УДК 519.95

О СТРУКТУРЕ КЛАССИФИКАЦИЙ, РАСПОЗНАВАЕМЫХ АЛГОРИТМОМ "КОРА-n"

В.А. Желиговский
Школа математики и статистики,
Сиднейский университет, Австралия;

Международный институт теории прогноза землетрясений
и математической геофизики Российской академии наук, Москва

С. Шварц
Школа математики и статистики,
Сиднейский университет, Австралия

Алгоритм распознавания образов "Кора-n" успешно применялся в различных областях знаний, начиная от геофизики и кончая общественными науками. Статья содержит результаты о геометрии в пространстве объектов классификаций, которые могут быть распознаны этим алгоритмом, включая результаты компьютерных экспериментов для задач малой размерности.

ON THE STRUCTURE OF CLASSIFICATIONS RECOGNIZABLE BY THE ALGORITHM "CORA-n"

V.A. Zheligovsky
School of Mathematics & Statistics,
University of Sydney, Australia;

International Institute of Earthquake Prediction Theory
and Mathematical Geophysics, Russian Academy of Sciences, Moscow

S. Schwarz
School of Mathematics & Statistics,
University of Sydney, Australia

The pattern recognition algorithm "Cora-n" has been successfully applied in different areas ranging from geophysics to social sciences. We present results on the geometry in the object space of classifications recognizable by the algorithm, including those obtained in computer experiments with low-dimensional problems.

ВВЕДЕНИЕ

Алгоритм распознавания образов "Кора-n" [1,2] (см. также [3,4]), предложенный в шестидесятых годах, успешно применялся в таких различных областях знаний, как финансы [5], медицина [6–8], геология и поиск полезных ископаемых [9–15] и геофизика. Использование этого алгоритма для прогноза мест возможных сильных землетрясений и определения критериев высокой сейсмичности мест пересечений морфоструктурных линеаментов, впервые осуществленное в [16–19] – возможно, наиболее эффектное приложение. Впоследствии прогнозы мест сильных землетрясений для всех значительных сейсмоактивных регионов были построены в Отделе вычислительной геофизики Института физики Земли АН СССР, преобразованном в 1989 г. в Международный институт теории прогноза землетрясений и математической геофизики АН СССР. Начиная с [17], эти прогнозы составили серию статей в ежегодном сборнике публикаций этого Отдела / Института "Вычислительная сейсмология"; в этой серии на настоящий момент последней является работа [20]. Проверка результатов этих прогнозов выполнена, например, в [21].

Широкий круг приложений оправдывает систематическое математическое изучение свойств указанного алгоритма. Однако, насколько нам известно, до настоящего времени оно не было предпринято. В статье представлены результаты такого исследования, включая результаты исчерпывающего компьютерного анализа задач малой размерности. Рассматривались следующие вопросы:

i. Какими свойствами характеризуются распознаваемые классификации? Какие классификации не могут быть распознаны?

Рассмотрение нераспознаваемых классификаций важно, в частности, так как их обилие характеризует "разрешающую способность" алгоритма.

ii. Каковы геометрические критерии устойчивости распознавания?

Подобные критерии представляют практический интерес, поскольку при проверке устойчивости данного распознавания они могут позволить избежать проведения исчерпывающих переборов всех комбинаций свободных параметров, требующих, как правило, значительных затрат вычислительных ресурсов.

iii. Как изменяются результаты распознавания, когда кодировка объектов дополняется "избыточными" дескрипторами?

Этот вопрос практически важен, так как обычно заранее неизвестно, какие из признаков, используемых для представления объектов, отвечают природе рассматриваемой задачи классифицирования.

1. Описание алгоритма распознавания образов "Кора-n"

Алгоритм "Кора-n" осуществляет *дихотомию* (т.е. разбиение на два класса, C_1 и C_2; $C_1 \cap C_2 = \emptyset$) ***объектов***, отображенных посредством ***кодирования*** на множество ***бинарных векторов*** $\mathbf{v}$ заданной размерности M (т.е. векторов, каждая из M компонент $\mathbf{v}_p$ которых равна 0 или 1). На практике исследователь подготавливает вопросник, состоящий из M вопросов, и каждая компонента вектора – *дескриптор* – представляет собой утвердительный (например, $\mathbf{v}_p = 1$) или отрицательный ($\mathbf{v}_p = 0$) ответ на соответствующий вопрос для данного объекта. Таким образом, объекты отождествляются с бинарными векторами $\mathbf{v}$, являющимися вершинами M-мерного гиперкуба $X^M = \{0,1\}^M$. Для успешного приложения алгоритма вопросник должен быть полным в том смысле, что совокупность ответов на вопросы должна содержать достаточно информации для правильного причисления каждого объекта к одному из классов.

Если не указано обратное, ниже не предполагается, что кодирование объектов *сюръективно*, т.е. некоторые векторы $\mathbf{v} \in X^M$ могут не соответствовать никаким объектам. Образ кодирования объектов обозначен O.

Как будет описано ниже, действие алгоритма зависит от значения шести *свободных параметров*: n, A_1, R_1, A_2, R_2 и Δ. Выбор величин этих параметров – прерогатива пользователя.

Работа алгоритма происходит в две стадии: *обучение* и *голосование*. Для выполнения первой стадии необходимо задать два подмножества, $L_1 \subset C_1$ и $L_2 \subset C_2$; $L_1 \sqcup L_2$ называется *материалом обучения*.

Определение 1. Признак $\sigma = \{(p_i, b_i) \mid 1 \leq i \leq k\}$ ранга $\mathrm{rank}\ \sigma = k \leq M$ – это множество из k пар, состоящих из взаимно различных дескрипторов p_i, $1 \leq p_i \leq M$, и ассоциированных бинарных величин $b_i \in \{0, 1\}$.

Определение 2. Вектор $\mathbf{v}$ *имеет* признак $\sigma = \{(p_i, b_i) \mid 1 \leq i \leq k\}$, если и только если $\mathbf{v}_{p_i} = b_i\ \forall i,\ \ 1 \leq i \leq k$.

Определение 3. Функция $\sigma : X^M \to \{0, 1\}$ называется *характеристической функцией признака* σ, если $\sigma(\mathbf{v}) = 1$, когда вектор $\mathbf{v}$ имеет признак σ, и $\sigma(\mathbf{v}) = 0$ в противном случае.

На стадии обучения для каждого признака σ ранга $\mathrm{rank}\ \sigma \leq n$ вычисляются величины

$$N_i(\sigma) = \sum_{\mathbf{v} \in L_i} \sigma(\mathbf{v}) = |\{\mathbf{v} \in L_i \mid \mathbf{v} \text{ имеет } \sigma\}| \quad (i = 1, 2).$$

В результате обучения компилируются два списка признаков $\mathcal{S}^i = \{\sigma^i_j\}$ $(i = 1, 2)$, характеризующих, соответственно, первый и второй класс:

Определение 4. Признак σ *характеризует* класс i $(i = 1, 2)$, если и только если

$$\mathrm{rank}\ \sigma \leq n, \quad N_i(\sigma) \geq A_i \quad \text{и} \quad N_{3-i}(\sigma) \leq R_i.$$

Замечание 1. Без ограничения общности можно считать, что $A_i \geq 0,\ R_i \geq 0$, и, если кодировка сюръективна, либо $A_i = \infty \Rightarrow \mathcal{S}^i = \emptyset$, либо $2^{M-n} \leq A_i + R_i \leq 2^{M-1}$.

На второй стадии векторы классифицируются при помощи "голосования": алгоритм относит $\mathbf{v} \in O$ к первому классу, если $E_1(\mathbf{v}) - E_2(\mathbf{v}) \geq \Delta$, и ко второму – в противном случае, здесь

$$E_i(\mathbf{v}) = \sum_{\sigma \in \mathcal{S}^i} \sigma(\mathbf{v}) = |\{\sigma \in \mathcal{S}^i \mid \mathbf{v} \text{ имеет } \sigma\}| \quad (i = 1, 2). \tag{1}$$

При выборе свободных параметров обычно руководствуются следующими соображениями "здравого смысла":

i. Никакой признак не должен характеризовать одновременно оба класса. (В компьютерных экспериментах мы не требовали выполнения этого условия.)

ii. Все векторы из материала обучения $L_1 \sqcup L_2$ должны быть распознаны правильно.

Замечание 2. Все векторы $\mathbf{v} \in L_1$ распознаются правильно, если и только если

$$\Delta \leq \min_{\mathbf{v} \in L_1} (E_1(\mathbf{v}) - E_2(\mathbf{v})).$$

Замечание 3. В приложениях для распознавания мест возникновения возможных землетрясений только векторы из L_1 (высокосейсмичные области) должны быть распознаны правильно: в этом контексте разрешается переход векторов из L_2 (низкосейсмичные области) в C_1 (сейсмически-опасные области). (См. ссылки, цитированные во Введении.)

2. Устойчивость распознавания

Понятие устойчивости распознавания было сформулировано в работе [22]. Устойчивость распознавания является следствием внутренней логической непротиворечивости процедуры. Пусть в некоторой задаче кодировка представляет объекты без потери существенной информации, материал обучения адекватно представляет классы, и алгоритм распознавания образов правильно воспроизводит искомую классификацию. Тогда дальнейшее увеличение объема материала обучения не должно привести к изменению результата распознавания.

Определение 5. Классификация $C_1 \sqcup C_2$ называется *устойчиво распознаваемой с материалом обучения* $L_1 \sqcup L_2$, если для любого материала обучения $L_1' \sqcup L_2'$, такого, что $L_1 \subset L_1' \subset C_1$ и $L_2 \subset L_2' \subset C_2$, существует набор свободных параметров $A_1(L_1', L_2')$, $R_1(L_1', L_2')$, $A_2(L_1', L_2')$, $R_2(L_1', L_2')$ и $\Delta(L_1', L_2')$, при которых результат распознавания совпадает с первоначальной классификацией $C_1 \sqcup C_2$.

Проверка устойчивости распознавания, следуя определению, требует значительных затрат вычислительных ресурсов, так как она включает в себя рассмотрение всех подмножеств множества $(C_1 \sqcup C_2) \setminus (L_1 \sqcup L_2)$ и циклы по всем возможным комбинациям свободных параметров. (В некоторых геометрически простых случаях теоремы 1 и 2, приведенные ниже, гарантируют устойчивость распознавания алгоритмом Кора-n.)

Из устойчивости распознавания некоторой классификации $C_1 \sqcup C_2$, в частности, следует, что она *самораспознаваема*, т.е. может быть распознана при использовании в качестве материала обучения самой этой классификации (при $L_1 = C_1$, $L_2 = C_2$). Это оправдывает особое внимание к самораспознаваемым классификациям.

I. НЕКОТОРЫЕ ДОСТАТОЧНЫЕ УСЛОВИЯ РАСПОЗНАВАЕМОСТИ КЛАССИФИКАЦИЙ

Область применимости теорем 1–4 может быть расширена посредством взаимной перестановки номеров классов C_1 и C_2.

Определение 6. $H \subset X^M$ называется *гипергранью коразмерности* k ($\mathrm{codim} H = k$), если существует признак σ ранга k такой, что вектор $\mathbf{v}$ имеет σ тогда и только тогда, когда $\mathbf{v} \in H$. Тогда $H = H_\sigma$ называется гипергранью, ассоциированной с признаком σ, и $\sigma = \sigma_H$ называется признаком, ассоциированным с гипергранью H.

Определение 6 согласуется со стандартным определением гиперграни в геометрии. Соответствие между гипергранями коразмерности k и признаками ранга k взаимно однозначно. Гипергрань коразмерности k состоит из 2^{M-k} вершин гиперкуба X^M (векторов). Гипергрань коразмерности M – это вершина (вектор).

Теорема 1. Пусть $C_1 = O \cap H$, где H - гипергрань. Если

$$|L_1| \geq \frac{1}{2}(|H| + 1), \tag{2}$$

то классификация $O = C_1 \sqcup C_2$ устойчиво распознаваема с материалом обучения $L_1 \sqcup L_2$.

Замечание 4. Если $\operatorname{codim} H \geq M - 1$, то из (2) следует $L_1 = C_1 = H$.

Доказательство. Рассмотрим произвольный материал обучения $L'_1 \sqcup L'_2$, где $L_1 \subset L'_1 \subset C_1$, $L_2 \subset L'_2 \subset C_2$. Положим $A_2 = \infty$, $R_2 = 0$, при которых ни один признак не характеризует C_2.

Выберем

$$A_1 = |L_1|, \qquad R_1 = \infty. \tag{3}$$

Тогда в силу (2) ранг любого признака, характеризующего C_1, не превышает $k = \operatorname{codim} H$. Пусть $\sigma_H = \{(p_i, b_i) \mid 1 \leq i \leq k\}$. Обозначим $\mathcal{P}_H = \{p_i,\ 1 \leq i \leq k\}$. Рассмотрим признак $\sigma = \{(p'_i, b'_i) \mid 1 \leq i \leq k'\}$ ранга $k' \leq k$ и обозначим $\mathcal{P}_\sigma = \{p'_i,\ 1 \leq i \leq k'\}$. Тогда, очевидно, $|L'_1 \cap H_\sigma| \leq |H \cap H_\sigma| \leq 2^{-|\mathcal{P}_\sigma \setminus \mathcal{P}_H|}|H|$. Следовательно, согласно (2) и (3) признак σ характеризует C_1, если и только если $H \subset H_\sigma$ и $\operatorname{rank} \sigma \leq n$. Число таких признаков

$$r = \begin{cases} 2^k - 1, & \text{если } k \leq n; \\ \sum_{s=1}^{n} \binom{k}{s}, & \text{если } k > n. \end{cases}$$

Очевидно, $H = \cap_{m=1}^{r} H^m$, где пересечение выполнено по всем гиперграням $H^m \supset H$ коразмерности $\operatorname{codim} H^m \leq \min(n, k)$. Таким образом, при голосовании с $\Delta = r$ будет построена классификация $C_1 \sqcup C_2$.▷

Теорема 2. Пусть

$$C_1 = O \bigcap \bigcup_{i=1}^{p} H^i, \tag{4}$$

где каждое H^i - гипергрань коразмерности $k \leq n$.

Рассмотрим $L_1 \subset C_1$, $L_2 \subset C_2 = O \setminus C_1$ и обозначим

$$Q = \min_{1 \leq i \leq p} |L_1 \cap H^i|. \tag{5}$$

Если

$$k = n \quad \text{или} \quad Q \geq \frac{1}{2}(2^{M-k} + 1) \tag{6}$$

и

$$|L_2 \cap H| > 0 \tag{7}$$

для любой гиперграни H коразмерности k такой, что $H \cap C_2 \neq \emptyset$, то классификация $O = C_1 \sqcup C_2$ устойчиво распознаваема с материалом обучения $L_1 \sqcup L_2$.

