Math-Net.Ru

Общероссийский математический портал

Е. Д. Лившиц, Об n -членном приближении с неотрицательными коэффициентами, *Матем. заметки*, 2007, том 82, выпуск 3, 373–382

DOI: <https://doi.org/10.4213/mzm3852>

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

Параметры загрузки:

IP: 54.80.97.219

26 апреля 2023 г., 12:55:06

Об n -членном приближении с неотрицательными коэффициентами

Е. Д. Лившиц

В работе рассматриваются алгоритмы построения n -членных приближений с неотрицательными коэффициентами. Доказывается теорема сходимости для “положительного” аналога чисто жадного алгоритма. Получено достаточное условие на последовательность коэффициентов слабости для сходимости положительного слабого жадного алгоритма, которое на классе монотонных последовательностей является необходимым.

Библиография: 12 названий.

Введение и основные результаты. В последние годы во многих приложениях возникла потребность в построении n -членных линейных комбинаций элементов фиксированной переполненной системы, хорошо приближающих заданный элемент гильбертова пространства. Это привело к развитию раздела теории приближения, изучающего жадные алгоритмы. Напомним некоторые понятия и определения из этой области.

Пусть H – действительное сепарабельное гильбертово пространство со скалярным произведением $\langle \cdot, \cdot \rangle$ и нормой $\| \cdot \|$. Фиксируем множество $\mathcal{D}$, $\mathcal{D} \subset H$, состоящее из векторов с единичной нормой, замыкание 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 накладывается, словари могут быть как базисами, так и переполненными системами.

Жадные алгоритмы определяют для заданной целевой функции $f \in H$ разложение

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

частичные суммы которого

$$\sum_{k=1}^n c_k g_k \quad (2)$$

достаточно хорошо приближают функцию f , т.е. являются n -членным приближением f (вообще говоря не наилучшим).

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты №№ 05-01-000620, 06-01-00160).

Существует довольно много видов жадных алгоритмов (см. [1]), здесь мы определим два наиболее известных – чисто жадный алгоритм и слабо жадный алгоритм.

СЛАБО ЖАДНЫЙ АЛГОРИТМ (СЖА). Пусть фиксирована последовательность коэффициентов слабости $\tau = \{t_m\}_{m=1}^{\infty} \subset [0, 1]$. Для целевой функции $f_0 \in H$ и словаря $\mathcal{D}$ положим $G_0^{\text{WGA}}(f_0, \mathcal{D}, \tau) := 0$. Для каждого $m \geq 0$ последовательно выбираем вектор $g_{m+1} \in \mathcal{D}$ такой, что

$$|\langle f_m, g_{m+1} \rangle| \geq t_{m+1} \sup_{g \in \mathcal{D}} |\langle f_m, g \rangle|, \quad (3)$$

и определяем

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

Если все $t_m = 1$, $m \in \mathbb{N}$, слабый жадный алгоритм превращается в чисто жадный алгоритм (ЧЖА). Отметим, что при этом необходимо предполагать, что супремум в формуле (3) достигается на каком-либо элементе словаря (что и делается в большинстве работах), в противном случае алгоритм реализован быть не может.

Джонс [2] доказал сходимость ЧЖА и СЖА с параметрами $t_m = t$, $m \geq 1$, $0 < t \leq 1$, для всех словарей $\mathcal{D}$, целевых функций f_0 :

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

Темляков [3] доказал, что достаточным условием на коэффициенты слабости для сходимости всех возможных реализаций СЖА для всех словарей и всех целевых функций является расходимость ряда $\sum_{m=1}^{\infty} t_m/m$. В работе [4] доказывается, что это условие является необходимым в классе монотонных последовательностей коэффициентов слабости. В статье [5] приводится необходимое и достаточное условие для сходимости СЖА для произвольных последовательностей коэффициентов слабости.

Во многих приложениях требуется, чтобы коэффициенты c_k в n -членном приближении (2) были неотрицательными. В настоящей работе мы исследуем возможность применения жадных алгоритмов для получения подобных разложений.

