Math-Net.Ru

Общероссийский математический портал

Е. Д. Лившиц, Об эффективности ортогонального жадного алгоритма в задаче о сжатых измерениях, *Матем. сб.*, 2012, том 203, номер 2, 33–44

DOI: <https://doi.org/10.4213/sm7827>

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

Параметры загрузки:

IP: 3.82.47.9

26 апреля 2023 г., 16:20:53

УДК 517.518.8

Е. Д. Лившиц

Об эффективности ортогонального жадного алгоритма в задаче о сжатых измерениях

В работе показано, что если матрица Φ обладает свойством ограниченной изометрии (СОИ) порядка $[CK^{1.2}]$ с изометрической константой $\delta = cK^{-0.2}$ и имеет когерентность меньшую, чем $1/(20K^{0.8})$, то произвольный K -разреженный сигнал может быть точно восстановлен по сжатым измерениям $y = \Phi x$ с помощью ортогонального жадного алгоритма (Orthogonal Matching Pursuit) за не более чем $[CK^{1.2}]$ итераций. Из этого результата вытекает, что произвольный K -разреженный сигнал может быть восстановлен с помощью жадного алгоритма по $M = O(K^{1.6} \log N)$ измерениям.

Библиография: 23 названия.

Ключевые слова: ортогональный жадный алгоритм, сжатые измерения, когерентность, свойство ограниченной изометрии, разреженность.

§ 1. Введение

В последние годы внимание многих специалистов по теории приближения привлекла задача о сжатых измерениях (compressed sensing), относящаяся к теории передачи сигналов. Оказалось, что если передаваемый сигнал является разреженным (sparse), то можно передать информацию о нем с помощью значительно меньшего числа измерений, чем в классических методах.

Результаты, полученные в этом направлении, вызвали широкий резонанс, потому что, во-первых, во многих важнейших приложениях обрабатываемые сигналы являются разреженными, а во-вторых, при разработке методов сжатия/восстановления были использованы глубокие результаты из многих математических дисциплин. К настоящему моменту различным аспектам задачи о сжатых измерениях посвящены более ста работ, в частности обзорные статьи [1]–[3]. В российских математических журналах сжатым измерениям посвящена работа Б. С. Кашина и В. Н. Темлякова [4].

В настоящей статье показана возможность применения жадных алгоритмов в задаче о сжатых измерениях. Начнем с постановки задачи и основных определений.

Пусть заданы числа $M, N \in \mathbb{N}$, $M < N$. По сигналу $x \in \mathbb{R}^N$ вычисляются M неадаптивных линейных измерений, иными словами, фиксируется $M \times N$ матрица $\Phi = \{\varphi_{i,j}\}_{1 \leq i \leq M, 1 \leq j \leq N}$, не зависящая от x , и вычисляется $y = \Phi x$, $y \in \mathbb{R}^M$. Восстановление элемента $x \in \mathbb{R}^N$ по его измерениям $y \in \mathbb{R}^M$ осуществляется с помощью (нелинейной) функции восстановления $\hat{x}_\Phi: \mathbb{R}^M \rightarrow \mathbb{R}^N$. Ясно, что, так как $M < N$, невозможно осуществить точное восстановление

Работа выполнена при поддержке Российского фонда фундаментальных исследований (гранты № 11-01-00476 и № 10-01-00442).

произвольного сигнала $x \in \mathbb{R}^N$. То есть для произвольных Φ и $\hat{x}_\Phi$ найдется $x \in \mathbb{R}^N$, $x \neq 0$, для которого $y = \Phi x = 0$ и $\hat{x}_\Phi(\Phi x) \neq x$. В то же время, если наложить на сигнал x некоторые ограничения, то точное восстановление становится возможным.

Обозначим через $\|x\|_0$ число ненулевых компонент сигнала $x \in \mathbb{R}^N$. Будем говорить, что сигнал x является K -разреженным (K -sparse), если $\|x\|_0 \leq K$.

Основной задачей сжатых измерений является построение матрицы Φ и быстрого и устойчивого алгоритма восстановления $\hat{x}_\Phi$, которые бы обеспечили точное восстановление K -разреженных сигналов, $K < M$:

$$\text{для любого } x \in \mathbb{R}^N, \quad \|x\|_0 \leq K, \quad \hat{x}_\Phi(y) = x, \quad \text{где } y = \Phi x,$$

и определение троек (M, N, K) , для которых такое восстановление возможно.

Будем говорить, что матрица Φ размера $M \times N$, $M < N$, является невырожденной, если любая квадратная подматрица Φ размера $M \times M$ (состоящая из M произвольных столбцов матрицы Φ) имеет максимальный ранг.

Легко заметить, что для фиксированной матрицы Φ максимальную точность на разреженных векторах обеспечивает следующий алгоритм:

$$\hat{x}_\Phi^0(y) = \operatorname{argmin}\{\|z\|_0 : \Phi z = y, z \in \mathbb{R}^N\}.$$

В случае невырожденной матрицы Φ для любого $x \in \mathbb{R}^N$, $\|x\|_0 = K \leq M/2$, выполняется равенство

$$x = \hat{x}_\Phi^0(y), \quad y := \Phi x.$$

Действительно, в противном случае нашелся бы $x' \neq x$, $\|x'\|_0 \leq K$ такой, что $\Phi x' = y = \Phi x$. Тогда $\Phi(x - x') = 0$, $\operatorname{supp}(x' - x) \leq 2K \leq M$, что противоречило бы невырожденности Φ .

