Sequential Monte Carlo methods

Lecture 6 – Auxiliary particle filters

Fredrik Lindsten, Linköping University

2025-02-04

Outline – Lecture 6

Aim: Show how we can improve the proposals for the particle filter by using auxiliary variables.

Outline:

1. Summary of day 1
2. Auxiliary variables
3. Ancestor indices as auxiliary variables
4. Improving the proposal distributions

Summary of day 1

Summary of day 1

A state space model can be expressed using probability densities as.

$$X_t \mid (X_{t-1} = x_{t-1}) \sim p(x_t \mid x_{t-1}),$$
$$Y_t \mid (X_t = x_t) \sim p(y_t \mid x_t),$$
$$X_0 \sim p(x_0).$$

The filtering problem amounts to computing the filter PDF $p(x_t \mid y_{1:t})$.

Solution conceptually given by,

$$p(x_t \mid y_{1:t}) = \frac{p(y_t \mid x_t)p(x_t \mid y_{1:t-1})}{p(y_t \mid y_{1:t-1})},$$
$$p(x_t \mid y_{1:t-1}) = \int p(x_t \mid x_{t-1})p(x_{t-1} \mid y_{1:t-1})\mathrm{d}x_{t-1}.$$

Monte Carlo: approximate an unknown distribution of interest by

$$\pi(x) \approx \widehat{\pi}^N(x) = \sum_{i=1}^{N} w^i \delta_{x^i}(x).$$

Summary of day 1

Monte Carlo: approximate an unknown distribution of interest by

$$\pi(x) \approx \widehat{\pi}^N(x) = \sum_{i=1}^{N} w^i \delta_{x^i}(x).$$

Importance sampling: For $i = 1, \ldots, N$,

1. Sample $x^i \sim q(x)$,
2. Compute $\widetilde{w}^i = \widetilde{\pi}(x^i)/q(x^i)$,
3. Normalize $w^i = \frac{\widetilde{w}^i}{\sum_{j=1}^{N} \widetilde{w}^j}$.

Summary of day 1

Bootstrap particle filter: sequentially using importance sampling to approximate the filter update equations, with proposal distribution at time t given by

$$q(x_t \mid y_{1:t}) = \sum_{i=1}^{N} w_{t-1}^{i} p(x_t \mid x_{t-1}^{i}).$$

The resulting importance weights are,

$$\widetilde{w}_t^i = p(y_t \mid x_t^i).$$

Summary of day 1

Bootstrap particle filter: sequentially using importance sampling to approximate the filter update equations, with proposal distribution at time t given by

$$q(x_t \mid y_{1:t}) = \sum_{i=1}^{N} w_{t-1}^i p(x_t \mid x_{t-1}^i).$$

The resulting importance weights are,

$$\widetilde{w}_t^i = p(y_t \mid x_t^i).$$

1. Resample particles according to their importance weights
2. Simulate new particles according to the system dynamics
3. Compute weights according to the measurement likelihood

Summary of day 1

Bootstrap particle filter: sequentially using importance sampling to approximate the filter update equations, with proposal distribution at time t given by

$$q(x_t \mid y_{1:t}) = \sum_{i=1}^{N} w_{t-1}^{i} p(x_t \mid x_{t-1}^{i}).$$

The resulting importance weights are,

$$\widetilde{w}_t^i = p(y_t \mid x_t^i).$$

1. Resample particles according to their importance weights
2. Simulate new particles according to the system dynamics
3. Compute weights according to the measurement likelihood

Under forgetting conditions, errors **do not** accumulate unboundedly with time — the bootstrap particle filter is **stable**

Auxiliary variables

Auxiliary variables

Target distribution: $\pi(x)$, difficult to sample from

Idea: Introduce another variable U with conditional distribution $\pi(u \mid x)$

Auxiliary variables

Target distribution: $\pi(x)$, difficult to sample from

Idea: Introduce another variable U with conditional distribution $\pi(u \mid x)$

> The joint distribution $\pi(x, u) = \pi(u \mid x)\pi(x)$ admits $\pi(x)$ as a marginal by construction, i.e., $\int \pi(x, u) \mathrm{d}u = \pi(x)$.

Auxiliary variables

Target distribution: $\pi(x)$, difficult to sample from

Idea: Introduce another variable U with conditional distribution $\pi(u \mid x)$

The joint distribution $\pi(x, u) = \pi(u \mid x)\pi(x)$ admits $\pi(x)$ as a marginal by construction, i.e., $\int \pi(x, u) \mathrm{d}u = \pi(x)$.

Sampling from the joint $\pi(x, u)$ may be easier than directly sampling from the marginal $\pi(x)$!

Auxiliary variables

Target distribution: $\pi(x)$, difficult to sample from

Idea: Introduce another variable U with conditional distribution $\pi(u \mid x)$

The joint distribution $\pi(x, u) = \pi(u \mid x)\pi(x)$ admits $\pi(x)$ as a marginal by construction, i.e., $\int \pi(x, u)\mathrm{d}u = \pi(x)$.

