The RIP for Random Matrices with Complex Gaussian Entries

Kuo Xu, Jian Wang, and Byonghyo Shim

School of Electrical Engineering, Korea University
Anam-dong, Seongbuk-gu, Seoul, Korea, 136-713
xukuo2006@korea.ac.kr

Abstract. In this paper, we show that complex Gaussian random matrix satisfies the restricted isometric property (RIP) with overwhelming probability. We also show that for compressive sensing (CS) applications, complex Gaussian random matrix outperforms its real number equivalent in the sense that it requires fewer measurements for exact recovery of sparse signals. Numerical results confirm our analysis.

Keywords: Sparse recovery, restricted isometric property, complex Gaussian entries.

1 Introduction

In recent years, compressive sensing (CS) has attracted much attention in the academic world [1]. In CS, a high dimensional vector $x \in R^n$, which is k-sparse (i.e., $\|x\|_0 \le k < n$), is sensed by a fat random matrix $A \in R^{m\times n}$ ($m < n$), yielding a low dimensional measurement vector

$$y = Ax\,. \tag{1}$$

Although (1) is an underdetermined system and has infinitely many solutions, the CS theory promises to achieve the perfect recovery of x by exploiting the sparsity. In analyzing the recovery performance, the restricted isometry property (RIP) for the sensing matrix A has been widely used. It has been a standard tool for studying how efficiently the measurement matrix A captures information about sparse signal. Letting A_T denote a submatrix of A with columns listed in set T, the matrix A is said to satisfy the k-RIP if there exists $\delta_k \in (0, 1)$ such that [2]

$$1-\delta_k \le \lambda\left(A_T' A_T\right) \le 1+\delta_k \tag{2}$$

for any T with cardinality $|T| \le k$. In particular, δ_k is called isometry constant. It has been shown that many types of random matrices have excellent restricted isometry behavior. For example, a matrix $A \in R^{m\times n}$, which has i.i.d. entries with Gaussian distribution $N(0,\ 1/m)$, obeys the k-RIP with $\delta_k < \varepsilon$ with overwhelming probability if [3]

James J. (Jong Hyuk) Park et al. (eds.), *Future Information Technology*,
Lecture Notes in Electrical Engineering 276,
DOI: 10.1007/978-3-642-40861-8_3, © Springer-Verlag Berlin Heidelberg 2014

$$m = o\left(k \log \frac{n}{k} \Big/ \varepsilon^2 \right) \tag{3}$$

In many CS applications, a more general setting is that the sensing matrix $A \in C^{m\times n}$ is a random matrix with complex Gaussian entries. There has been empirical evidence that CS works well under this setting, see, for instance, Shim [4]. However, difficulties remain in analyzing the theoretical recovery performance since little has been known about the RIP for random matrix with complex Gaussian entries.

In this work, we study the RIP for random complex Gaussian matrix. For simple description, we define GRM(m, n, σ^2) as a class of $m \times n$ random matrices, for which the real part and imaginary part of entries form a set of $2mn$ i.i.d. random variables with distribution $N(0, \sigma^2/2)$. We argue that RIP holds for complex Gaussian Matrices.

Theorem 1. For any random matrix $A \in$ GRM($m, n, 1/m$), we show that the matrix A satisfies the k-RIP with isometry constant

$$\delta_k < \alpha + 2\sqrt{\alpha} + \sqrt{\frac{4n}{m} H(\gamma)} \tag{4}$$

with overwhelming probability, where $\alpha = k/m$, $\gamma = k/n$ and H is the entropy function $H(\gamma) = -\gamma \log \gamma - (1-\gamma) \log (1 - \gamma)$.

2 Proof of Theorem 1

This section is devoted to the proof of Theorem 1. We stress that the technique we used in the proof is similar to that in [2]. We first consider the eigenvalue of $A_T^* A_T$ where A_T is a submatrix of A with k columns, and then extend the result to all such submatrices.

Lemma 1. [Lemma 4 in [5]] For any random matrix $B \in$ GRM($m, n,$ 1), it satisfies

$$E\left[Tr\left(\exp\left(tB^* B\right)\right)\right] \le n \exp\left(\left(\sqrt{m} + \sqrt{n}\right)^2 t + (m+n)t^2 \right) \ \forall \mathrm{t} \in \left[0, \frac{1}{2}\right], \tag{5}$$

And

$$E\left[Tr\left(\exp\left(-tB^* B\right)\right)\right] \le n \exp\left(-\left(\sqrt{m} - \sqrt{n}\right)^2 t + (m+n)t^2 \right) \ \forall \mathrm{t} \in [0, \infty], \tag{6}$$

where $E[\cdot]$ represents expectation and $Tr[\cdot]$ is the trace.

