Министерство образования и науки Российской Федерации
Московский физико-технический институт (государственный университет)

Физтех-школа радиотехники и компьютерных технологий
Кафедра системного программирования ИСП РАН
Лаборатория (laboratory name)

Выпускная квалификационная работа бакалавра

Исследование и разработка методов машинного обучения

Автор:
Студент 082 группы
Иванов Иван Иванович

Научный руководитель:
научная степень
Денисов Денис Денисович

Научный консультант:
научная степень
Сергеев Сергей Сергеевич

Москва 2024

Аннотация

Исследование и разработка методов машинного обучения
Иванов Иван Иванович

Краткое описание задачи и основных результатов, мотивирующее прочитать весь текст.

Abstract

Research and development of machine learning methods

Содержание

1 Введение

Генеративные модели широко используются во многих областях искусственного интеллекта и машинного обучения. Недавние достижения в параметризации этих моделей с использованием глубоких нейронных сетей, а также прогресс в методах стохастической оптимизации позволили масштабировать моделирование на сложных многомерных данных, включая изображения, текст, речь и другие.

В генеративном моделировании ставят задачу вероятностно, то есть любой наблюдаемый набор данных $\mathcal{D}$ рассматривают, как конечный набор исследуемых объектов $d \in \mathcal{D}$, принадлежащие базовому распределению $d \sim p_{data}$. Цель любой генеративной модели состоит в том, чтобы аппроксимировать это распределение данных при наличии доступа к набору данных $\mathcal{D}$. Обучая такую модель мы в дальнейшем можем ее использовать для того, чтобы генерировать новые объекты $\hat{d} \notin \mathcal{D}$.

Современными моделями в генеративном моделировании являются диффузионные модели [1], применяемые в различных сферах: от преобразования текста в изображения, 3D, видео и синтез белка. Они полагаются на итеративную процедуру, при которой распределение данных p_{data} сначала искажается с помощью процесса прямого зашумления, сходящегося к нормальному распределению. Затем изучается обратный во времени процесс, который восстанавливает данные из гауссовского распределения обратно в распределение данных, с использованием score matching подходов.

Недавняя практика и новые усовершенствования диффузионных моделей показывают, что зачастую нормальное распределение является не самым оптимальным выбором скрытого пространства. Например, для задачи восстановления зашумленных изображение, имеея начальное изображение, имеет смысл восстанавливать начаная с данного изображения. Однако, это требует изменения целевого распределения прямого процесса, что часто неразрешимо и приводит к численным трудностям.

Эту проблему решает два обобщения диффузионных моделей: Flow matching [2] и Bridge matching [**?**], которые связывают два распределения дискрименативно и стохастически, соответсвенно. Таким образом данный подход расширяет спектр задач и упрощает применение к уже существующим задачам. Например, преобразовывать синтетические изображения с камеры видеорегистратора в реальные (Рисунок 1).

Однако для вышеуказанных подходов требуются специализированные наборы данных, объекты которых являются пары из двух заданных конечных распределений, т.е. во время обучения оба элемента пары должны соответсвовать одному объекту, например нарисованный самолет и фотография самолета (в данном случая самолет должны быть одним и тем же).

В статье [3] доказывается, что вышеуказанное требование может быть опущено, если ввести дополнительные критерии оптимальности для прямого и обратного процесса. Такими требованиями обладают мосты Шредингера [4], [5]. Задача Мостов Шредингера находит стохастический процесс (Рис 2), ограниченный двумя заданными конечными распределениями, который наиболее похож с винеревским процессом с точки зрения дивергенции Кульбака — Лейблера. Кроме того, это позволяет вычислять вероятность этой стохастической процесса, что позволяет нам сравнивать два набора данных, то есть распределения, что может быть полезно для проверки гипотез и семантического сходства.

Вдохновленными результатами дифузионных генеративных моделей, и родства их с Мостами Шрёдингера было предложено несколько подходов [6], [7], [8], [9].