Замечание 5. Представление (4) класса C_1, как правило, не единственно.

Замечание 6. В (4) H^i могут пересекаться.

Замечание 7. При $n = M$ и $k = M - 1$ из (6) следует $L_1 = C_1$.

Замечание 8. Если $k = M$, то из $n = M$ и (7) следует $L_2 = C_2$. C_1 и C_2 тогда произвольны, но как показывает доказательство, в этом случае алгоритм распознает каждый объект независимо, и причина успеха распознавания – априорное точное знание классов, а не подлинная идентификация их структуры.

Доказательство. Рассмотрим некоторый материал обучения $L_1' \sqcup L_2'$, где $L_1 \subset L_1' \subset C_1$, $L_2 \subset L_2' \subset C_2$. Выберем $A_2 = \infty$, $R_2 = 0$, при которых ни один признак не характеризует C_2.

Положим $A_1 = Q$, $R_1 = 0$. Ввиду (6) ранг любого признака, характеризующего C_1, не превышает k. Рассмотрим признак σ с rank $\sigma \leq k$. Либо $O \cap H_\sigma \subset C_1$, либо существует гипергрань $H \subset H_\sigma$ коразмерности k такая, что $H \cap C_2 \neq \emptyset$. В последнем случае в силу (7) $|L_2' \cap H_\sigma| \geq |L_2 \cap H| > 0$ и, следовательно, σ не характеризует C_1. Таким образом, если признак σ характеризует C_1, то $\sigma(\mathbf{v}) = 1 \Rightarrow \mathbf{v} \notin C_2$.

Ввиду (5) каждый из признаков σ_{H^i} характеризует C_1. Следовательно, в силу (4) любой вектор $\mathbf{v} \in C_1$ имеет по крайней мере один признак, характеризующий C_1, и при голосовании с $\Delta = 1$ будет построена желаемая классификация $C_1 \sqcup C_2$.$\triangleright$

Теорема 3. Если классификация $C_1 \sqcup C_2 = X^M$ самораспознаваема алгоритмом Кора-n при

$$A_i \geq \min(2^{M-n-1},\ |C_i|,\ 1 + \max_H |C_i \cap H|) \quad \forall i = 1,2 \tag{8}$$

(здесь максимумы взяты по всем гиперграням H коразмерности $n+1$), то она также самораспознается алгоритмом Кора-n', $\forall n' > n$, при тех же величинах параметров A_i, R_i.

Доказательство. Классификация распознается при некоторых A_i, R_i, если и только если

$$\min_{\mathbf{v} \in C_1} (E_1(\mathbf{v}) - E_2(\mathbf{v})) > \max_{\mathbf{v} \in C_2} (E_1(\mathbf{v}) - E_2(\mathbf{v})) \tag{9}$$

(см. замечание 2). В частности, при распознавании алгоритмом Кора-n первоначальные величины параметров A_i, R_i удовлетворяют (9).

Рассмотрим распознавание алгоритмом Кора-n', $n' > n$ с первоначальными величинами параметров A_i, R_i. Некоторый признак будет характеризовать какой-либо из классов, либо если он характеризовал этот класс при первоначальном распознавании, либо, возможно, это новый признак ранга строго больше n. Пусть $E_i'(\mathbf{v})$ обозначает число признаков, характеризующих i–й класс при новом распознавании, которые имеет вектор $\mathbf{v}$. Сначала покажем, что

$$\min_{\mathbf{v} \in C_1} (E_1'(\mathbf{v}) - E_2(\mathbf{v})) > \max_{\mathbf{v} \in C_2} (E_1'(\mathbf{v}) - E_2(\mathbf{v})). \tag{10}$$

Согласно (8), имеют место три возможности:

i. $A_1 > \max_{\mathrm{codim} H = n+1} |C_1 \cap H|$. Тогда при распознавании с $n' > n$ новые признаки, характеризующие C_1, не появляются. Таким образом, $E_1'(\mathbf{v}) = E_1(\mathbf{v})\ \ \forall \mathbf{v}$, и (10) следует из (9).

ii. $A_1 = 2^{M-n-1}$. В силу этого равенства никакой $\mathbf{v} \in C_2$ не может иметь ни одного нового признака, характеризующего C_1. (В этом случае любой новый признак, характеризующий C_1, имеет ранг $n+1$.) Таким образом,

$$\min_{\mathbf{v}\in C_1}(E_1'(\mathbf{v})-E_2(\mathbf{v}))\geq \min_{\mathbf{v}\in C_1}(E_1(\mathbf{v})-E_2(\mathbf{v})),$$

$$\max_{\mathbf{v}\in C_2}(E_1'(\mathbf{v})-E_2(\mathbf{v}))= \max_{\mathbf{v}\in C_2}(E_1(\mathbf{v})-E_2(\mathbf{v})),$$

и (10) следует из (9).

iii. $A_1=|C_1|$. Предположим, что возникли e новых признаков, характеризующих C_1. В силу этого равенства их имеет каждый вектор из C_1. Следовательно,

$$\min_{\mathbf{v}\in C_1}(E_1'(\mathbf{v})-E_2(\mathbf{v}))=e+ \min_{\mathbf{v}\in C_1}(E_1(\mathbf{v})-E_2(\mathbf{v})),$$

$$\max_{\mathbf{v}\in C_2}(E_1'(\mathbf{v})-E_2(\mathbf{v}))\leq e+ \max_{\mathbf{v}\in C_2}(E_1(\mathbf{v})-E_2(\mathbf{v})),$$

и (10) опять вытекает из (9).

Аналог неравенства (9) для нового распознавания с $n'>n$ можно теперь вывести из (10), рассматривая подобным образом три возможности для A_2.▷

II. НЕКОТОРЫЕ ДОСТАТОЧНЫЕ УСЛОВИЯ НЕРАСПОЗНАВАЕМОСТИ КЛАССИФИКАЦИЙ

Теорема 4. Пусть $\mathcal{P}=\{p_i|\ 1\leq p_i\leq M\}$ – некоторое множество дескрипторов, $M\geq|\mathcal{P}|>n$, и

$$C_1=\left\{\mathbf{v}\in X^M\ \middle|\ \sum_{p\in\mathcal{P}}\mathbf{v}_p\ \textit{нечетно}\right\},\quad C_2=X^M\backslash C_1.$$

Тогда классификация $C_1\sqcup C_2$ не может быть распознана алгоритмом Кора-n, какой бы ни был использован материал обучения $L_1\sqcup L_2$.

Доказательство. Предположим, что указанная классификация распознана. Тогда $E_1(\mathbf{v})-E_2(\mathbf{v})\geq\Delta\ \forall\mathbf{v}\in C_1$ и, следовательно,

$$\sum_{\mathbf{v}\in C_1}E_1(\mathbf{v})-E_2(\mathbf{v})\geq\Delta|C_1|. \tag{11}$$

Аналогично, $\forall\mathbf{v}\in C_2,\ E_1(\mathbf{v})-E_2(\mathbf{v})<\Delta$, и поэтому

$$\sum_{\mathbf{v}\in C_2}E_2(\mathbf{v})-E_1(\mathbf{v})>-\Delta|C_2|. \tag{12}$$

В рассматриваемой классификации $|C_1|=|C_2|$, и, складывая (11) и (12), получаем

$$\sum_{\mathbf{v}\in C_1}(E_1(\mathbf{v})-E_2(\mathbf{v}))+\sum_{\mathbf{v}\in C_2}(E_2(\mathbf{v})-E_1(\mathbf{v}))>0,$$

или

$$\sum_{\sigma\in\mathcal{S}^1}\left(\sum_{\mathbf{v}\in C_1}\sigma(\mathbf{v})-\sum_{\mathbf{v}\in C_2}\sigma(\mathbf{v})\right)-\sum_{\sigma\in\mathcal{S}^2}\left(\sum_{\mathbf{v}\in C_1}\sigma(\mathbf{v})-\sum_{\mathbf{v}\in C_2}\sigma(\mathbf{v})\right)>0$$

(см. (1)). Тем самым получено противоречие, так как при условиях данной теоремы обе суммы в левой части этого неравенства равны нулю.

Для доказательства этого достаточно показать, что для любого признака $\sigma = \{(p'_i, b'_i) \mid 1 \le i \le k\}$ ранга $k \le n$

$$\sum_{\mathbf{v}\in C_1} \sigma(\mathbf{v}) - \sum_{\mathbf{v}\in C_2} \sigma(\mathbf{v}) = 0. \tag{13}$$

Обозначим $\mathcal{P}_\sigma = \{p'_i \mid 1 \le i \le k\}$. Вектор $\mathbf{v}$, имеющий σ, принадлежит C_1, если и только если суммы $\sum_{p\in\mathcal{P}\setminus\mathcal{P}_\sigma} \mathbf{v}_p$ и $\sum_{p\in\mathcal{P}\cap\mathcal{P}_\sigma} \mathbf{v}_p$ имеют одинаковую четность. Поскольку $|\mathcal{P}| > n$, $\mathcal{P}\setminus\mathcal{P}_\sigma \ne \emptyset$. Таким образом, каждому классу принадлежит равное число векторов, имеющих σ, откуда следует (13).▷

Классификации, рассмотренные в теореме 4, имеют хорошо определенную структуру, неспособность распознать которую – явный недостаток алгоритма. Этот пример указывает на важность правильности выбора кодирования объектов.

Теорема 5. Пусть $\varphi : X^M \to X^M$ – биекция гиперкуба X^M, при которой $\varphi(H)$ является гипергранью для любой гиперграни H (для любого множества C, $\varphi(C)$ обозначает образ C при отображении φ), множество O φ-инвариантно, и $|\varphi(H)\cap L_i| = |H\cap L_i|$ $(i=1,2)$ для любой гиперграни H коразмерности $\mathrm{codim} H = n$. Если $\varphi(C_1) \ne C_1$ ($\Leftrightarrow C_2 \ne \varphi(C_2)$), то классификация $C_1 \sqcup C_2$ не может быть распознана с материалом обучения $L_1 \sqcup L_2$.

Доказательство. Любая гипергрань коразмерности меньше n является объединением непересекающихся гиперграней коразмерности n. Поэтому при условиях данной теоремы $N_i(\sigma) = N_i(\sigma_{\varphi(H_\sigma)})$ $(i = 1, 2)$ для любого признака σ ранга rank $\sigma \le n$. Следовательно, C_i характеризуется признаком σ, если и только если $\sigma_{\varphi(H_\sigma)}$ характеризует C_i. Таким образом, всякий вектор $\mathbf{v} \in O$ окажется отнесен к тому же классу, что и $\varphi(\mathbf{v})$. Это противоречит условию $C_i \ne \varphi(C_i)$ для каждого из классов.▷

III. РАСПОЗНАВАНИЕ ГИПЕРЦИЛИНДРОВ

Определение 7. Множество

$$C = \left\{ \mathbf{v} \in X^M \left| \sum_{s=1}^{m} \left(\frac{\mathbf{v}_{p_s} - \mathbf{c}_{p_s}}{a_s} \right)^2 \le 1 \right.\right\} \tag{14}$$

называется *гиперцилиндром* с *длинами полуосей* $a_s > 0$, *гиперось* которого проходит через вершину $\mathbf{c}$; $\{p_s | 1 \le p_1 < \dots < p_s < \dots < p_m \le M\}$ называются *базисными* дескрипторами гиперцилиндра.

Замечание 9. Строго говоря, множество C из определения 7 – пересечение гиперцилиндра с гиперкубом X^M. Однако в соответствии с дискретной природой рассматриваемой задачи мы принимаем терминологию определения 7, игнорируя указанную формальную неточность.

Замечание 10. Гиперось гиперцилиндра – это гипергрань $\{\mathbf{v} \mid \mathbf{v}_{p_s} = \mathbf{c}_{p_s}\ \forall s,\ 1 \le s \le m\}$ коразмерности m. При $m = M$ гиперцилиндр вырождается в эллипсоид, гиперось которого – единственная вершина $\mathbf{c}$ – является центром эллипсоида. Поскольку гиперцилиндр – конечное множество, его гиперось не обязательно единственна (например, любая вершина является центром сферы радиуса $\sqrt{M}$).

Замечание 11. Если в (14) неравенство изменено на строгое, то определяемое им множество вершин – тоже гиперцилиндр: поскольку и гиперцилиндр и его дополнение в X^M конечны, то указанное множество определяется (14) с $a_i' = a_i(1-\epsilon)$ при любом достаточно малом $\epsilon > 0$.

Замечание 12. Вершины в дополнении в X^M к гиперцилиндру (14) удовлетворяют неравенству

$$\sum_{s=1}^{m} \left(\frac{\mathbf{v}_{p_s} - (1 - \mathbf{c}_{p_s})}{a_s} \right)^2 < \sum_{s=1}^{m} \frac{1}{a_s^2} - 1.$$

Таким образом, согласно предыдущему замечанию, это дополнение является гиперцилиндром с тем же набором базисных дескрипторов и с длинами полуосей $a_s' = a_s(1-\epsilon)\left(\sum_{q=1}^{m} \frac{1}{a_q^2} - 1\right)^{-1/2}$ (где $\epsilon > 0$ достаточно мало), гиперось которого симметрична относительно центра гиперкуба X^M гипероси исходного гиперцилиндра. Поскольку рассматриваются только бинарные векторы, из (14) легко получить уравнение гиперплоскости, разбивающей X^M на два гиперцилиндра.

Теорема 6.

i. Классификация $O = C_1 \sqcup C_2$ может быть распознана алгоритмом Кора-1, только если хотя бы один из классов является пересечением образа кодирования объектов O и гиперцилиндра, причем $\exists \rho \geq 0$ такое, что каждая из полуосей имеет длину ρ или $\sqrt{2}\rho$ (и согласно замечанию 12 это также верно по отношению к другому классу).

ii. Если кодирование объектов сюръективно, то всякая классификация, рассмотренная в i, самораспознаваема алгоритмом Кора-n, $\forall n \leq M$.

