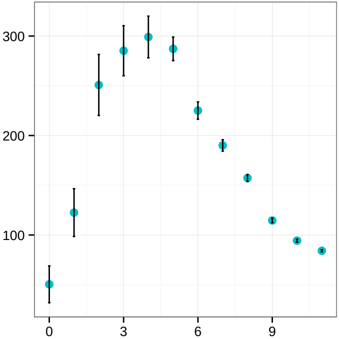
(a) $\varphi \equiv 1$ (b) $\varphi = Id - \eta_n^N(Id)$

Figure: Estimated asymptotic variances $NV_n^N(\varphi)$ (blue dots and error bars for the mean $\pm$ one standard deviation) against $\log_2 N$ for the SMC sampler example.

SMC sampler example: $v_{p,n}^N(\varphi)$, 1 iteration per kernel



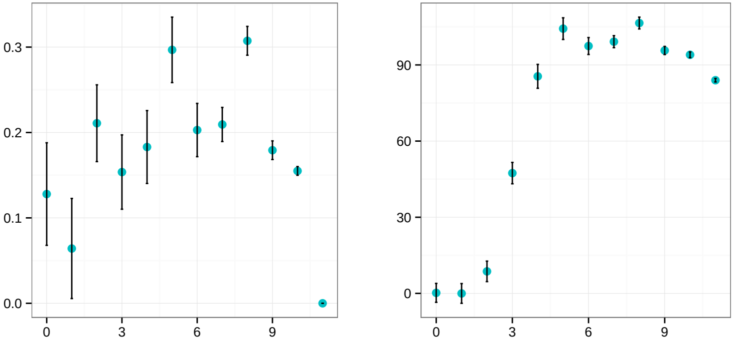
(a) $\varphi \equiv 1$



(b) $\varphi = Id - \eta_n(Id)$

Figure: Plot of $v_{p,n}^N(\varphi)$ (blue dots and error bars for the mean $\pm$ one standard deviation) at each $p \in \{0, \ldots, n\}$.

SMC sampler example: $v^N_{p,n}(\varphi)$, 10 iterations per kernel



(a) $\varphi \equiv 1$ (b) $\varphi = Id - \eta_n(Id)$

Figure: Plot of $v^N_{p,n}(\varphi)$ (blue dots and error bars for the mean $\pm$ one standard deviation) at each $p \in \{0, \ldots, n\}$.

Outline

Particle filters

Variance estimators

Applications of the estimators

Asymptotically optimal allocation

- Consider a particle filter with $\mathbf{N} = (N_p)_{p \geq 0}$ particles.
- At time p, for $i \in \{1, \ldots, N_p\}$,

$$A_{p-1}^i \sim \mathrm{Cat}\left(\frac{G_{p-1}(\zeta_{p-1}^1)}{\sum_{j=1}^{N_{p-1}} G_{p-1}(\zeta_{p-1}^j)}, \ldots, \frac{G_{p-1}(\zeta_{p-1}^{N_{p-1}})}{\sum_{j=1}^{N_{p-1}} G_{p-1}(\zeta_{p-1}^j)}\right),$$

  and $\zeta_p^i \sim M_p(\zeta_{p-1}^{A_{p-1}^i}, \cdot)$.
- In the regime $N_p = \lceil c_p N \rceil$, $(c_p)_{p \geq 0}$ fixed, and $N \to \infty$,

$$\sigma^2_{\gamma_n^{\mathbf{N}}}(\varphi) = \gamma_n(1)^2 \sum_{p=0}^{n} \frac{v_{p,n}(\varphi)}{c_p}, \qquad \sigma^2_{\eta_n^{\mathbf{N}}}(\varphi) = \bar{\sigma}^2_{\gamma_n^{\mathbf{N}}}(\varphi - \eta_n(\varphi)).$$

- Asymptotically optimal allocation: $c_p \propto \sqrt{v_{p,n}(\varphi)}$ or $c_p \propto \sqrt{v_{p,n}(\varphi - \eta_n(\varphi))}$ [see also Bhadra and Ionides, 2014].

Approximation of the asymptotically optimal allocation

- For some N run a particle filter with N particles.
- Set $c_p = \max\left\{v^N_{p,n}(\varphi), g(N)\right\}$, where $g \searrow 0$ as $N \to \infty$.
- Run a particle filter with $\mathbf{N} = (\lceil c_p N_p \rceil)_{p \geq 0}$ particles.

