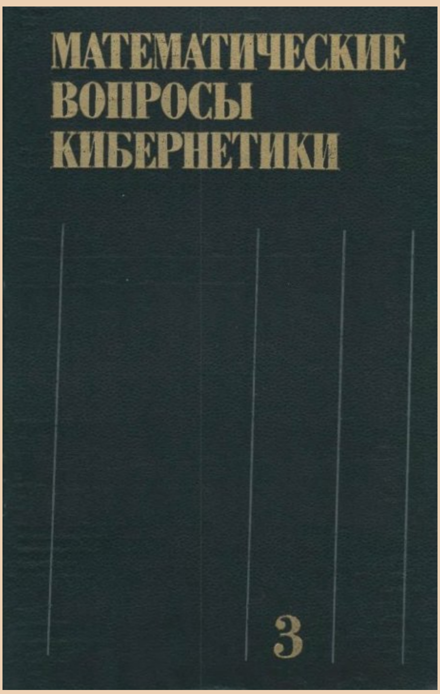

С. Е. Кузнецов,
Н. Н. Нурмеев,
Ф. И. Салимов

Задача о минимальном имплицирующем векторе

Рекомендуемая форма библиографической ссылки: Кузнецов С. Е., Нурмеев Н. Н., Салимов Ф. И. Задача о минимальном имплицирующем векторе // Математические вопросы кибернетики. Вып. 3. — М.: Наука, 1991. — С. 199–216. URL: http://library.keldysh.ru/mvk.asp?id=1991-199

ЗАДАЧА О МИНИМАЛЬНОМ ИМПЛИЦИРУЮЩЕМ ВЕКТОРЕ

С. Е. КУЗНЕЦОВ, Н. Н. НУРМЕЕВ, Ф. И. САЛИМОВ

(КАЗАНЬ)

Введение

Любую стохастическую $(m \times n)$-матрицу $\mathbf{P}$ можно представить в виде выпуклой комбинации вырожденных стохастических матриц, а именно:

$$\mathbf{P} = \sum_{i=1} p_i \mathbf{G}_i, \tag{1}$$

где $\mathbf{G}_i \in \mathcal{G}$, $i = 1, 2, \ldots, L$, $\mathcal{G}$ — множество стохастических $(m \times n)$-матриц из нулей и единиц, $\tilde{p} = (p_1, p_2, \ldots, p_L)$ — стохастический вектор, $p_i > 0$, $i = 1, 2, \ldots, L$. Вектор $\tilde{p}$ называется *имплицирующим для матрицы* $\mathbf{P}$. Понятие имплицируемости для векторов было введено Р. Г. Бухараевым. Им же [1] была поставлена задача нахождения минимального (по длине L) имплицирующего вектора. Эта задача возникла в теории вероятностных автоматов.

Рассмотрим автономный конечный вероятностный автомат с детерминированной функцией выхода, вероятности переходов состояний которого описываются матрицей $\mathbf{P}$ (матрицей переходов однородной цепи Маркова). Применяя к матрице $\mathbf{P}$ разложение (1), можно свести задачу синтеза этого автомата к задаче синтеза бернуллиевского источника случайности, выдающего букву x_i из некоторого алфавита X с вероятностью p_i, и детерминированного автомата, матрица переходов которого для входного символа x_i из X описывается матрицей $\mathbf{G}_i$. Пионерской в этом направлении была работа [15]. В дальнейшем усилиями ряда исследователей [2, 3, 11, 16] этот подход был распространен на конечные вероятностные автоматы общего вида. Поскольку число состояний детерминированного автомата совпадает с числом состояний исходного конечного вероятностного автомата, то оптимизация такой схемы связана с поиском разложения (1) с минимальным значением параметра L.

Эта задача имеет также и самостоятельный математический интерес как задача представления точки выпуклого многогранника через выпуклую комбинацию минимального числа его крайних точек. Близкой к задаче о минимальном имплицирующем векторе является классическая задача разложения дважды стохастической матрицы в выпуклую комбинацию матриц перестановок (теорема Биркгофа [12]).

Заметим, что задача о минимальном имплицирующем векторе легко обобщается на случай матриц с неотрицательными элементами и равными суммами элементов строк. Такие матрицы, следуя [9], назовем квазистохастическими. В этом случае сумма компонент вектора $\tilde{p}$ в разложении (1) совпадает с суммой элементов строки матрицы $\mathbf{P}$. Поскольку между решениями задачи о минимальном имплицирующем векторе для стохастической и построенной по ней квазистохастической матрице легко устанавливается взаимно однозначное соответствие, то в дальнейшем без дополнительных оговорок в (1) вместо матрицы $\mathbf{P}$ будем подставлять матрицу $c\mathbf{P}$, где $c > 0$ — некоторое число.

Хорошо известны следующие оценки длины минимального имплицирующего вектора матрицы $\mathbf{P}$ [2]:

$$\max_{1 \leqslant i \leqslant m} n_i \leqslant L(\mathbf{P}) \leqslant \sum_{i=1}^{m} n_i - m + 1, \tag{2}$$

где через n_i обозначено число ненулевых элементов в i-й строке матрицы $\mathbf{P}$, $i = 1, 2, \ldots, m$. Отметим, что эти оценки можно получить исходя из размерности линейного пространства, которому принадлежит матрица $\mathbf{P}$, с использованием теоремы Каратеодори о выпуклой оболочке. Пример матрицы, на которой достигается верхняя оценка соотношения (2) в частном случае при $n_i = n$, $i = 1, 2, \ldots, n$, приведен в [9]. Заметим, что подобные (2) оценки для дважды стохастических матриц даны в [12]. Из (2) следует простой переборный алгоритм нахождения минимального имплицирующего вектора матрицы $\mathbf{P}$ [10]. А именно, достаточно перебрать все подмножества A множества $\mathcal{G}$ в порядке неубывания их мощностей, решая для каждого A соответствующую (1) систему линейных уравнений. Тогда, с учетом (2), для матрицы $\mathbf{P}$ достаточно перебрать не более $\exp\left(\left(\sum_{i=1}^{m} n_i - m + 1\right) \ln \prod_{i=1}^{m} n_i\right)$ таких подмножеств. В [4] показывается, что перебор достаточно вести по подмножествам A, состоящим из линейно независимых матриц. Используя свойства выпуклого многогранника стохастических матриц, можно и далее пытаться уменьшить этот перебор. Однако на этом пути вряд ли можно рассчитывать на существенное понижение объема перебора. Поэтому для разработки алгоритмов нахождения минимального имплицирующего вектора в § 1 авторы предлагают другой подход: задача о минимальном имплицирующем векторе формулируется в виде задачи частично целочисленного линейного программирования, что позволяет перенести результаты, методы и алгоритмы теории частично целочисленного линейного программирования на задачу о минимальном имплицирующем векторе.

NP-полнота задачи о минимальном имплицирующем векторе отмечена в [5]. В связи с этим важное значение приобретает построение хороших приближенных алгоритмов решения задачи. Назовем алгоритм построения имплицирующего вектора разностным, если он строит имплицирующий вектор, компоненты которого или равны элементам исходной матрицы, или получаются из элементов матрицы путем применения к ним конечного числа операций сложения и вычитания. Все описанные в литературе до настоящего времени приближенные алгоритмы относятся к классу разностных алгоритмов. Наиболее известным является градиентный алгоритм [2, 3, 9, 13], описание которого приведено в § 2. Различные приближенные алгоритмы, включающие в себя частичный перебор, были предложены в работах [2, 7, 14]. В частности, в [7] был предложен переборный разностный алгоритм нахождения минимального имплицирующего вектора для стохастической матрицы, состоящий из двух строк, и явно описаны параметры матрицы, от которых зависит длина минимального имплицирующего вектора. В общем случае И. А. Метра [9] привел примеры матриц, минимальные имплицирующие вектора которых не строятся ни одним разностным алгоритмом.

Во всех упомянутых выше работах практически не исследовались оценки длины получаемых имплицируемых векторов и не проводилось сопоставление длины построенных имплицирующих векторов с длиной минимального имплицирующего вектора. В то же время представляет интерес задача нахождения оценок, подобных (2), в зависимости от структурных свойств исходной матрицы **P**.

В данной работе строится иерархия классов стохастических матриц, объединение которых совпадает с классом всех стохастических матриц некоторой размерности, и оценивается функция Шеннона длины минимального имплицирующего вектора для каждого класса. Получены верхние и нижние оценки функции Шеннона, незначительно отличающиеся друг от друга, а в некоторых случаях совпадающие. Верхняя оценка получается из анализа градиентного алгоритма. Для получения нижней оценки функции Шеннона применяется метод получения эффективных нижних оценок длины минимального имплицирующего вектора для конкретных матриц, излагаемый в § 3. Эти оценки позволили установить зависимость длины минимального имплицирующего вектора матрицы от точности задания ее компонент. Из результатов § 4 следует, что градиентный алгоритм для самых сложных матриц из введенных классов строит имплицирующий вектор по длине близкий к минимальному, причем для широкого класса матриц имплицирующий вектор, построенный градиентным алгоритмом, является минимальным. В то же время в § 5 приведены примеры матриц, для которых длина имплицирующего вектора, построенного градиентным алгоритмом, существенно отличается от минимальной. В § 6 приводятся оценки типичных значений длины минимального имплицирующего вектора для рациональных стохастических матриц. Частично эти вопросы исследовались в [6].

Результаты данной работы могут быть использованы при разработке управляемых генераторов случайных кодов, в структурной теории вероятностных автоматов. Кроме того, как представляется авторам, методология и отдельные результаты работы имеют более широкую область применимости (теория неотрицательных матриц, теория выпуклых многогранников и т. п.). Краткое содержание работы было опубликовано в [8, 17].

