

# Simplex LoRA: Expand important adapters

Давыденко Григорий, Шалыгин Игорь

Научный руководитель: Безносиков А. Н.

Научный консультант: Веприков А. С.

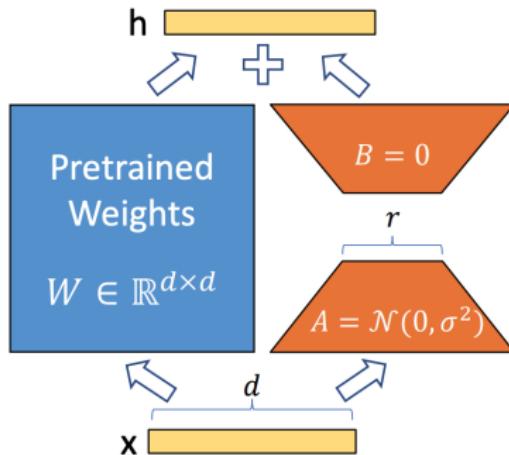
Московский физико-технический институт

18 марта 2025 г.

# Содержание доклада

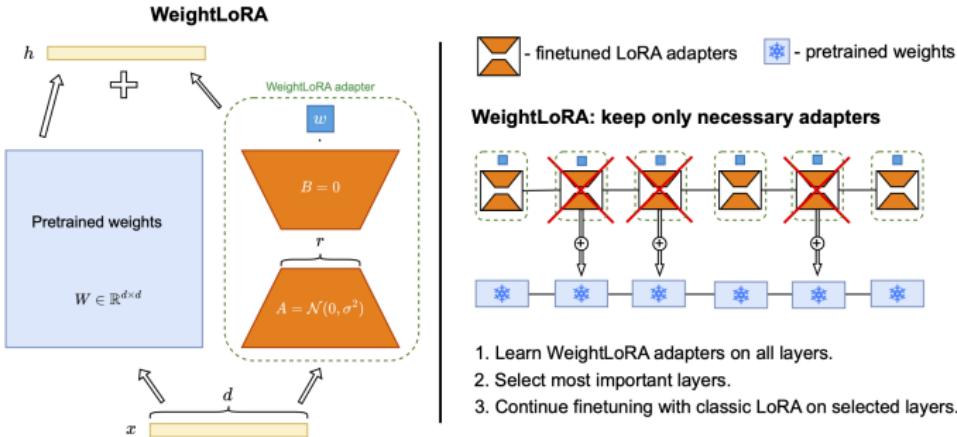
- ① LoRA: Low-Rank Adaptation
- ② Предыстория: WeightLoRA и WeightLoRA+
- ③ SimplexLoRA
- ④ Структура репозитория
- ⑤ Научный и технический продукт
- ⑥ Модели, датасеты и бенчмаркинг

# LoRA: Low-Rank Adaptation



- Тюнинг линейных слоев больших моделей через низкоранговые матрицы  $A \in \mathbb{R}^{d \times r}$  и  $B \in \mathbb{R}^{r \times d}$
- $r \ll d$  (например,  $r = 8$  для моделей с  $d = 1024$ )

# Предыстория: WeightLoRA и WeightLoRA+



- Основная идея WeightLoRA: избавиться с помощью параметров важности  $\omega$  от лишних LoRA
- Основная идея WeightLoRA+: увеличить ранги оставшихся LoRA

# SimplexLoRA

Теперь  $\omega$  - это веса на симплексе ( $\sum_{i=1}^n \omega_i = 1$ ).

