LOW-RANK MATRIX COMPLETION FOR ARRAY SIGNAL PROCESSING

Zhiyuan Weng and Xin Wang *

Department of Electrical and Computer Engineering, Stony Brook University
Stony Brook, New York, USA

ABSTRACT

In this paper, we propose the application of low-rank matrix completion techniques for array signal processing. Specifically, under the assumption that the number of targets is generally much smaller than the number of antennas, the received signals can form a low-rank matrix with noise. According to the recently proposed matrix completion theory, only a subset of the entries are enough to recover the whole matrix as long as certain conditions are met, thus the implementation cost of obtaining a matrix could be reduced. We prove that the matrix formed by the received signals satisfies the condition for matrix recovery. Moreover, a uniform spatial sampling (USS) method is proposed, which is easy for hardware implementation and also could take advantage of the available number of front-end elements to achieve a better performance. We analytically prove that the probability of matrix recovery failure under the USS model is asymptotically equal to that under the Bernoulli model. Simulation results demonstrate that the matrix recovery performance under the USS model is very close to that using the uniform model.

Index Terms— Matrix completion, Array signal processing

1. INTRODUCTION

Array signal processing techniques have been widely used in radar, sonar and other applications. With the increasing number of antenna elements, the cost of signal acquisition and processing becomes higher and higher. For example, an array with hundreds or even a thousand antenna elements would require the same number of frond-end hardware units (sampling filter, analog-digital converter etc.) each corresponds to one element. The need of a huge number of front-end units presents a serious challenge on the design of the system, especially on the design of mobile systems such as air-born radars where the available space and power are limited.

In this paper, we propose the use of matrix completion to conquer the challenge. Matrix completion is a new technique which can be applied to recover a low-rank matrix from a subset of the matrix entries [1, 2, 3, 4, 5, 6, 7]. More specifically, with some prior knowledge that a matrix has low rank, if some conditions are satisfied, the matrix can be recovered from partial matrix entries by minimizing the nuclear norm of the matrix.

In the scenarios of radar or sonar detection, the number of targets is generally small. As a result, the received signals, i.e. the superposition of the reflected signals from the targets plus noise can form a low-rank matrix. According to the matrix completion theory, only a few randomly chosen entries of the matrix are enough to recover the whole matrix. Thus we don't need to observe all the entries of the matrix. This means it is not necessary to sample all the signals on each element of the array in each time slot. Therefore the number of front-end hardware units can be significantly reduced. It also helps to reduce the energy consumption of the radar dramatically.

*This research was sponsored by the Office of Naval Research under the grant number N000141110575.

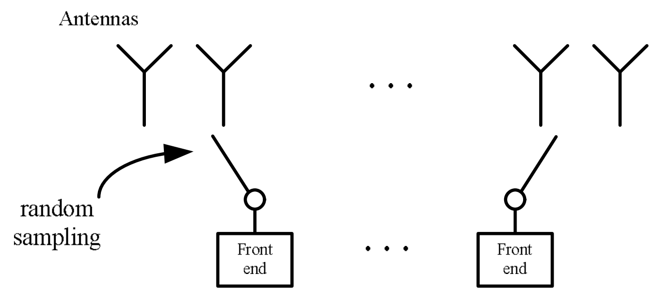

Fig. 1. structure of signal sampling

The subset of entries must meet certain condition, and there are samples from each row and column. Sampling according to two models have been analyzed and shown to be able to recover the matrix satisfactorily [1], the *Bernoulli model* and the *uniform model* as shown in Section 5. In our envisioned application scenarios of array signals, the number of front-end elements is fixed. To take the full advantage of the available array elements and for the implementation simplicity, we propose a *uniform spatial sampling model* (USS), and show its performance is comparable to the Bernoulli model and the uniform model.

The paper is organized as follows. We formulate the problem in section II, and prove the received signals naturally satisfy the condition for matrix completion in section III. We analyze the USS model in section IV, and verify the model through simulations in Section V. Section VI concludes the work. In this paper, we use the uppercase letters to denote matrices, bold lowercase letters to denote vectors and non-bold lowercase letters to denote scalars.