Несмотря на качество полученных результатов такие подходы являются вычислительно дорогостоящими из-за этого возникает потребность в использовнии других методов построения мостов Шредингера. Одним из примеров решения данной проблемы

Рис. 1: Два домена: синтетические и реальные виды города

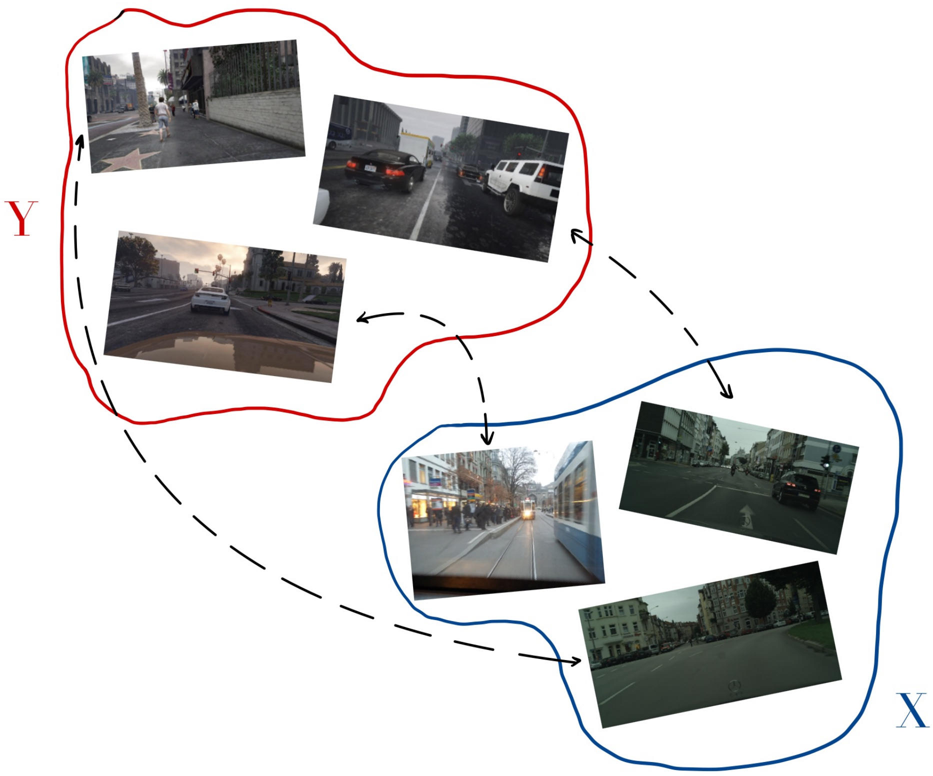

Рис. 2: Иллюстрация мостов Шрёдингера

является подход описанный в следующей статье [10]. Однако, несмотря на значительную скорость обучения, данный подход плохо справляется с данными большой размерности.

Решением данной проблемы служит использовние предшествуещего лидера в глубоких генеративных моделях (до диффузионных моделей) – состязательные генеративные сети [11]. Состязательные генеративные сети внесли большой вклад в развитие глубокого генеративного моделирования. Разработанный подход позволяет обучать неявным образом генеративные распределения, использую только сэмплы из реального распределения данных p_{data}.

Вдохновившись результатами состязательных генеративных сетей, данный подход был применен на задаче адаптации домена в работе [12]. Важной работой, посвященной исследованию состязательных генеративных сетей, является [13]. Еще Гудфеллоу заметил в работе [11], что целевая функция минимизирует девергенцию Йенсена-Шеннона. Благодаря этому факту авторы [13] обобщают целевую функцию состязательных генеративных сетей до f-семейства дивергенций, в которую входят прямая и обратная дивергенции Кульбака — Лейблера, общая вариация и в том числе девергенция Йенсена-Шеннона. То есть авторы связывают состязательное обучение с вариационным представлением f-дивергенций.

Унаследовав вариационный подход к разложению дивергенций, предложенный в [13], в работе предлагается новый подход к обучению мостов Шредингера, который наследует преимущества Bridge Matching, а также состязательных генеративных сетей. Помимо этого, можно сделать вывод, что мосты Шредингера являются актуальной исследовательским направлением в сфере машинного обучения и требуют активного участия научных деятелей.

Далее вводятся определения понятий, которые понадобятся в постановке задачи.

...

2 Постановка задачи

Необходимо формально изложить суть задачи в данной секции, предоставив такие ясные и точные описания, которые позволят в последующем оценить, насколько разработанное решение соответствует поставленной задаче. Текст главы должен следовать структуре технического задания, включая как описание самой задачи, так и набор требований к ее решению.