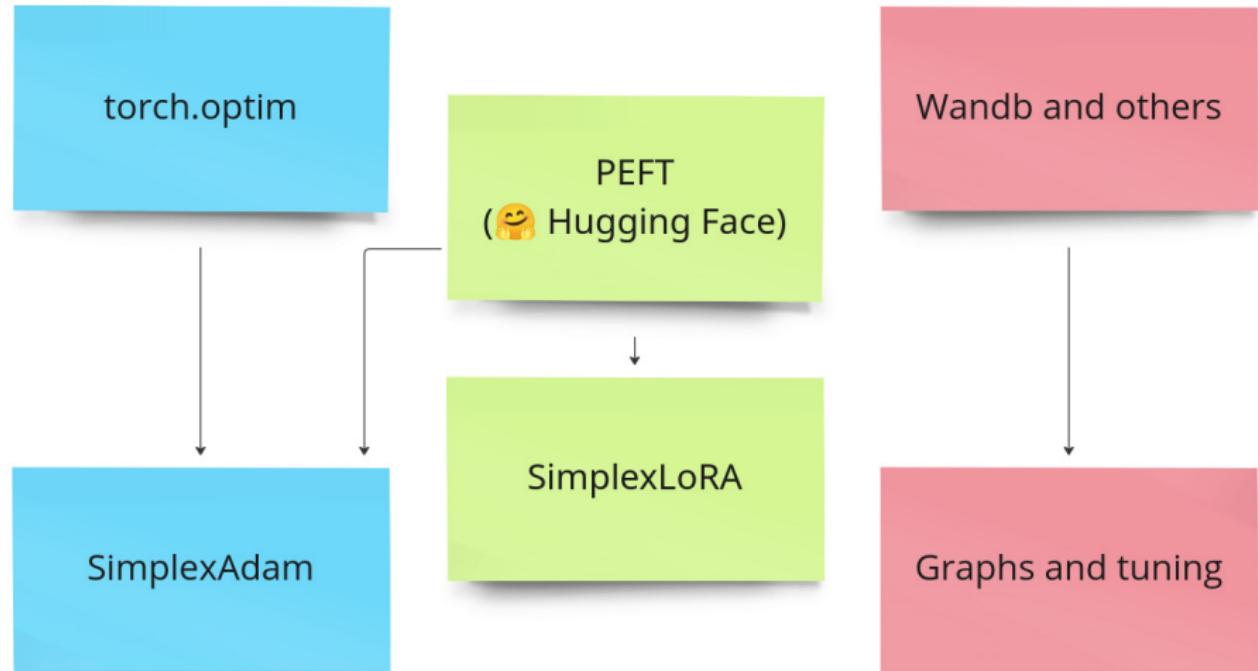
За время обучения мы совершаём раз в несколько шагов изменения рангов LoRA в соответствии с их обучаемыми весами:

- ➊ Сначала все ранги равны гиперпараметру  $r_i^{(0)} = r^0$ , веса -  $\omega_i^{(0)} = \frac{1}{n}$
- ➋ После шага в оптимизаторе  $\omega^{(k)} = proj_{simplex}(\omega^{(k)})$
- ➌ Через фиксированное количество шагов  $x$  меняем ранги LoRA согласно новым весам:

$$r_i^{(u)} = \text{ceil}(r^0 \cdot n \cdot \omega_i^{(u \cdot x)})$$

- ➍ Пункты 2 и 3 повторяем несколько раз.
- ➎ Запускаем стандартное обучение LoRA с подобранными рангами.

# Структура репозитория



# Научный и технический продукт

## Научный продукт

- Эффективный метод выбора наиболее важных лор
- Балансировка между эффективностью алгоритма и затратами по памяти и FLOP

## Технический продукт

- Реализация SimplexLoRA в библиотеке PEFT
- Оптимизатор SimplexAdam в библиотеке PEFT

# Модели, датасеты и бенчмаркинг

microsoft/  
deberta-v3-base

facebook/  
bart-large

Llama



GLUE + SQuAD

XSum

?

## Приложение: ссылки

- LoRA: Low-Rank Adaptation of Large Language Models
- DeBERTa model
- General Language Understanding Evaluation (GLUE) benchmark
- Stanford Question Answering Dataset (SQuAD)
- BART model
- The Extreme Summarization (XSum) dataset