Ясно, что при наложении дополнительного условия (неотрицательность коэффициентов) на n -членное приближение, мы должны также наложить дополнительное условие на словарь $\mathcal{D}$, чтобы построение соответствующего приближения было возможным. Множество $\mathcal{D} \subset H$ будем называть *положительным словарем*, если

$$g \in \mathcal{D} \Rightarrow \|g\| = 1, \quad \overline{\left\{ \sum_{\lambda \in \Lambda} c_{\lambda} g_{\lambda} \mid c_{\lambda} > 0, g_{\lambda} \in \mathcal{D}, \#\Lambda < \infty \right\}} = H. \quad (4)$$

Бэрроном [6] был предложен способ построения неотрицательных n -членных приближений; в последние годы его подход развивался и обобщался многими авторами [7]–[9]. Приведенное ниже определение соответствует работе [7].

РЕЛАКСАЦИОННЫЙ ЖАДНЫЙ АЛГОРИТМ (РЖА). Пусть заданы целевая функция $f_0 \in H$ и словарь $\mathcal{D}$. Положим $G_0^{\text{RGA}}(f_0, \mathcal{D}) := 0$. Для каждого $m \geq 0$ последовательно выбираем вектор $g_{m+1} \in \mathcal{D}$ и $c_{m+1} \geq 0$ такие, что

$$\left\| f_0 - \frac{m}{m+1} G_m^{\text{RGA}}(f_0, \mathcal{D}) - c_{m+1} g_{m+1} \right\| = \inf_{c \geq 0, g \in \mathcal{D}} \left\| f_0 - \frac{m}{m+1} G_m^{\text{RGA}}(f_0, \mathcal{D}) - c g \right\|, \quad (5)$$

и определяем

$$G_{m+1}^{\text{RGA}}(f_0, \mathcal{D}) := \frac{m}{m+1} G_m^{\text{RGA}}(f_0, \mathcal{D}) + c_{m+1} g_{m+1}.$$

Из результатов работы [7] вытекает, что РЖА сходится для произвольного положительного словаря $\mathcal{D}$ и целевой функции f_0 : $\lim_{m \rightarrow \infty} \|f_0 - G_m^{\text{RGA}}(f_0, \mathcal{D})\| = 0$. РЖА также обладает хорошей скоростью сходимости, оптимальной для некоторых классов функций. С другой стороны, РЖА дает только n -членные приближения (2), но не дает разложения (1); кроме того, последовательность $\{\|f_0 - G_m^{\text{RGA}}(f_0, \mathcal{D})\|\}$ не обязана монотонно убывать. В этой работе мы рассматриваем “положительные аналоги” чистого жадного и слабо жадного алгоритмов, которые обеспечивают построение разложений (1), т.е. обладают свойством оперативности (on-line свойством), и частичные суммы которых (2) монотонно (в смысле нормы остатка) стремятся к целевой функции. (При этом с точки зрения скорости сходимости они могут уступать РЖА).

Перед определением “положительных” версий ЧЖА и СЖА (ПЧЖА и ПСЖА) введем некоторые обозначения. Предположим, что с помощью ПЧЖА и ПСЖА для положительного словаря $\mathcal{D}$ и целевой функции f_0 построено разложение

$$\sum_{k=1}^{\infty} c_k g_k = \sum_{k=1}^{\infty} c_k(f_0, \mathcal{D}) g_k(f_0, \mathcal{D}).$$

Обозначим

$$D_n := D_n(f_0, \mathcal{D}) := \{g_k\}_{k=1}^n,$$

$$v_n(g) := v_n(g, f_0, \mathcal{D}) := \begin{cases} \sum_{k \leq n, g_k = g} c_k, & g \in D_n, \\ 0, & g \notin D_n. \end{cases}$$

Некоторые элементы g_k могут быть равны между собой, поэтому $\#D_n$ может быть меньше n . ПЧЖА и ПСЖА не гарантируют неотрицательность всех коэффициентов c_k , но любая частичная сумма ряда может быть рассмотрена как n -членное приближение с неотрицательными коэффициентами,

$$\sum_{k=1}^n c_k g_k = \sum_{g \in D_n} v_n(g) g, \quad v_n(g) \geq 0, \quad g \in \mathcal{D}_n.$$

