Computations of the Expected Euler Characteristic for the Largest Eigenvalue of a Real Wishart Matrix

Nobuki Takayama [*], Lin Jiu [†], Satoshi Kuriki [‡], Yi Zhang [§]

March 20, 2019

Abstract

We give an approximate formula of the distribution of the largest eigenvalue of real Wishart matrices by the expected Euler characteristic method for the general dimension. The formula is expressed in terms of a definite integral with parameters. We derive a differential equation satisfied by the integral for the 2×2 matrix case and perform a numerical analysis of it.

1 Introduction

For $i = 1, \dots, n$, let $\xi_i \in \mathbb{R}^{m \times 1}$ be independently distributed as the m-dimensional (real) Gaussian distribution $N_m(\mu_i, \Sigma)$, where μ_i and Σ are the mean vector and the covariance matrix of ξ_i, respectively. The (real) Wishart distribution $W_m(n, \Sigma; \Omega)$ is the probability measure on the cone of $m \times m$ positive semi-definite matrices induced by the random matrix

$$W = \Xi\Xi^\top, \quad \Xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^{m \times n}.$$

Here $\Omega = \Sigma^{-1} \sum_{i=1}^n \mu_i \mu_i^\top$ is non-central parameter matrix. Unless Ω vanishes, this distribution is referred to as the non-central (real) Wishart distribution.

The largest eigenvalue $\lambda_1(W)$ of W is used as a test statistic for testing $\Sigma = I_m$ and/or $\Omega \neq 0$ under the assumption that $\Sigma - I_m$ is positive semi-definite. This test statistic is expected to have a good power when the matrices $\Sigma - I_m$ and Ω are of low rank.

In the setting of testing hypotheses, the distribution of $\lambda_1(W)$ is of particular interest; It corresponds to the power of the test. When $\Omega = 0$, the celebrated works by A. T. James and other authors in the last century show (see e.g., the book by Muirhead [21]) that the

[*]Department of Mathematics, Kobe University, Japan. Email: takayama@math.kobe-u.ac.jp

[†]Department of Mathematics and Statistics, Dalhousie University, Canada. Email: lin.jiu@dal.ca

[‡]The Institute of Statistical Mathematics (ISM), Research Organization of Information and Systems (ROIS), Japan. Supported by JSPS KAKENHI Grant Number 16H02792. Email: kuriki@ism.ac.jp

[§]Corresponding author, Department of Mathematical Sciences, The University of Texas at Dallas (UTD), USA. Supported by the Austrian Science Fund (FWF): P29467-N32 and the UT Dallas Program: P-1-03246. Email: zhangy@amss.ac.cn

cumulative distribution function of $\lambda_1(W)$ can be written as a hypergeometric function of matrix arguments:

$$\Pr(\lambda_1(W) < x) = c_{m,n}\det\left(\frac{1}{2}nx\Sigma^{-1}\right)^{n/2} {}_1F_1\left(\frac{1}{n};\frac{1}{2}(n+m+1);-\frac{1}{2}nx\Sigma^{-1}\right),$$

where $c_{m,n}$ is a known constant [21, Corollary 9.7.2]. It is well-known that the hypergeometric function ${}_1F_1$ has a series expression in the zonal polynomial C_κ with index κ, which is a partition of an integer. However, in view of numerical calculation, this is less useful, because the explicit form of $C_\kappa(X)$ is not known unless the rank of matrix X is 1 or 2. Because of this difficulty, Hashiguchi, et al. [6] recently proposed a holonomic gradient method (HGM) for numerical evaluation, which utilizes a holonomic system of differential equations for computation. However, when $\Omega \neq 0$, the situation is getting worse. The cumulative distribution function $\Pr(\lambda_1(W) < x)$ can not be expressed as a simple series of the zonal polynomial but a series of the Hermite polynomial H_κ defined by the Laplace transform of C_κ:

$$\mathrm{etr}\big(-TT^\top\big)H_\kappa(T) = \frac{(-1)^{|\kappa|}}{\pi^{mn/2}}\int \mathrm{etr}\big(-2iTU^\top\big)\mathrm{etr}\big(-UU^\top\big)C_\kappa\big(UU^\top\big)\,dU, \quad T,U\in\mathbb{R}^{m\times n}$$

([7, Corollary 10]). The Hermite polynomial H_κ can be written as a linear combination of the zonal polynomial C_κ, but the coefficients are not given explicitly [3].