Доказательство.

i. Пусть классификация $O = C_1 \sqcup C_2$ распознана алгоритмом Кора-1 при использовании некоторого материала обучения. Все дескрипторы можно разбить на 16 множеств:

$$\mathcal{P}, \mathcal{P}^1, \mathcal{P}^2, \mathcal{P}^{12}, \mathcal{P}_1, \mathcal{P}_1^1, \mathcal{P}_1^2, \mathcal{P}_1^{12}, \mathcal{P}_2, \mathcal{P}_2^1, \mathcal{P}_2^2, \mathcal{P}_2^{12}, \mathcal{P}_{12}, \mathcal{P}_{12}^1, \mathcal{P}_{12}^2, \mathcal{P}_{12}^{12}.$$

Дескриптор p относится при этом ко множеству, верхний индекс которого является списком классов, характеризуемых признаком $(p,1)$, а нижний индекс – списком классов, характеризуемых признаком $(p,0)$ (для максимальной общности не запрещая ситуации, когда один признак одновременно характеризует оба класса). Взаимная перестановка значений 0 и 1, кодирующих[1] дескрипторы из $\mathcal{P}_1 \sqcup \mathcal{P}^2$, преобразует множества $\mathcal{P}_1, \mathcal{P}^2, \mathcal{P}_1^2, \mathcal{P}^{12}, \mathcal{P}_1^{12}$ и $\mathcal{P}_{12}^2$, соответственно, в подмножества $\mathcal{P}^1, \mathcal{P}_2, \mathcal{P}_2^1, \mathcal{P}_{12}, \mathcal{P}_{12}^1, \mathcal{P}_2^{12}$. Таким образом, список возможных множеств дескрипторов сокращается до 12:

$$\mathcal{P}, \mathcal{P}^1, \mathcal{P}_1^1, \mathcal{P}_2, \mathcal{P}_2^1, \mathcal{P}_2^2, \mathcal{P}_2^{12}, \mathcal{P}_{12}, \mathcal{P}_{12}^1, \mathcal{P}_{12}^{12}.$$

Для произвольного вектора $\mathbf{v}$ введем обозначения:

$$e_\beta^\alpha = |\{p \in \mathcal{P}_\beta^\alpha | \mathbf{v}_p = 1\}|; \quad n_\beta^\alpha = |\{p \in \mathcal{P}_\beta^\alpha | \mathbf{v}_p = 0\}|;$$

[1]Эта замена координат не изменяет внутреннюю геометрию классов; (см. раздел V.1).

очевидно, $e^\alpha_\beta + n^\alpha_\beta = |\mathcal{P}^\alpha_\beta| \quad \forall \mathbf{v} \in X^M$. Тогда

$$E_1(\mathbf{v}) = e^1 + e^1_1 + n^1_1 + e^1_2 + e^{12}_2 + n_{12} + e^1_{12} + n^1_{12} + e^{12}_{12} + n^{12}_{12} = \\ = e^1 + |\mathcal{P}^1_1| + e^1_2 + e^{12}_2 + n_{12} + |\mathcal{P}^1_{12}| + |\mathcal{P}^{12}_{12}|,$$

и

$$E_2(\mathbf{v}) = n_2 + n^1_2 + e^2_2 + n^2_2 + e^{12}_2 + n^{12}_2 + n_{12} + n^1_{12} + e^{12}_{12} + n^{12}_{12} = \\ = n_2 + n^1_2 + |\mathcal{P}^2_2| + |\mathcal{P}^{12}_2| + n_{12} + n^1_{12} + |\mathcal{P}^{12}_{12}|.$$

Соответственно, некоторый вектор будет отнесен к C_1, если и только если

$$E_1(\mathbf{v}) - E_2(\mathbf{v}) = e^1 + |\mathcal{P}^1_1| + e^1_2 + e^{12}_2 + n_{12} + \\ + |\mathcal{P}^1_{12}| + |\mathcal{P}^{12}_{12}| - (n_2 + n^1_2 + |\mathcal{P}^2_2| + |\mathcal{P}^{12}_2| + n_{12} + n^1_{12} + |\mathcal{P}^{12}_{12}|) = \\ = e^1 + |\mathcal{P}^1_1| + 2e^1_2 + e^{12}_2 - |\mathcal{P}_2| + e_2 - |\mathcal{P}^1_2| - |\mathcal{P}^2_2| - |\mathcal{P}^{12}_2| + e^1_{12} \geq \Delta.$$

Это неравенство эквивалентно (14) с $\mathbf{c}_p = 1,\ 1 \leq p \leq M$,

$$\rho^2 = \frac{1}{2}(\Delta - |\mathcal{P}^1_1| + |\mathcal{P}_2| + |\mathcal{P}^1_2| + |\mathcal{P}^2_2| + |\mathcal{P}^{12}_2|),$$

$$a_s = \begin{cases} \rho, & \text{если } p_s \in \mathcal{P}^1_2, \\ \sqrt{2}\rho, & \text{если } p_s \in \mathcal{P}^1 \sqcup \mathcal{P}^{12}_2 \sqcup \mathcal{P}_2 \sqcup \mathcal{P}^1_{12}. \end{cases}$$

Таким образом, C_1 представлено в виде искомого пересечения образа кодирования объектов и указанного гиперцилиндра.

ii. Пусть первый класс C_1 является гиперцилиндром (14), m_1 полуосей которого имеют длину $\sqrt{2}\rho$ и m_2 полуосей – длину ρ. Обозначим через $\mathcal{P}_1$ множество дескрипторов, соответствующих полуосям длины $\sqrt{2}\rho$, через $\mathcal{P}_2$ – полуосям длины ρ, и через $\mathcal{P}_0$ – оставшиеся $m_0 = M - m_1 - m_2$ дескрипторов. Без ограничения общности будем считать, что гиперось рассматриваемого гиперцилиндра проходит через вершину $\mathbf{c}_p = 1,\ 1 \leq p \leq M$. Тогда $\mathbf{v} \in C_1$, если и только если

$$e_1 + 2e_2 \geq \delta, \tag{15}$$

где $e_i = |\{p \in \mathcal{P}_i | \mathbf{v}_p = 1\}|$ $(i = 1, 2)$ и $\delta = m_1 + 2m_2 - 2\rho^2$. Можно считать, что в (15) $\delta > 0$ – целое (случай $\delta = 0$ тривиален). Теорема 1 доказывает *ii* при $\delta = 1$ (C_2 – гипергрань) и при $\delta = m_1 + 2m_2$ (C_1 – гипергрань). Соответственно, дальнейшее доказательство приведено для

$$2 \leq \delta \leq m_1 + 2m_2 - 1. \tag{16}$$

Без ограничения общности можно считать, что $m_1 > 0$ (случай $m_1 = 0,\ m_2 > 0$ эквивалентен случаю $m'_1 = m_2 > 0,\ m'_2 = 0$ с новой величиной радиуса $\rho' = \sqrt{2}\rho$; случай $m_1 = m_2 = 0$ – вырожденный). Более того, при $m_1 = 1,\ m_2 = 0$ (14) определяет гипергрань, и теорема 1 опять доказывает *ii*. При $m_1 = 1,\ m_2 > 0$ и нечетном δ множество целочисленных решений (15), удовлетворяющих $0 \leq e_i \leq m_i$ $(i = 1, 2)$, совпадает со множеством решений неравенства $e_1 + e_2 \geq \frac{1}{2}(\delta + 1)$, удовлетворяющих тем же условиям, и, следовательно, этот случай эквивалентен случаю $m'_0 = m_0,\ m'_1 = m_2 + 1,\ m'_2 = 0$. Наконец, при $m_1 = 1,\ m_2 > 0$ и четном δ

рассматриваемые решения (15) – это $e_1 = 0$ или 1, $m_2 \geq e_2 \geq \frac{\delta}{2}$, и этот случай эквивалентен случаю $m'_0 = m_0 + 1$, $m'_1 = m_2$ и $m'_2 = 0$. Соответственно, ниже предполагается $m_1 > 1$.

Дальнейшее доказательство утверждения *ii* разбито на четыре леммы.

Лемма 1. Пусть кодирование объектов сюръективно, $m_1 > 1$, $m_2 > 0$ и выполнено (16). Тогда любая классификация, рассмотренная в части i доказываемой теоремы, самораспознаваема алгоритмом Кора-1.

Доказательство. Пусть $N_i(\mathcal{P}, b)$ обозначает число векторов $\mathbf{v} \in C_i$, имеющих какой-либо (фиксированный) признак (p, b) ранга 1, где $p \in \mathcal{P} \in \{\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2\}$ (это число не зависит от конкретного $p \in \mathcal{P}$). Тогда

$$N_i(\mathcal{P}_0, 0) = N_i(\mathcal{P}_0, 1) = \frac{1}{2}|C_i| \quad \forall i = 1,2$$

(для $m_0 = 0$, $i = 1,2$ положим $N_i(\emptyset, 0) = N_i(\emptyset, 1) = \frac{1}{2}|C_i|$ по определению),

$$N_1(\mathcal{P}_1, 0) = 2^{m_0} \sum_{e_1 \geq 0, e_2 \geq 0:\ e_1 + 2e_2 \geq \delta} \binom{m_1 - 1}{e_1} \binom{m_2}{e_2},$$

$$N_1(\mathcal{P}_1, 1) = 2^{m_0} \sum_{e_1 \geq 0, e_2 \geq 0:\ e_1 + 2e_2 \geq \delta - 1} \binom{m_1 - 1}{e_1} \binom{m_2}{e_2},$$

$$N_1(\mathcal{P}_2, 0) = 2^{m_0} \sum_{e_1 \geq 0, e_2 \geq 0:\ e_1 + 2e_2 \geq \delta} \binom{m_1}{e_1} \binom{m_2 - 1}{e_2},$$

$$N_1(\mathcal{P}_2, 1) = 2^{m_0} \sum_{e_1 \geq 0, e_2 \geq 0:\ e_1 + 2e_2 \geq \delta - 2} \binom{m_1}{e_1} \binom{m_2 - 1}{e_2}.$$

Из комбинаторных результатов, приведенных выше, следует, что

$$N_1(\mathcal{P}_1, 0) < N_1(\mathcal{P}_1, 1), \quad N_1(\mathcal{P}_2, 0) < N_1(\mathcal{P}_2, 1).$$

(Тогда как все положительные слагаемые суммы, выражающей $N_1(\mathcal{P}_1, 0)$, присутствуют в сумме, выражающей $N_1(\mathcal{P}_1, 1)$, последняя содержит также дополнительные члены, соответствующие $e_1 + 2e_2 = \delta - 1$, например,$e_2 = \frac{1}{2}(\delta - 1 - e_1)$. Аналогично, все положительные слагаемые суммы, выражающей $N_1(\mathcal{P}_2, 0)$, присутствуют в сумме, выражающей $N_1(\mathcal{P}_2, 1)$, и последняя содержит дополнительные члены, соответствующие $\delta - 2 \leq e_1 + 2e_2 = \delta - 1$, например, $e_1 = \min(\delta - 2, m_1 - \mathrm{mod}_2(\delta - m_1))$, $e_2 = \frac{1}{2}(\delta - 2 - e_1)$.) Поскольку

$$N_2(\mathcal{P}_i, 0) = 2^{M-1} - N_1(\mathcal{P}_i, 0), \quad N_2(\mathcal{P}_i, 1) = 2^{M-1} - N_1(\mathcal{P}_i, 1) \quad \forall i = 0,1,2,$$

отсюда заключаем

$$N_2(\mathcal{P}_1, 0) > N_2(\mathcal{P}_1, 1), \quad N_2(\mathcal{P}_2, 0) > N_2(\mathcal{P}_2, 1).$$

Следовательно, так как

$$N_i(\mathcal{P}_1, 0) = |C_i| - N_i(\mathcal{P}_1, 1), \quad N_i(\mathcal{P}_2, 0) = |C_i| - N_i(\mathcal{P}_2, 1) \quad \forall i = 1,2\ ,$$

то

$$N_1(\mathcal{P}_1, 1) > N_1(\mathcal{P}_0, 1) = N_1(\mathcal{P}_0, 0) > N_1(\mathcal{P}_1, 0),$$

$$N_1(\mathcal{P}_2, 1) > N_1(\mathcal{P}_0, 1) = N_1(\mathcal{P}_0, 0) > N_1(\mathcal{P}_2, 0)$$

и

$$N_2(\mathcal{P}_1, 0) > N_2(\mathcal{P}_0, 0) = N_2(\mathcal{P}_0, 1) > N_2(\mathcal{P}_1, 1),$$

$$N_2(\mathcal{P}_2, 0) > N_2(\mathcal{P}_0, 0) = N_2(\mathcal{P}_0, 1) > N_2(\mathcal{P}_2, 1).$$

Складывая

$$\sum_{e_1\geq 0, e_2\geq 0:\ e_1+2e_2\geq \delta-2} \frac{1}{e_1!(m_1-e_1-1)!}\frac{1}{e_2!(m_2-e_2-1)!} >$$

$$> \sum_{e_1\geq 0, e_2\geq 0:\ e_1+2e_2\geq \delta-1} \frac{1}{e_1!(m_1-e_1-1)!}\frac{1}{e_2!(m_2-e_2-1)!}$$

(вследствие (15) сумма в левой части этого неравенства содержит дополнительные члены, как и выше) с

$$\sum_{e_1\geq 0, e_2\geq 0:\ e_1+2e_2\geq \delta-2} \frac{1}{(e_1-1)!(m_1-e_1)!}\frac{1}{e_2!(m_2-e_2-1)!} =$$

$$= \sum_{e_1\geq 0, e_2\geq 0:\ e_1+2e_2\geq \delta-1} \frac{1}{e_1!(m_1-1-e_1)!}\frac{1}{(e_2-1)!(m_2-e_2)!}$$

(это тождество может быть доказано с использованием новых индексов суммирования, $e_2' = e_2+1$ в левой части, и $e_1' = e_1+1$ – в правой; как обычно, при $e<0$ $e! = \infty$), получаем

$$\sum_{e_1\geq 0, e_2\geq 0:\ e_1+2e_2\geq \delta-2} \frac{m_1}{e_1!(m_1-e_1)!e_2!(m_2-e_2-1)!} >$$

$$> \sum_{e_1\geq 0, e_2\geq 0:\ e_1+2e_2\geq \delta-1} \frac{m_2}{e_1!(m_1-1-e_1)!e_2!(m_2-e_2)!}.$$

Умножая полученное неравенство на $(m_1-1)!(m_2-1)!$, заключаем, что $N_1(\mathcal{P}_2, 1) > N_1(\mathcal{P}_1, 1)$, откуда

$$N_1(\mathcal{P}_2, 1) > N_1(\mathcal{P}_1, 1) > N_1(\mathcal{P}_0, 1) = N_1(\mathcal{P}_0, 0) > N_1(\mathcal{P}_1, 0) > N_1(\mathcal{P}_2, 0) \tag{17.1}$$

и

$$N_2(\mathcal{P}_2, 0) > N_2(\mathcal{P}_1, 0) > N_2(\mathcal{P}_0, 0) = N_2(\mathcal{P}_0, 1) > N_2(\mathcal{P}_1, 1) > N_2(\mathcal{P}_2, 1). \tag{17.2}$$

Ввиду (17.1), при $A_1 = N_1(\mathcal{P}_1, 1)$, $R_1 = \infty$ некоторый признак (p, b) характеризует первый класс, если и только если $p \in \mathcal{P}_1 \sqcup \mathcal{P}_2$ и $b = 1$. Аналогично, вследствие (17.2) при $A_2 = N_2(\mathcal{P}_2, 0)$, $R_2 = \infty$ второй класс характеризуется признаком (p, b), если и только если $p \in \mathcal{P}_2$ и $b = 0$. Для произвольного вектора $\mathbf{v}$ обозначим $n_i = |\{p \in \mathcal{P}_i | \mathbf{v}_p = 0\}|$, $i = 1, 2$. Таким образом, при выбранных величинах свободных параметров вектор будет отнесен к C_1, если и только если

$$E_1(\mathbf{v}) - E_2(\mathbf{v}) = e_1 + e_2 - n_2 = e_1 + 2e_2 - m_2 \geq \Delta,$$

и, следовательно, выбор $\Delta = \delta - m_2$ гарантирует правильное распознавание алгоритмом Кора-1.$\triangleright$

Замечание 13. C_2 состоит из векторов, для которых

$$n_1 + 2n_2 \geq \delta' = m_1 + 2m_2 - \delta + 1$$

(см. замечание 12), причем из (15) следует $2 \leq \delta' \leq m_1 + 2m_2 - 1$. Поэтому при условиях леммы 1 вследствие взаимозаменяемости классов при $A_1 = N_1(\mathcal{P}_2, 1)$, $R_1 = \infty$, $A_2 = N_2(\mathcal{P}_1, 0)$, $R_2 = \infty$ вектор $\mathbf{v}$ будет отнесен к C_2 если и только если $E_1(\mathbf{v}) - E_2(\mathbf{v}) = e_2 - n_1 - n_2 = m_1 - n_1 - 2n_2 < \Delta$, и, значит, алгоритм Кора-1 будет правильно распознавать рассматриваемую классификацию при $\Delta = \delta - 2m_2$.

