Asymptotic analysis of genealogies induced by sequential Monte Carlo algorithms

Suzie Brown

March 6, 2019

1 Introduction

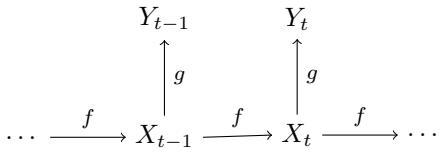
- what is SMC and what is it used for?
- basic SMC algorithm
- the three main tasks: prediction, filtering, smoothing
- the problem with smoothing: ancestral vs. weight degeneracy
- motivating plots
- organisation of the report

....

For our purposes it is sufficient to describe sequential Monte Carlo algorithms in the context of inference on a time-homogeneous hidden Markov model. Consider the following model:

$$
\begin{aligned}
&X_0 \sim \mu(\cdot) \\
&X_{t+1} \mid (X_t = x_t) \sim f(\cdot|x_t) \qquad t = 0, \dots, T-1 \\
&Y_t \mid (X_t = x_t) \sim g(\cdot|x_t) \qquad t = 0, \dots, T
\end{aligned}
$$

where $\{X_t\}_{t=0}^T$ is an unobservable Markov chain and the observables $\{Y_t\}_{t=0}^T$ satisfy $Y_t \perp \{Y_{-t}, X_{-t}\} \mid X_t$. The conditional independence structure is described by the following graphical model.



We assume that the *transition* and *emission* kernels have densities which are denoted by f and g respectively, but this is not necessary in general. We only require that we can sample from $\mu(\cdot)$ and $f(\cdot|x)$, and calculate *unnormalised* potentials $g(y|x)$, for all x, y.

2 SMC as a coalescent

- k-coalescent & Kingman coalescent
- pop gen literature about large population cts time limits of various models
- resampling viewed backwards in time: branching process $\to$ coalescent process
- asymptotic properties of SMC lit review: CLT, path storage, coalescence etc.
- the gap in knowledge that we aim to fill
- (remark: although SMC has other problems in high dimension, the coalescence rate doesn't depend on the dimension...)

3 Conditional SMC

- motivation: particle MCMC, need for multiple lineages
- conditional (multinomial) SMC algorithm and its context within particle Gibbs

- result: coalescence rate etc in terms of standard multinomial one; verification of assumptions of KJJS theorem (but exile horrible calculations to appendix)
- simulations & conclusions thence

4 Alternative resampling schemes

- multinomial not really used in practice, but other schemes hard to analyse
- overview of the main variance-reducing schemes
- results: theorem for residual resampling (hopefully)
- maybe results for other schemes
- simulation comparing all of them, and conclusions thence

5 Discussion

- results so far
- impact of this work: to practitioners, to enriching the SMC literature, interpretation within pop gen.
- future directions