Random Matrices: Theory and Applications

The Densities and Distributions of the Largest Eigenvalue and the Trace of a Beta-Wishart Matrix

VESSELIN DRENSKY

Institute of Mathematics and Informatics, Bulgarian Academy of Sciences
1113 Sofia, Bulgaria
drensky@math.bas.bg

ALAN EDELMAN

Department of Mathematics, Massachusetts Institute of Technology,
77 Massachusetts Avenue, Cambridge, MA 02139, U.S.A.
edelman@math.mit.edu

TIERNEY GENOAR

Department of Mathematics, San Jose State University,
1 Washington Square, San Jose, CA 95192, U.S.A.
tierney.genoar@sjsu.edu

RAYMOND KAN

Joseph L. Rotman School of Management, University of Toronto
105 St. George Street, Toronto, Ontario, Canada M5S 3E6
kan@chass.utoronto.ca

PLAMEN KOEV

Department of Mathematics, San Jose State University,
1 Washington Square, San Jose, CA 95192, U.S.A.
plamen.koev@sjsu.edu

Received (January 31, 2019)
Revised (Day Month Year)

Keywords: Wishart; eigenvalue; trace; hypergeometric function of a matrix argument

Mathematics Subject Classification 2000: 15A52, 60E05, 62H10, 65F15
We present new expressions for the densities and distributions of the largest eigenvalue and the trace of a Beta–Wishart matrix. The series expansions for these expressions involve fewer terms than previously known results. For the trace we also present a new algorithm that is linear in the size of the matrix and the degree of truncation, which is optimal.

1. Introduction

The Beta–Wishart ensemble was introduced recently in [5,8]. The eigenvalues of a Wishart matrix and its trace have long been used in multivariate statistical analysis for a variety of analyses and applications [11]. The only known expressions from [6] ((3.1) and (4.3) below), however, are in terms of infinite series of Jack functions, and in particular, the hypergeometric function of a matrix argument. These series are notoriously slow to converge and have been a computational challenge for decades despite recent progress [10]. The main issue is the exponential number of terms in (a finite truncation of) the expansion of hypergeometric function as a series of Jack functions. For an $m \times m$ matrix argument truncated for integer partitions of size not exceeding some M, that number grows exponentially as $O(M^m)$. In turn, each Jack function presents a computational challenge of its own [2].

In this paper we improve on previously known results for the largest eigenvalue and the trace as follows. For the largest eigenvalue, we exploit a connection with the Beta–MANOVA ensemble [4] to derive new expressions for the density and the distribution of the largest eigenvalue that only requires summation over integer partitions of size not exceeding $m-1$ parts instead of m. This modest improvement results in substantial computational savings given the exponential nature of the current best algorithm for computing the hypergeometric function.

For the trace, we present new expressions for its density and distribution in terms of the hypergeometric function of a matrix argument, but whose series expansion only involves Jack functions corresponding to partitions in only one part. We derive a new algorithm for computing each of these Jack functions in only $O(m)$ time, which is optimal.

Furthermore, since there is only one partition in only one part for any integer, the density (and in turn, the distribution) of the trace can be computed in time that is linear in both the size of the matrix, m, and the degree of the truncation, M, an optimal result. Our final expressions for trace are subtraction-free, meaning they are guaranteed to be computed to high relative accuracy in the presence of roundoff errors.

We finish our presentation with numerical tests.

2. Preliminaries

All results in this paper are for Beta random matrices, therefore we drop the Beta prefix when referring to the Wishart or MANOVA ensembles.

The sequence $\kappa = (\kappa_1, \kappa_2, \ldots)$ is a *partition* of an integer $k \geq 0$ (denoted $\kappa \vdash k$) if $\kappa_1 \geq \kappa_2 \geq \cdots \geq 0$ are integers such that $|\kappa| \equiv \kappa_1 + \kappa_2 + \cdots = k$.

For a partition $\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_m)$ and parameter $\beta > 0$, the *generalized Pochhammer symbol* is defined as

$$(a)_\kappa^{(\beta)} \equiv \prod_{i=1}^{m} \left(a - \tfrac{i-1}{2}\beta\right)_{\kappa_i}, \tag{2.1}$$

where $(a)_k = a(a+1)\cdots(a+k-1)$ is the *rising factorial*. In particular, $\left(\frac{m\beta}{2}\right)^{(\beta)}_\kappa = 0$ for partitions κ in more than m (nonzero) parts. For a partition $\kappa = (k)$ in only one part, $(a)^{(\beta)}_\kappa = (a)_k$ is independent of β.