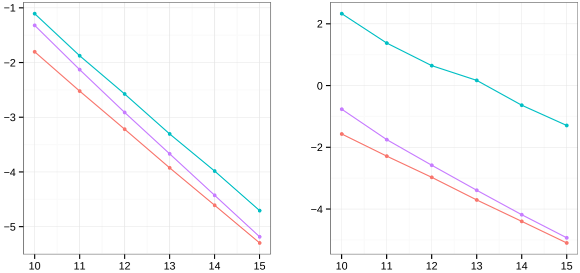


Figure: Linear Gaussian example. $\log_2 \mathrm{var}(\gamma^{\mathbf{N}}_n(1)/\gamma_n(1))$ against $\log_2 N$, using (blue) a constant N particle filter, (purple) the approximation of the a.o. particle filter, and (red) the a.o. particle filter. On the right a very simple observation seq. with an outlier is used.

An adaptive N particle filter

- Run particle filters, doubling N each time, until $V_n^N(\varphi) \in [0, \delta]$
- Run a particle filter with the successful N and report $\gamma_n^N(\varphi)$.
- May be useful in some applications, but it's possible for $\mathbb{P}(V_0^N(\varphi) \in [0, \delta]) \approx 1$ but $\mathrm{var}(\gamma_0^N(\varphi)) \gg \delta$.

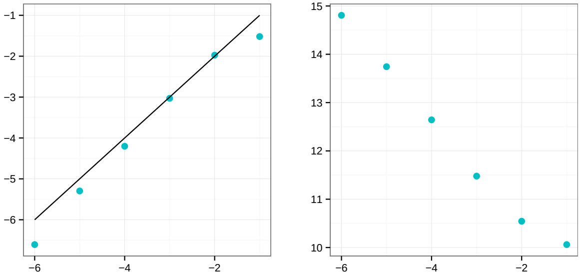


Figure: Linear Gaussian example. Left: $\log_2 \mathrm{var}(\hat{\gamma}_n^N(1)/\hat{\gamma}_n(1))$ against $\log_2 \delta$, with the straight line $y = x$. Right: $\log_2 N$ against $\log_2 \delta$, where N is the average number of particles used by the final particle filter.

What about i.i.d. replicates?

- For fixed N, consistent estimation of $\mathrm{var}(\gamma_n^N(\varphi)/\gamma_n(1))$ using sample variance and mean of i.i.d. replicates is straightforward.
- Lack-of-bias of $\gamma_n^N(1)^2 V_n^N(\varphi)$ allows an alternative estimate using replicates of $\gamma_n^N(1)$ and $V_n^N(\varphi)$.

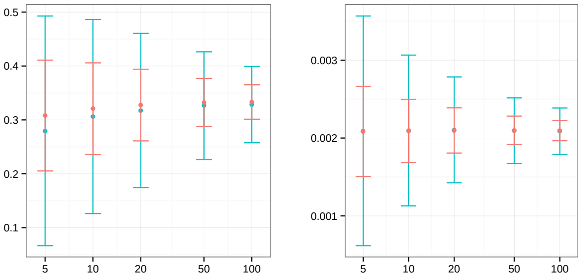


Figure: Plot of the standard estimate of $\mathrm{var}\left[\hat{\gamma}_n^N(\varphi)/\hat{\gamma}_n(1)\right]$ (blue) and the alternative estimate based on $\hat{V}_n^N(1)$ (red) against no. of replicates in the two examples, with $N = 10^3$.

References I

A. Bhadra and E. L. Ionides. Adaptive particle allocation in iterated sequential Monte Carlo via approximating meta-models. *Statistics and Computing*, pages 1–15, 2014.

F. Cérou, P. Del Moral, and A. Guyader. A nonasymptotic theorem for unnormalized Feynman–Kac particle models. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 47(3):629–649, 2011.

H. P. Chan and T. L. Lai. A general theory of particle filters in hidden Markov models and some applications. *Annals of Statistics*, 41(6):2877–2904, 2013.

N. Chopin. Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *Annals of Statistics*, 32(6):2385–2411, 2004.

P. Del Moral and A. Guionnet. Central limit theorem for nonlinear filtering and interacting particle systems. *Ann. Appl. Probab.*, 9(2):275–297, 1999.

P. Del Moral and L. Miclo. Genealogies and increasing propagation of chaos for Feynman-Kac and genetic models. *Ann. Appl. Probab.*, pages 1166–1198, 2001.

References II

R. Douc and E. Moulines. Limit theorems for weighted samples with applications to sequential Monte Carlo methods. *Ann. Statist.*, 36(5): 2344–2376, 2008.

H. Künsch. Recursive Monte Carlo filters: algorithms and theoretical analysis. *Annals of Statistics*, 33(5):1983–2021, 2005.