Аддитивная регуляризация и ее метапараметры при выборе структуры сетей глубокого обучения

Кирилл Олегович Вайсер

Московский физико-технический институт

Курс: Численные методы обучения по прецедентам (практика, В. В. Стрижов)/Группа 774, весна 2020
Консультант: аспирант М. С. Потанин
Научный руководитель: д.ф.-м.н. В. В. Стрижов

Цель исследования

Цель работы

Предложить метод построения критерия качества модели, учитывающего его точность, сложность и устойчивость.

Проблема

Нейронные сети, обладающие большой точностью, обычно демонстрируют высокую сложность и низкую устойчивость.

Метод решения

Построить функцию ошибки, включающую аддитивную регуляризацию

Задана выборка, конечное множество пар

$$(\mathbf{x}, y) \in D, \quad \mathbf{x} \in \mathbb{R}^n, \quad y \in \mathbb{R}$$

Структура модели имеет следующий вид

$$f = \sigma_k \circ \underset{1\times 1}{\mathbf{w}_k^{\mathsf{T}}} \boldsymbol{\sigma}_{k-1} \circ \mathbf{W}_{k-1} \boldsymbol{\sigma}_{k-2} \circ \dots \circ \underset{n_2\times 1}{\mathbf{W}_2 \boldsymbol{\sigma}_1} \circ \underset{n_1\times n}{\mathbf{W}_1} \underset{n\times 1}{\mathbf{x}} \tag{1}$$

Функция ошибки имеет вид

$$L = \lambda_x E_{\mathbf{x}} + \lambda_y E_{\mathbf{y}} + \lambda_1 \mathcal{R}_1 + \dots + \lambda_k \mathcal{R}_k = \lambda_x E_x + \lambda_y E_y + \sum_{i=1}^{k} \lambda_i \mathcal{R}_i(\mathbf{W}), \tag{2}$$

где $\mathcal{R}_i = \mathcal{R}(\mathbf{W}) = \left(\mathfrak{r}_1(\mathbf{W}), \cdots, \mathfrak{r}_r(\mathbf{W})\right)^{\mathsf{T}}$ —вектор, состоящий из значений регуляризаторов i-ого слоя.
Требуется решить задачу

$$\begin{aligned} \mathbf{w} &= \arg\min L(f|\lambda) \\ \lambda &= L(f|\mathbf{w}) \end{aligned} \tag{3}$$

при условии минимизации дисперсии параметров и сложности модели.

$$\sum_{i=1}^{k} \frac{1}{|W_k|} W_k^{\mathsf{T}} W_k \to \min \qquad \sum_{i=1}^{k} |W_k| \to \min \tag{4}$$

Решение

Задание критерия качества

Вводится понятие сложности, устойчивости и точности модели.

Построение критерия качества

Предлагается построить функцию ошибки, включающей аддитивную регуляризацию. Задаются гиперпараметры модели и метапараметры аддитивной регуляризации.

Расписание оптимизации

Метапараметры оптимизации изменяются с течением итераций обучения сети. Порядок и правило их изменения задается экспертно.

Свойства модели

- Сложность —мера множества допустимых значений параметров модели.
- Устойчивость — дисперсия ошибки и параметров.
- Точность —качество аппроксимации, выражаемое через значение функции ошибки.

Роль в аддитивной регуляризации	Тип регуляризатора
Ошибка выхода нейронной сети	$\lvert\lvert\mathbf{y} - f(\mathbf{W})\rvert\rvert_2^2$
Ошибка восстановления на каждом слое	$\lvert\lvert\mathbf{x} - \mathbf{r}(\mathbf{x})\rvert\rvert_2^2$
L_1 и L_2 регуляризация	$\lvert\lvert\mathbf{w} - \mathbf{w}_0\rvert\rvert_1$, $\lvert\lvert\mathbf{w} - \mathbf{w}_0\rvert\rvert_2^2$
Штраф за отличие матрицы одного слоя от тождественного преобразования	$\lvert\lvert\mathbf{W} - \mathbf{I}\rvert\rvert$
Штраф за отличие матрицы одного слоя от метода главных компонент	$\lvert\lvert\mathbf{W}\mathbf{W}^T - \mathbf{I}\rvert\rvert$
Тихоновская регуляризация	$\lVert\mathbf{T}\mathbf{W}\rVert$