The *multivariate Gamma function* of parameter $\beta > 0$ is defined as

$$\Gamma^{(\beta)}_m(c) \equiv \pi^{\frac{m(m-1)}{4}\beta} \prod_{i=1}^{m} \Gamma\big(c - \tfrac{i-1}{2}\beta\big) \quad \text{for } \Re(c) > \tfrac{m-1}{2}\beta. \tag{2.2}$$

In particular, the above definition immediately implies

$$\lim_{x\to\infty} \frac{\Gamma^{(\beta)}_m(x+a)}{\Gamma^{(\beta)}_m(x)x^{ma}} = 1, \tag{2.3}$$

which we will utilize below.

For an $m \times m$ Hermitian matrix X, the *Jack function*

$$C^{(\beta)}_\kappa(X) = C^{(\beta)}_\kappa(x_1, x_2, \ldots, x_m)$$

is a symmetric, homogeneous polynomial of degree $|\kappa|$ in the eigenvalues $x_1, x_2, \ldots, x_m$ of X [13, Proposition 4.2].

For integers $p \ge 0$ and $q \ge 0$, and $m \times m$ Hermitian matrices X and Y the *hypergeometric function of two matrix arguments* X and Y and parameter $\beta > 0$ is

$${}_pF^{(\beta)}_q(a_1, \ldots, a_p; b_1, \ldots, b_q; X, Y) \equiv \sum_{k=0}^{\infty} \sum_{\kappa\vdash k} \frac{1}{k!} \cdot \frac{(a_1)^{(\beta)}_\kappa \cdots (a_p)^{(\beta)}_\kappa}{(b_1)^{(\beta)}_\kappa \cdots (b_q)^{(\beta)}_\kappa} \cdot \frac{C^{(\beta)}_\kappa(X)C^{(\beta)}_\kappa(Y)}{C^{(\beta)}_\kappa(I_m)}. \tag{2.4}$$

For *one matrix argument*,

$${}_pF^{(\beta)}_q(a_1, \ldots, a_p; b_1, \ldots, b_q; X) \equiv {}_pF^{(\beta)}_q(a_1, \ldots, a_p; b_1, \ldots, b_q; X, I_m).$$

We will utilize the following identity for the hypergeometric function of a matrix argument [7, (13.4)]

$${}_1F^{(\beta)}_0(a; X) = |I_m - X|^{-a}, \qquad \text{for } X < I_m. \tag{2.5}$$

3. The density of the largest eigenvalue of the Wishart ensemble

We obtain this result as a limiting argument from the density of the largest eigenvalue of the MANOVA ensemble.

Proposition 3.1. *Let $\mu_i, i = 1, 2, \ldots, m$, be the eigenvalues of a MANOVA matrix with parameters a, b and covariance Σ^{-1}, then*

$$\lambda_i \equiv \lim_{b\to\infty} \frac{2b}{\beta} \cdot \frac{\mu_i}{1-\mu_i},$$

$i = 1, 2, \ldots, m$, are the eigenvalues of a Wishart matrix with $n = \frac{a\beta}{2}$ degrees of freedom and covariance Σ.

Proof. We start with the joint density of $M = \mathrm{diag}\,(\mu_1, \mu_2, \ldots, \mu_m)$ from [4]

$$\frac{1}{S_m^{(\beta)}(a,b)}|\Sigma|^{-a}|M|^{a-r}|I_m - M|^{-a-r}\ {}_1F_0^{(\beta)}(a+b; M(M-I_m)^{-1}, \Sigma^{-1})\mathrm{d}\mu(M),$$

where $r \equiv \frac{m-1}{2}\beta + 1$,

$$\mathrm{d}\mu(M) = \prod_{i<j}|\mu_i - \mu_j|^\beta \mathrm{d}\mu_1\mathrm{d}\mu_2\cdots\mathrm{d}\mu_m$$

and $S_m^{(\beta)}(a,b)$ is the value of the Selberg Integral [12]

$$S_m^{(\beta)}(a,b) \equiv \frac{m!\Gamma_m^{(\beta)}\left(\frac{m}{2}\beta\right)}{\pi^{\frac{m(m-1)}{2}\beta}\left(\Gamma\left(\frac{\beta}{2}\right)\right)^m} \cdot \frac{\Gamma_m^{(\beta)}(a)\,\Gamma_m^{(\beta)}(b)}{\Gamma_m^{(\beta)}(a+b)}.$$

