⌛

Flow Matching

参考

Continuous Normalizing Flows

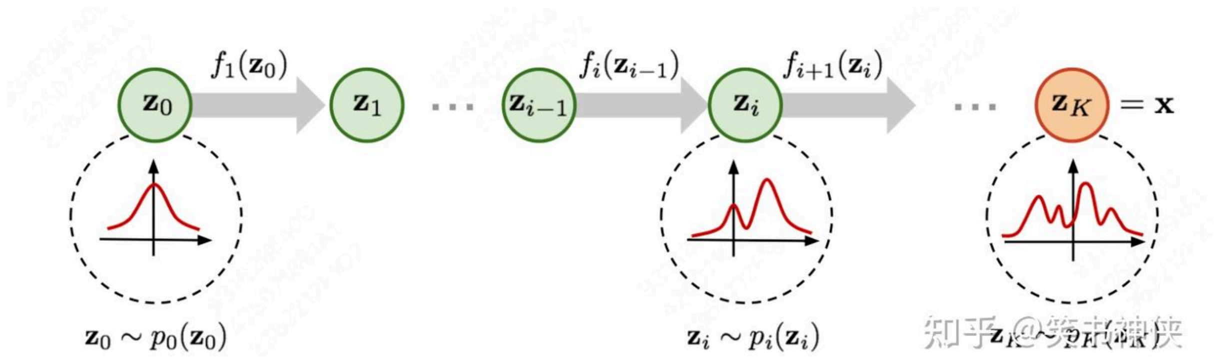
CNFs是一种用于描述连续的变量变换的数学工具

Normalizing Flows

给定随机变量z，其pdf为$\pi(z)$，设$x = f(z)$，则x的pdf为

$$p(x) = \pi(f^{-1}(x))\mathrm{det}\frac{\mathrm{d}f^{-1}}{\mathrm{d}x}$$

通过一系列可逆的概率密度变换，将简单的分布映射到复杂的分布



设p_0为原始分布，经过一系列可逆变换$\{f_i\}$:

$$x = z_K = f_K \circ f_{K-1} \circ \cdots \circ f_1(z_0)$$

也就是说

$$z_i = f_i(z_{i-1})$$

于是，我们有

$$\begin{aligned} p_i(z_i) &= p_{i-1}(f_i^{-1}(z_i)) \left| \det \frac{\mathrm{d} f_i^{-1}}{\mathrm{d} z_i} \right| \\ &= p_{i-1}(z_{i-1}) \left| \det \left(\frac{\mathrm{d} f_i}{\mathrm{d} z_{i-1}} \right)^{-1} \right| \\ &= p_{i-1}(z_{i-1}) \left| \det \frac{\mathrm{d} f_i}{\mathrm{d} z_{i-1}} \right|^{-1} \end{aligned}$$

取对数得

$$\log p_i(z_i) = \log p_{i-1}(z_{i-1}) - \log \left| \det \frac{\mathrm{d} f_i}{\mathrm{d} z_{i-1}} \right|$$

展开直到原始分布

$$\log p(x) = \log p_K(z_K) = \log \pi_0(z_0) - \sum_{i=1}^{K} \log \left| \det \frac{\mathrm{d} f_i}{\mathrm{d} z_{i-1}} \right|$$