ПОЛОЖИТЕЛЬНЫЙ ЧИСТО ЖАДНЫЙ АЛГОРИТМ (ПЧЖА). Пусть заданы целевая функция $f_0 \in H$ и положительный словарь $\mathcal{D}$. Положим $G_0^{\text{PGA}+}(f_0, \mathcal{D}) := 0$. Для каждого $m \geq 0$ последовательно выбираем вектор $g_{m+1} \in \mathcal{D}$ и $c_{m+1} \geq -v_m(g)$ так, чтобы

$$\|f_m - c_{m+1} g_{m+1}\| = \inf_{g \in \mathcal{D}, c \geq -v_m(g)} \|f_m - c g\|, \quad (6)$$

и определяем

$$G_{m+1}^{\text{PGA}+}(f_0, \mathcal{D}) := G_m^{\text{PGA}+}(f_0, \mathcal{D}) + c_{m+1} g_{m+1},$$

$$f_{m+1} := f_0 - G_{m+1}^{\text{PGA}+}(f_0, \mathcal{D}) = f_m - c_{m+1} g_{m+1}.$$

Отметим, что для фиксированных $g \in \mathcal{D}$ и $f \in H$ инфимум $\inf_{c \geq -v} \|f - cg\|$ достигается при

$$c = \max(-v, \langle f, g \rangle). \quad (7)$$

Для $g \in \mathcal{D}$ определим

$$\begin{aligned} K_m(g) &:= \sup_{c \geq -v_m(g)} \|f_m\|^2 - \|f_m - cg\|^2 \\ &= \langle f_m, f_m \rangle - \langle f_m - cg, f_m - cg \rangle \\ &= 2c\langle f_m, g \rangle - c^2 = \begin{cases} \langle f_m, g \rangle^2, & \langle f_m, g \rangle \geq -v_m(g), \\ v_m(g)(-2\langle f_m, g \rangle - v_m(g)), & \langle f_m, g \rangle < -v_m(g). \end{cases} \end{aligned} \quad (8)$$

Таким образом, легко видеть, что определение элемента g_{m+1} и числа c_{m+1} по формуле (6) можно переписать

$$g_{m+1} : K_m(g_{m+1}) = \sup_{g \in \mathcal{D}} K_m(g), \quad (9)$$

$$c_{m+1} = \max(-v_m(g_{m+1}), \langle f_m, g_{m+1} \rangle), \quad (10)$$

$$K_m(g_{m+1}) = \|f_m\|^2 - \|f_{m+1}\|^2. \quad (11)$$

Так как $v_m(g) \geq 0$ для всех $m \geq 0$ и $g \in \mathcal{D}$, то если $\langle f_m, g_{m+1} \rangle < -v_m(g_{m+1}) \leq 0$, то

$$\begin{aligned} K_m(g_{m+1}) &= v_m(g_{m+1})(-2\langle f_m, g_{m+1} \rangle - v_m(g_{m+1})) \\ &\in [-v_m(g_{m+1})\langle f_m, g_{m+1} \rangle, -2v_m(g_{m+1})\langle f_m, g_{m+1} \rangle]. \end{aligned}$$

Слабый жадный алгоритм отличается от чисто жадного тем, что на очередном шаге алгоритма при выборе вектора, максимизирующего скалярное произведение, нам разрешается ошибаться при вычислении скалярного произведения в t_m^{-1} раз. Таким образом, мы приходим к определению.

ПОЛОЖИТЕЛЬНЫЙ СЛАБО ЖАДНЫЙ АЛГОРИТМ (ПСЖА). Пусть фиксирована последовательность коэффициентов слабости $\tau = \{t_m\}_{m=1}^\infty \subset [0, 1]$. Для целевой функции $f_0 \in H$ и положительного словаря $\mathcal{D}$ положим $G_0^{\text{WGA}+}(f_0, \mathcal{D}, \tau) := 0$. Тогда для каждого $m \geq 0$ последовательно выбираем вектор $g_{m+1} \in \mathcal{D}$ такой, что

$$K_m(g_{m+1}) \geq \sup_{g \in \mathcal{D}} \begin{cases} t_{m+1}^2 \langle f_m, g \rangle^2, & t_{m+1} \langle f_m, g \rangle \geq -v_m(g), \\ t_{m+1} v_m(g)(-2\langle f_m, g \rangle - v_m(g)), & t_{m+1} \langle f_m, g \rangle < -v_m(g), \end{cases} \quad (12)$$

