

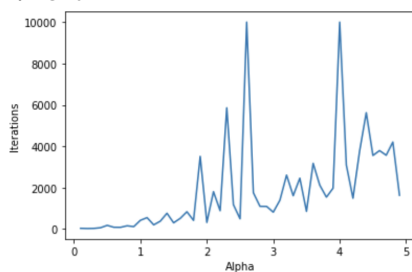
Методы оптимизации. Лабораторная работа 3

Daniil Bakushkin

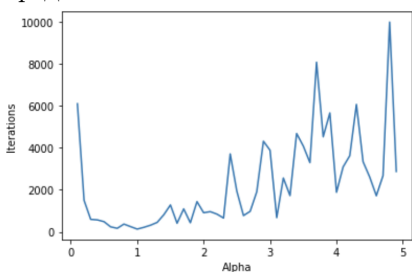
25 июня 2023 г.

Эксперимент 1: Выбор длины шага в субградиентном методе

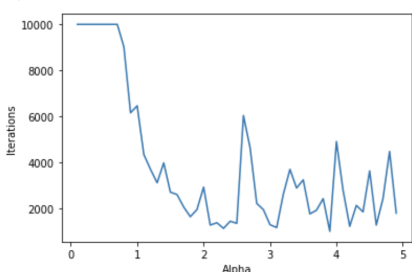
Близко



Средне



Далеко

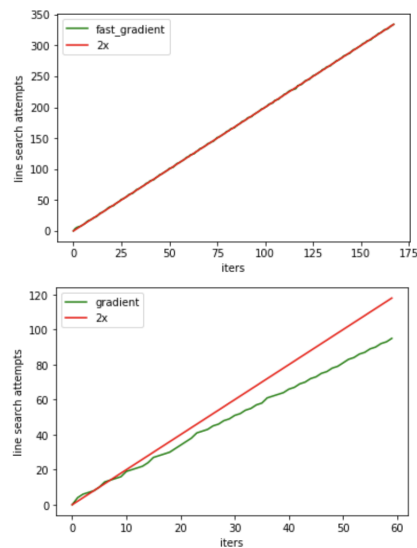


Как можно видеть, маленький размер шага выгоден для небольших расстояний, для больших расстояний обычно выбирается шаг больше.

Эксперимент 2: Среднее число итераций одномерного поиска в градиентных методах

Построим для градиентного и быстрого градиентного методов график зависимости суммарного числа итераций одномерного поиска от номера итерации метода.

Вывод: кол-во итераций линейного поиска практически идентично удвоенному кол-ву итераций, следовательно, кол-во итераций меняется всего примерно в 2 раза.



Эксперимент: Сравнение методов

Сравнение будет проводиться на задачах Lasso. Перебор по различным значениям размерности пр-ва n , размера выборки m , коэффициента регуляризации.

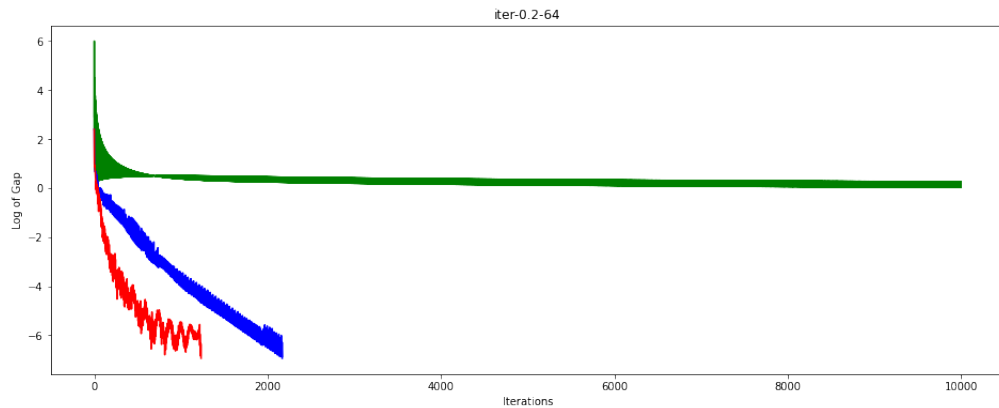
Ниже приведена сетка параметров, по которой совершался перебор:

$$n \in \{64, 1024\}$$

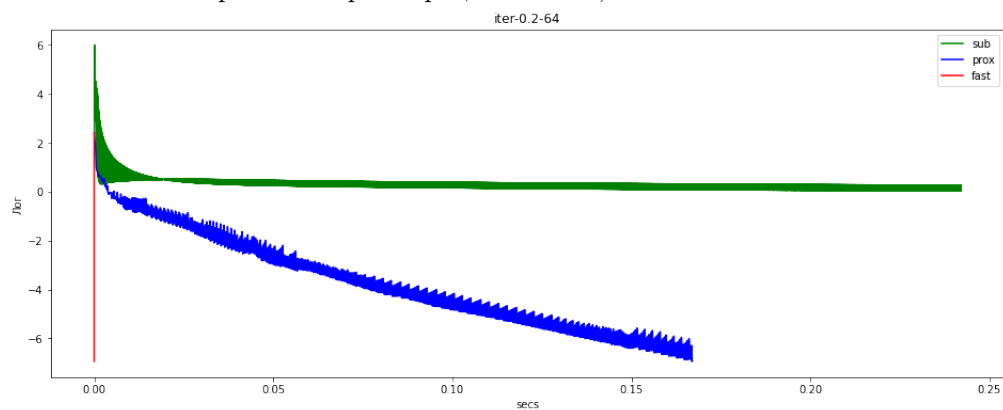
$$m \in \{64, 1024\}$$

Эксперименты были проведены только для $n == m$, чтобы не плодить большое число экспериментов.

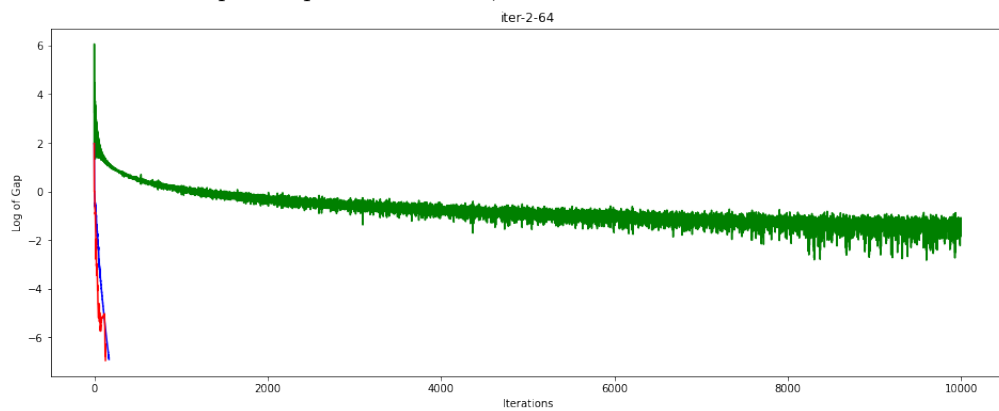
$$\lambda \in \{0.1, 1, 10\}$$



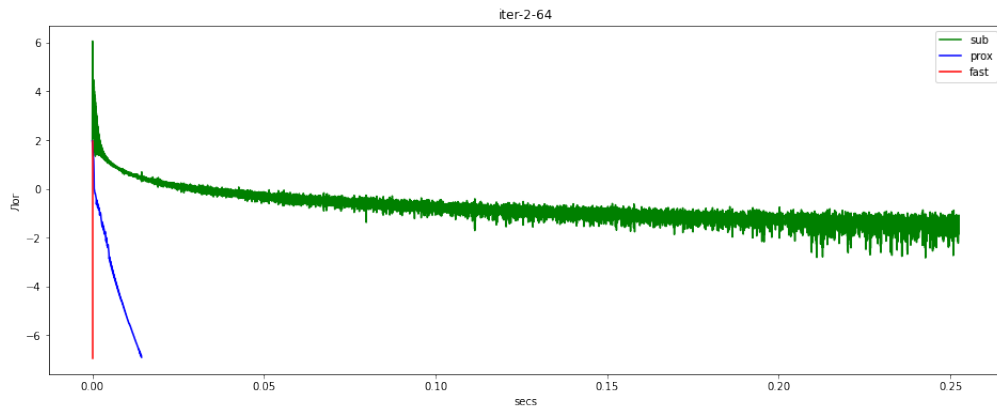
Зависимость зазора от номера итерации. $\lambda = 0.2, n = m = 64$



