Параметрические методы решения задачи о мосте Шредингера

Забарянская Ирина

Московский физико-технический институт
кафедра интеллектуальных систем

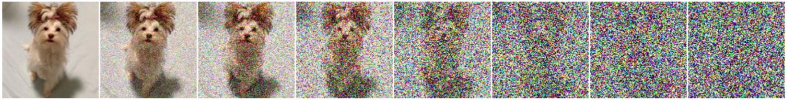
Научно-исследовательская работа
17 мая 2025 года

Диффузионные модели (DMs)

Основная идея: обратить процесс зашумления данных.

Forward SDE (data → noise)

$$\mathbf{x}(0) \longrightarrow \mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}, t)\mathrm{d}t + g(t)\mathrm{d}\mathbf{w} \longrightarrow \mathbf{x}(T)$$



score function

$$\mathbf{x}(0) \longleftarrow \mathrm{d}\mathbf{x} = \left[\mathbf{f}(\mathbf{x}, t) - g^2(t)\boxed{\nabla_{\mathbf{x}} \log p_t(\mathbf{x})}\right] \mathrm{d}t + g(t)\mathrm{d}\bar{\mathbf{w}} \longleftarrow \mathbf{x}(T)$$

Reverse SDE (noise → data)

Аппроксимируем $\nabla_x \log p_t(x) \implies$ генеративная модель!

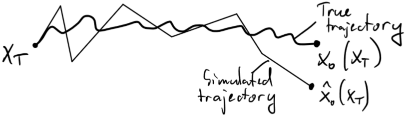
Ограничение DM №1: трудоемкий вывод

Процесс расшумления:

$$dx_t = x_t - [f(x_t, t) - g^2(t)\nabla_x \log p(x_t, t)]dt + g(t)d\overline{W}_t.$$

Дискретизация процесса расшумления (метод Эйлера-Маруямы):

$$x_{t-\Delta t} = x_t - [f(x_t, t) - g^2(t)\nabla_x \log p(x_t, t)]\Delta t + g(t)\sqrt{\Delta t \xi_t}, \quad \xi_t \sim \mathcal{N}(0, I).$$



Малое Δt :

- Более точная симуляция процесса расшумления;
- Много раз пересчитывать $\nabla_x \log p(x_t, t) \implies$ долго.

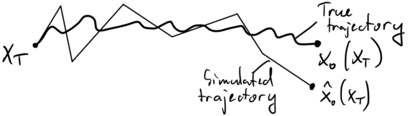
Ограничение DM №1: трудоемкий вывод

Процесс расшумления:

$$dx_t = x_t - [f(x_t, t) - g^2(t)\nabla_x \log p(x_t, t)]dt + g(t)d\overline{W}_t.$$

Дискретизация процесса расшумления (метод Эйлера-Маруямы):

$$x_{t-\Delta t} = x_t - [f(x_t, t) - g^2(t)\nabla_x \log p(x_t, t)]\Delta t + g(t)\sqrt{\Delta t \xi_t}, \quad \xi_t \sim \mathcal{N}(0, I).$$

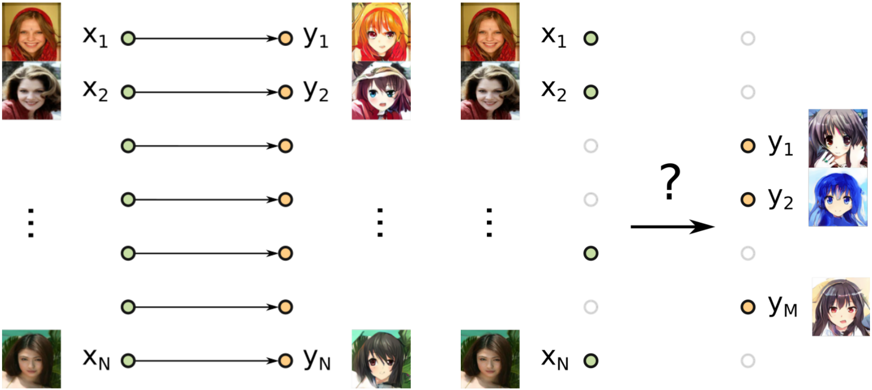


Малое Δt :

- Более точная симуляция процесса расшумления;
- Много раз пересчитывать $\nabla_x \log p(x_t, t) \implies$ долго.

Ограничение DM №2: задача непарного перевода домена

С учителем: **Без учителя:**



Применимы обусловленные DMs.

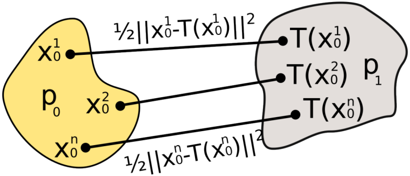
Не применимы DMs!

