

10: Introduction to Bayesian Computation

10/28/19

1 Importance Sampling with Resampling

2 Sequential Monte Carlo

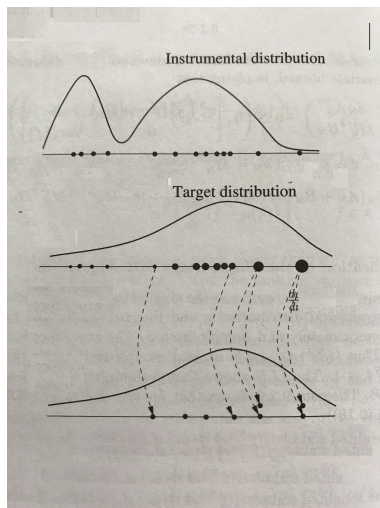
Adding Resampling

Importance Sampling gives you weighted draws $(\theta^1, w(\theta^1)), (\theta^2, w(\theta^2)), \dots$

You can draw from these, with replacement. At the expense of more variance, it will give you unweighted draws from your target distribution: $\tilde{\theta}^1, \tilde{\theta}^2, \dots$

This is known as **factored sampling** or **importance sampling with resampling** or **sampling importance resampling (SIR)**.

Adding Resampling



Original, unedited image is from

<https://www.springer.com/us/book/9780387402642>

Stage 1: do importance sampling to get $\{\theta^i, w(\theta^i)\}_{i=1}^S$.

Stage 2: for $i = 1, \dots, S$, select

$$P[\tilde{\theta}^i = \theta^j \mid \theta^1, w(\theta^1), \dots, \theta^S, w(\theta^S)] = \frac{w(\theta^j)}{\sum_k w(\theta^k)}.$$

Example 2 revisited

```
y <- 2 # fake data
log_unnorm_weight <- function(theta){
  # can ignore sqrt(2pi) because it will cancel out
  -.5*(y - theta)^2 }
num_samples <- 10000
theta_draws <- rt(num_samples , 1)
lunws <- log_unnorm_weight(theta_draws)
# note: prob arg automatically normalizes
random_indexes <- sample(x = num_samples,
                        size = num_samples,
                        replace = T,
                        prob = exp(lunws))
sort(random_indexes) # see there are repeats!
resampled_draws <- theta_draws[random_indexes]
hist(resampled_draws) # can't do this unless we resample
```

Other uses of importance (re)sampling

- Use importance resampling for obtaining starting points for an iterative simulation of the posterior distribution, as in Chapter 11.

Other uses of importance (re)sampling

- Use importance resampling for obtaining starting points for an iterative simulation of the posterior distribution, as in Chapter 11.
- Calculation of intractable quantity under a change of model.

Posterior under the previous model: $p_1(\theta|y)$

$$\theta^1, \dots, \theta^S \sim p_1(\theta|y)$$

Posterior after a change of model: $p_2(\theta|y) \propto q_2(\theta|y)$

We need $E_2(h(\theta)|y)$

Other uses of importance (re)sampling

- Use importance resampling for obtaining starting points for an iterative simulation of the posterior distribution, as in Chapter 11.
- Calculation of intractable quantity under a change of model.

Posterior under the previous model: $p_1(\theta|y)$

$$\theta^1, \dots, \theta^S \sim p_1(\theta|y)$$

Posterior after a change of model: $p_2(\theta|y) \propto q_2(\theta|y)$

We need $E_2(h(\theta)|y)$

Importance sampling with $g(\theta) = q_2(\theta|y)$, $\tilde{w}(\theta) = q_2(\theta|y)/q_1(\theta|y)$,
 $w(\theta^s) = \tilde{w}(\theta^s) / \sum_{i=1}^S \tilde{w}(\theta^i)$

$$E_2(h(\theta)|y) = \sum_{i=1}^S h(\theta^s) w(\theta^s) /$$

Going sequential

Resampling adds variance, so why do it?

It throws away bad samples, and duplicates promising ones.

When you're looking at a sequence of distribution targets, this can have a good effect on future samples.

Going sequential

Resampling adds variance, so why do it?

It throws away bad samples, and duplicates promising ones.