Однако применение этого алгоритма на практике невозможно. Дело в том, что в случае, если Φ невырожденна и $y = \Phi x$, $\|x\|_0 = K$, то для каждого множества $W \subset \{1, \dots, N\}$, $\#W < K$, необходимо будет установить, что система $\Phi z = y$, $\operatorname{supp} z = W$, не имеет решений. Таким образом, нахождение $\hat{x}_\Phi^0(y)$ потребует не менее чем $C(K)N^K$ действий. Такие вычисления невозможно выполнить на современных компьютерах. Кроме того, алгоритм $\hat{x}_\Phi^0(\cdot)$ очень неустойчив.

В настоящей работе рассматриваются два наиболее распространенных метода восстановления разреженных сигналов: метод l_1 минимизации и ортогональный жадный алгоритм.

Метод l_1 минимизации (l_1 minimization) или преследования базиса (basis pursuit) был предложен Д. Донохо с соавторами [3], [5].

$$\hat{x}_\Phi^1(y) = \operatorname{argmin}\{\|z\|_1 : \Phi z = y\}.$$

В отличие от алгоритма $\hat{x}_\Phi^0(\cdot)$ для нахождения $\hat{x}_\Phi^1(y)$ требуется найти решение выпуклой задачи, которое может быть (приближенно) найдено за полиномиальное (от N) время.

Другим подходом к восстановлению разреженных сигналов является использование жадных алгоритмов. Жадные алгоритмы (Greedy Algorithms) интенсивно изучались в теории приближения начиная с середины 90-х годов прошлого века. С их современным состоянием можно познакомиться в обзоре В. Н. Темлякова [6]. В теорию передачи сигнала они были введены С. Маллатом и Ж. Жангом [7] под именем Matching Pursuit. Ниже приведено определение

ортогонального жадного алгоритма (Orthogonal Greedy Algorithm) в терминах теории передачи сигнала.

Обозначим j -й столбец матрицы Φ через $\varphi_j \in \mathbb{R}^M$, $1 \leq j \leq M$. Будем предполагать, что

$$\|\varphi_j\|_2 = 1, \quad 1 \leq i \leq N. \quad (1.1)$$

Ортогональный жадный алгоритм

Вход: Φ , y .

Инициализация: $r^0 := y$, $x^0 := 0$, $\Lambda^0 = \emptyset$, $l = 0$.

Итерация: положим

$$\begin{aligned} \Lambda^{l+1} &:= \Lambda^l \cup \operatorname{argmax}_j |\langle r^l, \varphi_j \rangle|, \\ x^{l+1} &:= \operatorname{argmin}_{z: \operatorname{supp}(z) \subset \Lambda^{l+1}} \|y - \Phi z\|, \quad r^{l+1} := y - \Phi x^{l+1}. \end{aligned}$$

Если $r_{l+1} = 0$, то алгоритм заканчивает свою работу. В противном случае полагаем $l := l + 1$ и проводим новую итерацию.

Результат: если алгоритм закончил свою работу на l -й итерации, то полагаем что $\hat{x}_\Phi^{\text{OMP}}(y) = x_l$.

Для восстановления разреженных сигналов этот алгоритм начал использоваться в работах А. Гильберт, М. Мутукришнана, Дж. Штраусса [8], Д. Донохо, М. Элада, В. Н. Темлякова [9]. Легко видеть, что для любого $y \in \mathbb{R}^M$ в случае невырожденной матрицы Φ ортогональный жадный алгоритм сойдется за не более чем M итераций, и его сложность линейно зависит от N и полиномиально от M . Несмотря на то, что в ряде аспектов ортогональный жадный алгоритм уступает методу l_1 минимизации (см. ниже), Дж. Троппом [10] было показано, что алгоритмическая простота вычисления $\hat{x}_\Phi^{\text{OMP}}(\cdot)$ позволяет достичь очень хороших результатов в приложениях.

Для того чтобы алгоритмы $\hat{x}_\Phi^1$ и $\hat{x}_\Phi^{\text{OMP}}$ эффективно восстанавливали разреженные сигналы, достаточно, чтобы матрица Φ обладала специальными свойствами: малой когерентностью и/или обладала свойством ограниченной изометрии (СОИ) большого порядка.

ОПРЕДЕЛЕНИЕ 1. Когерентностью матрицы Φ , удовлетворяющей (1.1) называют величину

$$\mu(\Phi) := \sup_{i \neq j} |\langle \varphi_i, \varphi_j \rangle|.$$

ОПРЕДЕЛЕНИЕ 2 (Э. Канделс, Т. Тао; см. [11], [12]). Матрица Φ обладает свойством ограниченной изометрии [Restricted Isometry Property] порядка K с изометрической константой $\delta > 0$, если для любого K -разреженного сигнала $x \in \mathbb{R}^N$ имеет место неравенство

$$(1 - \delta)\|x\|^2 \leq \|\Phi x\|^2 \leq (1 + \delta)\|x\|^2.$$

В качестве модельного примера мы будем рассматривать случайные бернуллиевские матрицы, нормированные по столбцам. Будем считать, что коэффициенты матрицы Φ независимы и принимают значение $\pm M^{-1/2}$ с вероятностью $1/2$. Тогда для Φ будет выполняться (1.1). Известно, что для больших c_μ