Оптимальный транспорт (OT, постановка Монжа)

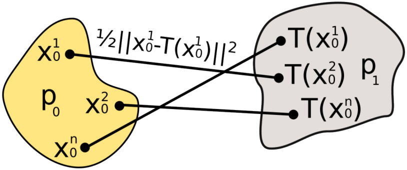
Стоимость транспорта между $p_0, p_1 \in \mathcal{P}_{2,ac}(\mathbb{R}^D)$:

$$Cost(p_0, p_1) = \inf_{T\sharp p_0 = p_1} \int_{\mathcal{X}} \frac{\|x_0 - T(x_0)\|^2}{2} p_0(x_0) dx_0.$$

Оптимальный транспорт T^* минимизирует стоимость.



Оптимальное транспортное отображение.



Не оптимальное транспортное отображение.

Энтропийный оптимальный транспорт (EOT)

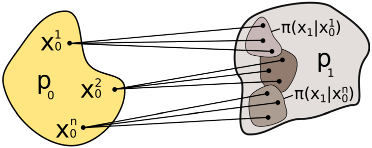
Стохастический OT (постановка Канторовича):

$$\inf_{\pi \in \Pi(p_0, p_1)} \int_{\mathbb{R}^D} C(x_0, \pi(.|x_0)) p_0(x_0) dx_0.$$

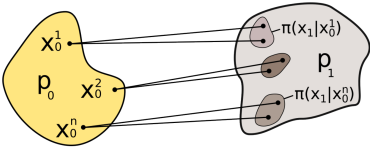
Стохастический OT с энтропийной регуляризацией:

$$C(x_0, \pi(.|x_0)) \triangleq \int_{\mathbb{R}^D} \frac{\|x_0 - x_1\|^2}{2} \pi(x_1|x_0) dx_1 - \epsilon H(\pi(.|x_0)).$$

Параметр регуляризации ϵ контролирует разнообразие.



EOT: большое ϵ.



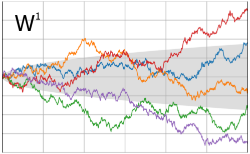
EOT: малое ϵ.

Задача о мосте Шредингера (SB)

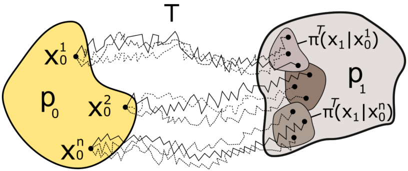
Обозначения:

$\mathcal{F}(p_0, p_1)$ - стохастические процессы с маргинальными p_0 и p_1;

W^{ϵ} - винеровский процесс с вариацией ϵ.



Задача о мосте Шредингера: $\boxed{\inf_{T \in \mathcal{F}(p_0, p_1)} KL(T \| W^{\epsilon})}$



Связь SB с EOT

Обозначение: π^T совместное распределение процесса T на концах.

Можно переписать:

$$\inf_{T \in \mathcal{F}(p_0, p_1)} KL(T \| W^\epsilon) = \inf_{T \in \mathcal{F}(p_0, p_1)} KL(\pi^T \| \pi^{W^\epsilon}) = \inf_{\pi \in \Pi(p_0, p_1)} KL(\pi \| \pi^{W^\epsilon}).$$

Связь SB с EOT:

Совместное распределение оптимального транспорта T^* – решение EOT с параметром регуляризации ϵ.

$$\boxed{\pi^{T^*} = \pi^*}$$

Yongxin Chen, et al. On the relation between optimal transport and Schrödinger bridges // Optimization Theory and Applications. 2016.

Связь SB с EOT

Обозначение: π^{T} совместное распределение процесса T на концах.

Можно переписать:

$$\inf_{T \in \mathcal{F}(p_0, p_1)} KL(T \| W^{\epsilon}) = \inf_{T \in \mathcal{F}(p_0, p_1)} KL(\pi^{T} \| \pi^{W^{\epsilon}}) = \inf_{\pi \in \Pi(p_0, p_1)} KL(\pi \| \pi^{W^{\epsilon}}).$$

Связь SB с EOT:

Совместное распределение оптимального транспорта T^{*} – решение EOT с параметром регуляризации ϵ.

$$\boxed{\pi^{T^{*}} = \pi^{*}}$$

Yongxin Chen, et al. On the relation between optimal transport and Schrödinger bridges // Optimization Theory and Applications. 2016.