2. PROBLEM FORMULATION

We consider an array with n elements. The discrete-time baseband transmitted signals can be written as a matrix $Z = [\mathbf{z}_1, \ldots, \mathbf{z}_n]^{\mathrm{T}}$ where $\mathbf{z}_i = [z_{i,1}, \ldots, z_{i,n}]$ corresponds to the waveform transmitted by the ith antenna, $(\cdot)^T$ denotes the transpose. Note that the number of samples of each transmitted signal pulse is also set equal to n. The transmitter steering vector can be represented as

$$\mathbf{a}(\alpha) = \left[e^{-j2\pi f_0 \tau_1(\alpha)}, \ldots, e^{-j2\pi f_0 \tau_n(\alpha)}\right]^{\mathrm{T}}, \tag{1}$$

where $\tau_i(\alpha)$ is the inter-element time delay difference between each element to the reference element when the target is at the direction

α and f_0 is the frequency of the carrier. We set the first element of the array as the reference and assume its $\tau_1(\alpha)$ is zero.

We adopt the classical Swerling model here and consider a target with many scatterers which exhibit random, independent and isotropic scintillation [8]. Assuming there are r targets, we denote the reflectivity factors of the kth target by η_k. Also we assume that $\{\eta\}_{k=1}^{r}$ are identically and independently distributed (i.i.d.) as complex Gaussian zero-mean variables with the variance ν_1^2.

As a result, the received signal can be described as

$$Y = \sum_{k=1}^{r} \eta_k \sqrt{\frac{t}{n}} \mathbf{a}(\alpha_k)\mathbf{a}^T(\alpha_k) Z + N \quad (2)$$

$$= W + N \quad (3)$$

where η_k and α_k are the parameters correspond to the kth target. t is the total energy used for the transmission of signals and N is an i.i.d. circular complex Gaussian noise matrix with variance ν_2^2 for each element.

3. MATRIX COMPLETION

Matrix completion theory says that one can recover an unknown $n \times n$ matrix of low rank r from just about $nr \log^2 n$ noisy samples with an error that is proportional to the noise level [4].

Suppose $M \in \mathbb{C}^{n \times n}$ is the matrix we would like to know. We use r to denote the rank of M. We only know a few entries of the matrix, say M_{ij} where $(i, j) \in \Omega$. Ω is the randomly chosen subset of the complete set of entries of the matrix. We use $\mathcal{P}_\Omega : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ to denote the sampling operator which is defined by

$$[\mathcal{P}_\Omega(X)]_{ij} = \begin{cases} X_{ij}, & (i, j) \in \Omega \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

In principle, if the singular vectors of M are sufficiently spread, one could recover the unknown matrix by solving

$$\begin{aligned} &\text{minimize} \quad \text{rank}(X) \\ &\text{subject to} \quad \mathcal{P}_\Omega(X) = \mathcal{P}_\Omega(M), \end{aligned} \quad (5)$$

where X is the variable matrix. Because the rank minimization problem is NP-hard, an alternative method has been proposed [2, 1] to solve the problem as follows:

$$\begin{aligned} &\text{minimize} \quad \|X\|_* \\ &\text{subject to} \quad \mathcal{P}_\Omega(X) = \mathcal{P}_\Omega(M) \end{aligned} \quad (6)$$

where $\|X\|_* := \sum_k \sigma_k$ and $\sigma_1, \cdots, \sigma_r \geq 0$ are the singular values of X. To formally state the assumption of 'sufficiently spread', we need to introduce more notations. Suppose the singular value decomposition (SVD) of M can be represented as:

$$M = \sum_{k \in \{1, \cdots, r\}} \sigma_k \mathbf{u}_k \mathbf{v}_k^H \quad (7)$$

where the singular vector $\mathbf{u}_1, \cdots, \mathbf{u}_r, \mathbf{v}_1, \cdots, \mathbf{v}_r \in \mathbb{C}^n$ are two sets of orthogonal vectors and $(\cdot)^H$ denotes Hermitian transpose. We denote by P_U and P_V the orthogonal projections onto the column and row space of M respectively. Note that

$$P_U = \sum_{1 \leq i \leq r} \mathbf{u}_i \mathbf{u}_i^H; \quad P_V = \sum_{1 \leq i \leq r} \mathbf{v}_i \mathbf{v}_i^H. \quad (8)$$

Define the matrix E as

$$E := \sum_{1 \leq i \leq r} \mathbf{u}_i \mathbf{v}_i^H. \quad (9)$$

To recover the matrix from part of the entries, the vectors u_i, v_i need to be 'incoherent' in some sense. Formally, the assumptions are as follows [1].

