Сжатое измерение (Compressed Sensing) и поперечник шара в ℓ_∞^N

11 мая 2019 г.

Пусть $x \in \mathbb{R}^N$, где размерность N очень велика. Предположим, нам не доступен вектор x, мы можем лишь *измерить* n скалярных произведений:

$$y = (\langle x, \varphi_1 \rangle, \langle x, \varphi_2 \rangle, \ldots, \langle x, \varphi_n \rangle) = \Phi x \in \mathbb{R}^n,$$

где $n < N$ (интересен случай, когда n намного меньше N), и хотим восстановить x по y. В общем случае это невозможно. Наложим условие, что x является s-разреженным:

$$\|x\|_0 := \#\{j \colon x_j \neq 0\} \leqslant s$$

или хотя бы хорошо приближается такими векторами:

$$\sigma_s(x)_1 := \inf_{x' \colon \|x'\|_0 \leqslant s} \|x - x'\|_1.$$

(Например, картинки хорошо приближаются разреженными векторами, в базисе всплесков.)

Рассмотрим задачу

$$\|z\|_0 \to \min, \quad \Phi z = y.$$

Если любые $2s$ столбцов матрицы Φ линейно независимы, то $z = x$ для s-разреженных x, поскольку $\|x - z\|_0 \leqslant \|x\|_0 + \|z\|_0 \leqslant 2s$, $\Phi(x - z) = 0$. Этого легко добиться для $n = 2s$ и любого $N \geqslant n$. Проблема, однако, в том, что для решения задачи придётся перебирать все возможные s-элементные подмножества $\{1, \ldots, N\}$, а это трудно.

Заменим $\|\cdot\|_0$ на $\|\cdot\|_1$:

$$\|z\|_1 \to \min, \quad \Phi z = y. \tag{1}$$

Задача (1) выпуклая и есть эффективные алгоритмы для её решения. (Почему ℓ_1, а не ℓ_2, скажем? Минимизация ℓ_1-нормы даёт разреженные решения.) Обозначим через $z_\Phi(y)$ решение задачи (1). При каких условиях на Φ имеем $z_\Phi(\Phi x) = x$ для разреженных векторов?

Всюду далее

$$\operatorname{Ker}\Phi = \{x\colon \Phi x = 0\}.$$

Теорема 1. *Пусть $\Phi \in \operatorname{Mat}(n \times N)$, $s \in \mathbb{N}$, и выполнено условие:*

$$\forall u \in \operatorname{Ker}\Phi \quad \frac{\|u\|_2}{\|u\|_1} \leqslant \frac{1}{4} s^{-1/2}.$$

Тогда для любого $x \in \mathbb{R}^N$ имеем

$$\|x - z_\Phi(\Phi x)\|_2 \leqslant s^{-1/2}\sigma_s(x)_1,$$

$$\|x - z_\Phi(\Phi x)\|_1 \leqslant 4\sigma_s(x)_1.$$

В частности, для s-разреженных x решение задачи (1) единственно и совпадает с x.

Замечание: в обратную сторону тоже верно: если $u \in \operatorname{Ker}\Phi$, то $z_\Phi(\Phi u) = z_\Phi(0) = 0$, $\|u\|_2 \leqslant s^{-1/2}\sigma_s(u)_1 \leqslant s^{-1/2}\|u\|_1$.

Вспомним, что величина $\sup_{u\in L^n} \|u\|_2/\|u\|_1$ возникает в поперечнике по Гельфанду:

$$d^n(B_1^N, \ell_2^N) = \inf_{\operatorname{codim} L^n \leqslant n} \sup_{u \in B_1^N \cap L^n} \|u\|_2 = \inf_{\operatorname{codim} L^n \leqslant n} \sup_{u \in L^n} \frac{\|u\|_2}{\|u\|_1}.$$

Значит, хорошие измерительные матрицы Φ — те, для которых $L^n = \operatorname{Ker}\Phi$ даёт оценку поперечника.

Теорема 2.

$$d^n(B_1^N, \ell_2^N) = d_n(B_2^N, \ell_\infty^N) \leqslant C\sqrt{\frac{\log\frac{2N}{n}}{n}}.$$

Замечание 1: верхняя оценка точна, однако мы не будем это доказывать.

