МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА
МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ

На правах рукописи

Орлова Анастасия Сергеевна

О сходимости и скорости сходимости жадных приближений в специальных случаях

1.1.1 — вещественный, комплексный и функциональный анализ

ДИССЕРТАЦИЯ
на соискание ученой степени
кандидата физико-математических наук

Научный руководитель:
доктор физико-математических наук,
профессор Лукашенко Тарас Павлович,
кандидат физико-математических наук
Галатенко Владимир Владимирович

Москва — 2023

Оглавление

Введение.

Актуальность проблемы.

Исследование рядов Фурье является классической областью математики. Начало изучения рядов Фурье относится к первой половине XIX века. Так, в 1807 году Ж.-Б. Ж. Фурье был сделан доклад перед Французской академией на тему представления функции в виде тригонометрического ряда, в 1811 году были опубликованы его записи [16], а в 1822 году опубликована работа [17] по этой теме. После этого последовали работы О. Л. Коши 1823–1828 годов [12], [11], [13], и дальнейшее развитие тема рядов Фурье получила в работах Г. Л. Дирихле, например, [15]. Первоначальной целью введения тригонометрических рядов было решение уравнения теплопроводности, однако позже было найдено множество различных приложений.

Ортогональные ряды Фурье являются обобщением тригонометрических рядов Фурье. Их активное изучение началось в первой половине прошлого века. В это время было написано много работ в этом направлении. Так, в 1909 году вышла статья Г. Вейля [31], позже были опубликованы работы А. Зигмунда 1927 и 1930 годов [32], [33], работы Г. Штейнгауза 1920–1934 годов [25], [27], [28], [26]. Отдельным направлением является изучение ортогональных рядов Фурье на конкретных системах. Тематике ортогональных рядов также посвящены работы А. Хаара [18], [20], [19] и Г. Радемахера [24].

Переупорядочение слагаемых в ряде Фурье по убыванию норм совершенно естественно с точки зрения приближения n-членными линейными комбинациями векторов ортогональной системы, так как при таком переупорядочении частичные суммы ряда Фурье и являются элементами наилучшего n-членного приближения (это утверждение сразу следует из экстремального свойства коэффициентов Фурье и тождества Бесселя [5, гл. 3, §4, п.4]). Ещё в работе С. Б. Стечкина 1955 года [9] установлено, что такое переупорядочение естественным образом возникает и при изучении вопроса абсолютной сходимости ряда Фурье.

Переупорядочение ряда Фурье по убыванию норм слагаемых эквивалентно применению к ортогональной системе и приближаемому элементу чисто жадного алгоритма [14], а также ортогонального жадного алгоритма [14].

В связи с активным развитием информационных технологий и необходимостью хранить и передавать данные в 90-х годах прошлого столетия жадные алгоритмы начали активно изучаться.

Сходимость чисто жадного алгоритма была доказана Л. Джонсом в 1992 году [21]. Более точно, была доказана следующая теорема.

Теорема А. *В гильбертовом пространстве H для произвольного словаря D и любого вектора $x \in H$ чисто жадный алгоритм сходится к приближаемому вектору x.*

Более подробное исследование сходимости алгоритма заключается в изучении скорости сходимости на индивидуальном векторе или на классе векторов. Классическим для изучения является класс векторов $\mathcal{A}_1(D)$, который можно считать обобщением класса ℓ_1 в пространстве ℓ_2 со стандартным ортогональным словарём D.

Когда речь идёт о скорости сходимости алгоритма на конкретном векторе x, вводится специальная последовательность $c_n(x)$, характеризующая наибольшее возможное отклонение жадного приближения от x на n-м шаге. Для оценки скорости сходимости на классе также вводится аналогичная последовательность. Так, в случае класса $\mathcal{A}_1(D)$ вводится последовательность c_n, которая характеризует наибольшее возможное отклонение жадного приближения от приближаемого вектора на n-м шаге для элементов из класса $\mathcal{A}_1(D)$.

В 1996 году в работе Р. ДеВора и В. Н. Темлякова [14] была получена оценка скорости сходимости для ортогонального жадного алгоритма на классе $\mathcal{A}_1(D)$.

Теорема B. *Для произвольного словаря D верно*

$$c_n^{OGA} \leqslant n^{-\frac{1}{2}}, \quad \forall n \in \mathbb{N}.$$

А. Бэррон в 1993 году [10] нашёл для класса $\mathcal{A}_1(D)$ скорость сходимости наилучших n-членных приближений, она совпадает с вышеприведённой оценкой скорости сходимости для ортогонального жадного алгоритма.

В той же работе Р. ДеВора и В. Н. Темлякова 1996 года [14] получена оценка скорости сходимости на классе $\mathcal{A}_1(D)$ для чисто жадного алгоритма.

Теорема C. *Для произвольного словаря D верно*

$$c_n^{PGA} \leqslant n^{-\frac{1}{6}}, \quad \forall n \in \mathbb{N}.$$

Данная оценка сверху уточнялась в работах [23] и [8]. В работе А. В. Сильниченко 2004 года [8] был получен такой результат.

Теорема D. *Для произвольного словаря D верно*

$$c_n^{PGA} \leqslant 1,7\ n^{-0{,}182}, \quad \forall n \in \mathbb{N}.$$

Оценка снизу также последовательно улучшалась в работах [14], [7] и [6]. Так, в работе [6] Е. Д. Лившицем доказана следующая оценка.

Теорема E. *Существуют словарь D и элемент $x \in \mathcal{A}_1(D)$, такие что для некоторого числа $C > 0$ верно*

$$c_n^{PGA}(x) > Cn^{-0{,}1898}, \quad \forall n \in \mathbb{N}.$$

В 2023 году был анонсирован [22] новый результат о скорости сходимости чисто жадного алгоритма, который пока не опубликован в рецензируемом издании. Данный результат уточняет нижнюю оценку скорости сходимости, причём показатель в новой оценке совпадает с показателем в оценке сверху, полученной А. В. Сильниченко в 2004 году.

Таким образом, скорость сходимости чисто жадного алгоритма на классе $\mathcal{A}_1(D)$ медленнее, чем для ортогонального жадного алгоритма. С другой стороны, в 2012 году А. В. Деревенцовым [4] был приведён пример вектора, на котором сходимость чисто жадного алгоритма существенно быстрее, чем ортогонального жадного алгоритма.

Наиболее сложным в реализации чисто жадного алгоритма и ортогонального жадного алгоритма является выбор очередного приближающего элемента $e_{n+1}(x)$. В. Н. Темляковым в работе [30] было предложено ввести так называемую ослабляющую последовательность $\{t_n\}_{n=1}^{\infty} \subset (0,1]$ и в качестве очередного элемента разложения $e_{n+1}(x)$ выбирать произвольный элемент словаря, удовлетворяющий ослабленному условию. Применение слабых жадных алгоритмов к ортогональным словарям приводит к рядам Фурье, в которых слагаемые переупорядочены, но требование монотонности существенно ослаблено (и ослабление тем больше, чем меньше значения t_n). Модификации жадных алгоритмов, в которых выбор очередного элемента разложения $e_{n+1}(x)$ осуществляется на основе такого ослабленного условия, называются слабыми жадными алгоритмами. В частности, соответствующая модификация чисто жадного алгоритма называется слабым жадным алгоритмом. При введении ослабляющей последовательности, ортогональному жадному алгоритму соответствует слабый ортогональный жадный алгоритм, который в случае ортогонального словаря эквивалентен слабому жадному алгоритму.

Сходимость слабых модификаций жадных алгоритмов зависит от ослабляющей последовательности, которая используется в данном алгоритме. Поэтому важным вопросом в изучении сходимости жадных алгоритмов является условие на ослабляющую последовательность, достаточное для сходимости алгоритма.

В работе 2000 года [30] были установлены следующие достаточные условия сходимости на ослабляющую последовательность для слабого жадного алгоритма и слабого ортогонального жадного алгоритма.

Теорема F. *Если для ослабляющей последовательности* $\{t_n\}_{n=1}^{\infty}$ *выполняется условие*

$$\sum_{k=1}^{\infty} \frac{t_k}{k} = +\infty,$$

то в гильбертовом пространстве H *для произвольного словаря* D *и любого вектора* $x \in H$ *слабый жадный алгоритм сходится к приближаемому вектору* x.

Теорема G. *Если для ослабляющей последовательности* $\{t_n\}_{n=1}^{\infty}$ *выполняется условие*

$$\sum_{k=1}^{\infty} t_k^2 = \infty,$$

то в гильбертовом пространстве H *для произвольного словаря* D *и любого вектора* $x \in H$ *слабый ортогональный жадный алгоритм сходится к приближаемому вектору* x.

Оценки скорости сходимости слабого ортогонального жадного алгоритма и слабого жадного алгоритма на классе $\mathcal{A}_1(D)$ были получены в той же работе [30] и могут быть сформулированы следующим образом.

Теорема H. *Для произвольных словаря* D *и ослабляющей последовательности* $\{t_n\}_{n=1}^{\infty}$ *верно*

$$c_n^{WOGA} \leqslant \left(1 + \sum_{k=1}^{n} t_k^2\right)^{-\frac{1}{2}}, \quad \forall n \in \mathbb{N}.$$

Теорема I. *Для произвольных словаря* D *и невозрастающей ослабляющей последовательности* $\{t_n\}_{n=1}^{\infty}$ *верно*

$$c_n^{WGA} \leqslant \left(1 + \sum_{k=1}^{n} t_k^2\right)^{-\frac{t_n}{2(2+t_n)}}, \quad \forall n \in \mathbb{N}.$$

Чисто жадные алгоритмы по системе возможно неполных словарей [2], в свою очередь, являются ещё одним из многих обобщений [29], [1], [3] чисто жадных алгоритмов. В работе П. А. Бородина и Е. Копецки [2] была доказана следующая теорема о слабой сходимости нового алгоритма.

Теорема J. *В гильбертовом пространстве H для произвольной системы возможно неполных словарей $\{D_1, D_2, \ldots, D_n\}$ таких, что для любого ненулевого вектора существует неортогональный вектор из $\cup_{i=1}^{n} D_i$, и любого вектора $x \in H$ чисто жадный алгоритм по системе возможно неполных словарей $\{D_1, D_2, \ldots, D_n\}$ слабо сходится к приближаемому вектору x.*

Постановка задач.

Сформулированные теоремы о достаточном условии сходимости слабых жадных приближений для векторов гильбертова пространства на ослабляющую последовательность приводят к вопросу о возможности ослабления данного достаточного условия сходимости, например, если рассматриваются векторы класса $\mathcal{A}_1(D)$, а на сам словарь D накладываются дополнительные ограничения. Так, интересен случай «минимального» словаря, когда в качестве словаря рассматривается ортогональный словарь. Если в данном случае такое ослабление возможно, то возможно ли аналогичное ослабление в случае векторов класса $\mathcal{A}_1(D)$ для произвольного словаря или хотя бы для некоторого расширения ортогонального словаря?

В связи с полученными оценками скорости сходимости жадных приближений на классе $\mathcal{A}_1(D)$ в случае произвольного словаря D возникает вопрос о возможности улучшения данных оценок. Возможно ли улучшить оценку скорости сходимости, когда словарь D является не произвольным, а, например, ортогональным?

Добавление к словарю дополнительных элементов расширяет аппарат приближения алгоритма. Можно ли утверждать, что в этом случае улучшается скорость сходимости? В частности, существует ли вектор такой, что приближение по ортогональному словарю быстрее, чем приближение по расширению ортогонального словаря?

При применении жадного алгоритма по паре словарей — частного случая жадного алгоритма по системе возможно неполных словарей — аппарат приближения также оказывается шире, чем при использовании жадного алгоритма по одному из словарей. С другой стороны, если речь идёт о сравнении жадного алгоритма по паре словарей и жадного алгоритма по объединению данных словарей, то не любая реализация второго алгоритма является реализацией первого. Таким образом, интересно сравнение скорости сходимости данных алгоритмов, хотя бы на индивидуальном векторе.

Здесь же возникает вопрос о сходимости ортогонального жадного алгоритма по системе возможно неполных словарей к приближаемому вектору гильбертова пространства.

Цель работы.

Цели работы включают:

- исследование классического вопроса достаточного условия сходимости слабых жадных алгоритмов на ослабляющую последовательность в случае ортогональных словарей и в случае словарей, являющихся расширениями ортогональных словарей;
- сравнение скорости сходимости чисто жадного алгоритма на ортогональном словаре и на расширении данного словаря;
- исследование вопроса сходимости ортогонального жадного алгоритма по системе возможно неполных словарей;
- сравнение скорости сходимости чисто жадных алгоритмов и соответствующих алгоритмов по паре словарей.

Научная новизна.

Основные результаты диссертации являются новыми и состоят в следующем.

1. Показано, что общие результаты о скорости сходимости слабых (ортогональных) жадных приближений в случае ортогонального словаря могут быть уточнены, причём полученное уточнение асимптотически неулучшаемо.

2. Показано, что достаточное условие сходимости слабого (ортогонального) жадного алгоритма в случае ортогонального словаря может быть ослаблено, причём полученное условие асимптотически неулучшаемо.

3. Показано, что достаточное условие сходимости в случае расширения ортогонального словаря одним вектором не может быть ослаблено для слабого жадного алгоритма и слабого ортогонального жадного алгоритма.

4. Показано, что скорость сходимости чисто жадного алгоритма по ортогональному словарю для индивидуального вектора может быть выше, чем скорость сходимости чисто жадного алгоритма по расширению данного словаря одним вектором.

5. Показано, что ортогональный жадный алгоритм по системе возможно неполных словарей сходится к приближаемому вектору гильбертова пространства.

6. Показано, что сходимость стандартного жадного алгоритма для индивидуального вектора может быть быстрее, чем сходимость жадного алгоритма по паре соответствующих словарей в случае чисто жадного алгоритма и в случае ортогонального жадного алгоритма; также доказано, что реализуема и обратная ситуация.

Методы исследования.

В работе использованы методы математического анализа, методы функционального анализа и специальные методы изучения жадных разложений.

Теоретическая и практическая ценность.

Диссертация носит теоретический характер. Её результаты могут быть использованы в теории жадных приближений, а также в прикладных вопросах обработки и передачи информации.

Соответствие паспорту научной специальности.

В диссертации изучается сходимость жадных алгоритмов, в силу чего диссертация соответствует паспорту специальности 1.1.1 “Вещественный, комплексный и функциональный анализ” по направлению “вещественный анализ”.

Положения, выносимые на защиту.

1. В случае ортогонального словаря получена оценка скорости сходимости слабого (ортогонального) жадного алгоритма и доказана асимптотическая неулучшаемость данной оценки.

2. Для слабого жадного алгоритма и слабого ортогонального жадного алгоритма в случае расширения ортогонального словаря одним вектором показана невозможность ослабления достаточного условия сходимости.

3. Для чисто жадного алгоритма показано, что скорость сходимости разложения по ортогональному словарю для индивидуального вектора может быть выше, чем скорость сходимости разложения по расширению данного словаря одним вектором.

4. Для ортогонального жадного алгоритма по системе возможно неполных словарей доказана сходимость разложения к приближаемому вектору гильбертова пространства.

5. В случае чисто жадного алгоритма и в случае ортогонального жадного алгоритма показано, что сходимость стандартного жадного алгоритма для индивидуального вектора может быть быстрее и может быть медленнее, чем сходимость жадного алгоритма по паре соответствующих словарей.

Апробация работы.

Результаты диссертации докладывались автором на следующих научных **конференциях**:

- Воронежская зимняя математическая школа «Современные методы теории функций и смежные проблемы» (Воронеж, 2017 г.);
- Воронежская зимняя математическая школа «Современные методы теории функций и смежные проблемы» (Воронеж, 2019 г.);
- XXVIII Международная конференция студентов, аспирантов и молодых учёных «Ломоносов» (Москва, 2021 г.);
- конференция «Аппроксимация и дискретизация» (Москва, 2021 г.);
- XXI Международная Саратовская зимняя школа «Современные проблемы теории функций и их приложения» (Саратов, 2022 г.);
- Воронежская зимняя математическая школа «Современные методы теории функций и смежные проблемы» (Воронеж, 2023 г.).

По теме диссертации были сделаны доклады на следующих **научно-исследовательских семинарах**:

- научный семинар «Ортоподобные системы» механико-математического факультета МГУ имени М. В. Ломоносова под руководством профессора Т. П. Лукашенко, доцента В. В. Галатенко, доцента Т. В. Родионова (многократно, 2014–2022 гг.);
- специальный семинар «Тригонометрические и ортогональные ряды» кафедры теории функций и функционального анализа механико-математического факультета МГУ имени М. В. Ломоносова под руководством профессора М. И. Дьяченко, профессора Т. П. Лукашенко, профессора В. А. Скворцова, профессора А. П. Солодова (2022 г.);
- специальный семинар «Геометрическая теория приближений» кафедры теории функций и функционального анализа механико-математического факультета МГУ имени М. В. Ломоносова под руководством профессора П. А. Бородина (2023 г.).

Публикации.

Основные результаты диссертации изложены в 8 публикациях автора [34]–[41]. Из них 3 статьи [34]–[36] опубликованы в рецензируемых научных изданиях, рекомендованных для защиты в диссертационном совете МГУ.011.3 по специальности 1.1.1 — “вещественный, комплексный и функциональный анализ” и входящих в базы цитирования Scopus, Web of Science и RSCI. Статьи [34]–[36] имеют две версии: на русском языке и на английском языке. В материалах международных конференций представлены 5 публикаций [37]–[41]. Работ в соавторстве нет. Список работ автора приведен в конце автореферата и диссертации.

Личный вклад автора.

Диссертантом совместно с научными руководителями проводился выбор темы, а также осуществлялось планирование всей работы. Автору диссертации принадлежат доказательства основных результатов, приведенных в диссертации.

Структура и объем работы.

Диссертация состоит из введения, 4 глав, заключения и списка литературы из 41 наименования. Общий объем работы составляет 80 страниц.

Краткое содержание диссертации.

Нумерация приводимых здесь результатов соответствует нумерации в основном тексте работы.

В **первой главе** содержатся основные понятия и определения. Так, приведены определения нормированного словаря, множества $\mathcal{A}_1(D)$ приближаемых векторов из гильбертова пространства H словарём D, определения рассматриваемых жадных алгоритмов: слабого жадного алгоритма и слабого ортогонального жадного алгоритма, чисто жадного алгоритма по системе возможно неполных словарей и ортогонального жадного алгоритма по системе возможно неполных словарей, чисто жадного алгоритма по паре словарей и ортогонального жадного алгоритма по паре словарей. Также в этой главе вводятся понятия скорости сходимости и сравнения алгоритмов по скорости сходимости, описывается рассматриваемый случай.

Во **второй главе** изучается скорость сходимости слабого ортогонального жадного алгоритма на подпространстве $\ell_1 \subset \ell_2$ в случае ортогонального словаря. Показано, что общие результаты о скорости сходимости слабых ортогональных жадных приближений в этом случае могут быть значительно уточнены, также установлено, что полученное уточнение асимптотически неулучшаемо.

Теорема 2.2. *Пусть для ослабляющей последовательности $\{t_n\}_{n=1}^{\infty}$ выполняется условие $\sum\limits_{k=1}^{\infty} t_k = +\infty$. Тогда для стандартного ортогонального словаря F имеет место эквивалентность*

$$c_n^{WOGA} \sim \frac{1}{2\sqrt{\sum\limits_{k=1}^{n} t_k}}, \quad n \to \infty.$$

В **третьей главе** изучается сходимость слабых жадных алгоритмов и слабых ортогональных жадных алгоритмов в случае некоторых словарей, являющихся расширениями ортогонального словаря. В частности, показано, что для словаря, полученного из ортогонального добавлением одного вектора, уже нельзя ослабить достаточное условие сходимости на ослабляющую последовательность также, как в случае ортогонального словаря. Более точно, получена следующая теорема для слабого ортогонального жадного алгоритма.

Теорема 3.2. *Для любой ослабляющей последовательности $\tau \in \ell_2 \setminus \ell_1$ существуют вектор $g \in \ell_2$ и вектор $x \in \ell_1$ такие, что для x существует реализация слабого ортогонального жадного алгоритма с ослабляющей последовательностью $t = c\tau$ ($c = \frac{\tau_1}{\|\tau\|_2^2}$, где $\tau = \{\tau_n\}_{n=1}^{\infty}$) по словарю, полученному из стандартного ортогонального добавлением вектора g, которая не сходится к приближаемому вектору x.*

Для слабого жадного алгоритма доказана аналогичная теорема.

Теорема 3.4. *Для любой ослабляющей последовательности $\tau \in \ell_2 \setminus \ell_1$ существуют вектор $g \in \ell_2$ и вектор $x \in \ell_1$ такие, что для x существует реализация слабого жадного алгоритма с ослабляющей последовательностью $t = c\tau$ ($c = \frac{\tau_1}{\|\tau\|_2^2}$, где $\tau = \{\tau_n\}_{n=1}^{\infty}$) по словарю, полученному из стандартного ортогонального добавлением вектора g, которая не сходится к приближаемому вектору x.*

Дополнительно показано, что добавление к стандартному ортогональному словарю одного вектора — даже из ℓ_1 — может значительно ухудшить скорость сходимости чисто жадного алгоритма.

