

Вопросы к коллоквиуму по курсу «Методы оптимизации», ФКН ВШЭ, 2017

1. Методы одномерной оптимизации без производной: методы золотого сечения, парабол и Брента. ([7], разделы 10.1, 10.2 + [8], раздел 5)
2. Метод градиентного спуска. Различные стратегии выбора длины шага. Скорость сходимости метода для сильно-выпуклых функций (с доказательством). ([1], разделы 2.2, 3.2, 3.3)
3. Метод Ньютона, его скорость сходимости (с доказательством). Модификации метода Ньютона для невыпуклых задач оптимизации. ([1], разделы 3.3, 3.4)
4. Метод сопряженных градиентов для решения системы линейных уравнений. Предобуславливание. ([1], раздел 5.1)
5. Неточный метод Ньютона (HFN), его скорость сходимости. Способы оценивания произведения гессиана на вектор. ([1], раздел 7.1 + [6])
6. Квазиньютоновские методы оптимизации. Метод L-BFGS. ([1], разделы 6.1, 6.2, 7.2)
7. Выпуклые множества и функции. Основные примеры. Операции, сохраняющие выпуклость. Простейшие свойства выпуклых функций. ([3], параграф 3.1; [2] разделы 2, 3, 4; [4], глава 5)
8. Субградиент и субдифференциал. Субдифференциальное исчисление. Условия оптимальности. ([3], разделы 3.1.5-3.1.6)
9. Задачи условной оптимизации. Условия Каруша-Куна-Таккера. Двойственность. Стандартные классы выпуклых задач: линейное программирование (LP), квадратичное программирование (QP, QCQP). ([4], глава 9; [2], разделы 4 и 5)
10. Метод Ньютона для минимизации на аффинном подпространстве. ([2], глава 10)
11. Метод барьеров для выпуклых задач условной оптимизации. ([2], глава 10)
12. Субградиентный метод, его скорость сходимости (с доказательством). ([3], раздел 3.2.3; [4], глава 5)
13. Метод проекции градиента и проксимальный градиентный метод. Процедура подбора длины шага. ([5], раздел 2.3)
14. Стохастический градиентный метод, его скорость сходимости (с доказательством). ([11], разделы 14.0, 14.1)
15. Методы SAG и SVRG для минимизации суммы функций. ([9, 10])

Литература

1. J. Nocedal, S.J. Wright. [Numerical Optimization](#). Springer, 2006.
2. S. Boyd, L. Vandenberghe. [Convex Optimization](#). Cambridge University Press, 2004.
3. Ю.Е. Нестеров [Методы выпуклой оптимизации](#). М.: МЦНМО. 2010.
4. Б.Т. Поляк. [Введение в оптимизацию](#). Наука, 1983.
5. [Optimization for Machine Learning](#). Edited by Suvrit Sra, Sebastian Nowozin and Stephen J. Wright. MIT Press, 2011.
6. M. Schmidt. Limited-Memory Quasi-Newton and Hessian-Free Newton Methods for NonSmooth Optimization. NIPS workshop on optimization for machine learning, 2010.

7. [Numerical Recipes. The Art of Scientific Computing](#). Cambridge University Press, 2007.
8. R. Brent. Algorithms for minimization without derivatives. Prentice-Hall, 1973.
9. M. Schmidt et al.. [Minimizing Finite Sums with the Stochastic Average Gradient](#). arXiv, 2015.
10. R. Johnson, T. Zhang. [Accelerating Stochastic Gradient Descent using Predictive Variance Reduction](#). NIPS, 2013.
11. A. Nemirovski. [Information-based complexity of convex programming](#). 1994

Теоретический минимум

Ниже перечислены вопросы, незнание ответа на которые (без подготовки) автоматически влечет ноль за коллоквиум.

1. Общее определение производной функции (через линейное отображение). Основные свойства производной. Определения второй производной, градиента и гессиана.
2. Определения функции с липшицевым градиентом, выпуклой и сильно выпуклой функции. Соответствующие глобальные верхние и нижние оценки на функцию.
3. Определение сублинейной, линейной, сверхлинейной и квадратичной скоростей сходимости. Оценки на количество требуемых итераций для достижения заданной точности.
4. Определения выпуклого множества и выпуклой функции.
5. Определение субградиента и субдифференциала выпуклой функции.
6. Необходимые и достаточные условия оптимальности в задачах безусловной минимизации.
7. Условия Армихо и Вульфа. Процедура бэктрекинга.
8. Схема методов градиентного спуска и Ньютона.
9. Основные матричные разложения: спектральное разложение, сингулярное разложение, LU разложение, QR разложение, разложение Холецкого, LDL разложение.
10. Примеры быстрой и медленной работы методов градиентного спуска и Ньютона.
11. Схема метода сопряженных градиентов.
12. Формулировка теоремы Каруша-Куна-Таккера для общей задачи нелинейной оптимизации.
13. Двойственная задача и ее основные свойства.
14. Проксимальное отображение. Связь проксимального отображения с евклидовой проекцией.
15. Общая схема квазиньютоновских методов.
16. Отличия метода L-BFGS от BFGS.
17. Общая схема метода барьеров.
18. Схема субградиентного метода.
19. Схема градиентного метода для композитной минимизации (проксимальный градиентный метод).
20. Схема стохастического градиентного метода.