

TAILS OF CONDITION NUMBER DISTRIBUTIONS

ALAN EDELMAN * AND BRIAN D. SUTTON †

Abstract. Let κ be the condition number of an m -by- n matrix with independent standard Gaussian entries, either real ($\beta = 1$) or complex ($\beta = 2$). The major result is the existence of a constant C (depending on m, n, β) such that $P[\kappa > x] < Cx^{-\beta}$ for all x . As $x \rightarrow \infty$, the bound is asymptotically tight. An analytic expression is given for the constant C , and simple estimates are given, one involving a Tracy-Widom largest eigenvalue distribution. All of the results extend beyond real and complex entries to general β .

Key words. condition number, random matrix

AMS subject classifications. 15A12, 15A52

1. Introduction. Take an m -by- n matrix G with iid standard Gaussian entries. What is the distribution of its condition number? The *condition number* κ is defined as the ratio of the largest to smallest singular values,

$$\kappa = \frac{\sigma_m(G)}{\sigma_1(G)} = \sqrt{\frac{\lambda_m(W)}{\lambda_1(W)}},$$

or equivalently as the square root of the ratio of the extreme eigenvalues of the *Wishart matrix* $W = GG^*$. This paper derives an upper bound on the tail $P[\kappa > x]$ for all x . As $x \rightarrow \infty$, the bound is tight. The results apply to the following cases.

- Square or rectangular,
- Finite or infinite,
- Real or complex, as well as extensions to “general β .”

Wishart matrices arise commonly in the statistics literature, for example in principal component analysis. The standard reference is [8]. Traditionally, a Wishart matrix of parameters m, n, β , is defined as GG^* , in which G is a matrix of iid standard Gaussian entries. G is m -by- n and has either real or complex Gaussian entries, depending on the value of β , either 1 or 2. (This paper assumes an identity covariance matrix.) The distributions of the largest and smallest eigenvalues of a Wishart matrix, expressed in terms of zonal polynomials and hypergeometric functions, can be found in [8].

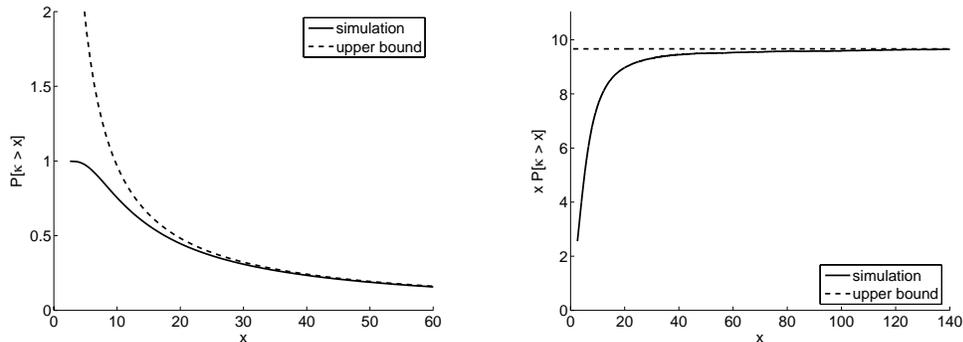
The class of Wishart matrices has been expanded beyond real and complex entries [3, 4]. The *generalized Wishart matrix* $W_{m,n}^\beta$ is defined for non-integral n and any $\beta > 0$. The generalized Wishart matrices are not always expressible as $W = GG^*$, so the condition number is defined in the general case to be

$$\kappa = \sqrt{\lambda_m/\lambda_1}.$$

Interest in random condition numbers first arose in the numerical linear algebra community, for their use in measuring the sensitivity of problems to perturbations. References include [1, 2, 11, 13]. The limiting distribution of the condition number,

*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139–4307, USA.

†Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139–4307, USA. Research supported by a National Science Foundation Graduate Research Fellowship.



(a) The cdf $P[\kappa > x]$ and an asymptotically tight upper bound.

(b) The same cdf and upper bound, after rescaling (multiplying by x).

FIG. 1.1. The condition number of a 6-by-6 matrix of independent real Gaussians.

as the size of the matrix m approaches infinity, can be found in [5]. The present paper considers finite m as well as large m limits.

The large m asymptotics take advantage of the recent observation that the largest eigenvalue of a Wishart matrix approaches a Tracy-Widom distribution as $m \rightarrow \infty$ [6]. This is another example of the convergence of the approaches to Wishart matrices and Gaussian ensembles.

Figure 1.1 plots the cdf of κ , taken from $W_{6,6}^1$, before and after rescaling. By cdf, we mean $P[\kappa > x]$, rather than the more common $P[\kappa < x]$. In this particular case, the tail of the cdf is asymptotic to x^{-1} , so multiplying by x produces a nonzero limiting value at infinity, shown in the second plot. The dashed curve is an upper bound that is asymptotically tight, the main focus of this paper.

In the special case $W_{m,m}^1$, the main theorem is

THEOREM 1.1. *Let κ be the condition number of an m -by- m matrix of iid real standard Gaussian entries. Then*

$$P[\kappa > x] \sim \frac{\sqrt{2}\Gamma(\frac{1}{2}(m+1))}{\Gamma(\frac{1}{2}m)} \mu x^{-1} \quad (x \rightarrow \infty),$$

and the right hand side is an upper bound for all values of x . The constant μ is

$$\mu = E[\sigma_{\max}(G_{m-1,m+1})] = E[\lambda_{\max}(W_{m-1,m+1}^1)^{1/2}],$$

the mean of the largest singular value of an $(m-1)$ -by- $(m+1)$ matrix of independent real standard Gaussian entries.

