О сложности перестройки формальных нейронов

А. П. Соколов

Работа посвящена вопросам сложности обучения нейронов. В качестве математической модели нейронов рассматриваются пороговые функции алгебры логики. Рассматривается вопрос сложности взаимной перестройки (обучаемости) пар пороговых функций в самом сложном случае, в большинстве случаев и внутри классов пороговых функций, инвариантных относительно групп перестановок. В качестве средства задания пороговых функций выстапают целочисленные линейные формы. В качестве меры сложности процесса перестройки рассматривается число элементарных операций над весовыми коэффициентами линейной формы. Получен ряд верхних и нижних оценок сложности взаимной перестройки формальных нейронов для различных случаев. Построен алгоритм, осуществляющий перестройку заданного формального нейрона в желаемый. Получены оценки временной сложности данного алгоритма.

Ключевые слова: пороговые функции, сложность обучения нейросетей.

1. Введение

Пороговые функции алгебры логики являются простейшей математической моделью функционирования биологических нейронов. Данная модель была впервые предложена американскими учеными Маккаллоком и Питтсом в 1943 году [6]. В соответствии с данной моделью отдельная нервная клетка функционирует как логическое устройство, вычисляющее функцию взвешенного голосования входов.

Более строго, функция $f : \{0,1\}^n \to \{0,1\}$ называется пороговой, если существует линейное неравенство $x_1w_1 + \ldots + x_nw_n - \sigma \geqslant 0$, с действительными коэффициентами и порогом, выполненное на тех и только тех наборах $(x_1, \ldots, x_n)$, на которых функция f принимает значение 1. Коэффициенты $w_1, \ldots, w_n$ принято называть весовыми коэффициентами, а свободный член σ — порогом. Функцию $\sum\limits_{i=1}^{n} x_iw_i - \sigma$ в дальнейшем будем называть линейной формой. Если линейная форма не обращается в ноль ни на одном из наборов $(a_1, \ldots, a_n) \in \{0,1\}^n$, то говорят, что она задает пороговую функцию строгим образом.

Несмотря на свою простоту, пороговые функции обладают значительными вычислительными возможностями и, при этом, весьма просты с точки зрения технической реализации.

При задании пороговых функций могут рассматриваться линейные формы с целочисленными коэффициентами и свободным членом. В этой связи можно ввести понятие размаха линейной формы

$$\max(|w_1|, \ldots, |w_n|, |\sigma|).$$

Также можно ввести понятие веса линейной формы

$$\sum_{i=1}^{n} |w_i| + |\sigma|.$$

Среди всех линейных форм, задающих пороговую функцию f строгим образом, очевидно, существует линейная форма минимального размаха. Ее размах назовем размахом функции f и обозначим $L(f)$. Аналогичным образом вводится понятие веса пороговой функции. В 1961 году Мурога [22] показал, что размах произвольной пороговой функции от n переменных, обозначаемый $L(n)$, можно оценить так

$$L(n) \leqslant 2^{-n} (n+1)^{\frac{n+1}{2}}.$$

Так как существует по крайней мере $2^{n^2+O(n\log n)}$ различных пороговых функций [4], то существуют пороговые функции, обладающие размахом не менее $2^{n+o(n)}$. С другой стороны, число пороговых функций не превосходит 2^{n^2}, поэтому более точных нижних оценок на величину размаха таким образом получить не удается.

В 1994 году Дж. Хастад [16] доказал существование пороговой функции, обладающей размахом $2^{\frac{1}{4}n\log n+o(n\log n)}$, что дает порядок логарифма при стремлении числа переменных к бесконечности. В 2002 году Алон и Ву [12] нашли асимптотику логарифма размаха при стремлении числа переменных к бесконечности.

Понятия размаха и веса оказываются весьма полезным при изучении одного из важнейших свойств пороговых функций — способности к обучению. Обучаемость пороговых функций — это их способность в результате многократной коррекции весовых коэффициентов и порога изменять свое функционирование на желаемое. Данное свойство было впервые отмечено в 1957 году Ф. Розенблаттом, разработавшим распознающую систему «персептрон» и предложившим алгоритм ее обучения [7].

После работ Розенблатта было предложено множество различных архитектур нейросетей и алгоритмов их обучения, детальный обзор которых приведен в [9]. Несмотря на то, что многие из данных алгоритмов обучения с успехом применяются в инженерной практике, для большинства из них нет математически строгих оценок на время работы. Более того, у такого наиболее распространенного алгоритма обучения как «метод обратного распространения ошибки» [27] было даже обнаружено свойство неустойчивости по отношению к начальным значениям весовых коэффициентов [25].

Из теоретических работ по сложности обучения схем из пороговых функциональных элементов отметим работы С. Джадда [17, 18] и [19]. В данных работах было показано, что общая задача обучения нейронных схем является NP-полной. Данная задача рассматривалась в различных постановках, в которых в качестве нейронов выступали как пороговые функции, так и функции, принимающие действительные значения на отрезке $[0, 1]$.

1.1. Краткая постановка задачи

В данной работе, также как и в [22], рассматриваются пороговые функции, задаваемые линейными формами с целочисленными коэффициентами и свободным членом.

Исследуется сложность преобразования одной пороговой функции, заданной линейной формой, к другой, путем последовательного изменения коэффициентов линейной формы. Данный процесс может интерпретироваться как процесс обучения нейрона с пороговой функцией активации.

В процессе обучения некая управляющая система получает в качестве исходных данных описание исходной и желаемой пороговых функций. Управляющая система осуществляет пошаговое изменение коэффициентов исходной линейной формы для получения линейной формы, задающей желаемую пороговую функцию. За один шаг управляющая система осуществляет изменение некоторого весового коэффициента или свободного члена линейной формы. При этом, степень данного изменения ограничена. В самом простейшем случае величина изменения ограничена единицей. Минимально достаточное количество таких шагов определяет сложность процесса обучения.

Элементарные операции перестройки линейных форм, степень воздействия которых ограничена, представляют интерес в связи с тем, что они хорошо согласуются с существующими в биологии моделями синаптической пластичности [29, 13]. Эти модели описывают механизмы изменения во времени силы синапсов биологических нейронов. Механизм синаптической пластичности считается основным способом, с помощью которого реализуется феномен памяти и обучения в живых организмах [15].

Ограничение множества допустимых операций при перестройке линейных форм изменениями весовых коэффициентов и порога на единицу является несколько неестественным с точки зрения процессов перестройки, происходящих в биологических нейронах. Однако, без рассмотрения данного случая, переход к более общей ситации невозможен.

Для характеризации сложности обучения нейронов в работе рассматривается следующая метрическая характеристика. Для каждой пары пороговых функций f' и f'' вводится функция близости $\rho\,(f', f'')$, характеризующая минимально достаточное число единичных изменений некоторой линейной формы, задающей функцию f', для получения линейной формы, задающей функцию f''. Таким образом, оценивается минимально возможная сложность перестройки исходной пороговой функции в желаемую вне зависимости от «стартовой» линейной формы.

Поведение функции близости $\rho\,(f', f'')$ изучается в трех постановках.

В первой постановке сложность взаимной перестройки изучается в самом сложном случае, то есть на всем классе пороговых функций от n переменных. Для характеризации сложности обучения в данном случае вводится шенноновская функция $\rho\,(n)$. Она говорит о том, сколько минимально достаточно выполнить единичных модификаций исходной линейной формы от n переменных для задания желаемой пороговой функции. Данная величина характеризует максимально возможное время обучения, с которым может столкнуться управляющая система, обучающая нейроны с n входами.

Помимо самого сложного случая, который может возникнуть при обучении нейронов, несомненный интерес представляет задача оценки сложности обучения нейронов для почти всех случаев. Рассматривается вопрос, как ведет себя в большинстве случаев функция близости $\rho\,(l, f)$, где l — линейная форма, а f — пороговая функция от n переменных. При этом полагается, что размах линейной формы l не превосходит величину $L\,(n)$.

Наконец, в третьей постановке сложность взаимной перестройки исследуется для классов пороговых функций, инвариантных относительно групп перестановок переменных. Данный случай представляет интерес, так как он выявляет особенности процесса обучения нейронов, обладающих симметрией входов. Наличие симметричных входов у нейрона может, с одной стороны, выступать в качестве способа повышения надежности его функционирования, а с другой стороны, обуславливать особенности функционирования нейрона при работе с симметричными входными векторами.

Далее исследуется задача характеризации сложности перестройки линейных форм для более мощных множеств допустимых операций. В работе рассмотрены две модификации понятия близости между пороговыми функциями. Данные модификации задаются множествами допустимых элементарных операций перестройки линейных форм.

В первом случае к операции единичного изменения коэффициентов добавляется операция инвертирования (умножения на -1) коэффициента или порога. Во втором случае вводится понятие операции ограниченной мощности. Данные операции позволяют за одно применение увеличить или уменьшить значение весового коэффициента или порога линейной формы в произвольное число раз, не превышающее заданный предел. К множеству допустимых операций добавляются все операции ограниченной мощности. Данная модификация является обобщением двух предыдущих случаев и в большей степени согласуется с существующими в биологии моделями синаптической пластичности [29, 13].

Для каждой модификации понятия близости сложность обучения исследуется в трех постановках: в самом сложном случае, для почти всех случаев и для классов пороговых функций, инвариантных относительно групп перестановок.

Помимо оценок на количество элементарных операций, необходимых для перехода от одной линейной формы к другой, рассматривается задача явного построения алгоритма для осуществления подобного перехода.

Введенные ранее характеристики веса и размаха пороговой функции определяют размер памяти, необходимый для хранения коэффициентов линейной формы. В связи с этим особый интерес представляет задача явного построения линейных форм минимального веса и размаха, обладающих экспоненциальным весом и размахом. Эта задача также исследуется в работе.

1.2. Основные результаты

Основные результаты работы заключаются в следующем:
1. Введены в рассмотрение классы $T_{m,k}$ пороговых функций от n переменных, где параметр m означает максимальный размер класса симметричных переменных, а параметр k — число независимых классов симметрии. Для данных классов получены верхняя и нижняя оценки сложности взаимной перестройки пороговых функций. Для этого введена величина $\rho\,(m,k)$, характеризующая близость между наиболее удаленными пороговыми функциями, не более чем с m взаимно симметричными переменными, и числом классов симметрии — k.

Теорема 1. *Для достаточно больших k выполнено*

$$m \cdot 2^{\frac{1}{2}k\log k + o(k\log k)} \leqslant \rho\,(m,k) \leqslant m^{k+2} \cdot 2^{\frac{1}{2}k\log k + o(k\log k)}.$$

При определенных соотношениях между параметрами симметрии m и k из теоремы 1 следует асимптотика логарифма функции близости $\rho\,(m,k)$.

Следствие. *Если $k \to \infty$ и $\log(m) = o(\log(k))$, то*

$$\log\rho\,(m,k) \sim \frac{1}{2}k\log k.$$

В частном случае, когда $m = 1$ и $k = n$, функция близости $\rho\,(m,k)$ совпадает с величиной $\rho\,(n)$. В этом случае из теоремы 1 вытекает асимптотика логарифма шенноновской функции сложности взаимной перестройки пороговых функций — $\rho\,(n)$.

Следствие. *Если $n \to \infty$, то $\log_2\rho\,(n) \sim \frac{1}{2}n\log_2 n$.*

С содержательной точки зрения данный результат означает, что в случае использования лишь операции единичного изменения весовых коэффициентов при обучении нейронов с n входами управляющей системе может потребоваться вплоть до $2^{\frac{1}{2}n\log n + o(n\log n)}$ элементарных шагов.

Отметим, что последнее следствие может быть получено при помощи известных оценок на величину размаха, полученных в работах

Муроги [22] и Н. Алона и В. Ву [12]. Теорема 1 обобщает данный результат для классов пороговый функций, инвариантных относительно групп перестановок, а также расширяет множество случаев, для которых верна асимптотика логарифма функции $\rho(m,k)$, указанная в первом следствии.
2. Показано, что для почти всех случаев перестройки «стартовой» линейной формы в желаемую пороговую функцию от n переменных, сложность перестройки с ростом n растет экспоненциально.

Теорема 2. *Если $n \to \infty$, то для почти всех пар (l,f), где f — пороговая фукнция, а l — линейная форма от n переменных такая, что $L(l) \leqslant L(n)$, выполнено $2^{n+o(n)} \leqslant \rho(l,f) \leqslant 2^{\frac{1}{2}n\log n+o(n\log n)}$.*

Таким образом, при достаточно большом числе входов n сложность обучения нейронов почти всегда велика. Отметим также, что данный результат имеет место в случае использования только лишь операции единичного изменения весовых коэффициентов. Далее будет показано, что при некотором расширении перечня допустимых операций над весовыми коэффициентами, сложность обучения может быть существенно снижена.
3. Получены верхние и нижние оценки сложности перестройки для двух модификаций понятия близости между пороговыми функциями.

В первом случае помимо операции единичного изменения коэффициентов разрешается использование операция инвертирования (умножения на -1) коэффициента или порога.

Во втором случае вводится понятие операции ограниченной мощности. Данные операции позволяют за одно применение увеличить или уменьшить значение весового коэффициента или порога линейной формы в произвольное число раз, не превышающее заданную величину — a. При этом, полагается, что a — натуральное число, не меньшее 2. К множеству допустимых операций добавляются все операции, мощность которых не превышает a.

Как и ранее, под близостью между пороговыми функциями f' и f'' понимается минимально достаточное число операций над некоторой линейной формой, задающей функцию f', для получения линейной формы, задающей функцию f''. Функцию близости, получаемую в первом случае обозначим ρ_{II}, а во втором — ρ_a.

Получены следующие теоремы, характеризующие поведение функций ρ_{II} и ρ_a в самом сложном случае, для почти всех случаев и для классов пороговых функций, инвариантных относительно групп перестановок.

Теорема 3. *Если $n \to \infty$, то*

$$\log \rho_{II}(n) \sim \frac{1}{2} n \log n;$$
$$\rho_a(n) \asymp n^2 \log n.$$

Теорема 4. *Если $n \to \infty$, то для почти всех пар (l, f), где f — пороговая фукнция, а l — линейная форма от n переменных такая, что $L(l) \leqslant L(n)$, выполнено*

$$2^{n+o(n)} \leqslant \rho_{II}(l, f) \leqslant 2^{\frac{1}{2} n \log n + o(n \log n)};$$
$$n \lesssim \rho_a(l, f) \lesssim n^2 \log n.$$

Теорема 5. *а) Для достаточно больших k выполнено*

$$m \cdot 2^{\frac{1}{4} k \log k + o(k \log k)} \leqslant \rho_{II}(m, k) \leqslant m^{k+2} \cdot 2^{\frac{1}{2} k \log k + o(k \log k)}.$$

б) Если $k \to \infty$ и $\log(m) = o(\log(k))$, то

$$mk \log k \lesssim \rho_a(m, k) \lesssim mk^2 \log k.$$

Утверждения теорем 1,. . . ,5 дают нижние и верхние оценки на количество элементарных операций перестройки линейных форм для перехода от одной пороговой функции к другой. Оценки даны для различных множеств допустимых операций и для различных случаев: для самого сложного случая, для почти всех случаев и для классов пороговых функций, инвариантных относительно групп перестановок. Данные оценки характеризуют сложность, с которой столкнется всякий алгоритм, решающий задачу перестройки, с использованием рассмотренных множеств допустимых операций. При этом, данные результаты не дают описания непосредственного самого алгоритма, осуществляющего перестройку одной линейной формы в линейную форму, задающую желаемую пороговую функцию. Этот вопрос рассматривается в главе 6.

В главе 6 вводится понятие множества допустимых линейных форм, весовые коэффициенты которых при стремлении числа переменных n к бесконечности по модулю ограничены величиной $O(n)\, L(n)$. Получены оценки сложности нахождения допустимой линейной формы для пороговой функции заданной в виде множеств единиц и нулей, то есть наборов, на которых функция принимает значения 1 и 0, соответственно. Показано, что в этом случае задача нахождения допустимой линейной формы имеет полиномиальную сложность от числа единиц и нулей желаемой функции. Получены оценки сложности взаимной перестройки допустимых линейных форм для различных функций близости в самом сложном случае. Описан алгоритм перестройки линейных форм, который имеет полиномиальную сложность как в терминах числа вспомогательных арифметических операций, так и в смысле числа элементарных шагов перестройки.

1.3. Структура работы

Изложение построено следующим образом.

В **главе 2** даются основные определения и сведения вспомогательного характера: вводятся понятия линейной формы и пороговой функции, определяется задание пороговой функции линейной формой. Вводится понятие близости между пороговыми функциями $\rho(f', f'')$.

Вводится шенноновская функция $\rho(n)$, характеризующая близость между наиболее удаленными пороговыми функциями от n переменных. Ставится задача исследования сложности обучения в самом сложном случае.

Для более общей характеризации типичного близости между пороговыми функциями ставится задача поиска верхних и нижних оценок на функцию близости $\rho(f', f'')$, имеющих место для «почти всех» пар пороговых функций f', f''.

Вводится понятие инвариантности пороговой функции относительно перестановки π. Ставится задача описания классов пороговых функций, инвариантных относительно групп перестановок. Также ставится задача поиска верхних и нижних оценок на функцию близости $\rho(f', f'')$, имеющих место внутри данных классов.

Далее, рассмотрены две модификации введенного понятия близости между пороговыми функциями. Для модификацированных функций близости ставится задача оценки сложности обучения в самом сложном случае, для почти всех случаев и для классов пороговых функций, инвариантных относительно групп перестановок.

Заключение главы посвящено общей характеризации класса пороговых функций с точки зрения задачи обучения. Как отмечалось ранее, при задании пороговых функций могут рассматриваться линейные формы, как с действительными, так и с целочисленными коэффициентами. Отмечается, что выразительные возможности данных способов эквивалентны, то есть любая пороговая функция f может быть задана некоторой линейной формой с целочисленными коэффициентами.

Далее, вводится понятие сигнатуры пороговой функции f как набор знаков коэффициентов некоторой линейной формы, задающей f. Оказывается, что если f существенно зависит от всех своих переменных, то сигнатура f определяется однозначным образом. Более того, отношение равенства сигнатур разбивает множество существенных пороговых функций на 2^n взаимно непересекающихся равномощных класса, одним из которых является класс монотонных пороговых функций. Таким образом, структура класса монотонных пороговых функций полностью определяет структуру класса всех пороговых функций.

Для любой пороговой функции существует бесконечное множество задающих ее линейных форм. Линейные формы назовем эквивалентными, если они задают одну и ту же пороговую функцию. Множество всех существенных линейных форм с целочисленными коэффициентами и свободным членом, задающих пороговую функцию

f, обозначим $U(f)$. Легко видеть, что множество $U(f)$ замкнуто относительно операции сложения линейных форм. В данной главе показано, что всякое множество $U(f)$ содержит единственный базис относительно операции сложения линейных форм, который является счетным и разрешимым, а также описан алгоритм, который последовательно строит базис множества $U(f)$ для заданной пороговой функции f.

В **главе 3** рассматриваются классы пороговых функций, инвариантных относительно групп перестановок. Получено полное описание данных классов в терминах групп перестановок, сохраняющих разбиение на множестве переменных.

Далее рассмотрен вопрос о сложности взаимной перестройки внутри классов пороговых функций, инвариантных относительно групп перестановок. Получены верхняя и нижняя оценки сложности перестройки внутри данных классов, что составляет теорему 1.

В заключении главы приводятся доказательства следствий 1 и 2 теоремы 1.

Глава 4 посвящена рассмотрению задачи обучения в большинстве случаев.

В данной главе показано, что для почти всех пар пороговых функций от n переменных сложность перестройки «стартовой» линейной формы в линейную форму, задающую желаемую пороговую функцию, с ростом n растет экспоненциально. Данное утверждение составляет теорему 2.

В **главе 5** обобщается понятие близости между пороговыми функциями. Вводятся различные множества операций над коэффициентами и порогом линейных форм. Для каждого множества операций вводится соответствующая функция близости, характеризующая минимально достаточное число операций над некоторой линейной формой, задающей исходную пороговую функцию, для получения линейной формы, задающей желаемую функцию.

Показано, что добавление операции инвертирования весовых коэффициентов и порога не дает ощутимого снижения сложности обучения как в самом сложном случае, так и для почти всех случаев.

Добавление же операций ограниченной мощности, напротив, снижает сложность взаимной перестройки пороговых функций с экспоненциальной до полиномиальной. Данные результаты составляют утверждения теорем 3, 4, 5.

В **главе 6** рассмотрен вопрос алгоритмической сложности задачи перехода от «стартовой» линейной формы к «целевой», задающей желаемую пороговую функцию. Для этого вводится понятие множества допустимых линейных форм, весовые коэффициенты которых при стремлении числа переменных n к бесконечности по модулю ограничены величиной $O(n)\,L(n)$. Получены оценки сложности нахождения допустимой линейной формы для пороговой функции заданной виде множеств единиц и нулей, то есть наборов, на которых функция принимает значения 1 и 0, соответственно. Показано, что в этом случае задача нахождения допустимой линейной формы имеет полиномиальную сложность от числа единиц и нулей желаемой функции. Получены оценки сложности взаимной перестройки допустимых линейных форм для различных функций близости в худшем случае. Описан алгоритм перестройки линейных форм, который имеет полиномиальную сложность как в терминах числа вспомогательных арифметических операций, так и в смысле числа элементарных шагов перестройки.

Приводятся конструктивное построение такого класса пороговых фукнций, что для почти всех пар функций из данного класса близость между ними зависит экспоненциально от числа переменных, при этом мощность данного класса в некотором смысле сопоставима с мощностью класса всех пороговых функций. В противоположность данному классу далее приводится конструктивное построение такого класса пороговых фукнций, что близость между функциями из данного класса ограничена произвольной заранее заданной величиной, лежащей в диапазоне от $3(n+1)$ до $3 \cdot 2^{\frac{1}{2}n}$, где n — число переменных. При этом, мощность данного класса экспоненциально зависит от n.

Далее, построено множество операций над линейными формами, которые сохраняют свойство минимальности размаха и веса, что составляет утверждение теоремы 16. Путем применения данных операций к элементарным линейным формам можно построить бесконечное многообразие линейных форм минимального размаха и веса. Одним из подмножеств данного семейства является последовательность линейных форм, весовые коэффициенты которых представляют собой числа Фибоначчи, что составляет известный пример, полученный в работах Мурога [23] и Парберри [26]. Отметим также, что из утверждения теоремы 16 следует одновременно минимальность веса и размаха элементов многообразия, полученного в результате применения данных операций.

Автор выражает глубокую благодарность своему научному руководителю академику В. Б. Кудрявцеву за постановку задачи и постоянное внимание к ней, Э. Э. Гасанову, А. А. Ирматову, А. А. Часовских и А. В. Галатенко за ряд ценных замечаний, а также коллективу кафедры математической теории интеллектуальных систем механико-математического факультета Московского государственного университета имени М. В. Ломоносова за всесторонние помощь и поддержку.

2. Постановки задач, вспомогательные утверждения и конструктивная характеризация пороговых функций

2.1. Постановки задач

Пусть $U = \{u_1, u_2, \ldots\}$ — счетный алфавит переменных. Каждое из переменных u_i может принимать значения из множества $E_2 = \{0, 1\}$. В дальнейшем во избежание употребления сложных индексов мы будем использовать для обозначения букв алфавита U метасимволы x_i с индексами или без них.

Введем определения линейной формы и пороговой функции.

Линейной формой назовем функцию вида

$$l_{\vec{w},\sigma}(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i w_i - \sigma,$$

где w_i и σ суть действительные числа при $i = 1, \ldots, n$.

Вектор $\vec{w} = (w_1, \ldots, w_n)$ называют вектором весовых коэффициентов, а σ — порогом. Для простоты иногда будем обозначать линейную форму $l_{\vec{w},\sigma}(x_1, \ldots, x_n)$ соответствующим набором весовых коэффициентов и порога — $(w_1, \ldots, w_n, \sigma)$.

Функция $f(x_1, \ldots, x_n) : E_2^n \to E_2$, называется *пороговой*, если существует линейная форма $l_{\vec{w},\sigma}$ такая, что

$$f(x_1, \ldots, x_n) = \begin{cases} 1, & \text{если } \sum\limits_{i=1}^{n} x_i w_i - \sigma \geqslant 0; \\ 0, & \text{иначе.} \end{cases}$$

В этом случае говорим, что *линейная форма $l_{\vec{w},\sigma}$ задает пороговую функцию $f(x_1, \ldots, x_n)$*, и записывается это так

$$l_{\vec{w},\sigma} \to f(x_1, \ldots, x_n),$$

или просто $f_{\vec{w},\sigma}$.

Будем говорить, что *линейная форма $l_{\vec{w},\sigma}$ задает пороговую функцию $f(x_1, \ldots, x_n)$ строгим образом*, если $l_{\vec{w},\sigma}$ задает f и $l_{\vec{w},\sigma}(a_1, \ldots, a_n) \neq 0$ ни на одном из наборов $(a_1, \ldots, a_n) \in E_2^n$. Задание пороговой функции строгим образом записывается так

$$l_{\vec{w},\sigma} \Rightarrow f(x_1, \ldots, x_n).$$

Множество всех пороговых функций обозначим T, а множество пороговых функций от n переменных — T^n.

Можно показать, что всякая пороговая функция может быть задана линейной формой с целочисленными коэффиентами и порогом. В связи с этим далее, если это не будет оговариваться отдельно, будем рассматривать линейные формы только с целочисленными коэффициентами и порогом.

Введем понятие близости между линейными формами и пороговыми функциями.