Замечание 2: оценка поперечника шара играет ключевую роль в нахождении порядков поперечников классов Соболева $d_n(W_p^r, L_q)$ в той области, где тригонометрические полиномы не оптимальны.

Замечание 3: все известные способы построения оптимальных подпространств — случайные.

Замечание 4: при $n \asymp N$ получаем, что ℓ_1 и ℓ_2-нормы пропорциональны на L^n:

$$N^{-1/2} \leqslant \frac{\|u\|_2}{\|u\|_1} \leqslant cn^{-1/2} \asymp N^{-1/2}, \quad u \in L^n.$$

Следствие 1. *Для любых $n < N$ существует матрица $\Phi \in \mathrm{Mat}(n \times N)$, позволяющая эффективно восстанавливать s-разреженные вектора $x \in \mathbb{R}^N$ по n измерениям, где $s \asymp \frac{n}{\log \frac{2N}{n}}$.*

Скажем, что матрица Φ обладает свойством ограниченной изометрии (RIP — Restricted Isometry Property) с параметрами $\delta \in (0,1)$ и $s \in \mathbb{N}$, если

$$(1-\delta)\|x\|_2 \leqslant \|\Phi x\|_2 \leqslant (1+\delta)\|x\|_2, \quad \forall x\colon \|x\|_0 \leqslant s.$$

Теорема 3. *Если $\Phi \sim \mathrm{RIP}\,(\delta, s)$, то*

$$\forall u \in \operatorname{Ker}\Phi \quad \frac{\|u\|_2}{\|u\|_1} \leqslant \frac{2}{1-\delta} s^{-1/2}.$$

Теорема 4. *Пусть элементы матрицы Φ — независимые нормальные случайные величины $\Phi_{i,j} \sim N(0, \frac{1}{n})$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant N$, $\delta \in (0,1)$. Тогда для $s \leqslant c_1(\delta) n / \log(2N/n)$ с вероятностью $> 1 - \exp(-c_2(\delta) n)$ матрица Φ обладает свойством $\mathrm{RIP}(\delta, s)$.*

Следствие 2. *В качестве оптимальной измерительной матрицы Φ с большой вероятностью можно взять реализацию матрицы из независимых нормальных величин $\Phi_{i,j} \sim \mathcal{N}(0, \frac{1}{n})$.*

Замечание: вместо нормальных величин можно брать $\Phi_{i,j} = \pm n^{-1/2}$ со случайными независимыми знаками.

Замечание: нормировка $\mathsf{E}\Phi_{i,j}^2 = \frac{1}{n}$ обеспечивает равенство

$$\mathsf{E}\|\Phi x\|_2^2 = \sum_{i=1}^{n} \mathsf{E} \sum_{j=1}^{N} \xi_{i,j}^2 x_j^2 = \|x\|_2^2.$$

Теорема 2 была доказана с лишним логарифмом в работе Кашина 1977 года, уточнена Глускиным и Гарнаевым. Теорию сжатого измерения разработали Donoho, Tao, Candes в 2000-х.

Перейдём к доказательствам.

Доказательство теоремы 1. Обозначим $u = x - z_\Phi(\Phi x)$ и оценим $\|u\|$. Мы имеем $u \in \operatorname{Ker}\Phi$ (по определению z_Φ) и $\|x-u\|_1 = \|z\|_1 \leqslant \|x\|_1$. Оценка на $\|u\|_2$ будет следовать из оценки $\|u\|_1$ и свойства пространства $\operatorname{Ker}\Phi$.

Оценим $\|u\|_1$. Возьмём T множество s наибольших координат x. Имеем $\|u\|_1 \leqslant \|u_T\|_1 + \|u_{T^c}\|_1$,

$$\|u_T\|_1 \leqslant s^{1/2}\|u_T\|_2 \leqslant s^{1/2}\|u\|_2 \leqslant s^{1/2}\frac{1}{4}s^{-1/2}\|u\|_1 = \frac{1}{4}\|u\|_1.$$

Через x_T обозначим вектор x, в котором координаты $j \notin T$ заменены на 0. Через T^c обозначаем дополнение T.