In this paper, instead of the direct calculation approach, we will approximate the distribution function by means of the expected Euler characteristic heuristic or the Euler characteristic method proposed in 2000's by Adler and Tayler [1] or by Kuriki and Takemura [17]. This is a methodology to approximate the tail upper probability of a random field. In our problem, since the square root of the largest eigenvalue $\lambda_1(W)^{1/2}$ is the maximum of a Gaussian field

$$\big\{u^\top \Xi v \mid \|u\|_{\mathbb{R}^m} = \|v\|_{\mathbb{R}^n} = 1\big\},$$

this method actually works for our purpose ([15], [16]). As we show later, the Euler-characteristics method evaluates the quantity

$$\Pr(\lambda_1(W)\geq x) - \Pr(\lambda_2(W)\geq x) + \cdots + (-1)^{m-1}\Pr(\lambda_m(W)\geq x)$$

instead of $\Pr(\lambda_1(W)\geq x)$. This formula approximates $\Pr(\lambda_1(W)\geq x)$ well when x is large, because $\Pr(\lambda_i(W)\geq x)$ $(i\geq 2)$ are negligible when x is large. This is practically sufficient for our purpose, since only the upper tail probability is required in testing hypothesis.

In this paper, we deal with the non-central real Wishart matrix. In the multiple-input multiple-output (MIMO) problem, the non-central complex Wishart matrix also plays an important role. The largest eigenvalue of the non-central complex Wishart is much easier to handle in that case, because the explicit formula for the cumulative distribution is given by Kang and Alouini [12]. The HGM based on Kang and Alouini's formula has been proposed in [5].

The organization of the paper is as follows. In Section 2, we give an integral representation formula of the expectation of the Euler characteristic for random matrices of a general size. In Section 3, we restrict to the case of 2×2 random matrices and study the integral representation derived in Section 2 with the polar coordinate system and study it from numerical point of view. In Section 4, we give a closed formula of the expectation of the Euler characteristic for random matrices of a general size for the central and scalar covariance case. The formula is expressed in terms of the Laguerre polynomial.

By virtue of the theory of holonomic systems (e.g., [8]), the integral representation given in Section 2 satisfies a holonomic system of linear differential equations. However, its explicit form is not known in general. In Section 5, we come back to the case of 2×2 random matrices. We derive a differential equation satisfied by the integral representation of the expectation of the Euler characteristic with a help of computer algebra algorithms, systems and perform a numerical analysis of the differential equation. This gives a new efficient method to numerically evaluate the Euler expectation when the numerical integration is hard to perform.

2 Expectation of an Euler characteristic number

Let $A = (a_{ij})$ be a real $m \times n$ matrix valued random variable (random matrix) with the density

$$p(A)dA, \quad dA = \prod da_{ij}.$$

We assume that $p(A)$ is smooth and $n \geq m \geq 2$. Define a manifold

$$M = \{hg^T \mid g \in S^{m-1}, h \in S \in S^{n-1}\} \simeq S^{m-1} \times S^{n-1}/\sim$$

where $(h, g) \sim (-h, -g)$ and h and g are regarded as column vectors and hg^T is a rank 1 $m \times n$ matrix. Set

$$f(U) = \mathrm{tr}(UA) = g^T Ah, \quad U \in M,$$

and

$$M_x = \{hg^T \in M \mid f(U) = g^T Ah \geq x\}.$$

Proposition 1. *Let A be a random matrix as above. Then the following claims are equivalent.*

1. *The function $f(U)$ has a critical point at $U = hg^T$.*
2. *The vectors g^T, h are left and right eigenvectors of A, respectively. In other words, there exists a constant c such that $g^T A = ch^T$, $Ah = cg$.*

Moreover, the function f takes the value c at the critical point (g, h).

