Методы оптимизации
«Программа коллоквиума»

Коллоквиум проводится во время официального семинара на неделе предшествующей зачетной (15-ой) и на зачетной неделе: часть группы сдает в один день, другая часть – через неделю. Коллоквиум принимает преподаватель семинарских занятий, а также приглашенные преподаватели. Коллоквиум проводится в формате экзамена: ответ на билет, 1-2 вопроса по программе и 1-2 вопроса на подумать. Время на подготовку билета 40 минут. Программа разбита на 3 части: 1 балл (удовлетворительно), 2 балла (хорошо), 3 балла (отлично). В билетах буду присутствовать вопросы из всех трех частей, студент вправе делать только ту часть, которая соответствует желаемой оценке, но никто не запрещает и отвечать на более продвинутые вопрос. Студент может озвучить преподавателю оценку, на которую рассчитывает, исходя из этого преподаватель задаст более релевантные дополнительные вопросы.

Возможные оценки за коллоквиум: 0 (неудовлетворительно), 0.5 (неудовлетворительно +), 1 (удовлетворительно), 1.5 (удовлетворительно +), 2 (хорошо), 2.5 (хорошо +), 3 (отлично), 3.5 (отлично +).

На подготовку вопроса дается 40 минут.

1. Определение выпуклого множества. Определение выпуклой функции (2): для непрерывно дифференцируемой функции и произвольной функции. Критерий сильной выпуклости дважды непрерывно дифференцируемой функции. Определение сопряженной функции (двойственность по Фенхелю). Лагранжиан. Двойственность по Лагранжу. Двойственная задача. Условия Слейтера. ККТ. Определение классов задач: от LP до SDP. Определение субградиента и субдифференциала.

2. Формулировка условия локального минимума на $\mathbb{R}^d$ для произвольной непрерывно дифференцируемой функции. Формулировка условия глобального минимума на $\mathbb{R}^d$ для выпуклой непрерывно дифференцируемой функции. Формулировка условия глобального минимума на выпуклом множестве $\mathcal{X}$ для выпуклой непрерывно дифференцируемой функции. Доказательство условия локального минимума на $\mathbb{R}^d$. Доказательство условия глобального минимума на $\mathbb{R}^d$ выпуклой функции. Доказательство условия глобального минимума на выпуклом множестве $\mathcal{X}$ выпуклой функции. Доказательство свойства гладкой непрерывно дифференцируемой функции. Доказательство свойства ограниченности субградиента выпуклой Липшицевой функции.

3. Итерация метода градиентного спуска. Интуиция: почему такой метод, зачем нужен параметр (шаг). Характер сходимости (линейная / субли-

нейная / ... / локальная / глобальная) градиентного спуска для гладких сильно выпуклых задач. Формулировка оценки сходимости градиентного спуска для гладких сильно выпуклых задач. Доказательство сходимости градиентного спуска для гладких сильно выпуклых задач.

4. Итерации метода тяжелого шарика и ускоренного градиентного метода (метода Нестерова). Интуиция: почему может быть лучше, чем градиентный спуск, как подбирать моментнумный параметр. Характер сходимости для гладких сильно выпуклых задач. Особенности сходимости по сравнению с градиентным спуском. Формулировка оценки сходимости ускоренного градиентного метода для гладких сильно выпуклых задач. Нижние оценки сложности методов первого порядка для решения гладких сильно выпуклых задач. Доказательство нижних оценок сложности методов первого порядка для решения гладких сильно выпуклых задач..

5. Сопряженные направления. Интуиция метода сопряженных градиентов: как работает, чего хотим добиться, почему именно так строим метод. Характер сходимости для систем линейных уравнение с положительно определенной матрицей. Формулировка оценки сходимости метода сопряженных градиентов (2 результата): в зависимости от d (размерности задачи), в зависимости от κ (числа обусловленности матрицы A). Доказательство сходимости метода сопряженных градиентов для системы линейных уравнений с положительно определенной матрицей за d итераций.

6. Итерация метода Ньютона. Интуиция метода Ньютона: почему берем именно такую итерацию. Характер сходимости для сильно выпуклых задач с Липшицевым гессианом. Квазиньютоновское уравнение, интуиция. Формулировка оценки сходимости метода Ньютона для сильно выпуклых задач с Липшицевым гессианом. Способы получения глобальной сходимости для метода Ньютона. Правила обновления матриц H или B для SR1 и BFGS. Доказательство сходимости метода Ньютона для сильно выпуклых задач с Липшицевым гессианом.

7. Итерация метода субградиентного спуска. Интуиция метода. Характер сходимости. Адаптивные методы: AdaGradNorm, AdaGrad, RMSProp, Adam, AdamW. Интуиции методов. Композитная задача. Итерация проксимального метода. Интуиция метода. Характер сходимости. Формулировка оценки сходимости субградиентного спуска для выпуклых задач с Липшицевой функцией. Формулировка свойств проксимального метода. Доказательство сходимости для выпуклых задач с Липшицевой функцией.

Доказательство сходимости проксимального метода для выпуклых композитных задач с гладким и проксимально дружественным слагаемыми (вместе с доказательствами свойств проксимального оператора).

8. Евклидова проекция. Итерация метода градиентного спуска с проекцией. Интуиция метода. Характер сходимости для гладких сильно выпуклых задач. Итерация метода Франк-Вульфа. Интуиция метода. Характер сходимости для гладких выпуклых задач. Формулировка свойств оператора проекции. Формулировка оценки сходимости метода Франк-Вульфа. Доказательство сходимости метода с проекцией (вместе с доказательствами необходимых свойств проекции). Доказательство сходимости метода Франк-Вульфа для гладких выпуклых задач.

9. Итерация метода зеркального спуска. Интуиция метода. Характер сходимости для выпуклых гладких задач. Шаг зеркального спуска в случае симплекса и KL-дивергенции (с доказательством). Доказательство сходимости метода зеркального спуска для выпуклых гладких задач.

10. Седловая задача. Итерация метода экстраградиента. Интуиция метода: почему лучше, чем градиентный спуск–подъем. Характер сходимости для сильно выпуклых – сильно вогнутых гладких задач. Формулировка оценки сходимости метода экстраградиента для сильно выпуклых – сильно вогнутых гладких задач.

11. Барьерная функция. Метод внутренней точки. Свойства решения барьерной задачи. Итерация метода внутренней точки для самосогласованных барьеров. Доказательство теоремы о множестве решений барьерной задачи и ϵ-окрестности множества решений исходной задачи.

12. Штрафная функция. Метод штрафных функций. Постановка задачи и итерация метода ADMM. Формулировка свойств решения штрафной задачи. Доказательства свойств решения штрафной задачи. Доказательство теоремы о множестве решений штрафной задачи и ϵ-окрестности множества решений исходной задачи. Доказательство сходимости ADMM.

13. Различные постановки задачи стохастической оптимизации. Итерация метода SGD. Интуиция метода. Характер сходимости в условиях ограниченной дисперсии стохастического градиента. Итерация метода SAGA. Интуиция метода. Характер сходимости. Оценки сходимости SGD для сильно выпуклых гладких задач в условиях ограниченной дисперсии стохастического градиента. Оценки сходимости SAGA для сильно выпуклых гладких задач вида конечной суммы. Доказательство сходимости SGD.