Math-Net.Ru

Общероссийский математический портал

Е. Д. Лившиц, Об оптимальности жадного алгоритма для некоторых классов функций, *Матем. сб.*, 2007, том 198, номер 5, 95–114

DOI: <https://doi.org/10.4213/sm1566>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением <http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.166.219.16

26 апреля 2023 г., 13:48:56

УДК 517.518.8 + 519.651.3

Е. Д. Лившиц

Об оптимальности жадного алгоритма для некоторых классов функций

В работе исследуется скорость сходимости чисто жадного алгоритма (ЧЖА). Получены оценки сверху на скорость сходимости ЧЖА в случае, если целевая функция принадлежит классам $\widehat{\mathcal{A}}_\gamma(\mathcal{D})$, $\gamma \geq 0$, которые являются расширениями класса $\widehat{\mathcal{A}}_1(\mathcal{D})$. Доказано, что полученная оценка сверху является точной по порядку при $\gamma \geq 2$.

Библиография: 14 названий.

§ 1. Введение

В последнее время в теории приближений увеличился интерес к нелинейным методам аппроксимации. Это обусловлено тем, что во многих прикладных задачах необходимо приблизить заданную целевую функцию линейной комбинацией функций из сильно переполненной системы и линейные методы аппроксимации в данной ситуации не всегда работают. Кроме того, возникает необходимость не только получать оценки скорости приближения, но и явно строить аппроксиманты, на которых эта скорость достигается. Эти два обстоятельства мотивировали изучение приближения функций с помощью жадных алгоритмов (greedy algorithms), которым и посвящена настоящая статья.

Рассмотрим действительное сепарабельное гильбертово пространство H со скалярным произведением $\langle \cdot, \cdot \rangle$; фиксируем множество $\mathcal{D}$, $\mathcal{D} \subset H$, состоящее из векторов единичной нормы, замыкание линейной оболочки которого совпадает со всем H :

$$g \in \mathcal{D} \implies \|g\| = 1, \quad \overline{\text{span}} \mathcal{D} = H.$$

Такие множества будем называть *словарями*. Как уже отмечалось, никаких условий на минимальность словаря $\mathcal{D}$ не налагается. В качестве словарей могут выступать как базисы, так и переполненные системы, например множество билинейных функций

$$H = L_2([0, 1]^2), \quad \mathcal{D} = \{f(x, y) = f_1(x)f_2(y) : f_1, f_2 \in L_2([0, 1]), \|f_1\| = \|f_2\| = 1\}$$

или множество плоских волн (ridge-функции)

$$H = L_2(\mathbb{R}^n), \quad \mathcal{D} = \{f(\bar{x}) = f_1(\bar{x} \bar{w}) : \bar{w} \in \mathbb{R}^n, f_1 \in L_2(\mathbb{R}), \|f\| = 1\}.$$

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты №№ 05-01-00062, 06-01-00160).

Рассматриваемой задачей является нахождение разложения целевой функции $f \in H$ по элементам словаря:

$$f \sim \sum_{k=1}^{\infty} c_k(f) g_k(f), \quad c_k(f) \in \mathbb{R}, \quad g_k(f) \in \mathcal{D},$$

частичные суммы которого $\sum_{k=1}^m c_k(f) g_k(f)$ достаточно хорошо приближают f . Искомое разложение может быть адаптивным: от функции зависят не только коэффициенты $c_k(f)$, но и элементы словаря $g_k(f)$.

Чисто жадный алгоритм (pure greedy algorithm) предлагает способ построения таких разложений.

Чисто жадный алгоритм (ЧЖА). Пусть заданы словарь $\mathcal{D}$ и функция $f_0 \in H$. Положим $G_0(f_0, \mathcal{D}) := 0$. Для каждого $m \geq 0$ последовательно выбираем вектор $g_{m+1} \in \mathcal{D}$ такой, что

$$|\langle f_m, g_{m+1} \rangle| = \sup_{g \in \mathcal{D}} |\langle f_m, g \rangle|, \quad (1)$$

и определяем

$$\begin{aligned} G_{m+1}(f_0, \mathcal{D}) &:= G_m(f_0, \mathcal{D}) + \langle f_m, g_{m+1} \rangle g_{m+1}, \\ f_{m+1} &:= f - G_{m+1}(f_0, \mathcal{D}) = f_m - \langle f_m, g_{m+1} \rangle g_{m+1}. \end{aligned} \quad (2)$$

Ряд

$$\sum_{m=0}^{\infty} \langle f_m, g_{m+1} \rangle g_{m+1}$$

называют PGA-разложением вектора f_0 по словарю $\mathcal{D}$.

Идеи, близкие к чисто жадному алгоритму, использовались при решении конкретных задач в различных областях математики. Вероятно, что впервые они были применены Е. Шмидтом [1], который в 1907 г. использовал для приближения функций из $L_2([0, 1]^2)$ билинейными функциями процесс, совпадающий с ЧЖА. В 80-х годах жадные алгоритмы стали активно использоваться в прикладных задачах, они вошли в теорию передачи сигналов под именем matching pursuit [2] и в статистику как projection pursuit regression [3] (в качестве словарей рассматривалось множество плоских волн). В 90-е годы соответствующие прикладные методы были формализованы и начали изучаться в рамках теории функций [4]–[6]. Но при прямом переносе алгоритмов из приложений, где пространство конечномерно, а словари конечны, в теорию функций, где изучаются бесконечные словари в бесконечномерных пространствах, возникают две трудности. Первая состоит в том, что супремум в формуле (1) может не достигаться ни на каком элементе словаря. Эта проблема может быть обойдена с помощью слабого жадного алгоритма [7]. В нем условие (1) заменяется на следующее:

$$g_{m+1} - \text{любой элемент из } \mathcal{D}, \text{ для которого } |\langle f_m, g_{m+1} \rangle| \geq t_m \sup_{g \in \mathcal{D}} |\langle f_m, g \rangle|, \quad (3)$$

где $\{t_m\}_{m=1}^{\infty} \subset [0, 1)$ – фиксированная последовательность.

Вторая трудность состоит в том, что в формуле (1) супремум берется по бесконечному множеству. Она может быть преодолена с помощью жадных алгоритмов с ограниченной глубиной поиска [8], в них (1) заменяется на следующее:

$$g_{m+1} - \text{любой элемент из } \mathcal{D}_{m+1}, \text{ для которого } |\langle f_m, g_{m+1} \rangle| \geq t_m \sup_{g \in \mathcal{D}_{m+1}} |\langle f_m, g \rangle|,$$

где $\{\mathcal{D}_m\}_{m=1}^\infty \subset \mathcal{D}$ – система расширяющихся подмножеств $\mathcal{D}$, $\bigcup_1^\infty \mathcal{D}_m = \mathcal{D}$. (Ясно, что если словарь $\mathcal{D}$ счетен, то можно выбрать конечные подмножества $\mathcal{D}_m$.)

В настоящей статье будет исследоваться исходный чисто жадный алгоритм, задаваемый соотношениями (1) и (2). В [4] Л. Джонс доказал, что ЧЖА сходится для всех словарей $\mathcal{D}$ и целевых функций f_0 , т.е.

$$G_m(f_0, \mathcal{D}) \rightarrow f_0, \quad f_m \rightarrow 0 \quad \text{при } m \rightarrow \infty.$$

Мы будем изучать вопрос скорости сходимости ЧЖА.

Для понимания того, каким может быть приближение заданной функции линейной комбинацией m элементов словаря, удобно использовать величину наилучшего m -членного приближения

$$\sigma_m(f, \mathcal{D}) = \inf_{c_1, \dots, c_m \in \mathbb{R}, g_1, \dots, g_m \in \mathcal{D}} \left\| f - \sum_{k=1}^m c_k g_k \right\|.$$

Ясно, что нельзя дать общих оценок на скорость сходимости ЧЖА (PGA-разложений) для произвольных словаря $\mathcal{D}$ и целевой функции f_0 . Поэтому обычно ставится вопрос о получении оценок скорости сходимости ЧЖА в случае, если целевая функция f_0 принадлежит некоторому классу, зависящему от $\mathcal{D}$. В идеале было бы получить результаты, которые по скорости убывания наилучшего m -членного приближения давали оценку на скорость убывания m -членного приближения, полученного с помощью ЧЖА. Но выяснилось, что класс функций с фиксированной скоростью убывания наилучшего m -членного приближения является слишком широким.

ТЕОРЕМА А [9]. *Для произвольных $\gamma, \delta > 0$ существуют словарь $\mathcal{D}$, функция $f_0 \in H$ и $C > 0$ такие, что для всех $m > 1$ выполняются неравенства*

$$\sigma_m(f, \mathcal{D}) \leq m^{-\gamma}, \quad \|f - G_m(f, \mathcal{D})\| \geq C m^{-\delta}.$$

В большинстве работ по жадным алгоритмам в качестве класса целевых функций выбирается класс $\mathcal{A}_1(\mathcal{D})$ – замыкание выпуклой оболочки векторов, пропорциональных векторам из словаря $\mathcal{D}$.