The constant

$$\lim_{x \rightarrow \infty} xP[\kappa > x] = \frac{\sqrt{2}\Gamma(\frac{1}{2}(m+1))}{\Gamma(\frac{1}{2}m)} \mu$$

is expressed analytically in terms of a multivariate hypergeometric function. Because multivariate hypergeometrics are difficult to compute, we provide estimates based on large m asymptotics. As can be seen in Figure 1.2(a), a combination of exact computation and large m asymptotics produces an estimate for every value of m that

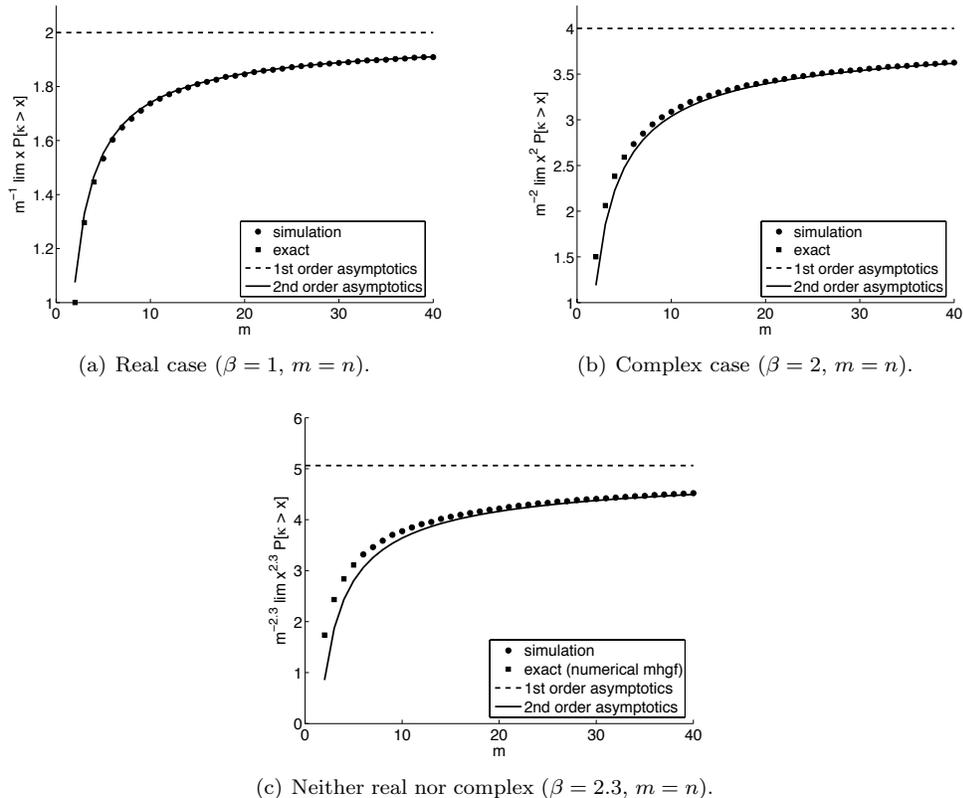


FIG. 1.2. Dependence on the size of the matrix m . The large x limiting value of $(\frac{x}{m})^\beta P[\kappa > x]$ is plotted against m . The asymptotics are developed in Section 4.

agrees well with simulations in the real square case. Figures 1.2(b) and 1.2(c) show similar estimates for $\beta = 2$ and $\beta = 2.3$, respectively. The asymptotic analysis, based on the Tracy-Widom largest eigenvalue distributions, is heuristic and is supported numerically by Figure 1.2.

Our results can be compared with the literature, specifically [1]. First, the present article enjoys greater generality; the cited paper restricts itself to real square matrices. In this restricted case, the reference shows that $\frac{x}{m} P[\kappa > x] < 5.60$. As evident from Figure 1.2(a), the value 5.60 can likely be replaced by 2, and even smaller numbers for finite m . The present article provides, for each m , n , and β , an exact expression for the constant, not just a universal upper bound.

Section 2 describes generalized Wishart matrices in more detail. Section 3 presents the main theorem in full generality, for all values of β . Section 4 provides estimates on the constant in the asymptotic relation. Section 5 considers asymptotics of κ near 1. Section 6 derives the exact cdf for the condition number of 2-by-2 matrices.

2. Generalized Wishart matrices. Traditionally, a real Wishart matrix is defined as the symmetrization GG^T of an m -by- n matrix G of independent Gaussian entries (mean 0, variance 1). The m -by- m Wishart matrix is said to have n degrees of freedom, $n > m - 1$. Letting $\lambda_1 \leq \dots \leq \lambda_m$ be the eigenvalues of GG^T , it follows that the condition number of G is $\kappa = \sqrt{\lambda_m/\lambda_1}$.

LEMMA 3.1. Let $f(\lambda_1, \lambda_m)$ be the joint density of the extreme eigenvalues of $W_{m,n}^\beta$. The density is bounded above by

$$f(\lambda_1, \lambda_m) \leq pC\lambda_1^{p-1}e^{-\lambda_1/2}g_{m-1,n+1}^\beta(\lambda_m),$$

in which $g_{m-1,n+1}^\beta$ is the density of the largest eigenvalue of $W_{m-1,n+1}^\beta$. Also, for fixed λ_m ,

$$f(\lambda_1, \lambda_m) \sim pC\lambda_1^{p-1}g_{m-1,n+1}^\beta(\lambda_m) \quad (\lambda_1 \rightarrow 0).$$

The constant C is

$$C = \frac{m}{p} \frac{c_{m,n}^\beta}{c_{m-1,n+1}^\beta},$$

in which $c_{m,n}^\beta$ is defined in (2.2).