We change variables $N = \frac{2b}{\beta}M(I_m - M)^{-1}$. The Jacobian is $\left(\frac{2b}{\beta}\right)^m|I_m - M|^{-2}$ and $\mathrm{d}\mu(M) = \left(\frac{2b}{\beta}\right)^{m(m-1)\beta/2}|I_m - M|^{(m-1)\beta}\mathrm{d}\mu(N)$, so the density of N is

$$\frac{\pi^{\frac{m(m-1)}{2}\beta}\left(\Gamma\left(\frac{\beta}{2}\right)\right)^m\Gamma_m^{(\beta)}(a+b)}{m!\Gamma_m^{(\beta)}\left(\frac{m}{2}\beta\right)\Gamma_m^{(\beta)}(a)\,\Gamma_m^{(\beta)}(b)}|\Sigma|^{-a}\left(\frac{\beta}{2b}\right)^{ma}|N|^{a-r}\ {}_1F_0^{(\beta)}\left(a+b; -\frac{\beta}{2b}N, \Sigma^{-1}\right)\mathrm{d}\mu(N),$$

which equals

$$\frac{\pi^{\frac{m(m-1)}{2}\beta}\left(\Gamma\left(\frac{\beta}{2}\right)\right)^m\left(\frac{\beta}{2}\right)^{ma}}{m!\Gamma_m^{(\beta)}\left(\frac{m}{2}\beta\right)\Gamma_m^{(\beta)}(a)} \cdot \frac{\Gamma_m^{(\beta)}(a+b)}{\Gamma_m^{(\beta)}(b)b^{ma}}|\Sigma|^{-a}|N|^{a-r}\ {}_1F_0^{(\beta)}\left(a+b; -\frac{1}{b}\cdot\frac{\beta}{2}N, \Sigma^{-1}\right)\mathrm{d}\mu(N).$$

Now, directly from the definition (2.1), $\lim_{b\to\infty}\frac{(a+b)_\kappa^{(\beta)}}{b^{|\kappa|}} = 1$, thus

$$\lim_{b\to\infty}\ {}_1F_0^{(\beta)}(a+b; -\tfrac{1}{b}\cdot\tfrac{\beta}{2}N, \Sigma^{-1}) = {}_0F_0^{(\beta)}(-\tfrac{\beta}{2}N, \Sigma^{-1}).$$

Using also (2.3), the density of $\Lambda = \mathrm{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_m) = \lim_{b\to\infty} N$ becomes

$$\frac{\pi^{\frac{m(m-1)}{2}\beta}\left(\Gamma\left(\frac{\beta}{2}\right)\right)^m\left(\frac{\beta}{2}\right)^{ma}}{m!\Gamma_m^{(\beta)}\left(\frac{m}{2}\beta\right)\Gamma_m^{(\beta)}(a)} \cdot |\Sigma|^{-a}|\Lambda|^{a-r}\ {}_0F_0^{(\beta)}(-\tfrac{\beta}{2}\Lambda, \Sigma^{-1})\mathrm{d}\mu(\Lambda),$$

which is joint eigenvalue density of a Wishart matrix with $n = \frac{2a}{\beta}$ degrees of freedom and covariance Σ [5]. □

Using the above result we immediately obtain the density of the largest eigenvalue of the Wishart matrix as a limit of that of a MANOVA matrix.

Theorem 3.1. *The density of the largest eigenvalue of an $m \times m$ Wishart matrix with $n = \frac{2a}{\beta}$ degrees of freedom and covariance Σ is*

$$\frac{ma\beta}{2} \cdot \frac{\Gamma_m^{(\beta)}(r)}{\Gamma_m^{(\beta)}(a+r)}|\Sigma|^{-a}e^{-\mathrm{tr}(\frac{x\beta}{2}\Sigma^{-1})}\left(\tfrac{x\beta}{2}\right)^{ma-1}{}_2F_2^{(\beta)}\left(r+\tfrac{\beta}{2}, \tfrac{m-1}{2}\beta; a+r, \tfrac{m}{2}\beta; \tfrac{x\beta}{2}\Sigma^{-1}\right),$$

where $r \equiv \frac{m-1}{2}\beta + 1$.

