EFM

Title	ENERGY-WEIGHTED FLOW MATCHING FOR OFFLINE REINFORCEMENT LEARNING
日期	2024.09
发表单位	THU
github	暂无
上次编辑	@2025年5月7日 14:59
备注	Sec.5 QIPO还没看
状态	To see again
重要程度	⭐⭐⭐

PRELIMINARIES

In order to ensure that the vector field $\mathbf{v}$ *generates* the probability density path p_t, the following *continuity equation* (Villani et al., 2009) is required:

连续性方程

用于确定向量场和概率路径的生成关系

$$\frac{\mathrm{d}}{\mathrm{d}t} p_t(\mathbf{x}) + \mathrm{div} \cdot [p_t(\mathbf{x}) \mathbf{v}_t(\mathbf{x})] = 0, \quad \forall \mathbf{x} \in \mathbb{R}^d. \tag{3.1}$$

The objective of flow matching is to learn a neural network $\mathbf{v}_t^{\boldsymbol{\theta}}$ to learn the ground truth vector field $\mathbf{u}_t$ by minimizing their differences, i.e., $\mathcal{L}_{\mathrm{FM}}(\boldsymbol{\theta}) = \mathbb{E}_{t,p_t(\mathbf{x})} \|\mathbf{v}_t^{\boldsymbol{\theta}}(\mathbf{x}) - \mathbf{u}_t(\mathbf{x})\|_2^2$ with respect to the network parameter $\boldsymbol{\theta}$. However, it is infeasible to calculate the ground truth vector field $\mathbf{u}_t$. To address this issue, Lipman et al. (2022) suggests to match the *conditional vector field* $\mathbf{u}_{t0}(\mathbf{x}|\mathbf{x}_0)$ instead of the vector field $\mathbf{u}_t(\mathbf{x})$, as presented by the following theorem:

Theorem 3.1 (Theorem 1, 2; Lipman et al. 2022)**.** Given the conditional vector field $\mathbf{u}_{t0}(\mathbf{x}|\mathbf{x}_0)$ that generates the conditional distribution $p_{t0}(\mathbf{x}|\mathbf{x}_0)$, then the "marginal" vector field $\mathbf{u}_t(\mathbf{x}) = \int_{\mathbf{x}_0} p_{0t}(\mathbf{x}_0|\mathbf{x}) \mathbf{u}_{t0}(\mathbf{x}|\mathbf{x}_0) \mathrm{d}\mathbf{x}_0$ generates the marginal distribution $p_t(\mathbf{x})$. In addition, up to a constant factor independent of $\boldsymbol{\theta}$, *Flow Matching* loss $\mathcal{L}_{\mathrm{FM}}(\boldsymbol{\theta})$ and *Conditional Flow Matching* loss $\mathcal{L}_{\mathrm{CFM}}(\boldsymbol{\theta})$ are equal, where

FM<=>CFM

$$\mathcal{L}_{\mathrm{FM}}(\boldsymbol{\theta}) = \mathbb{E}_{t,\mathbf{x}} \|\mathbf{v}_t^{\boldsymbol{\theta}}(\mathbf{x}) - \mathbf{u}_t(\mathbf{x})\|_2^2, \; \mathcal{L}_{\mathrm{CFM}}(\boldsymbol{\theta}) = \mathbb{E}_{t,\mathbf{x}_0,\mathbf{x}} \|\mathbf{v}_t^{\boldsymbol{\theta}}(\mathbf{x}) - \mathbf{u}_{t0}(\mathbf{x}|\mathbf{x}_0)\|_2^2, \tag{3.2}$$

where $t \sim \lambda(t)$, $\mathbf{x}_0$ follows the data distribution $p_0(\cdot)$ and $\mathbf{x} \sim p_{t0}(\cdot|\mathbf{x}_0)$ where p_{t0} is generated by conditional vector field $\mathbf{u}_{t0}$. Hence $\nabla_{\boldsymbol{\theta}} \mathcal{L}_{\mathrm{FM}}(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}} \mathcal{L}_{\mathrm{CFM}}(\boldsymbol{\theta})$.

