Вопрос 1

Решение

Несложно заметить, что последовательность x_n знакочередующееся и стремится к нулю. Строго докажем это утверждение. Пусть

$$a_n = (-1)^{n+1} x_n.$$

Тогда

$$a_{n+1} = a_n - a_n^3, \quad a_1 = 1/2$$

Докажем индукцией, что $0 < a_n \le 1/2$

$$a_{n+1} \quad = a_n(1 - a_n^2) > 0 \cdot (1 - 1/4) > 0;$$

$$a_{n+1} = a_n - a_n^3 < a_n \le 1/2$$

Мы также показали, что $a_{n+1} < a_n$, то есть последовательность a_n монотонно убывает.

По теореме Вейерштрасса у ограниченной монотонной последовательности a_n существует конечный предел. Пусть

$$A = \lim_{n \to +\infty} a_n$$

Тогда переходя к пределу в рекуррентном соотношении получаем равенство:

$$A = A - A^3 \quad \Rightarrow \quad A = 0.$$

Таким образом, ряд принимает вид ряда Лейбница:

$$\sum_{n=1}^{+\infty} x_n = \sum_{n=1}^{+\infty} (-1)^{n+1} a_n.$$

Из монотонного стремления a_n к нулю, следует, что данный ряд сходится.

Критерии оценивания

- **10 баллов**: Полностью обоснованное решение.
- **8–9 баллов**: Правильное решение с несущественными погрешностями в его обосновании. Например, если нет полного доказательства по индукции.
- **6–7 баллов**: Правильно выполнены основные действия, но есть ошибки в выкладках или отсутствует обоснование выполненных действий.
- **4–5 балла**: Задача решена лишь частично. Например, проверено лишь стремлении к нулю общего члена ряда.
- **1–3 баллов**: Различные попытки решения. Высказаны идеи, ведущие к решению.

Вопрос 2

Решение

Раз образ имеет размерность 9, то ядро имеет размерность 17 - 9 = 8. Значит, оператор имеет собственное значение 0 кратности хотя бы 8. Сумма кратностей всех собственных значений не превосходит размерности пространства; при этом 8 слотов уже занято. Следовательно, собственных значений может быть не больше 9 + 1 = 10. Осталось привести пример --- диагональную матрицу с числами

$$\underbrace{0,\ldots,0}_{8},1,2\ldots,9$$

на диагонали.

Минимальное число равно 1 (нет собственных значений, кроме нуля). Но нам нужно привести пример. Подойдёт верхнетреугольная матрица 17x17, у которой на диагонали нули, а ядро совпадает с линейной оболочкой, скажем, первых 8 базисных векторов. Она может иметь блочный вид

$$\begin{pmatrix} 0 & E \\ 0 & 0 \end{pmatrix},$$

где E --- единичная матрица размера 8x9.

Критерии оценки

- Объяснено, что больше 10 нельзя - **4 балла**;
- Приведён пример - **1 балл**;
- Приведён пример матрицы с только нулевым собственным значением и образом нужной размерности - **5 баллов**;
- Задача решалась без учёта слова "различных" в условии - **0 баллов** за задачу в целом;
- Игнорируются нулевые собственные значения - для оценки сверху может быть поставлен **1 балл** за пример; в случае, если имеется верный по сути пример для оценки снизу, за него ставится **4 балла**.

Вопрос 3

Ответ

$$\frac{1}{C_{10}^3}=\frac{1}{120}$$

Решение

Вариант 1

С точки зрения вероятности получить ту или иную последовательность шаров, вытащить по-очереди 10 шаров – то же самое, что вытащить случайный набор из 10 шаров, а затем случайным образом упорядочить. Поэтому искомая условная вероятность встретить 3 белых шара на первых местах равна условной вероятности встретить их на любых других местах. Всего таких вероятностей C_{10}^3 штук, в сумме они дают единицу. Следовательно, искомая вероятность равна $1/C_{10}^3$.

Вариант 2

Пусть событие *A* --- ``Первые три шара - белые", событие *B* --- ``Из 10 вытащенных шаров 3 белые". Посчитаем условную вероятность по формуле $P(A|B)=P(A\cap B)/P(B)$. Вероятность встретить сначала 3 белых, а потом 7 черных равна:

$$P(A\cap B)=\frac{50}{100}\cdot\frac{49}{99}\cdot\frac{48}{98}\cdot\frac{50}{97}\cdot\frac{49}{96}\cdot\ldots\cdot\frac{44}{91}=\frac{(50!/47!)\cdot(50!/43!)}{100!/90!}=\frac{50!\ 50!\ 90!}{47!\ 43!\ 100!}.$$

