Тропическая линейная алгебра

Никита Шапошник, Б05-025
научный руководитель: А. Э. Гутерман

1 Определения

Определение 1.1. *Тропическая алгебра ([2], [3]) — это множество $\mathbb{R}_{\max} = \mathbb{R} \cup \{-\infty\}$ с операциями сложения $\oplus$ и умножения $\odot$:*

$$a \oplus b = \max(a, b)$$
$$a \odot b = a + b.$$

Тропическая алгебра является полукольцом.

Множество матриц размера $n \times m$ над $\mathbb{R}_{\max}$ будем обозначать через $\mathbb{R}_{\max}^{n\times m}$.

2 Матрицы и графы

Определение 2.1. *Вещественная матрица A называется примитивной, если существует натуральное число m такое, что A^m положительна, то есть все числа в ней положительны. При этом наименьшее такое m называется экспонентой матрицы и обозначается через $exp(A)$.*

Теорема 2.2 (ЧТО ТАКОЕ ГРАФ СМЕЖНОСТИ Критерий примитивности матрицы, [10])**.** *Неотрицательная квадратная матрица порядка n над $\mathbb{R}$ примитивна тогда и только тогда, когда граф смежности этой матрицы сильно связен и НОК всех длин замкнутых путей (циклов) равно 1.*

Теорема 2.3 (Виландта, [5])**.** *Если неотрицательная квадратная матрица порядка n над полем вещественных чисел примитивна, то ее экспонента не превосходит число Виландта $Wi(n) = n^2 - 2n + 2$.*

Примитивность и экспонента тропической матрицы определяется так же, как и в вещественном случае, с отличием лишь в том, что в степени матрицы не должно быть нулей тропического полукольца, т.е. $-\infty$.

Рассмотрим тропическую матрицу $A \in \mathbb{R}_{\max}^{n\times n}$. По ней можно построить ориентированный взвешенный граф $\mathcal{G}(A) = (V, E)$, где $V = \{1, 2, \ldots, n\}$, а $E \subseteq V \times V$, где $(i, j) \in E$ тогда и только тогда, когда $a_{ij} \neq -\infty$. Веса рёбер определяются функцией $w : E \to \mathbb{R}$, $(i, j) \mapsto a_{ij}$. Говорят, что A является матрицей смежности графа $\mathcal{G}(A)$.

Наоборот, по взвешенному ориентированному графу аналогично можно построить матрицу смежности. Для этого нужно пронумеровать вершины и поставить в соответствующие ячейки матрицы веса рёбер.

3 Индекс цикличности

Определение 3.1. *Индекс цикличности (или просто цикличность) ориентированного графа $\mathcal{G}$ обозначается через $\sigma_{\mathcal{G}}$ и определяется следующим образом:*

1. *Если $\mathcal{G}$ сильно связен, и $|V(\mathcal{G})| \geq 2$, то цикличность равна НОД всех длин ориентированных циклов в $\mathcal{G}$.*

2. *Если в $\mathcal{G}$ есть только одна вершина (с петлей или без), то $\sigma_{\mathcal{G}} = 1$.*

3. *Если $\mathcal{G}$ не сильно связен, то его цикличность равна НОК цикличностей всех максимальных его сильно связных подграфов.*

Замечание 3.2 (Переформулировка критерия примитивности, см.[10])**.** *Тропическая матрица $A \in \mathbb{R}^{n\times n}_{\max}$ примитивна тогда и только тогда, когда $\mathcal{G}(A)$ сильно связен и его индекс цикличности равен* 1.

Заметим, что в сильно связном графе $\mathcal{G}$ с цикличностью γ любые 2 пути, соединяющий 2 фиксированные вершины, имеют одинаковые длины по модулю γ. Из этого следует, что на множестве $V(\mathcal{G})$ можно ввести отношение эквивалентности: 2 вершины лежат в одном классе эквивалентности тогда и только тогда, когда длина пути от одной к другой кратна γ. Эти классы эквивалентности называются циклическими классами.

Пусть $\mathcal{G} = (V, E)$ — взвешенный ориентированный граф с матрицей смежности $A = (a_{ij}) \in \mathbb{R}^{n\times n}_{\max}$. Пусть C — это ориентированный цикл в $\mathcal{G}$ с весами ребер $a_{i_1}, a_{i_2}, \ldots, a_{i_l}$. Средний вес ребра в C — это тропическое среднее геометрическое весов ребер в C:

$$w_a(C) = \sqrt[\odot l]{a_{i_1} \odot a_{i_2} \odot \cdots \odot a_{i_l}} = \frac{1}{l}(a_{i_1} + a_{i_2} + \cdots + a_{i_l})$$

Определение 3.3. *Ориентированный цикл называется критическим, если у него максимальный средний вес. Критический подграф $\mathcal{G}^c$ графа $\mathcal{G}$ — это объединение всех критических циклов в $\mathcal{G}$.*

Обозначим максимальный средний вес цикла в $\mathcal{G}(A)$ через $\lambda(A)$, т.е.

$$\lambda(A) = \bigoplus_{k=1}^{d} \bigoplus_{i_1,\ldots,i_k} (a_{i_1 i_2} \odot \cdots \odot a_{i_{k-1} i_k})^{\odot 1/k} =$$
$$= \max_{k=1}^{d} \max_{i_1,\ldots,i_k} \frac{(a_{i_1 i_2} + \cdots + a_{i_{k-1} i_k})}{k}$$

4 CSR-декомпозиция

4.1 Необходимые определения

Назовем тропическую матрицу A (или соответствующий ей граф) неразложимой, если граф $\mathcal{G}(A)$ сильно связен, иначе — разложимой.

Назовем тропическую матрицу A (или соответствующий ей граф) полностью разложимой, если в графе $\mathcal{G}(A)$ нет ребер между различными компонентами сильной связности.

Рассмотрим тропическую матрицу $A \in \mathbb{R}^{n\times n}_{\max}$. Тогда звездой Клини матрицы A называется следующая матрица:

$$A^* = \bigoplus_{i=0}^{\infty} A^i = \bigoplus_{i=0}^{n-1} A^i$$

В матрице A^* в ячейке под номером i и j лежит длина оптимального пути от вершины i к вершине j по всему графу, без ограничения на длину пути. Условие $\lambda(A) \leq 0$ необходимо, так как иначе этот ряд расходится: можно идти по циклу с положительным средним весом и улучшать ответ. Так как дважды проходить через одну и ту же вершину не имеет смысла, можно ограничиться первыми n матрицами.

4.2 Матрицы CSR

Рассмотрим неразложимую $A \in \mathbb{R}_{\max}^{n\times n}$. Введем обозначения: $\sigma = \sigma(\mathcal{G}^c(A))$ – индекс цикличности критического подграфа, $M = ((\lambda(A)^- \odot A^\sigma)^*$. Здесь и далее для $a \in \mathbb{R}_{\max}$, $a \neq -\infty$ через a^- будем обозначать обратное по умножению к a, т.е. $a^- = -a$.