с большой вероятностью выполняется неравенство

$$\mu(\Phi) \leq c_\mu M^{-1/2} \log^{1/2} N. \quad (1.2)$$

С большой вероятностью бернуллиевская $M \times N$ матрица Φ будет обладать СОИ порядка K с изометрической константой $\delta > 0$ для

$$K \asymp \frac{\delta^2 M}{\log(N/M)}. \quad (1.3)$$

Впервые неравенство (1.3) в случае $\delta \asymp 1$ было (в других обозначениях) установлено Б. С. Кашиным [13] для оценки поперечников. Спустя 30 лет Э. Кандес и Т. Тао [12] передоказали этот результат и успешно использовали его в своей работе о декодировании с помощью l_1 минимизации. В общем виде (для всех δ) неравенство (1.3) было получено Р. Баранюком, М. Дэвенпортом, Р. Девором и М. Уокиным [14].

ТЕОРЕМА 1 (Э. Кандес, Т. Тао; см. [11]). *Если матрица Φ обладает СОИ порядка K с изометрической константой $\delta_K > 0$, СОИ порядка $2K$ с изометрической константой $\delta_{2K} > 0$ и СОИ порядка $3K$ с изометрической константой δ_{3K} , и при этом $\delta_K + \delta_{2K} + \delta_{3K} < 1$, то для любого K -разреженного сигнала $x \in \mathbb{R}^N$ имеет место равенство*

$$\hat{x}_\Phi^1(\Phi x) = x.$$

Из этой теоремы и оценки (1.3) следует, что с помощью M измерений можно точно восстановить за полиномиальное время K -разреженный сигнал с

$$K \asymp \frac{M}{\log(N/M)}.$$

Перейдем к результатам о восстановлении разреженных сигналов с помощью ортогонального жадного алгоритма $\hat{x}_\Phi^{\text{OGA}}$. Хорошо известен следующий результат.

ТЕОРЕМА 2 (см. [8]–[10], [15]). *Если для матрицы Φ , удовлетворяющей (1.1), и $K \in \mathbb{N}$ выполняется неравенство*

$$K < \frac{1}{2}(\mu(\Phi)^{-1} + 1), \quad (1.4)$$

то для любого K -разреженного сигнала $x \in \mathbb{R}^N$ ортогональный жадный алгоритм сходится не более чем за K итераций и гарантирует точное восстановление x по сжатым измерениям $y := \Phi x$:

$$\hat{x}_\Phi^{\text{OGA}}(y) = x.$$

В. Н. Темляков и П. Желтов [16] показали, что в (1.4) нельзя строгое неравенство ($<$) заменить нестрогим ($\leq$). В дальнейшем многими авторами была исследована устойчивость этого результата [8]–[10], [16]–[18].

В своих недавних работах М. Дэвенпорт и М. Уокин [19], а также Э. Лю и В. Н. Темляков [20] исследовали $\hat{x}_\Phi^{\text{OGA}}$ в предположении, что Φ обладает СОИ.

ТЕОРЕМА 3 (см. [19], [20]). Если матрица Φ обладает СОИ порядка $K + 1$ с изометрической константой

$$\delta = \frac{1}{3K^{1/2}} \text{ (см. [19])}, \quad \delta = \frac{1}{(1 + 2^{1/2})K^{1/2}} \text{ (см. [20])},$$

то для любого K -разреженного сигнала $x \in \mathbb{R}^N$ ортогональный жадный алгоритм сходится не более чем за K итераций и имеет место равенство

$$\hat{x}_{\Phi}^{\text{OGA}}(y) = x, \quad y := \Phi x.$$

Основным результатом настоящей статьи является

ТЕОРЕМА 4. Существуют абсолютные константы $C > 0$ и $c > 0$ такие, что если матрица Φ обладает СОИ порядка $[CK^{1.2}]$ с изометрической константой $\delta = cK^{-0.2}$ и когерентностью $\mu(\Phi) \leq 1/(20K^{0.8})$, то для любого K -разреженного сигнала $x \in \mathbb{R}^N$ ортогональный жадный алгоритм сходится не более чем за $[CK^{1.2}]$ итераций и гарантирует точное восстановление x по сжатым измерениям $y := \Phi x$:

$$\hat{x}_{\Phi}^{\text{OGA}}(y) = x.$$

Применим результаты теорем 2, 3 и 4 к случайным бернуллиевским матрицам в случае $N \gtrsim M^{\alpha}$, $\alpha > 1$. Теорема 2 вместе с (1.2), а также теорема 3 с (1.3) влекут одну и ту же оценку: если фиксировать случайную $M \times N$ бернуллиевскую матрицу Φ , то с большой вероятностью восстановление $\hat{x}_{\Phi}^{\text{OGA}}(\Phi x)$ будет точно для всех K -разреженных $x \in \mathbb{R}^N$, где

$$K \lesssim \left(\frac{M}{\log N} \right)^{1/2},$$

или, другими словами, для восстановления K -разреженных сигналов с помощью ортогонального жадного алгоритма будет достаточно M измерений, где

$$M \lesssim K^2 \log N.$$

Совместное применение теоремы 4 с неравенствами (1.2) и (1.3) позволяет улучшить приведенные выше оценки: восстановление K -разреженных сигналов с помощью ортогонального жадного алгоритма и случайных бернуллиевских матриц будет точно (с большой вероятностью) для

$$K \lesssim \left(\frac{M}{\log N} \right)^{5/8},$$

для восстановления K -разреженных сигналов достаточно

$$M \lesssim K^{1.6} \log N$$

измерений.

