

TransportFilter for Target Tracking

Junjie Wang

Contents

1	local couplings nonlinear filter	1
1.1	Transport maps from samples	2
1.2	LocLF	2
1.2.1	local assimilation	2
1.2.2	propagation	3
2	SVGD For Target Tracking	3

1 local couplings nonlinear filter

In section 6 of [1], the author has proposed a nonlinear filter named **LocLF** and **LocNLF** rely on the low-dimensional deterministic couplings. I will give a brief introduction about this filter in the following. The idear of **LocLF** is a generalization of the EnKF philosophy. [The locLF seeks to transorm the forecast ensemble into samples from the filtering distribution by means of a sequence of local, low-dimensional, and possibly nonlinear deterministic couplings \[1\].](#)

1.1 Transport maps from samples

Assume that we have M samples $(z^i)_{i=1}^M$ from a target distribution π on \mathbb{R}^n . Given a reference density η on \mathbb{R}^n , our goal is to compute a triangular transport map S , that pushes forward the target to the reference density, i.e., $S_\# \pi = \eta$. Given a map S , we can easily sample from π by inverting the map at samples from η .

1.2 LocLF

Assumed that we have a collection of random variables, $Z = (Z_1, \dots, Z_n)$ and Y , with joint distribution $\pi_{Z_{1:n}, Y}$, and local likelihood function, i.e., $\pi_{Y|Z_{1:n}} = \pi_{Y|Z_1}$. The aim of **LocLF** is to generate approximate samples from the conditional $\pi_{Z|Y}$ given $Y = y$. The filtering distribution $\pi_{Z_{1:n}|Y}$ can be factorized as

$$\pi_{Z_{1:n}|Y} \propto \pi_{Z_1|Y} \pi_{Z_{2:n}|Z_1} \quad (1)$$

The **LocLF** contains two steps: local assimilation and propagation.

1.2.1 local assimilation

To sample from $\pi_{Z_1|Y}$, we need to estimate the marginal density π_{Z_1} in order to evaluate the local posterior

$$\pi_{Z_1|Y} \propto \pi_Y |_{Z_1} \pi_{Z_1} \quad (2)$$

The **LocLF** seeks an inverse map S^1 that pushes forward π_{Z_1} to a one dimensional reference density η_{X_1} . Next, the **LocLF** compute a transport map T^1 that samples the approximation of $\pi_{Z_1|Y}$. Given the pair of S^1 and T^1 , the approximate samples $(\alpha_1^i)_{i=1}^M$ from $\pi_{Z_1|Y}$ by

$$\alpha_1^i = T^1(S^1(z_1^i)) \quad (3)$$

*The **LocLF** does not sample $\pi_{Z_1|Y}$ by pushing forward samples from η_{X_1} through T^1 , but rather update each forecast particle directly using the composition $T^1 \circ S^1$.*

1.2.2 propagation

The goal of propagation steps is to sample from $Z_{2:n|Z_1}$ given the events $\{Z_1^i = \alpha_1^i\}_{i=1}^M$. In other words, the propagations steps need to propagate information from the local filtering marginal $\pi_{Z_1|Y}$ to every remaining state variable. A simple solution is to compute a triangular map S ,

$$S(x) = \begin{bmatrix} S^1(x_1) \\ S^2(x_1, x_2) \\ \dots \\ S^n(x_1, x_2, \dots, x_n) \end{bmatrix} \quad (4)$$

that pushes forward $\pi_{Z_{1:n}}$ to a reference density and compute the " marginal filtering samples $\alpha_{2:n}^i$ as

$$\alpha_{2:n}^i = S_{\alpha_1^i}^{-1}(S_{z_1^i}(z_{2:n}^i)) \quad (5)$$

Finally, by concatenating α_1^i and $\alpha_{2:n}^i$, the approximate filtering sample as

$$\alpha^i = \begin{bmatrix} \alpha_1^i \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n^i \end{bmatrix} \quad (6)$$

The **locLF** is more like a decompose method such as RBPF filter.

2 SVGD For Target Tracking

Stein Variational Gradient Descent (SVGD) [2] can be used to estimate a distribution by a set of particles. By iteratively transporting samples from an initial simple distribution in the direction of the likelihood, the SVGD can generate the posterior distribution.

The particles are subject to the following gradient ascent procedure.

$$x_i^{l+i} \leftarrow x_i^l + \epsilon_l \phi^*(\hat{x}_i^l)$$

$$\hat{\phi}^*(x) = \frac{1}{n} \sum_{j=1}^n [k(x_j^l, x) \nabla_{x_j^l} \log p(x_j^l) + \nabla_{x_j^l} k(x_j^l, x)]$$

for an arbitrary positive definite kernel function $k(.,.)$ usually chosen to be a Gaussian kernel.

Suppose we are given a time series Y_1, Y_2, \dots, Y_t for $Y \in \mathbb{R}$. We model the sequence as a state-space model parameterized by an observation density $p(y_t|x_t)$ and a transition density $p(x_t|x_{t-1})$.

We are interested in the filtering distribution $p(x_1, \dots, x_n|y_1, \dots, y_n)$ which by Bayes formula is

$$p(x_1, \dots, x_n|y_1, \dots, y_n) = \frac{p(y_1, \dots, y_n|x_1, \dots, x_n)p(x_1, \dots, x_n)}{Z}$$

Because computing the normalizing constant Z is intractable for many choices of $p(y_t|x_t)$ and $p(x_t|x_{t-1})$, we must resort to Monte Carlo algorithms. The classic approach that incorporates the sequential nature of the data is given by the particle filtering algorithm. Particle filtering approximates the filtering density using sequential importance sampling. We instead focus on the following recursion.

$$\begin{aligned} p(x_t|y_{1:t}) &= \int p(x_{0:t}|y_{1:t}) d_{x_{0:t-1}} \\ &= \frac{p(y_t|x_t)}{\int p(y_t|x_t)p(x_t|y_{1:t-1}) dx_t} p(x_t|y_{1:t-1}) \\ &\propto p(y_t|x_t)p(x_t|y_{1:t-1}) \\ &\propto p(y_t|x_t)p(x_t|y_{1:t-1}) \\ &\propto p(y_t|x_t) \int_{x_{t-1}} p(x_t, x_{t-1}|y_{1:t-1}) d_{x_{t-1}} \\ &\propto p(y_t|x_t) \int_{x_{t-1}} p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1}) d_{x_{t-1}} \end{aligned}$$

which we can approximate using svgd as

$$\approx p(y_t|x_t) \frac{1}{n} \sum_{i=1}^n p(x_t|x_{t-1}^{(i)})$$

References

- [1] Alessio Spantini. *On the low-dimensional structure of Bayesian inference*. PhD thesis, Massachusetts Institute of Technology, 2017.
- [2] Qiang Liu and Dilin Wang. Stein variational gradient descent: A general purpose bayesian inference algorithm. In *Advances In Neural Information Processing Systems*, pages 2378–2386, 2016.