§ 1. Задача о минимальном имплицирующем векторе как задача частично целочисленного линейного программирования

Через $\mathscr{P}(\tilde{n})$, где $\tilde{n}=(n_1, n_2, \ldots, n_m)$, обозначим класс стохастических $(m\times n)$-матриц с n_i ненулевыми элементами в i-й строке, $i=1, 2, \ldots$ $\ldots, m$. Далее без потери общности считаем $n_1\geqslant n_2\geqslant\ldots\geqslant n_m$. Через $L(\mathbf{P})$ обозначим длину минимального имплицирующего вектора для матрицы **P**. Положим $L^*(\tilde{n})=\sum\limits_{i=1}^{m} n_i - m + 1.$

Пусть $\mathbf{P}=\|p_{i,j}\|$ — исходная матрица, $\mathbf{P}\in\mathscr{P}(\tilde{n})$. Без потери общности будем считать, что для $i=1, 2, \ldots, m$, в i-й строке не равны нулю элементы $p_{i,1}, p_{i,2}, \ldots, p_{i,n_i}$. Для $i=1, 2, \ldots, m$, $j=1, 2, \ldots, n_i$, $s=1, 2, \ldots, L^*(\tilde{n})$ через p_s, y_s, $t_{i,j,s}$, $z_{i,j,s}$ обозначим переменные, $y_s\in\{0, 1\}$, $z_{i,j,s}\in\{0, 1\}$. Требуется решить задачу:

$$\sum_{s=1}^{L^*(\tilde{n})} y_s \Rightarrow \min \tag{3}$$

при ограничениях

$$\sum_{s=1}^{L^*(\tilde{n})} t_{i,j,s} = p_{i,j}, \quad i=1,2,\ldots,m, j=1,2,\ldots,n_i, \tag{4}$$

$$\sum_{j=1}^{n_i} t_{i,j,s} = p_s, \quad i=1,2,\ldots,m, \quad s=1,2,\ldots,L^*(\tilde{n}), \tag{5}$$

$$0 \leqslant t_{i,j,s} \leqslant z_{i,j,s}, \quad i=1, 2, \ldots, m, \; j=1, 2, \ldots, n_i, \; s=1, 2, \ldots, L^*(\tilde{n}), \tag{6}$$

$$0 \leqslant p_s \leqslant y_s, \quad s=1, 2, \ldots, L^*(\tilde{n}), \tag{7}$$

$$\sum_{j=1}^{n_i} z_{i,j,s} \leqslant y_s, \quad i=1, 2, \ldots, m, \quad s=1, 2, \ldots, L^*(\tilde{n}), \tag{8}$$

$$y_s \geqslant y_{s+1}, \quad s=1, 2, \ldots, L^*(\tilde{n})-1. \tag{9}$$

Вектор значений переменных $\tilde{r}=\big(p_1, p_2, \ldots, p_{L^*(\tilde{n})}; y_1, y_2, \ldots, y_{L^*(\tilde{n})}; t_{1,1,1}, t_{1,1,2}, \ldots, t_{m,n_m,L^*(\tilde{n})}; z_{1,1,1}, z_{1,1,2}, \ldots, z_{m,n_m,L^*(\tilde{n})}\big)$, удовлетворяющий условиям (3)—(9), назовем решением задачи (3)—(9). Через $\mathbf{Z}_s(\tilde{r})$ обозначим $(m \times n)$-матрицу $\|z_{i,j}\|$, где

$$z_{i,j} = \begin{cases} z_{i,j,s}, & \text{если} \quad j \leqslant n_i, \\ 0 & \text{в противном случае.} \end{cases}$$

Положим

$$l(\tilde{r}) = \begin{cases} L^*(\tilde{n}), & \text{если} \quad p_s > 0, \quad s=1, 2, \ldots, L^*(\tilde{n}), \\ \min\{t \mid p_{t+1}=0\} & \text{в противном случае.} \end{cases}$$

Теорема 1. *Пусть $\tilde{r}$ — решение задачи* (3)—(9). *Тогда* $\mathbf{Z}_k(\tilde{r})$ — *стохастическая матрица из нулей и единиц, где* $k=1, 2, \ldots, l(\tilde{r})$, $(p_1, p_2, \ldots, p_{l(\tilde{r})})$ — *минимальный имплицирующий вектор для матрицы* $\mathbf{P}$ *и* $\mathbf{P} = \sum_{k=1}^{l(\tilde{r})} p_k \mathbf{Z}_k(\tilde{r})$.

Доказательство. Для $k=1, 2, \ldots, l(\tilde{r})$ из соотношения (7) и определения $l(\tilde{r})$ следует, что $p_k > 0$ и $y_k = 1$. Из (5), (6) и $p_k > 0$ следует, что для любого i, $i=1, \ldots, m$, найдется j_i такое, что $1 \leqslant j_i \leqslant n_i$, $t_{i,j_i,k} > 0$ и $z_{i,j_i,k}=1$. Из (8) получаем, что при $j' \neq j_i$ выполняется $z_{i,j',k}=0$. Следовательно, $\mathbf{Z}_k(\tilde{r})$ — стохастическая (0, 1)-матрица. Из $z_{i,j',k}=0$ и (6) следует $t_{i,j',k}=0$, $p_k = t_{i,j_i,k}$, $t_{i,j_i,k} = p_k z_{i,j_i,k}$. Теперь (4) можно переписать так: $\sum_{k=1}^{l(\tilde{r})} p_k z_{i,j,k} = p_{i,j}$. Отсюда следует разложение $\mathbf{P} = \sum_{k=1}^{l(\tilde{r})} p_k \mathbf{Z}_k(\tilde{r})$ и, учитывая минимальность $l(\tilde{r})$ (см. (3)), получаем, что $(p_1, p_2, \ldots, p_{l(\tilde{r})})$ — минимальный имплицирующий вектор для матрицы $\mathbf{P}$.

§ 2. Верхние оценки

Для матрицы $\mathbf{P} = \|p_{i,j}\|$, $\mathbf{P} \in \mathscr{P}(\tilde{n})$, положим $\alpha_0(\mathbf{P}) = \{p_{i,j} \mid p_{i,j} > 0\}$. Пусть $\alpha_{i-1}(\mathbf{P})$ определено, $1 \leqslant i \leqslant L^*(\tilde{n})-1$. Тогда положим

$$\alpha_i(\mathbf{P}) = \alpha_{i-1}(\mathbf{P}) \cup \{|a-b| \mid a \neq b, \{a, b\} \subseteq \alpha_{i-1}(\mathbf{P})\}.$$

Обозначим $\alpha(\mathbf{P}) = \min\{a \mid a \in \alpha_{L^*(\tilde{n})-1}(\mathbf{P})\}$. Для $0 < \varepsilon \leqslant 1$ введем класс матриц $\mathscr{P}(\tilde{n}, \varepsilon)$:

$$\mathscr{P}(\tilde{n}, \varepsilon) = \{\mathbf{P} \mid \mathbf{P} \in \mathscr{P}(\tilde{n}), \alpha(\mathbf{P}) \geqslant \varepsilon\}.$$

Лемма 1. *Для произвольного $\tilde{n}$ класс $\mathscr{P}(\tilde{n}, \varepsilon)$ непуст тогда и только тогда, когда* $\varepsilon \in (0, 1/n_1]$.

Доказательство. Действительно, для $\mathbf{P} \in \mathscr{P}(\tilde{n})$ имеем $\alpha(\mathbf{P}) \leqslant \min_{1 \leqslant j \leqslant n_1} p_{1,j} \leqslant 1/n_1$. Поэтому при $\varepsilon > 1/n_1$ класс $\mathscr{P}(\tilde{n}, \varepsilon)$ пуст. С другой стороны матрица из класса $\mathscr{P}(\tilde{n})$ с рациональными элементами вида $a_{i,j}/n_1$ ($a_{i,j}$ — целые) принадлежит каждому классу $\mathscr{P}(\tilde{n}, \varepsilon)$ при $\varepsilon \in (0, 1/n_1]$.

В дальнейшем без дополнительных оговорок рассматриваются лишь непустые классы $\mathscr{P}(\tilde{n}, \varepsilon)$.

Для матриц из класса $\mathscr{P}(\tilde{n}, \varepsilon)$ обычным образом определяется функция Шеннона:

$$L(\tilde{n}, \varepsilon) = \max\{L(\mathbf{P}) \mid \mathbf{P} \in \mathscr{P}(\tilde{n}, \varepsilon)\}.$$

Наша цель — получение верхних и нижних оценок величины $L(\tilde{n}, \varepsilon)$. Определение принадлежности матрицы $\mathbf{P}$ к классу $\mathscr{P}(\tilde{n}, \varepsilon)$ через нахождение множеств $\alpha_0(\mathbf{P}), \ldots, \alpha_{L^*(\tilde{n})-1}(\mathbf{P})$ может потребовать большого перебора. Однако, для конкретных матриц этот вопрос может решаться более простым способом. Например, в случае рациональных матриц. Пусть $\mathbf{P}$ — матрица с рациональными элементами, $\mathbf{P} \in \mathscr{P}(\tilde{n})$, и N — наименьшее общее кратное знаменателей ее элементов. Легко видеть, что $\mathbf{P} \in \mathscr{P}(\tilde{n}, 1/N)$. Поскольку на практике мы имеем дело с рациональными матрицами, то оценки функции $L(\tilde{n}, \varepsilon)$ помимо теоретического интереса имеют и практическое приложение.

Для получения верхних оценок мы рассмотрим следующий алгоритм построения имплицирующего вектора, который по аналогии с задачей о покрытии назовем градиентным. Этот алгоритм под другими названиями хорошо известен (см. например [2, 13]).

Градиентный алгоритм. Алгоритм применяется к квазистохастической $(m \times n)$-матрице $\mathbf{P} = \|p_{i,j}\|$ и строит разложение вида (1). Для описания h-итерации ($h \geqslant 1$) градиентного алгоритма будем использовать матрицу $\mathbf{B} = \|b_{i,j}\|$. Положим $\mathbf{B} = \mathbf{P}$. Для $h = 1, 2, \ldots$ до тех пор, пока не получим нулевую матрицу $\mathbf{B}$, следует повторять основной цикл (h-итерацию):

Находим $p_h = \min\limits_{1 \leqslant i \leqslant m} \max\limits_{1 \leqslant j \leqslant n} b_{i,j}$ и $j_i = \min\left\{j \mid b_{i,j} = \max\limits_{1 \leqslant r \leqslant n_i} b_{i,r}\right\}$, $i = 1, 2, \ldots, m$. Строим стохастическую $(m \times n)$-матрицу $\mathbf{G}_h = \|g_{i,j}\|$, полагая $g_{i,j_i} = 1$, $i = 1, 2, \ldots, m$. Далее находится квазистохастическая матрица $\mathbf{B} - p_h\mathbf{G}_h$, которая обозначается через $\mathbf{B}$ и в случае, если $\mathbf{B}$ отлична от нулевой матрицы, вновь повторяется основной цикл с новым значением h, равным $h + 1$.