Определим матрицы $C, S, R \in \mathbb{R}_{\max}^{n\times n}$ следующим образом:

$$c_{ij} = \begin{cases} m_{ij}, \text{ если } j \in V(\mathcal{G}^c(A)) \\ -\infty, \text{ иначе,} \end{cases} \qquad r_{ij} = \begin{cases} m_{ij}, \text{ если } i \in V(\mathcal{G}^c(A)) \\ -\infty, \text{ иначе,} \end{cases}$$

$$s_{ij} = \begin{cases} \lambda(A)^- \odot a_{ij}, \text{ если } (i,j) \in E(\mathcal{G}^c(A)) \\ -\infty, \text{ иначе.} \end{cases}$$

Если матрицы C, S, R определены по матрице A, будем писать $CS^tR[A]$ для произвольного t.

Теорема 4.1 ([7], [8])**.** *Пусть $A \in \mathbb{R}_{\max}^{n\times n}$ неразложима. Тогда существует неотрицательное целое $T(A)$ такое, что для любого $t \geq T(A)$:*

$$A^t = \lambda(A)^{\odot t} \odot CS^tR[A]. \tag{1}$$

Заметим, что если $\lambda(A) = 0$, то (1) записывается в виде:

$$A^t = CS^tR[A].$$

Замечание 4.2 (Инвариантность относительно умножения на скаляр)**.** *Если $A' = A \odot \mu$, где $\mu \in \mathbb{R}$, то $\lambda(A') = \lambda(A) \odot \mu$ и $CSR[A'] = CSR[A]$.*

Значит, $T(A)$ инвариантно относительно умножения матрицы на скаляр, что позволяет нам без разграничения общности говорить, что $\lambda(A) = 0$.

Утверждение 4.3 (Периодичность, см. [9])**.** *Для любого $t \geq 0$ верно, что $CS^{t+\sigma}R[A] = CS^tR[A]$, где σ — это цикличность $\mathcal{G}^c(A)$. Иначе говоря, последовательность матриц $\{CS^tR[A]\}_{t\geq 0}$ периодична с периодом σ.*

Значит, в силу равенства $A^t = CS^tR$ при $t \geq T(A)$, последовательность матриц A^t при $t \geq T(A)$ является периодической с периодом σ.

Введем несколько новых обозначений:

1. Через $\mathcal{W}^{t,l}(i \to j)$ обозначим множество путей от вершины i к вершине j, имеющих длину t по модулю l;

2. Через $\mathcal{W}(i \xrightarrow{\mathcal{G}} j)$ обозначим множество путей от вершины i к вершине j, проходящих хотя бы через одну вершину из $\mathcal{G}$. Аналогично определяются $\mathcal{W}^t(i \xrightarrow{\mathcal{G}} j)$, $\mathcal{W}^{t,l}(i \xrightarrow{\mathcal{G}} j)$ — граф над стрелкой добавляет ограничение на пути в множестве.

3. Для множества $\mathcal{W}$ через $p(\mathcal{W})$ обозначим максимальный вес пути из множества $\mathcal{W}$.

Утверждение 4.4 ([8])**.** *Если $\lambda(A) = 0$, то верно следующее равенство:*

$$(CS^tR[A])_{ij} = p(\mathcal{W}^{t,\sigma}(i \xrightarrow{\mathcal{G}^c(A)} j)), \tag{2}$$

где σ обозначает цикличность $\mathcal{G}^c(A)$.

Определим несколько характеристик графа.

Обхватом графа $\mathcal{G}$ называется наименьшая длина цикла в $\mathcal{G}$ и обозначается как $g(\mathcal{G})$. Через $\hat{g}(\mathcal{G})$ обозначается максимальный обхват среди всех компонент сильной связности графа $\mathcal{G}$.

Окружностью графа $\mathcal{G}$ называется наибольшая длина цикла в $\mathcal{G}$ и обозначается как $cr(\mathcal{G})$ (от английского circumference).

Максимальную длину простого пути в графе $\mathcal{G}$ будем обозначать через $cd(\mathcal{G})$ (от английского cab-driver's diameter).

ДОБАВИТЬ ССЫЛКУ, учесть, что там было для T_1

Теорема 4.5 (Некоторые верхние оценки $T(A)$, см. [8])**.** *Для любой неразложимой $A \in \mathbb{R}_{\max}^{n\times n}$ имеем:*

1. $T(A) \leq Wi(n)$;
2. $T(A) \leq \hat{g}(n-2) + n$;
3. $T(A) \leq (\hat{g}-1)(cr-1) + (\hat{g}+1)cd$,

где $\hat{g} = \hat{g}(\mathcal{G}^c(A))$, $cr = cr(\mathcal{G}(A))$, а $cd = cd(\mathcal{G}(A))$.

5 Примеры

5.1 Полный граф

Рассмотрим матрицу $A \in \mathbb{R}_{\max}^{n\times n}$, где $a_{ij} = 0$ для любых индексов i, j. Граф $\mathcal{G}(A)$ является полным, веса всех ребер в нём равны 0. Значит, критический подграф $\mathcal{G}^c$ совпадает со всем графом $\mathcal{G}$.

Найдем матрицы C, S, R. Индекс цикличности полного графа $\sigma = 1$ (т.к. в нём есть циклы длины 1), следовательно $C = R = M = A^*$, $S = A$.

Так как для любого положительного t верно, что $A^t = A$, то $A^* = A$ и равенство $A^*A^tA^* = A^t$ выполняется для любого положительного t.

Следовательно, $T = 1$.

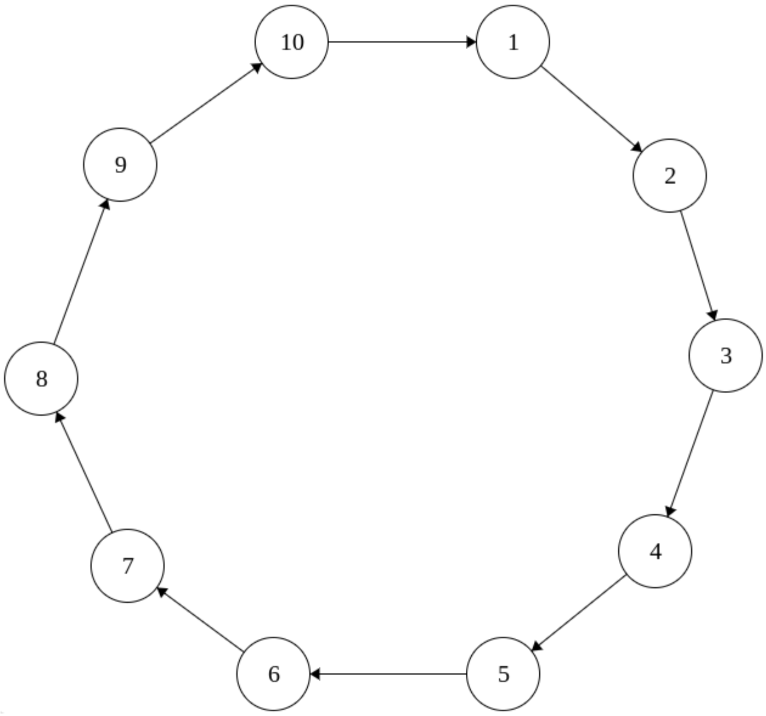
5.2 Односторонний цикл

Рассмотрим матрицу смежности $A \in \mathbb{R}_{\max}^{n\times n}$ одностороннего цикла на n вершинах.