Более точно, сначала определяется класс функций

$$\mathcal{A}_1^0(\mathcal{D}, M) := \left\{ f \in H : f = \sum_{\lambda \in \Lambda} c_\lambda w_\lambda, \quad w_\lambda \in \mathcal{D}, \quad \#\Lambda < \infty, \quad \sum_{\lambda \in \Lambda} |c_\lambda| \leq M \right\}$$

и класс $\mathcal{A}_1(\mathcal{D}, M)$, являющийся замыканием (в H) класса $\mathcal{A}_1^0(\mathcal{D}, M)$. Далее, определим $\mathcal{A}_1(\mathcal{D}) := \bigcup_{M \geq 0} \mathcal{A}_1(\mathcal{D}, M)$. Для $f \in \mathcal{A}_1(\mathcal{D})$ введем норму

$$|f|_{\mathcal{A}_1(\mathcal{D})} = \inf \{M \geq 0 : f \in \mathcal{A}_1(\mathcal{D}, M)\}.$$

Для этого класса известна скорость убывания m -членных приближений. Оценка сверху была получена С. Б. Стечкиным [10] в случае ортонормированного словаря и А. Бэрроном [11] для произвольного словаря.

ТЕОРЕМА В [11]. Для произвольных словаря $\mathcal{D}$, элемента $f \in \mathcal{A}_1(\mathcal{D})$ и $m \geq 1$ выполняется оценка

$$\sigma_m(f, \mathcal{D}) \leq |f|_{\mathcal{A}_1(\mathcal{D})} \frac{1}{m^{1/2}}. \quad (4)$$

При этом, как следует из результатов С. Б. Стечкина, показатель $1/2$ не может быть увеличен даже для ортонормированного словаря.

ТЕОРЕМА С [10]. Пусть $\mathcal{D}$ – ортонормированный словарь. Для произвольного $\varepsilon > 0$ существуют такие $C > 0$ и элемент $f \in \mathcal{A}_1(\mathcal{D})$, что для всех $m \geq 1$ имеет место неравенство

$$\sigma_m(f, \mathcal{D}) > C |f|_{\mathcal{A}_1(\mathcal{D})} \frac{1}{m^{1/2+\varepsilon}}.$$

Вопрос о скорости сходимости ЧЖА в классе $\mathcal{A}_1(\mathcal{D})$ впервые начали изучать Р. ДеВор и В. Н. Темляков.

ТЕОРЕМА D [5]. Для произвольных словаря $\mathcal{D}$, элемента $f \in \mathcal{A}_1(\mathcal{D})$ и $m \geq 1$ выполняется оценка

$$\|f - G_m(f, \mathcal{D})\| \leq |f|_{\mathcal{A}_1(\mathcal{D})} m^{-1/6}.$$

Далее эта оценка была улучшена С. В. Конягиным и В. Н. Темляковым.

ТЕОРЕМА Е [12]. Для произвольных словаря $\mathcal{D}$, элемента $f \in \mathcal{A}_1(\mathcal{D})$ и $m \geq 1$ справедливо

$$\|f - G_m(f, \mathcal{D})\| \leq 4 |f|_{\mathcal{A}_1(\mathcal{D})} m^{-11/62}.$$

Наилучшая в настоящий момент оценка скорости сходимости ЧЖА для класса $\mathcal{A}_1(\mathcal{D})$ принадлежит А. В. Сильниченко.

ТЕОРЕМА F [13]. Существует такое $C > 0$, что для произвольных словаря $\mathcal{D}$, элемента $f \in \mathcal{A}_1(\mathcal{D})$ и $m \geq 1$ имеет место неравенство

$$\|f - G_m(f, \mathcal{D})\| \leq C |f|_{\mathcal{A}_1(\mathcal{D})} m^{-0.182}.$$

В то же время гипотеза о том, что скорость сходимости ЧЖА будет близка к скорости убывания m -членных приближений (4), не подтвердилась. Оценка снизу была получена автором и В. Н. Темляковым.

ТЕОРЕМА G [14]. *Существуют такой словарь $\mathcal{D}$, элемент $f \in \mathcal{A}_1(\mathcal{D})$ и $C > 0$, что для произвольного $m \geq 1$ справедлива оценка*

$$\|f - G_m(f, \mathcal{D})\| > C m^{-0.27} |f|_{\mathcal{A}_1(\mathcal{D})}.$$

В работе [9] рассматривались расширения класса $\mathcal{A}_1(\mathcal{D})$, в которых ограничивались как порядок приближения целевой функции некоторой m -членной комбинацией элементов $\mathcal{D}$, так и $\mathcal{A}_1(\mathcal{D})$ -норма $(|\cdot|_{\mathcal{A}_1(\mathcal{D})})$ этой комбинации. Для указанного 2-параметрического семейства классов были получены оценки сверху на скорость сходимости ЧЖА. (Эти оценки не были точны по порядку.)

В этой работе мы развиваем технику из [9]. Рассматривается однопараметрическое семейство классов $\widehat{\mathcal{A}}_\gamma(\mathcal{D})$, для которых получены оценки сверху, оказывающиеся при некоторых γ точными по порядку.

Пусть заданы $\gamma \geq 0$, $M \geq 0$. Рассмотрим множества

$$\widehat{\mathcal{A}}_\gamma(\mathcal{D}, M) = \left\{ f \in H : \forall m \geq 1 \exists \text{ конечное } \Lambda, c_\lambda \in \mathbb{R}, g_\lambda \in \mathcal{D}, \lambda \in \Lambda, \text{ такие,} \right. \\ \left. \text{что } \left\| f - \sum_{\lambda \in \Lambda} c_\lambda g_\lambda \right\| \leq \frac{M}{m}, \sum_{\lambda \in \Lambda} |c_\lambda| \leq M m^\gamma \right\}.$$

Далее, определим $\widehat{\mathcal{A}}_\gamma(\mathcal{D})$ и норму $|f|_{\widehat{\mathcal{A}}_\gamma(\mathcal{D})}$ на нем:

$$\widehat{\mathcal{A}}_\gamma(\mathcal{D}) = \bigcup_{M \geq 0} \widehat{\mathcal{A}}_\gamma(\mathcal{D}, M), \\ |f|_{\widehat{\mathcal{A}}_\gamma(\mathcal{D})} = \inf \{ M \geq 0 : f \in \widehat{\mathcal{A}}_\gamma(\mathcal{D}, M) \}, \quad f \in \widehat{\mathcal{A}}_\gamma(\mathcal{D}).$$

Легко видеть, что $\widehat{\mathcal{A}}_0(\mathcal{D}) = \mathcal{A}_1(\mathcal{D})$ и $|f|_{\widehat{\mathcal{A}}_0(\mathcal{D})} = |f|_{\mathcal{A}_1(\mathcal{D})}$ для всех $f \in \mathcal{A}_1(\mathcal{D})$. Имеет место следующая оценка сверху на скорость сходимости ЧЖА в классах $\widehat{\mathcal{A}}_\gamma(\mathcal{D})$.

ТЕОРЕМА 1. *Пусть даны $\gamma \geq 0$ и ε , $0 < \varepsilon < 1/(2(1+\gamma))$. Тогда существует такое $C > 0$, что для всех $f \in \widehat{\mathcal{A}}_\gamma(\mathcal{D})$ и $m \geq 1$ выполняется неравенство*

$$\|f - G_m(f, \mathcal{D})\| \leq C |f|_{\widehat{\mathcal{A}}_\gamma(\mathcal{D})} m^{-\min\{1/6, 1/(2(1+\gamma))-\varepsilon\}}. \quad (5)$$

Отметим, что в случае $\gamma = 0$ оценка (5) совпадает по порядку с оценкой из теоремы D и уступает более сильным оценкам из теорем E и F. Для $\gamma > 2$ теорема 1 дает оценку точную по порядку. Доказан аналог теоремы C.

ТЕОРЕМА 2. *Пусть $\mathcal{D}$ – ортонормированный словарь. Для произвольного $\gamma > 0$ существуют такие $C > 0$ и элемент $f \in \widehat{\mathcal{A}}_\gamma(\mathcal{D})$, что для всех $m \geq 1$ имеет место неравенство*

$$\sigma_m(f, \mathcal{D}) > C |f|_{\widehat{\mathcal{A}}_\gamma(\mathcal{D})} m^{-1/(2(1+\gamma))}.$$

ЗАМЕЧАНИЕ 1. Легко видеть, что оценки сверху и снизу на скорость сходимости ЧЖА совпадают по порядку при $\gamma \geq 2$.

