

# Методы оптимизации

## «Программа экзамена»

Формат проведения 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 билеты будут 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 вопросы. Студент может озвучить преподавателю оценку, на которую рассчитывает, исходя из этого преподаватель задаст более релевантные дополнительные вопросы.

Возможные оценки за экзамен: 0 (неудовлетворительно), 1 (удовлетворительно), 2 (хорошо), 3 (отлично).

1. 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.
2. Итерация метода градиентного спуска. Интуиция: почему такой метод, зачем нужен параметр (шаг). Характер сходимости (линейная/сублинейная/..., локальная/глобальная) градиентного спуска для гладких сильно выпуклых задач. Формулировка оценки сходимости градиентного спуска для гладких сильно выпуклых задач, гладких выпуклых задач и гладких невыпуклых задач. Доказательство сходимости градиентного спуска для гладких сильно выпуклых задач, гладких выпуклых задач и гладких невыпуклых задач.
3. 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.

4. 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.
5. 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.
6. 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.
7. Итерация метода субградиентного спуска. Интуиция метода. Характер сходимости. Итерации методов AdaGradNorm, AdaGrad, RMSProp, Adam. Интуиция методов, их последовательное появление и связь с предшественниками. Характер сходимости. Формулировка оценки сходимости субградиентного спуска для выпуклых задач с Липшицевой функцией. Формулировка оценки сходимости AdaGrad для выпуклых задач с Липшицевой функцией. Доказательство сходимости для выпуклых задач с Липшицевой функцией. Доказательство сходимости AdaGrad для выпуклых задач с Липшицевой функцией.
8. 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. Итерация метода SVRG. Интуиция метода. Характер сходимости. Итерация метода SARAH. 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 для сильно выпуклых гладких задач в условиях ограниченной дисперсии стохастического градиента. Доказательство сходимости SAGA для сильно выпуклых гладких задач в условиях ограниченной дисперсии стохастического градиента.

9. Несмещенные и смещенные операторы сжатия. Примеры операторов сжатия. Итерация метода распределенного градиентного спуска с сжатием. Идея метода. Характер и особенности сходимости. Техника компенсации ошибки. Итерация градиентного спуска с сжатием и техникой компенсации ошибки. Характер и особенности сходимости. Техника памяти. Итерация метода EF21. Характер и особенности сходимости.

Теоретические особенности сходимости распределенного градиентного спуска с сжатием для смещенных и несмещенных операторов сжатия. Теоретические особенности сходимости метода EF21. Способ организации коммуникаций и общения между собой в случае отсутствия координирующего сервера: AllReduce и AllGather процедуры.

10. Итерация метода Local GD. Идея метода. Характер и особенности сходимости. Итерация методов FedProx и Scaffold. Идея методов. Характер сходимости. Сравнение с Local SGD.

Теоретические особенности сходимости метода Local SGD.

11. Интуиция метода центров тяжести. Интуиция метода центров эллипсоидов. Отделяющий оракул. Скорость сходимости метода центров тяжести по функционалу.
12. 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  $\kappa$  (числа обусловленности матрицы  $A$ ).
13. Метод множителей Лагранжа. Седловая задача. Итерация метода экстраградиента. Интуиция метода: почему лучше, чем градиентный спуск–подъем. Характер сходимости для сильно выпуклых – сильно вогнутых гладких задач.

14. Штрафная функция. Метод штрафных функций. **Формулировка свойств решения штрафной задачи.** Барьерная функция. Метод внутренней точки. **Формулировка свойств решения штрафной задачи.**