Proof. We utilize Proposition 3.1 above. The density of the maximum eigenvalue of an $m \times m$ MANOVA matrix with parameters a, b and covariance Σ^{-1} is [9]:

$$ma D_m(a,b)|\Sigma|^{-a}|I_m + \tfrac{x}{1-x}\Sigma^{-1}|^{-a-b}\big(\tfrac{x}{1-x}\big)^{ma-1}(1-x)^{-2} \\ \times {}_3F_2^{(\beta)}\big(a+b, r+\tfrac{\beta}{2}, r-1; a+r, \tfrac{m}{2}\beta; (\tfrac{1-x}{x}\Sigma + I_m)^{-1}\big),$$

where

$$D_m^{(\beta)}(a,b) \equiv \frac{\Gamma_m^{(\beta)}(a+b)\Gamma_m^{(\beta)}(r)}{\Gamma_m^{(\beta)}(a+r)\Gamma_m^{(\beta)}(b)}.$$

We change variables $\frac{2b}{\beta}\cdot\frac{x}{1-x} \to x$. The Jacobian is $\frac{\beta(1-x)^2}{2b}$ and the density becomes

$$ma D_m(a,b)|\Sigma|^{-a}|I_m + \tfrac{x\beta}{2b}\Sigma^{-1}|^{-a-b}\big(\tfrac{x\beta}{2b}\big)^{ma-1}\tfrac{\beta}{2b} \\ \times {}_3F_2^{(\beta)}\big(a+b, r+\tfrac{\beta}{2}, r-1; a+r, \tfrac{m}{2}\beta; \big(\tfrac{2b}{x\beta}\Sigma + I_m\big)^{-1}\big)$$

which equals

$$ma\frac{\Gamma_m^{(\beta)}(a+b)\Gamma_m^{(\beta)}(r)}{\Gamma_m^{(\beta)}(a+r)\Gamma_m^{(\beta)}(b)b^{ma}}|\Sigma|^{-a}|I_m + \tfrac{x\beta}{2b}\Sigma^{-1}|^{-a}|I_m + \tfrac{x\beta}{2b}\Sigma^{-1}|^{-b}\big(\tfrac{x\beta}{2}\big)^{ma-1}\tfrac{\beta}{2} \\ \times {}_3F_2^{(\beta)}\big(a+b, r+\tfrac{\beta}{2}, r-1; a+r, \tfrac{m}{2}\beta; \tfrac{1}{b}\big(\tfrac{2}{x\beta}\Sigma + \tfrac{1}{b}I_m\big)^{-1}\big).$$

Using again that $\lim_{b\to\infty}\frac{(a+b)_\kappa^{(\beta)}}{b^{|\kappa|}} = 1$,

$$\lim_{b\to\infty} {}_3F_2^{(\beta)}\big(a+b, r+\tfrac{\beta}{2}, r-1; a+r, \tfrac{m}{2}\beta; \tfrac{1}{b}\big(\tfrac{2}{x\beta}\Sigma + \tfrac{1}{b}I_m\big)^{-1}\big) \\ = {}_2F_2^{(\beta)}\big(r+\tfrac{\beta}{2}, r-1; a+r, \tfrac{m}{2}\beta; \tfrac{x\beta}{2}\Sigma^{-1}\big).$$

Also, using (2.3) and

$$\lim_{b\to\infty}|I_m + \tfrac{x\beta}{2b}\Sigma^{-1}|^{-b} = e^{-\mathrm{tr}\left(\frac{x\beta}{2}\Sigma^{-1}\right)}$$

the density becomes

$$\frac{ma\beta}{2}\cdot\frac{\Gamma_m^{(\beta)}(r)}{\Gamma_m^{(\beta)}(a+r)}|\Sigma|^{-a}e^{-\mathrm{tr}\left(\frac{x\beta}{2}\Sigma^{-1}\right)}\big(\tfrac{x\beta}{2}\big)^{ma-1}{}_2F_2^{(\beta)}\big(r+\tfrac{\beta}{2}, \tfrac{m-1}{2}\beta; a+r, \tfrac{m}{2}\beta; \tfrac{x\beta}{2}\Sigma^{-1}\big).$$

as claimed. □

Computationally, the significance of the above result is that the parameter $\frac{m-1}{2}\beta$ implies that the ${}_2F_2$ hypergeometric function need only be summed over partitions of $m-1$ parts instead of m. Since the number of partitions of M in not more than k parts for $M \gg k$ grows as $O(M^k)$, we can compute the density of the largest eigenvalue of a Wishart matrix about $O(M)$ times faster than a direct differentiation of the expression for the distribution from [6]:

$$P[\lambda_{\max} < x] = \frac{\Gamma_m^{(\beta)}(r)}{\Gamma_m^{(\beta)}(a+r)}\cdot\frac{\left|\frac{x}{2}\beta\Sigma^{-1}\right|^a}{e^{\mathrm{tr}(\frac{x}{2}\beta\Sigma^{-1})}}{}_1F_1^{(\beta)}(r; a+r; \tfrac{x}{2}\beta\Sigma^{-1}). \tag{3.1}$$

