Math-Net.Ru

Общероссийский математический портал

Е. Д. Лившиц, О возвратном жадном алгоритме, *Изв. РАН. Сер. матем.*, 2006, том 70, выпуск 1, 95–116

DOI: <https://doi.org/10.4213/im559>

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: 18.234.197.8

26 апреля 2023 г., 11:45:46

УДК 517.518.8

Е. Д. Лившиц

О возвратном жадном алгоритме

Исследуется возвратный жадный алгоритм (ВЖА). Доказывается его сходимость для произвольных начальной функции и словаря разложения. Получены точные в степенной шкале оценки скорости сходимости ВЖА в случае, если начальная функция принадлежит классу $A_1(\mathcal{D})$. Эти оценки распространяются на более широкие классы начальных функций и используются для сравнения некоторых классов функций, определяемых по заданному словарю.

Библиография: 21 наименование.

§ 1. Введение

В последние годы в теории приближения возрос интерес к методам аппроксимации, которые не только дают хорошую скорость приближения, но и указывают способ (алгоритм) нахождения аппроксимантов, на которых указанная скорость приближения достигается. Одним из подходов к решению данной задачи является использование *жадных алгоритмов*, исследованию которых 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 с единичной нормой, замыкание линейной оболочки которого совпадает со всем H :

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

Такие множества будем называть *словарями*. Никаких условий на минимальность $\mathcal{D}$ не налагается. Словарями могут быть как базисы, так и переполненные системы, например множество билинейных функций

$$H = L_2(\mathbb{R}^2), \quad \mathcal{D} = \{f(x, y) = f_1(x)f_2(y), \ f_1, f_2 \in L_2(\mathbb{R}), \ \|f_1\| = \|f_2\| = 1\}$$

или функций плоских волн

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

Для произвольной функции $f \in H$ будем искать разложение по элементам словаря:

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

Работа выполнена при финансовой поддержке РФФИ (грант № 05-01-00062).

частичные суммы которого $\sum_{k=1}^m c_k g_k$ достаточно хорошо ее приближают (предполагается, что векторы g_k с различными индексами могут совпадать). Для оценки приближения заданной функции линейными комбинациями m элементов словаря будем использовать величину наилучшего m -членного приближения

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

Однако знание величин $\sigma_m(f, \mathcal{D})$ не дает какого-либо способа нахождения аппроксиманта. Жадные алгоритмы предлагают некий путь нахождения аппроксиманта, но ошибка приближения при этом может отличаться от величины $\sigma_m(f, \mathcal{D})$. В настоящее время существует довольно много жадных алгоритмов в гильбертовом пространстве (см. [1], [9], [11], [12], [19], [21]), мы приведем здесь описание лишь двух из них.

СЛАБЫЙ ЖАДНЫЙ АЛГОРИТМ (СЖА) С ПАРАМЕТРОМ t . Пусть фиксировано число t , $0 < t \leq 1$. Положим $R_0^{\text{WGA}}(f_0, \mathcal{D}) := f_0$ и $G_0^{\text{WGA}}(f_0, \mathcal{D}) := 0$. Тогда для каждого $m \geq 0$ последовательно выбираем вектор $g_{m+1} \in \mathcal{D}$ такой, что

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

и определяем

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

В. Н. Темляков [19] определил СЖА в более общем виде: в его работе “коэффициенты слабости” t в формуле (1) могут отличаться на различных шагах алгоритма (t_m вместо t).

При $t = 1$ слабый жадный алгоритм становится *чисто жадным алгоритмом* (ЧЖА). В этом случае аппроксимант $G_m^{\text{WGA}}(f_0, \mathcal{D})$ и остаток $R_m^{\text{WGA}}(f_0, \mathcal{D})$ будут обозначаться соответственно $G_m^{\text{PGA}}(f_0, \mathcal{D})$ и $R_m^{\text{PGA}}(f_0, \mathcal{D})$. Отметим, что при определении ЧЖА необходимо предполагать, что супремум в формуле (1) достигается на какой-то функции из словаря (что и делается в большинстве работ), в противном случае алгоритм реализован быть не может. Эта некорректность послужила одной из причин для создания слабого жадного алгоритма.

Несмотря на то, что термин “жадный (гриди) алгоритм” стали использовать в теории приближения достаточно недавно [8], [9], в неявном виде жадный алгоритм в применении к конкретным пространствам и словарям начали использовать достаточно давно. Например, Е. Шмидт [18] при приближении функций из $L_2([0, 1]^2)$ билинейными функциями использовал процесс, совпадающий с ЧЖА. Известная в математической статистике процедура Projection Pursuit Regression (см. [10]) фактически совпадает с ЧЖА в применении к функциям плоских волн.

Отметим также, что приведенные выше СЖА и ЧЖА не являются алгоритмами в обычном смысле этого слова. Дело в том, что на шаге (1) супремум берется по бесконечному множеству. На практике это можно сделать приближенно, переходя к конечномерным пространствам большой размерности (см. [7]).

В настоящее время в математической литературе на русском языке нет устоявшегося перевода термина Pure (Week) Greedy Algorithm. Некоторые авторы переводят это как *чисто (слабый) гриди-алгоритм*, или PGA-разложение (WGA-разложение).

Возникают естественные вопросы о сходимости и скорости сходимости жадных алгоритмов. Л. Джонс [13], [14] доказал сходимость СЖА с параметром t для всех словарей $\mathcal{D}$, начальных функций f_0 и чисел $t > 0$:

$$\lim_{m \rightarrow \infty} \|f_0 - G_m^{\text{WGA}}(f_0, \mathcal{D})\| = 0.$$

(О сходимости жадных алгоритмов в гильбертовых пространствах см. также работы [1], [3], [9], [11], [12], [19], [20].)

Рассмотрим более подробно вопрос о скорости сходимости СЖА и ЧЖА. Ясно, что в случае произвольного словаря $\mathcal{D}$ нельзя априори гарантировать какую-либо скорость сходимости алгоритма для всех функций из пространства H . Для получения нетривиальных оценок скорости сходимости алгоритма необходимо потребовать, чтобы начальная функция принадлежала некоторому подмножеству H , вообще говоря, зависящему от словаря $\mathcal{D}$. Достаточно естественными являются классы начальных функций, состоящие из функций с фиксированной скоростью убывания m -членных приближений:

$$\mathcal{A}^\gamma(\mathcal{D}) = \{f \in H \mid \exists M: \forall m \ \sigma_m(f, \mathcal{D}) \leq Mm^{-\gamma}\}, \quad \gamma > 0.$$

Однако принадлежность начальной функции классу $\mathcal{A}^\gamma(\mathcal{D})$ является слабым предположением для получения полиномиальной скорости сходимости ЧЖА: как было показано в работе [2], для произвольных $\gamma, \delta > 0$ существуют словарь $\mathcal{D}$, функция $f_0 \in \mathcal{A}^\gamma(\mathcal{D})$ и $C > 0$ такие, что для всех $m > 1$ выполняется неравенство

$$\|f - G_m^{\text{PGA}}(f, \mathcal{D})\| \geq Cm^{-\delta}.$$

Часто сходимость СЖА и ЧЖА исследуется, когда начальная функция принадлежит классу $\mathcal{A}_1(\mathcal{D})$, определенному следующим образом:

$$\begin{aligned} \mathcal{A}_1^0(\mathcal{D}, M) &:= \left\{ f \in H: f = \sum_{k \in \Lambda} c_k w_k, \ w_k \in \mathcal{D}, \ \#\Lambda < \infty, \ \sum_{k \in \Lambda} |c_k| \leq M \right\}, \\ \mathcal{A}_1(\mathcal{D}, M) &= \overline{\mathcal{A}_1^0(\mathcal{D}, M)}, \\ \mathcal{A}_1(\mathcal{D}) &= \bigcup_{M > 0} \mathcal{A}_1(\mathcal{D}, M). \end{aligned}$$

Для $f \in \mathcal{A}_1(\mathcal{D})$ введем норму

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

С. Б. Стечкин [5] охарактеризовал функции, принадлежащие классу $\mathcal{A}_1(\mathcal{D})$, в терминах m -членных приближений в случае, когда $\mathcal{D}$ является ортонормированной системой. В случае произвольного словаря имеет место включение [6]

$$\mathcal{A}_1(\mathcal{D}) \subset A^{1/2}(\mathcal{D}),$$

причем известно, что показатель $1/2$ неулучшаем. Возник вопрос о скорости сходимости ЧЖА в классе $\mathcal{A}_1(\mathcal{D})$, в частности выдвигалась гипотеза, что скорость сходимости имеет порядок $m^{-1/2}$. Р. Девор и В. Н. Темляков [9] доказали оценку