Lemma 2. For any complex matrix B, all the eigenvalues of $\exp(tB^*B)$ satisfy

$$\lambda_i\left(\exp\left(tB^* B\right)\right) = \exp\left(t\lambda_i\left(B^* B\right)\right), \tag{7}$$

and are always positive.

Proof. Since B^*B is Hermitian and has full rank, its eigenvalues are all real numbers. According to the definition of exponential,

$$\exp\left(tB^*B\right)= I + tB^*B + \frac{1}{2!}\left(tB^*B\right)^2 + \cdots + \frac{1}{n!}\left(tB^*B\right)^n, \tag{8}$$

as $n \to \infty$. For a matrix Q consisted of the eigenvectors of B^*B, $\exp(tB^*B)$ can be diagonalized by Q.

$$Q^{-1}\exp\left(tB^*B\right)Q = I + t\Lambda + \frac{1}{2!}(t\Lambda)^2 + \cdots + \frac{1}{n!}(t\Lambda)^n. \tag{9}$$

By the definition of the exponential function, the i-th eigenvalue on the diagonal is

$$\lambda_i\left(\exp\left(tB^*B\right)\right) = \exp\left(t\lambda_i\left(B^*B\right)\right) > 0. \tag{10}$$

Lemma 3. For a matrix $A \in \mathrm{GRM}(m, n, 1/m)$, singular value concentration inequalities of A_T satisfy

$$P\left(\lambda_{\max}\left(A_T^*A_T\right) \geq \left(\sqrt{\alpha}+1\right)^2 + \varepsilon\right) \leq k\exp\left(-\frac{m\varepsilon^2}{4(\alpha+1)}\right), \tag{11}$$

$$P\left(\lambda_{\min}\left(A_T^*A_T\right) \leq \left(\sqrt{\alpha}-1\right)^2 - \varepsilon\right) \leq k\exp\left(-\frac{m\varepsilon^2}{4(\alpha+1)}\right), \tag{12}$$

where A_T is a submatrix consisted of k columns randomly selected from A.

Proof. For each A_T out of A, A_T times sqrt m belongs to $\mathrm{GRM}(m, n, 1)$. Let $\tau = mt$. Then from Lemma 1,

$$E\left[Tr\left(\exp\left(\tau A_T^*A_T\right)\right)\right] \leq k\exp\left(\left(\sqrt{\alpha}+1\right)^2\tau + m^{-1}(\alpha+1)\tau^2\right), \tag{13}$$

for $\tau \in [0, m/2]$. Since all eigenvalues of $\exp(\tau A_T^*A_T)$ are positive (from Lemma 2),

$$Tr\left(\exp\left(\tau A_T^*A_T\right)\right) \geq \lambda_{\max}\left(\exp\left(\tau A_T^*A_T\right)\right) = \exp\left(\tau\lambda_{\max}\left(A_T^*A_T\right)\right). \tag{14}$$

For $t \in [0, m/2]$ and a small deviation ε, we get

$$\begin{aligned} &P\left(\lambda_{\max}\left(A_T^*A_T\right) \geq \left(\sqrt{\alpha}+1\right)^2 + \varepsilon\right) \\ &= P\left(\exp\left(t\lambda_{\max}\left(A_T^*A_T\right) - t\left(\sqrt{\alpha}+1\right)^2 - t\varepsilon\right) \geq 1\right) \end{aligned} \tag{15}$$

$$\leq E\left[\exp\left(t\lambda_{\max}\left(A_T^* A_T\right)-t\left(\sqrt{\alpha}+1\right)^2-t\varepsilon\right)\right] \tag{16}$$

$$\leq \exp\left(-t\left(\sqrt{\alpha}+1\right)^2-t\varepsilon\right)E\left[Tr\left(\exp\left(tA_T^* A_T\right)\right)\right] \tag{17}$$

$$\leq k\exp\left(-t\varepsilon+m^{-1}(\alpha+1)t^2\right), \tag{18}$$

where (16) uses the Markov's inequality. For a quadratic function, it is obvious that $f(t) = -t\varepsilon + m^{1}(\alpha + 1)t^2$ attains the minimum at $t_0 = m\varepsilon / 2(\alpha + 1)$.

In a similar way, the lower bound in (12) can be proved.

