

Методы оптимизации. Лабораторная работа 2

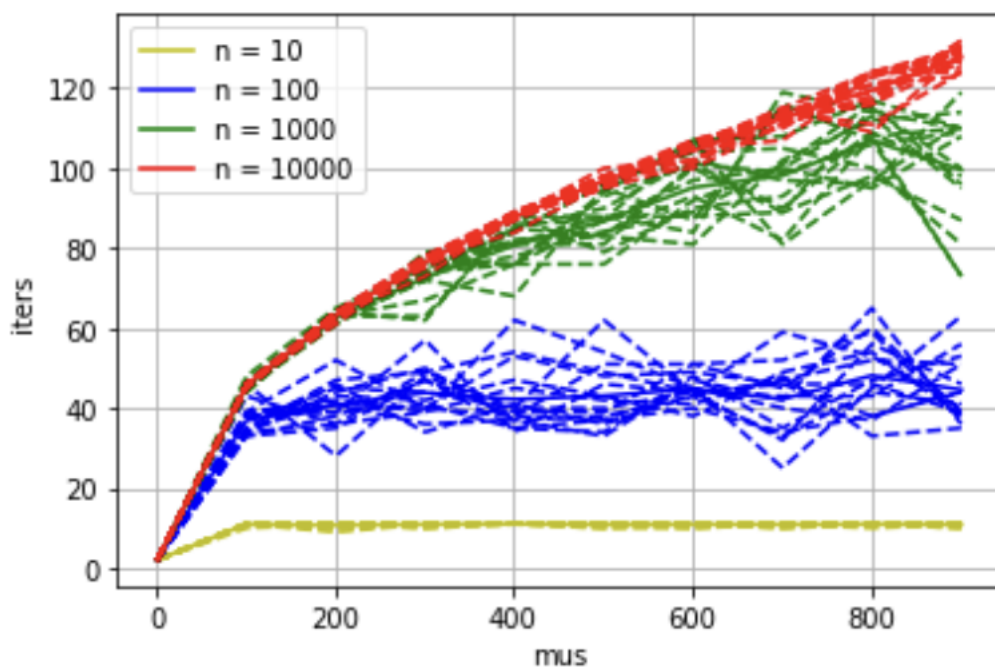
Daniil Bakushkin

4 июня 2023 г.

Зависимость числа итераций метода сопряженных градиентов от числа обусловленности и размерности пространства

Как и в методе градиентного спуска кол-во итераций сопряженных градиентов растет с ростом числа обусловленности, но в отличие от прошлой лабы, где рост был линейным, сейчас рост пропорционален корню.

Исходя из теории кол-во итераций $\propto \sqrt{\kappa}$ размерность пр-ва, что мы и видим на графике. Это связано с быстрой сходимостью алгоритма.



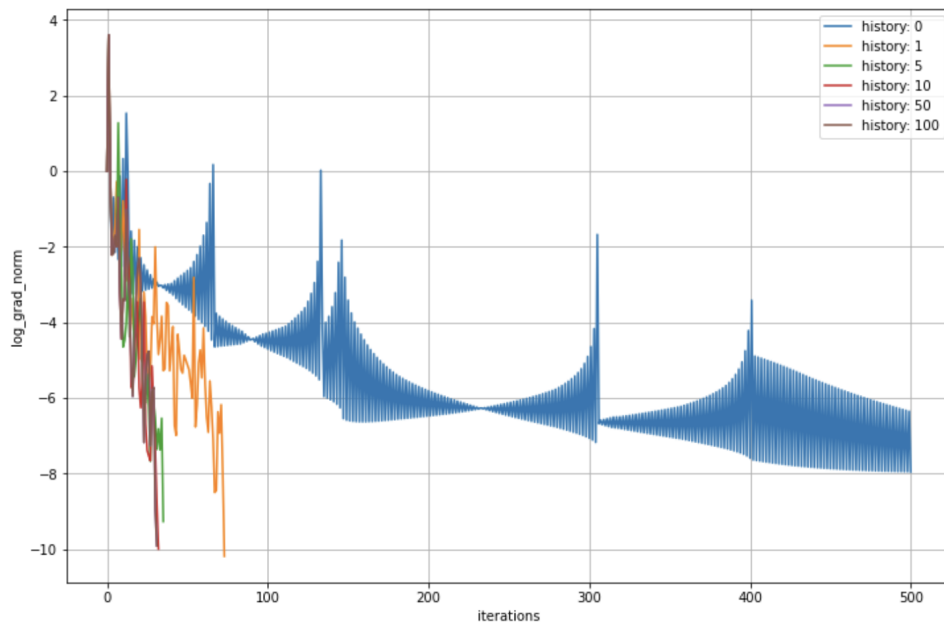
Зависимость числа итераций метода сопряженных градиентов от числа обусловленности и размерности пространства.