В силу инвариантности границ относительно домножения на скаляр из $\mathbb{R}$ (замечание 4.2), можно рассматривать только тот случай, в котором $\lambda(A) = 0$. Тогда $\mathcal{G}^c(A) = \mathcal{G}(A)$, $\sigma = n$.

$$M = (A^n)^* = E^* = E = \begin{pmatrix} 0 & -\infty & ... & -\infty \\ -\infty & 0 & ... & -\infty \\ ... & ... & ... & ... \\ -\infty & -\infty & ... & 0 \end{pmatrix} = diag(0, 0, \dots, 0)$$

Значит, $C = R = E$, $S = A$, и для любого неотрицательного t верно $CS^tR[A] = A^t$. Следовательно, $T = 0$.



5.3 Двусторонний цикл

Рассмотрим матрицу смежности $A \in \mathbb{R}_{\max}^{n\times n}$ двустороннего цикла на n вершинах, все рёбра в котором имеют нулевой вес, тогда $\lambda(A) = 0$ и $\mathcal{G}^c(A) = \mathcal{G}(A)$. Пронумеруем вершины так, чтобы первый цикл состоял из вершин $1, 2, \dots n$(в порядке обхода), а второй — из $n, n-1, \dots, 1$ (в порядке обхода). Чтобы избежать кратных рёбер, будем работать с $n \geq 3$.

Необходимо рассмотреть два случая: когда n нечётно и когда n чётно.

n нечетно. В этом случае цикличность критического графа $\sigma = 1$, т.е. граф примитивен. Значит, $T(A) = exp(A)$.

Утверждение 5.1. *Экспонента данного графа равна $n-1$.*

Доказательство. Заметим, что в A^{n-2} на главной диагонали стоят $-\infty$: $n-2$ нечётно, поэтому, чтобы вернуться в исходную вершину за $n-2$ шага, надо сменить чётность — пройти весь круг, так как остальные циклы имеют чётную длину. Но цикл имеет длину n, поэтому его пройти не получится. Значит, $exp(\mathcal{G}) \geq n-1$.

Покажем, что $A^{n-1} > -\infty$.

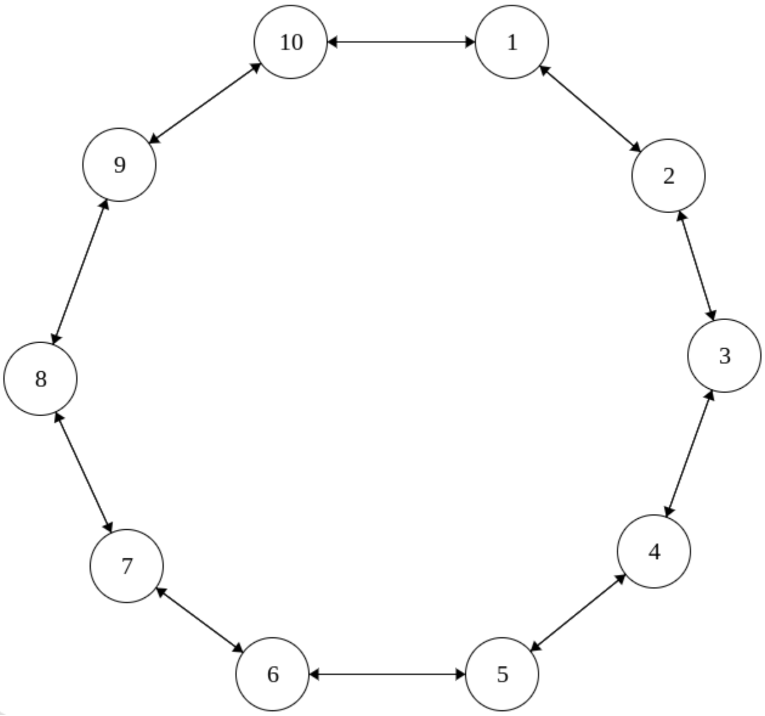
Зафиксируем произвольную вершину v графа. Назовем вершину *четной*, если до нее можно дойти из v за чётное число шагов. Заметим, что тогда все вершины графа четные, так как n нечетно и идти можно как по, так и против часовой стрелки. Наибольшая длина такого пути равна $n-1$. Значит, $A^{n-1} > -\infty$. □

n четно. В этом случае $\sigma = 2$ и граф не примитивен. $C = R = M = (A^2)^*$, $S = A$.

Так как последовательность матриц CS^tR периодична с периодом $\sigma = 2$ (см. [8]), то при $t \geq T(A)$

$$A^t = CS^tR = \begin{cases} (A^2)^*, \text{ если } t \text{ четно.} \\ A \odot (A^2)^*, \text{ если } t \text{ нечетно.} \end{cases}$$

В матрице $(A^2)^*$ небесконечные элементы стоят в клетках (i, j), если вершины i и j находятся на четном расстоянии друг от друга. Наибольшее расстояние между вершинами с одинаковой четностью равно $\frac{n}{2}$. Значит, условие при четном t выполняется при $t \geq \frac{n}{2}$, а при прочих t не выполняется.



В матрице $A \odot (A^2)^*$ небесконечные элементы стоят в клетках (i, j), если вершины i и j находятся на нечетном расстоянии друг от друга. Наибольшее расстояние между вершинами с разной четностью равно $\frac{n}{2} - 1$. Значит, условие при четном t выполняется при $t \geq \frac{n}{2} - 1$, а при прочих t — не выполняется.

Следовательно, $T(A) = \frac{n}{2}$.

5.4 Графы с нулевыми рёбрами

Функция T является обобщением экспоненты на непримитивные графы. Рассмотрим несколько примеров.

Рассмотрим примитивный граф с матрицей смежности A, в котором вес каждого ребра равен 0. Тогда критический подграф совпадает со всем графом: $\mathcal{G}^c(A) = \mathcal{G}(A) = \mathcal{G}$. Индекс цикличности примитивного графа $\sigma = 1$.

По утверждению 4.3 последовательность матриц $CS^tR[A]$ периодична с периодом $\sigma = 1$, то есть в этой последовательности все члены равны. Из утверждения 4.4 и примитивности A следует, что матрица $CS^tR[A]$ целиком состоит из 0 при любом t.

