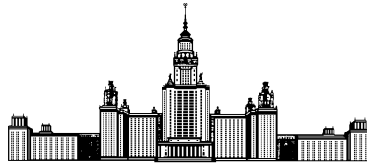


Московский государственный университет имени М. В. Ломоносова



Факультет Вычислительной Математики и Кибернетики

Кафедра Математических Методов Прогнозирования

## **ДИПЛОМНАЯ РАБОТА СТУДЕНТА 517 ГРУППЫ**

### **«Системы точек с вырожденными матрицами попарных расстояний»**

Выполнил:

студент 5 курса 517 группы

*Фигурнов Михаил Викторович*

Научный руководитель:

д.ф-м.н., профессор

*Дьяконов Александр Геннадьевич*

# Содержание

<b>1</b>	<b>Введение</b>	<b>2</b>
1.1	Определения и обозначения . . . . .	3
1.2	Обзор литературы . . . . .	6
<b>2</b>	<b>Результаты</b>	<b>8</b>
2.1	Метрический критерий $k$ -сингулярности . . . . .	8
2.2	Обобщение метрического критерия $k$ -сингулярности . . . . .	13
2.3	Обобщение на случай метрик разрезного конуса $CUT^q$ . . . . .	18
2.4	Лемма о равенстве линейных замыканий . . . . .	20
<b>3</b>	<b>Заключение</b>	<b>22</b>
<b>4</b>	<b>Список литературы</b>	<b>24</b>

# 1 Введение

Данная работа посвящена изучению свойства  $k$ -сингулярности конечной системы точек, впервые введённого в статье А.Г. Дьяконова [2]. Система  $q$  попарно различных точек пространства  $\mathbb{R}^m$  называется  $k$ -сингулярной, если размерность пространства значений полиномов (с поэлементным умножением) степени не выше  $k$  от столбцов матрицы попарных  $l_1$ -расстояний меньше  $q$ .

Ясно, что 1-сингулярность эквивалентна вырожденности матрицы попарных  $l_1$ -расстояний. Поскольку пространство полиномов степени 1 вложено в пространство полиномов степени не выше  $k$ , из  $k$ -сингулярности следует 1-сингулярность, то есть вырожденность матрицы попарных  $l_1$ -расстояний.

Отметим важный критерий из статьи [2], который связывает понятие  $k$ -сингулярности и теорию интерполяции функций, заданных на конечном множестве точек. Система точек не является  $k$ -сингулярной тогда и только тогда, когда любая функция на этой системе точек может быть представлена в виде суммы функций от  $k$  переменных.

Свойство  $k$ -сингулярности возникает и при изучении модели алгоритмов вычисления оценок (АВО) [5, 6]. Большую роль при изучении данной модели играет понятие корректности семейства, то есть возможности получить произвольную матрицу оценок, а, следовательно, произвольную классификацию объектов. В статье [3] показано, что можно построить систему точек, линейные комбинации столбцов матрицы  $l_1$ -расстояний которой полностью описывают матрицы оценок линейных комбинаций модели АВО. Таким образом, корректность множества полиномов степени не выше  $k$  над моделью АВО (или алгебраического замыкания степени  $k$ ) сводится к исследованию  $k$ -сингулярности соответствующей системы точек.

В статье [2] А.Г. Дьяконов доказал метрический критерий 1-сингулярности: 1-сингулярность эквивалентна линейной зависимости системы функций  $\rho(\tilde{s}, \tilde{s}_i)$ , где вместо второго аргумента подставляются точки системы, а  $\rho$  — метрика Хэмминга или  $l_1$ -метрика. Затем П.А. Карпович в статье [7] обобщил критерий на случай  $k$ -сингулярности. Оказывается, в этом случае нужно брать систему функций  $\rho^k(\tilde{s}, \tilde{s}_i)$ .

Возникает вопрос, верен ли критерий  $k$ -сингулярности для других метрик, например, для суммы  $l_1$ -метрики и метрики Хэмминга. Цель данной работы — исследование возможности обобщения упомянутого результата на другие метрики.

Сначала будет приведено авторское доказательство критерия  $k$ -сингулярности для метрики Хэмминга и  $l_1$ -метрики. С использованием предложенной методики будет рассмотрен случай, когда  $\rho$  — метрика Хэмминга или  $l_1$ -метрика, в которых значения близости по различным координатам оцениваются с разным весом. Мы будем называть такие метрики обобщёнными. Кроме того, будет рассмотрена сумма обобщённой метрики Хэмминга и обобщённой  $l_1$ -метрики. Затем мы перейдём к исследованию конуса разрезных полуметрик  $CUT^q$  [1]. Полуметрика принадлежит  $CUT^q$  тогда и только тогда, когда она изометрично вложима в  $l_1$ . В конце работы приведены исследования, связанные с устранением неточностей в доказательстве одной из теорем в статье [2].

## 1.1 Определения и обозначения

Будем следовать в обозначениях работе [2]. Приведём для удобства эти обозначения. Размерность векторов и матриц будут ясны из контекста.

Через  $\mathbb{N}$  обозначается множество натуральных чисел, через  $\mathbb{Z}$  — множество целых чисел, через  $\mathbb{R}$  — множество вещественных чисел.

Надстрочной тильдой обозначаются вектор-столбцы. У вектора  $\tilde{x}$  будем обозначать через  $(\tilde{x})_i$  его  $i$ -ю координату.

Введём удобные обозначения:

- $\tilde{1} = (1, \dots, 1)^T$  — вектор, состоящий из единиц;
- $\tilde{0} = (0, \dots, 0)^T$  — вектор, состоящий из нулей;
- $\tilde{e}_i = (0, \dots, 0, 1, 0, \dots, 0)^T$  — бинарный вектор,  $i$ -я координата которого равна единице, остальные координаты равны нулю.

**Определение 1.** Линейным замыканием  $L(X)$  множества векторов одинаковой размерности  $X$  называется множество всевозможных линейных комбинаций этих век-

торов:

$$L(X) = \left\{ \sum_{i=1}^n c_i \tilde{x}_i \mid c_1, \dots, c_n \in \mathbb{R}, x_1, \dots, x_n \in X, n \in \mathbb{N} \right\}$$

Пространство, ортогональное к пространству  $L(X)$ , будем обозначать через  $L^\perp(X)$ .

**Определение 2.** Адамаровым произведением  $\tilde{a} \circ \tilde{b}$  векторов  $\tilde{a}, \tilde{b} \in \mathbb{R}^n$  называется поэлементное произведение векторов:

$$\tilde{a} \circ \tilde{b} = ((\tilde{a})_1 \cdot (\tilde{b})_1, \dots, (\tilde{a})_n \cdot (\tilde{b})_n)^T$$

Операцию « $\circ$ » будем называть адамаровым умножением.

**Определение 3.** Алгебраическим замыканием  $k$ -й степени  $U^k(X)$  множества векторов одинаковой размерности  $X$  называется множество полиномов степени не выше  $k$  над множеством векторов с нулевым свободным членом относительно операции адамарова умножения:

$$U^k(X) = L(\{\tilde{x}_1 \circ \dots \circ \tilde{x}_s \mid \tilde{x}_1, \dots, \tilde{x}_s \in X, s \leq k\})$$

**Замечание.** Очевидно, что  $U^1(H) = L(H)$

Матрицы обозначаются заглавными латинскими буквами. Обозначим через  $E$  матрицу, состоящую из единиц.

**Определение 4.** Пусть  $H \in \mathbb{R}^{n \times m}$ ,  $H = [\tilde{h}_1, \dots, \tilde{h}_m]$ , то есть  $\tilde{h}_1, \dots, \tilde{h}_m$  —  $n$ -мерные вектор-столбцы матрицы. Обозначим

$$\begin{aligned} L(H) &\equiv L(\{\tilde{h}_1, \dots, \tilde{h}_m\}), \\ U^k(H) &\equiv U^k(\{\tilde{h}_1, \dots, \tilde{h}_m\}). \end{aligned}$$

Введём обозначение, позволяющее компактно записывать логические условия.