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

Она была улучшена С. В. Конягиным и В. Н. Темляковым [15]:

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

а затем А. В. Сильниченко [4], который показал, что существует такая константа $C_1 > 0$, что для всех словарей $\mathcal{D}$ и начальных функций f имеет место оценка

$$\|f - G_m^{\text{PGA}}(f, \mathcal{D})\| \leq C_1 |f|_{\mathcal{A}_1(\mathcal{D})} m^{-0.182}.$$

Автор настоящей статьи и В. Н. Темляков [16] привели пример такого словаря $\mathcal{D}$ и такой функции f_0 , что для некоторого $C_2 > 0$ выполняется

$$\|f_0 - G_m^{\text{PGA}}(f_0, \mathcal{D})\| \geq C_2 m^{-0.27}.$$

Скорость сходимости СЖА была оценена В. Н. Темляковым [19] (мы приводим ее здесь лишь в случае, когда все “коэффициенты слабости” равны): для произвольного словаря $\mathcal{D}$, числа t , $0 < t \leq 1$, и начальной функции $f \in H$ справедливо неравенство

$$\|f - G_m^{\text{WGA}}(f, \mathcal{D})\| \leq |f|_{\mathcal{A}_1(\mathcal{D})} (mt^2 + 1)^{-\frac{t}{2(2+t)}} \quad (2)$$

(при $t = 1$ эта оценка совпадает с оценкой из работы [9]). Таким образом, в классе функций $\mathcal{A}_1(\mathcal{D})$ СЖА и даже ЧЖА не могут гарантировать построение

m -членных приближений, сходящихся с оптимальной скоростью (т. е. $m^{-1/2}$). В работе [9] приведено описание релаксационного жадного алгоритма (см. также [6]) и ортогонального жадного алгоритма и доказано, что m -членные приближения, получаемые с их помощью, сходятся к исходной функции с оптимальной (в полиномиальной шкале) скоростью (т. е. $m^{-1/2}$). Однако в отличие от СЖА и ЧЖА эти алгоритмы не позволяют получить разложение вектора из H по элементам словаря $\mathcal{D}$: $(m + 1)$ -й аппроксимант получается из m -го не только добавлением очередного элемента словаря с каким-то коэффициентом, но и изменением коэффициентов при всех, уже участвовавших в разложении, элементах словаря.

Изучающийся в настоящей работе *возвратный жадный алгоритм*, с одной стороны, дает разложение начальной функции в ряд по элементам словаря

(как ЧЖА и СЖА), с другой – дает правильную в полиномиальной шкале скорость сходимости в классе $\mathcal{A}_1(\mathcal{D})$ (как релаксационный и ортогональный жадные алгоритмы).

Пусть заданы вектор $f_0 \in H$, словарь $\mathcal{D}$ и число $0 < t \leq 1$. Положим $R_0(f_0, \mathcal{D}) := f_0$, $G_0(f_0, \mathcal{D}) := 0$. Предположим, что для $n \geq 0$ уже определены функции $f_1, \dots, f_n \in H$, элементы $g_1, \dots, g_n \in \mathcal{D}$ и числа $c_1, \dots, c_n \in \mathbb{R}$, причем

$$f_k = f_0 - \sum_{i=1}^k c_i g_i \quad \forall k, \quad 1 \leq k \leq n. \quad (3)$$

Введем обозначения:

$$\begin{aligned} D_n &:= D_n(f_0, \mathcal{D}) := \{g_k\}_{k=1}^n, \\ v_n(g) &:= v_n(g, f_0, \mathcal{D}) := \sum_{k \leq n, g_k = g} c_k, \\ b_n &:= b_n(f_0, \mathcal{D}) := \sum_{g \in D_n} |v_n(g)|, \\ a_n &= \langle f_n, f_n \rangle. \end{aligned} \quad (4)$$

Если $n \geq 1$ и существует хотя бы один вектор $g \in D_n$ такой, что выполняются следующие соотношения:

$$\begin{aligned} \min(|v_n(g)|, |\langle f_n, g \rangle|) &\geq \frac{a_n t^2}{9b_n}, \\ \text{sign}(v_n(g)) &= -\text{sign}(\langle f_n, g \rangle), \end{aligned} \quad (5)$$

ТО ПОЛОЖИМ

$$g_{n+1} = g, \quad c_{n+1} = \text{sign}(\langle f_n, g \rangle) \min(|v_n(g)|, |\langle f_n, g \rangle|). \quad (6)$$

Будем называть такие шаги *возвратными*. В противном случае выберем произвольный элемент $g_{n+1} \in \mathcal{D}$, удовлетворяющий неравенству

$$|\langle f_n, g_{n+1} \rangle| \geq t \sup_{g \in \mathcal{D}} |\langle f_n, g \rangle|, \quad (7)$$

и положим

$$c_{n+1} = \langle f_n, g_{n+1} \rangle. \quad (8)$$

Будем называть такие шаги *жадными*. Определим аппроксимант и остаток:

$$\begin{aligned} G_{n+1}(f_0, \mathcal{D}) &:= \sum_{k=1}^{n+1} c_k g_k = G_n(f_0, \mathcal{D}) + c_{n+1} g_{n+1}, \\ f_{n+1} &:= f_n - c_{n+1} g_{n+1} = f_0 - G_{n+1}(f_0, \mathcal{D}), \end{aligned}$$

что гарантирует выполнение (3) на следующем шаге алгоритма.

Отметим, что последовательность $\{a_n\}$ будет убывающей. Справедлива следующая теорема о сходимости возвратных разложений.

ТЕОРЕМА 1. Для произвольного словаря $\mathcal{D}$, начального вектора f_0 и числа t , $0 < t \leq 1$, возвратный жадный алгоритм сходится:

$$\lim_{n \rightarrow \infty} \|f_0 - G_n(f_0, \mathcal{D})\| = 0.$$

ТЕОРЕМА 2. Для произвольного t , $0 < t \leq 1$, существует такая константа $K = K(t) > 0$, что для произвольного словаря $\mathcal{D}$, начального вектора $f_0 \in \mathcal{A}_1(\mathcal{D})$ и всех $n \geq 1$ имеет место неравенство

$$\|f_0 - G_n(f_0, \mathcal{D})\| \leq K |f_0|_{\mathcal{A}_1(\mathcal{D})} n^{-1/2} \ln n. \quad (9)$$

Отметим, что в отличие от оценки (2) теорема 2 гарантирует скорость сходимости ВЖА, не зависящую от параметра t . Ниже неравенство (9) распространяется на более широкий класс функций $F_\beta(\mathcal{D})$, $0 \leq \beta < 1/2$:

$$F_\beta(\mathcal{D}) = \bigcup_{\substack{M_\alpha, M_\beta > 0 \\ \alpha > 1/2 - \beta}} F_{\alpha, \beta}^{M_\alpha, M_\beta}(\mathcal{D}),$$

где

$$F_{\alpha, \beta}^{M_\alpha, M_\beta}(\mathcal{D}) = \{f \in H \mid \forall n \exists f^n: \|f - f^n\| \leq M_\alpha n^{-\alpha}, |f^n|_{\mathcal{A}_1(\mathcal{D})} \leq M_\beta n^\beta\}.$$

Определим норму

$$|f|_{F_\beta(\mathcal{D})} = \max \left(\|f_0\|, \inf \left\{ M \mid \exists \alpha > \frac{1}{2} - \beta, M_\alpha, M_\beta \leq M: f \in F_{\alpha, \beta}^{M_\alpha, M_\beta}(\mathcal{D}) \right\} \right).$$

Легко видеть, что для любых α , $M_\alpha > 0$ выполняется

$$F_{\alpha, 0}^{M_\alpha, M}(\mathcal{D}) = \mathcal{A}_1(\mathcal{D}, M),$$

откуда имеем равенство

$$\mathcal{A}_1(\mathcal{D}) = F_0(\mathcal{D}), \quad (10)$$

и с учетом неравенства $\|f\| \leq |f|_{\mathcal{A}_1(\mathcal{D})}$ получаем соотношение

$$|f|_{\mathcal{A}_1(\mathcal{D})} = |f|_{F_0(\mathcal{D})}, \quad f \in \mathcal{A}_1(\mathcal{D}). \quad (11)$$