Вероятность встретить из 10 шаров 3 белых равна:

$$P(B) = \frac{C_{50}^{3} \cdot C_{50}^{7}}{C_{100}^{10}} = \frac{50!\ 50!\ 10!\ 90!}{3!\ 47!\ 7!\ 43!\ 100!}.$$

Следовательно, искомая вероятность:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{50!\ 50!\ 90!}{47!\ 43!\ 100!} \cdot \frac{3!\ 47!\ 7!\ 43!\ 100!}{50!\ 50!\ 10!\ 90!} = \frac{3!\ 7!}{10!} = \frac{1}{120}.$$

Критерии

- **10 баллов**: Верное решение и правильный ответ в форме $1/C_{10}^{3}, \frac{1}{120}$, или аналогичной;
- **9 баллов**: Верное решение и незначительная ошибка, не повлиявшая на ответ, либо арифметическая ошибка;
- **5 баллов**: Верное решение, но не посчитанный ответ, например в форме несокращенной дроби $\frac{(50!/47!) \cdot (50!/43!) \cdot C_{100}^{10}}{(100!/90!) \cdot C_{50}^{3} \cdot C_{50}^{7}}$;
- **2 балла**: В целом верное решение, но с одной содержательной ошибкой в вычислении какого-то из множителей;
- **0 баллов**: Считается не та вероятность, про которую говорилось в задаче.

Вопрос **4**

Ответ

$$\frac{10!}{12}$$

Решение

Вариант 1

Выберем 3 места из 10. Это можно сделать (10,3)=120 способами. Расположим на них a, b и c (порядок определяется однозначно). Для каждого из этих вариантов выберем (7,2)=21 способов поставить d и e на 2 из оставшихся 7 мест. Для каждого из этих вариантов оставшиеся 5 элементов можно поставить на 5 мест в произвольном порядке (5!). Итого 120*21*120=302400 порядков. Возможны аналогичные трехходовые решения, отличающиеся порядком выбора.

Вариант 2

Рассмотрим случайный линейный порядок. Вероятность того, что выполняется первое условие, равна 1/6 (все относительные расположения a, b и c равновероятны), вероятность второго --- 1/2, условия независимы (это нужно доказать в работе!), поэтому вероятность того, что случайный линейный порядок удовлетворяет условию задачи равна 1/12. Поэтому нам подходят 10!/12 порядков.

Критерии оценивания

- Арифметическая ошибка в ответе (при правильном решении) - 9 **баллов**;
- Переборное решение при неверном ответе - **0 баллов**;
- Решение (как связный текст на русском языке) отсутствует - **0 баллов**;
- Не доказана независимость событий при правильном решении 2-го типа - 7 **баллов**;
- Одна ошибка при решении 1-го типа - не более 5 **баллов**;
- Две ошибки при решении 1-го типа - не более 3 **баллов**;
- Неверное понимание, что такое нестрогий линейный порядок (например, считается число

слабых или частичных) – 0 **баллов**;
- Комбинаторная часть решения пропущена – 3 **балла**;
- Решение второго типа с использованием независимости событий, таковыми не являющихся – 3 **балла**;
- Правильное решение, предполагающее, что a, b, c и d, e стоят рядом – 3 **балла**.

Общие критерии (используются, если решение не подходит под частные)

- Решение содержит пробелы или ошибки, однако его можно признать верным или почти верным – 7 **баллов**;
- Либо решена половина задачи, либо недочеты слишком существенны, чтобы признать решение верным (но само рассуждение основано на здравых идеях) – 5 **баллов**;
- Задача не решена, но в тексте есть идеи, которые могут при должном развитии привести к решению – 3 **балла**.

Вопрос 5

Ответ

8 матриц следующего вида

$$B(\epsilon_1,\epsilon_2,\epsilon_3)=\begin{pmatrix} 3\epsilon_1+8\epsilon_3-6\epsilon_2\sqrt{2} & 3\epsilon_1+12\epsilon_3-9\epsilon_2\sqrt{2} & -6\epsilon_1-18\epsilon_3+15\epsilon_2\sqrt{2} \\ -2\epsilon_1+8\epsilon_3-2\epsilon_2\sqrt{2} & -2\epsilon_1+12\epsilon_3-3\epsilon_2\sqrt{2} & 4\epsilon_1-18\epsilon_3+5\epsilon_2\sqrt{2} \\ 8\epsilon_3-4\epsilon_2\sqrt{2} & 12\epsilon_3-6\epsilon_2\sqrt{2} & -18\epsilon_3+10\epsilon_2\sqrt{2} \end{pmatrix}$$

Решение

Матрица A имеет три различных положительных собственных значения $\lambda_1=1,\lambda_2=2,\lambda_3=4$. Соответствующие им фундаментальные собственные вектора имеют вид (-3,2,0), (3,1,2), (1,1,1).

