Math-Net.Ru

Общероссийский математический портал

Е. Д. Лившиц, О нижних оценках скорости сходимости жадных алгоритмов, *Изв. РАН. Сер. матем.*, 2009, том 73, выпуск 6, 125–144

DOI: <https://doi.org/10.4213/im2742>

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: 18.209.158.208

26 апреля 2023 г., 13:26:48

УДК 517.518.8+519.651.3

Е. Д. Лившиц

О нижних оценках скорости сходимости жадных алгоритмов

Получена нижняя оценка скорости сходимости чисто жадного алгоритма в пространствах $\mathcal{A}_0(\mathcal{D})$ и $\mathcal{A}_1(\mathcal{D})$, которая оказывается весьма близкой к наилучшей известной верхней оценке. Также получена точная нижняя оценка скорости сходимости ортогонального жадного алгоритма в пространстве $\mathcal{A}_0(\mathcal{D})$.

Библиография: 16 наименований.

Ключевые слова: чисто жадный алгоритм, наилучшее n -членное приближение, интерполяционные классы, скорость сходимости.

§ 1. Введение

Настоящая работа посвящена изучению скорости сходимости жадных алгоритмов. Напомним основные определения, постановки задач и полученные ранее результаты.

Рассмотрим действительное сепарабельное гильбертово пространство H со скалярным произведением $\langle \cdot, \cdot \rangle$. Множество $\mathcal{D}$, $\mathcal{D} \subset H$, будем называть *словарем*, если оно состоит из векторов с единичной нормой и замыкание линейной оболочки $\mathcal{D}$ совпадает со всем H :

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

Для произвольной функции $f \in H$ требуется найти конечную линейную комбинацию элементов словаря, достаточно хорошо ее приближающую:

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

а также вычислить величину *наилучшего m -членного приближения*:

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

В последние годы для решения прикладных задач возникла необходимость использования переполненных словарей $\mathcal{D}$ таких, как словарь плоских волн (ridge-функций), как RBF-словари, переполненные системы случайных векторов и т. д. Для этих систем невозможно использовать разложение по базису, и

Работа выполнена при финансовой поддержке РФФИ (гранты № 08-01-00799, 06-01-00160).

получение m -членных приближений (1), пусть даже не наилучших, является весьма актуальной задачей.

Жадные алгоритмы предлагают эффективную процедуру получения подобных разложений. В настоящее время разработано множество модификаций жадных алгоритмов (см. [1], [2]). В этой работе будут рассматриваться две наиболее ранние и известные [3] его версии: чисто жадный алгоритм и ортогональный жадный алгоритм.

ОПРЕДЕЛЕНИЕ 1. Пусть задан словарь $\mathcal{D}$, и пусть $f_0^{\text{PGA}} := f_0 \in H$ – целевая функция. Определим последовательность аппроксимантов $G_m^{\text{PGA}}(f_0, \mathcal{D})$, $m \geq 0$, и последовательность остатков f_m^{PGA} , $m \geq 0$. Положим $G_0^{\text{PGA}}(f_0, \mathcal{D}) := 0$. Для каждого $m \geq 0$ последовательно выбираем вектор $g_{m+1} \in \mathcal{D}$ такой, что

$$|\langle f_m, g_{m+1} \rangle| = \sup_{g \in \mathcal{D}} |\langle f_m, g \rangle|, \quad (2)$$

и определяем

$$\begin{aligned} G_{m+1}^{\text{PGA}}(f_0, \mathcal{D}) &:= G_m^{\text{PGA}}(f_0, \mathcal{D}) + \langle f_m, g_{m+1} \rangle g_{m+1}, \\ f_{m+1}^{\text{PGA}} &:= f_{m+1} := f - G_{m+1}^{\text{PGA}}(f_0, \mathcal{D}) = f_m - \langle f_m, g_{m+1} \rangle g_{m+1}. \end{aligned} \quad (3)$$

Этот процесс называется *чисто жадным алгоритмом* (ЧЖА).

Отметим, что ЧЖА строит разложение элемента f_0 в ряд по словарю $\mathcal{D}$, уменьшая на каждом шаге норму остатка максимально возможным способом:

$$\|f_{m+1}\| = \inf_{c \in \mathbb{R}, g \in \mathcal{D}} \|f_m - cg\|, \quad m \geq 0.$$

Необходимо сделать несколько замечаний о формуле (2).

1. Если супремум в (2) достигается на нескольких элементах из словаря, то в качестве g_{m+1} можно взять любой из них.

2. Супремум в (2) может не достигаться ни на одном элементе словаря. Тогда, следуя [4], мы должны применить шаг слабого жадного алгоритма (СЖА), выбирая вектор $g_{m+1} \in \mathcal{D}$, удовлетворяющий неравенству

$$|\langle f_m, g_{m+1} \rangle| \geq t \sup_{g \in \mathcal{D}} |\langle f_m, g \rangle|, \quad (4)$$

где $t \in (0, 1)$ – фиксированный “параметр слабости”. Для уменьшения размера выкладок в настоящей работе предполагается, что g_{m+1} на всех шагах алгоритма удовлетворяет равенству (2). Если параметр t в шаге СЖА устремить к единице, то в пределе скорость сходимости СЖА совпадет со скоростью сходимости ЧЖА, которая и будет изучаться.

3. Супремум в (2) формально берется по бесконечному множеству, но в приложениях, как правило, H конечномерно, а $\mathcal{D}$ конечно. В случае бесконечного $\mathcal{D}$ проблема бесконечного перебора может быть решена с помощью перехода к жадным алгоритмам с ограниченной глубиной поиска [5], [6], но это также остается за рамками настоящей работы.

ОПРЕДЕЛЕНИЕ 2. Пусть задан словарь $\mathcal{D}$, и пусть $f_0^{\text{OGA}} := f_0 \in H$ – целевая функция. Положим $G_0(f_0, \mathcal{D}) := 0$. Для каждого $m \geq 0$ последовательно выбираем вектор $g_{m+1} \in \mathcal{D}$ такой, что

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

и определяем

$$G_{m+1}^{\text{OGA}}(f_0, \mathcal{D}) := \text{Proj}_{g_1, \dots, g_{m+1}} f_0, \quad f_{m+1}^{\text{OGA}} := f_{m+1} := f_0 - G_{m+1}^{\text{OGA}}(f_0, \mathcal{D}).$$

Этот процесс называется *ортгональным жадным алгоритмом* (ОЖА).

Л. Джонс [7] доказал сходимость ЧЖА для всех словарей $\mathcal{D}$ и целевых функций f_0 :

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

Для получения оценок скорости сходимости жадного алгоритма необходимо потребовать, чтобы целевая функция f_0 принадлежала некоторому зависящему от словаря классу $\mathcal{A}(\mathcal{D})$. Как было показано в [8], простого ограничения скорости убывания m -членных приближений оказывается недостаточно: для произвольных $\gamma, \delta > 0$ существуют словарь $\mathcal{D}$, функция $f \in H$ и $C > 0$ такие, что для всех $m > 1$ выполняются неравенства

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

Таким образом, для получения нетривиальных оценок сверху скорости сходимости жадных алгоритмов необходимо налагать более жесткие ограничения на класс $\mathcal{A}(\mathcal{D})$. Наиболее популярным в теории жадных алгоритмов является класс $\mathcal{A}_1(\mathcal{D})$, определенный ниже.

Для словаря $\mathcal{D}$ и $M > 0$ рассмотрим класс

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

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

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

Отметим, что шар $\mathcal{A}_1(\mathcal{D}, M)$ является замыканием выпуклой оболочки $\mathcal{D}$.

Для класса $\mathcal{A}_1(\mathcal{D})$ известна скорость убывания m -членных приближений $\sigma_m(f, \mathcal{D})$. Оценка сверху была получена С.Б. Стечкиным [9] в случае ортонормированного словаря и А. Бэрроном [10] для произвольного словаря.

ТЕОРЕМА А [10]. Для произвольного словаря $\mathcal{D}$, элемента $f \in \mathcal{A}_1(\mathcal{D})$ и $m \geq 1$ выполняется оценка

$$\sigma_m(f, \mathcal{D}) \leq |f|_{\mathcal{A}_1(\mathcal{D})} m^{-1/2}. \quad (5)$$

При этом, как следует из результатов, полученных С. Б. Стечкиным [9], показатель $-1/2$ не может быть уменьшен даже для ортонормированного словаря: 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

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

Первые оценки сверху скорости сходимости жадных алгоритмов были получены Р. ДеВором и В. Н. Темляковым.