Если $\mathbf{P}$ — стохастическая матрица, то $(p_1, p_2, \ldots, p_L)$ — имплицирующий вектор для $\mathbf{P}$. Обозначим $L^{\text{г}}(\mathbf{P})$ длину имплицирующего вектора матрицы $\mathbf{P}$, построенного градиентным алгоритмом. Тогда из описания работы алгоритма следуют оценки.

Лемма 2. *Для всякой матрицы* $\mathbf{P}$ *из* $\mathscr{P}(\tilde{n})$ *выполняется*

$$n_1 \leqslant L^{\text{г}}(\mathbf{P}) \leqslant L^*(\tilde{n}).$$

Через S_h обозначим сумму элементов строки матрицы $\mathbf{B}$, полученной после выполнения h-итерации ($S_0 = 1$).

Лемма 3. *Для любых* $\tilde{n}$ *и* l ($1 \leqslant l \leqslant L^*(\tilde{n})$) *найдется единственная тройка целых* t, r, k, *удовлетворяющая следующим условиям*:

$$1 \leqslant r \leqslant k \leqslant m, \tag{10}$$

$$n_{k+1} \leqslant t < n_k \quad (\textit{считаем } n_{m+1} = 0), \tag{11}$$

$$\sum_{i=1}^{k} (n_i - t) + r - k = l. \tag{12}$$

Доказательство. Обозначим через I мультимножество $\{1, 2, \ldots, n_1, 1, 2, \ldots, n_2, \ldots, 1, 2, \ldots, n_m\}$. В I удалим l наибольших элементов (это возможно, так как $l \leqslant \sum\limits_{i=1}^{m} n_i$). Пусть d — наименьшее число, удаляемое из I. Положим $t = d - 1$, r — равным количеству удаленных из I чисел d, а k выберем так, чтобы выполнялось условие (11). Тогда выполнение условий (10) и (12) легко проверяется. Докажем единственность выбора. Допустим, что найдутся два набора чисел (t_1, r_1, k_1) и (t_2, r_2, k_2), удовлетворяющим условиям (10)—(12). Если $t_1=t_2$, то из (11) следует $k_1=k_2$ и из (12) получаем $r_1=r_2$. Пусть $t_1>t_2$. Тогда из (11) следует, что $k_1\leqslant k_2$ и $l=\sum_{i=1}^{k_1}(n_i-t_1)+r_1-k_1\leqslant\sum_{i=1}^{k_1}(n_i-t_2)-$ $-2k_1+r_1\leqslant\sum_{i=1}^{k_1}(n_i-t_2)-k_1\leqslant\sum_{i=1}^{k_2}(n_i-t_2)-k_2<$ (поскольку $n_j>t_2$ при $j\leqslant$ $\leqslant k_2)<\sum_{i=1}^{k_2}(n_i-t_2)-k_2+r_2=l$. Пришли к противоречию. Аналогично разбирается случай $t_2>t_1$.

Тройку (множество) чисел t, r, k, удовлетворяющих лемме 3, обозначим через $\varphi(\tilde{n}, l)$. Ключевой для получения верхней оценки является

Лемма 4. *Для любых $\tilde{n}$ и l $(1\leqslant l\leqslant L^*(\tilde{n}))$ выполняется*

$$S_l\leqslant(t^r(t+1)^{k-r})\Big/\prod_{i=1}^{k}n_i,$$

где числа t, r, k из $\varphi(\tilde{n}, l)$.

Доказательство. Пусть i_h — номер строки матрицы **B**, из которой на h-итерации градиентного алгоритма выбирается элемент p_h. Через d_h обозначим число ненулевых элементов строки i_h матрицы **B** перед выполнением h-итерации. Тогда $p_h\geqslant S_{h-1}/d_h$ и $S_h=S_{h-1}-p_h\leqslant$ $\leqslant(S_{h-1}(d_h-1))/d_h$. Отсюда следует

$$S_l\leqslant\prod_{i=1}^{l}(d_i-1)/d_i. \tag{13}$$

Поскольку в строке i_h матрицы **B** после h-итерации будет (d_h-1) ненулевых элементов, то $d_1, d_2, \ldots, d_l$ суть l чисел из мультимножества I, введенного в лемме 3. Пусть $c_1, c_2, \ldots, c_l$ — l наибольших чисел в I. Из (13) следует

$$S_l\leqslant\prod_{i=1}^{l}(c_i-1)/c_i=\left(\prod_{j=1}^{r}\prod_{i=t+1}^{n_j}(i-1)/i\right)\left(\prod_{j=r+1}^{k}\prod_{i=t+2}^{n_j}(i-1)/i\right)=$$
$$=(t^r(t+1)^{k-r})\Big/\prod_{i=1}^{k}n_i$$

(здесь и далее считаем, что $\prod_{i=s}^{n}a_i=1$ при $n<s$). Лемма доказана.

Положим $\mathscr{D}_0=1$, а для всякого l $(1\leqslant l\leqslant L^*(\tilde{n}))$ $\mathscr{D}_l=$ $=\prod_{i=1}^{l}(c_i-1)/c_i$, где $c_1, c_2, \ldots, c_l$ — числа, введенные в лемме 4. Из доказательства этой леммы имеем, что

$$\mathscr{D}_l=(t^r(t+1)^{k-r})\Big/\prod_{i=1}^{k}n_i.$$

Лемма 5. *Для данных $\tilde{n}$ и ε таких, что*

$$1+n_1\leqslant\varepsilon^{-1}\leqslant\prod_{i=1}^{m}n_i,$$

найдется число l такое, что $1\leqslant l\leqslant L^*(\tilde{n})$ и

$$[\mathscr{D}_l/\varepsilon]\leqslant t+1<[\mathscr{D}_{l-1}/\varepsilon], \tag{14}$$

где t из $\varphi(\tilde{n}, l)$ и через $[x]$ обозначено наибольшее целое число z такое, что $z \leqslant x$ (соответственно через $]x[$ обозначается минимальное целое z такое, что $z > x$).

Доказательство. Число t из $\varphi(\tilde{n}, l)$ обозначим через t_l. Поскольку последовательности $t_1, t_2, \ldots, t_{L^*(\tilde{n})-1}$ и $[\mathscr{D}_0/\varepsilon]$, $[\mathscr{D}_1/\varepsilon]$, ..., $[\mathscr{D}_{L^*(\tilde{n})-1}/\varepsilon]$ невозрастающие, а $[\mathscr{D}_0/\varepsilon] \geqslant n_1 + 1 > t_1 + 1 = n_1$ и $[\mathscr{D}_{L^*(\tilde{n})-1}/\varepsilon] = \left[(1^m \cdot 2^{m-m}) \Big/ \left(\varepsilon \prod\limits_{i=1}^{m} n_i\right)\right] \leqslant 1 < 2 = t_{L^*(\tilde{n})-1} + 1$, то l, удовлетворяющее (14), существует. Лемма доказана.

Через l^* обозначим минимальное l, удовлетворяющее лемме 5. Положим $c = [\mathscr{D}_{l^*}/\varepsilon]$. Для чисел t, r, k из $\varphi(\tilde{n}, l^*)$ положим

$$\Phi(\tilde{n}, \varepsilon) = \begin{cases} n_1, & \text{если } n_1 \leqslant \varepsilon^{-1} < n_1 + 1, \\ \sum\limits_{i=1}^{k} n_i - (t+1)k + r + c, & \text{если } n_1 + 1 \leqslant \varepsilon^{-1} \leqslant \prod\limits_{i=1}^{m} n_i, \\ L^*(\tilde{n}), & \text{если } \varepsilon^{-1} > \prod\limits_{i=1}^{m} n_i. \end{cases}$$

Теорема 2. *Если класс $\mathscr{P}(\tilde{n}, \varepsilon)$ непуст, то*

$$L(\tilde{n}, \varepsilon) \leqslant \Phi(\tilde{n}, \varepsilon).$$

Доказательство. Из определения класса $\mathscr{P}(\tilde{n}, \varepsilon)$ следует, что каждая компонента имплицирующего вектора для матрицы $\mathbf{P}$ из $\mathscr{P}(\tilde{n}, \varepsilon)$, построенного градиентным алгоритмом, не менее ε. При ε, удовлетворяющем неравенствам $n_1 \leqslant \varepsilon^{-1} < n_1 + 1$, имеем $S_{n_1} \leqslant 1 - n_1\varepsilon < 1/(n_1 + 1)$, а это означает, что $S_{n_1} = 0$ и длина вектора, построенного градиентным алгоритмом, не превосходит n_1. Учитывая (2), в этом случае получаем $L(\tilde{n}, \varepsilon) = \Phi(\tilde{n}, \varepsilon)$. Пусть $n_1 + 1 \leqslant \varepsilon^{-1} \leqslant \prod\limits_{i=1}^{m} n_i$. Тогда в силу лемм 3—5,

$$L^{\Gamma}(\mathbf{P}) \leqslant l^* + [S_{l^*}/\varepsilon] \leqslant l^* + \left[t^r (t+1)^{k-r} \Big/ \left(\varepsilon \prod_{i=1}^{k} n_i\right)\right] = l^* + c = \Phi(\tilde{n}, \varepsilon),$$

где t, r, k лежат в $\varphi(\tilde{n}, l^*)$. При $\varepsilon^{-1} > \prod\limits_{i=1}^{m} n_i$ из (2) получаем $L(\tilde{n}, \varepsilon) \leqslant \Phi(\tilde{n}, \varepsilon)$. С учетом леммы 1, теорема доказана.

Следствие 1. *Если $n_1 + 1 \leqslant \varepsilon^{-1} < \prod\limits_{i=1}^{m} n_i$, то*

$$L(\tilde{n}, \varepsilon) < L^*(\tilde{n}).$$

Доказательство. Из определения $\mathscr{D}_l$ следует, что $\mathscr{D}_{L^*(\tilde{n})-1} = \left(\prod\limits_{i=1}^{m} n_i\right)^{-1}$. Тогда для любой матрицы $\mathbf{P}$ из класса $\mathscr{P}(\tilde{n}, \varepsilon)$ выполняется $L^{\Gamma}(\mathbf{P}) \leqslant L^*(\tilde{n}) - 1 + [S_{L^*(\tilde{n})-1}/\varepsilon] \leqslant L^*(\tilde{n}) - 1 + [\mathscr{D}_{L^*(\tilde{n})-1}/\varepsilon] = L^*(\tilde{n}) - 1 < L^*(\tilde{n})$.

