Random Matrices: Theory and Applications

Densities of the Extreme Eigenvalues of Beta–MANOVA Matrices

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 (May 30, 2018)
Revised (Day Month Year)

Keywords: MANOVA; random matrix; eigenvalue; hypergeometric function of a matrix argument

Mathematics Subject Classification 2000: 15A52, 60E05, 62H10, 65F15
We derive new expressions for the densities of extreme eigenvalues of a Beta–MANOVA matrix, which generalize the classical result of Khatri for the real case. We also present new expressions for the distributions of those eigenvalues which are valid for any values of the parameters.

1. Introduction

The Beta–MANOVA random matrix ensemble was introduced recently by Dubbs and Edelman in [2]. It generalizes the classical ($\beta = 1, 2, 4$) MANOVA models to any $\beta > 0$: For a given covariance Σ, if A is $m \times m$ Beta–Wishart [3] with n degrees of freedom and covariance Σ and G is Beta–Wishart with p degrees of freedom and covariance A^{-1}, then $M \equiv (I_m + G^{-1})^{-1}$ is Beta–MANOVA with parameters n, p and covariance Σ. In this paper we use parameters $a = \frac{p\beta}{2}$ and $b = \frac{n\beta}{2}$ instead of n and p making the presentation consistent with Beta–Jacobi [4] and simplifying most expressions.

In [2] the authors derive two expressions for the distribution of the largest eigenvalue $P(\lambda_{\max}(M) < x)$. The first, (2.7) below, is not valid for all $x \in [0, 1]$ (and for example, for identity covariance it is only valid for $x \in [0, \frac{1}{2}]$). The second, (2.9) below, is only valid when certain parameters are nonnegative integers. We utilize a simple formula to derive an expression that is valid for all values of the parameters and all $x \in [0, 1]$. This new expression is readily computable with existing software.

Our main result is a new expression for the *density* of the largest eigenvalue, which generalizes to any $\beta > 0$ the classical result of Khatri [8] for the real, $\beta = 1$, case. From a computational standpoint this expression has the attractive property that, for an $m \times m$ matrix, the series expansion of the hypergeometric function of a matrix argument be summed over partitions of only $m-1$ parts instead of m. In the process we obtain the density ((3.6) below) and the distribution ((3.8) below) of the largest eigenvalue of G in the Beta–MANOVA construction.

Because the largest and the smallest eigenvalues are simply related, we immediately obtain the density and distribution of the smallest eigenvalue as well.

We present preliminaries on partitions, Jack functions, the hypergeometric function of a matrix argument, and the Beta–MANOVA ensemble in section 2. Our main result on the density of the largest eigenvalue is in section 3. Finally, we present numerical experiments in section 4.

2. Preliminaries

We follow Dubbs and Edelman [2] and introduce the notions pertinent to the Beta–MANOVA ensemble.

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

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

where $(a)_k = a(a+1)\cdots(a+k-1)$ is the *rising factorial.*

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

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

For a matrix $X = \operatorname{diag}(x_1, x_2, \ldots, x_m)$, 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. We refer to [5,7] for a detailed treatment and properties.

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

$${}_pF^{(\beta)}_q(a_1, \ldots, a_p; b_1, \ldots, b_q; X, Y) \equiv \sum_{k=0}^{\infty} \sum_{\kappa \vdash k} \frac{1}{k!} \cdot \frac{(a_1)^{(\beta)}_\kappa \cdots (a_p)^{(\beta)}_\kappa}{(b_1)^{(\beta)}_\kappa \cdots (b_q)^{(\beta)}_\kappa} \cdot \frac{C^{(\beta)}_\kappa(X) C^{(\beta)}_\kappa(Y)}{C^{(\beta)}_\kappa(I_m)}. \tag{2.3}$$