A1.There exists $\mu_1 > 0$ such that for all pairs (a, a') and (b, b')

$$\left| \langle P_U \mathbf{e}_a, P_U \mathbf{e}_{a'} \rangle - \frac{r}{n} 1_{a=a'} \right| \leq \mu_1 \frac{\sqrt{r}}{n} \quad (10)$$

$$\left| \langle P_V \mathbf{e}_b, P_V \mathbf{e}_{b'} \rangle - \frac{r}{n} 1_{b=b'} \right| \leq \mu_1 \frac{\sqrt{r}}{n} \quad (11)$$

where $1 \leq a, b, a', b' \leq n$. Here $\mathbf{e}_a$ denotes the vector with the ath element equal to 1 and others equal to zero. $|\langle A, B \rangle|$ is defined as $\text{trace}(A^H B)$. $1_{a=a'}$ is the indicator function which is equal to 1 when $a = a'$ and 0 otherwise.

A2.Also there exists $\mu_2 > 0$ such that for all (a, b)

$$|E_{ab}| \leq \mu_2 \frac{\sqrt{r}}{n}, \quad (12)$$

where $1 \leq a, b \leq n$. E_{ab} is the (a, b) entry of the matrix defined by (9). If the above assumption holds, we say that the matrix M obeys the strong incoherence property with parameter $\mu = \max\{\mu_1, \mu_2\}$. According to the matrix completion theory, for a fixed rank-r matrix M with strong incoherence parameter μ, if we observe m entries of M with their positions taken uniformly at random from the matrix, M is the unique solution to (6) if there is a numerical constant c such that

$$m \geq c\mu^2 nr \log^6 n. \quad (13)$$

4. MATRIX COMPLETION FOR ARRAY SIGNAL MODEL

To determine if matrix completion theory can be applied to array signal model, we need to check if the matrix formed by the received signals, i.e., W in (2), meet the matrix recovery conditions in equations (10), (11) and (12). We mainly follow the ideas in [1]. To proceed, we will use the following Takagi's factorization theorem [9].

Takagi's factorization: If $A \in \mathbb{C}^{n \times n}$ is symmetric ($A = A^T$), then there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ and a real nonnegative diagonal matrix $\Sigma = \text{diag}(\sigma_1, \cdots, \sigma_n)$ such that $A = U \Sigma U^T$. The columns of U are an orthonormal set of eigenvectors for AA^H and the corresponding diagonal entries of Σ are the nonnegative square roots of the corresponding eigenvalues of AA^H.

Now we will examine equation (2). To simplify the notation, let

$$A = \sum_{k=1}^{r} \eta_k \sqrt{\frac{t}{n}} \mathbf{a}(\alpha_k)\mathbf{a}^T(\alpha_k). \quad (14)$$

Obviously A is a symmetric matrix. By Takagi's factorization theorem, we know that A can be decomposed into $U \Sigma U^T$. Obviously, it is the singular value decomposition of A. Since Z is usually an orthogonal matrix, we thus obtain the singular value decomposition of Y as $Y = U_Y \Sigma_Y V_Y^H$. Firstly, it is easy to verify that

$$\|\mathbf{u}_k\|_{\ell_\infty}, \|\mathbf{v}_k\|_{\ell_\infty} \leq \sqrt{c/n} < \infty \quad (15)$$

for a constant c. Then by the Cauchy-Schwarz inequality

$$\left| \langle P_U \mathbf{e}_a, P_U \mathbf{e}_{a'} \rangle - 1_{a=a'} \frac{r}{n} \right| \leq \max_{1 \leq a \leq n} \|P_U \mathbf{e}_a\|^2 \quad (16)$$