Комментарий. Из следствия 1 вытекает, что лишь матрицы $\mathbf{P}$, у которых $\alpha(\mathbf{P}) \leqslant \left(\prod\limits_{i=1}^{m} n_i\right)^{-1}$, могут иметь минимальный вектор длины $L^*(\tilde{n})$. В дальнейшем для каждого набора $\tilde{n}$ мы построим матрицу $\mathbf{P}$ из $\mathscr{P}\left(\tilde{n}, \left(\prod\limits_{i=1}^{m} n_i\right)^{-1}\right)$ с $\alpha(\mathbf{P}) = \left(\prod\limits_{i=1}^{m} n_i\right)^{-1}$ и $L(\mathbf{P}) = L^*(\tilde{n})$. Из определения $\Phi(\tilde{n}, \varepsilon)$ легко следует алгоритм вычисления $\Phi(\tilde{n}, \varepsilon)$ по данным $\tilde{n}$ и ε. Легко вывести и аналитические оценки, например, при $n = n_1 = n_2 = \ldots$

$\ldots = n_m$, $\varepsilon \leqslant (n_1+1)^{-1}$ имеем

$$L(\tilde{n}, \varepsilon) \leqslant nm - \left[n \sqrt[m-1]{\varepsilon n}\right](m-1)$$

или, более грубо, при $n = m$

$$L(\tilde{n}, \varepsilon) < n(2 - \ln n - \ln \varepsilon).$$

Эти оценки показывают содержательность задачи о минимальном имплицирующем векторе с точки зрения практики, поскольку ε можно трактовать как точность задания элементов рациональной матрицы.

§ 3. Нижние оценки

В этом параграфе мы излагаем метод доказательства нижних оценок длины минимального имплицирующего вектора для конкретных матриц. В дальнейшем этот метод применяется для получения нижних оценок для $L(\tilde{n}, \varepsilon)$.

Зафиксируем матрицу $\mathbf{P}$ из $\mathscr{P}(\tilde{n})$. Без потери общности далее будем считать, что в строке i не равны нулю первые n_i элементов, $i = 1, 2, \ldots$ $\ldots, m$. Для каждого i, $1 \leqslant i \leqslant m$, зафиксируем собственное подмножество $\mathscr{I}_i$ ненулевых элементов i-й строки, $\mathscr{I}_i \subset \{p_{i,1}, p_{i,2}, \ldots, p_{i,n_i}\}$. Заметим, что $|\mathscr{I}_i| < n_i$. Положим $c_{i,d} = \min\left\{\sum\limits_{p\in\mathscr{I}} p \,\middle|\, \mathscr{I} \subseteq \mathscr{I}_i,\ |\mathscr{I}| = d\right\}$, $d_m = |\mathscr{I}_m|$, $M_m = \sum\limits_{p\in\mathscr{I}_m} p$. Определим по индукции числа d_i и M_i, $i = 1, 2, \ldots, m-1$. Если d_{t+1} и M_{t+1} уже определены $(1 \leqslant t \leqslant m-1)$ и $\sum\limits_{p\in\mathscr{I}_t} p \leqslant M_{t+1}$, то положим $M_t = M_{t+1}$ и $d_t = 0$, в противном случае (если $\sum\limits_{p\in\mathscr{I}_t} p > M_{t+1}$) выберем наибольшее d такое, что $c_{t,d} \leqslant M_{t+1}$, и положим $d_t = |\mathscr{I}_t| - d$ и $M_t = \sum\limits_{p\in\mathscr{I}_t} p + M_{t+1} - c_{t,d}$.

Теорема 3. *Для любой матрицы* $\mathbf{P}$ *из класса* $\mathscr{P}(\tilde{n})$ *и определенных выше чисел* $d_1, d_2, \ldots, d_m$, M_1 *выполняется*

$$L(\mathbf{P}) \geqslant \begin{cases} \sum\limits_{i=1}^{m} d_i + 1, & \text{если}\quad M_1 < 1, \\ \sum\limits_{i=1}^{m} d_i, & \text{если}\quad M_1 \geqslant 1. \end{cases}$$

Доказательство. Пусть $\tilde{p} = (p_1, p_2, \ldots, p_L)$ — минимальный имплицирующий вектор матрицы $\mathbf{P}$. Будем говорить, что компонента p_j участвует в образовании компоненты $p_{s,t}$ матрицы $\mathbf{P}$, если в разложении (1) элемент $g_{s,t}$ матрицы $\mathbf{G}_j$ равен единице. Будем говорить, что компонента $p_{s,t}$ покрывается подмножеством $A = \{p_{i_1}, p_{i_2}, \ldots, p_{i_v}\}$ элементов вектора $\tilde{p}$, если в образовании $p_{s,t}$ участвуют только элементы из A.

Положим $\mu_t = \sum\limits_{i=t}^{m} d_i$ и $\mathscr{I}^t = \bigcup\limits_{i=t}^{m} \mathscr{I}_i$, $t = 1, 2, \ldots, m$. Индукцией по t $(t = m, m-1, \ldots, 1)$ докажем, что в образовании элементов множества $\mathscr{I}^t$ участвует не менее μ_t компонент минимального имплицирующего вектора $\tilde{p}$ и сумма этих элементов не превосходит M_t.

Базис индукции. Очевидно, что в образовании элементов множества $\mathscr{I}_m$ участвует не менее d_m компонент. Без потери общности будем считать, что этими компонентами являются $p_1, p_2, \ldots, p_{d_m}$. Легко видеть, что $\sum\limits_{i=1}^{d_m} p_i \leqslant \sum\limits_{p\in\mathscr{I}_m} p = M_m$.

Шаг индукции. Для некоторого t, $1 \leqslant t \leqslant m-1$, предположим, что в образовании элементов из $\mathcal{I}^{t+1}$ участвуют μ_{t+1} компонент минимального имплицирующего вектора $\tilde{p}$ (пусть этими компонентами являются $p_1, p_2, \ldots, p_{\mu_{t+1}}$) и $\sum_{i=1}^{\mu_{t+1}} p_i \leqslant M_{t+1}$. Если $d_t = 0$, то $\mu_t = \mu_{t+1}$, $M_t = M_{t+1}$ и $\sum_{i=1}^{\mu_t} p_i \leqslant M_t$. Пусть $d_t > 0$. Тогда компоненты $p_1, p_2, \ldots, p_{\mu_{t+1}}$ вектора $\tilde{p}$ покрывают не более $|\mathcal{I}_t| - d_t$ элементов из $\mathcal{I}_t$. (Действительно, из определения d_t следует, что сумма любых $|\mathcal{I}_t| - d_t + 1$ элементов множества $\mathcal{I}_t$ больше M_{t+1}.) Поэтому в образовании элементов множества $\mathcal{I}_t$ должны участвовать не менее d_t новых компонент (обозначим их через $p_{\mu_{t+1}+1}, p_{\mu_{t+1}+2}, \ldots, p_{\mu_{t+1}+d_t}$) вектора $\tilde{p}$. Тогда $\sum_{i=1}^{\mu_t} p_i \leqslant M_{t+1} + \sum_{p \in \mathcal{I}_t} p - c_{t,|\mathcal{I}_t|-d_t} = M_t$.

Таким образом, в минимальном имплицирующем векторе $\tilde{p}$ матрицы $\mathbf{P}$ должно быть не менее $\sum_{i=1}^{m} d_i$ компонент, причем, если $M_1 < 1$, то с учетом стохастичности матрицы $\mathcal{P}$ вектор $\tilde{p}$ должен содержать хотя бы еще один элемент. Теорема доказана.

Отметим, что теорема 3 остается справедливой и в том случае, когда условие $n_1 \geqslant n_2 \geqslant \ldots \geqslant n_m$ не выполняется.

Далее мы определяем конкретные матрицы, на которых продемонстрируем применение теоремы. Пусть $n \geqslant m$ и

$$\mathbf{B} = \begin{bmatrix} m-1 & m-1 & \ldots & m-1 & 0 \\ m-2 & m-2 & \ldots & m-2 & n-1 \\ m-3 & m-3 & \ldots & m-3 & 2(n-1) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 2 & 2 & \ldots & 2 & (m-3)(n-1) \\ 1 & 1 & \ldots & 1 & (m-2)(n-1) \end{bmatrix}$$

— квазистохастическая целочисленная $((m-1)\times n)$-матрица.

Следствие 2. $L(\mathbf{B}) \geqslant (n+1)\ln m - m$.

Доказательство. Положим $\mathcal{I}_i = \{b_{i,1}, b_{i,2}, \ldots, b_{i,n-1}\}$, $i = 1, 2, \ldots, m$, где через $b_{i,j}$ обозначен (i, j) элемент матрицы $\mathbf{B}$. Очевидно, $d_{m-1} = M_{m-1} = n-1$. Индукцией по t, $t = m-2, m-3, \ldots, 1$, покажем, что

$$d_t \geqslant (n-m+t+1)/(m-t), \tag{15}$$

$$M_t \leqslant (m-t)n - 1. \tag{16}$$

Базис индукции. Используя определения, данные перед теоремой 3, имеем $d_{m-2} = n-1-[(n-1)/2] \geqslant (n-1)/2$ и $M_{m-2} = n-1+2]\,(n-1)/2\,[\leqslant 2n-1$.

Шаг индукции. Предположим, что соотношения (15) и (16) верны при $t = s+1$, $1 \leqslant s \leqslant m-3$. Покажем, что они справедливы при $t = s$.

Имеем $\sum_{j=1}^{n-1} b_{s,j} = (n-1)(m-s) > (m-s-1)n - 1 \geqslant M_{s+1}$. Тогда из правил построения чисел d_s следует, что $d_s \geqslant n-1-[M_{s+1}/(m-s)] \geqslant n-1-((m-s-1)n-1)/(m-s) = (n+s-m+1)/(m-s)$. Кроме того, $M_s = (m-s)(n-1) + M_{s+1} - [M_{s+1}/(m-s)](m-s)$. Положим $\Delta = M_{s+1} - [M_{s+1}/(m-s)](m-s)$. Очевидно, $0 \leqslant \Delta \leqslant m-s-1$. Поэтому $M_s \leqslant (m-s)(n-1) + m-s-1 = (m-s)n-1$. Утверждение шага индукции доказано. Применяя теорему 3, получим

$$L(\mathrm{B}) \geqslant \sum_{i=1}^{m-1} d_i \geqslant \sum_{z=0}^{m-2} (n-z)/(z+1) - 1 = \sum_{z=0}^{m-2} (n+1)/(z+1) - m \geqslant$$
$$\geqslant (n+1) \sum_{z=1}^{m-1} 1/z - m \geqslant (n+1) \ln m - m.$$

Следствие доказано.

