## Вопросы к экзамену

## по спецкурсу «Методы оптимизации в машинном обучении», осень 2015

1. Методы одномерной оптимизации без производной: метод золотого сечения, парабол и Брента. Поиск ограничивающего сегмента (Текст + [6], разделы 10.1, 10.2 + [7], раздел 5)
2. Методы одномерной оптимизации с производной: метод деления отрезка пополам, секущей и Брента (Текст + [6], разделы 9.1, 9.2 + [7], раздел 4)
3. Метод градиентного спуска. Различные стратегии выбора длины шага. Понятие о скорости сходимости метода (Текст + [1], разделы 2.2, 3.2, 3.3)
4. Метод Ньютона, его скорость сходимости. Недостатки метода Ньютона и пути их преодоления. (Текст + [1], разделы 3.3, 3.4)
5. Метод сопряженных градиентов для решения СЛАУ, предобуславливание (Текст + [1], раздел 5.1)
6. Метод сопряжённых градиентов для произвольной функции. Схемы Флетчера-Ривса, Полака-Рибье, стратегия рестарта (Текст + [1], раздел 5.2)
7. Неточный и безгессианный метод Ньютона. Способы оценивания произведения гессиана на вектор ([1], раздел 7.1 + [3])
8. Квази-ньютоновские методы оптимизации, методы SR1 и BFGS, метод L-BFGS (Текст + [1], разделы 6.1, 6.2, 7.2)
9. Метод Ньютона для выпуклых задач оптимизации с ограничениями вида равенства. Прямо-двойственный вариант метода (Текст + [2], глава 10)
10. Метод логарифмических барьеров. Прямо-двойственный метод внутренней точки. Методы первой фазы (Текст + [2], глава 11)
11. Разреженные линейные модели для задач регрессии/классификации. Проксимальный метод оптимизации (Текст + [5], раздел 2.3)
12. Стохастический градиентный спуск, его скорость сходимости. Метод SAG ([8] + [5], глава 13 + [4], раздел 8.2)
13. Суррогатная оптимизация (оптимизация с использованием глобальных верхних оценок). Способы получения оценок. Примеры оценок. Пример применения метода для задачи LASSO (Текст + [10])
14. Метод Левенберга-Марквардта для задачи обучения нелинейной регрессии (Текст + [9])
15. Автоматическое дифференцирование и его применение в задаче обучения глубинного автокодировщика ([1], раздел 8.2 + [9]).

**Литература:**


1. J. Nocedal, S.J. Wright. Numerical Optimization. Springer, 2006.
2. S. Boyd, L. Vandenberghe. Convex Optimization, Cambridge University Press, 2004.
3. M. Schmidt. Limited-Memory Quasi-Newton and Hessian-Free Newton Methods for Non-Smooth Optimization // NIPS workshop on optimization for machine learning, 2010.
4. D. Bertsekas. Convex Analysis and Optimization, Athena Scientific, 2003.
5. Optimization for Machine Learning. Edited by Suvrit Sra, Sebastian Nowozin and Stephen J. Wright, MIT Press, 2011.
6. Numerical Recipes. The Art of Scientific Computing, Cambridge University Press, 2007.
7. R. Brent. Algorithms for minimization without derivatives. Prentice-Hall, 1973.
8. M. Schmidt. Notes on Big-n Problems, 2012.
9. J. Martens. Deep Learning via Hessian-free Optimization // ICML, 2010.
10. J. Mairal. Optimization with First-Order Surrogate Functions // Arxiv, 1305.3120, 2013.

## Теоретический минимум

Ниже перечислены вопросы, незнание ответа на которые во время экзамена автоматически влечет неудовлетворительную итоговую оценку.

1. Определение градиента, субградиента и гессиана функции многих переменных.
2. Необходимые и достаточные условия оптимальности в задачах безусловной и условной оптимизации.
3. Определение функции с липшицевым градиентом, выпуклой и сильно выпуклой функции.
4. Условия Армихо и Вольфа для неточной одномерной оптимизации. Процедура backtracking.
5. Схема методов покоординатного и градиентного спуска, Ньютона. Что такое неточный метод Ньютона?
6. Примеры задач быстрой и медленной работы методов покоординатного и градиентного спуска, Ньютона.
7. Схема метода сопряжённых градиентов. Что такое предобуславливание?
8. Двойственная задача условной оптимизации, её свойства.
9. Модели разреженной линейной/логистической регрессии, метода опорных векторов, байесовские варианты моделей.
10. Определение сублинейной, линейной и квадратичной скорости сходимости, оценки на количество требуемых итераций.
11. Общая схема BFGS и L-BFGS.
12. Общая схема метода логарифмических барьеров.
13. Общая схема прямо-двойственного метода внутренней точки.