模型训练时，目标为最大化对数似然$\log p(x)$

Continuous Normalizing Flows

CNFs为NFs的连续推广，变量的连续变换通过ODE来描述

$$\frac{\mathrm{d}z_t}{\mathrm{d}t} = v(z_t, t)$$

其中$t \in [0,1]$，z_t为Flow Map，$v(z_t, t)$为向量场，用于描述数据点的变化趋势，通常由神经网络预测

给定一个初始分布，$\mathbf{v(z_t, t)}$可以给出这个分布随时间的演变情况，最终达到目标分布

该ODE可通过数值方法估计

$$z_{t+\Delta t} = z_t + \Delta t \cdot v(z_t, t)$$

Continuity Equation

用于描述守恒量的传输行为

> 设有流体，密度为ρ，速度场为$\boldsymbol{v}$，则任取体积$\mathcal{V}$，单位时间流出$\mathcal{V}$的流体质量等于其内流体质量的减少量，结合散度定理有
>
> $$\int_{\mathcal{S}} \rho \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} + \frac{\partial}{\partial t} \int_{\mathcal{V}} \rho \mathrm{d}V = 0$$
>
> $$\int_{\mathcal{V}} \nabla \cdot (\rho \boldsymbol{v}) \mathrm{d}V + \int_{\mathcal{V}} \frac{\partial}{\partial t} \rho \mathrm{d}V = 0$$
>
> $$\frac{\partial}{\partial t} \rho + \nabla \cdot (\rho \boldsymbol{v}) = 0$$

由于概率密度总和总为1，因此**概率守恒**，将ρ置换为pdf p_t，流速场置换为概率密度流v_t，得

$$\frac{\partial p_t(x)}{\partial t} + \nabla \cdot [p_t(x) v_t(x)] = 0 \tag{1}$$

注意：Fokker-Plank Equation为Continuity Equation的推广！

性质：Vector Field $\mathbf{v_t(x)}$与$\mathbf{p_t(x)}$满足(1)当且仅当$\mathbf{v_t(x)}$生成概率密度路径$\mathbf{p_t(x)}$

Conditional and Marginal Probability Paths and Vector Fields

Conditional and Marginal Probability Paths and Vector Fields

现在的问题来到给定初始分布p_0，如何求向量场u_t，使得p_0生成$p_t \to q, t \to 1$

考虑条件概率流$u(x|x_1)$，其中x_1为数据空间中具体的数据样本，其生成条件概率密度路径$p_t(x|x_1)$[两者满足(1)]，定义边缘概率密度路径

$$p_t(x) = \int p_t(x|x_1)q(x_1)\mathrm{d}x_1 \tag{2}$$

再通过对条件向量场的积分得到边缘向量场

$$u_t(x) = \int u_t(x|x_1)\frac{p_t(x|x_1)q(x_1)}{p_t(x)}\mathrm{d}x_1 \tag{3}$$

则$u_t(x)$生成$p_t(x)$

现在，只需要设计$u(x|x_1)$或$p_t(x|x_1)$，使得$p_1(x) \approx q(x)$即可解决上述问题

由于

$$p_1(x) = \int p_1(x|x_1)q(x_1)\mathrm{d}x_1$$

令$p_1(x|x_1) = \delta(x_1 - x)$，则有

$$p_1(x) = \int \delta(x_1 - x)q(x_1)\mathrm{d}x_1 = q(x)$$

此时，u_t生成的概率密度p_t路径将收敛于q

Summary:

1. 定义$p_t(x|x_1)$，使得$p_1(x|x_1) = \delta(x_1 - x), p_0(x|x_1) = p_0(x), \forall x_1 \in \mathcal{X}$
2. 利用(1)求出对应条件向量场$u_t(x|x_1)$
3. 利用(2)(3)求出边缘向量场$u_t(x)$
4. p_0在u_t下演变至q

Calculate Conditional Probability Paths and Conditional Vector Fields

考虑高斯条件概率路径

$$p(x|x_1) = \mathcal{N}(\mu_t(x_1), \sigma_t(x_1))$$

满足

$$\mu_0 = 0, \sigma_0 = 1$$
$$\mu_1 = 0, \sigma_1 \approx 0$$

则其满足上述边值条件$p_1(x|x_1) = \delta(x_1 - x), p_0(x|x_1) = p_0(x), \forall x_1 \in \mathcal{X}$

其对应的Flow Map为

$$\psi_t(x|x_1) = \sigma_t(x_1)x + \mu_t(x_1)$$

代入(1)$\left(z_t \leftarrow \psi_t(x|x_1)\right)$得到闭式解

$$u_t(x|x_1) = \frac{\sigma_t'(x_1)}{\sigma_t(x_1)}[x - \mu_t(x_1)] + \mu_t'(x_1)$$