**Определение 5.** Индикаторной (характеристической) функцией предиката  $x$  называется функция

$$[x] = \begin{cases} 1, & x = \text{истина}, \\ 0, & x = \text{ложь}. \end{cases}$$

Определим ключевые для данной работы метрики, см. [1].

**Определение 6.** Метрикой Хэмминга называется функция  $\rho : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ :

$$\rho(\tilde{x}, \tilde{y}) = \sum_{i=1}^m [(\tilde{x})_i \neq (\tilde{y})_i].$$

**Замечание.** Формула записана в немного необычном виде. Это нужно для обобщения метрики в дальнейшем.

**Определение 7.**  $l_1$ -метрикой называется функция  $\rho : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ :

$$\rho(\tilde{x}, \tilde{y}) = \sum_{i=1}^m |(\tilde{x})_i - (\tilde{y})_i|.$$

В данной работе мы будем рассматривать систему попарно различных точек  $S = \{\tilde{s}_i\}_{i=1}^q$  пространства  $\mathbb{R}^m$ ,  $q \geq 2$ . Будем считать, что для каждой координаты существует две точки с различными значениями этой координаты. Если это не так, то отбросим соответствующую координату и понизим  $m$ . Также будем предполагать, что множество  $S$  упорядочено (порядок не имеет значения).

Обозначим через  $\{a_{t,0}, a_{t,1}, \dots, a_{t,p(t)}\}$  упорядоченное множество значений  $t$ -й координаты:

$$\{(\tilde{s}_i)_t \mid i \in \{1, 2, \dots, q\}\} = \{a_{t,0}, a_{t,1}, \dots, a_{t,p(t)}\}, a_{t,0} < a_{t,1} < \dots < a_{t,p(t)}.$$

В силу оговоренного выше,  $p(t) \geq 1$ ,  $t \in \{1, 2, \dots, m\}$ .

Пусть  $T^t = \|\theta_{ij}\|_{q \times (p(t)+1)}$  — бинарная матрица, в которой

$$\theta_{i,j} = [(\tilde{s}_i)_t = a_{t,j-1}], i \in \{1, 2, \dots, q\}, j \in \{1, 2, \dots, p(t) + 1\}.$$

Здесь было использовано определение 5 индикаторной функции.

**Определение 8.**  $T_S = [T^1 \dots T^m]$  (конкатенация по горизонтали матриц  $T^1, \dots, T^m$ ).

Введём обозначение:  $U^k[S] \equiv U^k(T_S)$ .

**Определение 9.** Столбцы матрицы  $T_S$  называются характеристическими векторами системы точек  $S$ .

**Определение 10** (Reid, Sun [8]). Характеристической матрицей  $H_S$  системы точек  $S$  называется матрица

$$H_S = T_S \cdot T_S^T,$$

умножение матричное.

**Определение 11.**  $P_S$  — матрица попарных расстояний Хэмминга системы точек  $S$ .

**Определение 12.**  $P'_S$  — матрица попарных  $l_1$ -расстояний системы точек  $S$ .

Введём ключевое для данной работы определение:

**Определение 13** (Дьяконов [2]). Система точек  $S$  называется  $k$ -сингулярной, если размерность пространства  $U^k[S]$  меньше  $q$ .

## 1.2 Обзор литературы

Приведём обзор результатов, полученных по данной теме ранее в работе [2].

Доказана простая лемма, связывающая характеристическую матрицу и матрицу попарных расстояний Хэмминга:

**Лемма 1** (Дьяконов [2]).  $P_S = tE - H_S$ .

Доказана лемма, из которой следует, что матрица попарных расстояний Хэмминга вырождена тогда и только тогда, когда вырождена матрица попарных  $l_1$ -расстояний:

**Лемма 2** (Дьяконов [2]). *Для функций*

$$F_k(x) = \frac{1}{k!} \prod_{j=1}^k (x - j + 1), \quad G_k(x) = x^k$$

*и системы попарно различных точек  $S = \{\tilde{s}_i\}_{i=1}^q \subseteq \mathbb{R}^m, q \geq 2$ , при  $k \leq m$  справедливы равенства*

$$U^k(P_S) = L(F_k(H_S)) = U^k(T_S) = L(G_k(H'_S)) = U^k(P'_S).$$

Учитывая определения  $k$ -сингулярности, получаем из этой леммы, что 1-сингулярность системы точек соответствует вырожденности матрицы попарных  $l_1$ -расстояний. Также можно сказать, что  $k$ -сингулярность формализует «полноту» пространства полиномов степени не выше  $k$  над столбцами матрицы попарных  $l_1$ -расстояний: система из  $q$  точек является  $k$ -сингулярной, если размерность пространства таких полиномов меньше  $q$ .

Доказана теорема, показывающая, что по системе точек строится новая система, которая 1-сингулярна тогда и только тогда, когда  $k$ -сингулярна исходная:

**Теорема 1** (Дьяконов [2]). Пусть  $k \leq m$  и  $z$  — биективное отображение множества  $\{1, 2, \dots, C_m^k\}$  на множество  $\{\{i_1, \dots, i_k\} \mid 1 \leq i_1 < \dots < i_k \leq m\}$ . Система точек  $S = \{\tilde{s}_i\}_{i=1}^q$  является  $k$ -сингулярной тогда и только тогда, когда является 1-сингулярной система точек  $D = \{\tilde{d}_i\}_{i=1}^q$  пространства  $\mathbb{R}^{C_m^k}$  такая, что для всех  $t \in \{1, 2, \dots, C_m^k\}$  справедливо

$$(\tilde{d}_i)_t = (\tilde{d}_j)_t \Leftrightarrow \forall r \in z(t) (\tilde{s}_i)_r = (\tilde{s}_j)_r.$$

Получен критерий  $k$ -сингулярности, связывающий  $k$ -сингулярность с представимостью функции многих переменных, заданной на конечном множестве точек  $\mathbb{R}^m$ , в виде суммы функций  $k$  переменных:

**Теорема 2** (Дьяконов [2]). Система точек  $S$  не является  $k$ -сингулярной при  $k \leq m$  тогда и только тогда, когда любая функция  $f(x_1, x_2, \dots, x_m)$  на точках системы  $S$  может быть представлена в виде конечной суммы функций, каждая из которых зависит от  $k$  переменных.

Также получен геометрический критерий  $k$ -сингулярности: система точек не является  $k$ -сингулярной тогда и только тогда, когда при любом её разбиении на две непересекающиеся подсистемы они разделимы при помощи гиперплоскости с предварительным действием элемента некоторой группы преобразований пространства.

Доказан метрический критерий 1-сингулярности:

**Теорема 3** (Дьяконов [2]). Система точек  $S = \{\tilde{s}_i\}_{i=1}^q$  пространства  $\mathbb{R}^m$  является 1-сингулярной тогда и только тогда, когда существует ненулевой вектор



$(c_1, \dots, c_q)$  такой, что для всех  $\tilde{s} \in \mathbb{R}^m$  справедливо равенство

$$\sum_{i=1}^q c_i \rho(\tilde{s}, \tilde{s}_i) = 0, \quad (1)$$

где  $\rho$  – метрика Хэмминга или  $l_1$ -метрика.

Также получен основной критерий  $k$ -сингулярности: система точек  $k$ -сингулярна тогда и только тогда, когда содержит непустой носитель суммы размеченных параллелепипедов (это обобщение понятия «размеченный прямоугольник»).