Характеристика решений EOT

Решения EOT π^* характеризуются начальным распредением p_0 и скалярной функцией $v^* : \mathbb{R}^D \to \mathbb{R}_+$:

$$\pi^*(x_0, x_1) = \pi^*(x_0)\pi^*(x_1|x_0) = p_0(x_0)\frac{\exp(\langle x_0, x_1\rangle\epsilon)v^*(x_1)}{c_{v^*}(x_0)},$$

$$c_{v^*}(x_0) \triangleq \int_{\mathbb{R}^D} \exp(\langle x_0, x_1\rangle\epsilon)v^*(x_1)dx_1.$$

$c_{v^*}(x_0)$ - нормализующая константа;

v^* - потенциал Шредингера.

Christian Leonard. A survey of the Schrödinger problem and some of its connections with optimal transport // Discrete and Continuous Dynamical Systems-A. 2014.

Light SB

Вместо решения EOT $\rightarrow$ минимизируем KL с искомым планом π^*:

$$\underset{\pi \in \Pi(p_0, p_1)}{\arg\min} KL(\pi \| \pi^{W^\epsilon}) \rightarrow \underset{\pi}{\arg\min} KL(\pi^* \| \pi)$$

Проблема: не знаем π^*.

Решение: выбрать хорошую параметризацию для π.

$$\pi_\theta(x_0, x_1) = p_0(x_0)\pi_\theta(x_1|x_0) = p_0(x_0)\frac{\exp(\langle x_0, x_1\rangle \epsilon)v_\theta(x_1)}{c_\theta(x_0)},$$

$c_\theta(x_0)$ - нормализующая константа, v_θ - параметризация v^*.

Функция потерь Light SB:

$$\min_\theta KL(\pi^* \| \pi_\theta) - C = \min_\theta \int_{\mathbb{R}^D} p_0(x_0) \log c_\theta(x_0) dx_0 - \int_{\mathbb{R}^D} p_1(x_1) \log v_\theta(x_1) dx_1.$$

Правая часть не зависит от π^*!

Light SB

Вместо решения EOT $\rightarrow$ минимизируем KL с искомым планом π^*:

$$\arg\min_{\pi \in \Pi(\boldsymbol{p_0}, \boldsymbol{p_1})} KL(\pi \| \pi^{W^\epsilon}) \rightarrow \arg\min_{\pi} KL(\pi^* \| \pi)$$

Проблема: не знаем π^*.

Решение: выбрать хорошую параметризацию для π.

$$\pi_\theta(x_0, x_1) = p_0(x_0)\pi_\theta(x_1 | x_0) = p_0(x_0)\frac{\exp(\langle x_0, x_1 \rangle \epsilon) v_\theta(x_1)}{c_\theta(x_0)},$$

$c_\theta(x_0)$ - нормализующая константа, v_θ - параметризация v^*.

Функция потерь Light SB:

$$\min_\theta KL(\pi^* \| \pi_\theta) - C = \min_\theta \int_{\mathbb{R}^D} p_0(x_0) \log c_\theta(x_0) dx_0 - \int_{\mathbb{R}^D} p_1(x_1) \log v_\theta(x_1) dx_1.$$

Правая часть не зависит от π^*!

Light SB

Вместо решения EOT $\to$ минимизируем KL с искомым планом π^*:

$$\arg\min_{\pi \in \Pi(p_0, p_1)} KL(\pi \| \pi^{W^\epsilon}) \to \arg\min_{\pi} KL(\pi^* \| \pi)$$

Проблема: не знаем π^*.

Решение: выбрать хорошую параметризацию для π.

$$\pi_\theta(x_0, x_1) = p_0(x_0)\pi_\theta(x_1|x_0) = p_0(x_0)\frac{\exp(\langle x_0, x_1\rangle \epsilon) v_\theta(x_1)}{c_\theta(x_0)},$$

$c_\theta(x_0)$ - нормализующая константа, v_θ - параметризация v^*.

Функция потерь Light SB:

$$\min_\theta KL(\pi^* \| \pi_\theta) - C = \min_\theta \int_{\mathbb{R}^D} p_0(x_0) \log c_\theta(x_0) dx_0 - \int_{\mathbb{R}^D} p_1(x_1) \log v_\theta(x_1) dx_1.$$

Правая часть не зависит от π^*!

Light SB: гауссова параметризация

Проблема: сложно вычислить $c_\theta(x_0)$ для произвольной v_θ.

Решение: аппроксимируем v_θ смесью гауссиан:

$$v_\theta \triangleq \sum_{k=1}^{K} \alpha_k \mathcal{N}(x_1 | r_k, S_k), \quad \theta \triangleq \{\alpha_k, r_k, S_k\}_{k=1}^{K}.$$

Теперь $c_\theta(x_0)$ аналитически вычислима:

$$\pi_\theta(x_1|x_0) = \frac{1}{c_\theta(x_0)} \sum_{k=1}^{K} \widetilde{\alpha}_k(x_0) \mathcal{N}(x_1 | r_k(x_0), \epsilon S_k),$$

$$r_k(x_0) \triangleq r_k + S_k x_0, \quad \widetilde{\alpha}_k(x_0) \triangleq \exp \frac{x_0^T S_k x_0 + 2 r_k^T x_0}{2\epsilon}, \quad c_\theta(x_0) \triangleq \sum_{k=1}^{K} \widetilde{\alpha}_k(x_0).$$

Alexander Korotin, et al. Light Schrödinger Bridge // ICLR. 2024.