определяем c_{m+1} по формуле (10) и полагаем

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

Отметим, что согласно определению $K_m(g)$ имеет место равенство

$$K_m(g_{m+1}) = \|f_m\|^2 - \|f_{m+1}\|^2 \geq 0.$$

Определение ПСЖА корректно. Как и в случае СЖА, для любого $t_{m+1} \in (0, 1)$ мы всегда можем осуществить m -й шаг алгоритма, т.е. найти $g_{m+1} \in \mathcal{D}$, удовлетворяющий (12). Это вытекает из следующей леммы.

ЛЕММА 1. Для любых $g \in \mathcal{D}$ и $m \geq 0$ имеет место неравенство

$$K_m(g) = \begin{cases} \langle f_m, g \rangle^2, & \langle f_m, g \rangle \geq -v_m(g), \\ v_m(g)(-2\langle f_m, g \rangle - v_m(g)), & \langle f_m, g \rangle < -v_m(g), \end{cases}$$

$$\geq \begin{cases} t_{m+1}^2 \langle f_m, g \rangle^2, & t_{m+1} \langle f_m, g \rangle \geq -v_m(g), \\ t_{m+1} v_{m+1}(g)(-2\langle f_m, g \rangle - v_m(g)), & t_{m+1} \langle f_m, g \rangle < -v_m(g). \end{cases}$$

ДОКАЗАТЕЛЬСТВО. Так как $t_{m+1} \in [0, 1]$, то неравенство очевидным образом выполняется если $\langle f_m, g \rangle \leq t_{m+1} \langle f_m, g \rangle < -v_m(g)$, или если $t_{m+1} \langle f_m, g \rangle \geq -v_m(g)$ и $\langle f_m, g \rangle \geq -v_m(g)$. Если же $\langle f_m, g \rangle < -v_m(g) \leq t_{m+1} \langle f_m, g \rangle$, то

$$v_m(g)(-2\langle f_m, g \rangle - v_m(g)) \geq -t_{m+1} \langle f_m, g \rangle(-2\langle f_m, g \rangle + \langle f_m, g \rangle) \geq t_{m+1}^2 \langle f_m, g \rangle^2.$$

Справедливы следующие теоремы о сходимости положительных чисто жадного и слабо жадного алгоритмов.

ТЕОРЕМА 1. Для произвольного положительного словаря $\mathcal{D}$ и целевой функции f_0 выполняется

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

ТЕОРЕМА 2. Для произвольного положительного словаря $\mathcal{D}$, целевой функции f_0 и последовательности коэффициентов слабости $\{t_m\}_{m=1}^{\infty} \subset [0, 1]$, удовлетворяющей равенству

$$\sum_{m=1}^{\infty} \frac{t_m}{m} = \infty, \quad (13)$$

имеет место соотношение

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

Из (9), (12) и расходимости ряда $\sum_{m=1}^{\infty} 1/m$ вытекает, что теорема 1 является непосредственным следствием теоремы 2, поэтому ограничимся доказательством последней.

Доказательство теоремы 2. Положим для $m \geq 0$

$$a_m = \langle f_m, f_m \rangle = \|f_0 - G_m^{\text{WGA}+}(f_0, \mathcal{D}, \tau)\|^2,$$

$$b_m = \sum_{k=1}^m c_m = \sum_{g \in D_m} v_m(g) = \sum_{g \in \mathcal{D}} v_m(g). \quad (14)$$

Предположим, что ПСЖА не сходится для словаря $\mathcal{D}$ и целевой функции f_0 .