В работе [7] П.А. Карпович обобщил метрический критерий 1-сингулярности на случай  $k$ -сингулярности. Оказывается, что для этого достаточно заменить условие (1) на

$$\sum_{i=1}^q c_i \rho^k(\tilde{s}, \tilde{s}_i) = 0.$$

## 2 Результаты

### 2.1 Метрический критерий $k$ -сингулярности

В данном разделе приведено авторское доказательство метрического критерия  $k$ -сингулярности, впервые полученного П.А. Карповичем в [7].

**Лемма 3.** Для системы попарно различных точек  $S = \{\tilde{s}_i\}_{i=1}^q \subseteq \mathbb{R}^m, q \geq 2$ , справедливы равенства

$$U^k[S] = U^k(P_S) = U^k(P'_S).$$

**Доказательство.** При  $k \leq m$  данная лемма является следствием леммы 2. При  $k > m$  утверждение очевидно, поскольку  $U^m[S] = U^m(P_S) = U^m(P'_S) = \mathbb{R}^m$  (все точки в системе  $S$  попарно различны).  $\square$

**Лемма 4.** Для системы попарно различных точек  $S = \{\tilde{s}_i\}_{i=1}^q \subseteq \mathbb{R}^m, q \geq 2$ , при  $k \leq m$  справедливы равенства

$$U^k[S] = L(\{\tilde{t}_1 \circ \dots \circ \tilde{t}_k | \tilde{t}_i \in T^{z(i)}, z(i) \neq z(j) \text{ при } i \neq j, \forall i, j \in \{1, \dots, k\}\}). \quad (2)$$

**Доказательство.** По определению,  $U^k[S] \equiv L(\{\tilde{t}_1 \circ \dots \circ \tilde{t}_n | \tilde{t}_i \in T_S \forall i \in \{1, \dots, n\}, n \leq k\})$ , поэтому вложение  $\supseteq$  в (2) очевидно.

Докажем вложение  $\subseteq$ . В силу линейности достаточно доказать его для мономов из  $U^k[S]$ . Пусть  $\tilde{0} \neq \tilde{t} \in U^k[S]$  — моном, то есть  $\tilde{t} = \tilde{t}_1 \circ \dots \circ \tilde{t}_n, \tilde{t}_i \in T^{z(i)}, n \leq k$ . Если  $z(i) = z(j), i \neq j$ , то  $\tilde{t}_i = \tilde{t}_j$ , поскольку  $\tilde{t} \neq \tilde{0}$  по предположению, а различные столбцы  $T^{z(i)}$  соответствуют различным классам эквивалентности, то есть их произведение равно нулевому вектору. Таким образом, можно потребовать, чтобы  $z(i) \neq z(j)$  при  $i \neq j$ .

Пусть  $n < k$ . Поскольку сумма столбцов любой подматрицы  $T^i$  равна  $\tilde{1}, \forall \tilde{x} \in \mathbb{R}^m \forall i \in \{1, \dots, m\} \tilde{x} = \sum_{\tilde{t}' \in T^i} \tilde{x} \circ \tilde{t}'$ . Обозначим элементы множества  $\{\{1, 2, \dots, k\} \setminus \{z(1), \dots, z(n)\}\}$ , занумерованные в любом порядке, через  $z'(1), \dots, z'(k-n)$ . Получим:

$$\tilde{t} = \sum_{\tilde{t}'_1 \in T^{z'(1)}} \sum_{\tilde{t}'_2 \in T^{z'(2)}} \dots \sum_{\tilde{t}'_{k-n} \in T^{z'(k-n)}} \tilde{t}'_1 \circ \tilde{t}'_2 \circ \dots \circ \tilde{t}'_{k-n} \circ \tilde{t}_1 \circ \dots \circ \tilde{t}_n$$

Эта сумма принадлежит линейному замыканию из правой части (2).

□

**Определение 14.** Функция  $f(\tilde{x})$  называется полиномом степени  $n = 0, 1, 2, \dots$ , если она представима в виде  $f(\tilde{x}) = \sum_{i=0}^n a_i \cdot \tilde{x}^i$ , причём  $a_n \neq 0$ . Операция возведения в степень поэлементная.

**Лемма 5.** Пусть  $f(\tilde{x})$  — полином степени  $n \geq 1$ ,  $\tilde{y} \neq \tilde{0}$  — бинарный вектор. Тогда  $f(\tilde{x} - \tilde{y}) - f(\tilde{x}) = \tilde{y} \circ g(\tilde{x})$ , где  $g(\tilde{x})$  есть полином степени  $(n-1)$ .

**Доказательство.**

$$\begin{aligned}
f(\tilde{x} - \tilde{y}) - f(\tilde{x}) &= \sum_{i=0}^n a_i \cdot ((\tilde{x} - \tilde{y})^i - \tilde{x}^i) = \\
&= \sum_{i=0}^n a_i \cdot \left( \sum_{j=0}^i (-1)^{i-j} \binom{i}{j} \cdot \tilde{x}^j \circ \tilde{y}^{i-j} - \tilde{x}^i \right) = \\
&= a_0 \cdot (\tilde{1} - \tilde{1}) + \sum_{i=1}^n a_i \cdot \left( \sum_{j=0}^{i-1} (-1)^{i-j} \binom{i}{j} \cdot \tilde{x}^j \circ \tilde{y}^{i-j} + \tilde{x}^i - \tilde{x}^i \right) = \\
&= \sum_{i=1}^n a_i \cdot \left( \sum_{j=0}^{i-1} (-1)^{i-j} \binom{i}{j} \cdot \tilde{x}^j \circ \tilde{y}^{i-j} \right) \stackrel{(*)}{=} \{\tilde{y} \text{ бинарный вектор}\} = \\
&= \sum_{i=1}^n a_i \cdot \left( \sum_{j=0}^{i-1} (-1)^{i-j} \binom{i}{j} \cdot \tilde{x}^j \circ \tilde{y} \right) = \tilde{y} \circ \sum_{i=1}^n a_i \cdot \left( \sum_{j=0}^{i-1} (-1)^{i-j} \binom{i}{j} \cdot \tilde{x}^j \right) = \\
&= \tilde{y} \circ g(\tilde{x}).
\end{aligned}$$

Обозначение (\*) будет использовано при доказательстве леммы 7.

$g(\tilde{x})$  — полином степени  $(n - 1)$ , поскольку коэффициент при  $\tilde{x}^{n-1}$  равен

$$a_n (-1)^{n-(n-1)} \binom{n}{n-1} = -na_n \neq 0.$$

□

**Теорема 4** (метрический критерий  $k$ -сингулярности). Система точек  $S = \{\tilde{s}_i\}_{i=1}^q$  пространства  $\mathbb{R}^m$  является  $k$ -сингулярной тогда и только тогда, когда существует ненулевой вектор  $(c_1, \dots, c_q)$  такой, что для всех  $\tilde{s} \in \mathbb{R}^m$  справедливо равенство

$$\sum_{i=1}^q c_i \rho^k(\tilde{s}, \tilde{s}_i) = 0, \quad (3)$$

где  $\rho$  — метрика Хэмминга или  $l_1$ -метрика.

**Доказательство.** Введём вспомогательные обозначения:

$$\begin{aligned}
\tilde{\rho}^k(\tilde{s}, S) &= (\rho^k(\tilde{s}, \tilde{s}_1), \dots, \rho^k(\tilde{s}, \tilde{s}_q))^T, \\
R^k[S] &= L(\{\rho^k(\tilde{s}, S) | \tilde{s} \in \mathbb{R}^m\}).
\end{aligned}$$

**Необходимость.** Метрика Хэмминга. Пусть  $S$   $k$ -сингулярна, т.е.  $\exists \tilde{c} = (c_1, \dots, c_q) \neq \tilde{0}$ :  $\tilde{c} \in L^\perp(U^k[S])$ . Рассмотрим  $\tilde{s} \in S$ .  $\tilde{\rho}(\tilde{s}, S)$  есть столбец  $P_S$ , поэтому  $\tilde{\rho}^k(\tilde{s}, S) \in U^k(P_S)$ . В силу леммы 3,  $U^k(P_S) \subseteq U^k[S]$ . Следовательно,  $L^\perp(U^k[S]) \subseteq L^\perp(U^k(P_S))$ . Таким образом, выполнено (3).