ТЕОРЕМА 3. Для произвольного числа t , $0 < t \leq 1$, существует такая константа $K = K(t)$, что для произвольного β , $0 \leq \beta < 1/2$, словаря $\mathcal{D}$, начального вектора $f_0 \in F_\beta(\mathcal{D})$ и всех $n \geq 1$ имеет место неравенство

$$\|f_0 - G_n(f_0, \mathcal{D})\| \leq K |f_0|_{F_\beta(\mathcal{D})} n^{-1/2 + \beta} \ln n.$$

В работе [2] изучалась скорость сходимости ЧЖА для классов $A_{\alpha, \beta}(\mathcal{D})$, определенных следующим образом. Пусть $\alpha, \beta > 0$ и

$$A_{\alpha, \beta}(\mathcal{D}) = \bigcup_{M_\alpha, M_\beta > 1} A_{\alpha, \beta}^{M_\alpha, M_\beta}(\mathcal{D}),$$

где

$$A_{\alpha,\beta}^{M_\alpha,M_\beta}(\mathcal{D}) = \left\{ f \in H \mid \forall m \exists g_0, \dots, g_{m-1} \in \mathcal{D}, c_0, \dots, c_{m-1} \in \mathbb{R} : \right. \\ \left. \left\| f - \sum_{i=0}^{m-1} c_i g_i \right\| \leq \frac{M_\alpha}{m^\alpha}; \sum_{i=0}^{m-1} |c_i| \leq M_\beta m^\beta \right\}.$$

Далее, в § 6, с помощью теоремы 3 будет установлена связь между классами $A_{\alpha,\beta}(\mathcal{D})$ и $A^\gamma(\mathcal{D})$ и доказана

ТЕОРЕМА 4. Для чисел $\alpha, \beta > 0$, $\alpha + \beta < 1/2$, выполняется равенство

$$\sup\{\gamma > 0 \mid A_{\alpha,\beta}(\mathcal{D}) \subset A^\gamma(\mathcal{D}) \ \forall \mathcal{D}\} = \frac{\alpha}{2(\alpha + \beta)}.$$

§ 2. Доказательство теоремы 1

Предположим, что нашлись f_0 , $\mathcal{D}$ и $0 < t < 1$, для которых ВЖА не сходится. Тогда найдется такое $T_0 > 0$, что

$$a_n = \langle f_n, f_n \rangle \geq T_0, \quad n \geq 0. \quad (12)$$

Пусть $\bar{b}_n = \sum_{k=1}^n |c_k|$ для $n \geq 1$. Ясно, что

$$b_n \leq \bar{b}_n, \quad n \geq 1. \quad (13)$$

Положим

$$s := \liminf_{n \rightarrow \infty} \frac{\sup_{g \in \mathcal{D}} |\langle f_n, g \rangle|}{a_n / \bar{b}_n}.$$

Рассмотрим два случая.

Случай 1. Пусть $s = 0$. Тогда найдется такая подпоследовательность $\{n_k\}_{k=1}^\infty$, что

$$\lim_{k \rightarrow \infty} \frac{\sup_{g \in \mathcal{D}} |\langle f_{n_k}, g \rangle|}{a_{n_k} / \bar{b}_{n_k}} = 0. \quad (14)$$

Пусть $n \geq 1$. Тогда справедливы следующие неравенства:

$$\begin{aligned} |\langle f_n, f_n - f_0 \rangle| &= \left| \left\langle f_n, \sum_{k=1}^n c_k g_k \right\rangle \right| \leq \sum_{k=1}^n |c_k| \sup_{g \in \mathcal{D}} |\langle f_n, g \rangle| \\ &\leq \bar{b}_n \sup_{g \in \mathcal{D}} |\langle f_n, g \rangle| \frac{a_0}{a_n} \leq \frac{\sup_{g \in \mathcal{D}} |\langle f_n, g \rangle|}{a_n / \bar{b}_n} a_0 \end{aligned}$$

(в предпоследнем неравенстве была использована монотонность последовательности $\{a_n\}_{n=0}^\infty$). Отсюда, используя (14), получаем

$$\lim_{k \rightarrow \infty} |\langle f_{n_k}, f_{n_k} - f_0 \rangle| = 0.$$

Отметим, что из соотношения (14) также следует, что

$$\lim_{k \rightarrow \infty} \sup_{g \in \mathcal{D}} |\langle f_{n_k}, g \rangle| = 0. \quad (15)$$

Для этого достаточно заметить, что

$$\sup_k \frac{a_{n_k}}{\bar{b}_{n_k}} < \infty,$$

что, в свою очередь, следует из неравенства $a_{n_k} < a_0$ и неравенства треугольника

$$\bar{b}_{n_k} \geq \|f_0\| - \|f_{n_k}\| \geq a_0^{1/2} - a_1^{1/2}.$$

Соотношение (15) влечет слабую сходимость последовательности $\{f_{n_k}\}_{k=1}^\infty$ к нулю, в частности

$$\lim_{k \rightarrow \infty} |\langle f_{n_k}, f_0 \rangle| = 0.$$

Тогда

$$0 \leq \lim_{k \rightarrow \infty} a_{n_k} = \lim_{k \rightarrow \infty} \langle f_{n_k}, f_{n_k} \rangle \leq \lim_{k \rightarrow \infty} |\langle f_{n_k}, f_0 \rangle| + \lim_{k \rightarrow \infty} |\langle f_{n_k}, f_{n_k} - f_0 \rangle| = 0.$$

Получим противоречие с (12).

Случай 2. Пусть $s > 0$. Тогда найдется такое $s_0 > 0$, что для всех $n \geq 1$ выполняется

$$\sup_{g \in \mathcal{D}} |\langle f_n, g \rangle| \geq \frac{s_0 a_n}{\bar{b}_n}. \quad (16)$$

Предположим, что $n \geq 1$ шагов уже сделано. Если $(n+1)$ -й шаг возвратный, то согласно (5) и (6)

$$|c_{n+1}| = \min(|v_n(g_{n+1})|, |\langle f_n, g_{n+1} \rangle|) \geq \frac{a_n t^2}{9\bar{b}_n} \geq \frac{a_n t^2}{9\bar{b}_n}. \quad (17)$$

Имеем

$$\begin{aligned} a_{n+1} &= \langle f_{n+1}, f_{n+1} \rangle = \langle f_n - c_{n+1}g_{n+1}, f_n - c_{n+1}g_{n+1} \rangle \\ &= \langle f_n, f_n \rangle - 2c_{n+1} \langle f_n, g_{n+1} \rangle + c_{n+1}^2 \langle g_{n+1}, g_{n+1} \rangle \\ &\leq a_n - 2|c_{n+1} \langle f_n, g_{n+1} \rangle| + c_{n+1}^2, \\ a_{n+1} &\leq a_n - c_{n+1}^2, \end{aligned} \quad (18)$$

$$\bar{b}_{n+1} = \bar{b}_n + |c_{n+1}|. \quad (19)$$

Если $(n+1)$ -й шаг является жадным, то выполняется соотношение (19). Следовательно, имеем

$$a_{n+1} = \langle f_n, f_n \rangle - 2c_{n+1} \langle f_n, g_{n+1} \rangle + c_{n+1}^2 \langle g_{n+1}, g_{n+1} \rangle = a_n - c_{n+1}^2, \quad (20)$$

при этом согласно (7), (8) и (16) имеем

$$|c_{n+1}| \geq t \sup_{g \in \mathcal{D}} |\langle f_n, g \rangle| \geq t s_0 \frac{a_n}{\bar{b}_n}.$$

Пусть $T = \max\left(\frac{9}{t^2}, \frac{1}{ts_0}\right)$. Тогда

$$c_{n+1} \geq \frac{1}{T} \frac{a_n}{\bar{b}_n}. \quad (21)$$

Из (19) и неравенства

$$(1 + \gamma)^{1/T} \leq \left(1 + \frac{\gamma}{T}\right), \quad \gamma > 0, \quad T \geq 1,$$

следует, что

$$\bar{b}_{n+1}^{1/T} \leq \bar{b}_n^{1/T} \left(1 + \frac{|c_{n+1}|}{\bar{b}_n}\right)^{1/T} \leq \bar{b}_n^{1/T} \left(1 + \frac{|c_{n+1}|}{T\bar{b}_n}\right). \quad (22)$$

Далее по формулам (18), (20) и (21) имеем

$$a_{n+1} \leq a_n - c_{n+1}^2 = a_n \left(1 - \frac{c_{n+1}^2}{a_n}\right) \leq a_n \left(1 - \frac{c_{n+1}}{T\bar{b}_n}\right).$$

Отсюда и из соотношения (22) вытекает, что

$$a_{n+1}\bar{b}_{n+1}^{1/T} \leq a_n \left(1 - \frac{c_{n+1}}{T\bar{b}_n}\right) \bar{b}_n^{1/T} \left(1 + \frac{c_{n+1}}{T\bar{b}_n}\right) \leq a_n \bar{b}_n^{1/T}.$$

Следовательно, для всех $n \geq 1$ справедливо неравенство

$$a_n \bar{b}_n^{1/T} \leq a_1 \bar{b}_1^{1/T},$$

и в силу предположения о том, что алгоритм не сходится, т. е. последовательность $\{a_n\}$ отделена от нуля, получаем, что последовательность $\{\bar{b}_n\}$ ограничена. Из этого и неравенства (21) следует существование числа $c_0 > 0$, для которого выполняется неравенство

$$c_n \geq c_0, \quad n \geq 1.$$

Поэтому для любого $n \geq 1$ имеем

$$a_n \leq a_0 - nc_0^2,$$

что при больших n противоречит неотрицательности чисел a_n .