Значительно меньше известно об оценках снизу. Х. Раухут [21] доказал, что если $M \leq \tilde{c}K^{\beta}$, $\beta = 3/2$, то для большинства случайных $M \times N$ матриц существует K -разреженный сигнал $x \in \mathbb{R}^N$, который не может быть точно восстановлен за K итераций ортогонального жадного алгоритма. Более того, в работе [21] выдвигается гипотеза (см. также [22]), что при любом $\beta < 2$ с высокой вероятностью найдется “плохой” K -разреженный вектор, который не восстанавливается за K итераций ортогонального жадного алгоритма.

§ 2. Вспомогательные леммы

Мы будем использовать два результата о скорости сходимости ортогонального жадного алгоритма.

ТЕОРЕМА А (Р. А. ДеВор, В. Н. Темляков; см. [23]). Пусть $y = \Phi x$. Тогда для любого $l \geq 1$ выполняется неравенство

$$\|r^l\| \leq \|x\|_1 l^{-1/2}.$$

ТЕОРЕМА В (см. [18]). Для любого l , $1 \leq l \leq 1/(20\mu(\Phi))$ справедлива оценка

$$\|r^{2l}\| \leq 3\sigma_l(y, \Phi).$$

Для $l \geq 0$ положим

$$z^l := x - x^l.$$

Тогда по определению ортогонального жадного алгоритма

$$r^l = y - \Phi x^l = \Phi x - \Phi x^l = \Phi z^l, \quad l \geq 0. \quad (2.1)$$

Пусть

$$x = (x_1, \dots, x_N), \quad z^l = (z_1^l, \dots, z_N^l), \quad l \geq 0.$$

Положим

$$V_0 = \text{supp } x, \quad \#V_0 \leq K. \quad (2.2)$$

По $x|_V$, $V \subset V_0$ определим элемент $(\tilde{x}_1, \dots, \tilde{x}_N)$ пространства $\mathbb{R}^N$ такой, что $\tilde{x}_i = x_i$, $i \in V$, и $\tilde{x}_i = 0$, $i \notin V$. Для каждого $V \subset V_0$ обозначим

$$R(V) = \sum_{i \in V} x_i^2.$$

Пусть

$$C := 2 \times 10^5, \quad c := 10^{-6}, \quad \delta := cK^{-0.2}. \quad (2.3)$$

ЛЕММА 2.1. Предположим, что $l + K \leq CK^{1.2}$, тогда справедливы неравенства

$$\sum_{i \in \Lambda^l} (z_i^l)^2 \leq 3\delta R(V_0 \setminus \Lambda^l), \quad (2.4)$$

$$R(V_0 \setminus \Lambda^l) \leq (1 + 2\delta)\|r^l\|^2. \quad (2.5)$$

ДОКАЗАТЕЛЬСТВО. Ясно, что $|z^l|_0 \leq |x|_0 + |x^l|_0 \leq K + l \leq CK^{1.2}$, тогда согласно СОИ и (2.1) имеем

$$(1 - \delta) \sum_{i=1}^N (z_i^l)^2 \leq \|\Phi z^l\|^2 = \|r^l\|^2 \leq (1 + \delta) \sum_{i=1}^N (z_i^l)^2. \quad (2.6)$$

С другой стороны, используя определение $R(\cdot)$ и СОИ для $x|_{V_0 \setminus \Lambda^l}$, записываем

$$\begin{aligned} \|x|_{V_0 \setminus \Lambda^l}\|^2 &= R(V_0 \setminus \Lambda^l), \\ (1 - \delta)R(V_0 \setminus \Lambda^l) &\leq \|\Phi(x|_{V_0 \setminus \Lambda^l})\|^2 \leq (1 + \delta)R(V_0 \setminus \Lambda^l). \end{aligned} \quad (2.7)$$

Из определения ортогонального жадного алгоритма следует, что

$$\|\Phi z^l\|^2 = \|r^l\|^2 \leq \|\Phi(x|_{V_0 \setminus \Lambda^l})\|^2.$$

Следовательно, применяя (2.6) и (2.7), получаем, что

$$\begin{aligned}
 (1 - \delta) \sum_{i=1}^N (z_i^l)^2 &\leq \|r^l\|^2 \leq \|\Phi(x|_{V_0 \setminus \Lambda^l})\|^2 \leq (1 + \delta) R(V_0 \setminus \Lambda^l), \\
 (1 - \delta) \left(\sum_{i \in \Lambda^l} (z_i^l)^2 + \sum_{i \in V_0 \setminus \Lambda^l} (z_i^l)^2 \right) &\leq (1 + \delta) R(V_0 \setminus \Lambda^l). \\
 \sum_{i \in \Lambda^l} (z_i^l)^2 + R(V_0 \setminus \Lambda^l) &\leq \frac{1 + \delta}{1 - \delta} R(V_0 \setminus \Lambda^l). \\
 \sum_{i \in \Lambda^l} (z_i^l)^2 &\leq \left(\frac{1 + \delta}{1 - \delta} - 1 \right) R(V_0 \setminus \Lambda^l) \leq 3\delta R(V_0 \setminus \Lambda^l),
 \end{aligned}$$