Пусть $l_{\vec{w}',\sigma'}$ и $l_{\vec{w}'',\sigma''}$ — линейные формы от n переменных. Расстоянием между линейными формами $l_{\vec{w}',\sigma'}$ и $l_{\vec{w}'',\sigma''}$ назовем следующую величину:

$$\rho\left(l_{\vec{w}',\sigma'}; l_{\vec{w}'',\sigma''}\right) = \left|\sigma' - \sigma''\right| + \sum_{i=1}^{n} \left|w_i' - w_i''\right|.$$

Эту величину интерпретируем, как необходимость сделать ρ последовательных единичных изменений компонент одной линейной формы, чтобы получить другую.

Пусть заданы пороговые функции $f'(x_1,\ldots,x_n)$ и $f''(x_1,\ldots,x_n)$. Близостью между ними назовем величину

$$\rho\left(f', f''\right) = \min_{\substack{l' \to f' \\ l'' \to f''}} \rho\left(l', l''\right)$$

Здесь минимум берется по всем линейным формам $l_{\vec{w}',\sigma'}$ и $l_{\vec{w}'',\sigma''}$ с целыми коэффициентами, задающим функции f' и f'', соответственно.

Аналогичным образом вводится функция близости между линейными формами и пороговыми функциями. Пусть l — линейная форма, а f — пороговая функция от n переменных. Близостью между ними назовем величину

$$\rho\left(l, f\right) = \min_{l' \to f} \rho\left(l, l'\right).$$

Здесь минимум берется по всем линейным формам l', задающим функцию f.

Определим величину $\rho\left(n\right)$ следующим образом

$$\rho\left(n\right) = \max_{f', f'' \in T^n} \rho\left(f', f''\right).$$

Данная величина характеризует близость между наиболее удаленными пороговыми функциями от n переменных.

Задача исследования сложности обучения в самом сложном случае состоит в отыскании функции $\rho\left(n\right)$.

Рассмотрим следующие меры сложности линейных форм

$$L\left(l_{\vec{w},\sigma}\right) = \max\left(\left|w_1\right|, \ldots, \left|w_n\right|, \left|\sigma\right|\right),$$

$$\mu\left(l_{\vec{w},\sigma}\right) = \sum_{i=1}^{n} \left|w_i\right| + \left|\sigma\right|.$$

Величину $L\left(l_{\vec{w},\sigma}\right)$ назовем размахом линейной формы $l_{\vec{w},\sigma}$, а $\mu\left(l_{\vec{w},\sigma}\right)$ — ее весом.

Рассмотрим соответствующие им меры сложности задания пороговой функции f линейными формами с целочисленными коэффициентами

$$L(f) = \min_{l_{\vec{w},\sigma} \Rightarrow f} L\left(l_{\vec{w},\sigma}\right),$$

$$\mu(f) = \min_{l_{\vec{w},\sigma} \Rightarrow f} \mu\left(l_{\vec{w},\sigma}\right).$$

Здесь минимум берется по всем целочисленным линейным формам, задающим функцию f строгим образом.

Величину $L(f)$ назовем размахом функции f, а $\mu(f)$ — ее весом.

Линейную форму $l_{\vec{w},\sigma}$ назовем формой минимального размаха для пороговой функции f, если $l_{\vec{w},\sigma}$ задает f строгим образом и ее размах минимален среди всех линейных форм, задающих f строгим образом. Аналогичным образом вводится понятие линейной формы минимального веса для пороговой функции f.

В связи с результатом работы [16] возникает задача явного построения пороговой функции, обладающей экспоненциальным весом и размахом.

Далее, возникает следующий вопрос: насколько типична такая ситуация, когда близость между двумя пороговыми функциями зависит экспоненциально от числа переменных n, или же, наоборот, не превышает заранее заданной величины. В частности, ставится задача конструктивного описания классов пороговых функций, которые были бы максимально удалены или, наоборот, близки друг к другу в терминах близости ρ.

Для более общей характеризации типичной близости между пороговыми функциями введем следующее определение. Рассмотрим две последовательности множеств: $A_1, A_2, \ldots$ и $A'_1, A'_2, \ldots$, где $A_i \subseteq A'_i$ для всех i. Говорят, что «почти все» элементы A_i совпадают с элементами A'_i, если

$$\lim_{i \to \infty} \frac{|A_i|}{|A'_i|} = 1.$$

Ставится задача поиска верхних и нижних оценок на функцию близости $\rho\,(l,f)$, где l — линейная форма, а f — пороговая функция от n переменных, для почти всех случаев. При этом полагается, что размах линейной формы l не превосходит величину $L\,(n)$.

Еще одним важным свойством, на основании которого можно «расслоить» класс пороговых функций по сложности взаимной перестройки, является свойство симметрии переменных.

Рассмотрим перестановку $\pi \in S_n$, где S_n — группа перестановок над множеством $\Omega_n = \{1,\ldots,n\}$. Будем называть пороговую функцию f *инвариантной относительно перестановки* π, если

$$f\,(x_1,\ldots,x_n) = f\left(x_{\pi(1)},\ldots,x_{\pi(n)}\right).$$

Легко видеть, что если пороговая функция f инвариантна относительно перестановок π_1 и π_2, то она также инвариантна относительно перестановки $\pi_1 \cdot \pi_2$, получающейся в результате последовательного применения перестановок π_2 и π_1. Отсюда следует, что для каждой подгруппы G группы перестановок S_n существует множество пороговых функций T_G^n, инвариантных относительно перестановок из G. Возникает естественный вопрос: сколько существует различных классов T_G^n и как они описываются? Также ставится задача поиска верхних и нижних оценок на функцию близости $\rho\,(f',f'')$, имеющих место для всех пар f', f'' из T_G^n.

Рассмотрим некоторое множество операций над целыми числами — M. Говорим, что линейная форма $l_{\vec{w}',\sigma'}$ M-переводима в линейную форму $l_{\vec{w}'',\sigma''}$, если она может быть переведена в линейную форму $l_{\vec{w}'',\sigma''}$ при помощи одного применения операции из M к одному из весовых коэффициентов вектора $\vec{w}'$, или порогу σ'.

Пусть $l_{\vec{w}',\sigma'}$ и $l_{\vec{w}'',\sigma''}$ — линейные формы от n переменных. Близостью $\rho_M(l_{\vec{w}',\sigma'}, l_{\vec{w}'',\sigma''})$, между данными линейными формами назовем длину кратчайшей последовательности $l_1,\ldots,l_r$, где $l_1 = l_{\vec{w}',\sigma'}$, $l_r = l_{\vec{w}'',\sigma''}$ и для всякого $j = 1,\ldots,r-1$, линейная форма l_j M-переводима в линейную форму l_{j+1}. Полагаем, что длина последовательности $l_1,\ldots,l_r$ равна $r-1$.

Пусть f', f'' — пороговые функции от n переменных. Близостью $\rho_M(f', f'')$, между ними назовем величину

$$\rho_M\left(f', f''\right) = \min_{\substack{l' \to f' \\ l'' \to f''}} \rho_i\left(l', l''\right),$$

где минимум берется по всем линейным формам $l_{\vec{w}', \sigma'}$ и $l_{\vec{w}'', \sigma''}$, задающим функции f' и f'', соответственно.

Легко видеть, что если в качестве множества операций M принять увеличение или уменьшение весового коэффициента на единицу, то полученная функция близости $\rho_M\left(f', f''\right)$ будет в точности равна введенной ранее функции $\rho\left(f', f''\right)$.

Ставится следующая задача: построить нижние и верхние оценки на функцию близости ρ_M, если добавить в множество M такие операции, как умножение на -1, множество операций, мощность которых не превосходит a, где a — натуральное число, не меньшее 2.

Рассмотрим вопрос алгоритмической сложности задачи перехода от «стартовой» линейной формы к «целевой», задающей желаемую пороговую функцию.

Задача перестройки пороговых функций ставится следующим образом. Дана некая «стартовая» линейная форма. Задана «целевая» пороговая функция. Требуется перестроить «стартовую» линейную форму так, чтобы полученная линейная форма задавала «целевую» пороговую функцию. При этом, на каждом шаге перестройки разрешено выполнять лишь одну операцию из множества допустимых. В качестве таких множеств выступают M_I, M_{II} и M_a. Естественно, требуется выполнить как можно меньшее число таких операций. Очевидно, что введенные функции близости дают нижние оценки на число элементарных шагов в процессе перестройки.

Так как ближайшая линейная форма, задающая «целевую» пороговую функцию не задана, то для определения направления перестройки допускается выполнение неких вспомогательных вычислений. Сложность данных вычислений оценивается отдельно. При этом, оценивается число элементарных арифметических операций с ограниченной точностью, выполненных в процессе всей перестройки.

Полагаем, что целевая пороговая функция f задается при помощи множеств $S = \{s_1, \ldots, s_p\}$ и $C = \{c_1, \ldots, c_q\}$ наборов из E_2^n таких, что $f(s_i) = 1$ и $f(c_j) = 0$ для всех $i \in \{1, \ldots, p\}$ и $j \in \{1, \ldots, q\}$, соответственно. При этом, допускается частичное задание целевой функции. Иными словами, объединение множеств S и C не обязано содержать все наборы из E_2^n. В этом случае, под целевой функцией f понимается всякая пороговая функция, принимающая значение 1 на наборах из S и 0 — на наборах из C. Далее, множества S и C будем называть, соответственно, множеством единиц и нулей целевой функции.

Задание пороговой функции при помощи линейно-отделимых множеств наборов является классическим в теории распознавания. Одной из первых работ, в которой рассматривается данный способ, является работа Новикова [24], посвященная алгоритму обучения персептрона.

Алгоритмическая сложность перестройки оценивается внутри классов пороговых функций, инвариантных относительно групп перестановок.

В связи с тем, что всякую пороговую функцию можно задать бесконечным множеством линейных форм с целыми коэффициентами, то рассматриваемая задача перестройки «стартовой» линейной формы в «целевую» требует уточнения. Если не ограничить величину коэффициентов «стартовой» линейной формы, то сложность перестройки является неограниченной. Из результатов работы Мурога [22] следует, что всякую пороговую функцию f можно задать линейной формой с целыми коэффициентами, по модулю не превосходящими $L(f)$.

Множество линейных форм назовем допустимым, если при $n \to \infty$ размахи линейных форм, входящих в данное множество, ограничены величиной $O(n)\, L(n)$.

Множитель $O(n)$ дает некую свободу выбора линейных форм для задания пороговых функций и, как будет показано далее, существенно облегчает задачу поиска линейной формы для «целевой» пороговой функции. В связи с этим, будем рассматривать задачу перестройки на множестве допустимых линейных форм.

Дадим формальную постановку задачи перестройки пороговых функций. Пусть задана «стартовая» допустимая линейная форма. «Целевая» пороговая функция f задана в виде множеств $S = \{s_1, \ldots, s_p\}$ и $C = \{c_1, \ldots, c_q\}$ единиц и нулей, соответственно. Известно, что «стартовая» и «целевая» функции лежат в классе $T_{m,k}$. Требуется перестроить «стартовую» линейную форму так, чтобы полученная допустимая линейная форма задавала функцию f. При этом, на каждом шаге перестройки разрешено выполнять лишь одну операцию из множества допустимых. В качестве таких множеств рассматриваются множества M_I, M_{II} или M_a. Для определения направления перестройки разрешается выполнять вспомогательные арифметические операции. Их число оценивается отдельно.

2.2. Вспомогательные утверждения

2.2.1. Сопряженные пороговые функции

Пусть $x, d \in E_2$, тогда

$$x^d = \begin{cases} x, & \text{если } d = 1; \\ \bar{x}, & \text{если } d = 0. \end{cases}$$

Функции алгебры логики $f(x_1, \ldots, x_n)$ и $f'(x_1, \ldots, x_n)$ назовем сопряженными относительно переменных $x_{i_1}, \ldots, x_{i_k}$, $i_1, \ldots, i_k \in \{1, \ldots, n\}$, если

$$f(x_1, \ldots, x_n) = f'\left(x_1^{d_1}, \ldots, x_n^{d_n}\right),$$

где для всех $j \in \{1, \ldots, n\}$ выполнено

$$\begin{cases} d_j = 0, & \text{если } j \in \{i_1, \ldots, i_n\}; \\ d_j = 1, & \text{иначе.} \end{cases}$$

Пусть M и M' — множества пороговых функций, зависящих от переменных $x_1, \ldots, x_n$. Назовем множества M и M' сопряженными относительно переменных $x_{i_1}, \ldots, x_{i_k}$, $i_1, \ldots, i_k \in \{1, \ldots, n\}$, если для любой функции $f \in M$ найдется функция $f' \in M'$, сопряженная ей относительно переменных $x_{i_1}, \ldots, x_{i_k}$, и, наоборот, для любой функции $f' \in M'$ найдется сопряженная ей функция $f \in M$.

Назовем множества M и M' сопряженными, если они сопряжены относительно некоторого непустого множества переменных.

2.2.2. Изоморфные пороговые функции

Введенные ранее пороговые функции будем также называть $(0,1)$-пороговыми функциями. Введем понятие $(-1,1)$-пороговой функции.

Пусть $V = \{v_1, v_2 \ldots\}$ счетный алфавит переменных, каждое из которых может принимать значения из множества $\bar{E}_2 = \{-1, 1\}$. Для обозначения букв алфавита V будем использовать метасимволы y_i с индексами или без них.

Функция $g\left(y_1, \ldots, y_n\right) : \bar{E}_2^n \to \bar{E}_2$ называется $(-1,1)$-*пороговой*, если существует линейная форма $l_{\vec{w},\sigma}\left(y_1, \ldots, y_n\right) = y_1 w_1 + \ldots + y_n w_n - \sigma$ такая, что

$$g(y_1, \ldots, y_n) = \begin{cases} -1, \text{ если } \sum\limits_{i=1}^{n} y_i w_i - \sigma \geqslant 0; \\ 1, \text{ иначе.} \end{cases}$$

Множество $(-1,1)$-пороговых функций обозначим $\bar{T}^n$.

Для того, чтобы отличать $(0,1)$-пороговые и $(-1,1)$-пороговые функции, введем следующие обозначения: $f^{0,1}$ и $g^{-1,1}$. Иногда для $(0,1)$-пороговых функций верхний индекс будем опускать.

Сопоставим переменным алфавита U переменные алфавита V по следующему правилу: $\varphi\left(u_i\right) = v_i$, для всех i. Положим также

$$\begin{cases} \varphi\left(0\right) = -1; \\ \varphi\left(1\right) = 1. \end{cases}$$

Определим изоморфизм множеств T^n и $\bar{T}^n$ следующим образом: каждой $(0,1)$-пороговой функции $f^{0,1}\left(x_1, \ldots, x_n\right)$ поставим в соответствие $(-1,1)$-пороговую функцию $g^{-1,1}\left(y_1, \ldots, y_n\right)$ следующим образом

$$\varphi\left(f\left(x_1, \ldots, x_n\right)\right) = g\left(\varphi\left(x_1\right), \ldots, \varphi\left(x_n\right)\right).$$

Очевидно, что определенное соответствие является взаимно-однозначным. Далее, для краткости, соответствующие друг другу в терминах описанного изоморфизма функции f и g будем называть *изоморфными* и обозначать их $f^{0,1}$ и $f^{-1,1}$.

Следующее утверждение позволяет строить линейные формы, задающие изоморфные пороговые функции.

Лемма 2.1. *Имеют место следующие утверждения:*

1) если $l_{\vec{w},\sigma} \to f^{0,1}$ *и* $l_{\vec{w},2\sigma-\sum\limits_{i=1}^{n} w_i} \to g^{-1,1}$, *то* $f^{0,1} \sim g^{-1,1}$;

2) если $l_{\vec{w},\sigma} \to g^{-1,1}$ *и* $l_{\vec{w},\frac{1}{2}\left(\sum\limits_{i=1}^{n} w_i+\sigma\right)} \to f^{0,1}$, *то* $g^{-1,1} \sim f^{0,1}$.

Доказательство. 1) Рассмотрим набор $(a_1,\ldots,a_n) \in \{0,1\}^n$. Значение функции $f^{0,1}$ на этом наборе определяется знаком суммы

$$\varphi\left(a_1,\ldots,a_n\right) = \sum_{i=1}^{n} a_i w_i - \sigma.$$

При этом, значение функции $g^{-1,1}$ на наборе $(2a_1-1,\ldots,2a_n-1)$ определяется знаком следующего выражения

$$\psi\left(a_1,\ldots,a_n\right) = \sum_{i=1}^{n}\left(2a_i-1\right)w_i - 2\sigma + \sum_{i=1}^{n} w_i = 2\sum_{i=1}^{n} a_i w_i - 2\sigma.$$

Очевидно, что знаки $\varphi\left(a_1,\ldots,a_n\right)$ и $\psi\left(a_1,\ldots,a_n\right)$ совпадают, а, следовательно, функции $f^{0,1}$ и $g^{-1,1}$ изоморфны.

2) Рассмотрим набор $(b_1,\ldots,b_n) \in \{-1,1\}^n$. Значение функции $g^{-1,1}$ на наборе $(b_1,\ldots,b_n)$ определяется знаком следующей суммы

$$\varphi\left(b_1,\ldots,b_n\right) = \sum_{i=1}^{n} b_i w_i - \sigma.$$

Значение функции $f^{0,1}$ на наборе $\left(\frac{1}{2}(b_1+1),\ldots,\frac{1}{2}(b_n+1)\right)$ определяется знаком выражения

$$\psi(b_1,\ldots,b_n)=\sum_{i=1}^{n}\frac{1}{2}(b_i+1)w_i-\frac{1}{2}\left(\sum_{i=1}^{n}b_i+\sigma\right)=$$
$$=\frac{1}{2}\left(\sum_{i=1}^{n}b_iw_i+\sum_{i=1}^{n}w_i-\sum_{i=1}^{n}w_i-\sigma\right)=\frac{1}{2}\left(\sum_{i=1}^{n}b_iw_i-\sigma\right).$$

Очевидно, что знаки сумм $\varphi(b_1,\ldots,b_n)$ и $\psi(b_1,\ldots,b_n)$ совпадают. Следовательно, функции $g^{-1,1}$ и $f^{0,1}$ изоморфны. Лемма доказана.

2.3. Конструктивная характеризация пороговых функций

Назовем пороговую функцию $f(x_1,\ldots,x_n)$ *$\mathbb{Z}$-пороговой*, если существует линейная форма $x_1w_1+\ldots+x_nw_n-\sigma$ с целочисленными коэффициентами и порогом, задающая функцию $f(x_1,\ldots,x_n)$. Множество $\mathbb{Z}$-пороговых функций обозначим $\mathbb{Z}T$. Можно показать, что имеет место следующее утверждение.

Лемма 2.2. *Классы $\mathbb{Z}T$ и T совпадают.*

В связи с леммой 2.2 далее, если это не будет оговариваться отдельно, будем рассматривать линейные формы только с целочисленными коэффициентами и порогом.

Из леммы 2.2 также вытекает важное следствие.

Лемма 2.3. *Если линейная форма $l_{\vec{w},\sigma}$ задает пороговую функцию $f(x_1,\ldots,x_n)$ и гиперплоскость, задаваемая линейной формой $l_{\vec{w},\sigma}$, не проходит ни через одну точку множества E_2^n, то найдется такое $\varepsilon>0,\varepsilon\in\mathbb{R}$, что для всех $i\in\{1,\ldots,n\}$ выполнено*

$$l_{(w_1,\ldots,w_{i-1},w_i+\varepsilon,w_{i+1},\ldots,w_n),\sigma}\to f(x_1,\ldots,x_n),$$
$$l_{(w_1,\ldots,w_n),\sigma+\varepsilon}\to f(x_1,\ldots,x_n).$$

Сигнатурой линейной формы $l_{\vec{w},\sigma} = x_1 w_1 + \ldots + x_n w_n - \sigma$ называем вектор $s\left(l_{\vec{w},\sigma}\right) = (s_1, \ldots, s_n)$, такой что:

$$s_i = \begin{cases} 1, & \text{если } w_i > 0; \\ 0, & \text{иначе,} \end{cases}$$

для всех $i \in \{1, \ldots, n\}$.

Будем говорить, что функция $f(x_1, \ldots, x_n)$ *существенно зависит от переменного* x_i, $i \in \{1, \ldots, n\}$, если найдутся два набора

$$\alpha = (a_1, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_n),$$
$$\alpha' = (a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_n),$$

такие, что $f(\alpha) \neq f(\alpha')$.

Пороговую функцию считаем *существенной*, если она существенно зависит от всех своих переменных.

Линейную форму $l_{\vec{w},\sigma}$ будем называть *существенной*, если задаваемая ею пороговая функция существенна. Множество существенных пороговых функций от n переменных обозначим через $\tilde{T}^n$.

Следующее утверждение характеризует сигнатуру существенной пороговой функции.

Лемма 2.4. *Если $l_{\vec{w},\sigma}$ и $l_{\vec{w}',\sigma'}$ задают существенную пороговую функцию f, то $s\left(l_{\vec{w},\sigma}\right) = s\left(l_{\vec{w}',\sigma'}\right)$.*

Из утверждения 2.4 следует, что сигнатура линейной формы существенной пороговой функции f однозначно определяется этой функцией, что обозначаем $s(f)$. Отношение равенства сигнатур является отношением эквивалентности и задает разбиение множества существенных пороговых функций на классы эквивалентности. Возникает вопрос: сколько этих классов и как они устроены? Следующее утверждение дает ответ на данный вопрос.

Введем отношение частичного порядка на множестве E_2^n. Говорят, что набор $\alpha = (a_1, \ldots, a_n)$ меньше или равен набору $\alpha' = (a'_1, \ldots, a'_n)$, где $\alpha, \alpha' \in E_2^n$, если $a_i \leqslant a'_i$ для всех $i = 1, \ldots, n$.

Функция алгебры логики $f : \{0,1\}^n \to \{0,1\}$ называется монотонной, если из условия $\alpha \leqslant \beta$, где $\alpha, \beta \in \{0,1\}^n$, следует $f(\alpha) \leqslant f(\beta)$.

Имеет место следующее утверждение.

Лемма 2.5. *Отношение равенства сигнатур разбивает множество $\tilde{T}^n$ на 2^n взаимно сопряженных равномощных множества, одним из которых является множество монотонных существенных пороговых функций от n переменных.*

Пусть $f(x_1, \ldots, x_n)$ — пороговая функция, а l_1 и l_2 — линейные формы от n переменных, задающие f. Легко видеть, что линейная форма $l_1 + l_2$ также задает функцию f. Обозначим $U(f)$ — множество линейных форм от n переменных, задающих f. Как уже отмечалось ранее, множество $U(f)$ замкнуто относительно операции сложения. Возникает вопрос, содержится ли в $U(f)$ базис относительно операции сложения? Если да, то является ли он конечным, и сколько различных базисов содержится в $U(f)$. Следующее утверждение дает ответы на эти вопросы.

Теорема 6. *Для любой пороговой функции $f(x_1, \ldots, x_n)$ множество $U(f)$ содержит единственный базис относительно операции сложения, который является счетным и разрешимым.*

Отметим, что в доказательстве теоремы 6 явно описан алгоритм, который последовательно строит базис множества $U(f)$.

2.4. Доказательства утверждений

Докажем лемму 2.2. Из определения классов $\mathbb{Z}T$ и T, очевидно, следует, что $\mathbb{Z}T \subseteq T$. Докажем обратное вложение. Очевидно, что если $l_{\vec{w},\sigma} \to f(x_1, \ldots, x_n)$, то для любого положительного числа k

$$l_{k\cdot\vec{w},k\cdot\sigma} \to f(x_1, \ldots, x_n).$$

Обозначим $\rho_{l_{\vec{w},\sigma}}(\alpha)$ — кратчайшее расстояние от точки α до гиперплоскости $l_{\vec{w},\sigma}$. Известно [8], что

$$\rho_{l_{\vec{w},\sigma}}(\alpha) = \frac{\left|\sum\limits_{i=1}^{n} a_i w_i - \sigma\right|}{\sqrt{\sum\limits_{i=1}^{n} w_i^2}}.$$

Покажем, что линейной формой с целочисленными коэффициентами можно сколь угодно точно аппроксимировать линейную форму с действительными коэффициентами.

Лемма 2.6. *Если $\vec{w} \in \mathbb{R}^n$, $\sigma \in \mathbb{R}$ и $\varepsilon > 0$, то найдутся $\vec{w}' \in \mathbb{Z}^n$ и $\sigma' \in \mathbb{Z}$ такие, что для угла γ между векторами $\vec{w}$ и $\vec{w}'$ и для расстояний $\rho_{l_{\vec{w},\sigma}}(\alpha)$, для всех $\alpha = (a_1, \ldots, a_n) \in E_2^n$, будет выполнено $\left|\rho_{l_{\vec{w},\sigma}}(\alpha) - \rho_{l_{\vec{w}',\sigma'}}(\alpha)\right| < \varepsilon$ и $\gamma < \varepsilon$.*

Доказательство. Для доказательства достаточно построить такую последовательность линейных форм $l_{\vec{w}_1,\sigma_1}, l_{\vec{w}_2,\sigma_2}, \ldots, l_{\vec{w}_k,\sigma_k}, \ldots$ с целыми коэффициентами и порогом, что для всех α

$$\lim_{k\to\infty}\Big(\rho_{l_{\vec{w},\sigma}}(\alpha) - \rho_{l_{\vec{w}_k,\sigma_k}}(\alpha)\Big) = 0,$$
$$\lim_{k\to\infty}\cos\gamma_k = \lim_{k\to\infty}\frac{\vec{w}\cdot\vec{w}_k}{|\vec{w}|\cdot|\vec{w}_k|} = 1,$$

где γ_k — угол между векторами $\vec{w}$ и $\vec{w}_k$.

Рассмотрим десятичное представление компонент w_i вектора $\vec{w}$ и порога σ

$$w_i = \langle w_{i0}, w_{i1}\ldots w_{ik}\ldots\rangle;$$
$$\sigma = \langle s_0, s_1 s_2 \ldots s_k \ldots\rangle.$$

Здесь $w_{10}, \ldots, w_{n0}, s_0$ — целые части компонент вектора $\vec{w}$ и порога σ.