Отметим, что техника, используемая при разборе случая 2, впервые была использована в работе [9].

§ 3. Доказательство теорем 2 и 3. Оценка $a_k \# D_k$

Поскольку из соотношений (10) и (11) следует, что теорема 2 является частным случаем теоремы 3, мы ограничимся доказательством последней. Так как при умножении начальной функции на константу аппроксимант и ошибка приближения умножаются на ту же константу, достаточно доказать теорему 3 для функции $\frac{f_0}{|f_0|_{F_\beta(\mathcal{D})}}$. Для уменьшения обозначений будем доказывать теорему 3 для начальной функции f_0 такой, что $|f_0|_{F_\beta(\mathcal{D})} = 1$. Тогда

$$\|f_0\| \leq 1 \quad (23)$$

и найдутся такие числа

$$0 < M_\alpha \leq 1, \quad 0 < M_\beta \leq 1 \quad (24)$$

и

$$\alpha > \frac{1}{2} - \beta, \quad (25)$$

что

$$f_0 \in F_{\alpha, \beta}^{M_\alpha, M_\beta}(\mathcal{D}). \quad (26)$$

Обозначим через $\#D_k$ мощность множества D_k .

ЛЕММА 1. Пусть $n \geq 1$. Тогда для всех l , $1 \leq l \leq n$, справедливо неравенство

$$a_l \#D_l \leq S n^{2\beta}, \quad (27)$$

где

$$S = \max \left(\max \left(\frac{51M_\alpha}{t^2}, \frac{30}{t} M_\beta \right)^2, 3a_0, 1 \right) < K_1(t). \quad (28)$$

ЗАМЕЧАНИЕ 1. В силу неравенств (23) и (24) найдется константа $K_1(t)$ такая, что

$$S \leq K_1(t). \quad (29)$$

ДОКАЗАТЕЛЬСТВО. Докажем неравенство (27) индукцией по k . База $k = 1$ непосредственно следует из соотношения (28):

$$a_1 \#D_1 = a_1 < a_0 \leq S.$$

Предположим, что для $1 \leq l = k < n$ неравенство (27) выполняется; будем доказывать его справедливость для $l = k + 1$. Если $(k + 1)$ -й шаг является возвратным, то $D_{k+1} \subset D_k$, поэтому

$$a_{k+1} \#D_{k+1} \leq a_k \#D_k \leq S n^{2\beta}.$$

Теперь предположим, что $(k + 1)$ -й шаг является жадным. В этом случае согласно (5) для всех $g \in D_k$ выполняется по крайней мере одно из следующих неравенств:

$$\min(|v_k(g)|, |\langle f_k, g \rangle|) < \frac{a_k t^2}{9b_k}, \quad (30)$$

$$\text{sign}(v_k(g)) \neq -\text{sign}(\langle f_k, g \rangle). \quad (31)$$

Если

$$a_k \leq \left(\frac{36M_\alpha}{t^2} k^{-\alpha} \right)^2,$$

то согласно (25) и (28) имеем

$$a_{k+1} \#D_{k+1} \leq a_k(k+1) \leq \left(\frac{36M_\alpha}{t^2} k^{-\alpha} \right)^2 2k \leq \left(\frac{51M_\alpha}{t^2} \right)^2 k^{-2\alpha+1} \leq S k^{2\beta} \leq S n^{2\beta}.$$

Таким образом, осталось рассмотреть случай, когда $(k + 1)$ -й шаг является жадным и

$$a_k > \left(\frac{36M_\alpha}{t^2} k^{-\alpha} \right)^2. \quad (32)$$

Соотношение (26) и определения классов $\mathcal{A}_1(\mathcal{D})$ и $F_{\alpha,\beta}^{M_\alpha,M_\beta}(\mathcal{D})$ гарантируют существование вектора r_k , числа N , элементов $\tilde{g}_i$ словаря и коэффициентов d_i , $1 \leq i \leq N$, таких, что

$$f_0 = r_k + \sum_{i=1}^N d_i \tilde{g}_i, \quad (33)$$

$$\|r_k\| < 2M_\alpha k^{-\alpha}, \quad (34)$$

$$\sum_{i=1}^N |d_i| \leq M_\beta k^\beta. \quad (35)$$

Из определения ВЖА вытекает, что

$$f_k = f_0 - \sum_{g \in D_k} v_k(g)g.$$

Используя соотношение (33), получим следующее представление для вектора f_k :

$$f_k = r_k + \sum_{i=1}^N d_i \tilde{g}_i - \sum_{g \in D_k} v_k(g)g.$$

Имеем

$$a_k = \langle f_k, f_k \rangle = \langle f_k, r_k \rangle + \sum_{i=1}^N d_i \langle f_k, \tilde{g}_i \rangle - \sum_{g \in D_k} v_k(g) \langle f_k, g \rangle. \quad (36)$$

Из неравенства Коши–Буняковского и соотношений (32) и (34) следует оценка

$$|\langle f_k, r_k \rangle| \leq \|r_k\| a_k^{1/2} \leq 2M_\alpha k^{-\alpha} \frac{a_k}{a_k^{1/2}} \leq \frac{2M_\alpha k^{-\alpha} a_k}{36M_\alpha k^{-\alpha} / t^2} = \frac{t^2}{18} a_k.$$

Отсюда, из неравенства (35) и равенства (36) вытекает, что

$$\begin{aligned} a_k &\leq |\langle f_k, r_k \rangle| + \sum_{i=1}^N |d_i| \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| - \sum_{g \in D_k} v_k(g) \langle f_k, g \rangle, \\ a_k \left(1 - \frac{t^2}{18} \right) &\leq M_\beta k^\beta \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| - \sum_{g \in D_k} v_k(g) \langle f_k, g \rangle. \end{aligned}$$

Учитывая справедливость хотя бы одного из неравенств (30) и (31), получаем

$$\begin{aligned}
 a_k \left(1 - \frac{t^2}{18}\right) &\leq M_\beta k^\beta \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| - \sum_{\substack{g \in D_k \\ \text{sign}(v_k(g)) \neq -\text{sign}(\langle f_k, g \rangle)}} v_k(g) \langle f_k, g \rangle \\
 &\quad - \sum_{\substack{g \in D_k \\ |v_k(g)| < a_k t^2 / (9b_k) \\ \text{sign}(v_k(g)) = -\text{sign}(\langle f_k, g \rangle)}} v_k(g) \langle f_k, g \rangle - \sum_{\substack{g \in D_k \\ |\langle f_k, g \rangle| < a_k t^2 / (9b_k) \\ \text{sign}(v_k(g)) = -\text{sign}(\langle f_k, g \rangle)}} v_k(g) \langle f_k, g \rangle \\
 &\leq M_\beta k^\beta \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| + \sum_{\substack{g \in D_k \\ |v_k(g)| < a_k t^2 / (9b_k)}} |v_k(g)| |\langle f_k, g \rangle| \\
 &\quad + \sum_{\substack{g \in D_k \\ |\langle f_k, g \rangle| < a_k t^2 / (9b_k)}} |v_k(g)| |\langle f_k, g \rangle| \\
 &\leq \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| \left(M_\beta k^\beta + \sum_{\substack{g \in D_k \\ |v_k(g)| < a_k t^2 / (9b_k)}} |v_k(g)| \right) + \sum_{\substack{g \in D_k \\ |\langle f_k, g \rangle| < a_k t^2 / (9b_k)}} |v_k(g)| \frac{a_k t^2}{9b_k}.
 \end{aligned}$$