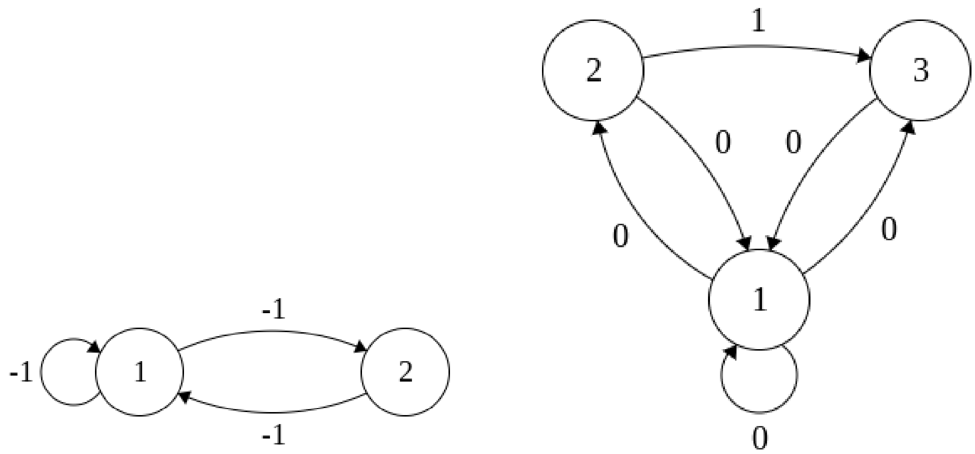
Заметим, что в любой степени матрицы A её элементы будут принимать только два значения: $-\infty$ и 0. Из определения $T(A)$ следует, что $A^t = CS^tR[A]$ тогда и только тогда, когда $t \geq T(A)$. Значит, матрица A^t не содержит $-\infty$ тогда и только тогда, когда $t \geq T(A)$. Значит, $T(A) = exp(A)$, если A примитивна.

Это приводит нас к более общему утверждению.

Утверждение 5.2. *Рассмотрим примитивную матрицу A, у которой $\mathcal{G}(A)$ совпадает со своим критическим подграфом, $\lambda(A) = 0$. Если для двух произвольных фиксированных вершин u и v верно, что все пути из u в v имеют одинаковый вес, то $T(A) = exp(A)$.*

Доказательство. В силу условия на одинаковый вес между любыми двумя вершинами матрицы вида $CS^tR[A]$ принимают только одно значение (по утверждению 4.4), а значение конкретной ячейки матрицы A^t либо равно $-\infty$, либо совпадает с соответстующей ячейкой $CS^tR[A]$. Значит, условие $A^t = CS^tR[A]$ равносильно условию $A^t > -\infty$. Следовательно, $T(A) = exp(A)$. $\square$

Заметим, что обратное утверждение неверно. Рассмотрим следующие графы:



(a) (b)

В обоих графах экспонента совпадает с T (в обоих графах экспонента равна 2), но в графе (a) максимальный средний вес цикла равен -1, а в графе (b) критический подграф не совпадает со всем графом.

Рассмотрим неразложимую матрицу A такую, что в $\mathcal{G}(A)$ все рёбра имеют нулевой вес. Пусть его индекс цикличности равен σ.

6 Разные ромашки

Определение 6.1. *Назовем ромашкой граф, состоящий из нескольких пересекающихся по одной вершине циклов.*

Здесь и далее будем рассматривать графы-ромашки, состоящие из циклов длины, кратной σ, все рёбра в которых имеют вес 0.

Определение 6.2. *Ромашку, состоящую из циклов длины $a_1\sigma, a_2\sigma, \ldots, a_n\sigma$, где числа $a_1, \ldots, a_n$ взаимно просты в совокупности, $a_1 \leq a_2 \leq \cdots \leq a_n$ назовем $(a_1, \ldots, a_n; \sigma)$-ромашкой.*

Границу T, определенную для такой ромашки, будем обозначать через $T(a_1, \ldots, a_n; \sigma)$.

Заметим, что индекс цикличности такой ромашки равен σ и всего в ней $N = \sum\limits_{i=1}^{n} a_i\sigma - n + 1$ вершин. Пусть вершина, в которой пересекаются все циклы, имеет номер 1. Пронумеруем вершины в порядке следующего обхода: начнем в вершине 1, далее пройдём по первому циклу, затем — по второму, и так далее до цикла с номером N (не изменяя номер у вершины 1).

Во всех примерах матрицу смежности рассматриваемого графа будем обозначать через $A \in \mathbb{R}_{\max}^{n \times n}$, а через C, S, R будем обозначать матрицы C, S, R, построенные по матрице A.

Теорема 6.3. $T(a_1, \ldots, a_n; \sigma) = (T(a_1, \ldots, a_n; 1) + 1)\sigma - 1$.

Доказательство. Обозначим граф, соответствующий $(a_1, \ldots, a_n; 1)$-ромашке через $\mathcal{G}$, а граф, соответствующий $(a_1, \ldots, a_n; \sigma)$-ромашке — через $\mathcal{G}_\sigma$. Граф $\mathcal{G}_\sigma$ получается из графа $\mathcal{G}$ разделением каждого ребра на σ более мелких рёбер. Вершины $\mathcal{G}_\sigma$, лежащие в одном циклическом классе с вершиной 1, будем называть начальными. Для краткости будем обозначать $T(a_1, \ldots, a_n; 1)$ через T^1, а $T(a_1, \ldots, a_n; \sigma)$ — через T^σ.

Покажем, что $T^\sigma > (T^1+1)\sigma - 2$. В $\mathcal{G}$ есть 2 вершины, между которыми нет пути длины T^1-1. Значит, в $\mathcal{G}_\sigma$ между соответствующими начальными вершинами нет пути длины $(T^1-1)\sigma$. Обозначим эти вершины через u и v. Но тогда между вершинами $\hat{u}$ и $\hat{v}$ не будет пути длины $(T^1-1)\sigma + 2(\sigma-1) = (T^1+1)\sigma - 2$, где $\hat{u}$ получается, если отойти от u на $\sigma-1$ шаг вперёд, а $\hat{v}$ — от вершины v на $\sigma-1$ шаг назад (обе новые вершины существуют, так как любая вершина в $\mathcal{G}$ лежит в цикле). Значит, $T^\sigma \geq (T^1+1)\sigma - 1$.

Покажем, что $T^\sigma \geq (T^1+1)\sigma - 1$. Для этого нужно доказать, что между любыми двумя вершинами u и v графа $\mathcal{G}_\sigma$ есть путь длины $(T^1+1)\sigma-1$ от u до v. Путь длины $(T^1+1)\sigma-1$ от u до v состоит из трех частей: путь от u до ближайшей начальной вершины, путь между начальными вершинами, и путь от ближайшей начальной вершины до v. Суммарная длина первой и третьей частей не превосходит $2\sigma-2$, значит, длина второй части не меньше $(T^1-1)\sigma+1$. Но длина пути между двумя начальными вершинами должна быть кратна σ, поэтому длина второй части не меньше $T^1 \cdot \sigma$. Но, по определению T^1, между любыми начальными вершинами есть путь длины $T^1 \cdot \sigma$. Значит, $T^\sigma \geq (T^1+1)\sigma - 1$, и утверждение доказано. □

Таким образом, при расчёте границы T для произвольной ромашки достаточно посчитать искомую границу при $\sigma = 1$, а затем получить ответ по формуле из утверждения 6.3.