Положим $\vec{w}_k = (\langle w_{10}\ldots w_{1k}\rangle, \langle w_{20}\ldots w_{2k}\rangle, \ldots, \langle w_{n0}\ldots w_{nk}\rangle)$ и $\sigma_k = \langle s_0 \ldots s_k\rangle$.

Из данного представления следует, что $\vec{w}_k \to 10^k\cdot\vec{w}$, $\sigma_k \to 10^k\cdot\sigma$ при $k\to\infty$, а, следовательно,

$$\lim_{k\to\infty}(\rho_{l_{\vec{w},\sigma}}(\alpha) - \rho_{l_{\vec{w}_k,\sigma_k}}(\alpha)) = \lim_{k\to\infty}\left(\frac{\left|\sum\limits_{j=1}^{n} a_j w_j - \sigma\right|}{\sqrt{\sum\limits_{j=1}^{n} w_j^2}} - \frac{\left|\sum\limits_{j=1}^{n} a_j w_{kj} - \sigma_k\right|}{\sqrt{\sum\limits_{j=1}^{n} w_{kj}^2}}\right)$$

$$\lim_{k\to\infty}\left(\frac{\left|\sum\limits_{j=1}^{n} a_j w_j - \sigma\right|}{\sqrt{\sum\limits_{j=1}^{n} w_j^2}} - \frac{10^k\cdot\left|\sum\limits_{j=1}^{n} a_j w_j - \sigma\right|}{10^k\cdot\sqrt{\sum\limits_{j=1}^{n} w_j^2}}\right) = 0,$$

следовательно

$$\lim_{k\to\infty}\frac{\vec{w}\cdot\vec{w}_k}{|\vec{w}|\cdot|\vec{w}_k|} = \lim_{k\to\infty}\frac{10^k\cdot\vec{w}\cdot\vec{w}}{10^k\cdot|\vec{w}|\cdot|\vec{w}|} = 1.$$

Лемма доказана.

Лемма 2.7. *Справедливо соотношение: $T \subseteq \mathbb{Z}T$.*

Доказательство. Рассмотрим произвольную пороговую функцию $f(x_1,\ldots,x_n)$, задаваемую линейной формой $l_{\vec{w},\sigma}$, где $w_i,\sigma\in R$, $i = 1\ldots n$. Линейная форма $l_{\vec{w},\sigma}$ задает в пространстве $\mathbb{R}^n$ такую гиперплоскость, что по разные стороны от нее функция $f(x_1,\ldots,x_n)$ принимает различные значения.

Рассмотрим три случая:

1) Все точки множества E_2^n лежат по одну сторону от гиперплоскости, задаваемой линейной формой $l_{\vec{w},\sigma}$.
   В таком случае $f(x_1,\ldots,x_n)$ является константой 0 или 1. В обоих случаях видно, что $f(x_1,\ldots,x_n)\in\mathbb{Z}T$. Для этого достаточно рассмотреть линейные формы $l_{\vec{w},\sigma} = 1$ и $l_{\vec{w}',\sigma'} = -1$, соответственно.

2) Точки множества E_2^n расположены по обе стороны от гиперплоскости, задаваемой линейной формой $l_{\vec{w},\sigma}$, и ни одна из них не лежит на данной гиперплоскости.
   Рассмотрим величину

$$\rho_* = \min_{\alpha\in E_2^n}\rho_{l_{\vec{w},\sigma}}(\alpha).$$

   Очевидно, что в окрестности $V\left(l_{\vec{w},\sigma}\right) = \left\{v\in\mathbb{R}^n : \rho_{l_{\vec{w},\sigma}}(v) < \rho_*\right\}$ нет ни одной точки булего куба. Пусть B — шар с центром в точке $\left(\frac{1}{2},\ldots,\frac{1}{2}\right)$ и радиусом $\frac{\sqrt{n}}{2}$. В таком случае все точки множества E_2^n лежат в пределах шара B. Из леммы 2.6 следует, что существует гиперплоскость, заданная линейной формой $l_{\vec{w}',\sigma'}$, где $w'_i, \sigma' \in \mathbb{Z}$, $i = 1\ldots n$, такая, что в пределах B она лежит в окрестности $V\left(l_{\vec{w},\sigma}\right)$. Следовательно, соответствующее ей разбиение булего куба будет совпадать с разбиением, задаваемым линейной формой $l_{\vec{w},\sigma}$. Следовательно, нашлась такая линейная форма $l_{\vec{w}',\sigma'}$, где $w'_i, \sigma' \in \mathbb{Z}$, $i = 1\ldots n$, что задаваемая ей пороговая функция совпадает с $f(x_1, \ldots, x_n)$.

3) Точки множества E_2^n расположены по обе стороны от гиперплоскости, задаваемой линейной формой $l_{\vec{w},\sigma}$, а также существуют точки булего куба, лежащие на данной гиперплоскости. Рассмотрим величину $\rho_* = \min\limits_{\alpha \in E_2^n} \rho_{l_{\vec{w},\sigma}}(\alpha)$.

   Легко видеть, что линейная форма $l_{\vec{w},\sigma-\frac{\rho_*}{2}}$ задает такое же разбиение, как $l_{\vec{w},\sigma}$, при этом, ни одна из точек множества E_2^n не лежит на задаваемой ею гиперплоскости. Таким образом, данный случай сводится к предыдущему.

Лемма доказана.

Из леммы 2.7 следует очевидно следует лемма 2.2.

Докажем леммы 2.4 и 2.5. Очевидно, что если $f(x_1, \ldots, x_n)$ — существенная пороговая функция, то ни один из коэффициентов линейной формы $l_{\vec{w},\sigma}$, задающей f, не может равняться нулю. Докажем следующее вспомогательное утверждение.

Лемма 2.8. *Если $f(x_1, \ldots, x_n)$ — существенная пороговая функция и $l_{\vec{w},\sigma} \to f(x_1, \ldots, x_n)$, то функция f является монотонной точно тогда, когда $s\left(l_{\vec{w},\sigma}\right) = (1, \ldots, 1)$.*

Доказательство. Достаточность очевидна. Докажем необходимость.

Назовем набор $\alpha \in E_2^n$ нижней единицей монотонной функции $f(x_1, \ldots, x_n)$, если $f(\alpha) = 1$, а на любом наборе $\alpha' \in E_2^n$, $\alpha' \leqslant \alpha$, имеет место $f(\alpha') = 0$. Так как функция f монотонная, то она однозначно задается множеством своих нижних единиц $\alpha_1, \ldots, \alpha_m$, где $\alpha_k \in E_2^n$, $k = 1\ldots m$. Рассмотрим сокращенную дизъюнктивную нормальную форму

$$B = A_1 \vee A_2 \vee \ldots \vee A_m,$$

для функции f, где A_i — простая импликанта, соответствующая нижней единице α_i, $i = 1 \ldots m$. Переменная x_j, $j = 1 \ldots n$, входит в импликанту A_i точно тогда, когда на j-й позиции набор α_i принимает значение 1.

Так как f существенным образом зависит от всех своих переменных, то для всякого переменного x_i, $i = 1 \ldots n$, найдется хотя бы одна импликанта A_j, $j \in \{1, \ldots, m\}$, зависящая от x_i. Рассмотрим наборы

$$\begin{aligned} \alpha_j &= (a_1, \ldots a_{i-1}, 1, a_{i+1}, \ldots, a_n), \\ \alpha'_j &= (a_1, \ldots a_{i-1}, 0, a_{i+1}, \ldots, a_n). \end{aligned}$$

Пусть некоторая линейная форма $l_{\vec{w},\sigma}$ задает f. Так как α_j — нижняя единица f, то имеют место неравенства

$$\begin{cases} w_1 a_1 + \ldots + w_{i-1} a_{i-1} + w_i \cdot 1 + w_{i+1} a_{i+1} + \ldots + w_n a_n - \sigma \geqslant 0, \\ w_1 a_1 + \ldots + w_{i-1} a_{i-1} + w_i \cdot 0 + w_{i+1} a_{i+1} + \ldots + w_n a_n - \sigma < 0, \end{cases}$$

что возможно только в случае $w_i > 0$.

Так как соответствующие наборы α_j и α'_j найдутся для каждого переменного x_i, то $s\left(l_{\vec{w},\sigma}\right) = (1, \ldots, 1)$. Лемма доказана.

Пусть $\delta = (d_1, \ldots, d_n)$, где $d_i \in \{0, 1\}$ для всех $i = 1, \ldots, n$. Назовем δ-преобразованием оператор, определенный на множестве линейных форм и принимающий значения на нем же. Этот оператор записывается так

$$\delta\left[l_{\vec{w},\sigma}\right](x_1, \ldots, x_n) = l_{\vec{w}',\sigma'}(x_1, \ldots, x_n),$$

где

$$\begin{cases} w'_i = w_i, \text{ если } d_i = 1, \\ w'_i = -w_i, \text{ иначе}, \\ \sigma' = \sigma + \sum\limits_{i=1}^{n} (d_i - 1) w_i, \end{cases}$$

для всех $i = 1 \ldots n$.

Отметим важное свойство δ-преобразования.

Лемма 2.9. *Если* $f_{\vec{w},\sigma}(x_1,\ldots,x_n) \in \mathbb{Z}T$, $\delta = (d_1,\ldots,d_n) \in E_2^n$ *и* $\alpha = (a_1,\ldots,a_n) \in E_2^n$, *то*

$$f_{\vec{w},\sigma}\left(a_1,\ldots,a_n\right) = f_{\delta[\vec{w},\sigma]}\left(a_1^{d_1},\ldots,a_n^{d_n}\right).$$

Доказательство. Утверждение леммы вытекает из следующих соображений.

Значение функции $f_{\vec{w},\sigma}\left(a_1,\ldots,a_n\right)$ определяется знаком выражения

$$\sum_{i=1}^{n} a_i w_i - \sigma.$$

Значение функции $f_{\delta[\vec{w},\sigma]}\left(a_1^{d_1},\ldots,a_n^{d_n}\right)$, в свою очередь, определяется знаком выражения

$$\sum_{i=1}^{n} (-1)^{d_i+1} \cdot a_i^{d_i} \cdot w_i + \sum_{i=1}^{n} \left(1-d_i\right) w_i - \sigma.$$

Перегруппировав слагаемые, данное выражение можно записать так

$$\sum_{i=1}^{n} w_i \cdot \left[(-1)^{d_i+1} \cdot a_i^{d_i} - d_i + 1\right] - \sigma.$$

Непосредственной проверкой всех возможных значений a_i и d_i можно убедиться в том, что слагаемое w_i участвует в сумме тогда и только тогда, когда $a_i = 1$. Таким образом на всех наборах $\alpha = (a_1,\ldots,a_n) \in E_2^n$ значения данных сумм равны. Лемма доказана.

Из лемм 2.8 и 2.9 следует, что сигнатура пороговой функции, получаемой в результате применения оператора δ-преобразования к линейной форме, задающей существенную пороговую функцию, определена однозначным образом. Таким образом, можно определить δ-преобразование существенной пороговой функции.

Если $f\left(x_1,\ldots,x_n\right)$ — существенная пороговая функция, $\delta = (d_1,\ldots,d_n)$, где $d_i \in \{0,1\}$, то

$$\delta\left[f_{\vec{w},\sigma}\right]\left(x_1,\ldots,x_n\right) = f_{\vec{w}',\sigma'}\left(x_1,\ldots,x_n\right),$$

где

$$\begin{cases} w_i' = w_i, \text{ если } d_i = 1, \\ w_i' = -w_i, \text{ иначе}, \\ \sigma' = \sigma + \sum\limits_{i=1}^{n} (d_i - 1)\, w_i, \end{cases}$$

для всех $i = 1 \ldots n$.

При этом, как следует из леммы 2.9, в результате применения оператора δ-преобразования к существенной пороговой функции получается сопряженная ей существенная пороговая функция.

Обозначим через $\tilde{T}_s^n$ множество всех существенных пороговых функций от n переменных, имеющих сигнатуру s. Имеют место следующие следствия леммы 2.9

1) если $s, \delta \in E_2^n$, $f, f' \in \tilde{T}_s^n$, $f \neq f'$, то $\delta\,[f] \neq \delta\,[f']$,
2) если $f \in \tilde{T}_s^n$, $\delta = s\,(f)$, то $\delta\,[f]$ — монотонная существенная пороговая функция.

Имеет место следующее утверждение.

Лемма 2.10. *Если $s, s' \in E_2^n$ и $s \neq s'$, то $\tilde{T}_s^n \cap \tilde{T}_{s'}^n = \varnothing$.*

Доказательство. Пусть $f\,(x_1, \ldots, x_n)$ — существенная пороговая функция такая, что $f \in \tilde{T}_s^n$, $f \in \tilde{T}_{s'}^n$. Тогда существуют линейные формы $l_{\vec{w},\sigma}$ и $l_{\vec{w}',\sigma'}$, задающие f, такие, что:

$$\begin{cases} w_i > 0, \text{ если } s_i = 1; \\ w_i < 0, \text{ если } s_i = 0; \end{cases}$$

$$\begin{cases} w_i' > 0, \text{ если } s_i' = 1; \\ w_i' < 0, \text{ если } s_i' = 0; \end{cases}$$

соответственно.

Положим $\delta = s$ и выполним δ-преобразование функции f, заданной формами $l_{\vec{w},\sigma}$ и $l_{\vec{w}',\sigma'}$. По лемме 2.9 имеем

$$\delta\left[f_{\vec{w},\sigma}\,(x_1, \ldots, x_n)\right] = \delta\left[f_{\vec{w}',\sigma'}\,(x_1, \ldots, x_n)\right].$$

Так как вектор $\delta = s$, то из следствия леммы 2.9 в левой части этого выражения имеем монотонную функцию от n-переменных, а в правой части — функцию, не являющуюся монотонной, так как $s \neq s'$. Полученное противоречие завершает доказательство леммы.

Лемма 2.4 очевидным образом следует из леммы 2.10.

Доказательство леммы 2.5. Из леммы 2.10 следует, что

$$\tilde{T}^n = \bigcup_{i=1}^{2^n} \tilde{T}^n_{s_i},$$

где $s_1 = (0, \ldots, 0)$; $s_2 = (0, \ldots, 0, 1); \ldots; s_{2^n} = (1, \ldots, 1)$; при этом, если $i \neq j$, то $\tilde{T}^n_{s_i} \cap \tilde{T}^n_{s_j} = \varnothing$.

Пусть $s \neq (1, \ldots, 1)$, $\delta = s$. Рассмотрим множество

$$\delta \left[\tilde{T}^n_s\right] = \left\{\delta\left[f\right] : f \in \tilde{T}^n_s\right\}.$$

По следствию из леммы 2.9 данное множество будет состоять только из монотонных пороговых функций. Верно также и обратное: если $\delta' = \bar{\delta}$, то для любой монотонной пороговой функции f из $\tilde{T}_{(1,\ldots,1)}$ верно, что $\delta'\left[f\right] \in \tilde{T}^n_{(0,\ldots,0)}$. Следовательно, если $\delta = s$, то

$$\delta \left[\tilde{T}^n_s\right] = \tilde{T}^n_{(1,\ldots,1)}.$$

Следовательно, классы $\tilde{T}^n_{s_i}$, $i \in \{1, \ldots, 2^n\}$, взаимно сопряженые и равномощны. Лемма доказана.

Докажем теорему 6. Две существенные линейные формы $l_{\vec{w},\sigma}$ и $l_{\vec{w}',\sigma'}$ будем называть *эквивалентными* и обозначать $l_{\vec{w},\sigma} \sim l_{\vec{w}',\sigma'}$, если задаваемые ими пороговые функции равны, то есть $f_{\vec{w},\sigma} = f_{\vec{w}',\sigma'}$.

Как уже отмечалось ранее, множество $U\left(f\right)$ замкнуто относительно операции сложения линейных форм. Складывая эквивалентные формы сами с собой, получаем, что в нашем распоряжении есть также операция умножения линейной формы на натуральную константу.

Пусть L — множество линейных форм, обозначим через $[L]$ — замыкание множества L относительно операций сложения и умножения на натуральную константу, то есть множество всех линейных форм, которые можно получить из L при помощи данных операций.

Лемма 2.11. *Если f — существенная пороговая функция, то множество $U(f)$ является точно счетно-порожденным относительно операций сложения и умножения на константу из $\mathbb{Z}$.*

Доказательство. Предположим противное. Без ограничения общности будем полагать, что f — монотонная функция. Пусть найдется конечное множество линейных форм

$$L = \{l_1(x_1, \ldots, x_n), \ldots, l_k(x_1, \ldots, x_n)\}$$

такое, что $[L] = U(f)$.

Если найдется такая форма $l_i \in L$, которая представляется в виде линейной комбинации остальных форм из L, тогда, очевидно, $[L \backslash l_i] = [L]$. Поэтому естественно полагать, что множество L несократимо, то есть из него нельзя выбросить ни одной линейной формы, не сузив замыкание. Иными словами, можно считать, что множество L состоит только из линейно-независимых элементов.

Пусть $l_{\vec{w},\sigma} \in U(f)$ и ни одна из точек множества E_2^n не лежит на плоскости, задаваемой $l_{\vec{w},\sigma}$.

Если $[L] = U(f)$, то $l_{\vec{w},\sigma}$ представима в виде линейной комбинации элементов из L, то есть

$$l_{\vec{w},\sigma} = \sum_{j=1}^{k} a_j l_j,\ a_j \in \mathbb{N}_0, l_j \in L,\ j \in \{1, \ldots, k\}.$$

В таком случае, если $c \in \mathbb{N}$, то

$$\sum_{j=1}^{k} a_j l_j \sim c \cdot \left(\sum_{j=1}^{k} a_j l_j \right).$$

Рассмотрим линейную форму

$$l' = c \cdot \left(\sum_{j=1}^{k} a_j l_j \right) + p x_i,\ i \in \{1, \ldots, n\},$$

где $p \in \mathbb{N}$. Возможно два случая: $l_{\vec{w},\sigma} \sim l'$ и $l_{\vec{w},\sigma} \nsim l'$.

Покажем, что для любого $i \in \{1, \ldots, n\}$ найдется такой множитель $c \in N$, что $l_{\vec{w},\sigma} \sim l'$.

Из леммы 2.3 следует, что для произвольной линейной формы с целыми коэффициентами $l_{\vec{w},\sigma} = w_1x_1 + \ldots + w_nx_n - \sigma$, которая не проходит ни через одну точку множества E_2^n, существует линейная форма

$$l_\varepsilon = w_1x_1 + \ldots + w_{i-1}x_{i-1} + (w_i + \varepsilon)x_i + w_{i+1}x_{i+1} + \ldots + w_nx_n - \sigma,$$

где $\varepsilon = \frac{1}{q}$, $q \in \mathbb{N}$, такая, что $l_{\vec{w},\sigma} \sim l_\varepsilon$. При этом, форма l_ε существует для любого $i \in \{1, \ldots, n\}$.

Положим $c = p \cdot q$, тогда

$$l_{\vec{w},\sigma} \sim c \cdot l_\varepsilon = \sum_{j=1}^{i-1} p \cdot q \cdot w_jx_j + p \cdot q \cdot \left(w_i + \frac{1}{q}\right) x_i + \sum_{j=i+1}^{n} p \cdot q \cdot w_jx_j - \sigma =$$

$$= p \cdot q \cdot \left(\sum_{j=1}^{n} w_jx_j - \sigma\right) + p \cdot x_i = l'.$$

Таким образом, всегда возможно подобрать такой натуральный множитель c, что $l_{\vec{w},\sigma} \sim l'$.

Следовательно, $l' \in U(f)$ и, значит, l' представима в виде линейной комбинации элементов из L, то есть

$$l' = \sum_{j=1}^{k} a'_jl_j = \sum_{j=1}^{k} a_jl_j + px_i.$$

Пусть $\tilde{l} = x_i$. Возможны два случая: $\tilde{l} \in L$ и $\tilde{l} \notin L$. Покажем, что второй случай невозможен. Предположим противное. Линейную форму $l_{\vec{w},\sigma}$ будем называть положительно-определенной, если все ее весовые коэффициенты и порог положительны. Концентрацией переменного x_i в положительно-определенной линейной форме $l_{\vec{w},\sigma} = \sum\limits_{j=1}^{n} w_jx_j - \sigma$ назовем отношение

$$v(x_i, l_{\vec{w},\sigma}) = \frac{w_i}{\sum\limits_{j=1}^{n} w_j}.$$

Очевидны следующие свойства концентрации:

1) $0 \leqslant v(x_i, l_{\vec{w},\sigma}) \leqslant 1$;
2) если $\lambda \in \mathbb{N}$, то $v(x_i, l_{\vec{w},\sigma}) = v(x_i, l_{\lambda\vec{w},\lambda\sigma})$;
3) если заданы линейные формы $l_{\vec{w},\sigma}, l_{\vec{w}',\sigma'}$ и $\lambda, \lambda' \in N$, то

$$v(x_i, \lambda l_{\vec{w},\sigma} + \lambda' l_{\vec{w}',\sigma'}) \leqslant \max(v(x_i, l_{\vec{w},\sigma}), v(x_i, l_{\vec{w}',\sigma'})).$$

Так как множитель p выбирается произвольно, можно утверждать, что найдется такой множитель c, что

$$v(x_i, l') > \max(v(x_i, l_1), \ldots, v(x_i, l_k)),$$

что невозможно. Следовательно, $\tilde{l} \in L$. В связи с тем, что переменная x_i выбиралась произвольно, получаем, что все формы вида x_i содержатся в L. Если в качестве l' рассмотреть форму

$$l' = c \cdot (a_1 l_1 + \ldots + a_k l_k) + p, \text{ где } p \in \mathbb{N},$$

то получим, что линейная форма, тождественно равная единице, содержится в L.

Полученного набора линейных форм достаточно для получения произвольной положительно-определенной линейной формы. Получаем, что все положительно-определенные линейные формы содержатся в $[L]$ и, соответственно, в $U(f)$, что неверно. Лемма доказана.

Пусть L и B — множества линейных форм. Будем говорить, что *B полно в L* относительно операций сложения и умножения на натуральную константу, если $[B] = L$. Назовем *B базисом L* относительно данных операций, если B полно в L и для всякой линейной формы $l \in B$ выполнено $[B \backslash l] \neq L$.

Из леммы 2.11 следует, что $U(f)$ не может обладать конечным базисом. Остается вопрос о существовании в $U(f)$ бесконечных, а точнее, счетных базисов. Приведем примеры пороговых функций, для которых множество $U(f)$ содержит счетный базис.

Лемма 2.12. *Существует такая функция $f \in \tilde{T}^1$, что множество $U(f)$ обладает счетным базисом относительно операции сложения двух линейных форм.*

Доказательство. Рассмотрим пороговую функцию $f(x) = x$. Так как f — монотонная, то по лемме 2.4 множество $U(f)$ состоит только из положительно-определенных линейных форм. Следовательно, любая линейная форма $l \in U(f)$ имеет вид

$$l(x) = wx - \sigma,$$

где $w > 0$ и $\sigma > 0$. Таким образом, выполнено

$$U(f) = \{w \cdot x - \sigma : w \geqslant \sigma;\ w, \sigma \in N\}.$$

Множество $U(f)$ изображено на рис. 1.

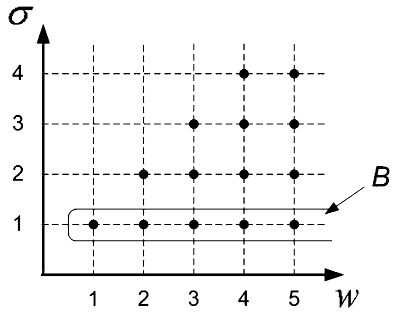

Рис. 1. Множество $U(f)$ при $f(x) = x$.

Рассмотрим множество линейных форм $B = \{(1,1), (2,1), (3,1), \ldots\}$. Здесь первая компонента соответствует значению w, а вторая σ. Данная система является полной относительно операции сложения, так как любая линейная форма $l_{w,\sigma}(x) = wx - \sigma$ из множества $U(f)$ может быть представлена в виде

$$l_{w,\sigma}(x) = l_{(w-\sigma+1),1}(x) + (\sigma - 1) \cdot l_{1,1}(x).$$

При этом, если из B отбросить произвольный элемент $l_{w,1}(x)$, то его невозможно будет представить в виде линейной комбинации остальных, так как при сумме любых двух элементов значение порога σ будет не менее 2. Поэтому B — базис множества $U(x)$.

Лемма 2.13. *Существует такая $f \in \tilde{T}^2$, что множество $U(f)$ обладает счетным базисом относительно операции сложения линейных форм.*

Доказательство. Рассмотрим пороговую функцию $f(x_1, x_2) = x_1 \& x_2$. Так как f монотонная, то по лемме 2.4 множество $U(f)$ состоит только из положительно-определенных линейных форм. Следовательно, любая линейная форма $l \in U(f)$ имеет вид

$$l(x_1, x_2) = w_1 x_1 + w_2 x_2 - \sigma,$$

где $w_1, w_2, \sigma > 0$.

Таким образом,

$$U(f) = \{w_1 \cdot x_1 + w_2 \cdot x_2 - \sigma : w_1 + w_2 \geqslant \sigma;\ w_1 < \sigma;\\ w_2 < \sigma;\ w_1, w_2, \sigma \in N\}.$$

Множество $U(f)$ можно представить в виде

$$U(f) = \bigcup_{i=2}^{\infty} U_{\sigma=i}(f),$$

где

$$U_i(f) = \left\{ l_{\vec{w},\sigma} \in \widetilde{ZL}^n : l_{\vec{w},\sigma} \to f;\ \sigma = i \right\}.$$

Здесь $\widetilde{ZL}^n$ — множество линейных форм с целыми коэффициентами и порогом от n переменных.