When you're looking at a sequence of distribution targets, this can have a good effect on future samples.

sequential monte carlo or **particle filtering** methods are basically doing SIR over and over again.

At “time” $t - 1$ you just resampled, so you have draws $\tilde{\theta}_{t-1}^1, \tilde{\theta}_{t-1}^2, \dots, \tilde{\theta}_{t-1}^S$, and you want to turn them into draws for the next “time” period:

$$\tilde{\theta}_t^1, \tilde{\theta}_t^2, \dots, \tilde{\theta}_t^S.$$

Examples of sequences of distributions

Data annealing¹

$$p(\theta), p(\theta \mid y_1), p(\theta \mid y_{1:2}), \dots, p(\theta \mid y_{1:n}),$$

Temperature annealing²

$$p(y \mid \theta)^{a_0} p(\theta), p(y \mid \theta)^{a_1} p(\theta), \dots p(y \mid \theta)^{a_n} p(\theta)$$

with $0 = a_0 < a_1 < \dots < a_n = 1$.

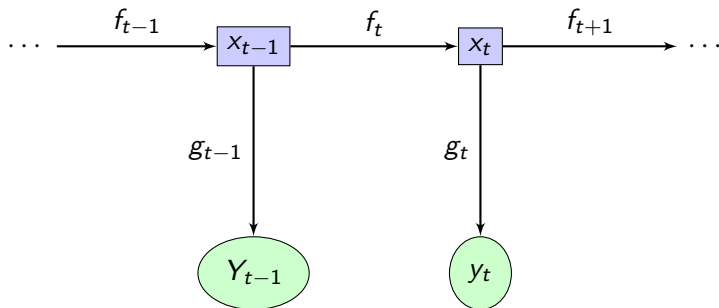
filtering and smoothing in state space models

$$p(x_1 \mid y_1, \theta), \dots, p(x_n \mid y_{1:n}, \theta)$$

¹Chopin, *A Sequential Particle Filter Method for Static Models*.

²Neal, "Annealed Importance Sampling".

state space models



$$p(x_{1:T}, y_{1:T} \mid \theta)$$

$$= g(y_1 \mid x_1, \theta) f(x_1 \mid \theta) \prod_{t=2}^T g(y_t \mid x_t, \theta) f(x_t \mid x_{t-1}, \theta)$$

Example: filtering in state space models

Here's an example of a state space model. y_t is a univariate time series, and x_t is a hidden/unobserved/latent time series.

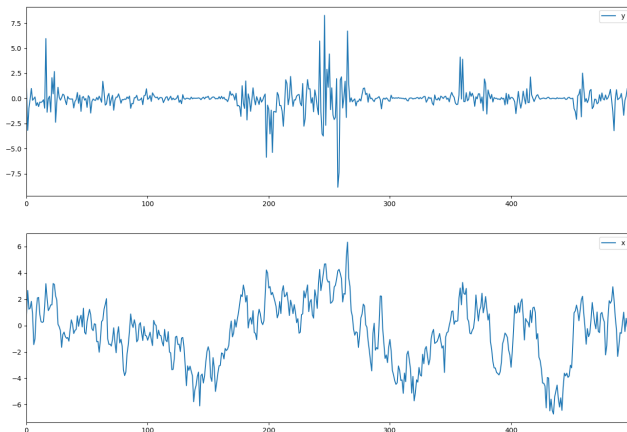
$$y_t = \exp(x_t/2)\epsilon_t \quad (1)$$

$$x_t = c + \phi x_{t-1} + v_t \quad (2)$$

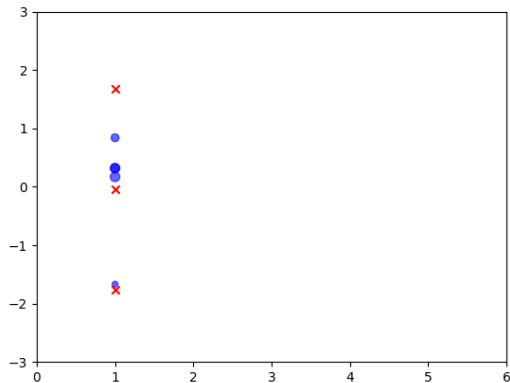
We sometimes refer to (1) as $g(y_t | x_t, \theta)$ or the observation equation, and (2) as the state transition equation or $f(x_t | x_{t-1}, \theta)$.