for $a \neq a'$ and by the Frobenius norm bound

$$\max_{1\leq a\leq n} \|P_U \mathbf{e}_a\|^2 \geq \frac{1}{n}\|P_U\|_F^2 = \frac{r}{n} \tag{17}$$

for $a = a'$. Also, the assumption (15) leads to

$$\max_{1\leq a\leq n} \|P_U \mathbf{e}_a\|^2 \leq \frac{cr}{n}. \tag{18}$$

So there exists μ_1 such that $\mu_1 \leq c\sqrt{r}$. Next, we check with (12). Again by the Cauchy-Schwarz inequality, we have

$$|E_{ab}| \leq \max_{1\leq a\leq n} \|P_U \mathbf{e}_a\| \max_{1\leq b\leq n} \|P_V \mathbf{e}_b\| \leq \frac{cr}{n}. \tag{19}$$

Thus there exists μ_2 such that $\mu_2 \leq c\sqrt{r}$. Therefore we obtain, $\mu \leq c\sqrt{r}$. Therefore, we proves the received signal in array signal processing as a matrix obeys the strong incoherence property.

5. UNIFORM SPATIAL SAMPLING

In [4], the authors analyze two models for the sample set Ω. One is Bernoulli model, and the other is uniform model. Under the Bernoulli model, each entry in the matrix is observed with a probability $p = \frac{m}{n^2}$, while under the uniform model Ω is sampled uniformly at random among all subsets of the matrix with the cardinality m. The two models were shown to have the equivalent performance. However neither of them is suitable to be used in our structure. If we fix the number of front-end modules and still apply the uniform model or the Bernoulli model, some samples will be discarded when the number of samples is greater than the number of front-end modules and some front-end modules will be left unused when the number of samples is less than the number of front-end modules. In either case, the performance will not be optimal. We call such type of model truncated uniform model.

The desired sampling model should have an equal number of samples (e.g., the number of front-end elements in the system) in each time slot so that every front-end module will work in all the time slots, which allows us to take full advantage of the hardware resources for better performance and simplicity of the algorithm implementation. To achieve the objective, we propose the *uniform spatial sampling model*. Under this model, we take an equal number of samples in spatial domain in each time slot. In other words, we uniformly take $\frac{m}{n}$ samples from each column of the matrix. If $\frac{m}{n}$ is not an integer, we will round it to the smallest integer above it. Thus, the number of front-end units is set to $\frac{m}{n}$, rather than n. With USS sampling, all front-end hardware modules are always in use.

The key difference between the USS model and the other two models is that under the USS model, every column is guaranteed to be sampled at least one entry, but the chance that every row is sampled is lower than the other two. It is clear that if we fail to observe at least one entry in a row (or a column) of the matrix, we have no way of recovering the matrix. We now investigate the difference quantitatively. We will show that the probability of missing an entire row under the USS model is asympototically equal to that under the Bernoulli model.

Let F be the event that we miss an entire row. Under the Bernoulli model, the probability $\mathrm{P}_{\mathrm{Ber}}(F)$ is $(1-p)^n$ as each sample is taken independently. Because $p = \frac{m}{n^2}$, we have

$$\mathrm{P}_{\mathrm{Ber}}(F) = \left(\frac{n^2-m}{n^2}\right)^n. \tag{20}$$

Under the uniform spatial sampling model, the probability is

$$\mathrm{P}_{\mathrm{USS}}(F) = \left(\frac{\binom{n-1}{m/n}}{\binom{n}{m/n}}\right)^n. \tag{21}$$