Теорема 3.5. *Существуют вектор $g \in \ell_1$ и финитный вектор x такие, что любая реализация чисто жадного алгоритма по словарю, полученному из стандартного ортогонального добавлением вектора g, не сходится к приближаемому вектору x за конечное число шагов.*

Четвёртая глава посвящена сравнению стандартного жадного алгоритма и соответствующего жадного алгоритма по паре словарей. В первой части данной главы рассматривается случай чисто жадного алгоритма.

Показано, что на индивидуальном векторе новый чисто жадный алгоритм по паре словарей может быть как быстрее, так и медленнее стандартного чисто жадного алгоритма. Более точно, доказывается следующая теорема.

Теорема 4.1. *Существуют*

I. *ортонормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что*

1) *чисто жадный алгоритм по словарю D_1 и чисто жадный алгоритм по словарю D_2 не сходятся за конечное число шагов,*
2) *чисто жадный алгоритм по паре словарей D_1 и D_2 сходится за конечное число шагов.*

II. *ортонормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что*

1) *чисто жадный алгоритм по словарю D_1 и чисто жадный алгоритм по словарю D_2 сходятся за конечное число шагов,*
2) *чисто жадный алгоритм по паре словарей D_1 и D_2 не сходится за конечное число шагов.*

Во второй части четвёртой главы речь идёт об обобщении ортогонального жадного алгоритма. Так, получен положительный результат о сходимости ортогонального жадного алгоритма по системе возможно неполных словарей.

Теорема 4.4. *В гильбертовом пространстве H для произвольной системы возможно неполных словарей $\{D_1, D_2, \ldots, D_N\}$ и любого вектора $x \in H$ ортогональный жадный алгоритм по системе возможно неполных словарей $\{D_1, D_2, \ldots, D_N\}$ сходится к приближаемому вектору x.*

При сравнении ортогонального жадного алгоритма и ортогонального жадного алгоритма по паре словарей получена теорема, аналогичная теореме, сформулированной выше для чисто жадного алгоритма.

Теорема 4.5. *Существуют*

I. *ортонормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что*

1) *ортогональный жадный алгоритм по словарю* D_1 *и ортогональный жадный алгоритм по словарю* D_2 *не сходятся за конечное число шагов,*

2) *ортогональный жадный алгоритм по паре словарей* D_1 *и* D_2 *сходится за конечное число шагов.*

II. *нормированные словари* D_1 *и* D_2 *и вектор* $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ *такие, что*

1) *ортогональный жадный алгоритм по словарю* D_1 *и ортогональный жадный алгоритм по словарю* D_2 *сходятся за конечное число шагов,*

2) *ортогональный жадный алгоритм по паре словарей* D_1 *и* D_2 *не сходится за конечное число шагов.*

Глава 1

Основные определения.

Напомним определения *слабого жадного алгоритма* (*WGA*) [30] и *слабого ортогонального жадного алгоритма* (*WOGA*) [30].

Пусть H — гильбертово пространство, $D \subset H$ — *нормированный словарь* (т. е. линейная оболочка D всюду плотна в H и все элементы D имеют единичную норму в H), $\{t_n\}_{n=1}^{\infty} \subset (0,1]$ — *ослабляющая последовательность*.

Слабый жадный алгоритм. Для вектора $x \in H$ индуктивно определим *последовательность остатков* $\left\{r_n^{WGA}(x)\right\}_{n=0}^{\infty}$, *последовательность коэффициентов* $\left\{\hat{x}_n^{WGA}\right\}_{n=1}^{\infty}$ и *последовательность элементов словаря* $\left\{e_n^{WGA}(x)\right\}_{n=1}^{\infty} \subset D$:

$$r_0^{WGA}(x) = x;$$

$$e_{n+1}^{WGA}(x) \in D: \quad \left|\left(r_n^{WGA}(x), e_{n+1}^{WGA}(x)\right)\right| \geqslant t_{n+1} \sup_{d \in D} \left|\left(r_n^{WGA}(x), d\right)\right|,$$

$$\hat{x}_{n+1}^{WGA} = (r_n^{WGA}(x), e_{n+1}^{WGA}(x)), \quad r_{n+1}^{WGA}(x) = r_n^{WGA}(x) - \hat{x}_{n+1}^{WGA} e_{n+1}^{WGA}(x),$$

$$n = 0,\, 1,\, 2,\, \ldots\ .$$

Слабым жадным разложением вектора x по словарю D называется ряд $\sum\limits_{n=1}^{\infty} \hat{x}_n^{WGA} e_n^{WGA}(x)$.

Из определения сразу следует, что для остатка разложения $r_N^{WGA}(x)$ верно представление

$$r_N^{WGA}(x) = x - \sum_{n=1}^{N} \hat{x}_n^{WGA} e_n^{WGA}(x).$$

Разложение всегда возможно, если все $t_n < 1$ при $n \in \mathbb{N}$, но не всегда определено однозначно.

Слабый ортогональный жадный алгоритм. Для $x \in H$ определим индуктивно *последовательность остатков* $\left\{r_n^{WOGA}(x)\right\}_{n=0}^{\infty}$, *последовательность приближений* $\left\{G_n^{WOGA}(x)\right\}_{n=0}^{\infty}$ и *последовательность элементов словаря* $\left\{e_n^{WOGA}(x)\right\}_{n=1}^{\infty} \subset D$:

$$G_0^{WOGA}(x) = 0, \quad r_0^{WOGA}(x) = x;$$

$$e_{n+1}^{WOGA}(x) \in D :$$

$$\left|\left(r_n^{WOGA}(x), e_{n+1}^{WOGA}(x)\right)\right| \geqslant t_{n+1} \sup_{d \in D} \left|\left(r_n^{WOGA}(x), d\right)\right|, \qquad (1.1)$$

$$G_{n+1}^{WOGA}(x) = \operatorname{Proj}_{\left\{e_k^{WOGA}(x)\right\}_{k=1}^{n+1}} x, \quad r_{n+1}^{WOGA}(x) = x - G_{n+1}^{WOGA}(x),$$

$$n = 0,\, 1,\, 2,\, \ldots\, .$$

Здесь $\operatorname{Proj}_{\{e_k^{WOGA}(x)\}_{k=1}^{n+1}}$ — оператор ортогонального проектирования на линейную оболочку $\{e_k^{WOGA}(x)\}_{k=1}^{n+1}$.

Последовательность $\{G_n^{WOGA}(x)\}_{n=0}^{\infty}$ называется *последовательностью слабых ортогональных жадных приближений вектора x по словарю D*, а элемент $G_n^{WOGA}(x)$ — *n-м слабым ортогональным жадным приближением x по словарю D*.

Если $t_n \equiv 1$, то слабый жадный алгоритм совпадает с *чисто жадным алгоритмом (PGA)*, а слабый ортогональный жадный алгоритм — с *ортогональным жадным алгоритмом (OGA)*.

Отметим, что для ортогонального словаря $\sup\limits_{d \in D} \left|\left(r_n^{PGA}(x), d\right)\right|$ заведомо достигается и вместо точной верхней грани можно говорить о максимуме. С другой стороны, в отличие от слабых жадных алгоритмов, чисто жадные алгоритмы не всегда реализуемы.

Пусть $D_1, D_2, \ldots, D_N$ — множества элементов единичной нормы в H таких, что объединение данных множеств $D_1 \cup D_2 \cup \cdots \cup D_N = D$ является словарём (такой набор множеств называется *системой возможно неполных словарей*). Тогда аналогично чисто жадному разложению по словарю D, можно определить **чисто жадное разложение вектора по системе возможно неполных словарей** $\{D_1, D_2, \ldots, D_N\}$, где на очередном шаге n в качестве словаря берётся множество $D_{i(n)}$ для заранее определённых $i(n) \in \overline{1, N}$.

Более точно, зафиксируем последовательность $i(n) \in \overline{1, N}$ такую, что $i(n) \neq i(n+1)$ при $\forall n \in \mathbb{N}$, причём каждое число $k \in \overline{1, N}$ встречается в последовательности $i(n)$ бесконечное число раз. Для вектора $x \in H$ определим *чисто жадное разложение вектора по системе возможно неполных словарей* $\mathcal{D} = \{D_1, D_2, \ldots, D_N\}$ [2] следующим образом. Индуктивно определим *последовательность остатков* $\left\{r_n^{PGA\ \mathcal{D}}(x)\right\}_{n=0}^{\infty}$, *последовательность коэффициентов* $\left\{\hat{x}_n^{PGA\ \mathcal{D}}\right\}_{n=1}^{\infty}$ и *последовательность элементов словарей* $\left\{e_n^{PGA\ \mathcal{D}}(x)\right\}_{n=1}^{\infty}$:

$$r_0^{PGA\,\mathcal{D}}(x) = x;$$

$$e_{n+1}^{PGA\,\mathcal{D}}(x) \in D_{i(n+1)}: \quad \left|\left(r_n^{PGA\,\mathcal{D}}(x), e_{n+1}^{PGA\,\mathcal{D}}(x)\right)\right| = \sup_{d \in D_{i(n+1)}} \left|\left(r_n^{PGA\,\mathcal{D}}(x), d\right)\right|,$$

$$\hat{x}_{n+1}^{PGA\,\mathcal{D}} = (r_n^{PGA\,\mathcal{D}}(x), e_{n+1}^{PGA\,\mathcal{D}}(x)),$$

$$r_{n+1}^{PGA\,\mathcal{D}}(x) = r_n^{PGA\,\mathcal{D}}(x) - \hat{x}_{n+1}^{PGA\,\mathcal{D}} e_{n+1}^{PGA\,\mathcal{D}}(x), \quad n = 0,\, 1,\, 2,\, \ldots\ .$$

Чисто жадным разложением вектора x по системе возможно неполных словарей $\mathcal{D} = \{D_1, D_2, \ldots, D_k\}$ называется ряд $\sum\limits_{n=1}^{\infty} \hat{x}_n^{PGA\,\mathcal{D}} e_n^{PGA\,\mathcal{D}}(x)$.

Как и в случае чисто жадного алгоритма разложение не всегда реализуемо и не всегда определено однозначно, для остатка разложения также верно соотношение $r_N^{PGA\,\mathcal{D}}(x) = x - \sum\limits_{n=1}^{N} \hat{x}_n^{PGA\,\mathcal{D}} e_n^{PGA\,\mathcal{D}}(x)$.

Аналогично определяется **ортогональный жадный алгоритм по системе возможно неполных словарей** $\mathcal{D} = \{D_1, D_2, \ldots, D_N\}$. А именно, для вектора $x \in H$ индуктивно определим *последовательность остатков* $\left\{r_n^{OGA\,\mathcal{D}}(x)\right\}_{n=0}^{\infty}$, *последовательность приближений* $\left\{G_n^{OGA\,\mathcal{D}}(x)\right\}_{n=0}^{\infty}$ и *последовательность элементов словарей* $\left\{e_n^{OGA\,\mathcal{D}}(x)\right\}_{n=1}^{\infty}$:

$$G_0^{OGA\,\mathcal{D}}(x) = 0, \quad r_0^{OGA\,\mathcal{D}}(x) = x;$$

$$e_{n+1}^{OGA\,\mathcal{D}}(x) \in D_{i(n+1)}: \quad \left|\left(r_n^{OGA\,\mathcal{D}}(x), e_{n+1}^{OGA\,\mathcal{D}}(x)\right)\right| = \sup_{d \in D_{i(n+1)}} \left|\left(r_n^{OGA\,\mathcal{D}}(x), d\right)\right|,$$

$$G_{n+1}^{OGA\,\mathcal{D}}(x) = \mathrm{Proj}_{\left\{e_k^{OGA\,\mathcal{D}}(x)\right\}_{k=1}^{n+1}} x, \quad r_{n+1}^{OGA\,\mathcal{D}}(x) = x - G_{n+1}^{OGA\,\mathcal{D}}(x),$$

$$n = 0,\, 1,\, 2,\, \ldots\ .$$

Последовательность $\{G_n^{OGA\,\mathcal{D}}(x)\}_{n=0}^{\infty}$ называется *последовательностью ортогональных жадных приближений вектора x по системе возможно неполных словарей $\mathcal{D} = \{D_1, D_2, \ldots, D_k\}$*, элемент $G_n^{OGA\,\mathcal{D}}(x)$ — *n-м ортогональным жадным приближением x по системе возможно неполных словарей $\mathcal{D} = \{D_1, D_2, \ldots, D_k\}$*.

Данный алгоритм также не всегда реализуем и не всегда определен однозначно.

Чисто жадное разложение вектора по паре словарей D_1 и D_2 является частным случаем чисто жадного разложения вектора по системе возможно неполных словарей при $k = 2$ и $i(1) = 1$. Для вектора $x \in H$ индуктивно определяются *последовательность остатков* $\left\{r_n^{PGA(D_1,D_2)}(x)\right\}_{n=0}^{\infty}$, *последовательность коэффициентов* $\left\{\hat{x}_n^{PGA(D_1,D_2)}\right\}_{n=1}^{\infty}$ и *последовательность элементов словарей* $\left\{e_n^{PGA(D_1,D_2)}(x)\right\}_{n=1}^{\infty}$. При этом на нечётных шагах чисто жадного алгоритма в качестве словаря берётся словарь D_1, а на чётных шагах — словарь D_2. Аналогично, *чисто жадным разложением по паре словарей D_1 и D_2* является ряд $\sum\limits_{n=1}^{\infty} \hat{x}_n^{PGA(D_1,D_2)} e_n^{PGA(D_1,D_2)}(x)$.

Частным случаем ортогонального жадного разложения вектора по системе возможно неполных словарей при $k = 2$ и $i(1) = 1$ является **ортогональный жадный алгоритм по паре словарей D_1 и D_2**. Для вектора $x \in H$ аналогично определяется $\left\{G_n^{OGA(D_1,D_2)}(x)\right\}_{n=0}^{\infty}$ — *последовательность ортогональных жадных приближений вектора x по паре словарей D_1 и D_2*, а также *последовательность остатков* $\left\{r_n^{OGA(D_1,D_2)}(x)\right\}_{n=0}^{\infty}$ и *последовательность элементов словарей* $\left\{e_n^{OGA(D_1,D_2)}(x)\right\}_{n=1}^{\infty}$.

Жадные алгоритмы по словарям, которые не являются нормированными.

В случае, когда рассматривается жадный алгоритм по словарю D, не являющемуся нормированным (т. е. линейная оболочка D всюду плотна в H и все элементы D имеют ненулевую норму в H), будем считать, что рассматривается соответствующий жадный алгоритм по нормированному словарю $\tilde{D}$, который однозначно получается из словаря D нормированием элементов.

Множество $\mathcal{A}_1(D)$ приближаемых словарём D векторов из гильбертова пространства H.

Для произвольного нормированного словаря D и положительного числа M определим *множество* $\mathcal{A}_1^0(D, M)$ следующим образом:

$$\mathcal{A}_1^0(D, M) = \left\{ x \in H \mid x = \sum_{i=1}^{N} c_i d_i,\ N \in \mathbb{N},\ d_i \in D\ \forall i \in \mathbb{N} \ \text{ и } \ \sum_{i=1}^{N} |c_i| \leqslant M \right\}.$$

Множеством $\mathcal{A}_1(D, M)$ назовём замыкание $\mathcal{A}_1^0(D, M)$ в H. Далее, определим *множество $\mathcal{A}_1(D)$ приближаемых словарём D векторов из гильбертова пространства H* [7] как объединение $\cup_M \mathcal{A}_1(D, M)$.

Для вектора $x \in \mathcal{A}_1(D)$ положим

$$\|x\|_{\mathcal{A}_1(D)} = \inf_{x \in \mathcal{A}_1(D,M)} M.$$

Скорость сходимости.

Будем говорить, что алгоритм на векторе x (гарантированно) *сходится за конечное число шагов*, если каждая реализация данного алгоритма сходится за конечное число шагов. Будем говорить, что алгоритм на векторе x (заведомо) *не сходится за конечное число шагов*, если никакая реализация данного алгоритма не сходится за конечное число шагов. Если на векторе x один алгоритм сходится за конечное число шагов, а другой не сходится за конечное число шагов, то будем считать, что на векторе x первый алгоритм *быстрее*, чем второй (соответственно, второй алгоритм *медленнее*, чем первый).

Сходимость алгоритма по конкретному словарю D за конечное число шагов возможна лишь для достаточно узкого класса векторов. Конкретнее, сходимость алгоритма по словарю D или по паре словарей D_1 и D_2 за конечное число шагов возможна лишь для векторов из $< D >$ или $< D_1 \cup D_2 >$, соответственно. С учётом этого введём понятие финитного вектора: *финитным вектором x относительно словаря D* будем называть вектор x, который является конечной линейной комбинацией элементов словаря D, финитный вектор x относительно стандартного ортогонального словаря F пространства ℓ_2 будем называть *финитным*.

Если остаток на каждом шаге ненулевой (т.е. нет сходимости за конечное число шагов), рассматривается понятие скорости сходимости, которое для конкретного вектора $x \in \mathcal{A}_1(D)$ можно охарактеризовать последовательностью

$$c_n(x) = \frac{\|r_n(x)\|_2}{\|x\|_{\mathcal{A}_1(D)}}, \quad \text{где} \quad \|a\|_2 = \sqrt{(a, a)}.$$

Если алгоритм реализуется по паре словарей D_1 и D_2, то под $\|x\|_{\mathcal{A}_1(D)}$ здесь будем понимать $\|x\|_{\mathcal{A}_1(D_1, D_2)} = \max\{\|x\|_{\mathcal{A}_1(D_1)}, \|x\|_{\mathcal{A}_1(D_2)}\}$. Для вектора x и вектора αx такие алгоритмы будут эквивалентны, а последовательности будут совпадать, поскольку остатки также будут отличаться на коэффициент α.

В случае отсутствия сходимости за конечное число шагов будем говорить, что первый алгоритм *быстрее*, чем второй (соответственно, второй алгоритм *медленнее*, чем первый) на векторе x, если

$$\lim_{n \to \infty} \frac{c_n(x)}{\tilde{c}_n(x)} = 0,$$

где последовательность $c_n(x)$ соответствует первому алгоритму, а последовательность $\tilde{c}_n(x)$ — второму.

Для некоторого класса векторов $K \subset \mathcal{A}_1(D)$ можно ввести последовательность, аналогичную последовательности $c_n(x)$, следующим образом:

$$c_n(K) = \sup_{x \in K \setminus \{0\}} \frac{\|r_n(x)\|_2}{\|x\|_{\mathcal{A}_1(D)}}.$$

В частности, если $K = \mathcal{A}_1(D)$, то будем обозначать $c_n(K)$ через c_n.

Величина $c_n(K)$ характеризует наибольшее возможное отклонение жадного приближения от приближаемого вектора на n-м шаге для элементов из класса K.

В работе изучается последовательность $c_n(x)$ для индивидуального вектора x и последовательность c_n для класса $\mathcal{A}_1(D)$ для различных жадных алгоритмов.

Рассматриваемый случай.

В данной работе будут рассматриваться бесконечномерные сепарабельные гильбертовы пространства, поэтому, в силу изоморфизма данных пространств, можно считать, что $H = \ell_2$.

Для удобства выкладок будем обозначать стандартный базис для ℓ_2 через $F = \{f_n\}_n$, то есть f_n — это последовательность, у которой на n-м месте 1, а на остальных — 0.

В случае пространства $H = \ell_2$, если F — стандартный базис ℓ_2, то множеством приближаемых векторов $\mathcal{A}_1(F)$ является подпространство ℓ_1.

Глава 2

Скорость сходимости слабых жадных алгоримов по ортогональным словарям.

Результаты данной главы изложены в статье автора [34].

Поскольку на ортогональных словарях слабый жадный алгоритм и слабый ортогональный жадный алгоритм эквивалентны, то при обсуждении скорости сходимости жадных приближений по ортогональным словарям корректнее осуществлять сравнение с общими результатами именно для ортогонального жадного алгоритма.

Общий результат о скорости сходимости слабых ортогональных жадных приближений [30] был приведён во введении (Теорема H). В случае ортогонального словаря данная оценка скорости сходимости может быть уточнена. А именно, справедлива следующая

Теорема 2.1. *Для ортогонального словаря D верно*

$$c_n^{WOGA} \leqslant \frac{1}{2\sqrt{\sum\limits_{k=1}^{n} t_k}}, \quad \forall n \in \mathbb{N}.$$

Доказательство.