ТЕОРЕМА В [3]. *Для произвольного словаря $\mathcal{D}$, элемента $f \in \mathcal{A}_1(\mathcal{D})$ и $m \geq 1$ выполняются оценки*

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

$$\|f - G_m^{\text{OGA}}(f, \mathcal{D})\| \leq |f|_{\mathcal{A}_1(\mathcal{D})}m^{-1/2}. \quad (7)$$

Неравенство (7) показывает оптимальность ОЖА в пространстве $\mathcal{A}_1(\mathcal{D})$, для ЧЖА ситуация оказалось более сложной. Оценка (6) была улучшена С. В. Княгининым и В. Н. Темляковым.

ТЕОРЕМА С [11]. *Для произвольного словаря $\mathcal{D}$, элемента $f \in \mathcal{A}_1(\mathcal{D})$ и $m \geq 1$ справедливо*

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

Этот результат, в свою очередь, был улучшен А. В. Сильниченко.

ТЕОРЕМА D [12]. *Существует такое $C > 0$, что для произвольного словаря $\mathcal{D}$, элемента $f \in \mathcal{A}_1(\mathcal{D})$ и $m \geq 1$ имеет место неравенство*

$$\|f - G_m^{\text{PGA}}(f, \mathcal{D})\| \leq C|f|_{\mathcal{A}_1(\mathcal{D})}m^{-0.182}. \quad (8)$$

После получения оценок скорости сходимости для класса $\mathcal{A}_1(\mathcal{D})$ стали изучать вопрос о получении аналогичных оценок классов, являющихся расширениями или сужениями $\mathcal{A}_1(\mathcal{D})$.

Ряд шагов в первом направлении был сделан в работе [13], где были получены оценки скорости сходимости ЧЖА в пространствах $\widehat{A}_\gamma(\mathcal{D})$ (они будут определены в § 2). А. Бэррон, А. Коэн, В. Дамени и Р. ДеВор изучали скорость сходимости ОЖА (а также релаксационного жадного алгоритма) в интерполяционных пространствах $[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$ (см. § 2). В § 2 доказывается лемма, устанавливающая эквивалентность этих пространств, здесь же мы приведем формулировки этих результатов в терминах интерполяционных пространств.

Ортогональный жадный алгоритм, как и в случае класса $\mathcal{A}_1(\mathcal{D})$, оказался оптимальным.

ТЕОРЕМА Е [14]. *Для произвольного θ , $0 < \theta \leq 1$, существует такое $C > 0$, что для произвольного словаря $\mathcal{D}$, элемента $f \in [H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$ и $m \geq 1$ имеет место неравенство*

$$\|f - G_m^{\text{OGA}}(f, \mathcal{D})\| \leq C|f|_{[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}}m^{-\theta/2}.$$

Как известно (см., например, [13]), показатель $\theta/2$ не может быть уменьшен даже в случае ортонормированного словаря $\mathcal{D}$.

В случае малых θ чисто жадный алгоритм также является почти оптимальным.

ТЕОРЕМА 1. *Для произвольных $\theta, 0 < \theta \leq 1/3$, и $\varepsilon, 0 < \varepsilon < \theta/2$, существует такое $C > 0$, что для всех $f \in [H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$ и $m \geq 1$ справедлива оценка*

$$\|f - G_m^{\text{PGA}}(f, \mathcal{D})\| \leq C \|f\|_{[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}} m^{-\theta/2+\varepsilon}. \quad (9)$$

С другой стороны, вопрос о том, можно ли за счет сужения класса рассматриваемых функций добиться существенного прогресса в оценках сверху скорости сходимости ЧЖА и ОЖА, получает отрицательный ответ. Это связано с тем, что при доказательстве нижних оценок скорости сходимости жадных алгоритмов в пространстве $\mathcal{A}_1(\mathcal{D})$ функция, на которой нижняя оценка достигается довольно часто, является конечно линейной комбинацией элементов словаря, т. е. является элементом максимально узкого класса $\mathcal{A}_0(\mathcal{D})$,

$$\mathcal{A}_0(\mathcal{D}) := \left\{ f \in H : f = \sum_{\lambda \in \Lambda} c_\lambda g_\lambda, \ g_\lambda \in \mathcal{D}, \ \#\Lambda < \infty \right\}.$$

В [3] доказывается следующая оценка снизу скорости сходимости ЧЖА.

ТЕОРЕМА F [3]. *Существуют такие словарь $\mathcal{D}$, элемент $f \in \mathcal{A}_0(\mathcal{D})$ и $C > 0$, что для всех $m \geq 1$ выполняется оценка*

$$\|f - G_m^{\text{PGA}}(f, \mathcal{D})\| \geq C m^{-1/2}.$$

В дальнейшем было показано, что скорость сходимости ЧЖА значительно уступает скорости сходимости ОЖА (7). Соответствующая оценка снизу была получена автором и В. Н. Темляковым.

ТЕОРЕМА G [15]. *Существуют такие словарь $\mathcal{D}$, элемент $f \in \mathcal{A}_0(\mathcal{D})$ и $C > 0$, что для произвольного $m \geq 1$ справедлива оценка*

$$\|f - G_m^{\text{PGA}}(f, \mathcal{D})\| > C m^{-0.27}.$$

В настоящей статье развивается техника, используемая в работе [15], и доказывается новая оценка снизу скорости сходимости ЧЖА, которая оказывается весьма близкой к оценке сверху (8).

ТЕОРЕМА 2. *Существуют такие словарь $\mathcal{D}$, элемент $f \in \mathcal{A}_1(\mathcal{D})$ и $C > 0$, что для произвольного $m \geq 1$ выполняется неравенство*

$$\|f - G_m^{\text{PGA}}(f, \mathcal{D})\| = \|f_m^{\text{PGA}}\| > C m^{-0.1898} \|f\|_{\mathcal{A}_0(\mathcal{D})}.$$

Кроме того, доказывается, что оценка (7) скорости сходимости ортогонального жадного алгоритма не улучшаема не только в пространстве $\mathcal{A}_1(\mathcal{D})$, но и в более узком пространстве $\mathcal{A}_0(\mathcal{D})$.

ТЕОРЕМА 3. *Существуют такие словарь $\mathcal{D}$, элемент $f \in \mathcal{A}_0(\mathcal{D})$ и $C > 0$, что для произвольного $m \geq 1$ выполняется неравенство*

$$\|f - G_m^{\text{OGA}}(f, \mathcal{D})\| > Cm^{-1/2}.$$

Статья имеет следующую структуру: в § 2 установлена связь между классами $\hat{\mathcal{A}}_\gamma(\mathcal{D})$ и интерполяционными пространствами $[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$ и доказана теорема 1. В § 3–7 доказан основной результат работы – теорема 2. В § 8 содержится доказательство теоремы 3.

§ 2. Интерполяционные классы

Теорема 1 является прямым следствием теоремы 1 из работы [13] и эквивалентности нормированных пространств $\hat{\mathcal{A}}_\gamma(\mathcal{D})$ и $[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$, которая устанавливается в настоящем параграфе.

Следуя [13], определим классы $\hat{\mathcal{A}}_\gamma(\mathcal{D})$. Пусть заданы $\gamma \geq 0$, $M \geq 0$. Рассмотрим множества

$$\hat{\mathcal{A}}_\gamma(\mathcal{D}, M) = \left\{ f \in H \mid \forall m \geq 1 \exists \text{ конечное } \Lambda, c_\lambda \in \mathbb{R}, g_\lambda \in \mathcal{D}, \lambda \in \Lambda : \right. \\ \left. \left\| f - \sum_{\lambda \in \Lambda} c_\lambda g_\lambda \right\| \leq \frac{M}{m}, \sum_{\lambda \in \Lambda} |c_\lambda| \leq Mm^\gamma \right\}. \quad (10)$$

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

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

Пусть X и Y – действительные нормированные пространства. Напомним (см. [16]) определение интерполяционных пространств $[X, Y]_{\theta, \infty}$, $0 < \theta < 1$. По определению $f \in [X, Y]_{\theta, \infty}$ тогда и только тогда, когда существует такое $C > 0$, что для всех $t > 0$ выполняется

$$K(f, t) \leq Ct^\theta, \quad (11)$$

где

$$K(f, t) = K(f, t, X, Y) := \inf_{h \in Y} \{ \|f - h\|_X + t\|h\|_Y \}.$$

В качестве нормы $|f|_{[X, Y]_{\theta, \infty}}$ берется инфимум по числам C , для которых выполняется неравенство (11).