На рис. 2 изображены первые шесть членов данного разложения множества $U(x_1 \& x_2)$. Обозначим

$$B_i = \left\{ l_{\vec{w},\sigma} \in U_i(f) : (w_1 = i - 1) \vee (w_2 = i - 1) \right\},\ i = 2, 3, \ldots.$$

Рассмотрим множество линейных форм $B = B_2 \cup B_3 \cup B_4 \cup \ldots$. Покажем, что B является базисом $U(x_1 \& x_2)$.

Полнота B следует из того, что для любого $i \geqslant 4$ элементы множества $U_i(x_1 \& x_2) \backslash B_i$ (на рис. 2 данные множества обозначены серым цветом) могут быть представлены в виде $l_{(1,1,2)} + l'$, где

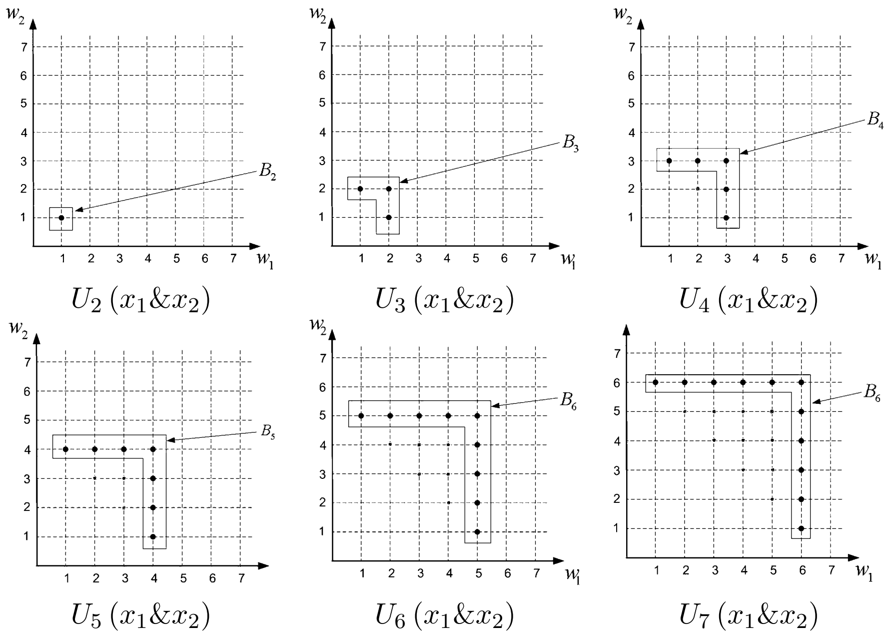

Рис. 2. Множество $U\,(x_1\&x_2)$.

$l_{(1,1,2)} = x_1 + x_2 - 2$, а l' — соответствующий элемент множества $U_{i-2}\,(x_1\&x_2)$.

Покажем, что ни один из элементов множества B не может быть выражен через остальные при помощи операций суммы и умножения на константу. Отметим, что для элементов множества $U_i\,(f)$ выполнено

$$w_1 < i,\ w_2 < i.$$

Таким образом, для любой линейной комбинации с положительными коэффициентами элементов из B, состоящей из более, чем одного элемента, выполнено

$$\begin{cases} w_1 < \sigma - 2; \\ w_2 \leqslant \sigma - 2. \end{cases}$$

В то же время для элементов множества B имеет место

$$\begin{cases} w_1 = \sigma - 1; \\ w_2 = \sigma - 1. \end{cases}$$

Следовательно, B является базисом $U\left(x_1 \& x_2\right)$. Лемма доказана.

Обобщая леммы 2.12 и 2.13, докажем теорему 6.

Доказательство теоремы 6. Без ограничения общности будем считать, что f — произвольная монотонная пороговая функция от n переменных. Для остальных пороговых функций утверждение будет следовать из леммы 2.5. Так как f монотонная, следовательно, множество $U\left(f\right)$ целиком содержится в множестве $\mathbb{N}^{n+1}$. Рассмотрим следующее разбиение множества $\mathbb{N}^{n+1}$:

$$\mathbb{N}^{n+1} = K_1 \cup \left(K_2 \backslash K_1\right) \cup \left(K_3 \backslash K_2\right) \cup \ldots,$$

где

$$K_i = \left\{\left(a_1, \ldots, a_{n+1}\right) : a_j \in \mathbb{N},\, a_j \leqslant i,\, j = 1 \ldots n+1\right\},\, i = 1 \ldots \infty.$$

Иными словами, K_i — куб в пространстве $\mathbb{N}^{n+1}$ со стороной i.

Пусть B — множество линейных форм, которое задается следующим образом:

1) На множестве K_1 ищем пересечение $K_1 \cap U\left(f\right)$. Так как в любом множестве K_i содержится $i^{(n+1)}$ элементов, то задача нахождения $K_1 \cap U\left(f\right)$ является алгоритмически разрешимой. Множество $K_1 \cap U\left(f\right)$ целиком заносим в B.
2) Переходим к рассмотрению множества $\left(K_2 \backslash K_1\right)$. Аналогично предыдущему шагу, находим множество $\left(K_2 \backslash K_1\right) \cap U\left(f\right)$. Из данного множества отбрасываем те элементы, которые могут быть получены из множества $K_1 \cap U\left(f\right)$ при помощи операций сложения и умножения на константу — обозначим данное множество L_2. Задача нахождения множества L_2 также является алгоритмически разрешимой, так как всевозможных линейных комбинаций элементов из $K_1 \cap U\left(f\right)$, координаты которых не превосходят 2, конечное число. Множество $\left[\left(K_2 \backslash K_1\right) \cap U\left(f\right)\right] \backslash L_2$ целиком заносим в B.

3) Переходим к рассмотрению множества $(K_3 \backslash K_2)$. Аналогично находим множество $[(K_3 \backslash K_2) \cap U(f)] \backslash L_3$. Здесь L_3 — множество всех элементов $(K_3 \backslash K_2) \cap U(f)$, которые могут быть получены при помощи операций сложения и умножения из множества $K_2 \cap U(f)$.
4) и т. д.
5) Переходим к рассмотрению множества $(K_{i+1} \backslash K_i)$. Находим множество $[(K_{i+1} \backslash K_i) \cap U(f)] \backslash L_{i+1}$, где L_{i+1} — множество всех элементов $(K_{i+1} \backslash K_i) \cap U(f)$, которые могут быть получены при помощи операций сложения и умножения из множества $K_i \cap U(f)$.
6) и т. д.

Полученное множество B имеет вид:

$$B = \bigcup_{i=1}^{\infty} [(K_i \backslash K_{i-1}) \cap U(f)] \backslash L_i,$$

где L_i — те элементы множества $(K_i \backslash K_{i-1}) \cap U(f)$, которые могут быть получены при помощи операций сложения и умножения из множества $K_{i-1} \cap U(f)$. Полагаем $L_1 = \varnothing$.

Покажем, что B — базис множества $U(f)$. Полнота B следует из построения — любой элемент $U(f)$ либо находится в B, либо линейно выражется через его элементы. При этом, из B нельзя отбросить ни один элемент. Это следует из того, что элементы множества

$$[(K_i \backslash K_{i-1}) \cap U(f)] \backslash L_i$$

не могут быть линейно выражены через элементы множества

$$[(K_j \backslash K_{j-1}) \cap U(f)] \backslash L_j,$$

если $j > i$. В противном случае любая линейная комбинация элементов из $[(K_j \backslash K_{j-1}) \cap U(f)] \backslash L_j$ имеет координаты больше, чем i. Поэтому B — базис. Единственность B следует из того, что ни один из элементнов множества $[(K_i \backslash K_{i-1}) \cap U(f)] \backslash L_i$, $i = 1, 2, \ldots$, не может быть выражен через остальные. Утверждение доказано.

3. Обучение нейронов, инвариантных относительно групп перестановок переменных

3.1. Основные понятия, постановки и результаты

Рассмотрим множество $\Omega_n = \{1, \ldots, n\}$. Элементы $a, b \in \Omega_n$ назовем π-*эквивалентными*, если $a = \pi^r(b)$ для некоторого целого r. Отношение π-эквивалентности рабивает множество Ω_n на попарно непересекающиеся классы $O_1, \ldots, O_k$, которые принято называть π-*орбитами*.

Далее, пусть $R = \{R_1, \ldots, R_k\}$ — некоторое разбиение множества Ω_n. Говорят, что элементы $a, b \in \Omega_n$ *эквивалентны относительно разбиения* R, если a и b принадлежат одному и тому же подмножеству R_i разбиения R, и обозначается это так: $a \sim b \pmod R$.

Будем говорить, что перестановка π сохраняет разбиение R, если для всякого $a \in \Omega_n$ выполнено $a \sim \pi(a) \pmod R$.

Теорема 7. *Если пороговая функция $f \in T^n$ инвариантна относительно перестановки $\pi \in S_n$, и $O_1, \ldots, O_k$ — π-орбиты, то f инвариантна относительно всякой перестановки $\pi' \in S_n$, сохраняющей разбиение $O_1, \ldots, O_k$.*

Пусть $R = \{R_1, \ldots, R_k\}$ — разбиение множества Ω_n. Пороговую функцию f назовем R-симметрической, если она инвариантна относительно всякой перестановки, сохраняющей разбиение R. Из теоремы 7 следует, что каждый класс пороговых функций, инвариантых относительно некоторой группы перестановок, однозначно задается соответствующим разбиением R множества переменных. Верно и обратное: всякому разбиению R множества переменных соответствует класс R-симметрических пороговых функций, инвариантных относительно перестановок, сохраняющих разбиение R.

Переменные R-симметрической функции, принадлежащие одному классу эквивалентности R_i, будем называть *симметричными*.

Рассмотрим множество $T_{m,k}$ пороговых функций от n переменных, где $n \leqslant m \cdot k$, таких, что для каждой функции f из $T_{m,k}$ существует разбиение $R = \{R_1, \ldots, R_k\}$ множества $\Omega_n = \{1, \ldots, n\}$ такое, что $\max(|R_1|, \ldots, |R_k|) = m$ и f является R-симметрической. Параметр m характеризует максимальный размер класса симметрии: если $m = 1$, то функция — антисимметрическая, если же $m = n$, то симметрическая. Параметр k характеризует число независимых классов симметрии.

Определим величину $\rho(m, k)$ следующим образом

$$\rho(m, k) = \max_{f', f'' \in T_{m,k}} \rho\left(f', f''\right),$$

где максимум берется по всем парам пороговых функций f', f'' из класса $T_{m,k}$. Данная величина характеризует близость между наиболее удаленными пороговыми функциями, не более чем с m взаимно симметричными переменными, и числом классов симметрии — k.

Имеет место следующая теорема, характеризующая сложность взаимной перестройки внутри классов пороговых функций, инвариантных относительно групп перестановок.

Теорема 1. *1 Для достаточно больших k выполнено*

$$m \cdot 2^{\frac{1}{2} k \log k + o(k \log k)} \leqslant \rho(m, k) \leqslant m^{k+2} \cdot 2^{\frac{1}{2} k \log k + o(k \log k)}.$$

3.2. Доказательства утверждений

Доказательство теоремы 7. Рассмотрим случай, когда

$$\begin{cases} O_1 = \{1, \ldots, k\}, \\ O_2 = \{k+1\}, \\ \ldots \\ O_1 = \{n\}. \end{cases}$$

Иными словами, перестановка π содержит единственный цикл, состоящий из элементов множества $\{1, \ldots, k\}$. Покажем, что в таком случае для всякой перестановки $\pi' \in S_k$ выполнено

$$f(x_1, \ldots, x_n) = f\left(x_{\pi'(1)}, \ldots, x_{\pi'(k)}, x_{k+1}, \ldots, x_n\right).$$

Так как всякая перестановка $\pi \in S_n$ однозначно раскладывается в произведение независимых циклов, то отсюда будет следовать утверждение теоремы.

Положим

$$\begin{cases} x_{k+1} = a_{k+1}, \\ \quad \dots \\ x_n = a_n, \end{cases}$$

и рассмотрим пороговую функцию

$$g\left(x_1, \dots, x_k\right) = f\left(x_1, \dots, x_k, a_{k+1}, \dots, a_n\right).$$

Для доказательства теоремы достаточно показать, что g является симметрической функцией. То есть для всякой перестановки $\pi' \in S_k$ выполнено

$$g\left(x_1, \dots, x_k\right) = g\left(x_{\pi'(1)}, \dots, x_{\pi'(k)}\right).$$

Предположим противное: пусть найдется перестановка $\pi' \in S_k$ такая, что

$$g\left(x_1, \dots, x_k\right) \neq g\left(x_{\pi'(1)}, \dots, x_{\pi'(k)}\right).$$

То есть для некоторого набора $\alpha = \left(a_1, \dots, a_k\right)$

$$g\left(a_1, \dots, a_n\right) \neq g\left(a_{\pi'(1)}, \dots, a_{\pi'(k)}\right).$$

Без ограничения общности положим, что $g\left(\alpha\right) = 1$.

По условию теоремы имеем

$$g\left(x_1, \dots, x_k\right) = g\left(x_{\pi(1)}, \dots, x_{\pi(k)}\right).$$

Рассмотрим множество наборов $P = \left\{\alpha, \pi\left(\alpha\right), \dots, \pi^{s-1}\left(\alpha\right)\right\}$, где s — порядок перестановки π. Множество P представляет собой все наборы, получающихся из α применением перестановки π во всех возможных степенях. Так как g инвариантна относительно перестановки π, то для всякого набора $\gamma \in P$ выполнено $g\left(\gamma\right) = 1$.

Следовательно, для линейной формы $\left(w_1, \dots, w_n, \sigma\right)$, задающей функцию f, имеет место следующая система неравенств

$$\begin{cases} \bar{w} \cdot \alpha - \sigma \geqslant 0, \\ \bar{w} \cdot \pi\left(\alpha\right) - \sigma \geqslant 0, \\ \dots \\ \bar{w} \cdot \pi^{s-1}\left(\alpha\right) - \sigma \geqslant 0. \end{cases}$$

Сложив все неравенства вместе, получим

$$\bar{w} \cdot \sum_{i=0}^{s-1} \pi^i(\alpha) - s \cdot \sigma \geqslant 0.$$

Легко видеть, что

$$\sum_{i=0}^{s-1} \pi^i(\alpha) = (l(\alpha), \ldots, l(\alpha)),$$

где $l(\alpha)$ — число единиц в наборе α.

Следовательно,

$$\bar{w} \cdot (l(\alpha), \ldots, l(\alpha)) - s \cdot \sigma \geqslant 0.$$

Далее, рассмотрим набор $\beta = \left(a_{\pi'(1)}, \ldots, a_{\pi'(k)}\right)$. По предположению имеем $g(\beta) = 0$.

Рассмотрим множество наборов $P' = \left\{\beta, \pi(\beta), \ldots, \pi^{s-1}(\beta)\right\}$. Так как функция g инвариантна относительно перестановки π, то для всякого набора $\gamma \in P'$ верно

$$g(\gamma) = 0.$$

Следовательно, имеет место система неравенств

$$\begin{cases} \bar{w} \cdot \beta - \sigma < 0, \\ \bar{w} \cdot \pi(\beta) - \sigma < 0, \\ \ldots \\ \bar{w} \cdot \pi^{s-1}(\beta) - \sigma < 0. \end{cases}$$

Аналогичным образом приходим к выводу, что

$$\bar{w} \cdot (l(\beta), \ldots, l(\beta)) - s \cdot \sigma < 0.$$

Осталось заметить, что $l(\beta) = l(\alpha)$, так как β получается из α при помощи перестановки π'. Пришли к противоречию. Следовательно функция g — симметрическая.

Теорема доказана.

Рассмотрим разбиение $R = \{R_1, \ldots, R_k\}$ множества $\Omega_n = \{1, \ldots, n\}$. Множество всех R-симметрических пороговых функций от n переменных обозначим T_R^n. Имеет место следующая лемма, характеризующая особенности задания R-симметрических пороговых функций линейными формами.

Лемма 3.1. *Если $R = \{R_1, \ldots, R_k\}$ — разбиение множества Ω_n, пороговая функция f является R-симметрической и линейная форма $l_{\bar{w},\sigma} = (w_1, \ldots, w_n, \sigma)$ задает f, то линейная форма*

$$l_{\bar{w}',\sigma} = \left(w'_{h(1)}, \ldots, w'_{h(n)}, \sigma\right),$$

также задает функцию f. Здесь $h(a) = i$, если $a \in R_i$, и

$$w'_i = \frac{1}{|R_i|} \sum_{r \in R_i} w_r.$$

Доказательство. Рассмотрим случай, когда

$$R = \{\{1, \ldots, l\}, \{l+1\}, \{l+2\}, \ldots, \{n\}\}.$$

Положим

$$\begin{cases} x_{l+1} = a_{l+1}, \\ \ldots \\ x_n = a_n. \end{cases}$$

В таком случае

$$l_{\bar{w},\sigma}(x_1, \ldots, x_l, a_{l+1}, \ldots, a_n) = \sum_{i=1}^{l} x_i \cdot w_i + \sum_{j=l+1}^{n} a_j \cdot w_j - \sigma.$$

Рассмотрим линейную форму $l_{\bar{w}'',\sigma}(x_1, \ldots, x_l) = (w_1, \ldots, w_l, \theta)$, где

$$\theta = -\sum_{j=l+1}^{n} a_j \cdot w_j + \sigma,$$

и задаваемую ей пороговую функцию f''. Очевидно, что

$$f''(x_1, \ldots, x_l) = f(x_1, \ldots, x_l, a_{l+1}, \ldots, a_n).$$

Так как пороговая функция f является R-симметрической, то по теореме 7 функция f'' является симметрической. Покажем, что линейная форма $l_{\bar{w}',\sigma}(x_1, \ldots, x_l) = (w', \ldots, w', \theta)$, где $w' = \frac{1}{l} \sum\limits_{j=1}^{l} w_j$, также задает пороговую функцию f''.

Если $f''(b_1, \ldots, b_l) = 1$ на некотором наборе $(b_1, \ldots, b_l)$, то

$$\sum_{j=1}^{l} b_j \cdot w_j - \theta \geqslant 0.$$

Покажем, что

$$\sum_{j=1}^{l} b_j \cdot w' - \theta \geqslant 0.$$

Рассмотрим набор $(b'_1, \ldots, b'_l)$, такой что количество единиц в нем совпадает с количеством единиц в наборе $(b_1, \ldots, b_l)$ и, при этом, сумма

$$\sum_{j=1}^{l} b'_j \cdot w_i$$

принимает минимальное значение среди всех двоичных наборов с таким же числом единиц. Так как функция f'' — симметрическая, то $f''(b'_1, \ldots, b'_l) = 1$.

Очевидно, что

$$\sum_{j=1}^{l} b'_j \cdot w' \geqslant \sum_{j=1}^{l} b'_j \cdot w_i.$$

Следовательно, $\sum\limits_{j=1}^{l} b'_j \cdot w' - \theta \geqslant 0$.

Случай, когда $f''(b_1, \ldots, b_l) = 0$, рассматривается аналогичным образом, за тем лишь исключением, что набор $(b'_1, \ldots, b'_l)$ выбирается так, чтобы сумма

$$\sum_{j=1}^{l} b'_j \cdot w_i$$

принимала максимальное значение.

Таким образом, линейная форма

$$l_{\bar{w}',\sigma}(x_1,\ldots,x_n)=\Bigg(\underbrace{w',\ldots,w'}_{l},w_{l+1},\ldots,w_n,\sigma\Bigg)$$

задает исходную функцию f.

Для доказательства леммы осталось заметить, что описанная процедура осреднения коэффициентов линейной формы может быть проведена независимо для каждого класса эквивалентности R_i. Лемма доказана.

Пусть $M=\|m_{ij}\|_{n\times n}$ — обратимая матрица и $M^{-1}=\left\|m'_{ij}\right\|_{n\times n}$. Обозначим

$$\varphi(M)=\max_{i,j}\left|m'_{ij}\right|.$$

Приведем без доказательства следующий известный результат [12], который позволяет по обратимой матрице M построить $(-1,1)$-пороговую функцию, размах которой не менее, чем $\varphi(M)$.

Лемма 3.2. *Если* $M=\|m_{ij}\|_{n\times n}$ *— обратимая матрица и* $m_{ij}\in\{-1,1\}$, *то существует* $(-1,1)$*-пороговая функция* $f^{-1,1}$ *такая, что*

$$L\left(f^{-1,1}\right)\geqslant\varphi(M).$$

Обозначим $L_i\left(l_{\vec{w},\sigma}\right)=|w_i|$, где $i\in\{1,\ldots,n\}$. Введем следующую характеристику сложности задания пороговой функции линейными формами

$$L_i\left(f^{0,1}\right)=\min_{l_{\vec{w},\sigma}\to f^{0,1}}L_i\left(l_{\vec{w},\sigma}\right).$$

Здесь минимум берется по всем линейным формам, задающим $f^{0,1}$.

Аналогичная характеристика вводится для $(-1,1)$-пороговых функций

$$L_i\left(f^{-1,1}\right)=\min_{l_{\vec{w},\sigma}\to f^{-1,1}}L_i\left(l_{\vec{w},\sigma}\right).$$

Здесь минимум берется по всем линейным формам, задающим $f^{-1,1}$.

Имеет место следующее утверждение.

Лемма 3.3. *Если $f^{0,1}$ пороговая функция, то для всех $i \in \{1, \ldots, n\}$ выполнено*

$$L_i\left(f^{0,1}\right) \geqslant L_i\left(f^{-1,1}\right).$$

Доказательство. Предположим противное: для некоторого i выполнено

$$L_i\left(f^{0,1}\right) < L_i\left(f^{-1,1}\right).$$

Пусть $L_i\left(f^{0,1}\right)$ достигается на линейной форме $l_{\vec{w},\sigma}$. В таком случае по лемме 2.1 имеем

$$l_{\vec{w},2\sigma-\sum\limits_i w_i} \to f^{-1,1}.$$

Следовательно,

$$L_i\left(f^{0,1}\right) = L_i\left(l_{\vec{w},\sigma}\right) = L_i\left(l_{\vec{w},2\sigma-\sum\limits_i w_i}\right) \geqslant L_i\left(f^{-1,1}\right).$$

Противоречие, лемма доказана.

Сформулируем один известный результат [12].

Теорема 8. *Для достаточно больших n выполнено*

$$2^{\frac{1}{2}n\log_2 n-2n-o(n)} \leqslant \varphi(n) \leqslant 2^{\frac{1}{2}n\log_2 n-n-o(n)}.$$

Определим величину $L(m,k)$ следующим образом

$$L(m,k) = \max_{f\in T_{m,k}} L(f),$$

где максимум берется по всем пороговым функциям f из класса $T_{m,k}$. Содержательно данная величина характеризует величину размаха пороговой функций с максимальным размером класса симметрии — m и числом независимых классов симметрии — k.

Имеет место следующее утверждение, характеризующее величину $L(m,k)$.

Лемма 3.4. *Для достаточно больших k выполнено*

$$2^{\frac{1}{2}k\log k+o(k\log k)} \leqslant L(m,k) \leqslant m^{k+1}\cdot 2^{\frac{1}{2}k\log k+o(k\log k)}.$$

Доказательство. Докажем верхнюю оценку.

Заметим, что всякая существенная пороговая функция f от n переменных однозначно определяется своими значениями на наборах из E_2^n. Всего таких наборов имеется 2^n. Рассмотрим следующую систему из 2^n неравенств

$$\begin{cases} \underset{\alpha_i \in N_f}{\&} (\alpha_i \cdot \vec{w} - \sigma \geqslant 0), \\ \underset{\alpha_i \in E_2^n \setminus N_f}{\&} (\alpha_i \cdot \vec{w} - \sigma < 0). \end{cases} \tag{3.1}$$

Здесь $\alpha_i = (a_1, \ldots, a_n) \in E_2^n$, $i = 1, \ldots, 2^n$, и $N_f = \{\alpha \in E_2^n : f(\alpha) = 1\}$ — множество наборов, на которых функция f принимает значение 1.

Очевидно, что всякая линейная форма $(w_1, \ldots, w_n, \sigma)$, удовлетворяющая данной системе, будет задавать f.

Рассмотрим вектора $\alpha'_i, \alpha''_i \in E_2^{n+1}$ такие, что

$$\begin{cases} \alpha'_i = (\alpha_{i1}, \ldots, \alpha_{in}, 1) \text{ для всех } \alpha_i \in E_2^n, \\ \alpha''_i = \alpha'_i, \text{ если } \alpha_i \in N_f \text{ и} \\ \alpha''_i = -\alpha'_i, \text{ иначе.} \end{cases}$$

Рассмотрим следующую систему из 2^n неравенств

$$\underset{\alpha_i \in E_2^n}{\&} \left(\alpha''_i \cdot \vec{w}' \geqslant 1\right), \tag{3.2}$$

где $\vec{w}' = \left(w'_1, \ldots, w'_{n+1}\right)$.

Легко видеть, что всякий вектор $\vec{w}'$, являющийся решением системы (3.2), задает также решение системы (3.1). Для этого достаточно положить $\vec{w} = (w'_1, \ldots, w'_n)$, а $\sigma = w'_{n+1}$. Верно и обратное: если решение системы (3.1) существует, то существует также и решение системы (3.2). Это выполнено так как произвольная пороговая функция может быть задана целочисленной линейной формой строгим образом. Для обоснования этого достаточно заметить, что если целочисленная линейная форма $l_{\vec{w},\sigma}$ задает пороговую функцию f, то линейная форма $l_{2\cdot\vec{w},2\cdot\sigma+1}$ задает f строгим образом.

Поэтому переход к рассмотрению системы (3.2) является допустимым.

Множество всех решений системы (3.2) представляет собой выпуклый многогранник M в пространстве R^{n+1}. При этом, так как решения системы 3.2 существуют, то многогранник M не пуст. Покажем, что в M существует хотя бы одна экстремальная точка $\vec{w}'$, в которой некоторые $(n+1)$ неравенств системы (3.2) обращаются в строгие равенства.