Замечание 6.4. *При $\sigma = 1$ $(a_1, \ldots, a_n; 1)$-ромашка примитивна, все рёбра в ней имеют нулевой вес. Значит, граница T ромашки совпадает её экспонентой.*

Введём вспомогательную функцию P:

Определение 6.5. *Для взаимно простых в совокупности натуральных чисел $a_1 \leq \cdots \leq a_n$ обозначим через $P(a_1, \ldots, a_n)$ минимальное целое неотрицательное число, удовлетворяющее следующему свойству: любое $p \geq P(a_1, \ldots, a_n)$ выражается в виде линейной комбинации чисел $a_1, \ldots, a_n$ с целыми неотрицательными коэффициентами $\lambda_1, \ldots, \lambda_n$, то есть*

$$p = a_1\lambda_1 + \ldots a_n\lambda_n \tag{3}$$

. Число, выражающееся в виде линейной комбинации чисел $a_1, \ldots, a_n$ с целыми неотрицательными коэффициентами, назовём выразимым.

Здесь и далее под линейной комбинацией будем понимать линейную комбинацию с целыми неотрицательными коэффициентами.

Теорема 6.6. $T(a_1, \ldots, a_n; 1) = P(a_1, \ldots, a_n) + 2a_n - 2.$

Доказательство. Предположим, что в $(a_1, \ldots, a_n; 1)$-ромашке между любыми двумя вершинами существует путь длины t. Рассмотрим две произвольные вершины u и v. Любой путь длины хотя бы $a_n - 1$ проходит через вершину 1, и $t \geq a_n - 1$. Поэтому путь длины t от u до v состоит из трёх частей: пути от u до 1 (обозначим длину этой части через $\hat{u}$), λ_i циклов длины a_i для $i = 1 \ldots n$, и пути от 1 до v (обозначим длину этой части через $\hat{v}$). Тогда имеет место равенство:

$$t = \hat{u} + a_1\lambda_1 + \cdots + a_n\lambda_n + \hat{v} \quad \Longleftrightarrow \quad t - \hat{u} - \hat{v} = a_1\lambda_1 + \cdots + a_n\lambda_n.$$

Сумма $\hat{u} + \hat{v}$ принимает любые значения от 0 до $2a_n - 2$ (так как $0 \leq \hat{u}, \hat{v} \leq a_n - 1$). Следовательно, для любого $t - 2a_n + 2 \leq p \leq t$ должны существовать коэффициенты $\lambda_1, \ldots, \lambda_n$, удовлетворяющие уравнению

$$p = a_1\lambda_1 + \cdots + a_n\lambda_n. \tag{4}$$

При $t < P(a_1, \dots, a_n) + 2a_n - 2$ минимальное значение p не превосходит $P(a_1, \dots, a_n) - 1$, и, по определению $P(a_1, \dots, a_n)$, при наименьшем значении p уравнение 4 решений не имеет — противоречие с наличием пути между u и v.

Напротив, при $t \geq P(a_1, \dots, a_n) + 2a_n - 2$ наименьшее значение p не меньше $P(a_1, \dots, a_n)$, и, в силу определения $P(a_1, \dots, a_n)$, коэффициенты λ_i найдутся для любого возможного значения p.

Значит, $T(a_1, \dots, a_n; 1) = P(a_1, \dots, a_n) + 2a_n - 2$. □

Следствие 6.7 (Корректность функции P)**.** *Функция P определена корректно: её значение существует для любых возможных аргументов.*

Доказательство. Рассмотрим $(a_1, \dots, a_n; 1)$-ромашку. По замечанию 6.4 этот граф примитивен и, следовательно, имеет экспоненту, которая, в свою очередь, совпадает с границей T для данной ромашки. По формуле из теоремы 6.6 имеем $P(a_1, \dots, a_n) = T(a_1, \dots, a_n; 1) - 2a_n + 2$. □

Утверждение 6.8 (Свойства функции P)**.**

1. *Если $a_1 = 1$, то $P(1, \dots, a_n) = 0$.*
2. *$P(a_1, \dots, a_n) \leq P(a_{i_1}, a_{i_2}, \dots, a_{i_k})$, где $1 \leq i_1 < i_2 < \dots < i_k \leq n$ — возрастающая последовательность индексов.*
3. *$P(a_1, \dots, a_n) = P(b_1, \dots, b_m)$, где набор $b_1, \dots, b_m$ получается из набора $a_1, \dots, a_n$ удалением повторяющихся элементов.*
4. *Если a_j делится на a_i, то $P(a_1, \dots, a_n) = P(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n)$.*
5. *Если a_j представляется в виде линейной комбинации меньших элементов, то $P(a_1, \dots, a_n) = P(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n)$.*

Доказательство. 1) Действительно, если $a_1 = 1$, то любое неотрицательное число k выражается как $1 \cdot k$. Следовательно, $P = 0$.

2) Свойство следует из следующего факта: сумма $a_{i_1}\lambda_{i_1} + \dots + a_{i_k}\lambda_{i_k}$ является частным случаем суммы $a_1\lambda_1 + \dots + a_n\lambda_n$.

3) При приведении подобных членов в сумме $a_1\lambda_1 + \dots + a_n\lambda_n$ получается корректная сумма $b_1\mu_1 + \dots b_m\mu_m$. С другой стороны, сумма $b_1\mu_1 + \dots b_m\mu_m$ является корректной суммой вида $a_1\lambda_1 + \dots + a_n\lambda_n$.

4) Очевидно, что любая сумма $a_1\lambda_1 + \dots + a_{j-1}\lambda_{j-1} + a_{j+1}\lambda_{j+1} + \dots + a_n\lambda_n$ является суммой вида $a_1\lambda_1 + \dots + a_n\lambda_n$, где $\lambda_j = 0$. С другой стороны, заменив a_j на $a_i \cdot \frac{a_j}{a_i}$, можно избавиться от слагаемого $a_j\lambda_j$ в сумме $a_1\lambda_1 + \dots + a_n\lambda_n$, что доказывает утверждение.

5) Доказетельство этого свойства аналогично предыдущему. □

Утверждение 6.9. $P(a, b) = (a-1)(b-1)$.

Доказательство. Покажем, что $p = ab - a - b \neq ma + nb$ для любых целых неотрицательных m, n.