ЛЕММА 1. *Пусть заданы словарь $\mathcal{D}$ и положительные числа γ и θ , удовлетворяющие равенству*

$$\frac{1}{1 + \gamma} = \theta.$$

Тогда нормированные пространства $\hat{\mathcal{A}}_\gamma(\mathcal{D})$ и $[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$ эквивалентны.

ДОКАЗАТЕЛЬСТВО. Докажем, что для произвольного $M > 0$ найдется такое $C = Mc(\gamma)$, что для произвольной функции f из $\widehat{\mathcal{A}}_\gamma(\mathcal{D}, M)$ и $t > 0$ будет выполняться неравенство (11). Полагая $m = 1$ в определении (10), получим, что

$$\|f\| \leq 2M, \\ K(f, t) = K(f, t, H, \mathcal{A}_1(\mathcal{D})) \leq \|f - 0\| + t|0|_{\mathcal{A}_1(\mathcal{D})} \leq \|f\| \leq 2M.$$

Таким образом, если выбрать $C = Mc(\gamma) \geq 2M$, то неравенство (11) будет выполняться для всех $t \geq 1$. Если $0 < t < 1$, то выберем $m \geq 1$, для которого выполняются неравенства

$$(m+1)^{-1} < t^{\frac{1}{1+\gamma}} \leq m^{-1}. \quad (12)$$

Тогда найдется конечная комбинация $\sum_{\lambda \in \Lambda} c_\lambda g_\lambda$, $g_\lambda \in \mathcal{D}$, удовлетворяющая определению (10). Для $h = \sum_{\lambda \in \Lambda} c_\lambda g_\lambda$ будут выполняться неравенства

$$\|f - h\| \leq Mm^{-1}, \quad |h|_{\mathcal{A}_1(\mathcal{D})} \leq Mm^\gamma.$$

Следовательно, с учетом (12) имеем

$$K(f, t) \leq \|f - h\| + t|h|_{\mathcal{A}_1(\mathcal{D})} \leq Mm^{-1} + tMm^\gamma \\ \leq c_1(\gamma)Mt^{\frac{1}{1+\gamma}} + tMt^{-\frac{\gamma}{1+\gamma}} = Mt^{\frac{1}{1+\gamma}}(c_1(\gamma) + 1) =: Mc(\gamma)t^\theta = Ct^\theta.$$

Обратно, пусть для функции $f \in [H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$ и числа $C > 0$ неравенство (11) выполняется при всех $t > 0$. Тогда для $t = m^{-\frac{1}{1+\gamma}}$ найдется такая конечная комбинация $h = \sum_{\lambda \in \Lambda} c_\lambda g_\lambda$, $g_\lambda \in \mathcal{D}$, что

$$\|f - h\| \leq 2K(f, t) \leq 2Ct^\theta = 2Cm^{-1}, \\ t|h|_{\mathcal{A}_1(\mathcal{D})} \leq 2K(f, t) \leq 2Ct^\theta = 2tCt^{-\frac{\gamma}{1+\gamma}} = 2tCm^\gamma.$$

Следовательно, $f \in \widehat{\mathcal{A}}_\gamma(\mathcal{D}, M)$ для $M = 2C$, что завершает доказательство леммы.

§ 3. Общие формулы

Для доказательства теоремы 2 мы будем следовать технике, разработанной в [15]. Более точно, нам придется воспользоваться большей частью шагов из доказательства теоремы 1.1 работы [15], и лишь на последнем этапе будут сделаны более аккуратные оценки, которые позволят получить более точную итоговую оценку. Для полноты изложения нам придется в § 4, 5 в значительной степени использовать технику из [15, § 2, 3, пп. 3.1, 3.2].

Предположим, что в пространстве H задан ортонормированный базис $\{e_j\}_{j=1}^\infty$. Будем строить две последовательности $\{x_n\}_{n=N}^\infty \subset H$ и $\{g_n\}_{n=N}^\infty \subset H$, где N – достаточно большое число, которое будет определено далее. Пусть элементы последовательности $\{x_n\}$ имеют следующий вид:

$$x_n = \sum_{k=1}^n a_{n,k} e_k, \quad n = N, N+1, \dots, \quad (13)$$

а элементы последовательности $\{g_n\}$ удовлетворяют соотношению

$$g_{n+1} = \gamma_{n+1}x_n + h_{n+1} + \xi_{n+1}e_{n+1}, \quad n = N, N+1, \dots, \quad (14)$$

где последовательности $\{\gamma_k\} \subset \mathbb{R}$, $\{h_k\} \subset H$ и $\{\xi_k\}$, $0 \leq \xi_k \leq 1$, так же, как и элемент $g_N \in H$, будут определены далее. При этом будет гарантироваться выполнение включения $h_{n+1} \in \text{span}\{e_1, \dots, e_n\}$ для всех $n \geq N$, а также равенства

$$\|g_k\| = 1, \quad k = N, N+1, \dots, \quad (15)$$

$$x_{n+1} := x_n - \langle x_n, g_{n+1} \rangle g_{n+1}, \quad n = N, N+1, \dots. \quad (16)$$

В настоящем параграфе последовательности $\{a_{n,k}\}$, $\{\gamma_k\}$, $\{h_k\}$, $\{\xi_k\}$, $0 \leq \xi_k \leq 1$, определены не будут, но будет показано, что при определенных условиях для них выполняется ряд соотношений.

Введем следующее обозначение:

$$q_{n+1} := \langle x_n, g_{n+1} \rangle, \quad n = N, N+1, \dots. \quad (17)$$

Предположения $h_{n+1} \in \text{span}\{e_1, \dots, e_n\}$ и (16) влекут равенство

$$a_{n+1, n+1} = -q_{n+1}\xi_{n+1}. \quad (18)$$

Предположим также, что для последовательностей $\{\gamma_k\}$ и $\{q_k\}$ выполняются следующие равенства:

$$1 - \gamma_{n+1}q_{n+1} = \frac{q_{n+1}}{q_n}, \quad n = N+1, N+2, \dots. \quad (19)$$

Осуществим ряд преобразований, которые будут полезны далее. Согласно соотношениям (14) и (17) из формулы (16) следует, что

$$\begin{aligned} x_{n+1} &= x_n - q_{n+1}g_{n+1} = (1 - \gamma_{n+1}q_{n+1})x_n - q_{n+1}h_{n+1} - q_{n+1}\xi_{n+1}e_{n+1} \\ &= \frac{q_{n+1}}{q_n}x_n - q_{n+1}h_{n+1} - q_{n+1}\xi_{n+1}e_{n+1}. \end{aligned} \quad (20)$$

Из соотношений (13), (14) и (17) вытекает, что

$$q_{n+1} = \langle x_n, g_{n+1} \rangle = \gamma_{n+1}\|x_n\|^2 + \langle x_n, h_{n+1} \rangle. \quad (21)$$

Нашей ближайшей целью является получение выражений для $\langle x_n, g_k \rangle$. Ясно, что из соотношений (15) и (16) следует

$$\begin{aligned} \langle x_n, g_n \rangle &= \langle x_{n-1} - \langle x_{n-1}, g_n \rangle g_n, g_n \rangle = \langle x_{n-1}, g_n \rangle (1 - \|g_n\|^2) = 0, \\ n &= N+1, N+2, \dots \end{aligned} \quad (22)$$

I. Случай $N \leq k < n$. Используя (20) и включение $g_k \in \text{span}\{e_1, \dots, e_k\}$, для $n \geq N+2$ получим

$$\langle x_n, g_k \rangle = \left\langle \frac{q_n}{q_{n-1}}x_{n-1} - q_n h_n, g_k \right\rangle = \frac{q_n}{q_{n-1}} \langle x_{n-1}, g_k \rangle - q_n \langle h_n, g_k \rangle. \quad (23)$$

