Prelim topic: Likelihood-free MCMC

Eric Kernfeld[1]

[1]University of Washington Department of Statistics

Paper

Darren Wilkinson's "Parameter inference for stochastic kinetic models of bacterial gene regulation," a book chapter in [1].

Likelihood Free MCMC

Cast of Characters

- r_j is a menu item in a restaurant.
- x_t is the amount of money in the cash register t.
- $\mathcal{D}_t$ incomplete observation of x_t with error.
- c is the popularity of menu items.
- τ governs measurement error.
- θ is τ and c together.

Likelihood Free MCMC

To produce a chain of samples from $P(\theta|D)$, using a proposal $q(\theta^*|\theta)$, accept with probability $p_{rej}(\theta^*|\theta) \equiv \min\{1, A\}$ if

$$A = \frac{q(\theta, x|\theta^*, x^*)}{q(\theta^*, x^*|\theta, x)} \times \frac{P(\theta^*, x^*|\mathcal{D})}{P(\theta, x|\mathcal{D})}$$

Likelihood Free MCMC

Cast of Characters (things DW can't evaluate are in red)

- $P(\theta)$
- $P(x|\theta)$
- $P(\mathcal{D}|x,\theta)$
- $P(x,\theta|\mathcal{D})$

$$A = \frac{q(\theta, x|\theta^*, x^*)}{q(\theta^*, x^*|\theta, x)} \times \frac{P(\theta^*, x^*|\mathcal{D})}{P(\theta, x|\mathcal{D})}.$$

Quote: "Conditional on discrete-time observations, the Markov process breaks up into a collection of independent bridge processes that appear not to be analytically tractable."

Likelihood Free MCMC

$$
\begin{aligned}
&\frac{q(\theta^*, x^*|\theta, x)}{q(\theta, x|\theta^*, x^*)} \times \frac{P(x|\theta)}{P(x^*|\theta^*)} \times \frac{P(\theta)}{P(\theta^*)} \times \frac{P(\mathcal{D}|x, \theta)}{P(\mathcal{D}|x^*, \theta^*)} \\
&= \frac{f(\theta^*|\theta)}{f(\theta|\theta^*)} \times \frac{P(x^*|\theta^*)}{P(x|\theta)} \times \frac{P(x|\theta)}{P(x^*|\theta^*)} \times \frac{P(\theta)}{P(\theta^*)} \times \frac{P(\mathcal{D}|x, \theta)}{P(\mathcal{D}|x^*, \theta^*)} \\
&= \frac{f(\theta^*|\theta)}{f(\theta|\theta^*)} \times \frac{P(\theta)}{P(\theta^*)} \times \frac{P(\mathcal{D}|x, \theta)}{P(\mathcal{D}|x^*, \theta^*)}.
\end{aligned}
$$

Sequential Importance Sampling

I learned about SIS here [2].
For (non-sequential) importance sampling, get $\int g(x,\theta)P(x,\theta|\mathcal{D})dx$ by doing this.

- draw $(x^{(i)},\theta^{(i)})$ from Q
- compute $w^{(i)} = \frac{P(x^{(i)},\theta^{(i)}|\mathcal{D})}{Q(x^{(i)},\theta^{(i)})}$
- average out $\frac{\sum w^{(i)} g(x^{(i)},\theta^{(i)})}{\sum w^{(i)}}$

Again, you need the ratios $\frac{P(x^{(i)},\theta^{(i)})}{P(x^{(j)},\theta^{(j)})}$ and $\frac{Q(x^{(j)},\theta^{(j)})}{Q(x^{(i)},\theta^{(i)})}$.

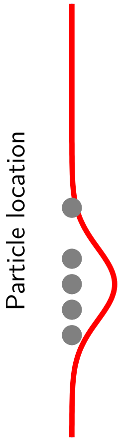
Sequential Monte Carlo

Importance sampling has low effective sample size when Q is dis-similar to P. (The location and form of g affects the variance also.)