Следовательно, матрица A может быть записана в виде $A=PDP^{-1}$, где

$$P=\begin{pmatrix} -3 & 3 & 1 \\ 2 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix},\quad D=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix},\quad P^{-1}=\begin{pmatrix} -1 & -1 & 2 \\ -2 & -3 & 5 \\ 4 & 6 & -9 \end{pmatrix}.$$

Известно, что искомая матрица B имеет вид $B=PD^{1/2}P^{-1}$, где

$$D=\begin{pmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_2\sqrt{2} & 0 \\ 0 & 0 & 2\epsilon_3 \end{pmatrix}$$

и $\epsilon_j=\pm1$. Таким образом, существует 8 разных матриц B, удовлетворяющих условиям задачи.

Критерии оценивания

- **10 баллов**: Полностью обоснованное решение с правильным ответом;
- **8–9 баллов**: Верно найдены матрицы P, D. Указано, что $B=PD^{1/2}P^{-1}$
- **7 баллов**: Верно найдены матрицы P, D. Но получен лишь один ответ, отвечающий случаю $e_j=1$;
- **6 баллов**: Неверно найдены матрицы P, D. Указано, что $B=PD^{1/2}P^{-1}$ и что задача имеет 8 разных ответов;
- **4–5 баллов**: Неверно найдены матрицы P, D. Указано, что $B=PD^{1/2}P^{-1}$. Не указано, что задача имеет 8 разных ответов;
- **1–3 баллов**: Указано, что необходимо диагонализировать матрицу A, но найдены лишь собственные значения и некоторые собственные вектора.

Вопрос 6

Решение

Нам будет удобно соединить начала и концы в один массив `S[1:n]` пар `[B[i], E[i]]`.

Отсортируем отрезки по их началам; сделать это можно, например, с помощью сортировки слиянием (это потребует $O(n \log n)$ операций и $O(n)$ дополнительной памяти. Теперь будем итерироваться по индексу `i`. Чтобы найти нужный отрезок для `S[i]`, мы возьмём его конец `S[i][2]` и найдём первый элемент `S[j]`, для которого `S[j][1] >= S[i][2] + 1` с помощью немного модифицированного бинпоиска. А именно, мы будем делить отрезок `[i + 1:n]` пополам, и если в серединке k имеем `S[k][1] < S[i][2] + 1`, переходим в правую половинку, если `S[k][1] = S[i][2] + 1`, предъявляем индекс k, а если `S[k][1] > S[i][2] + 1`, то в левую, но при этом мы поддерживаем в памяти последнее `k_last`, для которого имело место неравенство `S[k_last][1] > S[i][2] + 1`, и если оказалось, что числа `S[i][2] + 1` мы не нашли, то предъявляем `k_last`.

Описанный алгоритм имеет сложность $O(n \log n + n \log n)$ (сортировка + бинарные поиски) и требует $O(n)$ дополнительной памяти (дополнительный массив для хранения отрезков + дополнительный массив при сортировке + массив для записи ответов).

Критерии оценки

- описан неверный алгоритм – **0 баллов** за всю задачу вне зависимости от остального;
- алгоритм верен, но не удовлетворяет ограничениям по числу операций – **2 балла**;
- алгоритм верен и удовлетворяет ограничениям по числу операций – **8 баллов**;
- оценка числа операций – **1 балл**;
- оценка объёма дополнительной памяти – **1 балл**;
- не обоснована работоспособность алгоритма – **может вычитается до 3 баллов**;
- решение использует сложные структуры данных без пояснений их внутренней структуры – **вычитается 2 балла**;
- в решении используются конкретные реализации классов алгоритмов и структур данных из тех или иных языков программирования без пояснения того, какие именно представители этих классов имплементированы – **вычитается 2 балла**;
- алгоритм верен, но удовлетворяет ограничениям по числу операций лишь в среднем, а не в худшем случае – **6 баллов**;
- алгоритм верен, но использует не поименованную сортировку, для которой безапелляционно постулированы некоторые свойства – **5 баллов**;
- модифицированный бинарный поиск не описан явно – **снижается до 2 баллов**.

Вопрос 7

Решение

1. Граф планарен, это булев куб, в каждой грани которого проведена одна диагональ из двух. Как правило, в работах плоский граф был построен в явном виде, этого достаточно.
2. В графе 8 вершин и 18 ребер, поэтому при добавлении ребра их станет 19 и нарушится условие $E \le 3V-6$, т.е. граф перестанет быть планарным.

Внимание! Выполнение неравенства не помогает решить 1) – это работает только в одну сторону.