Без ограничения общности будем считать, что словарь D совпадает со стандартным словарём F. Зафиксируем произвольное натуральное n и произвольный приближаемый вектор $x = (x_1, x_2, x_3, \ldots) \in \ell_1 \subset H$

и докажем, что

$$\frac{\|r_n^{WOGA}(x)\|_2}{\|x\|_1} \leqslant \frac{1}{2\sqrt{\sum\limits_{k=1}^{n} t_k}}.$$

Для доказательства требуемой оценки достаточно рассмотреть случай, когда $r_n^{WOGA}(x) \neq 0$. В этом случае на каждом из первых n шагов слабого ортогонального жадного алгоритма при переходе к очередному остатку осуществляется замена одной из ненулевых компонент предыдущего остатка нулем. Не ограничивая общности можно считать, что на j-м шаге нулем заменяется j-я компонента ($j = 1, 2, \ldots, n$) — этого можно добиться переупорядочением координат. Тогда остаток $r_n^{WOGA}(x)$ равен $(0, 0, \ldots, 0, x_{n+1}, x_{n+2}, \ldots)$. Также без ограничения общности можно считать, что все компоненты вектора x — неотрицательные числа.

В силу однородности устанавливаемой оценки (т. е. в силу того, что при умножении приближаемого вектора x на положительное число на это же число одновременно умножаются числитель и знаменатель в левой части оценки) при фиксированном n достаточно ограничиться рассмотрением случая $\left\|r_n^{WOGA}(x)\right\|_\infty = 1$. При таком ограничении условия выбора следующего вектора словаря (1.1) гарантируют, что для всех $j \in \{1, 2, \ldots, n\}$ справедлива оценка $x_j \geqslant t_j$.

Для удобства дальнейших выкладок будем использовать обозначение $T_n = \sum\limits_{k=1}^{n} t_k$.

Обозначим через y "хвост" n-го остатка (совпадающий с n-м "хвостом" приближаемого вектора x): $y = (x_{n+1}, x_{n+2}, x_{n+3}, \ldots)$. Заметим, что

$$\frac{\left\|r_n^{WOGA}(x)\right\|_2}{\|x\|_1} = \frac{\|y\|_2}{\|y\|_1 + \sum\limits_{k=1}^{n} x_k} \leqslant \frac{\|y\|_2}{\|y\|_1 + \sum\limits_{k=1}^{n} t_k} = \frac{\|y\|_2}{\|y\|_1 + T_n}.$$

Так как все компоненты вектора y принадлежат отрезку $[0, 1]$, справедливо неравенство

$$\|y\|_2 = \sqrt{\sum_{k=n+1}^{\infty} x_k^2} \leqslant \sqrt{\sum_{k=n+1}^{\infty} x_k} = \sqrt{\|y\|_1},$$

с учетом которого получаем,

$$\frac{\left\|r_n^{WOGA}(x)\right\|_2}{\|x\|_1} \leqslant \frac{\sqrt{\|y\|_1}}{\|y\|_1+T_n} = f\left(\|y\|_1\right),$$

где $f(\alpha)=\frac{\sqrt{\alpha}}{\alpha+T_n}$, $T_n>0$. При этом $\|y\|_1 \geqslant 1$. Так как

$$f'(\alpha)=\frac{T_n-\alpha}{2\sqrt{\alpha}(\alpha+T_n)^2},$$

то для $T_n<1$ функция f принимает наибольшее значение на луче $[1,+\infty)$ при значении аргумента, равном единице, а для $T_n \geqslant 1$ — при значении аргумента, равном T_n. Тогда в первом случае

$$\frac{\left\|r_n^{WOGA}(x)\right\|_2}{\|x\|_1} \leqslant f(1)=\frac{1}{1+T_n} \leqslant \frac{1}{2\sqrt{T_n}},$$

а во втором случае

$$\frac{\left\|r_n^{WOGA}(x)\right\|_2}{\|x\|_1} \leqslant f\left(T_n\right)=\frac{\sqrt{T_n}}{T_n+T_n}=\frac{1}{2\sqrt{T_n}},$$

т. е. доказываемая оценка справедлива в каждом из случаев.

Таким образом, *теорема доказана.*

Как легко видеть, в случае t_n, стремящихся к нулю, это уточнение является существенным. Например, в случае $t_n=\frac{1}{\sqrt{n}}$ общий результат гарантирует, что порядок скорости сходимости не хуже $\frac{1}{\sqrt{\ln n}}$, в то время как результат для ортогональных систем гарантирует, что порядок скорости сходимости не хуже $\frac{1}{\sqrt[4]{n}}$.

Полученная для ортогональных систем оценка асимптотически неулучшаема при $\sum\limits_{k=1}^{\infty} t_k=+\infty$. Так, справедливо следующее утверждение.

Теорема 2.2. *Пусть для ослабляющей последовательности* $\{t_n\}_{n=1}^{\infty}$ *выполняется условие* $\sum\limits_{k=1}^{\infty} t_k = +\infty$. *Тогда для ортогонального словаря* D *имеет место эквивалентность*

$$c_n^{WOGA} \sim \frac{1}{2\sqrt{\sum\limits_{k=1}^{n} t_k}}, \quad n \to \infty.$$

Доказательство.

Для доказательства теоремы достаточно показать, что при всех натуральных n величина c_n^{WOGA} заключена между $\frac{\beta_n}{2\sqrt{T_n}}$ и $\frac{1}{2\sqrt{T_n}}$, где $T_n = \sum\limits_{k=1}^{n} t_k$, а $\beta_n = \left(1 - (1+2T_n)^{-1}\right)$: легко видеть, что $\beta_n \to 1$ при $n \to \infty$.

Верхняя оценка $c_n \leqslant \frac{1}{2\sqrt{T_n}}$ сразу следует из установленной теоремы 2.1. Для доказательства нижней оценки $c_n \geqslant \frac{\beta_n}{2\sqrt{T_n}}$, где $\beta_n = (1 - (1+2T_n)^{-1})$, достаточно рассмотреть вектор

$$x = (t_1, t_2, ..., t_n, \underbrace{1, 1, ..., 1}_{1+[T_n]}, \{T_n\}, 0, 0, ...).$$

Здесь квадратными скобками обозначена целая часть, а фигурными — дробная часть. Возможной является реализация слабого ортогонального жадного алгоритма, при которой на первом шаге в качестве элемента разложения выбирается первый базисный вектор, на втором шаге — второй базисный вектор и т. д. Для этой реализации

$$r_n^{WOGA}(x) = (\underbrace{0, 0, ..., 0}_{n}, \underbrace{1, 1, ..., 1}_{1+[T_n]}, \{T_n\}, 0, 0, ...).$$

Соответственно

$$\frac{\left\|r_n^{WOGA}(x)\right\|_2}{\|x\|_1} = \frac{\sqrt{1+[T_n]+\{T_n\}^2}}{T_n+1+[T_n]+\{T_n\}} = \frac{\sqrt{1+[T_n]+\{T_n\}^2}}{1+2T_n} \geqslant$$

$$\geqslant \frac{\sqrt{T_n}}{1+2T_n} = \frac{1}{2\sqrt{T_n}}\left(1 - \frac{1}{1+2T_n}\right).$$

Теорема доказана.

Глава 3

Сходимость слабых жадных алгоритмов по одноэлементным расширениям ортогональных словарей.

Основные результаты этой главы опубликованы в статье автора [35].

Условием на ослабляющую последовательность $\{t_n\}_{n=1}^{\infty}$, гарантирующим сходимость слабых ортогональных жадных приближений к приближаемому вектору, является расходимость ряда $\sum\limits_{k=1}^{\infty} t_k^2$ [30]. Причём при отсутствии дополнительных ограничений на словарь и на приближаемые векторы ослабить это условие нельзя. Однако в случае ортогонального словаря $D = \{d_n\}_{n\in\mathbb{N}}$ и приближения векторов из подпространства $\mathcal{A}_1(D)$ сходимость гарантируется при выполнении более слабого условия расходимости ряда $\sum\limits_{k=1}^{\infty} t_k$. Это сразу следует из теоремы 2.1. При этом дальнейшее ослабление условия невозможно. Чтобы это обосновать, достаточно для произвольной ослабляющей последовательности $\{t_n\}_{n=1}^{\infty}$ с $\sum\limits_{k=1}^{\infty} t_k < +\infty$ рассмотреть слабые ортогональные жадные приближения вектора $x = d_1 + \sum\limits_{n=1}^{\infty} t_n d_{n+1} \in \mathcal{A}_1(D)$. Возможна реализация слабого ортогонального жадного алгоритма, при которой ни на одном шаге в качестве элемента разложения не выбирается первый базисный вектор (на первом шаге выбирается второй базисный вектор, на втором шаге — третий базисный вектор, на третьем шаге — четвертый базисный вектор и т. д.). Соответственно, для этой реализации $\left\|r_n^{WOGA}(x)\right\|_2 \geqslant 1$ и алгоритм не сходится к приближаемому вектору.

Расширение словаря влияет на реализацию жадных алгоритмов, но при этом увеличивает или, как минимум, не ухудшает их общие аппроксимационные свойства. Тем не менее аналогичным образом ослабить общее достаточное условие сходимости слабых жадных разложений для словарей, являющихся расширениями ортогональных, нельзя. Так, например, для единичной сферы в качестве "экстремального" словаря ($D = S$) общее достаточное условие сходимости будет неослабляемо по сравнению с расходимостью ряда $\sum\limits_{n=1}^{\infty} t_n^2$. Более точно, верно следующее утверждение.

Теорема 3.1. *Для любой ослабляющей последовательности $t \in \ell_2$ и любого ненулевого вектора $x \in H$ существуют реализации*

1) *слабого ортогонального жадного алгоритма*

2) *слабого жадного алгоритма*

с ослабляющей последовательностью t по словарю, равному единичной сфере, которые не сходятся к приближаемому вектору x.

Необходимость расходимости ряда $\sum\limits_{n=1}^{\infty} t_n^2$ для гарантированной сходимости слабых ортогональных жадных приближений к приближаемому элементу была также отмечена в работе [30] (Замечание 2), однако в другом контексте, не покрывающем обсуждаемый здесь случай единичной сферы как словаря.

Доказательство.

Будем одновременно строить реализацию слабого жадного алгоритма и слабого ортогонального жадного алгоритма.

Заметим, что в случае словаря $D = S$ для любого вектора $r \in H$ верно $\sup\limits_{d \in D} |(r, d)| = \|r\|_2$.

Выберем в качестве приближающего вектора $e_1^{W(O)GA}(x)$ такой единичный вектор, что

$$|(x, e_1^{W(O)GA}(x))| = t_1 \sup_{d \in D} |(x, d)| = t_1 \|x\|_2.$$

На каждом следующем шаге разложения в качестве вектора $e_n^{W(O)GA}(x)$ при $n > 1$ выберем единичный вектор, ортогональный уже выбранным векторам $e_1^{W(O)GA}(x), \ldots, e_{n-1}^{W(O)GA}(x)$, такой, что

$$|(r_{n-1}^{W(O)GA}(x), e_n^{W(O)GA}(x))| = t_n \sup_{d\in D} |(r_{n-1}^{W(O)GA}(x), d)| = t_n \|r_{n-1}^{W(O)GA}(x)\|_2.$$

Такие векторы всегда будут существовать в словаре $D = S$.

Для данной реализации слабого (ортогонального) жадного алгоритма справедливы равенства

$$\|r_n^{W(O)GA}(x)\|_2^2 = \|r_{n-1}^{W(O)GA}(x)\|_2^2 - |(r_{n-1}^{W(O)GA}(x), e_n^{W(O)GA}(x))|^2 =$$

$$= \|r_{n-1}^{W(O)GA}(x)\|_2^2 - (t_n \sup_{d\in D} |(r_{n-1}^{W(O)GA}(x), d)|)^2 = (1 - t_n^2)\|r_{n-1}^{W(O)GA}(x)\|_2^2 =$$

$$= \prod_{i=1}^{n} (1 - t_i^2)\|x\|_2^2.$$

Отсюда

$$\lim_{n\to\infty} \|r_n^{W(O)GA}(x)\|_2^2 = \|x\|_2^2 \prod_{i=1}^{\infty} (1 - t_i^2).$$

Сходимость данного бесконечного произведения к ненулевому значению эквивалентна сходимости ряда $\sum\limits_{i=1}^{\infty} t_i^2$. То есть в случае $t \in \ell_2$ верно неравенство $\lim\limits_{n\to\infty} \|r_n^{W(O)GA}(x)\|_2^2 > 0$, и данная реализация слабого (ортогонального) жадного алгоритма с ослабляющей последовательностью t по словарю $D = S$ не сходится к приближаемому вектору x.

Теорема доказана.

Возникает вопрос, где при расширении ортогонального словаря наступает переход от возможности ослабления общего достаточного условия сходимости при приближении элементов из соответствующего класса $\mathcal{A}_1$ к невозможности: например, можно ли хотя бы в случае словарей, получающихся из ортогонального добавлением ровно одного элемента, ослабить достаточное условие с расходимости ряда $\sum\limits_{n=1}^{\infty} t_n^2$ до расходимости ряда $\sum\limits_{n=1}^{\infty} t_n$. Оказывается, ответ на этот вопрос — отрицательный: добавление даже одного вектора (если этот вектор “плохой”) уже достаточно для невозможности такого ослабления. Так, для слабого ортогонального жадного алгоритма получена следующая теорема.

Теорема 3.2. *Для любой ослабляющей последовательности $\tau \in \ell_2 \setminus \ell_1$ существуют вектор $g \in \ell_2$ и вектор $x \in \ell_1$ такие, что для x существует реализация слабого ортогонального жадного алгоритма с ослабляющей последовательностью $t = c\tau$ ($c = \frac{\tau_1}{\|\tau\|_2^2}$, где $\tau = \{\tau_n\}_{n=1}^{\infty}$) по словарю, полученному из стандартного ортогонального добавлением вектора g, которая не сходится к приближаемому вектору x.*

Доказательство.

Зафиксируем ослабляющую последовательность $\tau = \{\tau_n\} \in \ell_2 \setminus \ell_1$ и в качестве g выберем вектор $(0, g_1, g_2, g_3, ...)$, где $g_n = \dfrac{\tau_n}{\|\tau\|_2}$ при $n \in \mathbb{N}$. Это единичный вектор из $\ell_2 = H$. В качестве ослабляющей последовательности для слабого ортогонального жадного алгоритма выберем последовательность $t = \dfrac{\tau_1}{\|\tau\|_2^2}\tau$, тогда $t_n = \tau_n \dfrac{\tau_1}{\|\tau\|_2^2} = g_1 g_n$ при $n \in \mathbb{N}$ и $\{t_n\} \subset (0, 1]$.

Для удобства дальнейших выкладок определим векторы d_n при $n \in \mathbb{N}$ как $(0, g_1, \underbrace{0, \ldots, 0}_{n-1}, g_{n+1}, g_{n+2}, \ldots)$, и положим $D_n = ||d_n||_2$ при $n \in \mathbb{N}$.

Теперь рассмотрим слабый ортогональный жадный алгоритм с ослабляющей последовательностью $t = \frac{\tau_1}{\|\tau\|_2^2}\tau$ по словарю $D = F \cup \{g\}$ для вектора $x = (1, 1, 0, 0, 0, ...)$ и докажем индуктивно существование реализации слабого ортогонального жадного алгоритма, для которой при $n \geqslant 1$ верно

$$r_n^{WOGA}(x) = \left(1, 1 - g_1 \frac{g_1}{D_n^2}, \underbrace{0, \ldots, 0}_{n-1}, -g_{n+1}\frac{g_1}{D_n^2}, -g_{n+2}\frac{g_1}{D_n^2}, \ldots\right). \tag{3.1}$$

Шаг 1. Имеют место равенства

$$r_0^{WOGA}(x) = x,$$

$$\sup_{d \in D} |(r_0^{WOGA}(x), d)| = \max\left(\sup_{f \in F} |(r_0^{WOGA}(x), f)|, |(r_0^{WOGA}(x), g)|\right) =$$
$$= \max(1, g_1) = 1.$$

Тогда

$$t_1 \sup_{d \in D} |(r_0^{WOGA}(x), d)| = t_1 = g_1^2 \leqslant g_1 = |(r_0^{WOGA}(x), g)|$$

и, значит, можно выбрать в качестве приближающего вектора $e_1^{WOGA}(x)$ вектор g. В этом случае

$$G_1^{WOGA}(x) = \mathrm{Proj}_{<e_1^{WOGA}(x)>} x = (x, g)g = g_1 d_1,$$

$$r_1^{WOGA}(x) = \left(1, 1 - g_1^2, -g_1 g_2, -g_1 g_3, \dots\right).$$

Шаг 2. В силу равенства $(r_1^{WOGA}(x), g) = 0$ верно

$$\sup_{d \in D} |(r_1^{WOGA}(x), d)| = \sup_{f \in F} |(r_1^{WOGA}(x), f)| = 1,$$

откуда получаем, что

$$t_2 \sup_{d \in D} |(r_1^{WOGA}(x), d)| = t_2 = g_1 g_2 = |(r_1^{WOGA}(x), f_3)|,$$

и в качестве $e_2^{WOGA}(x)$ можно выбрать вектор f_3. Тогда

$$G_2^{WOGA}(x) = \mathrm{Proj}_{<e_1^{WOGA}(x), e_2^{WOGA}(x)>} x = \mathrm{Proj}_{<d_1, f_3>} x = \mathrm{Proj}_{<d_2, f_3>} x =$$

$$= \left(x, \frac{d_2}{D_2}\right) \frac{d_2}{D_2} + (x, f_3) f_3 = \frac{g_1}{D_2^2} d_2,$$

$$r_2^{WOGA}(x) = \left(1, 1 - g_1 \frac{g_1}{D_2^2}, 0, -g_3 \frac{g_1}{D_2^2}, -g_4 \frac{g_1}{D_2^2}, \dots\right).$$

Шаг $n > 2$: в силу предположения индукции $r_{n-1}^{WOGA}(x)$ имеет вид

$$\left(1, 1 - g_1 \frac{g_1}{D_{n-1}^2}, \underbrace{0, \dots, 0}_{n-2}, -g_n \frac{g_1}{D_{n-1}^2}, -g_{n+1} \frac{g_1}{D_{n-1}^2}, \dots\right).$$

Тогда, поскольку $(r_{n-1}^{WOGA}(x), g) = 0$,

$$\sup_{d \in D} |(r_{n-1}^{WOGA}(x), d)| = \sup_{f \in F} |(r_{n-1}^{WOGA}(x), f)| = 1,$$

$$t_n \sup_{d \in D} |(r_{n-1}^{WOGA}(x), d)| = t_n = g_1 g_n \leqslant g_n \frac{g_1}{D_{n-1}} = |(r_{n-1}^{WOGA}(x), f_{n+1})|.$$

Значит, можно выбрать $e_n^{WOGA}(x) = f_{n+1}$. В этом случае

$$G_n^{WOGA}(x) = \mathrm{Proj}_{<e_1^{WOGA}(x), e_2^{WOGA}(x), \dots, e_n^{WOGA}(x)>} x = \mathrm{Proj}_{<d_1, f_3, \dots, f_{n+1}>} x =$$

$$= \mathrm{Proj}_{<d_n, f_3, \dots, f_{n+1}>} x = \left(x, \frac{d_n}{D_n} \right) \frac{d_n}{D_n} = \frac{g_1}{D_n^2} d_n,$$

$$r_n^{WOGA}(x) = \left(1, 1 - g_1 \frac{g_1}{D_n^2}, \underbrace{0, \dots, 0}_{n-1}, -g_{n+1} \frac{g_1}{D_n^2}, -g_{n+2} \frac{g_1}{D_n^2}, \dots \right).$$

Таким образом, индукция завершена, и равенство (3.1) доказано. Соответственно, для этой реализации на каждом шаге $||r_n^{WOGA}(x)||_2 > 1$, и алгоритм не сходится к приближаемому элементу (более точно, $r_n^{WOGA}(x) \to f_1$ при $n \to \infty$).

Теорема доказана.

Замечание. При $c \leqslant 1$ (это выполняется, например, при $\tau_1 = 1$ независимо от остальных τ_n) разложение с ослабляющей последовательностью t автоматически является разложением с ослабляющей последовательностью τ.

Теорема 3.2 показывает невозможность ослабления достаточного условия сходимости при добавлении к ортогональному словарю "плохого" вектора. Однако при добавлении финитного вектора условие сходимости, ослабленное до расходимости ряда $\sum\limits_{n=1}^{\infty} t_n$, остаётся достаточным.

Теорема 3.3. *Для любых финитного вектора g, ослабляющей последовательности $t \notin \ell_1$ и вектора $x \in \mathcal{A}_1(F)$ любая реализация слабого ортогонального жадного алгоритма с ослабляющей последовательностью t по словарю, полученному из стандартного ортогонального добавлением вектора g, сходится к приближаемому вектору x.*

Замечание. В случае финитного вектора g имеет место соотношение $\mathcal{A}_1(F) = \mathcal{A}_1(F \cup \{g\})$.

Доказательство теоремы 3.3.