Example: filtering in state space models

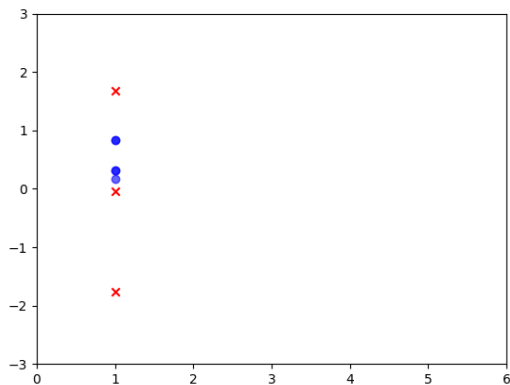
$y_{1:t}$ observed, $x_{1:t}$ hidden. Goal: $p(x_t \mid y_{1:t})$ in real-time.



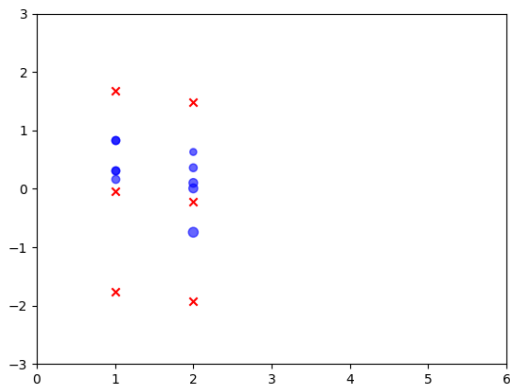
Example: filtering in state space models