По определению (4) чисел b_k преобразуем последнее слагаемое:

$$\sum_{\substack{g \in D_k \\ |\langle f_k, g \rangle| < a_k t^2 / (9b_k)}} |v_k(g)| \frac{a_k t^2}{9b_k} \leq \frac{a_k t^2}{9b_k} \sum_{g \in D_k} |v_k(g)| = \frac{a_k t^2}{9b_k} b_k = \frac{a_k t^2}{9}.$$

Имеем

$$a_k \left(1 - \frac{t^2}{6}\right) = a_k \left(1 - \frac{t^2}{9} - \frac{t^2}{18}\right) \leq \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| \left(M_\beta k^\beta + \sum_{\substack{g \in D_k \\ |v_k(g)| < a_k t^2 / (9b_k)}} |v_k(g)| \right).$$

Это неравенство влечет два следствия:

$$\begin{aligned}
 a_k \left(1 - \frac{t^2}{6}\right) &\leq \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| \left(M_\beta k^\beta + \sum_{\substack{g \in D_k \\ |v_k(g)| < a_k t^2 / (9b_k)}} |v_k(g)| \right) \\
 &\leq \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| \left(M_\beta k^\beta + \sum_{g \in D_k} |v_k(g)| \right) = \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| (M_\beta k^\beta + b_k)
 \end{aligned}$$

и

$$\begin{aligned}
 a_k \left(1 - \frac{t^2}{6}\right) &\leq \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| \left(M_\beta k^\beta + \sum_{\substack{g \in D_k \\ |v_k(g)| < a_k t^2 / (9b_k)}} |v_k(g)| \right) \\
 &\leq \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| \left(M_\beta k^\beta + \frac{a_k t^2}{9b_k} \# D_k \right).
 \end{aligned}$$

Используя эти неравенства, получаем

$$\begin{aligned}
 a_k \left(1 - \frac{t^2}{6}\right) &\leq \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| \left(M_\beta k^\beta + \min \left(b_k, \frac{a_k t^2}{9b_k} \# D_k \right) \right) \\
 &\leq \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| \left(M_\beta k^\beta + \sqrt{b_k \frac{a_k t^2}{9b_k} \# D_k} \right), \\
 \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| &\geq \frac{a_k (1 - \frac{t^2}{6})}{M_\beta k^\beta + \frac{t}{3} (a_k \# D_k)^{1/2}}, \\
 \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| &\geq \frac{\frac{5}{6} a_k}{M_\beta n^\beta + \frac{t}{3} (a_k \# D_k)^{1/2}}. \tag{37}
 \end{aligned}$$

Используя индукционное предположение (27) для $l = k$ и определение (28), преобразуем неравенство (37):

$$\begin{aligned}
 \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| &\geq \frac{\frac{5}{6} a_k}{M_\beta n^\beta + \frac{t}{3} (a_k \# D_k)^{1/2}} \geq \frac{\frac{5}{6} a_k}{n^\beta (M_\beta + \frac{t}{3} S^{1/2})} \\
 &\geq \frac{\frac{5}{6} a_k}{n^\beta (S^{1/2} \frac{t}{30} + \frac{t}{3} S^{1/2})} \geq \frac{1.2 a_k}{t n^\beta S^{1/2}}. \tag{38}
 \end{aligned}$$

Поскольку $(k + 1)$ -й шаг является жадным, то

$$\begin{aligned}
 |c_{k+1}| &\geq t \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle|, \\
 a_{k+1} = a_k - c_{k+1}^2 &\leq a_k - t^2 \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle|^2 \leq a_k - \frac{1.44 a_k^2}{S n^{2\beta}}.
 \end{aligned}$$

В силу определения (28) выполняется $S \geq 3a_0$. Легко видеть, что функция

$$h(x) = x - \frac{1.44 x^2}{S n^{2\beta}}$$

возрастает на отрезке $[0, a_0]$. Действительно,

$$\frac{dh(x)}{dx} = 1 - \frac{2.88x}{S n^{2\beta}} \geq 1 - \frac{2.88x}{3a_0} > 0, \quad 0 \leq x \leq 3a_0.$$

Поэтому, используя индукционное предположение, можно получить

$$\begin{aligned}
 a_{k+1} \leq h(a_k) \leq h \left(\frac{S n^{2\beta}}{\# D_k} \right) &= S n^{2\beta} \left(\frac{1}{\# D_k} - \frac{1.44}{\# D_k^2} \right) \leq \frac{S n^{2\beta}}{\# D_{k+1}} \leq \frac{S n^{2\beta}}{\# D_{k+1}}, \\
 a_{k+1} \# D_{k+1} &\leq S n^{2\beta},
 \end{aligned}$$

что и доказывает его истинность для всех k , $1 \leq k \leq n$.

Лемма 1 утверждает, что неравенство (27) верно для $k = l = n - 1$. Тогда при сделанных предположениях о том, что $(k + 1)$ -й шаг является жадным и выполнено (32), верна оценка (38). Отсюда вытекает

СЛЕДСТВИЕ 1. Если $(k+1)$ -й шаг является жадным, $k \geq 1$, и

$$a_k > \left(\frac{36M_\alpha}{t^2} k^{-\alpha} \right)^2,$$

то справедливо неравенство

$$c_{k+1} \geq t \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| \geq \frac{1.2a_k}{(k+1)^\beta S^{1/2}} \geq \frac{a_k}{2k^\beta S^{1/2}}.$$

§ 4. Доказательство теорем 2 и 3. Основная лемма

Приведем некоторые числовые неравенства, которые будут использованы в дальнейшем. Для чисел γ и κ , $0 \leq \gamma \leq 1$, $0 \leq \kappa \leq 1$, справедливы неравенства

$$\frac{1}{1+\gamma} \leq 1 - \frac{\gamma}{2}, \quad (39)$$

$$\frac{1-\kappa}{(1+\gamma)^2} \geq \frac{1-\kappa}{1+3\gamma} \geq 1-\kappa-3\gamma. \quad (40)$$

Рассмотрим функцию

$$\varphi(u) = \frac{u}{\ln^2(u/(3a_0))}.$$

Ясно, что она обладает следующими свойствами:

- 1) $\varphi(u)$ возрастает при $u \in (0, a_0]$;
- 2) $\frac{\varphi(u)}{u}$ возрастает при $u \in (0, a_0]$;
- 3) $\varphi(u) < u$, $u \in (0, a_0]$.

Предположим, что числа $u \in (0, a_0]$ и v связаны соотношением

$$v = (1-\kappa)u, \quad 0 \leq \kappa < 1.$$

Тогда имеют место оценки

$$\begin{aligned} \frac{\varphi(v)}{\varphi(u)} &= \frac{v \ln^2(u/(3a_0))}{u \ln^2(v/(3a_0))} = (1-\kappa) \frac{\ln^2(u/(3a_0))}{(-\ln(u/(3a_0)) - \ln(1-\kappa))^2} \\ &\leq (1-\kappa) \frac{\ln^2(u/(3a_0))}{(-\ln(u/(3a_0)) + \kappa)^2} \leq (1-\kappa) \left(\frac{1}{1 + \frac{\kappa}{-\ln(u/(3a_0))}} \right)^2 \\ &\leq (1-\kappa) \frac{1}{1 + \frac{\kappa}{-\ln(u/(3a_0))}}. \end{aligned}$$

Из неравенства (39) следует, что

$$\frac{\varphi(v)}{\varphi(u)} \leq (1-\kappa) \frac{1}{1 + \frac{\kappa}{-\ln(u/(3a_0))}} \leq (1-\kappa) \left(1 - \frac{\kappa}{-2\ln(u/(3a_0))} \right).$$

Если $0 \leq \kappa \leq 1/2$, то

$$\frac{\varphi(v)}{\varphi(u)} \leq 1 - \kappa - \frac{\kappa}{-4\ln(u/(3a_0))}. \quad (41)$$

Если $1/2 < \kappa < 1$, то

$$\begin{aligned} \frac{\varphi(v)}{\varphi(u)} \frac{u}{v} &\leq 1 - \frac{\kappa}{-2 \ln(u/(3a_0))} \leq 1 - \frac{1}{-4 \ln(u/(3a_0))}, \\ \varphi(u) \frac{v}{u} &\geq \varphi(v) \frac{1}{1 - \frac{1}{-4 \ln(u/(3a_0))}}. \end{aligned} \quad (42)$$

Пусть M – минимальное число, которое удовлетворяет неравенствам

$$\frac{M}{2} \geq 10a_0^{1/2}, \quad (43)$$

$$\frac{12 \left(\frac{36M_\alpha}{t^2} \right)^2}{\beta M^2/4} \leq 1, \quad (44)$$

$$48S^{1/2} \leq M. \quad (45)$$

Тогда в силу неравенств (23), (24) и (29) найдется $K_2(t)$ такое, что

$$M \leq K_2(t). \quad (46)$$

Имеет место следующая лемма.