Lemma 3 demonstrates the lower and upper bounds of $\lambda(A_T^*A_T)$ for some A_T. Note that the isometry constant $\delta_k \in (0, 1)$ is defined as the minimum constant such that for all $T \in \{1, \cdots, n\}$ and $|T| = k$,

$$1-\delta_k \leq \lambda\left(A_T^* A_T\right) \leq 1+\delta_k\,. \tag{19}$$

For notational simplicity, denote $\lambda_{\max} = \lambda_{\max}(A_T^*A_T)$ and observe that

$$P\left(\forall_{A_T}, \lambda_{\max} \leq \left(\sqrt{\alpha}+1\right)^2+\varepsilon\right) \geq 1-k\binom{n}{k}\exp\left(-\frac{m\varepsilon^2}{4(\alpha+1)}\right). \tag{20}$$

From Stirling's approximation, we know the combination number k out of n approximates to $\exp(nH(\gamma))$. Then it follows that

$$P\left(\forall_{A_T}, \lambda_{\max} \leq \left(\sqrt{\alpha}+1\right)^2+\varepsilon\right) \geq 1-k\exp\left(-\frac{m}{4\alpha}\left(\varepsilon^2-\frac{4nH(\gamma)}{m}\right)\right). \tag{21}$$

As thus, for m goes to infinite,

$$\lambda_{\max} < (\alpha+1)^2+\sqrt{4nH(\gamma)/m}\,. \tag{22}$$

Similar results hold for $\lambda_{\min}$. The proof of Theorem 1 is established.

3 Simulation and Discusssion

From Theorem 1, one can show that upper bound in (4) is more stringent than that in the real number situation [2]. Indeed, let A^r denote a $m \times n$ random matrix with real number entries satisfying $N(0, 1/m)$. Then for $n \to \infty$,

$$\delta_k(A) \to 2\sqrt{\alpha\log(n)} \text{ and } \delta_k\left(A^r\right) \to 2\sqrt{2\alpha\log(n)}. \tag{23}$$

The isometry constant for A is 0.7 times as much as that for A^r and hence is more stringent. To see the difference between $\delta_k(A)$ and $\delta_k(A^r)$, we perform simulations to provide an empirical comparison. To be specific, we generate a number of real and complex Gaussian random matrices. For each matrix, we calculate the isometry constant by an exhaustive search. The distributions of $\delta_k(A)$ and $\delta_k(A^r)$ are displayed in Fig. 1. One can easily observe that $\delta_k(A)$ is uniformly smaller than $\delta_k(A^r)$.

The reason why the complex Gaussian random matrix has more stringent isometry constant than the real Gaussian random matrix is perhaps that the extreme singular values of any submatrix formed by k (or fewer) columns from A has stronger concentration property. Indeed, the probability of violation for real Gaussian random matrix decreases at a speed of $\exp(-m\varepsilon^2/8)$ as ε increases [6], whereas, as shown in Lemma 3, the probability of violation for complex case decreases at $\exp(-m\varepsilon^2/4)$. In other words, the distribution of the extreme singular value of complex Gaussian matrix has a smaller tail, and therefore, the complex Gaussian matrix has a smaller δ_k (for the same γ and α), when compared to the real case. For compressive sensing applications, this result implies that fewer measurements are required [7].

To illustrates the advantage of complex Gaussian sensing matrix over the real case in compressive sensing, we perform simulations on sparse signal recovery with complex and real Gaussian random matrices. In our simulation, we employ orthogonal matching pursuit (OMP) algorithm as the recovery algorithm to recover k-sparse signals with complex entries. Two kinds of recovery are performed. First, we directly employ OMP to recover the complex signal x in the complex number signal model (1). Second, we reformulate model (1) to a real number signal model [8]:

$$y' = \begin{bmatrix} \mathrm{Re}(y) \\ \mathrm{Im}(y) \end{bmatrix}, x' = \begin{bmatrix} \mathrm{Re}(x) \\ \mathrm{Im}(x) \end{bmatrix}, A' = \begin{bmatrix} \mathrm{Re}(A) & -\mathrm{Im}(A) \\ \mathrm{Im}(A) & \mathrm{Re}(A) \end{bmatrix}, \tag{24}$$

and then employ OMP to recover x'. Note that the recovery of the second case is performed in the real domain. We compare the minimally required measurements y guaranteeing exact recovery of sparse signals.