Повторяя (23) и используя (22), для $k \geq N + 1$ имеем

$$\begin{aligned} \langle x_n, g_k \rangle &= \frac{q_n}{q_{n-1}} \langle x_{n-1}, g_k \rangle - q_n \langle h_n, g_k \rangle \\ &= \frac{q_n}{q_{n-1}} \left(\frac{q_{n-1}}{q_{n-2}} \langle x_{n-2}, g_k \rangle - q_{n-1} \langle h_{n-1}, g_k \rangle \right) - q_n \langle h_n, g_k \rangle = \dots \\ &\dots = \frac{q_n}{q_k} \langle x_k, g_k \rangle - q_n \sum_{l=k+1}^n \langle h_l, g_k \rangle = -q_n \sum_{l=k+1}^n \langle h_l, g_k \rangle. \end{aligned} \quad (24)$$

Равенство (14) задает g_n при $n \geq N + 1$, теперь определим g_N . Положим $h_{N+1} := 0$ и для $\varepsilon = \varepsilon(N)$, $0 < \varepsilon < \|x_{N+1}\|$, положим

$$g_N := \varepsilon x_{N+1} \|x_{N+1}\|^{-2} + \xi g_{N+1}, \quad (25)$$

где $\xi = \xi(\varepsilon)$ гарантирует выполнение равенства $\|g_N\| = 1$. Тогда согласно соотношению (22) получим выполнение равенства

$$\langle x_{N+1}, g_N \rangle = \varepsilon$$

и включения $x_{N+1} \in \text{span}(g_N, g_{N+1})$. Согласно (23), учитывая $g_N \in \text{span}\{e_1, \dots, e_{N+1}\}$, получим аналогично равенству (24) выполнение соотношения

$$\langle x_n, g_N \rangle = \frac{q_n}{q_{N+1}} \langle x_{N+1}, g_N \rangle - q_n \sum_{l=N+2}^n \langle h_l, g_N \rangle = -q_n \left(\sum_{l=N+2}^n \langle h_l, g_N \rangle - \frac{\varepsilon}{q_{N+1}} \right). \quad (26)$$

II. Случай $k \geq n + 1 \geq N + 2$. Согласно (14) для $k \geq n + 1$ выполняется

$$\langle x_n, g_{k+1} \rangle = \langle x_n, \gamma_{k+1} x_k + h_{k+1} \rangle = \gamma_{k+1} \langle x_n, x_k \rangle + \langle x_n, h_{k+1} \rangle. \quad (27)$$

Далее,

$$\langle x_n, x_k \rangle = \langle x_n, x_{k-1} - q_k g_k \rangle = \langle x_n, x_{k-1} \rangle - q_k \langle x_n, g_k \rangle. \quad (28)$$

Также справедливы равенства

$$\begin{aligned} \langle x_n, g_k \rangle &= \langle x_n, \gamma_k x_{k-1} + h_k \rangle = \gamma_k \langle x_n, x_{k-1} \rangle + \langle x_n, h_k \rangle, \\ \langle x_n, x_{k-1} \rangle &= \gamma_k^{-1} (\langle x_n, g_k \rangle - \langle x_n, h_k \rangle). \end{aligned} \quad (29)$$

Комбинируя (28) и (29), из соотношения (27) получаем

$$\langle x_n, g_{k+1} \rangle = (\gamma_k^{-1} - q_k) \gamma_{k+1} \langle x_n, g_k \rangle + \left\langle x_n, h_{k+1} - \frac{\gamma_{k+1}}{\gamma_k} h_k \right\rangle. \quad (30)$$

Используя (19), переписываем (30) в следующем виде:

$$\begin{aligned} \langle x_n, g_{k+1} \rangle &= \frac{1 - \gamma_k q_k}{\gamma_k} \gamma_{k+1} \langle x_n, g_k \rangle + \left\langle x_n, h_{k+1} - \frac{\gamma_{k+1}}{\gamma_k} h_k \right\rangle \\ &= \frac{\gamma_{k+1}}{\gamma_k} \frac{q_k}{q_{k-1}} \langle x_n, g_k \rangle + \left\langle x_n, h_{k+1} - \frac{\gamma_{k+1}}{\gamma_k} h_k \right\rangle. \end{aligned} \quad (31)$$

Заметим, что из (16) и (17) следуют равенства

$$\|x_n\|^2 = \|x_{n+1}\|^2 + \langle x_n, g_{n+1} \rangle^2, \quad q_{n+1}^2 = \|x_n\|^2 - \|x_{n+1}\|^2, \quad n = N, N + 1, \dots \quad (32)$$

§ 4. Конструкция

Пусть заданы числа $N \in \mathbb{N}$, $\varepsilon > 0$ и последовательность положительных чисел $\{v_n\}_{n=1}^\infty$. Положим

$$x_N := -N^{-1/2} v_N \sum_{i=1}^N e_i.$$

С помощью формул (14), (16), а также (25) и числа ε определим элементы последовательностей $\{x_n\}_{n=N+1}^\infty$ и $\{g_n\}_{n=N}^\infty$. Если нам удастся для достаточно большого N подобрать свободные параметры $(\gamma_n, h_n, \varepsilon(N))$ из § 2 так, чтобы выполнялись равенство

$$v_n = \|x_n\|, \quad n \geq N+1, \quad (33)$$

и неравенство

$$|\langle x_n, g_m \rangle| < q_{n+1} = \langle x_n, g_{n+1} \rangle, \quad n \geq N+1, \quad m \geq N, \quad m \neq n+1, \quad (34)$$

то, рассматривая чисто жадный алгоритм для гильбертова пространства $H' = \overline{\text{span}}\{g_n\}_{n=N}^\infty$, словаря $\mathcal{D} = \{g_n\}_{n=N}^\infty$ и начального элемента

$$f_0^{\text{PGA}} := f_0 := x_{N+1} \in \text{span}(g_N, g_{N+1}) \subset \mathcal{A}_0(\mathcal{D})$$

и сравнивая определение ЧЖА с соотношением (16), мы получим

$$f_l^{\text{PGA}} = x_{N+1+l}, \quad l \geq 0,$$

и согласно (33) имеем

$$\|f_l^{\text{PGA}}\| = v_{N+1+l}, \quad l \geq 0. \quad (35)$$

§ 5. Подбор параметров

Введем следующие обозначения. Для последовательности $\{y_n\}$ будем записывать $y_n \simeq bn^\alpha$, $b \neq 0$, если $y_n = bn^\alpha(1 + c/n + O(1/n^2))$, $n \rightarrow \infty$, и $y_n \approx bn^\alpha$, если $\lim_{n \rightarrow \infty} y_n/(bn^\alpha) = 1$.

Пусть задано число β , $0 < \beta < 1/2$. Для $n \geq N$ положим

$$\|x_n\|^2 := v_n^2 := (1 - \beta)^2(1 - 2\beta)^{-1}n^{-1+2\beta}. \quad (36)$$

Тогда согласно (32) имеем

$$\begin{aligned} q_n^2 &:= \|x_{n-1}\|^2 - \|x_n\|^2 \simeq (1 - \beta)^2 n^{-2+2\beta}, & n \geq N+1, \\ q_n &:= (\|x_{n-1}\|^2 - \|x_n\|^2)^{1/2} \simeq (1 - \beta) n^{-1+\beta}, & n \geq N+1. \end{aligned} \quad (37)$$

Из (19) вытекает, что

$$\gamma_n := q_n^{-1} - q_{n-1}^{-1} \simeq n^{-\beta}. \quad (38)$$

Далее, имеем

$$\gamma_{n+1}\|x_n\|^2 - q_{n+1} \simeq \left(\frac{(1-\beta)^2}{1-2\beta} - (1-\beta) \right) n^{-1+\beta} = \beta(1-\beta)(1-2\beta)^{-1} n^{-1+\beta}, \quad (39)$$

$$\frac{\gamma_{n+1}\|x_n\|^2 - q_{n+1}}{q_n} \simeq \frac{\beta}{1-2\beta}. \quad (40)$$

Заметим, что для $y_n \simeq bn^\alpha$ выполняется равенство

$$\frac{y_{n+1}}{y_n} = 1 + \frac{\alpha}{n} + O\left(\frac{1}{n^2}\right).$$

Используя это, можно вывести следующее соотношение:

$$\frac{\gamma_{n+1}}{\gamma_n} \frac{q_n}{q_{n-1}} = \left(1 - \frac{\beta}{n} + o(n^{-1})\right) \left(1 + \frac{-1+\beta}{n} + o(n^{-1})\right) = 1 - (1+o(1))n^{-1}. \quad (41)$$

