FFBS

Step 1. Run particle filtering to obatin, $\{x_{1:T}^{1:K}, w_{1:T}^{1:K}\}$

Step 2. Run the following algorithm to obtain M sample realizations and their associated weights$\{\tilde{x}_{1:T}^{1:M}, \tilde{w}^{1:M}\}$

See section 3 in ref

Algorithm 1 (Sample realizations).

1. Choose $\widetilde{x}_T = x_T^{(i)}$ with probability $w_T^{(i)}$.
2. For $t = T - 1$ to 1:
    - Calculate $w_{t|t+1}^{(i)} \propto w_t^{(i)} f(\widetilde{x}_{t+1}|x_t^{(i)})$ for each $i = 1, \ldots, N$.
    - Choose $\widetilde{x}_t = x_t^{(i)}$ with probability $w_{t|t+1}^{(i)}$.
3. $\widetilde{\mathbf{x}}_{1:T} = (\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_T)$ is an approximate realization from $p(x_{1:T}|y_{1:T})$.

Note that the weights are computed as follows:

$p(x_{1:T}|y_{1:T}) = p(x_T|y_{1:T}) \prod_{t=1}^{T-1} p(x_t|x_{t+1:T}, y_{1:T})$.

Then the weight asccociated with $\tilde{x}_{1:T}$ is $\tilde{w} = w_T^{(i_T)} \prod_{t=T-1}^{1} w_{t|t+1}^{t_i}$, where t_i denotes the index of the selected particle.

Step 3.

Method 1, FFBS_score_loss

(In runner_flag.py, set the flag FFBS_score_loss=True)

Compute the surrogate loss as $\frac{1}{M} \sum_{i=1}^{M} \log p_\theta(\tilde{x}_{1:T}^i, y_{1:T})$.

Then the gradient is $\frac{1}{M} \sum_{i=1}^{M} \nabla \log p_\theta(\tilde{x}_{1:T}^i, y_{1:T})$

Method 2. ELBO-stype

(In runner_flag.py, set the flag FFBS_score_loss=False)

Compute the surrogate loss as $\frac{1}{M}\sum_{i=1}^{M}[\log p_\theta(\tilde{x}^i_{1:T}, y_{1:T}) - \log \tilde{w}^i]$

Then directly evaluate the gradient.

Question:

surrogate loss should be computed as $\tilde{w}^i \sum_{i=1}^{M}[\log p_\theta(\tilde{x}^i_{1:T}, y_{1:T}) - \log \tilde{w}^i]$?

But if it is this case, using Jensen's inequality,

$\tilde{w}^i \sum_{i=1}^{M}[\log p_\theta(\tilde{x}^i_{1:T}, y_{1:T}) - \log \tilde{w}^i] \leq \log \sum_{i=1}^{m} \tilde{w}^i \frac{p_\theta(\tilde{x}^i_{1:T}, y_{1:T})}{\tilde{w}^i}$

the RHS is just FFBS_score_loss