Предположим противное. Тогда:

$$ab - a - b = am + bn \iff ab = (m+1)a + (n+1)b$$

В силу взаимной простоты a и b получим, что $n + 1 \,\vdots\, a$, и $m + 1 \,\vdots\, b$. Тогда, в силу того, что $m, n \geq 0$, имеем 2 случая:

$$\begin{cases} n+1 = a \\ m+1 = 0 \end{cases} \quad \begin{cases} n+1 = 0 \\ m+1 = b. \end{cases}$$

В обоих случаях получаем противоречие. Следовательно, $P(a, b) \geq (a-1)(b-1)$.

Теперь покажем, что $P(a, b) \leq ab + b - a - 1$. Для любого $p \geq ab - b - a + 1$ решим уравнение:

$$am + bn = p$$

Так как a и b взаимно просты, числа из набора $0, b, 2b, \ldots, (a-1)b$ дают все a остатков по модулю a. Значит, существует единственное $0 \leq n \leq a - 1$, что $bn \equiv p \pmod a$, причём $p - bn \geq 0$, так как

$$p - bn \geq ab - b - a + 1 - (a-1)b = -a + 1 > -a \Longrightarrow p - bn \geq 0.$$

Значит, $m = \frac{p-bn}{a} \geq 0$.

Таким образом, нами были найдены целые $m \geq 0$, $n \geq 0$. Следовательно, $P(a, b) = (a-1)(b-1)$. □

Следствие 6.10. $T(a, b; \sigma) = (ab + b - a)\sigma - 1$.

Утверждение 6.11. $P(2, a, b) = \begin{cases} P(2, b) = b - 1, & \textit{если } a \textit{ чётно,} \\ P(2, a) = a - 1, & \textit{иначе.} \end{cases}$

Доказательство. Первый случай следует из свойства 4 утверждения 6.8.

Разберём второй случай: a нечётно. Неравенство $P(2, a, b) \leq P(2, a)$ следует из свойства 2 утверждения 6.8. Докажем обратное неравенство: необходимо показать, что с помощью слагаемых $2, a, b$ невозможно получить сумму $a - 2$. Действительно, из трёх слагаемых можно использовать только одно: 2. Но $a - 2$ нечётно — противоречие. Следовательно, $P(2, a, b) = P(2, a)$. □

Следствие 6.12. $T(2, a, b; \sigma) = \begin{cases} T(2, b; \sigma) = (3b - 2)\sigma - 1, & \textit{если } a \textit{ нечётно,} \\ (2b + a - 2)\sigma - 1, & \textit{иначе.} \end{cases}$

Утверждение 6.13. $P(3, a, b) = \begin{cases} P(3, b) = 2(b - 1), & \textit{если } a \vdots 3, \\ b - 2, & \textit{если } a \not\vdots\, 3, a + b \vdots 3 \textit{ и } b < P(3, a) = 2a - 2, \\ P(3, a) = 2(a - 1), & \textit{иначе.} \end{cases}$

Доказательство. Первый случай следует из свойства 4 утверждения 6.8.

Разберём второй случай. Покажем, что $P(3, a, b) \geq b - 2$. Предположим противное. Тогда число $b - 3$ должно выражаться в виде линейной комбинации $2, a$ и b:

$$b - 3 = 3\lambda_1 + a\lambda_2 + b\lambda_3$$

Тогда $\lambda_3 = 0$ и $\lambda_2 \leq 1$. При $\lambda_2 = 0$ имеем $b = 3\lambda_1 + 3 \vdots 3$. При $\lambda_2 = 1$ имеем $b - a = 3\lambda_1 + 3 \vdots 3$. В обоих случаях $a \vdots 3$, так как $a + b \vdots 3$, что противоречит условию второго случая. Следовательно, $P(3, a, b) \geq b - 2$.

Докажем обратное неравенство: для любого $p \geq b - 2$ решим уравнение:

$$p = 3\lambda_1 + a\lambda_2 + b\lambda_3$$

Так как в правой части есть слагаемое $3\lambda_1$, то достаточно решить уравнение для $p = b - 2$, $p = b - 1$ и $p = b$ — тогда линейные комбинации для больших p получатся увеличением λ_1.

- $p = b - 2$. Если $b \equiv 2 \pmod 3$, то $\lambda_1 = \frac{b-2}{3}, \lambda_2 = \lambda_3 = 0$.

  Если $b \equiv 1 \pmod 3$, то $a \equiv 2 \pmod 3$, $b - 2 = (b - a - 2) + a$ и $\lambda_1 = \frac{b-a-2}{3}, \lambda_2 = 1, \lambda_3 = 0$.

- $p = b - 1$. Если $b \equiv 1 \pmod 3$, то $\lambda_1 = \frac{b-1}{3}, \lambda_2 = \lambda_3 = 0$.

  Если $b \equiv 2 \pmod 3$, то $a \equiv 1 \pmod 3$, $b-2 = (b-a-1)+a$ и $\lambda_1 = \frac{b-a-1}{3}, \lambda_2 = 1, \lambda_3 = 0$.

- $p = b$. Тогда $\lambda_1 = \lambda_2 = 0, \lambda_3 = 1$.

Таким образом, $P(3, a, b) = b - 2$.

Перейдём к третьему случаю: если $b \geq P(3, a)$, то наличие слагаемого $b\lambda_3$ не повлияет на значение функции P: если некое p выражается в виде линейной комбинации с участием b, то $p \geq P(3, a)$ и, следовательно, выражается и без участия b. Следовательно, $P(3, a, b) = P(a, b)$.

Рассмотрим последний случай: $a \not\vdots\, 3,\ a + b \not\vdots\, 3,\ b < P(3, a)$. Неравенство $P(3, a, b) \geq P(a, b)$ следует из свойства 2 утверждения 6.8. Докажем обратное неравенство. Для этого покажем, что следующее уравнение не имеет решений:

$$2a - 3 = 3\lambda_1 + a\lambda_2 + b\lambda_3$$

Заметим, что $\lambda_3 = 0$, так как $b < 2a - a$. Также, $\lambda_2 \leq 1$. Тогда $(2 - \lambda_2)a = 3\lambda_1 + 3 \,\vdots\, 3$ — противоречие с $a \not\vdots\, 3$. Значит, $P(3, a, b) = P(a, b)$. $\square$

Следствие 6.14. $T(3, a, b; 1) = \begin{cases} T(3, b; 1) = 4b - 4, & \text{если } a \,\vdots\, 3, \\ 3b - 4, & \text{если } a \not\vdots\, 3, a + b \,\vdots\, 3 \text{ и } m < 2a - 2, \\ 2a + 2b - 4, & \text{иначе.} \end{cases}$

Алгоритм вычисления функции P

Рассмотрим массив M длины a_1, где в $M[i]$ лежит минимальное выразимое число, сравнимое с i по модулю a_1. Заметим, что $M[0] = 0$ и что $M[i] \equiv i \pmod{a_1}$.

Утверждение 6.15. $P(a_1, \ldots, a_n) = \max\limits_{i=0}^{a_0-1} M[i] - a_1 + 1$.

Доказательство. Пусть $\max\limits_{i=0}^{a_0-1} M[i] - a_1 + 1 = M[k] - a_1 + 1$.

