LP DECODING WITH BOUNDED-WIDTH MATRICES

NOTES BY PETAR MAYMOUNKOV

1. The width property

Definition 1.1 — *A subspace $\Gamma \subset \mathbb{R}^m$ of dimension $m-n$, where $n \leqslant m$, is said to have the width property iff all $x \in \Gamma$ are such that*

$$(*) \qquad \|x\|_2 \leqslant \frac{\omega}{n^{1/2}} \|x\|_1,$$

where $\omega = O(\ln(em/n)^{1/2})$.

Definition 1.2 — *A linear transformation $A \in \mathbb{R}^{n\times m}$, where $n < m$, of rank n has the width property iff $\ker(A)$ has the width property.*

Matrices with the width property exist (See [1] for references), e.g. a random matrix has this property with high probability. Explicit constructions with slightly bigger ω are found in [2]. It is convenient to think of ω as being "small" compared to m.

Remark 1.1 — The subspace Γ provides a distortion-$\omega(m/n)^{1/2}$ embedding of ℓ_2^{m-n} into ℓ_1^m by the map

$$(1.1) \qquad x \mapsto m^{1/2} \cdot Ux,$$

where $U \in \mathbb{R}^{m\times(m-n)}$ is a column-wise orthonormal basis for Γ. To verify the distortion properties, use

$$\frac{1}{m^{1/2}}\|x\|_1 \leqslant \|x\|_2 \leqslant \frac{\omega}{n^{1/2}}\|x\|_1.$$

Let $S = n/\omega^2$. The following two lemmas show that (in this order):

(a) Vectors in Γ have large support, and
(b) Their mass is evenly distributed across their support

Lemma 1.1 — *Let $0 \neq x \in \Gamma$, then $\|x\|_0 \geqslant S$.*

Proof. — Set $k = \|x\|_0$

$$\|x\|_1 \leqslant k^{1/2}\|x\|_2 \leqslant k^{1/2}\frac{\omega}{n^{1/2}}\|x\|_1$$

The first inequality is Cauchy-Schwarz, the second follows from (*). ∎

Lemma 1.2 — *Let $0 \neq x \in \Gamma$, then for any index set $\Lambda \subseteq [m]$ with $|\Lambda| < S/4$ one has $\|x_\Lambda\|_1 < \|x\|_1/2$.*

Proof. — Set $k = |\Lambda|$

$$\begin{aligned} \|x_\Lambda\|_1 &\leqslant k^{1/2}\|x_\Lambda\|_2 && \text{(using Cauchy-Schwarz)} \\ &\leqslant k^{1/2}\|x\|_2 \\ &\leqslant k^{1/2}\omega\frac{\|x\|_1}{n^{1/2}} && \text{(using (*))} \\ &< \|x\|_1/2 && \text{(using } k < S/4\text{)} \end{aligned}$$

∎

2. LP decoding

Note that all results in this section are more or less obvious when thinking of $\delta \in \mathbf{ker}(A)$ as an evenly spread out vector.

LEMMA 2.1 — *If $u \in \mathbb{R}^m$ with $\|u\|_0 < S/4$, then for all $0 \neq x \in \Gamma$ we have $\|u + x\|_1 > \|u\|_1$.*

Proof. —

$$\begin{aligned} \|u + x\|_1 &= \|u_\Lambda + x_\Lambda\|_1 + \|x_{\overline{\Lambda}}\|_1 \\ &\geqslant \|u_\Lambda\|_1 - \|x_\Lambda\|_1 + \|x_{\overline{\Lambda}}\|_1 \\ &= \|u\|_1 + \|x\|_1 - 2\|x_\Lambda\|_1 \\ &> \|u\|_1 \end{aligned}$$

In the last derivation we use that $\|x\|_1 - 2\|x_\Lambda\|_1 > 0$, which follows from Lemma 1.2. ∎

Remark 2.1 — Recall that for a signal u the *Basis Pursuit* algorithm computes an approximation u_A of u

$$u_A := u + \underset{\delta \in \mathbf{ker}(A)}{\arg\min} \|u + \delta\|_1 \tag{2.1}$$

by solving the Linear Program (LP)

$$\text{minimize } \|u^*\|_1 \text{ subject to } Au^* = Au$$

Here $u^* = u + \delta$. Lemma 2.1 thus ensures that when $\|u\|_0 < S/4$ we can recover the signal *exactly*, i.e. $u = u_A$, when A is a matrix with the width property. The next theorem provides recovery guarantees for the case when the signal is not necessarily sparse.

THEOREM 2.1 — *For any u and u^* such that $\|u^*\|_1 \leqslant \|u\|_1$, $u^* - u \in \Gamma$ and $k \leqslant S/16$ we have*

$$\|u^* - u\|_1 \leqslant 4 \cdot \mathbf{Err}_1^k(u), \text{ and} \tag{2.2}$$

$$\|u^* - u\|_2 \leqslant k^{-1/2} \cdot \mathbf{Err}_1^k(u) \tag{2.3}$$

Remark 2.2 — Recall that

$$\mathbf{Err}_p^k(u) = \min_{\substack{w \\ \|w\|_0 \leqslant k}} \|w - u\|_p \tag{2.4}$$

and

$$u_\Lambda = \operatorname*{arg\,min}_{\substack{w \\ \|w\|_0 \leqslant k}} \|w - u\|_p, \tag{2.5}$$

where u_Λ is the restriction of u to the index set Λ, and Λ is the index set of u's k heaviest (in absolute value) coordinates.