Без ограничения общности положим, что функция $f(x_1, \ldots, x_n)$ — монотонная, а следовательно, многогранник M целиком содержится на множестве $K = R_+^{n+1}$, где R_+ — множество неотрицательных действительных чисел. Множество K будем называть квадрантом.

Очевидно, что всякая прямая l, проходящая через точки A и B, лежащие в квадранте K, пересекает хотя бы одну из его граней, задаваемых уравнениями вида $w_i' = 0$, $i = 1, \ldots, n+1$. Так как многогранник M целиком содержится в K, то всякая прямая l_1, проходящая через точки A_1 и B_1, лежащие в M, пересекает хотя бы одну из граней многогранника M, задаваемых неравенствами системы (3.2). Полученную грань обозначим G_1. Очевидно, что размерность грани G_1 строго меньше $n+1$. Если размерность грани G_1 равна 1, то в качестве экстремальной точки выберем точку пересечения l_1 и G_1. В противном случае построим прямую l_2, проходящую через некоторые точки A_2 и B_2, лежащие на G_1. Аналогичным образом приходим к выводу, что прямая l_2 пересекает хотя бы одну из граней G_2 многогранника M. При этом, размерность грани G_2 строго меньше размерности грани G_1. Продолжаем описанные построения до тех пор, пока размерность грани G_k не станет равна 1 и не будет получена экстремальная точка $\vec{w}'$.

По условию теоремы f является R-симметрической пороговой функцией, где $R = \{R_1, \ldots, R_k\}$ — некоторое разбиение множества $\Omega_n = \{1, \ldots, n\}$. При этом,

$$\max\left(|R_1|, \ldots, |R_k|\right) = m.$$

Без ограничения общности положим, что

$$\begin{cases} R_1 = \{1, \ldots, r_1\}, \\ \ldots \\ R_k = \{n - r_k + 1, \ldots, n\}, \end{cases}$$

где $r_i = |R_i|$, $i = 1, \ldots, k$.

В таком случае, по лемме 3.1 найдется линейная форма

$$\left(\underbrace{w_1'', \ldots, w_1''}_{r_1}, \ldots, \underbrace{w_k'', \ldots, w_k''}_{r_k}, \sigma \right),$$

где

$$w_i'' = \frac{1}{|R_i|} \sum_{j \in R_i} w_j,$$

также задающая f.

Следовательно, всякое решение системы

$$\begin{cases} \underset{\alpha_i \in E_2^n}{\&} (\alpha_i'' \cdot \bar{w}'' \geqslant 1); \\ w_i'' = \frac{1}{|R_i|} \sum\limits_{j \in R_i} w_j; \end{cases} \tag{3.3}$$

где $\bar{w}'' = \left(w_1'', \ldots, w_k'', w_{k+1}''\right)$, также задает решение системы (3.1). Для этого достаточно положить $\vec{w} = (w_1'', \ldots, w_k'')$, а $\sigma = -w_{k+1}''$. Отметим также, что по лемме 3.1, система (3.3) имеет хотя бы одно решение.

Заметим, что вектор $\bar{w}''$ содержит не более $(k+1)$ различных компонент. Следовательно, множество всех решений системы (3.3) представляет собой выпуклый многогранник M в пространстве $\mathbb{R}^{k+1}$. При этом, так как решения системы (3.3) существуют, то многогранник M не пуст. В экстремальных точках M некоторые $(k+1)$ неравенств обращаются в строгие равенства. Пусть $\left(w_1'', \ldots, w_{k+1}''\right)$ — некоторая экстремальная точка многогранника M. Как было показано ранее, такая точка существует. В таком случае набор $\bar{w}''$ удовлетворяет следующей системе уравнений

$$A \cdot \left(w_1'', \ldots, w_{k+1}''\right)^T = (1, \ldots, 1)^T,$$

где $A = \|a_{ij}\|_{(k+1)\times(k+1)}$ — матрица, коэффициенты a_{ij} которой по модулю не превосходят m — мощность максимального класса эквивалентности R_i. Заметим, что так как набор $\left(w''_1, \ldots, w''_{k+1}\right)$ задает экстремальную точку многогранника M, то данная система уравнений имеет единственное решение, а, следовательно, определитель матрицы A отличен от нуля. По правилу Крамера получим

$$w''_i = \frac{\Delta_i}{\Delta};\ i = 1, \ldots, k+1,$$

где Δ — определитель матрицы A, а Δ_i — определитель матрицы A_i, получаемой из A заменой i-го столбца на единичный вектор. Видно, что при делении на Δ компоненты w''_i могут получиться нецелыми, однако, легко видеть, что при домножении всех компонент линейной формы $\left(w''_1, \ldots, w''_{k+1}\right)$ на Δ задаваемая ей пороговая функция не изменится, а значения компонент w''_i станут целыми. Поэтому можно полагать, что

$$w''_i = \Delta_i,\ i = 1, \ldots, k+1.$$

Как уже отмечалось ранее, элементы матрицы A по модулю не превосходят m, поэтому по теореме Адамара [1] имеем

$$\Delta_i^2 \leqslant \prod_{i=1}^{k+1} \sum_{j=1}^{k+1} a_{ij}^2.$$

Следовательно,

$$\Delta_i^2 \leqslant \left((k+1)\cdot m^2\right)^{k+1} = (k+1)^{k+1} \cdot m^{2(k+1)}.$$

В итоге, получим

$$w''_i \leqslant m^{k+1} \cdot 2^{\frac{1}{2} k \log k + o(k \log k)}.$$

Верхняя оценка доказана, докажем нижнюю оценку.

Из теоремы 8 следует, что для достаточно больших k найдется такая пороговая функция $g \in T^k$, что

$$L(g) \geqslant 2^{\frac{1}{2} k \log k + o(k \log k)}.$$

Рассмотрим R-симметрическую пороговую функцию $f \in T^n$, такую, что $n = k \cdot m$,

$$\begin{cases} R_1 = \{1, \ldots, m\}, \\ \ldots \\ R_k = \{(k-1) \cdot m + 1, \ldots, k \cdot m\}, \end{cases}$$

и

$$f\left(\underbrace{x_1, 0, \ldots, 0}_{|R_1|}, \ldots, \underbrace{x_k, 0, \ldots, 0}_{|R_k|}\right) = g(x_1, \ldots, x_k).$$

Очевидно, что $L(f) \geqslant L(g)$. Утверждение доказано.

Из леммы 3.2 вытекает важное следствие, которое позволяет по обратимой матрице M построить $(-1,1)$-пороговые функции, удаленные друг от друга на расстрояние не менее, чем $\left[\frac{\varphi(M)+1}{2}\right]$.

Лемма 3.5. *Если $M = \|m_{ij}\|_{n \times n}$ — обратимая матрица и $m_{ij} \in \{-1, 1\}$, то существуют $(-1,1)$-пороговые функции $f_1^{-1,1}$ и $f_2^{-1,1}$ такие, что*

$$\rho\left(f_1^{-1,1}, f_2^{-1,1}\right) \geqslant \left[\frac{\varphi(M)+1}{2}\right].$$

Доказательство. Выберем в качестве функции $f_1^{-1,1}$ пороговую функцию, существование которой доказано в лемме 3.2. Заметим, что если линейная форма $l_{\vec{w}',\sigma'}$ задает $f_1^{-1,1}$ не строгим образом, то

$$w'_1 \geqslant \left[\frac{\varphi(M)+1}{2}\right].$$

Это выполнено так как произвольная пороговая функция может быть задана целочисленной линейной формой строгим образом. Для обоснования этого достаточно заметить, что если целочисленная линейная форма $l_{\vec{w},\sigma}$ задает пороговую функцию f, то линейная форма $l_{2\cdot\vec{w},2\cdot\sigma+1}$ задает f строгим образом.

В качестве функции $f_2^{-1,1}$ выберем функцию $f_1^{-1,1}$, у которой все переменные домножены на -1. В таком случае, по лемме 2.5, выполнено

$$\rho\left(f_1^{-1,1}, f_2^{-1,1}\right) \geqslant \left[\frac{\varphi(M)+1}{2}\right].$$

Утверждение доказано.

Из леммы 3.3 вытекает аналог леммы 3.5 для $(0,1)$-пороговых функций.

Лемма 3.6. *Если* $M = \|m_{ij}\|_{n\times n}$ *— обратимая матрица и* $m_{ij} \in \{-1,1\}$*, то существуют* $(0,1)$*-пороговые функции* $f_1^{0,1}$ *и* $f_2^{0,1}$*, такие что*

$$\rho\left(f_1^{0,1}, f_2^{0,1}\right) \geqslant \varphi\left(M\right).$$

Докажем утверждение теоремы 1.

Для доказательства верхней оценки достаточно заметить, что $n \leqslant k \cdot m$,

$$\rho\left(m,k\right) \leqslant 2\cdot\left(n+1\right)\cdot L\left(m,k\right)$$

и воспользоваться леммой 3.4.

Таким образом,

$$\rho\left(m,k\right) \leqslant 2\cdot\left(m\cdot k+1\right)\cdot m^{k+1}\cdot 2^{\frac{1}{2}k\log k+o(k\log k)}.$$

Раскрыв скобки, получим

$$\rho\left(m,k\right) \leqslant m^{k+2}\cdot 2^{\frac{1}{2}k\log k+o(k\log k)}.$$

Докажем нижнюю оценку.

Из леммы 3.6 и теоремы 8 следует, что для достаточно больших k найдется такая пара пороговых функций $g', g'' \in T^k$, что

$$\rho\left(g', g''\right) \geqslant 2^{\frac{1}{2}k\log k+o(k\log k)}.$$

Рассмотрим R-симметрические пороговые функции $g', g'' \in T^n$, где $n = k\cdot m$,

$$R_1 = \left\{1,\ldots,m\right\},$$
$$\ldots$$
$$R_k = \left\{(k-1)\cdot m+1,\ldots,k\cdot m\right\},$$

такие, что

$$f'\left(\underbrace{x_1, 0, \ldots, 0}_{|R_1|}, \ldots, \underbrace{x_k, 0, \ldots, 0}_{|R_k|}\right) = g'(x_1, \ldots, x_k),$$

$$f''\left(\underbrace{x_1, 0, \ldots, 0}_{|R_1|}, \ldots, \underbrace{x_k, 0, \ldots, 0}_{|R_k|}\right) = g''(x_1, \ldots, x_k).$$

Очевидно, что $\rho(f', f'') \geqslant m \cdot \rho(g', g'')$. Следовательно,

$$\rho(m, k) \geqslant m \cdot 2^{\frac{1}{2}k\log k + o(k\log k)}.$$

Теорема доказана.

4. Обучение в большинстве случаев

4.1. Основные понятия, постановки и результаты

Следующие утверждения, характеризуют размах и вес почти всех пороговых функций от n переменных.

Теорема 9. *Если $n \to \infty$, $c > 1$, то для почти всех пороговых функций $f \in T^n$ выполнено*

$$2^{n - c\cdot\log n} \leqslant L(f) \leqslant 2^{\frac{1}{2}n\log n + o(n\log n)}.$$

Теорема 10. *Если $n \to \infty$, то для почти всех пороговых функций $f \in T^n$ выполнено*

$$2^{n + o(n)} \leqslant \mu(f) \leqslant 2^{\frac{1}{2}n\log n + o(n)}.$$

Следующая теорема характеризует близость $\rho(l, f)$ для почти всех пар (l, f) при $n \to \infty$.

Теорема 2. *2 Если $n \to \infty$, то для почти всех пар (l, f), где f — пороговая фукнция, а l — линейная форма от n переменных такая, что $L(l) \leqslant L(n)$, выполнено*

$$2^{n + o(n)} \leqslant \rho(l, f) \leqslant 2^{\frac{1}{2}n\log n + o(n\log n)}.$$

4.2. Доказательства утверждений

Докажем теорему 9.

Ранее в разделе 2 была введена величина $\rho\,(n)$, характеризующая близость между наиболее удаленными пороговыми функциями от n переменных.

Легко видеть, что $L\,(n) < \rho\,(n)$, поэтому верхняя оценка теоремы 9 вытекает из следствия 2 теоремы 1. Таким образом, для доказательства теоремы 9 достаточно показать нижнюю оценку.

Приведем без доказательства следующий известный результат [20].

Теорема 11. *Если $n \to \infty$, то*

$$|T^n| = 2^{n^2 - n\log n + o(n\log n)}.$$

Доказательство теоремы 9. Очевидно, что существует ровно k^{n+1} наборов $(w_1, \ldots, w_n, \sigma)$ таких, что $0 \leqslant w_i < k$ для всех $i = 1, \ldots, n$ и $0 \leqslant \sigma < k$. Обозначим $r\,(n, k) = k^{n+1}$. Определим наибольшее значение k при фиксированном n такое, что

$$r\,(n, k) = o\,(|MT^n|)\,,$$

где MT^n — множество монотонных пороговых функций от n переменных.

Очевидно, что при таком k для почти всех монотонных пороговых функций будет выполнено

$$L\,(f) > k.$$

По лемме 2.5 данная оценка будет верна и для произвольной пороговой функции. Оценим мощность множества MT^n. По лемме 2.5 она может быть оценена следующим образом

$$|MT^n| = \frac{|T^n|}{2^n}.$$

Тогда по теореме 11 имеем

$$|MT^n| = 2^{n^2 - n\log n + o(n\log n)}.$$

Следовательно, задача оценки k может быть сформулирована так: найти наибольшее k такое, что

$$\lim_{n\to\infty}\frac{2^{n^2-n\log n+o(n\log n)}}{k^{n+1}}=\infty.$$

Упростив выражение, получим

$$\lim_{n\to\infty}2^{n^2-n\log n-(n+1)\log k+o(n\log n)}=\infty.$$

$$\lim_{n\to\infty}\left(n^2-n\log n-(n+1)\log k+o\left(n\log n\right)\right)=\infty.$$

Пусть $k=2^{n-\varepsilon(n)}$, тогда

$$\lim_{n\to\infty}\left(n^2-n\log n-(n+1)\left(n-\varepsilon\left(n\right)\right)+o\left(n\log n\right)\right)=\infty.$$

$$\lim_{n\to\infty}\left((n+1)\cdot\varepsilon\left(n\right)-n\log n+o\left(n\log n\right)\right)=\infty.$$

Пусть $\varepsilon\left(n\right)=\left(1+c'\right)\log n$, где $c'>0$, тогда

$$\lim_{n\to\infty}\left(c'\cdot(n+1)\log n+o\left(n\log n\right)\right)=\infty.$$

Это верно при любом $c'>0$. Следовательно, $k\geqslant 2^{n-c\log n}$, где $c>1$. Утверждение доказано.

Докажем теорему 10.

Легко видеть, что $\mu\left(n\right)<\rho\left(n\right)$, поэтому верхняя оценка теоремы 10 вытекает из следствия 2 теоремы 1. Таким образом, для доказательства теоремы 10 достаточно показать нижнюю оценку.

Рассмотрим следующее обозначение

$$x^{\underline{m}}=x\cdot(x-1)\cdot\ldots\cdot(x-m+1)\,,$$

где m — целое число. В таком случае говорим, что величина $x^{\underline{m}}$ равна «x в убывающей степени m». Очевидно, что $x^{\underline{m}}=x^m+O\left(x^{m-1}\right)$. Приведем без доказательства два известных результата.

Лемма 4.1 ([2]). *Если n и m целые, то*

$$\sum_{x=0}^{n-1}x^{\underline{m}}=\frac{1}{m+1}\cdot n^{\underline{m+1}}.$$

Лемма 4.2 ([2]). *Если n целое, то*

$$x^n = \sum_{k=0}^{n} \left\{ \begin{array}{c} n \\ k \end{array} \right\} x^{\underline{k}},$$

где $\left\{ \begin{array}{c} n \\ k \end{array} \right\}$ *— числа Стирлинга второго рода.*

Имеет место следующее утверждение.

Лемма 4.3. *Если* $p \in \mathbb{N}$, *то*

$$\sum_{x=0}^{n-1} x^m = \frac{1}{m+1} \cdot n^{m+1} + O\left(n^m\right).$$

Доказательство. По лемме 4.2 выполнено

$$\sum_{x=0}^{n-1} x^m = \sum_{x=0}^{n-1} \sum_{k=0}^{m} \left\{ \begin{array}{c} m \\ k \end{array} \right\} x^{\underline{k}} = \sum_{k=0}^{m} \sum_{x=0}^{n-1} \left\{ \begin{array}{c} m \\ k \end{array} \right\} x^{\underline{k}} = \sum_{k=0}^{m} \left\{ \begin{array}{c} m \\ k \end{array} \right\} \sum_{x=0}^{n-1} x^{\underline{k}}.$$

По лемме 4.1 получим

$$\sum_{x=0}^{n-1} x^m = \sum_{k=0}^{m} \left\{ \begin{array}{c} m \\ k \end{array} \right\} \cdot \frac{1}{k+1} \cdot n^{\underline{k+1}} =$$

$$= \sum_{k=0}^{m} \left[\left\{ \begin{array}{c} m \\ k \end{array} \right\} \cdot \frac{1}{k+1} \cdot n^{k+1} + O\left(n^k\right) \right] =$$

$$= \left\{ \begin{array}{c} m \\ m \end{array} \right\} \cdot \frac{1}{m+1} \cdot n^{m+1} + O\left(n^m\right).$$

Так как $\left\{ \begin{array}{c} m \\ m \end{array} \right\} = 1$, то

$$\sum_{x=0}^{n-1} x^m = \frac{1}{m+1} \cdot n^{m+1} + O\left(n^m\right).$$

Лемма доказана.

Рассмотрим набор $l = (w_1, \ldots, w_n)$, где w_i — натуральные числа или ноль. Рассмотрим множество

$$R_{n,k} = \left\{ l,\ \sum_{i=1}^{n} w_i < k \right\},$$

где k — натуральное число.

Данное множество состоит из всех наборов l, сумма компонент которых меньше k.

Обозначим $R\,(n, k)$ мощность множества $R_{n,k}$.

Докажем вспомогательный результат, характеризующий величину $R\,(n, k)$.

Лемма 4.4. *Если $n \to \infty$, то*

$$R\,(n, k) = \frac{1}{n!}\left(k^n + O\left(k^{n-1}\right)\right).$$

Доказательство. Легко видеть, что

$$R\,(2, k) = \frac{k^2 + k}{2} = \frac{1}{2}\left(k^2 + k\right).$$

Также отметим, что

$$R\,(n+1, k) = \sum_{i=1}^{k} R\,(n, i).$$

Таким образом, для того чтобы доказать утверждение леммы достаточно показать, что

$$\sum_{i=1}^{k} i^p = \frac{1}{p+1} \cdot k^{p+1} + O\left(k^p\right).$$

А это верно по лемме 4.3. Лемма доказана.

Доказательство теоремы 10. Нетрудно видеть, что количество монотонных пороговых функций от $n-1$ переменного, для которых $\mu\,(f) < k$, не превосходит величину $R\,(n, k)$.

По лемме 2.5 количество всех пороговых функций от $n-1$ переменного, для которых $\mu(f) < k$, не превосходит величину $2^{n-1} \cdot R(n,k)$.

Далее, если при некотором k выполнено

$$2^{n-1} \cdot R(n,k) = o\left(T^{n-1}\right),$$

то можно утверждать, что для почти всех пороговых функций выполнено $\mu(f) \geqslant k$. Таким образом, необходимо найти наибольшее k, при котором выполнено данное асимптотическое равенство. По лемме 4.4 имеем

$$R(n,k) = \frac{2^{n \log k} + O\left(k^{n-1}\right)}{2^{n \log n - n \log e + o(n)}} \sim 2^{n \log k - n \log n + n \log e + o(n)}.$$

Теперь найдем наибольшее k, при котором выполнено

$$\lim_{n \to \infty} \frac{2^{n^2 - n \log n + O(n)}}{2^{n \log k - n \log n + O(n)}} = \infty,$$

или, что то же самое,

$$\lim_{n \to \infty} \left(n^2 - n \log k + O(n)\right) = \infty.$$

Пусть $k = 2^{n-g(n)}$, тогда

$$\lim_{n \to \infty} (n \cdot g(n) + O(n)) = \infty.$$

Данное выражение верно при всех $g(n) = o(n)$ таких, что $g(n) \to \infty$ при $n \to \infty$. Теорема доказана.

Докажем утверждение теоремы 2.

Верхняя оценка теоремы 2 очевидно вытекает из следствия 2 теоремы 1. Таким образом, достаточно показать нижнюю оценку.

Зафиксируем линейную форму l, зависящую от n переменных. Определим число $p(n)$ пороговых функций от $n-1$ переменного, близость которых до линейной формы l не превосходит величины $k = 2^{n - \log n}$.

По лемме 2.5 количество всех пороговых функций от $n-1$ переменного, для которых $\rho(l, f) < k$, не превосходит величину $2^{n-1} \cdot R(n, k)$.

Следовательно, $p(n) \leqslant 2^{n-1} \cdot R\left(n, 2^{n-\log n}\right)$. По лемме 4.4 имеем

$$p(n) \lesssim 2^{n(n-\log n)-n\log n+o(n\log n)} = 2^{n^2-2n\log n+O(n)}.$$

По теореме 11 при $n \to \infty$ следует, что $p(n) = o\left(T^{n-1}\right)$, так как

$$\lim_{n\to\infty} \frac{\left|T^{n-1}\right|}{p(n)} = 2^{n\log n+o(n\log n)} = \infty.$$

Пробегая по конечному множеству линейных форм l таких, что $L(l) \leqslant L(n)$, получаем, что для почти всех пар (l, f) близость $\rho(l, f)$ не менее, чем $2^{n-\log n}$. Теорема доказана.

5. Модификации понятия близости

5.1. Основные понятия, постановки и результаты

Рассмотрим M — множество операций над целыми числами. Говорим, что линейная форма $l_{\vec{w}',\sigma'}$ является M-переводимой в линейную форму $l_{\vec{w}'',\sigma''}$, если она может быть преобразована к линейной форме $l_{\vec{w}'',\sigma''}$ при помощи одного применения операции из M к одному из весовых коэффициентов вектора $\vec{w}'$, или порогу σ'.

Пусть $l_{\vec{w}',\sigma'}$ и $l_{\vec{w}'',\sigma''}$ — линейные формы от n переменных. Близостью $\rho_M(l_{\vec{w}',\sigma'}, l_{\vec{w}'',\sigma''})$, между данными линейными формами назовем длину кратчайшей последовательности $l_1, \ldots, l_r$, где $l_1 = l_{\vec{w}',\sigma'}$, $l_r = l_{\vec{w}'',\sigma''}$ и для всякого $j = 1, \ldots, r-1$, линейная форма l_j является M-переводимой к линейной форме l_{j+1}. Полагаем, что длина последовательности, состоящей из единственного элемента, равна нулю.

Пусть f', f'' — пороговые функции от n переменных. Близостью $\rho_M(f', f'')$, между ними назовем величину

$$\rho_M\left(f', f''\right) = \min_{\substack{l'\to f' \\ l''\to f''}} \rho_i\left(l', l''\right),$$

где минимум берется по всем линейным формам $l_{\vec{w}',\sigma'}$ и $l_{\vec{w}'',\sigma''}$, задающим функции f' и f'', соответственно.

Аналогичным образом вводится функция близости между линейными формами и пороговыми функциями $\rho_M(l, f)$.

Естественным образом вводится шенноновская функция $\rho_M(n)$, характеризующая близость ρ_M между наиболее удаленными пороговыми функциями от n переменных.

Также вводится функция $\rho_M(m, k)$, характеризующая близость между наиболее удаленными пороговыми функциями, не более чем с m взаимно симметричными переменными, и числом классов симметрии — k.

Рассмотрим следующие множества операций над целыми числами

$$M_I = \{+1, -1\};$$
$$M_{II} = \{+1, -1, \cdot(-1)\}.$$

Здесь «± 1» — операция увеличения или уменьшения числа на единицу, «$\cdot(-1)$» — умножение числа на -1.

Рассмотрим функцию $\varphi : Z \to Z$. Будем говорить, мощность функции φ ограничена величиной a, где a — натуральное, если для всех x выполнено $\left[\frac{x}{a}\right] \leqslant \varphi(x) \leqslant a \cdot x$. Множество всех функций, мощность которых ограничена величиной $a \in N$, обозначим R_a. Обозначим $M_a = M_{II} \cup R_a$. Легко видеть, что если $a = 1$, то $M_a = M_{II}$, поэтому далее будем рассматривать случай, когда $a > 1$.

С множествами допустимых операций M_I, M_{II}, M_a связаны соответствующие им функции близости ρ_I, ρ_{II} и ρ_a. Легко видеть, что введенная в главе 2 функция близости ρ в точности равна близости ρ_I. Поэтому следствие 2 теоремы 1, теоремы 2 и 1 характеризуют близость ρ_I в самом сложном случае, для почти всех случаев и для классов пороговых функций, инвариантных относительно групп перестановок.

Нетрудно видеть, что $\rho_I(f', f'') \geqslant \rho_{II}(f', f'') \geqslant \rho_a(f', f'')$.

Имеют место следующие утверждения, характеризующие функции ρ_{II} и ρ_a в самом сложном случае, для почти всех случаев и для классов пороговых функций, инвариантных относительно групп перестановок.

Теорема 3. *3 Если $n \to \infty$, то*

$$\log \rho_{II}(n) \sim \frac{1}{2} n \log n;$$
$$\rho_a(n) \asymp n^2 \log n.$$

Теорема 4. *4 Если $n \to \infty$, то для почти всех пар (l, f), где f — пороговая фукнция, а l — линейная форма от n переменных такая, что $L(l) \leqslant L(n)$, выполнено*

$$2^{n+o(n)} \leqslant \rho_{II}(l, f) \leqslant 2^{\frac{1}{2} n \log n + o(n \log n)};$$
$$n \lesssim \rho_a(l, f) \lesssim n^2 \log n.$$

Теорема 5. *5 а) Для достаточно больших k выполнено*

$$m \cdot 2^{\frac{1}{4} k \log k + o(k \log k)} \leqslant \rho_{II}(m, k) \leqslant m^{k+2} \cdot 2^{\frac{1}{2} k \log k + o(k \log k)}.$$

б) Если $k \to \infty$ и $\log(m) = o(\log(k))$, то

$$mk \log k \lesssim \rho_a(m, k) \lesssim mk^2 \log k.$$

5.2. Доказательства утверждений

Для доказательства теоремы 3 нам понадобятся некоторые вспомогательные утверждения.