Положим $h_{N+1} = 0$. Для $n \geq N+2$ определим h_n следующим образом:

$$h_n := \frac{\alpha_n}{n-1} \Phi_n, \quad 0 \leq \alpha_n \leq C(a, \beta), \quad (42)$$

при

$$\Phi_n := \sum_{k=1}^n \phi_n(k) e_k,$$

где

$$\phi_n(k) = \begin{cases} 1, & \text{если } an < k < n-1, \\ 1 - \{an\}, & \text{если } k = [an], \\ 0, & \text{если } 1 \leq k < [an], \end{cases}$$

с некоторым фиксированным параметром a , $0 < a < 1$. Число α_{n+1} выбирается на каждом шаге таким образом, чтобы с учетом определения (14) выполнялось равенство

$$\begin{aligned} q_{n+1} &= \langle x_n, g_{n+1} \rangle = \langle x_n, \gamma_{n+1}x_n + h_{n+1} + \xi_{n+1}e_{n+1} \rangle \\ &= \gamma_{n+1}\|x_n\|^2 + \frac{\alpha_{n+1}}{n} \langle x_n, \Phi_{n+1} \rangle, \quad n \geq N. \end{aligned} \quad (43)$$

Отметим, что

$$\begin{aligned} \|\gamma_{n+1}x_n + h_{n+1}\| &\leq \|\gamma_{n+1}x_n\| + \|h_{n+1}\| = O(n^{-1/2}), \\ \gamma_{n+1}x_n + h_{n+1} &\in \text{span}\{e_1, \dots, e_n\}, \quad \langle \gamma_{n+1}x_n + h_{n+1}, e_{n+1} \rangle = 0. \end{aligned}$$

Тогда из формул (14), (15) и включения $h_{n+1} \in \text{span}\{e_1, \dots, e_n\}$ вытекает равенство

$$\xi_n = (1 - \|\gamma_{n+1}x_n + h_{n+1}\|^2)^{1/2} \approx 1. \quad (44)$$

§ 6. Дальнейшие оценки

Улучшение нижней оценки для скорости сходимости ЧЖА, излагаемое в настоящей работе, основывается на более точной оценке чисел $a_{n,k} = \langle x_n, e_k \rangle$, $N+1 \leq k \leq n$. Имеет место равенство

$$a_{n,k} = -q_n \left(\xi_k + \sum_{l=k+1}^n \langle h_l, e_k \rangle \right). \quad (45)$$

Действительно, при $k = n$ оно является прямым следствием (18). При $N+1 \leq k \leq n$ согласно (18) и (20) имеем

$$\begin{aligned} a_{n,k} = \langle x_n, e_k \rangle &= \left\langle \frac{q_n}{q_{n-1}} x_{n-1} - q_n h_n - q_n \xi_n e_n, e_k \right\rangle \\ &= \frac{q_n}{q_{n-1}} \langle x_{n-1}, e_k \rangle - q_n \langle h_n, e_k \rangle = \dots = \frac{q_n}{q_k} \langle x_k, e_k \rangle - q_n \sum_{l=k+1}^n \langle h_l, e_k \rangle \\ &= -q_n \left(\xi_k + \sum_{l=k+1}^n \langle h_l, e_k \rangle \right). \end{aligned}$$

Используя (42), переписываем (45) в следующем виде:

$$a_{n,k} = -q_n \left(\xi_k + \sum_{l=k+1}^n \langle h_l, e_k \rangle \right) = -q_n \left(\xi_k + \sum_{l=k+1}^n \frac{\alpha_l \phi_l(k)}{l-1} \right). \quad (46)$$

ЛЕММА 2. Если число $\bar{\alpha} > 0$ удовлетворяет соотношениям

$$\bar{\alpha}((1-a) + \bar{\alpha}(1-a+a \ln a)) = \frac{\beta}{1-2\beta}, \quad \bar{\alpha}^2(1-a+a \ln a) \frac{1-2\beta}{\beta} < 1, \quad (47)$$

то последовательность $\{\alpha_n\}$ имеет предел

$$\lim_{n \rightarrow \infty} \alpha_n = \bar{\alpha}. \quad (48)$$

ДОКАЗАТЕЛЬСТВО. Предположим, что

$$\liminf_{n \rightarrow \infty} \alpha_n \geq \alpha_0. \quad (49)$$

Тогда с учетом равенства (44) и определения $\phi_n(k)$ из (46) вытекает, что

$$\begin{aligned} -\frac{a_{n,k}}{q_n} &\geq 1 + \alpha_0 \sum_{l=k}^{\min(n, [k/a]+1)} \frac{1}{l} + o(1), \\ -\frac{a_{n,k}}{q_n} &\geq 1 + \alpha_0 \left(\ln \left(\min \left(n, \left[\frac{k}{a} \right] + 1 \right) \right) - \ln k \right) + o(1), \end{aligned} \quad (50)$$

где $o(1)$ соответствует базе $k, n \rightarrow \infty$. Следовательно,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{-\langle x_n, \Phi_{n+1} \rangle}{nq_n} &= \liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{-a_{n,k}}{nq_n} \phi_k(n) = \liminf_{n \rightarrow \infty} \sum_{k=[an]}^n \frac{-a_{n,k}}{nq_n} \\ &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} \left((n - an) + \alpha_0 \sum_{k=[an]}^n (\ln n - \ln k) \right) \\ &= \liminf_{n \rightarrow \infty} \frac{1}{n} (n - an + \alpha_0((n - an) \ln n - n \ln n + n + an \ln(an) - an)) \\ &= 1 - a + \alpha_0(1 - a + a \ln a). \end{aligned}$$

Используя (40), из равенства (43) выводим

$$\begin{aligned} \limsup_{n \rightarrow \infty} \alpha_n &\leq \limsup_{n \rightarrow \infty} \frac{\gamma_{n+1} \|x_n\|^2 - q_{n+1}}{\frac{-\langle x_n, \Phi_{n+1} \rangle}{n}} \\ &= \frac{\lim_{n \rightarrow \infty} \frac{\gamma_{n+1} \|x_n\|^2 - q_{n+1}}{q_n}}{\liminf_{n \rightarrow \infty} \frac{-\langle x_n, \Phi_{n+1} \rangle}{nq_n}} \leq \frac{\beta}{1 - 2\beta} (1 - a + \alpha_0(1 - a + a \ln a))^{-1}. \end{aligned} \quad (51)$$

Аналогично можно доказать, что из неравенства $\limsup_{n \rightarrow \infty} \alpha_n \leq \alpha_0$ вытекает оценка

$$\liminf_{n \rightarrow \infty} \alpha_n \geq \frac{\beta}{1 - 2\beta} (1 - a + \alpha_0(1 - a + a \ln a))^{-1}. \quad (52)$$

С помощью неравенства $\gamma_{n+1} \|x_n\|^2 - q_{n+1} > 0$ (см. (39)) и (43) легко по индукции доказать, что $\alpha_n \geq 0$ и $a_{n,k} \leq 0$ для $1 \leq k \leq n$. Тогда из (51) и (52) вытекает, что для функции

$$F: \alpha \rightarrow \frac{\beta}{1 - 2\beta} (1 - a + \alpha(1 - a + a \ln a))^{-1}$$

будут выполняться следующие неравенства:

$$\begin{aligned} \liminf_{n \rightarrow \infty} \alpha_n &\geq 0, & \limsup_{n \rightarrow \infty} \alpha_n &\leq F(0), & \liminf_{n \rightarrow \infty} \alpha_n &\geq F(F(0)), \\ \limsup_{n \rightarrow \infty} \alpha_n &\leq F(F(F(0))), & \liminf_{n \rightarrow \infty} \alpha_n &\geq F(F(F(F(0)))), & \dots \end{aligned}$$

Таким образом, для доказательства леммы 2 достаточно проверить, что

$$\lim_{n \rightarrow \infty} \underbrace{F(\dots(F(0)))}_n = \bar{\alpha}. \quad (53)$$

Для проверки (53) мы воспользуемся стандартными фактами о “притягивающих неподвижных точках”. В силу того, что функция F является гиперболой с положительными коэффициентами $F(\alpha) = \frac{w_1}{w_2 + w_3 \alpha}$, $w_1, w_2, w_3 > 0$, то равенство (53) для $\bar{\alpha} \geq 0$ будет вытекать из соотношений

$$F(\bar{\alpha}) = \bar{\alpha}, \quad |F'(\bar{\alpha})| < 1,$$

которые в силу равенства

$$F'(t) = \left(\frac{w_1}{w_2 + w_3 t} \right)' = -\frac{w_3}{w_1} \left(\frac{w_1}{w_2 + w_3 t} \right)^2$$

эквивалентны соотношениям

$$F(\bar{\alpha}) = \bar{\alpha}, \quad \left| \frac{w_3}{w_1} \right| \bar{\alpha}^2 < 1,$$

которые, в свою очередь, эквивалентны (47). Лемма доказана.

