SAND2011-7103C

On Discovering the Compressive Sensing Matrix From Few Signal/Measurement Pairs

Hyrum S. Anderson
Sandia National Laboratories
Albuquerque, NM 87185

***Abstract*—Recently, it has been shown that a randomly-generated compressive sensing (CS) measurement matrix may double as a long cryptographic key that provides a degree of secrecy: an eavesdropper that inspects only the compressed samples can learn almost nothing about the encoded sparse vector. In practical CS applications, however, the measurement matrix is often structured in order to simplify CS reconstruction, and this structure effectively reduces the length of the cryptographic key. This work addresses the scenario in which the eavesdropper is allowed to observe a few signal/measurement pairs, and investigates how many pairs must be observed in order to exactly recover the CS measurement matrix. Two matrix classes are studied that are common in practice: Toeplitz, and Fourier measurement matrices.**

I. Introduction

Compressive sensing (CS) has begun to receive attention from an information-theoretic viewpoint [1]–[3]. In canonical CS, one aims to estimate K-sparse $\mathbf{x} \in \mathbb{R}^N$ from samples

$$\mathbf{y} = \Phi\mathbf{x}, \quad \mathbf{y} \in \mathbb{R}^M, \Phi \in \mathbb{R}^{M\times N}, M \ll N. \tag{1}$$

Candes, Romberg and Tao [4] and Donoho [5] provided conditions on Φ that allow perfect recovery of $\mathbf{x}$ from only $M = O(K \log N)$ measurements using a linear program:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad s.t. \quad \mathbf{y} = \Phi\mathbf{x}. \tag{2}$$

Importantly, Candes and Tao [6] showed that Φ generated from iid Gaussian draws allows recovery of $\mathbf{x}$ with high probability. Recently cryptography researchers have addressed the secrecy that Φ provides when generated with iid Gaussian entries. In [1], it was shown that if an eavesdropper observes $\mathbf{y}$ and knows that $\mathbf{x}$ is K-sparse, then Φ provides computational secrecy (NP-hard discovery). Agrawal and Vishwanath demonstrated Wolkowitz secrecy in the context of a wiretap channel [2], while Reeves et al. study Φ's secrecy capacity [3].

Despite the secrecy results for randomly-generated Φ, current research in compressive sensing is driving towards structured Φ to improve reconstruction speed and ease of implementation. The structure imposed on Φ reduces its secrecy capacity to a degree which has yet to be investigated.

In this work, two important classes of structured matrices are considered: Toeplitz [7] and random Fourier [4]. Rather than a blind frontal attack as in [1]–[3], this work addresses a much less difficult scenario for discovering Φ. Here, the eavesdropper knows the general structure of Φ, observes P measurement vectors $\{\mathbf{y}_i\}_{i=1}^P$ and obtains the first $Q \leq P$ source vectors $\{\mathbf{x}_i : \mathbf{y}_i = \Phi\mathbf{x}_i\}_{i=1}^Q$. Further, it is known that the true $\{\mathbf{x}_i\}_{i=1}^P$ live near a low-dimensional subspace, such that the matrix $X = [\mathbf{x}_1 \cdots \mathbf{x}_P]$ can be decomposed as $X = L + S$, where L is low rank (eavesdropper knows rank), and S is a sparse matrix (eavesdropper knows sparsity). Such a decomposition is common for modeling video frames with slowly-varying background in the columns of L, and foreground represented in columns of S [8]. With these assumptions, the following algorithm to discover Φ is proposed:

Input: $\{\mathbf{y}_i\}_{i=1}^P$, $\{\mathbf{x}_i : \mathbf{y}_i = \Phi\mathbf{x}_i\}_{i=1}^Q$, structure of Φ
Output: estimated measurement matrix $\hat{\Phi}$
Initialize: $\{\hat{\mathbf{x}}_i\}_{i=Q+1}^P$ as random draws from $\{\mathbf{x}_i\}_{i=1}^Q$.
(1) Estimate $\hat{\Phi}$ from $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^Q$, $\{(\hat{\mathbf{x}}_i, \mathbf{y}_i)\}_{i=Q+1}^P$
(2) Solve (2) for $\{\hat{\mathbf{x}}_i\}_{i=Q+1}^P$ and fixed $\hat{\Phi}$
(3) Decompose $\hat{X} = [\mathbf{x}_1 \cdots \mathbf{x}_Q\, \hat{\mathbf{x}}_{Q+1} \cdots \hat{\mathbf{x}}_P]$ into $\hat{X} = L + S + G$ with noise matrix G [8] s.t. columns $j = 1, \ldots, Q$ of G are zero; reassign $\hat{X} = L + S$
(4) Iterate **(1)**-**(3)** until negligible change in $\hat{\Phi}$

This work investigates how large P and Q must be to recover Φ using the above algorithm. It is demonstrated that a structured CS measurement matrix Φ can be discovered for $Q < P$. Section II addresses the number P needed to recover Φ for Toeplitz and Fourier CS measurement matrices. Section III shows empirical results for Q.