Лемма 2. Пусть кодирование объектов сюръективно, $m_1 > 1$, $m_2 > 0$ *и*

$$2 < \delta < m_1 + 2m_2 - 1. \tag{18}$$

Тогда любая классификация, рассмотренная в части i теоремы 6, самораспознаваема алгоритмом Кора-n, $\forall n > 1$.

Доказательство. Обозначим через $N_i(\mathcal{P}_{n_1}, \mathcal{P}_{n_2}, b_1, b_2)$ число векторов, имеющих признак $\{(p_1, b_1), (p_2, b_2)\}$ ранга 2, где $p_i \in \mathcal{P}_{n_i}$, $0 \leq n_i \leq 2$, $i = 1, 2$ и $p_1 \neq p_2$. Очевидно, $\forall b_1, b_2 \in \{0, 1\}$

$$N_1(\mathcal{P}_0, \mathcal{P}_i, b_1, b_2) \leq N_1(\mathcal{P}_0, b_1) \quad (0 \leq i \leq 2, \text{ или } i = 1, 2 \text{ если } m_0 = 1), \tag{19.1}$$

$$N_1(\mathcal{P}_i, \mathcal{P}_j, 0, b_2) \leq N_1(\mathcal{P}_i, 0) \quad (i = 1, 2 \text{ и } j = 1, 2), \tag{19.2}$$

$$N_1(\mathcal{P}_1, \mathcal{P}_1, 1, 1) < N_1(\mathcal{P}_1, 1) \tag{19.3}$$

((19.3) может быть нарушено лишь, если $e_1 + 2e_2 \geq \delta \Rightarrow e_1 = m_1$; это имеет место только при $\delta = m_1 + 2m_2$, что несовместимо с (18)),

$$N_1(\mathcal{P}_1, \mathcal{P}_2, 1, 1) < N_1(\mathcal{P}_1, 1) \tag{19.4}$$

((19.4) может быть нарушено лишь, если $e_1 + 2e_2 \geq \delta \Rightarrow e_2 = m_2$; это имеет место только при $\delta \geq m_1 + 2m_2 - 1$, что несовместимо с (18)).

Возникают две возможности:

1. $m_2 = 1$, или

$$\sum_{e_1 \geq 0, e_2 \geq 0:\ e_1 + 2e_2 \geq \delta - 4} \binom{m_1}{e_1} \binom{m_2 - 2}{e_2} < \sum_{e_1 \geq 0, e_2 \geq 0:\ e_1 + 2e_2 \geq \delta - 1} \binom{m_1 - 1}{e_1} \binom{m_2}{e_2},$$

что эквивалентно

$$N_1(\mathcal{P}_2, \mathcal{P}_2, 1, 1) < N_1(\mathcal{P}_1, 1). \tag{19.5}$$

Ввиду (17.1), пять неравенств (19) показывают, что любой признак ранга 2 или выше имеют строго меньше $N_1(\mathcal{P}_1, 1)$ векторов из C_1.

Аналогично, всегда, кроме случая $i = j = 2$ и $b_1 = b_2 = 0$, в силу (17.2)

$$N_2(\mathcal{P}_i, \mathcal{P}_j, b_1, b_2) < N_2(\mathcal{P}_2, 0) \quad \forall b_1, b_2 \in \{0, 1\} \text{ и } 0 \leq i, j \leq 2.$$

При $m_2 > 1$ неравенство $N_2(\mathcal{P}_2, \mathcal{P}_2, 0, 0) < N_2(\mathcal{P}_2, 0)$ может быть нарушено лишь, если $e_1 + 2e_2 < \delta \Rightarrow e_2 = 0$; это имеет место только при $\delta < 2$, что несовместимо с (18). Следовательно, любой признак ранга 2 или выше имеют строго меньше $N_2(\mathcal{P}_2, 0)$ векторов из C_2.

Таким образом, согласно теореме 3 рассматриваемая классификация самораспознается алгоритмом Кора-n, $\forall n > 1$, при A_1, R_1, A_2, R_2 и Δ таких же, как в лемме 1.

2. $m_2 > 1$ и

$$\sum_{e_1 \geq 0, e_2 \geq 0:\ e_1 + 2e_2 \geq \delta - 4} \binom{m_1}{e_1} \binom{m_2 - 2}{e_2} \geq \sum_{e_1 \geq 0, e_2 \geq 0:\ e_1 + 2e_2 \geq \delta - 1} \binom{m_1 - 1}{e_1} \binom{m_2}{e_2}. \qquad (20)$$

В этом случае в доказательстве меняются роли классов.

C_2 состоит из векторов, для которых $n_1 + 2n_2 \geq \delta' = m_1 + 2m_2 - \delta + 1$, и из (18) следует $2 < \delta' < m_1 + 2m_2 - 1$. Такими же аргументами, как и выше, устанавливается, что

$$N_2(\mathcal{P}_0, \mathcal{P}_i, b_1, b_2) \leq N_2(\mathcal{P}_0, b_1) \quad (0 \leq i \leq 2, \text{ или } i = 1, 2, \text{ если } m_0 = 1), \qquad (21.1)$$

$$N_2(\mathcal{P}_i, \mathcal{P}_j, 1, b_2) \leq N_2(\mathcal{P}_i, 1) \quad (i = 1, 2 \text{ и } j = 1, 2), \qquad (21.2)$$

$$N_2(\mathcal{P}_1, \mathcal{P}_i, 0, 0) < N_2(\mathcal{P}_1, 0) \quad \forall i = 1, 2. \qquad (21.3)$$

Также

$$\sum_{n_1 \geq 0, n_2 \geq 0:\ n_1 + 2n_2 \geq \delta' - 4} \binom{m_1}{n_1} \binom{m_2 - 2}{n_2} =$$

$$= 2^{m_1 + m_2 - 2} - \sum_{n_1 \geq 0, n_2 \geq 0:\ n_1 + 2n_2 < \delta' - 4} \binom{m_1}{n_1} \binom{m_2 - 2}{n_2} =$$

(используются новые индексы суммирования $e_1 = m_1 - n_1$, $e_2 = m_2 - 2 - n_2$)

$$= 2^{m_1 + m_2 - 2} - \sum_{e_1 \geq 0, e_2 \geq 0:\ e_1 + 2e_2 \geq \delta} \binom{m_1}{e_1} \binom{m_2 - 2}{e_2} =$$

$$= 2^{m_1 + m_2 - 2} - \sum_{e_1 \geq 0, e_2 \geq 0:\ e_1 + 2e_2 \geq \delta - 4} \binom{m_1}{e_1} \binom{m_2 - 2}{e_2} +$$

$$+ \sum_{e_1 \geq 0, e_2 \geq 0:\ \delta - 4 \leq e_1 + 2e_2 < \delta} \binom{m_1}{e_1} \binom{m_2 - 2}{e_2} \leq$$

$$\leq 2^{m_1+m_2-1} - \sum_{e_1\geq 0, e_2\geq 0:\ e_1+2e_2\geq \delta-4} \binom{m_1}{e_1}\binom{m_2-2}{e_2} \leq$$

$$\leq 2^{m_1+m_2-1} - \sum_{e_1\geq 0, e_2\geq 0:\ e_1+2e_2\geq \delta-1} \binom{m_1-1}{e_1}\binom{m_2}{e_2} = \qquad \text{(в силу (20))}$$

$$= \sum_{e_1\geq 0, e_2\geq 0:\ e_1+2e_2< \delta-1} \binom{m_1-1}{e_1}\binom{m_2}{e_2} =$$

(подстановка новых индексов суммирования $n_1 = m_1 - 1 - e_1,\ n_2 = m_2 - e_2$)

$$= \sum_{n_1\geq 0, n_2\geq 0:\ n_1+2n_2> \delta'-1} \binom{m_1-1}{n_1}\binom{m_2}{n_2} < \sum_{n_1\geq 0, n_2\geq 0:\ n_1+2n_2\geq \delta'-1} \binom{m_1-1}{n_1}\binom{m_2}{n_2},$$

что эквивалентно

$$N_2(\mathcal{P}_2, \mathcal{P}_2, 0, 0) < N_2(\mathcal{P}_1, 0). \tag{21.4}$$

Из неравенств (21) вместе с (17.2) следует, что любой признак ранга 2 или выше имеют строго меньше $N_2(\mathcal{P}_1, 0)$ векторов из C_2. Как и выше, любой признак ранга 2 или выше имеют строго меньше $N_1(\mathcal{P}_2, 1)$ векторов из C_1. Следовательно, и в этом случае по теореме 3 рассматриваемая классификация самораспознается алгоритмом Кора-n, $\forall n > 1$, при A_1, R_1, A_2, R_2 и Δ таких же, как в замечании 13.▷

Лемма 3. Пусть кодирование объектов сюръективно, $m_1 > 1$, $m_2 > 0$ *и* $\delta = 2$. *Тогда любая классификация, удовлетворяющая (15), самораспознаваема алгоритмом Кора-n* $\forall n > 1$.

Доказательство. Выберем $A_2 = \infty, R_2 = 0$, при которых ни один признак не характеризует C_2. При $\delta = 2$, $\mathbf{v} \in C_2$, если и только если $\mathbf{v}_p = 0\ \forall p \in \mathcal{P}_2$ и $\mathbf{v}_p = 1$ не более, чем для одного $p \in \mathcal{P}_1$ (см. (15)). Следовательно, при $A_1 = R_1 = 0$ признак $\{(p_i, b_i) | 1 \leq i \leq k\}$ ранга k характеризует C_1, если и только если выполнено хотя бы одно из следующих условий: либо $\exists p_i \in \mathcal{P}_2$ такой, что $b_i = 1$, либо $\exists p_i, p_j \in \mathcal{P}_1$ такие, что $p_i \neq p_j$ и $b_i = b_j = 1$. Таким образом, поскольку $n \geq 2$, любой $\mathbf{v} \in C_1$ имеет по крайней мере один признак, характеризующий C_1, и желаемое распознавание получается при $\Delta = 1$.▷

Очевидно, случай $m_2 > 0$ и $\delta = m_1 + 2m_2 - 1$ может быть приведен к случаю, рассмотренному в лемме 3, перестановкой классов C_1 и C_2.

Лемма 4. Пусть кодирование объектов сюръективно, $m_2 = 0$ *и*

$$2 \leq \delta \leq m_1 - 1. \tag{22}$$

Тогда любая классификация, рассмотренная в части i теоремы, самораспознаваема алгоритмом Кора-n, $\forall n \geq 1$.

Доказательство. Класс C_1 состоит из векторов, у которых не более $m_1 - \delta$ компонент из $\mathcal{P}_1$ – нули. Поэтому ввиду (22)

$$N_1(\mathcal{P}_1, 1) = 2^{m_0} \sum_{n=0}^{m_1-\delta} \binom{m_1-1}{n} > 2^{m_0} \sum_{n=0}^{m_1-\delta-1} \binom{m_1-1}{n} = N_1(\mathcal{P}_1, 0).$$

Так как $N_1(\mathcal{P}_1, 1) + N_1(\mathcal{P}_1, 0) = |C_1|$ и $N_1(\mathcal{P}_0, b) = \frac{1}{2}|C_1|$ $\forall b = 0, 1$, то

$$N_1(\mathcal{P}_1, 1) > N_1(\mathcal{P}_0, 1) = N_1(\mathcal{P}_0, 0) > N_1(\mathcal{P}_1, 0). \tag{23}$$

Следовательно, при $A_1 = N_1(\mathcal{P}_1, 1)$ и $R_1 = \infty$ любой признак ранга 1, характеризующий первый класс, имеет вид $(p, 1)$, $p \in \mathcal{P}_1$. Выберем $A_2 = \infty$, $R_2 = 0$, при которых ни один признак не характеризует C_2. Поскольку $\mathbf{v} \in C_1 \Leftrightarrow e_1 \geq \delta$, алгоритм Кора-1 распознает желаемую классификацию при $\Delta = \delta$.