Another advantage of the density expression of Theorem 3.1 is that it allows us to write the density of $\lambda_{\max}$ as an infinite mixture of chi-squared densities. To see that, let $\tilde{\Sigma} = \Sigma^{-1}/\mathrm{tr}(\Sigma^{-1})$ and the density of $\lambda_{\max}$ becomes

$$f_{\lambda_{\max}}(x) = ma \cdot \frac{\Gamma_m^{(\beta)}(r)}{\Gamma_m^{(\beta)}(a+r)} |\tilde{\Sigma}|^a \sum_{k=0}^{\infty} \beta \mathrm{tr}(\Sigma^{-1}) f_{2(ma+k)}(\beta \mathrm{tr}(\Sigma^{-1})x) \Gamma(ma+k)$$
$$\times \sum_{\kappa \vdash k} \frac{\left(r+\frac{\beta}{2}\right)_\kappa^{(\beta)} \left(\frac{m-1}{2}\beta\right)_\kappa^{(\beta)}}{(a+r)_\kappa^{(\beta)} \left(\frac{m}{2}\beta\right)_\kappa^{(\beta)}} \frac{C_\kappa^{(\beta)}(\tilde{\Sigma})}{k!},$$

where

$$f_\nu(z) = \frac{e^{-\frac{z}{2}}}{2\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{z}{2}\right)^{\frac{\nu}{2}-1}$$

is the density of a chi-squared random variable with ν degrees of freedom. Now, let

$$w_k = ma \cdot \frac{\Gamma_m^{(\beta)}(r)}{\Gamma_m^{(\beta)}(a+r)} |\tilde{\Sigma}|^a \Gamma(ma+k) \sum_{\kappa \vdash k} \frac{\left(r+\frac{\beta}{2}\right)_\kappa^{(\beta)} \left(\frac{m-1}{2}\beta\right)_\kappa^{(\beta)}}{(a+r)_\kappa^{(\beta)} \left(\frac{m}{2}\beta\right)_\kappa^{(\beta)}} \frac{C_\kappa^{(\beta)}(\tilde{\Sigma})}{k!}$$

and we express the density of $\lambda_{\max}$ as

$$f_{\lambda_{\max}}(x) = \beta \mathrm{tr}(\Sigma^{-1}) \sum_{k=0}^{\infty} w_k f_{2(ma+k)}(\beta \mathrm{tr}(\Sigma^{-1})x).$$

Since the w_k's are independent of x, we only need to compute them once for all values of x. In addition, integrating the above density expression gives us an alternative expression for the cumulative distribution of $\lambda_{\max}$ as

$$P[\lambda_{\max} < x] = \sum_{k=0}^{\infty} w_k F_{2(ma+k)}(\beta \mathrm{tr}(\Sigma^{-1})x),$$

where $F_\nu(z)$ is the distribution of a chi-squared random variable with ν degrees of freedom. By taking limit of $x \to \infty$ on both sides, we see that $\sum_{k=0}^{\infty} w_k = 1$. As w_k only involves partitions of $m-1$ parts, this is numerically more efficient than the expression (3.1), which involves partitions of m parts.

Finally, the infinite mixture of chi-squared representation of the density of $\lambda_{\max}$ also allows us to obtain an explicit expression for the moment of $\lambda_{\max}$:

$$\mathrm{E}[\lambda_{\max}^p] = \left[\frac{2}{\beta \mathrm{tr}(\Sigma^{-1})}\right]^p \sum_{k=0}^{\infty} w_k (ma+k)_p.$$

4. The trace of a Wishart matrix

We present new expressions for the density and distribution of the trace of a Wishart matrix that can be computed in time that is linear in the size of the matrix and the degree of the truncation of its series expansion. This complexity is optimal and is an exponential improvement over the previous result for the trace from [6].