Conditional Flow Matching

尽管已经求得$u_t(x|x_1)$，但求$u_t(x)$依旧是Intractable的(积分包含未知分布q)，此时，就需要依赖作者提出的**Conditional Flow Matching技术**

Flow Matching

拟合q已经被转化为可以生成q的概率密度向量场u_t，令v_t为由NN参数化的向量场，目标为估计u_t

$$\mathcal{L}_{\text{FM}}(\theta) = \mathbb{E}_{t, p_t(x)} \|u_t(x) - v_t(x)\|^2$$

问题是：$u_t(x)$是Intractable的

Conditional Flow Matching

$u_t(x)$ Intractable，但$u_t(x|x_1)$确是已知

考虑**Conditional Flow Matching Loss**

$$\mathcal{L}_{\text{CFM}}(\theta) = \mathbb{E}_{t, q(x_1), p_t(x|x_1)} \|u_t(x|x_1) - v_t(x)\|^2$$

$$
\begin{aligned}
&\nabla_\theta L_{CFM}(\theta) = \nabla_\theta E_{t,p_t(x|z),q(z)} \left\| v_\theta(t,x) - u_t(x|z) \right\|^2 \\
&= \nabla_\theta E_{t,p_t(x|z),q(z)} \left[v_\theta(t,x)^2 - 2 \cdot v_\theta(t,x) \cdot u_t(x|z) + u_t(x|z)^2 \right] \\
&= \nabla_\theta E_{t,p_t(x|z),q(z)} \left[v_\theta(t,x)^2 \right] - 2\nabla_\theta E_{t,p_t(x|z),q(z)} \left[v_\theta(t,x) \cdot u_t(x|z) \right]
\end{aligned}
$$

类似的，我们考察(4)式对于 θ 的导数，有

$$
\begin{aligned}
&\nabla_\theta L_{FM}(\theta) = \nabla_\theta E_{t,p_t(x)} \left\| v_\theta(t,x) - u_t(x) \right\|^2 \\
&= \nabla_\theta E_{t,p_t(x)} \left[v_\theta(t,x)^2 \right] - 2\nabla_\theta E_{t,p_t(x)} \left[v_\theta(t,x) \cdot u_t(x) \right]
\end{aligned}
$$

所以现在的关键在于，以上两个式子的右边第二项有什么联系？我们继续化简

$$
\begin{aligned}
&\nabla_\theta E_{t,p_t(x)} \left[v_\theta(t,x) \cdot u_t(x) \right] \\
&= \nabla_\theta \int \int v_\theta(t,x)\, u_t(x)\, p_t(x)\, dxdt \\
&= \nabla_\theta \int \int v_\theta(t,x) \left[\int \frac{u_t(x|z)p_t(x|z)}{p_t(x)} q(z)\, dz \right] p_t(x)\, dxdt \\
&= \nabla_\theta \int \int \int v_\theta(t,x)\, u_t(x|z)\, p_t(x|z)\, q(z)\, dxdtdz \\
&= \nabla_\theta E_{t,p_t(x|z),q(z)} \left[v_\theta(t,x) \cdot u_t(x|z) \right]
\end{aligned}
$$