Ранее, в разделе 2, было введено понятие линейной формы минимального веса. Имеет место следующее свойство таких линейных форм.

Лемма 5.1. *Если $l_{\vec{w},\sigma}$ — линейная форма минимального веса, $\vec{w} = (w_1, \ldots, w_n)$, то линейная форма $l_{\vec{w}',\sigma}$, где $\vec{w}' = (w_1, \ldots, w_n, 2)$, также является линейной формой минимального веса.*

Доказательство. Предположим противное: найдется линейная форма $l_{\vec{w}'',\sigma''} \sim l_{\vec{w}',\sigma}$ такая, что $\mu\left(l_{\vec{w}'',\sigma''}\right) < \mu\left(l_{\vec{w}',\sigma}\right)$.

Обозначим $f(x_1, \ldots, x_n)$ и $f'(x_1, \ldots, x_n)$ — пороговые функции, задаваемые линейными формами $l_{\vec{w},\sigma}$ и $l_{\vec{w}',\sigma}$, соответственно. Легко видеть, что $f'(x_1, \ldots, x_n, 0) = f(x_1, \ldots, x_n)$. Таким образом, так как линейная форма $l_{\vec{w}'',\sigma''}$ задает ту же пороговую функцию, что и $l_{\vec{w}',\sigma}$, то

$$\sum_{i=1}^{n} \left|w_i''\right| + \left|\sigma''\right| \geqslant \sum_{i=1}^{n} |w_i| + |\sigma| \,.$$

В противном случае линейная форма $(w_1'', \ldots, w_n'', \sigma'')$ задавала бы функцию f и, при этом, обладала бы меньшим весом, чем $l_{\vec{w},\sigma}$. Что противоречит условию минимальности веса $l_{\vec{w},\sigma}$.

С другой стороны, так как $\mu\left(l_{\vec{w}'',\sigma''}\right) < \mu\left(l_{\vec{w}',\sigma}\right)$, то

$$\sum_{i=1}^{n+1} \left|w_i''\right| + \left|\sigma''\right| < \sum_{i=1}^{n} |w_i| + |\sigma| + 2.$$

Следовательно, либо $w_{n+1}'' = 0$, либо $w_{n+1}'' = 1$. Рассмотрим оба случая.

Если $w_{n+1}'' = 0$, то функция f' не зависит существенным образом от переменной x_{n+1}. В таком случае, так как $l_{\vec{w}',\sigma}$ задает f' строгим образом, то для всякого набора $(x_1, \ldots, x_n, 0) \in E_2^{n+1}$, на котором f' обращается в ноль, было бы выполнено $\sum\limits_{i=1}^{n} x_i w_i + \sigma \leqslant -3$. В противном случае подстановка единицы вместо x_{n+1} могла бы изменить значение функции f', что противоречит несущественности переменной x_{n+1}. В таком случае линейная форма $l_{\vec{w},\sigma-2}$ также задавала бы функцию f и, при этом, имела бы меньший вес, чем линейная форма $l_{\vec{w},\sigma}$. Пришли к противоречию минимальности веса $l_{\vec{w},\sigma}$.

Если $w_{n+1}'' = 1$, то аналогичным образом приходим к выводу, что для всякого набора $(x_1, \ldots, x_n, 0) \in E_2^{n+1}$, на котором f' обращается в ноль, было бы выполнено $\sum\limits_{i=1}^{n} x_i w_i + \sigma \leqslant -2$. В противном случае подстановка единицы вместо x_{n+1} могла бы обращать сумму $\sum\limits_{i=1}^{n+1} x_i w_i + \sigma$ в ноль, что противоречит заданию функции f' строгим образом. В таком случае линейная форма $l_{\vec{w},\sigma-1}$ также задает функцию f и, при этом, имеет меньший вес, чем линейная форма $l_{\vec{w},\sigma}$. Пришли к противоречию. Лемма доказана.

Рассмотрим некоторое множество операций над целыми числами $M = \{\varphi_1, \ldots, \varphi_k\}$. Пусть a, b — целые числа. Говорим, что a является M-переводимым в b, если $b = \varphi_j(a)$ для некоторого $j \in \{1, \ldots, k\}$. Близостью $\rho_M(a, b)$, между данными числами назовем длину кратчайшей последовательности $a_1, \ldots, a_r$, где $a_1 = a$, $a_r = b$ и для всякого $j = 1, \ldots, r-1$, число a_j является M-переводимым в a_{j+1}. Полагаем, что длина последовательности, состоящей из единственного элемента, равна нулю.

Имеет место следующая лемма, характеризующая близость ρ_a между целыми числами.

Лемма 5.2. *Если p, q — целые числа, отличные от нуля, $|p| < |q|$, то*

$$\left|[\log_a |p|] - [\log_a |q|]\right| \leqslant \rho_a(p, q) \leqslant [\log_a |p|] + [\log_a |q|] + 3.$$

Доказательство. Заметим, что a-ичная запись модуля целого числа p содержит $[\log_a |p|] + 1$ бит. Операция умножения на a эквивалентна операции единичного сдвига a-ичной записи числа в сторону старших разрядов. Операция прибавления единицы позволяет изменить младший разряд a-ичной записи. Таким образом, для преобразования числа 0 в $|p|$ достаточно выполнить не более $[\log_a |p|] + 2$ операций из множества M_a. Следовательно,

$$[\log_a |p|] + 1 \leqslant \rho_a(0, p) \leqslant [\log_a |p|] + 2.$$

Также легко видеть, что

$$\rho_a(p, 0) = [\log_a |p|] + 1.$$

Близость $\rho_a(p, q)$ может быть оценена так

$$\left|[\log_a |p|] - [\log_a |q|]\right| \leqslant \rho_a(p, q) \leqslant \rho_a(q, 0) + \rho_a(0, p).$$

Нижняя оценка вытекает из следующих соображений. При перестройке числа p в q число бит a-ичного представления p должно сравняться с числом бит a-ичного представления q. При этом, заметим, что за одно применение любой операции из M_a число бит в a-ичном представлении не может измениться более, чем на единицу.

Используя полученные выше оценки для $\rho_a(p,0)$ и $\rho_a(0,p)$, имеем

$$|[\log_a|p|]-[\log_a|q|]|\leqslant\rho_a(p,q)\leqslant[\log_a|p|]+[\log_a|q|]+3.$$

Лемма доказана.

Булеву функцию $f(x_1,\ldots,x_n)$ называют *самодвойственной*, если

$$f(x_1,\ldots,x_n)=\bar f(\bar x_1,\ldots,\bar x_n).$$

Если f — самодвойственная пороговая функция, то соответствующую ей $(-1,1)$-пороговую функцию $f^{-1,1}$ будем называть самодвойственной $(-1,1)$-пороговой функцией.

Ранее, в разделе 2 было введено понятие размаха пороговой функций — $L(f)$. Данная величина характеризует величину весовых коэффициентов линейной формы, задающей функцию f.

Характеристика $L(f)$ естественным образом обобщается на $(-1,1)$-пороговые функции. Напомним известный результат [16], характеризующий величину $L(f)$ для самодвойственных $(-1,1)$-пороговых функций.

Теорема 12. *Если $n=2^k$, $k\in N$, $k\geqslant 3$, то найдется существенная самодвойственная пороговая функция $f^{-1,1}$ такая, что для всех $i\in\{2,\ldots,n\}$ имеет место*

$$L_i\left(f^{-1,1}\right)\geqslant L_1\left(f^{-1,1}\right)\cdot\frac{1}{2n}e^{-4n^{\log_2\left(\frac{3}{2}\right)}}2^{\frac{1}{2}n\log_2 n-n}.$$

Доказательство теоремы 3.

Докажем первое утверждение. Заметим, что по лемме 3.6 и теореме 8 найдется линейная форма минимального веса $l_{\vec w,\sigma}$ такая, что

$$\mu\left(l_{\vec w,\sigma}\right)\geqslant 2^{\frac{1}{2}n\log n+o(n\log n)}.$$

В таком случае, по лемме 5.1, линейная форма $l_{\vec w',\sigma}$, где $\vec w'=(w_1,\ldots,w_n,2)$, также является линейной формой минимального веса. Пороговую функцию, задаваемую линейной формой $l_{\vec w',\sigma}$, обозначим f'.

Так как
$$\mu\left(l_{\vec{w},\sigma}\right)=\sum_{i=1}^{n}|w_i|+|\sigma|,$$
то среди весовых коэффициентов $w_1,\ldots,w_n$ и порога σ найдется такой элемент r, что
$$|r|\geqslant 2^{\frac{1}{2}n\log n+o(n\log n)}.$$

Без ограничения общности можно полагать, что r есть некоторый весовой коэффициент w_i, $i\in\{1,\ldots,n\}$. В противном случае, если $r=\sigma$, то в качестве линейной формы $l_{\vec{w},\sigma}$ рассмотрим линейную форму $(w_1,\ldots,w_n,w_{n+1},\sigma)$, получающуюся из нее добавлением еще одного весового коэффициента w_{n+1}. При этом, положим, что задаваемая ей пороговая функция f на наборе $(0,\ldots,0,1)$ принимает значение 1. В таком случае получим, что $w_{n+1}>\sigma$. Следовательно, r — есть некоторых весовой коэффициент.

Таким образом, получили линейную форму минимального веса такую, что один из ее весовых коэффициентов — w_i, является «очень большим», а другой — w_{n+1}, равен двум.

Рассмотрим пороговую функцию f'', получаемую из f' перестановкой переменных x_i и x_{n+1}. Иными словами,
$$f'(x_1,\ldots,x_i,\ldots,x_{n+1})=f''(x_1,\ldots,x_{n+1},\ldots,x_i).$$

Легко видеть, что
$$\rho_{II}\left(f',f''\right)\geqslant 2^{\frac{1}{2}n\log n+o(n\log n)}.$$

Так как $\rho_{II}\left(f',f''\right)\leqslant\rho_{I}\left(f',f''\right)$ для всех пороговых функций f',f'', то верхняя оценка для $\rho_{II}(n)$ вытекает из следствия 2 теоремы 1. Первое утверждение теоремы доказано.

Докажем второе утверждение теоремы. Заметим, что по теореме 12 и лемме 5.1 найдется линейная форма минимального веса $l_{\vec{w},\sigma}$ такая, что
$$\begin{cases} w_1\geqslant 2;\\ |w_i|\geqslant w_1\cdot 2^{\frac{1}{4}n\log n+o(n\log n)}, \text{ если } i\in\left\{2,\ldots,\left[\frac{n}{2}\right]\right\};\\ w_i=2, \text{ иначе.}\end{cases}$$

Пороговую функцию, задаваемую линейной формой $l_{\vec{w},\sigma}$, обозначим f'. Рассмотрим пороговую функцию f'' такую, что

$$f''(x_{[\frac{n}{2}]+1}, \ldots, x_n, x_1, \ldots, x_{[\frac{n}{2}]}) = f'(x_1, \ldots, x_{[\frac{n}{2}]}, x_{[\frac{n}{2}]+1}, \ldots, x_n).$$

Иными словами, функция f'' получается из f' перестановкой первой и второй половин переменных. Таким образом, для всех линейных форм $l_{\vec{w}',\sigma'}$ и $l_{\vec{w}'',\sigma''}$, задающих функции f' и f'', выполнено

$$\left|\left[\log_a |w_i'|\right] - \left[\log_a |w_i''|\right]\right| \geqslant \frac{1}{4\log a} n \log n + o(n \log n),$$

для всех $i \in \{2, \ldots, n\}$.

Следовательно,

$$\rho_a\left(f', f''\right) \geqslant (n-1) \cdot \left(\frac{1}{4\log a} n \log n + o(n\log n)\right) = \\ = \frac{1}{4\log a} n^2 \log n + o(n^2 \log n).$$

Докажем верхнюю оценку. Ранее, в разделе 3, была введена шенноновская функция $L(n)$, характеризующая максимальную величину размаха пороговой функций от n переменных. Там же была получена следующая оценка

$$\log L(n) \leqslant \frac{1}{2} n \log n + o\left(n \log n\right).$$

Легко видеть, что если f', f'' — пороговые функции от n переменных, то по лемме 5.2 выполнено

$$\rho_a\left(f', f''\right) \leqslant 2n \log_a L(n) + 3n.$$

В итоге получим

$$\frac{1}{4\log a} n^2 \log n + o(n^2 \log n) \leqslant \rho_a\left(f', f''\right) \leqslant \frac{1}{\log a} n^2 \log n + o(n^2 \log n).$$

Теорема доказана.

Обозначим $\widetilde{T}^n$ — множество существенных пороговых функций от n переменных.

Оказывается почти все пороговые функции существенные.

Лемма 5.3. *Если $n \to \infty$, то*

$$\left|\tilde{T}^n\right| \sim |T^n|.$$

Доказательство. Обозначим $\hat{T}^n = T^n \backslash \tilde{T}^n$ множество пороговых функций, не являющихся существенными. Утверждение теоремы эквивалентно тому, что при $n \to \infty$ выполнено

$$\left|\hat{T}^n\right| = o\left(|T^n|\right).$$

Покажем, что всякой существенной пороговой функции f от $(n-1)$ переменных можно поставить в соответствие множество $E(f)$ из 2^{n-1} функции от n переменного, такое, что если $f_1 \neq f_2$, то $E(f_1) \cap E(f_2) = \varnothing$. Отсюда будет следовать, что мощность множества T^n по крайней мере в $\frac{2^{n-1}}{n}$ раз больше, чем мощность множества $\hat{T}^n$, а следовательно, так как

$$\frac{\left|\hat{T}^n\right|}{|T^n|} < \frac{n}{2^{n-1}},$$

то

$$\lim_{n \to \infty} \frac{\left|\hat{T}^n\right|}{|T^n|} = 0.$$

Итак, рассмотрим существенную пороговую функцию f от $(n-1)$ переменного. В связи с леммой 2.5 можно без ограничения общности считать, что f — монотонная. Пусть $l_{\vec{w},\sigma}$ задает f строгим образом.

Обозначим $R(\vec{w},\sigma) = \left\{l_{\vec{w},\sigma}(\alpha),\ \alpha \in E_2^{n-1}\right\}$ — множество всех значений, которые принимает линейная форма $l_{\vec{w},\sigma}$ на наборах из E_2^{n-1}. Из леммы 2.3 следует, что найдется такая линейная форма $l_{\vec{w},\sigma}$, задающая f, что множество $R(\vec{w},\sigma)$ состоит из 2^{n-1} элемента. Иными словами на всех наборах из E_2^{n-1} линейная форма $l_{\vec{w},\sigma}$ принимает различные значения. Пронумеруем элементы множества $R(\vec{w},\sigma)$ по возрастанию: $R(\vec{w},\sigma) = \{r_1, \ldots, r_{2^{n-1}}\}$. Соответствующую нумерацию введем на множестве наборов E_2^{n-1}, то есть $l_{\vec{w},\sigma}(\alpha_i) = r_i$ для всех $i = 1, \ldots, 2^{n-1}$.

Рассмотрим последовательность пороговых функций $f_1, \ldots, f_{2^{n-1}}$ от n переменных, заданных линейными формами $l_1, \ldots, l_{2^{n-1}}$, где

$$l_i(x_1, \ldots, x_n) = l_{\vec{w},\sigma}(x_1, \ldots, x_{n-1}) - r_i x_n.$$

Отметим, что для всякого $i \in \left\{1, \ldots, 2^{n-1}\right\}$ имеет место

$$f_i(x_1, \ldots, x_{n-1}, 0) = f(x_1, \ldots, x_{n-1}).$$

При этом все функции f_i различны, так как

$$\begin{cases} f_i(\alpha_i, 1) = 1; \\ f_i(\alpha_{i+1}, 1) = 0; \\ f_{i+1}(\alpha_{i+1}, 1) = 1; \end{cases}$$

для всех $i \in \left\{1, \ldots, 2^{n-1}\right\}$.

Так как переменная x_n существенная, то каждая из функций $f_1, \ldots, f_{2^{n-1}}$ является существенной. Лемма доказана.

Таким образом, «почти все» пороговые функции от n переменных — существенные.

Доказательство теоремы 4.

Заметим, что верхние оценки очевидно следуют из теоремы 3 и леммы 5.2. Более того, нижняя оценка второго утверждения вытекает из нижней оценки первого утверждения и леммы 5.2. Таким образом, для доказательства теоремы достаточно доказать нижнюю оценку первого утверждения.

Ранее, в разделе 2, было введено понятие сигнатуры пороговой функции. Легко видеть, что операция инвертирования изменяет сигнатуру функции в одной позиции. Следовательно, всякая пороговая функция не более, чем за n операций инвертирования, может быть преобразована к симметричной ей монотонной пороговой функции. Из лемм 2.5 и 5.3 следует, что почти все монотонные пороговые функции являются существенными. Следовательно, для получения нижней оценки на близость ρ_{II} для почти всех случаев с точностью до n достаточно оценить близость ρ для почти всех пар (l, f), где l — линейная форма с натуральными коэффициентами и свободным членом, а f — монотонная существенная пороговая функция.

Зафиксируем линейную форму l с натуральными коэффициентами и порогом, зависящую от n переменных. Определим число $p\,(n)$ монотонных пороговых функций от $n-1$ переменного, близость которых до линейной формы l не превосходит величины $k = 2^{n-\log n}$.

Нетрудно видеть, что число монотонных пороговых функций, близость которых до линейной формы l не превосходит k, ограничено сверху величиной $2^{n-1} \cdot R\,(n, k)$, где $R_{n,k}$ — число наборов натуральных чисел длины n, сумма компонент которых меньше k. Следовательно, $p\,(n) \leqslant 2^{n-1} \cdot R\left(n, 2^{n-\log n}\right)$. По лемме 4.4 имеем

$$p\,(n) \lesssim 2^{n(n-\log n)-n\log n+o(n\log n)} = 2^{n^2-2n\log n+o(n\log n)}.$$

Рассмотрим как соотносятся величины $p\,(n)$ и $\left|\widetilde{\mathrm{MT}}^{\mathrm{n-1}}\right|$, где $\widetilde{\mathrm{MT}}^{\mathrm{n-1}}$ — множество существенных монотонных пороговых функций от $n-1$ переменного.

Из теоремы 11 и лемм 2.5, 5.3 при $n \to \infty$ следует

$$\lim_{n\to\infty} \frac{\left|\widetilde{MT}^{n-1}\right|}{p\,(n)} = \lim_{n\to\infty} \frac{\left|T^{n-1}\right|}{n \cdot p\,(n)} = 2^{n\log n+o(n\log n)} = \infty.$$

Пробегая по конечному множеству линейных форм l с натуральными коэффициентами таких, что $L\,(l) \leqslant L\,(n)$, получаем, что для почти всех пар (l, f) близость $\rho_{II}\,(l, f)$ не менее, чем $2^{n-\log n}$. Теорема доказана.

Доказательство теоремы 5. Докажем первое утверждение теоремы. Заметим, что верхняя оценка первого утверждения очевидно следует из теоремы 1, так как $\rho_{II}\,(m, k) \leqslant \rho\,(m, k)$.

Докажем верхнюю оценку второго утверждения. Легко видеть, что число переменных у пороговых функций из класса $T_{m,k}$ не превосходит $m \cdot k$. Следовательно, по лемме 5.2, величина $\rho_a\,(m, k)$ может быть оценена так $\rho_a\,(m, k) \leqslant 2mk\log_a L\,(m, k) + 3mk$.

Откуда, по лемме 3.4, получаем

$$\rho_a(m,k) \lesssim \frac{2mk}{\log a}\cdot\left(k\log m + \frac{1}{2}k\log k + o(k\log k)\right) = \\ = \frac{mk^2}{\log a}\left(\log k + o(k\log k)\right) \lesssim mk^2\log k.$$

Верхние оценки доказаны. Докажем нижние оценки первого и второго утверждений. Заметим, что по теореме 12 найдется линейная форма минимального веса $l_{\vec{w},\sigma}$ такая, что

$$\begin{cases} w_1 \geqslant 2; \\ |w_i| \geqslant w_1\cdot 2^{\frac{1}{4}k\log k + o(k\log k)}, \text{ если } i\in\{2,\ldots,k\}. \end{cases}$$

По лемме 5.1, линейная форма $l_{\vec{w}',\sigma}$, где

$$\vec{w}' = \left(w_1,\ldots,w_{k-1},\underbrace{2,\ldots,2}_{m \text{ раз}}\right),$$

обладает минимальным весом. Также легко видеть, что минимальным весом обладает линейная форма $l_{\vec{w}'',\sigma}$, где

$$\vec{w}'' = \left(w_1,\ldots,w_{k-1},\underbrace{w_k,\ldots,w_k}_{m \text{ раз}}\right).$$

Пороговые функции, задаваемые линейными формами $l_{\vec{w}',\sigma}$ и $l_{\vec{w}'',\sigma}$, обозначим f' и f'', соответственно. По лемме 5.2 получим

$$\rho_{II}(m,k) \geqslant \rho_{II}\left(f',f''\right) \geqslant m\cdot 2^{\frac{1}{4}k\log k + o(k\log k)},$$
$$\rho_a(m,k) \geqslant \rho_a\left(f',f''\right) \geqslant \frac{1}{4\log a}mk\log k + o(k\log k).$$

Теорема доказана.

6. Алгоритмическая сложность обучения и некоторые классы пороговых функций

6.1. Основные понятия, постановки и результаты

Рассмотрим вопрос алгоритмической сложности задачи перехода от «стартовой» линейной формы к «целевой», задающей желаемую пороговую функцию.

Для решения данной задачи предлагается алгоритм, состоящий из двух этапов. На первом этапе на основании множеств единиц и нулей целевой функции f осуществляется поиск допустимой линейной формы для функции f, которая в дальнейшем выступает в качестве «целевой». На втором шаге осуществляется непосредственная перестройка «стартовой» формы в «целевую».

Далее, под арифметическими операциями будем понимать следующие: сложение, вычитание, умножение, деление, извлечение квадратного корня и нахождение наибольшего из двух чисел. Операции выполняются над рациональными числами в двоичном представлении. Число разрядов в двоичном представлении каждого из операндов будем называть точностью операции.

Имеет место следующее утверждение, характеризующее сложность поиска «целевой» допустимой линейной формы.

Теорема 13. *Если $S = \{s_1, \ldots, s_p\}$, $C = \{c_1, \ldots, c_q\}$ — множества, соответственно, единиц и нулей функции f, то достаточно выполнить $O\left((p+q)\, n^4 \log n\right)$ арифметических операций с точностью $O\left(n \log n\right)$ бит, чтобы найти допустимую линейную форму для f, если f-пороговая, или же показать, что такой линейной формы нет.*

Рассмотрим следующую величину

$$d_i\left(f', f''\right) = \max_{\substack{l' \to f' \\ l'' \to f''}} \rho_i\left(l', l''\right),$$

где максимум берется по всем допустимым линейным формам $l_{\vec{w}', \sigma'}$ и $l_{\vec{w}'', \sigma''}$, задающим функции f' и f'', соответственно.

Введем более общую характеристику для классов пороговых функций, инвариантных относительно групп перестановок. Обозначим

$$d_i\left(m,k\right) = \max_{f',f''\in T_{m,k}} d_i\left(f',f''\right),$$

где максимум берется по всем пороговым функциям $f', f'' \in T_{m,k}$.

Величина $d_i\left(m,k\right)$ характеризует максимально-возможную сложность взаимной перестройки допустимых линейных форм, задающих пороговые функции из класса $T_{m,k}$, при использовании множества допустимых операций M_i. Имеет место следующее утверждение.

Лемма 6.1. *Если M — множество допустимых операций, $k \to \infty$, $\log m = o\left(\log k\right)$, то*

$$\begin{cases} d_i\left(m,k\right) \leqslant 2^{\frac{1}{2}k\log k + o(k\log k)}, & \text{если } M = M_I, M_{II};\\ d_i\left(m,k\right) \leqslant O\left(mk^2\log k\right), & \text{если } M = M_a. \end{cases}$$

Теорема 13 и лемма 6.1 дают описание алгоритма, который решает задачу перестройки так, что возникающая при этом траектория перестройки по длинне близка к минимально-возможной, а сложность нахождения этой траектории зависит полиномиальным образом от числа единиц и нулей в задании «целевой» пороговой функции. Сформулируем основное следствие теоремы 13 и леммы 6.1.

Следствие. *Если $k \to \infty$, $\log m = o\left(\log k\right)$, то для решения задачи перестройки «стартовой» допустимой формы в допустимую форму, задающую «целевую» пороговую функцию, представленную в виде множеств единиц — $\{s_1,\ldots,s_p\}$ и нулей — $\{c_1,\ldots,c_q\}$, достаточно выполнить следующее число элементарных операций из множества M*

$$\begin{cases} 2^{\frac{1}{2}k\log k + o(k\log k)}, & \text{если } M = M_I, M_{II};\\ O\left(mk^2\log k\right), & \text{если } M = M_a. \end{cases}$$

При этом, число вспомогательных арифметических операций не превосходит $O\left(\left(p+q\right)k^4\log k\right)$, а точность каждой операции не более, чем $O\left(k\log k\right)$ бит.

Таким образом, в случае задания «целевой» функции в виде множеств единиц и нулей и использования множества допустимых операций M_a задача перестройки пороговых функций может быть решена за полиномиальное время как с точки зрения числа вспомогательных арифметических операций для нахождения направления перестройки, так и с точки зрения числа элементарных шагов в процессе самой перестройки.