Выбор размера истории в методе L-BFGS

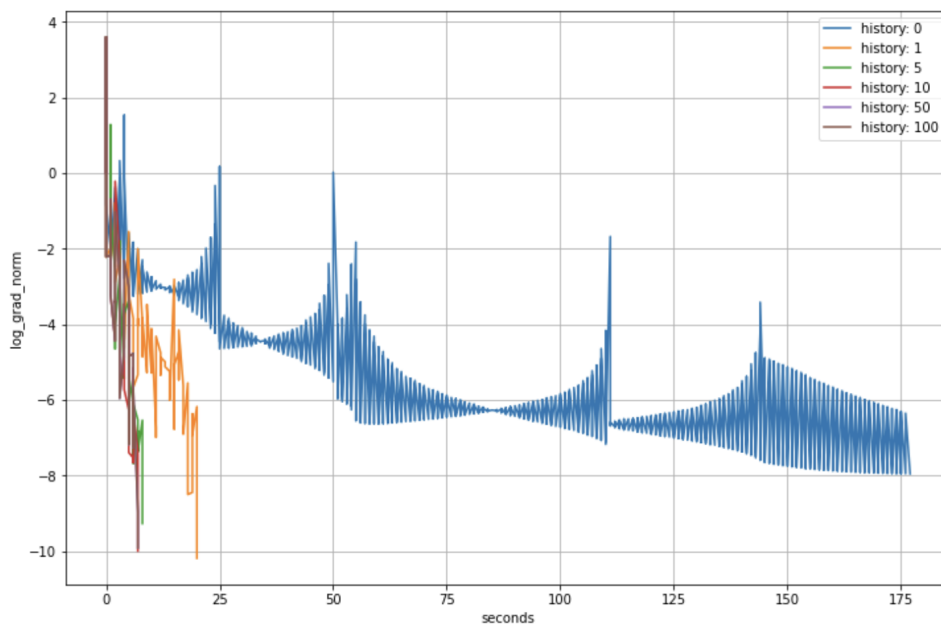
Т.к. реализация без истории аналогична градиентному спуску, то она и работает медленнее всех.

Для размера 0, 1 еще есть разница, а все остальное \pm в пределах погрешности.

Из наблюдений оптимальная история ± 10 , большее увеличение не приводит к росту скорости сходимости.



Норма градиента относительно итераций

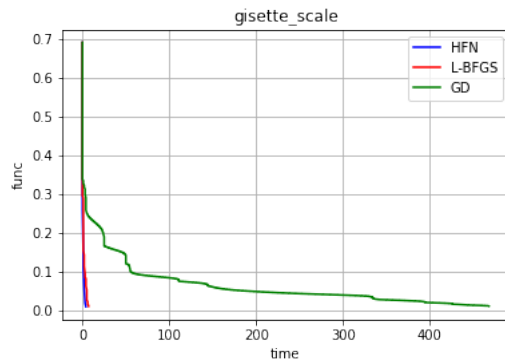


Норма градиента относительно времени

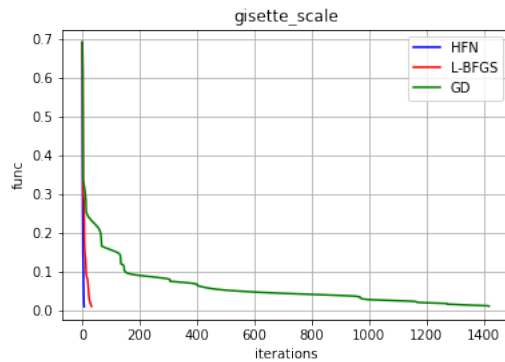
Сравнение методов на реальной задаче логистической регрессии

Вне зависимости от датасета, результат алгоритмов по норме невязки к итерациям \pm одинаковые. Градиентный спуск сходится гораздо медленнее, а L-BFGS и Hessian Free Newton примерно одинаково.

