# On the Distribution of a Scaled Condition Number

*Alan Edelman**
Department of Mathematics
University of California
Berkeley, California 94720

**Abstract**

In this note, we give the exact distribution of a scaled condition number used by Demmel to model the probability that matrix inversion is difficult. Specifically, consider a random matrix $A$ and the scaled condition number $\kappa_D(A) = \|A\|_F \cdot \|A^{-1}\|$. Demmel provided bounds for the condition number distribution when $A$ has real or complex normally distributed elements. Here, we give the exact formula.

---

Key words and phrases: condition number, ill-conditioning, multivariate statistics, numerical analysis, random matrix
*This work was started while the author was a Visiting Scholar in the Department of Mathematics, MIT, and completed at CERFACS with the support of the North Atlantic Treaty Organization under a Grant awarded in 1989.

## 1 Introduction and statement of results

In [4], Demmel investigates the probability that numerical analysis problems are difficult by unifying the common algebraic and geometric structures underlying the notion of ill-conditioning. As an application of his theory, he constructs a probabilistic model to examine the probability that matrix inversion is difficult. It is our goal in this note to work exclusively within his framework and derive exact distributions for the condition numbers that he considers. This condition number is a measure of difficulty in that the larger the value of the condition number, the more "difficult" matrix inversion becomes. The limitations of the model are discussed in [4]. We consider it a rather remarkable accident of mathematics that these distributions can be written down in a closed form at all. Although, as Demmel states, the assumption that matrices are uniformly distributed spherically is rather strong, the mathematics stands on its own, and indeed might have further applications to the "tubular neighborhoods" that he uses and perhaps also in multivariate statistics.

The objects of study are a scaled version $\kappa_D(A) \equiv \|A\|_F \cdot \|A^{-1}\|$ of the usual condition number, and its distribution when considering random real and complex $n$ by $n$ matrices with elements distributed uniformly on the sphere $1 = \|A\|_F^2 = \sum a_{ij}^2$. Because of the scale invariance of the condition number and special properties of the normal distribution, it is equivalent to assume that the random matrices are generated with independent elements from a real or complex standard normal distribution. Demmel concludes that for real matrices

$$\frac{C(1-1/x)^{n^2-1}}{x} \leq \mathrm{Prob}(\kappa_D(A) \geq x) \leq \sum_{k=1}^{n^2} 2\binom{n^2}{k}\left(\frac{2n}{x}\right)^k,$$

where $C$ is a constant. For complex matrices, on the other hand, he concludes that

$$\frac{(1-x^{-1})^{2n^2-2}}{2n^4x^2} \leq \mathrm{Prob}(\kappa_D(A) \geq x) \leq \frac{e^2n^5(1+n^2/x)^{2n^2-2}}{x^2}, \tag{1.1}$$

and that asymptotically

$$\mathrm{Prob}(\kappa_D(A) \geq x) = \frac{n(n^2-1)}{x^2} + o(\frac{1}{x^2}),$$

as $x \to \infty$ for fixed $n$.

In this note, we derive the exact probability distribution by combining exact distribution expressions for the smallest singular values [5, 6] of these random matrices with equations from [3, 8] which relate to $\kappa_D$. Our results are:

| Probability Densities for $\kappa_D(A)$: |
|---|
| Real $n$ by $n$ matrices:<br>$\mu x^{1-n^2}(x^2-n)^{\frac{n(n+1)}{2}-2} {}_2F_1\left(\frac{n}{2}-\frac{1}{2}, \frac{n}{2}+1; \frac{n^2}{2}+\frac{n}{2}-1; -(x^2-n)\right)$<br>$\mu = \frac{2n\Gamma(\frac{n+1}{2})\Gamma(\frac{n^2}{2})}{\sqrt{\pi}\Gamma(\frac{n(n+1)}{2}-1)}$ |
| Complex $n$ by $n$ matrices:<br>$2n(n^2-1)x^{1-2n^2}(x^2-n)^{n^2-2}$ |

Here ${}_2F_1$ denotes the Gauss hypergeometric function. Since obviously $\sqrt{n} \leq \kappa_D(A)$, these formulas are valid only for $x \geq \sqrt{n}$.