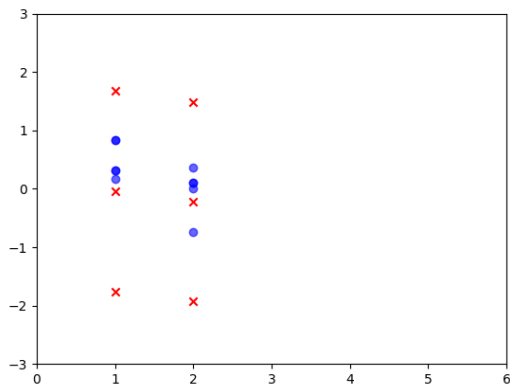
Going sequential



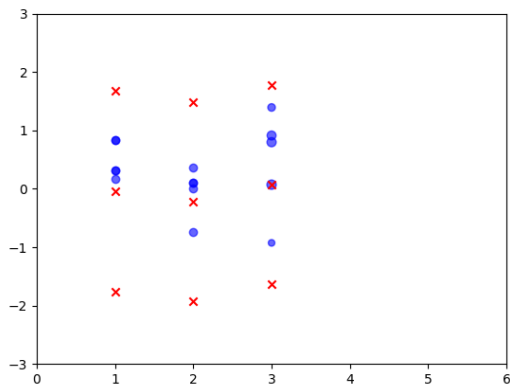
Going sequential



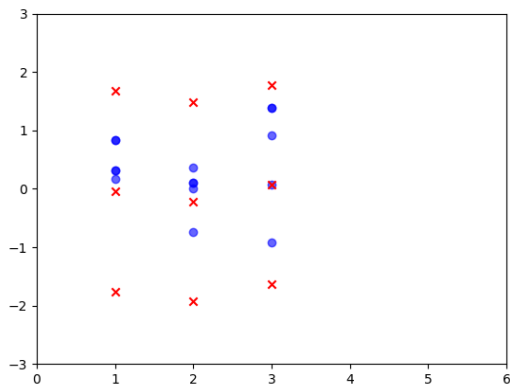
Going sequential



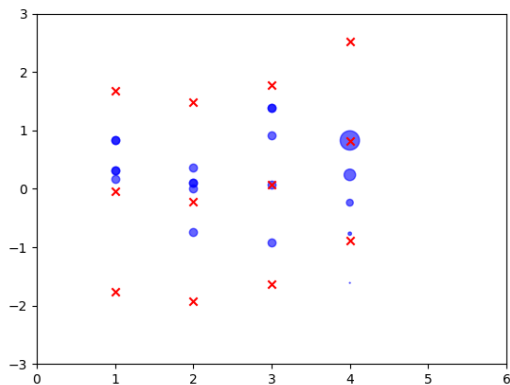
Going sequential



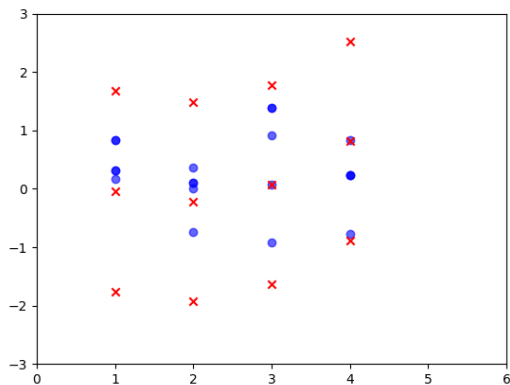
Going sequential



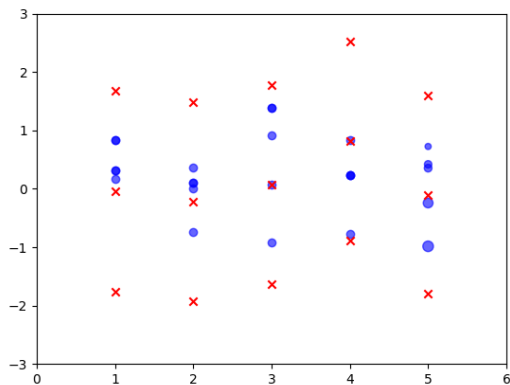
Going sequential



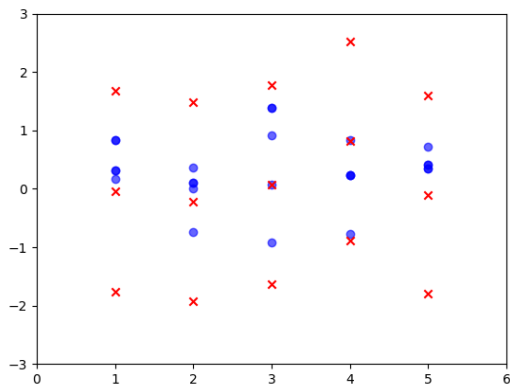
Going sequential



Going sequential



Going sequential



Filtering recursions

Drop dependence on θ from the notation...

$$\begin{aligned} p(x_{1:t}|y_{1:t}) &= C_t^{-1} p(x_t, y_t \mid x_{t-1}) p(x_{1:t-1} \mid y_{1:t-1}) \\ &= C_t^{-1} \frac{p(x_t, y_t \mid x_{t-1})}{q_t(x_t \mid y_t, x_{t-1})} \times \\ &\quad q_t(x_t \mid y_t, x_{t-1}) p(x_{1:t-1} \mid y_{1:t-1}) \\ &= C_t^{-1} \frac{g(y_t|x_t)f(x_t|x_{t-1})}{q_t(x_t|x_{t-1}, y_t)} \times \\ &\quad q_t(x_t \mid x_{t-1}, y_t) p(x_{1:t-1} \mid y_{1:t-1}) \end{aligned}$$

Repeat through time:

- 1 start with samples from $p(x_{1:t-1} \mid y_{1:t-1})$
- 2 mutate/propagate/extend using $q_t(x_t \mid x_{t-1}, y_t)$
- 3 adjust weights by multiplying by $\frac{g(y_t|x_t)f(x_t|x_{t-1})}{q_t(x_t|x_{t-1}, y_t)}$
- 4 resample, giving you particles distributed as $p(x_{1:t} \mid y_{1:t})$

Chopin, Nicolas. *A Sequential Particle Filter Method for Static Models*. 2000.

Neal, Radford M. "Annealed Importance Sampling". In: *Statistics and Computing* 11.2 (Apr. 2001), pp. 125–139. ISSN: 0960-3174. DOI: 10.1023/A:1008923215028. URL: <https://doi.org/10.1023/A:1008923215028>.