Тогда в силу монотонности чисел $\{a_m\}$ найдется $s > 0$, для которого

$$a_m \geq s > 0, \quad m \geq 0. \quad (15)$$

Вследствие того, что $\mathcal{D}$ является положительным словарем, найдется конечное множество индексов Λ , $\tilde{c}_\lambda \geq 0$, $\tilde{g}_\lambda \in \mathcal{D}$, $\lambda \in \Lambda$ и $r \in H$ таких, что

$$f_0 = \sum_{\lambda \in \Lambda} \tilde{c}_\lambda \tilde{g}_\lambda + r, \quad \|r\| \leq \frac{s}{5\|f_0\|}, \quad \sum_{\lambda \in \Lambda} \tilde{c}_\lambda =: A \quad (16)$$

для некоторого $A \geq 0$.

Разобьем доказательство теоремы на два случая.

Случай 1. Найдется $m \geq 0$, для которого выполняются неравенство и импликация:

$$\sup_{g \in \mathcal{D}} \langle f_m, g \rangle \leq \frac{s}{4A}, \quad (17)$$

$$\forall g \in \mathcal{D} \quad (t_{m+1} \langle f_m, g \rangle \geq -v_m(g)) \Rightarrow \left(|\langle f_m, g \rangle| \leq \frac{s}{4b_m} \right), \quad (18)$$

$$\forall g \in \mathcal{D} \quad (t_{m+1} \langle f_m, g \rangle < -v_m(g)) \Rightarrow \left(-v_m(g) \langle f_m, g \rangle \leq \frac{s}{4m} \right). \quad (19)$$

Используя (18), (19), неотрицательность $v_m(g)$, определение (14) чисел b_m и неравенство $\sharp D_m \leq m$, оценим

$$\begin{aligned} \langle f_m, f_m - f_0 \rangle &= \langle f_m, f_0 - \sum_{k=1}^m c_k g_k - f_0 \rangle = \langle f_m, - \sum_{g \in D_m} v_m(g) g \rangle \\ &= \sum_{g \in D_m, t_{m+1} \langle f_m, g \rangle \geq -v_m(g)} -v_m(g) \langle f_m, g \rangle \\ &\quad + \sum_{g \in D_m, t_{m+1} \langle f_m, g \rangle < -v_m(g)} -v_m(g) \langle f_m, g \rangle \\ &\leq \frac{s}{4b_m} \left(\sum_{g \in D_m} v_m(g) \right) + \sum_{g \in D_m} \frac{s}{4m} = \frac{s}{4b_m} b_m + \sharp D_m \frac{s}{4m} \leq \frac{s}{2}. \end{aligned}$$

Применяя (16), (17) и неравенство Коши–Буняковского, получим

$$\begin{aligned} \langle f_m, f_0 \rangle &= \langle f_m, \sum_{\lambda \in \Lambda} \tilde{c}_\lambda \tilde{g}_\lambda + r \rangle \\ &\leq \sum_{\lambda \in \Lambda} \tilde{c}_\lambda \sup_{g \in \mathcal{D}} \langle f_m, g \rangle + \|f_m\| \|r\| \leq A \frac{s}{4A} + \|f_0\| \frac{s}{5\|f_0\|} = \frac{9}{20} s. \end{aligned}$$

Два последних неравенства влекут, что

$$a_m = \langle f_m, f_m \rangle = \langle f_m, f_m - f_0 \rangle + \langle f_m, f_0 \rangle \leq \frac{19}{20} s,$$

что противоречит предположению (15).

Случай 2. Для всех $m \geq 0$ не выполняется хотя бы одно из условий (17), (18) или (19). Обозначим через M_1 , M_2 и M_3 подмножества $\mathbb{N}$, для которых не выполнены условия (17), (18) и (19) соответственно. Тогда $M_1 \cup M_2 \cup M_3 = \mathbb{N}$. Поэтому

$$\sum_{m \in M_1} \frac{t_{m+1}}{m} + \sum_{m \in M_2} \frac{t_{m+1}}{m} + \sum_{m \in M_3} \frac{t_{m+1}}{m} = \infty.$$