Критерии оценивания

- Правильное выполнение пункта 1) – **5 баллов**;
- Правильное выполнение пункта 2) – **5 баллов**;
- Решение (как связный текст на русском языке) отсутствует – **0 баллов**;
- Неверное понимание понятия гомеоморфности в теореме Понтрягина – Куратовского (например, K_5 и K_3,3 не обязаны быть подграфами исходного графа) – **0 баллов**;

- Попытки доказать непланарность исходного графа – 0 **баллов**;
- Неверное понимание планарности (например, предположение, что K_4 не планарен) – 0 **баллов**;
- 1) Использование утверждения E<=3V-6, теоремы Эйлера или аналогичных в качестве критерия планарности – 0 **баллов**;
- 2) Попытка перебрать варианты, как добавить ребро, без приведённого перебора – 0 **баллов**;
- 2) Решение основывается на конкретном способе построить плоский граф, а их может быть (бесконечно) много. Утверждение, о том, что в остальных случаях будет то же самое, требует доказательства – 0 **баллов**;
- 2) Очевидно доводящийся до конца, но не сделанный перебор – **не более** 2 **баллов**;
- 2) Используется, но не доказано, что все грани любого плоского графа, соответствующего исходному, будут треугольниками – 2 **балла**;

Общие критерии (используются, если решение не подходит под частные)

- Решение содержит пробелы или ошибки, однако его можно признать верным или почти верным – 7 **баллов**;
- Либо решена половина задачи, либо недочеты слишком существенны, чтобы признать решение верным (но само рассуждение основано на здравых идеях) – 5 **баллов**;
- Задача не решена, но в тексте есть идеи, которые могут при должном развитии привести к решению – 3 **балла**.

Вопрос 8

Решение

Пусть D – событие, что выиграл Дима. Обозначим через K_i – событие, что в результате раунда получилось число i. Очевидно, что $2 \le i \le 12$ и что

$$P(D) = \sum_{i=2}^{12} P(K_i)P(D|K_i).$$

Несложным перебором получаем

$$P(K_2) = P(K_{12}) = \frac{1}{36}, \quad P(K_3) = P(K_{11}) = \frac{2}{36}, \quad P(K_4) = P(K_{10}) = \frac{3}{36},$$
$$P(K_5) = P(K_9) = \frac{4}{36}, \quad P(K_6) = P(K_8) = \frac{5}{36}, \quad P(K_7) = \frac{6}{36}.$$

Из условий задачи сразу следует, что

$$P(D|K_2) = P(D|K_3) = P(D|K_{12}) = 0, \quad P(D|K_7) = P(D|K_{11}) = 1.$$

Следовательно,

$$P(D) = \frac{6}{36} + \frac{2}{36} + \sum_{i \in W} P(K_i)P(D|K_i),$$

где W={4,5,6,8,9,10}. Осталось заметить, что при $i \in \{4,5,6,8,9,10\}$ выполнено

$$P(D|K_i) = \frac{P(K_i)}{P(K_i) + P(K_7)}.$$

Эту формулу можно получить несколькими способами. Например, вероятность того, что Дима выигрывает в отдельно рассматриваемом j-ом раунде равна P(K_i), а вероятность перехода к следующему раунду равна 1-P(K_i)-P(K_7). Следовательно,

$$P(D|K_i) = P(K_i) + (1 - P(K_i) - P(K_7))P(K_i) + (1 - P(K_i) - P(K_7))^2 P(K_i) + \ldots =$$
$$P(K_i)\sum_{j=0}^{\infty}(1 - P(K_i) - P(K_7))^j = \frac{P(K_i)}{P(K_i) + P(K_7)}.$$

Можно и просто заметить, что справедливо равенство

$$P(D|K_i) = P(K_i) + (1 - P(K_i) - P(K_7))P(D|K_i)$$

из которого опять же выводится требуемая формула.

В итоге мы получаем, что

$$P(D|K_4) = P(D|K_{10}) = \frac{1}{3}, \quad P(D|K_5) = P(D|K_9) = \frac{2}{5}, \quad P(D|K_6) = P(D|K_8) = \frac{5}{11}.$$

В итоге получаем

$$P(D) = \frac{8}{36} + \frac{3}{36} \cdot \frac{2}{3} + \frac{4}{36} \cdot \frac{4}{5} + \frac{5}{36} \cdot \frac{10}{11} = \frac{244}{495} = 0,49(29).$$

Критерии оценивания

- **10 баллов**: Полностью обоснованное решение с правильным ответом;
- **8–9 баллов**: Верное решение с правильным ответом, но с недостаточно подробным обоснованием;
- **7 баллов**: Верное решение с несущественными арифметическими ошибками (например при вычислении итоговой суммы в P(D)), при устранении которых получается абсолютно верное решение;
- **4–6 баллов**: Получено существенное продвижение при решении задачи, но либо присутствуют серьезные ошибки, повлекшие неверный ответ, либо решение не доведено до конца.
  Например, не найдена (или получен ее неверный аналог) формула $P(D|K_i) = \dfrac{P(K_i)}{P(K_i) + P(K_7)}$, или сумма $\sum\limits_{i \in W}$ умножена на 2/3 – вероятность того, что $i \in W$ (т.е. что игра продолжилась после первого раунда);
- **1–3 баллов**: Различные попытки решения.

