Construction of Sparse Binary Sensing Matrices Using Set Systems

R. Ramu Naidu

Abstract Recent developments at the intersection of algebra and optimization theory—by the name of compressed sensing (CS)—aim at providing linear systems with sparse descriptions. The deterministic construction of the sensing matrices is now an active directions in CS. The sparse sensing matrix contributes to fast processing with low computational complexity. The present work attempts to relate the notion of set systems to CS. In particular, we show that the set system theory may be adopted to designing a binary CS matrix of high sparsity from the existing binary CS matrices.

Keywords Compressed sensing · Restricted isometry property · Deterministic construction · Set systems

1 Introduction

In recent years, sparse representations have become a powerful tool for efficiently processing data in nontraditional ways. Compressed sensing (CS) is an emerging area potential for sparsity-based representations. Since the problem of sparse recovery through l_0 norm minimization is generally NP-hard, Donoho et al. [1], Candes [2] and Cohen et al. [3] have made several pioneering contributions and have reposed the problem as an l_1-minimization problem. It is known that restricted isometry property (RIP) is a sufficient condition to ensure the equivalence between l_0 and l_1 norm problems. As verifying RIP is computationally hard, there is much interest in construction of RIP matrices.

Of late, the deterministic construction of binary CS matrices has attracted significant attention. Devore [4], Li et al. [5], Amini et al. [6], Indyk [7] have constructed deterministic binary sensing matrices using ideas from algebra, graph theory, and

R.R. Naidu (✉)
Department of Mathematics, Indian Institute of Technology, Medak,
Hyderabad 502205, A.P., India
e-mail: ma11p003@iith.ac.in

P.N. Agrawal et al. (eds.), *Mathematical Analysis and its Applications*,
Springer Proceedings in Mathematics & Statistics 143,
DOI 10.1007/978-81-322-2485-3_21

coding theory. Devore [4] has been first to construct deterministic binary sensing matrix of size $p^2 \times p^{r+1}$ with p number of ones in each column and coherence being at most $\frac{r}{p}$, for every fixed r and prime power p such that $r < p$. In the present work, using the results from set systems we construct a binary sensing matrix from a given binary sensing matrix in such a way that the resulting matrix is more sparser than the input matrix. Consequently, the new matrix has potential for resulting in fast algorithms.

The paper is organized into several sections. In Sect. 2, we present basic CS theory and the conditions that ensure the equivalence between l_0-norm problem and l_1-norm problem. In Sect. 3, we use the ideas from the set system theory and construct binary sensing matrices of higher sparsity from the existing ones. We present our concluding remarks in the last section.

2 Sparse Recovery from Linear Measurements

CS refers to the problem of reconstruction of an unknown vector $u \in \mathscr{R}^M$ from the linear measurements $y = (\langle u, \phi_1 \rangle, \ldots, \langle u, \phi_M \rangle) \in \mathscr{R}^m$ with $\langle u, \phi_j \rangle$ being the innerproduct between u and ϕ_j. The basic objective in CS is to design a recovery procedure based on the sparsity assumption on u when the number of measurements m is much small compared to M. Sparse representations seem to have merit for various applications in areas such as image/signal processing and numerical computation.

A vector $u \in \mathscr{R}^M$ is said to be k-sparse, if it has at most k nonzero coordinates. One can find the sparse vector from its linear measurements by solving the following l_0-norm optimization problem:

$$\min_{v} \|v\|_0 \text{ subject to } \phi v = y. \tag{1}$$

Here, $\|v\|_0 = |\{i \mid v_i \neq 0\}|$. The l_0-norm problem (1) is an NP-hard problem [2]. Candes et al. [2] have proposed the following l_1-norm minimization problem instead of l_0-norm problem, making it computationally tractable LPP problem:

$$\min_{v} \|v\|_1 \text{ subject to } \phi v = y. \tag{2}$$

Here, $\|v\|_1$ denotes the l_1-norm of the vector $v \in \mathscr{R}^M$.

Donoho et al. [1] and Kashin et al. [8], have provided the conditions under which the solution to l_0-norm problem (1) is the same as that of l_1-norm problem (2). For later use, we denote the solution to l_1-norm problem by $f_\phi(y)$ and solution to l_0-norm problem by $u^0_\phi(y) \in \mathscr{R}^M$.