Очевидно, $\forall b_1, b_2 \in \{0, 1\}$

$$N_1(\mathcal{P}_0, \mathcal{P}_i, b_1, b_2) \leq N_1(\mathcal{P}_0, b_1) \quad (i = 0, 1 \text{ или } i = 1 \text{ если } m_0 = 1),$$

$$N_1(\mathcal{P}_1, \mathcal{P}_i, 0, b_2) \leq N_1(\mathcal{P}_1, 0) \quad (i = 0, 1)$$

и вследствие (22)

$$N_1(\mathcal{P}_1, \mathcal{P}_1, 1, 1) < N_1(\mathcal{P}_1, 1).$$

Поэтому ввиду (23)

$$N_1(\mathcal{P}_i, \mathcal{P}_j, b_1, b_2) < N_1(\mathcal{P}_0, 1), \quad \forall b_1, b_2 \in \{0, 1\},\ i = 0, 1,\ j = 0, 1$$

и по теореме 3 рассматриваемая классификация самораспознается алгоритмом Кора-n, $\forall n > 1$, при том же наборе свободных параметров, что и при $n = 1$. ▷

Замечание 14. Так как C_1 и C_2, рассмотренные в лемме 4, взаимно перестановочны (см. замечание 12), то эта же классификация самораспознается алгоритмом Кора-n, $\forall n \geq 1$, при $A_1 = \infty$, $R_1 = 0$, $A_2 = N_2(\mathcal{P}_1, 0)$, $R_2 = \infty$ и $\Delta = m_1 - \delta$.

Доказательство теоремы завершено.▷

IV. НЕОБХОДИМОЕ УСЛОВИЕ РАСПОЗНАВАЕМОСТИ КЛАССИФИКАЦИЙ АЛГОРИТМОМ КОРА-2

Если классификация $O = C_1 \sqcup C_2$ распознаваема с некоторым материалом обучения алгоритмом Кора-n, то C_1 состоит из $\mathbf{v} \in O$ таких, что

$$\sum_{k=1}^{n} \sum_{\{(p_i, b_i)|1 \leq i \leq k\}} \epsilon_{p_1,\ldots,p_k}^{b_1,\ldots,b_k} \sigma_{p_1,\ldots,p_k}^{b_1,\ldots,b_k}(\mathbf{v}) \geq \Delta. \tag{24}$$

Здесь внешняя сумма взята по всем комбинациям из k бинарных значений $b_i \in \{0, 1\}$ и k дескрипторов $1 \leq p_1 < \ldots < p_k \leq M$; $\sigma_{p_1,\ldots,p_k}^{b_1,\ldots,b_k}(\mathbf{v})$ – характеристическая функция признака $\{(p_i, b_i)|1 \leq i \leq k\}$; $\epsilon_{p_1,\ldots,p_k}^{b_1,\ldots,b_k} = 1$, если признак $\{(p_i, b_i)|1 \leq i \leq k\}$ характеризует только первый класс, $\epsilon_{p_1,\ldots,p_k}^{b_1,\ldots,b_k} = -1$, если признак $\{(p_i, b_i)|1 \leq i \leq k\}$ характеризует только второй класс, и $\epsilon_{p_1,\ldots,p_k}^{b_1,\ldots,b_k} = 0$ в противном случае. В принципе, в различных задачах эти коэффициенты могут принимать любые комбинации указанных значений.

Поскольку

$$\sum_{\{b_i \in \{0,1\}|1 \leq i \leq k\}} \sigma_{p_1,\ldots,p_k}^{b_1,\ldots,b_k}(\mathbf{v}) = 1 \quad \forall \mathbf{v} \in X^M \tag{25}$$

(здесь сумма взята по всем комбинациям из k бинарных значений b_i) для любого множества $k \geq 1$ попарно различных дескрипторов p_i, (24) эквивалентно неравенству

$$\sum_{k=1}^{n} \sum_{\{(p_i,b_i)|1\leq i\leq k\}} \bar{\epsilon}_{p_1,\ldots,p_k}^{b_1,\ldots,b_k} \sigma_{p_1,\ldots,p_k}^{b_1,\ldots,b_k}(\mathbf{v}) \geq \Delta', \tag{26}$$

где $0 \leq \bar{\epsilon}_{p_1,\ldots,p_k}^{b_1,\ldots,b_k} \leq 2$. ((26) является обобщением (15) на случай произвольного n.) Используя тождество

$$\sigma_{p_1,\ldots,p_{k-1},p_k}^{b_1,\ldots,b_{k-1},0}(\mathbf{v}) = \sigma_{p_1,\ldots,p_{k-1}}^{b_1,\ldots,b_{k-1}}(\mathbf{v}) - \sigma_{p_1,\ldots,p_{k-1},p_k}^{b_1,\ldots,b_{k-1},1}(\mathbf{v}) \quad \forall \mathbf{v} \in X^M \tag{27}$$

(верное для любого множества $k > 1$ попарно различных дескрипторов p_i и любой комбинации $k-1$ бинарных значений $b_i \in \{0,1\}$) и (25) для $k=1$, (26) можно переписать в виде

$$\sum_{k=1}^{n} \sum_{\{p_i|1\leq i\leq k\}} \hat{\epsilon}_{p_1,\ldots,p_k} \sigma_{p_1,\ldots,p_k}^{1,\ldots,1}(\mathbf{v}) \geq \delta, \tag{28}$$

где константы $\hat{\epsilon}_{p_1,\ldots,p_k}$ и δ определяются значениями $\epsilon_{p_1,\ldots,p_k}^{b_1,\ldots,b_k}$ и Δ. Тождество

$$\sigma_{p_1,\ldots,p_k}^{1,\ldots,1}(\mathbf{v}) = \prod_{i=1}^{k} \mathbf{v}_{p_i} \tag{29}$$

приводит (28) к виду

$$\sum_{k=1}^{n} \sum_{\{p_i|1\leq i\leq k\}} \left(\hat{\epsilon}_{p_1,\ldots,p_k} \prod_{i=1}^{k} \mathbf{v}_{p_i} \right) \geq \delta. \tag{30}$$

Цель настоящего раздела – выяснить, какие наборы коэффициентов $\hat{\epsilon}_{p_1,\ldots,p_k}$ *приемлемы* при $n=2$, т.е. для каких коэффициентов (30) эквивалентно исходному неравенству (24) с некоторыми целыми $\epsilon_{p_1,\ldots,p_k}^{b_1,\ldots,b_k}$, $-1 \leq \epsilon_{p_1,\ldots,p_k}^{b_1,\ldots,b_k} \leq 1$.

При $n=2$ (24) имеет вид

$$\sum_{p_1=1}^{M-1} \sum_{p_2=p_1+1}^{M} (\epsilon_{p_1,p_2}^{0,0}\sigma_{p_1,p_2}^{0,0}(\mathbf{v}) + \epsilon_{p_1,p_2}^{0,1}\sigma_{p_1,p_2}^{0,1}(\mathbf{v}) + \epsilon_{p_1,p_2}^{1,0}\sigma_{p_1,p_2}^{1,0}(\mathbf{v}) +$$

$$+\epsilon_{p_1,p_2}^{1,1}\sigma_{p_1,p_2}^{1,1}(\mathbf{v})) + \sum_{p=1}^{M}(\epsilon_p^0\sigma_p^0(\mathbf{v}) + \epsilon_p^1\sigma_p^0(\mathbf{v})) \geq \Delta.$$

Преобразуя левую часть этого неравенства, используя (27) для $k=2$, (25) для $k=1$ и (29), получаем

$$\sum_{p=1}^{M}\left(\epsilon_p^1-\epsilon_p^0+\sum_{p_2=1}^{p-1}(\epsilon_{p_2,p}^{0,1}-\epsilon_{p_2,p}^{0,0})+\sum_{p_2=p+1}^{M}(\epsilon_{p,p_2}^{1,0}-\epsilon_{p,p_2}^{0,0})\right)\mathbf{v}_p+$$
$$+\sum_{p_1=1}^{M-1}\sum_{p_2=p_1+1}^{M}(\epsilon_{p_1,p_2}^{0,0}-\epsilon_{p_1,p_2}^{0,1}-\epsilon_{p_1,p_2}^{1,0}+\epsilon_{p_1,p_2}^{1,1})\mathbf{v}_{p_1}\mathbf{v}_{p_2}\geq$$
$$\geq\Delta-\sum_{p=1}^{M}\epsilon_p^1-\sum_{p_1=1}^{M-1}\sum_{p_2=p_1+1}^{M}\epsilon_{p_1,p_2}^{0,0}. \tag{31}$$

Это неравенство имеет форму

$$\sum_{p=1}^{M}\epsilon_p\mathbf{v}_p+\sum_{p_1=1}^{M-1}\sum_{p_2=p_1+1}^{M}\epsilon_{p_1,p_2}\mathbf{v}_{p_1}\mathbf{v}_{p_2}\geq\delta, \tag{32}$$

которую (30) принимает при $n=2$. Сравнение (31) и (32) показывает, что набор коэффициентов (31) приемлем, если они допускают представление

$$\epsilon_p=\epsilon_p^1-\epsilon_p^0+\sum_{p_2=1}^{p-1}(\epsilon_{p_2,p}^{0,1}-\epsilon_{p_2,p}^{0,0})+\sum_{p_2=p+1}^{M}(\epsilon_{p,p_2}^{1,0}-\epsilon_{p,p_2}^{0,0}) \tag{33}$$

и

$$\epsilon_{p_1,p_2}=\epsilon_{p_1,p_2}^{0,0}-\epsilon_{p_1,p_2}^{0,1}-\epsilon_{p_1,p_2}^{1,0}+\epsilon_{p_1,p_2}^{1,1}\qquad(p_1<p_2) \tag{34}$$

при некоторых допустимых коэффициентах

$$-1\leq\epsilon_p^b\leq1,\quad -1\leq\epsilon_{p_1,p_2}^{b_1,b_2}\leq1\qquad(p_1<p_2). \tag{35}$$

Обозначим

$$\tau_{p_1,p_2}=\begin{cases}\epsilon_{p_1,p_2}^{1,0}-\epsilon_{p_1,p_2}^{0,0}, & \text{если } p_1<p_2, \qquad (36.1)\\ 0, & \text{если } p_1=p_2, \qquad (36.2)\\ \epsilon_{p_2,p_1}^{0,1}-\epsilon_{p_2,p_1}^{0,0}, & \text{если } p_1>p_2. \qquad (36.3)\end{cases}$$

В новых переменных

$$\epsilon_p=\epsilon_p^1-\epsilon_p^0+\sum_{p_2=1}^{M}\tau_{p,p_2}. \tag{37}$$

Таким образом, коэффициенты (32), ϵ_p и ϵ_{p_1,p_2}, приемлемы, если они задаются равенствами (34), (36) и (37) при некоторых целых $\epsilon_{p_1,\dots,p_k}^{b_1,\dots,b_k}$, удовлетворяющих (35). Значения $\epsilon_{p_1,p_2}^{1,0}$, $\epsilon_{p_1,p_2}^{0,1}$ и $\epsilon_{p_2,p_1}^{1,1}$, полученные из (34), (36.1) и (36.3), допустимы, если и только если

$$-1\leq\epsilon_{p_1,p_2}^{1,0}=\tau_{p_1,p_2}+\epsilon_{p_1,p_2}^{0,0}\leq1, \tag{38.1}$$
$$-1\leq\epsilon_{p_1,p_2}^{0,1}=\tau_{p_2,p_1}+\epsilon_{p_1,p_2}^{0,0}\leq1, \tag{38.2}$$
$$-1\leq\epsilon_{p_1,p_2}^{1,1}=\epsilon_{p_1,p_2}^{0,0}+\tau_{p_1,p_2}+\tau_{p_2,p_1}+\epsilon_{p_1,p_2}\leq1. \tag{38.3}$$

Неравенства (38) совместны и определяют допустимое $\epsilon^{0,0}_{p_1,p_2}$, если и только если

$$\max(-1, -1-\tau_{p_1,p_2}, -1-\tau_{p_2,p_1}, -1-\epsilon_{p_1,p_2}-\tau_{p_1,p_2}-\tau_{p_2,p_1}) \leq$$

$$\leq \min(1, 1-\tau_{p_1,p_2}, 1-\tau_{p_2,p_1}, 1-\epsilon_{p_1,p_2}-\tau_{p_1,p_2}-\tau_{p_2,p_1}), \tag{39}$$

где $p_1 < p_2$. Следовательно, набор коэффициентов ϵ_p и ϵ_{p_1,p_2} приемлем, если и только если ϵ_p представимы в виде (37), где ϵ^b_p удовлетворяют (35) и целые τ_{p_1,p_2} удовлетворяют (36.2) и (39).

Неравенство (39) представляет собой систему из 12 нетривиальных индивидуальных неравенств, эквивалентных следующим двум:

$$\max(-2, -2-\epsilon_{p_1,p_2}) \leq \tau_{p_1,p_2} \leq \min(2, 2-\epsilon_{p_1,p_2}), \tag{40.1}$$

$$\begin{aligned}&\max(-2, -2-\epsilon_{p_1,p_2}, -2+\tau_{p_1,p_2}, -2-\epsilon_{p_1,p_2}-\tau_{p_1,p_2}) \leq \\ &\leq \tau_{p_2,p_1} \leq \min(2, 2-\epsilon_{p_1,p_2}, 2+\tau_{p_1,p_2}, 2-\epsilon_{p_1,p_2}-\tau_{p_1,p_2}),\end{aligned} \tag{40.2}$$

где $p_1 < p_2$. В новых переменных

$$\xi_{p_1,p_2} = \begin{cases} \tau_{p_1,p_2} - \frac{1}{2}\epsilon_{p_1,p_2}, & \text{если } \ p_1 < p_2, \\ 0, & \text{если } p_1 = p_2, \\ \tau_{p_1,p_2} - \frac{1}{2}\epsilon_{p_2,p_1}, & \text{если } \ p_1 > p_2 \end{cases}$$

(40) приводится к эквивалентному виду

$$|\xi_{p_1,p_2}| \leq 2 - \frac{1}{2}|\epsilon_{p_1,p_2}|, \tag{41.1}$$

$$|\xi_{p_2,p_1}| \leq 2 - \frac{1}{2}|\epsilon_{p_1,p_2}|, \tag{41.2}$$

$$|\xi_{p_1,p_2}| + |\xi_{p_2,p_1}| \leq 2, \tag{41.3}$$

где $p_1 < p_2$. Из (41.1) и (41.2) следует

$$|\epsilon_{p_1,p_2}| \leq 4 \tag{42}$$

(ср. (34)). Следовательно, набор коэффициентов ϵ_p и ϵ_{p_1,p_2} приемлем, если и только если выполнены (42) и (37), которое в терминах ξ принимает вид

$$\epsilon_p = \epsilon^1_p - \epsilon^0_p - \frac{1}{2}\left(\sum_{p_2=1}^{p-1} \epsilon_{p_2,p} + \sum_{p_2=p+1}^{M} \epsilon_{p,p_2}\right) + \sum_{p_2=1}^{M} \xi_{p,p_2}, \tag{43}$$

где ξ_{p_1,p_2} и ξ_{p_2,p_1} – пары любых чисел, удовлетворяющих (41), одновременно целых или полуцелых в зависимости от того, четно или нечетно ϵ_{p_1,p_2}.