ЗАМЕЧАНИЕ 2. Техника, используемая при доказательстве теоремы 1, может быть применена для получения оценок скорости сходимости слабого жадного алгоритма (он отличается от ЧЖА заменой условия (1) на условие (3)). Получаемые оценки будут зависеть от параметров $\{t_m\}_{m=1}^{\infty}$ и будут ухудшаться при уменьшении t_m . Тем не менее, по заданному $\varepsilon > 0$ можно подобрать такое близкое к 1 число $t(\varepsilon)$, что для слабого жадного алгоритма с параметрами $t_m = t(\varepsilon)$, $m \geq 1$, будет также иметь место оценка (5).

§ 2. Численные неравенства

Пусть задана убывающая последовательность таких положительных чисел $\{a_n\}_{n=0}^{\infty}$, что выполняются неравенства

$$1 \geq a_0 \geq a_1 \geq \dots > 0. \quad (6)$$

Зададим последовательности $\{d_n\}_{n=0}^{\infty}$ и $\{b_n\}_{n=0}^{\infty}$ следующими формулами:

$$d_n = (a_n - a_{n+1})^{1/2}, \quad n \geq 0, \quad (7)$$

$$b_0 = 0, \quad b_{n+1} = \sum_{k=0}^n d_k, \quad n \geq 0. \quad (8)$$

ТЕОРЕМА 3. Пусть задано число α , $0 < \alpha < 1$. Тогда существует $C(\alpha) > 0$, обладающее следующим свойством: для любых $C \geq C(\alpha)$, $k, l \geq 2$, $k+1 < l$, и последовательностей $\{a_n\}_{n=0}^{\infty}$, $\{d_n\}_{n=0}^{\infty}$ и $\{b_n\}_{n=0}^{\infty}$, удовлетворяющих соотношениям (6), (7) и (8) соответственно и неравенствам

$$a_k \geq k^{-\alpha}, \quad (9)$$

$$a_j < j^{-\alpha}, \quad k+1 \leq j \leq l-1, \quad (10)$$

$$a_l \geq l^{-\alpha}, \quad (11)$$

имеет место импликация

$$a_k b_k \leq C k^{\max\{0, 1/2-3\alpha/2\}} \implies a_l b_l \leq C l^{\max\{0, 1/2-3\alpha/2\}}. \quad (12)$$

Рассмотрим произвольные числа $u > v$ и функцию $\psi_{u,v} \in C[v, u]$,

$$\psi_{u,v}(t) = (u-t)^{1/2} + (t-v)^{1/2}, \quad t \in [v, u]. \quad (13)$$

Ясно, что

$$\psi_{u,v}(\cdot) \text{ возрастает на } \left[v, \frac{u+v}{2}\right) \text{ и убывает на } \left(\frac{u+v}{2}, u\right]. \quad (14)$$

Докажем сначала следующую лемму.

ЛЕММА 1. Пусть заданы число α , $0 < \alpha < 1$, числа $k, l \geq 1$, $k + 1 < l$, и a_k , $k^{-\alpha} \leq a_k \leq 2$. Тогда

$$\begin{aligned} & \sup \left\{ \sum_{j=k}^{l-1} (a_j - a_{j+1})^{1/2} : a_k \geq a_{k+1} \geq \dots \geq a_l, \right. \\ & \quad \left. a_j \leq j^{-\alpha}, \quad k+1 \leq j \leq l-1, \quad a_l \geq l^{-\alpha} \right\} \\ & = (a_k - (k+1)^{-\alpha})^{1/2} + \sum_{j=k+1}^{l-1} (j^{-\alpha} - (j+1)^{-\alpha})^{1/2}. \end{aligned}$$

ДОКАЗАТЕЛЬСТВО. Ясно, что множество допустимых значений наборов $(a_{k+1}, \dots, a_l)$ является компактом и по теореме Вейерштрасса непрерывная функция

$$\sum_{j=k}^{l-1} (a_j - a_{j+1})^{1/2} \quad (15)$$

достигает на нем максимума в точке $(\hat{a}_{k+1}, \dots, \hat{a}_l)$. Положим $\hat{a}_k = a_k$. Из невозрастания a_j и неравенства $a_l \geq l^{-\alpha}$ следует, что $\hat{a}_l = l^{-\alpha}$. Докажем, что

$$\hat{a}_j = j^{-\alpha}, \quad k+1 \leq j \leq l. \quad (16)$$

Если это не так, то рассмотрим наибольший индекс p , $p < l$, где (16) не выполняется и $\hat{a}_p < p^{-\alpha}$. Тогда в силу выпуклости функции $t^{-\alpha}$ имеем, что

$$0 \leq s := \hat{a}_p - \hat{a}_{p+1} < p^{-\alpha} - \hat{a}_{p+1} = p^{-\alpha} - (p+1)^{-\alpha} \leq q^{-\alpha} - (q+1)^{-\alpha} \quad (17)$$

для всех q , $k \leq q \leq p$. Докажем индукцией по q , что при $q \leq p$ имеют место соотношения $\hat{a}_q < q^{-\alpha}$ и $\hat{a}_q - \hat{a}_{q+1} = s$. База для $q = p$ непосредственно следует из (17). Предположим, что индукционная гипотеза верна для q , $k < q \leq p$, проверим ее для $q-1$. Фиксируем в функции (15) переменные $a_j = \hat{a}_j$ при $j \neq q$. Тогда получается, что $\hat{a}_q$ максимизирует функцию $\psi_{\hat{a}_{q-1}, \hat{a}_{q+1}}(t)$ (см. (13)) при ограничениях $\hat{a}_{q-1} \geq t \geq \hat{a}_{q+1}$, $t \leq q^{-\alpha}$. Но так как по предположению индукции $\hat{a}_q < q^{-\alpha}$, то согласно (14) и индукционному предположению выполняются равенства

$$\hat{a}_{q-1} - \hat{a}_q = \hat{a}_q - \hat{a}_{q+1} = s.$$

И наконец, в силу индукционного предположения и (17) имеем

$$\hat{a}_{q-1} = \hat{a}_q + s < q^{-\alpha} + ((q-1)^{-\alpha} - q^{-\alpha}) = (q-1)^{-\alpha}.$$

Это доказывает индукционную гипотезу для всех q , $k \leq q \leq p$, и, следовательно, имеет место неравенство $a_k = \hat{a}_k < k^{-\alpha}$, что противоречит условиям леммы. Следовательно, соотношения (16) истинны. Из них непосредственно вытекает утверждение леммы.

ЛЕММА 2. Пусть задано число α , $0 < \alpha < 1$. Тогда существует $C_1(\alpha) > 0$ такое, что для всех $C_1 \geq C_1(\alpha)$ и $k \geq 2$

$$\max_{t \in [k^{-\alpha}, 1]} \chi(t) = \chi(k^{-\alpha}), \quad (18)$$

зде

$$\chi(t) = C_1 k^{\max\{0, 1/2-3\alpha/2\}} \frac{1}{t} + (t - (k+1)^{-\alpha})^{1/2}.$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим функцию $\tilde{\chi} \in C[\mathbb{R}]$

$$\tilde{\chi}(v) = C_1 k^{\max\{0, 1/2-3\alpha/2\}} \frac{1}{v^2 + (k+1)^{-\alpha}} + v. \quad (19)$$

Ясно, что $\chi(v^2 + (k+1)^{-\alpha}) = \tilde{\chi}(v)$, $v \in [0, 1]$. Поэтому равенство (18) вытекает из равенства

$$\max_{v \in [v_0, 1]} \tilde{\chi}(v) - \tilde{\chi}(v_0) = 0,$$

где

$$v_0 = (k^{-\alpha} - (k+1)^{-\alpha})^{1/2}, \quad (20)$$

что, в свою очередь, эквивалентно

$$\hat{\chi}(v) := C_1 k^{\max\{0, 1/2-3\alpha/2\}} + (v^2 + (k+1)^{-\alpha})v - \tilde{\chi}(v_0)(v^2 + (k+1)^{-\alpha}) \leq 0, \quad v \in [v_0, 1]. \quad (21)$$

Так как $\hat{\chi}(\cdot)$ является полиномом третьей степени с положительным коэффициентом при старшем мономе и при этом $\hat{\chi}(v_0) = 0$, то неравенство (21) (и тем самым утверждение леммы) будет вытекать из неравенств

$$\frac{d}{dv} \hat{\chi}(v_0) \leq 0, \quad \hat{\chi}(1) \leq 0. \quad (22)$$

Из (19) и (20) непосредственно следует, что

$$\tilde{\chi}(v_0) = C_1 k^{\max\{0, 1/2-3\alpha/2\}} k^\alpha + v_0 \geq C_1 k^{\max\{0, 1/2-3\alpha/2\}} k^\alpha. \quad (23)$$