2.1 Equivalence Between l_0-Norm and l_1-Norm Problems

Definition 1 The mutual coherence $\mu(\phi)$ of a given matrix ϕ is the largest absolute normalized innerproduct between the pairs of columns of ϕ, that is,

$$\mu(\phi) = \max_{1\leq i,j\leq M,\, i\neq j} \frac{|\phi_i^T \phi_j|}{\|\phi_i\|_2 \|\phi_j\|_2}, \tag{3}$$

where ϕ_i is the ith column of ϕ. It is known [1] that for μ-coherent matrices ϕ, one has

$$u_\phi^0(y) = f_\phi(y) = u, \tag{4}$$

provided u is k-sparse with $k < \frac{1}{2}\,(1 + \frac{1}{\mu})$.

Candes et al. ([2] and the references therein) have introduced the following isometry condition on matrices ϕ and have established its important role in CS. An $m \times M$ matrix ϕ is said to satisfy the restricted isometry property (RIP) of order k with constant δ_k if for all k-sparse vectors $x \in \mathscr{R}^M$, we have

$$(1-\delta_k)\,\|x\|_{l_2}^2 \leq \|\phi x\|_{l_2}^2 \leq (1+\delta_k)\,\|x\|_{l_2}^2\,. \tag{5}$$

The following proposition [9] relates the RIP constant δ_k and μ.

Proposition 1 *Suppose that $\phi_1, \ldots, \phi_M$ are the unit-normed columns of the matrix ϕ with coherence μ. Then ϕ satisfies the RIP of order k with constant $\delta_k = (k-1)\mu$.*

Candes [2] has shown that whenever ϕ satisfies RIP of order $3k$ with $\delta_{3k} < 1$, the CS reconstruction error satisfies the following estimate

$$\left\| u - f_\phi(\phi u) \right\|_{l_2^M} \leq C k^{\frac{-1}{2}} \sigma_k(u)_{l_1^M}, \tag{6}$$

where $\sigma_k(u)_{l_1^M}$ denotes the l_1 error of the best k—term approximation, and the constant C depends only on δ_{3k}. This implies that the bigger the value of k for which we can verify the RIP then better the guarantee we have on the performance of ϕ.

One of the important problems in CS theory deals with constructing CS matrices that satisfy the RIP for the largest possible range of k. It is known that the widest range possible is $k \leq C\frac{m}{\log(\frac{M}{m})}$ [4, 10–12]. However, the only known matrices that satisfy the RIP for this range are based on random constructions [10]. To the best of our knowledge, designing the good deterministic constructions of RIP matrices is still an open problem.

Since the sparsity of the sensing matrix is key to minimizing the computational complexity associated with the matrix vector multiplication, it is desirable that the CS matrix has smaller density. The sparse sensing matrix may contribute to fast processing with low computational complexity in compressed sensing [13].

Definition 2 [14] The density of a matrix is the ratio of the number of its nonzero entries to the total number of its entries.

It may be noted that the density of the sensing matrix constructed by Devore [4] is $\frac{1}{p}$. The sensing matrix constructed by Li et al. [5] have $\frac{1}{q}$ as density. This matrix, a generalization of [4], is of size $|\mathscr{P}|q \times q^{\mathscr{L}(G)}$, where q is any prime power and $\mathscr{P}$ is the set of all rational points on algebraic curve $\mathscr{X}$ over finite field $\mathscr{F}_q$. Amini et al. [6] have constructed binary sensing matrices using OOC codes. The density of this matrix is $\frac{\lambda}{m}$, where m is row size and λ is the number of ones in each column.

Many data mining tasks can be concerned with identifying a small number of interesting items from a tremendously large group without exceeding certain resource constraints. Specific examples [15] include the sketching and monitoring of heavy hitters in high-volume data streams, source localization in sensor networks, multiplier-less data compression and tomography. Note that, all these applications naturally correspond to binary matrices. Furthermore, binary matrices with small density are generally better. Thus, we focus on designing sparse binary matrices herein.