For *one matrix argument*,

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

Definition 2.1. An $m \times m$ matrix M is *Beta–MANOVA* of parameters $a, b > r-1$ for $\beta > 0$ if the joint density of its eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ is:

$$\frac{1}{S_m^{(\beta)}(a,b)} |\Sigma|^a |\Lambda|^{a-r} |I_m - \Lambda|^{-a-r} \ {}_1F_0^{(\beta)}(a+b; \Lambda(\Lambda - I_m)^{-1}, \Sigma) \mathrm{d}\mu(\Lambda), \tag{2.4}$$

where $\Lambda = \mathrm{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_m)$,

$$r \equiv \frac{m-1}{2}\beta + 1, \tag{2.5}$$

$|A|$ denotes the determinant of a matrix A, $S_m^{(\beta)}(a,b)$ is the value of the *Selberg Integral* [11]:

$$S_m^{(\beta)}(a,b) \equiv \frac{m!\Gamma_m^{(\beta)}\left(\frac{m}{2}\beta\right)}{\pi^{\frac{m(m-1)}{2}\beta}\left(\Gamma\left(\frac{\beta}{2}\right)\right)^m} \cdot \frac{\Gamma_m^{(\beta)}(a)\,\Gamma_m^{(\beta)}(b)}{\Gamma_m^{(\beta)}(a+b)}, \tag{2.6}$$

and the Vandermonde determinant is incorporated into a new measure to prevent it from appearing everywhere

$$\mathrm{d}\mu(\Lambda) = \prod_{i<j} |\lambda_i - \lambda_j|^\beta \mathrm{d}\lambda_1 \mathrm{d}\lambda_2 \cdots \mathrm{d}\lambda_m.$$

The joint density (2.4) is positive by Liu [10, Prop. 2.3] (since $0 < \lambda_i < 1, i = 1, 2, \ldots, m$, and the covariance Σ is positive definite). Liu also establishes an alternative expression for it [10, Prop. 2.4] (based on (3.3) below), which reduces directly to the one for the β-Jacobi ensemble, [4, (2.1)], when $\Sigma = I$.

For the largest eigenvalue we have from [2, (3.6), (1.1)],

$$P(\lambda_{\max}(M) < x) = |\Sigma|^a D_m^{(\beta)}(a,b) \left(\tfrac{x}{1-x}\right)^{ma} {}_2F_1^{(\beta)}\left(a+b, a; a+r; -\tfrac{x}{1-x}\Sigma\right), \tag{2.7}$$

where

$$D_m^{(\beta)}(a,b) \equiv \frac{\Gamma_m^{(\beta)}(a+b)\Gamma_m^{(\beta)}(r)}{\Gamma_m^{(\beta)}(a+r)\Gamma_m^{(\beta)}(b)}, \tag{2.8}$$

and also, when $t \equiv b - r$ is a nonnegative integer,

$$P(\lambda_{\max}(M) < x) = \left|\tfrac{1-x}{x}\Sigma^{-1} + I_m\right|^{-a} \sum_{k=0}^{mt} \sum_{\kappa \vdash k, \kappa_1 \le t} \frac{1}{k!} (a)_\kappa^{(\beta)} C_\kappa^{(\beta)}\left((I_m + \tfrac{x}{1-x}\Sigma)^{-1}\right). \tag{2.9}$$

Since the hypergeometric function ${}_2F_1^{(\beta)}$ only converges when the eigenvalues of the matrix argument do not exceed 1 in absolute value, the first expression (2.7) only converges when

$$\tfrac{x}{1-x}\sigma_i \le 1, \tag{2.10}$$

where $\sigma_i > 0$, $i = 1, 2, \ldots, m$, are the eigenvalues of Σ. The inequality (2.10) is equivalent to $x \le \min_i \left(\frac{1}{1+\sigma_i}\right)$. For $\Sigma = I_m$, for example, it becomes $x \le \frac{1}{2}$, which is unfortunate since the largest eigenvalue appears to take values mostly in $[\frac{1}{2}, 1]$ in our numerical tests.

On the other side, the second expression, (2.9), is only valid when $b-r$ is a nonnegative integer, which restricts its usefulness.

As we will see in the next section, the formula (3.2) yields an expression for the distribution of the largest eigenvalue, which is valid for all values of the parameters and all $x \in [0,1]$.

3. New results

Before we present our main theorem we need several identities. Throughout this section r is defined as in (2.5).

