

МЕТОДЫ ИССЛЕДОВАНИЯ ЛАНДШАФТА ФУНКЦИИ ПОТЕРЬ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ

Находнов Максим Сергеевич

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: nakhodnov17@gmail.com

Научный руководитель — Кропотов Дмитрий Александрович

Информация о топологии поверхности функции потерь оказывается полезной для анализа архитектур нейронных сетей. Исследование ландшафта лосс функции позволило понять влияние батч нормализации на качество и обобщающую способность в глубоких нейронных сетях [2]. Анализ и визуализация статистик гессiana показывает зависимость между архитектурными особенностями, такими как соединения быстрого доступа, гладкостью поверхности в окрестности точек минимума и обобщающей способностью [4].

В данной работе рассмотрены подходы для анализа геометрии нейронной сети как функции от входа и как функции от своих параметров. Также было произведено сравнение способов оценки кривизны функции потерь и разработка регуляризаторов на их основе.

В качестве прокси-объекта для изучения геометрии нейронной сети можно ввести понятие отступа объекта. Проведённые исследования показали, что распределение отступов на обучающей выборке имеет сильную взаимосвязь с зазором генерализации (GG), то есть разницей между качеством на обучении и качеством на тестовой выборке. Линейная модель, обученная на 377 примерах показывает коэффициент детерминации $\hat{R}^2 = 0.89$. Однако, статистики распределения отступов коррелируют с генерализацией только на последних итерациях обучения. Более того, модель предсказаний GG не обобщается на модели, которые использовали меньшее число итераций для обучения.

Изучение лосс функции как функции от своих параметров позволяет выявить характеристики имеющие высокую корреляцию с генерализацией. Показано, что оценка кривизны в окрестности точки в пространстве весов коррелирует с обобщающей способностью — более плоские и гладкие минимумы означают лучшую генерализацию [3]. Были рассмотрены несколько способов учитывать кривизну в процессе обучения. Во-первых, было исследовано использование локальных метрик кривизны, таких как след гессiana и норма градиента функции потерь в качестве регуляризаторов. Во-вторых,

был рассмотрен метод SAM [1], который неявно минимизирует ширину функции потерь в малой окрестности текущей точки. В-третьих, рассматривалась модификация SAM, которая использует риманову оптимизацию внутри данного алгоритма.

Было показано, что методы основанные на локальных оценках кривизны проигрывают SAM по качеству генерализации, несмотря на то, что они позволяют существенно уменьшить кривизну функции потерь в окрестности оптимума. Использование римановой оптимизации в алгоритме SAM позволило незначительно превзойти базовый алгоритм по качеству на тестовой выборке.

Литература

1. Foret Pierre. Sharpness-Aware Minimization for Efficiently Improving Generalization. — 2020.
2. Shibani Santurkar. How Does Batch Normalization Help Optimization? — 2019.
3. Shuofeng Zhang. Why Flatness Correlates With Generalization For Deep Neural Networks. — 2019.
4. Zhewei Yao. PyHessian: Neural Networks Through the Lens of the Hessian. — 2020.