The present work attempts to address the deterministic construction of new binary sensing matrix of smaller density from a given binary sensing matrix. Suppose ϕ is a binary CS matrix of size $m \times M$ with $\frac{m(m+1)}{2} < M$. In the next section, using the results from set systems, we construct a binary sensing matrix ψ from ϕ in such a way that the resulting matrix ψ is more sparse compared to the given matrix ϕ.

3 Construction of Binary CS Matrix of Smaller Density from Existing Binary Matrix

Before presenting the main result, we discuss the definitions and results [16] relevant to our construction methodology. Let $V = \{v_1, v_2, \ldots, v_m\}$ be a set of m elements (treated as "universe"). A set system $\mathscr{S}$ on V is simply some subset chosen from all of the subsets of V, that is, $\mathscr{S} \subset 2^V$, the power set. A hypergraph is a collection of several subsets of V, where some subsets may be present with a multiplicity greater than 1. A set system may, however, contain each subset of V at most once.

Definition 3 Let $\mathscr{H} = \{H_1, H_2, \ldots, H_M\}$ be a hypergraph of M sets over the universe V, and let $\phi = \{\phi_{ij}\}$ be the $m \times M$ binary sensing incidence matrix of hypergraph $\mathscr{H}$, that is, the columns of ϕ correspond to the sets of H. The characteristic vector on each H_j gives the j^th column in ϕ, that is, $\phi_{ij} = 1$ if $x_i \in H_j$ otherwise $\phi_{ij} = 0$.

Definition 4 Let $A = \{a_{ij}\}$ and $B = \{b_{ij}\}$ be the two $m \times M$ matrices over a ring R. Their dream product is an $m \times M$ matrix $C = \{c_{ij}\}$, denoted by $A \odot B$, and is defined as $c_{ij} = a_{ij} b_{ij}$ for $1 \leq i \leq m$, $1 \leq j \leq M$.

Definition 5 Let $f(x_1, x_2, \ldots, x_m) = \sum_{I \subseteq \{1,2,\ldots,m\}} a_I x_I$ be a multilinear polynomial, where $x_I = \prod_{i \in I} x_i$. Let $w(f) = |\{a_I : a_I \neq 0\}|$ and let $L_1(f) = \sum_{I \subseteq \{1,2,\ldots,m\}} |a_I|$.

Definition 6 Let $\mathscr{H}$ be a set system on the universe X with $m \times M$ incidence matrix ϕ. Let $f(x_1, x_2, \ldots, x_m) = \sum_{I \subseteq \{1,2,\ldots,m\}} a_I x_I$ be a multilinear polynomial with nonnegative integer coefficients or coefficients from $\mathscr{Z}_r$. Then $f(\mathscr{H}_\phi)$ is a hypergraph on the $L_1(f)$-element vertex set, and its incidence matrix is the $L_1(f) \times M$ matrix ψ. The rows of ψ correspond to x_I's of f; there are a_I identical rows of ψ corresponding to the same x_I. The row, corresponding to x_I is defined as the dream product of those rows of ϕ that correspond to $v_i, i \in I$.

Lemma 1 [16] *Suppose in the Definition 6, the coefficients of $x_1, x_2, \ldots, x_m$ are nonzeros in f. Then the resulting Hypergraph $f(\mathscr{H}_\phi)$ is a set system [16].*

The most remarkable property of $f(\mathscr{H}_\phi)$ is given by the following theorem:

Theorem 1 [16] *Let $\mathscr{H} = \{H_1, H_2, \ldots, H_M\}$ be a set system and ϕ its $m \times M$ incidence matrix. Let f be a multilinear polynomial with nonnegative integer coefficients or coefficients from $\mathscr{Z}_r$. Let $f(\mathscr{H}_\phi) = \{\widehat{H}_1, \widehat{H}_2, \ldots, \widehat{H}_M\}$. Then for any $1 \le k \le M$ and for any $1 \le i_1 < i_2 < \cdots < i_k \le M$: $f(H_{i_1} \cap H_{i_2} \cap \ldots \cap H_{i_k}) = |\widehat{H}_{i_1} \cap \widehat{H}_{i_2} \cap \ldots \widehat{H}_{i_k}|$.*

Following theorem discusses the construction of a new set system from a given set system using the Definition 6, Lemma 1 and Theorem 1.