Kadell's generalization of Selberg's integral [6] is

$$\int_{[0,1]^m} C_\kappa^{(\beta)}(X)|X|^{a-r}|I_m - X|^{b-a-r}\mathrm{d}\mu(X) = \frac{(a)_\kappa^{(\beta)}}{(b)_\kappa^{(\beta)}} S_m^{(\beta)}(a, b-a) C_\kappa^{(\beta)}(I_m). \quad (3.1)$$

We will also utilize the formulas [5, Prop. 13.1.6], [1, (2.15)]

$${}_2F_1^{(\beta)}(a,b;c;X) = {}_2F_1^{(\beta)}\left(c-a,b;c;-X(I_m-X)^{-1}\right) \cdot |I_m - X|^{-b}, \quad (3.2)$$

$${}_1F_0^{(\beta)}(t;-T,\Sigma) = |I_m + x\Sigma|^{-t}\,{}_1F_0^{(\beta)}(t;(\Sigma^{-1}+xI_m)^{-1}, xI_m - T). \quad (3.3)$$

We also simplify the following expression, which we will need in Theorem 3.1.

Lemma 3.1. *With S_m defined as in* (2.6),

$$\frac{S_{m-1}^{(\beta)}(r+\frac{\beta}{2}, a-\frac{\beta}{2})}{S_m^{(\beta)}(a,b)} = a\frac{\Gamma_m^{(\beta)}(a+b)}{\Gamma_m^{(\beta)}(b)} \cdot \frac{\Gamma_m^{(\beta)}(r)}{\Gamma_m^{(\beta)}(a+r)}. \quad (3.4)$$

Proof. We start by noting that from definition (2.2), for $z \geq r-1$,

$$\frac{\Gamma_{m-1}^{(\beta)}\left(z-\frac{\beta}{2}\right)}{\Gamma_m^{(\beta)}(z)} = \frac{1}{\pi^{\frac{m-1}{2}\beta}\Gamma(z)}. \quad (3.5)$$

Then

$$\frac{S_{m-1}^{(\beta)}(r+\frac{\beta}{2}, a-\frac{\beta}{2})}{S_m^{(\beta)}(a,b)}$$

$$= \frac{\dfrac{(m-1)!\Gamma_{m-1}^{(\beta)}\left(\frac{m-1}{2}\beta\right)}{\pi^{\frac{(m-1)(m-2)}{2}\beta}\left(\Gamma\left(\frac{\beta}{2}\right)\right)^{m-1}} \cdot \dfrac{\Gamma_{m-1}^{(\beta)}\left(r+\frac{\beta}{2}\right)\Gamma_{m-1}^{(\beta)}\left(a-\frac{\beta}{2}\right)}{\Gamma_{m-1}^{(\beta)}(a+r)}}{\dfrac{m!\Gamma_m^{(\beta)}\left(\frac{m}{2}\beta\right)}{\pi^{\frac{m(m-1)}{2}\beta}\left(\Gamma\left(\frac{\beta}{2}\right)\right)^{m}} \cdot \dfrac{\Gamma_m^{(\beta)}(a)\Gamma_m^{(\beta)}(b)}{\Gamma_m^{(\beta)}(a+b)}}$$

$$= \frac{\Gamma_m^{(\beta)}(a+b)}{\Gamma_m^{(\beta)}(b)} \cdot \frac{\pi^{\frac{m-1}{2}\beta}\Gamma\left(\frac{\beta}{2}\right)}{m} \cdot \frac{\Gamma_{m-1}^{(\beta)}\left(\frac{m-1}{2}\beta\right)}{\Gamma_m^{(\beta)}\left(\frac{m}{2}\beta\right)} \cdot \frac{\Gamma_{m-1}^{(\beta)}\left(a-\frac{\beta}{2}\right)}{\Gamma_m^{(\beta)}(a)} \cdot \frac{\Gamma_{m-1}^{(\beta)}\left(r+\frac{\beta}{2}\right)}{\Gamma_{m-1}^{(\beta)}(a+r)},$$