Выразимость $M[k] - a_1$ вела бы к противоречию с определением массива M, так как $M[k] - a_1 \equiv M[k] \pmod{a_1}$. Значит, $P(a_1, \ldots, a_n) \geq \max\limits_{i=0}^{a_0-1} M[i] - a_1 + 1$.

Заметим, что если произвольное x выразимо, то и число $x + a_1$ выразимо. Из этого следует, что любое число, сравнимое с i по модулю a_1 и не меньшее $M[i]$ выразимо. Значит, все числа, начиная с $M[k] - a_1 + 1$ выразимы — иначе $M[k]$ было бы не максимальным числом в массиве M.

Следовательно, $P(a_1, \ldots, a_n) = \max\limits_{i=0}^{a_0-1} M[i] - a_1 + 1$. $\square$

Используя массив M, можно легко посчитать $P(4, a, b)$. Здесь и далее через $x \ rem \ y$ будем обозначать остаток при делении x на y.

Утверждение 6.16 (Формула для $P(4, a, b)$)**.**

1. *$a \,\vdots\, 4, b \not\vdots\, 2$. Тогда $P(4, a, b) = P(4, b)$.*

2. *$a \not\vdots\, 2, b \,\vdots\, 4$, или $0 \not\equiv a \equiv b \pmod 4$, или $a \not\vdots\, 2, b \geq P(4, a)$. Тогда $P(4, a, b) = P(4, a)$.*

3. *$a \equiv 2 \pmod 4), b \not\vdots\, 2$. Тогда $P(4, a, b) = a + b - 3$.*

4. $a \not\vdots 2, b \equiv 2 \pmod 4$. *Тогда*

$$P(4,a,b) = \begin{cases} a+b-3, & \text{если } b < 2a \\ 3a-3, & \text{иначе.} \end{cases}$$

5. $a, b \not\vdots 2$, $a+b \vdots 4, b < P(4,a)$. *Тогда*

$$P(4,a,b) = \begin{cases} 2a-3, & \text{если } b \le 2a \\ b-3, & \text{иначе.} \end{cases}$$

Доказательство. Из свойства 4 утверждения 6.8 можно вывести случай $a \vdots 4, b \not\vdots 2$ и случай $a \not\vdots 2, b \vdots 4$, а из свойства 5 того же утверждения — случай $0 \not\equiv a \equiv b \pmod 4$.

Во всех остальных случаях посчитаем массив M, и по утверждению 6.15 найдём ответ. Заметим, $M[0]$ всегда равен 0.

Докажем случай $a \not\vdots 2, b \ge P(4,a)$. Тогда $M[a\ rem\ 4] = a$, $M[2] = 2a$, и $M[4 - a\ rem\ 4] = 3a$ — число b слишком большое, чтобы повлиять на этот массив. Таким образом, максиму этого массива равен $3a$, и ответом будет число $3a - 3 = P(4,a)$.

Разберём случай $a \equiv 2 \pmod 4), b \not\vdots 2$. Заметим, что $M[2] = a$, $M[b\ rem\ 4] = b$, $M[4 - b\ rem\ 4] = a + b$. Максимум этого массива — $a + b$, поэтому ответ равен $a + b - 3$.

Разберём случай $a \not\vdots 2, b \equiv 2 \pmod 4$. Тогда $M[a\ rem\ 4] = a$. На место $M[2]$ есть два кандидата: $2a$ и b. Если $b < 2a$, то $M[2] = b$, и иначе — $2a$. Далее, для $M[4 - a\ rem\ 4]$ имеем два варианта: $3a$ и $a + b$, и если $b < 2a$, то $M[4 - a\ rem\ 4] = a + b$, и иначе — $3a$. Таким образом, если $b < 2a$, то ответ равен $a + b - 3$, а иначе — $3a - 3 = P(4,a)$.

Разберём последний случай: $a, b \not\vdots 2, a + b \vdots 4, b < 3a - 3$. Тогда $M[a\ rem\ 4] = a$, $M[b\ rem\ 4] = b$ и $M[2] = 2a$. В зависимости от относительного расположения $2a$ и b имеем 2 различных возможных максимума массива M, откуда, по утверждению 6.18 находим ответ. □

Приведём алгоритм, вычисляющий функцию P. На вход ему подаётся число n числа $a_1, \ldots, a_n$.

Алгоритм вычисляет массив M, а затем, по формуле из леммы 6.15, вычисляет ответ на поставленную задачу. Массив M вычисляется постепенно: изначально в каждой ячейке $M[i]$ значения ∞ из $\mathbb{R}_{\min}$ — это значит, что пока не было найдено ни одного выразимого числа, сравнимого с i по модулю a_1. Если при последующем переборе было найдено некоторое p, сравнимое с i по модулю a_1 и меньшее $M[i]$, то необходимо перезаписать в ячейку $M[i]$ значение p.

Перебор начинается с рассмотрения всех линейных комбинаций с одним слагаемым (здесь и далее через количество слагаемых будем обозначать количество ненулевых коэффициентов λ_i в линейной комбинации вида 3). Затем будем перебирать линейные комбинации, на каждом шаге увеличивая максимальное количество слагаемых вдвое. Таким образом, необходимо сделать $\lceil log_2 n \rceil$ итераций, где $\lceil x \rceil$ — это округление числа x вверх.

Алгоритм 6.17.

1. *Создадим массив M длины a_1 содержащий числа из $\mathbb{R}_{\min}$. Запишем во все ячейки значения ∞.*

2. *На нулевой итерации переберём все линейные комбинации с одним слагаемым. Для этого для каждого a_i и для каждого множителя $0 \le k < a_1$ проверим, можем ли мы улучшить ответ: сравним $a_i^{\odot k} = a_i \cdot k$ с $M[a_i \cdot k \text{ } rem \text{ } a_1]$, и если в массиве записано большее число, то улучшим ответ: запишем в ячейку $a_i \cdot k \text{ } rem \text{ } a_1$ значение $a_i^{\odot k} = a_i \cdot k$.*

3. *На каждой следующей итерации будем перебирать все пары ячеек $M[i]$ и $M[j]$ и пытаться улучшить ответ: сравним $M[(i+j) \text{ } rem \text{ } a_1]$ с $M[i] \odot M[j]$ (т.е. $M[i] + M[j]$, если оба эти числа меньше ∞, и ∞ иначе), и если в массиве записано большее число, то улучшим ответ: запишем в ячейку $(i+j) \text{ } rem \text{ } a_1$ значение $M[i] \odot M[j]$.*

4. *Всего необходимо сделать $\lceil log_2(n) \rceil + 1$ итераций. После этого ответом будет $\bigoplus\limits_{i=0}^{a_0-1} M[i] - a_1 + 1 = \max\limits_{i=0}^{a_0-1} M[i] - a_1 + 1$.*

Для доказательства корректности докажем следующее утверждение.