Пусть теперь  $\tilde{s} \notin S$ . Очевидно, что  $(c_1, \dots, c_q, 0) \in L^\perp(U^k[S \cup \{\tilde{s}\}]) \subseteq L^\perp(U^k(P_{S \cup \{\tilde{s}\}}))$ . Следовательно, выполнено равенство  $(c_1, \dots, c_q, 0)P_{S \cup \{\tilde{s}\}}^k = \tilde{0}$ . Рассматривая умножение строки на столбец, соответствующий  $\tilde{s}$ , получаем требуемое утверждение.

*l<sub>1</sub>-метрика.* Доказывается аналогично случаю метрики Хэмминга при замене матрицы  $P_S$  на  $P'_S$ .

**Достаточность.** Нужно доказать, что  $L^\perp(R^k[S]) \subseteq L^\perp(U^k[S])$ . Докажем эквивалентное утверждение:  $U^k[S] \subseteq R^k[S]$ . Пусть  $k \leq m$  и  $\tilde{0} \neq \tilde{w} = \tilde{w}_1 \circ \dots \circ \tilde{w}_k \in U^k[S]$ , где  $\tilde{w}_i \in T^{z(i)}$ ,  $z(i) \neq z(j)$  при  $i \neq j$ . В силу леммы 4 достаточно показать, что  $\tilde{w} \in R^k[S]$ .

Поскольку  $\tilde{w} \neq \tilde{0}$ , выполнено  $\tilde{w}_i \neq \tilde{0}$ ,  $i = 1, \dots, k$  и  $\exists \alpha: (\tilde{w}_1)_\alpha \neq 0, \dots, (\tilde{w}_k)_\alpha \neq 0$ . Это означает, что вектор  $\tilde{w}$  представим в виде произведения столбцов  $T_S$ , соответствующих  $k$  координатам точки  $\tilde{s}_\alpha$ .

*Метрика Хэмминга.* Пусть  $\tilde{s} = \tilde{s}_\alpha$ ,  $\tilde{s}' = \tilde{s} + \varepsilon \tilde{e}_{z(1)}$ , причём  $\varepsilon$  возьмём таким, что  $(\tilde{s}')_{z(1)} \notin \{(\tilde{s}_i)_{z(1)} | i \in \{1, \dots, m\}\}$ .

Заметим, что

$$\forall \tilde{u} \in \mathbb{R}^m \quad \tilde{\rho}^k(\tilde{u}, S) = \left( \sum_{i=1}^m (\tilde{1} - \tilde{t}_i) \right)^k = \left( m \cdot \tilde{1} - \sum_{i=1}^m \tilde{t}_i \right)^k, \quad (4)$$

где  $\tilde{t}_i$  — соответствующий столбец матрицы  $T^i$ , если координата  $(\tilde{u})_i$  встречается среди  $i$ -х координат точек  $S$ , иначе  $\tilde{0}$ . Это выражение есть полином степени  $k$  от  $\sum_{i=1}^m \tilde{t}_i$ .

Ясно, что в введённых обозначениях  $\tilde{u}$  и  $\tilde{s}'$  будут совпадать все векторы  $\tilde{t}_i$ , кроме  $z(1)$ -ого, который у  $\tilde{s}'$  будет нулевым. Применим лемму 5 при  $\tilde{x} = \sum_{i=1}^m \tilde{t}_i$ ,  $\tilde{y} = \tilde{t}_{z(1)} \equiv \tilde{w}_1$ ,  $f(\tilde{x}) = (m\tilde{1} - \tilde{x})^k$ . Получим, что  $\tilde{\rho}^k(\tilde{s}, S) - \tilde{\rho}^k(\tilde{s}', S) = \tilde{w}_1 \circ h(\tilde{x}) \in R^k[S]$ , где  $h(\tilde{x})$  — полином степени  $(k-1)$ .

На следующем шаге возьмём  $\tilde{s} = \tilde{w}_1 \circ h(\tilde{x})$ ,  $\tilde{s}' = \tilde{s} + \varepsilon \tilde{e}_{z(2)}$ ,  $\varepsilon$  выберем аналогично. Применяя лемму 5, получим, что  $\tilde{w}_1 \circ \tilde{w}_2 \circ h(\tilde{x}) \in R^k[S]$ , где  $h(\tilde{x})$  есть полином степени  $(k-2)$ .

Повторим процедуру  $(k-2)$  раз для  $\tilde{w}_3, \dots, \tilde{w}_k$ . Получим, что  $\tilde{w}_1 \circ \dots \circ \tilde{w}_k \circ h(\tilde{x}) \in R^k[S]$ , где  $h(\tilde{x})$  — полином 0 степени, то есть ненулевая константа (см. определение 14). Значит,  $\tilde{w} \in R^k[S]$ .

*l<sub>1</sub>-метрика.* Очевидно, что

$$\forall \tilde{u} \in \mathbb{R}^m \quad \tilde{\rho}^k(\tilde{u}, S) = \left( \sum_{i=1}^m \sum_{\tilde{t} \in T^i} |(\tilde{u})_i - \text{coord}(\tilde{t})| \cdot \tilde{t} \right)^k, \quad (5)$$

где  $coord(\tilde{t})$  — значение координаты, которой соответствует столбец  $\tilde{t}$  матрицы  $T^i$ . Таким образом,  $\tilde{\rho}^k(\tilde{u}, S)$  имеет производную любого порядка везде, кроме конечного числа точек, где выражение под одним из модулей меняет знак.

Распишем производную  $k$  порядка по координатам  $z(1), \dots, z(k)$  в точке дифференцируемости:

$$\frac{\partial^k}{\partial u_{z(1)} \partial u_{z(2)} \dots \partial u_{z(k)}} \tilde{\rho}^k(\tilde{u}, S) = k! \prod_{i=1}^k \sum_{\tilde{t} \in T^{z(i)}} \text{sgn}((\tilde{u})_{z(i)} - coord(\tilde{t})) \cdot \tilde{t}$$

Рассмотрим  $k$ -мерный булев куб и сопоставим каждой его точке  $x_1, x_2, \dots, x_k$  точку пространства  $\mathbb{R}^m$

$$\tilde{s}_{x_1, x_2, \dots, x_k} = \tilde{s}_\alpha + (-1)^{x_1} \varepsilon \tilde{e}_{z(1)} + (-1)^{x_2} \varepsilon \tilde{e}_{z(2)} + \dots + (-1)^{x_k} \varepsilon \tilde{e}_{z(k)},$$

где  $\varepsilon > 0$  — малое число, такое, что внутри «куба» находится лишь точка  $\tilde{s}_\alpha$  и во всех точках  $\tilde{s}_{x_1, x_2, \dots, x_k}$  функция  $\tilde{\rho}^k(\tilde{u}, S)$  дифференцируема. Обозначим производную  $\tilde{\rho}^k(\tilde{u}, S)$  в точке  $x_1, x_2, \dots, x_k$  через  $p(x_1, x_2, \dots, x_k)$ .

Докажем, что  $p(x_1, x_2, \dots, x_k) \in R^k[S]$ . Для этого возьмём шар  $A$  достаточно малого радиуса, содержащий точку<sup>1</sup>  $s_{x_1, x_2, \dots, x_k}$  и не содержащий другие точки системы  $S$ . В силу непрерывности отображения  $\tilde{\rho}^k(\tilde{u}, S)$ , образ шара  $A$  при воздействии этим отображением есть замкнутое множество. Его линейное замыкание  $L = L(\{\tilde{\rho}^k(\tilde{u}, S) \mid \tilde{u} \in A\})$  также будет замкнутым множеством. Но производная  $p(x_1, x_2, \dots, x_k)$  есть предельная точка  $L$ . Осталось вспомнить, что по определению  $\tilde{\rho}^k(\tilde{u}, S) \in R^k[S]$ , следовательно,  $L \subseteq R^k[S]$ .