which using (3.5) becomes

$$= \frac{\Gamma_m^{(\beta)}(a+b)}{\Gamma_m^{(\beta)}(b)} \cdot \frac{\Gamma\left(\frac{\beta}{2}\right)}{m} \cdot \frac{1}{\Gamma\left(\frac{m}{2}\beta\right)} \cdot \frac{1}{\Gamma(a)} \cdot \frac{\Gamma_{m-1}^{(\beta)}\left(r+\frac{\beta}{2}\right)}{\Gamma_{m-1}^{(\beta)}(a+r)}$$

and using that $\Gamma(z+1) = z\Gamma(z)$ for any z and $\Gamma(1) = 1$,

$$\begin{aligned}
&= \frac{\Gamma_m^{(\beta)}(a+b)}{\Gamma_m^{(\beta)}(b)} \cdot a \cdot \frac{\Gamma\left(\frac{\beta}{2}+1\right)}{\Gamma\left(\frac{m}{2}\beta+1\right)} \cdot \frac{\Gamma(1)}{\Gamma(a+1)} \cdot \prod_{i=1}^{m-1} \frac{\Gamma\left(\frac{m-1}{2}\beta+1+\frac{\beta}{2}-\frac{i-1}{2}\beta\right)}{\Gamma\left(a+\frac{m-1}{2}\beta+1+\frac{i-1}{2}\beta\right)} \\
&= \frac{\Gamma_m^{(\beta)}(a+b)}{\Gamma_m^{(\beta)}(b)} \cdot a \cdot \prod_{i=1}^{m} \frac{\Gamma\left(\frac{m-1}{2}\beta+1-\frac{i-1}{2}\beta\right)}{\Gamma\left(a+\frac{m-1}{2}\beta+1+\frac{i-1}{2}\beta\right)} \\
&= a \cdot \frac{\Gamma_m^{(\beta)}(a+b)}{\Gamma_m^{(\beta)}(b)} \cdot \frac{\Gamma_m^{(\beta)}(r)}{\Gamma_m^{(\beta)}(a+r)}.
\end{aligned}$$

□

Finally, we prove our main result.

Theorem 3.1. *With notation as in Definition 2.1, the density of the largest eigenvalue of an $m \times m$ Beta–MANOVA matrix is*

$$maD_m(a,b)|\Sigma|^{-b}|C|^{a+b}x^{ma-1}(1-x)^{mb-1}{}_3F_2^{(\beta)}\left(a+b, r+\tfrac{\beta}{2}, r-1; a+r, \tfrac{m}{2}\beta; xC\right)\mathrm{d}x,$$

where $C \equiv ((1-x)\Sigma^{-1} + xI_m)^{-1}$.

Proof. We start with the joint density (2.4) of the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$:

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

We need to get rid of the $|I_m - \Lambda|^{-a-r}$ factor so that we can use (3.1) to integrate all but the largest eigenvalue out. We change variables $T = \Lambda(I_m - \Lambda)^{-1}$. The Jacobian is $|I_m - \Lambda|^{-2}$ and $\mathrm{d}\mu(\Lambda) = |I_m - \Lambda|^{(m-1)\beta}\mathrm{d}\mu(T)$, so the density of T is

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

Using (3.3) the above becomes

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

where x is arbitrary. We choose x to be t_m, the largest eigenvalue of T, and change variables $t_i = x_i x$, $i = 1, 2, \ldots, m-1$. Let $X = \mathrm{diag}\,(x_1, x_2, \ldots, x_{m-1})$. The Jacobian is x^{m-1} so the term $|T|^{a-r}\mathrm{d}\mu(T)$ becomes

$$|X|^{a-r}x^{m(a-r)+\frac{m(m-1)}{2}\beta+m-1}\mathrm{d}\mu(X)\,\mathrm{d}x = |X|^{a-r}x^{ma-1}\mathrm{d}\mu(X)\,\mathrm{d}x.$$

Also, from the properties of the Jack function [12], since one of the eigenvalues of $xI_m - T$ is zero, $C_\kappa^{(\beta)}(xI_m - T) = C_\kappa^{(\beta)}(x(I_{m-1} - X))$. The joint density of $x, x_1, x_2, \ldots, x_{m-1}$ is thus

$$\frac{1}{S_m^{(\beta)}(a,b)}|\Sigma|^a|I_m+x\Sigma|^{-a-b}x^{ma-1}|X|^{a-r}|I_{m-1}-X|^{\beta} \\ \times {}_1F_0^{(\beta)}(a+b; x(\Sigma^{-1}+xI_m)^{-1}, I_{m-1}-X)\mathrm{d}\mu(X)\mathrm{d}x.$$

