Suzie Brown

2 November 2018

Genealogies of Sequential Monte Carlo Algorithms

Jere Koskela

Adam Johansen

Paul Jenkins

Hidden Markov model

$$
\begin{tikzcd}
& Y_{t-1} & Y_t & \\
\cdots \arrow[r, "f"'] & X_{t-1} \arrow[u, "g"] \arrow[r, "f"'] & X_t \arrow[u, "g"] \arrow[r, "f"'] & \cdots
\end{tikzcd}
$$

- $\{X_1, X_2, \dots\}$ hidden Markov states
- Y_i noisy observation of X_i
- Markov transition kernel f
- ‘emission distribution’ g

Inference on a HMM

- **filtering distribution** $p(x_T \mid y_{1:T})$
  'current state'
- **smoothing distribution** $p(x_{1:T} \mid y_{1:T})$
  'trajectory of all previous states'

Sequential Monte Carlo

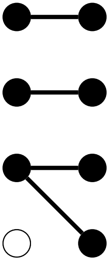
- Approximate these distributions using N particles
- Initialise, then iterate the steps:
  1. **propagate:** update positions of particles by applying the Markov kernel f
  2. **calculate weights:** weight the particles according to how well they agree with the observations
  3. **resample** resample particles proportionally to their weights
     ('good' particles multiply, 'bad' particles die out)

Genealogical interpretation

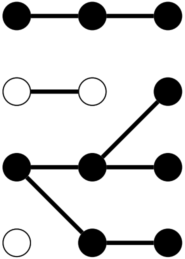
- Resampling step induces a genealogical (family tree) structure
- **Ancestral degeneracy:** genealogies of all current particles necessarily coalesce at some past time step
- Bad news for estimating smoothing distribution!

Genealogical interpretation

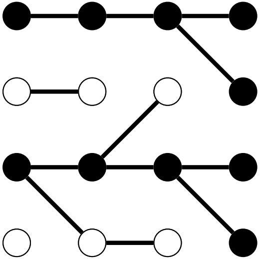
Genealogical interpretation



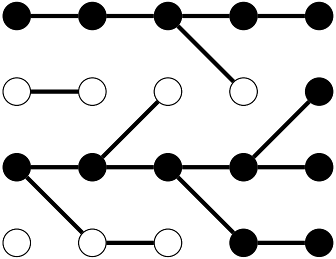
Genealogical interpretation



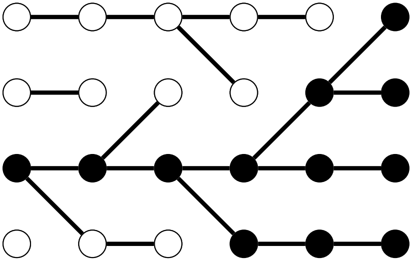
Genealogical interpretation



Genealogical interpretation



Genealogical interpretation



Genealogical interpretation

- How many particles do I need to ensure n distinct lineages remain in generation $T-t$ with probability greater than $1-\alpha$?

Genealogical interpretation

- How many particles do I need to ensure n distinct lineages remain in generation $T - t$ with probability greater than $1 - \alpha$?
- Easiest case (multinomial resampling) addressed in
  Koskela, Jere, et al. "Asymptotic genealogies of interacting particle systems with an application to sequential Monte Carlo." arXiv preprint arXiv:1804.01811 (2018).
- **Next:** relax assumptions, generalise to other resampling schemes and conditional SMC