METHODOLOGY

ENERGY-WEIGHTED FLOW MATCHING

给定Conditional Flow $u_{t0}(x|x_0)$以及生成的条件概率路径$p_{t0}(x|x_0)$

设Energy-Guided的目标分布为$q_t(x) \propto p_t(x)\exp(-\mathcal{E}_t(x))$，则$q_t(x)$由如下Flow生成

$$\hat{\mathbf{u}}_t(\mathbf{x}) = \int_{\mathbf{x}_0} p_{0t}(\mathbf{x}_0|\mathbf{x}) \mathbf{u}_t(\mathbf{x}|\mathbf{x}_0) \frac{\exp(-\beta\mathcal{E}(\mathbf{x}_0))}{\exp(-\mathcal{E}_t(\mathbf{x}))} d\mathbf{x}_0$$

为了拟合上述Flow，有如下定理

令

$$\mathcal{L}_{\mathrm{EFM}}(\theta) = \mathbb{E}_{t,\mathbf{x}} \left[\frac{\exp(-\mathcal{E}_t(\mathbf{x}))}{\mathbb{E}_{p_t(\tilde{\mathbf{x}})}[\exp(-\mathcal{E}_t(\tilde{\mathbf{x}}))]} \|\mathbf{v}_t^\theta(\mathbf{x}) - \hat{\mathbf{u}}_t(\mathbf{x})\|_2^2 \right]$$

以及

$$\mathcal{L}_{\mathrm{CEFM}}(\theta) = \mathbb{E}_{t,\mathbf{x},\mathbf{x}_0} \left[\frac{\exp(-\beta\mathcal{E}(\mathbf{x}_0))}{\mathbb{E}_{p_0(\widetilde{\mathbf{x}_0})}[\exp(-\beta\mathcal{E}(\widetilde{\mathbf{x}_0}))]} \|\mathbf{v}_t^\theta(\mathbf{x}) - \mathbf{u}_{t0}(\mathbf{x}|\mathbf{x}_0)\|_2^2 \right]$$

则有

$$\nabla_\theta \mathcal{L}_{\mathrm{EFM}}(\theta) = \nabla_\theta \mathcal{L}_{\mathrm{CEFM}}(\theta)$$

也就是说，两者有相同的极值点

将$\mathcal{L}_{\text{CEFM}}(\theta)$中的$t, x_0$固定得到

$$\mathcal{L}_{\text{CEFM}}(\theta; t, \mathbf{x}) = \mathbb{E}_{p_{0t}(\mathbf{x}_0|\mathbf{x})}\left[\frac{\exp(-\beta\mathcal{E}(\mathbf{x}_0))}{\mathbb{E}_{p_0(\widetilde{\mathbf{x}_0})}[\exp(-\beta\mathcal{E}(\widetilde{\mathbf{x}_0}))]}\|\mathbf{v}_t^\theta(\mathbf{x}) - \mathbf{u}_{t0}(\mathbf{x}|\mathbf{x}_0)\|_2^2\right]$$

直观上来说，能量值越低的样本在Loss里的权重越高，因此$\mathbf{v}_t^\theta(\mathbf{x})$会更倾向于生成能量值低的样本；而原始的CFM则平等对待每一个样本，无先验偏好

此Loss亦可从**Importance Sampling**推导得到

Remark 4.6 (Connection with the importance sampling). The conditional weighted energy guided loss $\mathcal{L}_{\text{CEFM}}$ can be also interpreted from the importance sampling techniques. Suppose we can sample directly from the data $q_0(\mathbf{x}) \propto p_0(\mathbf{x})\exp(-\beta\mathcal{E}(\mathbf{x}))$, minimizing the following loss $\mathcal{L}_q$ will get a velocity field $\mathbf{v}_t$ for generating distribution q_0

$$\mathcal{L}_q(\theta) = \mathbb{E}_{t,\mathbf{x}_0\sim q_0(\mathbf{x}),\mathbf{x}\sim q_{t0}(\mathbf{x}|\mathbf{x}_0)}[\|\mathbf{v}_t^{\boldsymbol{\theta}}(\mathbf{x}) - \mathbf{u}_{t0}(\mathbf{x}|\mathbf{x}_0)\|_2^2],$$

where $q_{t0}(\mathbf{x}|\mathbf{x}_0) = p_{t0}(\mathbf{x}|\mathbf{x}_0)$. Since , where Z is a constant, changing the data distribution from q_0 to p_0 yields that