Proof. The factors of $f(\lambda_1, \lambda_m)$ may be grouped conveniently,

$$\begin{aligned} f(\lambda_1, \lambda_m) &= pC \left[\lambda_1^{p-1} e^{-\lambda_1/2} \right] \cdot (m-1)! c_{m-1,n+1}^\beta \\ &\quad \cdot \int \dots \int \prod_{2 \leq j < i \leq m} |\lambda_i - \lambda_j|^\beta \prod_{i=2}^m \left[|\lambda_i - \lambda_1|^\beta \lambda_i^{p-1} \right] e^{-\sum_{i=2}^m \lambda_i/2} d\lambda_2 \dots d\lambda_{m-1}, \end{aligned}$$

where the integration is taken over the region $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{m-1} \leq \lambda_m$.

To obtain the upper bound, use the estimate $|\lambda_i - \lambda_1|^\beta \leq \lambda_i^\beta$ and expand the region of integration. Compare with (2.1).

For the asymptotic result, we need to show that

$$\begin{aligned} (m-1)! c_{m-1,n+1}^\beta \\ \cdot \int \dots \int \prod_{2 \leq j < i \leq m} |\lambda_i - \lambda_j|^\beta \prod_{i=2}^m \left[|\lambda_i - \lambda_1|^\beta \lambda_i^{p-1} \right] e^{-\sum_{i=2}^m \lambda_i/2} d\lambda_2 \dots d\lambda_{m-1} \end{aligned}$$

approaches $g_{m-1,n+1}^\beta(\lambda_m)$, since this quantity is a nonzero constant. To see this, use the dominated convergence theorem. As $\lambda_1 \rightarrow 0$, the integrand converges upwards to the joint density of the eigenvalues of $W_{m-1,n+1}^\beta$. All eigenvalues except for the largest are integrated out, leaving the marginal density of the largest eigenvalue. \square

Now we can present our main theorem. One of the constants involved will be a moment of the largest eigenvalue of a Wishart matrix. Specifically, μ will denote the expected value of λ_{\max}^p , in which λ_{\max} is the largest eigenvalue of a Wishart matrix of parameters $m-1, n+1, \beta$:

$$\mu = E[\lambda_{\max}(W_{m-1,n+1}^\beta)^p]. \quad (3.1)$$

(To clarify this expression, $p = \frac{\beta}{2}(n-m+1)$, not $\frac{\beta}{2}((n+1)-(m-1)+1)$.) μ can be expressed as an integral involving a hypergeometric function; see Section 4.

THEOREM 3.2. Let κ be the condition number associated with a Wishart matrix of parameters m, n, β . Then for $x \geq 1$,

$$P[\kappa > x] \leq C\mu x^{-2p}.$$

As x tends to infinity, the right hand side gives the first order asymptotics for the tail of the condition number,

$$P[\kappa > x] \sim C\mu x^{-2p} \quad (x \rightarrow \infty).$$

C is the same constant from Lemma 3.1, and depends on m, n, β . μ is defined in (3.1).

Proof. Denoting the joint density of λ_1, λ_m by $f(\lambda_1, \lambda_m)$ as in the lemma, we have

$$\begin{aligned} \int_0^{x^{-2}\lambda_m} f(\lambda_1, \lambda_m) d\lambda_1 &\leq pCg_{m-1, n+1}^\beta(\lambda_m) \int_0^{x^{-2}\lambda_m} \lambda_1^{p-1} d\lambda_1 \\ &= C\lambda_m^p g_{m-1, n+1}^\beta(\lambda_m) x^{-2p}. \end{aligned}$$

Therefore, the upper bound is immediate,

$$\begin{aligned} P[\kappa > x] &= \int_0^\infty \int_0^{x^{-2}\lambda_m} f(\lambda_1, \lambda_m) d\lambda_1 d\lambda_m \\ &\leq Cx^{-2p} \int_0^\infty \lambda_m^p g_{m-1, n+1}^\beta(\lambda_m) d\lambda_m \\ &= C\mu x^{-2p}. \end{aligned}$$

For the asymptotics, we need to compute the limit of

$$\frac{1}{C} \frac{1}{\mu} x^{2p} P[\kappa > x] = \frac{1}{\mu} \int_0^\infty \left[\frac{1}{C} x^{2p} \int_0^{x^{-2}\lambda_m} f(\lambda_1, \lambda_m) d\lambda_1 \right] d\lambda_m$$

as $x \rightarrow \infty$. By the second part of the lemma, the integrand w.r.t. $d\lambda_m$ is dominated by and converges pointwise to

$$\lambda_m^p g_{m-1, n+1}^\beta(\lambda_m)$$

as $x \rightarrow \infty$. By the dominated convergence theorem, the entire integral converges to μ , so that

$$\frac{1}{C} \frac{1}{\mu} x^{2p} P[\kappa > x] \rightarrow 1,$$

as desired. \square

The following corollary applies to square matrices of real or complex Gaussians, in particular.

COROLLARY 3.3. *If $m = n$, then*

$$P[\kappa > x] \sim \frac{1}{2^{\beta/2-1}\Gamma(1+\beta)} \frac{\Gamma(\frac{\beta}{2}(m+1))}{\Gamma(\frac{\beta}{2}m)} \mu x^{-\beta} \quad (x \rightarrow \infty).$$

From numerical evidence, e.g., Figure 1.1(b), the rescaled tail appears to be monotonic, but a proof remains an open question.