以上推导使用了 $u_t(x)$ 的定义(7)以及期望的积分定义，所以这就巧了，我们发现

$$
\nabla_\theta L_{FM}(\theta) = \nabla_\theta L_{CFM}(\theta)
$$

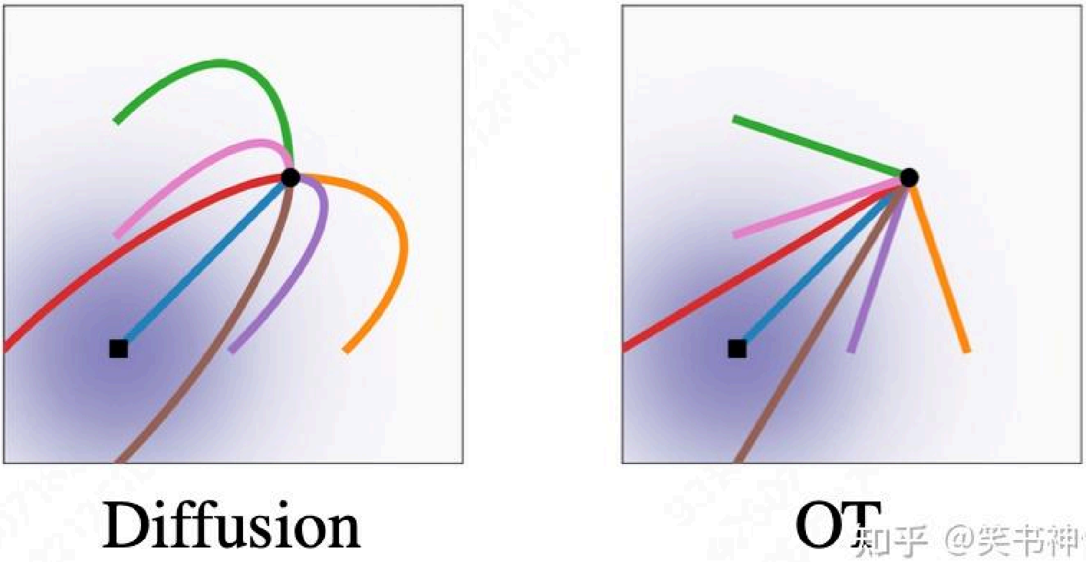
也就是说，$\mathcal{L}_{\mathrm{FM}}$与$\mathcal{L}_{\mathrm{CFM}}$有相同的极小值！！！

因此，我们将Intractable的FM转化为CFM，只要$v_t(x)$最小化$\mathcal{L}_{\mathrm{CFM}}$，则v_t(x)成功拟合$u_t(x)$

于是，在得到$v_t(x)$后，即可通过数值求解对应的概率密度变换ODE来将p_0变换为目标分布q

Diffusion conditional Vector Fields and Optimal Transport conditional Vector Fields

通过设计不同的Flow Map(也就是μ_t, σ_t)，可以得到不同的概率密度路径



Diffusion conditional Vector Fields

- **Variance Exploding (VE)**

  在扩散过程中逐步增大数据噪声(能量变大)，也就是增大方差，数据会逐渐变得嘈杂
  有利于探索更广阔的隐空间，使生成的样本更加多样化

  $$p_t(x|x_1) = \mathcal{N}(x_1, \sigma_{1-t}^2 \mathbf{I})$$

  其中σ_t为满足$\sigma_0 = 0, \sigma_1 \gg 1$的单增函数

  对应

  $$u_t(x|x_1) = -\frac{\sigma'_{1-t}}{\sigma_{1-t}}(x - x_1)$$

- **Variance Preserving (VP)**

  保持数据整体方差不变，用于需要保持数据分布稳定性的场景，如保持图像清晰度和结构特征

  $$p_t(x|x_1) = \mathcal{N}(\alpha_{1-t}x_1, (1 - \alpha_{1-t}^2)\mathbf{I})$$

  对应

  $$u_t(x \mid x_1) = \frac{\alpha'_{1-t}}{1 - \alpha_{1-t}^2}(\alpha_{1-t}x - x_1) = -\frac{T'(1-t)}{2}\left[\frac{e^{-T(1-t)}x - e^{-\frac{1}{2}T(1-t)}x_1}{1 - e^{-T(1-t)}}\right]$$

训练更稳定

Optimal Transport conditional Vector Fields

简单定义条件概率路径为均值与标准差为时间的线性函数

$$\mu_t(x_1) = tx_1, \sigma_t(x) = 1 - (1 - \sigma_{\min})t$$

对应Flow Map

$$\psi_t(x) = (1 - (1 - \sigma_{\min})t)x + tx_1$$

条件向量场闭式解为

$$u_t(x|x_1) = \frac{x_1 - (1 - \sigma_{\min})x}{1 - (1 - \sigma_{\min})t}$$

训练速度更快