В силу финитности g лишь конечный набор векторов $f_k \in F$ неортогонален g. Тогда возьмем N достаточно большим, чтобы все f_k из этого набора, выбираемые в качестве элементов разложения при реализации слабого ортогонального жадного алгоритма для элемента x, выбирались на шагах с номерами, меньшими N. Иными словами, зафиксируем N

такое, что

$$N > \max\{n \in \mathbb{N} \mid \exists k \in \mathbb{N}\ f_k \not\perp g,\ e_n^{WOGA}(x) = f_k\}.$$

Также будем считать, что $N > N'$, если g выбирается в качестве вектора разложения $e_{N'}^{WOGA}(x)$.

Тогда для $n \geqslant N$ получаем следующее условие на вектор $e_{n+1}^{WOGA}(x)$:

$$\left|\left(r_n^{WOGA}(x), e_{n+1}^{WOGA}(x)\right)\right| \geqslant t_{n+1} \sup_{d \in D} \left|\left(r_n^{WOGA}(x), d\right)\right| \geqslant$$

$$\geqslant t_{n+1} \sup_{d \in F} \left|\left(r_n^{WOGA}(x), d\right)\right|.$$

Следовательно, данная реализация слабого ортогонального жадного алгоритма с ослабляющей последовательностью $\{t_n\}_{n=1}^{\infty}$ по расширенному словарю $F \cup \{g\}$ для вектора $x \in \mathcal{A}_1(F)$ также является реализацией слабого ортогонального жадного алгоритма с ослабляющей последовательностью $\{t_{n+N}\}_{n=1}^{\infty}$ по ортогональному словарю F для вектора $r_N^{WOGA}(x) \in \mathcal{A}_1(F)$, а значит, сходится к приближаемому вектору.

Таким образом, *теорема доказана.*

Замечание. Аналогично можно доказать следующее обобщение теоремы 3.3.

Теорема 3.3*. *Для любых финитных векторов $g^1, \ldots, g^n$, ослабляющей последовательности $t \notin \ell_1$ и вектора $x \in \mathcal{A}_1(F)$ любая реализация слабого ортогонального жадного алгоритма с ослабляющей последовательностью t по словарю, полученному из стандартного ортогонального добавлением векторов $g^1, \ldots, g^n$, сходится к приближаемому вектору x.*

Замечание. Для финитных векторов $g^1, \ldots, g^n$ имеет место равенство $\mathcal{A}_1(F) = \mathcal{A}_1(F \cup \{g^1, \ldots, g^n\})$.

Аналогичные результаты можно получить и для слабого жадного алгоритма. Если словарь получен из ортогонального добавлением произвольного вектора, ослабить достаточное условие сходимости (так же, как и в случае ортогонального словаря) нельзя. А именно, справедлива следующая теорема.

Теорема 3.4. *Для любой ослабляющей последовательности* $\tau \in \ell_2 \setminus \ell_1$ *существуют вектор* $g \in \ell_2$ *и вектор* $x \in \ell_1$ *такие, что для* x *существует реализация слабого жадного алгоритма с ослабляющей последовательностью* $t = c\tau$ ($c = \frac{\tau_1}{\|\tau\|_2^2}$, *где* $\tau = \{\tau_n\}_{n=1}^{\infty}$) *по словарю, полученному из стандартного ортогонального добавлением вектора* g, *которая не сходится к приближаемому вектору* x.

Доказательство.

Аналогично доказательству теоремы 3.2, зафиксируем ослабляющую последовательность $\tau = \{\tau_n\} \in \ell_2 \setminus \ell_1$. Также выберем в качестве g вектор $(0, g_1, g_2, g_3, ...)$, где $g_n = \dfrac{\tau_n}{\|\tau\|_2}$ при $n \in \mathbb{N}$, — это единичный вектор из $\ell_2 = H$.

Ещё раз отметим, что если $t = \tau \dfrac{\tau_1}{\|\tau\|_2^2}$, то $t_n = \tau_n \dfrac{\tau_1}{\|\tau\|_2^2} = g_1 g_n$ при $n \in \mathbb{N}$ и $\{t_n\} \subset (0, 1]$.

Рассмотрим слабый жадный алгоритм с ослабляющей последовательностью $\{t_n\}$ по словарю $D = F \cup \{g\}$ для вектора $x = (1, 1, 0, 0, 0, ...)$ и докажем индуктивно существование реализации жадного алгоритма, для которой

$$r_n^{WGA}(x) = (1, 1 - g_1^2, \underbrace{0, ..., 0}_{n-1}, -g_1 g_{n+1}, -g_1 g_{n+2}, \dots). \tag{3.2}$$

Шаг 1. Поскольку справедливы равенства

$$r_0^{WGA}(x) = x,$$

$$\sup_{d \in D} |(r_0^{WGA}(x), d)| = \max\left(\sup_{f \in F} |(r_0^{WGA}(x), f)|, |(r_0^{WGA}(x), g)|\right) = \\ = \max(1, g_1) = 1,$$

то

$$t_1 \sup_{d \in D} |(r_0^{WGA}(x), d)| = t_1 = g_1^2 \leqslant g_1 = |(r_0^{WGA}(x), g)|,$$

и на шаге 1 можно выбрать в качестве элемента разложения $e_1^{WGA}(x)$ вектор g. Тогда

$$r_1^{WGA}(x) = \left(1, 1 - g_1^2, -g_1 g_2, -g_1 g_3, ...\right).$$

Шаг 2. В силу соотношения

$$\sup_{d\in D} |(r_1^{WGA}(x), d)| = \sup_{f\in F} |(r_1^{WGA}(x), f)| = 1$$

имеет место равенство

$$t_2 \sup_{d\in D} |(r_1^{WGA}(x), d)| = t_2 = g_1 g_2 = |(r_1^{WGA}(x), f_3)|,$$

и здесь можно выбрать вектор $e_2^{WGA}(x) = f_3$. В этом случае

$$r_2^{WGA}(x) = \left(1, 1 - g_1^2, 0, -g_1 g_3, -g_1 g_4, ...\right).$$

Шаг $n > 2$. По предположению индукции

$$r_{n-1}^{WGA}(x) = (1, 1 - g_1^2, 0, ..., 0, -g_1 g_n, -g_1 g_{n+1}, \dots).$$

Значит, выполняется равенство

$$\sup_{d\in D} |(r_{n-1}^{WGA}(x), d)| = \max\left(\sup_{f\in F} |(r_{n-1}^{WGA}(x), f)|, |(r_{n-1}^{WGA}(x), g)|\right) =$$

$$= \max\left(1, \left|(1 - g_1^2)g_1 + \sum_{k=n}^{\infty}(-g_1 g_k)g_k\right|\right) = \max\left(1, g_1\left|1 - (g_1^2 + \sum_{k=n}^{\infty} g_k^2)\right|\right).$$

Поскольку $g_1^2 + \sum\limits_{k=n}^{\infty} g_k^2 < \sum\limits_{k=1}^{\infty} g_k^2 = \|g\|^2 = 1$, верна оценка

$$\left|1 - (g_1^2 + \sum_{k=n}^{\infty} g_k^2)\right| < 1$$

и второй аргумент max меньше $g_1 < 1$. Таким образом, получаем, что $\sup\limits_{d\in D} |(r_{n-1}^{WGA}(x), d)| = 1$. Следовательно,

$$t_n \sup_{d\in D} |(r_{n-1}^{WGA}(x), d)| = t_n = g_1 g_n = |(r_{n-1}^{WGA}(x), f_{n+1})|,$$

и, значит, в качестве $e_n^{WGA}(x)$ можно выбрать вектор f_{n+1}. Тогда

$$r_n^{WGA}(x) = (1, 1 - g_1^2, \underbrace{0, ..., 0}_{n-1}, -g_1 g_{n+1}, -g_1 g_{n+2}, \dots).$$

Итак, индукция завершена и равенство (3.2) доказано. На каждом шаге для данной реализации верно $||r_n^{WGA}(x)||_2 > 1$ и разложение не сходится к разлагаемому элементу (более точно, $r_n^{WGA}(x) \to f_1$ при $n \to \infty$).

Теорема 3.4 доказана.

Следует также отметить, что добавление к ортогональному словарю нефинитного вектора, даже вектора из класса $\ell_1 = \mathcal{A}_1(F)$, может значительно ухудшить скорость сходимости алгоритма.

Теорема 3.5. *Существуют вектор $g \in \ell_1$ и финитный вектор x такие, что любая реализация чисто жадного алгоритма по словарю, полученному из стандартного ортогонального добавлением вектора g, не сходится к приближаемому вектору x за конечное число шагов.*

Доказательство.

Сначала опишем общую идею доказательства. Если вектор $g \in \ell_1$ выбирается в качестве элемента разложения $e_n^{PGA}(x)$, например, на первом шаге, то задача разложения финитного вектора x трансформируется в разложение нефинитного относительно F вектора $r_1^{PGA}(x)$. Причём в случае, когда g выбирается в качестве $e_n^{PGA}(x)$ только один раз, остальные $e_n^{PGA}(x)$ являются элементами ортогонального словаря и реализация чисто жадного алгоритма не сходится к приближаемому вектору x за конечное число шагов.

Приведём теперь строгое доказательство.

Зафиксируем произвольное $\varepsilon \in (0, \frac{\sqrt{2}-1}{2})$, положим $Q = \frac{\frac{1}{2}-2\varepsilon-2\varepsilon^2}{1+2\varepsilon} > 0$ и выберем $h = \{h_n\}_n$ — последовательность из ℓ_1 такую, что выполняются следующие условия:

(1) $h_n > h_{n+1}$ при всех натуральных n,

(2) $||h||_2 = \sqrt{\frac{1}{2} - 2\varepsilon - 2\varepsilon^2} = \sqrt{(1+2\varepsilon)Q}$,

(3) $h_n > h_n^2 + h_{n+1}^2 + \ldots$ при всех натуральных n,

(4) $\max\{h_1^2, h_2\} < Q < h_1$.

Например, для $\varepsilon = \frac{1}{2}\left(\sqrt{\frac{61}{60}} - 1\right)$ можно выбрать в качестве h последовательность $2^{-\frac{n+1}{2}}$, из которой “выкинуты” члены вида $2^{-\frac{4k+1}{2}-3}$. Несложно проверить, что условия (1)–(4) выполняются.

Пусть $g = (\frac{1}{2}+\varepsilon, \frac{1}{2}+\varepsilon, h_1, ..., h_n, ...)$. Это единичный вектор из $\ell_2 = H$, принадлежащий ℓ_1. Рассмотрим чисто жадный алгоритм по словарю $D = F \cup \{g\}$ для вектора $x = \left(\frac{1}{1+2\varepsilon}, \frac{1}{1+2\varepsilon}, 0, 0, ...\right)$. По индукции докажем, что для любой реализации чисто жадного алгоритма будет верно

$$r_n^{PGA}(x) = (\underbrace{0, ..., 0}_{n-1}, -h_{n-2}, -h_{n-1}, -h_n, ...), \quad n > 4.$$

Шаг 1. Справедливы соотношения

$$r_0^{PGA}(x) = x,$$

$$\sup_{d \in D} |(r_0^{PGA}(x), d)| = \max\left(\sup_{f \in F} |(r_0^{PGA}(x), f)|, |(r_0^{PGA}(x), g)|\right) =$$

$$= \max\left(\frac{1}{1+2\varepsilon}, |(x, g)|\right) = \max\left(\frac{1}{1+2\varepsilon}, 1\right) = 1 = |(x, g)|.$$

Супремум достигается на единственном векторе — g. Тогда, $e_1^{PGA}(x) = g$ и

$$r_1^{PGA}(x) = \left(\frac{1}{1+2\varepsilon} - \left(\frac{1}{2}+\varepsilon\right), \frac{1}{1+2\varepsilon} - \left(\frac{1}{2}+\varepsilon\right), -h_1, -h_2, ...\right) =$$

$$= (Q, Q, -h_1, -h_2, ...).$$

Шаг 2. Верно равенство

$$\sup_{d \in D} |(r_1^{PGA}(x), d)| = \sup_{d \in F} |(r_1^{PGA}(x), d)| = \max(Q, h_1) = h_1 = |(r_1^{PGA}(x), f_3)|,$$

откуда однозначно получаем, что $e_2^{PGA}(x) = f_3$. Следовательно,

$$r_2^{PGA}(x) = (Q, Q, 0, -h_2, -h_3, ...).$$

Шаг 3. Имеет место равенство

$$\sup_{d \in D} |(r_2^{PGA}(x), d)| = \max\left(\sup_{d \in F} |(r_2^{PGA}(x), d)|, |(r_2^{PGA}(x), g)|\right) =$$

$$= \max\left(Q, h_2, 2 \cdot \left(\frac{1}{2}+\varepsilon\right) Q - \sum_{k=2}^{\infty} h_k^2\right) = \max\left(Q, h_2, h_1^2\right) = Q =$$

$$= |(r_1^{PGA}(x), f_1)| = |(r_1^{PGA}(x), f_2)|,$$

и в качестве приближающего вектора $e_3^{PGA}(x)$ можно выбрать вектор f_1 или f_2. Без ограничения общности можно считать, что $e_3^{PGA}(x) = f_1$. В этом случае

$$r_3^{PGA}(x) = (0, Q, 0, -h_2, -h_3, ...)\,.$$

Шаг 4. Верно равенство

$$\sup_{d\in D} |(r_3^{PGA}(x), d)| = \max\left(\sup_{d\in F} |(r_3^{PGA}(x), d)|, |(r_3^{PGA}(x), g)|\right) =$$

$$= \max\left(Q, h_2, \left|Q\left(\frac{1}{2}+\varepsilon\right) - \sum_{k=2}^{+\infty} h_k^2\right|\right) =$$

$$= \max\left(Q, h_2, \left|\frac{1}{2}\|h\|_2^2 - \left(\|h\|_2^2 - h_1^2\right)\right|\right) = \max\left(Q, h_2, \left|\frac{1}{2}\|h\|_2^2 - h_1^2\right|\right) =$$

$$= Q = |(r_2, f_2)|,$$

и, значит, $e_4^{PGA}(x) = f_2$. Тогда

$$r_4^{PGA}(x) = (0, 0, 0, -h_2, -h_3, ...).$$

Заметим, что независимо от выбора вектора $e_3^{PGA}(x)$ на шаге 3, остаток $r_4^{PGA}(x)$ определяется однозначно.

Шаг $n > 4$. По предположению индукции

$$r_{n-1}^{PGA}(x) = (\underbrace{0, ..., 0}_{n-2}, -h_{n-3}, -h_{n-2}, -h_{n-1}, ...).$$

Тогда верно соотношение

$$\sup_{d\in D} |(r_{n-1}^{PGA}(x), d)| = \max\left(\sup_{f\in F} |(r_{n-1}^{PGA}(x), f)|, |(r_{n-1}^{PGA}(x), g)|\right) =$$

$$= \max\left(h_{n-3}, \sum_{k=n-3}^{+\infty} h_k^2\right) = h_{n-3} = |(r_{n-1}^{PGA}(x), f_{n-1})|,$$

причём супремум достигается только на векторе f_{n-1}. Следовательно, получаем, что $e_n^{PGA}(x) = f_{n-1}$. Тогда

$$r_n^{PGA}(x) = (\underbrace{0, ..., 0}_{n-1}, -h_{n-2}, -h_{n-1}, -h_n, ...).$$

Таким образом, на каждом следующем шаге в качестве элемента разложения $e_n^{PGA}(x)$ выбирается вектор f_{n-1}, и алгоритм не сходится за конечное число шагов к приближаемому вектору.

Теорема доказана.

Глава 4

Сравнение стандартного жадного алгоритма и жадного алгоритма по паре словарей.

Часть результатов данной главы приведены в работе автора [36].

4.1 Сравнение чисто жадного алгоритма и чисто жадного алгоритма по паре словарей.

Если $D_1 \not\subseteq D_2$ и $D_2 \not\subseteq D_1$, то при разложении по паре словарей множество векторов, по которому происходит приближение, больше, чем при разложении по каждому из словарей по отдельности. С другой стороны, несмотря на то, что аппарат приближения в алгоритме по одному словарю оказывается у́же, чем в алгоритме по паре словарей, во втором случае присутствует некоторое дополнительное ограничение в выборе элемента словаря, накладываемое алгоритмом по паре словарей. Эти факты позволяют установить следующую теорему.

Теорема 4.1. *Существуют*

I. *ортонормированные словари* D_1 *и* D_2 *и вектор* $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ *такие, что*

  1) *чисто жадный алгоритм по словарю* D_1 *и чисто жадный алгоритм по словарю* D_2 *не сходятся за конечное число шагов,*

2) *чисто жадный алгоритм по паре словарей* D_1 *и* D_2 *сходится за конечное число шагов.*

II. *ортонормированные словари* D_1 *и* D_2 *и вектор* $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ *такие, что*

1) *чисто жадный алгоритм по словарю* D_1 *и чисто жадный алгоритм по словарю* D_2 *сходятся за конечное число шагов,*
2) *чисто жадный алгоритм по паре словарей* D_1 *и* D_2 *не сходится за конечное число шагов.*

Для удобства дальнейших выкладок в данном разделе потребуется следующая несложная конструкция. Будем сопостовлять произвольному вектору $x = (x_1, x_2, x_3, x_4, \dots)$ два вектора:

$$x_o = (x_1, 0, x_2, 0, x_3, 0, x_4, 0, \dots) \text{ и } x_e = (0, x_1, 0, x_2, 0, x_3, 0, x_4, \dots).$$

В этом случае пространство оказывается "поделено" на два подпространства, и рассматривая вектор $a = b_o + c_e$ можно считать, что рассматриваются два вектора b и c, которые составленны из нечётных и из чётных координат вектора a соответственно.

Доказательство пункта I теоремы 4.1.

Выберем ортонормированный базис $\{d_i\}_{i=1}^{\infty}$ такой, что $d_1 \in \ell_1$ и d_1 является нефинитным вектором относительно стандартного базиса F.

В качестве словарей выберем

$$D_1 = \{(d_i)_o\}_{i=1}^{\infty} \cup \{f_{2i}\}_{i=1}^{\infty} \quad \text{и} \quad D_2 = \{(d_i)_e\}_{i=1}^{\infty} \cup \{f_{2i-1}\}_{i=1}^{\infty}.$$

То есть, когда мы рассматриваем словарь D_1, можно считать, что для "нечётной части" пространства рассматривается словарь D, а для "чётной части" пространства — словарь F. Для словаря D_2 — наоборот.

Рассмотрим "симметричный" вектор $x = (d_1)_o + (d_1)_e$, для него при разложении по одному словарю, например D_1, "нечётная" часть вектора за один шаг приблизится словарём D, а "чётная" — будет приближаться словарём F за бесконечное число шагов. При разложении данного вектора по паре словарей алгоритм сойдётся за два шага. Теперь приведём строгое доказательство.

1) Чисто жадное разложение по паре словарей D_1 и D_2.

На шаге 1 разложение проводится по словарю D_1.

$$r_0^{PGA(D_1,D_2)}(x) = x, \quad \sup_{d \in D_1} \left| \left(r_0^{PGA(D_1,D_2)}(x), d \right) \right| =$$

$$= \max \left(\sup_i \left| \left(r_0^{PGA(D_1,D_2)}(x), (d_i)_o \right) \right|, \sup_i \left| \left(r_0^{PGA(D_1,D_2)}(x), f_{2i} \right) \right| \right) =$$

$$= \max(\|d_1\|_2^2, \|d_1\|_\infty) = 1,$$

и супремум достигается на единственном векторе $(d_1)_o$. Тогда получаем, что $e_1^{PGA(D_1,D_2)}(x) = (d_1)_o$ и $\hat{x}_1^{PGA(D_1,D_2)} = 1$. Отсюда следует, что

$$r_1^{PGA(D_1,D_2)}(x) = r_0^{PGA(D_1,D_2)}(x) - \hat{x}_1^{PGA(D_1,D_2)} e_1^{PGA(D_1,D_2)}(x) = (d_1)_e.$$

Шаг 2. Разложение по словарю D_2. Супремум

$$\sup_{d \in D_2} \left| \left(r_1^{PGA(D_1,D_2)}(x), d \right) \right| =$$

$$= \max \left(\sup_i \left| \left(r_1^{PGA(D_1,D_2)}(x), (d_i)_e \right) \right|, \sup_i \left| \left(r_1^{PGA(D_1,D_2)}(x), f_{2i-1} \right) \right| \right) =$$

$$= \max \left(\|d_1\|_2^2, 0 \right) = 1,$$

достигается только на векторе $(d_1)_e$. Следовательно, $e_2^{PGA(D_1,D_2)}(x) = (d_1)_e$, $\hat{x}_2^{PGA(D_1,D_2)} = 1$ и

$$r_2^{PGA(D_1,D_2)}(x) = r_1^{PGA(D_1,D_2)}(x) - \hat{x}_2^{PGA(D_1,D_2)} e_2^{PGA(D_1,D_2)}(x) = 0.$$

Таким образом, чисто жадный алгоритм по паре словарей D_1 и D_2 для вектора x сходится за два шага.

2) Первый шаг чисто жадного разложения по словарю D_1 совпадает с первым шагом алгоритма по паре словарей. Далее алгоритм будет эквивалентен чисто жадному алгоритму по стандартному словарю F для нефинитного вектора d_1. Значит, алгоритм не будет сходиться за конечное число шагов.