Light SB: гауссова параметризация

Проблема: сложно вычислить $c_\theta(x_0)$ для произвольной v_θ.

Решение: аппроксимируем v_θ смесью гауссиан:

$$v_\theta \triangleq \sum_{k=1}^{K} \alpha_k \mathcal{N}(x_1 | r_k, S_k), \quad \theta \triangleq \{\alpha_k, r_k, S_k\}_{k=1}^{K}.$$

Теперь $c_\theta(x_0)$ аналитически вычислима:

$$\pi_\theta(x_1|x_0) = \frac{1}{c_\theta(x_0)} \sum_{k=1}^{K} \widetilde{\alpha}_k(x_0) \mathcal{N}(x_1 | r_k(x_0), \epsilon S_k),$$

$$r_k(x_0) \triangleq r_k + S_k x_0, \quad \widetilde{\alpha}_k(x_0) \triangleq \exp \frac{x_0^T S_k x_0 + 2 r_k^T x_0}{2\epsilon}, \quad c_\theta(x_0) \triangleq \sum_{k=1}^{K} \widetilde{\alpha}_k(x_0).$$

Alexander Korotin, et al. Light Schrödinger Bridge // ICLR. 2024.

Вариационная оценка: Bohning vs Jaakkola

Функция потерь Light SB:

$$\min_{\theta} \int_{\mathbb{R}^D} p_0(x_0) \log c_\theta(x_0) dx_0 - \int_{\mathbb{R}^D} p_1(x_1) \log v_\theta(x_1) dx_1.$$

Как оптимизировать? Исходная статья: градиентный спуск.

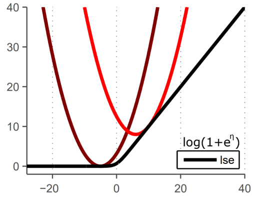
Мой результат: вывод вариационной оценки (с Jaakkola).

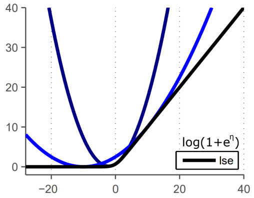
Оценка Bohning

- Менее точная;
- Быстрее;
- Фиксированная кривизна.

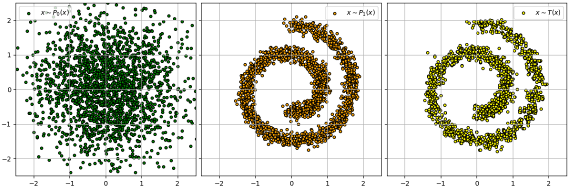
Оценка Jaakkola

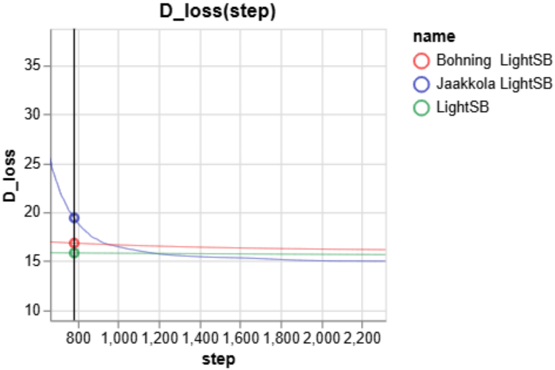
- Более точная;
- Медленнее;
- Переменная кривизна.





Эксперимент: Swiss Rolls



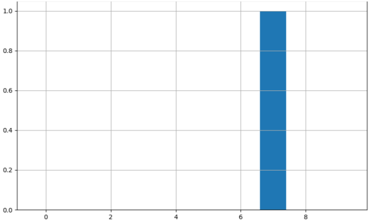


Эксперимент: ALAE

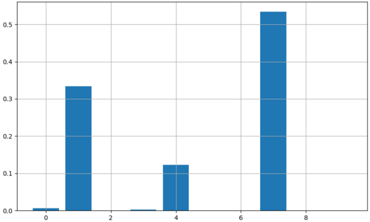
Мужчина $\to$ женщина.

Взрослый $\to$ ребенок.

Jaakkola Light SB: решение проблемы затухания весов?



Bohning Light SB: $\mathbb{E}_{p_0(x)}[\alpha_k(x)]$.



Jaakkola Light SB: $\mathbb{E}_{p_0(x)}[\alpha_k(x)]$.