Вопрос 9

Ответ

Да, существует

Решение (одно из)

Сначала определим предикат – быть "чётным" элементом: E(x) как $\exists y : x = y + y$. Теперь основное соображение: в *Z+Z* любые три элемента "линейно зависимы по модулю 2", т.е. для **x,y,z** одна из сумм $x, y, z, x + y, x + z, y + z, x + y + z$ "чётна". В **Z+Z+Z** это не так: можно взять **x=(1,0,0), y=(0,1,0), z=(0,0,1)**. Формула, таким образом:

$$\forall x, y, z (E(x) \vee E(y) \vee E(z) \vee E(x+y) \vee E(x+z) \vee E(y+z) \vee E(x+y+z)) .$$

Критерии оценивания

- **10 баллов**: полностью верное решение, приведена правильная формула (в т.ч. отличная от указанного решения);
- **8 баллов**: правильно описаны все идеи построения формулы, но сама финальнаая формула отсутствует;
- **0 баллов**: идеи построения формулы не указаны или указаны неправильно; бинарный ответ ("да, существует") без обоснований не принимается, так как его можно легко получить простым угадыванием.

Вопрос 10

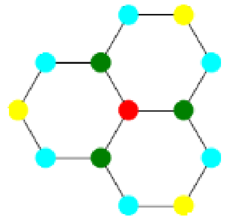
Ответ

$$3(\frac{N*(N+1)}{2})+1$$

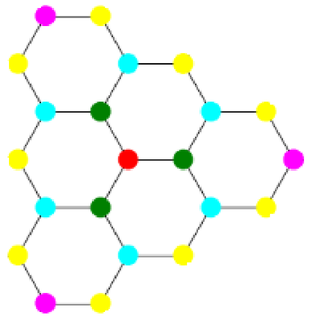
Решение

Шаг 1. Визуализация (2 балла)

Нарисуем первые три квартала. На рисунке ниже красная вершина – центр города, к ней премыкает три квартала.



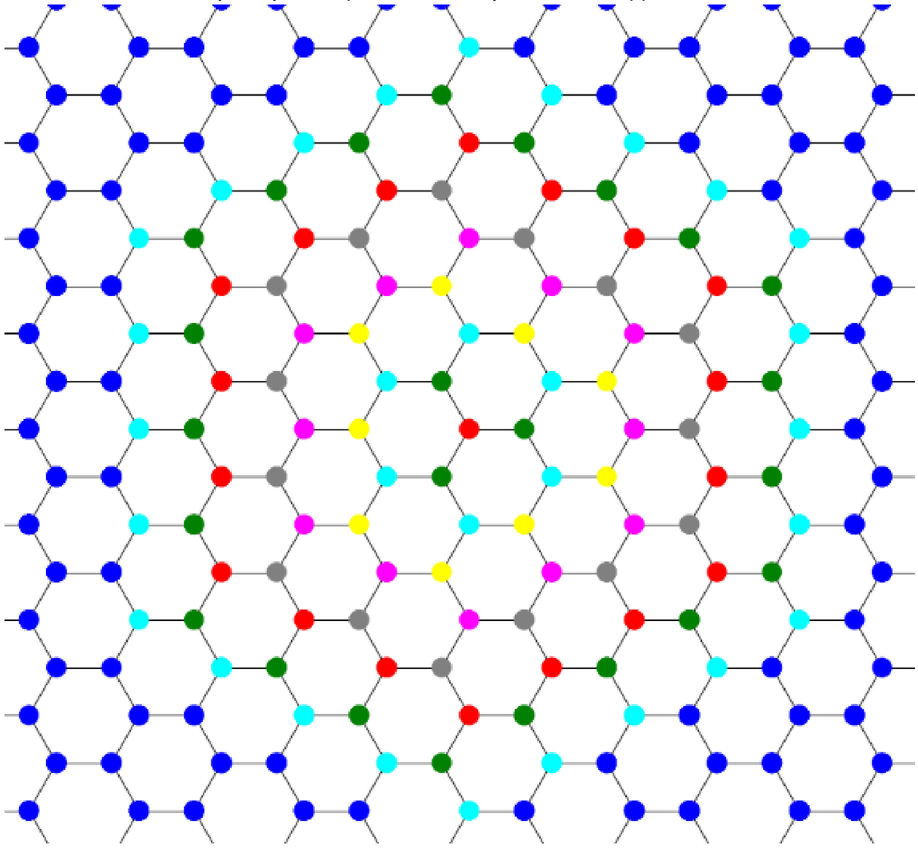
Нарисуем следующие три квартала из развязок на стыке двух кварталов (зеленые вершины)



Шаг 2. Достижимость (3 балла)

При условии N < K, где K – степень графа, видим, что достижимые вершины прирастают гораздо медленнее чем новые кварталы, а значит при оценке достижимости не нужно заботиться о том, какие узлы на окраинах города будут достижимы из центра.