Рассмотрим еще один пример. Для данных $\tilde{n}$ и l $(1 \leqslant l < L^*(\tilde{n}))$ определим целочисленную неотрицательную $(m \times n)$-матрицу $\mathrm{A}(\tilde{n}, l) = \|a_{i,j}\|$. Далее в этом параграфе через t, r и k обозначаются соответствующие числа из набора $\varphi(\tilde{n}, l)$.

Положим $R_k = n_k$ и $a_{i,j} = 1$ для i и j таких, что $k \leqslant i \leqslant m$ и $r \leqslant j \leqslant n_i$. Пусть элементы $a_{i+1,2}, a_{i+1,3}, \ldots, a_{i+1,n_{i+1}}$ $(1 \leqslant i < k)$ определены. Тогда

а) если $i > r$, то положим

$$b_i = R_{i+1} - (t+1)\,[R_{i+1}/(t+1)],$$
$$a_{i,j} = \begin{cases} [R_{i+1}/(t+1)], & \text{если} \quad 2 \leqslant j \leqslant t+1-b_i, \\ ]R_{i+1}/(t+1)[, & \text{если} \quad t+1-b_i < j \leqslant n_i, \end{cases}$$
$$R_i = R_{i+1} + (n_i - t - 1)]\,R_{i+1}/(t+1)[,$$

б) если $i \leqslant r$, то положим

$$b_i = R_{i+1} - t\,[R_{i+1}/t],$$
$$a_{i,j} = \begin{cases} [R_{i+1}/t], & \text{если} \quad 2 \leqslant j \leqslant t - b_i, \\ ]R_{i+1}/t[, & \text{если} \quad t - b_i < j \leqslant n_i, \end{cases}$$
$$R_i = R_{i+1} + (n_i - t)]\,R_{i+1}/t[.$$

Полагаем $a_{1,1} = [R_2/t]$, $a_{i,1} = R_1 - \sum_{j=2}^{n_i} a_{i,j}$, $i = 2, 3, \ldots, m$, и $a_{i,j} = 0$, $i = 1, 2, \ldots, m$, $j = n_i + 1, n_i + 2, \ldots, n$. Матрица $\mathrm{A}(\tilde{n}, l)$ построена.

Обозначим

$$f_j = \begin{cases} n_j - t, & \text{если} \quad 1 \leqslant j \leqslant r, \\ n_j - t - 1, & \text{если} \quad r < j \leqslant k, \end{cases}$$
$$h_j = n_j - f_j, \quad j = 1, 2, \ldots, k.$$

Далее нам понадобятся следующие технические леммы.

Лемма 6.

$$a_{i,1} = \begin{cases} [R_{i+1}/h_i] + \sum_{j=1}^{i-1} f_j a_{j,n_j}, & \textit{если} \quad 1 \leqslant i \leqslant k-1, \\ 1 + \sum_{j=1}^{k-1} f_j a_{j,n_j}, & \textit{если} \quad i = k, \\ 1 + \sum_{j=1}^{k-1} f_j a_{j,n_j} + n_k - n_i, & \textit{если} \quad k+1 \leqslant i \leqslant m. \end{cases}$$

Доказательство. Легко видеть, что $R_1 = R_2 + f_1 a_{1,n_1} = R_3 + f_2 a_{2,n_2} + f_1 a_{1,n_1} = \ldots = R_{i+1} + \sum_{j=1}^{i} f_j a_{j,n_j}$, $1 \leqslant i \leqslant k-1$. Далее, $a_{i,1} = R_{i+1} + \sum_{j=1}^{i} f_j a_{j,n_j} - \sum_{j=2}^{n_i} a_{i,j} = R_{i+1} + \sum_{j=1}^{i-1} f_j a_{j,n_j} - \sum_{j=2}^{h_i} a_{i,j}$. Из определения элементов $a_{i,2}, a_{i,3} \ldots, a_{i,h_i}$ видно, что $R_{i+1} = [R_{i+1}/h_i] + \sum_{j=2}^{h_i} a_{i,j}$. Аналогично доказываются и остальные утверждения.

Положим

$$\psi(\tilde{n}, l) = \begin{cases} \left(\prod_{i=1}^{k} n_i\right) \Big/ \left(t^r (t+1)^{k-r-1}\right), & \text{если } 1 \leqslant r \leqslant k-1, \\ \left(\prod_{i=1}^{k} n_i\right) \Big/ t^{k-1}, & \text{если } r = k. \end{cases}$$

Лемма 7.

$$\psi(\tilde{n}, l) \leqslant R_1 < \psi(\tilde{n}, l)(1 + h_{k-1}/n_k).$$

Доказательство. По определению имеем $R_k = n_k$ и $(n_k n_{k-1})/n_{k-1} \leqslant \leqslant R_{k-1} < n_k + (n_{k-1} - h_{k-1})] n_k/h_{k-1} [\leqslant (n_k n_{k-1})/h_{k-1} + n_{k-1} - h_{k-1}$. Предположим, что для j, $2 \leqslant j \leqslant k-1$, выполняется

$$\prod_{i=j}^{k} n_i \Big/ \prod_{i=j}^{k-1} h_i \leqslant R_j < (1 + h_{k-1}/n_k) \prod_{i=j}^{k} n_i \Big/ \prod_{i=j}^{k-1} h_i - h_j.$$

Тогда из определения R_{j-1} имеем

$$(n_{j-1}R_j)/h_{j-1} \leqslant R_{j-1} < n_{j-1}R_j/h_{j-1} + n_{j-1} - h_{j-1} <$$
$$< (1 + h_{k-1}/n_k) \prod_{i=j-1}^{k} n_i \Big/ \prod_{i=j-1}^{k-1} h_i - n_{j-1}h_j/h_{j-1} + n_{j-1} - h_{j-1}.$$

Замечая, что $h_j \geqslant h_{j-1}$, и используя принцип математической индукции, мы выводим справедливость леммы.

Поскольку сумма элементов каждой строки матрицы $\mathbf{A}(\tilde{n}, l)$ равна R_1, то поделив каждый элемент $a_{i,j}$ на R_1, мы получим стохастическую $(m \times n)$-матрицу $\mathbf{P}(\tilde{n}, l) = \|p_{i,j}\|$ из класса $\mathcal{P}(\tilde{n}, R_1^{-1})$.

Теорема 4.

$$L(\mathbf{P}(\tilde{n}, l)) = \begin{cases} \sum_{i=1}^{k} n_i - (t+1)(k-1) + r, & \textit{если } k > r, \\ \sum_{i=1}^{k} n_i - t(k-1), & \textit{если } k = r. \end{cases}$$

Доказательство. Нижняя оценка. Положим

$$\mathcal{I}_i = \begin{cases} \{p_{i,2}, p_{i,3}, \ldots, p_{i,n_i}\}, & \text{если } 1 \leqslant i \leqslant k, \\ \varnothing, & \text{если } k < i \leqslant m. \end{cases}$$

По определению имеем $d_m = d_{m-1} = \ldots = d_{k+1} = M_m = M_{m-1} = \ldots = M_{k+1} = 0$, $d_k = n_k - 1$, $M_k = (R_k - 1)/R_1$. Индукцией по i $(i = k-1, k-2, \ldots \ldots, 1)$ докажем, что $d_i = n_i - h_i$, $M_i = (R_i - 1)/R_1$.

Базис. $i = k-1$. Согласно определению матрицы $\mathbf{P}(\tilde{n}, l)$ выполняются следующие соотношения: $p_{i,2} \leqslant p_{i,3} \leqslant \ldots \leqslant p_{i,n_i}$, $\sum_{j=2}^{h_{k-1}} p_{k-1,j} \leqslant M_k$ и $\sum_{j=2}^{h_{k-1}+1} p_{k-1,j} \geqslant R_k/R_1 > M_k$. Поэтому $d_{k-1} = n_{k-1} - h_{k-1}$, $M_{k-1} = \sum_{p \in \mathcal{I}_{k-1}} p + M_k - \sum_{j=2}^{h_i} p_{k-1,j} = ((n_{k-1} - h_{k-1})]R_k/h_{k-1}[+ R_k - 1)/R_1 = (R_{k-1} - 1)/R_1$.

Шаг индукции. Пусть $d_{i+1} = n_{i+1} - h_{i+1}$, $M_{i+1} = (R_{i+1} - 1)/R_1$ для некоторого i, $1 \leqslant i \leqslant k-2$. Тогда $\sum_{j=2}^{h_i} p_{i,j} \leqslant M_{i+1} < R_{i+1}/R_1 < \sum_{j=2}^{h_i+1} p_{i,j}$ и поэтому $d_{i+1} = n_{i+1} - h_{i+1}$, $M_i = \sum_{j=2}^{n_i} p_{i,j} - \sum_{j=2}^{h_i} p_{i,j} + M_{i+1} = \sum_{j=h_i+1}^{n_i} p_{i,j} + M_{i+1} = ((n_i - h_i)]\, R_{i+1}/h_i\, [+ R_{i-1} - 1)/R_1 = (R_{i-1} - 1)/R_1$. Тогда имеем $M_1 < 1$. Применяя теорему 3, получаем нижнюю оценку.

