

Inequalities for Non-Standard Gaussian Matrix Norms

Arvind Rathnashyam
RPI CS and Math
rathna@rpi.edu

September 2, 2023

Abstract

In this paper we explore upper bounds on the spectral norm for Gaussian Matrices with columns standard from Central Correlated Multivariate Normal Distributions. We utilize a lemma from [Chi17] and extend the analysis from [CD05]. These bounds find applications in the generalization of the randomized SVD given in [BT22] and wireless network science.

1 Introduction

The study of the expectation of the norms of the pseudoinverse of standard normal gaussian matrices first appeared in [HMT11] when analyzing the error bounds for the Randomized SVD algorithm. The bounds developed in [HMT11] used theory developed in analyzing the condition numbers of standard normal matrices in [CD05]. In a generalization of the Randomized SVD, the need for bounds on the expectation of the spectral norm for correlated Gaussian matrices appears in [BT22].

2 Relevant Work in Standard Uncorrelated Matrices

In this section we will briefly discuss bounds developed for the inequalities of standard normal matrices.

Proposition 1. (HMT Proposition 10.2). Draw a $k \times (k + p)$ standard Gaussian matrix \mathbf{G} with $k \geq 2$ and $p \geq 2$. Then

$$\mathbb{E} \|\mathbf{G}^\dagger\| \leq \frac{e\sqrt{k+p}}{p} \quad (1)$$

From our search in the literature, there is no bound on equation 1 when the columns are not sampled from a multiple of the identity.

3 Theory

We will first introduce the necessary lemmas needed to prove our main results.

3.1 Necessary Lemmas

Lemma 2. [Jam64, Eq. (58,59)]. If $\lambda_1 \geq \dots \geq \lambda_m$ are the eigenvalues of \mathbf{W} s.t. $\mathbf{W} \sim \mathcal{W}_m(n, \mathbf{C})$ s.t. $n > m - 1$, then the joint PDF of eigenvalues is

$$f(\lambda_1, \dots, \lambda_m) = K_{m,n} (\det \mathbf{C})^{-n/2} \exp\left(-\frac{1}{2} \text{Tr}(\mathbf{C}^{-1}\mathbf{W})\right) \prod_{i=1}^m \lambda_i^{(n-m-1)/2} \prod_{i < j} (\lambda_i - \lambda_j) \quad (2)$$

where

$$K_{m,n} = \frac{\pi^{m^2/2}}{\Gamma_m(\frac{1}{2}m) \Gamma_m(\frac{1}{2}n)} \quad (3)$$

Lemma 3. [WLRT08, Lemma 3.6]. Let $m, n \in \mathbb{N}$ s.t. $n \geq m$. Suppose $\mathbf{A} \in \mathbb{R}^{n \times m}$, then if $(\mathbf{A}^\top \mathbf{A})$ is invertible

$$\left\| (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \right\| = \frac{1}{\sigma_m(\mathbf{A})} \quad (4)$$

Lemma 4. [Chi17, Lemma 1]. Draw a $m \times n$ matrix \mathbf{G} s.t. the columns of \mathbf{G} are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{C})$ where the eigenvalues of \mathbf{C} are represented as $\sigma_1 > \sigma_2 > \dots > \sigma_m$. Let $\mathbf{W} \sim \mathcal{W}_m(n, \mathbf{C})$. The eigenvalue distribution is given as

$$f(x_1, \dots, x_n) = K_{\mathbf{C}} |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i=1}^{m-1} \prod_{j=i+1}^m (x_i - x_j) \prod_{i=1}^n x_i^{n-m} \quad (5)$$

where $\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma}) = \{e^{-x_i/\sigma_j}\}_{i,j=1}^m = \begin{bmatrix} e^{-\frac{x_1}{\sigma_1}} & \dots & e^{-\frac{x_1}{\sigma_m}} \\ \vdots & \ddots & \vdots \\ e^{-\frac{x_m}{\sigma_1}} & \dots & e^{-\frac{x_m}{\sigma_m}} \end{bmatrix}$ and

$$K_{\mathbf{C}}^{-1} = \prod_{i=1}^{m-1} \prod_{j=i+1}^m (\sigma_i - \sigma_j) \prod_{i=1}^m \sigma_i^{n-m+1} (n-i)! \quad (6)$$

With these lemmas we will go to proving the main results.

