A Negative Result Concerning Explicit Matrices With The Restricted Isometry Property

Venkat Chandar

March 1, 2008

Abstract

In this note, we prove that matrices whose entries are all 0 or 1 cannot achieve good performance with respect to the Restricted Isometry Property (RIP). Most currently known deterministic constructions of matrices satisfying the RIP fall into this category, and hence these constructions suffer inherent limitations. In particular, we show that DeVore's construction of matrices satisfying the RIP is close to optimal once we add the constraint that all entries of the matrix are 0 or 1.

1 Introduction

Compressive sensing [CRT04,CRT06,Don06] has been put forward as a new paradigm for data acquisition. Compressive sensing is based on the observation that if a vector $x \in \mathcal{R}^n$ is sparse, in the sense that only $k \ll n$ entries of x are nonzero, then we can recover x *exactly* using far fewer than n linear measurements. Formally, we represent the measurements by an $m \times n$ measurement matrix A. If x is the input signal, we measure $Ax \in \mathcal{R}^m$, and our goal is to recover the original vector x from Ax. Of course, if A is invertible then this is simple - the goal is to make m substantially smaller than n. The price we pay is that we will not be able to recover all $x \in \mathcal{R}^n$. Instead, we will only be able to recover k-sparse vectors, i.e., vectors with at most k nonzero entries.

One of the remarkable results of compressive sensing is that there exist matrices A with only $m = O(k \log(n/k))$ rows such that for all k-sparse x, we can recover x exactly from Ax. In [CT05], the *restricted isometry property* (RIP) was introduced as a useful tool for proving this result. In this paper, we will use the following definition of the RIP.

Definition 1. *Let A be an $m \times n$ matrix. Then, A satisfies $(k, D) - RIP$ if there exists $c > 0$ such that for all k-sparse $x \in \mathbb{R}^n$,*

$$c||Ax||_2 \leq ||x||_2 \leq cD||Ax||_2.$$

Here $||\cdot||_2$ denotes the L2 norm. c is a scaling constant that we include for convenience in what follows.

The RIP is useful because it can be shown that if a matrix A satisfies the RIP, then linear programming can be used recover a k-sparse signal x from Ax. We note that there are other properties besides the RIP that can be used to prove that a matrix can be used for compressive sensing (cf. [KT07]). However, one nice aspect of the RIP is that if a matrix satisfies the RIP, it can be shown that recovery is possible even if there is some noise in the measurements.

Given that matrices satisfying the RIP are good for compressive sensing, it is natural to ask how to construct such matrices, and what type of performance we can achieve, e.g., how few rows can A have. It is well-known that there exist matrices satisfying the RIP with $O(k \log(n/k))$ rows, and this is within a constant factor of a lower bound which states that $\Omega(k \log(n/k))$ rows are necessary. However, the proof that such matrices exist uses the probabilistic method, and hence an explicit construction of such a matrix remains elusive.

The currently known explicit constructions ([M06], [CM06], [DeV07]) use many more rows than random matrices. Specifically, these constructions requires $\Omega(k^2)$ rows. One would hope that these constructions could be improved to get explicit matrices with close to $O(k \log(n/k))$ rows. For example, the results of [GIKS07,BI08] show that if we modify the definition of the RIP by using an Lp norm with p close to 1 instead of the L2 norm, then explicit matrices can be constructed that satisfy the modified RIP. Specifically, there exist (non-explicit) matrices with only $O(k \log(n/k))$ rows which satisfy the modified RIP. In addition, there are explicit matrices with $O(k 2^{(\log \log n)^{O(1)}})$ rows that satisfy the modified RIP.

In this note, we show that for L2, a fundamentally different approach is needed. In particular, the previously mentioned explicit constructions all use matrices whose entries are 0 or 1. We show that $0,1$-matrices require substantially more than $O(k \log(n/k))$ rows to satisfy the RIP.

2 $0,1$-Matrices Are Bad With Respect to the RIP

In this section we prove our main result, which shows that $0,1$-matrices cannot achieve good performance with respect to the RIP.

First, we make a couple of definitions that will be used in what follows. Given a $0,1$-matrix A, let f be the minimum fraction of 1's in any column of A, i.e., fm is the minimum number of ones in a column of A. Also, let r be the maximum number of 1's in any row of A. By permuting the rows and columns of A, we may assume that $A_{1,i} = 1$ for $1 \le i \le r$, i.e., that the first row of A has a 1 in the first r entries. In the following proofs, we assume that this reordering has already been done.

The following theorem shows that $0,1$-matrices require substantially more rows that optimal RIP matrices.

Theorem 1. *Let A be an $m \times n$ $0,1$-matrix that satisfies $(k, D) - RIP$. Then,* $m \ge \min\left\{\frac{k^2}{D^4}, \frac{n}{D^2}\right\}$.

Before we prove Theorem 1, we need the following lemma.