$$\begin{aligned}\mathcal{L}_q(\theta) &= \mathbb{E}_{t,\mathbf{x}_0\sim p_0(\mathbf{x}),\mathbf{x}\sim q_{t0}(\mathbf{x}|\mathbf{x}_0)}\left[\frac{q_0(\mathbf{x})}{p_0(\mathbf{x})}\|\mathbf{v}_t^{\boldsymbol{\theta}}(\mathbf{x}) - \mathbf{u}_{t0}(\mathbf{x}|\mathbf{x}_0)\|_2^2\right] \\ &= \mathbb{E}_{t,\mathbf{x}_0\sim p_0(\mathbf{x}),\mathbf{x}\sim p_{t0}(\mathbf{x}|\mathbf{x}_0)}\left[\frac{\exp(-\beta\mathcal{E}(\mathbf{x}_0))}{\mathbb{E}_{p_0(\tilde{\mathbf{x}}_0)}[\exp(-\beta\mathcal{E}(\tilde{\mathbf{x}}_0))]}\|\mathbf{v}_t^{\boldsymbol{\theta}}(\mathbf{x}) - \mathbf{u}_{t0}(\mathbf{x}|\mathbf{x}_0)\|_2^2\right] = \mathcal{L}_{\text{CEFM}}(\theta),\end{aligned}$$

where the second equation is given by $q_0(\mathbf{x}) = p_0(\mathbf{x})\exp(-\beta\mathcal{E}(\mathbf{x}))/\mathbb{E}_{\mathbf{x}_0}[\exp(-\beta\mathcal{E}(\mathbf{x}_0))]$ according to Lemma B.1.

WEIGHTED DIFFUSION MODELS

回顾CEP，其Intermediate Energy-guided Term需要计算对能量函数的梯度，而如下推论使得我们不再需要计算这个梯度

设$\nabla_{\mathbf{x}}\log q_t(\mathbf{x}) = \nabla_{\mathbf{x}}\log p_t(\mathbf{x}) - \nabla_{\mathbf{x}}\mathcal{E}_t(\mathbf{x})$且$p_{t0}(\mathbf{x}|\mathbf{x}_0)$ Gaussian，则有

$$\nabla_{\mathbf{x}}\log q_t(\mathbf{x}) = \int_{\mathbf{x}_0} p_{0t}(\mathbf{x}_0|\mathbf{x})\nabla_{\mathbf{x}}\log p_{t0}(\mathbf{x}|\mathbf{x}_0)\frac{\exp(-\beta\mathcal{E}(\mathbf{x}_0))}{\exp(-\mathcal{E}_t(\mathbf{x}))}d\mathbf{x}_0$$

利用上述推论可以得到拟合$\nabla_{\mathbf{x}}\log q_t(\mathbf{x})$的Loss

$$\mathcal{L}_{\mathrm{ED}}(\theta) = \mathbb{E}_{t,\mathbf{x}}\left[\frac{\exp(-\mathcal{E}_t(\mathbf{x}))}{\mathbb{E}_{p_t(\tilde{\mathbf{x}})}[\exp(-\mathcal{E}_t(\tilde{\mathbf{x}}))]}\|\mathbf{s}_t^\theta(\mathbf{x}) - \nabla_{\mathbf{x}}\log q_t(\mathbf{x})\|_2^2\right],$$

$$\mathcal{L}_{\mathrm{CED}}(\theta) = \mathbb{E}_{t,\mathbf{x},\mathbf{x}_0}\left[\frac{\exp(-\beta\mathcal{E}(\mathbf{x}_0))}{\mathbb{E}_{p_0(\tilde{\mathbf{x}}_0)}[\exp(-\beta\mathcal{E}(\tilde{\mathbf{x}}_0))]}\|\mathbf{s}_t^\theta(\mathbf{x}) - \nabla_{\mathbf{x}}\log p_{t0}(\mathbf{x}|\mathbf{x}_0)\|_2^2\right],$$

满足

$$\nabla_\theta\mathcal{L}_{\mathrm{ED}}(\theta) = \nabla_\theta\mathcal{L}_{\mathrm{CED}}(\theta)$$

