

Определения и формулировки

1. Положительно определённая матрица.
2. Евклидова норма вектора.
3. p -норма вектора.
4. Как выглядит единичный шар в p - норме на плоскости для $p = 1, 2, \infty$?
5. Норма Фробениуса для матрицы.
6. Спектральная норма матрицы.
7. Скалярное произведение двух векторов.
8. Скалярное произведение двух матриц, согласованное с нормой Фробениуса.
9. Собственные значения матрицы. Спектр матрицы.
10. Связь спектра матрицы и её определенности.
11. Спектральное разложение матрицы.
12. Сингулярное разложение матрицы.
13. Связь определителя и собственных чисел для квадратной матрицы.
14. Связь следа и собственных чисел для квадратной матрицы.
15. Линейная сходимость последовательности.
16. Сублинейная сходимость последовательности.
17. Сверхлинейная сходимость последовательности.
18. Квадратичная сходимость последовательности.
19. Рассмотрим итеративный процесс $x_k \in \mathbb{R}$, сходящийся к решению $x^* \in \mathbb{R}$ квадратично. Как изменяется количество верных значащих цифр в решении после одной итерации метода?
20. Формулировка теста корней для определения скорости сходимости последовательности.
21. Формулировка теста отношений для определения скорости сходимости последовательности.
22. Унимодальная функция.
23. Метод дихотомии.
24. Метод золотого сечения.
25. Метод параболической интерполяции.
26. Условие достаточного убывания для неточного линейного поиска.
27. Условия Гольдштейна для неточного линейного поиска.
28. Условие ограничения на кривизну для неточного линейного поиска.
29. Градиент функции $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$.
30. Гессиан функции $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$.
31. Якобиан функции $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$.
32. Формула для аппроксимации Тейлора первого порядка $f_{x_0}^I(x)$ функции $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ в точке x_0 .
33. Формула для аппроксимации Тейлора второго порядка $f_{x_0}^{II}(x)$ функции $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ в точке x_0 .
34. Связь дифференциала функции df и градиента ∇f для функции $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$.
35. Связь второго дифференциала функции d^2f и гессиана $\nabla^2 f$ для функции $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$.
36. Формула для приближенного вычисления производной функции $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ по k -ой координате с помощью метода конечных разностей.
37. Пусть $f = f(x_1(t), \dots, x_n(t))$. Формула для вычисления $\frac{\partial f}{\partial t}$ через $\frac{\partial x_i}{\partial t}$ (Forward chain rule).
38. Пусть L - функция, возвращающая скаляр, а v_k - функция, возвращающая вектор $x \in \mathbb{R}^t$. Формула для вычисления $\frac{\partial L}{\partial v_k}$ через $\frac{\partial L}{\partial x_i}$ (Backward chain rule).
39. Афинное множество. Аффинная комбинация. Аффинная оболочка.
40. Выпуклое множество. Выпуклая комбинация. Выпуклая оболочка.
41. Конус. Выпуклый конус. Коническая комбинация. Коническая оболочка.
42. Сумма Минковского.
43. Любые 2 операции с множествами, сохраняющие выпуклость.
44. Выпуклая функция.

45. Надграфик функции $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$.
46. Дифференциальный критерий выпуклости первого порядка.
47. Дифференциальный критерий выпуклости второго порядка.
48. Связь выпуклости функции и её надграфика.
49. μ -сильно выпуклая функция.
50. Дифференциальный критерий сильной выпуклости первого порядка.
51. Дифференциальный критерий сильной выпуклости второго порядка.
52. Любые 2 операции с функциями, сохраняющие выпуклость.
53. Необходимые условия локального экстремума.
54. Достаточные условия локального экстремума.
55. Общая задача математического программирования. Функция Лагранжа.
56. Теорема Каруша - Куна - Таккера в форме необходимых условий решения задачи математического программирования.
57. Условие Слейтера.
58. Задача выпуклого программирования.
59. Задача линейного программирования. Задача линейного программирования в стандартной форме.
60. Идея симплекс метода.
61. Сходимость симплекс метода.
62. Показать, что направление антиградиента - направление наискорейшего локального убывания функции.
63. Метод градиентного спуска.
64. Наискорейший спуск.
65. Как направлены две соседние итерации метода наискорейшего спуска по отношению друг к другу?
66. Липшицева парабола для гладкой функции.
67. Размер шага наискорейшего спуска для квадратичной функции.
68. Характер сходимости градиентного спуска для гладких выпуклых функций в терминах \mathcal{O} от числа итераций метода.
69. Характер сходимости градиентного спуска для гладких и сильно выпуклых функций в терминах \mathcal{O} от числа итераций метода.
70. Связь спектра гессиана с константами сильной выпуклости и гладкости функции.
71. Условие Поляка-Лоясиевича (градиентного доминирования) для функций.
72. Сходимость градиентного спуска для сильно выпуклых квадратичных функций. Оптимальные гиперпараметры.
73. Связь PL-функций и сильно выпуклых функций.
74. Привести пример выпуклой, но не сильно выпуклой задачи линейных наименьших квадратов (возможно, с регуляризацией).
75. Привести пример сильно выпуклой задачи линейных наименьших квадратов (возможно, с регуляризацией).
76. Нижние оценки для гладкой выпуклой оптимизации с помощью методов первого порядка в терминах \mathcal{O} от числа итераций метода.
77. Нижние оценки для гладкой сильно выпуклой оптимизации с помощью методов первого порядка в терминах \mathcal{O} от числа итераций метода.
78. Отличие ускоренной и неускоренной линейной сходимости для методов первого порядка.
79. Метод тяжелого шарика (Поляка).
80. Понятие локальной и глобальной сходимости численного метода оптимизации.
81. Ускоренный градиентный метод Нестерова для выпуклых гладких функций.
82. Ускоренный градиентный метод Нестерова для сильно выпуклых гладких функций.
83. A-сопряженность двух векторов. A-ортогональность. Скалярное произведение $\langle \cdot, \cdot \rangle_A$.