2.1 Статическая постановка задачи мостов Шредингера

Пусть также как и динамической постановке задачи заданы два множества X и Y, но вместо Винеровского процесса $\mathbb{W}^{\gamma}$ задано его совместное распределение $p^{\mathbb{W}^{\gamma}}(x,y)$, для которого выполняется следующее равенство:

$$p^{\mathbb{W}^{\gamma}}(x,y) = p_1(y)p(x|y,\gamma) = p_1(y)\mathcal{N}(x|y,\gamma\mathbb{I}),$$

где $p_1(y)$ - любое распределение.

Эквивалентной задачей будет являтся нахождение наиболее близкого совместного распределения $q(x,y)$, которое ограниченно заданными распределениями $\pi_0(x)$ и $\pi_1(y)$, то есть:

$$\pi_0(x) = \int_{\mathcal{Y}} q(x,y)dy,$$
$$\pi_1(y) = \int_{\mathcal{X}} q(x,y)dx,$$

к распределению $p^{\mathbb{W}^{\gamma}}(x,y)$ с точки зрения дивергенции Кульбака - Лейблера. То есть:

$$\begin{cases} \hat{q}(x,y) = \arg\min_{q(x,y)} D_{KL}(q(x,y)||p^{\mathbb{W}^{\gamma}}(x,y)), \\ \pi_0(x) = \int q(x,y)dy, \\ \pi_1(y) = \int q(x,y)dx. \end{cases} \quad (1)$$

Для решения данной задачи в подавляющем большинстве для трансформации используют стохастические процессы, решение которых в свою очередь требуют сложных вычислений.

2.2 Проблема

В подавляющем большинстве разработанные алгортимы и способы параметризации основываются на стохастичесих процессах, а трансформация из одного маргинального распределения получается за счет решения стохастических дифференциальных уравнений. В связи с этим возникает потребность в более простых подходах, способных не

Результатом работы проблемы мостов Шредингера являтся некоторые функции $q: x \to y$ и обратная $p: y \to x$, которое переносят $x \sim \pi_0(x)$ в $y \sim \pi_1(y)$ и наоборот. Однако подавляющее число таких оторбражений являются стохатически

Цель данной работы - предложить решение проблемы сложности инференса модели, которое заключается в использовании состязательных сетей в качестве параметризации моделей трансформации, и экспериментально проверить предложенное решение для коллекций с различным балансом тем и для моделей с разными параметрами.

3 Обзор существующих решений

Здесь надо рассмотреть все существующие решения поставленной задачи, но не просто пересказать, в чем там дело, а оценить степень их соответствия тем ограничениям, которые были сформулированы в постановке задачи.

3.1 Состязательные генеративные сети

Состязательные [11] являлись новым, после вариационных автоколировщиков, подходом обучения генеративных моделей, которые обучаются состязательно. Такой метод обучает генеративную модель G (а генеративное распределение p_g изучается неявно) состязательно, внедряя в обучение дискриминативную модель D, которая пытается определить является данный сэмпл реальным или сгенерированным:

$$\begin{aligned}\min_G \max_D V(G,D) = \min_G \max_D \mathbb{E}_{x\sim p_{data}}[\log D(x)] - \mathbb{E}_{x\sim p_g}[1-\log D(x)] = \\ = \min_G \max_D \mathbb{E}_{x\sim p_{data}}[\log D(x)] - \mathbb{E}_{z\sim p(z)}[1-\log D(G(z))],\end{aligned}$$

Также, в статье доказывается, что при оптимальном дискриминаторе D^* минимизируемая целевая функция соответствует дивергенции Йенсена - Шэннона между p_{data} и p_g, то есть:

$$\begin{gathered}\min_G \max_D \mathbb{E}_{x\sim p_{data}}[\log D(x)] - \mathbb{E}_{x\sim p_g}[1-\log D(x)] = \\ \min_G \mathbb{E}_{x\sim p_{data}}[\log D^*(x)] - \mathbb{E}_{x\sim p_g}[1-\log D^*(x)] = \min_G D_{JS}(p_{data}||p_g) - 2\log 2,\end{gathered} \tag{2}$$

Такой подход оказался одним из первых подходов, который позволил качественно генерировать многомерные данные, такие как изображения выского разрешения [**?**].

Однако, несмотря на это состязательные сети являются не самыми простыми моделями для обучения. Из-за минимаксной задачи возникают проблема схлапывания моды и другое.