3.2 Main Results

Theorem 5. Draw a $m \times m$ matrix \mathbf{G} s.t. the columns of \mathbf{G} are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{C})$ where the eigenvalues of \mathbf{C} are represented as $\sigma_1 > \sigma_2 > \dots > \sigma_m$. Then

$$\mathbb{E} \|\mathbf{G}^\dagger\| \approx \sqrt{\pi \sum_{k=1}^m \frac{1}{\sigma_k}} \quad (7)$$

Proof. We will first note

$$\|\mathbf{G}^\dagger\| \stackrel{\text{lem. 3}}{=} \frac{1}{\sigma_m(\mathbf{G})} = \frac{1}{\sqrt{\lambda_{\min}(\mathbf{G}\mathbf{G}^\top)}} \quad (8)$$

For \mathbf{W} sampled from $\mathcal{W}_m(m, \mathbf{C})$. We will now derive the distribution for minimum eigenvalue of \mathbf{W} similar to [NZYY08].

$$f_{\lambda_{\min}}(x_m) = \int_{x_2}^{\infty} \dots \int_{x_{m-1}}^{\infty} K_{\mathbf{C}} |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i < j}^m (x_i - x_j) \prod_{i=1}^m x_j^{m-m} \prod_{i=1}^{m-1} dx_i \quad (9)$$

$$= K_{\mathbf{C}} \int_{x_2}^{\infty} \dots \int_{x_{m-1}}^{\infty} |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i=1}^{m-2} \prod_{j=i+1}^{m-1} (x_i - x_j) \prod_{i=1}^{m-1} (x_i - x_m) \prod_{i=1}^{m-1} dx_i \quad (10)$$

$$\stackrel{(a)}{=} e^{-\sum_{i=1}^m \frac{x_m}{\sigma_i}} \left(\int_{y_2}^{\infty} \dots \int_{y_{m-1}}^{\infty} \sum_{i=1}^m (-1)^{i+m} K_{\mathbf{C}} |\mathbf{E}_i(\mathbf{x} - \mathbf{x}_m, \boldsymbol{\sigma})| \prod_{i=1}^{m-2} \prod_{j=i+1}^{m-1} (y_i - y_j) \prod_{i=1}^{m-1} dy_i \right) \quad (11)$$

$$\stackrel{(b)}{=} \Xi e^{-\sum_{i=1}^m \frac{x_m}{\sigma_i}} \quad (12)$$

(a) follows due to the properties of the determinant. (b) follows as the intergral expresion in Equation (11) no longer integrates over x_m and thus integrates to some constant we define as Ξ . Since the PDF must integrate to 1, we thus have,

$$f_{\lambda_{\min}}(x) = \left(\sum_{k=1}^m \frac{1}{\sigma_k} \right) e^{-x \sum_{k=1}^m \frac{1}{\sigma_k}} \quad (13)$$

The Expected Value follows from a simple integration.

$$\mathbb{E} \|\mathbf{G}^\dagger\| = \int_0^\infty \frac{1}{\sqrt{x}} e^{-x \sum_{k=1}^m \sigma_k^{-1}} dx \quad (14)$$

$$= \sqrt{\pi \sum_{k=1}^m \frac{1}{\sigma_k}} \operatorname{erf} \left(\sqrt{\pi \sum_{k=1}^m \frac{1}{\sigma_k}} \right) \lesssim \sqrt{\pi \sum_{k=1}^m \frac{1}{\sigma_k}} \quad (15)$$

The proof is complete. ■

Theorem 6. Draw a $m \times n$ matrix \mathbf{G} s.t. the columns of \mathbf{G} are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{C})$ where the eigenvalues of \mathbf{C} are represented as $\sigma_1 > \sigma_2 > \dots > \sigma_m$. Let $\mathbf{W} \sim \mathcal{W}_m(n, \mathbf{C})$. Let $\gamma \triangleq \min_{k \in [m-1]} (\sigma_k - \sigma_{k+1})$ and $p \triangleq n - m + 1$, then

$$\mathbb{E} \|\mathbf{G}^\dagger\| \leq \gamma^m 2 \sqrt{\frac{\pi}{e}} \left(\frac{p}{e}\right)^p \sum_{i=1}^m \frac{1}{\sqrt{\sigma_i}} \quad (16)$$

