

---

# Метрический анализ пространства параметров глубоких нейросетей.

---

A Preprint

Эрнест Р. Насыров\*

ФПМИ

МФТИ

Долгопрудный

`nasyrov.rr@phystech.edu`

28 февраля 2023 г.

Abstract

В работе исследуется проблема снижения размерности пространства параметров модели. В ее рамках решается задача восстановления временного ряда. В качестве модели восстановления ряда используются различные автоэнкодеры. В работе проводится метрический анализ пространства параметров автоэнкодера. Новизна заключается в том, что отдельные параметры модели - случайные величины - собираются в векторы – многомерные случайные величины, анализ взаимного расположения которых в пространстве и представляет предмет исследования нашей работы. Этот анализ позволит снизить количество параметров модели, сделать выводы о значимости параметров, произвести их отбор. Для определения положения вектора параметров в пространстве оцениваются его матожидание и матрица ковариации с помощью методов bootstrap и variational inference. Эксперименты проводятся на моделях SSA, RNN и VAE на задачах предсказания синтетических временных рядов, квазипериодических показаний акселерометра, периодических видеоданных.

---

\*Use footnote for providing further information about author (webpage, alternative address)

## 1 Introduction

С развитием технологий скорость обработки данных растет, они становятся более сложными, большей размерности. Такие высокоразмерные данные часто избыточны, хотя и содержат много полезной информации, что представляет сложность для их эффективной обработки и использования. Отсюда возникает задача снижения размерности признакового описания объекта, которая привлекла большое внимание ученых. Ее базовый принцип состоит в том, чтобы отобразить высокоразмерное признаковое пространство в низкоразмерное, сохраняя важную информацию о данных [Jia et al., 2022].

На текущий момент известно много методов снижения размерности данных. В работе [Örnek and Vural, 2019] снижения размерности достигается за счет построения дифференцируемой функции эмбединга в низкоразмерное представление, а в [Cunningham and Yu, 2014] обсуждаются линейные методы.

В работе [Isachenko and Strijov, 2022] задача снижения размерности решается для предсказания движения конечностей человека по электрокортикограмме с использованием метода QPFS, учитывающем мультикоррелированность и входных, и целевых признаков.

Наряду с задачей снижения размерности входных данных стоит задача выбора оптимальной структуры модели. В случае оптимизации структуры нейросети, большое внимание уделено изучению признакового пространства модели. В работах [Hassibi et al., 1993] и [Dong et al., 2017] применяется метод OBS (Optimal Brain Surgeon), состоящий в удалении весов сети с сохранением ее качества аппроксимации, причем выбор удаляемых весов производится с помощью вычисления гессиана функции ошибки по весам.

В статье [Грабовой et al., 2019] приводится метод первого порядка, решающий задачу удаления весов, основанный на нахождении дисперсии градиента функции ошибки по параметру и анализе ковариационной матрицы параметров, а в статье [Грабовой et al., 2020] нерелевантные веса не удаляют, а прекращают их обучение.

Приведенные выше задачи понижения размерности данных и выбора оптимальной структуры нейросети основаны на исследовании пространства входных данных и пространства признаков соответственно. На взгляд авторов статьи, существенный недостаток предыдущих работ состоит в том, что в них анализируются отдельные параметры (скаляры) моделей и их взаимозависимость. Тем самым не учитывается, что на входные данные действуют вектора параметров посредством скалярных произведений, то есть упускается из виду простая структура преобразования.

В данной работе мы решаем задачу восстановления временного ряда, в рамках которой занимаемся проблемой снижения пространства параметров модели, основанном на анализе сопряженного пространства ко входному, которое связывает входное пространство и пространство параметров.

Наше исследование в большой степени полагается на простоту устройства глубоких нейросетей, которые являются композицией линейных и простых нелинейных функций (функций активации). Составной блок нейросети может быть описан формулой  $y = \sigma(Wx)$ ,  $y \in \mathbb{R}^m$ ,  $x \in \mathbb{R}^n$ ,  $W \in \mathbb{R}^{m \times n}$ ,  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ . Тогда если раньше элементы  $W_{ij}$  исследовались по-отдельности, как скаляры, то в нашей работе изучаются

векторы-строки  $w_1, \dots, w_m : W = \begin{pmatrix} w_1^T \\ \dots \\ w_m^T \end{pmatrix}$ . В нейросети эти строки обычно называются нейронами. В

SSA  $\sigma = Id$ , а матрица  $W = W_k$  это приближение истинной матрицы фазовых траекторий  $X$  (матрицы Ганкеля) суммой  $k$  элементарных матриц.