что завершает доказательство (2.4). Из неравенства (2.6) вытекает, что

$$R(V_0 \setminus \Lambda^l) = \sum_{i \in V_0 \setminus \Lambda^l} (z_i^l)^2 \leq \sum_{i=1}^N (z_i^l)^2 \leq (1 - \delta)^{-1} \|r^l\|^2 \leq (1 + 2\delta) \|r^l\|^2.$$

Для возрастающей последовательности $0 = l_0 < l_1 < \dots < l_s$, $s \geq 1$, обозначим

$$V_k := V_0 \setminus \Lambda^{l_k}, \quad R_k = R(V_k), \quad 0 \leq k \leq s. \quad (2.8)$$

ЛЕММА 2.2. *Предположим, что $l_k + K \leq CK^{1.2}$, $1 \leq k \leq s$, тогда для произвольного $p \in \mathbb{N}$ имеет место оценка*

$$\|r^{l_k+p}\|^2 \leq \frac{R_k}{p} (6\delta CK^{1.2} + 2K).$$

ДОКАЗАТЕЛЬСТВО. В силу того, что $r^{l_k} = \Phi z^{l_k}$, оцениваем с помощью теоремы A

$$\|r^{l_k+p}\|^2 \leq \frac{\|z^{l_k}\|_1^2}{p}. \quad (2.9)$$

Таким образом, для доказательства леммы достаточно оценить

$$\|z^{l_k}\|_1^2 = \left(\sum_{i=1}^N |z_i^{l_k}| \right)^2 = \left(\sum_{i \in V_0 \cup \Lambda^{l_k}} |z_i^{l_k}| \right)^2 \leq 2 \left(\left(\sum_{i \in V_0 \setminus \Lambda^{l_k}} |z_i^{l_k}| \right)^2 + \left(\sum_{i \in \Lambda^{l_k}} |z_i^{l_k}| \right)^2 \right).$$

Применяя (2.8) и (2.2), имеем

$$\begin{aligned}
 \left(\sum_{i \in V_0 \setminus \Lambda^{l_k}} |z_i^{l_k}| \right)^2 &= \left(\sum_{i \in V_0 \setminus \Lambda^{l_k}} |x_i| \right)^2 \\
 &= \left(\sum_{i \in V_k} |x_i| \right)^2 \leq \#V_k \sum_{i \in V_k} |x_i|^2 \leq \#V_0 R_k \leq R_k K.
 \end{aligned} \quad (2.10)$$

Используя (2.4) из леммы 2.1 получаем

$$\begin{aligned}
 \left(\sum_{i \in \Lambda^{l_k}} |z_i^{l_k}| \right)^2 &\leq \# \Lambda^{l_k} \sum_{i \in \Lambda^{l_k}} (z_i^{l_k})^2 = l_k \sum_{i \in \Lambda^{l_k}} (z_i^{l_k})^2 \\
 &\leq CK^{1.2} 3\delta R(V_0 \setminus \Lambda^{l_k}) \leq CK^{1.2} 3\delta R_k.
 \end{aligned}$$

Комбинируя с (2.10) достигаем желаемое неравенство

$$\|z^{l_k}\|_1^2 \leq R_k(6\delta CK^{1.2} + 2K),$$

что вместе с неравенством (2.9) завершает доказательство леммы.

ЛЕММА 2.3. Пусть $1 \leq p \leq K^{0.8}$ и $l_k + 2p \leq CK^{1.2}$, $1 \leq k \leq s$. Тогда для произвольного $W \subset V_k$, для которого $\#W = p$, будет выполняться неравенство

$$R(V_k \setminus \Lambda^{l_k+2p}) \leq 10R(V_k \setminus W) + 30\delta R_k.$$

ДОКАЗАТЕЛЬСТВО. Согласно СОИ, (2.1), (2.4) и (2.8) оцениваем

$$\begin{aligned} \left(\sigma_p(r^{l_k})\right)^2 &\leq \|r^{l_k} - \Phi(x|_W)\|^2 = \|\Phi(z^{l_k}) - \Phi(z^{l_k}|_W)\|^2 \\ &= \|\Phi(z^{l_k} - z^{l_k}|_W)\|^2 \leq (1 + \delta) \sum_{1 \leq i \leq N, i \notin W} (z_i^{l_k})^2 \\ &\leq (1 + \delta) \left(\sum_{i \in V_k \setminus W} (z_i^{l_k})^2 + \sum_{1 \leq i \leq N, i \notin V_k} (z_i^{l_k})^2 \right) \\ &\leq (1 + \delta) \left(\sum_{i \in V_k \setminus W} (x_i)^2 + \sum_{i \in \Lambda^{l_k}} (z_i^{l_k})^2 \right) \\ &\leq (1 + \delta) (R(V_k \setminus W) + 3\delta R(V_0 \setminus \Lambda^{l_k})) \leq (1 + \delta) (R(V_k \setminus W) + 3\delta R_k). \end{aligned}$$