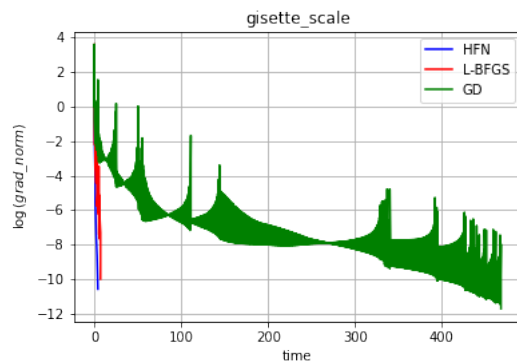
Результаты на датасете *gisettescale*:



Значение оптимизируемой функции от времени

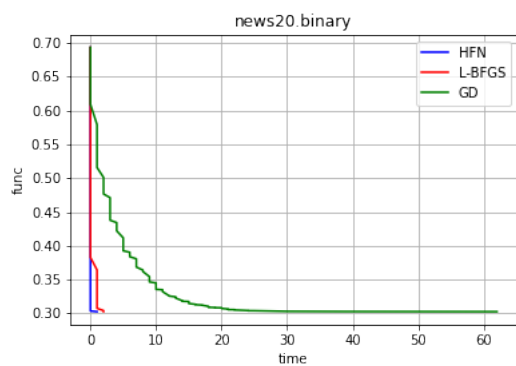


Значение оптимизируемой функции от итераций

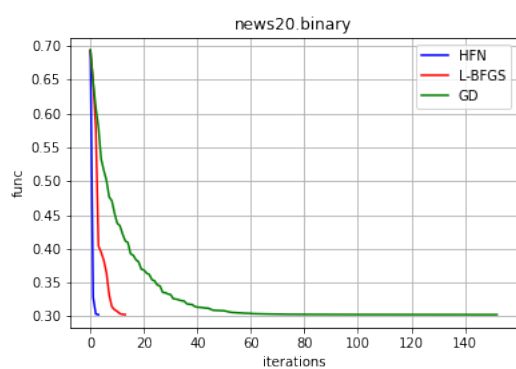


Значение нормы градиента от времени

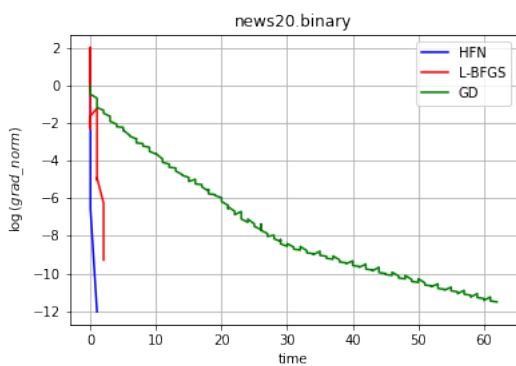
Результаты на датасете *news20*:



Значение оптимизируемой функции от времени

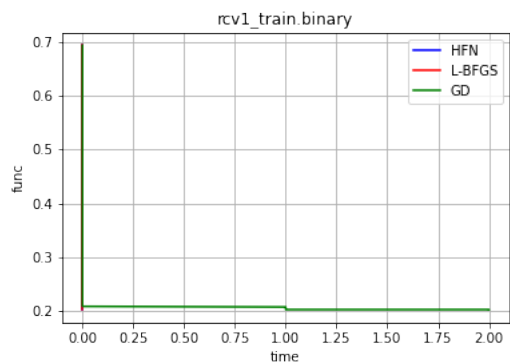


Значение оптимизируемой функции от итераций

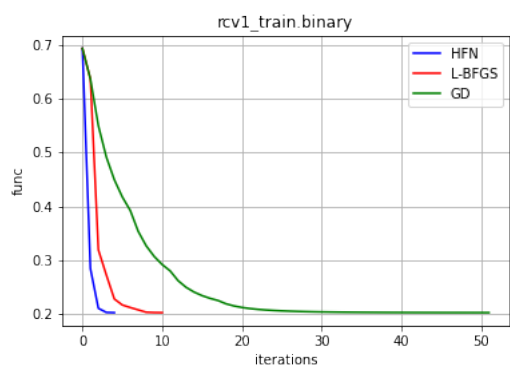


Значение нормы градиента от времени

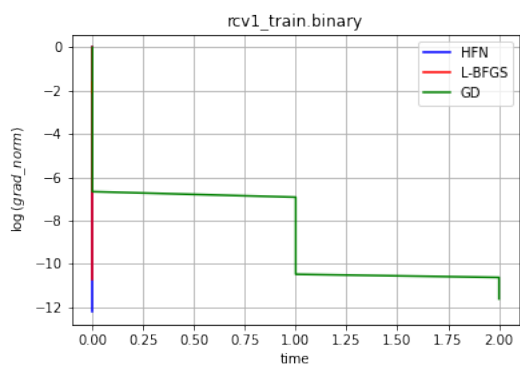
Результаты на датасете *rcv1*:



Значение оптимизируемой функции от времени

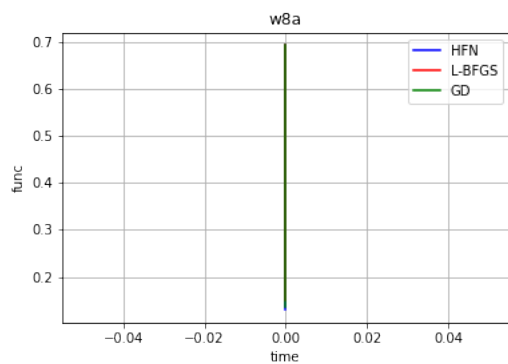


Значение оптимизируемой функции от итераций

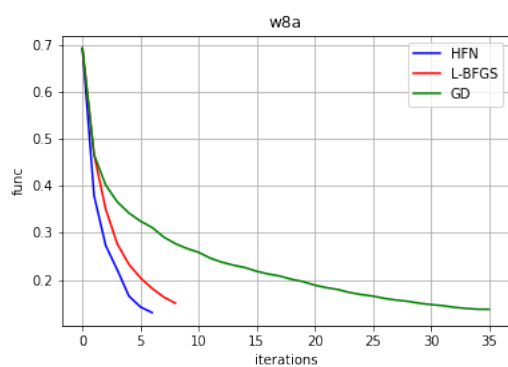


Значение нормы градиента от времени

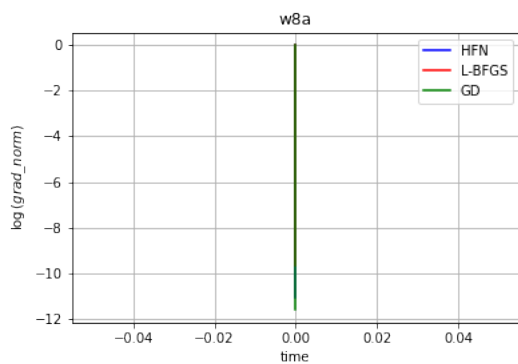
Результаты на датасете *w8a*:



Значение оптимизируемой функции от времени

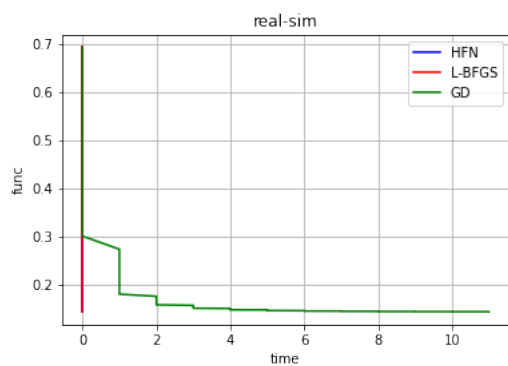


Значение оптимизируемой функции от итераций

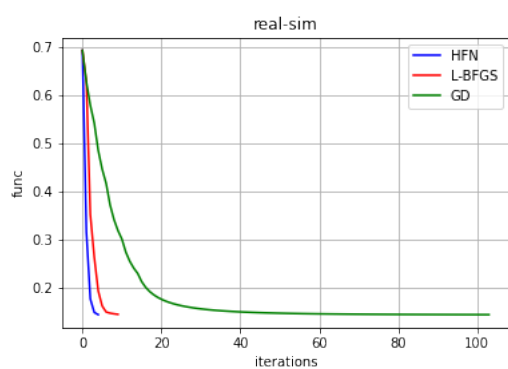


Значение нормы градиента от времени

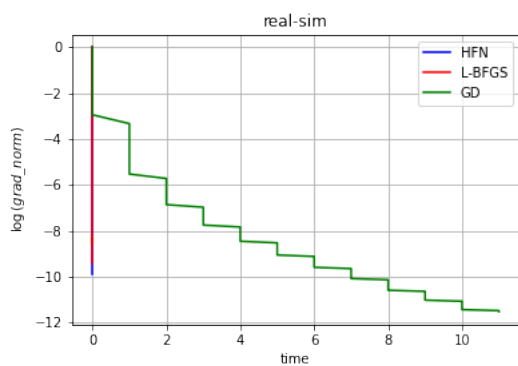
Результаты на датасете *real - sim*:



Значение оптимизируемой функции от времени



Значение оптимизируемой функции от итераций



Значение нормы градиента от времени