С помощью леммы 2 и формулы (46) легко установить (аналогично выводу формулы (50)) следующее равенство:

$$-\frac{a_{n,k}}{q_n} = 1 + \bar{\alpha} \left(\ln \left(\min \left(n, \left\lceil \frac{k}{a} \right\rceil + 1 \right) \right) - \ln k \right) + o(1), \quad (54)$$

где $o(1)$ соответствует базе $k, n \rightarrow \infty$.

СЛЕДСТВИЕ 1. Если число $\bar{\alpha} > 0$ удовлетворяет условиям (47) леммы 2, то выполнено

$$\frac{\alpha_{n+1}}{\alpha_n} = 1 + o\left(\frac{1}{n}\right). \quad (55)$$

ДОКАЗАТЕЛЬСТВО. Из равенства (43) вытекает, что

$$\frac{\alpha_{n+1}}{\alpha_n} = \frac{q_{n+1} - \gamma_{n+1} \|x_n\|^2}{q_n - \gamma_n \|x_{n-1}\|^2} \frac{n}{n-1} \left(\frac{\langle x_n, \Phi_{n+1} \rangle}{\langle x_{n-1}, \Phi_n \rangle} \right)^{-1}. \quad (56)$$

Применяя равенство (20), записываем

$$\begin{aligned} \frac{\langle x_n, \Phi_{n+1} \rangle}{\langle x_{n-1}, \Phi_n \rangle} &= \frac{\langle x_n, \Phi_n \rangle}{\langle x_{n-1}, \Phi_n \rangle} + \frac{\langle x_n, \Phi_{n+1} - \Phi_n \rangle}{\langle x_{n-1}, \Phi_n \rangle} \\ &= \frac{\langle \frac{q_n}{q_{n-1}} x_{n-1} - q_n h_n, \Phi_n \rangle}{\langle x_{n-1}, \Phi_n \rangle} + \frac{\langle x_n, \Phi_{n+1} - \Phi_n \rangle}{\langle x_{n-1}, \Phi_n \rangle} \\ &= \frac{q_n}{q_{n-1}} - \frac{q_n \alpha_n \langle \Phi_n, \Phi_n \rangle}{(n-1) \langle x_{n-1}, \Phi_n \rangle} + \frac{\langle x_n, \Phi_{n+1} - \Phi_n \rangle}{\langle x_{n-1}, \Phi_n \rangle}. \end{aligned} \quad (57)$$

Учитывая существование предела (48) (лемма 2), из двух последних равенств вытекает, что имеет место следующее представление:

$$\frac{\alpha_{n+1}}{\alpha_n} = 1 + \frac{d(\alpha, \beta, a)}{n} + o\left(\frac{1}{n}\right).$$

Для доказательства следствия достаточно проверить, что $d(\alpha, \beta, a) = 0$. Этот результат может быть выведен из формул (56) и (57) прямым вычислением, но мы докажем его неявно: если бы $d(\alpha, \beta, a) \neq 0$, то бесконечное произведение $\prod_{n=N+1}^{\infty} \frac{\alpha_{n+1}}{\alpha_n}$ расходилось бы, но согласно соотношению (48) мы имеем

$$\prod_{n=N+1}^{\infty} \frac{\alpha_{n+1}}{\alpha_n} = \frac{\bar{\alpha}}{\alpha_{N+1}}.$$

Следствие доказано.

Пусть α , β и a удовлетворяют условиям (47). Найдем достаточные условия для этих параметров, которые обеспечат выполнение неравенства (34). Рассмотрим сначала случай $m > n + 1$. Будем доказывать (34) индукцией по $m \geq n + 1$. По определению $|\langle x_n, g_{n+1} \rangle| = q_{n+1}$, и для проверки шага индукции достаточно установить, что из неравенства $|\langle x_n, g_m \rangle| \leq q_{n+1}$ будет вытекать неравенство $|\langle x_n, g_{m+1} \rangle| < q_{n+1}$ для $m \geq n + 1$. Воспользуемся соотношениями (31) и (41):

$$\begin{aligned} |\langle x_n, g_{m+1} \rangle| &\leq \left| \frac{\gamma_{m+1}}{\gamma_m} \frac{q_m}{q_{m-1}} \langle x_n, g_m \rangle \right| + \left| \left\langle x_n, h_{m+1} - \frac{\gamma_{m+1}}{\gamma_m} h_m \right\rangle \right| \\ &\leq (1 - (1 + o(1))m^{-1}) |\langle x_n, g_m \rangle| + \left| \left\langle x_n, h_{m+1} - \frac{\gamma_{m+1}}{\gamma_m} h_m \right\rangle \right| \\ &\leq (1 - (1 + o(1))m^{-1}) q_{n+1} + \left| \left\langle x_n, h_{m+1} - \frac{\gamma_{m+1}}{\gamma_m} h_m \right\rangle \right|. \end{aligned}$$

Таким образом, для проверки шага индукции и доказательства (34) в случае $m > n + 1$ достаточно для $m \geq n + 1$ проверить выполнение неравенства

$$I_{n,m}^1 := \left| \left\langle x_n, h_{m+1} - \frac{\gamma_{m+1}}{\gamma_m} h_m \right\rangle \right| \leq (1 - \delta) m^{-1} q_{n+1} \quad (58)$$

с некоторым $\delta > 0$. Обозначим

$$p_m = h_{m+1} - \frac{\gamma_{m+1}}{\gamma_m} h_m.$$

Ясно, что для $k < [am]$ будет выполняться равенство $\langle p_m, e_k \rangle = 0$. Из (42) и леммы 2 следует, что

$$\langle p_m, e_{[am]} \rangle + \langle p_m, e_{[am]+1} \rangle \approx -\frac{a\bar{\alpha}}{m}.$$

Далее, из соотношения (38) и следствия 1 вытекает, что

$$\begin{aligned} \langle p_m, e_k \rangle &\approx \frac{\alpha_{m+1}}{m} - \left(\frac{m+1}{m} \right)^{-\beta} \frac{\alpha_m}{m-1} \\ &= \frac{\alpha_m}{m} \left(1 + o\left(\frac{1}{m} \right) \right) - \left(1 - \frac{\beta}{m} + o\left(\frac{1}{m} \right) \right) \frac{\alpha_m}{m} \left(1 + \frac{1}{m} + o\left(\frac{1}{m} \right) \right) \\ &= \frac{\alpha_m}{m} \left(\frac{\beta-1}{m} + o\left(\frac{1}{m} \right) \right) \approx \frac{\beta\bar{\alpha} - \bar{\alpha}}{m^2}, \quad [am] + 1 < k \leq m. \end{aligned}$$

Обозначим для действительных $x \geq y > 0$

$$S_{x,y} := x - ay \ln x - ay + ay \ln(ay),$$

тогда будет выполняться равенство

$$S_{y,y} := y - ay \ln y - ay + ay \ln(ay) = y(1 - a + a \ln a).$$

Заметим, что

$$\begin{aligned} \sum_{k=[am]+1}^n \ln n - \ln k &= (n - am) \ln n - n \ln n + n + am \ln(am) - am + o(n) \\ &= n - am \ln n - am + am \ln(am) + o(n) = S_{n,m} + o(n). \end{aligned} \quad (59)$$