В силу того, что

$$p \leq K^{0.8} \leq \frac{1}{20\mu(\Phi)},$$

мы можем применить теорему В и достичь оценку

$$\|r^{l_k+2p}\| \leq 3\sigma_p(r^{l_k}).$$

Используя (2.5) из леммы 2.1, получаем

$$\begin{aligned} R(V_k \setminus \Lambda^{l_k+2p}) &= R(V_0 \setminus \Lambda^{l_k+2p}) \leq (1 + 2\delta) \|r^{l_k+2p}\|^2 \leq (1 + 2\delta) 9(\sigma_p(r^{l_k}))^2 \\ &\leq (1 + 2\delta) 9(1 + \delta) (R(V_k \setminus W) + 3\delta R_k) \leq 10(R(V_k \setminus W) + 3\delta R_k). \end{aligned}$$

§ 3. Доказательство теоремы 4

Докажем индукцией по k , что если $l_k + K \leq CK^{1.2}$ и $V_k \neq \emptyset$, то можно определить $l_{k+1} > l_k$, V_{k+1} и R_{k+1} , удовлетворяющее (2.8), для которого

$$l_{k+1} + K \leq CK^{1.2}, \quad (3.1)$$

и верно по меньшей мере одно из следующих утверждений:

$$(A) \quad l_{k+1} - l_k \leq 3 \times 10^4 (6\delta CK^{1.2} + 2K), \quad \#(V_k \setminus V_{k+1}) \geq K^{0.8}; \quad (3.2)$$

$$(B) \quad l_{k+1} - l_k \leq 2K^{0.8}, \quad \#(V_k \setminus V_{k+1}) \geq K^{0.6}; \quad (3.3)$$

$$(C) \quad l_{k+1} - l_k \leq 2K^{0.6}, \quad \#(V_k \setminus V_{k+1}) \geq K^{0.4}; \quad (3.4)$$

$$(D) \quad l_{k+1} - l_k \leq 2K^{0.4}, \quad \#(V_k \setminus V_{k+1}) \geq K^{0.2}; \quad (3.5)$$

$$(E) \quad l_{k+1} - l_k \leq 2K^{0.2}, \quad \#(V_k \setminus V_{k+1}) \geq 1. \quad (3.6)$$

Положим

$$p_A = [3 \times 10^4(6\delta CK^{1.2} + 2K)], \quad W_A = V_k \cap \Lambda^{l_k + p_A}.$$

Если $\#W_A \geq K^{0.8}$, определим

$$l_{k+1} := l_k + p_A, \quad V_{k+1} := V_0 \setminus \Lambda^{l_{k+1}} = V_k \setminus W_A, \quad R_{k+1} := R(V_{k+1}).$$

Тогда $\#(V_k \setminus V_{k+1}) = \#W_A$ и утверждение (3.2) истинно. Неравенство

$$l_k + p_A + K \leq CK^{1.2}$$

будет проверено ниже. Предположим, что

$$\#W_A < K^{0.8}. \quad (3.7)$$

Применяя лемму 2.2, имеем

$$\|r^{l_k + p_A}\|^2 \leq \frac{R_k(6\delta CK^{1.2} + 2K)}{[3 \times 10^4(6\delta CK^{1.2} + 2K)]} \leq \frac{R_k}{2.5 \times 10^4}.$$

Используя неравенство (2.5) из леммы 2.1, оцениваем

$$R(V_k \setminus W_A) = R(V_0 \setminus \Lambda^{l_k + p_A}) \leq (1 + 2\delta)\|r^{l_k + p_A}\|^2 \leq (1 + 2\delta)\frac{R_k}{2.5 \times 10^4} \leq \frac{R_k}{2 \times 10^4}. \quad (3.8)$$

Положим

$$p_B := \#W_A, \quad W_B = V_k \cap \Lambda^{l_k + 2p_B}.$$

Если $\#W_B \geq K^{0.6}$, определим

$$l_{k+1} = l_k + 2p_B, \quad V_{k+1} = V_k \setminus \Lambda^{l_{k+1}} = V_k \setminus W_B, \quad R_{k+1} = R(V_{k+1}).$$

Тогда $\#(V_k \setminus V_{k+1}) = \#W_B \geq K^{0.6}$ и, принимая во внимание (3.7), получаем (3.3). Неравенство

$$l_k + 2p_B + K \leq CK^{1.2}$$

будет проверено ниже.

Предположим, что

$$\#W_B < K^{0.6}. \quad (3.9)$$

Применяя лемму 2.3 для $W = W_A$ и $p = p_B$, а также неравенство (3.8), оцениваем

$$R(V_k \setminus W_B) = R(V_k \setminus \Lambda^{l_k + 2p_B}) \leq 10R(V_k \setminus W_A) + 30\delta R_k \leq \frac{R_k}{2 \times 10^3} + 30\delta R_k. \quad (3.10)$$

Эти рассуждения будут повторены трижды.

Положим

$$p_C := \#W_B, \quad W_C = V_k \cap \Lambda^{l_k + 2p_C}.$$

Если $\#W_C \geq K^{0.4}$, определим

$$l_{k+1} = l_k + 2p_C, \quad V_{k+1} = V_k \setminus \Lambda^{l_{k+1}} = V_k \setminus W_C, \quad R_{k+1} = R(V_{k+1}).$$

Тогда $\#(V_k \setminus V_{k+1}) = \#W_C \geq K^{0.4}$, и с учетом (3.9) получаем (3.4). Неравенство

$$l_k + 2p_C + K \leq CK^{1.2}$$

будет проверено ниже.