Proof. Let $K_{\mathbf{C}}$ and $\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})$ be defined as in Lemma 4.

$$f_{\lambda_{\min}}(x_m) = \int_{x_2}^\infty \dots \int_{x_{m-1}}^\infty K_{\mathbf{C}} |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i < j}^m (x_i - x_j) \prod_{i=1}^m x_j^{n-m} \prod_{i=1}^{m-1} dx_i \quad (17)$$

$$= K_{\mathbf{C}} x_m^{n-m} \int_{x_2}^\infty \dots \int_{x_{m-1}}^\infty |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i < j}^{m-1} (x_i - x_j) \prod_{i=1}^{m-1} (x_i - x_m) \prod_{i=1}^m x_i^{n-m} \prod_{i=1}^{m-1} dx_i \quad (18)$$

$$\leq K_{\mathbf{C}} x_m^{n-m} \int_{x_2}^\infty \dots \int_{x_{m-1}}^\infty |\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i < j}^{m-1} (x_i - x_j) \prod_{i=1}^{m-1} x_i^{n-m+1} \prod_{i=1}^{m-1} dx_i \quad (19)$$

$$\leq K_{\mathbf{C}} x_m^{n-m} \sum_{i=1}^m \left((-1)^{i+m} e^{-\frac{x_m}{\sigma_i}} \int_{x_2}^\infty \dots \int_{x_{m-1}}^\infty |\mathbf{E}_{m,i}(\mathbf{x}, \boldsymbol{\sigma})| \cdot \prod_{i < j}^{m-1} (x_i - x_j) \prod_{i=1}^{m-1} x_i^{n-m+1} \prod_{i=1}^{m-1} dx_i \right) \quad (20)$$

$$= \underbrace{x_m^{n-m} K_{\mathbf{C}} \sum_{i=1}^m (-1)^{i+m} e^{-\frac{x_m}{\sigma_i}} K_{\mathbf{C},i}^{-1}}_{\Xi} \quad (21)$$

We will now upper bound Ξ .

$$\Xi \triangleq K_{\mathbf{C}} \sum_{i=1}^m (-1)^{i+m} e^{-\frac{x_m}{\sigma_i}} K_{\mathbf{C},i}^{-1} \quad (22)$$

$$= \sum_{k=1}^m (-1)^{k+m} e^{-\frac{x_m}{\sigma_k}} \frac{\prod_{i=1}^{m-1} \prod_{j=i+1}^m \mathbb{1}_{i,j \neq k} (\sigma_i - \sigma_j) \prod_{i=1}^m \mathbb{1}_{i \neq k} \sigma_i^{n-m+1} (n-i)!}{\prod_{i=1}^{m-1} \prod_{j=i+1}^m (\sigma_i - \sigma_j) \prod_{i=1}^m \sigma_i^{n-m+1} (n-i)!} \quad (23)$$

$$= \sum_{k=1}^m (-1)^{k+m} e^{-\frac{x_m}{\sigma_k}} \left(\prod_{i > k}^m (\sigma_i - \sigma_k) \prod_{i < k}^m (\sigma_k - \sigma_i) \cdot \sigma_k^{n-m+1} (n-k)! \right)^{-1} \quad (24)$$

$$\leq \sum_{k=1}^m e^{-\frac{x_m}{\sigma_k}} \underbrace{\left(\prod_{i > k}^m (\sigma_i - \sigma_k) \prod_{i < k}^m (\sigma_k - \sigma_i) \cdot \sigma_k^{n-m+1} (n-k)! \right)^{-1}}_K \quad (25)$$

Now we will lower bound K .

$$K \triangleq \prod_{i>k}^m (\sigma_i - \sigma_k) \prod_{i<k}^m (\sigma_k - \sigma_i) \cdot \sigma_k^{n-m+1} (n-k)! \quad (26)$$

$$\geq \prod_{i>k}^m (\sigma_i - \sigma_k) \prod_{i<k}^m (\sigma_k - \sigma_i) \cdot \sigma_k^{n-m+1} \quad (27)$$

$$\geq \min_{k \in [m-1]} (\sigma_k - \sigma_{k+1})^m \sigma_1^{n-m+1} \quad (28)$$