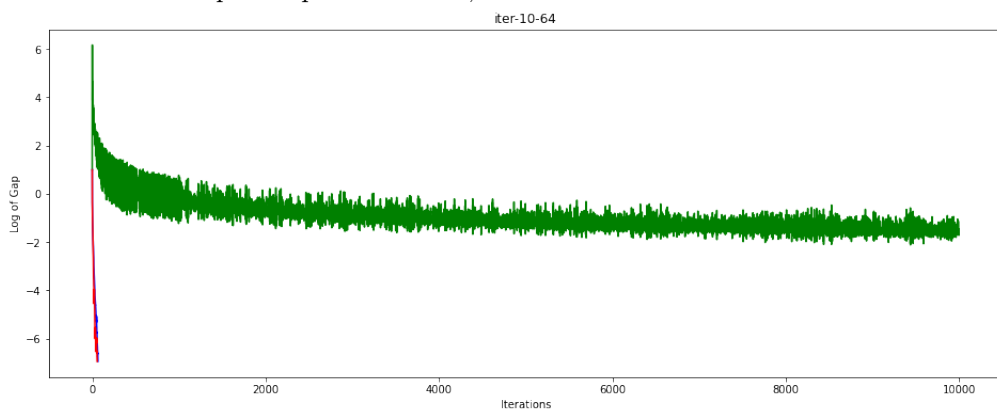
Зависимость зазора от времени. $\lambda = 0.2, n = m = 64$



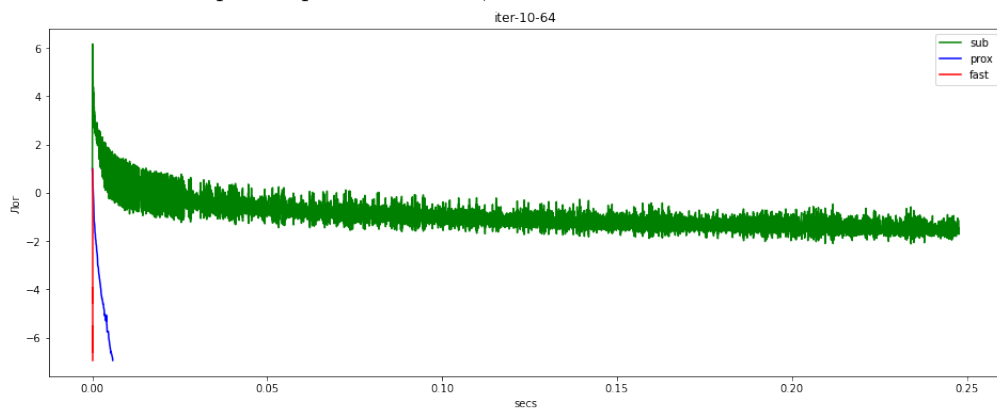
Зависимость зазора от номера итерации. $\lambda = 2, n = m = 64$



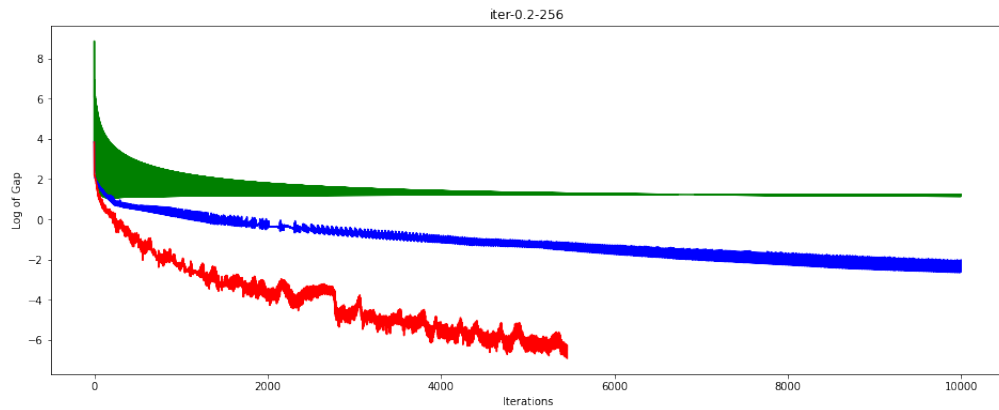
Зависимость зазора от времени. $\lambda = 2, n = m = 64$