Предположим, что

$$\#W_C < K^{0.4}. \quad (3.11)$$

Применяя лемму 2.3 для $W = W_B$ и $p = p_C$, а также неравенство (3.10), оцениваем

$$R(V_k \setminus W_C) = R(V_k \setminus \Lambda^{l_k+2p_C}) \leq 10R(V_k \setminus W_B) + 30\delta R_k \leq \frac{R_k}{2 \times 10^2} + 3(10^2 + 10)\delta R_k. \quad (3.12)$$

Положим

$$p_D := \#W_C, \quad W_D = V_k \cap \Lambda^{l_k+2p_D}.$$

Если $\#W_D \geq K^{0.2}$, определим

$$l_{k+1} = l_k + 2p_D, \quad V_{k+1} = V_k \setminus \Lambda^{l_{k+1}} = V_k \setminus W_D, \quad R_{k+1} = R(V_{k+1}).$$

Тогда $\#(V_k \setminus V_{k+1}) = \#W_D \geq K^{0.2}$ и, принимая во внимание (3.11), мы получаем (3.5). Неравенство

$$l_k + 2p_D + K \leq CK^{1.2}$$

будет проверено ниже.

Предположим, что

$$\#W_D < K^{0.2}. \quad (3.13)$$

Применяя лемму 2.3 для $W = W_C$ и $p = p_D$, а также неравенство (3.12), оцениваем

$$R(V_k \setminus W_D) = R(V_k \setminus \Lambda^{l_k+2p_D}) \leq 10R(V_k \setminus W_C) + 30\delta R_k \leq \frac{R_k}{20} + 3(10^3 + 10^2 + 10)\delta R_k. \quad (3.14)$$

Положим

$$p_E := \#W_D, \quad W_E = V_k \cap \Lambda^{l_k+2p_E},$$

$$l_{k+1} = l_k + 2p_E, \quad V_{k+1} = V_k \setminus \Lambda^{l_{k+1}} = V_k \setminus W_E, \quad R_{k+1} = R(V_{k+1}).$$

С учетом неравенства (3.13) получаем

$$l_{k+1} - l_k \leq 2K^{0.2}.$$

Неравенство

$$l_k + 2p_E + K \leq CK^{1.2}$$

будет проверено ниже.

Применяя лемму 2.3 для $W = W_D$ и $p = p_E$, а также неравенства (3.14) и (2.3), имеем

$$\begin{aligned} R_{k+1} &= R(V_k \setminus W_E) = R(V_k \setminus \Lambda^{l_k+2p_E}) \leq 10R(V_k \setminus W_D) + 30\delta R_k \\ &\leq \frac{R_k}{2} + 3(10^4 + 10^3 + 10^2 + 10)\delta R_k < R_k(1/2 + 4 \times 10^4\delta) < R_k. \end{aligned}$$

Следовательно, $W_E \neq \emptyset$ и $\#(V_k \setminus V_{k+1}) \geq 1$, и утверждение (3.6) выполняется.

Таким образом, для окончания доказательства индукционного предположения остается оценить

$$l_k + p_A + K, \quad l_k + 2p_B + K, \quad l_k + 2p_B + K, \quad l_k + 2p_C + K, \quad l_k + 2p_D + K.$$

Ясно, что первое из этих чисел является наибольшим.

Используя индукционное предположение, включение $V_0 \supset V_1 \supset \dots \supset V_k$, и равенство $\#V_0 = K$, мы получаем, что утверждение (3.2) (случай А) может выполняться не более чем $K^{0.2} = K/K^{0.8}$ раз, утверждение (3.3) (случай В) – не более чем $K^{0.4} = K/K^{0.6}$ раз, утверждение (3.4) (случай С) – не более чем $K^{0.6} = K/K^{0.4}$ раз, утверждение (3.5) (случай D) – не более чем $K^{0.8} = K/K^{0.2}$ раз, и утверждение (3.6) (случай Е) – не более чем K раз. Следовательно, используя определение (2.3), мы получаем оценку

$$\begin{aligned} l_k &\leq 3 \times 10^4 (6\delta CK^{1.2} + 2K)K^{0.2} + 2K^{0.8}K^{0.4} \\ &\quad + 2K^{0.6}K^{0.6} + 2K^{0.4}K^{0.8} + 2K^{0.2}K \\ &= (18 \times 10^4 \delta CK^{0.2} + 6 \times 10^4 + 8)K^{1.2}, \\ l_k + p_A + K &\leq (18 \times 10^4 \delta CK^{0.2} + 6 \times 10^4 + 8)K^{1.2} + 3 \times 10^4 (6\delta CK^{1.2} + 2K) + K \\ &\leq (36 \times 10^4 \delta CK^{0.2} + 12 \times 10^4 + 9)K^{1.2} \\ &\leq (36 \times 10^4 \times 10^{-6} K^{-0.2} 2 \times 10^5 K^{0.2} + 12 \times 10^4 + 9)K^{1.2} \\ &< 20 \times 10^4 K^{1.2} = CK^{1.2}. \end{aligned}$$

В силу того, что $\#(V_k \setminus V_{k+1}) \geq 1$, существует $s \in \mathbb{N}$, для которого $V_s = \emptyset$ и $l_s + K \leq CK^{1.2}$. Поэтому согласно (2.8) и (2.1) имеем

$$V_0 \subset \Lambda^{l_s} \quad \text{и} \quad r^{l_s} = \Phi z^{l_s} = \Phi(x - x^{l_s}) = 0.$$

Используя СОИ, получаем, что

$$x = x^{l_s}.$$

ЗАМЕЧАНИЕ. Мы предполагаем, что константа 3 в теореме В не является оптимальной и, следовательно, константы C и c^{-1} из (2.3) могут быть уменьшены.