Верхняя оценка. Рассмотрим работу градиентного алгоритма на матрице $\mathbf{P}(\tilde{n}, l)$. Положим $g_0 = 0$, $g_j = \sum_{i=1}^{j} f_i$, $1 \leqslant j \leqslant k-1$. Наша цель показать, что имплицирующий вектор $\tilde{p} = (p_1, p_2, \ldots, p_L)$, построенный градиентным алгоритмом, содержит $L = g_{k-1} + n_k$ элементов, первые f_1 из которых равны p_{1,n_1}, следующие f_2 равны p_{2,n_2}, и так далее, последние n_k элементов равны p_{k,n_k}. Пусть $h = g_{i-1} + c$, где $1 \leqslant c \leqslant f_i$, $1 \leqslant i \leqslant k$. Индукцией по h докажем, что $(m \times n)$-матрица $\mathbf{B} = \|b_{i,j}\|$, получающаяся после h-й итерации градиентного алгоритма, удовлетворяет следующим условиям:

а) в j-й строке матрицы $\mathbf{B}$ при $1 \leqslant j \leqslant i-1$ имеется не более h_j ненулевых элементов;

б) $b_{j,u} = p_{j,u}$ для $j = i+1, i+2, \ldots, m$, $u = 2, 3, \ldots, n$ и $b_{j,1} = p_{j,1} - \sum_{u=1}^{i-1} p_{u,n_u} f_u - c p_{i,n_i}$ для $j = i+1, i+2, \ldots, m$;

в) $b_{i,1} = p_{j,1} - \sum_{u=1}^{i-1} p_{u,n_i} f_u$, $b_{i,u} = p_{i,u}$, $u = 2, 3, \ldots, n_i - c$, и $b_{i,u} = 0$, $u = n_i - c + 1, n_i - c + 2, \ldots, n$;

г) все ненулевые элементы матрицы $\mathbf{B}$ представляются в виде D_1/D_2, где D_i — целые положительные числа, $i = 1, 2$, и R_1 кратно D_2.

Базис. $h = 1$. Легко видеть, что $p_1 = p_{1,n_1}$ и после первой итерации градиентного алгоритма получим следующую матрицу $\mathbf{B}$: $b_{1,n_1} = 0$, $b_{j,1} = p_{j,1} - p_{1,n_1}$, $j = 2, 3, \ldots, m$, остальные элементы $b_{i,j}$ матрицы $\mathbf{B}$ равны $p_{i,j}$, т. е. условия а) — г) для $\mathbf{B}$ выполняются.

Шаг индукции. Пусть условия а)—г) выполняются для матрицы $\mathbf{B}$, полученной после h-й итерации, где $h = g_{i-1} + c$ и $1 \leqslant h < g_{k-1}$.

Пусть $c < f_i$. Из леммы 6 и предположения индукции следует, что $b_{i,1} = [R_{i+1}/h_i]/R_1$, $b_{i,j} = p_{i,j}$, $j = 2, 3, \ldots, n_i - c$, $b_{i,j} = 0$, $j = n_i - c + 1, n_i - c + 2, \ldots, n$, и согласно правилу построения матрицы $\mathbf{P}(\tilde{n}, l)$ имеем $b_{i,1} \leqslant b_{i,2} \leqslant \ldots \leqslant b_{i,n_i-c} =]R_{i+1}/h_i[/R_1 = p_{i,n_i}$. Для всех j таких, что $i+1 \leqslant j \leqslant m$, также имеем

$$b_{i,1} = [R_{j+1}/h_j]/R_1 + \sum_{d=1}^{j-1} f_d p_{d,n_d} - \sum_{d=1}^{i-1} f_d p_{d,n_d} - c p_{i,n_i} \geqslant$$
$$\geqslant [R_{i+1}/h_j]/R_1 + p_{i,n_i} = b_{i,n_i-c}.$$

Поскольку суммы элементов строк матрицы $\mathbf{B}$ совпадают и в строке с номером j, где $1 \leqslant j \leqslant i-1$, содержится не более h_i ненулевых элементов, то в этой строке есть элемент не меньший b_{i,n_i-c}. Поэтому на $(h+1)$-й итерации $p_{h+1} = b_{i,n_i-c} = p_{i,n_i}$ и в матрице $\mathcal{B}$ после $(h+1)$-й итерации выполняются условия а) — г), так как изменения в строках $1, 2, \ldots, i-1$ на пункты а) — г) не влияют, в i-й строке $b_{i,n_i-c} = 0$, а в строках с номерами j, где $j = i+1, i+2, \ldots, m$ от элемента $b_{j,1}$ отнимается p_{i,n_i}. Практически аналогично рассматривается и случай $c = f_i$.

Таким образом, после g_{k-1} итераций k-я строка матрицы $\mathbf{B}$ состоит из n_k ненулевых элементов, равных $1/R_1$. Поэтому последние n_k элементов имплицирующего вектора равны p_{k,n_k}.

Комментарий. Заметим, что если при построении матрицы $\mathbf{A}(\tilde{n}, l)$ в качестве элемента a_{1,n_1} взять любое действительное число $\beta >]R_2/t[$ и положить $R_1 = R_2 + (n_1 - t - 1)]R_2/t[+ \beta$, а остальные формулы в определении оставить без изменения, то для получившейся при этом из матрицы $\mathbf{A}(\tilde{n}, l)$ (делением всех ее элементов на R_1) матрицы $\mathbf{P}(\tilde{n}, l)$ оценки теоремы 4 сохраняются. Тем самым, для каждой пары $\tilde{n}$ и l таких, что $n_1 \leqslant l \leqslant L^*(\tilde{n})$, мы построили континуальное семейство $\mathscr{P}1(\tilde{n}, l) \subset \mathscr{P}(\tilde{n})$ стохастических матриц с длиной минимального имплицирующего вектора равной l. Если при построении матрицы $\mathbf{P}(\tilde{n}, l)$ мы положим $n_1 = n_2 = \ldots = n_m = n = r = k$, $t = 1$, то получим в точности матрицу из [9], у которой $R_1 = n^n$ и длина минимального имплицирующего вектора равна $n^2 - n + 1$.

§ 4. Функция Шеннона

В этом параграфе доказываются нижние и верхние оценки функции Шеннона $L(\tilde{n}, \varepsilon)$. Положим $R(1) = n_1 + 1$. Через $R(l)$ обозначим сумму элементов строки матрицы $\mathbf{A}(\tilde{n}, l)$, $l = 2, 3, \ldots, L^*(\tilde{n}) - 1$ (напомним, что суммы элементов строк матрицы $\mathbf{A}(\tilde{n}, l)$ совпадают). Из правила построения $\mathbf{A}(\tilde{n}, l)$ (см. § 3) легко следует, что $n_1 \leqslant R(2) \leqslant R(3) \leqslant \ldots \ldots \leqslant R(L^*(\tilde{n}) - 1) = \prod_{i=1}^{m} n_i$. Разобьем $(0, 1/n_1]$ на интервалы*): $U_1 = (1/(n_1+1), 1/n_1]$, $U_i = (1/R(i), 1/R(i-1)]$, $i = 2, 3, \ldots, L^*(\tilde{n}) - 1$, $U_{L^*(\tilde{n})} = \left(0, \left(\prod_{i=1}^{m} n_i\right)^{-1}\right)$.

Теорема 5. а) *Если* $\varepsilon \in U_1$, *то* $L(\tilde{n}, \varepsilon) = n_1$;

б) *если* $\varepsilon \in U_l$, $l = 2, 3, \ldots, L^*(\tilde{n}) - 1$ и $t \in \varphi(\tilde{n}, l)$, *то* $\Phi(\tilde{n}, \varepsilon) - t - 1 \leqslant L(\tilde{n}, \varepsilon) \leqslant \Phi(\tilde{n}, \varepsilon)$;

в) *если* $\varepsilon \in U_{L^*(\tilde{n})}$, *то* $L(\tilde{n}, \varepsilon) = L^*(\tilde{n})$.

Доказательство. Верхняя оценка функции $L(\tilde{n}, \varepsilon)$ в п. а)—в) следует из теоремы 2. Нижняя оценка в а) очевидна. Для доказательства нижней оценки в п. в) заметим, что набор $\varphi(\tilde{n}, L^*(\tilde{n}) - 1)$ состоит из чисел $t = 1$, $r = k = m$. Легко подсчитать, что сумма элементов любой строки матрицы $\mathbf{A}(\tilde{n}, L^*(\tilde{n}) - 1)$ равна $\prod_{i=1}^{m} n_i$. Поэтому $\mathbf{P}(\tilde{n}, L^*(\tilde{n}) - 1) \in \mathscr{P}\left(\tilde{n}, \left(\prod_{i=1}^{m} n_i\right)^{-1}\right)$ и при $\varepsilon \in U_{L^*(\tilde{n})}$ имеем

$$L(\tilde{n}, \varepsilon) \geqslant L(\mathbf{P}(\tilde{n}, L^*(\tilde{n}) - 1)) = \sum_{i=1}^{m} n_i - m + 1 = L^*(\tilde{n}).$$

Перейдем к доказательству нижней оценки в б). Зафиксируем l, $2 \leqslant l \leqslant L^*(\tilde{n}) - 1$. Пусть $\varepsilon \in U_l$ и $(t, r, k) \in \varphi(\tilde{n}, l)$. Далее считаем, что $r \leqslant k - 1$, поскольку при $r = k$ интервал U_l пуст. Поэтому без ограничения общности, в дальнейшем полагаем $k \geqslant 2$. Так как $\mathbf{P}(\tilde{n}, l - 1) \in \mathscr{P}(\tilde{n}, \varepsilon)$, то с учетом теорем 2 и 4 имеем

$$\Phi(\tilde{n}, \varepsilon) \geqslant L(\tilde{n}, \varepsilon) \geqslant L(\mathbf{P}(\tilde{n}, l - 1)) = l + t. \tag{17}$$

Оценим снизу функцию $L(\tilde{n}, \varepsilon)$ через $\Phi(\tilde{n}, \varepsilon)$. Как и ранее через l^* обозначено минимальное число l, удовлетворяющее лемме 5. Рассмотрим два случая.

*) Интервал $(a, b]$ при $a \geqslant b$ считаем пустым.