Theorem 4.1. *For a Wishart matrix $A \sim W_m^{(\beta)}(n, \Sigma)$, the distribution and density of its trace for an arbitrary z and $a \equiv \frac{n\beta}{2}$ are*

$$P[\mathrm{tr}A \le x] = \left|\tfrac{x\beta}{2}\Sigma^{-1}\right|^a \frac{1}{\Gamma(ma+1)}\, {}_1F_1^{(2a)}\left(a; ma+1; -\tfrac{x\beta}{2}\Sigma^{-1}\right), \tag{4.1}$$

$$f_{\mathrm{tr}A}(x) = \left|\tfrac{x\beta}{2}\Sigma^{-1}\right|^a \frac{e^{-z}}{x\Gamma(ma)}\, {}_1F_1^{(2a)}\left(a; ma; zI_m - \tfrac{x\beta}{2}\Sigma^{-1}\right). \tag{4.2}$$

Proof. The expression for the density of the trace from [6] is

$$f_{\mathrm{tr}(A)}(x) = \left|\tfrac{x\beta}{2}\Sigma^{-1}\right|^a e^{-\frac{x\beta}{2\lambda}} \frac{1}{x} \sum_{k=0}^{\infty} \frac{\left(\frac{x\beta}{2\lambda}\right)^k}{\Gamma(ma+k)} \sum_{\kappa \vdash k} \frac{(a)_\kappa^{(\beta)}}{k!} C_\kappa^{(\beta)}(I_m - \lambda\Sigma^{-1}), \tag{4.3}$$

where $\lambda > 0$ is arbitrary.

From the identity (2.5) we see that for an arbitrary matrix X and a scalar y, the value of ${}_1F_0^{(\beta)}(a; yX)$ does not depend on β. Using also the fact that $(a)_\kappa^{(2a)} = 0$ for partitions κ in more than one part,

$$\sum_{k=0}^{\infty} \sum_{\kappa \vdash k} \frac{y^k}{k!} (a)_\kappa^{(\beta)} C_\kappa^{(\beta)}(X) = {}_1F_0^{(\beta)}(a; yX) = {}_1F_0^{(2a)}(a; yX) = \sum_{k=0}^{\infty} \frac{y^k}{k!} (a)_k\, C_k^{(2a)}(X).$$

By comparing the coefficients in front of y^k, we obtain

$$\sum_{\kappa \vdash k} \frac{(a)_\kappa^{(\beta)}}{k!} C_\kappa^{(\beta)}(X) = \frac{(a)_k}{k!} C_k^{(2a)}(X).$$

In particular, the above is true for $X = I_m - \lambda\Sigma^{-1}$. Thus

$$\begin{aligned} f_{\mathrm{tr}(A)}(x) &= \left|\tfrac{x}{2}\beta\Sigma^{-1}\right|^a e^{-\frac{x\beta}{2\lambda}} \frac{1}{x} \sum_{k=0}^{\infty} \frac{1}{\Gamma(ma+k)} \left(\tfrac{x\beta}{2\lambda}\right)^k \frac{(a)_k}{k!} C_\kappa^{(2a)}(I_m - \lambda\Sigma^{-1}) & (4.4)\\ &= \left|\tfrac{x}{2}\beta\Sigma^{-1}\right|^a \frac{e^{-\frac{x\beta}{2\lambda}}}{x\Gamma(ma)} \sum_{k=0}^{\infty} \frac{(a)_k}{k!(ma)_k} \left(\tfrac{x\beta}{2\lambda}\right)^k C_\kappa^{(2a)}(I_m - \lambda\Sigma^{-1}) \\ &= \left|\tfrac{x}{2}\beta\Sigma^{-1}\right|^a \frac{e^{-\frac{x\beta}{2\lambda}}}{x\Gamma(ma)}\ {}_1F_1^{(2a)}\left(a; ma; \tfrac{x\beta}{2\lambda} I_m - \tfrac{x\beta}{2}\Sigma^{-1}\right) \\ &= \left|\tfrac{x}{2}\beta\Sigma^{-1}\right|^a \frac{e^{-z}}{x\Gamma(ma)}\ {}_1F_1^{(2a)}\left(a; ma; zI_m - \tfrac{x\beta}{2}\Sigma^{-1}\right), & (4.5) \end{aligned}$$

where $z \equiv \frac{x\beta}{2\lambda}$. Because $\lambda > 0$ we have $z > 0$, but otherwise arbitrary. Since the ${}_1F_1^{(2a)}$ function in (4.5) converges for any value of the matrix argument and is defined for all z, (4.2) is true for any value of z (this fact is also implied by Corollary 2.3 in [3]). Setting $z = 0$ and integrating (4.5), we get (4.1). □