Proof. We assume that the vector $g \in S^{n-1}$ is expressed by a local coordinate u_i, $1 \leq i \leq m-1$ and the vector $h \in S^{m-1}$ is expressed by a coordinate v_j, $1 \leq j \leq n-1$. We denote $\partial/\partial u_i$ by ∂_i and $\partial/\partial v_a$ by ∂_a. Since $g^T g = 1$, we have $g_i^T g = 0$ where $g_i = \partial_i \bullet g$. We will omit $\bullet$, which means the action, as long as no confusion arises. Analogously, we have $h_a^T h_a = 0$, where $h_a = \partial_a h$.

Assume that A is a $m \times n$ (random real) matrix. Let us consider the function $f(U)$ expressed by the local coordinate $(g(u), h(v))$

$$f(g, h) = g^T Ah, \quad g \in S^{n-1}, h \in S^{m-1}. \tag{1}$$

At the critical point of f, we have

$$\partial_i f = g_i A h = 0, \quad \partial_a f = g A h_a = 0.$$

Since it holds for all i and u is a local coordinate of S^{n-1}, g_i's are linearly independent. Therefore, there exists a constant c such that $Ah = cg$ at the critical point. Analogously, we can see that there exists a constant d such that $A^T g = dh$. Let us show $c = d$. We have

$$(g^T A)h = (dh^T)h = d(h^T h) = d$$

and

$$g^T(Ah) = g^T(cg) = c(g^T g) = c.$$

Therefore, we have $d = c = f(f, g)$ at the critical point.

Conversely, $Ah = cg$ and $A^T g = dh$ at a point (u, v) imply that $(g(u), h(v))$ is a critical point of $f(g(u), h(v))$. //

We take a continuous family of elements of $SO(m)$ parametrized by the first column vector g. In other words, we take a continuous family of orthogonal frames of $\mathbb{R}^m$ parametrized by $g \in S^{m-1}$. The element of $SO(m)$ is denoted by $(g, G) \in O(m)$ where G is an $m \times (m-1)$ matrix. Analogously, we take a family $(h, H) \in SO(n)$ parametrized by $h \in S^{n-1}$ where H is an $n \times (n-1)$ matrix parametrized by g. Then, by putting

$$\sigma = g^T A h, \ B = G^T(g) A H(h), \tag{2}$$

the matrix A can be expressed as

$$A = \sigma g h^T + G(g) B H(h)^T, \tag{3}$$

which is, intuitive speaking, a partial singular value decomposition. We denote the set of the $(m-1) \times (n-1)$ matrices by $M(m-1, n-1)$.

This decomposition above gives coordinate systems for the space of matrices A. Let us introduce these coordinate systems. Without loss of generality, we assume that $m \leq n$. We sort the singular values of B by the descending order. We denote by $\lambda_j(B)$ the j-th singular value of the matrix B. For a real number σ, we define

$$\begin{aligned}\mathcal{B}(i, \sigma) \ &= \ \{B \in M(m-1, n-1) \,|\, \text{all the singular values of } B \text{ are different and non-zero.} \\ &\qquad \lambda_j(B) > \sigma \text{ for all } j < i, \ \lambda_j(B) \leq \sigma \text{ for all } j \geq i\}.\end{aligned}$$

Set

$$\mathcal{A} = \{A \in M(m, n) \,|\, \text{all the singular values of } A \text{ are different and non-zero.}\},$$

and

$$\mathcal{A}_i = \{(\sigma, g, h, B) \,|\, \sigma \in \mathbb{R}_{>0}, (g, h) \in S^{m-1} \times S^{n-1}/\sim, B \in \mathcal{B}(i, \sigma)\}.$$