For real matrices, the formula is cumbersome. For large $x$ and $n > 20$, say, the formula below is quite adequate:

$$\mathrm{Prob}(\kappa_D(A) \geq x) \approx n^{3/2}/x, \quad x \gg \sqrt{n}, n \gg 1.$$

For complex matrices the exact distribution is a simple expression:

$$\mathrm{Prob}(\kappa_D(A) \geq x) = 1 - (1-n/x^2)^{n^2-1}, \quad x > \sqrt{n}.$$

For large $n$, the condition numbers of real and complex matrices scale like $n^{3/2}$. To be precise let $\kappa'$ be the random variable $2\kappa_D/n^{3/2}$. Then as $n \to \infty$, for real matrices,

$$\text{Prob}(\kappa' < x) \to e^{-2/x-2/x^2}. \tag{1.2}$$

For complex matrices,

$$\text{Prob}(\kappa' < x) \to e^{-4/x^2}. \tag{1.3}$$

In fact, for large $n$ the Demmel condition number $\kappa_D$ of random uniformly distributed matrices is roughly $\sqrt{n}/2$ times as big as the ordinary 2-norm condition number $\kappa_2$. To be precise, as $n \to \infty$, $\frac{2}{\sqrt{n}}\kappa_D/\kappa_2$ converges almost surely to 1. Thus for large $n$, $\kappa_D$ truly deserves to be called a scaled condition number, and the distribution of $\kappa'$ is the same as the distribution of $\kappa_2/n$ which we have presented in [5, 6].

# 2 The distribution of $\kappa_D$ (real case)

Let $A$ be a real random $n \times n$ matrix with independent and identically distributed (**iid**) elements from a standard normal distribution. The matrix $W = AA^T$ is said to be a Wishart matrix or to have the Wishart distribution. Our goal is to study the random quantity

$$\kappa_D(A) = \sqrt{\frac{\sum_{i=1}^n \lambda_i}{\lambda_n}},$$

where $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$ are the eigenvalues of $AA^T$. Clearly $\kappa_D(A) \geq \sqrt{n}$.

Let $f_n$ be the probability density function (pdf) of $(\kappa_D(A))^{-2} = \lambda_n / \sum_{i=1}^n \lambda_i$, and let $g_n$ be the pdf of $\lambda_n$. The distribution function for $\kappa_D(A)$ will be derived from two lemmas regarding $f_n$ and $g_n$.

**Lemma 2.1** *(Davis) The pdfs $f_n$ and $g_n$ are related by*

$$L\left((1+w)^{\frac{1}{2}n^2-2} f_n\left(\frac{1}{1+w}\right)\right)(s) = 2\Gamma(n^2/2)e^s s^{-\frac{1}{2}n^2+1} g_n(2s),$$

*where $L$ denotes the Laplace transform.*

This lemma was proved in [3], where the more general case of $\lambda_j / \sum_{i=1}^n \lambda_i$ is examined. These ratios arise in the multivariate analysis of variance (MANOVA) as described in multivariate analysis books such as [2].

The density function $g_n$ is known exactly (see [5] or [6]):

**Lemma 2.2** *The density of the smallest eigenvalue of a Wishart matrix is*

$$g_n(x) = \frac{n}{\sqrt{2\pi}} \Gamma\left(\frac{n+1}{2}\right) x^{-1/2} e^{-xn/2} U\left(\frac{n-1}{2}, -\frac{1}{2}, x/2\right).$$

*When $a > 0$ and $b < 1$, the Tricomi function, $U(a,b,z)$, is the unique solution to Kummer's equation*

$$z\frac{d^2w}{dz^2} + (b-z)\frac{dw}{dz} - aw = 0, \tag{2.1}$$

*satisfying $U(a,b,0) = \Gamma(1-b)/\Gamma(1+a-b)$ and $U(a,b,\infty) = 0$.*

Combining Lemma 2.1 and Lemma 2.2, we obtain

**Theorem 2.1** *The density of $(\kappa_D(A))^{-2}$ is*

$$f_n(x) = \alpha x^{n^2/2-2}(x^{-1}-n)^{\frac{n(n+1)}{2}-2} {}_2F_1\left(\frac{n}{2}-\frac{1}{2}, \frac{n}{2}+1; \frac{n^2}{2}+\frac{n}{2}-1; -(x^{-1}-n)\right),$$