Лемма 6.18. *После итерации с номером d в ячейке $M[i]$ лежит минимальное число, сравнимое с i по модулю a_1, которое может быть представлено в виде линейной комбинации с не более чем 2^d слагаемыми, или ∞, если такого числа не существует.*

Доказательство. Докажем утверждение по индукции.

База: $d = 0$. В шаге 1 перебираются все линейные комбинации вида $a_j \cdot k$, где $0 \le k < a_1$. Рассмотрим линейную комбинацию, которую мы не перебрали: $a_i \cdot m$. Так как мы не перебрали эту комбинацию, то $m \ge a_1$. Но тогда $a_i \cdot m \equiv a_i \cdot (m - a_1) \pmod{a_1}$ и $a_i \cdot m > a_i \cdot (m - a_1) \ge 0$ — эта линейная комбинация не может улучшить ответ. Значит, база верна.

Докажем переход. Предположим, утверждение доказано для $d-1$, докажем его для d. Обозначим массив M в состоянии до итерации с номером d через M'.

Рассмотрим произвольную ячейку $M[i]$, в которой записано число, меньшее ∞. Тогда существуют два индекса j и k такие, что $i = (j+k) \text{ } rem \text{ } a_1$ и $M[i] = M'[j] + M'[k]$. По предположению индукции в каждой ячейке массива M' лежит число, которое может быть представлено в виде линейной комбинации с не более чем 2^{d-1} слагаемыми. Значит, в $M[i]$ лежит число, представимое в виде линейной комбинации с не более чем 2^d слагаемыми. По предположению индукции $M[i] = M'[j] + M'[k] \equiv j + k \equiv i \pmod{a_1}$.

Осталось доказать минимальность $M[i]$. Предположим противное: пусть существует число $x < M[i]$, сравнимое с i по модулю a_1 и представимое в виде линейной комбинации с не более чем 2^d слагаемыми. Тогда эту комбинацию можно разбить на две меньших, в каждой из которых будет не более 2^{d-1} слагаемых. Обозначим суммы этих линейных комбинаций через S_1 и S_2. Пусть $S_1 \equiv j \pmod{a_1}$, а $S_2 \equiv k \pmod{a_1}$.

Тогда $S_1 + S_2 = x < M[i] \le M'[j] + M'[k]$ и или $S_1 < M'[j]$, или $S_2 < M'[k]$. В обоих случаях имеем противоречие с предположением индукции. Значит, предположение индукции верно и для d, что и требовалось доказать. □

Утверждение 6.19. *Алгоритм 6.17 корректен. Время его работы — $O(n \cdot a_1 + a_1^2 \cdot log \text{ } n)$. Объем затраченной памяти — $O(a_1)$.*

Доказательство. Докажем асимптотики. Первый шаг работает за $O(a_1)$, второй — за $O(a_1 \cdot n)$ (надо перебрать все $1 \le j \le n$ и все $0 \le k < a_1$). Третий работает за $O(a_1^2 \cdot log \text{ } n)$, так как всего $O(log \text{ } n)$ итераций, в каждой из которых надо перебрать пары (i, j), где $0 \le i, j \le a_1$. Четвертый — за $O(a_1)$. Итоговая сложность алгоритма: $O(n \cdot a_1 + a_1^2 \cdot log \text{ } n)$.

Память тратится только на массив M длины a_1. Значит, алгоритм требует $O(a_1)$ памяти.

Докажем корректность. По лемме 6.18 после итерации с номером d в ячейках массива M лежит информация об оптимальных линейных комбинациях с не более чем 2^d слагаемыми. Следовательно, после итерации с номером $\lceil log_2(n) \rceil$ в массиве M лежит информация об оптимальных линейных комбинациях из n слагаемых, то есть массив M будет наконец посчитан.

Во время работы алгоритма каждая ячейка массива M изменит своё значение хотя бы раз: это следует из корректности функции P. Значит, после последней итерации в массиве M не останется ∞.

Далее ответ может быть получен по лемме 6.15. □

На моём компьютере при $n = 100, a_1 = 100$ алгоритм ни разу не показывал время, большее 0.2 с. При $n = 1000, a_1 = 1000$ алгоритм работал не дольше 0.3 с. При $n = 10000, a_1 = 10000$ алгоритм работает существенно медленнее: в районе 40 с.

6.1 Верхние оценки функции P

Утверждение 6.20. *Функция $P(a_1, \ldots, a_n)$ оценивается сверху следующими функциями:*

1. $Wi(N) - 2a_n + 2$,
2. $(a_1 + 1)N - 2a_1 - 2a_n + 2$,
3. $(a_1 - 1)(a_n - 1) + a_1(2a_n - 2)$,

где $N = \sum_{i=1}^{n} a_i - n + 1$ *— количество вершин в $(a_1, \ldots, a_n; 1)$-ромашке.*

Доказательство. По замечанию 6.4 граница T данной ромашки совпадает с её экспонентой, которая по теореме 4.5 оценивается сверху числом Виландта от количества вершин $Wi(N)$, функцией $\hat{g}(N-2) + N$ и функцией $(\hat{g} - 1)(cr - 1) + (\hat{g} + 1)cd$.

Обхват $(a_1, \ldots, a_n)$-ромашки равен a_1, её окружность равна a_n, а длина наибольшего простого пути не превышает $2a_n - 2$.

Далее достаточно применить теорему 6.6. □

References

[1] Imre Simon *On semigroups of matrices over the tropical semiring* Theoretical Informaties and Applications (Tome 28 (1994) no. 3-4, pp. 277-294)

[2] Semere Tsehaye Tesfay. *A Glance at Tropical Operations and Tropical Linear Algebra* Eastern Illinois University, 2015.

[3] David Speyer, Bernd Sturmfels. *Tropical Mathematics* Mathematics Magazine, vol. 82, №3, June 2009.

[4] Ю.М. Волченко *Max-plus алгебра и ее применение*, декабрь 2017

[5] Hans Schneider. *Wielandt's proof of the exponent inequality for primitive nonnegative matrices* Department of Mathematics, University of Wisconsin at Madison, 2002.

[6] Alexander Guterman, Elena Kreines, and Carsten Thomassen. *Linear transformations of tropical matrices preserving the cyclicity index* Special Matrices Volume 9, 2021.

[7] Arthur Kennedy-Cochran-Patrick, Glenn Merlet, Thomas Nowak, Sergei Sergeev. *New bounds on the periodicity transient of the powers of a tropical matrix: Using cyclicity and factor rank* Linear Algebra and its Applications, 2020

[8] Glenn Merlet, Thomas Nowak, Sergei Sergeev.
https://www.sciencedirect.com/science/article/pii/S0024379514004777

[9] Sergei Sergeev, Hans Schneider. *CSR expansions of matrix powers in max algebra* Transactions of the American Mathematical Society, December 2009

[10] Brualdi RA, Ryser HJ. *Combinatorial matrix theory.* Cambridge: Cambridge University Press; 1991. (Encyclopedia of mathematics and its applications; 39).