Таблица: Каталог регуляризаторов аддитивной функции ошибки

Пример расписания регуляризации

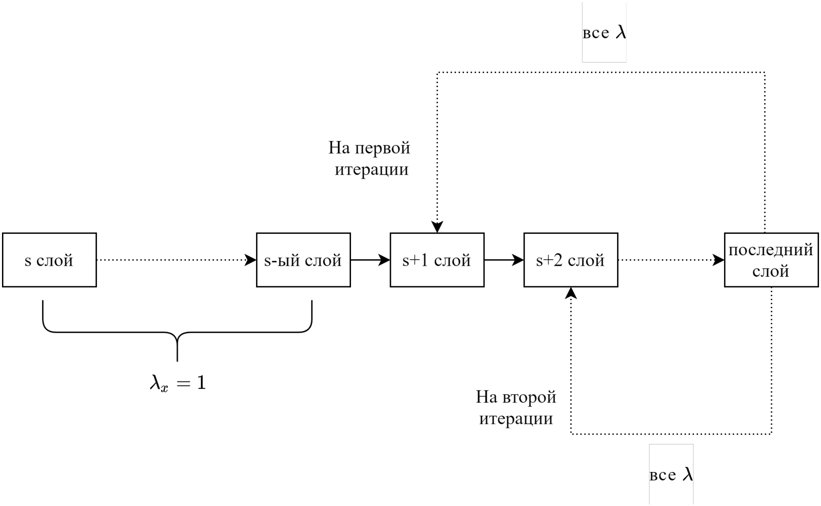

Рис.: Экспертное задание расписания оптимизации, вар 1

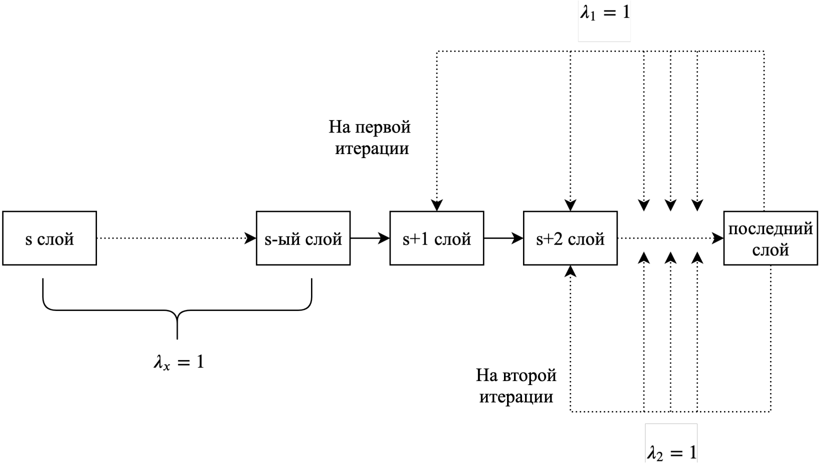

Рис.: Экспертное задание расписания оптимизации, вар 2

Цель эксперимента

- Проверить работоспособность предложенного метода и его соответствие целям исследования.
- Продемонстрировать эффект от использования аддитивной регуляризации

Дисперсия параметров во время обучения

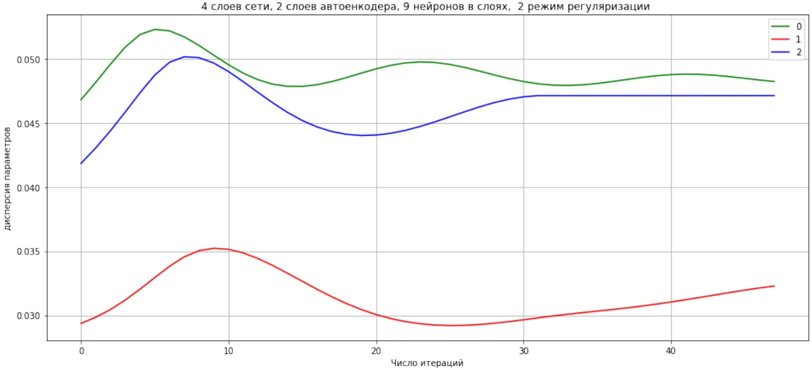

Рис.: Обычное поведение дисперсии параметров : в регуляризованных моделях она меньше.