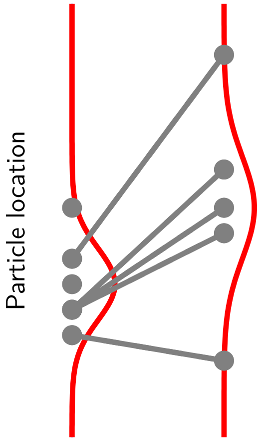
tex_says_fuck_you-eps-converted-to.pdf

The graphic on the following slides is from [3].

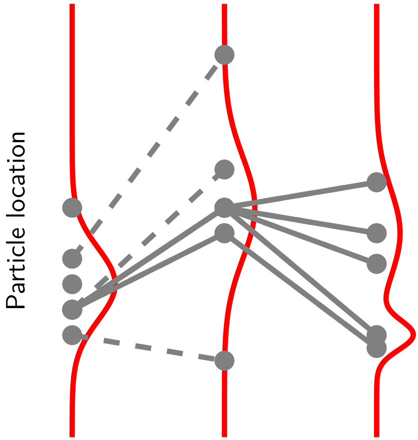
SMC: Sketch



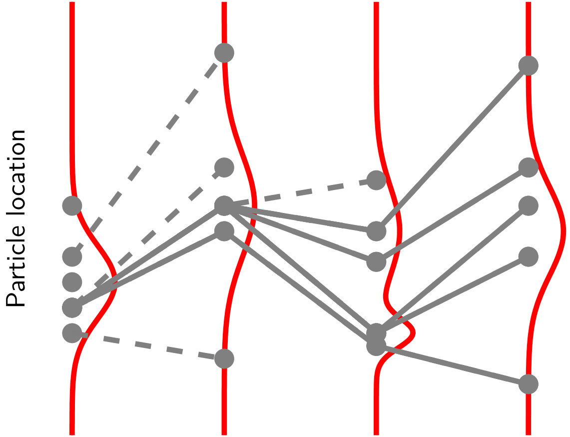
SMC: Sketch



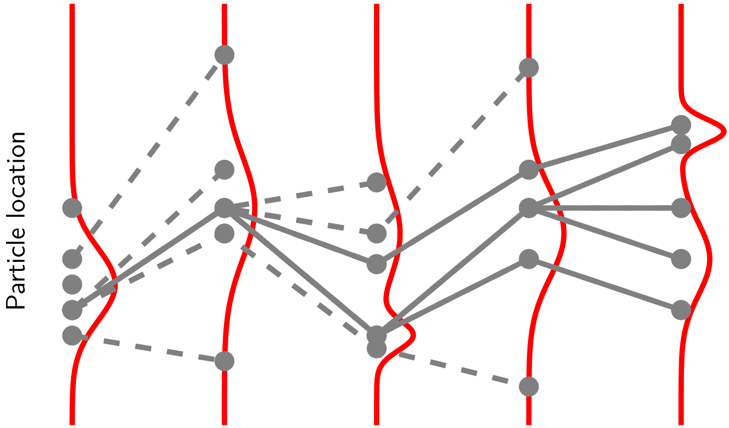
SMC: Sketch



SMC: Sketch



SMC: Sketch



Likelihood-free Sequential Importance Sampling

From an earlier slide: "You need the ratios $\frac{P(x^{(i)},\theta^{(i)})}{P(x^{(j)},\theta^{(j)})}$ and $\frac{Q(x^{(j)},\theta^{(j)})}{Q(x^{(i)},\theta^{(i)})}$." Not (quite) true.

From the intro: "If you can't compute $\frac{P(x^{(i)},\theta^{(i)})}{P(x^{(j)},\theta^{(j)})}$, try to get the nasty part to appear in the proposal ratio."

Questions?

Wilkinson's paper is a chapter from this book:

- Bernardo, J.M., Bayarri, M.J., Berger, J.O., Dawid, A.P., Heckerman, D., Smith, A.F.M., West, M.:
  Ninth Valencia international meeting on Bayesian statistics, Benidorm, Spain, 03-08.06.2010.
  Oxford U.P., Oxford (2012)

- Doucet, A., de Freitas, N., Gordon, N.
  Statistics for Engineering and Information Science. In: An Introduction to Sequential Monte Carlo Methods. Springer (2001) 3–14

- Finke, A.:
  Introduction to sequential monte carlo and particle mcmc methods (July 2013)