3) Аналогично, чисто жадное разложение по словарю D_2 не будет сходиться за конечное число шагов.

Утверждение пункта I доказано.

Замечание. Из доказательства сразу следует, что в данном случае порядок словарей D_1 и D_2 в паре неважен.

Доказательство пункта II теоремы 4.1.

Возьмём словари $D_1 = F$ и $D_2 = \left\{\frac{1}{\sqrt{2}}(f_1 + f_2), \frac{1}{\sqrt{2}}(f_1 - f_2)\right\} \cup \{f_i\}_{i=3}^{\infty}$ и рассмотрим для вектора $x = f_1 + f_2$ чисто жадное разложение по паре словарей D_1 и D_2, а также чисто жадные разложения по словарям D_1 и D_2 по отдельности.

Очевидно, что любая реализация чисто жадного алгоритма по словарю D_1 сойдётся за два шага, а любая реализация чисто жадного алгоритма по словарю D_2 сойдётся за один шаг.

Для чисто жадного алгоритма по паре словарей D_1 и D_2 докажем, что в случае $r_{2n}^{PGA(D_1,D_2)}(x) = c(f_1 \pm f_2)$ верно $r_{2n+2}^{PGA(D_1,D_2)}(x) = \pm\dfrac{c}{2}(f_1 \pm f_2)$, где знаки $\pm$ не зависят друг от друга. Тогда, учитывая, что изначально $r_0^{PGA(D_1,D_2)}(x) = f_1 + f_2$, получим, что $r_{2n}^{PGA(D_1,D_2)}(x) = \pm\dfrac{1}{2^n}(f_1 \pm f_2)$, а значит, чисто жадный алгоритм по паре словарей D_1 и D_2 для вектора x не будет сходиться за конечное число шагов.

Итак, пусть $r_{2n}^{PGA(D_1,D_2)}(x) = c(f_1 \pm f_2)$. Тогда на шаге $2n + 1$

$$\sup_{d \in D_1} \left| \left(r_{2n}^{PGA(D_1,D_2)}(x), d \right) \right| = |c|$$

достигается на двух векторах f_1 и f_2. Выбрав в качестве элемента разложения $e_{2n+1}^{PGA(D_1,D_2)}(x)$ вектор f_1, получим $\hat{x}_{2n+1}^{PGA(D_1,D_2)} = c$ и

$$r_{2n+1}^{PGA(D_1,D_2)}(x) = c(f_1 \pm f_2) - cf_1 = \pm cf_2 = Cf_2.$$

Тогда на шаге $2n + 2$ получаем, что

$$\sup_{d \in D_2} \left| \left(r_{2n+1}^{PGA(D_1,D_2)}(x), d \right) \right| = \frac{|C|}{\sqrt{2}}$$

достигается на векторах $\frac{1}{\sqrt{2}}(f_1 \pm f_2)$. Значит, $e_{2n+2}^{PGA(D_1,D_2)}(x) = \frac{1}{\sqrt{2}}(f_1 \pm f_2)$, отсюда $\hat{x}_{2n+2}^{PGA(D_1,D_2)} = \pm\frac{C}{\sqrt{2}}$ и

$$r_{2n+2}^{PGA(D_1,D_2)}(x) = Cf_2 - \left(\pm\frac{C}{\sqrt{2}}\right)\frac{1}{\sqrt{2}}(f_1 \pm f_2) = \frac{C}{2}f_1 \mp \frac{C}{2}f_2 = \pm\frac{c}{2}(f_1 \mp f_2).$$

Рассмотрим второй случай. Если $e_{2n+1}^{PGA(D_1,D_2)}(x) = f_2$, то, аналогично, $\hat{x}_{2n+1}^{PGA(D_1,D_2)} = \pm c$, остаток

$$r_{2n+1}^{PGA(D_1,D_2)}(x) = c(f_1 \pm f_2) - (\pm c)f_2 = cf_1$$

и на шаге $2n+2$

$$\sup_{d\in D_2}\left|\left(r_{2n+1}^{PGA(D_1,D_2)}(x), d\right)\right| = \frac{|c|}{\sqrt{2}}$$

достигается на $\frac{1}{\sqrt{2}}(f_1 \pm f_2)$. Значит, $e_{2n+2}^{PGA(D_1,D_2)}(x) = \dfrac{1}{\sqrt{2}}(f_1 \pm f_2)$, тогда $\hat{x}_{2n+2}^{PGA(D_1,D_2)} = \dfrac{c}{\sqrt{2}}$ и

$$r_{2n+2}^{PGA(D_1,D_2)}(x) = cf_1 - \frac{c}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}}(f_1 \pm f_2) = \frac{c}{2}(f_1 \mp f_2).$$

Таким образом, доказано, что $r_{2n+2}^{PGA(D_1,D_2)}(x) = \pm\dfrac{c}{2}(f_1 \pm f_2)$, и *доказательство теоремы 4.1 завершено.*

Для скорости сходимости можно получить результат, аналогичный теореме 4.1. Также за счёт большего множества приближающих векторов алгоритм по паре словарей ожидаемо может оказываться быстрее стандартного алгоритма по каждому словарю по отдельности, но дополнительное ограничение в отдельных случаях влечёт за собой обратную ситуацию. Так, имеет место следующая теорема.

Теорема 4.2. *Существуют*

I. *ортонормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что для некоторой монотонно убывающей к нулю положительной последовательности $\{s_n\}$ при всех $n \in \mathbb{N}$ верны соотношения*

$$c_{2^n-1}^{PGA\ D_1}(x) \geqslant s_n, \quad c_{2^n-1}^{PGA\ D_2}(x) \geqslant s_n, \quad c_{2n+2}^{PGA\ (D_1,D_2)}(x) = s_n,$$

II. *ортонормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что для некоторой монотонно убывающей к нулю положительной последовательности $\{s_n\}$ при всех $n \in \mathbb{N}$ верны соотношения*

$$c_{6n}^{PGA\ D_1}(x) = s_n, \quad c_{3n}^{PGA\ D_2}(x) = \sqrt{2} s_n, \quad c_{2^{n+2}}^{PGA\ (D_1,D_2)}(x) \geqslant s_n.$$

Доказательство пункта I.

Выберем ортонормированный базис $D = \{d_i\}_{i=1}^{\infty}$ такой, что d_{2k} являются линейными комбинациями с равными коэффициентами очередных 2^k элементов стандартного базиса F, то есть $d_{2k} = \dfrac{1}{\sqrt{2^k}} \sum\limits_{n=1}^{2^k} f_{2^k-1+n}$ при $k \in \mathbb{N}$, и рассмотрим вектор $x' = \sum\limits_{k=1}^{\infty} \dfrac{1}{2^k} \cdot d_{2k} \in \mathcal{A}_1(F) \cap \mathcal{A}_1(D)$.

Далее аналогично доказательству теоремы 4.1 в качестве словарей выберем

$$D_1 = \{(d_i)_o\}_{i=1}^{\infty} \cup \{f_{2i}\}_{i=1}^{\infty} \quad \text{и} \quad D_2 = \{(d_i)_e\}_{i=1}^{\infty} \cup \{f_{2i-1}\}_{i=1}^{\infty}.$$

В этом случае при рассмотрении разложения по словарю D_1 можно считать, что для нечётных координат рассматривается словарь D, а для чётных — словарь F. Приближаемый вектор x выберем следующим образом: $x = x'_o + x'_e$. Тогда при разложении по словарю D_1 на разложение вектора $\dfrac{1}{2^k}(d_{2k})_e$, который входит в “чётную часть” вектора x, будет затрачиваться всё большее количество шагов в зависимости от k, в то время как при разложении по паре словарей данный вектор будет приближаться за один шаг. Аналогично для словаря D_2.

Заметим, что, поскольку вектор x' принадлежит $\mathcal{A}_1(F)$ и $\mathcal{A}_1(D)$, то вектор x принадлежит $\mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$.

Так как для вектора x' и словарей D и F верно $\|x'\|_{\mathcal{A}_1(D)} = \sum\limits_{k=1}^{\infty} \frac{1}{2^k} = 1$ и $\|x'\|_{\mathcal{A}_1(F)} = \sum\limits_{k=1}^{\infty} \frac{1}{2^k} \cdot \frac{1}{\sqrt{2^k}} \sum\limits_{n=1}^{2^k} 1 = \sum\limits_{k=1}^{\infty} \frac{1}{\sqrt{2^k}} = \frac{\frac{1}{\sqrt{2}}}{1-\frac{1}{\sqrt{2}}} = 1 + \sqrt{2}$, то для выбранных вектора x и словарей D_1 и D_2 справедливо равенство

$$\|x\|_{\mathcal{A}_1(D_1)} = \|x\|_{\mathcal{A}_1(D_2)} = \|x\|_{\mathcal{A}_1(D_1,D_2)} = \|x'\|_{\mathcal{A}_1(D)} + \|x'\|_{\mathcal{A}_1(F)} = 2 + \sqrt{2}.$$

Перейдём теперь к рассмотрению реализаций чисто жадных алгоритмов и оценке скорости сходимости.

1) Для чисто жадного разложения по паре словарей D_1 и D_2 по индукции докажем, что

$$r_{2(N-1)}^{PGA(D_1,D_2)}(x) = \sum_{k=N}^{\infty} \frac{1}{2^k} \cdot ((d_{2k})_o + (d_{2k})_e)\,.$$

Очевидно, это верно для $N = 1$. Пусть данное соотношение верно на шаге $2(N-1)$, тогда на шаге $2N-1$ получаем

$$\sup_{d\in D_1} \left| \left(r_{2(N-1)}^{PGA(D_1,D_2)}(x), d \right) \right| =$$

$$= \max\left(\sup_i \left| \left(r_{2(N-1)}^{PGA(D_1,D_2)}(x), (d_i)_o \right) \right|, \sup_i \left| \left(r_{2(N-1)}^{PGA(D_1,D_2)}(x), f_{2i} \right) \right| \right) =$$

$$= \max\left(\frac{1}{2^N} \cdot \|d_{2N}\|_2^2, \frac{1}{2^N} \cdot \|d_{2N}\|_\infty \right) = \frac{1}{2^N},$$

то есть супремум достигается на единственном векторе $(d_{2N})_o$. Значит, $e_{2N-1}^{PGA(D_1,D_2)}(x) = (d_{2N})_o$, отсюда $\hat{x}_{2N-1}^{PGA(D_1,D_2)} = \dfrac{1}{2^N}$ и

$$r_{2N-1}^{PGA(D_1,D_2)}(x) = \sum_{k=N+1}^{\infty} \frac{1}{2^k}((d_{2k})_o + (d_{2k})_e) + \frac{1}{2^N}(d_{2N})_e.$$

На шаге $2N$

$$\sup_{d\in D_2} \left| \left(r_{2N-1}^{PGA(D_1,D_2)}(x), d \right) \right| =$$

$$= \max\left(\sup_i \left| \left(r_{2N-1}^{PGA(D_1,D_2)}(x), (d_i)_e \right) \right|, \sup_i \left| \left(r_{2N-1}^{PGA(D_1,D_2)}(x), f_{2i-1} \right) \right| \right) =$$

$$= \max\left(\frac{1}{2^N} \cdot \|d_{2N}\|_2^2, \frac{1}{2^N} \cdot \|d_{2N}\|_\infty \right) = \frac{1}{2^N},$$

причём супремум достигается на только на векторе $(d_{2N})_e$. Следовательно, $e_{2N}^{PGA(D_1,D_2)}(x) = (d_{2N})_e$ и $\hat{x}_{2N}^{PGA(D_1,D_2)} = \dfrac{1}{2^N}$. В итоге на шаге $2N$ остаток, действительно, имеет требуемый вид

$$r_{2N}^{PGA(D_1,D_2)}(x) = \sum_{k=N+1}^{\infty} \frac{1}{2^k} \cdot ((d_{2k})_o + (d_{2k})_e).$$

Таким образом, при всех $N \in \mathbb{N}$ верно равенство

$$\|r_{2N}^{PGA(D_1,D_2)}(x)\|_2^2 = \sum_{k=N+1}^{\infty} 2 \cdot \frac{1}{2^{2k}} = 2 \cdot \frac{1}{2^{2(N+1)}} \cdot \frac{1}{1-\frac{1}{4}} = \frac{2}{3} \cdot \frac{1}{2^{2N}},$$

значит,

$$c_{2n+2}^{PGA(D_1,D_2)}(x) = \frac{\|r_{2(n+1)}^{PGA(D_1,D_2)}(x)\|_2}{\|x\|_{\mathcal{A}_1(D_1,D_2)}} = \frac{\sqrt{2}}{\sqrt{3}\cdot(2+\sqrt{2})} \cdot \frac{1}{2^{n+1}}.$$

С учётом этого положим $s_n = \frac{1}{\sqrt{3}(1+\sqrt{2})} \cdot \frac{1}{2^{n+1}}$.

2) Рассмотрим чисто жадный алгоритм по словарю D_1. Сравнивая данный алгоритм с чисто жадным алгоритмом по стандартному словарю F для вектора x', получаем:

$$\|r_{2^n-1}^{PGA\ D_1}(x)\|_2^2 \geqslant \|r_{2^n-1}^{PGA\ F}(x')\|_2^2 = \sum_{k=n+1}^{\infty} \frac{1}{2^{2k}} = \frac{1}{3} \cdot \frac{1}{2^{2n}}.$$

Отсюда

$$c_{2^n-1}^{PGA\ D_1}(x) = \frac{\|r_{2^n-1}^{PGA\ D_1}(x)\|_2}{\|x\|_{\mathcal{A}_1(D_1)}} \geqslant \frac{\frac{1}{\sqrt{3}} \cdot \frac{1}{2^n}}{(2+\sqrt{2})} > \frac{1}{\sqrt{3}(1+\sqrt{2})} \cdot \frac{1}{2^{n+1}} = s_n.$$

3) Для чисто жадного алгоритма по словарю D_2 аналогично доказывается, что $c_{2^n-1}^{PGA\ D_2}(x) > s_n$.

Итак, *пункт I теоремы 4.2 доказан.*

Замечание. Порядок словарей D_1 и D_2 в паре неважен.

Замечание. При выборе в качестве x' вектора $\sum\limits_{k=1}^{\infty} \sqrt{\frac{1}{k^2} - \frac{1}{(k+1)^2}} \cdot d_{2k}$ имеют место следующие соотношения для нормы остатка на соответствующем шаге:

$$\|r_{2^n-1}^{PGA\ D_1}(x)\|_2 \geqslant S_n,\ \|r_{2^n-1}^{PGA\ D_2}(x)\|_2 \geqslant S_n,\ \|r_{2n+2}^{PGA\ (D_1,D_2)}(x)\|_2 = S_n,$$

где последовательность S_n убывает медленнее, а именно $S_n = \frac{1}{\sqrt{2}(n+1)}$. Однако в этом случае $x \notin \mathcal{A}_1(D_i),\ i = \overline{1,2}$.

Доказательство пункта II теоремы 4.2.

Выберем в качестве словарей $D_1 = F$ и $D_2 = \left\{\frac{1}{\sqrt{2}}(f_{2n+1} \pm f_{2n+2})\right\}_{n=0}^{\infty}$, а в качестве x вектор $\sum\limits_{k=0}^{\infty}\frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2}) \in \mathcal{A}_1(D_1)\cap\mathcal{A}_1(D_2)$. В этом случае для приближения вектора $\frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2})$, который является "составляющей частью" вектора x, словарём D_1 или словарём D_2 по отдельности потребуется один или два шага. Если разложение реализуется по паре словарей, то для приближения трёх таких подряд идущих векторов потребуется тем больше векторов, чем больше номер k. Далее приводится строгое доказательство.

Для вектора x верно $\|x\|_{\mathcal{A}_1(D_1)} = \sum\limits_{k=0}^{\infty}\frac{1}{2^{2^k}}\cdot 2$. Для сокращения записи обозначим через s число $\sum\limits_{k=0}^{\infty}\frac{1}{2^{2^k}}$. Тогда

$$\|x\|_{\mathcal{A}_1(D_1)} = 2s, \quad \|x\|_{\mathcal{A}_1(D_2)} = \sum_{k=0}^{\infty}\frac{1}{2^{2^k}}\cdot\sqrt{2} = \sqrt{2}s \quad \text{и} \quad \|x\|_{\mathcal{A}_1(D_1,D_2)} = 2s.$$

Теперь рассмотрим скорость сходимости алгоритмов для выбранного вектора x.

1) Для чисто жадного алгоритма по словарю D_1, очевидно, верно

$$r_{2n}^{PGA\ D_1}(x) = \sum_{k=n}^{\infty}\frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2}) \quad \text{и} \quad \|r_{2n}^{PGA\ D_1}(x)\|_2^2 = \sum_{k=n}^{\infty} 2\cdot\frac{1}{2^{2^{k+1}}}.$$

С учётом этого, определим последовательность s_n следующим образом:

$$s_n = c_{6n}^{PGA\ D_1}(x) = \frac{\|r_{6n}^{PGA\ D_1}(x)\|_2}{\|x\|_{\mathcal{A}_1(D_1)}} = \frac{1}{2s}\sqrt{\sum_{k=3n}^{\infty} 2\cdot\frac{1}{2^{2^{k+1}}}}.$$

2) Аналогично, для чисто жадного алгоритма по словарю D_2 имеют место равенства

$$r_{n}^{PGA\ D_2}(x) = \sum_{k=n}^{\infty}\frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2}) \quad \text{и} \quad \|r_{n}^{PGA\ D_2}(x)\|_2^2 = \sum_{k=n}^{\infty} 2\cdot\frac{1}{2^{2^{k+1}}}.$$

Значит,

$$c_{3n}^{PGA\ D_2}(x) = \frac{\|r_{3n}^{PGA\ D_2}(x)\|_2}{\|x\|_{\mathcal{A}_1(D_2)}} = \frac{1}{\sqrt{2}s}\sqrt{\sum_{k=3n}^{\infty} 2\cdot\frac{1}{2^{2^{k+1}}}} = \sqrt{2}s_n.$$

3) Перейдём к рассмотрению чисто жадного алгоритма по паре словарей D_1 и D_2.

Для упрощения изложения без ограничения общности ниже будет рассмотрена только часть возможных реализаций чисто жадного алгоритма по паре словарей, остальные его реализации будут отличаться лишь знаками координат векторов $r_n^{PGA\ (D_1,D_2)}(x)$ и коэффициентов разложения $\hat{x}_n^{PGA\ (D_1,D_2)}$.