Рассчитаем зависимость количества узлов, достижимых из центральной вершины за N шагов. Под достижимостью здесь и далее будем понимать достижимость любым наикратчайшим путём на графе. Отметим все вершины, достижимые за 1 шаг зеленым цветом, достижимые за 2 шага – бирюзовым, достижимые за 3 шага – желтым и т.д.



Шаг 3. Радиусы (3 балла)

Обратим внимание, что на каждом следующем "радиусе" на *три* вершины больше, чем на предыдущем. Таким образом зависимость количества вершин, достижимых ровно за N шагов можно выписать в виде таблицы:

Зависимость количества вершин

N	Вершин на i радиусе	Вершин на всех радиусах
0	1	1
1	3	4
2	6	10
3	9	19
4	12	31
5	15	46
6	18	66
7	21	85
8	24	109
9	27	136
10	30	166
11	33	199
12	36	235
13	39	274
14	42	316

Шаг 4: Прогрессия (2 балла)

Тогда общее количество вершин, достижимых из центра за N шагов будет

$$3N + 3(N-1) + 3(N-2) + .. + 3 + 1$$

или

$3(N + (N-1) + (N-2) + ... + 1) + 1$

или

$$3(\frac{N*(N+1)}{2}) + 1 \quad 3(N*(N+1)2)+1$$

$$3(\frac{N*(N+1)}{2}) + 1$$

Вопрос 11

Ответ

e

Решение

Заметим, что распределение непрерывное, вероятность совпадения каких-то двух значений равна 0. Удобно считать, что все случайные величины $l_i, i \in \mathbb{N}$ принимают какое-то значение, даже когда процесс обучения остановлен.

Обозначим случайные события:

$A_n, n \geqslant 1$ --- функция потерь была посчитана ровно n раз

$B_n = \neg(A_1 \sqcup A_2 \sqcup \ldots \sqcup A_n), n \geqslant 0$ --- не менее n+1 раза.

Заметим, что событие $B_n, n \geq 1$, случается в том и только том случае, когда последовательность из первых n значений функции потерь --- неубывающая. Но вероятность такого события равна 1/ (n!). Действительно, $l_1, l_2, ..., l_n$ --- независимые случайные величины из одного расредeления, поэтому вероятность получить при их упорядочивании по возрастанию именно последовательность $l_n, l_{n-1}, ..., l_1$ такая же, как и любую другую из n! перестановок. Отдельно заметим, что та же формула выполняется и для n=0.

$P(B_0) = 1 = 1/0!$

$P(B_n) = P(l_1 > l_2 > \ldots > l_n) = 1/n!, \quad n \geqslant 1$

Теперь посчитаем искомое мат.ожидание $\sum nP(A_n)$, поменяв порядок суммирования и подставив значение стандартного ряда $\sum \frac{1}{k!}$

$$\sum_{i=1}^{\infty} i \cdot P(A_i) \quad = \sum_{k=1}^{\infty} (\sum_{i=k}^{\infty} P(A_i)) = \sum_{k=1}^{\infty} P\left(\bigsqcup_{i=k}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(B_{k-1}) =$$
$$= \sum_{k=0}^{\infty} P(B_k) = \sum_{k=0}^{\infty} \frac{1}{k!} = e$$

Критерии

- 10 **баллов**: Верное решение и ответ;
- 8 **баллов**: 1/(n!) не сокращён до e;
- 7 **баллов**: В целом верное решение, но l_1 не посчитано за отдельное число (что приводит к ответу e-1) или путаница, где n, а где n+1, приводящая к не совсем верному ответу вроде e+1 или e-1.
- 0 **баллов**: Решение предполагает, что события $l_1 > l_2, l_2 > l_3$ независимы

Вопрос 12

Решение

Найдём плотность распределения $\widehat{\theta}$. Начнём с функции распределения:

$$F_{\widehat{\theta}}(x) = \mathbb{P}\{X_1 \leqslant x, \dots, X_n \leqslant x\}$$