Lemma 1. *Let A be an $m \times n$ $0,1$-matrix that satisfies $(k, D) -$ RIP. Then, $f \leq \frac{D^2}{k}$.*

Proof. To start, we square the inequalities defining $(k, D) -$ RIP to obtain

$$c^2||Ax||_2^2 \leq ||x||_2^2 \leq c^2 D^2 ||Ax||_2^2.$$

We can bound c by considering the 1-sparse vector $x = e_i$, where e_i is a standard basis vector with a 1 in a coordinate corresponding to a column of A with fm 1's. Then, the inequalities above give $c^2 fm \leq 1 \leq c^2 D^2 fm$. We will only need the second inequality, which we rewrite as

$$\frac{1}{c^2} \leq D^2 fm. \tag{1}$$

Now, consider the vector x with 1's in the first k positions and 0's elsewhere, i.e., $x = \sum_{i=1}^k e_i$. Let $w(i)$ denote the number of 1's in row i and in the first k columns of A. From the definition of f, the first k columns of A contain at least fmk 1's, so $\sum w(i) \geq fmk$. Applying the Cauchy-Schwarz inequality, we obtain

$$||Ax||_2^2 = \sum_{i=1}^m w(i)^2 \geq \frac{\left(\sum_{i=1}^m w(i)\right)^2}{m} \geq (fk)^2 m.$$

But

$$||Ax||_2^2 \leq \frac{||x||_2^2}{c^2} \leq D^2 fmk,$$

so putting the two inequalities together gives $(fk)^2 m \leq D^2 fmk$. Cancelling out fmk from both sides gives $f \leq \frac{D^2}{k}$. □

Now, we prove Theorem 1.

Proof. Lemma 1 gives us a bound on f that will be useful when $r > k$. We can obtain a second bound on f from the obvious inequality

$$fmn \leq \text{number of 1's in } A \leq rm.$$

Thus, $f \leq \frac{r}{n}$. As we will see, this bound is useful when $r \leq k$.

For notational convenience, let $s = \min\{r, k\}$. Consider the vector x with 1's in the first s positions and 0's elsewhere, i.e., $x = \sum_{i=1}^s e_i$. For this choice of x, we see that

$$||Ax||_2^2 \geq s^2$$

because the first entry of Ax is s. Note that because $s \leq k$, the inequalities defining $(k, D) -$ RIP apply to x. Thus, we can apply equation 1 to get

$$s^2 \leq ||Ax||_2^2 \leq \frac{||x||_2^2}{c^2} \leq D^2 fms.$$

We now plug in our two bounds on f. If $r > k$, then $s = k$, and we use the bound from Lemma 1. This gives

$$k^2 \leq D^2 \frac{D^2}{k} mk,$$

so $m \geq \frac{k^2}{D^4}$. Similarly, if $r \leq k$, then $s = r$, so using our second bound on f gives

$$r^2 \leq D^2 \frac{r}{n} mr.$$

Thus, in this case $m \geq \frac{n}{D^2}$.

Putting the two bounds together, $m \geq \min\{\frac{k^2}{D^4}, \frac{n}{D^2}\}$, completing the proof. □

3 Conclusion

We have shown that $0,1$-matrices cannot be optimal with respect to the RIP. Note that our lower bounds on m are essentially tight for a large range of parameters. Specifically, assume that $D = O(1)$ and that $k = \Omega(n^\alpha)$ for some constant $0 < \alpha < .5$. Then, the matrices constructed in [DeV07] achieve $m = O((\frac{k}{\alpha})^2)$, so these matrices are within a constant factor of our lower bound.

References

[BI08] R. Berinde and P. Indyk. Sparse recovery using sparse random matrices. MIT CSAIL Technical Report, available at http://hdl.handle.net/1721.1/40089.

[CRT04] E. Candès, J. Romberg, and T. Tao. Exact signal reconstruction from highly incomplete frequency information. Submitted for publication, June 2004.

[CRT06] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(1):489509, February 2006.

[CT05] E. Candès and T. Tao. Decoding by linear programming. *IEEE Trans. on Information Theory*, 51(12), pp. 4203 - 4215, December 2005.

[CM06] G. Cormode and S. Muthukrishnan. Combinatorial algorithms for compressed sensing. In Paola Flocchini and Leszek Gasieniec, editors, *Structural Information and Communication Complexity, 13th International Colloquium, SIROCCO 2006, Chester, UK, July 2-5, 2006, Proceedings*, volume 4056 of *Lecture Notes in Computer Science*, pages 280294. Springer, 2006.

[DeV07] R. DeVore. Deterministic Constructions Of Compressed Sensing Matrices. *Preprint*, 2007.

[Don06] D. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):12891306, 2006.

[GIKS07] A. Gilbert, P. Indyk, H. Karloff, and M. Strauss. Combining geometry and combinatorics: a unified approach to sparse signal recovery. Manuscript, 2007.

[KT07] B.Kashin and V. Temlyakov. A remark on compressed sensing. *Preprint*, 2007.

[M06] S. Muthukrishnan. Some algorithmic problems and results in compressed sensing. In *Allerton Conference*, 2006.