CONJECTURE 3.4. *For any m, n, β , the rescaled tail $x^{2p} P[\kappa > x]$ is an increasing function of x .*

The following theorem provides a lower bound on the tail for some values of m , n , and β . A lower bound of a different flavor which applies to the real square case ($\beta = 1$, $m = n$) is given in [1].

THEOREM 3.5. *When $p \geq 1$, the condition number is bounded below, for all $x \geq 1$, by*

$$P[\kappa > x] \geq C\mu x^{-2p} - \frac{m}{2} \frac{pC}{p+1} \nu x^{-2p-2},$$

where C , p , and μ are defined as before, and

$$\nu = E[\lambda_{\max}(W_{m-1, n+1}^\beta)^{p+2}].$$

Proof. Perform the change of variables $\rho_i = \lambda_{i+1} - \lambda_1$, $i = 1, \dots, m-1$, and use the estimate $(\rho_i + \lambda_1)^{p-1} \geq \rho_i^{p-1}$ to find the following bound on the joint density of the extreme eigenvalues:

$$\begin{aligned} f(\lambda_1, \lambda_m) &= pC \left[\lambda_1^{p-1} e^{-m\lambda_1/2} \right] \cdot (m-1)! c_{m-1, n+1}^\beta \\ &\quad \cdot \int \dots \int \prod_{1 \leq j < i \leq m-1} |\rho_i - \rho_j|^\beta \prod_{i=1}^{m-1} \left[\rho_i^\beta (\rho_i + \lambda_1)^{p-1} \right] e^{-\sum_{i=1}^{m-1} \rho_i/2} d\rho_1 \dots d\rho_{m-2} \\ &\geq pC \left[\lambda_1^{p-1} e^{-m\lambda_1/2} \right] g_{m-1, n+1}^\beta(\rho_{m-1}) \\ &= pC \left[\lambda_1^{p-1} e^{-m\lambda_1/2} \right] g_{m-1, n+1}^\beta(\lambda_m - \lambda_1), \end{aligned}$$

in which the integral is taken over the region $0 \leq \rho_1 \leq \dots \leq \rho_2 \leq \rho_{m-1} = \lambda_m - \lambda_1$. Then, the tail is bounded by

$$\begin{aligned} P[\kappa > x] &= \int_0^\infty \int_0^{x^{-2}\lambda_m} f(\lambda_1, \lambda_m) d\lambda_1 d\lambda_m \\ &\geq pC \int_0^\infty \left(\int_0^{x^{-2}\lambda_m} \lambda_1^{p-1} e^{-m\lambda_1/2} d\lambda_1 \right) g_{m-1, n+1}^\beta(\lambda_m - \lambda_1) d\lambda_m. \end{aligned}$$

Applying the change of variables $s = \lambda_1$, $t = \lambda_m - \lambda_1$, and estimating $e^{-m\lambda_1/2} \geq 1 - \frac{m}{2}\lambda_1$, the lower bound becomes

$$\begin{aligned} P[\kappa > x] &\geq pC \int_0^\infty \left[\int_0^{x^{-2}(1-x^{-2})^{-1}t} s^{p-1} \left(1 - \frac{m}{2}s \right) ds \right] g_{m-1, n+1}^\beta(t) dt \\ &\geq pC \int_0^\infty \left[\int_0^{x^{-2}t} s^{p-1} \left(1 - \frac{m}{2}s \right) ds \right] g_{m-1, n+1}^\beta(t) dt \\ &= C\mu x^{-2p} - \frac{m}{2} \frac{pC}{p+1} \nu x^{-2p-2}. \end{aligned}$$

□

4. Estimates for the constant. Recall the main theorem,

$$P[\kappa > x] \sim C\mu x^{-2p} \quad (x \rightarrow \infty).$$

In this section, we give an exact formula and heuristic estimates for the constant $C\mu$. Then we compare the estimates with numerical data in the cases $\beta = 1, 2, 2.3$. The value 2.3 was chosen arbitrarily, in order to illustrate the approach when $\beta \neq 1, 2$.

The estimates are based on large m asymptotics. They will be restricted to the “square” case, $m = n$, for ease of presentation. The arguments work just as well for $m \neq n$.

4.1. Exact value. The constant C was specified earlier, and

$$\mu = E[\lambda_{\max}(W_{m-1,n+1}^\beta)^p].$$

The cdf for $\lambda_{\max}(W_{m-1,n+1}^\beta)$ is

$$\frac{\Gamma_{m-1}^\beta(\frac{\beta}{2}(m-2)+1)}{\Gamma_{m-1}^\beta(\frac{\beta}{2}(n+m-1)+1)} \cdot \left(\frac{x}{2}\right)^{(m-1)(n+1)\beta/2} {}_1F_1^\beta\left(\frac{\beta}{2}(n+1), \frac{\beta}{2}(n+m-1), -\frac{x}{2}I_{m-1}\right),$$

in which Γ_{m-1}^β is a multivariate gamma function and ${}_1F_1^\beta$ is a multivariate hypergeometric function [3]. Therefore, the p th moment is an integral over the corresponding measure.