Мы предполагаем, что каждая точка фазовой траектории распределена нормально вокруг своего матожидания. Тогда обучающая выборка  $\mathcal{D} = (\mathbf{X}, \mathbf{y})$  это набор случайных величин, поэтому и результат обучения модели на ней, то есть веса модели тоже будут случайными.

В работе исследуется положение случайных векторов параметров модели  $w_i$  в метрическом пространстве. С помощью методов bootstrap и Variational Inference [Hastie et al., 2009] оцениваются их матожидания  $e_i = Ew_i$  и ковариационные матрицы  $var(w_i) = \mathbf{A}_i^{-1}$ . Мы работаем в гипотезе, что эти векторы  $w_i$  распределены нормально, таким образом пара  $(e_i, \mathbf{A}_i^{-1})$  полностью описывает вероятностное распределение вектора  $w_i$ .

В качестве графического анализа пространства производится изображение положения этих векторов как смеси гауссианов 1, а также изображение 95% доверительной области каждого вектора 2.

Уменьшение размерности достигается за счет метрического анализа пространства векторов-параметров путем отбора релевантных строк (с малой дисперсией), замены мультикоррелирующих строк на их линейную композицию с помощью обобщения алгоритма QPFS, изучения структуры сообществ строк.

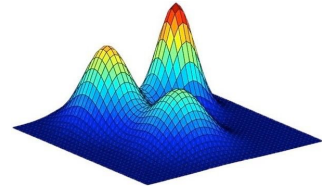


Рис. 1: Пример смеси гауссианов для 2-х мерных векторов.

В качестве базовых моделей используются SSA ([Golyandina et al., 2001]), нелинейный PCA???(А как назвать?, ссылка на источник!), RNN ([Bronstein et al., 2021]) и NeuroODE ([Chen et al., 2018]).

Задача восстановления временного ряда решается на синтетических данных зашумленного *sin*, данных показания акселерометра в датасете MotionSense3 [Malekzadeh et al., 2018] (НЕТ ДОСТУПА К ЭТИМ ДАННЫМ: ДОСТУПНЫ ТОЛЬКО ПО ЗАПРОСУ), периодических видеоданных (НАЙТИ ИХ).

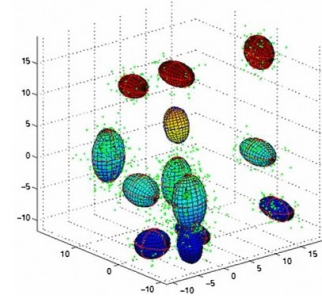


Рис. 2: Пример доверительных областей для 3-х мерных векторов.

## Список литературы

- Weikuan Jia, Meili Sun, Jian Lian, and Sujuan Hou. Feature dimensionality reduction: a review. *Complex & Intelligent Systems*, 8(3):2663–2693, 2022.
- Cem Örneк and Elif Vural. Nonlinear supervised dimensionality reduction via smooth regular embeddings. *Pattern Recognition*, 87:55–66, 2019.
- John P Cunningham and Byron M Yu. Dimensionality reduction for large-scale neural recordings. *Nature neuroscience*, 17(11):1500–1509, 2014.
- RV Isachenko and VV Strijov. Quadratic programming feature selection for multicorrelated signal decoding with partial least squares. *Expert Systems with Applications*, 207:117967, 2022.
- Babak Hassibi, David G Stork, and Gregory J Wolff. Optimal brain surgeon and general network pruning. In *IEEE international conference on neural networks*, pages 293–299. IEEE, 1993.
- Xin Dong, Shangyu Chen, and Sinno Pan. Learning to prune deep neural networks via layer-wise optimal brain surgeon. *Advances in Neural Information Processing Systems*, 30, 2017.
- Андрей Валериевич Грабовой, Олег Юрьевич Бахтеев, and Вадим Викторович Стрижов. Определение релевантности параметров нейросети. *Информатика и её применения*, 13(2):62–70, 2019.
- Андрей Валериевич Грабовой, Олег Юрьевич Бахтеев, and Вадим Викторович Стрижов. Введение отношения порядка на множестве параметров аппроксимирующих моделей. *Информатика и её применения*, 14(2):58–65, 2020.
- Trevor Hastie, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.
- Nina Golyandina, Vladimir Nekrutkin, and Anatoly A Zhigljavsky. *Analysis of time series structure: SSA and related techniques*. CRC press, 2001.

Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković.

Geometric deep learning: Grids, groups, graphs, geodesics, and gauges.

arXiv preprint arXiv:2104.13478, pages 89–95, 2021.

Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K

Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

Mohammad Malekzadeh, Richard G Clegg, Andrea Cavallaro, and Hamed

Haddadi. Protecting sensory data against sensitive inferences. In *Proceedings of the 1st Workshop on Privacy by Design in Distributed Systems*, pages 1–6, 2018.