Предположим сначала, что x является s-разреженным, $x = x_T$. Тогда $u = 0$, иначе

$$\|x-u\|_1 = \|x_T - u_T\|_1 + \|u_{T^c}\|_1 \geqslant \|x\|_1 - \|u_T\|_1 + \|u_{T^c}\|_1 > \|x\|_1,$$

что невозможно (т.е.: u не сконцентрирован, поэтому вычитание u увеличивает норму x). В общем случае $\|x_{T^c}\|_1 = \sigma_s(x)_1$,

$$\|u_{T^c}\|_1 \leqslant \|x_{T^c}\|_1 + \|(x-u)_{T^c}\|_1,$$

$$\|(x-u)_{T^c}\|_1 = \|x-u\|_1 - \|(x-u)_T\|_1 \leqslant \|x\|_1 + \|u_T\|_1 - \|x_T\|_1 = \|x_{T^c}\|_1 + \|u_T\|_1.$$

Следовательно. $\|u_{T^c}\|_1 \leqslant 2\sigma_s(x)_1 + \|u_T\|_1$. Отсюда: $\|u\|_1 \leqslant 2\|u_T\|_1 + 2\sigma_s(x)_1 \leqslant \frac{1}{2}\|u\|_1 + 2\sigma_s(x)_1$, ч.т.д. $\square$

Заметим, что теоремы 3 и 4 влекут теорему 2. Действительно, фиксируем $\delta = 1/2$, построим RIP-матрицу (δ, s), $s \asymp n/\log(2N/n)$, её нуль-пространство $\operatorname{Ker}\Phi$ даст нужную оценку для поперечника.

Доказательство теоремы 3. Возьмём $u \in \operatorname{Ker}\Phi$, $\|u\|_1 = 1$. Разобьём координаты в $\mathbb{R}^N$ на группы T_j по s штук, в T_0 включим самые большие по модулю координаты, в T_1 — следующие, и т.д.

Очевидно, $\|u\|_2 \leqslant \|u_{T_0}\|_2 + \|u_{T_0^c}\|_2$. Оценим второе слагаемое: пусть $a_0 = \min_{j\in T_0}|u_j|$, тогда $\|u\|_1 \geqslant sa_0$, $a_0 \leqslant 1/s$,

$$\|u_{T_0^c}\|_2^2 = \sum_{j\notin T_0}|u_j|^2 \leqslant a_0\sum_j |u_j| \leqslant 1/s.$$

Первое слагаемое оценивается, используя RIP:

$$\|u_{T_0}\|_2 \leqslant (1-\delta)^{-1}\|\Phi u_{T_0}\|_2.$$

В силу $\Phi u = 0$, имеем $\Phi u_{T_0} = -\sum_{j\geqslant 1}\Phi u_{T_j}$, норма этого вектора не превосходит $(1+\delta)\sum_{j\geqslant 1}\|u_{T_j}\|_2$. В силу монотонности координат,

$$\|u_{T_{j+1}}\|_2 \leqslant s^{-1/2}\|u_{T_j}\|_1, \quad j\geqslant 0.$$

Действительно, если $a_j := \min_{i\in T_j}|u_i|$, то норма слева не больше $s^{1/2}a_j$, а норма справа не меньше sa_j. Отсюда

$$\sum_{j\geqslant 1}\|u_{T_j}\|_2 \leqslant s^{-1/2}\sum_{j\geqslant 0}\|u_{T_j}\|_1 \leqslant s^{-1/2}.$$

Окончательно получаем

$$\|u\|_2 \leqslant (1+\frac{1+\delta}{1-\delta})s^{-1/2}.$$

□

Для доказательства теоремы 4 нам потребуется утверждение из теории концентрации меры.