4.2. First order asymptotics. First, we will develop a crude estimate, which will motivate a more refined estimate. Start by applying Stirling’s formula to Corollary 3.3. If $m = n$, then

$$\begin{aligned} C &= \frac{1}{2^{\beta/2-1}\Gamma(1+\beta)} \frac{\Gamma(\frac{\beta}{2}(m+1))}{\Gamma(\frac{\beta}{2}m)} \\ &\sim \frac{\beta^{\beta/2-1}}{2^{\beta-1}\Gamma(\beta)} m^{\beta/2} \quad (m \rightarrow \infty). \end{aligned}$$

We shall argue heuristically in a moment that the mean of the largest eigenvalue of $W_{m-1,m+1}^\beta$ is asymptotic to $4\beta m$, and the standard deviation is dominated by the mean. Hence, an immediate estimate for μ is

$$\begin{aligned} \mu &= E[\lambda_{\max}(W_{m-1,m+1}^\beta)^{\beta/2}] \\ &\sim 2^\beta \beta^{\beta/2} m^{\beta/2} \quad (m \rightarrow \infty), \end{aligned}$$

giving

$$\lim_{x \rightarrow \infty} x^\beta P[\kappa > x] = C\mu \sim \frac{2\beta^{\beta-1}}{\Gamma(\beta)} m^\beta \quad (m \rightarrow \infty). \quad (4.1)$$

The estimate (4.1) is plotted for $\beta = 1, 2, 2.3$ in Figure 1.2. For $\beta = 1$, the estimate is the horizontal line at 2. For $\beta = 2$, it is the horizontal line at 4. For $\beta = 2.3$, it is the horizontal line at ≈ 5.1 . It is clear that this asymptotic estimate is only valid for very large matrices.

As an extra sanity check, this estimate can be compared with the large m limiting distribution of the condition number, when $\beta = 1, 2$.

PROPOSITION 4.1 ([5]). *Let κ be the condition number associated with $W_{m,m}^1$. Then*

$$\lim_{m \rightarrow \infty} P[\kappa/m > x] = 1 - e^{-2/x - 2/x^2}.$$

Let $\tilde{\kappa}$ be the condition number associated with $W_{m,m}^2$. Then

$$\lim_{m \rightarrow \infty} P[\tilde{\kappa}/m > x] = 1 - e^{-4/x^2}.$$

Taking series expansions,

$$\lim_{m \rightarrow \infty} \frac{1}{m} P[\kappa > x] \sim 2/x \quad (x \rightarrow \infty),$$

and

$$\lim_{m \rightarrow \infty} \frac{1}{m^2} P[\tilde{\kappa} > x] \sim 4/x^2 \quad (x \rightarrow \infty).$$

In these two cases, at least, the limits $x \rightarrow \infty$, $m \rightarrow \infty$ can be interchanged.

4.3. The Tracy-Widom distributions. For a more refined estimate, the result of Johnstone on the largest eigenvalue of a Wishart matrix will be key. It states that as $m \rightarrow \infty$, the largest eigenvalue of an m -by- m Wishart matrix, appropriately scaled, converges in distribution to a Tracy-Widom largest eigenvalue distribution.

The Tracy-Widom distributions were originally discovered as the limiting distributions of the largest eigenvalues of Gaussian ensembles. Their cdf's are expressed in terms of a Painlevé transcendent and can be found in [6, 12]. For our purposes, just the means of the distributions will suffice:

$$E[TW_1] \approx -1.21, \quad E[TW_2] \approx -1.77, \quad E[TW_{2.3}] \approx -1.8,$$

in which TW_β is a random variable following the Tracy-Widom distribution of parameter β . (The means for $\beta = 1$ and $\beta = 2$ were computed using a numerical ode solver and the Tracy-Widom densities. The mean for $\beta = 2.3$ was computed by Monte Carlo simulation.)

Although Johnstone's theorem handles "rectangular" cases ($m \neq n$), we present only the "square" case. Note that Johnstone's complex Wishart matrix differs from ours by a factor of $\frac{1}{2}$.

THEOREM 4.2 ([6]). *Fix $\beta = 1$ or $\beta = 2$. Let l_m be the largest eigenvalue of $W_{m-1, m+1}^\beta$, and scale each random variable in the sequence by*

$$\tilde{l}_m = \frac{l_m - 4\beta m}{2^{4/3} \beta m^{1/3}}. \quad (4.2)$$

Then the sequence \tilde{l}_m converges in distribution to the Tracy-Widom largest eigenvalue distribution for β .

CONJECTURE 4.3. *The theorem also holds for general β .* Although the generalized Tracy-Widom distributions ($\beta \neq 1, 2, 4$) have no known closed form solutions, the distributions can be sampled quite efficiently when $\beta \geq 1$ [4, 9].

TABLE 4.1

Large m asymptotic approximation to $\lim_{x \rightarrow \infty} x^\beta P[\kappa > x]$ when $m = n$. These estimates are based on Conjecture 4.3.