Оценим $d\hat{\chi}(v_0)/dv$:

$$\frac{d}{dv} \hat{\chi}(v_0) = 3v_0^2 + (k+1)^{-\alpha} - \tilde{\chi}(v_0)2v_0. \quad (24)$$

Из (20) легко следует, что найдется такая постоянная $\check{C}(\alpha) > 0$, что для всех $k \geq 2$ имеет место неравенство $v_0 \geq \check{C}(\alpha)k^{-1/2-\alpha/2}$. Подставляя это неравенство вместе с (20) и (23) в (24), получаем

$$\begin{aligned} \frac{d}{dv} \hat{\chi}(v_0) &\leq 3k^{-\alpha} - 2(k+1)^{-\alpha} - C_1 k^{\max\{0, 1/2-3\alpha/2\}} k^\alpha 2\check{C}(\alpha)k^{-1/2-\alpha/2} \\ &\leq 3k^{-\alpha} - 2C_1 \check{C}(\alpha)k^{\max\{-1/2+\alpha/2, -\alpha\}} \leq k^{-\alpha}(3 - 2C_1 \check{C}(\alpha)). \end{aligned} \quad (25)$$

Далее, с учетом (23) оценим $\hat{\chi}(1)$:

$$\begin{aligned} \hat{\chi}(1) &= C_1 k^{\max\{0, 1/2-3\alpha/2\}} + (1 + (k+1)^{-\alpha}) - \tilde{\chi}(v_0)(1 + (k+1)^{-\alpha}) \\ &\leq C_1 k^{\max\{0, 1/2-3\alpha/2\}} + 2 - C_1 k^{\max\{0, 1/2-3\alpha/2\}} k^\alpha \\ &= 2 + C_1 k^{\max\{0, 1/2-3\alpha/2\}}(1 - k^\alpha). \end{aligned} \quad (26)$$

Так как $k \geq 2$, то найдется такое $C_1(\alpha) > 0$, что для всех $C_1 \geq C_1(\alpha)$ из неравенств (25) и (26) будут вытекать неравенства (22), что доказывает лемму.

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 3. Пусть $a_k b_k \leq C k^{\max\{0, 1/2-3\alpha/2\}}$, тогда

$$b_k \leq C k^{\max\{0, 1/2-3\alpha/2\}} \frac{1}{a_k},$$

$$b_l \leq C k^{\max\{0, 1/2-3\alpha/2\}} \frac{1}{a_k} + \sum_{j=k}^{l-1} (a_j - a_{j+1})^{1/2}.$$

Из соотношений (9)–(11) и леммы 1 вытекает, что

$$b_l \leq C k^{\max\{0, 1/2-3\alpha/2\}} \frac{1}{a_k} + (a_k - (k+1)^{-\alpha})^{1/2} + \sum_{j=k+1}^{l-1} (j^{-\alpha} - (j+1)^{-\alpha})^{1/2}.$$

Далее, по лемме 2 для $C > C_1(\alpha)$

$$b_l \leq C k^{\max\{0, 1/2-3\alpha/2\}} \frac{1}{k^{-\alpha}} + (k^{-\alpha} - (k+1)^{-\alpha})^{1/2} + \sum_{j=k+1}^{l-1} (j^{-\alpha} - (j+1)^{-\alpha})^{1/2}$$

$$= C k^{\max\{\alpha, 1/2-\alpha/2\}} + \sum_{j=k}^{l-1} (j^{-\alpha} - (j+1)^{-\alpha})^{1/2}.$$

Из (10) для $j = l-1$ вытекает, что $a_l \leq a_{l-1} \leq (l-1)^{-\alpha}$. Таким образом, для доказательства теоремы 3 надо проверить, что для больших C выполняется оценка

$$a_l b_l \leq (l-1)^{-\alpha} \left(C k^{\max\{\alpha, 1/2-\alpha/2\}} + \sum_{j=k}^{l-1} (j^{-\alpha} - (j+1)^{-\alpha})^{1/2} \right) \leq C l^{\max\{0, 1/2-3\alpha/2\}}. \quad (27)$$

Легко видеть, что найдется такое $C_2 = C_2(\alpha) > 0$, что для любых $k, l, 2 \leq k \leq l-2 < l$, выполняется

$$\sum_{j=k}^{l-1} (j^{-\alpha} - (j+1)^{-\alpha})^{1/2} \leq C_2 (l^{1/2-\alpha/2} - k^{1/2-\alpha/2}).$$

Тогда для $C > C_2$

$$C k^{\max\{\alpha, 1/2-\alpha/2\}} + \sum_{j=k}^{l-1} (j^{-\alpha} - (j+1)^{-\alpha})^{1/2}$$

$$\leq C k^{\max\{\alpha, 1/2-\alpha/2\}} + C_2 (l^{1/2-\alpha/2} - k^{1/2-\alpha/2})$$

$$= C_2 l^{1/2-\alpha/2} + (C - C_2) k^{\max\{\alpha, 1/2-\alpha/2\}}$$

$$\leq C_2 l^{1/2-\alpha/2} + (C - C_2) (l-2)^{\max\{\alpha, 1/2-\alpha/2\}}.$$

Для доказательства неравенства (27) достаточно проверить оценки

$$\frac{(l-1)^{-\alpha} (C_2 l^{1/2-\alpha/2} + (C - C_2) (l-2)^{\max\{\alpha, 1/2-\alpha/2\}})}{C l^{\max\{0, 1/2-3\alpha/2\}}} \leq 1,$$

$$\left(1 - \frac{1}{l}\right)^{-\alpha} \left(\frac{C_2 l^{1/2-\alpha/2}}{C l^{\max\{\alpha, 1/2-\alpha/2\}}} + \frac{C - C_2}{C} \left(1 - \frac{2}{l}\right)^{\max\{\alpha, 1/2-\alpha/2\}} \right) \leq 1,$$

которые вытекают из неравенства

$$\left(1 - \frac{1}{l}\right)^{-\alpha} \left(\frac{C_2}{C} + \frac{C - C_2}{C} \left(1 - \frac{2}{l}\right)^\alpha\right) \leq 1.$$

Ясно, что последнее неравенство выполняется для $C \geq C_3(\alpha, C_2(\alpha))$. Тем самым, $C(\alpha) = \max\{C_1(\alpha), C_2(\alpha), C_3(\alpha)\}$ удовлетворяет условиям теоремы.

В качестве следствия теоремы 3 докажем следующую теорему.

ТЕОРЕМА 4. Пусть заданы число α , $0 < \alpha < 1$, и последовательности $\{a_n\}_{n=0}^\infty$, $\{d_n\}_{n=0}^\infty$ и $\{b_n\}_{n=0}^\infty$, удовлетворяющие соотношениям (6), (7) и (8) соответственно. Тогда существует такое $\bar{C}(\alpha) > 0$, что для всех $l \geq 4$ из неравенств

$$a_0 \leq 1, \quad a_j < j^{-\alpha}, \quad 1 \leq j \leq l-1, \quad (28)$$

$$a_l \geq l^{-\alpha}, \quad (29)$$

следует оценка

$$a_l b_l \leq \bar{C}(\alpha) l^{\max\{0, 1/2-3\alpha/2\}}. \quad (30)$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим последовательность $\{\tilde{a}_n\}_{n=0}^\infty$, удовлетворяющую равенствам $\tilde{a}_0 = \tilde{a}_1 = \tilde{a}_2 = 2^{-\alpha}$, $\tilde{a}_j = a_j$, $j \geq 3$. Определим последовательности $\{\tilde{d}_n\}_{n=0}^\infty$ и $\{\tilde{b}_n\}_{n=0}^\infty$, используя соотношения (7) и (8). Тогда

$$\tilde{b}_2 = 0, \quad b_l = \tilde{b}_l + d_0 + d_1 + d_2 \leq \tilde{b}_l + 3, \quad l \geq 3.$$

Применяя теорему 3 к последовательностям $\{\tilde{a}_n\}_{n=0}^\infty$, $\{\tilde{d}_n\}_{n=0}^\infty$, $\{\tilde{b}_n\}_{n=0}^\infty$, $k = 2$ и $C = C(\alpha)$, получаем

$$a_l \tilde{b}_l = \tilde{a}_l \tilde{b}_l \leq C(\alpha) l^{\max\{0, 1/2-3\alpha/2\}}.$$

Следовательно,

$$a_l b_l \leq a_l (\tilde{b}_l + 3) \leq C(\alpha) l^{\max\{0, 1/2-3\alpha/2\}} + 3a_l \leq (C(\alpha) + 3) l^{\max\{0, 1/2-3\alpha/2\}},$$

что доказывает теорему для $\bar{C}(\alpha) = C(\alpha) + 3$.

