



Московский государственный университет имени М. В. Ломоносова  
Факультет Вычислительной математики и кибернетики  
Кафедра Математических методов прогнозирования

Рецензия на курсовую работу

**First order implicit optimization methods for linear  
models in machine learning**

студента 317 группы  
Чабоненко Владислава  
Дмитриевича

**Рецензент:**  
студент 317 группы  
Кудрявцев Георгий Алексеевич

Москва, 2015

## Актуальность

Работа посвящена неявным методам оптимизации для линейных моделей машинного обучения. На сегодняшний день методы оптимизации являются основой для большинства методов машинного обучения.

## Краткое содержание

Сначала автор описывает несколько явных разновидностей градиентного спуска (FG, SG, SAG). Затем он поясняет, что такое неявный метод и его отличие от явного, на примере решения дифференциального уравнения. После этого автор показывает вывод формул для неявных методов градиентного спуска (IFG, ISG, ISAG). В конце работы проводятся эксперименты, которые показывают разницу работы между явными и неявными версиями градиентного спуска и BFGS, а также обсуждается выбор шага для ISG и IFG.

Автор смог реализовать неявные методы IFG, ISAG и ISG для линейной и логистической регрессии. Также автор выяснил, что неявность не изменяет качество работы SAG, но улучшает FG.

## Оценка работы

В целом работа является хорошей.

Последовательное изложение материала. Нет ничего лишнего, и все основные моменты прояснены. Культура речи оценена быть не может, так как рецензент не обладает достаточными знаниями английского языка, но переводить было не сложно. Эксперименты логичны и хорошо показывают интересующие автора характеристики.

Выводы автора дают полезную информацию в исследовании вопроса использования неявных методов оптимизации.

## Рекомендации автору

В подглаве Stochastic Gradient Descent в предпоследнем параметре автор сравнивает сходимости FG и SG. Общий смысл понимается с трудом с третьего перечитывания. Советую написать более понятно и последовательно.

Когда автор объясняет, что такое неявный метод, автор говорит, что  $y_{k+1}$  нельзя неявно выразить. Но на самом деле переменная явно выражается. Следует или поменять пример, или не говорить об этом.

Также встречалась пара опечаток. Автору имеет смысл проверить правильность написания слов еще раз.

В экспериментах ничего не сказано про то, как выбиралось начальное приближение. Хотелось бы пояснение автора на этот счет.