Обзор на статью (Bernstein and Sheldon, 2019) "Differentially Private Bayesian Linear Regression"

Автор: Власенков Кирилл

Аннотация

В данной статье предлагается взглянуть на классическую модель байесовской регрессии, но с точки зрения дифференциальной приватности. Это связано с тем, что регрессия по своей природе - очень хорошо интерпретируемая модель, поэтому часто ее используют, например, в экономическом или банковском секторах. Появляется тогда проблема: может ли при желании злоумышленник вынести какую-то приватную информацию из данных, посчитав, например, апостериорное распределение? Оказывается, что при базовом подходе такой риск есть, поэтому и предлагается некоторая обновленная версия нестареющей классики.

Содержание

Мне данная статья понравилась. Хотя дифференциальная приватность является весьма популярной техникой в прикладной математике, хорошо, что авторы статьи все же уделили базовым понятиям отдельный параграф. Математическая основа получения совместного распределения - методы Монте-Карло для марковских цепей (MCMC), а в частности - сэмплирование Гиббса, что является классическим методом рассчета совместного распределния итерационно, высчитывая условные распределения. Распределение достаточной статистики при условии параметров раскладывается в произведение нормального распределения и Лапласовского, поэтому вычислительно требуется вычислить обратную матрицу, а это в случае тензора порядка 2 - O(n^3), а в случае порядка 3 - O(n^6). Так что вычислительно данные алгоритмы очень медленные и имеют применение только к небольшим датасетам (батчам). Тем не менее, задача авторов была не в реализации эффективного алгоритма, а в имплементации идеи дифф. приватности, поэтому минусом это не является. Более того, были выбраны довольно удачные и логичные метрики качества для полученного алгоритма, результатами которых авторы очень наглядно поделились на графиках. На них видно, что приватный метод имеет схожесть с неприватным.

Вывод

Могу рекомендовать к прочтению данную статью!