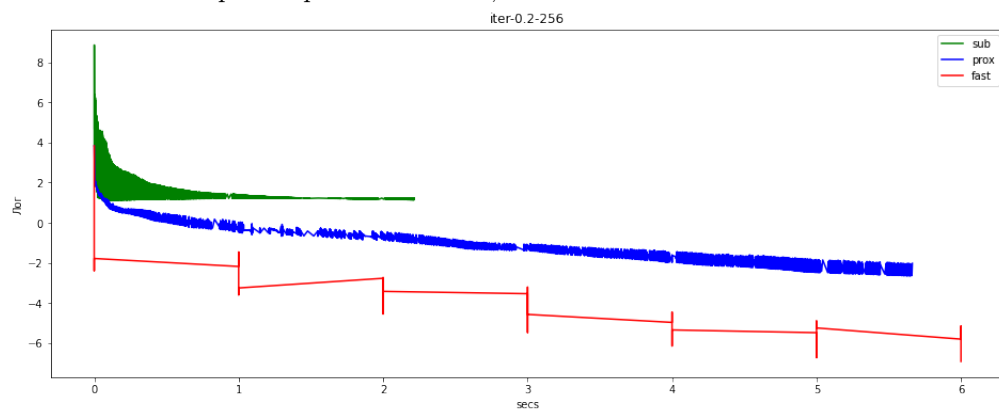
Зависимость зазора от времени. $\lambda = 10, n = m = 64$



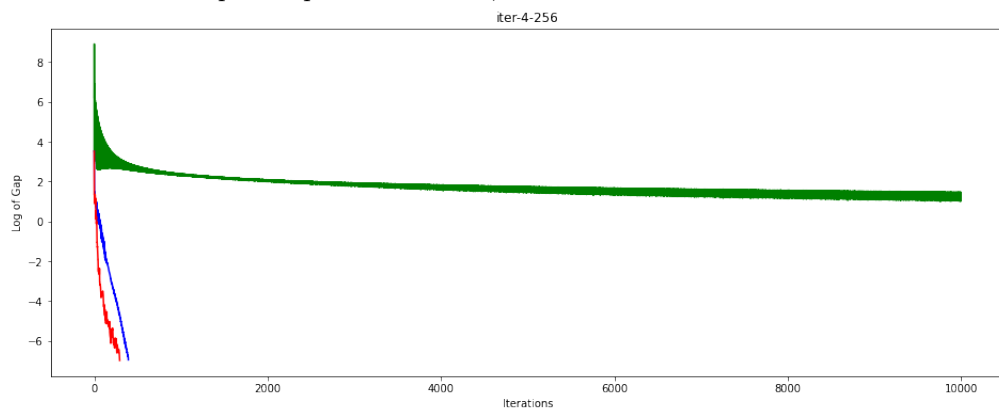
Зависимость зазора от времени. $\lambda = 10, n = m = 64$



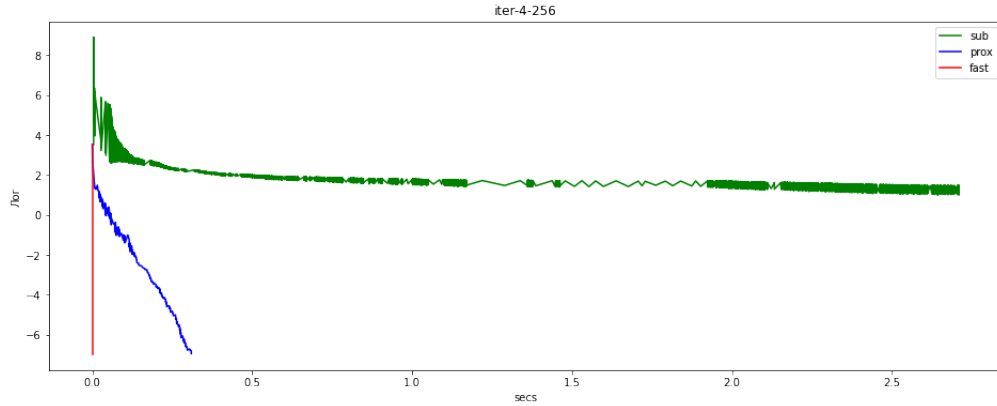
Зависимость зазора от времени. $\lambda = 0.1$, $n = m = 256$



Зависимость зазора от времени. $\lambda = 0.1$, $n = m = 256$



Зависимость зазора от времени. $\lambda = 4$, $n = m = 256$



Зависимость зазора от времени. $\lambda = 10$, $n = m = 256$

Тесты проводились на синтетических данных.

Вывод: proximal fast градиентный спуск всегда выигрывает вне зависимости от размерности пр-ва / размерности выборки / коэффициента регуляризации. Можно отметить, что регуляризация сильно влияет на скорость сходимости при маленьких значениях 0.2, но выше 4 – 10 разница уже незаметна.