1) $l^* < l$. Пусть $t^* \in \varphi(\tilde{n}, l^*)$ и $t_{l-1} \in \varphi(\tilde{n}, l-1)$. Поскольку для чисел l_1, l_2, t_1 и t_2, удовлетворяющих условиям $1 \leqslant l_1 < l_2 \leqslant L^*(\tilde{n}) - 1$, $t_1 \in \varphi(\tilde{n}, l_1)$, $t_2 \in \varphi(\tilde{n}, l_2)$, выполняется соотношение $l_1 + t_1 \leqslant l_2 + t_2$, то имеем: $\Phi(\tilde{n}, \varepsilon) = l^* + [\mathscr{D}_{l^*}/\varepsilon] \leqslant l^* + t^* + 1 = l - 1 + t^* - (l - l^* - 1) + 1 \leqslant l + t_{l-1} \leqslant l + t + 1$. Поэтому $t + l \geqslant \Phi(\tilde{n}, \varepsilon) - 1$ и, учитывая (17), имеем

$$\Phi(\tilde{n}, \varepsilon) - 1 \leqslant L(\tilde{n}, \varepsilon) \leqslant \Phi(\tilde{n}, \varepsilon).$$

2) $l^* \geqslant l$. Из леммы 7 следует, что $1/\varepsilon \leqslant R(l) < \psi(\tilde{n}, l)(1 + h_{k-1}/n_k)$. Тогда $[\mathscr{D}_l/\varepsilon] < (t+1)(1 + h_{k-1}/n_k) \leqslant (t+1)(1 + (t+1)/n_k) \leqslant 2t + 2$. С учетом, что $[\mathscr{D}_l/\varepsilon]$ — целое, имеем $[\mathscr{D}_l/\varepsilon] \leqslant 2t + 1$. С другой стороны из определения $\mathscr{D}_l$ и l^* следует, что $\mathscr{D}_l - \mathscr{D}_{l^*} \geqslant (l - l^*)\varepsilon$. Тогда

$$l^* + [\mathscr{D}_{l^*}/\varepsilon] = (l^* - l) + l + [\mathscr{D}_{l^*}/\varepsilon] \leqslant$$
$$\leqslant \mathscr{D}_l/\varepsilon - \mathscr{D}_{l^*}/\varepsilon + l + [\mathscr{D}_{l^*}/\varepsilon] \leqslant l + \mathscr{D}_l/\varepsilon.$$

Учитывая, что $l^* + [\mathscr{D}_{l^*}/\varepsilon]$ целое, имеем $l^* + [\mathscr{D}_{l^*}/\varepsilon] \leqslant l + [\mathscr{D}_l/\varepsilon]$. Объединяя вышесказанное, получим $\Phi(\tilde{n}, \varepsilon) = l^* + [\mathscr{D}_{l^*}/\varepsilon] \leqslant l + [\mathscr{D}_l/\varepsilon] \leqslant \leqslant l + 2t + 1$. Поэтому $l + t \geqslant \Phi(\tilde{n}, \varepsilon) - t - 1$ и с учетом (17) получаем соотношение $L(\tilde{n}, \varepsilon) \geqslant \Phi(\tilde{n}, \varepsilon) - t - 1$.

Комментарий. Нижние оценки теоремы 5 могут быть усилены, но их доказательство требует больших технических выкладок. Поэтому мы ограничились такой формулировкой теоремы. Отметим, что функция $\Phi(\tilde{n}, \varepsilon)$ в ряде случаев совпадает с $L(\tilde{n}, \varepsilon)$. Например, если $t \in \varphi(\tilde{n}, l)$, $\psi(\tilde{n}, l) \leqslant \varepsilon^{-1} < (t+1)\psi(\tilde{n}, l)/(t+2)$ и $n_2, n_3, \ldots, n_{r+1}$ нацело делятся на t, а $n_{r+2}, n_{r+3}, \ldots, n_k$ на $t+1$, то $L(\tilde{n}, \varepsilon) = \Phi(\tilde{n}, \varepsilon)$.

§ 5. Погрешность градиентного алгоритма

В предыдущих пунктах мы показали, что градиентный алгоритм для широкого класса матриц строит имплицирующий вектор, близкий к минимальному. Поэтому представляет интерес оценка его погрешности. Через $\mathscr{P}(m, n)$ обозначим класс $(m \times n)$-стохастических матриц и положим $\lambda(m, n) = \max\limits_{\mathbf{P} \in \mathscr{P}(m,n)} \left(L^{\text{г}}(\mathbf{P})/L(\mathbf{P})\right)$. Тогда верна

Теорема 6. *Для* $m \geqslant 2$ *и* $n \geqslant 2$

$$\lambda(m, n) \geqslant \begin{cases} ((n+1)\ln m + n - m)/(n + m - 1), & \text{если} \quad m \leqslant n, \\ (n+1)\ln n/(2n-1), & \text{если} \quad m > n. \end{cases}$$

Сначала рассмотрим два вспомогательных результата. Пусть $\Delta > 0$ и

$$\mathbf{A} = \begin{bmatrix} a & a-\Delta & \ldots & a-(n-2)\Delta & b \\ a-\Delta & a-2\Delta & \ldots & a-(n-1)\Delta & b+(n-1)\Delta \\ a-2\Delta & a-3\Delta & \ldots & a-n\Delta & b+2(n-1)\Delta \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a-(m-1)\Delta & a-m\Delta & \ldots & a-(m+n-3)\Delta & b+(m-1)(n-1)\Delta \end{bmatrix} \qquad (18)$$

— неотрицательная матрица с равными суммами строк.

Лемма 8. $L(\mathbf{A}) \leqslant m + n - 1$.

Доказательство. Для любого числа r, $2 \leqslant r \leqslant n + m - 1$, введем множество $\mathbf{Z}_{n-1}(r) = \{\alpha \mid 1 \leqslant \alpha \leqslant r,\ \alpha \equiv r(\mathrm{mod}(n-1))\}$. Для любых целых r, i, j таких, что $2 \leqslant r \leqslant m + n - 1$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, положим

$$g_{i,j}^{(r)} = \begin{cases} 1, & \text{если} \quad 1 \leqslant i \leqslant m \leqslant r,\ 1 \leqslant j \leqslant n-1,\ i + j \in \mathbf{Z}_{n-1}(r), \text{ или } r < \\ & < i \leqslant m,\ j = n, \quad \text{или} \quad 1 \leqslant i \leqslant r \text{ и } m > r,\ 1 \leqslant j \leqslant n-1, \\ & i + j \in \mathbf{Z}_{n-1}(r), \\ 0 & \text{в противном случае,} \end{cases}$$

$$g_{i,j}^{(1)} = \begin{cases} 1, & \text{если} \quad j = n, \\ 0 & \text{в противном случае,} \end{cases}$$

и определим вырожденные стохастические $(m \times n)$-матрицы $\mathbf{G}_r = \|g_{i,j}^{(r)}\|$, $r = 1, 2, \ldots, m + n - 1$. Тогда имеет место разложение

$$\mathbf{A} = b\mathbf{G}_1 + (n-1)\Delta \sum_{j=2}^{m} \mathbf{G}_j + \sum_{j=1}^{n-1} (a - (m-2)\Delta + j\Delta)\mathbf{G}_{m+j}. \qquad (19)$$

Коэффициенты перед матрицами $\mathbf{G}_j$, $j = 1, 2, \ldots, m + n - 1$ в разложении (19) дают имплицируюший вектор для матрицы $\mathbf{A}$. Лемма доказана.

Пусть $n \geqslant m \geqslant 2$ и $a > (mn + m - 2)\Delta$. Положим в матрице из предыдущей леммы $b = m\Delta$ и обозначим полученную матрицу через $\mathbf{D}$.

Лемма 9. $L^r(\mathbf{D}) \geqslant (n+1)\ln m + n - m$.

Доказательство. Вырожденную стохастическую $(m \times n)$-матрицу $\mathbf{G} = \|g_{i,j}\|$ условимся задавать вектором $\bar{b} = (b_1, b_2, \ldots, b_m)$, где $b_i = j$, если $g_{i,j} = 1$. Поскольку $a > (mn + m - 2)\Delta$, то на первых n итерациях градиентный алгоритм отберет в имплицирующий вектор элементы $p_i = a - (m - 2 + i)\Delta$, $i = 1, 2, \ldots, n-1$, $p_n = m\Delta$. Обозначая через $\mathbf{G}_j$ матрицу, определяемую вектором $(j, j, \ldots, j)$, положим $\mathbf{B}_1 = \left(\mathbf{A} - \sum_{j=1}^{n} p_j \mathbf{G}_j\right)\Big/\Delta$. Тогда $\mathbf{B}_1$ имеет вид

$$\mathbf{B}_1 = \begin{bmatrix} m-1 & m-1 & \ldots & m-1 & 0 \\ m-2 & m-2 & \ldots & m-2 & n-1 \\ m-3 & m-3 & \ldots & m-3 & 2(n-1) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & & 1 & (m-2)(n-1) \\ 0 & 0 & & 0 & (m-1)(n-1) \end{bmatrix}.$$

Легко видеть, что $L(\mathbf{B}_1) = L(\mathbf{B})$, где $\mathbf{B}$ — матрица, определенная следствием 2. Применяя это следствие, получим $L^r(\mathbf{D}) \geqslant n + L(\mathbf{B}_1) \geqslant (n + 1)\ln m + n - m$.

Перейдем к доказательству теоремы 6. При $m \leqslant n$ утверждение теоремы непосредственно следует из лемм 8 и 9.

Пусть $m > n$. Рассмотрим $(n \times n)$-матрицу $\mathbf{D}$, определенную в лемме 9, и дополним ее до $(m \times n)$-матрицы, добавив к ней $(m - n)$ раз последнюю строку матрицы $\mathbf{D}$. Вновь полученную матрицу обозначим через $\mathbf{D}'$. Поскольку добавление уже существующей строки не меняет длину минимального имплицирующего вектора матрицы, то $L(\mathbf{D}') = L(\mathbf{D}) \leqslant 2n - 1$. В то же время повторяя рассуждения, приведенные в лемме 7, можно показать, что $L^r(\mathbf{D}') = L^r(\mathbf{D}) \geqslant (n+1)\ln n$. Теорема доказана.

Комментарий. Заметим, что из возможностей выбора параметров a и Δ в матрице $\mathbf{A}$ следует, что класс стохастических $(m \times n)$-матриц, для которых выполняется условие (18), имеет континуальную мощность.

§ 6. Типичные значения

Из предыдущих результатов следует, что класс $\mathcal{P}(\bar{n}, \varepsilon)$ содержит матрицы сложности $L^*(\bar{n})$ только при $\varepsilon \leqslant \left(\prod_{i=1}^{m} n_i\right)^{-1}$. Поэтому представляет интерес как ведут себя длины имплицирующего вектора для почти всех матриц из $\mathcal{P}(\bar{n}, \varepsilon)$ при $\varepsilon \to 0$. В этом параграфе исследуется длина минимального имплицирующего вектора для почти всех матриц из классов матриц с рациональными элементами. Поскольку каждой такой стохастической матрице $\mathbf{P}$ можно сопоставить целочисленную квазистохастическую матрицу $\mathbf{P}'$ с суммой элементов строки равной N, где N — наименьшее общее кратное знаменателей элементов матрицы $\mathbf{P}$ и $L(\mathbf{P}) = L(\mathbf{P}')$, то в дальнейшем изложении вместо матрицы $\mathbf{P}$ будем рассматривать матрицу $\mathbf{P}'$ (напомним, что для $\mathbf{P}'$ в качестве ε можно взять число $1/N$).