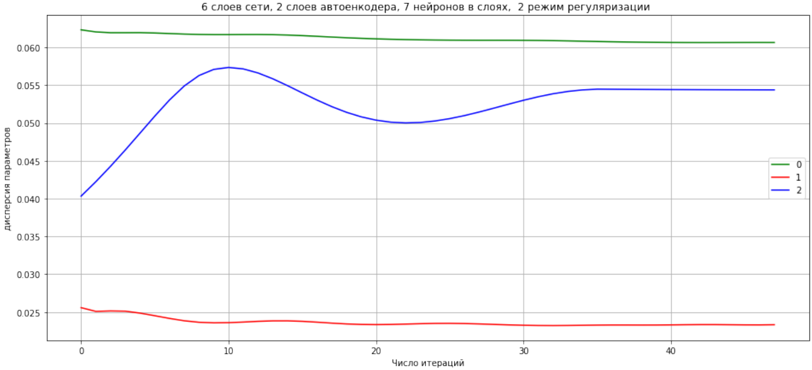

Рис.: Первое расписание оптимизации демонстрирует большую эффективность

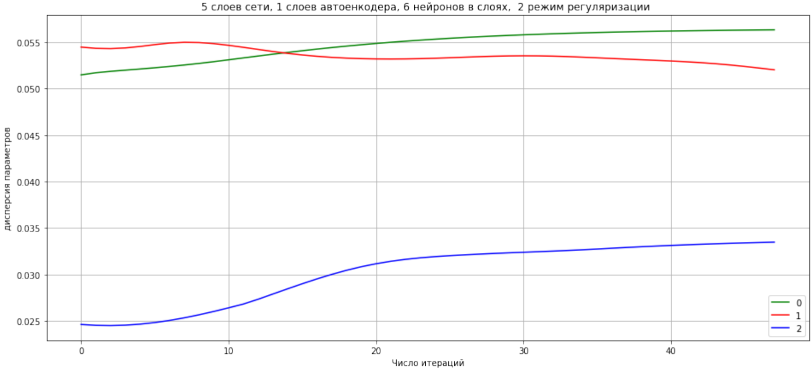

Рис.: Первое расписание оптимизации демонстрирует большую эффективность

Зависимость точности от сложности

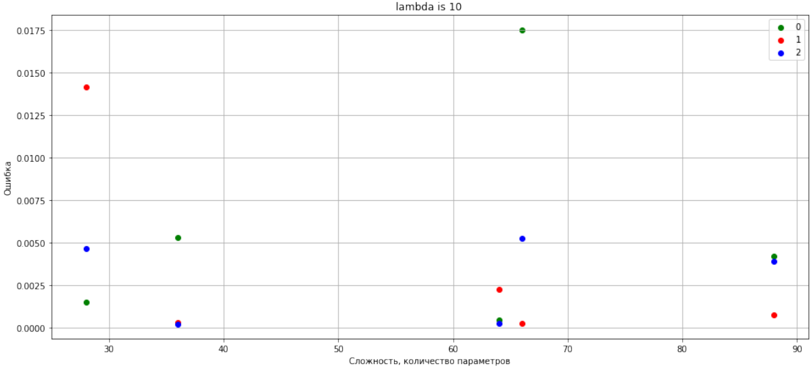

Рис.: Зависимость между сложностью и точностью для разных расписаний оптимизации

Зависимость между устойчивостью и сложностью

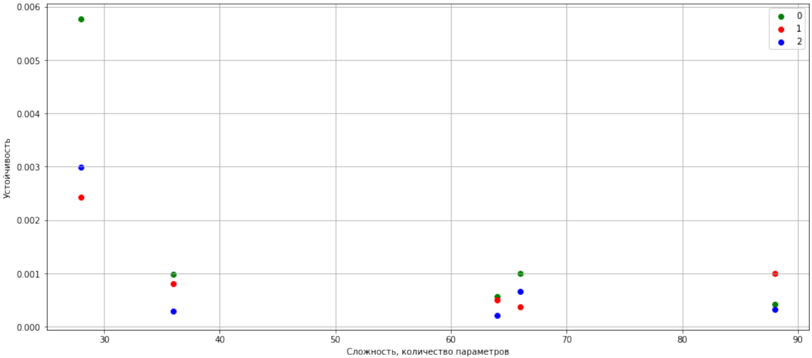

Рис.: Зависимость между сложностью и устойчивостью для разных расписаний оптимизации

1. Показано, что регуляризованная модель имеет большую точность при равной сложности.
2. Показано, что регуляризованная модель имеет лучшую устойчивость при равной сложности.
3. Это предварительные результаты, по окончательным будет построена таблица с сравнением точности, сложности и устойчивости по разным выборкам и конфигурациям модели.

Список литературы:

- Tikhonov, Andrey N and Arsenin, Vasiliy Y Solutions of ill-posed problems // New York, 1977
- Svensén, Markus and Bishop, Christopher M Pattern Recognition and Machine Learning
- Tibshirani, Robert Regression shrinkage and selection via the lasso // Journal of the Royal Statistical Society: Series B (Methodological), 1966