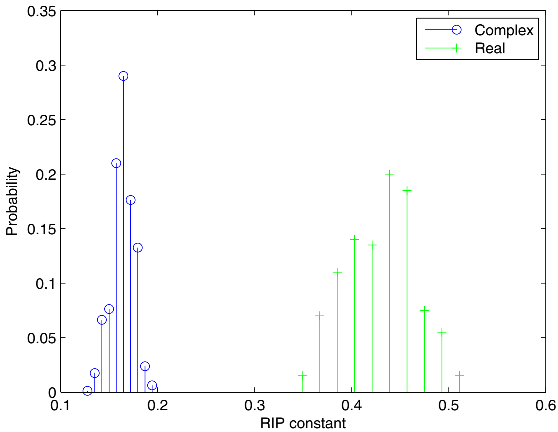

Fig. 1. Distribution of $\delta_k(A)$ and $\delta_k(A^r)$, with $k = 4$, $m = 20$, and $n = 128$

For a fixed sparsity ratio γ, $k = 10$, and $n = 256$, the exact recovery ratio by OMP algorithm is simulated with different measurement number m. Note that α and γ remain the same after reformulation. We calculate the exact recovery ratio for x and $x^{'}$. The result is shown in Fig. 2. where cOMP represents the recovery result using the first method (i.e., direct recovery in the complex domain). It is easily observed that the first method outperforms the second method, as it uniformly requires fewer measurements for exact reconstruction.

It is interesting to note that the superior numerical performance of A over $A^{'}$ can also be explained as follows. For an $n \times 1$ k-sparse complex signal, its real number equivalent is an $2n \times 1$ $2k$-sparse signal. In the recovery process, one sparse signal is selected with a candidates number of k out of n. Whereas, in the real equivalent case, candidates number $2k$ out of $2n$. By Stirling's approximation, we know

$$\binom{n}{k} \to e^{nH\left(\frac{k}{n}\right)} \text{ and } \binom{2n}{2k} \to e^{2nH\left(\frac{k}{n}\right)}. \tag{25}$$

Thus it is easier to solve the complex sparse signal recovery problem than the reformulated real sparse signal recovery problem.

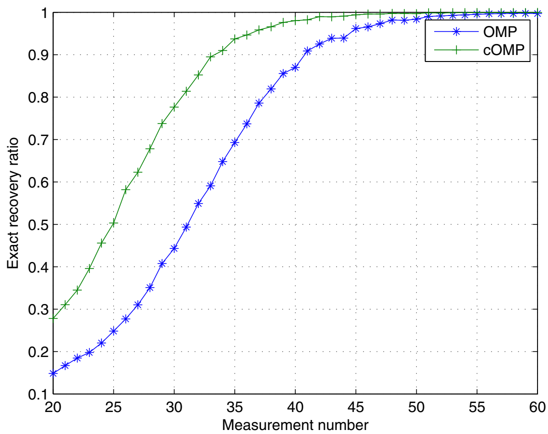

Fig. 2. Exact recovery of sparse signals via OMP for complex Gaussian measurement matrix and its real number equivalent

4 Conclusion

This paper presented the RIP for Gaussian random matrix with complex entries. The result demonstrated that compared to the isometry constant for real Gaussian random matrices, the isometry constant for complex Gaussian random matrices are more stringent. This implies that for CS applications, the required number of measurements guaranteeing exact recovery can be fewer when the complex Gaussian random measurements are used.

References

1. Donoho, D.L.: Compressed sensing. IEEE Trans. Inform. Theory 52, 1289–1306 (2006)
2. Candes, E.J., et al.: Decoding by linear programming. IEEE Trans. Inform. Theory 51, 4203–4215 (2005)
3. Baraniuk, R., et al.: A simple proof of the restricted isometry property for random matrices. Constructive Approximation 28, 253–263 (2008)
4. Shim, B., et al.: Multiuser Detection via Compressive Sensing. IEEE Communications Letters 16, 972–974 (2012)
5. Haagerup, U., et al.: Random matrices with complex Gaussian entries. Expositiones Mathematicae 21, 293–337 (2003)
6. Ledoux, M.: The concentration of measure phenomenon. American Mathematical Soc. (2001)
7. Wang, J., et al.: On Recovery Limit of Orthogonal Matching Pursuit using Restricted Isometric Property. IEEE Trans. Signal Process 60, 4973–4976 (2012)
8. Shim, B., et al.: Sphere Decoding With a Probabilistic Tree Pruning. IEEE Trans. Signal Process. 56, 4867–4878 (2008)