Применяя (54) и используя неравенство $\beta < 1/2$, получаем

$$\begin{aligned} I_{n,m}^1 &= \left| \sum_{k=1}^n \langle x_n, e_k \rangle \langle p_m, e_k \rangle \right| = \left| |a_{n,[am]+1}| \frac{a\bar{\alpha}}{m} + \sum_{k=[am]+1}^n |a_{n,k}| \frac{|\beta\bar{\alpha} - \bar{\alpha}|}{m^2} \right| + o\left(\frac{q_n}{m}\right) \\ &= q_n \left| \frac{a\bar{\alpha}}{m} (1 + \bar{\alpha}(\ln n - \ln(am))) + \frac{\bar{\alpha} - \beta\bar{\alpha}}{m^2} \sum_{k=[am]+1}^n (1 + \bar{\alpha}(\ln n - \ln k)) \right| + o\left(\frac{q_n}{m}\right) \\ &= \frac{q_n}{m} \left| a\bar{\alpha}(1 + \bar{\alpha}(\ln n - \ln(am))) + \frac{\bar{\alpha} - \beta\bar{\alpha}}{m} (n - am + \bar{\alpha}S_{n,m}) \right| + o\left(\frac{q_n}{m}\right) \\ &\leq \frac{q_n}{m} \left| a\bar{\alpha}(1 + \bar{\alpha}(\ln m - \ln(am))) + \frac{\bar{\alpha} - \beta\bar{\alpha}}{m} (m - am + \bar{\alpha}S_{m,m}) \right| + o\left(\frac{q_n}{m}\right) \\ &= \frac{q_n}{m} \left| a\bar{\alpha}(1 - \bar{\alpha} \ln a) + (\bar{\alpha} - \beta\bar{\alpha})(1 - a + \bar{\alpha}(1 - a + a \ln a)) \right| + o\left(\frac{q_n}{m}\right). \end{aligned}$$

Тогда согласно (58) для справедливости оценки (34) в случае $m > n$ достаточно выполнения неравенства

$$\left| a\bar{\alpha}(1 - \bar{\alpha} \ln a) + (\bar{\alpha} - \beta\bar{\alpha})(1 - a + \bar{\alpha}(1 - a + a \ln a)) \right| < 1. \quad (60)$$

Перейдем к доказательству (34) в случае $m \leq n$. Из равенств (24), (26) и произвольности выбора ε (сначала фиксируется N , а по нему подбирается ε) следует, что достаточно проверить выполнение неравенства

$$I_{k,n}^2 = \left| \sum_{l=k+1}^n \langle h_l, g_k \rangle \right| < 1 - \tau \quad (61)$$

для всех $N + 1 \leq k < n$ и некоторого $\tau > 0$.

Используя определение (14), соотношение (48) и асимптотику (54), получим равенства

$$\begin{aligned} \langle g_k, e_k \rangle &= 1 + o(1), \\ \langle g_k, e_p \rangle &= \gamma_k(-q_k)(1 + \bar{\alpha}(\ln k - \ln p)) + \frac{\bar{\alpha}}{k} + o(k^{-1}), \quad [ak] + 1 < p < k. \end{aligned}$$

Если $l \geq k + 1$ таково, что $[al] + 1 < k$, то, используя (59), а также соотношение $\gamma_k q_k \simeq (1 - \beta)k^{-1}$, вытекающее из (37) и (38), получим

$$\begin{aligned} \langle h_l, g_k \rangle &= \sum_{p=1}^{\infty} \langle h_l, e_p \rangle \langle g_k, e_p \rangle = \sum_{p=[al]}^k \langle h_l, e_p \rangle \langle g_k, e_p \rangle \\ &= \sum_{p=[al]+2}^{k-1} \langle h_l, e_p \rangle \langle g_k, e_p \rangle + \langle h_l, e_k \rangle \langle g_k, e_k \rangle + o(k^{-1}) \end{aligned}$$

$$\begin{aligned}
&= \frac{\bar{\alpha}}{l} \sum_{p=[al]+2}^{k-1} \left(\gamma_k(-q_k)(1 + \bar{\alpha}(\ln k - \ln p)) + \frac{\bar{\alpha}}{k} \right) + \frac{\bar{\alpha}}{l} + o(k^{-1}) \\
&= \frac{\bar{\alpha}}{l} \left(1 + \sum_{p=[al]+2}^{k-1} \left(\frac{\bar{\alpha}}{k} - \gamma_k g_k \right) - \gamma_k g_k \bar{\alpha} \sum_{p=[al]+1}^k (\ln k - \ln p) \right) + o(k^{-1}) \\
&= \frac{\bar{\alpha}}{l} \left(1 + (k - al) \left(\frac{\bar{\alpha}}{k} - \gamma_k g_k \right) - \gamma_k g_k \bar{\alpha} S_{k,l} \right) + o(k^{-1}) \\
&= \frac{\bar{\alpha}}{l} \left(1 + (k - al) \left(\frac{\bar{\alpha}}{k} - \gamma_k g_k \right) - \gamma_k g_k \bar{\alpha} (k - al \ln k - al + al \ln(al)) \right) + o(k^{-1}) \\
&= \frac{\bar{\alpha}}{l} \left(1 + k \left(\frac{\bar{\alpha}}{k} - \gamma_k g_k \right) - \gamma_k g_k \bar{\alpha} k \right) + \bar{\alpha} \left(-a \left(\frac{\bar{\alpha}}{k} - \gamma_k g_k \right) - \gamma_k g_k \bar{\alpha} (-a) \right) \\
&\quad + \bar{\alpha}^2 \gamma_k g_k a (\ln k - \ln(al)) + o(k^{-1}) \\
&= \frac{\bar{\alpha}}{l} (1 + (\bar{\alpha} - 1 + \beta) - (1 - \beta)\bar{\alpha}) + \bar{\alpha} \left(a \left(\frac{1 - \beta}{k} - \frac{\bar{\alpha}}{k} \right) + \frac{1 - \beta}{k} \bar{\alpha} a \right) \\
&\quad + \frac{\bar{\alpha}^2 a (1 - \beta)}{k} (\ln k - \ln a - \ln l) + o(k^{-1}) \\
&= \frac{\bar{\alpha}}{l} (\beta + \beta \bar{\alpha}) + \frac{\bar{\alpha} a}{k} (1 - \beta - \bar{\alpha} \beta) + \frac{\bar{\alpha}^2 a (1 - \beta)}{k} \left(\ln \frac{k}{a} - \ln l \right) + o(k^{-1}). \tag{62}
\end{aligned}$$

Заметим, что при $l \geq (k+1)/a$ выполняется равенство $\langle h_l, g_k \rangle = 0$, а при $l < (k-1)/a$ выполняется неравенство $[al] + 1 < k$. Кроме того, легко видеть, что для любой фиксированной константы $C = C(a) > 0$ и всех l , $|l - k/a| < C$, выполняется равенство

$$\langle h_l, g_k \rangle = o(1).$$

Имеем

$$I_{k,n}^2 = \left| \sum_{l=k+1}^n \langle h_l, g_k \rangle \right| = \left| \sum_{l=k+1}^{[(k+1)/a]+1} \langle h_l, g_k \rangle \right| = \left| \sum_{l=k+1}^{[(k-1)/a]-1} \langle h_l, g_k \rangle \right| + o(1).$$

Оценим отдельно суммы составляющих $\langle h_l, g_k \rangle$:

$$\begin{aligned}
&\sum_{l=k+1}^{[(k-1)/a]-1} \frac{\bar{\alpha}}{l} (\beta + \beta \bar{\alpha}) = \bar{\alpha} (-\ln a) (\beta + \beta \bar{\alpha}) + o(1), \\
&\sum_{l=k+1}^{[(k-1)/a]-1} \frac{\bar{\alpha} a}{k} (1 - \beta - \bar{\alpha} \beta) = \left(\frac{k}{a} - k \right) \frac{\bar{\alpha} a}{k} (1 - \beta - \bar{\alpha} \beta) + o(1) \\
&\quad = (1 - a) \bar{\alpha} (1 - \beta - \bar{\alpha} \beta) + o(1), \\
&\sum_{l=k+1}^{[(k-1)/a]-1} \frac{\bar{\alpha}^2 a (1 - \beta)}{k} \left(\ln \frac{k}{a} - \ln l \right) = \frac{\bar{\alpha} a (1 - \beta)}{k} S_{k/a, k/a} + o(1) \\
&\quad = \frac{\bar{\alpha}^2 a (1 - \beta)}{k} \frac{k}{a} (1 - a + a \ln a) + o(1) = \bar{\alpha}^2 (1 - \beta) (1 - a + a \ln a) + o(1).
\end{aligned}$$