β	Large $m = n$ asymptotics for $C\mu = \lim_{x \rightarrow \infty} x^\beta P[\kappa > x]$
1	$\frac{1}{\left(\sqrt{m} - \frac{1}{4\sqrt{m}}\right) \left(\sqrt{m} + \sqrt{m-1} + \frac{1}{2^{2/3}} E[TW_1] m^{-1/6}\right)}$
general	$\frac{2\beta^\beta}{\Gamma(1+\beta)} m^\beta \left(1 + \frac{\beta-2}{4} m^{-1}\right) \left(1 + 2^{-5/3} \beta E[TW_\beta] m^{-2/3}\right)$

4.4. Second order asymptotics. Now, we can compute a second order approximation to the constant $C\mu$. Stirling's second order approximation gives

$$C = \frac{1}{2^{\beta/2-1} \Gamma(1+\beta)} \frac{\Gamma\left(\frac{\beta}{2}(m+1)\right)}{\Gamma\left(\frac{\beta}{2}m\right)}$$

$$\sim \frac{\beta^{\beta/2}}{2^{\beta-1} \Gamma(1+\beta)} m^{\beta/2} \left(1 + \frac{\beta-2}{4} m^{-1}\right) \quad (m \rightarrow \infty),$$

and Johnstone's result suggests

$$\begin{aligned} \mu &= E[\lambda_{\max}(W_{m-1, m+1}^\beta)^{\beta/2}] \\ &\sim E[(4\beta m + 2^{4/3} \beta m^{1/3} TW_\beta)^{\beta/2}] \\ &\sim E\left[(4\beta m)^{\beta/2} + \frac{\beta}{2} (4\beta m)^{\beta/2-1} 2^{4/3} \beta m^{1/3} TW_\beta\right] \\ &\sim 2^\beta \beta^{\beta/2} m^{\beta/2} \left(1 + 2^{-5/3} \beta E[TW_\beta] m^{-2/3}\right) \quad (m \rightarrow \infty). \end{aligned}$$

To avoid questions of convergence, we simply label this estimate as heuristic. When $\beta = 1$ and m is small, the moment is better approximated by

$$\begin{aligned} \mu &\sim \beta^{\beta/2} (\sqrt{m} + \sqrt{m-1})^\beta + 2^{\beta-5/3} \beta^{\beta/2+1} E[TW_\beta] m^{\beta/2-2/3} \\ &\sim \sqrt{m} + \sqrt{m-1} + 2^{-2/3} E[TW_1] m^{-1/6} \quad (m \rightarrow \infty), \end{aligned}$$

as suggested by [6]. (Of course, this estimate agrees with the previous one asymptotically. There may be a generalization of this formula to general β , providing a reasonably accurate estimate for small m without modifying the large m asymptotics, but we are not aware of one.)

Multiplying the estimates gives Table 4.1.

4.5. Small m . For small values of m , the constant $C\mu$ can be computed directly. In the cases $\beta = 1, 2$, the constant can be computed symbolically in Mathematica; see Tables 4.2(a) and 4.2(b). For other values of β , the moment can be evaluated numerically using equation (4.1); see Table 4.2(c). The numerical entries in this table were computed with the help of Plamen Koev's software for multivariate hypergeometric functions [7].

5. Asymptotics at $\kappa = 1$. Asymptotics can also be derived near $\kappa = 1$. The proofs are broken into two cases, $p \leq 1$ and $p \geq 1$.

Throughout this section,

$$D = m! c_{m,n}^\beta \frac{I}{\beta \binom{m}{2} + m - 1} \left(\frac{2}{m}\right)^{\beta \binom{m}{2} + mp} \Gamma\left(\beta \binom{m}{2} + mp\right),$$

TABLE 4.2
Preferred approximation to $\lim_{x \rightarrow \infty} x^\beta P[\kappa > x]$ when $m = n$.

(a) Real case.

$m = n$	$\beta = 1$: Preferred approximation to $C\mu = \lim_{x \rightarrow \infty} xP[\kappa > x]$
2	2
3	$11/(2\sqrt{2})$
4	$2^{-13/2}(2510\sqrt{2}/9 + 105\pi - 210 \tan^{-1}(\sqrt{2}))$
≥ 5	$(\sqrt{m} - \frac{1}{4\sqrt{m}}) \left(\sqrt{m} + \sqrt{m-1} + \frac{1}{2^{2/3}} E[TW_1] m^{-1/6} \right)$

(b) Complex case.

$m = n$	$\beta = 2$: Preferred approximation to $C\mu = \lim_{x \rightarrow \infty} x^2 P[\kappa > x]$
2	6
3	297/16
4	1332815/34992
5	4512225660125/69657034752
≥ 6	$4m^2 (1 + 2^{-2/3} E[TW_2] m^{-2/3})$

(c) Neither real nor complex.

$m = n$	$\beta = 2.3$: Preferred approximation to $C\mu = \lim_{x \rightarrow \infty} x^{2.3} P[\kappa > x]$
2	8.54
3	30.44
4	68.81
5	126.03
≥ 6	$4m^2 (1 + 2^{-2/3} E[TW_{2.3}] m^{-2/3})$

where

$$I = \int \dots \int \prod_{1 \leq j < i \leq m} |t_i - t_j|^\beta dt_2 \dots dt_{m-1},$$

the integral taken over the region $0 = t_1 \leq t_2 \leq \dots \leq t_{m-1} \leq t_m = 1$.

LEMMA 5.1. *Let κ be the condition number associated with a Wishart matrix of parameters m, n, β , and suppose that $p \leq 1$. Then for all $x \geq 1$, the following upper and lower bounds hold:*

$$P[\kappa \leq x] \leq D(x^2 - 1)^{\beta \binom{m}{2} + m - 1},$$

$$P[\kappa \leq x] \geq D \left(\frac{x^2 - 1}{x^2} \right)^{\beta \binom{m}{2} + m - 1}.$$

