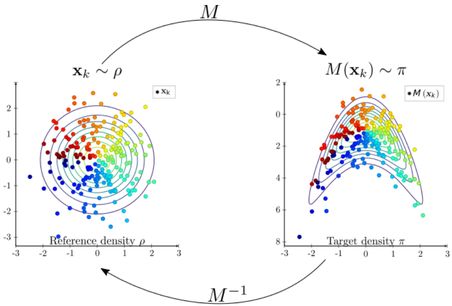
Transport Map

Dawen Wu

CNRS@CREATE

07/02/2024

Transport Map



- Target density π is unknown.
- Reference density ρ is a standard normal distribution.
- Some samples $\{x_k\} \sim \pi$ are given.
- We aim to build a transport map M, such that

$$Y = M(X), \tag{1}$$

  where $X \sim \rho$ and $Y \sim \pi$.
- We can generate samples from π using the reference ρ.

Approximated PDF $M_{\#}\rho$

Denote $M_{\#}\rho$ as the probability density function (PDF) induced by M and ρ,

$$M_{\#}\rho(x) = \rho(M^{-1}(x))|\det \nabla M^{-1}(x)|, \tag{2}$$

- ρ is the reference PDF, $\mathcal{N}$.
- M^{-1}: Inverse function of M.
- $\nabla M^{-1}(x)$: Jacobian of M^{-1} at x.
- $\det \nabla M^{-1}(x)$: Deteminant of the Jacobian.

Problem definition

The goal now becomes finding a transport map $M \in \mathcal{M}$, such that:

$$M_{\#}\rho(x) = \pi(x), \quad \forall x \in \Omega. \tag{3}$$

Given an arbitrary M, how to measure the distance between $M_{\#}\rho$ and π?

Objective function: KL divergence

Measure the gap: $M_{\#}\rho \approx \pi$

$$
\begin{aligned}
\mathcal{D}_{\mathrm{KL}}\left(\pi \| M_{\#}\rho\right) &= \mathbb{E}_{\pi}\left(\log \frac{\pi}{M_{\#}\rho}\right) \\
&= \int_X \left(\log \pi(x) - \log M_{\#}\rho(x)\right)\pi(x)dx \\
&= \int_X \left(\log \pi(x) - \log \rho(M^{-1}(x)) - \log|\det \nabla M^{-1}(x)|\right)\pi(x)dx \\
&\approx \frac{1}{n}\sum_{i=1}^{n}\left(\log \pi(x_i) - \log \rho(M^{-1}(x_i)) - \log|\det \nabla M^{-1}(x_i)|\right)
\end{aligned}
\tag{4}
$$

Objective function:

$$
\min_{M \in \mathcal{M}} -\frac{1}{n}\sum_{i}\left(\log \rho(M^{-1}(x_i)) + \log|\det \nabla M^{-1}(x_i)|\right), \tag{5}
$$

- $\mathcal{M}$ is a considered function class.
- $\{x_i\}_{i=1}^{n}$ are given samples drawn according to π.

Existing method

Parameterize the transport map $M(x)$ by $M(x; w)$.

Objective function:

$$\min_{w} -\frac{1}{n}\sum_{i}\left(\log\rho(M^{-1}(x_i; w)) + \log|\det\nabla M^{-1}(x_i; w)|\right), \tag{6}$$

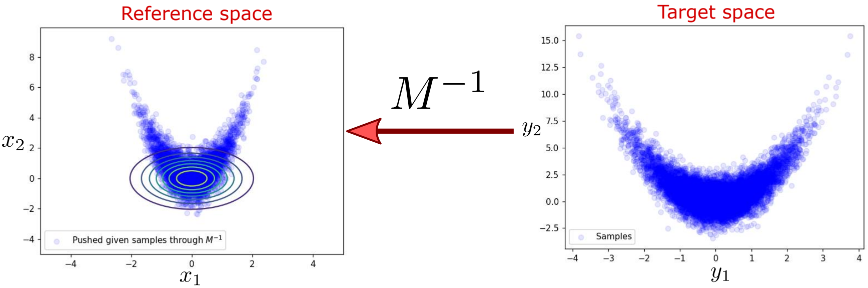
Existing method: Polynomial approximation

- El Moselhy, Tarek A., and Youssef M. Marzouk. "Bayesian inference with optimal maps." Journal of Computational Physics (2012).
- Rubio P B, Louf F, Chamoin L. Transport Map sampling with PGD model reduction for fast dynamical Bayesian data assimilation. International Journal for Numerical Methods in Engineering (2019).
- Grashorn J, Urrea-Quintero J H, Broggi M, et al. Transport map Bayesian parameter estimation for dynamical systems. PAMM (2023).

$M(x; w)$ takes the form of

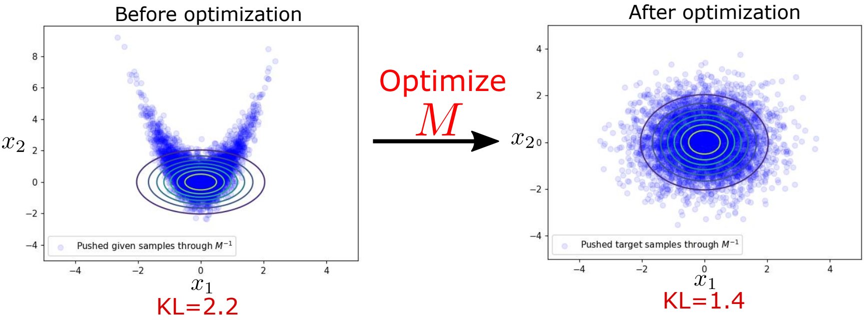
$$M(x; w) = \sum_{i\in\mathcal{J}} w_i\psi_i(x), \quad \psi_i(x) = \prod_{j=1}^{n}\varphi_{i_j}(x_j) \tag{7}$$

Experiment: Data



Push backward the given samples through M^{-1}

Experiment: Optimization



Optimize the transport map $M(x; w)$ w.r.t. the objective function.

Experiment: Approximated and Target PDF