В силу независимости это можно переписать в виде:

$$F_{\widehat{\theta}}(x) = \mathbb{P}\{X_1 \leqslant x\} \cdot \dots \cdot \mathbb{P}\{X_n \leqslant x\} = F_{X_1}(x)^n$$

Дифференцируя это выражение и используя то, что $F'(x) = \frac{1}{\theta}\mathbb{I}[0 < x < \theta]$, получаем

$$p_{\widehat{\theta}}(x) = N\frac{x^{N-1}}{\theta^N}\mathbb{I}[0 < x < \theta]$$

Матожидание равно:

$$\mathbb{E}_\theta(\widehat{\theta}) = \int_0^\theta x \cdot N\frac{x^{N-1}}{\theta^N}dx = \frac{N}{N+1}\theta$$

Найдём bias нашей оценки:

$$\text{bias}_\theta(\widehat{\theta}) = -\frac{1}{N+1}\theta \to 0 \quad N \to \infty$$

Это означает, что наша оценка является асимптотически несмещённой (но не несмещённой!).

Чтобы проверить состоятельность, нам осталось разобраться с $\mathbb{V}_\theta(\widehat{\theta})$. Найдём её и убедимся, что она стремится к нулю при $N \to \infty$.

Вначале посчитаем второй момент нашей оценки:

$$\mathbb{E}_\theta(\widehat{\theta})^2 = \frac{N}{N+2}\theta^2$$

Теперь посчитаем дисперсию:

$$\mathbb{V}_\theta(\widehat{\theta}) = \mathbb{E}_\theta(\widehat{\theta})^2 - (\mathbb{E}_\theta\widehat{\theta})^2 = \\ \frac{N}{N+2}\theta^2 - \left(\frac{N}{N+1}\theta\right)^2 = \\ \frac{N}{(N+1)^2(N+2)}\theta^2 \xrightarrow[N\to\infty]{} 0$$

Получается, что оценка является состоятельной, так как и bias, и дисперсия нашей оценки стремятся к 0 при $N \to \infty$.

Критерии оценивания

- Проверка несмещённости – 4 **балла**;
- Проверка состоятельности – 6 **баллов**;
- Из асимптотической несмещённости делается вывод о том, что оценка несмещённая – 2 **балла** за проверку несмещённости;
- Арифметические ошибки, не повлиявшие на суть задачи, при условии корректности рассуждений – **вычитается 1 балл за каждую арифметическую ошибку**, но не более 5 баллов;
- Не обоснованные логические переходы – **вычитается 2 балла за каждый переход**, но не более 8 баллов.

Вопрос 13

Решение

Найдём градиент функции потерь. Обозначим для краткости

$$g(t) = \log(\cosh(t))$$

Имеем

$$\frac{\partial}{\partial w_j}\mathcal{L}(w, X, y) = \frac{1}{N}\sum_{i=1}^{N} g'(\langle x_i, w\rangle - y_i)x_{ij}$$

$$\frac{\partial}{\partial w}\mathcal{L}(w, X, y) = \frac{1}{N}\sum_{i=1}^{N} g'(\langle x_i, w\rangle - y_i)x_i$$

В условии сказано, что x_i линейно независимы, и это существенно упрощает задачу, потому что в этом случае можно сразу сказать, что если их линейная комбинация равна нулю, то коэффициенты в ней равны нулю. Мы получаем систему

$$g'(\langle x_i, w\rangle - y_i) = 0$$

Нетрудно убедиться, что $g'(t) = \tanh(t)$, так что отсюда

$$\langle x_i, w\rangle - y_i = 0$$

Это система линейных уравнений $X^T w = y$ с линейно независимыми строками, но неизвестным количеством столбцов. Она вполне может иметь бесконечно много решений.

Множитель 1/N действительно не так важен для оптимизации, но (1) при сравнении качества на выборках различного размера позволяет отнормировать значение функции потерь и (2) без него на разных размерах батча могут быть де факто разные величины шага оптимизации.

В окрестности начала координат функция $g(t) = \log(\cosh(t))$ ведёт себя как $const \cdot t^2$, а с ростом t --- как $const' \cdot |t|$. Таким образом, эта функция, как и MAE, не очень сильно штрафует за выбросы, но при этом дифференцируема (и даже дважды дифференцируема).

Критерии оценки

- Верный подсчёт градиента – 2 **балла**;
- Выводы про единственность – 2 **балла** (из них 1 за выводы из линейной независимости и ещё 1 за дальнейшее рассуждение про систему уравнений);
- Верные рассуждения насчёт 1/N – **до** 3 **баллов**;
- Верные соображения по поводу сравнения с другими функциями потерь – **до** 3 **баллов**.