Следовательно, хотя бы одно из трех слагаемых бесконечно. Предположим, что

$$\sum_{m \in M_1} \frac{t_m}{m} = \infty. \quad (20)$$

Для всех $m \in M_1$ не выполняется условие (17). Это значит, что найдется $g \in \mathcal{D}$, для которого $\langle f_m, g \rangle \geq s/(5A)$. Следовательно, $t_{m+1}\langle f_m, g \rangle \geq 0 \geq -v_m(g)$ и по определению (12)

$$K_m(g_{m+1}) \geq t_{m+1}^2 \left(\frac{s}{5A} \right)^2.$$

С учетом (11) ясно, что

$$\begin{aligned} a_0 = \langle f_0, f_0 \rangle &\geq \sum_{m=0}^{\infty} \langle f_m, f_m \rangle - \langle f_{m+1}, f_{m+1} \rangle \\ &= \sum_{m=0}^{\infty} K_m(g_{m+1}) \geq \sum_{m \in M_1} K_m(g_{m+1}) \geq \sum_{m \in M_1} t_{m+1}^2 \left(\frac{s}{5A} \right)^2. \end{aligned}$$

Следовательно,

$$\sum_{m \in M_1} t_{m+1}^2 < \infty.$$

Поэтому по неравенству Коши–Буняковского

$$\sum_{m \in M_1} \frac{t_{m+1}}{m} \leq \left(\sum_{m \in M_1} t_{m+1}^2 \right)^{1/2} \left(\sum_{m \in M_1} m^{-2} \right)^{1/2} < \infty,$$

что противоречит предположению (20).

Предположим, что

$$\sum_{m \in M_2} \frac{t_m}{m} = \infty. \quad (21)$$

Из формул (7), (8) и (10) вытекает, что $K_m(g_{m+1}) \geq c_{m+1}^2$. Поэтому,

$$a_0 = \langle f_0, f_0 \rangle \geq \sum_{m=0}^{\infty} \langle f_m, f_m \rangle - \langle f_{m+1}, f_{m+1} \rangle = \sum_{m=0}^{\infty} K_m(g_{m+1}) \geq \sum_{m=1}^{\infty} c_m^2,$$

т.е. последовательности $\{K_m^{1/2}(g_{m+1})\}_{m=0}^{\infty}, \{c_m\}_{m=1}^{\infty} \in l_2$. Так как согласно (14) выполняется $b_m = \sum_{k=1}^m c_k$, то в силу [10; гл. 1, п. 9] можно заключить, что последовательность $\{b_m/m\}_{m=1}^{\infty}$ также принадлежит l_2 . (Здесь мы следуем доказательству леммы Конягина и Темлякова из работы [3].) Так как для произвольного $m \in M_2$ условие (18) не выполняется, найдется $g \in \mathcal{D}$, для которого верны оценки $t_{m+1}\langle f_m, g \rangle \geq -v_m(g)$ и $|\langle f_m, g \rangle| > s/(4b_m)$. Тогда согласно (12) имеем

$$\begin{aligned} K_m(g_{m+1}) &\geq t_{m+1}^2 \left(\frac{s}{4b_m} \right)^2, \\ K_m^{1/2}(g_{m+1}) \frac{b_m}{m} &\geq s \frac{t_{m+1}}{4m}. \end{aligned}$$

Суммируя по всем $m \in M_2$, имеем

$$\sum_{m \in M_2} K_m^{1/2}(g_{m+1}) \frac{b_m}{m} \geq \frac{s}{4} \sum_{m \in M_2} \frac{t_{m+1}}{m}.$$

Так как $\{K_m^{1/2}(g_{m+1})\}_{m=0}^\infty \in l_2$ и $\{b_m/m\}_{m=1}^\infty \in l_2$, левая часть последнего неравенство ограничена; следовательно, правая часть также ограничена, что противоречит (21).