*where*

$$\alpha = \frac{n\Gamma(\frac{n+1}{2})\Gamma(\frac{n^2}{2})}{\sqrt{\pi}\Gamma(\frac{n(n+1)}{2}-1)}$$

*and ${}_2F_1$ is the Gauss (hypergeometric) function.*

**Proof** According to [7] (formula $7.522.4^6$, p. 850),[1] if $b > 0$,

$$L\left(w^{b-1}{}_2F_1(a, a-c+1; b; -w)\right)(s) = \Gamma(b)s^{a-b}U(a,c,s), \tag{2.2}$$

where ${}_2F_1$ is the Gauss (hypergeometric) function.

With $a = \frac{n}{2} - \frac{1}{2}$, $b = \frac{n^2}{2} + \frac{n}{2} - 1$, $c = -\frac{1}{2}$, and $d = \frac{n}{2} + 1$, we have from Lemmas 2.1 and 2.2 that

$$L\left((1+w)^{\frac{1}{2}n^2-2} f_n\left(\frac{1}{1+w}\right)\right)(s) = \alpha e^{s(1-n)}\Gamma(b)s^{a-b}U(a,c,s). \tag{2.3}$$

From 2.2 we have

$$L\left(\alpha w^{b-1}{}_2F_1(a, d; b; -w)\right)(s) = \alpha\Gamma(b)s^{a-b}U(a,c,s).$$

Using familiar results concerning the Laplace transform, we then obtain

$$L\left(\alpha(w-n+1)^{b-1}{}_2F_1(a, d; b; -(w-n+1))\right)(s) = \alpha e^{s(1-n)}\Gamma(b)s^{a-b}U(a,c,s). \tag{2.4}$$

Combining 2.3 and 2.4, we obtain

$$(1+w)^{\frac{1}{2}n^2-2} f_n\left(\frac{1}{1+w}\right) = \alpha(w-n+1)^{b-1}{}_2F_1(a, d; b; -(w-n+1))$$

from which the theorem follows. □

**Corollary 2.1** *Let $h_n$ $(x \geq \sqrt{n})$ be the density of $\kappa_D(A)$ for real matrices. Then*

$$h_n(x) = \mu x^{1-n^2}(x^2-n)^{\frac{n(n+1)}{2}-2}{}_2F_1\left(\frac{n}{2} - \frac{1}{2}, \frac{n}{2} + 1; \frac{n^2}{2} + \frac{n}{2} - 1; -(x^2-n)\right),$$

*where*

$$\mu = \frac{2n\Gamma(\frac{n+1}{2})\Gamma(\frac{n^2}{2})}{\sqrt{\pi}\Gamma(\frac{n(n+1)}{2} - 1)}$$

*and ${}_2F_1$ is the Gauss (hypergeometric) function.*

[1] This formula is incorrect in older editions of [7]. We have verified that the formula as listed in our edition of [7] is indeed correct.

**Proof** This follows from Theorem 2.1 using the standard change of variable formula for probability densities.

**Corollary 2.2** *For fixed $n$, as $x \to \infty$,*

$$h_n(x) \sim \nu_n x^{-2},$$

*where*

$$\nu_n = \frac{n\Gamma(\frac{n+1}{2})\Gamma(\frac{n^2}{2})}{\Gamma(\frac{n^2-1}{2})\Gamma(\frac{n+2}{2})}.$$

*For $n > 20$, $\nu_n \approx n^{3/2}$.*

**Proof** The asymptotic formula for $h_n$ follows from 15.3.4 and 15.1.20 of [1].

## 3 The distribution of $\kappa_D$ (complex case)

The complex case is much easier than the real case, and the resulting formulas are considerably simpler. In [6], we gave a complete derivation of the exact density from first principles, but here we will proceed in an analogous manner to the real case.

Let $A$ be a complex $n \times n$ matrix with independent and identically distributed (**iid**) elements from a complex standard normal distribution. A complex standard normal distribution can be defined as $u + vi$, where $u$ and $v$ are independent standard normals.