---

<sup>1</sup>В доказательстве из курсовой работы шаром накрывались сразу все точки. Но этот вариант не подходит для случая суммы метрик Хэмминга и  $l_1$ , поскольку такая метрика не является всюду непрерывной.

Заметим теперь, что  $\forall x_2, \dots, x_k \in B^{k-1}$ :

$$\begin{aligned}
& p(0, x_2, \dots, x_k) - p(1, x_2, \dots, x_k) = \\
& = k! \sum_{\tilde{t} \in T^{z(1)}} \left( \operatorname{sgn} \left( (\tilde{s}_\alpha)_{z(1)} + \varepsilon \tilde{e}_{z(1)} - \operatorname{coord}(\tilde{t}) \right) - \operatorname{sgn} \left( (\tilde{s}_\alpha)_{z(1)} - \varepsilon \tilde{e}_{z(1)} - \operatorname{coord}(\tilde{t}) \right) \right) \cdot \tilde{t} \circ \\
& \circ \prod_{i=2}^k \sum_{\tilde{t} \in T^{z(i)}} \operatorname{sgn} \left( (\tilde{s}_\alpha)_{z(i)} + (-1)^{x_i} \varepsilon \tilde{e}_{z(i)} - \operatorname{coord}(\tilde{t}) \right) \cdot \tilde{t} = \\
& = k! \cdot 2\tilde{w}_1 \circ \prod_{i=2}^k \sum_{\tilde{t} \in T^{z(i)}} \operatorname{sgn} \left( (\tilde{s}_\alpha)_{z(i)} + (-1)^{x_i} \varepsilon \tilde{e}_{z(i)} - \operatorname{coord}(\tilde{t}) \right) \cdot \tilde{t} = \\
& = p_1(x_2, \dots, x_k)
\end{aligned}$$

Итак,  $p_1(x_2, \dots, x_k)$  задана на  $B^{k-1}$ , причём её значения представимы в виде  $\tilde{w}_1 \circ \hat{p}_1(x_2 \dots x_k)$ . Легко убедиться, что  $p_1(0, x_3, \dots, x_k) - p_1(1, x_3, \dots, x_k) = p_2(x_3, \dots, x_k) = \tilde{w}_1 \circ \tilde{w}_2 \circ \hat{p}_2(x_3 \dots x_k)$ . Продолжая, получим, что  $\tilde{w}_1 \circ \dots \circ \tilde{w}_k = \tilde{w} \in R^k[S]$ .

Опишем теперь построение, позволяющее доказать теорему при  $k > m$  для обеих метрик. Пусть  $k > m, i \in \{1, \dots, m\}$ . Покажем, что  $\tilde{e}_i \in R^k[S]$ . Для этого возьмём  $\tilde{s}_\alpha = \tilde{s}_i$ . Из свойств матрицы  $T_S$  и того, что в системе точек  $S$  нет двух точек с одинаковыми координатами, вектор  $\tilde{e}_i$  равен произведению  $m$  столбцов матрицы  $T_S$ , соответствующих этой точке. Используя вышеописанную процедуру, убеждаемся, что  $\tilde{e}_i \in R^k[S]$ . Итак,  $\tilde{e}_1, \dots, \tilde{e}_m \in R^k[S]$ , следовательно,  $R^k[S] = \mathbb{R}^q$ . Отсюда  $\tilde{w} \in R^k[S]$ .  $\square$

## 2.2 Обобщение метрического критерия $k$ -сингулярности

Попробуем обобщить метрический критерий  $k$ -сингулярности (теорема 4) на более общий случай. Мотивацией к этому является то, что метрика Хэмминга и  $l_1$ -метрика, заданные на  $q$  точках, принадлежат разрезному конусу  $CUT^q$  [1].

Введём соответствующие определения.

**Определение 15.** Пусть  $\tilde{d} \in \mathbb{R}^m$ . Функцию

$$\rho(\tilde{x}, \tilde{y}) = \sum_{i=1}^m (\tilde{d})_i [(\tilde{x})_i \neq (\tilde{y})_i], \quad (\tilde{d})_i \geq 0 \quad \forall i \in \{1, \dots, m\}$$

назовём *обобщённой полуметрикой Хэмминга*. Если же  $(\tilde{d})_i > 0 \quad \forall i \in \{1, \dots, m\}$ , то будем называть её *обобщённой метрикой Хэмминга*.

В предыдущем определении использовалось определение 5 (индикаторная функция  $[x]$ ).

**Определение 16.** Пусть  $\tilde{d}' \in \mathbb{R}^m$ . Функцию

$$\rho(\tilde{x}, \tilde{y}) = \sum_{i=1}^m (\tilde{d}')_i |(\tilde{x})_i - (\tilde{y})_i|, \quad (\tilde{d}')_i \geq 0 \quad \forall i \in \{1, \dots, m\},$$

назовём *обобщённой  $l_1$ -полуметрикой*. Если же  $(\tilde{d}')_i > 0 \quad \forall i \in \{1, \dots, m\}$ , то будем называть её *обобщённой  $l_1$ -метрикой*.

Матрицу попарных расстояний обобщённой метрики Хэмминга будем обозначать  $P_{S, \tilde{d}}$ , обобщённой  $l_1$ -метрики —  $P'_{S, \tilde{d}'}$ .

Покажем, что метрический критерий  $k$ -сингулярности обобщается на случай этих двух метрик, а также случай их суммы.

Сначала докажем аналог леммы 3 для обобщённых полуметрик Хэмминга и  $l_1$ .

**Лемма 6.** Для системы попарно различных точек  $S = \{\tilde{s}_i\}_{i=1}^q \subseteq \mathbb{R}^m$ ,  $q \geq 2$ , а также произвольных векторов  $\tilde{d}, \tilde{d}' \in \mathbb{R}^m$ ,  $(\tilde{d})_i \geq 0, (\tilde{d}')_i \geq 0 \quad \forall i \in \{1, \dots, m\}$  справедливы равенства

$$U^k(P_{S, \tilde{d}}) \subseteq U^k[S] \tag{6}$$

$$U^k(P'_{S, \tilde{d}'}) \subseteq U^k[S] \tag{7}$$

**Доказательство.** Рассмотрим случай обобщённой метрики Хэмминга. В начале раздела 4 работы [4] доказано, что  $L(P_{S, \tilde{d}}) = L(H)$ , где  $H$  — обобщённая характеристическая матрица системы точек. Для неё известно, что  $L(H) \subseteq U^1[S]$ , поскольку  $H$  есть матрица Грама для матрицы  $T_S$ , столбцы которой были домножены на неположительные коэффициенты  $(\tilde{d})_i$  (если все компоненты  $\tilde{d}$  положительны, то  $L(H) = U^1[S]$ ). Из этих двух равенств следует первое утверждение леммы.

Для случая обобщённой  $l_1$ -метрики проведём следующее рассуждение. Рассмотрим вместе с исходной системой точек  $S = \{\tilde{s}_i\}_{i=1}^q$  вспомогательную «отмасштабированную» систему точек  $S' = \{\tilde{d}' \circ \tilde{s}_i\}_{i=1}^q$ . У координат, соответствующих ненулевым компонентам  $(\tilde{d}')_i$ , характеристические векторы не изменились. У координат, соответствующих нулевым компонентам характеристические векторы стали равными  $\tilde{1}$ , то есть вырожденными. Но  $\tilde{1} \in L(T_S)$  как сумма компонент любого из блоков  $T^i$ .