Автор благодарит Б. С. Кашина, С. В. Конягина и В. Н. Темлякова за полезные обсуждения.

Список литературы

- [1] R. G. Baraniuk, “Compressive sensing”, *IEEE Signal Processing Magazine*, **24**:4 (2007), 118–121.
- [2] E. J. Candès, “Compressive sampling”, *Proceedings of the international congress of mathematicians*, v. III: *Invited lectures* (Madrid, Spain, 2006), Eur. Math. Soc., Zürich, 2006, 1433–1452.
- [3] D. L. Donoho, “Compressed sensing”, *IEEE Trans. Inform. Theory*, **52**:4 (2006), 1289–1306.
- [4] Б. С. Кашин, В. Н. Темляков, “Замечание о задаче сжатого измерения”, *Матем. заметки*, **82**:6 (2007), 829–837; англ. пер.: B. S. Kashin, V. N. Temlyakov, “A remark on compressed sensing”, *Math. Notes*, **82**:5–6 (2007), 748–755.
- [5] S. S. Chen, D. L. Donoho, M. A. Saunders, “Atomic decomposition by basis pursuit”, *SIAM Rev.*, **43**:1 (2001), 129–159.
- [6] V. N. Temlyakov, “Greedy approximation”, *Acta Numer.*, **17** (2008), 235–409.
- [7] S. G. Mallat, Z. Zhang, “Matching pursuit with time-frequency dictionaries”, *IEEE Trans. Signal Process.*, **41**:12 (1993), 3397–3415.

- [8] A. C. Gilbert, M. Muthukrishnan, M. J. Strauss, "Approximation of functions over redundant dictionaries using coherence", *Proceedings of the fourteenth annual ACM-SIAM symposium on discrete algorithms* (Baltimore, MD, USA, 2003), Association for Computing Machinery, New York, 2003, 243–252.
- [9] D. L. Donoho, M. Elad, V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise", *IEEE Trans. Inform. Theory*, **52**:1 (2006), 6–18.
- [10] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation", *IEEE Trans. Inform. Theory*, **50**:10 (2004), 2231–2242.
- [11] E. J. Candes, T. Tao, "Decoding by linear programming", *IEEE Trans. Inform. Theory*, **51**:12 (2005), 4203–4215.
- [12] E. J. Candes, T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?", *IEEE Trans. Inform. Theory*, **52**:12 (2005), 5406–5425.
- [13] Б. С. Кашин, "Поперечники некоторых конечномерных множеств и классов гладких функций", *Изв. АН СССР. Сер. матем.*, **41**:2 (1977), 334–351; англ. пер.: B. S. Kašin, "Diameters of some finite-dimensional sets and classes of smooth functions", *Math. USSR-Izv.*, **11**:2 (1977), 317–333.
- [14] R. Baraniuk, M. Davenport, R. DeVore, M. Wakin, "A simple proof of the restricted isometry property for random matrices", *Constr. Approx.*, **28**:3 (2008), 253–263.
- [15] R. Gribonval, M. Nielsen, "On the strong uniqueness of highly sparse expansions from redundant dictionaries", *Proc. Int Conf. Independent Component Anal. (ICA 04)*, 2004.
- [16] V. N. Temlyakov, P. Zheltov, "On performance of greedy algorithms", *J. Approx. Theory*, **163**:9 (2011), 1134–1145.
- [17] D. L. Donoho, M. Elad, V. N. Temlyakov, "On Lebesgue-type inequalities for greedy approximation", *J. Approx. Theory*, **147**:2 (2007), 185–195.
- [18] E. D. Livshits, *On the optimality of Orthogonal Greedy Algorithm for M-coherent dictionaries*, arXiv: abs/1003.5349.
- [19] M. A. Davenport, M. B. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property", *IEEE Trans. Inform. Theory*, **56**:9 (2010), 4395–4401.
- [20] E. Liu, V. M. Temlyakov, *Orthogonal super greedy algorithm and applications in compressed sensing*, 2010, <http://dsp.rice.edu/sites/dsp.rice.edu/files/cs/LiuTemlyakov.pdf>.
- [21] H. Rauhut, "On the impossibility of uniform sparse reconstruction using greedy methods", *Sampl. Theory Signal Image Process.*, **7**:2 (2008), 197–215.
- [22] W. Dai, O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction", *IEEE Trans. Inform. Theory*, **55**:5 (2009), 2230–2249.
- [23] R. A. DeVore, V. N. Temlyakov, "Some remarks on greedy algorithms", *Adv. Comput. Math.*, **5**:1 (1996), 173–187.

Е. Д. Лившиц (E. D. Livshits)
 Эверноут Корпорейшн, г. Москва
 E-mail: livshitz@rambler.ru

Поступила в редакцию
 03.12.2010 и 11.02.2011