

# Потоки для одновременных manifold learning и оценки плотности

Плетнев Никита Вячеславович

Московский физико-технический институт  
Факультет управления и прикладной математики  
Кафедра интеллектуальных систем

15 декабря 2020 г.

# Постановка задачи

Решается задача оценки плотности распределения  $p^*(x)$

при условии, что входные данные принадлежат многообразию меньшей размерности, чем всё пространство.

## Предположения

- данные  $x \in \mathcal{M}^* \subset X = \mathbb{R}^d$  семплируются из распределения  $x \sim p^*(x)$ , где  $\mathcal{M}^*$  —  $n$ -мерное Риманово многообразие,  $n < d$ ;
- это многообразие топологически эквивалентно  $\mathbb{R}^n$ , причём размерность  $n$  известна.

Для упрощения описываем все генеративные модели в терминах двух векторов латентных переменных:  $u \in U = \mathbb{R}^n$  (латентное пространство, отображаемое в многообразие  $\mathcal{M}$ ) и  $v \in V$  — параметры «вне многообразия».

## Потоки в окружающем пространстве (ambient flows, AF)

$$f : U \times V \rightarrow X; \quad u, v \rightarrow f(u, v)$$

с вычислимой плотностью  $p_{uv}(u, v)$ . Результат:

$$p_x(x) = p_{uv}(f^{-1}(x)) |\det J_f(f^{-1}(x))|^{-1}.$$

Плотность получается «размазанной» по всему пространству.

## Flows on a manifold (FOM)

$$g^* : U \rightarrow \mathcal{M}^* \subset X; \quad u \rightarrow g^*(u),$$

$p_u(u)$  моделируется normalizing flow. Результат:

$$p_x(x) = p_u(g^{*-1}(x)) |\det[J_g^T(g^{*-1}(x))J_g(g^{*-1}(x))]|^{-\frac{1}{2}}.$$

Многообразие считается априори известным.

# Возможные пути решения и их проблемы

## Генеративно-состязательные сети (GAN)

$$g : U \rightarrow \mathcal{M} \subset X; \quad u \rightarrow g(u),$$

где  $g$  — обучаемое вложение; поэтому плотность невычислима.

## Вариационные автокодировщики (VAE)

Плотность невычислима, только ELBO.

## Псевдообратимые кодировщики (PIE) и их последовательности

Уравнения как для AF, но для разных латентных переменных выбираются разные базовые плотности: для  $u$  —  $n$ -мерный евклидов поток, для  $v$  — с резким пиком около 0.

Однако для PIE не достигается ограничение плотности требуемым многообразием, а для их последовательности плотность вычислима лишь с точностью до нормировки.

Beitler, Sosnovik, Smeulders (2019): Pseudo-Invertible Encoder.

# Предлагаемое решение: $\mathcal{M}$ -поток

## Основная идея

Соединить вычислимую оценку плотности, присущую PIE, со связью латентного пространства и многообразия, которая есть у GAN, и добавить оценку близости точки к многообразию.

## Алгоритм

Связь между латентным пространством и данными:

$$f : U \times V \rightarrow X; \quad u, v \rightarrow f(u, v).$$

Многообразие для модели:

$$g : U \rightarrow \mathcal{M} \subset X; \quad u \rightarrow g(u) = f(u, 0),$$

На практике  $g = f_k \circ \dots \circ f_1 \circ \text{Pad}$

Базовая плотность  $p_u(u)$  моделируется  $n$ -мерным потоком  $h$ , отображающим  $u$  в другое латентное пространство с  $p_{\tilde{u}}(\tilde{u})$ .

## Оценка плотности

$$\begin{aligned} p_{\mathcal{M}}(x) &= p_u(g^{-1}(x)) |\det[J_g^T(g^{-1}(x))J_g(g^{-1}(x))]|^{-\frac{1}{2}} = \\ &= p_{\tilde{u}}(h^{-1}(g^{-1}(x))) |\det J_h(h^{-1}(g^{-1}(x)))|^{-1} \times \\ &\quad \times |\det[J_g^T(g^{-1}(x))J_g(g^{-1}(x))]|^{-\frac{1}{2}}. \end{aligned}$$

## Сэмплирование

$$\tilde{u} \sim p_{\tilde{u}}(\tilde{u}), \quad u = h(\tilde{u}), \quad x = g(u) = f(u, 0).$$

## Оценка принадлежности многообразию

$g$  отображает из пространства меньшей размерности и, по сути, является декодировщиком. Определим кодировщик.

$$g^{-1} : X \rightarrow U, \quad x \rightarrow g^{-1}(x) = Proj(f^{-1}(x))$$

Здесь  $Proj(u, v) = u$ .