Объединяя последние три равенства и используя (62), получим

$$I_{k,n}^2 = |\bar{\alpha}(-\ln a)(\beta + \beta \bar{\alpha}) + (1-a)\bar{\alpha}(1-\beta - \bar{\alpha}\beta) + \bar{\alpha}^2(1-\beta)(1-a+a\ln a)| + o(1).$$

Таким образом, согласно (61) для справедливости неравенства (34) в случае $m \leq n$ достаточно выполнения неравенства

$$|\bar{\alpha}(-\ln a)(\beta + \beta \bar{\alpha}) + (1-a)\bar{\alpha}(1-\beta - \bar{\alpha}\beta) + \bar{\alpha}^2(1-\beta)(1-a+a\ln a)| < 1. \quad (63)$$

§ 7. Окончание доказательства теоремы 2

В § 6 было показано, что для выполнения неравенства (34) достаточно, чтобы β , a и $\bar{\alpha}$ удовлетворяли соотношениям (47), (60) и (63), а именно

$$\bar{\alpha}((1-a) + \bar{\alpha}(1-a+a\ln a)) = \frac{\beta}{1-2\beta}, \quad \alpha^2(1-a+a\ln a)\frac{1-2\beta}{\beta} < 1,$$

$$|a\bar{\alpha}(1-\bar{\alpha}\ln a) + (\bar{\alpha}-\beta\bar{\alpha})(1-a+\bar{\alpha}(1-a+a\ln a))| < 1,$$

$$|\bar{\alpha}(-\ln a)(\beta + \beta \bar{\alpha}) + (1-a)\bar{\alpha}(1-\beta - \bar{\alpha}\beta) + \bar{\alpha}^2(1-\beta)(1-a+a\ln a)| < 1.$$

Тогда согласно (35) и (36) для конструкции, приведенной в § 3, найдется $C > 0$ такое, что для всех $n \geq 0$ выполняется неравенство

$$\|f_n^{\text{PGA}}\| \geq Cn^{-1/2+\beta}.$$

Прямое вычисление показывает, что для

$$\beta = 0.3102, \quad a = 0.2693,$$

$$\bar{\alpha} = \frac{-(1-a) + ((1-a)^2 + 4(1-a+a\ln a)\frac{\beta}{1-2\beta})^{1/2}}{2(1-a+a\ln a)} = 0.793305\dots$$

приведенные выше соотношения ((47), (60) и (63)) будут выполняться. Это завершает доказательство теоремы 1.

§ 8. Доказательство теоремы 3

Рассмотрим ортонормированный базис $\mathcal{D} = \{e_n\}_{n=0}^\infty$ в пространстве H . Фиксируя c , $0 < c < 1/2$, определим элемент g_c по формуле

$$g_c = \xi_c e_0 + c \sum_{n=1}^{\infty} (n+1)^{-1} e_n,$$

где ξ_c подбирается таким образом, чтобы выполнялось равенство $\|g_c\| = 1$. Докажем, что для словаря $\mathcal{D}_c = \mathcal{D} \cup \{g_c\}$ и целевой функции $f_0 = g_c - \xi_c e_0 \in \mathcal{A}_1(\mathcal{D}_c)$ будет выполняться неравенство

$$\|f_m\| = \|f_0 - G_m^{\text{OGA}}(f_0, \mathcal{D})\| > c(m+2)^{-1/2}. \quad (64)$$

Для этого достаточно проверить, что при проведении ортогонального жадного алгоритма для всех $m \geq 1$ будет выполняться равенство $g_m = e_m$, а остаток будет удовлетворять равенству

$$f_m = c \sum_{n=m+1}^{\infty} (n+1)^{-1} e_n. \quad (65)$$

Проведем доказательство по индукции. В силу того, что f_0 удовлетворяет равенству (65) для $m = 0$, нам достаточно проверить переход от m к $m+1$. Вычислим скалярное произведение f_m со всеми элементами словаря:

$$\begin{aligned} \langle f_m, e_n \rangle &= 0, \quad 0 \leq n \leq m, \quad \langle f_m, e_n \rangle = c(n+1)^{-1}, \quad n \geq m+1, \\ \langle f_m, g_c \rangle &= \left\langle c \sum_{n=m+1}^{\infty} (n+1)^{-1} e_n, \xi_c e_0 + c \sum_{n=1}^{\infty} (n+1)^{-1} e_n \right\rangle = c^2 \sum_{n=m+1}^{\infty} (n+1)^{-2} \\ &\leq c^2 \sum_{n=m+1}^{\infty} n^{-1} (n+1)^{-1} \leq c^2 (m+1)^{-1} \leq 2c^2 (m+2)^{-1} < c(m+2)^{-1}. \end{aligned}$$

Отсюда по определению ОЖА следует, что $g_{m+1} = e_{m+1}$. Далее,

$$f_{m+1} := f_0 - \text{Proj}_{g_1, \dots, g_{m+1}} f_0 = f_0 - \text{Proj}_{e_1, \dots, e_{m+1}} f_0 = c \sum_{n=m+2}^{\infty} (n+1)^{-1} e_n,$$

что доказывает справедливость (65) для всех $m \geq 1$. Следовательно,

$$\|f_m\| = c^2 \sum_{n=m+1}^{\infty} (n+1)^{-2} \geq c^2 \sum_{n=m+1}^{\infty} (n+1)^{-1} (n+2)^{-1} = c^2 (m+2)^{-1},$$

что завершает доказательство (64) и теоремы 3.

Список литературы

1. В. В. Галатенко, Е. Д. Лившиц, “Обобщенные приближенные слабые жадные алгоритмы”, *Матем. заметки*, **78**:2 (2005), 186–201; англ. пер.: V. V. Galatenko, E. D. Livshits, “Generalized approximate weak greedy algorithms”, *Math. Notes*, **78**:1–2 (2005), 170–184.
2. V. N. Temlyakov, “Nonlinear methods of approximation”, *Found. Comput. Math.*, **3**:1 (2003), 33–107.
3. R. A. DeVore, V. N. Temlyakov, “Some remarks on greedy algorithms”, *Adv. Comput. Math.*, **5**:1 (1996), 173–187.
4. V. N. Temlyakov, “Weak greedy algorithms”, *Adv. Comput. Math.*, **12**:2–3 (2000), 213–227.
5. 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.
6. В. Н. Темляков, “О гриди-алгоритмах с ограниченной глубиной поиска”, *Исследования по теории функций и дифференциальным уравнениям*, Тр. МИАН, **248**, Наука, М., 2005, 262–274; англ. пер.: V. N. Temlyakov, “Greedy algorithms with restricted depth search”, *Proc. Steklov Inst. Math.*, **248**:1 (2005), 255–267.

7. 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.
8. 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.
9. С. Б. Стечкин, “Об абсолютной сходимости ортогональных рядов”, *Докл. АН СССР*, **102**:1 (1955), 37–40.
10. A. R. Barron, “Universal approximation bounds for superpositions of a sigmoidal function”, *IEEE Trans. Inform. Theory*, **39**:3 (1993), 930–945.
11. S. V. Konyagin, V. N. Temlyakov, “Rate of convergence of pure greedy algorithm”, *East J. Approx.*, **5**:4 (1999), 493–499.
12. А. 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.
13. Е. Д. Лившиц, “Об оптимальности жадного алгоритма для некоторых классов функций”, *Матем. сб.*, **198**:5 (2007), 95–114; англ. пер.: E. D. Livshits, “Optimality of the greedy algorithm for some function classes”, *Sb. Math.*, **198**:5 (2007), 691–709.
14. A. R. Barron, A. Cohen, W. Dahmen, R. A. DeVore, “Approximation and learning by greedy algorithms”, *Ann. Statist.*, **36**:1 (2008), 64–94.
15. E. D. Livshitz, V. N. Temlyakov, “Two lower estimates in greedy approximation”, *Constr. Approx.*, **19**:4 (2003), 509–523.
16. Й. Берг, Й. Лефстрем, *Интерполяционные пространства. Введение*, Мир, М., 1980; пер. с англ.: J. Bergh, J. Löfström, *Interpolation spaces. An introduction*, Grundlehren der Mathematischen Wissenschaften, **223**, Springer-Verlag, Berlin–Heidelberg–New York, 1976.

Е. Д. Лившиц (E. D. LIVSHITS)
Механико-математический факультет
Московского государственного университета
им. М. В. Ломоносова
E-mail: livshitz@rambler.ru

Поступило в редакцию
02.11.2007