Отсюда  $L(T_{S'}) \subseteq L(T_S)$ . Кроме того, матрица  $P'_{S'}$  попарных  $l_1$ -расстояний системы точек  $S'$  совпадает с матрицей  $P'_{S, \tilde{d}'}$ . Поэтому  $U^k(P'_{S, \tilde{d}'}) = U^k(P'_{S'}) = U^k[S'] \subseteq U^k[S]$ , откуда следует второе утверждение леммы.  $\square$

Докажем теперь обобщение леммы 5 на случай вектора  $\tilde{y}$ , компоненты которого принимают два значения: 0 и  $b > 0$ .

**Лемма 7.** Пусть  $f(\tilde{x})$  — полином степени  $n \geq 1$ ,  $\tilde{u} \neq \tilde{0}$  — бинарный вектор,  $\tilde{y} = b \cdot \tilde{u}$ ,  $b > 0$ . Тогда  $f(\tilde{x} - \tilde{y}) - f(\tilde{x}) = \tilde{u} \circ g(\tilde{x})$ , где  $g(\tilde{x})$  есть полином степени  $(n - 1)$ .

**Доказательство.** Доказательство проводится аналогично лемме 5 до равенства, обозначенного (\*). Часть доказательства после этого равенства заменяется на следующее:

$$\begin{aligned} &\stackrel{(*)}{=} \sum_{i=1}^n a_i \cdot \left( \sum_{j=0}^{i-1} (-1)^{i-j} \binom{i}{j} b^{i-j} \cdot \tilde{x}^j \circ \tilde{u} \right) = \\ &= \tilde{u} \circ \sum_{i=1}^n a_i \cdot \left( \sum_{j=0}^{i-1} (-b)^{i-j} \binom{i}{j} \cdot \tilde{x}^j \right) = \\ &= \tilde{u} \circ g(\tilde{x}). \end{aligned}$$

$g(\tilde{x})$  — полином степени  $(n - 1)$ , поскольку коэффициент при  $\tilde{x}^{n-1}$  равен

$$a_n (-b)^{n-(n-1)} \binom{n}{n-1} = -n a_n b \neq 0.$$

$\square$

**Лемма 8.** Пусть  $A, B$  — матрицы одинакового размера, такие что  $L(A) = L(B)$ . Тогда  $L(A + B) \subseteq L(A)$ .

**Доказательство.** Докажем два простых факта.

1.  $L(A + B) \subseteq L(A \cup B)$ . Это следует из того, что произвольный вектор из  $L(A + B)$  представим в виде линейной комбинации столбцов  $A$  и  $B$ .
2.  $L(A) = L(B) \Rightarrow L(A \cup B) = L(A)$ . Вложение  $L(A \cup B) \supseteq L(A)$  очевидно в силу монотонности линейного замыкания. Обратное вложение верно, так как любой вектор  $\tilde{x}$  из  $L(A \cup B)$  представим в виде  $\tilde{y} + \tilde{z}$ , где  $\tilde{y} \in L(A)$ ,  $\tilde{z} \in L(B)$ , но по предположению  $L(A) = L(B)$ , то есть  $\tilde{x} = \tilde{y} + \tilde{z} \in L(A)$ .



Объединяя эти утверждения, получаем требуемое.  $\square$

**Лемма 9.** Для системы попарно различных точек  $S = \{\tilde{s}_i\}_{i=1}^q \subseteq \mathbb{R}^m$ ,  $q \geq 2$ , при  $k \leq m$ , а также произвольных векторов  $\tilde{d}, \tilde{d}' \in \mathbb{R}^m$ ,  $(\tilde{d})_i > 0$ ,  $(\tilde{d}')_i > 0 \forall i \in \{1, \dots, m\}$  справедливо

$$U^k(P_{S,\tilde{d}} + P'_{S,\tilde{d}'}) \subseteq U^k[S]$$

**Доказательство.** Применим предыдущую лемму при  $A = P_{S,\tilde{d}}$ ,  $B = P'_{S,\tilde{d}'}$ . Получим:

$$L(P_{S,\tilde{d}} + P'_{S,\tilde{d}'}) \subseteq L(P_{S,\tilde{d}})$$

Откуда, в силу монотонности операции линейного замыкания,

$$U^k(P_{S,\tilde{d}} + P'_{S,\tilde{d}'}) \subseteq U^k(P_{S,\tilde{d}}).$$

По лемме 6,  $U^k(P_{S,\tilde{d}}) = U^k[S]$ .  $\square$

**Теорема 5** (обобщённый метрический критерий  $k$ -сингулярности). Система точек  $S = \{\tilde{s}_i\}_{i=1}^q$  пространства  $\mathbb{R}^m$  является  $k$ -сингулярной тогда и только тогда, когда существует ненулевой вектор  $(c_1, \dots, c_q)$  такой, что для всех  $\tilde{s} \in \mathbb{R}^m$  справедливо равенство

$$\sum_{i=1}^q c_i \rho^k(\tilde{s}, \tilde{s}_i) = 0, \quad (8)$$

где  $\rho$  — обобщённая метрика Хэмминга, обобщённая  $l_1$ -метрика, либо их сумма.

**Замечание.** Случай суммы многих обобщённых метрик Хэмминга и обобщённых  $l_1$ -метрик, возможно, с неотрицательными коэффициентами при них, сводится к случаю суммы одной обобщённой метрики Хэмминга и одной обобщённой  $l_1$ -метрики. поскольку коэффициенты при метриках можно соответствующим образом сгруппировать.

**Доказательство.** Проводится по схеме доказательства теоремы 4. Укажем, какие изменения в нём нужно произвести.

**Необходимость.** Для обобщённой метрики Хэмминга и обобщённой  $l_1$ -метрики достаточно заменить матрицу  $P_S$  на соответствующую матрицу попарных расстояний и воспользоваться вместо леммы 3 леммой 6.

Для случая суммы метрик нужно заметить, что в силу леммы 9:

$$L^\perp(U^k[S]) \subseteq L^\perp(U^k(P_{S,\tilde{d}} + P'_{S,\tilde{d}'})).$$

**Достаточность.** *Обобщённая метрика Хэмминга.* Пусть  $\tilde{d} \in \mathbb{R}^m$ ,  $(\tilde{d})_i > 0 \forall i \in \{1, \dots, m\}$  — вектор коэффициентов обобщённой метрики Хэмминга. Формула (4) заменяется на следующую формулу:

$$\forall \tilde{u} \in \mathbb{R}^m \quad \tilde{\rho}^k(\tilde{u}, S) = \left( \sum_{i=1}^m (\tilde{d})_i (\tilde{1} - \tilde{t}_i) \right)^k = \left( \sum_{i=1}^m (\tilde{d})_i \cdot \tilde{1} - \sum_{i=1}^m (\tilde{d})_i \cdot \tilde{t}_i \right)^k. \quad (9)$$

Эта формула есть полином степени  $k$  от  $\sum_{i=1}^m (\tilde{d})_i \cdot \tilde{t}_i$ .

Теперь вместо леммы 5 мы применим лемму 7 при  $\sum_{i=1}^m (\tilde{d})_i \cdot \tilde{t}_i, \tilde{y} = (\tilde{d})_{z(1)} \cdot \tilde{t}_{z(1)} \equiv \text{const} \cdot \tilde{w}_1, f(\tilde{x}) = (\sum_{i=1}^m (\tilde{d})_i - \tilde{x})^k$ . Дальнейшее доказательство остаётся без изменений.