Thus, we just need to compare $\frac{n^2-m}{n^2}$ and $\frac{\binom{n-1}{m/n}}{\binom{n}{m/n}}$. By Sterling approximation, we obtain

$$\frac{\binom{n-1}{m/n}}{\binom{n}{m/n}} \approx \frac{\sqrt{\frac{n-1}{\frac{m}{n}2\pi\left(n-1-\frac{m}{n}\right)}}\frac{(n-1)^{n-1}}{\left(\frac{m}{n}\right)^{\frac{m}{n}}\left(n-1-\frac{m}{n}\right)^{n-1-\frac{m}{n}}}}{\sqrt{\frac{n}{\frac{m}{n}2\pi\left(n-\frac{m}{n}\right)}}\frac{n^n}{\left(\frac{m}{n}\right)^{\frac{m}{n}}\left(n-\frac{m}{n}\right)^{n-\frac{m}{n}}}} \tag{22}$$

$$= \sqrt{\frac{(n-1)(n-\frac{m}{n})}{n(n-1-\frac{m}{n})}} \cdot \frac{(n-1)^{n-1}(n-\frac{m}{n})^{n-\frac{m}{n}}}{n^n(n-1-\frac{m}{n})^{n-1-\frac{m}{n}}}. \tag{23}$$

Because m is at least $n\log n$, the two items that we want to compare become $\frac{n^2-m}{n^2} = \frac{n-\log n}{n}$ and

$$\frac{\binom{n-1}{m/n}}{\binom{n}{m/n}} \approx \sqrt{\frac{(n-1)(n-\log n)}{n(n-1-\log n)}} \tag{24}$$

$$\cdot \frac{(n-1)^{n-1}(n-\log n)^{n-\log n}}{n^n(n-1-\log n)^{n-1-\log n}}. \tag{25}$$

We first show that (25) is smaller than $\frac{n-\log n}{n}$. Since $n > n - \log n$, by the fact that $f(x) = \left(\frac{x+1}{x}\right)^x$ is an increasing function, we have

$$\left(\frac{n}{n-1}\right)^{n-1} \geq \left(\frac{n-\log n}{n-\log n-1}\right)^{n-\log n-1}. \tag{26}$$

Then move the left term to the right

$$1 \geq \left(\frac{n-1}{n}\right)^{n-1}\left(\frac{n-\log n}{n-\log n-1}\right)^{n-\log n-1}. \tag{27}$$

By further algebraic manipulation, we have

$$\frac{n-\log n}{n} \geq \frac{(n-1)^{n-1}}{n^n}\frac{(n-\log n)^{n-\log n}}{(n-\log n-1)^{n-\log n-1}} \tag{28}$$

as expected. Next we will show that for any infinitely small ε, we can find an n' such that (24) raised to the n-th power is at most greater than 1 by ε.

$$\left(\sqrt{\frac{(n-1)(n-\log n)}{n(n-1-\log n)}}\right)^n \tag{29}$$

$$= \left(1+\frac{\log n}{n^2-n-n\log n}\right)^{n/2} \tag{30}$$

$$< \left(1+\frac{n^{1/3}}{n^{5/3}}\right)^{n/2} \tag{31}$$

$$= \left(1+\frac{1}{n^{4/3}}\right)^{n/2}. \tag{32}$$

Taking logarithm

$$\frac{n}{2}\log\left(1+\frac{1}{n^{4/3}}\right) < \frac{n}{2}\frac{1}{n^{4/3}} < \frac{n^{-1/3}}{2}. \tag{33}$$

Therefore, to make $\frac{n^{-1/3}}{2} < \log(1+\varepsilon)$ for any infinitely small ε, we set

$$n' = \left(\frac{1}{2\log(1+\varepsilon)}\right)^3 \tag{34}$$

such that for all $n > n'$, we have

$$\frac{n^{-1/3}}{2} < \log(1+\varepsilon). \tag{35}$$

Finally combining (28) and (35), we obtain $\mathrm{P}_{USS}(F) < (1+\varepsilon)\mathrm{P}_{Ber}(F)$.