Докажем, что на то, чтобы "занулить" очередные 6 координат (в стандартном базисе F), требуется $2^{n+1}+2$ или $2^{n+1}+4$ шагов, при этом остальные координаты остаются неизменными. В этом случае, поскольку верно соотношение

$$\sum_{k=0}^{n}(2^{k+1}+2) = 2\cdot\frac{1-2^{n+1}}{1-2} + 2(n+1) = 2^{n+2}+2n \geqslant 2^{n+2},$$

на шаге 2^{n+2} все координаты вектора x и остатка $r_{2^{n+2}}^{PGA(D_1,D_2)}(x)$, корме первых $6n$ координат, будут совпадать. Следовательно, получаем неравенство

$$\|r_{2^{n+2}}^{PGA\ (D_1,D_2)}(x)\|_2^2 \geqslant \left\|\sum_{k=3n}^{\infty}\frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2})\right\|_2^2 = \sum_{k=3n}^{\infty} 2\cdot\frac{1}{2^{2^{k+1}}},$$

из которого сразу следует, что

$$c_{2^{n+2}}^{PGA\ (D_1,D_2)}(x) = \frac{\|r_{2^{n+2}}^{PGA\ (D_1,D_2)}(x)\|_2}{\|x\|_{\mathcal{A}_1(D_1,D_2)}} \geqslant \frac{1}{2s}\sqrt{\sum_{k=3n}^{\infty} 2\cdot\frac{1}{2^{2^{k+1}}}} = s_n.$$

Итак, пусть на некотором чётном шаге $2N$ верно

$$r_{2N}^{PGA\ (D_1,D_2)}(x) = \sum_{k=n}^{\infty}\frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2}),$$

тогда покажем, что при $0 \leqslant m \leqslant 2^n-1$ имеет место равенство

$$r_{2N+2m}^{PGA\ (D_1,D_2)}(x) = \frac{1}{2^m}\cdot\frac{1}{2^{2^n}}(f_{2n+1}+f_{2n+2}) + \sum_{k=n+1}^{\infty}\frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2}).$$

При $m=0$ получаем шаг $2N$ и по первому предположению данное утверждение верно. Предположив, что для некоторого $0 \leqslant m < 2^n - 1$ остаток $r_{2N+2m}^{PGA\ (D_1,D_2)}(x)$ имеет требуемый вид, докажем, что

$$r_{2N+2(m+1)}^{PGA\ (D_1,D_2)}(x) = \frac{1}{2^{m+1}} \cdot \frac{1}{2^{2^n}}(f_{2n+1} + f_{2n+2}) + \sum_{k=n+1}^{\infty} \frac{1}{2^{2^k}}(f_{2k+1} + f_{2k+2}).$$

Рассмотрим шаг $2N + 2m + 1$. Поскольку $m < 2^n - 1$, то имеет место неравенство $\dfrac{1}{2^m} \cdot \dfrac{1}{2^{2^n}} > \dfrac{1}{2^{2^{n+1}}}$, откуда получаем, что

$$\sup_{d \in D_1} \left| \left(r_{2N+2m}^{PGA\ (D_1,D_2)}(x), d \right) \right| = \frac{1}{2^m} \cdot \frac{1}{2^{2^n}}$$

достигается на двух векторах f_{2n+1} и f_{2n+2}. Выберем в качестве элемента разложения $e_{2N+2m+1}^{PGA\ (D_1,D_2)}(x)$ вектор f_{2n+1} (второй случай аналогичен), тогда верно $\hat{x}_{2N+2m+1}^{PGA\ (D_1,D_2)} = \dfrac{1}{2^m} \cdot \dfrac{1}{2^{2^n}}$ и

$$r_{2N+2m+1}^{PGA\ (D_1,D_2)}(x) = \frac{1}{2^m} \cdot \frac{1}{2^{2^n}} \cdot f_{2n+2} + \sum_{k=n+1}^{\infty} \frac{1}{2^{2^k}}(f_{2k+1} + f_{2k+2}).$$

Шаг $2N + 2m + 2$. Учитывая, что $m < 2^n - 1$, получаем неравенство $\dfrac{1}{\sqrt{2}} \cdot \dfrac{1}{2^m} \cdot \dfrac{1}{2^{2^n}} > 2 \cdot \dfrac{1}{\sqrt{2}} \cdot \dfrac{1}{2^{2^{n+1}}}$. Следовательно,

$$\sup_{d \in D_2} \left| \left(r_{2N+2m+1}^{PGA\ (D_1,D_2)}(x), d \right) \right| = \frac{1}{\sqrt{2}} \cdot \frac{1}{2^m} \cdot \frac{1}{2^{2^n}},$$

причём супремум достигается только на векторах $\frac{1}{\sqrt{2}}(f_{2n+1} \pm f_{2n+2})$. Например, будем считать, что $e_{2N+2k+2}^{PGA\ (D_1,D_2)}(x) = \frac{1}{\sqrt{2}}(f_{2n+1} - f_{2n+2})$ (второй случай аналогичен). Тогда $\hat{x}_{2N+2k+2}^{PGA\ (D_1,D_2)} = -\dfrac{1}{\sqrt{2}} \cdot \dfrac{1}{2^k} \cdot \dfrac{1}{2^{2^n}}$ и, опуская выкладки, получаем, что

$$r_{2N+2m+2}^{PGA\ (D_1,D_2)}(x) = \frac{1}{2^{m+1}} \cdot \frac{1}{2^{2^n}}(f_{2n+1} + f_{2n+2}) + \sum_{k=n+1}^{\infty} \frac{1}{2^{2^k}}(f_{2k+1} + f_{2k+2}).$$

Таким образом, получаем, что при $m = 2^n - 1$ для остатка справедливо требуемое равенство:

$$r_{2N+2^{n+1}-2}^{PGA\ (D_1,D_2)}(x) = \frac{1}{2^{2^{n+1}-1}}(f_{2n+1} + f_{2n+2}) + \sum_{k=n+1}^{\infty} \frac{1}{2^{2^k}}(f_{2k+1} + f_{2k+2}).$$

На следующем шаге $2N+2^{n+1}-1$ всё ещё верно $e_{2N+2^{n+1}-1}^{PGA\ (D_1,D_2)}(x) = f_{2n+1}$ или $e_{2N+2^{n+1}-1}^{PGA\ (D_1,D_2)}(x) = f_{2n+2}$ и, например, выбрав первый вариант (второй аналогичен), получаем

$$r_{2N+2^{n+1}-1}^{PGA\ (D_1,D_2)}(x) = \frac{1}{2^{2^{n+1}-1}} \cdot f_{2n+2} + \sum_{k=n+1}^{\infty} \frac{1}{2^{2^k}}(f_{2k+1} + f_{2k+2}).$$

Шаг $2N + 2^{n+1}$. Супремум

$$\sup_{d\in D_2} \left|\left(r_{2N+2^{n+1}-1}^{PGA\ (D_1,D_2)}(x), d\right)\right| = \max\left(\frac{1}{\sqrt{2}} \cdot \frac{1}{2^{2^{n+1}-1}}, 2 \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{2^{2^{n+1}}}\right) = \frac{1}{\sqrt{2}} \cdot \frac{1}{2^{2^{n+1}}}$$

достигается на трёх векторах, а именно, $\frac{1}{\sqrt{2}}(f_{2n+1} \pm f_{2n+2})$ и $\frac{1}{\sqrt{2}}(f_{2n+3}+f_{2n+4})$. В данном случае не все случаи будут аналогичными, поэтому разберём их отдельно.

Случай, когда $e_{2N+2^{n+1}}^{PGA\ (D_1,D_2)}(x) = \frac{1}{\sqrt{2}}(f_{2n+1} + f_{2n+2})$, аналогичен случаю $e_{2N+2^{n+1}}^{PGA\ (D_1,D_2)}(x) = \frac{1}{\sqrt{2}}(f_{2n+1} - f_{2n+2})$, поэтому, достаточно рассмотреть, например, только второй случай. Здесь $\hat{x}_{2N+2^{n+1}}^{PGA\ (D_1,D_2)} = -\dfrac{1}{\sqrt{2}} \cdot \dfrac{1}{2^{2^n}}$ и, опуская выкладки, получаем

$$r_{2N+2^{n+1}}^{PGA\ (D_1,D_2)}(x) = \frac{1}{2^{2^{n+1}}}(f_{2n+1}+f_{2n+2}+f_{2n+3}+f_{2n+4}) + \sum_{k=n+2}^{\infty} \frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2}).$$

На шаге $2N + 2^{n+1} + 1$

$$\sup_{d\in D_1} \left|\left(r_{2N+2^{n+1}}^{PGA\ (D_1,D_2)}(x), d\right)\right| = \frac{1}{2^{2^{n+1}}},$$

при этом супремум достигается на четырёх векторах f_{2n+i}, где $i = 1, 2, 3, 4$. Здесь без ограничения общности можно считать, что на данном шаге в качестве приближающего вектора $e_{2N+2^{n+1}+1}^{PGA\ (D_1,D_2)}(x)$ выбирается вектор f_{2n+1}, тогда $\hat{x}_{2N+2^{n+1}+1}^{PGA\ (D_1,D_2)}(x) = \dfrac{1}{2^{2^{n+1}}}$ и

$$r_{2N+2^{n+1}+1}^{PGA\ (D_1,D_2)}(x) = \frac{1}{2^{2^{n+1}}}(f_{2n+2} + f_{2n+3} + f_{2n+4}) + \sum_{k=n+2}^{\infty} \frac{1}{2^{2^k}}(f_{2k+1} + f_{2k+2}).$$

Тогда на шаге $2N + 2^{n+1} + 2$ супремум

$$\sup_{d\in D_2}\left|\left(r_{2N+2^{n+1}+1}^{PGA\ (D_1,D_2)}(x),d\right)\right| = 2\cdot\frac{1}{\sqrt{2}}\cdot\frac{1}{2^{2^{n+1}}}$$

достигается на $\frac{1}{\sqrt{2}}(f_{2n+3}+f_{2n+4})$, то есть $e_{2N+2^{n+1}+2}^{PGA\ (D_1,D_2)}(x) = \frac{1}{\sqrt{2}}(f_{2n+3}+f_{2n+4})$ и $\hat{x}_{2N+2^{n+1}+2}^{PGA\ (D_1,D_2)}(x) = 2\cdot\dfrac{1}{\sqrt{2}}\cdot\dfrac{1}{2^{2^{n+1}}}$, отсюда

$$r_{2N+2^{n+1}+2}^{PGA\ (D_1,D_2)}(x) = \frac{1}{2^{2^{n+1}}}f_{2n+2} + \sum_{k=n+2}^{\infty}\frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2}).$$

Шаг $2N+2^{n+1}+3$.

$$\sup_{d\in D_1}\left|\left(r_{2N+2^{n+1}+2}^{PGA\ (D_1,D_2)}(x),d\right)\right| = \frac{1}{2^{2^{n+1}}}$$

достигается только на векторе f_{2n+2}, откуда однозначно следует, что $e_{2N+2^{n+1}+3}^{PGA\ (D_1,D_2)}(x) = f_{2n+2}$, $\hat{x}_{2N+2^{n+1}+3}^{PGA\ (D_1,D_2)}(x) = \dfrac{1}{2^{2^{n+1}}}$ и

$$r_{2N+2^{n+1}+1}^{PGA\ (D_1,D_2)}(x) = \sum_{k=n+2}^{\infty}\frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2}).$$

Но, поскольку цикл должен завершаться на чётном шаге, рассмотрим ещё один шаг, то есть шаг $2N+2^{n+1}+4$. Здесь

$$\sup_{d\in D_2}\left|\left(r_{2N+2^{n+1}+3}^{PGA\ (D_1,D_2)}(x),d\right)\right| = 2\cdot\frac{1}{\sqrt{2}}\cdot\frac{1}{2^{2^{n+2}}}$$

достигается только на векторе $\frac{1}{\sqrt{2}}(f_{2n+5}+f_{2n+6})$. Отсюда получаем, что $e_{2N+2^{n+1}+4}^{PGA\ (D_1,D_2)}(x) = \frac{1}{\sqrt{2}}(f_{2n+5}+f_{2n+6})$ и $\hat{x}_{2N+2^{n+1}+4}^{PGA\ (D_1,D_2)}(x) = 2\cdot\dfrac{1}{\sqrt{2}}\cdot\dfrac{1}{2^{2^{n+2}}}$. Таким образом,

$$r_{2N+2^{n+1}+4}^{PGA\ (D_1,D_2)}(x) = \sum_{k=n+3}^{\infty}\frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2}).$$

Рассмотрим оставшийся случай, когда $e_{2N+2^{n+1}}(x) = \frac{1}{\sqrt{2}}(f_{2n+3}+f_{2n+4})$. Здесь $\hat{x}_{2N+2^{n+1}} = 2\cdot\dfrac{1}{\sqrt{2}}\cdot\dfrac{1}{2^{2^{n+1}}}$ и

$$r_{2N+2^{n+1}}(x) = \frac{1}{2^{2^{n+1}-1}}\cdot f_{2n+2} + \sum_{k=n+2}^{\infty}\frac{1}{2^{2^k}}(f_{2k+1}+f_{2k+2}).$$

Далее, шаг $2N + 2^{n+1} + 1$:

$$\sup_{d \in D_1} \left| \left(r_{2N+2^{n+1}}^{PGA\ (D_1,D_2)}(x), d \right) \right| = \frac{1}{2^{2^{n+1}-1}}$$

достигается только на векторе f_{2n+2}, следовательно, $e_{2N+2^{n+1}+1}^{PGA\ (D_1,D_2)}(x) = f_{2n+2}$ и $\hat{x}_{2N+2^{n+1}+1}^{PGA\ (D_1,D_2)}(x) = \dfrac{1}{2^{2^{n+1}-1}}$, тогда

$$r_{2N+2^{n+1}+1}^{PGA\ (D_1,D_2)}(x) = \sum_{k=n+2}^{\infty} \frac{1}{2^{2^k}} (f_{2k+1} + f_{2k+2}).$$

Чтобы завершить цикл на чётном шаге, рассмотрим дополнительно шаг $2N + 2^{n+1} + 2$.

$$\sup_{d \in D_2} \left| \left(r_{2N+2^{n+1}+1}^{PGA\ (D_1,D_2)}(x), d \right) \right| = 2 \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{2^{2^{n+2}}}$$

достигается на единственном векторе $\frac{1}{\sqrt{2}}(f_{2n+5} + f_{2n+6})$. По определению, $e_{2N+2^{n+1}+2}^{PGA\ (D_1,D_2)}(x) = \frac{1}{\sqrt{2}}(f_{2n+5} + f_{2n+6})$ и $\hat{x}_{2N+2^{n+1}+2}^{PGA\ (D_1,D_2)}(x) = 2 \cdot \dfrac{1}{\sqrt{2}} \cdot \dfrac{1}{2^{2^{n+2}}}$. Отсюда

$$r_{2N+2^{n+1}+2}^{PGA\ (D_1,D_2)}(x) = \sum_{k=n+3}^{\infty} \frac{1}{2^{2^k}} (f_{2k+1} + f_{2k+2}).$$

Таким образом, действительно, за очередные $2^{n+1} + 2$ или $2^{n+1} + 4$ шагов только "зануляются" очередные 6 координат вектора x и *теорема 4.2 доказана.*

Замечание. Результат второго пункта данной теоремы можно усилить в следующем смысле: найдутся словари и вектор такие, что скорость сходимости чисто жадного алгоритма по паре словарей на шаге с номером $f(n)$ будет больше, чем скорости сходимости стандартных чисто жадных алгоритмов на шаге $3n$ по одному словарю и на шаге $6n$ по другому, где $f(n)$ растёт быстрее, чем 2^n. Для этого будет необходимо выбрать вектор x, координаты которого убывают быстрее, чем в приведённом доказательстве.

Чтобы "скомпенсировать" разницу между множествами, по которым происходит приближение, дополнительно к рассмотрению разложения по словарям D_1 и D_2 по отдельности сравним чисто жадный алгоритм по паре словарей D_1 и D_2 с чисто жадным алгоритмом по их объединению $D = D_1 \cup D_2$. В этом случае словарь для стандартного алгоритма уже не является меньше, чем множество приближающих векторов в случае алгоритма по паре словарей.

Здесь, конечно, верен аналог второго пункта теоремы 4.1. Однако разделение словаря $D = D_1 \cup D_2$ на два и проведение алгоритма по паре словарей накладывает определённое ограничение, благодаря чему чисто жадный алгоритм по паре словарей также может оказываться быстрее стандартного алгоритма.

Теорема 4.3. *Существуют*

I. *ортонормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что*

   1) *чисто жадный алгоритм по словарю $D = D_1 \cup D_2$ сходится за конечное число шагов,*

   2) *чисто жадный алгоритм по паре словарей D_1 и D_2 не сходится за конечное число шагов.*

II. *нормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что*

   1) *чисто жадный алгоритм по словарю $D = D_1 \cup D_2$ не сходится за конечное число шагов,*

   2) *чисто жадный алгоритм по паре словарей D_1 и D_2 сходится за конечное число шагов.*

Доказательство.

Для *доказательства пункта I*, как и в доказательстве второго пункта теоремы 4.1 в качестве словарей выберем следующие словари: $D_1 = F$ и $D_2 = \left\{\frac{1}{\sqrt{2}}(f_1 + f_2), \frac{1}{\sqrt{2}}(f_1 - f_2)\right\} \cup \{f_i\}_{i=3}^{\infty}$. Рассмотрим алгоритмы для вектора $x = f_1 + f_2$.

Очевидно, что чисто жадное разложение по словарю $D = D_1 \cup D_2$ сходится за 1 шаг; с другой стороны, чисто жадное разложение по паре словарей D_1 и D_2 не сходится за конечное число шагов (см. доказательство теоремы 4.1).

Пункт I теоремы доказан.

Доказательство пункта II.

Определим вектор f следующим образом:

$$f = \frac{1}{\sqrt{2}} \cdot f_1 - \sum_{k=1}^{\infty} \frac{1}{\sqrt{3}^k} f_{k+1}, \text{ тогда } \|f\|_2^2 = \frac{1}{2} + \sum_{k=1}^{\infty} \frac{1}{3^k} = \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{1-\frac{1}{3}} = 1,$$

и f — единичный вектор в ℓ_2.

В качестве словарей выберем $D_1 = F$ и $D_2 = F \cup \{f\}$ и рассмотрим чисто жадные алгоритмы для вектора $x = \sqrt{2} f_1 - \sqrt{3} f_2$. Покажем, что алгоритм по паре словарей сходится за два шага, а при разложении по объединению двух словарей сначала получается нефинитный остаток, а далее на каждом шаге зануляется наибольшая по модулю координата.

1) Чисто жадное разложение по словарю $D = D_1 \cup D_2$.

Шаг 1. Легко видеть, что

$$\sup_{d \in D} \left| \left(r_0^{PGA\ D_1 \cup D_2}(x), d \right) \right| = \max\left(\sqrt{3}, 2 \right) = 2,$$

и при этом супремум достигается только на одном векторе f. Следовательно, $e_1^{PGA\ D_1 \cup D_2}(x) = f$ и $\hat{x}_1^{PGA\ D_1 \cup D_2} = 2$, отсюда

$$r_1^{PGA\ D_1 \cup D_2}(x) = \sqrt{2} \cdot f_1 - \sqrt{3} \cdot f_2 - 2 \cdot \left(\frac{1}{\sqrt{2}} \cdot f_1 - \sum_{k=1}^{\infty} \frac{1}{\sqrt{3}^k} f_{k+1} \right) =$$

$$= -\frac{1}{\sqrt{3}} \cdot f_2 + \sum_{k=2}^{\infty} \frac{2}{\sqrt{3}^k} f_{k+1}.$$

Шаг 2. Очевидно, что

$$\sup_{d \in D} \left| \left(r_1^{PGA\ D_1 \cup D_2}(x), d \right) \right| = \frac{2}{\sqrt{3}^2},$$

и супремум достигается на единственном векторе — f_3. Тогда получаем, что $e_2^{PGA\ D_1 \cup D_2}(x) = f_3$, $\hat{x}_2^{PGA\ D_1 \cup D_2} = \frac{2}{\sqrt{3}^2}$ и

$$r_2^{PGA\ D_1\cup D_2}(x) = -\frac{1}{\sqrt{3}}\cdot f_2 + \sum_{k=3}^{\infty}\frac{2}{\sqrt{3}^k}f_{k+1}.$$

Шаг 3. Поскольку

$$(r_2^{PGA\ D_1\cup D_2}(x), f) = \frac{1}{3} - 2\sum_{k=3}^{\infty}\frac{1}{3^k} = \frac{1}{3} - 2\cdot\frac{1}{27}\cdot\frac{1}{1-\frac{1}{3}} = \frac{2}{9},$$

то

$$\sup_{d\in D}\left|\left(r_2^{PGA\ D_1\cup D_2}(x), d\right)\right| = \frac{1}{\sqrt{3}},$$

и супремум достигается только на векторе f_2. Значит, $e_3^{PGA\ D_1\cup D_2}(x) = f_2$ и $\hat{x}_3^{PGA\ D_1\cup D_2} = -\frac{1}{\sqrt{3}}$, следовательно,

$$r_3^{PGA\ D_1\cup D_2}(x) = \sum_{k=3}^{\infty}\frac{2}{\sqrt{3}^k}f_{k+1}.$$

Докажем, что при $n \geqslant 3$ для остатка верно равенство

$$r_n^{PGA\ D_1\cup D_2}(x) = \sum_{k=n}^{\infty}\frac{2}{\sqrt{3}^k}f_{k+1}.$$

При $n = 3$ данное соотношение выполнено. При $n > 3$ предположим, что верно

$$r_{n-1}^{PGA\ D_1\cup D_2}(x) = \sum_{k=n-1}^{\infty}\frac{2}{\sqrt{3}^k}f_{k+1}.$$

Тогда на шаге n имеет место равенство

$$(r_{n-1}^{PGA\ D_1\cup D_2}(x), f) = 2\sum_{k=n-1}^{\infty}\frac{1}{3^k} = 2\cdot\frac{1}{3^{n-1}}\cdot\frac{1}{1-\frac{1}{3}} = \frac{1}{3^{n-2}}$$

и

$$\sup_{d\in D}\left|\left(r_{n-1}^{PGA\ D_1\cup D_2}(x), d\right)\right| = \frac{2}{\sqrt{3}^{n-1}}$$

достигается только на f_n. Отсюда $e_n^{PGA\ D_1\cup D_2}(x) = f_n$ и $\hat{x}_n^{PGA\ D_1\cup D_2} = \frac{2}{\sqrt{3}^{n-1}}$. Тогда, действительно,

$$r_n^{PGA\ D_1\cup D_2}(x) = \sum_{k=n}^{\infty}\frac{2}{\sqrt{3}^k}f_{k+1}.$$

Итак, алгоритм по объединению двух словарей за конечное число шагов не сходится.

2) Рассмотрим чисто жадное разложение по паре словарей D_1 и D_2.