Утверждение. *Пусть функция $f\colon \mathbb{R}^k \to \mathbb{R}$ липшицева с константой 1:*

$$|f(x) - f(y)| \leqslant \|x-y\|_2, \quad \forall x, y \in \mathbb{R}^k.$$

Тогда для нормального вектора $\xi = (\xi_1, \dots, \xi_k)$, где $\xi_j \sim \mathcal{N}(0,1)$ и независимы, и $t > 0$, имеем

$$\mathsf{P}(|f(\xi) - Mf(\xi)| > t) \leqslant 2\exp(-t^2/2),$$

где $Mf(\xi)$ — медиана величины $f(\xi)$.

Из данного утверждения следует, что $f(\xi)$ сильно сконцентрирована вокруг своей медианы M; следовательно, $|\mathsf{E}f(\xi) - M| \leqslant C$, а для $f \geqslant 0$ имеем $|(\mathsf{E}f(\xi)^p)^{1/p} - M| \leqslant C_p$, $p \in [1, +\infty)$. (Упражнение!) Отсюда, в частности, получаем следующее неравенство для неотрицательных f:

$$\mathsf{P}(|f(\xi) - (\mathsf{E}f(\xi)^2)^{1/2}| > t) \leqslant c_1 \exp(-c_2 t^2). \tag{2}$$

Доказательство теоремы 4. Нужно обеспечить неравенства $1-\delta \leqslant \|\Phi x\|_2 \leqslant 1+\delta$ для всех x из множества

$$\Sigma_s^N := \{x \in \mathbb{R}^N \colon \|x\|_2 = 1,\ \|x\|_0 \leqslant s\}.$$

Зафиксируем вектор x^* единичной длины. Отождествляя пространство матриц с $\mathbb{R}^{nN}$, мы видим, что функция $f(M) = \|Mx^*\|_2$ является 1-липшицевой по M:

$$|\|M_1 x^*\|_2 - \|M_2 x^*\|_2| \leqslant \|(M_1 - M_2)x^*\|_2 \leqslant \|M_1 - M_2\|_{2\to 2} \leqslant \|M_1 - M_2\|_F.$$

Элементы матрицы $\Xi = \sqrt{n}\Phi$ — случайные величины из $\mathcal{N}(0,1)$. Имеем

$$\mathsf{E}f(\Xi)^2 = \sum_{i=1}^n \mathsf{E} \sum_{j=1}^N \xi_{i,j}^2 (x_j^*)^2 = n.$$

Следовательно, имеет место концентрация

$$\mathsf{P}(|f(\Xi) - n^{1/2}| > t) \leqslant c\exp(-ct^2),$$

$$\mathsf{P}(|\|\Phi x^*\|_2 - 1| > t) \leqslant c\exp(-cnt^2).$$

Мы положим $t := \delta/4$.

Далее мы построим $(\delta/4)$-сеть $\mathcal{M}$ для множества Σ_s^N и обеспечим неравенство

$$(1-\delta/4)\|x\|_2 \leqslant \|\Phi x\|_2 \leqslant (1+\delta/4)\|x\|_2, \quad \forall x \in \mathcal{N}. \tag{3}$$

Для каждого набора из координат $T \subset \{1, \dots, N\}$, $|T| = s$, нужно построить $(\delta/4)$-сеть в сфере в $\mathbb{R}^T$. Как хорошо известно, для этого достаточно $(3/(\delta/4))^s \leqslant (c/\delta)^s$ точек. Поскольку всего наборов координат $\binom{N}{s} \leqslant (eN/s)^s$, получаем

$$|\mathcal{M}| \leqslant (eN/s)^s (c/\delta)^s \leqslant (\frac{cN}{s\delta})^s.$$

Вероятность, что неравенство в (3) нарушится хотя бы для одного $x \in \mathcal{M}$, не превосходит

$$|\mathcal{M}| \cdot c_1 \exp(-c_2 n\delta^2) \leqslant c_1 \exp(-c_2 n\delta^2 + s\log(\frac{c_3 N}{s\delta})).$$

Эта величина меньше $\exp(-c(\delta)n)$ при выбранном нами s, следовательно, с большой вероятностью выполнено (3).

Из (3) и того, что $\mathcal{M}$ — $(\delta/4)$-сеть для Σ_s^N, вытекает нужное нам неравенство для всего Σ_s^N (упражнение!). $\square$