3.2 f-GAN

Авторы [13] обоснованно расширяют теорию состязательных сетей [11] до более широкого подхода, используя вариационное (двойственное) представление f-девергенций:

$$D_f(P||Q) = \sup_{g:\mathcal{X}\to\mathbb{R}} \mathbb{E}_P[g(X)] - \mathbb{E}_Q[f^*(g(X))],$$

Наример, выбирая $f^*(t) = -\log(1-\exp(t))$ and $g(x) = \log\frac{p(x)}{p(x)+q(x)}$ получим 2.

Авторы выводя целевые функции для всех f-дивергенций и упрощяют алгоритм нахождения седловой точки, изначально предложенный в [11].

3.3 Diffusion Schrodinger Bridge Matching (DSBM)

Авторы статьи [7] предлагают новый метод решения статических мостов Шредингера, используя предложенный Iterative Markovian Fitting.

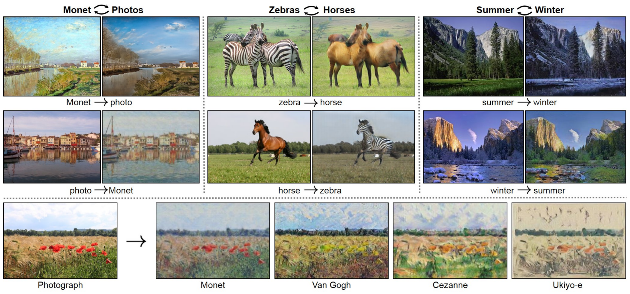

Рис. 3: Пример работы Cycle GAN

3.4 Cycle GAN

Вдохновившись качеством состязательных сетей, авторы [12] предлагают способ применеия GAN для задачи трансформации из изображения в изображение. Авторы обучают две генеративне сети G и F, которые трансформируют изображения из одного в другое, и наоборот. Для того, чтобы добиться лучшего сходимости целвой функции, авторы внедряют дополнительную регуляризацию за непоследовательность модели:

$$smth.$$

В результате удается значительно упростить обучение и добиться хорошего качетва генерируемых изображений 3.

3.5

3.6 Diffusion Schrodinger Bridge with Applications to Score-Based Generative Modeling

dsb proposed using SBP in dynamic way to build generative model

Подход предложенный в статье [8] обладает значительным недостатком

4 Исследование и построение решения задачи

Требуется разбить большую задачу, описанную в постановке, на более мелкие подзадачи. Процесс декомпозиции следует продолжать до тех пор, пока подзадачи не станут достаточно простыми для решения непосредственно. Это может быть достигнуто, например, путем проведения эксперимента, доказательства теоремы или поиска готового решения.

В итоге получается модель которая параметризованна с помощью нормального распределения. В силу того, что начальные условия распределения $q(x|y) = \mathcal{N}(x|y, \gamma\mathbb{I})$ остаются неизбежными, чтобы полностью получить преимущество созтязательных генеративных моделей, необходимо избавиться от такого вида параметризации. Для решения этой проблемы прибегнем к предобучению.

Обучим генератор $q_0(x|y)$ с помощью созтязательного обучения, где реальными данными будут сэмплы из распределения $\mathcal{N}(y, \gamma\mathbb{I})$.

5 Описание практической части

Если в рамках работы писался какой-то код, здесь должно быть его описание: выбранный язык и библиотеки и мотивы выбора, архитектура, схема функционирования, теоретическая сложность алгоритма, характеристики функционирования (скорость/-память).

6 Заключение

Здесь надо перечислить все результаты, полученные в ходе работы. Из текста должно быть понятно, в какой мере решена поставленная задача.

Список литературы

[1] *Ho, Jonathan.* Denoising Diffusion Probabilistic Models / Jonathan Ho, Ajay Jain, Pieter Abbeel. — 2020.

[2] Flow Matching for Generative Modeling / Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu et al. — 2023.

[3] Augmented Bridge Matching / Valentin De Bortoli, Guan-Horng Liu, Tianrong Chen et al. — 2023.

[4] *Schrödinger, E.* Über die Umkehrung der Naturgesetze / E. Schrödinger. — Sitzungsberichte der Preussischen Akademie der Wissenschaften. Physikalisch-mathematische Klasse. Verlag der Akademie der Wissenschaften in Kommission bei Walter De Gruyter u. Company, 1931. `https://books.google.ru/books?id=FTYYtwAACAAJ`.

