Вопросы к экзамену

по спецкурсу «Методы оптимизации в машинном обучении», 2014

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. Неточная одномерная оптимизация, условия Армихо, Голдштайна, Вольфа. Процедура backtracking и алгоритм минимизации для условий Вольфа (Текст + [1], разделы 3.1, 3.5 + [8])
4. Методы покоординатного и градиентного спуска. Различные стратегии выбора длины шага. Понятие о скорости сходимости методов (Текст + [1], разделы 2.2, 3.2, 3.3)
5. Метод Ньютона, его скорость сходимости. Недостатки метода Ньютона и пути их преодоления. Метод Левенберга-Марквардта для задачи обучения нелинейной регрессии (Текст + [1], разделы 3.3, 3.4, 10.3)
6. Метод сопряженных градиентов для решения СЛАУ, предобуславливание. Схемы Флетчера-Ривса и Полака-Рибье (Текст + [1], глава 5)
7. Неточный и безгессианный метод Ньютона. Способы оценивания произведения гессиана на вектор ([1], раздел 7.1 + [3])
8. Квази-ньютоновские методы оптимизации, метод L-BFGS (Текст + [1], разделы 6.1, 7.2)
9. Методы оптимизации с использованием глобальных верхних оценок. Способы получения верхних оценок. Примеры оценок. Пример применения метода для задачи LASSO (Текст)
10. Метод Ньютона для выпуклых задач оптимизации с ограничениями вида равенства. Прямо-двойственный вариант метода (Текст + [2], глава 10)
11. Метод логарифмических барьеров. Прямо-двойственный метод внутренней точки (Текст + [2], глава 11)
12. Применение прямо-двойственного метода внутренней точки для задачи обучения метода опорных векторов (Текст + [5], глава 12)
13. Разреженные линейные модели для задач регрессии/классификации. Проксимальный метод оптимизации (Текст + [5], раздел 2.3)
14. Разреженные линейные модели для задач регрессии/классификации. Метод покоординатного спуска (Текст + [5], раздел 2.4)
15. Стохастический градиентный спуск, его скорость сходимости. Метод SAG ([9] + [5], глава 13 + [4], раздел 8.2)

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. J. J. Mor´e, D. J. Thuente, Line search algorithms with guaranteed sufficient decrease // ACM Transactions on Mathematical Software, 20 (1994), pp. 286–307.
9. M. Schmidt. Notes on Big-n Problems, 2012.

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. Основные стратегии неточной одномерной оптимизации.
4. Схема методов покоординатного и градиентного спуска, Ньютона. Что такое неточный метод Ньютона?
5. Примеры задач быстрой и медленной работы методов покоординатного и градиентного спуска, Ньютона.
6. Схема метода сопряжённых градиентов. Что такое предобуславливание?
7. Двойственная задача условной оптимизации, её свойства.
8. Модели разреженной линейной/логистической регрессии, метода опорных векторов, байесовские варианты моделей.
9. Определение сублинейной, линейной и квадратичной скорости сходимости, оценки на количество требуемых итераций.
10. Общая схема BFGS и L-BFGS.
11. Общая схема метода логарифмических барьеров.
12. Общая схема прямо-двойственного метода внутренней точки.