The expression (4.2) also allows us to write the density of tr$A$ as a mixture of chi-squared densities. From (4.4)

$$f_{\mathrm{tr}A}(x) = \left|\lambda\Sigma^{-1}\right|^a \sum_{k=0}^{\infty} \left(\frac{x\beta}{2\lambda}\right)^{ma+k} \frac{e^{-\frac{x\beta}{2\lambda}}}{x\Gamma(ma+k)} \frac{(a)_k}{k!} C_k^{(2a)}(I_m - \lambda\Sigma^{-1})$$
$$= \frac{\beta}{\lambda}|\lambda\Sigma^{-1}|^a \sum_{k=0}^{\infty} d_k f_{2(ma+k)}\left(\frac{x\beta}{\lambda}\right), \tag{4.6}$$

where $\lambda > 0$ and

$$d_k \equiv \frac{(a)_k}{k!} C_k^{(2a)}(I_m - \lambda\Sigma^{-1}).$$

Integrating this with respect to x allows us to get an alternative expression of the distribution of $\mathrm{tr}A$ as

$$P[\mathrm{tr}A \le x] = |\lambda\Sigma^{-1}|^a \sum_{k=0}^{\infty} d_k F_{2(ma+k)}\left(\frac{x\beta}{\lambda}\right). \tag{4.7}$$

Although (4.6) and (4.7) are valid for any positive λ, the speed of convergence and numerical stability of these expressions could crucially depend on the choice of λ. For numerical stability, we can choose $\lambda = \sigma_m$, where σ_m is the smallest eigenvalue of Σ. This ensures that $I_m - \lambda\Sigma^{-1}$ has at most $m-1$ nonzero eigenvalues all of which are positive. This choice eliminates all subtractions meaning the density and distribution will be computed to high relative accuracy in floating point arithmetic.

5. Computing the Jack function of partition in one part in $O(m)$ time

The algorithm of [10] can be used to compute $C_k^{(2a)}(B)$, $B = \mathrm{diag}(b_1, b_2, \ldots, b_m)$, in $O(km)$ time. Now we present an alternative algorithm, based on an idea in [1], which only costs $O(m)$. This is optimal since each Jack function depends on m variables.

Define d_k via

$$D(t) = |I_m - tB|^{-a} = {}_1F_0^{(2a)}(a; tB) = \sum_{k=0}^{\infty} \frac{(a)_k t^k}{k!} C_k^{(2a)}(B) = \sum_{k=0}^{\infty} d_k t^k,$$

i.e., $d_k = \frac{(a)_k}{k!} C_k^{(2a)}(B)$. Let

$$P(t) \equiv \sum_{i=1}^{m} \frac{tb_i}{1 - tb_i} = \sum_{k=1}^{\infty} p_k t^k,$$

where $p_k = \mathrm{tr}(B^k) = \sum_{i=1}^m b_i^k$. Differentiating $D(t)$,

$$tD'(t) = aD(t)P(t).$$

Comparing the coefficients of t^k on both sides, we obtain

$$kd_k = a\sum_{r=1}^{k} p_r d_{k-r} = a\sum_{r=1}^{k}\sum_{i=1}^{m} b_i^r d_{k-r} = a\sum_{i=1}^{m}\sum_{r=1}^{k} d_{k-r} b_i^r = a\sum_{i=1}^{m} q_{ki},$$

where

$$q_{ki} \equiv \sum_{r=1}^{k} d_{k-r} b_i^r$$

is a polynomial in b_i. For q_{ki} we thus have the following recurrence relation

$$q_{ki} = b_i \left(d_{k-1} + \sum_{r=2}^{k} d_{k-r} b_i^{r-1} \right) = b_i (d_{k-1} + q_{k-1,i}).$$

With the initial conditions of $d_0 = 1$ and $q_{0i} = 0$ for $i = 1, 2, \ldots, m$, d_k can be obtained from

$$q_{ki} = (d_{k-1} + q_{k-1,i}) b_i, \quad i = 1, \ldots, m,$$
$$d_k = \frac{a}{k} \sum_{i=1}^{m} q_{ki}.$$

The cost of computing d_k is thus $O(m)$, which is optimal, and is exactly what we need to compute the density and distribution of the trace.

6. Numerical experiments

We performed extensive numerical tests to verify the correctness of the formulas in this paper and present four examples in Figure 1.

Acknowledgments

This research was supported by the National Science Foundation Grant DMS–1016086 and by the Woodward Fund for Applied Mathematics at San Jose State University. The Woodward Fund is a gift from the estate of Mrs. Marie Woodward in memory of her son, Henry Teynham Woodward. He was an alumnus of the Mathematics Department at San Jose State University and worked with research groups at NASA Ames.