Предположим, что

$$\sum_{m \in M_3} \frac{t_m}{m} = \infty. \quad (22)$$

Для всех $m \in M_3$ не выполняется условие (19). Это значит, что найдется $g \in \mathcal{D}$, для которого $t_{m+1}\langle f_m, g \rangle < -v_m(g)$ и $-v_m(g)\langle f_m, g \rangle > s/(4m)$. Следовательно, по определению (12) имеем

$$\begin{aligned} K_m(g_{m+1}) &\geq t_{m+1}v_m(g)(-2\langle f_m, g \rangle - v_m(g)) \\ &\geq t_{m+1}v_m(g)(-\langle f_m, g \rangle - t_{m+1}\langle f_m, g \rangle - v_m(g)) \\ &\geq -t_{m+1}v_m(g)\langle f_m, g \rangle \geq t_{m+1}\frac{s}{4m}. \end{aligned}$$

Тогда

$$a_0 \geq \sum_{m=0}^{\infty} K_m(g_{m+1}) \geq \sum_{m \in M_3} K_m(g_{m+1}) \geq \frac{s}{4} \sum_{m \in M_3} \frac{t_{m+1}}{m},$$

что противоречит неравенству (22). Это заканчивает разбор случая 2 и завершает доказательство теоремы.

Отметим, что при доказательстве теоремы 2 была использована техника из работ [11] и [12].

Связь между СЖА и ПСЖА. Необходимое условие сходимости ПСЖА.

Легко заметить, что если множество $\mathcal{D}$ является словарем в H , то множество $\hat{\mathcal{D}} = \mathcal{D} \cup (-\mathcal{D})$ является положительным словарем в H . Следующая лемма устанавливает связь между СЖА и ПСЖА.

ЛЕММА 2. Пусть фиксирована последовательность коэффициентов слабости $\tau = \{t_m\}_{m=1}^\infty \subset [0, 1]$, словарь $\mathcal{D}$ и целевая функция f_0 . Тогда любая реализация СЖА $\{G_m^{\text{WGA}}(f_0, \mathcal{D}, \tau)\}_{m=0}^\infty$ является одной из реализаций ПСЖА для последовательности коэффициентов слабости $\hat{\tau} = \{\hat{t}_m\}_{m=1}^\infty \subset [0, 1]$, $\hat{t}_m = t_m 2^{-1/2}$, $m \geq 1$, словаря $\hat{\mathcal{D}}$ и целевой функции f_0 :

$$G_m^{\text{WGA}}(f_0, \mathcal{D}, \tau) = G_m^{\text{WGA}+}(f_0, \hat{\mathcal{D}}, \hat{\tau}), \quad m \geq 0. \quad (23)$$

ДОКАЗАТЕЛЬСТВО. Докажем соотношение (23) по индукции. Для $m = 0$ выполняется $G_0^{\text{WGA}}(f_0, \mathcal{D}, \tau) = G_0^{\text{WGA}+}(f_0, \hat{\mathcal{D}}, \hat{\tau}) = 0$. Из справедливости (23) для m выведем справедливость (23) для $m + 1$. Пусть на m -м шаге обоих алгоритмов получен остаток f_m ,

$$f_m = f_0 - G_m^{\text{WGA}}(f_0, \mathcal{D}, \tau) = f_0 - G_m^{\text{WGA}+}(f_0, \hat{\mathcal{D}}, \hat{\tau}).$$

Предположим, что

$$G_{m+1}^{\text{WGA}}(f_0, \mathcal{D}, \tau) = G_m^{\text{WGA}}(f_0, \mathcal{D}, \tau) + \langle f_m, g_{m+1} \rangle g_{m+1}.$$

Тогда

$$|\langle f_m, g_{m+1} \rangle| \geq t_{m+1} \sup_{g \in \mathcal{D}} |\langle f_m, g \rangle|. \quad (24)$$

В силу того, что оба вектора g_{m+1} и $-g_{m+1}$ принадлежат $\widehat{\mathcal{D}}$, можно без ограничения общности считать, что

$$\langle f_m, g_{m+1} \rangle \geq 0. \quad (25)$$

Тогда, если g_{m+1} будет удовлетворять соотношению (12), то согласно (10) будет выполняться $c_{m+1} = \langle f_m, g_{m+1} \rangle$, что докажет формулу (23) для $m + 1$. С учетом соотношений (8), (24) и (25) имеем