We thus have

$$f_{\lambda_{\min}}(x_m) \leq K x_m^{n-m} \sum_{i=1}^m e^{-\frac{x_m}{\sigma_i}} \leq \mathcal{O} \left(x_m^{n-m} \sum_{i=1}^m e^{-\frac{x_m}{\sigma_i}} \right) \quad (29)$$

Now we will integrate over $f_{\lambda_{\min}}(x_m)$.

$$\mathbb{E} \|\mathbf{G}^\dagger\| = \int_0^\infty \mathcal{O} \left(x^{n-m-\frac{1}{2}} \sum_{i=1}^m e^{-\frac{x}{\sigma_i}} \right) dx \quad (30)$$

$$= \sum_{i=1}^m \mathcal{O} \left(\sigma_i^{n-m+\frac{1}{2}} \Gamma \left(\frac{n-m+\frac{1}{2}}{2} \right) \right) \quad (31)$$

$$\stackrel{(a)}{\leq} \sum_{i=1}^m \mathcal{O} \left(\sigma_i^{n-m+\frac{1}{2}} \sqrt{\frac{4\pi}{n-m+\frac{1}{2}}} \left(\frac{n-m+\frac{1}{2}}{e} \right)^{n-m+\frac{1}{2}} \right) \quad (32)$$

$$= \sum_{i=1}^m \mathcal{O} \left(\sigma_i^{n-m+\frac{1}{2}} 2\sqrt{\frac{\pi}{e}} \left(\frac{n-m+\frac{1}{2}}{e} \right)^{n-m} \right) \quad (33)$$

(a) follows from an application of Stirling's Approximation [Rob55]. Now we will plug in the lower bound for K from Equation (28).

$$\mathbb{E} \|\mathbf{G}^\dagger\| \leq \frac{\sum_{i=1}^m \sigma_i^{n-m+\frac{1}{2}} 2\sqrt{\frac{\pi}{e}} \left(\frac{n-m+\frac{1}{2}}{e} \right)^{n-m}}{\min_{k \in [m-1]} (\sigma_k - \sigma_{k+1})^m \sigma_1^{n-m+1}} \quad (34)$$

Define $\gamma \triangleq \min_{k \in [m-1]} (\sigma_k - \sigma_{k+1})$ and the proof is complete. \blacksquare

Theorem 7. Draw a $m \times n$ matrix \mathbf{G} s.t. the columns of \mathbf{G} are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{C})$ where the eigenvalues of \mathbf{C} are represented as $\sigma_1 > \sigma_2 > \dots > \sigma_m$. Let $\mathbf{W} \sim \mathcal{W}_m(n, \mathbf{C})$. Then,

$$\mathbb{E} \|\mathbf{G}^\dagger\| \leq \sum_{i=1}^m \mathcal{O} \left(\sigma_i^{n-m+\frac{1}{2}} 2\sqrt{\frac{\pi}{e}} \left(\frac{n-m+\frac{1}{2}}{e} \right)^{n-m} \right) \quad (35)$$

4 Numerical Experiments

In Figure 1, we verify the results given in Theorem 5.

In Figure 2, we verify the results given in Theorem 6.

5 Conclusions

In this paper, we derive novel upper bounds for the spectral norm of Gaussian matrices with columns sampled from a central correlated multivariate normal distribution.