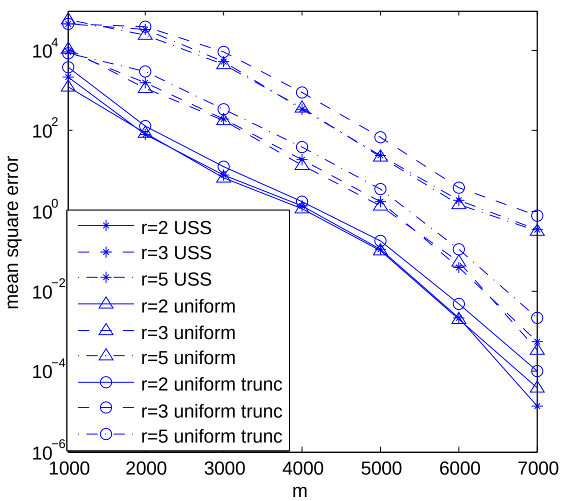

Fig. 2. performance of matrix recovery

6. SIMULATION

We provide simulation results in this section. We use the Augmented Lagrange Multiplier Method [10] to solve the problem. We set $n = 100$, so the total number of matrix entries is 10000. The signal is corrupted by noise with SNR at $10dB$. We keep the noise power at unity and change the signal strength at different SNR levels. m varies from 1000 to 7000 and $r = 2, 3, 5$. To evaluate the performance of the proposed USS model in matrix completion, we define the mean square error(MSE) of the recovered matrix as $\frac{1}{n^2}\sum_{i,j=1}^{n}(M_{ij} - X_{ij})^2$ where M_{ij} and X_{ij} are true matrix element and recovered element respectively. We compare the performance under the USS model, the standard uniform model and the truncated uniform model in Fig 2. We can see that the performance of the USS model and the standard uniform models are almost the same for every combination of m and r. As expected, a smaller rank r would require a lower number of m to achieve the same recovery accuracy. This demonstrates that our uniform spatial sampling method can be applied to taking signal samples of the matrix used in array signal processing, which would allow the matrix to be recovered. On the other hand, if we still apply the standard uniform model, as shown in fig 2, the performance of the *truncated* uniform model is about $0.3dB$ worse than the other two. This justifies the use of our proposed model.

7. CONCLUSION

In this paper, we propose to apply matrix completion theory in array signal processing to reduce the cost. Specifically, under the assumption that the number of targets is much smaller than the number of antennas, the received superposed received signals plus noise would form a low-rank matrix. By using matrix completion theory, we are able to recover the matrix from only a subset of the entries. Therefore, lower sampling rate and simplified hardware can be used to get the signals. We prove that the array signal model satisfy the condition for matrix recovery. In addition, we propose the uniform spatial sampling model which can easily be implemented in hardware and analyze its performance. Simulation shows that much smaller number of the entries would be enough to recover a 100×100 matrix. In addition, uniform spatial sampling achieves the same performance as the standard uniformly random model. Therefore, the proposed method can be used in array receiver to simplify the implementation and reduce the system cost and energy.

Acknowledgements

The authors would like to thank Dr. Rabinder Madan of the Office of Naval Research for sharing his ideas and motivating the work of applying sparsity theory to radar detection field for reducing the cost and increasing the quality.

8. REFERENCES

[1] E.J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *Information Theory, IEEE Transactions on*, vol. 56, no. 5, pp. 2053–2080, 2010.

[2] E.J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[3] R.H. Keshavan, A. Montanari, and S. Oh, "Matrix completion from a few entries," *Information Theory, IEEE Transactions on*, vol. 56, no. 6, pp. 2980–2998, 2010.

[4] E.J. Candes and Y. Plan, "Matrix completion with noise," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 925–936, 2010.

[5] J.F. Cai, E.J. Candes, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *Arxiv preprint arXiv:0810.3286*, 2008.

[6] B. Recht, "A simpler approach to matrix completion," *Arxiv preprint arXiv:0910.0651*, 2009.

[7] E.J. Candes, Y. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *Arxiv preprint arXiv:1109.0573*, 2011.

[8] E. Fishler, A. Haimovich, R.S. Blum, L.J. Cimini, D. Chizhik, and R.A. Valenzuela, "Spatial Diversity in Radars - Models and Detection Performance," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 823–838, Mar. 2006.

[9] R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge Univ Pr, 1990.

[10] Z. Lin, M. Chen, L. Wu, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," *Arxiv preprint arXiv:1009.5055*, 2010.