Ошибка восстановления:

$$\|x - x'\| = \|x - g(g^{-1}(x))\|$$

## Проверка входных данных $x \in X$

- отобразить в латентное пространство;
- спроектировать на многообразие (оценивается плотность);
- отобразить в исходное пространство (вычисляется ошибка восстановления).

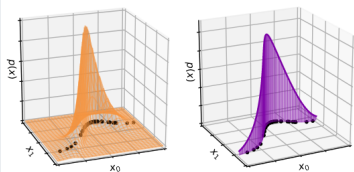
# Предлагаемое решение: $\mathcal{M}_e$ -потoki

## Отличие от $\mathcal{M}$ -потоков

Вместо  $g^{-1}$  используется обучаемый кодировщик  $e(x)$ .

## Сравнение подходов

Model	Manifold	Chart	Generative mode	Tractable density	Restricted to manifold
Ambient flow (AF)	no manifold	×	✓	✓	×
Flow on manifold (FOM)	prescribed	✓	✓	✓	✓
Generative adversarial network (GAN)	learned	×	✓	×	✓
Variational autoencoder (VAE)	learned	×	✓	only ELBO	(×)
Pseudo-invertible encoder (PIE)	learned	✓	✓	✓	(×)
Slice of PIE	learned	✓	×	up to normalization	✓
Manifold-learning flow ( $\mathcal{M}$ -flow)	learned	✓	✓	✓ (may be slow)	✓
Manifold-learning flow with sep. encoder ( $\mathcal{M}_e$ -flow)	learned	✓	✓	✓ (may be slow)	✓





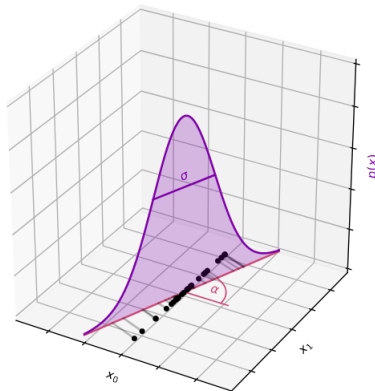
## Преимущества $\mathcal{M}$ -потоков

- более точная аппроксимация истинного распределения;
- увеличение производительности;
- возможность моделировать условную плотность на фиксированном носителе;
- скрытое пространство меньшей размерности снижает сложность модели;
- проекция на носитель обеспечивает уменьшение размерности и возможность шумоподавления;
- возможно использование для обнаружения образцов вне распределения.

## Сложности использования

Обычный подход к оценке плотности распределения — максимизация правдоподобия — в данном случае не работает.

## Проблема максимизации правдоподобия



(a) Setup. The model manifold is a straight line in 2D Euclidean space that passes through the origin and is rotated with respect to the  $x_1$ -axis by an angle  $\alpha$ . On this line, the density is a Gaussian with mean at the origin, its standard deviation is a model parameter  $\sigma$ . The training data (black dots) are generated with  $\alpha^* = \pi/2$  and  $\sigma^* = 1$ .

При обучении есть две цели: построение многообразия и оценка плотности. Соответственно, минимизация ошибки восстановления и максимизация правдоподобия на многообразии - разные шаги обучения.

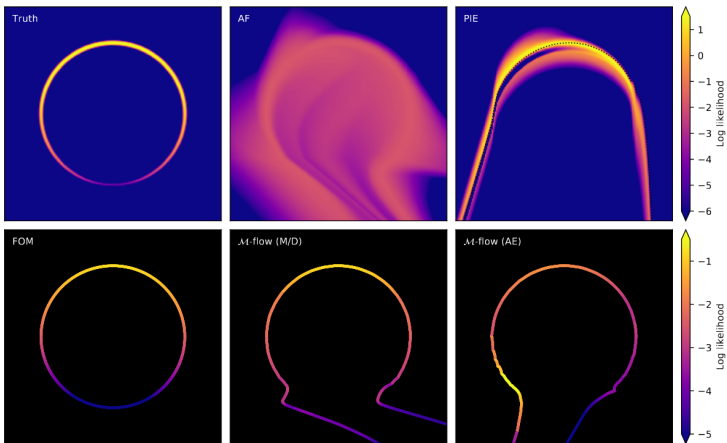
Manifold phase:

$$L_{manifold}[g] = \frac{1}{b} \sum_x \|x - g(g^{-1}(x))\|_2^2 \rightarrow \min.$$

Density phase:

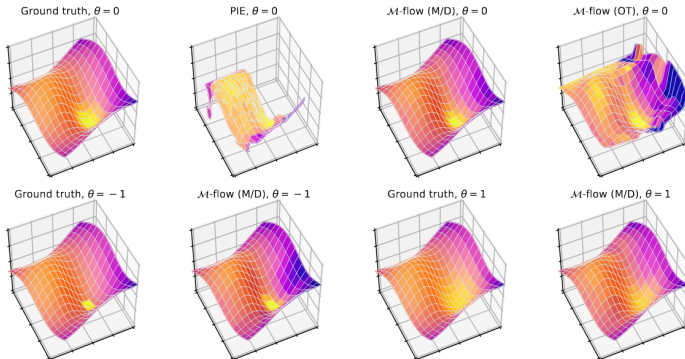
$$L_{density}[h] = -\frac{1}{b} \sum_x \log p_u(g^{-1}(x))$$

## «Игрушечный» эксперимент на синтетических данных



**Fig. 5.** Learning a Gaussian density on a circle. **Top left:** true density of the data-generating process. **Top middle and top right:** 2D density learned by a standard ambient flow (AF) and a PIE). **Bottom:** manifold and density learned by a manifold flow with specified true manifold (FOM), a manifold-learning flow ( $\mathcal{M}$ -flow (M/D)), and a manifold-learning flow that was only trained on the reconstruction error ( $\mathcal{M}$ -flow (AE)). To highlight the differences, we use simple, less expressive architectures (see text).

## Смеси моделей для полиномиальных поверхностей



**Fig. 6.** Mixture model on a polynomial surface. **Top:** the true data manifold as well as the manifolds learned by the PIE,  $\mathcal{M}$ -flow (M/D), and  $\mathcal{M}$ -flow (OT) models. The color shows the log likelihood for  $\theta = 0$  (bright yellow represents a high density, dark blue a low density). In order to increase the clarity of the PIE panel we have removed parts of that manifold which “fold” above and below the shown part. **Bottom:** ground truth and  $\mathcal{M}$ -flow (M/D) manifold for  $\theta = -1$  and  $\theta = 1$ .

## Обработка данных о столкновениях протонов с БАК

Model (algorithm)	Sample closure	Mean reconstruction error	Log posterior
AF	<b>0.0019</b> $\pm$ 0.0001	—	−3.94 $\pm$ 0.87
PIE (original)	0.0023 $\pm$ 0.0001	2.054 $\pm$ 0.076	−4.68 $\pm$ 1.56
PIE (unconditional manifold)	0.0022 $\pm$ 0.0001	1.681 $\pm$ 0.136	−1.82 $\pm$ 0.18
$\mathcal{M}$ -flow	0.0045 $\pm$ 0.0004	<b>0.012</b> $\pm$ 0.001	−1.71 $\pm$ 0.30
$\mathcal{M}_e$ -flow	0.0046 $\pm$ 0.0002	0.029 $\pm$ 0.001	−1.44 $\pm$ 0.34
AF (SCANDAL)	0.0565 $\pm$ 0.0059	0.000 $\pm$ 0.000	−0.40 $\pm$ 0.09
PIE (original, SCANDAL)	0.1293 $\pm$ 0.0218	3.090 $\pm$ 0.052	0.03 $\pm$ 0.17
PIE (uncond. manifold, SCANDAL)	0.1019 $\pm$ 0.0104	1.751 $\pm$ 0.064	<b>0.23</b> $\pm$ 0.05
$\mathcal{M}$ -flow (SCANDAL)	0.0371 $\pm$ 0.0030	<b>0.011</b> $\pm$ 0.001	0.11 $\pm$ 0.04
$\mathcal{M}_e$ -flow (SCANDAL)	0.0291 $\pm$ 0.0010	0.030 $\pm$ 0.002	<b>0.14</b> $\pm$ 0.09

SCANDAL — методика аугментации; видно, что ошибка восстановления у  $\mathcal{M}$ -потоков минимальна.

Предложены два типа потоков с использованием manifold learning:  $\mathcal{M}$ -потоки и  $\mathcal{M}_e$ -потоки.

Выявлена тонкость в наивной интерпретации плотности таких моделей и предложены более эффективные стратегии обучения.

Полученные результаты проверены экспериментально.

## Научный прогресс

Построенный метод позволит эффективнее извлекать знания из крупномасштабных экспериментов, как показывает эксперимент на данных об элементарных частицах.

## Генерация изображений

Существует риск злоупотребления для генерации и дальнейшего использования реалистичных поддельных данных.



- [1 ] Johann Brehmer and Kyle Cranmer: Flows for simultaneous manifold learning and density estimation, <https://arxiv.org/pdf/2003.13913.pdf>
- [2 ] Beitler JJ, Sosnovik I, Smeulders A (2019) PIE: Pseudo-Invertible Encoder.