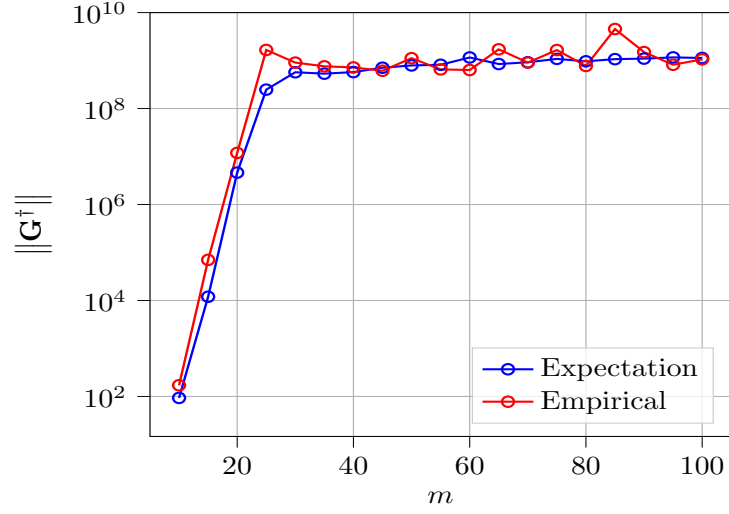


Figure 1: Comparing the expected norm upper bound on $\|\mathbf{G}^\dagger\|$ where $\mathbf{G} \in \mathbb{R}^{m \times m}$ and the columns of \mathbf{G} are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{K})$ with the average norm of \mathbf{G}^\dagger over 100 samples. The expected norm is calculated with Proposition 5.

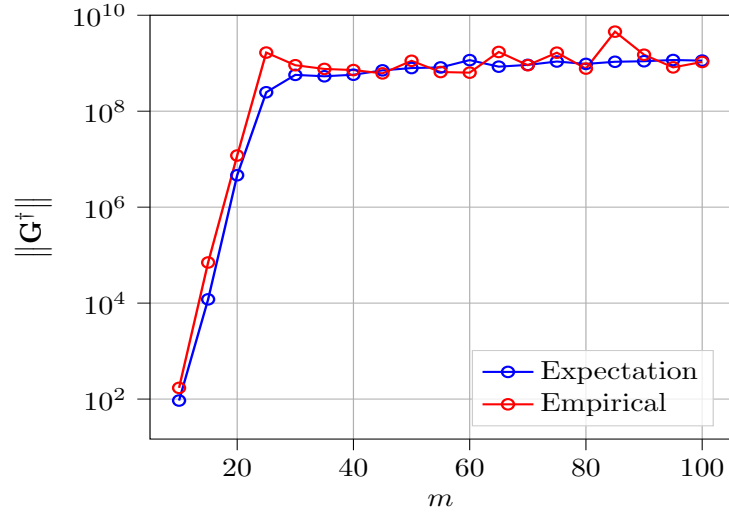


Figure 2: Comparing the expected norm upper bound on $\|\mathbf{G}^\dagger\|$ where $\mathbf{G} \in \mathbb{R}^{m \times m}$ and the columns of \mathbf{G} are sampled from $\mathcal{N}_m(\mathbf{0}, \mathbf{K})$ with the average norm of \mathbf{G}^\dagger over 100 samples. The expected norm is calculated with Proposition 5.

References

- [BT22] Nicolas Boule and Alex Townsend. A generalization of the randomized singular value decomposition. In *International Conference on Learning Representations*, 2022.
- [CD05] Zizhong Chen and Jack J. Dongarra. Condition numbers of gaussian random matrices. *SIAM Journal on Matrix Analysis and Applications*, 27(3):603–620, 2005.
- [Chi17] Marco Chiani. On the probability that all eigenvalues of gaussian, wishart, and double wishart random matrices lie within an interval, 2017.
- [HMT11] N. Halko, P. G. Martinsson, and J. A. Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*, 53(2):217–288, 2011.
- [Jam64] Alan T. James. Distributions of matrix variates and latent roots derived from normal samples. *The Annals of Mathematical Statistics*, 35(2):475–501, 1964.
- [NZYY08] Fangfang Niu, Haochuan Zhang, Hongwen Yang, and Dacheng Yang. Distribution of the smallest eigenvalue of complex central semi-correlated wishart matrices. In *2008 IEEE International Symposium on Information Theory*, pages 1788–1792, 2008.
- [Rob55] Herbert Robbins. A remark on stirling’s formula. *The American mathematical monthly*, 62(1):26–29, 1955.
- [WLRT08] Franco Woolfe, Edo Liberty, Vladimir Rokhlin, and Mark Tygert. A fast randomized algorithm for the approximation of matrices. *Applied and Computational Harmonic Analysis*, 25(3):335–366, 2008.