$$\begin{aligned} \sup_{g \in \widehat{D}} \begin{cases} \widehat{t}_{m+1}^2 \langle f_m, g \rangle^2, & \widehat{t}_{m+1} \langle f_m, g \rangle \geq -v_m(g), \\ \widehat{t}_{m+1} v_{m+1}(g) (-2 \langle f_m, g \rangle - v_m(g)), & \widehat{t}_{m+1} \langle f_m, g \rangle < -v_m(g), \end{cases} \\ \leq \sup_{g \in \widehat{D}} 2 \widehat{t}_{m+1}^2 \langle f_m, g \rangle^2 \leq t_{m+1}^2 \sup_{g \in \widehat{D}} \langle f_m, g \rangle^2 \\ = t_{m+1}^2 \sup_{g \in \mathcal{D}} \langle f_m, g \rangle^2 \leq \langle f_m, g_{m+1} \rangle^2 = K_m(g_{m+1}). \end{aligned}$$

Таким образом, g_{m+1} удовлетворяет соотношению (12), что доказывает лемму.

Лемма 2 позволяет перенести необходимое условие сходимости СЖА для монотонных последовательностей коэффициентов слабости [4; теорема 1] на случай ПСЖА.

ТЕОРЕМА 3. Пусть задана последовательность коэффициентов слабости $\tau = \{t_m\}_{m=1}^\infty \subset [0, 1]$, удовлетворяющая неравенству

$$\sum_{m=1}^{\infty} \frac{t_m}{m} < \infty. \quad (26)$$

Тогда найдутся положительный словарь $\mathcal{D} \subset H$ и целевая функция $f_0 \in H$ такие, что найдется реализация ПСЖА $\{G_m^{\text{WGA}+}(f_0, \mathcal{D}, \tau)\}_{m=0}^\infty$, не сходящаяся к f_0 при $m \rightarrow \infty$.

СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ

- [1] В. В. Галатенко, Е. Д. Лившиц, “Обобщенные приближенные слабые жадные алгоритмы”, *Матем. заметки*, **78**:2 (2005), 186–201.
- [2] L. Jones, “A simple lemma on greedy approximation in Hilbert space and convergence rates for projection pursuit regression and network training”, *Ann. Statist.*, **20**:1 (1992), 608–613.
- [3] V. N. Temlyakov, “Weak greedy algorithms”, *Adv. Comput. Math.*, **12**:2–3 (2000), 213–227.
- [4] Е. Д. Лившиц, В. Н. Темляков, “О сходимости слабого гриди-алгоритма”, *Функциональные пространства, гармонический анализ, дифференциальные уравнения*, Тр. МИАН, **232**, 2001, 236–247.
- [5] V. N. Temlyakov, “A criterion for convergence of Weak Greedy Algorithms”, *Adv. Comp. Math.*, **17**:3 (2002), 269–280.
- [6] A. Barron, “Approximation and estimation bounds for artificial neural networks”, *Proceedings of the Fourth Annula Workshop on Computational Learning Theory*, Kaufmann, 1991, 243–249.

- [7] A. Barron, A. Cohen, W. Dahmen, R. DeVore, *Approximation and Learning by Greedy Algorithms submitted to Annals of Statistics*, <http://www.igpm.rwth-aachen.de/~dahmen/rep-06.html>, 2006.
- [8] M. Donahue, L. Gurvits, C. Darken, E. Sontag, “Rate of convex approximation in non-Hilbert spaces”, *Constr. Approx.*, **13**:2 (1997), 187–220.
- [9] V.N. Temlyakov, *Relaxation in Greedy Approximation submitted to Constructive Approximation*, <http://www.math.sc.edu/IMI/technical/tech06.html>.
- [10] А. Зигмунд, *Тригонометрические ряды*, Мир, М., 1965.
- [11] R. A. DeVore, V. N. Temlyakov, “Some remarks on Greedy Algorithms”, *Adv. Comput. Math.*, **5**:1 (1996), 173–187.
- [12] Е. Д. Лившиц, “О возвратном жадном алгоритме”, *Изв. РАН. Сер. матем.*, **70**:1 (2006), 95–116.

Е. Д. Лившиц

Московский государственный университет

им. М. В. Ломоносова

E-mail: livshitz@rambler.ru

Поступило

03.11.2006

Исправленный вариант

29.01.2007