§ 3. Основные определения и неравенства для ЧЖА

Предположим, что заданы словарь $\mathcal{D}$ и целевая функция $f_0 := f$ и по ним проводится ЧЖА, т.е. по формулам (1) и (2) определяются последовательности $\{g_n\}_{n=0}^\infty \subset \mathcal{D}$ и $\{f_n\}_{n=1}^\infty \subset H$. Следуя [5], введем стандартные обозначения, $n \geq 0$:

$$a_n = \langle f_n, f_n \rangle, \quad (31)$$

$$d_n = |\langle f_n, g_{n+1} \rangle| = \sup_{g \in \mathcal{D}} |\langle f_n, g \rangle|, \quad (32)$$

$$b_{n+1} = \sum_{k=0}^n d_n. \quad (33)$$

Из (1) и (2) вытекают равенства

$$\begin{aligned} a_{n+1} &= \langle f_{n+1}, f_{n+1} \rangle = \langle f_n, f_n \rangle - (\langle f_n, g_{n+1} \rangle)^2 = a_n - d_n^2, \\ d_n^2 &= a_n - a_{n+1}, \end{aligned} \quad (34)$$

$$f_n = f_0 - G_n(f_0, \mathcal{D}) = f_0 - \sum_{k=0}^{n-1} \langle f_k, g_{k+1} \rangle g_{k+1}. \quad (35)$$

Отметим также, что в силу равенства $\overline{\text{span}} \mathcal{D} = H$ для всех $n \geq 0$ будет выполняться неравенство $d_n > 0$, и поэтому последовательность $\{a_n\}_{n=0}^{\infty}$ строго убывает.

Пусть $f_0 \in \widehat{\mathcal{A}}_1(\mathcal{D}, 1)$. Тогда согласно определению класса $\widehat{\mathcal{A}}_1(\mathcal{D}, 1)$ для $m = [n^{1/(2(1+\gamma))}]$ найдутся конечное множество индексов Λ , числа $c_\lambda \in \mathbb{R}$ и элементы $g_\lambda \in \mathcal{D}$, $\lambda \in \Lambda$, для которых

$$\begin{aligned} \|r_n\| &\leq m^{-1} \leq 2n^{-1/(2(1+\gamma))}, \quad r_n := f_0 - \sum_{\lambda \in \Lambda} c_\lambda g_\lambda; \\ \sum_{\lambda \in \Lambda} |c_\lambda| &\leq m^\gamma \leq n^{\gamma/(2(1+\gamma))}. \end{aligned}$$

Используя соотношения (31)–(33) и (35) вместе с последними двумя неравенствами, получаем

$$\begin{aligned} a_n &= \langle f_n, f_n \rangle = \left\langle f_n, f_0 - \sum_{k=0}^{n-1} \langle f_k, g_{k+1} \rangle g_{k+1} \right\rangle \\ &= \left\langle f_n, r_n + \sum_{\lambda \in \Lambda} c_\lambda g_\lambda - \sum_{k=0}^{n-1} \langle f_k, g_{k+1} \rangle g_{k+1} \right\rangle \\ &\leq |\langle f_n, r_n \rangle| + \sum_{\lambda \in \Lambda} |c_\lambda| |\langle f_n, g_\lambda \rangle| + \sum_{k=0}^{n-1} |\langle f_k, g_{k+1} \rangle| |\langle f_n, g_{k+1} \rangle| \\ &\leq \|f_n\| \|r_n\| + \sup_{g \in \mathcal{D}} |\langle f_n, g \rangle| \left(\sum_{\lambda \in \Lambda} |c_\lambda| + \sum_{k=0}^{n-1} |\langle f_k, g_{k+1} \rangle| \right) \\ &\leq a_n^{1/2} (2n^{-1/(2(1+\gamma))}) + d_n (n^{\gamma/(2(1+\gamma))}) + b_n. \end{aligned}$$

Отсюда следует справедливость неравенства

$$d_n \geq \frac{a_n - 2n^{-1/(2(1+\gamma))} a_n^{1/2}}{b_n + n^{\gamma/(2(1+\gamma))}}. \quad (36)$$

§ 4. Множества $\Delta(m_0)$

Рассмотрим множество

$$\mathcal{M}_0 = \{m_0 \in \mathbb{N} : a_{m_0} \geq m_0^{-1/(1+\gamma)+\varepsilon}, a_{m_0} < a_{m_0-1} < (m_0 - 1)^{-1/(1+\gamma)+\varepsilon}\}. \quad (37)$$

Для $m_0 \in \mathcal{M}_0$ рассмотрим промежуток натурального ряда $\Delta(m_0)$, равный

$$\Delta(m_0) = \{m_0, m_0 + 1, \dots, s - 1\},$$

где s – минимальное число, большее m_0 , для которого $a_s < s^{-1/(1+\gamma)+\varepsilon}$, или, если такого s не существует, равный

$$\Delta(m_0) = \{m_0, m_0 + 1, \dots\}.$$

Тогда для всех $k \in \Delta(m_0)$ выполняется неравенство

$$a_k \geq k^{-1/(1+\gamma)+\varepsilon}. \quad (38)$$

ЛЕММА 3. *Существует такое $N_0 \in \mathbb{N}$, что для любого $m_0 \in \mathcal{M}_0$, $m_0 \geq N_0$, неравенство*

$$b_{m_0} \leq m_0^{\gamma/(2(1+\gamma))+\varepsilon/3} \quad (39)$$

влечет справедливость соотношения

$$\Delta(m_0) = \{m_0\}. \quad (40)$$

ДОКАЗАТЕЛЬСТВО. Равенство (40) равносильно $m_0 + 1 \notin \Delta(m_0)$, что, в свою очередь, равносильно неравенству

$$a_{m_0+1} < (m_0 + 1)^{-1/(1+\gamma)+\varepsilon}.$$

С учетом равенства (34) и определения (37) последнее неравенство будет вытекать из оценки

$$\begin{aligned} d_{m_0} &= (a_{m_0} - a_{m_0+1})^{1/2} > ((m_0 - 1)^{-1/(1+\gamma)+\varepsilon} - (m_0 + 1)^{-1/(1+\gamma)+\varepsilon})^{1/2} \\ &\geq C_0 m_0^{-1/(2(1+\gamma))+\varepsilon/2-1/2} = C_0 m_0^{-(2+\gamma)/(2(1+\gamma))+\varepsilon/2} \end{aligned} \quad (41)$$

с некоторым положительным C_0 . Если выполнено неравенство (39), то согласно неравенству (36) и определению (37) имеем

$$\begin{aligned} d_{m_0} &\geq \frac{a_{m_0}(1 - 2m_0^{-1/(2(1+\gamma))}a_{m_0}^{-1/2})}{b_{m_0} + m_0^{\gamma/(2(1+\gamma))}} \geq \frac{m_0^{-1/(1+\gamma)+\varepsilon}(1 - 2m_0^{-\varepsilon/2})}{2m_0^{\gamma/(2(1+\gamma))+\varepsilon/3}} \\ &= m_0^{-(2+\gamma)/(2(1+\gamma))+2\varepsilon/3} \left(\frac{1 - 2m_0^{-\varepsilon/2}}{2} \right). \end{aligned}$$

Ясно, что существует такое $N_0 \in \mathbb{N}$, что для всех $m_0 \in \mathcal{M}_0$, $m_0 \geq N_0$, имеет место неравенство

$$d_{m_0} \geq m_0^{-(2+\gamma)/(2(1+\gamma))+2\varepsilon/3} \left(\frac{1 - 2m_0^{-\varepsilon/2}}{2} \right) \geq C_0 m_0^{-(2+\gamma)/(2(1+\gamma))+\varepsilon/2},$$

что доказывает неравенство (41) и вместе с ним лемму.