References

[1] Robert Brown, *The distribution function of positive definite quadratic forms in normal random variables*, SIAM J. Sci. Stat. Comput. **7** (1996), no. 4, 689–695.
[2] James Demmel and Plamen Koev, *Accurate and efficient evaluation of Schur and Jack functions*, Math. Comp. **75** (2006), no. 253, 223–239.
[3] Patrick Desrosiers and Dang-Zheng Liu, *Asymptotics for products of characteristic polynomials in classical β-ensembles*, Constr. Approx. **39** (2014), no. 2, 273–322.
[4] Alexander Dubbs and Alan Edelman, *The Beta-MANOVA ensemble with general covariance*, Random Matrices: Theory and Applications **03** (2014), 1450002.
[5] Alexander Dubbs, Alan Edelman, Plamen Koev, and Praveen Venkataramana, *The Beta–Wishart ensemble*, J. Math. Phys. **54** (2013), 083507.
[6] Alan Edelman and Plamen Koev, *Eigenvalue distributions of Beta–Wishart matrices*, Random Matrices: Theory and Applications **03** (2014), 1450009.

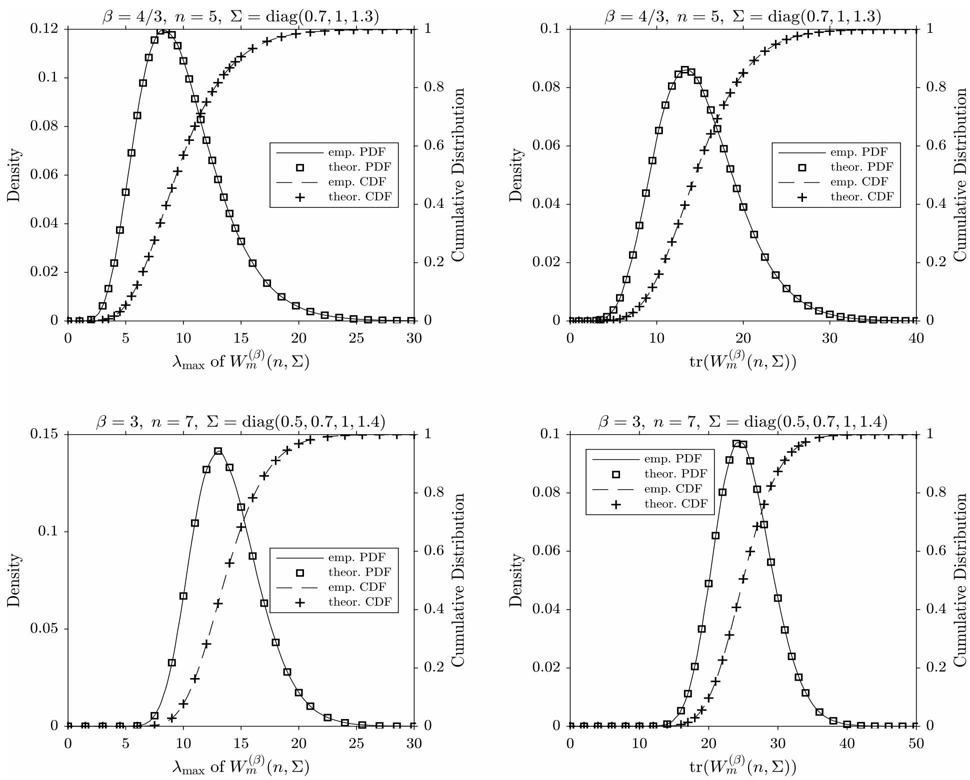

Fig. 1. Numerical experiments comparing the theoretical predictions of the eigenvalue distributions for various ensembles vs. the empirical results.

[7] Peter J. Forrester, *Log-gases and random matrices*, London Mathematical Society Monographs Series, vol. 34, Princeton University Press, Princeton, NJ, 2010.
[8] ———, *Probability densities and distributions for spiked and general variance Wishart β-ensembles*, Random Matrices: Theory and Applications **2** (2013), no. 4, 1350011.
[9] Raymond Kan and Plamen Koev, *Densities of the extreme eigenvalues of Beta–MANOVA matrices*, Random Matrices Theory Appl. **8** (2019), no. 1, 1950002, 10.
[10] Plamen Koev and Alan Edelman, *The efficient evaluation of the hypergeometric function of a matrix argument*, Math. Comp. **75** (2006), no. 254, 833–846.
[11] R. J. Muirhead, *Aspects of multivariate statistical theory*, John Wiley & Sons Inc., New York, 1982.
[12] Atle Selberg, *Remarks on a multiple integral*, Norsk Mat. Tidsskr. **26** (1944), 71–78.
[13] Richard P. Stanley, *Some combinatorial properties of Jack symmetric functions*, Adv. Math. **77** (1989), no. 1, 76–115.