The matrix $W = AA^H$ is said to be a complex Wishart matrix or have the complex Wishart distribution. Again our goal is to study the random quantity

$$\kappa_D(A) = \sqrt{\frac{\sum_{i=1}^n \lambda_i}{\lambda_n}},$$

where $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$ are the eigenvalues of the complex Wishart matrix $AA^H$. As in the real case, $\kappa_D(A) \geq \sqrt{n}$.

Using the same notation as in the real case, let $f_n$ be the probability density function (pdf) of $(\kappa_D(A))^{-2} = \lambda_n / \sum_{i=1}^n \lambda_i$ and let $g_n$ be the pdf of $\lambda_n$. The generalization of Lemma 2.1 for the complex case can be found in [8].

**Lemma 3.1** *(Krishnaiah and Schuurmann) The pdfs $f_n$ and $g_n$ are related by*

$$L\left((1+w)^{n^2-2} f_n\left(\frac{1}{1+w}\right)\right)(s) = \Gamma(n^2) e^s s^{1-n^2} g_n(s).$$

Again, we have the density function $g_n$ exactly (see [5] or [6]):

**Lemma 3.2** *The density of the smallest eigenvalue of a complex Wishart matrix is $g_n(x) = ne^{-xn}$, i.e, $n\lambda_{\min}$ is exponentially distributed.*

**Theorem 3.1** *The density of $(\kappa_D(A))^{-2}$ is*

$$f_n(x) = n(n^2-1)(1-nx)^{n^2-2}.$$

**Proof** This formula can be derived from the two lemmas, and the integral formula for the gamma function. □

**Corollary 3.1** *Let $h_n(x)$ $(x \geq \sqrt{n})$ be the density of the condition number $\kappa_D(A)$ for complex matrices. Then*

$$h_n(x) = 2n(n^2-1)x^{1-2n^2}(x^2-n)^{n^2-2}.$$

**Corollary 3.2** *The probability distribution of $\kappa_D$ is given in the complex case by*

$$P(\kappa_D \geq x) = 1 - (1 - n/x^2)^{n^2-1}, \quad x > \sqrt{n}.$$

The above result allows us to verify that indeed

**Corollary 3.3** *For fixed $n$, as $x \to \infty$,*

$$P(\kappa_D \geq x) \sim n(n^2 - 1)/x^2.$$

## Acknowledgements

I would like to thank Jim Demmel for the inspiration for studying this condition number, and Tim Davis for the large amount of time he spent showing me how to use the Alliant FX/80 to perform random matrix experiments to check this work on eight processors in parallel. Also I wish to thank Nick Trefethen's red pen for several useful suggestions.

## References

[1] M. Abramowitz and I.A. Stegun, eds., *Handbook of Mathematical Functions,* Dover Publications, New York, 1970.

[2] T.W. Anderson, *An Introduction to Multivariate Statistical Analysis,* John Wiley & Sons, New York, 1958.

[3] A.W. Davis, On the ratios of the individual latent roots to the trace of a Wishart matrix, *J. Mult. Anal. 2* (1972), 440–443.

[4] J.W. Demmel, The probability that a numerical analysis problem is difficult, *Math. Comput. 50* (1988), 449–480.

[5] A. Edelman, Eigenvalues and condition numbers of random matrices, *SIAM J. Matrix Anal. Appl. 9* (1988), 543–560.

[6] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices,* Ph.D. thesis, Dept. of Math., M.I.T., 1989.

[7] I.S. Gradshteyn and I.W. Ryzhik, *Table of Integrals, Series, and Products,* 6th printing, Academic Press, New York, 1980.

[8] P.R. Krishnaiah and F.J. Schuurmann, On the evaluation of some distributions that arise in simultaneous tests for the equality of the latent roots of the covariance matrix, *J. Mult. Anal. 4* (1974), 265–282.

[9] F.J. Schuurmann, P.R. Krishnaiah, and A.K. Chattopadhyay, On the distributions of the ratios of the extreme roots to the trace of the Wishart matrix, *J. Mult. Anal. 3* (1973), 445–453.