При $|\epsilon_{p_1,p_2}| \geq 2$ (41.3) следует из (41.1) и (41.2); при $\epsilon_{p_1,p_2} = 0$ (41.1) и (41.2) следуют из (41.3). Если $|\epsilon_{p_1,p_2}|=1$, любое полуцелое решение (41) допускает представление $\xi_{p_1,p_2} = \mu_{p_1,p_2} + \nu_{p_1,p_2}$ и $\xi_{p_2,p_1} = \mu_{p_2,p_1} + \nu_{p_2,p_1}$, где $|\mu_{p_1,p_2}| = |\mu_{p_2,p_1}| = \frac{1}{2}$ и их знаки произвольны, ν_{p_1,p_2} и ν_{p_2,p_1} – целые и $|\nu_{p_1,p_2}| + |\nu_{p_2,p_1}| \leq 1$.

Ввиду (35) и частичной взаимной независимости ξ_{p_1,p_2} и ξ_{p_2,p_1}, описанной в предыдущем абзаце, из (43) следует, что набор коэффициентов (32) приемлем, если и только если выполнено (42) и

$$\left|\epsilon_p + \frac{1}{2}\left(\sum_{p_2=1}^{p-1} \epsilon_{p_2,p} + \sum_{p_2=p+1}^{M} \epsilon_{p,p_2}\right)\right| \leq 2 + \frac{1}{2}\left(\sum_{p_2=1}^{p-1} \min(4 - |\epsilon_{p_2,p}|, |\epsilon_{p_2,p}|) + \right.$$
$$\left. + \sum_{p_2=p+1}^{M} \min(4 - |\epsilon_{p,p_2}|, |\epsilon_{p,p_2}|)\right) + \sum_{p_2=1}^{M} \nu_{p,p_2}. \tag{44}$$

Здесь ν_{p_1,p_2} – целые, удовлетворяющие (предполагая $p_1 < p_2$) неравенствам

$$\nu_{p_1,p_2} \geq 0, \quad \nu_{p_2,p_1} \geq 0, \quad \nu_{p_1,p_2} + \nu_{p_2,p_1} \leq \zeta_{p_1,p_2}, \tag{45}$$

где

$$\zeta_{p_1,p_2} = \begin{cases} 2, & \text{если } \epsilon_{p_1,p_2} = 0, \\ 1, & \text{если } |\epsilon_{p_1,p_2}| = 1, \\ 0, & \text{если } |\epsilon_{p_1,p_2}| \geq 2 \text{ или } p_1 = p_2. \end{cases} \tag{46}$$

Лемма 5. Целые числа $\Gamma_p \geq 0$, $1 \leq p \leq M$ допускают представление

$$\Gamma_p = \sum_{p_2=1}^{M} \nu_{p,p_2}, \tag{47}$$

где ν_{p_1,p_2} – целые числа, удовлетворяющие (45), если и только если

$$\sum_{p=1}^{M} \Gamma_p \mathbf{v}_p \leq \sum_{p_1=1}^{M-1} \sum_{p_2=p_1+1}^{M} \min(1, \mathbf{v}_{p_1} + \mathbf{v}_{p_2})\zeta_{p_1,p_2} \quad \forall \mathbf{v} \in X^M \ (\mathbf{v} \neq \mathbf{0}). \tag{48}$$

Доказательство. Неравенства (48) являются очевидными следствиями (45) и (47).

Задача о представлении чисел Γ_p будет решена итеративно так, что на каждом шагу $\sum_{p=1}^{M} \Gamma_p$ и/или размерность M строго уменьшаются. Поскольку для $\Gamma_p = 0$ $\forall p$ или $M \leq 2$ задача тривиальна, построение такой редукции доказывает лемму.

Рассмотрим две возможности.

i. Неравенства (48) являются строгими $\forall \mathbf{v} \neq \mathbf{0}$, кроме, быть может, для $\mathbf{v}_p = 1$ $\forall p$. Если итерации еще не сошлись к финальному состоянию $\sum_{p=1}^{M} \Gamma_p = 0$, существует пара индексов $p_1' < p_2'$ таких, что $\zeta_{p_1',p_2'} > 0$ и $\Gamma_{p_1'} > 0$ и/или $\Gamma_{p_2'} > 0$. Тогда задача сводится к эквивалентной подстановками

$$\Gamma_p' = \begin{cases} \Gamma_p, & \text{если } p \neq p_1', \\ \Gamma_{p_1'} - 1, & \text{если } p = p_1' \end{cases}$$

(если $\Gamma_{p'_1} > 0$) или

$$\Gamma'_p = \begin{cases} \Gamma_p, & \text{если } p \neq p'_2, \\ \Gamma_{p'_2} - 1, & \text{если } p = p'_2 \end{cases}$$

(в противном случае) и

$$\zeta'_{p_1,p_2} = \begin{cases} \zeta_{p_1,p_2}, & \text{если } p_1 \neq p'_1 \text{ или } p_2 \neq p'_2, \\ \zeta_{p'_1,p'_2} - 1, & \text{если } p_1 = p'_1 \text{ и } p_2 = p'_2. \end{cases}$$

Очевидно, неравенства (48) выполнены и для новых величин $\Gamma'_p \geq 0$ и ζ'_{p_1,p_2}, а $\sum_{p=1}^{M} \Gamma_p$ уменьшается на 1.

ii. $\exists \mathbf{v}^* \neq \mathbf{0}$, для которого (48) – равенство, и $\mathbf{v}^*_p = 0$ хотя бы для одного дескриптора. Предположим, что $\mathbf{v}^*_p = 1$ при $1 \leq p \leq p^*$ и $\mathbf{v}^*_p = 0$ при $p^* + 1 \leq p \leq M$ (при необходимости, изменив порядок дескрипторов). Заметим, что $\Gamma_p \geq \sum_{p_2=p^*+1}^{M} \zeta_{p,p_2} \quad \forall p,\ 1 \leq p \leq p^*$. (Пусть для некоторого индекса p', $1 \leq p' \leq p^*$, это неверно. Сложим неравенство $\Gamma_{p'} < \sum_{p_2=p^*+1}^{M} \zeta_{p',p_2}$ с (48), в котором $\mathbf{v}_p = 1$, если и только если $1 \leq p \leq p^*$ и $p \neq p'$. Из полученного неравенства тогда следует, что (48) для $\mathbf{v}^*$ – тоже строгое неравенство.) При $p_1 \leq p^* < p_2$ положим $\nu_{p_1,p_2} = \zeta_{p_1,p_2}$, $\nu_{p_2,p_1} = 0$. Тогда исходная задача распадается на две подзадачи меньших размерностей: $M'_1 = p^*$ для $\Gamma'_p = \Gamma_p - \sum_{p_2=p^*+1}^{M} \zeta_{p,p_2}$, $1 \leq p \leq M'_1$, и $M'_2 = M - p^*$ для Γ_p, $p^* < p \leq M$. Для первой подзадачи (48) выполнено, так как оно выполнено для исходной задачи $\forall \mathbf{v} \in X^M$ с $\mathbf{v}_p = 0\ \forall p > p^*$. Аналогично, для второй подзадачи (48) выполнено, так как оно выполнено для исходной задачи $\forall \mathbf{v} \in X^M$ с $\mathbf{v}_p = 1\ \ \forall p \leq p^*$ и поскольку для $\mathbf{v}^*$ (48) – равенство.$\triangleright$

Применение леммы 5 к последней сумме в (44) устанавливает необходимое условие распознаваемости классификаций алгоритмом Кора-2:

Теорема 7. Если классификация $O = C_1 \sqcup C_2$ распознаваема при использовании некоторого материала обучения алгоритмом Кора-2, то C_1 состоит из $\mathbf{v} \in O$, удовлетворяющих (32), где ϵ_{p_1,p_2} – произвольные целые числа, удовлетворяющие (42), и

$$\left| \epsilon_p + \frac{1}{2} \left(\sum_{p_2=1}^{p-1} \epsilon_{p_2,p} + \sum_{p_2=p+1}^{M} \epsilon_{p,p_2} \right) \right| \leq$$

$$\leq 2 + \Gamma_p + \frac{1}{2} \left(\sum_{p_2=1}^{p-1} \min(4 - |\epsilon_{p_2,p}|, |\epsilon_{p_2,p}|) + \sum_{p_2=p+1}^{M} \min(4 - |\epsilon_{p,p_2}|, |\epsilon_{p,p_2}|) \right). \quad (49)$$

Здесь Γ_p – произвольные целые неотрицательные числа, удовлетворяющие (48), в котором ζ_{p_1,p_2} задаются равенствами (46). Это утверждение также выполнено в отношении C_2.$\triangleright$

Замечание 15. В силу (33) в любом приемлемом наборе коэффициентов (32) $|\epsilon_p| \leq 2M$.

V. ПРОСТРАНСТВА ОБЪЕКТОВ МАЛОЙ РАЗМЕРНОСТИ

Для рассмотрения вопросов, сформулированных во Введении, был выполнен исчерпывающий анализ всех возможных классификаций с сюръективным кодированием объектов в пространствах размерности $3 \le M \le 5$.

V.1. Генерация изотипических классификаций

Генерировать все 2^{2^M} разбиений $X^M = C_1 \sqcup C_2$ не практично, так как следующие естественные эквивалентности классификаций позволяют существенно уменьшить число вариантов:

1. Классы C_1 и C_2 можно взаимно переставить (т.е. номера классов взаимозаменяемы).

2. Значения 0 и 1 каждого данного дескриптора можно взаимно переставить одновременно во всех векторах (утвердительный ответ на данный вопрос можно кодировать как нулем, так и единицей).

3. Порядок дескрипторов (т.е. вопросов в вопроснике для кодирования объектов) несущественен.

4. Порядок векторов в каждом классе C_1 и C_2 несущественен.

Иными словами, классы эквивалентности классификаций – это орбиты действия группы симметрий M-мерного куба на множестве разбиений вершин этого куба на два подмножества.

Пример. Существует 13 классов эквивалентности классификаций $X^M = C_1 \sqcup C_2$ при $M = 3$ (см. рис. 1).

Для $i = 1, 2$ пусть V_i обозначает матрицу размером $M \times |C_i|$, составленную из векторов-строк $\mathbf{v}^j \in C_i$.

Генерировались изотипические представители каждого класса эквивалентности, удовлетворяющие следующим условиям:

1. $|C_1| \le |C_2|$.

2. Вектор $\mathbf{v}$ с $\mathbf{v}_p = 1\ \forall p,\ 1 \le p \le M$, принадлежит C_1.

3. Столбцы матрицы $V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$, а также строки V_1 и V_2 лексикографически убывают.

Условие 1 разрешает эквивалентность 1, если $|C_1| \ne |C_2|$. Условие 2 частично разрешает эквивалентность 2 кодирования объектов: его можно удовлетворить, изменяя кодирование соответствующих дескрипторов таким образом, чтобы некоторый фиксированный объект из C_1 оказался бы закодирован строкой единиц. Если кодирование фиксировано, то при данном порядке векторов упорядочение столбцов разрешает эквивалентность 3, а при данном порядке дескрипторов упорядочение строк разрешает эквивалентность 4. Правомерно требовать, чтобы оба порядка были соблюдены одновременно:

Лемма 6. Множество разбиений, взаимно эквивалентных посредством эквивалентностей 3 и 4, включает представителя, у которого и строки матриц V_1 и V_2, и столбцы V лексикографически убывают.

Доказательство. Матрице V поставим в соответствие двоичное число $\mathcal{V}(V) = \mathbf{v}_1^1 ... \mathbf{v}_M^1 \mathbf{v}_1^2 ... \mathbf{v}_M^2 ... \mathbf{v}_1^{2^M} ... \mathbf{v}_M^{2^M}$, образованное слиянием строк V. Каждая из двух операций: упорядочение неупорядоченных строк матриц V_1 и/или V_2, или

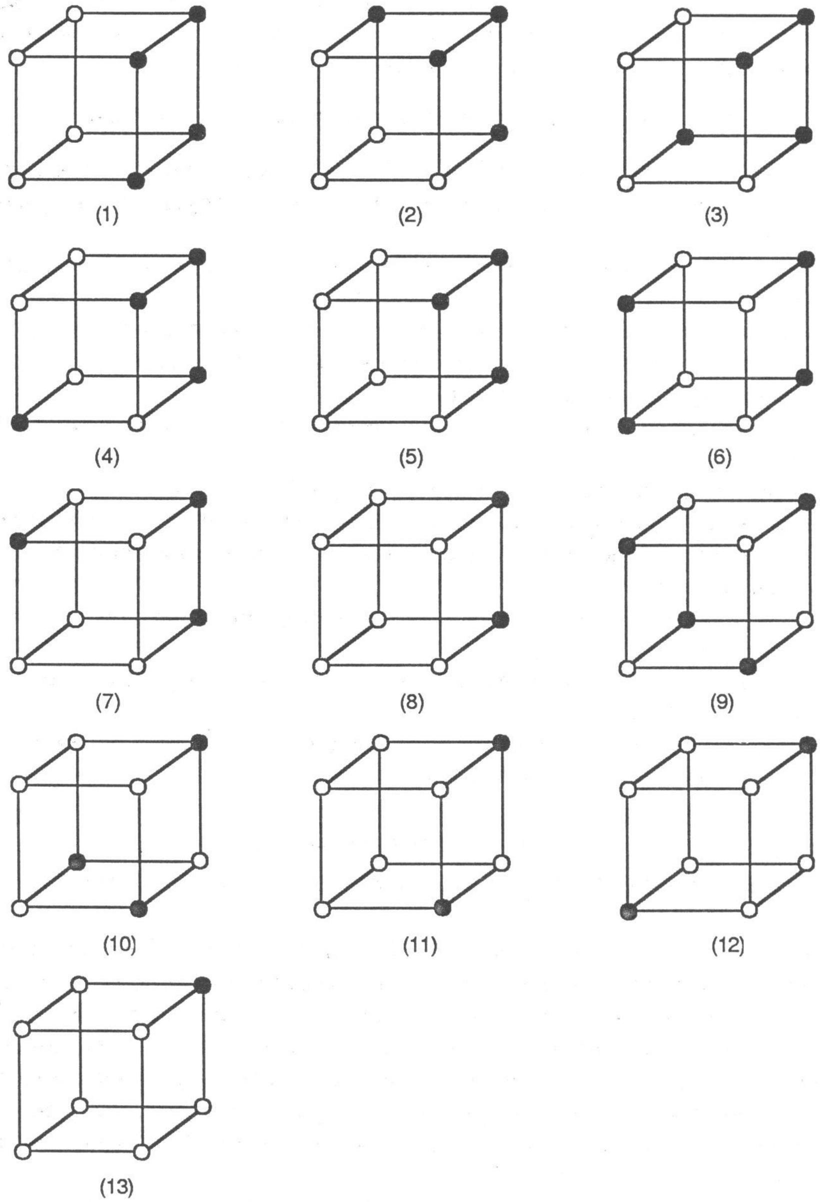

Рис. 1. **Изотипические представители каждого класса эквивалентности классификаций $C_1 \sqcup C_2 = X^3$ с максимальными индексами первого класса:**

(1) $\mathcal{I} = 11110000$, (2) $\mathcal{I} = 11101000$, (3) $\mathcal{I} = 11100100$, (4) $\mathcal{I} = 11100001$, (5) $\mathcal{I} = 11100000$, (6) $\mathcal{I} = 11000011$, (7) $\mathcal{I} = 11000010$, (8) $\mathcal{I} = 11000000$, (9) $\mathcal{I} = 10010110$, (10) $\mathcal{I} = 10010100$, (11) $\mathcal{I} = 10010000$, (12) $\mathcal{I} = 10000001$, (13) $\mathcal{I} = 10000000$; C_1 – черные вершины, C_2 – белые вершины

упорядочение неупорядоченных столбцов V – приводит к строгому увеличению числа $\mathcal{V}(V)$. Однако $\mathcal{V}(V) < 2^{M2^M}$, и, следовательно, после некоторого конечного числа поочередных применений этих операций последующее упорядочение ни строк, ни столбцов не может увеличить $\mathcal{V}(V)$. Соответственно, в построенной таким образом матрице V и строки, и столбцы упорядочены желаемым образом.▷

Замечание 16. При фиксированном кодировании каждого из дескрипторов и фиксированной нумерации классов C_1 и C_2 представитель данного класса эквивалентности, построенный в лемме 6, единственен.