ЛЕММА 2. Пусть $a_n > 0$, $n \geq 1$. Тогда выполняется неравенство

$$\frac{a_n}{b_n^2} \geq \frac{\varphi(a_n)}{M^2 n^{2\beta}}. \quad (47)$$

ДОКАЗАТЕЛЬСТВО. Докажем утверждение леммы по индукции. Справедливость базы $\frac{a_1}{b_1^2} \geq \frac{\varphi(a_1)}{M^2}$ вытекает из свойства 3), неравенства (43) и того, что

$$b_1^2 = c_1^2 \leq a_0.$$

Если $a_k = 0$, то $a_l = 0$, $l \geq k$. Предположим, что неравенство (47) выполняется для $n = k$:

$$\frac{a_k}{b_k^2} \geq \frac{\varphi(a_k)}{M^2 k^{2\beta}}. \quad (48)$$

Докажем, что неравенство (47) истинно и для $n = k + 1$. Если $(k + 1)$ -й шаг возвратный, то

$$\begin{aligned} a_{k+1} &< a_k, \\ v_{k+1}(g_{k+1}) &= v_k(g_{k+1}) + c_{k+1}, \quad |v_{k+1}(g_{k+1})| \leq |v_k(g_{k+1})|, \\ b_{k+1} &\leq b_k, \end{aligned}$$

и согласно (48) имеем

$$\frac{a_{k+1}}{b_{k+1}^2} \geq \frac{a_{k+1}}{b_k^2} \geq \frac{a_k}{b_k^2} \frac{a_{k+1}}{a_k} \geq \frac{1}{M^2 k^{2\beta}} \frac{\varphi(a_k)}{a_k} a_{k+1}.$$

Используя свойство 2), получаем

$$\frac{a_{k+1}}{b_{k+1}^2} \geq \frac{a_{k+1}}{M^2 k^{2\beta}} \frac{\varphi(a_k)}{a_k} \geq \frac{a_{k+1}}{M^2 (k+1)^{2\beta}} \frac{\varphi(a_{k+1})}{a_{k+1}} \geq \frac{\varphi(a_{k+1})}{M^2 (k+1)^{2\beta}}.$$

Теперь рассмотрим случай, когда $(k+1)$ -й шаг является жадным. Если

$$b_k \leq \frac{M}{2} k^\beta \left(-\ln \frac{a_k}{3a_0} \right),$$

то в силу неравенства (43) и того, что $(-\ln(a_k/(3a_0))) > 1$, получаем

$$\begin{aligned} b_{k+1} &\leq b_k + |c_{k+1}| \leq b_k + \sup_{g \in \mathcal{D}} |\langle f_k, g \rangle| \leq b_k + a_k^{1/2} \leq b_k + a_0^{1/2}, \\ b_{k+1} &\leq b_k + a_0^{1/2} \leq \frac{M}{2} k^\beta \left(-\ln \frac{a_k}{3a_0} \right) + a_0^{1/2} \\ &\leq M k^\beta \left(-\ln \frac{a_k}{3a_0} \right) \leq M(k+1)^\beta \left(-\ln \frac{a_{k+1}}{3a_0} \right), \\ \frac{a_{k+1}}{b_{k+1}^2} &\geq \frac{a_{k+1}}{(M(k+1)^\beta (-\ln(a_{k+1}/(3a_0))))^2} = \frac{\varphi(a_{k+1})}{M^2(k+1)^{2\beta}}. \end{aligned} \quad (49)$$

Поэтому достаточно рассмотреть случай

$$b_k > \frac{M}{2} k^\beta \left(-\ln \frac{a_k}{3a_0} \right). \quad (50)$$

Если $a_{k+1}/a_k < 1/2$, то найдется κ , для которого выполнено

$$a_{k+1} = a_k(1 - \kappa), \quad \frac{1}{2} < \kappa < 1.$$

В силу соотношений (50), (43) и (49) справедливы оценки

$$\begin{aligned} b_{k+1} &\leq b_k + a_0^{1/2} \leq b_k \left(1 + \frac{a_0^{1/2}}{b_k} \right) \\ &\leq b_k \left(1 + \frac{a_0^{1/2}}{\frac{M}{2} k^\beta \left(-\ln \frac{a_k}{3a_0} \right)} \right) \leq b_k \left(1 + \frac{1}{-10 \ln \frac{a_k}{3a_0}} \right), \\ \frac{b_k}{b_{k+1}} &\geq \frac{1}{1 + \frac{1}{-10 \ln(a_k/(3a_0))}} \geq 1 - \frac{1}{-10 \ln \frac{a_k}{3a_0}}, \\ \left(\frac{b_k}{b_{k+1}} \right)^2 &\geq 1 - \frac{1}{-5 \ln \frac{a_k}{3a_0}}. \end{aligned} \quad (51)$$

Применяя неравенство (48), получим

$$\frac{a_{k+1}}{b_{k+1}^2} = \frac{a_k}{b_k^2} \frac{a_{k+1}}{a_k} \left(\frac{b_k}{b_{k+1}} \right)^2 \geq \frac{1}{M^2 k^{2\beta}} \varphi(a_k) \frac{a_{k+1}}{a_k} \left(\frac{b_k}{b_{k+1}} \right)^2.$$

Используя неравенство (42) при $u = a_k$ и $v = a_{k+1}$ и неравенство (51), имеем

$$\begin{aligned} \frac{a_{k+1}}{b_{k+1}^2} &\geq \frac{1}{M^2 k^{2\beta}} \varphi(a_k) \frac{a_{k+1}}{a_k} \left(\frac{b_k}{b_{k+1}} \right)^2 \\ &\geq \frac{1}{M^2(k+1)^{2\beta}} \varphi(a_{k+1}) \frac{1}{1 - \frac{1}{-4 \ln(a_k/(3a_0))}} \left(1 - \frac{1}{-5 \ln \frac{a_k}{3a_0}} \right) \geq \frac{\varphi(a_{k+1})}{M^2(k+1)^{2\beta}}. \end{aligned}$$

Если $a_{k+1}/a_k \geq 1/2$, то в силу (20) получаем

$$a_{k+1} = a_k - c_{k+1}^2, \quad a_{k+1} = a_k(1 - \kappa), \quad \kappa = \frac{c_{k+1}^2}{a_k}, \quad 0 \leq \kappa \leq \frac{1}{2}. \quad (52)$$

Этот случай разбивается на два подслучая:

$$|c_{k+1}| \leq \frac{b_k \beta}{k}, \quad (53)$$

$$|c_{k+1}| > \frac{b_k \beta}{k}. \quad (54)$$

Если выполняется неравенство (53), то

$$\begin{aligned} \frac{b_{k+1}}{b_k} &\leq \frac{b_{k+1} + |c_{k+1}|}{b_k} \leq 1 + \frac{\beta}{k} \leq \left(1 + \frac{1}{k}\right)^\beta = \frac{(k+1)^\beta}{k^\beta}, \\ \frac{b_{k+1}^2}{b_k^2} &\leq \frac{(k+1)^{2\beta}}{k^{2\beta}}. \end{aligned}$$

Используя свойство 2) и неравенство (48), имеем

$$\frac{a_{k+1}}{b_{k+1}^2} = \frac{a_{k+1}}{\frac{b_{k+1}^2}{b_k^2} b_k^2} \geq \frac{k^{2\beta}}{(k+1)^{2\beta}} \frac{a_k}{b_k^2} \frac{a_{k+1}}{a_k} \geq \frac{k^{2\beta}}{(k+1)^{2\beta}} \frac{\varphi(a_k)}{M^2 k^{2\beta}} \frac{a_{k+1}}{a_k} \geq \frac{\varphi(a_{k+1})}{M^2 (k+1)^{2\beta}}.$$