ЛЕММА 4. *Существует такое $N_1 \in \mathbb{N}$, что для любого $m_0 \in \mathcal{M}_0$, $m_0 \geq N_1$, неравенство*

$$b_{m_0} \geq m_0^{\gamma/(2(1+\gamma))+\varepsilon/3} \quad (42)$$

влечет справедливость неравенств

$$b_k \geq k^{\gamma/(2(1+\gamma))+\varepsilon/4} \quad \forall k \in \Delta(m_0). \quad (43)$$

ДОКАЗАТЕЛЬСТВО. Докажем неравенство (43) по индукции. Фиксируем m , $m_0 < m \in \Delta(m_0)$. Предположим, что неравенство (43) справедливо для всех k , $m_0 \leq k < m$. Тогда, используя (36) и (38), получаем, что найдется такое число $N_{1,1} \in \mathbb{N}$, что если $m_0 \geq N_{1,1}$, то для $m_0 \leq k < m$ выполняется

$$\begin{aligned} d_k &\geq \frac{a_k(1 - 2k^{-1/(2(1+\gamma))})a_k^{-1/2}}{b_k(1 + k^{\gamma/(2(1+\gamma))})b_k^{-1}} \geq \frac{k^{-1/(1+\gamma)+\varepsilon}(1 - 2k^{-\varepsilon/2})}{k^{\gamma/(2(1+\gamma))+\varepsilon/4}(1 + k^{-\varepsilon/4})} \\ &\geq \frac{k^{-1/(1+\gamma)+\varepsilon}}{k^{\gamma/(2(1+\gamma))+\varepsilon/3}} \geq \frac{m^{-1/(1+\gamma)+\varepsilon}}{m^{\gamma/(2(1+\gamma))+\varepsilon/3}}. \end{aligned}$$

Из определения (33) следует, что

$$b_m = b_{m_0} + \sum_{k=m_0}^{m-1} d_k \geq m_0^{\gamma/(2(1+\gamma))+\varepsilon/3} + (m - m_0) \frac{m^{-1/(1+\gamma)+\varepsilon}}{m^{\gamma/(2(1+\gamma))+\varepsilon/3}}.$$

Тогда утверждение леммы будет следовать из неравенства $\Phi(m) \geq 0$, где

$$\Phi(t) := m_0^{\gamma/(2(1+\gamma))+\varepsilon/3} + (t - m_0) \frac{t^{-1/(1+\gamma)+\varepsilon}}{t^{\gamma/(2(1+\gamma))+\varepsilon/3}} - t^{\gamma/(2(1+\gamma))+\varepsilon/4}, \quad t \in [m_0, \infty).$$

Для проверки последнего неравенства достаточно заметить, что $\Phi(m_0) \geq 0$, и проверить, что $\Phi'(t) \geq 0$, $t \geq m_0$. Действительно, найдется такое $N_{1,2} \in \mathbb{N}$, что для всех $N_{1,2} \leq m_0 \leq t$ выполняется

$$\begin{aligned} \Phi'(t) &:= \left(\left(1 - \frac{1}{1+\gamma} + \varepsilon - \frac{\gamma}{2(1+\gamma)} - \frac{\varepsilon}{3} \right) t - \left(-\frac{1}{1+\gamma} + \varepsilon - \frac{\gamma}{2(1+\gamma)} - \frac{\varepsilon}{3} \right) m_0 \right) \\ &\quad \times t^{-1/(1+\gamma)+\varepsilon-\gamma/(2(1+\gamma))-\varepsilon/3-1} - \left(\frac{\gamma}{2(1+\gamma)} + \frac{\varepsilon}{4} \right) t^{\gamma/(2(1+\gamma))+\varepsilon/4-1} \\ &\geq \left(1 - \frac{1}{1+\gamma} + \varepsilon - \frac{\gamma}{2(1+\gamma)} - \frac{\varepsilon}{3} \right) t^{1-1/(1+\gamma)+\varepsilon-\gamma/(2(1+\gamma))-\varepsilon/3} \\ &\quad - \left(\frac{\gamma}{2(1+\gamma)} + \frac{\varepsilon}{4} \right) t^{\gamma/(2(1+\gamma))+\varepsilon/4-1} \\ &= \left(\frac{\gamma}{2(1+\gamma)} + \frac{2\varepsilon}{3} \right) t^{(-2-\gamma)/(2(1+\gamma))+2\varepsilon/3} - \left(\frac{\gamma}{2(1+\gamma)} + \frac{\varepsilon}{4} \right) t^{(-2-\gamma)/(2(1+\gamma))+\varepsilon/4} \\ &\geq 3t^{(-2-\gamma)/(2(1+\gamma))+\varepsilon/4} - \left(\frac{\gamma}{2(1+\gamma)} + \frac{\varepsilon}{4} \right) t^{(-2-\gamma)/(2(1+\gamma))+\varepsilon/4} \geq 0, \end{aligned}$$

что доказывает лемму с $N_1 := \max\{N_{1,1}, N_{1,2}\}$.

ЛЕММА 5. Существует такое $N_2 \geq 4$, что для любого $m_0 \in \mathcal{M}_0$, $m_0 \geq N_2$, неравенство (42) влечет справедливость неравенств

$$\frac{a_m}{b_m^2} \leq \frac{3}{m} \quad \text{для } m \in \Delta(m_0), \quad m \geq 2m_0. \quad (44)$$

ДОКАЗАТЕЛЬСТВО. Из неравенств (36), (38) и леммы 4 следует, что найдется такое число $N_2 \geq \max\{4, N_1\}$, что если $m_0 \geq N_2$, то для $m_0 \leq k < m$

выполняется

$$d_k \geq \frac{a_k(1 - 2k^{-1/(2(1+\gamma))})a_k^{-1/2}}{b_k(1 + k^{\gamma/(2(1+\gamma))}b_k^{-1})} \geq \frac{a_k(1 - 2k^{-\varepsilon/2})}{b_k(1 + k^{-\varepsilon/4})} \geq \frac{a_k}{b_k(1 + k^{-\varepsilon/5})} \geq \frac{a_k}{1.5b_k}. \quad (45)$$

Тогда для $m \geq 2m_0$ согласно (33) и монотонности последовательностей $\{a_k\}$ и $\{b_k\}$ имеем

$$b_m \geq \sum_{k=m_0}^{m-1} d_k \geq \sum_{k=m_0}^{m-1} \frac{a_k}{1.5b_k} \geq (m - m_0) \frac{a_m}{1.5b_m} \geq \frac{m}{2} \frac{a_m}{1.5b_m},$$

откуда вытекает неравенство (44).

§ 5. Оценка произведений $a_m b_m$

ЛЕММА 6. Для любых $m_0 \in \mathcal{M}_0$, $m_0 \geq N_3 := \max\{N_0, N_2\} \geq 4$, и $m \in \Delta(m_0)$ имеет место неравенство

$$a_m b_m \leq a_{m_0} b_{m_0} \prod_{k=m_0}^{m-1} \left(1 + \frac{d_k^2}{a_k} k^{-\varepsilon/5}\right).$$

ДОКАЗАТЕЛЬСТВО. Если справедливо неравенство (39), то в силу леммы 3 $\Delta(m_0) = \{m_0\}$ и утверждение леммы очевидно. Если выполняется (42), то согласно (45) имеем

$$d_k \geq \frac{a_k}{b_k(1 + k^{-\varepsilon/5})}.$$

Тогда, используя (33), получаем

$$b_{m+1} = b_m + d_m = b_m \left(1 + \frac{d_m^2}{d_m b_m}\right) \leq b_m \left(1 + \frac{d_m^2}{a_m} (1 + m^{-\varepsilon/5})\right).$$

С другой стороны, из неравенства (34) вытекает, что

$$a_{m+1} = a_m - d_m^2 = a_m \left(1 - \frac{d_m^2}{a_m}\right). \quad (46)$$

Почленно перемножая последние соотношения, имеем

$$a_{m+1} b_{m+1} \leq a_m b_m \left(1 + \frac{d_m^2}{a_m} m^{-\varepsilon/5}\right).$$

Следовательно, для всех $m \in \Delta(m_0)$

$$a_m b_m \leq a_{m_0} b_{m_0} \prod_{k=m_0}^{m-1} \left(1 + \frac{d_k^2}{a_k} k^{-\varepsilon/5}\right).$$

Пусть для $m \geq 3$

$$\Psi(m) = \exp\left(2\left(\frac{1}{1+\gamma} - \varepsilon\right) \sum_{n=2}^{m-1} n^{-\varepsilon/5} (\ln n - \ln(n-1))\right) \geq 1. \quad (47)$$

ТЕОРЕМА 5. Для любых $m_0 \in \mathcal{M}_0$, $m_0 \geq N_3 \geq 4$, и $m \in \Delta(m_0)$ имеет место неравенство

$$a_m b_m \leq \frac{\Psi(m)}{\Psi(m_0)} a_{m_0} b_{m_0}. \quad (48)$$

При доказательстве теоремы мы будем использовать преобразование Абеля:

$$\sum_{n=m_0}^{m-1} x_n y_n = \sum_{n=m_0}^{m-1} (x_n - x_{n+1}) \sum_{k=m_0}^n y_k + x_m \sum_{k=m_0}^{m-1} y_k.$$

ДОКАЗАТЕЛЬСТВО. Из равенства (46) вытекает

$$a_m = a_{m_0} \prod_{k=m_0}^{m-1} \left(1 - \frac{d_k^2}{a_k}\right).$$

Используя неравенства (37) и (38), получаем

$$\begin{aligned} \frac{m^{-1/(1+\gamma)+\varepsilon}}{(m_0-1)^{-1/(1+\gamma)+\varepsilon}} &\leq \frac{a_m}{a_{m_0}} = \prod_{k=m_0}^{m-1} \left(1 - \frac{d_k^2}{a_k}\right), \\ \left(-\frac{1}{1+\gamma} + \varepsilon\right) (\ln m - \ln(m_0-1)) &\leq \sum_{k=m_0}^{m-1} \ln \left(1 - \frac{d_k^2}{a_k}\right), \\ \sum_{k=m_0}^{m-1} \frac{d_k^2}{a_k} &\leq \left(\frac{1}{1+\gamma} - \varepsilon\right) (\ln m - \ln(m_0-1)) \\ &\leq 2 \left(\frac{1}{1+\gamma} - \varepsilon\right) (\ln(m-1) - \ln(m_0-1)). \end{aligned} \quad (49)$$