Expanding the hypergeometric function this is

$$\frac{1}{S_m^{(\beta)}(a,b)}|\Sigma|^a|I_m+x\Sigma|^{-a-b}x^{ma-1}|X|^{a-r}|I_{m-1}-X|^\beta \times\sum_{k=0}^{\infty}\sum_{\kappa\vdash k}\frac{(a+b)_\kappa^{(\beta)}}{k!}\cdot\frac{C_\kappa^{(\beta)}(x(\Sigma^{-1}+xI_m)^{-1})C_\kappa^{(\beta)}(I_{m-1}-X)}{C_\kappa^{(\beta)}(I_m)}.$$

In order to integrate X out of the joint density, we must evaluate the integral

$$\int_{[0,1]^{m-1}}|X|^{a-r}|I_{m-1}-X|^\beta C_\kappa^{(\beta)}(I_{m-1}-X)\mathrm{d}\mu(X),$$

which we achieve by changing variables $x_i\to 1-x_i$ and using (3.1):

$$\int_{[0,1]^{m-1}}|X|^\beta|I_{m-1}-X|^{a-r}C_\kappa^{(\beta)}(X)\mathrm{d}\mu(X)=\frac{(r+\frac{\beta}{2})_\kappa^{(\beta)}}{(a+r)_\kappa^{(\beta)}}S_{m-1}^{(\beta)}(r+\tfrac{\beta}{2},a-\tfrac{\beta}{2})C_\kappa^{(\beta)}(I_{m-1}).$$

Going back, for the density of x we have

$$m\frac{S_{m-1}^{(\beta)}(r+\frac{\beta}{2},a-\frac{\beta}{2})}{S_m^{(\beta)}(a,b)}|\Sigma|^a|I_m+x\Sigma|^{-a-b}x^{ma-1} \times\sum_{k=0}^{\infty}\sum_{\kappa\vdash k}\frac{1}{k!}(a+b)_\kappa^{(\beta)}\cdot C_\kappa^{(\beta)}(x(\Sigma^{-1}+xI_m)^{-1})\frac{(r+\frac{\beta}{2})_\kappa^{(\beta)}}{(a+r)_\kappa^{(\beta)}}\cdot\frac{C_\kappa^{(\beta)}(I_{m-1})}{C_\kappa^{(\beta)}(I_m)}.$$

The factor of m appears because the eigenvalues are unordered and the largest eigenvalue can be any of $t_1,t_2,\ldots,t_m$. Now, from Stanley [12, Theorem 5.4],

$$\frac{C_\kappa^{(\beta)}(I_{m-1})}{C_\kappa^{(\beta)}(I_m)}=\frac{(\frac{m-1}{2}\beta)_\kappa^{(\beta)}}{(\frac{m}{2}\beta)_\kappa^{(\beta)}}.$$

Using (3.4) and (2.8), we obtain the density of the largest eigenvalue of T to be

$$\begin{aligned}&maD_m(a,b)|\Sigma|^a|I_m+x\Sigma|^{-a-b}x^{ma-1}\\&\qquad\times\sum_{k=0}^{\infty}\sum_{\kappa\vdash k}\frac{1}{k!}(a+b)_\kappa^{(\beta)}\frac{(r+\frac{\beta}{2})_\kappa^{(\beta)}(\frac{m-1}{2}\beta)_\kappa^{(\beta)}}{(\frac{m}{2}\beta)_\kappa^{(\beta)}(a+r)_\kappa^{(\beta)}}C_\kappa^{(\beta)}(x(\Sigma^{-1}+xI_m)^{-1})\\&=maD_m(a,b)|\Sigma|^a|I_m+x\Sigma|^{-a-b}x^{ma-1}\\&\qquad\times{}_3F_2^{(\beta)}(a+b,r+\tfrac{\beta}{2},\tfrac{m-1}{2}\beta;a+r,\tfrac{m}{2}\beta;x(\Sigma^{-1}+xI_m)^{-1}).\end{aligned}\tag{3.6}$$