*Обобщённая  $l_1$ -метрика.* Пусть  $\tilde{d}' \in \mathbb{R}^m$ ,  $(\tilde{d}')_i > 0 \forall i \in \{1, \dots, m\}$  — вектор коэффициентов обобщённой  $l_1$ -метрики. Формула (5) заменяется на

$$\forall \tilde{u} \in \mathbb{R}^m \quad \tilde{\rho}^k(\tilde{u}, S) = \left( \sum_{i=1}^m \sum_{\tilde{t} \in T^i} (\tilde{d}')_i |(\tilde{u})_i - \text{coord}(\tilde{t})| \cdot \tilde{t} \right)^k. \quad (10)$$

Производная  $k$  порядка по сравнению со случаем обычной  $l_1$ -метрики на положительный скалярный множитель  $\prod_{i=1}^k (\tilde{d}')_{z(i)}$ . Этот множитель «объединяется» с множителем  $k!$  и не влияет на дальнейший ход доказательства.

*Сумма метрик.* Доказательство выполняется по схеме случая обобщённой  $l_1$ -метрики. Формула (5) заменяется на

$$\forall \tilde{u} \in \mathbb{R}^m \quad \tilde{\rho}^k(\tilde{u}, S) = \left( \sum_{i=1}^m \sum_{\tilde{t} \in T^i} ((\tilde{d})_i + (\tilde{d}')_i \cdot |(\tilde{u})_i - \text{coord}(\tilde{t})|) \cdot \tilde{t} \right)^k. \quad (11)$$

Производная  $k$  порядка функции  $\rho^k(\tilde{u}, S)$  остаётся прежней, что проверяется непосредственно.

Также заметим, что доказательство принадлежности  $p(x_1, x_2, \dots, x_k) \in R^k[S]$  остаётся верным, поскольку всегда можно найти достаточно малую окрестность точки  $s_{x_1, x_2, \dots, x_k}$ , в которой функция  $\tilde{\rho}^k(\tilde{u}, S)$  является непрерывной.

□

## 2.3 Обобщение на случай метрик разрезного конуса $CUT^q$

**Определение 17** ([1]). Множество всевозможных полуметрик на множестве  $\{1, 2, \dots, q\}$  с матрицами попарных расстояний вида

$$\sum_{\tilde{\delta} \in \{0,1\}^q} \lambda_{\tilde{\delta}} (E - \tilde{\delta} \cdot \tilde{\delta}^T - (\tilde{1} - \tilde{\delta}) \cdot (\tilde{1} - \tilde{\delta})^T),$$

где  $\lambda_{\tilde{\delta}} \in \mathbb{R}^+ = [0, +\infty]$  при  $\tilde{\delta} \in \{0, 1\}^q$ , называется разрезным конусом (или конусом Хэмминга) и обозначается как  $CUT^q$ .

Конус  $CUT^q$  описывает всевозможные полуметрики типа  $l_1$  на множестве  $\{1, 2, \dots, q\}$  [1], поскольку полуметрика  $\rho$  принадлежит конусу тогда и только тогда, когда существует натуральное  $r$  и точки  $\tilde{s}_1, \dots, \tilde{s}_q \in \mathbb{R}^r$  такие, что

$$\rho(i, j) = \sum_{t=1}^r |(\tilde{s}_i)_t - (\tilde{s}_j)_t|.$$

Конус разрезных полуметрик задаёт метрики, получаемые на всевозможных системах из  $q$  точек. Нас же интересует конкретная система точек  $S \subset \mathbb{R}^m$ . Поэтому определим подконус полуметрик, которые можно получить на заданной системе точек.

**Определение 18.** Пусть задана система точек  $S = \{\tilde{s}_i\}_{i=1}^q \subset \mathbb{R}^m$ . Множество всевозможных обобщённых полуметрик Хэмминга назовём разрезным конусом, согласованным с системой точек  $S$ , и обозначим как  $CUT_S^q$ .

**Лемма 10.** Любая полуметрика из  $CUT_S^q$  имеет матрицу попарных расстояний  $\|\rho(\tilde{s}_i, \tilde{s}_j)\|_{i,j=1}^q$  вида

$$\sum_{r=1}^m \lambda_r \sum_{\tilde{\delta} \in T^r} (E - \tilde{\delta} \cdot \tilde{\delta}^T - (\tilde{1} - \tilde{\delta}) \cdot (\tilde{1} - \tilde{\delta})^T). \quad (12)$$

**Доказательство.** Зафиксируем произвольную обобщённую полуметрику Хэмминга. Обозначим вектор её коэффициентов через  $\tilde{d}$ .

Рассмотрим  $ij$ -й элемент матрицы. Бинарная матрица  $T^r$  состоит из характеристических векторов системы точек по координате  $r$ . Внутренняя сумма в (12) равно расстоянию Хэмминга между  $i$ -й и  $j$ -й строками матрицы  $T^r$ . В каждой из этих строк лишь одна единица. Поэтому очевидно, что расстояние Хэмминга равно либо нулю,

если значение данной координаты у точек совпадает, либо двум, если оно различается. Сравнивая это с определением обобщённой полуметрики Хэмминга, получаем, что при  $\lambda_r = \frac{(\tilde{d})_r}{2}$  матрица попарных расстояний выбранной полуметрики задаётся формулой вида (12).  $\square$

**Следствие.** Любая полуметрика из  $CUT_S^q$  принадлежит  $CUT^q$ , то есть  $CUT_S^q$  является подконусом  $CUT^q$ .

**Замечание.** Можно выбрать такую систему из  $q$  точек, что  $T_S$  будет содержать всевозможные бинарные векторы из  $q$  элементов. В этом случае конусы  $CUT_S^q$  и  $CUT^q$  совпадают.

Обычно метрики из разрезного конуса  $CUT^q$  рассматриваются как функции, определённые на множестве  $\{1, 2, \dots, q\} \times \{1, 2, \dots, q\}$ . Это связано с тем, что данный конус определяется лишь через свойства матрицы попарных расстояний. Мы же определили конус  $CUT_S^q$  как множество полуметрик Хэмминга на системе точек, а не как множество полуметрик с матрицами попарных расстояний заданного вида. Поэтому мы будем считать, что метрики из  $CUT_S^q$  заданы на множестве  $\mathbb{R}^m \times \mathbb{R}^m$ . Это существенно для получения метрического критерия  $k$ -сингулярности, который требует, чтобы метрика было определена для произвольных точек пространства  $\mathbb{R}^m$ , а не только для точек системы  $S$ .

**Теорема 6** (метрический критерий  $k$ -сингулярности для конуса разрезных полуметрик). Система точек  $S = \{\tilde{s}_i\}_{i=1}^q$  пространства  $\mathbb{R}^m$  является  $k$ -сингулярной тогда и только тогда, когда существует ненулевой вектор  $(c_1, \dots, c_q)$  такой, что для всех метрик  $\rho$  из конуса разрезных полуметрик, согласованных с системой точек  $S$ , и для всех  $\tilde{s} \in \mathbb{R}^m$  справедливо равенство

$$\sum_{i=1}^q c_i \rho^k(\tilde{s}, \tilde{s}_i) = 0. \quad (13)$$

**Доказательство.**

**Необходимость.** Доказывается аналогично доказательству необходимости в теореме 4. Заменяем матрицу  $P_S$  на  $P_{S, \tilde{d}}$  и воспользуемся вместо леммы 3 леммой 6. Поскольку лемма 6 верна для любой обобщённой полуметрики Хэмминга, получим

требуемое утверждение. Заметим, что в качестве вектора  $(c_1, \dots, c_q)$  можно взять любой элемент из  $L \perp (U^k[S])$ , то есть этот вектор не зависит от выбора  $\rho$ .

**Достаточность.** Выберем в качестве  $\rho$  обычную (не обобщённую) метрику Хэмминга. Для всех  $\tilde{s} \in \mathbb{R}^m$  верно равенство (13). Используя теорему 4, получим, что система точек  $k$ -сингулярна.  $\square$

## 2.4 Лемма о равенстве линейных замыканий

Дмитрий Кондрашкин обнаружил неточность в доказательстве вложения  $L(H_S) \subseteq L(P_S)$  в лемме 2 в работе [2]. Мне было предложено устранить эту неточность.