[5] *Léonard, Christian.* A survey of the Schrödinger problem and some of its connections with optimal transport / Christian Léonard. — 2013.

[6] Diffusion Schrödinger Bridge with Applications to Score-Based Generative Modeling / Valentin De Bortoli, James Thornton, Jeremy Heng, Arnaud Doucet // *arXiv preprint arXiv:2106.01357.* — 2021.

[7] Diffusion Schrödinger Bridge Matching / Yuyang Shi, Valentin De Bortoli, Andrew Campbell, Arnaud Doucet. — 2023.

[8] Solving Schrödinger Bridges via Maximum Likelihood / Francisco Vargas, Pierre Thodoroff, Austen Lamacraft, Neil Lawrence // *Entropy.* — 2021. — . — Vol. 23, no. 9. — P. 1134. `http://dx.doi.org/10.3390/e23091134`.

[9] Unpaired Image-to-Image Translation via Neural Schrödinger Bridge / Beomsu Kim, Gihyun Kwon, Kwanyoung Kim, Jong Chul Ye. — 2024.

[10] *Korotin, Alexander.* Light Schrödinger Bridge. — 2024.

[11] Generative Adversarial Networks / Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza et al. — 2014.

[12] Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks / Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros. — 2020.

[13] *Nowozin, Sebastian.* f-GAN: Training Generative Neural Samplers using Variational Divergence Minimization / Sebastian Nowozin, Botond Cseke, Ryota Tomioka. — 2016.

[14] *Mott-Smith, H.* The theory of collectors in gaseous discharges / H. Mott-Smith, I. Langmuir // *Phys. Rev.* — 1926. — Vol. 28.

[15] *Морз, Р.* Бесстолкновительный PIC-метод / Р. Морз // Вычислительные методы в физике плазмы / Ed. by Б. Олдера, С. Фернбаха, М. Ротенберга. — М.: Мир, 1974.

[16] *Киселёв, А. А.* Численное моделирование захвата ионов бесстолкновительной плазмы электрическим полем поглощающей сферы / А. А. Киселёв, Долгоносов М. С., Красовский В. Л. // Девятая ежегодная конференция «Физика плазмы в Солнечной системе». — 2014.

[17] *Pavon, Michele*. The Data-Driven Schrödinger Bridge / Michele Pavon, Giulio Trigila, Esteban G. Tabak // *Communications on Pure and Applied Mathematics*. — 2021. — Vol. 74, no. 7. — Pp. 1545–1573. https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.21975.

[18] Schrödinger Bridge Samplers / Espen Bernton, Jeremy Heng, Arnaud Doucet, Pierre E. Jacob. — 2019.

[19] Deep Generative Learning via Schrödinger Bridge / Gefei Wang, Yuling Jiao, Qian Xu et al. — 2021. — 18–24 Jul. — Vol. 139. — Pp. 10794–10804. https://proceedings.mlr.press/v139/wang21l.html.

[20] *Vargas, Francisco*. Machine-learning approaches for the empirical Schrödinger bridge problem / Francisco Vargas. — No. UCAM-CL-TR-958. 2021. — . https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-958.pdf.

[21] *Fortet, Robert*. Résolution d'un système d'équations de M. Schrödinger / Robert Fortet // *Journal de Mathématiques Pures et Appliquées*. — 1940. — Vol. 9e série, 19, no. 1-4. http://www.numdam.org/item/JMPA_1940_9_19_1-4_83_0/.

[22] *Essid, Montacer*. Traversing the Schroedinger Bridge strait: Robert Fortet's marvelous proof redux / Montacer Essid, Michele Pavon. — 2018.

[23] *Kullback, S*. Probability Densities with Given Marginals / S. Kullback // *The Annals of Mathematical Statistics*. — 1968. — Vol. 39, no. 4. — Pp. 1236 – 1243. https://doi.org/10.1214/aoms/1177698249.

[24] *Rüschendorf, Ludger*. Convergence of the iterative proportional fitting procedure / Ludger Rüschendorf // *Annals of Statistics*. — 1995. — Vol. 23. — Pp. 1160–1174. https://api.semanticscholar.org/CorpusID:122665767.