Sampling from the joint $\pi(x, u)$ may be easier than directly sampling from the marginal $\pi(x)$!

The variable U is an auxiliary variable. It may have some “physical” interpretation (an unobserved measurement, unknown model parameter, ...) but this is not necessary.

ex) Auxiliary variables

Importance sampling setting with target $\pi(x)$ and proposal $q(x)$. We assume $\widetilde{\pi}(x) \leq q(x)$.

ex) Auxiliary variables

Importance sampling setting with **target** $\pi(x)$ and **proposal** $q(x)$. We **assume** $\widetilde{\pi}(x) \leq q(x)$.

To sample from $\pi(x)$, introduce an auxiliary variable

$$U \mid (X = x) \sim \mathcal{U}(0, \widetilde{\pi}(x)).$$

ex) Auxiliary variables

Importance sampling setting with target $\pi(x)$ and proposal $q(x)$. We assume $\widetilde{\pi}(x) \leq q(x)$.

To sample from $\pi(x)$, introduce an auxiliary variable

$$U \mid (X = x) \sim \mathcal{U}(0, \widetilde{\pi}(x)).$$

Joint target: $\pi(u, x) = \pi(u \mid x)\pi(x) = \mathcal{U}(u \mid 0, \widetilde{\pi}(x))\pi(x)$

ex) Auxiliary variables

Importance sampling setting with target $\pi(x)$ and proposal $q(x)$. We assume $\widetilde{\pi}(x) \leq q(x)$.

To sample from $\pi(x)$, introduce an auxiliary variable

$$U \mid (X = x) \sim \mathcal{U}(0, \widetilde{\pi}(x)).$$

Joint target: $\pi(u, x) = \pi(u \mid x)\pi(x) = \mathcal{U}(u \mid 0, \widetilde{\pi}(x))\pi(x)$
Joint proposal: $q(u, x) = q(u \mid x)q(x) = \mathcal{U}(u \mid 0, q(x))q(x)$

ex) Auxiliary variables

Importance sampling setting with target $\pi(x)$ and proposal $q(x)$. We assume $\widetilde{\pi}(x) \leq q(x)$.

To sample from $\pi(x)$, introduce an auxiliary variable

$$U \mid (X = x) \sim \mathcal{U}(0, \widetilde{\pi}(x)).$$

Joint target: $\pi(u, x) = \pi(u \mid x)\pi(x) = \mathcal{U}(u \mid 0, \widetilde{\pi}(x))\pi(x)$
Joint proposal: $q(u, x) = q(u \mid x)q(x) = \mathcal{U}(u \mid 0, q(x))q(x)$

The weights are,

$$\frac{\mathcal{U}(u \mid 0, \widetilde{\pi}(x))}{\mathcal{U}(u \mid 0, q(x))} \frac{\widetilde{\pi}(x)}{q(x)} = \mathbb{1}(u \leq \widetilde{\pi}(x)) \frac{q(x)}{\widetilde{\pi}(x)} \frac{\widetilde{\pi}(x)}{q(x)} = \mathbb{1}(u \leq \widetilde{\pi}(x))$$

ex) Auxiliary variables

Importance sampling setting with target $\pi(x)$ and proposal $q(x)$. We assume $\widetilde{\pi}(x) \le q(x)$.

To sample from $\pi(x)$, introduce an auxiliary variable

$$U \,|\, (X = x) \sim \mathcal{U}(0, \widetilde{\pi}(x)).$$

Joint target: $\pi(u, x) = \pi(u \,|\, x)\pi(x) = \mathcal{U}(u \,|\, 0, \widetilde{\pi}(x))\pi(x)$
Joint proposal: $q(u, x) = q(u \,|\, x)q(x) = \mathcal{U}(u \,|\, 0, q(x))q(x)$

The weights are,

$$\frac{\mathcal{U}(u \,|\, 0, \widetilde{\pi}(x))}{\mathcal{U}(u \,|\, 0, q(x))} \frac{\widetilde{\pi}(x)}{q(x)} = \mathbb{1}(u \le \widetilde{\pi}(x)) \frac{q(x)}{\widetilde{\pi}(x)} \frac{\widetilde{\pi}(x)}{q(x)} = \mathbb{1}(u \le \widetilde{\pi}(x))$$

In fact, conditionally on $\widetilde{w}^i = 1$, a sample x^i is an exact draw from $\pi(x)$ — referred to as **rejection sampling.**

Auxiliary particle filter

Sampling from the joint proposal

Sampling from the **joint proposal** $q(x_t, a_t \mid y_{1:t}) = \nu_{t-1}^{a_t} q(x_t \mid x_{t-1}^{a_t}, y_t)$:

1. Sample the auxiliary variable (**resampling**),

$$a_t^i \sim \mathcal{C}(\{\nu_{t-1}^j\}_{j=1}^N).$$

2. Sample x_t conditionally on the auxiliary variable (**propagation**),

$$x_t^i \sim q(x_t \mid x_{t-1}^{a_t^i}, y_t).$$

Repeat N times for $i = 1, \ldots, N$.