Proof. — (2.3) follows directly from (2.2) and (*). We now show (2.2). Let $\sigma = \mathbf{Err}_1^k(u)$.

$$\|u - u^*\|_1 = \|u_\Lambda - u^*_\Lambda\|_1 + \|u_{\overline{\Lambda}} - u^*_{\overline{\Lambda}}\|_1 \tag{2.6}$$

Consider the tail error first:

$$\|u_{\overline{\Lambda}} - u^*_{\overline{\Lambda}}\|_1 \leqslant \underbrace{\|u_{\overline{\Lambda}}\|_1}_{\sigma} + \|u^*_{\overline{\Lambda}}\|_1 \tag{2.7}$$

Just using $\|u^*\|_1 \leqslant \|u\|_1$ we bound

$$\begin{aligned} \|u^*_{\overline{\Lambda}}\|_1 &\leqslant \|u_{\overline{\Lambda}}\|_1 + \|u_\Lambda\|_1 - \|u^*_\Lambda\|_1 && \big(\text{rewriting } \|u^*\|_1 \leqslant \|u\|_1\big) \\ &\leqslant \|u_{\overline{\Lambda}}\|_1 + \|u_\Lambda - u^*_\Lambda\|_1 && \big(\text{triangular inequality}\big) \end{aligned} \tag{2.8}$$

Combine (2.7), (2.8) and (2.6)

$$\|u - u^*\|_1 \leqslant 2\|u_\Lambda - u^*_\Lambda\|_1 + 2\sigma \tag{2.9}$$

Let's examine the head error now:

$$\begin{aligned} \|u_\Lambda - u^*_\Lambda\|_1 &\leqslant k^{1/2}\|u_\Lambda - u^*_\Lambda\|_2 && \big(\text{Cauchy-Schwarz}\big) \\ &\leqslant k^{1/2}\|u - u^*\|_2 \\ &\leqslant k^{1/2}S^{-1/2}\|u - u^*\|_1 && \big(\text{width property}\big) \\ &\leqslant 1/4 \cdot \|u - u^*\|_1 && \big(\text{using } k \leqslant S/16\big) \end{aligned} \tag{2.10}$$

Combine (2.9) and (2.10) to obtain (2.2). ∎

3. Relation to RIP

Here we pursue the connection between RIP and width property of matrices.

Definition 3.1 — *A matrix $A \in \mathbb{R}^{n\times m}$ has the (k,δ)-Restricted Isometry Property (RIP) iff for all $x \in \mathbb{R}^m$ so that $\|x\|_0 \leqslant k$*

$$(1-\delta)\|x\|_2 \leqslant \|Ax\|_2 \leqslant (1+\delta)\|x\|_2. \tag{3.1}$$

It will be helpful to keep the following RIP theorems in mind:

Theorem 3.1 — *A random (with independent Gaussian entries) $A \in \mathbb{R}^{n\times m}$ with $n = \Theta(k\ln(m/k))$ has the $(k, 1/3)$-RIP with high probability.*

Corollary 3.1 — *A random $A \in \mathbb{R}^{n\times m}$ has the $(n/\ln m, 1/3)$-RIP w.h.p.*

Theorem 3.2 — *If $A \in \mathbb{R}^{n\times m}$ is $(O(k), 1/3)$-RIP then for all $u \in \mathbb{R}^m$*

$$\|u - u_A\|_2 \leqslant \frac{O(\mathbf{Err}_1^k(u))}{k^{1/2}} \leqslant \frac{O(\|u\|_1)}{k^{1/2}}, \tag{3.2}$$

where u_A is the recovered signal using Basis Pursuit, defined in (2.1).

The next lemma shows that if a matrix is good enough for LP decoding (e.g. if it is RIP), then it must have the width property.

LEMMA 3.1 (LP implies WP) — *Let $A \in \mathbb{R}^{n\times m}$ and k be so that*

$$\left\| \arg\min_{\delta\in\mathbf{ker}(A)} \|u+\delta\|_1 \right\|_2 \leqslant k^{-1/2}\|u\|_1, \tag{3.3}$$

then A has the width property with $\|x\|_2 \leqslant k^{-1/2}\|x\|_1$ for all $x \in \mathbf{ker}(A)$.

Remark 3.1 — Corollary 3.1 asserts the existence of matrices $A \in \mathbb{R}^{n\times m}$ with $(O(n/\ln m), 1/3)$-RIP, which then have the width property with $k = O(n/\ln m)$ according to Lemma 3.1. Note that this is slightly weaker than (*), where $k = O(n/\ln(m/n))$ is required.

Proof. — Set $\Gamma = \mathbf{ker}(A)$ and let $u \in \Gamma$. Then

$$\arg\min_{\delta\in\mathbf{ker}(A)} \|u+\delta\|_1 = -u$$

and (3.3) gives

$$\|u\|_2 \leqslant k^{-1/2}\|u\|_1.$$

∎

4. Unresolved: Decoding with noise

When recovering $u \in \mathbb{R}^m$ "in the presence of noise" (See [3]), it is assumed that an error $e \in \mathbb{R}^n$ occurs in the measurement process in which event the measured signal is $Au + e$. In this event, [3] consider a relaxed decoding procedure

$$\text{minimize } \|u^*\|_1 \text{ subject to } \|Au^* - (Au+e)\|_2 \leqslant \epsilon,$$

where ϵ is the size of the error term e. In this event, a decoding error guarantee linear in ϵ is obtained when A is RIP. It is unresolved whether a similar noise-resilience can be derived for matrices with the width property only.

References

[1] KASHIN AND TEMLYAKOV, *A remark on compressed sensing*, 2007

[2] LEE, RAZBOROV AND GURUSWAMI, *Almost Euclidean sections of ℓ_1^N using expander codes*, in SODA'08

[3] CANDS, ROMBERG AND TAO, *Stable signal recovery from incomplete and inaccurate measurements*, in Comm. Pure Appl. Math., 59 1207-1223