Рассмотрим задачу конструктивного описания классов классов пороговых функций, которые были бы максимально удалены или же, наоборот, близки друг другу в терминах функции близости ρ.

Перенумеруем пороговые фукнции от n переменных $T^n = \left\{f_1, \ldots, f_{|T^n|}\right\}$. Введем матрицу $D\left(n\right) = \left\|\rho_{ij}\right\|_{|T^n| \times |T^n|}$. Элементы матрицы $D\left(n\right)$ определим следующим образом

$$\rho_{ij} = \rho\left(f_i, f_j\right),$$

где $f_i, f_j \in T^n$. Будем называть $D\left(n\right)$ матрицей близости множества T^n.

Отметим очевидные свойства матрицы близости:

1) $\rho_{ii} = 0$,
2) $\rho_{ij} = \rho_{ji}$,

для всех $i, j \in \{1, \ldots, |T^n|\}$.

Имеет место следующая теорема, предъявляющая конструктивный пример класса пороговых функций, значительно удаленных друг от друга в терминах функции близости ρ.

Теорема 14. *Если $n \to \infty$, то в матрице близости $D\left(n\right)$ содержится такое подмножество $R\left(n\right)$, что*

$$\log_2 \left|R\left(n\right)\right| \asymp \log_2 \left|D\left(n\right)\right|,$$

и для всех $\rho_{ij} \in R\left(n\right)$ выполнено

$$\log_2 \rho_{ij} \asymp \log_2 \rho\left(n\right).$$

Следующая теорема описывает класс пороговых функций, мощность которого экспоненциально зависит от числа переменных, а сложность перестройки его элементов ограничена заранее заданной величиной.

Теорема 15. *Если $n + 1 \leqslant c \leqslant 2^{\frac{n}{2}}$, то существует класс M пороговых функций от n переменных, содержащий $(c - n - 1) \cdot 2^{n-2}$ элементов, такой, что для всех f', f'' из M выполнено*

$$\rho\left(f', f''\right) \leqslant 3 \cdot c.$$

Рассмотрим задачу явного построения линейных форм минимального размаха и веса. Напомним, что линейная форма $l_{\vec{w},\sigma}$ называется формой минимального размаха для пороговой функции f, если $l_{\vec{w},\sigma}$ задает f строгим образом и ее размах минимален среди всех линейных форм, задающих f строгим образом. Аналогичным образом вводится понятие линейной формы минимального веса для пороговой функции f.

Введенное понятие линейной формы минимального размаха (веса) является корректным, так как для всякой пороговой функции существует линейная форма минимального размаха (веса). Для обоснования этого достаточно заметить, что если целочисленная линейная форма $l_{\vec{w},\sigma}$ задает пороговую функцию f, то линейная форма $l_{2\cdot\vec{w},2\cdot\sigma+1}$ задает f строгим образом.

Будем обозначать линейную форму $x_1 w_1 + \ldots + x_n w_n - \sigma$ вектором $(w_1, \ldots, w_n, \sigma)$. Введем следующие операции над линейными формами.

Операции O_1, O_2 и O_3 применимы к произвольной линейной форме и определяются так

$$O_1\left[(w_1, \ldots, w_n, \sigma)\right] = \left(w_1, \ldots, w_n, \sum_{i=1}^{n} w_i - \sigma\right);$$

$$O_2\left[(w_1, \ldots, w_i, \ldots, w_n, \sigma)\right] = (w_1, \ldots, w_i, w_i, \ldots, w_n, \sigma);$$

$$O_3\left[(w_1, \ldots, w_n, \sigma)\right] = (2, w_1, \ldots, w_n, \sigma).$$

Операция O_4 применима к линейной форме $(w_1, \ldots, w_n, \sigma)$, задающей пороговую функцию f, если на некотором наборе $\alpha =$

$(a_1, \ldots, a_n) \in E_2^n$, $f(\alpha) = 0$ и $\sigma = \sum\limits_{i:a_i=1} w_i + 1$, где сумма берется по всем i, для которых $a_i = 1$. В этом случае

$$O_4 \left[(w_1, \ldots, w_n, \sigma)\right] = (w_1, \ldots, w_n, w_{n+1}, \sigma),$$

где $w_{n+1} = \sum\limits_{i:a_i=1} w_i + 2$.

Имеет место следующий результат.

Теорема 16. *Если $l_{\vec{w},\sigma} = (w_1, \ldots, w_n, \sigma)$ — линейная форма минимального размаха (веса), то $O_i\left[l_{\vec{w},\sigma}\right]$, где $i \in \{1, \ldots, 4\}$, является линейной формой минимального размаха (веса).*

Путем применения данных операций к элементарным линейным формам можно построить бесконечное многообразие линейных форм минимального размаха и веса. Одним из подмножеств данного семейства является последовательность линейных форм, весовые коэффициенты которых представляют собой числа Фибоначчи, что составляет известный пример, полученный в работах Мурога [23] и Парберри [26]. Отметим также, что из утверждения теоремы 16 следует одновременно минимальность веса и размаха элементов многообразия, полученного в результате применения операций O_i, где $i = 1, \ldots, 4$. Это усиливает результаты Мурога и Парберри.

Следствие. *Если $n \geqslant 1$, то линейная форма*

$$(2 \cdot F_1, 2 \cdot F_2, \ldots, 2 \cdot F_n, 2 \cdot F_{n+1} - 1),$$

где F_i — i-е число Фибоначчи, является линейной формой минимального размаха и веса для соответствующей пороговой функции от n переменных.

Следствие. *Существует последовательность пороговых функций $f_1, \ldots, f_n, \ldots$, таких что f_i зависит от i переменных и при $n \to \infty$ выполнено*

$$L(f_n) \sim \frac{2 \cdot \varphi^n}{\sqrt{5}},$$

$$\mu(f_n) \sim \frac{2 \cdot \varphi^{n+2}}{\sqrt{5}},$$

где $\varphi = \frac{1+\sqrt{5}}{2}$.

6.2. Доказательства утверждений

Докажем утверждения теоремы 13 и леммы 6.1.

Пусть $S = \{s_1, \ldots, s_p\}$ и $C = \{c_1, \ldots, c_q\}$ — множества наборов из E_2^n. Говорят, что множества S и C линейно-отделимы, если существует линейная форма $l_{\vec{w},\sigma}$, такая что

$$\begin{cases} l_{\bar{w},\sigma}(\alpha) \geqslant 0, \text{ если } \alpha \in S; \\ l_{\bar{w},\sigma}(\alpha) < 0, \text{ если } \alpha \in C. \end{cases}$$

Если S и C — множества единиц и нулей «целевой» пороговой функции, то, очевидно, S и C являются линейно-отделимыми.

Введем понятие квазиминимальной линейной формы множеств S и C. Далее этот класс линейных форм будет использован для доказательства теоремы 13, характеризующей сложность поиска «целевой» допустимой линейной формы.

Рассмотрим набор $\alpha = (a_1, \ldots, a_n) \in S \cup C$. Обозначим

$$\alpha' = \begin{cases} (a_1, \ldots, a_n, 1), \text{ если } \alpha \in S; \\ (-a_1, \ldots, -a_n, -1), \text{ иначе.} \end{cases}$$

Рассмотрим систему неравенств

$$\underset{\alpha_i \in S \cup C}{\&} \left(\alpha_i' \cdot \vec{v} \geqslant n + 2\right), \tag{6.1}$$

где $\vec{v} = (w_1, \ldots, w_{n+1}) \in R^{n+1}$.

Рассмотрим набор $\bar{v}' = ([v_1], \ldots, [v_{n+1}])$ такой, что набор $\bar{v}$ удовлетворяет системе 6.1 и доставляет минимум функционала $\mu(\bar{v}) = \sum\limits_{i=1}^{n+1} |v_i|$. В этом случае набор $\bar{v}'$ назовем квазиминимальной линейной формой множеств S и C. Заметим, что $\bar{v}' \in Z^{n+1}$.

Можно показать, что для всякой пары линейно-отделимых множеств существует квазиминимальная линейная форма. При этом, данная линейная форма также является отделяющей.

Лемма 6.2. *Если $S, C \subseteq E_2^n$ — линейно-отделимые множества, то существует квазиминимальная линейная форма l множеств S и C. При этом, l отделяет множества S и C.*

Доказательство. Заметим, что по определению квазиминимальная форма $\bar{v}'$ равна $([v_1], \ldots, [v_{n+1}])$, где набор $\bar{v}$ удовлетворяет системе неравенств 6.1.

Так как для всякого $x \in R$ выполнено $x - 1 \leqslant [x] \leqslant x + 1$, то на всех наборах $\alpha = (a_1, \ldots, a_n) \in S \cup C$ выполнено

$$\begin{cases} \sum\limits_{i=1}^{n} v_i' a_i + v_{n+1}' \geqslant 1, \text{ если } \alpha \in S; \\ \sum\limits_{i=1}^{n} v_i' a_i + v_{n+1}' \leqslant -1, \text{ иначе.} \end{cases}$$

Очевидно, что линейная форма $l_{\bar{w},\sigma}$, где $\bar{w} = (v_1', \ldots, v_n')$ и $\sigma = -v_{n+1}'$, отделяет множества S и C. Для того, чтобы доказать существование квазиминимальной формы $\bar{v}'$, достаточно показать, что множество решений системы 6.1 непусто и, при этом, минимум функционала $\mu(\bar{v}) = \sum\limits_{i=1}^{n+1} |v_i|$ достигается.

Легко видеть, что для всякой пары линейно-отделимых множеств S и C найдется линейная форма $l_{\bar{w},\sigma}$ такая, что для всех наборов $\alpha = (a_1, \ldots, a_n) \in S \cup C$ выполнено

$$\begin{cases} l_{\bar{w},\sigma}(\alpha) \geqslant 1, \text{ если } \alpha \in S; \\ l_{\bar{w},\sigma}(\alpha) \leqslant -1, \text{ иначе.} \end{cases}$$

Следовательно, для доказательства существования решения системы 6.1 достаточно рассмотреть линейную форму $l_{(n+2)\cdot\bar{w},(n+2)\cdot\sigma}$.

Существование решения системы 6.1, доставляющего минимум функционала $\mu(\bar{v})$, обосновывается следующим образом. Множество решений системы 6.1 представляет собой выпуклый многогранник в пространстве R^{n+1}. Выпуклость обосновывается следующим образом. Если l_1, l_2 — линейные формы, отделяющие множества S и C, $r \in [0, 1]$, то линейная форма вида $l = r \cdot l_1 + (1 - r) \cdot l_2$ также отделяет S и C. Функционал $\mu(\bar{v})$ представляет собой манхэттенское расстояние от точки пространства R^{n+1} до начала координат. Следовательно, в силу выпуклости многогранника решений, существует хотя бы одно решение системы 6.1, находящееся на кратчайшем расстоянии от начала координат. Лемма доказана.

Из леммы 6.2 следует, что если S и C — множества единиц и нулей «целевой» пороговой функции, то квазиминимальная форма множеств S и C задает «целевую» пороговую функцию.

Имеет место следующее утверждение, характеризующее все возможные квазиминимальные линейные формы.

Лемма 6.3. *Множество квазиминимальных форм является допустимым.*

Доказательство. Обоснуем допустимость множества квазиминимальных линейных форм. Для этого оценим величину их размаха. Рассмотрим линейную форму $l_{\bar{w},\sigma}$ с целочисленными весами и порогом, задающую пороговую функцию f, на которой достигается размах $L(f)$. Рассмотрим линейную форму $l_{\bar{w}',\sigma'}$ такую, что $\bar{w}' = (2n+4)\cdot\bar{w}$, $\sigma' = (2n+4)\cdot\sigma - (n+2)$. В таком случае

$$\begin{cases} l_{\bar{w}',\sigma'}(\alpha) \geqslant n+2, \text{ если } f(\alpha)=1; \\ l_{\bar{w}',\sigma'}(\alpha) \leqslant -(n+2)\,, \text{ иначе.} \end{cases}$$

Таким образом, линейная форма $l_{\bar{w}',\sigma'}$ задает f, удовлетворяет системе неравенств 6.1 для произвольных множеств S, C — единиц и нулей функции f, и, так как $L(l_{\bar{w},\sigma}) \geqslant 1$, то

$$L\left(l_{\bar{w}',\sigma'}\right) \leqslant (2n+4)\cdot L(l_{\bar{w},\sigma}) + n + 2 \leqslant (3n+6)\cdot L(f)\,.$$

Лемма доказана.

Доказательство теоремы 13. Из лемм 6.2 и 6.3 следует, что, если S и C — линейно-отделимы, то существует квазиминимальная форма множеств S и C, которая задает «целевую» пороговую функцию и, при этом, является допустимой.

Следовательно, для доказательства леммы достаточно найти квазиминимальную форму множеств S и C или показать, что такой формы нет.

Напомним, что по определению квазиминимальная форма $\bar{v}'$ равна $([v_1],\ldots,[v_{n+1}])$, где набор $\bar{v}$ удовлетворяет системе неравенств 6.1.

Ранее было также показано, что множество решений системы 6.1 является выпуклым. Следовательно, для нахождения квазиминимальной формы $\bar{v}'$, необходимо найти решение системы 6.1, минимизирующее функционал $\mu(\bar{v})$, или же показать, что решения системы 6.1 не существует.

Данная задача является задачей линейного программирования. Воспользовавшись методом Хачияна [10, 11], получим, что для решения данной задачи достаточно выполнить $O\left(mn^3L\right)$ арифметических операций с точностью $O(L)$ бит, где $m = p + q$ и $L = \log_2 K + \log_2(n+2) + \log_2(p+q+n+1)$.

Здесь K — максимальное абсолютное значение определителя произвольной подматрицы A, определенной следующим образом

$$A = \begin{pmatrix} s_1(1) & \dots & s_1(n) \\ \dots & \dots & \dots \\ s_p(1) & \dots & s_p(n) \\ -c_1(1) & \dots & -c_1(n) \\ \dots & \dots & \dots \\ -c_q(1) & \dots & -c_q(n) \end{pmatrix}.$$

Так как элементы матрицы A принимают значения из множества $\{-1, 0, 1\}$, то по теореме Адамара [1] получим, что $K \leqslant n^{\frac{1}{2}n\log n}$.

Так как $p + q \leqslant 2^n$, то $L = O(n\log n)$. В итоге, общее число арифметических операций всей процедуры поиска не превосходит

$$O\left((p+q)\,n^4\log n\right),$$

а точность операций составляет $O(n\log n)$ бит. Теорема доказана.

Оценим сложность взаимной перестройки допустимых линейных форм, задающих пороговые функции из класса $T_{m,k}$.

Докажем первое утверждение леммы 6.1. Так как $\log m = o(\log k)$, то по лемме 3.4 получим, что размах допустимой линейной формы, задающей функцию из $T_{m,k}$, ограничен величиной

$$L'(m,k) \leqslant O(n) \cdot L(m,k) \leqslant O(m \cdot k) \cdot 2^{\frac{1}{2}k\log k + o(k\log k)}.$$

Следовательно $L'(m,k) \leqslant 2^{\frac{1}{2}k\log k + o(k\log k)}$. Очевидно, что близость между допустимыми линейными формами в этом случае ограничена величиной $2mk \cdot L'(m,k)$. Подставляя выражение для $L'(m,k)$, получаем утверждение леммы.

Докажем второе утверждение леммы 6.1. По лемме 5.2, величина $\rho_a(m,k)$ может быть оценена так $\rho_a(m,k) \leqslant 2mk \log_a L(m,k) + 3mk$. Следовательно, близость $\rho'_a(m,k)$ между допустимыми линейными формами, задающими пороговые функции из класса $T_{m,k}$ удовлетворяет следующему неравенству

$$\rho'_a(m,k) \leqslant \frac{2mk}{\log a} \cdot (\log(O(mk)) + \log L(m,k)) + 3mk.$$

По лемме 3.4 получаем

$$\rho'_a(m,k) \leqslant O\left(mk^2 \log k\right).$$

Лемма доказана.

Докажем теорему 16. Для этого достаточно показать, что каждая из операций O_1, O_2, O_3 и O_4 сохраняет свойство минимальности размаха и веса.

Докажем, что операция O_1 сохраняет свойство минимальности размаха и веса. Заметим, что в случае строгого задания $(-1,1)$-пороговой функции для перехода к двойственной ей $(-1,1)$-пороговой функции достаточно изменить знак порога. Иными словами, если $l_{\vec{w},\sigma} \Rightarrow f^{-1,1}$, то $l_{\vec{w},-\sigma} \Rightarrow f^{*-1,1}$.

Далее легко видеть, что $L(f) = L(f^*)$ и $\mu(f) = \mu(f^*)$.

В таком случае из леммы 2.1 вытекает следующее утверждение.

Лемма 6.4. *Если $l_{\vec{w},\sigma}$ — форма минимального размаха (веса), то линейная форма*

$$l_{\vec{w},\sum\limits_{i=1}^{n} w_i - \sigma}$$

является формой минимального размаха (веса) для соответствующей пороговой функции.

Доказательство. По лемме 2.1 имеем

$$l_{\vec{w},2\sigma - \sum\limits_{i=1}^{n} w_i} \Rightarrow f^{-1,1}.$$

Следовательно,

$$l_{\vec{w}, \sum\limits_{i=1}^{n} w_i - 2\sigma} \Rightarrow f^{*-1,1}.$$

Наконец, снова применяя лемму 2.1, получаем

$$l_{\vec{w}, \frac{1}{2}\left(2\sum\limits_{i=1}^{n} w_i - 2\sigma\right)} = l_{\vec{w}, \sum\limits_{i=1}^{n} w_i - \sigma} \Rightarrow f^{*0,1}.$$

Лемма доказана.

Операция O_2 очевидно сохраняет свойство минимальности размаха и веса.

Тот факт, что операция O_3 сохраняет свойство минимальности размаха и веса, следует из леммы 5.1, доказанной ранее в разделе 5.

Докажем, что операция O_4 сохраняет свойство минимальности размаха и веса.

Лемма 6.5. *Если* $(w_1, \ldots, w_n, \sigma)$ *— форма минимального размаха (веса) для* $f \in T^n$, $f(\alpha) = 0$, *где* $\alpha = (a_1, \ldots, a_n) \in E_2^n$, *и* $\sigma = \sum\limits_{i: a_i = 1} w_i + 1$, *то линейная форма* $(w_1, \ldots, w_n, w_{n+1}, \sigma)$, *где* $w_{n+1} = \sum\limits_{i: a_i = 1} w_i + 2$, *является формой минимального размаха (веса) для соответствующей пороговой функции, где сумма берется по всем* i, *для которых* $a_i = 1$.

Доказательство. Так как $f(\alpha) = 0$, то

$$\sum_{i:\, a_i = 1} w_i < \sigma.$$

Рассмотрим линейную форму $(w_1, \ldots, w_n, w_{n+1}, \sigma)$ и задаваемую ей пороговую функцию $f' \in T^{n+1}$. Очевидно, что $f'(x_1, \ldots, x_n, 0) = f(x_1, \ldots, x_n)$ при любых значениях весового коэффициента w_{n+1}. Доопределим f' следующим образом $f'(0, \ldots, 0, 1) = 1$.

Очевидно, что пороговая функция f', обладающая таким свойством существует. В таком случае $w_{n+1} > \sigma$, а, следовательно, $w_{n+1} > \sigma > \sum\limits_{i:\, a_i = 1} w_i$ или же

$$w_{n+1} \geqslant \sum_{i:\, a_i = 1} w_i + 2.$$

Лемма доказана.

Таким образом, каждая из операций O_1, O_2, O_3 и O_4 сохраняет свойство минимальности размаха и веса. Теорема 16 доказана. Докажем следствия этой теоремы.

Доказательство следствия 1 теоремы 16. Для доказательства следствия построим последовательность пороговых функций $f_1, f_2, \ldots, f_n, \ldots$ и соответствующие им линейные формы минимального размаха и веса.

Положим $f_1(x_1) = x$. Легко видеть, что набор $(2,1)$ задает линейную форму минимального размаха и веса для f_1. Отметим также, что $f_1(0) = 0$.

Далее, если задана пороговая функция f_i, $i \geqslant 1$, задаваемая линейной формой минимального размаха (веса) — $(w_1, \ldots, w_i, \sigma)$, и

$$f_i\left(\underbrace{1, \ldots, 1}_{i-1}, 0\right) = 0,$$

тогда по лемме 6.5 линейная форма $(w_1, \ldots, w_i, w_{i+1}, \sigma)$, где $w_{i+1} = \sum\limits_{j=1}^{i-1} w_i + 2$, является линейной формой минимального размаха (веса) для некоторой пороговой функции g_{i+1} от $i+1$ переменных. При этом, будет выполнено

$$g_{i+1}\left(\underbrace{0, \ldots, 0}_{i}, 1\right) = 1.$$

Рассмотрим функцию f_{i+1} двойственную к g_{i+1}, то есть $f_{i+1} = g^*_{i+1}$. По лемме 6.4 линейная форма

$$\left(w_1, \ldots, w_i, w_{i+1}, \sum_{j=1}^{i+1} w_j - \sigma\right)$$

является формой минимального размаха (веса) для f_{i+1}. Отметим также, что по определению двойственности выполнено

$$f_{i+1}\left(\underbrace{1, \ldots, 1}_{i}, 0\right) = 0.$$

При этом, по лемме 6.4 имеем

$$L(f_{i+1}) = L(g) = L(f_i) + \sum_{j=1}^{i-1} w_j + 2.$$

Условия леммы 6.5 опять выполнены. Таким образом осуществляется построение последовательности пороговых функций $f_1, \ldots, f_n, \ldots$.

Для оценки веса и размаха функции f_n отметим, что для соответствующей линейной формы минимального размаха $(w_1, \ldots, w_n, \sigma)$ выполнено

$$w_n = \sum_{j=1}^{n-2} w_j + 2,$$

при этом, $w_1 = 2$.

Легко видеть, что $w_n = 2 \cdot F_n$, где F_n — n-е число Фибоначчи.

Следовательно, линейная форма $l_{\bar{w},\sigma}$, задающая функцию f_n строгим образом, имеет вид $(w_1, \ldots, w_n, \sigma)$, где $w_i = 2 \cdot F_n$, для всех $i = 1, \ldots, n$, и $\sigma = 2F_{n+1} - 1$.

Размах и вес линейной формы $l_{\bar{w},\sigma}$ выражаются следующим образом

$$L(f_n) = 2 \cdot F_n,$$

$$\mu(f_n) = 2\sum_{i=1}^{n} F_n = 2 \cdot F_{n+2} - 2.$$

Следствие доказано.

Доказательство следствия 2 теоремы 16. Известно [2], что при $n \to \infty$ имеет место асимптотика $F_n \sim \frac{\varphi^n}{\sqrt{5}}$, где $\varphi = \frac{1+\sqrt{5}}{2}$.

Следствие доказано.

Докажем теорему 14.

Из теоремы 12 следует, что существует $(-1, 1)$-пороговая функция, такая, что при ее задании линейной формой, все коэффициенты будут «большими».

Оказывается, существует монотонная $(0, 1)$-пороговая функция, обладающая этим же свойством.

Для доказательства данного утверждения, сначала докажем некоторое свойство введенного в разделе 2 оператора δ-преобразования.

Лемма 6.6. *Для любых $f \in T^n$, $\delta \in \{0,1\}^n$ и $i \in \{1,\ldots,n\}$ выполнено*

$$L_i(f) = L_i(\delta[f]).$$

Из лемм 6.6 и 2.1 вытекает следующее утверждение.

Лемма 6.7. *Существует монотонная пороговая функция $f^{0,1}$ такая, что для всех $i \in \{1,\ldots,n\}$ имеет место*

$$L_i\left(f^{0,1}\right) \geqslant \frac{1}{2n} e^{-4n^{\log_2\left(\frac{3}{2}\right)}} 2^{\frac{1}{2}n\log_2 n - n}.$$

Пороговую функцию, существование которой следует из леммы 6.7, обозначим $f_{H(n)}$.

Будем говорить, что *линейная форма $l_{\vec{w},\sigma}$ задает пороговую функцию $f(x_1,\ldots,x_n)$ граничным образом*, и записывать это так

$$l_{\vec{w},\sigma} \mapsto f(x_1,\ldots,x_n),$$

если $l_{\vec{w},\sigma}$ задает f и найдется такой набор $(x_1,\ldots,x_n) \in \{0,1\}^n$, что на нем

$$l_{\vec{w},\sigma}(x_1,\ldots,x_n) = 0.$$

Следующее утверждение говорит о том, что произвольную пороговую функцию можно задать целочисленной линейной формой граничным образом.

Лемма 6.8. *Если f — пороговая функция и $f \neq 0$, то найдется линейная форма $l_{\vec{w},\sigma}$, такая, что*

$$l_{\vec{w},\sigma} \mapsto f.$$

Доказательство. Пусть $l_{\vec{w},\sigma'} \to f$. Можно считать, что ни для какого набора $\alpha \in \{0,1\}^n$ не выполнено $l_{\vec{w},\sigma'}(\alpha) = 0$, так как в противном случае утверждение леммы очевидно.

Рассмотрим множество $t_f = \{\alpha \in \{0,1\}^n : f(\alpha) = 1\}$. Очевидно, что для всех $\alpha \in t_f$ имеет место $l_{\vec{w},\sigma'}(\alpha) > 0$. Рассмотрим такой набор α', что

$$l_{\vec{w},\sigma'}(\alpha') = \min_{\alpha \in t_f} l_{\vec{w},\sigma'}(\alpha).$$

Пусть $l_{\vec{w},\sigma'}(\alpha') = c$.

Так как $l_{\vec{w},\sigma'}$ — линейная форма с целочисленными коэффициентами, то c — натуральное. Действительно, $c > 0$, так как $f(\alpha') = 1$.