Theorem 2 [16] *Let f be an m—variable symmetric polynomial with nonnegative integer coefficients, and $\mathscr{H}$ a set system of size M on the m element universe with $m \times M$ incidence matrix ϕ. Suppose that*

$$L(\mathscr{H}) = \{|H_i \cap H_j|, H_i \neq H_j, H_i, H_j \in \mathscr{H}\} = \{l_1, l_2, \ldots, l_s\}.$$

Then one may construct in $O(L_1(f)mM)$ time a hypergraph $f(\mathscr{H}_\phi)$ of size M on the $L_1(f)$—vertex universe such that the sizes of the pairwise intersections of the sets of $f(\mathscr{H}_\phi)$ are

$$f(l_1), f(l_2), \ldots, f(l_s).$$

3.1 Set Systems for Designing CS Matrices

Using the afore-stated results [16] from set system theory, we construct a new binary sensing matrix from a given binary sensing matrix. The new matrix has small density as compared to the given one.

Theorem 3 *Suppose $f(x_1, x_2, \ldots, x_m) = x_1 + x_2 + \cdots + x_m + \sum_{i<j} x_i x_j$ is a symmetric polynomial. Let ϕ be a binary sensing matrix of size $m \times M$ such that $\frac{m(m+1)}{2} < M$ with the coherence being at most $\frac{r}{k}$. Here k represents the number of nonzero elements that each column of ϕ has. Then there exists a binary sensing matrix ψ of size $\frac{m(m+1)}{2} \times M$ whose coherence is at most $\frac{r+\binom{r}{2}}{k+\binom{k}{2}}$.*

Proof Define $\mathscr{H} = \{H_i : 1 \leq i \leq M, H_i = \text{supp}(\phi_i), \text{where } \phi_i \text{ is } i\text{th column of } \phi\}$. Since all columns of ϕ are distinct, $\mathscr{H}$ is a set system. Let $L(\mathscr{H}) = \{|H_i \cap H_j|, H_i \neq H_j, H_i, H_j \in \mathscr{H}\} = \{l_1, l_2, \ldots, l_s\}$. Since the coherence of the matrix ϕ is at most $\frac{r}{k}$, the cardinality of overlap between the supports of any two columns is at most r. Consequently, $l_i \leq r$ for all i. Let $X = \{1, 2, \ldots, m\}$.

Since $f(x_1, x_2, \ldots, x_m) = x_1 + x_2 + \cdots + x_m + \sum_{i<j} x_i x_j$ is a symmetric polynomial and $\mathscr{H}$ is a set system, we have $L_1(f) = m + \binom{m}{2}$ and $f(\mathscr{H}_\phi)$ is a set system of size M on $L_1(f)$-element universe, from Theorem 2 and Lemma 1.

Define $(v_{ij})_{m\times 1}$ to be the characteristic vector on $H_i \cap H_j$ in the universe X. Since each $l_s = |H_i \cap H_j|$ for some $i \neq j$, $f(l_s) = f((v_{ij})) \leq f(r) = r + \binom{r}{2}$. It follows that $f(l_i) \leq r + \binom{r}{2}$ for all i. Therefore, the sizes of the pairwise intersections of the sets of $f(\mathscr{H}_\phi)$ is at most $r + \binom{r}{2}$. The incidence matrix ψ of the set system $f(\mathscr{H}_\phi)$ is of size $L_1(f) \times M$, that is, $(m + \binom{m}{2}) \times M$. From the hypothesis of the theorem $\left(m + \binom{m}{2}\right) < M$, so it is an underdetermined system and its first m rows are same as ϕ and remaining $\binom{m}{2}$ rows are the dream products of first m rows. Each column in ψ contains $k + \binom{k}{2}$ number of ones. The cardinality of overlap between any two columns is at most $r + \binom{r}{2}$. It follows that coherence of the matrix ψ is $\mu(\psi)$, which is at most equal to $\frac{r+\binom{r}{2}}{k+\binom{k}{2}}$.

The following theorem concludes that the matrix $\psi_0 = \frac{1}{\sqrt{k+\binom{k}{2}}}\psi$ defined is RIP compliant.