Sampling from the joint proposal

Sampling from the joint proposal $q(x_t, a_t \mid y_{1:t}) = \nu_{t-1}^{a_t} q(x_t \mid x_{t-1}^{a_t}, y_t)$:

1. Sample the auxiliary variable (resampling),

$$a_t^i \sim \mathcal{C}(\{\nu_{t-1}^j\}_{j=1}^N).$$

2. Sample x_t conditionally on the auxiliary variable (propagation),

$$x_t^i \sim q(x_t \mid x_{t-1}^{a_t^i}, y_t).$$

Repeat N times for $i = 1, \ldots, N$.

Algorithmically, sampling from the proposal is done exactly in the same way as before!

Computing the weights

The importance weights are given by the ratio between the joint target and joint proposal,

$$\widetilde{w}_t^i = \frac{w_{t-1}^{a_t^i}}{\nu_{t-1}^{a_t^i}} \frac{p(y_t \mid x_t^i) p(x_t^i \mid x_{t-1}^{a_t^i})}{q(x_t^i \mid x_{t-1}^{a_t^i}, y_t)}, \qquad i = 1, \ldots, N.$$

The weights can be computed in $O(N)$ computational time for quite arbitrary choices of $\{\nu_{t-1}^i\}_{i=1}^N$ and $q(\cdot)$.

Computing the weights

The importance weights are given by the ratio between the joint target and joint proposal,

$$\widetilde{w}_t^i = \frac{w_{t-1}^{a_t^i}}{\nu_{t-1}^{a_t^i}} \frac{p(y_t \mid x_t^i) p(x_t^i \mid x_{t-1}^{a_t^i})}{q(x_t^i \mid x_{t-1}^{a_t^i}, y_t)}, \qquad i = 1, \ldots, N.$$

The weights can be computed in $O(N)$ computational time for quite arbitrary choices of $\{\nu_{t-1}^i\}_{i=1}^N$ and $q(\cdot)$.

Note that the resampling weights $\{\nu_{t-1}^i\}_{i=1}^N$

- can be different from the importance weights $\{w_{t-1}^i\}_{i=1}^N$,
- may depend on $\{x_{t-1}^i\}_{i=1}^N$ as well as on y_t.

Algorithm 1 Auxiliary particle filter (for $i = 1, \dots, N$)

1. **Initialization ($t = 0$):**
   (a) Sample $x_0^i \sim p(x_0)$.
   (b) Set initial weights: $w_0^i = 1/N$.
2. **for** $t = 1$ **to** T **do**
   (a) Resample: sample ancestor indices $a_t^i \sim \mathcal{C}(\{\nu_{t-1}^j\}_{j=1}^N)$.
   (b) Propagate: sample $x_t^i \sim q(x_t \mid x_{t-1}^{a_t^i}, y_t)$.
   (c) Weight: compute

$$\widetilde{w}_t^i = \frac{w_{t-1}^{a_t^i}}{\nu_{t-1}^{a_t^i}} \frac{p(y_t \mid x_t^i) p(x_t^i \mid x_{t-1}^{a_t^i})}{q(x_t^i \mid x_{t-1}^{a_t^i}, y_t)}$$

   and normalize $w_t^i = \widetilde{w}_t^i / \sum_{j=1}^N \widetilde{w}_t^j$.

Selecting the proposals

How do we select the proposals?

There is freedom in selecting the resampling weights $\{\nu_{t-1}^i\}_{i=1}^N$ and proposal $q(\cdot)$. **How are they chosen in practice?!**

How do we select the proposals?

There is freedom in selecting the resampling weights $\{\nu_{t-1}^i\}_{i=1}^N$ and proposal $q(\cdot)$. How are they chosen in practice?!

With $\nu_{t-1}^i = w_{t-1}^i$ and $q(x_t \,|\, x_{t-1}, y_t) = p(x_t \,|\, x_{t-1})$ we recover exactly the bootstrap particle filter.

How do we select the proposals?

There is freedom in selecting the resampling weights $\{\nu_{t-1}^i\}_{i=1}^N$ and proposal $q(\cdot)$. **How are they chosen in practice?!**

With $\nu_{t-1}^i = w_{t-1}^i$ and $q(x_t \mid x_{t-1}, y_t) = p(x_t \mid x_{t-1})$ we recover exactly the bootstrap particle filter.

Is it possible to select the proposals so that $w_t^i \equiv \frac{1}{N}$?

A few concepts to summarize lecture 6

Auxiliary variable: a variable by which the target distribution is extended to improve efficiency or enable sampling from the target.

Ancestor index: auxiliary variable used in the particle filter, representing one of the components in the mixture target distribution.

Auxiliary particle filter: particle filter explicitly using the ancestor indices as auxiliary variables.