Сначала приведём исходное доказательство из [2]:

Для доказательства вложения  $L(H_S) \subseteq L(P_S)$  достаточно показать, что

$$\tilde{\mathbf{1}} \in L(mE - H_S). \quad (14)$$

Пусть эта принадлежность не выполняется и  $\tilde{\mathbf{1}} = H_S \tilde{c}$ ,  $\tilde{c} = (c_1, \dots, c_q)$ , тогда

$$(mE - H_S)\tilde{c} = mE\tilde{c} - \tilde{\mathbf{1}} = m \sum_{i=1}^q c_i \tilde{\mathbf{1}} - \tilde{\mathbf{1}} = \left( m \sum_{i=1}^q (c_i - 1) \tilde{\mathbf{1}} \right).$$

Поскольку принадлежность (14) не выполняется, то  $m \sum_{i=1}^q c_i = 1$ , но тогда из  $\tilde{\mathbf{1}} = H_S \tilde{c}$  следует, что  $1 = h_{11}c_1 + \dots + h_{1q}c_q < m \sum_{i=1}^q c_i = 1$ , где  $H_S = \|h_{ij}\|_{q \times q}$  (мы пользуемся тем, что  $q \geq 2$ ,  $h_{11} = m$  и  $h_{1j} < m$  при  $j \in \{2, \dots, q\}$ , если точки попарно различны). Получили противоречие.

Неточность в этом доказательстве заключается в том, что для выполнения неравенства  $h_{11}c_1 + \dots + h_{1q}c_q < m \sum_{i=1}^q c_i$  помимо требуемых условий необходимо, чтобы все коэффициенты  $c_i$  были неотрицательными.

Была проверена гипотеза о том, что всегда можно найти такой вектор  $\tilde{c}$  с неотрицательными компонентами, что  $\tilde{\mathbf{1}} = H_S \tilde{c}$ . Тот факт, что всегда можно найти вектор  $\tilde{c}$  без ограничения на знак компонент, следует из того, что  $L(T_S) = L(H_S)$ , а вектор  $\tilde{\mathbf{1}}$  является суммой столбцов подматрицы  $T^1$  матрицы  $T^S$ . Покажем, что эта гипотеза не верна. Рассмотрим следующую « $\Gamma$ -образную» систему из  $q = 4$  точек в  $\mathbb{R}^2$  (в ячейках таблицы указан номер точки):

1		
2	3	4

Запишем её матрицы  $T_S$  и  $H_S$ :

$$T_S = \left( \begin{array}{cc|ccc} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array} \right), \quad H_S = \left( \begin{array}{cccc} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{array} \right)$$

Матрица  $H_S$  является невырожденной (проверяется непосредственно), поэтому система линейных уравнений  $H_S \tilde{c} = \tilde{1}$  имеет единственное решение

$$\tilde{c} = (0.6, -0.2, 0.4, 0.4)^T,$$

то есть для данной системы точек не существует вектора  $\tilde{c}$  с неотрицательными компонентами, такого, что  $H_S \tilde{c} = \tilde{1}$ . Гипотеза не верна.

Затем было решено проверить, существует ли система точек  $S$  и вектор  $\tilde{c}$ , т.ч.

$$\begin{cases} H_S \tilde{c} = \tilde{1} \\ m \sum_{i=1}^q c_i = 1 \end{cases} \quad (15)$$

Если это так, то  $\tilde{1} \notin L(P_S)$  и утверждение леммы не верно.

Заметим, что (15) есть система линейных алгебраических уравнений. Для проверки её совместности воспользуемся следующей теоремой:

**Теорема 7** (Кронекер-Капелли). *Система линейных алгебраических уравнений совместна тогда и только тогда, когда ранг её основной матрицы равен рангу её расширенной матрицы, причём система имеет единственное решение, если ранг равен числу неизвестных, и бесконечное множество решений, если ранг меньше числа неизвестных.*

Таким образом, нужно проверить, существует ли система точек  $S$ , т.ч. верно равенство

$$\operatorname{rg} \begin{pmatrix} H_S \\ m \tilde{1}^T \end{pmatrix} = \operatorname{rg} \left( \begin{array}{c|c} H_S & \tilde{1} \\ \hline m \tilde{1}^T & 1 \end{array} \right) \quad (16)$$

Заметим, что если матрица  $H_S = \|h_{ij}\|_{q \times q}$  есть характеристическая матрица, то для  $P_S$  выполняется неравенство треугольника. Учитывая, что  $P_S = mE - H_S$ , получаем, что для  $H_S$  должно быть верно

$$(m - h_{ij}) \leq (m - h_{ik}) + (m - h_{kj}) \quad \forall i, j, k \in \{1, \dots, q\} \quad (17)$$

Фиксируем  $q \geq 2$  и  $m$  и переберём всевозможные  $H_S$  как симметричные матрицы со значением  $m$  на диагонали и значениями из множества  $\{0, 1, \dots, m - 1\}$  вне диагонали. В ходе численного эксперимента было установлено, что находятся матрицы, удовлетворяющие условию (16). Но все эти матрицы нарушают неравенство треугольника в форме (17).

Итак, можно сделать вывод, что равенство линейных замыканий  $H_S$  и  $P_S$  не может следовать непосредственно из «простых» свойств матрицы  $H_S$ : неотрицательности, симметричности, значений  $m$  на диагонали и меньших целочисленных значений вне диагонали матрицы. Однако при добавлении неравенства треугольника в форме (17) утверждение о равенстве замыканий становится верным. Скорее всего, в доказательстве нужно использовать именно это неравенство. Поиск такого доказательства — направление дальнейших исследований.

### 3 Заключение

В данной работе был существенно обобщён метрический критерий  $k$ -сингулярности, полученный П.А. Карповичем в работе [7].

На защиту дипломной работы выносятся:

- авторское доказательство метрического критерия  $k$ -сингулярности системы точек для метрики Хэмминга и  $l_1$ -метрики;
- метрический критерий  $k$ -сингулярности для обобщённой метрики Хэмминга, обобщённой  $l_1$ -метрики, а также их суммы;
- метрический критерий  $k$ -сингулярности для конуса разрезных полуметрик;
- пример, показывающий, что не всегда существует вектор  $\tilde{c}$  с неотрицательными компонентами, такой, что  $H_S \tilde{c} = \tilde{1}$ .

Направлением дальнейших исследований является продолжение поиска способа устранить неточность в доказательстве вложения  $L(H_S) \subseteq L(P_S)$ .



## 4 Список литературы

- [1] М.М. Деза. М. Лоран, «Геометрия разрезов и метрик», М.: МЦНМО, 2001.
- [2] А.Г. Дьяконов, «Критерии вырожденности матрицы попарных  $l_1$  расстояний и их обобщения», Докл. РАН. 425:1 (2009), 11-14.
- [3] А.Г. Дьяконов, «Критерии корректности алгебраических замыканий модели алгоритмов вычисления оценок», Докл. АН, 420:6 (2008), 732-735.
- [4] А.Г. Дьяконов, «Метрики алгебраических замыканий в задачах распознавания образов с двумя непересекающимися классами», Ж. вычисл. матем. и матем. физ., 48:5 (2008), 916-927,
- [5] Ю.И. Журавлев, «Об алгебраическом подходе к решению задач распознавания или классификации», Пробл. кибернетики, 33 (1978), 5-68.
- [6] Ю.И. Журавлев, «Корректные алгоритмы над множествами некорректных (эвристических) алгоритмов. II», Кибернетика, 6 (1977), 21-27.
- [7] П.А. Карпович, «Критерии  $k$ -сингулярности и разделение 1-сингулярных систем», Вестник Московского университета. Секция, 15. «Вычислительная математика и кибернетика» 34:4 (2010), 164-171.
- [8] L. Reid, X. Sun, «Distance matrices and ridge function interpolation», Canadian Journal of Mathematics., 45 (1993), 1313-1323.