Вопрос **14**

Решение

Из условия задачи сразу следует, что

$$g'(t) = \frac{g(t)}{2} - \frac{g^2(t)}{2}.$$

Решением этого дифференциального уравнения являются функции

$$g(t) = \frac{1}{1 - C \cdot e^{-t/2}}.$$

Так как g(0)=100, то C=99/100 и, следовательно, искомым ответом является функция

$$g(t) = \frac{100}{100 - 99 \cdot e^{-t/2}}.$$

Критерии оценивания

- **10 баллов**: Полностью обоснованной решение с правильным ответом.
- **8–9 баллов**: Верное решение с правильным ответом, с несущественными погрешностями в обосновании.
- **7 баллов**: Верное решение с несущественными арифметическими ошибками (например при вычислении константы C), при устранении которых получается абсолютно верное решение.
- **4–6 баллов**: Получено верное дифференциальное уравнение, но при его решении допущены существенные ошибки.
- **1–3 баллов**: Различные попытки решения.

Вопрос 15

Решение

А) Расчетное значение t-статистики: (6.9-7)/0.3 = -0.33. Критическое значение распределения Стьюдента с n-k=20-3=17 степенями свободы на уровне значимости 1% равно 2,898.

Модуль значения расчетной t-статистики меньше, чем критическое значение, значит мы не можем отклонить гипотезу о том, что коэффициент при переменной x_{1t} равен 7.

Б) Во-первых, необходимо рассчитать R_R^2 исходной модели по формуле: 1 - RSS/TSS = 0.9.

Далее, необходимо использовать F-тест в следующей форме:

$$F_{n-k,q} = \frac{R_{UR}^2 - R_R^2}{1 - R_{UR}^2} \bullet \frac{n-k}{q} = \frac{0,95 - 0,9}{1 - 0,95} \bullet \frac{20-5}{2} = 7,5$$

Где q - разница между числом свободных регрессоров в новой и исходной модели, а k - число свободных регрессоров в новой модели. Критическое значение $F_{15,2}$ с соответствующими степенями свободы на уровне значимости 1% по таблице составляет 6,36.

Расчетная величина 7,5 больше критического значения – соответственно нулевая гипотеза о равном качестве моделей отвергается, т.е. переменные добавлять стоит.

Критерии оценивания

- **9–10 баллов**: верное решение без пробелов и ошибок;
- **7–8 баллов**: решение содержит пробелы или ошибки, однако его можно признать верным или почти верным;
- **4–6 баллов**: либо решена половина задачи, либо недочеты слишком существенны, чтобы признать решение верным (но само рассуждение основано на здравых идеях);
- **2–3 баллов**: задача не решена, но в тексте есть идеи, которые могут при должном развитии привести к решению.

Вопрос 16

Решение

Можно ввести следующие обозначения: x_0 и x_1 - два различных возможных значений случайной величины x .

Тогда логарифм функции максимального правдоподобия запишется в виде:

$$\log L = \log\left[\prod_{i=1}^{n} F(\alpha+\beta x_i)^{y_i}\left(1-F(\alpha+\beta x_i)^{1-y_i}\right)\right] \xrightarrow{\text{by } \alpha \text{ and } \beta} \max$$

Раскрывая произведение, выражение log L можно упростить до следующего вида:

$$\log L = n_{01}\log F(u_0) + \ n_{00}\log(1-F(u_0)) + \ n_{11}\log F(u_1) + \ n_{10}\log(1-F(u_1))$$

где $u_0 = \alpha + \beta x_0, u_1 = \alpha + \beta x_1$ - линейные комбинации соответствующих значений x, а n_{pq} – просто число наблюдений, для которых $x = x_p, y = q$, а индексы $p, q \in \{0, 1\}$.

Условия первого порядка (берем производные от log L по α, β и приравниваем к нулю), получаем после упрощения:

$$\begin{cases} \alpha + \beta x_0 = F^{-1}\left(\frac{n_{01}}{n_{01}+n_{00}}\right) \\ \alpha + \beta x_1 = F^{-1}\left(\frac{n_{11}}{n_{11}+n_{10}}\right) \end{cases}$$

Отсюда видно, что при подстановке этих выражений в log L будет произведена операция $F(F^{-1}(\bullet))$, что приведет отсутствию зависимости от конкретного вида функции $F(\bullet)$.

Критерии оценивания

- 9–10 **баллов**: верное решение без пробелов и ошибок;
- 7–8 **баллов**: решение содержит пробелы или ошибки, однако его можно признать верным или почти верным;
- 4–6 **баллов**: либо решена половина задачи, либо недочеты слишком существенны, чтобы признать решение верным (но само рассуждение основано на здравых идеях);
- 2–3 **баллов**: задача не решена, но в тексте есть идеи, которые могут при должном развитии привести к решению.