На шаге 1 получаем, что

$$\sup_{d \in D_1} \left| \left(r_0^{PGA(D_1,D_2)}(x), d \right) \right| = \sqrt{3},$$

причём супремум достигается только на одном векторе f_2. Следовательно, $e_1^{PGA(D_1,D_2)}(x) = f_2$ и, значит, $\hat{x}_1^{PGA(D_1,D_2)} = -\sqrt{3}$. Отсюда

$$r_1^{PGA(D_1,D_2)}(x) = r_0^{PGA(D_1,D_2)}(x) - \hat{x}_1^{PGA(D_1,D_2)} e_1^{PGA(D_1,D_2)}(x) = \sqrt{2} \cdot f_1.$$

Шаг 2. Верно, что

$$\sup_{d \in D_2} \left| \left(r_1^{PGA(D_1,D_2)}(x), d \right) \right| = \max \left(\sqrt{2}, \sqrt{2} \cdot \frac{1}{\sqrt{2}} \right) = \sqrt{2},$$

и при этом супремум достигается на единственном векторе — f_1. Здесь $e_2^{PGA(D_1,D_2)}(x) = f_1$ и $\hat{x}_2^{PGA(D_1,D_2)} = \sqrt{2}$, тогда

$$r_2^{PGA(D_1,D_2)}(x) = r_1^{PGA(D_1,D_2)}(x) - \hat{x}_2^{PGA(D_1,D_2)} e_2^{PGA(D_1,D_2)}(x) = 0.$$

Алгоритм сходится за 2 шага.

Теорема доказана.

Замечание. Можно доказать аналогичную теорему для неполных “словарей”, например, выбрав в каждом из пунктов соответствующие “словари” $D_1 \setminus \{f_3\}$ и $D_2 \setminus \{f_4\}$.

Таким образом, сравнивая чисто жадный алгоритм и чисто жадный алгоритм по паре словарей, можно сделать вывод о том, что в зависимости от словарей и приближаемого вектора стандартный алгоритм может быть как быстрее, так и медленнее своей модификации.

4.2 Сравнение ортогонального жадного алгоритма и ортогонального жадного алгоритма по паре словарей.

Для ортогонального жадного алгоритма по системе возможно неполных словарей имеет место положительный результат о сильной сходимости данного алгоритма на произвольном векторе гильбертова пространства, аналогичный теореме J о слабой сходимости чисто жадного алгоритма по системе возможно неполных словарей. Более точно, верно следующее утверждение.

Теорема 4.4. *В гильбертовом пространстве H для произвольной системы возможно неполных словарей $\{D_1, D_2, \dots, D_N\}$ и любого вектора $x \in H$ ортогональный жадный алгоритм по системе возможно неполных словарей $\{D_1, D_2, \dots, D_N\}$ сходится к приближаемому вектору x.*

В доказательстве данного утверждения используется идея доказательства теоремы G для слабого ортогонального жадного алгоритма, приведённонного в статье [30].

Доказательство теоремы 4.4.

Обозначим $\mathcal{D} = \{D_1, D_2, \dots, D_N\}$.

Пусть $H_n = \operatorname{span}(e_1^{OGA\,\mathcal{D}}(x), e_2^{OGA\,\mathcal{D}}(x), \dots e_n^{OGA\,\mathcal{D}}(x))$, тогда $H_n \subset H_{n+1}$ и, следовательно, $\{P_{H_n}\}$ сходится к некоторому вектору v. Допустим, что $v \neq x$, и обозначим $u = x - v \neq 0$. В этом случае существует $\delta > 0$ такое, что

$$\max_{k \in \overline{1,N}} \sup_{d \in D_k} |(u, d)| \geqslant 2\delta,$$

и max достигается на некотором $k = K$.

Тогда для некоторого n_0 при всех $n \geqslant n_0$ верно

$$\sup_{d \in D_K} |(r_n^{OGA\,\mathcal{D}}(x), d)| \geqslant \delta.$$

Для удобства дальнейшего обозначения определим последовательность $\kappa(i)$ следующим образом: $\kappa(1) = n_0,\ \kappa(i+1) = \min\{n > \kappa(i) :\ i(n) = K\}$ при $i \in \mathbb{N}$.

Заметим, что $x = r_{n-1}^{OGA\,\mathcal{D}}(x) + \mathrm{Proj}_{\left\{e_k^{OGA\,\mathcal{D}}(x)\right\}_{k=1}^{n-1}} x$, следовательно,

$r_n^{OGA\,\mathcal{D}}(x) = x - \mathrm{Proj}_{\left\{e_k^{OGA\,\mathcal{D}}(x)\right\}_{k=1}^{n}} x = r_{n-1}^{OGA\,\mathcal{D}}(x) - \mathrm{Proj}_{\left\{e_k^{OGA\,\mathcal{D}}(x)\right\}_{k=1}^{n}} r_{n-1}^{OGA\,\mathcal{D}}(x)$. Отсюда

$$\|r_n^{OGA\,\mathcal{D}}(x)\|_2^2 \leqslant \|r_{n-1}^{OGA\,\mathcal{D}}(x)\|_2^2 - (r_n^{OGA\,\mathcal{D}}(x), e_n^{OGA\,\mathcal{D}}(x))^2 =$$

$$= \|r_{n-1}^{OGA\,\mathcal{D}}(x)\|_2^2 - \sup_{d \in D_i} |(r_n^{OGA\,\mathcal{D}}(x), d)|^2.$$

Тогда при $n = \kappa(i)$ имеет место неравенство

$$\|r_n^{OGA\,\mathcal{D}}(x)\|_2^2 \leqslant \|r_{n-1}^{OGA\,\mathcal{D}}(x)\|_2^2 - \delta^2,$$

а при остальных n верно соотношение

$$\|r_n^{OGA\,\mathcal{D}}(x)\|_2^2 \leqslant \|r_{n-1}^{OGA\,\mathcal{D}}(x)\|_2^2.$$

Следовательно,

$$\|r_{\kappa(i+1)}^{OGA\,\mathcal{D}}(x)\|_2^2 \leqslant \|r_{\kappa(i)}^{OGA\,\mathcal{D}}(x)\|_2^2 - \delta^2 \leqslant \|r_{n_0}^{OGA\,\mathcal{D}}(x)\|_2^2 - \delta^2 i < 0,$$

при $i > n_0 + \left(\dfrac{\|r_{n_0}^{OGA\,\mathcal{D}}(x)\|_2}{\delta}\right)^2$.

Таким образом, получаем противоречие. Значит, $v = x$ и алгоритм сходится к приближаемому вектору x.

Утверждение доказано.

При сравнении ортогонального жадного алгоритма по паре словарей и ортогонального жадного алгоритма оказываются справедливыми теоремы, аналогичные теоремам 4.1–4.3, сформулированным для чисто жадного алгоритма по паре словарей в разделе 4.1.

Так, имеет место следующая теорема (заметим, что в данном случае словари не обязательно ортогональны).

Теорема 4.5. *Существуют*

I. *ортонормированные словари* D_1 *и* D_2 *и вектор* $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ *такие, что*

1) *ортогональный жадный алгоритм по словарю D_1 и ортогональный жадный алгоритм по словарю D_2 не сходятся за конечное число шагов,*

2) *ортогональный жадный алгоритм по паре словарей D_1 и D_2 сходится за конечное число шагов.*

II. *нормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что*

1) *ортогональный жадный алгоритм по словарю D_1 и ортогональный жадный алгоритм по словарю D_2 сходятся за конечное число шагов,*

2) *ортогональный жадный алгоритм по паре словарей D_1 и D_2 не сходится за конечное число шагов.*

Доказательство пункта I данной теоремы почти дословно совпадает с доказательством пункта I теоремы 4.1.

Замечание. Порядок словарей D_1 и D_2 в паре неважен.

Доказательство пункта II теоремы 4.5.

Определим два вектора следующим образом:

$$f = \sum_{k=1}^{\infty} \frac{1}{\sqrt{2}^k} f_k, \quad f' = \frac{\sqrt{3}}{2} f_2 + \sum_{k=3}^{\infty} \frac{1}{\sqrt{2}^k} f_k.$$

Легко проверить, что $\|f\|_2 = 1$ и $\|f'\|_2 = 1$. В качестве словарей выберем нормированные словари

$$D_1 = \{f'\} \cup F \setminus \{f_3\} \quad \text{и} \quad D_2 = \left\{ f, \sqrt{\frac{2}{3}} f_2 + \sqrt{\frac{1}{3}} f_3 \right\} \cup F \setminus \{f_2\},$$

а в качестве приближаемого вектора выберем $x = \sum\limits_{k=1}^{\infty} \frac{1}{\sqrt{2}^k} f_k = f$.

1) Рассмотрим ортогональный жадный алгоритм по словарю D_1.

Шаг 1. По определению $r_0^{OGA\ D_1}(x) = x$, и, поскольку

$$(x, f') = \frac{\sqrt{3}}{2} \cdot \frac{1}{2} + \sum_{k=3}^{\infty} \frac{1}{2^k} = \frac{\sqrt{3}}{4} + \frac{1}{4} = \frac{\sqrt{3}+1}{4},$$

имеет место равенство

$$\sup_{d\in D_1}\left(r_0^{OGA\ D_1}(x),d\right)=\max\left(\frac{1}{\sqrt{2}},\frac{\sqrt{3}+1}{4}\right)=\frac{1}{\sqrt{2}},$$

причём супремум достигается на единственном векторе — f_1. Таким образом, $e_1^{OGA\ D_1}(x)=f_1$ и

$$r_1^{OGA\ D_1}(x)=\sum_{k=2}^{\infty}\frac{1}{\sqrt{2}^k}f_k.$$

Шаг 2. Здесь $(r_1^{OGA\ D_1}(x),f')=(x,f')=\frac{\sqrt{3}+1}{4}$ и

$$\sup_{d\in D_1}\left(r_1^{OGA\ D_1}(x),d\right)=\max\left(\frac{1}{2},\frac{\sqrt{3}+1}{4}\right)=\frac{\sqrt{3}+1}{4},$$

при этом супремум достигается только на векторе f'. Отсюда следует, что $e_2^{OGA\ D_1}(x)=f'$ и

$$r_2^{OGA\ D_1}(x)=\sum_{k=2}^{\infty}\frac{1}{\sqrt{2}^k}f_k-\frac{\sqrt{3}+1}{4}\left(\frac{\sqrt{3}}{2}f_2+\sum_{k=3}^{\infty}\frac{1}{\sqrt{2}^k}f_k\right)=$$

$$=\left(\frac{1}{2}-\frac{\sqrt{3}+1}{4}\cdot\frac{\sqrt{3}}{2}\right)f_2+\left(1-\frac{\sqrt{3}+1}{4}\right)\sum_{k=3}^{\infty}\frac{1}{\sqrt{2}^k}f_k=$$

$$=\frac{1-\sqrt{3}}{8}f_2+\frac{3-\sqrt{3}}{4}\sum_{k=3}^{\infty}\frac{1}{\sqrt{2}^k}f_k.$$

На шаге 3 получаем, что

$$\sup_{d\in D_1}\left(r_2^{OGA\ D_1}(x),d\right)=\max\left(\frac{\sqrt{3}-1}{8},\frac{3-\sqrt{3}}{4}\cdot\frac{1}{\sqrt{2}^4}\right)=\frac{\sqrt{3}-1}{8},$$

и супремум достигается на единственном векторе — f_2. Следовательно, в качестве вектора разложения $e_3^{OGA\ D_1}(x)$ однозначно выбирается f_2, откуда получаем, что $r_3^{OGA\ D_1}(x)=x-\mathrm{Proj}_{<f_1,f_2,f'>}x=0$.

Итак, данный алгоритм сходится за три шага.

2) Ортогональный жадный алгоритм по словарю D_2, очевидно, сходится за один шаг.

3) Теперь рассмотрим ортогональный жадный алгоритм по паре словарей D_1 и D_2.

Первый шаг данного алгоритма совпадает с первым шагом ортогонального жадного алгоритма по словарю D_1, рассмотренного выше. Следовательно, $e_1^{OGA\ (D_1,D_2)}(x) = f_1$ и

$$r_1^{OGA\ (D_1,D_2)}(x) = \sum_{k=2}^{\infty} \frac{1}{\sqrt{2}^k} f_k.$$

Шаг 2 проводится по словарю D_2. Здесь

$$(r_1^{OGA\ (D_1,D_2)}(x), f) = \sum_{k=2}^{\infty} \frac{1}{2^k} = \frac{1}{2},$$

$$\left(r_1^{OGA\ (D_1,D_2)}(x), \sqrt{\frac{2}{3}} f_2 + \sqrt{\frac{1}{3}} f_3\right) = \frac{1}{\sqrt{2}^2} \cdot \sqrt{\frac{2}{3}} + \frac{1}{\sqrt{2}^3} \cdot \sqrt{\frac{1}{3}} = \sqrt{\frac{3}{8}},$$

поэтому

$$\sup_{d \in D_2} \left(r_1^{OGA\ (D_1,D_2)}(x), d\right) = \max\left(\frac{1}{2}, \sqrt{\frac{3}{8}}, \frac{1}{\sqrt{2}^3}\right) = \sqrt{\frac{3}{8}},$$

супремум достигается на единственном векторе $\sqrt{\frac{2}{3}} f_2 + \sqrt{\frac{1}{3}} f_3$. Значит, $e_2^{OGA\ (D_1,D_2)}(x) = \sqrt{\frac{2}{3}} f_2 + \sqrt{\frac{1}{3}} f_3$ и

$$r_2^{OGA\ (D_1,D_2)}(x) = \sum_{k=4}^{\infty} \frac{1}{\sqrt{2}^k} f_k.$$

Предположив, что

$$\{e_i^{OGA\ (D_1,D_2)(x)}\}_{i=1}^{n-1} = \left\{f_1, \sqrt{\frac{2}{3}} f_2 + \sqrt{\frac{1}{3}} f_3\right\} \cup \{f_i\}_{i=4}^{n}$$

(здесь при $n = 3$ второе множество в объединении подразумевается пустым) и, соответственно,

$$r_{n-1}^{OGA\ (D_1,D_2)}(x) = \sum_{k=n+1}^{\infty} \frac{1}{\sqrt{2}^k} f_k,$$

покажем, что на следующем шаге верно $e_n^{OGA\ (D_1,D_2)}(x) = f_{n+1}$ и

$$r_n^{OGA\ (D_1,D_2)}(x) = \sum_{k=n+2}^{\infty} \frac{1}{\sqrt{2}^k} f_k.$$

На шаге n по D_i

$$(r_{n-1}^{OGA\ (D_1,D_2)}(x), f) = (r_{n-1}^{OGA\ (D_1,D_2)}(x), f') = \sum_{k=n+1}^{\infty} \frac{1}{2^k} = \frac{1}{2^n},$$

откуда следует, что

$$\sup_{d \in D_i} \left(r_{n-1}^{OGA\ (D_1,D_2)}(x), d \right) = \max \left(\frac{1}{2^n}, \frac{1}{\sqrt{2}^{n+1}} \right) = \frac{1}{\sqrt{2}^{n+1}},$$

и супремум достигается только на векторе f_{n+1}. Значит, действительно, $e_n^{OGA\ (D_1,D_2)}(x) = f_{n+1}$ и

$$r_n^{OGA\ (D_1,D_2)}(x) = \sum_{k=n+2}^{\infty} \frac{1}{\sqrt{2}^k} f_k.$$

Таким образом, алгоритм не сходится за конечное число шагов.

Теорема 4.5 доказана.

Также справедлива следующая теорема для скорости сходимости (для неортогональных словарей).

Теорема 4.6. *Существуют*

I. *ортонормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что для некоторой монотонно убывающей к нулю положительной последовательности $\{s_n\}$ при всех $n \in \mathbb{N}$ верны соотношения*

$$c_{2^n-1}^{OGA\ D_1}(x) \geqslant s_n, \quad c_{2^n-1}^{OGA\ D_2}(x) \geqslant s_n, \quad c_{2n+2}^{OGA\ (D_1,D_2)}(x) = s_n,$$

II. *нормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что для некоторой монотонно убывающей к нулю положительной* *последовательности $\{s_n\}$ при всех $n \in \mathbb{N}$ верны соотношения*

$$c_{3n}^{OGA\ D_1}(x) = s_n, \quad c_n^{OGA\ D_2}(x) \geqslant s_n, \quad c_{2^{n+1}-2}^{OGA\ (D_1,D_2)}(x) = s_n.$$

Доказательство первого пункта данной теоремы также почти дословно совпадает с доказательством первого пункта теоремы 4.2.

Доказательство пункта II.

Для $i \in \mathbb{N}$ определим векторы d_i и d_i': положим

$$d_i = \sum_{k=1}^{2^i} \frac{1}{\sqrt{2}^k} f_{\sigma_i+k} + \frac{1}{\sqrt{2}^{2^i}} f_{\sigma_i+2^i+1},$$

$$d_i' = \frac{\sqrt{3}}{2} f_{\sigma_i+2} + \sum_{k=3}^{2^i} \frac{1}{\sqrt{2}^k} f_{\sigma_i+k} + \frac{1}{\sqrt{2}^{2^i}} f_{\sigma_i+2^i+1},$$

где $\sigma_i = \sum\limits_{k=1}^{i-1}(2^k+1) = 2^i + i - 2$ при $i \neq 1$ и $\sigma_1 = 0$. В этом случае $\|d_i\|_2 = 1$ и $\|d_i'\|_2 = 1$ при всех $i \in \mathbb{N}$, а системы $\{d_i'\}_{i\in\mathbb{N}}$ и $\{d_i\}_{i\in\mathbb{N}}$ являются ортонормированными.

Рассмотрим нормированные словари

$$D_1 = \{d_i'\}_{i\in\mathbb{N}} \cup F \setminus \{f_{\sigma_i+3}\}_{i\in\mathbb{N}},$$

$$D_2 = \{d_i\}_{i\in\mathbb{N}} \cup \left\{ \sqrt{\frac{2}{3}} f_{\sigma_i+2} + \sqrt{\frac{1}{3}} f_{\sigma_i+3} \right\}_{i\in\mathbb{N}} \cup F \setminus \{f_{\sigma_i+2}\}_{i\in\mathbb{N}}$$

и приближаемый вектор $x = \sum\limits_{i=0}^{\infty} \frac{1}{2^{2^i}} d_i \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$.

Векторы d_i имеют единственное представление словарём D_1:

$$d_i = d_i' + \frac{1}{\sqrt{2}} f_{\sigma_i+1} + \frac{1-\sqrt{3}}{2} f_{\sigma_i+2}, \quad i \in \mathbb{N}.$$

Значит, $\|d_i\|_{\mathcal{A}_1(D_1)} = 1 + \frac{1}{\sqrt{2}} + \frac{1-\sqrt{3}}{2} = \frac{3+\sqrt{2}-\sqrt{3}}{2}$ при $i \in \mathbb{N}$.

Несложно видеть, что

$$\|x\|_{\mathcal{A}_1(D_1)} = \sum_{i=0}^{\infty} \frac{1}{2^{2^i}} \|d_i\|_{\mathcal{A}_1(D_1)} = \frac{3+\sqrt{2}-\sqrt{3}}{2} \sum_{i=0}^{\infty} \frac{1}{2^{2^i}}.$$

Обозначим сумму ряда $\sum\limits_{i=0}^{\infty} \frac{1}{2^{2^i}}$ через s. Тогда для вектора x справедливы равенства

$$\|x\|_{\mathcal{A}_1(D_2)} = \sum_{i=0}^{\infty} \frac{1}{2^{2^i}} \|d_i\|_{\mathcal{A}_1(D_2)} = \sum_{i=0}^{\infty} \frac{1}{2^{2^i}} = s$$

и

$$\|x\|_{\mathcal{A}_1(D_1,D_2)} = \frac{3+\sqrt{2}-\sqrt{3}}{2} s.$$

Ортогональный жадный алгоритм по словарю D_1 на каждых трёх шагах вида $3i+1, 3i+2, 3i+3$ аналогичен ортогональному жадному алгоритму по первому словарю из доказательства второго пункта теоремы 4.5 и за каждые три таких шага приближается слагаемое $\frac{1}{2^{2^i}} d_i$. Таким образом, на шаге $3n$ получаем

$$r_{3n}^{OGA\ D_1}(x) = \sum_{i=n+1}^{\infty} \frac{1}{2^{2^i}} d_i.$$

С учётом этого, положим

$$s_n = \frac{\sqrt{\sum\limits_{i=n+1}^{\infty} \frac{1}{2^{2^{i+1}}}}}{\frac{3+\sqrt{2}-\sqrt{3}}{2} s} = \frac{\|r_{3n}^{OGA\ D_1}(x)\|_2}{\|x\|_{\mathcal{A}_1(D_1)}} = c_{3n}^{OGA\ D_1}(x).$$

Рассмотрим далее ортогональный жадный алгоритм по словарю D_2. Очевидно, на каждом шаге $e_n^{OGA\ D_2}(x) = d_{n-1}$. Тогда

$$r_n^{OGA\ D_2}(x) = \sum_{i=n+1}^{\infty} \frac{1}{2^{2^i}} d_i,$$

и для скорости сходимости получаем оценку

$$c_n^{OGA\ D_2}(x) = \frac{\sqrt{\sum\limits_{i=n+1}^{\infty} \frac{1}{2^{2^{i+1}}}}}{s} = \frac{3+\sqrt{2}-\sqrt{3}}{2} s_n > s_n.$$

Перейдём теперь к рассмотрению ортогонального жадного алгоритма по паре словарей D_1 и D_2. Чтобы сократить изложение доказательства, будем опираться на доказательство второго пункта теоремы 4.5, не рассматривая шаги.