Theorem 4 *The matrix* $\psi_0 = \frac{1}{\sqrt{k+\binom{k}{2}}}\psi$ *has the RIP with* $\delta = (k-1)\left(\frac{r+\binom{r}{2}}{k+\binom{k}{2}}\right)$ *whenever* $k - 1 < \frac{k+\binom{k}{2}}{r+\binom{r}{2}}$.

Proof Proof follows from the Proposition 1 and Theorem 3 □

Remark 1 The density of the new matrix ψ is $\frac{k+\binom{k}{2}}{m+\binom{m}{2}}$, which is smaller than $\frac{k}{m}$, the density of ϕ.

4 Concluding Remarks

As the sensing matrices of higher sparsity (or lower density) have potential for fast processing, the construction of such matrices is of relevance. In the present work, we have used the ideas from the set system theory and have showed that a CS matrix of higher sparsity can be generated from a given binary CS matrix.

Acknowledgments The author gratefully acknowledges the support (Ref No. 20-6/2009(i)EU-IV) that he receives from UGC, Govt of India. Author is thankful to Dr. C. S. Sastry and Dr. Phanindra Varma for fruitful discussions.

References

1. Donoho, D.L., Elad, M., Temlyakov, V.N.: Stable recovery of sparse overcomplete representations in the presence of noise. IEEE Trans. Inf. Theory **52**, 6–18 (2006)
2. Candes, E.: The restricted isometry property and its implications for compressed sensing. Comptes Rendus Mathematique **346**, 589–592 (2008)
3. Cohen, A., Dahmen, W., DeVore, R.: Compressed sensing and best k-term approximation. J. Amer. Math. Soc. **22**(1) 211–231 (2009)
4. DeVore, R.A.: Deterministic constructions of compressed sensing matrices. J. Complex. **23**, 918–925 (2007)
5. Li, S., Gao, F., Ge, G., Zhang, S.: Deterministic construction of compressed sensing matrices via algebraic curves. IEEE Trans. Inf. Theory **58**, 5035–5041 (2012)
6. Amini, A., Marvasti, F.: Deterministic construction of binary, bipolar and ternary compressed sensing matrices. IEEE Trans. Inf. Theory **57**, 2360–2370 (2011)
7. Indyk, P.: Explicit constructions for compressed sensing matrices. In: Proceedings 19th Annual ACM-SIAM Symposium Discrete Algorithms, pp. 30–33 (2008)
8. Kashin, B.S., Temlyakov, V.N.: A remark on compressed sensing. Math. Notes **82**(6), 748–755 (2007)
9. Bourgain, J., Dilworth, S., Ford, K., Konyagin, S., Kutzarova, D.: Explicit constructions of RIP matrices and related problems. Duke Math. J. **159**, 145–185 (2011)
10. Baraniuk, R., Davenpor, M., De Vore, R., Wakin, M.: A simple proof of the restricted isometry property for random matrices. Constr. Approx. **28**(3), 253–263 (2008)
11. Garnaev, A., Gluskin, E.: The widths of a Euclidean ball. Dokl. Akad. Nauk USSR **277**, 1048–1052 (1984); English transl. in Soviet Math. Dokl. **30**, 200–204 (1984)
12. Kashin, B.S.: Widths of certain finite-dimensional sets and classes of smooth functions. Izv. Akad. Nauk SSSR, Ser. Mat. **41**, 334–351 (1977). English transl. in Math. USSR IZV. **11**, 317–333 (1978)
13. Gilbert, A., Indyk, P.: Sparse recovery using sparsematrices. Proc. IEEE **98**(6), 937–947 (2010)
14. Fickus, M., Mixon, D.G., Tremain, J.C.: Steiner equiangular tight frames. Lin. Alg. Applicat. **436**(5), 1014–1027 (2012)
15. Iwen, M.: Compressed sensing with sparse binary matrices: instance optimal error guarantees in near-optimal time. J. Complex. **30**(1), 1–15 (2014)
16. Grolmusz, V.: Constructing set-systems with prescribed intersection sizes. J. Algorithms **44**, 321–337 (2001)