Если выполняется неравенство (54), то в силу (50) $b_k \geq M/2$, а согласно (43) получаем

$$\frac{|c_{k+1}|}{b_k} \leq \frac{a_0^{1/2}}{M/2} < 1.$$

Из неравенства $b_{k+1} \leq b_k + |c_{k+1}|$ и соотношений (52) следует, что

$$\frac{a_{k+1}}{b_{k+1}^2} \geq \frac{a_k}{b_k^2} \frac{1 - \frac{c_{k+1}^2}{a_k}}{\left(1 + \frac{|c_{k+1}|}{b_k}\right)^2}.$$

Применяя неравенства (40) и (48), получим

$$\frac{a_{k+1}}{b_{k+1}^2} \geq \frac{\varphi(a_k)}{M^2 k^{2\beta}} \left(1 - \frac{c_{k+1}^2}{a_k} - 3 \frac{c_{k+1}}{b_k}\right).$$

Используя (41) при $u = a_k$, $v = a_{k+1}$, $\kappa = \frac{c_{k+1}^2}{a_k}$, находим

$$\begin{aligned} \frac{a_{k+1}}{b_{k+1}^2} &\geq \frac{\varphi(a_k)}{M^2 k^{2\beta}} \left(1 - \frac{c_{k+1}^2}{a_k} - 3 \frac{c_{k+1}}{b_k}\right) \geq \frac{\varphi(a_{k+1})}{M^2 k^{2\beta}} \frac{1 - \frac{c_{k+1}^2}{a_k} - 3 \frac{c_{k+1}}{b_k} \frac{a_k}{b_k c_{k+1}}}{1 - \kappa - \frac{\kappa}{-4 \ln(a_k/(3a_0))}} \\ &\geq \frac{\varphi(a_{k+1})}{M^2 (k+1)^{2\beta}} \frac{1 - \kappa - 3\kappa \frac{a_k}{b_k c_{k+1}}}{1 - \kappa - \frac{\kappa}{-4 \ln(a_k/(3a_0))}}. \end{aligned}$$

Таким образом, для завершения доказательства леммы достаточно проверить, что

$$3 \frac{a_k}{b_k c_{k+1}} \leq \frac{1}{-4 \ln \frac{a_k}{3a_0}}. \quad (55)$$

Если

$$a_k \leq \left(\frac{36M_\alpha}{t^2} k^{-\alpha} \right)^2,$$

то, используя неравенства (54) и (50), получаем

$$3 \frac{a_k}{b_k c_{k+1}} \leq 3 \frac{k a_k}{\beta b_k^2} \leq 3 \frac{k \left(\frac{36M_\alpha}{t^2} k^{-\alpha} \right)^2}{\beta \frac{M^2}{4} k^{2\beta} \ln^2 \frac{a_k}{3a_0}}.$$

Учитывая соотношения (25) и (44), продолжим оценку:

$$3 \frac{a_k}{b_k c_{k+1}} \leq 3 \frac{k \left(\frac{36M_\alpha}{t^2} k^{-\alpha} \right)^2}{\beta \frac{M^2}{4} k^{2\beta} \ln^2 \frac{a_k}{3a_0}} \leq \frac{12 \left(\frac{36M_\alpha}{t^2} \right)^2}{\beta \frac{M^2}{4}} \frac{1}{-4 \ln \frac{a_k}{3a_0}} \leq \frac{1}{-4 \ln \frac{a_k}{3a_0}}.$$

Остается рассмотреть случай, когда

$$a_k > \left(\frac{36M_\alpha}{t^2} k^{-\alpha} \right)^2.$$

Из следствия 1 получаем

$$c_{k+1} \geq \frac{a_k}{2k^\beta S^{1/2}}.$$

Тогда, учитывая соотношения (50) и (45), имеем

$$3 \frac{a_k}{b_k c_{k+1}} \leq 3 \frac{2k^\beta S^{1/2}}{b_k} \leq \frac{6S^{1/2}}{\frac{M}{2} \left(-\ln \frac{a_k}{3a_0} \right)} = \frac{\frac{48}{M} S^{1/2}}{-4 \ln \frac{a_k}{3a_0}} \leq \frac{1}{-4 \ln \frac{a_k}{3a_0}},$$

что доказывает (55) и завершает доказательство леммы 2.

§ 5. Доказательство теорем 2 и 3. Окончание

Пусть $n \geq 1$. В силу неравенств (18) и (20) имеет место оценка

$$a_{n+1} \leq a_n - c_{n+1}^2 = a_n \left(1 - \frac{c_{n+1}^2}{a_n} \right).$$

Если $(n+1)$ -й шаг является возвратным, то согласно (5) и (6)

$$c_{n+1} \geq \frac{a_n t^2}{9b_n}, \quad a_{n+1} \leq a_n \left(1 - \frac{a_n t^4}{81b_n^2} \right),$$

и по лемме 2 получаем

$$a_{n+1} \leq a_n \left(1 - \frac{\varphi(a_n) t^4}{81M^2 n^{2\beta}} \right). \quad (56)$$

Если $(n+1)$ -й шаг является жадным и

$$a_n > \left(\frac{36M_\alpha}{t^2} n^{-\alpha} \right)^2, \quad (57)$$

то согласно следствию 1 справедливо неравенство

$$c_{n+1} \geq \frac{a_n}{2n^\beta S^{1/2}}.$$

Поэтому

$$a_{n+1} \leq a_n \left(1 - \frac{a_n}{4n^{2\beta} S} \right).$$

Объединяя это неравенство с неравенством (56) и учитывая свойство 3), имеем, что для всех n , удовлетворяющих (57), справедливы неравенства

$$\begin{aligned} a_{n+1} &\leq a_n \left(1 - \frac{\varphi(a_n)t^4}{n^{2\beta} \max(4S, 81M^2)} \right), \\ \frac{\frac{a_{n+1}}{\ln^2(a_{n+1}/(3a_0))}}{n^{2\beta}} &\leq \frac{\frac{a_n}{\ln^2(a_n/(3a_0))}}{n^{2\beta}} \left(1 - \frac{\varphi(a_n)t^4}{n^{2\beta} \max(4S, 81M^2)} \right), \\ \varphi(a_{n+1}) \frac{t^4}{\max(4S, 81M^2)(n+1)^{2\beta}} &\leq \varphi(a_n) \frac{t^4}{\max(4S, 81M^2)n^{2\beta}} \left(1 - \frac{\varphi(a_n)t^4}{n^{2\beta} \max(4S, 81M^2)} \right). \end{aligned}$$

Обозначая для всех $n \geq 1$

$$\psi_n = \frac{\varphi(a_n)t^4}{n^{2\beta} \max(4S, 81M^2)},$$

получаем

$$\psi_{n+1} \leq \psi_n (1 - \psi_n). \quad (58)$$

Если неравенство (57) не выполняется, то с учетом (25) имеем

$$\begin{aligned} \psi_{n+1} &= \frac{\varphi(a_{n+1})t^4}{n^{2\beta} \max(4S, 81M^2)} \leq \frac{\varphi(a_n)t^4}{n^{2\beta} \max(4S, 81M^2)} \leq \frac{a_n t^4}{n^{2\beta} \max(4S, 81M^2)} \\ &\leq \frac{\left(\frac{36M_\alpha}{t^2} n^{-\alpha} \right)^2 t^4}{n^{2\beta} \max(4S, 81M^2)} \leq \frac{1}{n} \frac{(36M_\alpha)^2}{\max(4S, 81M^2)} \leq \frac{\frac{2(36M_\alpha)^2}{\max(4S, 81M^2)}}{n+1}. \end{aligned} \quad (59)$$

Пусть

$$A = \max \left(1, \psi_1, \frac{2(36M_\alpha)^2}{\max(4S, 81M^2)} \right).$$

Учитывая (24) и (28), получим существование константы $K_3(t)$ такой, что

$$A \leq K_3(t). \quad (60)$$

Тогда, объединяя неравенства (59) и (58), получаем $\psi_1 \leq A$ и (если имеет место (57))

$$\psi_{n+1} \leq \psi_n \left(1 - \frac{\psi_n}{A} \right)$$

или (в противном случае)

$$\psi_{n+1} \leq \frac{A}{n+1}.$$

Используя лемму 3.4 из работы [2], имеем, что для всех $n \geq 1$ справедлива оценка

$$\psi_n \leq \frac{A}{n}.$$

Таким образом,

$$a_n \leq \varphi^{-1} \left(\frac{An^{2\beta} \max(4S, 81M^2)}{nt^4} \right).$$

Из этого неравенства с учетом определения функции $\varphi(\cdot)$ и неравенств (29), (46) и (60) получаем существование такой константы $K = K(t) > 0$, что для всех $n \geq 1$ выполнено неравенство

$$a_n \leq K^2 n^{-1+2\beta} \ln^2 n.$$

Таким образом, имеем

$$\|f_n\| = a_n^{1/2} \leq K n^{-1/2+\beta} \ln n.$$

§ 6. Доказательство теоремы 4

Разобьем доказательство теоремы 4 на две части.