Упорядочения достаточно производить в матрице V_1:

Лемма 7. Пусть в обозначениях, использованных выше,

i. кодирование объектов сюръективно,

ii. строки лексикографически убывают в обеих матрицах V_1 и V_2,

iii. столбцы лексикографически не возрастают в V_1.

Тогда столбцы матрицы V также лексикографически убывают.

Доказательство. Пусть P-й столбец в матрице V лексикографически предшествует $P+1$-му. Тогда $\exists J,\ 1 \le J \le 2^M$, такое, что

$$\mathbf{v}_P^j = \mathbf{v}_{P+1}^j \quad \forall j,\ 1 \le j < J, \tag{50}$$

и

$$\mathbf{v}_P^J < \mathbf{v}_{P+1}^J. \tag{51}$$

Поскольку V состоит из бинарных векторов, из (51) следует, что $\mathbf{v}_P^J = 0$ и $\mathbf{v}_{P+1}^J = 1$. Однако $\mathbf{v}^J \in V_2$ (в силу *iii*), поэтому любой вектор с $\mathbf{v}_p^{J'} = \mathbf{v}_p^J$ $\forall p,\ 1 \le p < P$, $\mathbf{v}_P^{J'} = 1$ и $\mathbf{v}_{P+1}^{J'} = 0$ является строкой матрицы V (в силу *i*), расположенной в V выше (в силу *ii*) $\mathbf{v}^J$: $J' < J$. Это противоречит (50).▷

Выбор представителя данного класса эквивалентности с максимальной величиной $\mathcal{V}(V)$ полностью разрешает эквивалентность 2 (неопределенность выбора объекта из C_1, закодированного строкой единиц) и эквивалентность 1, когда $|C_1| = |C_2|$. Изотипические классификации с максимальным $\mathcal{V}(V)$ реализуют также максимум в данном классе эквивалентности *индекса класса* C_1:

Определение 8. Пусть $\pi(\mathbf{v})$ обозначает порядковый номер вектора $\mathbf{v}$ в списке всех 2^M бинарных векторов из X^M в порядке лексикографического убывания. Индексом класса C_1 называется двоичное число $\mathcal{I} = i_1 i_2 \ldots i_{2^M}$, где $i_{\pi(\mathbf{v})} = 1$, если $\mathbf{v} \in C_1$, и $i_{\pi(\mathbf{v})} = 0$ в противном случае.

V.2. Результаты вычислений

Типы распознаваемости классификаций представлены в табл. 1, 2, 3 строками кодов длины $M-1$, где n-й символ описывает тип распознаваемости алгоритмом Кора-n. (Теорема 1 гарантирует, что любая классификация самораспознаваема посредством алгоритма Кора-M (см. замечание 8); таким образом, M-ый символ в строке был бы избыточен.) При $M = 3$ и 4 n-й символ строки обозначает следующее (эти коды основаны на теоремах 2–7):

h – хотя бы один из классов является объединением гиперграней коразмерности n, и, следовательно, по теореме 2 классификация самораспознается алгоритмом Кора-n', $\forall n' \geq n$;

s – ни один из классов не является объединением гиперграней коразмерности n, но классификация самораспознается алгоритмом Кора-n;

c – для данного n классификация имеет структуру "шахматной гипердоски", удовлетворяющую условиям теоремы 4, и, следовательно, не может быть распознана алгоритмом Кора-n' $\forall n' \leq n$, какой бы ни использовался материал обучения;

n – при $L_1 = C_1$, $L_2 = C_2$ классификация удовлетворяет условиям теоремы 5 и, следовательно, не самораспознается алгоритмом Кора-n', $\forall n' \geq n$, но она не удовлетворяет условиям теоремы 4;

r – классификация не самораспознаваема алгоритмом Кора-n, но распознается им при использовании некоторого материала обучения;

i – дополнительный код для $n = 2$: каждый класс C_i определяется неравенством (32) с приемлемыми коэффициентами, удовлетворяющими (49), но классификация не распознается алгоритмом Кора-2, какой бы ни использовался материал обучения;

x – классификация не принадлежит ни одному из типов, указанных выше: она не распознается алгоритмом Кора-n, какой бы ни использовался материал обучения, но для данного n не удовлетворяет условиям теоремы 5, и при $n = 2$ ни один из классов не определяется неравенством (32) с приемлемыми коэффициентами.

Пример 1. Если кодирование объектов сюръективно, то один из классов является объединением гиперграней коразмерности 1, если и только если второй из классов – гипергрань (такие классификации рассмотрены в теореме 1). Две классификации такого типа эквивалентны, если гиперграни, являющиеся вторыми из классов, имеют одинаковую коразмерность. Таким образом, распознаваемость ровно M изотипических классификаций такого типа кодируется строками "$\underbrace{\mathtt{h}\ldots\mathtt{h}}_{M-1}$" (первые строки в приведенных ниже таблицах).

Пример 2. Две классификации, удовлетворяющие условиям теоремы 4, эквивалентны, если и только если множества $\mathcal{P}$, определяющие их структуру "шахматной гипердоски", имеют одинаковое число элементов. В любой такой классификации каждый класс есть объединение гиперграней коразмерности $|\mathcal{P}|$. Следовательно, $\forall k,\ 0 < k < M$, распознаваемость единственной изотипической классификации кодируется строкой "$\underbrace{\mathtt{c}\ldots\mathtt{c}}_{k}\ \underbrace{\mathtt{h}\ldots\mathtt{h}}_{M-1-k}$" (последние M-1 строк в табл. 1-3).

ТАБЛИЦА 1. Типы распознаваемости классификаций при $M = 3$

Код типа распознаваемости	Число изотипических классификаций	Изотипические представители на рис. 1
"hh"	3	a,h,m
"sh"	2	b,e
"nh"	4	c,g,k,l
"ni"	1	d
"xi"	1	j
"ch"	1	f
"cc"	1	i

ТАБЛИЦА 2. Типы распознаваемости классификаций при $M = 4$

Код типа распознаваемости	Число изотипических классификаций
"hhh"	4
"shh"	6
"ssh"	1
"nhh"	25
"nsh"	17
"nrh"	78
"nrr"	2
"nnh"	13
"nnr"	11
"nnx"	2
"nih"	14
"nir"	4
"nxh"	6
"nxr"	10
"nxx"	1
"xhh"	4
"xsh"	1
"xrh"	11
"xir"	4
"xxh"	1
"xxr"	3
"chh"	1
"cch"	1
"ccc"	1

Теорема 8. При сюръективном кодировании объектов в X^3 существуют 13 различных классов эквивалентности классификаций (см. рис. 1), реализующих 7 типов распознаваемости (см. табл. 1). При сюръективном кодировании объектов в X^4 существуют 221 различных классов эквивалентности классификаций[2]*, реализующих 24 типа распознаваемости (см. табл. 2). В табл. 1, 2 указано число различных изотипических классификаций каждого типа.*

Для $M = 3, 4$ верны следующие утверждения:

i. Если некоторая классификация самораспознаваема алгоритмом Кора-n, то она самораспознается алгоритмом Кора-n с $A_i \geq \min(2^{M-n-1}, |C_i|)$, $\forall i = 1, 2$, и, следовательно, по теореме 3 она также самораспознаваема алгоритмом Кора-n', $\forall n' \geq n$.

ii. Если при некотором n и при $L_1 = C_1$, $L_2 = C_2$ некоторая классификация удовлетворяет условиям теоремы 5 (в частности, если она имеет структуру "шахматной гипердоски", удовлетворяющую при этом n условиям теоремы 4), то она не распознается алгоритмом Кора-n, какой бы ни использовался материал обучения; если при этом $n = 2$, то ни один из классов не определяется неравенством (32) с приемлемыми коэффициентами, описанными в теореме 7.

Доказательство теоремы было получено непосредственными вычислениями.▷

[2] В обоих случаях, $M = 3$ и $M = 4$, вырожденные классификации, у которых один из классов пуст, не рассматриваются.

ТАБЛИЦА 3. Типы самораспознаваемости классификаций при $M = 5$

Код типа самораспознаваемости	Число изотипических классификаций	Изменения типа самораспознаваемости при пополнении кодировки избыточными дескрипторами
"hhhh"	5	
"shhh"	11	
"sshh"	6	
"nhhh"	144	
"nshh"	307	"nshh"(265), "nShh"(6), "nxhh"(36)
"nssh"	53	"nssh"(42), "nSsh"(1), "nxsh"(9), "nxxh"(1)
"nShh"	42	"nShh"(24), "nxhh"(18)
"nSsh"	1	"nxsh"(1)
"nnhh"	1267	
"nnsh"	533	"nnsh"(218), "nnSh"(30), "nnxh"(285)
"nnSh"	4	"nnSh"(2), "nnxh"(2)
"nnnh"	147	
"nnnx"	459	
"nnxh"	61179	
"nnxs"	10	"nnxS"(10)
"nnxx"	12981	
"nxhh"	33608	
"nxsh"	8371	"nxsh"(2552), "nxSh"(481), "nxxh"(5338)
"nxSh"	102	"nxsh"(2), "nxSh"(66), "nxxh"(34)
"nxxh"	428487	"nxsh"(5), "nxSh"(1), "nxxh"(428481)
"nxxs"	50	"nxxs"(1), "nxxS"(49)
"nxxx"	65175	
"xhhh"	30	
"xshh"	61	"xshh"(57), "xShh"(1), "xxhh"(3)
"xssh"	3	
"xShh"	6	"xShh"(4), "xxhh"(2)
"xSsh"	1	"xxsh"(1)
"xxhh"	498	
"xxsh"	74	"xxsh"(36), "xxSh"(4), "xxxh"(34)
"xxSh"	7	"xxSh"(4), "xxxh"(3)
"xxxh"	1876	
"xxxs"	8	"xxxs"(1), "xxxS"(7)
"xxxx"	615	
"chhh"	1	
"cchh"	1	
"ccch"	1	
"cccc"	1	

Поскольку вычислительные ресурсы были ограничены, исследование распознаваемости классификаций при использовании произвольного материала обучения и генерация всех изотипических классификаций, удовлетворяющих неравенствам (32) с наборами приемлемых коэффициентов, при $M = 5$ не проводились. Соответственно, в табл. 3 использована модифицированная система кодов:

h – такой же, как и для $M = 3$ и $M = 4$;

s – ни один из классов не является объединением гиперграней коразмерности n, но классификация самораспознаваема алгоритмом Кора-n с параметрами A_1 и A_2, удовлетворяющими (8), и, следовательно, по теореме 3 самораспознается алгоритмом Кора-n', $\forall n' \geq n$;

S – классификация не принадлежит ни одному из двух типов, указанных выше, но она самораспознается алгоритмом Кора-n;

c, n – такие же, как и для $M = 3$ и $M = 4$;

x – классификация не принадлежит ни одному из типов, указанных выше: она не самораспознается алгоритмом Кора-n, но для данного n не удовлетворяет условиям теоремы 5.

Теорема 9. При сюръективном кодировании объектов в X^5 существуют 616125 различных классов эквивалентности классификаций[3]*. Они реализуют 37 типов распознаваемости, указанных в табл. 3 вместе с числом различных изотипических классификаций каждого типа.*

Для всех изотипических классификаций, кроме двух (с кодами типа распознаваемости "nSsh" *и* "Ssh"*, см. рис. 2) верно следующее: если классификация самораспознаваема алгоритмом Кора-n, то она самораспознается алгоритмом Кора-n с параметрами A_1 и A_2, удовлетворяющими (8), и/или хотя бы один из классов является объединением гиперграней коразмерности $n + 1$; соответственно, по теоремам 2 и 3 классификация также самораспознается алгоритмом Кора-n', $\forall n' \geq n$.*

Доказательство теоремы было получено непосредственными вычислениями.▷

Теоремы 1–3, 6, 8 и 9 позволяют высказать следующее предположение:

Гипотеза. Пусть кодирование объектов сюръективно. Если некоторая классификация самораспознаваема алгоритмом Кора-n, то она также самораспознаваема алгоритмом Кора-n', $\forall n' > n$.▷

Из приведенных таблиц следует, что число самораспознаваемых классификаций и классификаций, распознаваемых при использовании некоторого материала обучения, быстро растет с n; в частности, предположение о самораспознаваемости классификации алгоритмом Кора-n при малом n является сильным ограничением на ее структуру.

VI. О ВЛИЯНИИ МАЛОИНФОРМАТИВНЫХ ДЕСКРИПТОРОВ

Известно, что наличие малоинформативных дескрипторов может приводить к ухудшению качества результатов распознавания (для уменьшения их влияния можно выполнять исключение так называемых подчиненных признаков и использовать аналогичные процедуры (см. работу [2]), что, однако, приводит к зависимости работы алгоритма классификации от порядка дескрипторов и других субъективных факторов).