84. Зависимость сходимости метода сопряженных градиентнов от спектра матрицы.
85. Метод Ньютона.
86. Сходимость метода Ньютона для квадратичной функции.
87. Характер сходимости метода Ньютона для сильно выпуклых гладких функций - куда и как сходится.
88. Демпфированный метод Ньютона.
89. Идея квазиньютоновских методов. Метод SR-1.
90. Проекция.
91. Достаточное условие существования проекции точки на множество.
92. Достаточное условие единственности проекции точки на множество.
93. Метод проекции градиента.
94. Критерий проекции точки на выпуклое множество (Неравенство Бурбаки-Чейни-Гольдштейна).
95. Проекция как нерастягивающий оператор.
96. Метод Франк-Вульфа.
97. Характер сходимости метода проекции градиента для гладких выпуклых функций в терминах \mathcal{O} от числа итераций метода.
98. Характер сходимости метода проекции градиента для гладких сильно выпуклых функций в терминах \mathcal{O} от числа итераций метода.
99. Характер сходимости метода Франк-Вульфа для гладких выпуклых функций в терминах \mathcal{O} от числа итераций метода.
100. Характер сходимости метода Франк-Вульфа для гладких сильно выпуклых функций в терминах \mathcal{O} от числа итераций метода.
101. Привести пример выпуклой негладкой задачи линейных наименьших квадратов (возможно, с регуляризацией).
102. Субградиент. Субдифференциал.
103. Как считать субдифференциал поточечного максимума выпуклых функций.
104. Субградиентный метод.
105. Характер сходимости субградиентного метода для негладких выпуклых Липшицевых функций в терминах \mathcal{O} от числа итераций метода.
106. Какому условию должна удовлетворять стратегия выбора шага, чтобы субградиентный метод сходился для выпуклых Липшицевых функций?
107. Характер сходимости субградиентного метода для негладких сильно выпуклых Липшицевых функций в терминах \mathcal{O} от числа итераций метода. Стратегия выбора шага.
108. Метод стохастического градиентного спуска.
109. Идея мини-батча для метода стохастического градиентного спуска. Эпоха.
110. Характер сходимости стохастического градиентного спуска для гладких выпуклых функций в терминах \mathcal{O} от числа итераций метода.
111. Характер сходимости стохастического градиентного спуска для гладких PL-функций в терминах \mathcal{O} от числа итераций метода.
112. Характер работы стохастического градиентного спуска с постоянным шагом для гладких PL-функций.
113. Grokking.
114. Double Descent.
115. Взрыв/Затухание градиентов при обучении глубоких нейронных сетей.
116. Идея gradient checkpointing.

Теоремы с доказательствами

1. Тест корней для определения скорости сходимости последовательности.
2. Метод дихотомии и золотого сечения для унимодальных функций. Скорость сходимости.
3. Автоматическое дифференцирование. Вычислительный граф. Forward/ Backward mode (в этом вопросе нет доказательств, но необходимо подробно описать алгоритмы).
4. Выпуклость надграфика как критерий выпуклости функции.
5. Дифференциальный критерий сильной выпуклости первого порядка.
6. Дифференциальный критерий сильной выпуклости второго порядка.
7. Необходимые условия безусловного экстремума.
8. Достаточные условия безусловного экстремума.
9. Симплекс метод для задачи линейного программирования с ограничениями в форме неравенств.
Двухфазный симплекс метод.
10. Теорема сходимости градиентного спуска для гладких PL функций.
11. Теорема сходимости градиентного спуска для сильно выпуклых квадратичных функций.
Оптимальные гиперпараметры.