Докажем, что

$$\sum_{n=m_0}^{m-1} \frac{d_n^2}{a_n} n^{-\varepsilon/5} \leq 2 \left(\frac{1}{1+\gamma} - \varepsilon\right) \sum_{n=m_0}^{m-1} n^{-\varepsilon/5} (\ln n - \ln(n-1)). \quad (50)$$

Применяя преобразование Абеля с $x_n = n^{-\varepsilon/5}$ и $y_n = d_n^2/a_n$ с учетом неравенства (49), имеем

$$\begin{aligned} \sum_{n=m_0}^{m-1} \frac{d_n^2}{a_n} n^{-\varepsilon/5} &\leq \sum_{n=m_0}^{m-1} (n^{-\varepsilon/5} - (n+1)^{-\varepsilon/5}) \sum_{k=m_0}^n \frac{d_k^2}{a_k} + m^{-\varepsilon/5} \sum_{k=m_0}^{m-1} \frac{d_k^2}{a_k} \\ &\leq 2 \left(\frac{1}{1+\gamma} - \varepsilon\right) \left(\sum_{n=m_0}^{m-1} (n^{-\varepsilon/5} - (n+1)^{-\varepsilon/5}) (\ln n - \ln(m_0-1)) \right. \\ &\quad \left. + m^{-\varepsilon/5} (\ln(m-1) - \ln(m_0-1)) \right). \end{aligned}$$

Повторно применяя преобразование Абеля для $x_n = \ln n - \ln(m_0 - 1)$ и $y_n = n^{-\varepsilon/5} - (n+1)^{-\varepsilon/5}$, получаем

$$\begin{aligned}
& \frac{1}{2} \left(\frac{1}{1+\gamma} - \varepsilon \right)^{-1} \sum_{n=m_0}^{m-1} \frac{d_n^2}{a_n} n^{-\varepsilon/5} \\
& \leq \sum_{n=m_0}^{m-1} (n^{-\varepsilon/5} - (n+1)^{-\varepsilon/5}) (\ln n - \ln(m_0 - 1)) \\
& \quad + m^{-\varepsilon/5} (\ln(m-1) - \ln(m_0 - 1)) \\
& \leq \sum_{n=m_0}^{m-1} (\ln n - \ln(n+1)) (m_0^{-\varepsilon/5} - (n+1)^{-\varepsilon/5}) \\
& \quad + (\ln m - \ln(m_0 - 1)) (m_0^{-\varepsilon/5} - m^{-\varepsilon/5}) \\
& \quad + m^{-\varepsilon/5} (\ln(m-1) - \ln(m_0 - 1)) \\
& \leq \sum_{n=m_0}^{m-1} (\ln(n+1) - \ln n) (n+1)^{-\varepsilon/5} \\
& \quad + m_0^{-\varepsilon/5} \left(\ln m - \ln(m_0 - 1) + \sum_{n=m_0}^{m-1} (\ln n - \ln(n+1)) \right) \\
& \quad + m^{-\varepsilon/5} (-\ln m + \ln(m_0 - 1) + \ln(m-1) - \ln(m_0 - 1)) \\
& \leq \sum_{n=m_0}^{m-1} (\ln(n+1) - \ln n) (n+1)^{-\varepsilon/5} + m_0^{-\varepsilon/5} (\ln m_0 - \ln(m_0 - 1)) \\
& \quad + m^{-\varepsilon/5} (-\ln m + \ln(m-1)) \\
& = \sum_{n=m_0}^{m-1} n^{-\varepsilon/5} (\ln n - \ln(n-1)),
\end{aligned}$$

что доказывает (50). Согласно лемме 6, неравенству (50) и определению (47) выполняются оценки

$$\begin{aligned}
\frac{a_m b_m}{a_{m_0} b_{m_0}} & \leq \prod_{k=m_0}^{m-1} \left(1 + \frac{d_k^2}{a_k} k^{-\varepsilon/5} \right) = \exp \left(\sum_{k=m_0}^{m-1} \ln \left(1 + \frac{d_k^2}{a_k} k^{-\varepsilon/5} \right) \right) \\
& \leq \exp \left(\sum_{k=m_0}^{m-1} \frac{d_k^2}{a_k} k^{-\varepsilon/5} \right) \\
& \leq \exp \left(2 \left(\frac{1}{1+\gamma} - \varepsilon \right) \sum_{n=m_0}^{m-1} n^{-\varepsilon/5} (\ln n - \ln(n-1)) \right) = \frac{\Psi(m)}{\Psi(m_0)}.
\end{aligned}$$

Отсюда непосредственно следует утверждение теоремы.

ТЕОРЕМА 6. *Существует такое число $\hat{C} > 0$, что если*

$$a_0 < N_3^{-1/(1+\gamma)+\varepsilon}, \quad (51)$$

то для любых $m_0 \in \mathcal{M}_0$ и $m \in \Delta(m_0)$ имеет место неравенство

$$a_m b_m \leq \widehat{C} m^{\max\{0, 1/2-3(1/(1+\gamma)-\varepsilon)/2\}}. \quad (52)$$

ДОКАЗАТЕЛЬСТВО. Пусть $\alpha = 1/(1+\gamma) - \varepsilon$. Рассмотрим числа

$$\Theta(m) = \max\{\overline{C}(\alpha), C(\alpha), 1\} \Psi(m) m^{\max\{0, 1/2-3\alpha/2\}}, \quad m \geq 2. \quad (53)$$

Пусть

$$\mathcal{M}_0 = \{m_0^1 < m_0^2 < \dots < m_0^n < \dots\}. \quad (54)$$

В силу того, что последовательность $\{a_n\}_{n=0}^\infty$ убывает, и из соотношений (37) и (51) следует, что $m_0^1 \geq N_3 \geq 4$. Тогда докажем по индукции, что для любого n для всех $m \in \Delta(m_0^n)$ выполняется неравенство

$$a_m b_m \leq \Theta(m). \quad (55)$$

Фиксируем $n \in \mathbb{N}$. Тогда

$$a_{m_0^n} b_{m_0^n} \leq \Theta(m_0^n).$$

Действительно, если $n = 1$, то по теореме 4, примененной к $l = m_0^1$, имеем $a_{m_0^1} b_{m_0^1} \leq \overline{C}(\alpha) (m_0^1)^{\max\{0, 1/2-3\alpha/2\}} \leq \Theta(m_0^1)$. Если $n \geq 2$, то рассмотрим максимальное k из $\Delta(m_0^{n-1})$. Тогда $k+1 < m_0^n$, по индукционному предположению

$$a_k b_k \leq \max\{\overline{C}(\alpha), C(\alpha), 1\} \Psi(k) k^{\max\{0, 1/2-3\alpha/2\}}$$

и по теореме 3 для $l = m_0^n$

$$a_{m_0^n} b_{m_0^n} \leq \max\{\overline{C}(\alpha), C(\alpha), 1\} \Psi(m_0^n) (m_0^n)^{\max\{0, 1/2-3\alpha/2\}} \leq \Theta(m_0^n).$$

Если $m \in \Delta(m_0^n)$, то по теореме 5 имеем

$$a_m b_m \leq \frac{\Psi(m)}{\Psi(m_0^n)} a_{m_0^n} b_{m_0^n} \leq \max\{\overline{C}(\alpha), C(\alpha), 1\} \Psi(m) (m_0^n)^{\max\{0, 1/2-3\alpha/2\}} \leq \Theta(m),$$

что доказывает неравенство (55). Используя определения (47) и (53), а также монотонность чисел $\Psi(m)$, получаем справедливость неравенства (52) для

$$\begin{aligned} \widehat{C} &= \max\{\overline{C}(\alpha), C(\alpha), 1\} \lim_{m \rightarrow \infty} \Psi(m) \\ &\leq \max\{\overline{C}(\alpha), C(\alpha), 1\} \exp\left(2\left(\frac{1}{1+\gamma} - \varepsilon\right) \sum_{n=1}^{\infty} n^{-\varepsilon/5} (\ln n - \ln(n-1))\right) < \infty. \end{aligned}$$

§ 6. Доказательство теоремы 1

Будем использовать технику из работы [5], а именно, оценивать a_m через оценки на $a_m b_m$ и a_m/b_m^2 , а также использовать порядковое неравенство $a_m/b_m^2 \lesssim m^{-1}$.