[3] Вырожденные классификации, у которых один из классов пуст, не рассматриваются.

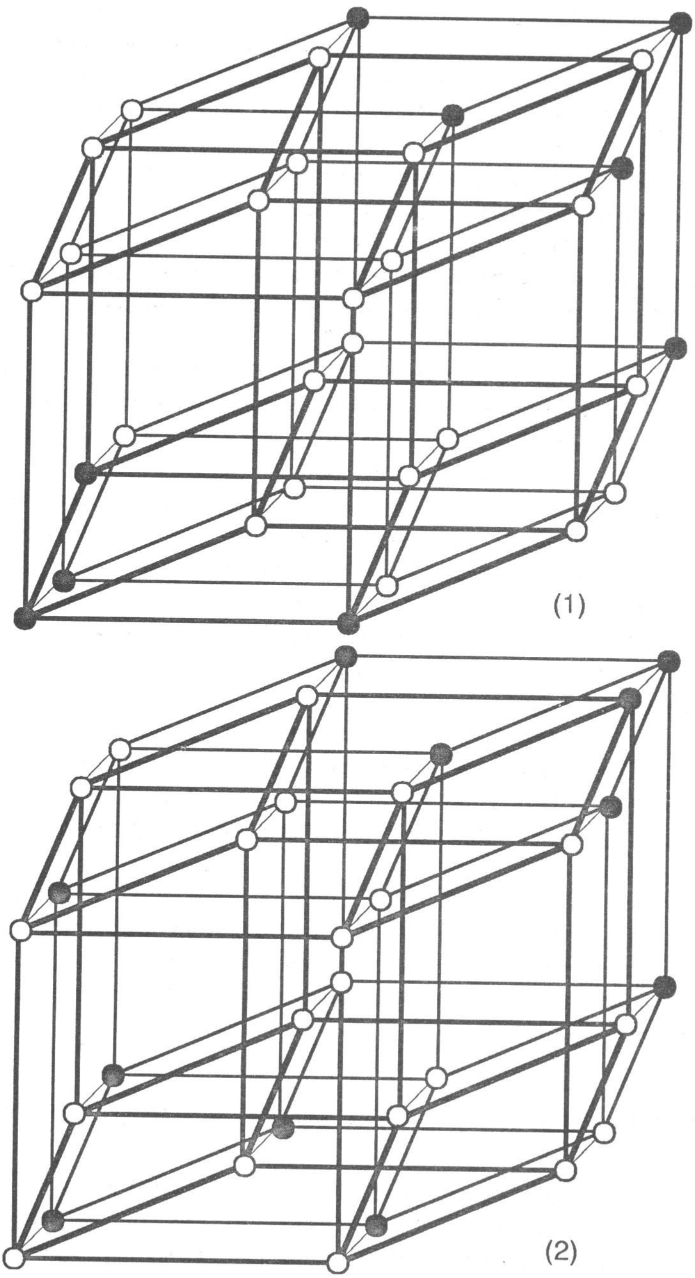

Рис. 2. Изотипические классификации $C_1 \sqcup C_2 = X^5$, самораспознаваемые алгоритмом Кора-n, $\forall n \geq 2$, в которых ни один из классов не является объединением гиперграней коразмерности 3 и которые не самораспознаются алгоритмом Кора-2 с параметрами A_1 и A_2, удовлетворяющими (8) при $n = 2$ (см. теорему 9).

Код типа распознаваемости "nSsh" (1), " Ssh" (2); C_1 – черные вершины, C_2 – белые вершины. Использована обычная проекция 4-мерного куба на плоскость для изображения гиперграней $x_5 = 0$ (жирные линии) и $x_5 = 1$ (тонкие линии)

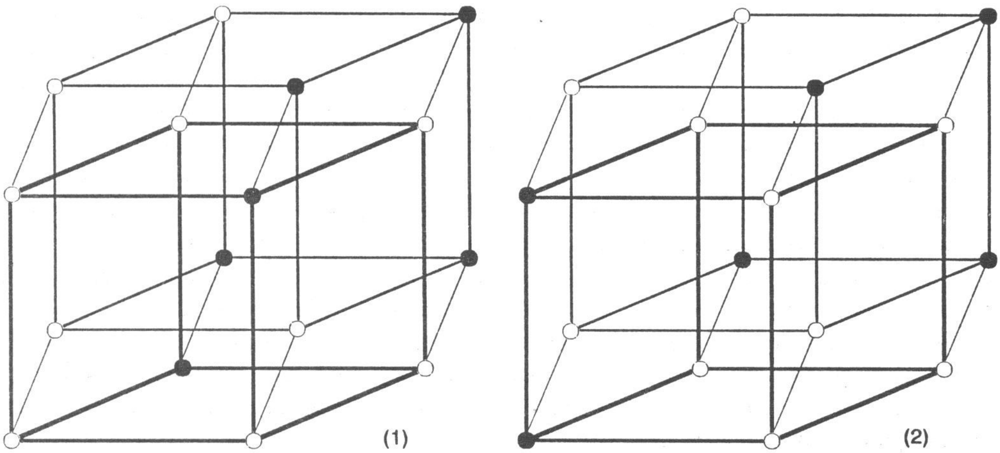

Рис. 3. Изотипические классификации $C_1 \sqcup C_2 = X^4$, не самораспознаваемые устойчиво при пополнении кодировки избыточным дескриптором:

(1) $\mathcal{I} = 1110010000100100$, (2) $\mathcal{I} = 1110010000000011$; C_1 – черные вершины, C_2 – белые вершины. Гиперграни $x_4 = 0$ и $x_4 = 1$ показаны соответственно жирными и тонкими линиями

Простые статистические критерии (см. [18]) могут быть использованы для исключения неинформативных дескрипторов (т.е. для изъятия некоторых вопросов из вопросника для кодирования объектов и соответствующего уменьшения размерности гиперкуба X^M). Тем не менее, при решении прикладных задач, как правило, векторы, представляющие собой объекты, содержат малоинформативные дескрипторы. В данном разделе рассмотрен численный эксперимент, иллюстрирующий их влияние. Пусть $p : X^{M+1} \to X^M$ – проектор: $p\mathbf{v} = (\mathbf{v}_1, ..., \mathbf{v}_M)$, где $\mathbf{v} = (\mathbf{v}_1, ..., \mathbf{v}_M, \mathbf{v}_{M+1})$.

Определение 9. Распознавание классификации $O = C_1 \sqcup C_2 \subset X^M$ с материалом обучения $L_1 \sqcup L_2$ алгоритмом Кора-n называется *устойчивым при пополнении кодировки избыточным дескриптором* $\mathbf{v}_{M+1}$, если классификация $p^{-1}C_1 \sqcup p^{-1}C_2 \subset X^{M+1}$ может быть распознана алгоритмом Кора-n с материалом обучения $p^{-1}L_1 \sqcup p^{-1}L_2$.

Согласно теоремам 1,2 и 6 самораспознаваемость классификации $C_1 \sqcup C_2$ заведомо устойчива при пополнении кодировки избыточным дескриптором, если хотя бы один из классов C_i является гипергранью, либо если C_i – объединение гиперграней коразмерности n, либо если кодирование объектов сюръективно и исходная классификация распознается алгоритмом Кора-1.

Вычисления показали, что при $M = 3$ самораспознаваемость классификаций при пополнении кодировки избыточным дескриптором всегда устойчива. При $M = 4$ только две изотипические классификации (с кодами типа распознаваемости "nsh", см. рис. 3) не самораспознаются устойчиво в указанном смысле: при пополнении кодировки избыточным дескриптором они перестают самораспознаваться алгоритмом Кора-2. (В частности, не возникают случаи самораспознавания типа S.) Однако при $M = 5$ изменения типа самораспознаваемости уже более существенны.

В тех случаях, когда тип изменяется, новые коды, усеченные до четырех символов, приведены в третьей колонке табл. 3 (число изотипических классификаций новых типов указано в табл. 3 в скобках). (В этом эксперименте не проверялось, как пополнение кодировки избыточным дескриптором влияет на распознаваемость классификаций при использовании произвольного материала обучения.)

Эти результаты показывают, что даже в задачах с умеренным числом дескрипторов присутствие малоинформативных дескрипторов может существенно изменить результаты распознавания алгоритмом Кора-n и их влияние может расти при увеличении размерности задачи. Таким образом, желательно удалять эти дескрипторы при подготовке исходных данных, если это возможно.

Тестирование устойчивости при пополнении кодировки избыточным дескриптором можно рекомендовать для проверки того, что результаты, полученные для некоторой задачи классифицирования, не являются ложными из-за присутствия малоинформативных дескрипторов.

Авторы благодарны Karl Wehrhahn за плодотворные обсуждения и рецензентам статьи В.Г.Кособокову, И.М.Ротвайн и А.А.Соловьеву за замечания, сделанные при чтении рукописи. С.Шварц был поддержан стипендиями Summer Vacation Scholarship Школы Математики и Статистики Сиднейского Университета, а также фонда Southern Cross Foundation (Австралия).

ЛИТЕРАТУРА

1. *Вайнцвайг М.Н.* Алгоритм обучения распознаванию образов "Кора"// Алгоритмы обучения распознаванию образов. М.: Сов. радио, 1973. С.8-12.
2. *Вапник В.Н., Червоненкис А.Я.* Теория распознавания образов. М.: Наука, 1974. 415 с.
3. *Браиловский В.Л.* Алгоритм распознавания объектов со многими параметрами и его приложения // Изв. АН СССР. Техническая кибернетика. 1964. N 2. С.30-39.
4. *Бонгард М.М.* Проблема узнавания. М.: Наука, 1967. 320 с.
5. *Caputo M., Kolari J.* Pattern recognition of the financial condition of banks//Accademia delle Scienze di Ferrara Atti. Vol.660, anno accademico 1660. 1989/90. P.1-13.
6. *Кунин П.Е., Таксар И.М., Бояджан В.А., Карп В.П., Марморштейн С.Я., Лихтенштейн Е.А., Агамова К.А., Гураль Е.И., Вольфсон Е.Б.* Использование электронных цифровых вычислительных машин для дифференциальной диагностики // Вычислительная техника в физиологии и медицине. М.: Наука, 1968.
7. *Гельфанд И.М., Гельштейн Г.Г., Губерман Ш.А., Ротвайн И.М., Файн Т.Л.* О предсказании давления в легочной артерии по данным электро- и фонокардиограмм // Кардиология. 1971а. N 5.
8. *Гельфанд И.М., Губерман Ш.А., Гельштейн Г.Г., Зорин А.Б., Силин В.А., Сухов В.Н., Ротвайн И.М., Файн Т.Л.* Распознавание степени легочной гипертонии при дефекте межжелудочковой перегородки с помощью ЭВМ // Кровообращение. АН Арм. ССР. 1976б. N 6.
9. *Куклин А.П.* Опыт изучения закономерностей локализации месторождений золота с помощью электронных цифровых машин//ДАН СССР. 1966. Т.171, N 2. С.429-430.
10. *Куклин А.П.* Содержательный анализ алгоритмов распознавания типа "Кора-3" и их улучшение при металлогенических исследованиях // Советская геология. 1972. N 10. С.55-66.

11. *Бонгард М.М., Вайнцвайг М.Н., Губерман Ш.А., Извекова М.Л., Смирнов М.С.* Использование обучающейся программы для выявления нефтеносных пластов // Геология и геофизика. 1966. N 6. С.96-105.

12. *Вайнцвайг М.Н., Губерман Ш.А., Чуринова И.М.* Обучение машины распознаванию нефтеносных пластов в залежах массивного типа // Нефтегазовая геология и геофизика. 1968. N 8.

13. *Губерман Ш.А.* Проблема распознавания в геологических исследованиях// Вычислительные методы в геофизике. М.: Радио и связь, 1981. С.58-65.

14. *Губерман Ш.А., Жидков М.П., Пиковский Ю.И., Ранцман Е.Я.* О некоторых критериях нефтегазоносности морфоструктурных узлов (Анды Южной Америки) // ДАН СССР. 1986. Т.291, N 6. С.1436-1440.

15. *Губерман Ш.А., Жидков М.П., Пиковский Ю.И., Ранцман Е.Я.* Распознавание нефтегазоносных морфоструктурных узлов Анд // Сквозные рудоконцентрирующие структуры. М.: Наука, 1989. С.78-85.

16. *Гельфанд И.М., Губерман Ш.А., Извекова М.Л., Кейлис-Борок В.И., Ранцман Е.Я.* О критериях высокой сейсмичности // ДАН СССР. 1972. Т.202, N 6. С.1317-1320.

17. *Гельфанд И.М., Губерман Ш.А., Извекова М.Л., Кейлис-Борок В.И., Ранцман Е.Я.* Распознавание мест возможного возникновения сильных землетрясений. I. Памир и Тянь-Шань // Вычислительные и статистические методы интерпретации сейсмических данных. М.: Наука, 1973. С.107-133. (Вычисл. сейсмология; Вып.6).

18. *Гельфанд И.М., Губерман Ш.А., Извекова М.Л., Кейлис-Борок В.И., Кнопов Л., Пресс Ф., Ранцман Е.Я., Ротвайн И.М., Садовский А.М.* Условия возникновения сильных землетрясений (Калифорния и некоторые другие регионы) // Исследование сейсмичности и моделей Земли. М.: Наука, 1976. С.3-91. (Вычисл. сейсмология; Вып.9).

19. *Gelfand I.M., Guberman Sh.A., Keilis-Borok V.I., Knopoff L., Press F., Ranzman E.Ya., Rotwain I.M., Sadovsky A.M.* Pattern recognition applied to earthquake epicenters in California // Phys. Earth and Planet. Inter. 1976б. Vol.11. P.227-283.

20. *Бхатия С.С., Рао М.Н., Четти Т.Р.К., Ранцман Е.Я., Горшков А.И., Филимонов М.Б., Шток Н.В.* Распознавание мест возможного возникновения сильных землетрясений. XIX. Гималаи, $M \geq 7.0$ // Теоретические проблемы геодинамики и сейсмологии. М.: Наука, 1994. С.280-287. (Вычисл. сейсмология; Вып.27).

21. *Губерман Ш.А., Ротвайн И.М.* Проверка результатов прогноза мест возникновения сильных землетрясений (1974-1984 гг.) // Изв. АН СССР. Физика Земли. 1986. N 12. С.72-74.

22. *Гвишиани А.Д., Гурвич В.А.* Динамические задачи распознавания образов. 1. Условия стабильности для прогноза мест сильных землетрясений // Математическое моделирование и интерпретация геофизических данных. М.: Наука, 1984. С.70-88. (Вычисл. сейсмология; Вып.16).