II. Solving for Φ from P input/output pairs

This section addresses step **(1)** of the proposed algorithm: discovering Φ for Toeplitz and Fourier measurement matrices from P input/output pairs, where P is determined below.

A. Toeplitz measurement matrix

A Toeplitz measurement matrix

$$\Phi = \begin{bmatrix} \phi_n & \phi_{n-1} & \cdots & \phi_1 \\ \phi_{n+1} & \phi_n & \cdots & \phi_2 \\ \vdots & \vdots & \vdots & \vdots \\ \phi_{n+m-1} & \phi_{n+m-2} & \cdots & \phi_m \end{bmatrix},$$

is completely characterized by its first row $\mathbf{r} \triangleq [\phi_n\, \phi_{n-1} \cdots \phi_1]^T$ and column $\mathbf{c} \triangleq [\phi_{n+1} \cdots \phi_{n+m-1}]^T$

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporatoin, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

(ignoring duplicate ϕ_n). Given $\mathbf{x}_i$ and $\mathbf{y}_i$, (1) may be re-written as $\tilde{X}_i \begin{bmatrix} \mathbf{r} \\ \mathbf{c} \end{bmatrix} = \mathbf{y}_i$, where $\tilde{X}_i$ is the $M \times (N+M-1)$ concatenation of a rectangular Hankel matrix (goes with $\mathbf{r}$) with a Toeplitz matrix (goes with $\mathbf{c}$):

$$\begin{bmatrix} x_{i,1} & x_{i,2} & \cdots & x_{i,N} & 0 & 0 & \cdots & 0 \\ x_{i,2} & x_{i,3} & \cdots & 0 & x_{i,1} & 0 & \cdots & 0 \\ x_{i,3} & x_{i,4} & \cdots & 0 & x_{i,2} & x_{i,1} & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ x_{i,M} & x_{i,M+1} & \cdots & 0 & x_{i,M-1} & x_{i,M-2} & \cdots & x_{i,1} \end{bmatrix}.$$

To solve for $\mathbf{r}$ and $\mathbf{c}$, stack P such equations together and solve the constrained least squares problem

$$\min_{r,c} \left\| \begin{bmatrix} \tilde{X}_1 \\ \vdots \\ \tilde{X}_P \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ \mathbf{c} \end{bmatrix} - \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_P \end{bmatrix} \right\|^2 \text{s.t.} \begin{bmatrix} \tilde{X}_1 \\ \vdots \\ \tilde{X}_Q \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ \mathbf{c} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_Q \end{bmatrix}. \tag{3}$$

Since each $\mathbf{x}_i$ is a K-sparse vector, unless P is large enough, some columns of the $PM \times (N+M-1)$ matrix in (3) may be entirely zero, and therefore uninformative about entries of $\mathbf{r}$ or $\mathbf{c}$. Determining P can be cast as a simple generalization of the coupon collector's problem by assuming that for each $\mathbf{x}_i$ the K non-zero elements are distributed uniformly, and that $\{\mathbf{x}_i\}_{i=1}^P$ are mutually independent draws. In particular, denote by random variable T the number of trials required for one to collect N coupons when collecting K coupons per trial. It has been shown that T is sharply concentrated around $\mathrm{E}[T] = \frac{N}{K}\sum_{n=1}^{N}\frac{1}{n}$ [9, Ch. 3], so that $P \geq \frac{N}{K}\sum_{n=1}^{N}\frac{1}{n}$ suffices for our purposes. A stricter bound based on vanishing probability may be derived from the variance of T and Chebyshev's inequality, but for brevity is left to future work.

B. Random Fourier measurements

A CS Fourier measurement matrix may be written as $\Phi = EF$, where F is an $N \times N$ Fourier matrix and E is an $M \times N$ matrix consisting of M randomly selected (without replacement) rows of an $N \times N$ identity matrix. Thus, to discover Φ one needs only learn the i, j where $E_{i,j} = 1$, for which there is exactly one per row.