Proof. Begin by changing to variables $\lambda_1, t_2, \dots, t_{m-1}, \Delta$:

$$\lambda_1 = \lambda_1$$

$$t_i = \frac{\lambda_i - \lambda_1}{\lambda_m - \lambda_1}, \quad i = 2, \dots, m-1$$

$$\Delta = \lambda_m - \lambda_1.$$

The associated Jacobian matrix is lower triangular, and the Jacobian is Δ^{m-2} . Introducing the shorthand $t_1 = 0, t_m = 1$, the joint density of $\lambda_1, t_2, \dots, t_{m-1}, \Delta$ is

$$f(\lambda_1, t_2, \dots, t_{m-1}, \Delta) =$$

$$m! c_{m,n}^\beta \Delta^{\beta \binom{m}{2} + m - 2} e^{-m\lambda_1/2} \prod_{1 \leq j < i \leq m} |t_i - t_j|^\beta \prod_{i=1}^m (\lambda_1 + t_i \Delta)^{p-1} e^{-(\Delta/2) \sum_{i=1}^m t_i}.$$

Because

$$P[\kappa \leq x] = \int_0^\infty \int_0^{\lambda_1(x^2-1)} \left[\int \dots \int f(\lambda_1, t_2, \dots, t_{m-1}, \Delta) dt_2 \dots dt_{m-1} \right] d\Delta d\lambda_1,$$

with the inner integral taken over the region $0 = t_1 \leq t_2 \leq \dots \leq t_{m-1} \leq t_m = 1$, we have

$$P[\kappa \leq x] = m! c_{m,n}^\beta \int_0^\infty \int_0^{\lambda_1(x^2-1)} \Delta^{\beta \binom{m}{2} + m - 2} g(\lambda_1, \Delta) d\Delta d\lambda_1, \quad (5.1)$$

in which

$$g(\lambda_1, \Delta) = \frac{\int \dots \int f(\lambda_1, t_2, \dots, t_{m-1}, \Delta) dt_2 \dots dt_{m-1}}{m! c_{m,n}^\beta \Delta^{\beta \binom{m}{2} + m - 2}}. \quad (5.2)$$

The asymptotic decay rate $(x^2 - 1)^{\beta \binom{m}{2} + m - 1}$ seems inevitable.

To obtain the bounds, observe

$$\begin{aligned} g(\lambda_1, \Delta) &\leq \lambda_1^{m(p-1)} e^{-m\lambda_1/2} I, \\ g(\lambda_1, \Delta) &\geq (\lambda_1 + \Delta)^{m(p-1)} e^{-m(\lambda_1 + \Delta)/2} I. \end{aligned}$$

Therefore,

$$\begin{aligned} P[\kappa \leq x] &\leq m! c_{m,n}^\beta I \int_0^\infty \lambda_1^{m(p-1)} e^{-m\lambda_1/2} \left(\int_0^{\lambda_1(x^2-1)} \Delta^{\beta \binom{m}{2} + m - 2} d\Delta \right) d\lambda_1 \\ &= D(x^2 - 1)^{\beta \binom{m}{2} + m - 1}, \\ P[\kappa \leq x] &\geq m! c_{m,n}^\beta I \int_0^\infty \int_0^{\lambda_1(x^2-1)} \Delta^{\beta \binom{m}{2} + m - 2} (\lambda_1 + \Delta)^{m(p-1)} e^{-m(\lambda_1 + \Delta)/2} d\Delta d\lambda_1 \\ &= m! c_{m,n}^\beta I \int_0^\infty \lambda_m^{m(p-1)} e^{-m\lambda_m/2} \left(\int_0^{\lambda_m(x^2-1)/x^2} \Delta^{\beta \binom{m}{2} + m - 2} d\Delta \right) d\lambda_m \\ &= D \left(\frac{x^2 - 1}{x^2} \right)^{\beta \binom{m}{2} + m - 1}. \end{aligned}$$

□

LEMMA 5.2. *Let κ be the condition number associated with a Wishart matrix of parameters m , n , β , and suppose that $p \geq 1$. Then for all $x \geq 1$, the following upper and lower bounds hold:*

$$\begin{aligned} P[\kappa \leq x] &\leq Dx^2(x^2 - 1)^{\beta \binom{m}{2} + m - 1} \\ P[\kappa \leq x] &\geq Dx^{-2(\beta \binom{m}{2} + mp)}(x^2 - 1)^{\beta \binom{m}{2} + m - 1}. \end{aligned}$$

Proof. Equations (5.1) and (5.2) from the proof of the previous lemma are still true. Now that $p \geq 1$, we have the bounds

$$\begin{aligned} g(\lambda_1, \Delta) &\leq (\lambda_1 + \Delta)^{m(p-1)} e^{-m\lambda_1/2} I, \\ g(\lambda_1, \Delta) &\geq \lambda_1^{m(p-1)} e^{-m(\lambda_1 + \Delta)/2} I. \end{aligned}$$

Therefore,

$$\begin{aligned}
P[\kappa \leq x] &\leq m! c_{m,n}^\beta I \int_0^\infty \int_0^{\lambda_1(x^2-1)} \Delta^{\beta(\frac{m}{2})+m-2} (\lambda_1 + \Delta)^{m(p-1)} e^{-m\lambda_1/2} d\Delta d\lambda_1 \\
&\leq m! c_{m,n}^\beta I \int_0^\infty \lambda_m^{m(p-1)} e^{-m(\lambda_m/x^2)/2} \left(\int_0^{\lambda_m(x^2-1)/x^2} \Delta^{\beta(\frac{m}{2})+m-2} d\Delta \right) d\lambda_m \\
&= Dx^2(x^2-1)^{\beta(\frac{m}{2})+m-1}, \\
P[\kappa \leq x] &\geq m! c_{m,n}^\beta I \int_0^\infty \int_0^{\lambda_1(x^2-1)} \Delta^{\beta(\frac{m}{2})+m-2} \lambda_1^{m(p-1)} e^{-m(\lambda_1+\Delta)/2} d\Delta d\lambda_1 \\
&\geq m! c_{m,n}^\beta I \int_0^\infty \lambda_1^{m(p-1)} e^{-m(\lambda_1 x^2)/2} \left(\int_0^{\lambda_1(x^2-1)} \Delta^{\beta(\frac{m}{2})+m-2} d\Delta \right) d\lambda_1, \\
&= Dx^{-2(\beta(\frac{m}{2})+mp)} (x^2-1)^{\beta(\frac{m}{2})+m-1}.
\end{aligned}$$