Несложно проверить, что при всех $n \in \mathbb{N}$ верно

$$\{e_k^{OGA\ (D_1,D_2)}(x)\}_{k=2^n-2+1}^{2^n-2+2^n} = \left\{ f_{\sigma_n+1}, \sqrt{\frac{2}{3}} f_{\sigma_n+2} + \sqrt{\frac{1}{3}} f_{\sigma_n+3} \right\} \cup \{f_{\sigma_n+k}\}_{k=4}^{2^n+1},$$

где $\sum\limits_{i=1}^{n} 2^i = 2^{n+1} - 2$, и для остатка выполняется равенство

$$r_{2^{n+1}-2}^{OGA\ (D_1,D_2)}(x) = \sum_{i=n+1}^{\infty} \frac{1}{2^{2^i}} d_i.$$

Следовательно,

$$c_{2^{n+1}-2}^{OGA\ (D_1,D_2)}(x) = s_n.$$

Итак, *теорема доказана.*

Сравнивая ортогональный жадный алгоритм по паре словарей D_1 и D_2 с ортогональным жадным алгоритмом по объединению данных словарей $D = D_1 \cup D_2$, получаем следующую теорему.

Теорема 4.7. *Существуют*

I. *нормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что*

   1) *ортогональный жадный алгоритм по словарю $D = D_1 \cup D_2$ сходится за конечное число шагов,*

   2) *ортогональный жадный алгоритм по паре словарей D_1 и D_2 не сходится за конечное число шагов.*

II. *нормированные словари D_1 и D_2 и вектор $x \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2)$ такие, что*

   1) *ортогональный жадный алгоритм по словарю $D = D_1 \cup D_2$ не сходится за конечное число шагов,*

   2) *ортогональный жадный алгоритм по паре словарей D_1 и D_2 сходится за конечное число шагов.*

Для *доказательства пункта I* выберем те же нормированные словари и приближаемый вектор, что и в доказательстве второго пункта теоремы 4.5:

$$D_1 = \{f'\} \cup F \setminus \{f_3\} \quad \text{и} \quad D_2 = \left\{ f, \sqrt{\frac{2}{3}} f_2 + \sqrt{\frac{1}{3}} f_3 \right\} \cup F \setminus \{f_2\},$$

где

$$f = \sum_{k=1}^{\infty} \frac{1}{\sqrt{2}^k} f_k, \quad f' = \frac{\sqrt{3}}{2} f_2 + \sum_{k=3}^{\infty} \frac{1}{\sqrt{2}^k} f_k,$$

и положим $x = f$.

Тогда стандартный ортогональный жадный алгоритм по словарю $D = D_1 \cup D_2$ сходится за один шаг, а ортогональный жадный алгоритм по паре словарей D_1 и D_2 не сходится за конечное число шагов (см. доказательство теоремы 4.5).

Доказательство пункта II.

Определим вектор f следующим образом:

$$f = \sum_{k=2}^{\infty} \frac{1}{\sqrt{2}^{k-1}} f_k.$$

Очевидно, что f — единичный вектор в ℓ_2. В качестве словарей выберем нормированные словари:

$$D_1 = \left\{ f,\ \sqrt{\frac{1}{3}} f_1 - \sqrt{\frac{2}{3}} f_2 \right\} \cup F \setminus \{f_1\} \text{ и } D_2 = \left\{ \sqrt{\frac{4}{7}} f_1 + \sqrt{\frac{2}{7}} f_2 + \sqrt{\frac{1}{7}} f_3 \right\} \cup F,$$

а в качестве приближаемого вектора возьмём

$$x = \sum_{k=1}^{\infty} \frac{1}{\sqrt{2}^k} f_k \in \mathcal{A}_1(D_1) \cap \mathcal{A}_1(D_2).$$

Рассмотрим ортогональный жадный алгоритм по словарю

$$D = D_1 \cup D_2 = \left\{ f,\ \sqrt{\frac{1}{3}} f_1 - \sqrt{\frac{2}{3}} f_2, \sqrt{\frac{4}{7}} f_1 + \sqrt{\frac{2}{7}} f_2 + \sqrt{\frac{1}{7}} f_3 \right\} \cup F.$$

Шаг 1. Для скалярных произведений в данном случае справедливы равенства

$$(x, f) = \sum_{k=2}^{\infty} \frac{1}{\sqrt{2}^{k-1}} \cdot \frac{1}{\sqrt{2}^{k}} = \frac{1}{\sqrt{2}},$$

$$\left(x, \sqrt{\frac{1}{3}} f_1 - \sqrt{\frac{2}{3}} f_2\right) = \sqrt{\frac{1}{3}} \cdot \frac{1}{\sqrt{2}} - \sqrt{\frac{2}{3}} \cdot \frac{1}{2} = 0,$$

$$\left(x, \sqrt{\frac{4}{7}} f_1 + \sqrt{\frac{2}{7}} f_2 + \sqrt{\frac{1}{7}} f_3\right) = \sqrt{\frac{4}{7}} \cdot \frac{1}{\sqrt{2}} + \sqrt{\frac{2}{7}} \cdot \frac{1}{2} + \sqrt{\frac{1}{7}} \cdot \frac{1}{2\sqrt{2}} = \frac{\sqrt{7}}{2\sqrt{2}}.$$

Следовательно, поскольку по определению $r_0^{OGA\ D_1 \cup D_2}(x) = x$, то справедливо равенство

$$\sup_{d \in D_1 \cup D_2} \left(r_0^{OGA\ D_1 \cup D_2}(x), d\right) = \max\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{\sqrt{7}}{2\sqrt{2}}\right) = \frac{\sqrt{7}}{2\sqrt{2}},$$

причём супремум достигается только на векторе $\sqrt{\frac{4}{7}} f_1 + \sqrt{\frac{2}{7}} f_2 + \sqrt{\frac{1}{7}} f_3$. Тогда $e_1^{OGA\ D_1 \cup D_2}(x) = \sqrt{\frac{4}{7}} f_1 + \sqrt{\frac{2}{7}} f_2 + \sqrt{\frac{1}{7}} f_3$ и

$$r_1^{OGA\ D_1 \cup D_2}(x) = \sum_{k=4}^{\infty} \frac{1}{\sqrt{2}^{k}} f_k.$$

При $n > 1$ предположим, что

$$\left\{e_i^{OGA\ D_1 \cup D_2}(x)\right\}_{i=1}^{n-1} = \left\{\sqrt{\frac{4}{7}} f_1 + \sqrt{\frac{2}{7}} f_2 + \sqrt{\frac{1}{7}} f_3\right\} \cup \{f_i\}_{i=4}^{n+1}$$

(здесь при $n = 2$ второе множество в объединении считаем пустым) и, соответственно,

$$r_{n-1}^{OGA\ D_1 \cup D_2}(x) = \sum_{k=n+2}^{\infty} \frac{1}{\sqrt{2}^{k}} f_k.$$

В этом случае покажем, что на следующем шаге $e_n^{OGA\ D_1 \cup D_2}(x) = f_{n+2}$ и

$$r_n^{OGA\ D_1 \cup D_2}(x) = \sum_{k=n+3}^{\infty} \frac{1}{\sqrt{2}^{k}} f_k.$$

То есть на каждом шаге остаток будет ненулевым и алгоритм не будет сходится за конечное число шагов.

Рассмотрим шаг n.

$$(r_{n-1}^{OGA\ D_1 \cup D_2}(x), f) = \sum_{k=n+2}^{\infty} \frac{1}{\sqrt{2}^k} \cdot \frac{1}{\sqrt{2}^{k-1}} = \frac{1}{\sqrt{2}^{2n+1}}$$

и при $n > 1$

$$\sup_{d \in D_1 \cup D_2} \left(r_1^{OGA\ D_1 \cup D_2}(x), d \right) = \left(\frac{1}{\sqrt{2}^{2n+1}}, \frac{1}{\sqrt{2}^{n+2}} \right) = \frac{1}{\sqrt{2}^{n+2}},$$

причём супремум достигается на единственном векторе — f_{n+2}. Отсюда $e_n^{OGA\ D_1 \cup D_2}(x) = f_{n+2}$ и

$$r_n^{OGA\ D_1 \cup D_2}(x) = \sum_{k=n+3}^{\infty} \frac{1}{\sqrt{2}^k} f_k.$$

Таким образом, ортогональный жадный алгоритм не сходится за конечное число шагов.

Перейдём к рассмотрению ортогонального жадного алгоритма по паре словарей D_1 и D_2.

На первом шаге, учитывая вычисленые выше скалярные произведения для $r_0^{OGA\ (D_1, D_2)}(x) = x$, получаем

$$\sup_{d \in D_1} \left(r_0^{OGA\ (D_1, D_2)}(x), d \right) = \max \left(\frac{1}{\sqrt{2}}, \frac{1}{2} \right) = \frac{1}{\sqrt{2}},$$

и супремум достигается только на векторе f. Следовательно, получаем, что $e_1^{OGA\ (D_1, D_2)}(x) = f$ и

$$r_1^{OGA\ (D_1, D_2)}(x) = \frac{1}{\sqrt{2}} f_1.$$

Шаг 2.

$$\sup_{d \in D_2} \left(r_1^{OGA\ (D_1, D_2)}(x), d \right) = \max \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \cdot \sqrt{\frac{4}{7}} \right) = \frac{1}{\sqrt{2}},$$

и при этом супремум достигается на единственном векторе — f_1. Значит, $e_2^{OGA\ (D_1, D_2)}(x) = f_1$ и

$$r_2^{OGA\ (D_1, D_2)}(x) = 0,$$

то есть алгоритм сходится за два шага.

Теорема доказана.

Замечание. Также можно доказать аналогичную теорему для неполных "словарей", например, выбрав в каждом из пунктов соответствующие "словари" $D_1 \setminus \{f_5\}$ и $D_2 \setminus \{f_6\}$.

Заключение.

Обзор проведенного исследования.

Тематика диссертации относится к области математического анализа. В диссертации изучаются вопросы сходимости и скорости сходимости жадных приближений. Установлены следующие основные результаты.

1. Показано, что общие результаты о скорости сходимости слабых (ортогональных) жадных приближений в случае ортогонального словаря могут быть уточнены, причём полученное уточнение асимптотически неулучшаемо.

2. Показано, что достаточное условие сходимости слабого (ортогонального) жадного алгоритма в случае ортогонального словаря может быть ослаблено, причём полученное условие асимптотически неулучшаемо.

3. Показано, что достаточное условие сходимости в случае расширения ортогонального словаря одним вектором не может быть ослаблено для слабого жадного алгоритма и слабого ортогонального жадного алгоритма.

4. Показано, что скорость сходимости чисто жадного алгоритма по ортогональному словарю для индивидуального вектора может быть выше, чем скорость сходимости чисто жадного алгоритма по расширению данного словаря одним вектором.

5. Показано, что ортогональный жадный алгоритм по системе возможно неполных словарей сходится к приближаемому вектору гильбертова пространства.

6. Показано, что сходимость стандартного жадного алгоритма для индивидуального вектора может быть быстрее, чем сходимость жадного алгоритма по паре соответствующих словарей в случае чисто жадного алгоритма и в случае ортогонального жадного алгоритма; также доказано, что реализуема и обратная ситуация.

Рекомендации и перспективы по дальнейшей разработке темы.

Полученные результаты создают основу для последующего изучения данной области. Дальнейшие исследования по тематике диссертации могут проводиться, в частности, в следующих направлениях.

1. Исследование возможности ослабления достаточного условия сходимости в случае словаря, не являющегося расширением ортогонального словаря, для слабого жадного алгоритма и для слабого ортогонального жадного алгоритма.

2. Сравнение скорости сходимости стандартного жадного алгоритма и скорости сходимости жадного алгоритма по паре соответствующих словарей на индивидуальном векторе для слабого жадного алгоритма и слабого ортогонального жадного алгоритма в случае различных ослабляющих последовательностей.

3. Исследование вопроса сильной сходимости чисто жадного алгоритма по системе возможно неполных словарей, а также исследование скорости сходимости жадных алгоритмов по системе возможно неполных словарей на классах сходимости в случае чисто жадного алгоритма и ортогонального жадного алгоритма.

4. Сравнение классовой оценки скорости сходимости стандартного жадного алгоритма и классовой оценки жадного алгоритма по паре словарей для чисто жадного алгоритма и для ортогонального жадного алгоритма.

Автор выражает **благодарность** своим научным руководителям доктору физико-математических наук профессору Тарасу Павловичу Лукашенко и кандидату физико-математических наук Владимиру Владимировичу Галатенко за постановку задач, плодотворные обсуждения, ценные советы и постоянное внимание к работе, а также сотрудникам кафедры математического анализа и кафедры теории функции и функционального анализа за доброжелательное отношение и поддержку.

Литература

[1] Бородин П. А. Жадные приближения произвольным множеством // Изв. РАН. Сер. матем. — 2020. — Т. 84, № 2. — С. 43–59.

[2] Бородин П. А., Копецка Е. Слабые пределы последовательных проекций и жадных шагов // Теория приближений, функциональный анализ и приложения, Сборник статей. К 70-летию академика Бориса Сергеевича Кашина, Труды МИАН. — 2022. — Т. 319. — С. 64–72.

[3] Валов М. А. Конический жадный алгоритм // Изв. РАН. Сер. матем. — 2022. — Т. 112, № 2. — С. 163–169.

[4] Деревенцов А. В. Сравнение скорости сходимости чисто жадного и ортогонального жадного алгоритмов // Матем. заметки. — 2012. — Т. 92, № 4. — С. 528–532.

[5] Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. — М.: Наука. — 1976.

[6] Лившиц Е. Д. О нижних оценках скорости сходимости жадных алгоритмов // Изв. РАН. Сер. матем. — 2009. — Т. 73, № 6. — С. 125–144.

[7] Лившиц Е. Д. О скорости сходимости чисто жадного алгоритма // Матем. заметки. — 2004. — Т. 76, № 4. — С. 539–552.

[8] Сильниченко А. В. О скорости сходимости жадных алгоритмов // Матем. заметки. — 2004. — Т. 76, № 4. — С. 628–632.

[9] Стечкин С. Б. Об абсолютной сходимости ортогональных рядов // Докл. АН СССР. — 1955. — Т. 102, № 1. — С. 37–40.

[10] Barron A. R. Universal approximation bounds for superpositions of a sigmoidal function // IEEE Trans. Inform. Theory. — 1993. — Vol. 39, № 3. — P. 930–945.

[11] Cauchy A.-L. Leçons sur les applications du calcul infinitésimal a la géométrie. — 1826–1828.

[12] Cauchy A.-L. Résumé des leçons données à l'École royale polytechnique sur le calcul infinitésimal. — 1823.

[13] Cauchy A.-L. Théorie de la propagation des ondes à la surface d'un fluide pesant d'une profondeur indéfinie // Mémoires présentés par divers savants à l'Académie royale des sciences de l'Institut de France et imprimés par son ordre. Sciences mathématiques et physiques. — 1827. — Vol. 1. — P. 157–169.

[14] DeVore R. A., Temlyakov V. N. Some remarks on greedy algorithms // Adv. Comput. Math. — 1996. — Vol. 5, № 1. — P. 173–187.

[15] Dirichlet L. P. G. Sur la convergence des séries trigonométriques qui servent à représenter une fonction arbitraire entre des limites données // J. für die reine und angewandte Mathematik. — 1829. — Vol. 4. — P. 157–169.

[16] Fourier J.-B. J. Mémoire sur la propagation de la chaleur dans les corps solides. — 1811.

[17] Fourier J.-B. J. Théorie analytique de la chaleur. — 1822.

[18] Haar A. Zur Theorie der orthogonalen Functionsysteme // Math. Ann. — 1910. — Vol. 69. — P. 331–371; — 1912. — Vol. 71. — P. 33–53.

[19] Haar A. Über die Multiplikationstabelle der orthogonalen Funktionensysteme // Math. Zeit. — 1930. — Vol. 32. — P. 769–798.

[20] Haar A. Über einige Eigenschaften der orthogonalen Funktionensysteme // Math. Zeit. — 1929. — Vol. 31. — P. 128–137.

[21] Jones L. K. A simple lemma on greedy approximation in Hilbert space and convergence rates for projection pursuit regression and neural network training // Ann. Statist. — 1992. — Vol. 20, № 1. — P. 608–613.

[22] Klusowski J. M., Siegel J. W. Sharp convergence rates for matching pursuit // arXiv:2307.07679 [stat.ML].

[23] Konyagin S. V., Temlyakov V. N. Rate of convergence of pure greedy algorithms // East J. Approx. — 1999. — Vol. 5, № 4. — P. 493–499.

[24] Rademacher H. Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen // Math. Ann. — 1922. — Vol. 87. — P. 112–138.

[25] Steinhaus H. An example of a thoroughly divergent orthogonal development // Proc. London Math. Soc. — 1920. — Vol. 20, № 2. — P. 123–126.

[26] Steinhaus H. Przyklady rozwiniec biortogonalnych // Mathesis Polska. — 1934. — Vol. 9. — P. 33–40.

[27] Steinhaus H. Sur les développements orthogonaux // Bull. Acad. Polonaise. — 1926. — P. 11–39.

[28] Steinhaus H. Sur quelques applications du calcul fonctionnel à la théorie de séries orthogonales // Studia Math. — 1929. — Vol. 1. — P. 191–200.

[29] Temlyakov V. Greedy Approximation (Cambridge Monographs on Applied and Computational Mathematics). — Cambridge: Cambridge University Press. — 2011.

[30] Temlyakov V. N. Weak greedy algorithms // Adv. Comput. Math. — 2000. — Vol. 12, № 2–3. — P. 213–227.

[31] Weyl H. Über die Konvergenz von Reihen, die nach Orthogonal-funktionen vortschreiten // Math. Ann. — 1909. — Vol. 67. — P. 225–245.

[32] Zygmund A. Sur l'application de la première moyenne arithmétique dans la théorie des séries de fonctions orthogonales // Fund. Math. — 1927. — Vol. 10. — P. 356–362.

[33] Zygmund A. Un theoreme sur les series orthogonales // Studia Math. — 1930. — Vol. 2. — P. 181–182.

Работы автора по теме диссертации

Статьи в рецензируемых научных изданиях, рекомендованных для защиты в диссертационном совете МГУ.011.3 по специальности 1.1.1 — “вещественный, комплексный и функциональный анализ” и входящих в базы цитирования Scopus, Web of Science и RSCI

[34] Орлова А.С. Скорость сходимости слабых жадных приближений по ортогональным словарям // Вестн. Моск. ун-та. Сер. 1. Матем. Механ. — 2017. — № 2. — С. 68–72.

[35] Орлова А.С. Сходимость слабого ортогонального жадного алгоритма при добавлении одного вектора к ортогональному словарю // Вестн. Моск. ун-та. Сер. 1. Матем. Механ. — 2022. — № 5. — С. 17–25.

[36] Орлова А.С. Сравнение чисто жадного алгоритма и чисто жадного алгоритма по паре словарей // Вестн. Моск. ун-та. Сер. 1. Матем. Механ. — 2023. — № 2. — С. 3–11.

Иные публикации

[37] Орлова А.С. Асимптотически точная оценка скорости сходимости слабых жадных приближений по ортогональным словарям // Современные методы теории функций и смежные проблемы : материалы Международной конференции : Воронежская зимняя математическая школа (26 января – 1 февраля 2017 г.). — Воронеж: Издательский дом ВГУ, 2017. — С. 159.

[38] Орлова А.С. Индивидуальные и классовые оценки скорости сходимости слабых жадных приближений по ортогональным словарям // Современные методы теории функций и смежные проблемы : материалы Международной конференции : Воронежская зимняя математическая школа (28 января – 2 февраля 2019 г.). — Воронеж: Издательский дом ВГУ, 2019. — С. 198–199.

[39] Орлова А.С. Сходимость слабых жадных приближений по расширениям ортогональных словарей // Материалы Международного молодеж-

ного научного форума «ЛОМОНОСОВ–2021» [Электронный ресурс]. — М.: МАКС Пресс, 2021.

[40] Орлова А.С. Сравнение скорости сходимости стандартного чисто жадного алгоритма и чисто жадного алгоритма по паре словарей // Современные проблемы теории функций и их приложения : Материалы 21–й международной Саратовской зимней школы (Саратов, 31 января – 4 февраля 2022 г.). — Саратов: Саратовский университет. Текст: электронный, 2022. — С. 212–214.

[41] Орлова А.С. Ортогональные жадные алгоритмы и их модификации по паре словарей // Современные методы теории функций и смежные проблемы : материалы Международной конференции : Воронежская зимняя математическая школа (27 января – 1 февраля 2023 г.). — Воронеж: Издательский дом ВГУ, 2023. — С. 276–278.