Concatenate the rows of E into a sparse vector $\mathbf{e} \in \mathbb{R}^{MN}$. Then, given $\mathbf{x}_i$ and $\mathbf{y}_i$, rewrite (1) as $\hat{F}_i\mathbf{e} = \mathbf{y}_i$, where $\hat{F}_i$ is the $M \times MN$ constant-block diagonal matrix:

$$\begin{bmatrix} (F\mathbf{x}_i)^T & 0 & \cdots & 0 \\ 0 & (F\mathbf{x}_i)^T & \cdots & 0 \\ 0 & 0 & \cdots & (F\mathbf{x}_i)^T \end{bmatrix}.$$

Since $\mathbf{e}$ is an M-sparse vector in $\mathbb{R}^{MN}$, one might employ ℓ_1 minimization as in (2) to find a sparse solution. However, the problem is further constrained in that exactly M elements of $\mathbf{e}$ are unity, and the remaining are zero. Thus $\|\mathbf{e}\|_1 = \sum_i |e_i| = \sum_i e_i = M$ is a non-informative objective when including the tighter constraints. Instead, solve for $\mathbf{e}$ (with matrix form E)

$$\min_e \sum_{i=Q+1}^{P} \|\hat{F}_i e - y_i\|^2 \text{ s.t. } \hat{F}_k e = y_k,\ k = 1, \ldots, Q \tag{4a}$$

$$0 \leq E_{ij} \leq 1, E\mathbf{1} = \mathbf{1},\ \mathbf{0} \leq E^T\mathbf{1} \leq \mathbf{1} \tag{4b}$$

where (4b) promotes M-sparsity with a single 1 per row of E. Defining $\hat{F} \in \mathbb{R}^{PM \times MN}$ as the matrix formed by stacking $\hat{F}_i, i = 1, \ldots, P$, atop one another, CS theory dictates the number or matrix rows PM needed to recover $\mathbf{e}$, which depends on properties of $\hat{F}$ [5]. Matrices with structure like $\hat{F}$ have yet to be studied; and is a topic for future work. Instead, we employ the empirical rule-of-thumb $(\#rows) \geq \frac{(\#cols)}{4}$ [5], which implies $P \geq \frac{N}{4}$.

III. Empirical Q for discovering Φ

A simulated dataset $\{\mathbf{x}_i \in \mathbb{R}^{100}\}_{i=1}^P$ is generated according to the model $\mathbf{x}_i = \mathbf{s}_0 + \mathbf{s}_i$, where $K_0 = \|\mathbf{s}_0\|_0$ and $K_1 = \|\mathbf{s}_i\|_0$, so that $\|\mathbf{x}_i\|_0 \leq K_0 + K_1$. This model ensures that $X = L + S$, where L is a (sparse) rank-1 matrix, and S is sparse. Each nonzero element of $\{\mathbf{s}_i\}_{i=0}^P$ is drawn from a standard normal distribution. The true measurement Toeplitz or Fourier matrix Φ is generated with $M = \frac{N}{4}$ rows. For step **(3)** of the proposed method, GoDec [8] is trivially modified to constrain $N_{i,j} = 0, i = 1, \ldots, M, j = 1, \ldots, Q$. Figure 1 reports the probability of recovering Φ as a function of Q, where recovery is successful if $\|\Phi - \hat{\Phi}\|_F \leq 10^{-3}$.

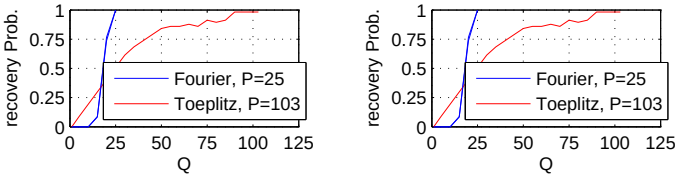

Fig. 1. Probability of recover for (left) $K_0 = 5, K_1 = 5$ and (right) $K_0 = 3, K_1 = 7$, where probability is measured over 100 trials.

References

[1] Y. Rachlin and D. Baron, "The secrecy of compressed sensing measurements," *Comm., Control, and Computing, 46th Annual Allerton Conf.*, pp. 813–817, 2009.

[2] S. Agrawal and S. Vishwanath, "Secrecy using compressive sensing," *Proc. Info. Theory Wkshp*, 2011.

[3] G. Reeves, N. Goela, N. Milosaviljevic, and M. Gastpar, "A compressed sensing wire-tap channel," *Proc. Info. Theory Wkshp*, 2011.

[4] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Info. Th.*, vol. 52, no. 2, pp. 489–509, 2006.

[5] D. Donoho, "Compressed sensing," *IEEE Trans. Info. Th.*, vol. 24, no. 4, pp. 1289–1306, 2006.

[6] E. Candes and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Info. Th.*, vol. 52, no. 12, pp. 5406–5425, 2006.

[7] W. Yin, S. P. Morgan, J. Yang, and Y. Zhang, "Practical compressive sensing with Toeplitz and circulant matrices," Rice University CAAM Tech Report TR10-01, Tech. Rep., 2010.

[8] T. Zhou and D. Tao, "Godec: Randomized low-rank & sparse matrix decomposition in noisy case," *Proc. Int'l Conf. Mach. Learn.*, 2011.

[9] R. Motwani and P. Raghavan, *Randomized algorithms*. Cambridge University Press, 1995.