1. Докажем, что для произвольных чисел $\alpha, \beta > 0$, $\alpha + \beta < 1/2$, словаря $\mathcal{D}$ и $\varepsilon > 0$ справедливо включение

$$A_{\alpha,\beta}(\mathcal{D}) \subset A^{\frac{\alpha}{2(\alpha+\beta)}-\varepsilon}. \quad (61)$$

Пусть $f \in A_{\alpha,\beta}(\mathcal{D})$. Тогда существуют $M_\alpha > 0$ и $M_\beta > 0$, для которых $f \in A_{\alpha,\beta}^{M_\alpha, M_\beta}(\mathcal{D})$. Проверим, что $f \in F_{\alpha',\beta'}^{M_\alpha, 2M_\beta}(\mathcal{D})$, где

$$\alpha' = \frac{\alpha}{2(\alpha+\beta)}, \quad \beta' = \frac{\beta}{2(\alpha+\beta)}.$$

Действительно, пусть $n \in \mathbb{N}$; по определению класса $A_{\alpha,\beta}^{M_\alpha, M_\beta}(\mathcal{D})$ для $m = \left\lceil n^{\frac{1}{2(\alpha+\beta)}} \right\rceil + 1$ найдется такой вектор

$$f_m = f - \sum_{i=0}^{m-1} c_i g_i,$$

что

$$\begin{aligned} \|f_m\| &\leq M^\alpha m^{-\alpha} \leq M^\alpha n^{-\alpha'}, \\ |f - f_m|_{\mathcal{A}_1} &\leq \sum_{i=0}^{m-1} |c_i| \leq M_\beta m^\beta \leq 2M_\beta n^{\beta'}. \end{aligned}$$

Таким образом, функция f принадлежит классу $F_{\beta'+\varepsilon/2}(\mathcal{D})$ для произвольного $\varepsilon > 0$. Тогда по теореме 3 выполняется оценка

$$\|f - G_n(f, \mathcal{D})\| \leq K n^{-1/2+\beta'+\varepsilon/2} \ln n \leq K_1 n^{-\frac{\alpha}{2(\alpha+\beta)}+\varepsilon},$$

откуда следует требуемая оценка n -членного приближения функции f и доказательство включения (61).

2. Докажем, что для произвольных чисел $\alpha, \beta > 0$, $\alpha + \beta < 1/2$, существуют словарь $\mathcal{D}$, функция $f_0 \in A_{\alpha, \beta}(\mathcal{D})$ и $C > 0$ такие, что для произвольного $m \in \mathbb{N}$ выполняется неравенство

$$\sigma_m(f, \mathcal{D}) \geq C m^{-\frac{\alpha}{2(\alpha+\beta)}}. \quad (62)$$

Пусть $D = \{e_n\}_{n=1}^\infty$ – ортонормированная система в H . Пусть

$$\begin{aligned} \delta &= \frac{2\alpha + \beta}{2(\alpha + \beta)} > \frac{1}{2}, \\ f_0 &= \sum_{n=1}^\infty n^{-\delta} e_n. \end{aligned} \quad (63)$$

Ясно, что для $m \in \mathbb{N}$

$$\sigma_m(f_0, \mathcal{D}) = \left(\sum_{n=m+1}^\infty n^{-2\delta} \right)^{1/2} \geq C_1 m^{\frac{1-2\delta}{2}} = C_1 m^{\frac{-\alpha}{2(\alpha+\beta)}}. \quad (64)$$

Проверим, что

$$f_0 \in A_{\alpha, \beta}(\mathcal{D}). \quad (65)$$

Пусть $m \in \mathbb{N}$. Положим $l = \lceil m^{\frac{\beta}{1-\delta}} \rceil$. Тогда, используя (63), получаем

$$\begin{aligned} \left\| f_0 - \sum_{k=1}^l k^{-\delta} e_k \right\| &= \left(\sum_{k=l+1}^\infty k^{-2\delta} \right)^{1/2} \leq C_2 l^{\frac{1-2\delta}{2}} \leq C_3 m^{\frac{(1-2\delta)\beta}{2(1-\delta)}} = C_3 m^{-\alpha}, \\ \left| \sum_{k=1}^l k^{-\delta} e_k \right|_{\mathcal{A}_1(\mathcal{D})} &= \sum_{k=1}^l k^{-\delta} \leq C_4 l^{1-\delta} \leq C_4 m^\beta. \end{aligned}$$

Это доказывает истинность включения (65) и вместе с неравенством (64) завершает доказательство оценки (62). Теорема 4 непосредственно следует из справедливости соотношений (61) и (62).

Список литературы

1. Галатенко В. В., Лившиц Е. Д., “Об устойчивости жадных разложений к ошибкам в вычислении коэффициентов”, *Тез. докладов 12-й Саратовской зимней школы “Современные проблемы теории функций и их приложения”*, УНЦ “Колледж”, Саратов, 2004, 52–53.
2. Лившиц Е. Д., “О скорости сходимости чисто жадного алгоритма”, *Матем. заметки*, **76:4** (2004), 539–552.
3. Лившиц Е. Д., Темляков В. Н., “О сходимости слабого гриди-алгоритма”, *Тр. Матем. ин-та им. В. А. Стеклова РАН*, **232** (2001), 236–247.
4. Сильниченко А., “О скорости сходимости жадного алгоритма”, *Матем. заметки*, **76:4** (2004), 628–632.
5. Стечкин С. Б., “Об абсолютной сходимости ортогональных рядов”, *Докл. АН СССР*, **102:1** (1955), 37–40.

6. Barron A., “Universal approximation bounds for superposition of n sigmoidal functions”, *IEEE Transactions on Information Theory*, **39** (1993), 930–945.
7. Buhlmann P., <http://stat.ethz.ch/~buhlmann/bibliog.html>, 2004.
8. Davis G., Mallat S., Avellaneda M., “Adaptive greedy approximations”, *Constructive Approximation*, **13** (1997), 57–98.
9. DeVore R. A., Temlyakov V. N., “Some remarks on Greedy Algorithms”, *Adv. Comput. Math.*, **5** (1996), 173–187.
10. Friedman J. H., Stueeuzle W., “Projection pursuit regression”, *J. Amer. Statist. Assoc.*, **76** (1981), 817–823.
11. Galatenko V. V., Livshitz E. D., “On the convergence of Approximate Weak Greedy Algorithms”, *East J. Approx.*, **9:1** (2003), 43–49.
12. Gribonval R., Nielsen M., “Approximate Weak Greedy Algorithms”, *Adv. Comput. Math.*, **14:4** (2001), 361–378.
13. Jones L. K., “On a conjecture of Huber concerning the convergence of PP-regression”, *Ann. Statist.*, **15** (1987), 880–882.
14. Jones L., “A simple lemma on greedy approximation in Hilbert space and convergence rates for projection pursuit regression and network training”, *Ann. Statist.*, **20** (1992), 608–613.
15. Konyagin S. V., Temlyakov V. N., “Rate of convergence of Pure Greedy Algorithm”, *East J. Approx.*, **5** (1999), 493–499.
16. Livshitz E. D., Temlyakov V. N., “Two lower estimates in greedy approximation”, *Constructive Approximation*, **19** (2003), 509–523.
17. Rejtő L., Walter G. G., “Remarks on projection pursuit regression and density estimation”, *Stochastic Analysis and Application*, **10** (1992), 213–222.
18. Schmidt E., “Zur Theorie der linearen und nichtlinearen integralgleichungen. I”, *Math. Ann.*, **63** (1906–1907), 433–476.
19. Temlyakov V. N., “Weak greedy algorithms”, *Adv. Comput. Math.*, **12:2**, 3 (2000), 213–227.
20. Temlyakov V. N., *A criterion for convergence of Weak Greedy Algorithms* <http://www.math.sc.edu/~imip/00.html> -- 00:21.
21. Temlyakov V. N., “Nonlinear Methods of Approximation”, *Found. Comput. Math.*, **3** (2003), 33–107.

Е. Д. Лившиц (E. D. LIVSHITS)
 E-mail: livshitz@mccme.ru

Поступило в редакцию
 26.11.2004