Going back to the largest eigenvalue of $\Lambda=(I_m+T^{-1})^{-1}$, the change of variables is $x\to\frac{x}{1-x}$ with Jacobian $(1-x)^{-2}$, so its density is

$$\begin{aligned}&maD_m(a,b)|\Sigma|^a|I_m+\tfrac{x}{1-x}\Sigma|^{-a-b}(\tfrac{x}{1-x})^{ma-1}(1-x)^{-2}\\&\qquad\times{}_3F_2^{(\beta)}\big(a+b,r+\tfrac{\beta}{2},\tfrac{m-1}{2}\beta;a+r,\tfrac{m}{2}\beta;x((1-x)\Sigma^{-1}+xI_m)^{-1}\big)\mathrm{d}x,\end{aligned}$$

and the claim follows. □

Since $\left(\frac{m-1}{2}\beta\right)_\kappa^{(\beta)} = 0$ for partitions of more than $m-1$ parts (see (2.1)), the series for ${}_3F_2^{(\beta)}$ in the density above needs only be computed for partitions of not more than $m-1$ parts. Since the number of partitions of an integer N in not more than k parts grows roughly as $\mathcal{O}(N^k)$ for $N \gg k$, this results in substantial computational savings.

Looking back at the construction $M = (I_m + G^{-1})^{-1}$ of the Beta–MANOVA matrix M, we obtain expressions for the distributions of the largest eigenvalues of M and G valid for all $x \in [0,1]$ and all $x \geq 0$, respectively, by applying (3.2) to [2, (3.5)] and (2.7).

Proposition 3.1. *With notation as in Definition 2.1, the largest eigenvalues of a Beta–MANOVA matrix M and the matrix $G = (M^{-1} - I_m)^{-1}$ have distributions*

$$P(\lambda_{\max}(M) < x) = D_m^{(\beta)}(a,b)|xC|^a{}_2F_1^{(\beta)}(a, r-b; a+r; xC); \tag{3.7}$$

$$P(\lambda_{\max}(G) < x) = D_m^{(\beta)}(a,b)|xE|^a{}_2F_1^{(\beta)}(a, r-b; a+r; xE). \tag{3.8}$$

where $C \equiv ((1-x)\Sigma^{-1} + xI_m)^{-1}$ and $E \equiv (\Sigma^{-1} + xI_m)^{-1}$. The above expressions are valid for all $x \in [0,1]$ and all $x \geq 0$, respectively, and any $a, b > \frac{m-1}{2}\beta$.

If $\sigma_1, \sigma_2, \ldots, \sigma_m > 0$ are the eigenvalues of Σ, the eigenvalues of the matrix argument xC in (3.7) are $((1-x)(x\sigma_i)^{-1} + 1)^{-1} \in (0,1]$ for all $x \in (0,1]$ and 0 for $x = 0$. The eigenvalues of the matrix argument xE in (3.8) are $(1 + (x\sigma_i)^{-1})^{-1} \in [0,1]$ for all $x > 0$ and 0 for $x = 0$. Thus both hypergeometric functions in (3.7) and (3.8) always converge.

For the smallest eigenvalue, it follows directly from the construction of Dubbs and Edelman in [2] that if M is Beta–MANOVA with parameters a, b and covariance Σ, then $I_m - M$ is Beta–MANOVA with parameters b, a and covariance Σ^{-1}.

Proposition 3.2. *If $F_{a,b,\Sigma}(x)$ is the distribution* (3.7) *of the largest eigenvalue of a Beta–MANOVA random matrix M with parameters $a, b,$ and covariance Σ and $f_{a,b,\Sigma}(x)$ is its density, then for the smallest eigenvalue $\lambda_{\min}$*

$$P(\lambda_{\min}(M) < x) = 1 - F_{b,a,\Sigma^{-1}}(1-x)$$

and its density is $f_{b,a,\Sigma^{-1}}(1-x)$.