Пусть $L^*(m, n) = m(n-1)+1$, $C(m, n, N)$ — множество всех целочисленных квазистохастических $(m \times n)$-матриц, сумма элементов каждой строки которых равна N, $C'(m, n, N)$ — множество всех матриц $\mathbf{A}$ из $C(m, n, N)$, для которых $L(\mathbf{A}) \leqslant L^*(m, n) - 1$.

Будем говорить, что почти все матрицы $\mathbf{A}$ из $C(m, n, N)$ обладают свойством E, если для множества $C_E(m, n, N)$ матриц из $C(m, n, N)$, обладающих свойством E, выполняется соотношение

$$\lim_{N \to \infty} (|C_E(m, n, N)|)/(|C(m, n, N)|) = 1.$$

Легко проверяется следующая

Лемма 10. $|C(m, n, N)| = \binom{N+n-1}{n-1}^m$.

В дальнейшем полагаем $m \geqslant 2$.

Теорема 7. а) *Если* $m = \mathrm{const}$, $n = \mathrm{const}$, $N \to \infty$, *то для почти всех* $\mathbf{A}$ *из* $C(m, n, N)$

$$L(A) = L^*(m, n).$$

б) *Если* $n = \mathrm{const}$, $m \to \infty$, $\ln N \geqslant m^2 n \ln n$, *то для почти всех* $\mathbf{A}$ *из* $C(m, n, N)$

$$L(\mathbf{A}) = L^*(m, n).$$

в) *Если* $m + n \to \infty$ *и* $\ln N \geqslant (m^2 + m) n \ln n$, *то для почти всех* $\mathbf{A}$ *из* $C(m, n, N)$

$$L(\mathbf{A}) = L^*(m, n).$$

Доказательство. Согласно (2) для любой матрицы $\mathbf{A}$ из $C(m, n, N)$ выполняется $L(\mathbf{A}) \leqslant L^*(m, n)$. Положим $R(m, n, N) = |C'(m, n, N)|/|C(m, n, N)|$. Для доказательства теоремы достаточно установить, что выполнение одного из условий теоремы влечет

$$R(m, n, N) \to 0 \tag{20}$$

при $N \to \infty$. Оценим сверху величину $|C'(m, n, N)|$. Для каждой матрицы $\mathbf{A}$ из $C'(m, n, N)$ можно указать имплицирующий вектор $\tilde{p} = (p_1, p_2, \ldots, p_L)$, где $L = m(n-1)$. Зафиксируем для i-й строки матрицы $\mathbf{A}$, $i = 1, 2, \ldots, m$ разбиение $T_i = \{T_{i,1}, T_{i,2}, \ldots, T_{i,n}\}$ множества $\{p_1, p_2, \ldots, p_L)$ элементов имплицирующего вектора $\tilde{p}$ такое, что

$$a_{i,j} = \sum_{p \in T_{i,j}} p, \quad j = 1, 2, \ldots, n.$$

Рассмотрим вектор $\tilde{p}' = (p_1', p_2', \ldots, p_L')$, полученный из $\tilde{p}$ отбрасыванием дробных частей его компонент $(p_i' = [p_i],\ i = 1, 2, \ldots, L)$. Очевидно, $\tilde{p}' \in C(1, L, N')$, где $N - L + 1 \leqslant N' \leqslant N$. Рассмотрим матрицу $\mathbf{A}' = \|a_{i,j}'\|$ из $C(m, n, N')$ с элементами

$$a_{i,j}' = \sum_{p' \in T_{i,j}} p'.$$

Пусть $\mathbf{A}'' = \mathbf{A} - \mathbf{A}'$ $(a_{i,j}'' = a_{i,j} - a_{i,j}')$. Очевидно, что $\mathbf{A}'' \in C(m, n, N - N')$. Поэтому для задания матрицы $\mathbf{A}$ достаточно задать разбиения $T_1, T_2, \ldots, T_m$ (не более n^{mL} способов), число N' (не более L способов), вектор $\tilde{p}'$ (не более $|C(1, L, N')|$ способов) и матрицу $\mathbf{A}''$ (не более

$|C(m, n, N-N')|$ способов). Используя лемму 10, получаем, что

$$|C'(m, n, N)| \leqslant n^{mL} \sum_{N-L+1<N'\leqslant N} \binom{N'+L-1}{L-1}\binom{N-N'+n-1}{n-1}^m =$$
$$= n^{mL}\sum_{s=0}^{L-1}\binom{N+L-1-s}{L-1}\binom{s+n-1}{n-1}^m.$$

Так как $\binom{a}{b} \leqslant \binom{a+1}{b}$, то любое слагаемое в последней сумме не превосходит $\binom{N+L-1}{L-1}\binom{L+n-2}{n-1}^m$. Поэтому

$$|C'(m, n, N)| \leqslant n^{mL} L \binom{N+L-1}{L-1}\binom{n+L-2}{n-1}^m. \tag{21}$$

Используя лемму 10 и очевидные неравенства

$$a^b(1-b/a)^b/b! \leqslant \binom{a}{b} \leqslant a^b/b!,$$

из (21) получаем

$$\ln R \leqslant mL\ln n + \ln L + (L-1)\ln(N+L-1) - \ln((L-1)!) +$$
$$+ m(n-1)\ln(L+n-2) - m\ln((n-1)!) - m(n-1)\ln(N+n-1) +$$
$$+ m\ln((n-1)!) - m(n-1)\ln(1-(n-1)/(N+n-1)). \tag{22}$$

Так как при $z \to 0$ $\ln(1+z) = z(1+o(1))$, то, используя условия доказываемой теоремы, имеем

$$\ln(N+L-1) = \ln N + \ln(1+(L-1)/(N+L-1)) = \ln N + O(L/N),$$
$$\ln(N+n-1) = \ln N + O(n/N),$$
$$-\ln(1-(n-1)/(N+n-1)) = O(n/N).$$

Поэтому из (22) следует, что

$$\ln R \leqslant mL\ln n + 2\ln L - \ln(L!) + L\ln(L+n-2) - \ln N + O(Ln/N).$$

Из последнего неравенства, используя вытекающее из формулы Стирлинга соотношение $\ln(a!) = a\ln a - a + o(a)$, получаем

$$\ln R \leqslant mL\ln n + 2\ln L + L + o(L) + L\ln(1+(n-2)/L) - \ln N + O(Ln/N).$$

Отсюда, используя неравенство $\ln(1+x) \leqslant x$, имеем

$$\ln R \leqslant mL\ln n + 2\ln L + L + o(L) + n - 2 - \ln N + O(Ln/N) \leqslant$$
$$\leqslant m^2 n\ln n - m^2\ln n + o(mn) + O(Ln/N) - \ln N.$$

Простой анализ полученного соотношения показывает, что при выполнении условий доказываемой теоремы $\ln(R(m, n, N)) \to -\infty$.

СПИСОК ЛИТЕРАТУРЫ

1. Бухараев Р. Г. Управляемые генераторы случайных величин // Вероятностные методы и кибернетика. Вып. 2.— Казань: Изд-во КГУ, 1963.— С. 68—87.
2. Бухараев Р. Г., Захаров В. М. Управляемые генераторы случайных кодов.— Казань: Изд-во КГУ, 1978.
3. Бухараев Р. Г. Основы теории вероятностных автоматов.— М.: Наука.— 1985.
4. Габбасов Н. З. О нахождении минимального имплицирующего вектора // Вероятностные методы и кибернетика.— Казань: Изд-во КГУ, 1984.— С. 29—40.
5. Габбасов Н. З. О минимальном имплицирующем векторе для линейных автоматов // Вероятностные автоматы и их приложения.— Казань: Изд-во КГУ, 1986.
6. Гиоргадзе А. Х., Джебашвили Т. Л., Сафиуллина А. Г. К вопросу определения имплицирующего вектора стохастических матриц // Сообщ. АН ГрузССР.— 1982.— Т. 108, № 1.— С. 49—52.

7. Костромин Г. Я. О нахождении имплицирующего вектора для стохастической матрицы // Деп. в ВИНИТИ 10.04.72, № 4248—72 Деп. Краткое сообщение: Автоматика и вычислительная техника.— 1972.— № 5.— С. 17—18.
8. Кузнецов С. Е, Нурмеев Н. Н., Салимов Ф. И. Задача о минимальном имплицирующем векторе // Тезисы межреспубл. научно-техн. конф. «Вероятностные автоматы и их приложения».— Тбилиси: Мецниереба, 1986.— С. 3.
9. Метра И. А. Замечания о минимальном имплицирующем векторе для стохастической матрицы // Автоматика и вычислительная техника.— 1970.— № 5.— С. 95—96.
10. Паршенков Н. Я., Ченцов В. М. Вопросы теории вероятностных автоматов // Автоматы и управление сетями связи.— М.: Наука.— 1971.— С. 180—202.
11. Паршенков Н. Я. Декомпозиционная модель вероятностного автомата // Построение управляющих устройств и систем.— М.: Наука, 1974.— С. 180—202.
12. Сачков В. Н. Введение в комбинаторные методы дискретной математики — М.: Наука, 1982.
13. Ченцов В. М. Об одном методе синтеза автономного автомата // Кибернетика.— 1968.— № 3 — С.32—35.
14. Чирков М. К., Шилкевич Т. П. О реализуемости вероятностных автоматов автоматами со случайными входами // Методы вычислений. Вып. 6.— Л : Изд-во ЛГУ, 1970.— С. 127—136.
15. Davis A. C. Markov chains as random input automata // Ann. Amer. Math. Monthly.— 1961.— V. 68, № 3 — P. 264—267.
16. Gelenbe S. E. A realizable model for stohastic sequential machines // IEEE Trans. on Comput.— 1971.— C — 20, № 2.— P. 199—204.
17. Kuznetsov S. E., Nurmeev N. N., Salimov F. I. The problem of minimal implicating vector // Lecture Notes in Comput. Sci.— 199: Fundamentals of Comput Th.— Berlin: Springer-Verlag.— 1987.— P. 273—278.