□

THEOREM 5.3. *Let κ be the condition number associated with a Wishart matrix of parameters m, n, β . Then*

$$P[\kappa \leq x] \sim D(x^2-1)^{\beta(\frac{m}{2})+m-1} \quad (x \rightarrow 1).$$

6. Two-by-two matrices. Performing the change of variables $\kappa = \sqrt{\lambda_2/\lambda_1}$ and integrating out λ_1 in the joint density (2.1), one arrives at the density of the condition number of a 2-by-2 Wishart matrix,

$$f(x) = \frac{4\Gamma(1 + \frac{\beta}{2})\Gamma(\beta n)}{\Gamma(1 + \beta)\Gamma(p)\Gamma(\frac{\beta}{2}n)} \frac{x^{2p-1}(x^2-1)^\beta}{(x^2+1)^{\beta n}}.$$

Near $\kappa = \infty$, the density is asymptotic to

$$f(x) \sim \frac{4\Gamma(1 + \frac{\beta}{2})\Gamma(\beta n)}{\Gamma(p)\Gamma(1 + \beta)\Gamma(\frac{\beta}{2}n)} x^{-2p-1} \quad (x \rightarrow \infty).$$

Integrating,

$$P[\kappa > x] \sim \frac{2\Gamma(1 + \frac{\beta}{2})\Gamma(\beta n)}{\Gamma(1 + p)\Gamma(1 + \beta)\Gamma(\frac{\beta}{2}n)} x^{-2p} \quad (x \rightarrow \infty).$$

If $n = 2 = m$, this simplifies to

$$P[\kappa > x] \sim \frac{2\Gamma(2\beta)}{\Gamma(1 + \beta)\Gamma(\beta)} x^{-\beta} \quad (x \rightarrow \infty).$$

The prediction from Corollary 3.3 is

$$P[\kappa > x] \sim \frac{\Gamma(\frac{3\beta}{2})}{2^{\beta/2-1}\Gamma(1 + \beta)\Gamma(\beta)} \mu x^{-\beta} \quad (x \rightarrow \infty),$$

with

$$\mu = E[\lambda_{\max}(W_{1,3}^\beta)^{\beta/2}] = E[\chi_{3\beta}^\beta] = \frac{2^{\beta/2}\Gamma(2\beta)}{\Gamma(3\beta/2)},$$

so that

$$P[\kappa > x] \sim \frac{2\Gamma(2\beta)}{\Gamma(1+\beta)\Gamma(\beta)}x^{-\beta} \quad (x \rightarrow \infty).$$

The main theorem is verified in this case.

REFERENCES

- [1] J.-M. AZAÏS AND M. WSCHEBOR, *Upper and lower bounds for the tails of the distribution of the condition number of a Gaussian matrix*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 426–440.
- [2] J. W. DEMMEL, *The probability that a numerical analysis problem is difficult*, Math. Comp., 50 (1988), pp. 449–480.
- [3] I. DUMITRIU, *Eigenvalue statistics for beta-ensembles*, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA 02139, June 2003.
- [4] I. DUMITRIU AND A. EDELMAN, *Matrix models for beta ensembles*, J. Math. Phys., 43 (2002), pp. 5830–5847.
- [5] A. EDELMAN, *Eigenvalues and condition numbers of random matrices*, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA 02139, May 1989.
- [6] I. M. JOHNSTONE, *On the distribution of the largest eigenvalue in principal components analysis*, Ann. Statist., 29 (2001), pp. 295–327.
- [7] P. KOEV AND A. EDELMAN, *Efficient evaluation of the hypergeometric function of matrix argument*. <http://math.mit.edu/~plamen/files/hyper.pdf>.
- [8] R. J. MUIRHEAD, *Aspects of multivariate statistical theory*, John Wiley & Sons Inc., New York, 1982. Wiley Series in Probability and Mathematical Statistics.
- [9] P.-O. PERSSON, *Numerical methods for eigenvalue distributions of random matrices*, (2005). <http://arxiv.org/abs/math-ph/0501068>.
- [10] J. W. SILVERSTEIN, *The smallest eigenvalue of a large-dimensional Wishart matrix*, Ann. Probab., 13 (1985), pp. 1364–1368.
- [11] S. SMALE, *On the efficiency of algorithms of analysis*, Bull. Amer. Math. Soc. (N.S.), 13 (1985), pp. 87–121.
- [12] C. A. TRACY AND H. WIDOM, *Airy kernel and Painlevé II*, in *Isomonodromic deformations and applications in physics (Montréal, QC, 2000)*, vol. 31 of CRM Proc. Lecture Notes, Amer. Math. Soc., Providence, RI, 2002, pp. 85–96.
- [13] J. VON NEUMANN, *Collected works. Vol. V: Design of computers, theory of automata and numerical analysis*, General editor: A. H. Taub. A Pergamon Press Book, The Macmillan Co., New York, 1963.