4. Numerical experiments

We performed extensive numerical tests to verify the correctness of the expressions for the densities and distributions of the exteme eigenvalues of the MANOVA ensemble using the software `mhg` [9]. We present examples in Figure 1.

Acknowledgments

This research was supported by NSF Grant DMS–1016086 and by the Woodward Fund for Applied Mathematics at San Jose State University. The Woodward Fund

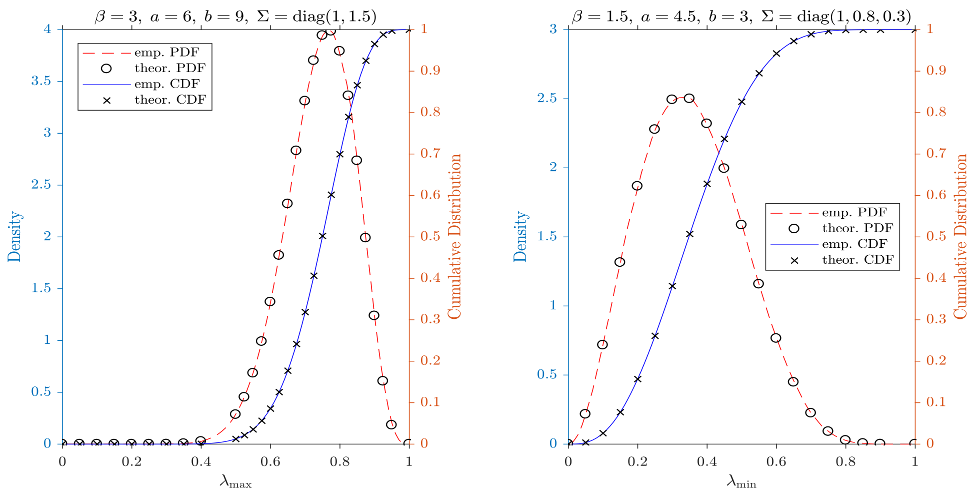

Fig. 1. The theoretical predictions of the extreme eigenvalues vs. the empirical results.

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] Patrick Desrosiers and Dang-Zheng Liu, *Asymptotics for products of characteristic polynomials in classical β-ensembles*, Constr. Approx. **39** (2014), no. 2, 273–322.

[2] Alexander Dubbs and Alan Edelman, *The Beta-MANOVA ensemble with general covariance*, Random Matrices: Theory and Applications **03** (2014), 1450002.

[3] Alexander Dubbs, Alan Edelman, Plamen Koev, and Praveen Venkataramana, *The Beta–Wishart ensemble*, J. Math. Phys. **54** (2013), 083507.

[4] Ioana Dumitriu and Plamen Koev, *The distributions of the extreme eigenvalues of beta–Jacobi random matrices*, SIAM J. Matrix Anal. Appl. **1** (2008), 1–6.

[5] Peter J. Forrester, *Log-gases and random matrices*, London Mathematical Society Monographs Series, vol. 34, Princeton University Press, Princeton, NJ, 2010.

[6] Kevin W. J. Kadell, *The Selberg-Jack symmetric functions*, Adv. Math. **130** (1997), no. 1, 33–102.

[7] Jyoichi Kaneko, *Selberg integrals and hypergeometric functions associated with Jack polynomials*, SIAM J. Math. Anal. **24** (1993), no. 4, 1086–1110.

[8] C. G. Khatri, *Some distribution problems connected with the characteristic roots of $S_1 {S_2}^{-1}$*, Ann. Math. Statist. **38** (1967), 944–948.

[9] Plamen Koev and Alan Edelman, *The efficient evaluation of the hypergeometric function of a matrix argument*, Math. Comp. **75** (2006), no. 254, 833–846.

[10] Dang-Zheng Liu, *Limits for circular Jacobi beta-ensembles*, J. Approx. Theory **215** (2017), 40–67.

[11] Atle Selberg, *Remarks on a multiple integral*, Norsk Mat. Tidsskr. **26** (1944), 71–78.

[12] Richard P. Stanley, *Some combinatorial properties of Jack symmetric functions*, Adv. Math. **77** (1989), no. 1, 76–115.