Так как для всех словарей $\mathcal{D}$, $f \in H$, $m \in \mathbb{N}$ и $\beta > 0$ выполняется равенство $G_m(\beta f, \mathcal{D}) = \beta G_m(f, \mathcal{D})$, то для доказательства теоремы достаточно найти такое η , $0 < \eta < 1$, что для всех $f \in \widehat{\mathcal{A}}_\gamma(\mathcal{D}, \eta)$ и $m \geq 0$ выполняется

$$a_{m+1}^{1/2} = \|f - G_m(f, D)\| \leq \tilde{C} m^{-\min\{1/6, 1/(2(1+\gamma))-\varepsilon\}} \quad (56)$$

с некоторой константой $\tilde{C} > 0$.

Из определения класса $\widehat{\mathcal{A}}_\gamma(\mathcal{D}, M)$ легко видеть, что для $f \in \widehat{\mathcal{A}}_\gamma(\mathcal{D}, M)$ выполняется неравенство $\|f\| \leq 2M$. Выберем произвольное η , удовлетворяющее неравенству

$$0 < \eta < \frac{1}{2} (N_3^{-1/(1+\gamma)+\varepsilon})^{1/2},$$

произвольную функцию $f \in \widehat{\mathcal{A}}_\gamma(\mathcal{D}, \eta)$, построим по ней и формулам (31) и (33) последовательности $\{a_n\}_{n=0}^\infty$ и $\{b_n\}_{n=0}^\infty$ и определим согласно (37) множество $\mathcal{M}_0 = \{m_0^n\}_{n=1}^\infty$. Ясно, что

$$a_0 = \|f_0\|^2 = \|f\|^2 < N_3^{-1/(1+\gamma)+\varepsilon}.$$

Пусть $m \in \mathbb{N}$. Если $m \notin \bigcup_{m_0 \in \mathcal{M}_0} \Delta(m_0)$, то согласно определению промежутков $\Delta(m_0)$

$$a_m \leq m^{-1/(1+\gamma)+\varepsilon}. \quad (57)$$

Если $m \in \bigcup_{m_0 \in \mathcal{M}_0} \Delta(m_0)$, то согласно представлению (54) найдется $n \in \mathbb{N}$, для которого $m \in \Delta(m_0^n)$. В случае, когда $m < 2m_0^n$, согласно определению (37) имеем

$$a_m \leq a_{m_0^n} \leq (m_0^n - 1)^{-1/(1+\gamma)+\varepsilon} \leq \left(\frac{m}{3}\right)^{-1/(1+\gamma)+\varepsilon}. \quad (58)$$

И, наконец, в случае, когда $m \geq 2m_0^n$, согласно теореме 6 и лемме 5 получаем

$$\begin{aligned} a_m &= \left((a_m b_m)^2 \frac{a_m}{b_m^2} \right)^{1/3} = \left(\widehat{C}^2 m^{\max\{0, 1-3/(1+\gamma)-\varepsilon\}} \frac{3}{m} \right)^{1/3} \\ &= \widehat{C}^{2/3} 3^{1/3} m^{\max\{-1/3, -1/(1+\gamma)+\varepsilon\}}. \end{aligned} \quad (59)$$

Комбинация оценок (57)–(59) доказывает неравенство (56) и вместе с ним теорему 1.

§ 7. Доказательство теоремы 2

Пусть $\mathcal{D} = \{e_n\}_{n=0}^\infty$ – ортонормированный базис в гильбертовом пространстве H . Для $\gamma > 0$ положим

$$\frac{1}{2} < \delta = \frac{2+\gamma}{2+2\gamma} < 1. \quad (60)$$

Пусть

$$f = \sum_{k=1}^{\infty} k^{-\delta} e_k.$$

Из неравенства $\delta > 1/2$ следует, что элемент $f \in H$ определен корректно. Докажем, что

$$f \in \widehat{\mathcal{A}}_{\gamma}(\mathcal{D}). \quad (61)$$

Пусть $m \geq 1$. Положим

$$n = n(m) = \lceil m^{2+2\gamma} \rceil. \quad (62)$$

Рассмотрим множество $\Lambda = \{1, \dots, n\}$ и числа $c_{\lambda} = \lambda^{-\delta}$, $\lambda = 1, \dots, n$. Тогда с учетом (60) и (62) для некоторых положительных постоянных $C_1, C_2, C_3, C_4 > 0$, зависящих только от γ (и не зависящих от m), справедливы неравенства

$$\begin{aligned} \left\| f - \sum_{\lambda \in \Lambda} c_{\lambda} g_{\lambda} \right\| &= \left\| f - \sum_{k=1}^n k^{-\delta} g_k \right\| \leq \left(\sum_{k=n+1}^{\infty} k^{-2\delta} \right)^{1/2} \\ &\leq (C_1 n^{1-2\delta})^{1/2} \leq C_2 m^{(2+2\gamma)(1/2-\delta)} = \frac{C_2}{m}, \\ \sum_{\lambda \in \Lambda} |c_{\lambda}| &= \sum_{k=1}^n k^{-\delta} \leq C_3 n^{1-\delta} \leq C_4 m^{(2+2\gamma)(1-\delta)} = C_4 m^{\gamma}. \end{aligned}$$

Последние два неравенства доказывают (61). Оценим $\sigma_m(f, \mathcal{D})$ снизу. В силу ортонормированности $D = \{e_n\}_{n=0}^{\infty}$ и монотонности $k^{-\delta}$ имеем

$$\sigma_m(f, \mathcal{D}) = \left\| f - \sum_{k=1}^m k^{-\delta} e_k \right\| = \left(\sum_{k=m+1}^{\infty} k^{-2\delta} \right)^{1/2} \geq C_5 m^{(1-2\delta)/2} = C_5 m^{-1/(2(1+\gamma))}, \quad (63)$$

где $C_5 > 0$ – некоторая постоянная, зависящая от γ и не зависящая от m . Объединение соотношений (61) и (63) доказывает теорему 2.

Список литературы

- [1] E. Schmidt, “Zur Theorie der linearen und nichtlinearen Integralgleichungen I. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener”, *Math. Ann.*, **63**:4 (1907), 433–476.
- [2] S. G. Mallat, Zhifeng Zhang, “Matching pursuits with time-frequency dictionaries”, *IEEE Trans. Signal Process.*, **41**:12 (1993), 3397–3415.
- [3] J. H. Friedman, W. Stuetzle, “Projection pursuit regression”, *J. Amer. Statist. Assoc.*, **76**:376 (1981), 817–823.
- [4] L. K. Jones, “A simple lemma on greedy approximation in Hilbert space and convergence rates for projection pursuit regression and neural network training”, *Ann. Statist.*, **20**:1 (1992), 608–613.
- [5] R. A. DeVore, V. N. Temlyakov, “Some remarks on greedy algorithms”, *Adv. Comput. Math.*, **5**:1 (1996), 173–187.
- [6] G. Davis, S. Mallat, M. Avellaneda, “Adaptive greedy approximations”, *Constr. Approx.*, **13**:1 (1997), 57–98.
- [7] V. N. Temlyakov, “Weak greedy algorithms”, *Adv. Comput. Math.*, **12**:2–3 (2000), 213–227.

- [8] В. В. Галатенко, “Об одной модификации жадных разложений”, *Современные методы теории функций и смежные проблемы. Материалы Воронежской зимней математической школы* (ВГУ, Воронеж, 2005), Изд-во ВГУ, Воронеж, 2005, 65–66.
- [9] Е. Д. Лившиц, “О скорости сходимости чисто жадного алгоритма”, *Матем. заметки*, **76**:4 (2004), 539–552; англ. пер.: E. D. Livshits, “Rate of convergence of pure greedy algorithms”, *Math. Notes*, **76**:3–4 (2004), 497–510.
- [10] С. Б. Стечкин, “Об абсолютной сходимости ортогональных рядов”, *Докл. АН СССР*, **102**:1 (1955), 37–40.
- [11] A. R. Barron, “Universal approximation bounds for superpositions of a sigmoidal function”, *IEEE Trans. Inform. Theory*, **39**:3 (1993), 930–945.
- [12] S. V. Konyagin, V. N. Temlyakov, “Rate of convergence of pure greedy algorithm”, *East J. Approx.*, **5**:4 (1999), 493–499.
- [13] А. В. Сильниченко, “О скорости сходимости жадных алгоритмов”, *Матем. заметки*, **76**:4 (2004), 628–632; англ. пер.: A. V. Sil’nichenko, “Rate of convergence of greedy algorithms”, *Math. Notes*, **76**:3–4 (2004), 582–586.
- [14] E. D. Livshitz, V. N. Temlyakov, “Two lower estimates in greedy approximation”, *Constr. Approx.*, **19**:4 (2003), 509–523.

Е. Д. Лившиц (E. D. Livshits)
Московский государственный
университет им. М. В. Ломоносова
E-mail: livshitz@mccme.ru

Поступила в редакцию
16.05.2006 и 09.03.2007