Рассмотрим линейную форму $l_{\vec{w},\sigma} = l_{\vec{w},\sigma'} - c$. Очевидно, что $l_{\vec{w},\sigma}(\alpha') = 0$, тогда как для остальных наборов $\alpha \in t_f$ имеет место $l_{\vec{w},\sigma}(\alpha) \geqslant 0$. Лемма доказана.

Введем операцию склейки пороговых функций, заданных граничным образом.

Лемма 6.9. *Если $f_1 \in T^n$, $f_2 \in T^m$, f_1 и f_2 не константа 0, тогда найдутся наборы*

$$\alpha = (a_1, \ldots, a_n) \in \{0,1\}^n;$$
$$\beta = (b_1, \ldots, b_m) \in \{0,1\}^m;$$

а также пороговая функция $f \in T^{n+m}$ такие, что

$$\begin{cases} f(x_1, \ldots, x_n, b_1, \ldots, b_m) = f_1(x_1, \ldots, x_n), \\ f(a_1, \ldots, a_n, x_{n+1}, \ldots, x_{n+m}) = f_2(x_{n+1}, \ldots, x_{n+m}). \end{cases}$$

Доказательство. По лемме 6.8 каждая из функций f_1 и f_2 может быть задана граничным образом. Следовательно, найдутся линейные формы $l_1 \mapsto f_1$, $l_2 \mapsto f_2$ и наборы α и β, такие, что $l_1(\alpha) = l_2(\beta) = 0$.

Рассмотрим линейную форму

$$l(x_1, \ldots, x_{n+m}) = l_1(x_1, \ldots, x_n) + l_2(x_{n+1}, \ldots, x_{n+m}).$$

Очевидно, что l задает некоторую пороговую функцию f от $n+m$ переменных. При этом

$$\begin{cases} l(x_1, \ldots, x_n, b_1, \ldots, b_m) = l_1(x_1, \ldots, x_n), \\ l(a_1, \ldots, a_n, x_{n+1}, \ldots, x_{n+m}) = l_2(x_{n+1}, \ldots, x_{n+m}). \end{cases}$$

Лемма доказана.

Данная лемма позволяет строить пороговые функции от большего числа переменных путем «склейки» пороговых функций от меньшего числа переменных.

Следующая лемма характеризует величины $L_i(f)$ пороговой функции, полученной в результате операции склейки.

Лемма 6.10. *Если* $f_1 \in T^n$, $f_2 \in T^m$, $f \in T^{n+m}$, $\alpha = (a_1, \ldots, a_n) \in \{0,1\}^n$, $\beta = (b_1, \ldots, b_m) \in \{0,1\}^m$ *и*

$$\begin{cases} f(x_1, \ldots, x_n, b_1, \ldots, b_m) = f_1(x_1, \ldots, x_n), \\ f(a_1, \ldots, a_n, x_{n+1}, \ldots, x_{n+m}) = f_2(x_{n+1}, \ldots, x_{n+m}), \end{cases}$$

то для всякого $i \in \{1, \ldots, n\}$ *имеет место*

$$L_i(f) \geqslant L_i(f_1),$$

а для всякого $i \in \{n+1, \ldots, n+m\}$

$$L_i(f) \geqslant L_i(f_2).$$

Доказательство. Пусть $L_i(f)$ достигается на линейной форме l. В таком случае

$$\begin{cases} l(x_1, \ldots, x_n, b_1, \ldots, b_m) = l_1(x_1, \ldots, x_n) \to f_1, \\ l(a_1, \ldots, a_n, x_{n+1}, \ldots, x_{n+m}) = l_2(x_{n+1}, \ldots, x_{n+m}) \to f_2. \end{cases}$$

Лемма доказана.

Оказывается, когда в склейке участвует существенная монотонная пороговая функция, то утверждение леммы 6.10 может быть усилено. Обозначим через MT^n множество монотонных пороговых функций от n переменных, а через $\widetilde{MT}^n$ — множество существенных функций из MT^n.

Лемма 6.11. *Если* $f_1 \in \widetilde{MT}^n$, $f_2 \in \widetilde{MT}^m$, $f_1(x_1, \ldots, x_n) \neq \mathop{\&}\limits_{i=1}^{n} x_i$, *то найдется такая пороговая функция* $f \in \widetilde{MT}^{n+m}$, *что:*

1) если $i \in \{1, \ldots, n\}$, *то* $L_i(f) \geqslant L_i(f_1)$, *2) если $i \in \{n+1, \ldots, n+m\}$, то найдется такое $j \in \{1, \ldots, n\}$, что*

$$L_i(f) \geqslant L_j(f_1).$$

Доказательство. Первое утверждение следует из леммы 6.10. Докажем второе утверждение.

В соответствии с леммой 6.8, каждая из функций f_1 и f_2 может быть задана граничным образом. Следовательно, найдутся линейные формы $l_1 \mapsto f_1$, $l_2 \mapsto f_2$ и наборы

$$\alpha = (a_1, \ldots, a_n) \in \{0,1\}^n;$$
$$\beta = (b_1, \ldots, b_m) \in \{0,1\}^m;$$

такие, что $l_1(\alpha) = l_2(\beta) = 0$.

Рассмотрим линейную форму

$$l(x_1, \ldots, x_{n+m}) = c_1 \cdot l_1(x_1, \ldots, x_n) + c_2 \cdot l_2(x_1, \ldots, x_m),$$

где $c_1, c_2 \in N$.

Очевидно, что вне зависимости от значений c_1 и c_2 линейная форма l задает некоторую пороговую функцию $f \in T^{n+m}$. При этом

$$\begin{cases} f(x_1, \ldots, x_n, b_1, \ldots, b_m) = f_1(x_1, \ldots, x_n), \\ f(a_1, \ldots, a_n, x_{n+1}, \ldots, x_{n+m}) = f_2(x_{n+1}, \ldots, x_{n+m}). \end{cases}$$

Напомним, что набор $\beta \in E_2^n$ называют нижней единицей монотонной функции $f(x_1, \ldots, x_n)$, если $f(\beta) = 1$, а на любом наборе $\beta' \in E_2^n$, $\beta' \leqslant \beta$, имеет место $f(\beta') = 0$.

Так как функция f_2 монотонная, то она однозначно задается множеством своих нижних единиц $\beta_1, \ldots, \beta_p$. Рассмотрим сокращенную дизъюнктивную нормальную форму

$$B = B_1 \vee B_2 \vee \ldots \vee B_p,$$

функции f_2, где B_i — простая импликанта, соответствующая нижней единице β_i. Импликанты дизъюнктивной нормальной формы B обладают следующим свойством: переменная x_j входит в импликанту B_i точно тогда, когда $\beta_j(i) = 1$.

Так как f_2 существенным образом зависит от всех своих переменных, то для всякого переменного x_i, $i = 1 \dots m$, найдется хотя бы одна импликанта B_j, $j \in \{1, \dots, p\}$, зависящая от x_i. Рассмотрим наборы

$$\beta_j = (b_1, \dots b_{i-1}, 1, b_{i+1}, \dots, b_m)\,,$$
$$\beta'_j = (b_1, \dots b_{i-1}, 0, b_{i+1}, \dots, b_m)\,.$$

Пусть $k \in \{1, \dots, m\}$. Рассмотрим произвольный набор γ из $\{\beta_1, \dots, \beta_p\}$ такой, что $\gamma\,(k) = 1$.

Положим γ' — набор соседний с γ, такой, что $\gamma\,(k) = 0$.

Возможны два случая: $l_2\,(\gamma) > 0$ и $l_2\,(\gamma) = 0$.

Рассмотрим первый случай. Пусть $l_2\,(\gamma) = d$. В таком случае всегда можно выбрать коэффициент c_2 достаточно большим так, чтобы для всякого набора переменных $x_1, \dots, x_n$ выполнялось следующее доопределение функции f

$$f\,(x_1, \dots, x_n, \gamma) = 1,$$
$$f\left(x_1, \dots, x_n, \gamma'\right) = 0.$$

В таком случае, для всякой линейной формы $l_{\vec{w},\sigma} \to f$ должно быть выполнено

$$\begin{cases} \left(\underbrace{0, \dots, 0}_{n \text{ раз}}, \gamma_1, \dots, \gamma_m\right) \cdot \vec{w} - \sigma \geqslant 0, \\ \left(\underbrace{1, \dots, 1}_{n \text{ раз}}, \gamma_1, \dots, \gamma_m\right) \cdot \vec{w} - \sigma - w_{n+k} < 0. \end{cases}$$

Следовательно, $w_{n+k} > \sum\limits_{i=1}^{n} w_i$ и получаем утверждение леммы.

Рассмотрим второй случай. Пусть $l_2\,(\gamma) = 0$. Можно выбрать коэффициент c_2 достаточно большим так, чтобы для всякого набора $x_1, \dots, x_n$ выполнялось следующее доопределение функции f

$$\begin{cases} f\,(\alpha, \gamma) = 1, \\ f\,(x_1, \dots, x_n, \gamma') = 0. \end{cases}$$

Здесь первое утверждение следует из того, что $l_1(\alpha) = l_2(\gamma) = 0$.

Таким образом, для всякой линейной формы $l_{\vec{w},\sigma} \to f$ должно быть выполнено

$$\begin{cases} ((\alpha_1,\ldots,\alpha_n,\gamma_1,\ldots,\gamma_m),\vec{w}) - \sigma \geqslant 0, \\ ((\alpha_1,\ldots,\alpha_n,\gamma_1,\ldots,\gamma_m),\vec{w}) - \sigma - w_{n+k} + \sum\limits_{j:\alpha_j=0} w_j < 0. \end{cases}$$

Так как $f_1 \neq \mathop{\&}\limits_{i=1}^{n} x_i$, то $\alpha \neq (1,\ldots,1)$, а, следовательно, найдется такое $j \in \{1,\ldots,n\}$, что $\alpha_j = 0$. Следовательно, $w_{n+k} > \sum\limits_{j:\alpha_j=0} w_j$ и лемма доказана.

Функцию f, существование которой было показано в лемме 6.11, обозначим через $f = f_1 \circ f_2$.

Пусть $n = 2^k$, где $k \geqslant 3$, тогда обозначим

$$MR_{nm} = \bigl\{ f(x_1,\ldots,x_{n+m}) = f_{H(n)}(x_1,\ldots,x_n) \circ g(x_{n+1},\ldots,x_{n+m}), \\ g \in \widetilde{MT}^m \bigr\}.$$

Данное определение корректно, так как $f_{H(n)}(x_1,\ldots,x_n) \neq \mathop{\&}\limits_{i=1}^{n} x_i$.

Отметим, что все функции, содержащиеся в классе MR_{nm}, суть монотонные.

Введем в рассмотрение следующий класс

$$R_{nm} = \bigcup_{\delta \in \{0,1\}^n} \delta\,[MR_{nm}].$$

Как видно, класс R_{nm} содержит класс MR_{nm}, а также все симметричные ему классы. Вычислим мощность класса R_{nm}.

Лемма 6.12. *При $m \to \infty$ имеет место следующее асимптотическое равенство*

$$|R_{nm}| \sim |T^m|.$$

Доказательство. По построению класса R_{nm} очевидно, что

$$|R_{nm}| = 2^m \cdot \left|\widetilde{MT}^m\right|.$$

Однако из лемм 2.5 и 5.3 следует, что при $m \to \infty$ выполнено

$$|R_{nm}| = 2^m \cdot \left|\widetilde{MT}^m\right| \sim 2^m \cdot |MT^m| = |T^m|\,.$$

Лемма доказана.

Оказывается, взаимная удаленность функций из класса R_{nm} «почти всегда» большая.

Лемма 6.13. *Если* $f_i, f_j \in R_{nm}$ *и* $s(f_i) \neq s(f_j)$, *то*

$$\rho_{ij} \geqslant \frac{1}{2n} e^{-4n^{\log_2\left(\frac{3}{2}\right)}} 2^{\frac{1}{2}n\log_2 n - n}.$$

Доказательство. Так как функции f_i и f_j существенные и их сигнатуры не равны, следовательно, для всех линейных форм $l_{\vec{w}',\sigma'} \to f_i$ и $l_{\vec{w}'',\sigma''} \to f_j$ найдется такое $k \in \{1, \dots, m\}$, что

$$sgn\left(w'_k\right) \neq sgn\left(w''_k\right).$$

А следовательно, либо

$$\rho\left(l_{\vec{w}',\sigma'}, l_{\vec{w}'',\sigma''}\right) \geqslant w'_k,$$

либо

$$\rho\left(l_{\vec{w}',\sigma'}, l_{\vec{w}'',\sigma''}\right) \geqslant w''_k.$$

В таком случае, из лемм 6.7 и 6.11 следует

$$\rho_{ij} = \rho\left(f_i, f_j\right) \geqslant \frac{1}{2n} e^{-4n^{\log_2\left(\frac{3}{2}\right)}} 2^{\frac{1}{2}n\log_2 n - n}.$$

Лемма доказана.

Оказывается, существует такой класс пороговых фукнций, что для почти всех пар функций из него близость между ними с ростом числа переменных растет экспоненциально. При этом мощность данного класса по порядку логарифма равна мощности класса всех пороговых функций.

Лемма 6.14. *Если n, m — натуральные числа, $n = 2^k$, где $k \geqslant 3$, то в матрице близости $D\,(n+m)$ содержится такое подмножество $R\,(n+m)$, что для всех $\rho_{ij} \in R\,(n+m)$ выполнено*

$$\rho_{ij} \geqslant \frac{1}{2n} e^{-4n^{\log_2\left(\frac{3}{2}\right)}} 2^{\frac{1}{2}n\log_2 n - n},$$

а при $m \to \infty$ имеет место следующее асимптотическое равенство

$$|R\,(n+m)| \sim |T^m|^2 .$$

Доказательство. В качестве множества $R\,(n+m)$ рассмотрим множество таких близостей ρ_{ij}, что $f_i, f_j \in R_{nm}$ и $s\,(f_i) \neq s\,(f_i)$. По лемме 6.13 для этих элементов будет выполнена оценка

$$\rho_{ij} \geqslant \frac{1}{2n} e^{-4n^{\log_2\left(\frac{3}{2}\right)}} 2^{\frac{1}{2}n\log_2 n - n}.$$

Для доказательства леммы осталось показать, что

$$|R\,(n+m)| \sim |T^m|^2 .$$

Так как всего существует 2^m сигнатур длины m, то мощность множества $R\,(n+m)$ можно оценить так

$$|R\,(n+m)| \sim \left(\frac{2^m - 1}{2^m} \cdot |T^m|\right)^2 = \left(1 - \frac{2}{2^m} + \frac{1}{2^{2m}}\right) \cdot |T^m|^2 \sim |T^m|^2 .$$

Лемма доказана.

Из леммы 6.14 вытекает важное следствие.

Следствие. *Если n, m — науральные числа, $n = 2^k$, то в матрице близости $D\,(n+m)$ содержится такое подмножество $R\,(n+m)$, что при $n, m \to \infty$*

$$|R\,(n+m)| \sim |T^m|^2,$$

для всех $\rho_{ij} \in R\,(n+m)$ выполнено

$$\log_2 \rho_{ij} \sim \log_2 L\,(n) .$$

Известна следующая оценка числа пороговых функций [4].

Теорема 17. *Если $n \to \infty$, то*

$$\log_2 |T^n| \sim n^2.$$

Из следствия леммы 6.14 и теоремы 17 вытекает утверждение теоремы 14.

Доказательство теоремы 14. Отметим, что n можно представить в виде $n = p + q$, где $p = 2^k$, и

$$p > \frac{1}{4}n,\ q > \frac{1}{2}n.$$

В таком случае в матрице близости $D(n)$ содержится такое подмножество $R(p+q)$, что

$$|R(p+q)| \sim |T^q|^2,$$

и для всех $\rho_{ij} \in R(p+q)$ выполнено

$$\log_2 \rho_{ij} \sim \log_2 L(p).$$

Оценим порядок логарифма мощности класса $R(p+q)$. Ясно, что

$$\log_2 |R(p+q)| > 2\log_2 \left|T^{\frac{1}{2}n}\right| \sim 2 \cdot \frac{1}{4}n^2 \asymp \log_2 |T^n| \asymp \log_2 |D(n)|.$$

Порядок логарифма величины элементов ρ_{ij} оценивается так:

$$\log_2 |\rho_{ij}| \sim \log_2 L(p) \sim \frac{1}{4} \cdot \frac{n}{4} \log_2 \frac{n}{4} \asymp n \log_2 n \asymp \log_2 \rho(n).$$

Теорема доказана.

Докажем утверждение теоремы 15.

Назовем $(-1,1)$-пороговую функцию $f^{-1,1}(y_1, \ldots, y_n)$ *центральной*, если существует линейная форма

$$l_{\vec{w}}(y_1, \ldots, y_n) = y_1 w_1 + \ldots + y_n w_n$$

такая, что

$$f^{-1,1}(y_1, \ldots, y_n) = \begin{cases} 1, & \text{если } \sum\limits_{i=1}^{n} y_i w_i \geqslant 0; \\ -1, & \text{иначе.} \end{cases}$$

Иными словами, центральная $(-1,1)$-пороговая функция может быть задана линейной формой с нулевым порогом. В связи с этим будем говорить, что вектор $\vec{w}$ задает $f^{-1,1}$ и обозначать это так $\vec{w} \to f^{-1,1}$. Линейные формы $l_{\vec{w}}$ с нулевым порогом будем называть *центральными.*

Ранее в разделе 2 было введено понятие веса пороговой функции. Введем аналогичную характеристику для центральных пороговых функций.

Назовем $l_{\vec{w}}$ минимальной центральной линейной формой, задающей $f^{0,1}$, если ее вес минимален среди всех центральных линейных форм, задающих $f^{0,1}$. Аналогичным образом вводится понятие минимальной центральной линейной формы, задающей $f^{-1,1}$.

Назовем *s-весом центральной пороговой функции* $f^{0,1}$ вес минимальной центральной линейной формы, задающей данную функцию. Обозначим s-вес пороговой функции так: $\mu_s\left(f^{0,1}\right)$. Аналогичным образом вводим понятие s-веса $(-1,1)$-пороговой функции $f^{-1,1}$ и обозначаем его $\mu_s\left(f^{-1,1}\right)$.

Следующее утверждение характеризует соотношение веса и s-веса центральных пороговых функций.

Лемма 6.15. *Если f — центральная пороговая функция, то*

$$\mu\left(f\right) \leqslant \mu_s\left(f\right).$$

Доказательство. Обозначим $U_s\left(f\right)$ — множество центральных линейных форм, задающих f. Очевидно, что $U_s\left(f\right) \subseteq U\left(f\right)$. Следовательно,

$$\mu\left(f\right) = \min_{l \in U(f)} \mu\left(l\right) \leqslant \min_{l \in U_s(f)} \mu\left(l\right) = \mu_s\left(f\right).$$

Лемма доказана.

Из лемм 6.15 и 2.1 вытекает следующее утверждение.

Лемма 6.16. *Если $f^{-1,1}\left(x_1, \ldots, x_n\right)$ центральная $(-1,1)$-пороговая функция, и $\mu_s\left(f^{-1,1}\right) = 2k$, где k — натуральное, то*

$$\mu\left(f^{0,1}\right) \leqslant \frac{3}{2}\mu_s\left(f^{-1,1}\right).$$

Напомним, что булеву функцию называют *существенной*, если она существенно зависит от всех своих переменных. Множество всех существенных пороговых функций от n переменных $x_1, \ldots, x_n$ обозначим $\widetilde{T}^n$.

В работе [14] был получен следующий результат, позволяющий строить существенные центральные пороговые функции с заранее заданным s-весом.

Лемма 6.17. *Если n и c четные, $n \leqslant c \leqslant 2^{\frac{n}{2}}$, то найдется существенная центральная $(-1,1)$-пороговая функция $f^{-1,1}$ такая, что $\mu_s\left(f^{-1,1}\right) = c$.*

Следующая лемма доказывает существование экспоненциального класса пороговых функций с близостью, ограниченной заранее заданной величиной.

Лемма 6.18. *Если n и c четные и $n \leqslant c \leqslant 2^{\frac{n}{2}}$, тогда существует класс M пороговых функций от n переменных, содержащий $(c-n)\cdot 2^{n-1}$ элементов, такой что для всех $f_i^{0,1}, f_j^{0,1}$ из M выполнено*

$$\rho\left(f_i^{0,1}, f_j^{0,1}\right) \leqslant 3c.$$

Доказательство. Если n и d удовлетворяют условиям леммы 6.17, то найдется такая центральная монотонная пороговая функция $f^{-1,1}$, что $\mu_s\left(f^{-1,1}\right) = d$, и, при этом, $\mu_s\left(f^{-1,1}\right)$ — четное.

По лемме 6.17 функция $f^{-1,1}$ является существенной. Тогда по лемме 6.16 получаем $\mu\left(f^{0,1}\right) \leqslant \frac{3}{2}d$, и $f^{0,1}$ также является монотонной и существенной.

Пусть M_d — множество функций, получающихся из $f^{0,1}$ применением всех возможных δ-преобразований. Так как $f^{0,1}$ существенная, то по лемме 2.4 все полученные таким образом функции различны. Поэтому мощность M_d равна 2^n.

Заметим также, что вес всякой δ-симметричной к $f^{0,1}$ пороговой функции не может превышать $\frac{3}{2}d$. А стало быть близость между произвольными δ-симметричными к $f^{0,1}$ пороговыми функциями не превосходит $3d$.

Зададим класс M как объединение классов M_d при всех возможных четных d в диапазоне от n до c. В таком случае мощность M равна $(c-n)\cdot 2^{n-1}$. Лемма доказана.

Обобщая лемму 6.18 на случай n и c произвольной четности, получаем теорему 15.

Доказательство теоремы 15. Так как число переменных n может быть нечетным, то для доказательства теоремы достаточно при построении класса M из леммы 6.18 заменить c на $(c-1)$ и добавить в функцию f одну несущественную переменную. Теорема доказана.

Список литературы

[1] Виноградов И. М. Математическая энциклопедия. М.: Советская энциклопедия, 1985.

[2] Грэхем Р., Кнут Д., Поташник О. Конкретная математика. М.: Мир, 1998.

[3] Золотых Н. Ю., Шевченко В. Н. Расшифровка пороговых функций k-значной логики // Дискретный анализ и исследование операций. Т. 2: 3. С. 18–23. 1995.

[4] Зуев Ю. А. Комбинаторно-вероятностные и геометрические методы в пороговой логике // Дискретная математика. Т. 3, вып. 2. С. 47–57. 1991.

[5] Кострикин А. И. Введение в алгебру. Линейная алгебра. М.: Физматлит, 2000.

[6] Маккаллок У. С., Питтс У. Логическое исчисление идей, относящихся к нервной активности // Автоматы. С. 362–384. 1956.

[7] Розенблатт Ф. Принципы нейродинамики (персептрон и теория механизмов мозга) // Автоматы. М.: Мир, 1965.

[8] Федорчук В. В. Курс аналитической геометрии и линейной алгебры. М.: МГУ, 1990.

[9] Хайкин С. Нейронные сети: полный курс. М.: Вильямс, 2006.

[10] Хачиян Л. Г. Полиномиальный алгоритм в линейном программировании // ДАН СССР. 244. С. 1093–1096. 1979.

[11] Хачиян Л. Г. Полиномиальные алгоритмы в линейном программировании // Вычисл. матем. и матем. физ. 20: 1. С. 51–68. 1980.

[12] Alon N., Vu V. H. Anti-Hadamar matrices, coin weighting, threshold gates and indecomposable hypergraphs // Journal of Combinatorial Theory. 79: 1. P. 133–160. 1997.

[13] Bliss T. V., Lomo T. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path // The Journal of Physiology. 232: 2. P. 331–356. 1973.

[14] Bohossian V., Bruck J. On Neural Networks with Minimal Weights // NIPS. P. 246–252. 1995.

[15] Citri A., Malenka R. C. Synaptic Plasticity: Multiple Forms, Functions, and Mechanisms // Neuropsychopharmacology. 33: 1. P. 1–24. 2008.

[16] Hastad J. On the size of weights for threshold gates // SIAM J. Discr. Math. 7. P. 484–492. 1994.

[17] Judd J. S. On the complexity of loading shallow neural networks // Journal of Complexity. 4. P. 177–192. 1988.

[18] Judd J. S. Neural Network Design and the Complexity of Learning. Cambridge, MA: MIT Press, 1990.

[19] Judd J. S. Why are Neural Networks so Wide // Aleksander, Taylor. 1. P. 45–52. 1992.

[20] Kahn J., Komlos J., Szemeredi E. On the probability that a random ± 1-matrix is singular // J. Amer. Math. Soc. 8: 1. P. 223–240. 1995.

[21] Maas W., Turan G. How fast can a threshold gate learn? // IIG-Report Ser. Rep. 321. Graz Univ. of Technology, 1991.

[22] Muroga S., Toda I., Takasu S. Theory of majority decision elements // J. Franklin Institute. 271: 5. P. 376–418. 1961.

[23] Muroga S. Threshold logic. New York: Wiley-Interscience, 1971.

[24] Novikoff A. B. On convergence proofs on perceptrons // Symposium on the Mathematical Theory of Automata. 12. P. 615–622. 1962.

[25] Kolen J. F., Pollack J. B. Back Propagation is Sensitive to Initial Conditions // Proceedings of the 1990 conference on Advances in neural information processing systems. P. 860–867. 1990.

[26] Parberry I. Circuit Complexity and Neural Networks. Cambridge, MA: MIT Press, 1994.

[27] Rumelhart D. E., Hinton G. E., Williams R. J. Learning Representation by Back-Propagating Errors // Nature. 323. P. 533–536. 1986.

[28] Yajima S., Ibaraki T. A lower bound on the number of threshold functions // IEEE Trans. Electronic Computers. 14. P. 926–929. 1965.

[29] Zucker R. S., Regehr W. G. Short-term synaptic plasticity // Annual Review of Physiology. 64. P. 355–405. 2002.