Expressive Power of Recurrent Neural Networks, II

Выполнил: Алкин Эмиль Венерович[1,2]

Научный руководитель: Оселедец Иван Валерьевич[2]

[1]Московский физико-технический институт
[2]Сколковский институт науки и технологий

16 декабря 2023 г.

Постановка задачи

Мотивация

Одним из актуальных направлений в исследовании теоретических основ глубоких нейронных сетей является изучение выразительной силы различных архитектур глубоких сетей. В значимых статьях ([1], [2]) по этой теме в качестве меры выразительности сети рассматривается канонический CP-ранг тензора, соответствующего определенной архитектуре сети. Такой подход оказался удобным и его более широкое применение имеет актуальность.

[1] Nadav Cohen et al. On the expressive power of deep learning: A tensor analysis. (2016)
[2] Valentin Khrulkov et al. Expressive power of recurrent neural networks (2017)

Постановка задачи

Проблема

В статье [2] показана связь между реккурентными нейронными сетями и Tensor Train разложением тензора, а также доказана теорема о "нижней оценке" CP-ранга тензора, соответствующего реккурентной сети, в которой каждый слой имеет собственный набор весов. Аналогичный результат для RNN с количеством весов, не зависящим от глубины сети, не доказан, а сформулирован в качестве гипотезы. Также, эксперименты в этой статье лишь частично подтверждают полученные теоритические результаты.

[2] Valentin Khrulkov et al. Expressive power of recurrent neural networks (2017)

Связь между Tensor Train разложением тензора и реккурентной архитектурой сетей

A tensor $\mathcal{X}$ is said to be represented in the Tensor Train (TT) format if each element of $\mathcal{X}$ can be computed as follows

$$\mathcal{X}^{i_1 i_2 \dots i_d} = \sum_{\alpha_1=1}^{r_1} \sum_{\alpha_2=1}^{r_2} \dots \sum_{\alpha_{d-1}=1}^{r_{d-1}} G_1^{i_1 \alpha_1} G_2^{\alpha_1 i_2 \alpha_2} \dots G_d^{\alpha_{d-1} i_d}$$

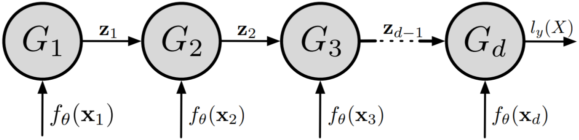

Рис.: Recurrent-type neural architecture that corresponds to the Tensor Train decomposition. Gray circles are bilinear maps.

Обозначения

Let us denote $\mathrm{n} = (n_1, n_2 \dots n_d)$. Set of all tensors $\mathcal{X}$ with mode sizes n representable in TT-format with

$$\operatorname{rank}_{TT} \mathcal{X} \leqslant \mathrm{r},$$

for some vector of positive integers r (inequality is understood entry-wise) forms an irreducible algebraic variety, which we denote by $\mathcal{M}_{\mathrm{n},\mathrm{r}}$.
By $\mathcal{M}^{eq}_{\mathrm{n},\mathrm{r}}$ we denote such tensors in $\mathcal{M}_{\mathrm{n},\mathrm{r}}$ whose TT-cores are equal (except the first core and the last core).

Основной результат

Гипотеза 1 из статьи [2], сформулированная в качестве теоремы:

Theorem

Suppose that $d = 2k$ is even. Define the following set

$$B := \left\{ \mathcal{X} \in \mathcal{M}^{eq}_{\mathrm{n,r}} : \operatorname{rank}_{CP} \mathcal{X} < q^{\frac{d}{2}} \right\},$$

where $q = \min\{n, r-1\}$.
Then

$$\mu(B) = 0,$$

where μ is the standard Lebesgue measure on $\mathcal{M}^{eq}_{\mathrm{n,r}}$.

[2] Valentin Khrulkov et al. Expressive power of recurrent neural networks (2017)

Заключение

- Доказана Гипотеза 1 из статьи [2];
- Повторены эксперименты на синтетических данных из статьи [2];
- В планах модифицировать архитектуру реккурентных сетей, предложеннную в статье [2], для получения высокого качества на датасете CIFAR-10.

[2] Valentin Khrulkov et al. Expressive power of recurrent neural networks (2017)