至此，我们不再需要估计$\nabla_{\mathbf{x}}\mathcal{E}_t(\mathbf{x}), \mathcal{E}_t(\mathbf{x})$

训练算法如下(非常EASY就可以部署)

Algorithm 1 Training Energy-Weighted Diffusion Model

Input: Score function $\mathbf{s}_t^{\boldsymbol{\theta}}(\cdot)$, schedule (μ_t, σ_t), guidance scale β, batch size B, time weight $\lambda(t)$
1: **for** batch $\{\mathbf{x}_0^i, \mathcal{E}(\mathbf{x}_0^i)\}_i$ **do**
2: **for** index $i \in [B]$ **do**
3: Calculate guidance $g_i = \mathrm{softmax}(-\beta\mathcal{E}(\mathbf{x}_0^i)) = \exp(-\beta\mathcal{E}(\mathbf{x}_0^i))/\sum_j \exp(-\beta\mathcal{E}(\mathbf{x}_0^j))$
4: Sample $t_i \sim U(0,1)$, calculate μ_{t_i}, σ_{t_i}, sample $\boldsymbol{\epsilon}_i \sim \mathcal{N}(0, \mathbf{I}_d)$ and $\mathbf{x}_{t_i} = \mu_{t_i}\mathbf{x}_0^i + \sigma_{t_i}\boldsymbol{\epsilon}_i$
5: **end for**
6: Calculate and take a gradient step using $\mathcal{L}_{\mathrm{CED}}(\boldsymbol{\theta}) = \sum_i \lambda(t_i) g_i \|\mathbf{s}_{t_i}^{\boldsymbol{\theta}}(\mathbf{x}_{t_i}) + \boldsymbol{\epsilon}_i/\sigma_{t_i}\|_2^2$.
7: **end for**

COMPARISON BETWEEN CEP AND CLASSIFIER (FREE) GUIDANCE

考虑$q_0(\mathbf{x}) \propto p_0(\mathbf{x})p^\beta(c|\mathbf{x})$，则有$\mathcal{E}(\mathbf{x}) = -\log p(c|\mathbf{x})$

有如下引理成立

Lemma 4.10. Given the same guidance scale β and the same diffusion process, let the energy function be defined by $\mathcal{E}(\mathbf{x}) = -\log p(c|\mathbf{x})$, the score function for CG and CFG are both:

$$\nabla_{\mathbf{x}}\log q_t(\mathbf{x}) = \nabla_{\mathbf{x}}\log p_t(\mathbf{x}) + \nabla_{\mathbf{x}}\log\left[\mathbb{E}_{p_{0t}(\mathbf{x}_0|\mathbf{x})}p(c|\mathbf{x}_0)\right]^\beta, \tag{4.6}$$

while the score function for energy-weighted diffusion and CEP are both

$$\nabla_{\mathbf{x}}\log q_t(\mathbf{x}) = \nabla_{\mathbf{x}}\log p_t(\mathbf{x}) + \nabla_{\mathbf{x}}\log\mathbb{E}_{p_{0t}(\mathbf{x}_0|\mathbf{x})}p^\beta(c|\mathbf{x}_0). \tag{4.7}$$

这里用到

$$
\begin{aligned}
p_t(c|\mathbf{x}_t) &= \int p_t(c|\mathbf{x}_0, \mathbf{x}_t) p_{0t}(\mathbf{x}_0|\mathbf{x}_t) \mathrm{d}\mathbf{x}_0 \\
&= \mathbb{E}_{p_{0t}(\mathbf{x}_0|\mathbf{x}_t)} p(c|\mathbf{x}_0)
\end{aligned}
$$

其关系总结如下

Table 1: Comparison between guidance methods. *Exact Guidance?* means if the model can generate $p(\mathbf{x})p^{\beta}(c|\mathbf{x})$ when $\beta \neq 1$. *w/o Auxiliary Model?* means if the method can direct learn the guidance without auxiliary model (✓) or not (×).

Guidance	Exact Guidance?	w/o Auxiliary Model?
Classifier-guidance (Dhariwal & Nichol, 2021)	×	×
Classifier-free guidance (Ho & Salimans, 2021)	×	✓
Contrastive energy prediction (Lu et al., 2023)	✓	×
Energy-weighted diffusion (ours)	✓	✓