For a matrix A in $\mathcal{A} \subset M(m, n)$, we sort the singular values of A as

$$\sigma^{(1)} > \sigma^{(2)} > \cdots > \sigma^{(m)} > 0.$$

Let $g^{(i)}$ be the left eigenvector of A for $\sigma^{(i)} \in S^{(m-1)}$ and $h^{(i)}$ the right eigenvector for A for $\sigma^{(i)} \in S^{(n-1)}$. Since $h^{(i)}$ is an eigenvector for $A^T A$ for the eigenvalue $\sigma^{(i)}$ and $g^{(i)}$ is an eigenvector for AA^T for the eigenvalue $\sigma^{(i)}$ and the eigenvalues are different, $g^{(i)}$ and $h^{(i)}$ are uniquely determined modulo the multiplication by ± 1. Define a map φ_i from $\mathcal{A}$ to $\mathcal{A}_i$ by

$$\varphi_i(A) = (\sigma^{(i)}, g^{(i)}, h^{(i)}, G(g^{(i)}) A H^T(h^{(i)})). \tag{4}$$

Note that the matrix $G(g^{(i)}) A H^T(h^{(i)})$ lies in $\mathcal{B}(i, \sigma^{(i)})$, because the singular values of $B^{(i)}$ agree with those of A excluding $\sigma^{(i)}$.

Lemma 1. *The map φ_i is smooth and isomorphic.*

Proof. Define a map ψ from $\mathcal{A}_i$ to $\mathcal{A}$ by

$$\psi(\sigma, g, h, B) = g\sigma h^T + G(g) B H(h)^T.$$

We can see that $\varphi_i \circ \psi$ and $\psi \circ \varphi_i$ are identity maps by a calculation. Then the map φ_i is one-to-one and surjective. Next, we show that the map ψ is smooth. Since we assume that all the singular values are different, the maps of taking i-th singular value of a given A and an eigenvector for the singular value are smooth on an open connected neighborhood $W \subset \mathcal{A}$ of A (Or check the Jacobian does not vanish.). Then the inverse map is locally smooth. Hence, φ_i is smooth and isomorphism. //

We are interested in the Euler characteristic number of M_x.

Theorem 1. *Assume $x > 0$ and suppose that $f(U)$ is a Morse function for almost all A's. We assume that if a set is measure 0 set with respect to the Lebesgue measure, then it is also a measure 0 set with respect to the measure $p(A)dA$. The expectation of the Euler characteristic number $E[\chi(M_x)]$ is equal to*

$$\frac{1}{2}\int_x^\infty \sigma^{n-m} d\sigma \int_{\mathbb{R}^{(m-1)(n-1)}} dB \int_{S^{m-1}} G^T dg \int_{S^{n-1}} H^T dh \, \det(\sigma I_{m-1} - BB^T) p(A). \tag{5}$$

Here, we set $G^T dg = \wedge_{i=1}^{m-1} G_i^T dg$, $H^T dh = \wedge_{i=1}^{n-1} H_i^T dh$, where G_i and H_i are the i-th column vectors of G and H, respectively, and $dg = (dg_1, \ldots, dg_m)^T$ and $dh = (dh_1, \ldots, dh_n)^T$.

We note that $G^T dg$ and $H^T dh$ are $O(m)$ and $O(n)$ invariant measures on S^{m-1} and S^{n-1}, respectively.

Proof. Without loss of generality, we assume that $m \le n$. According to the Morse theory, if $f(U)$ is a Morse function, which is a smooth function without a degenerated critical point, then we have

$$\begin{aligned}
\chi(M_x) &= \sum_{\text{critical point}} \mathbf{1}(f(U) \ge x) \operatorname{sgn} \det \begin{pmatrix} -\partial_i \partial_j f & -\partial_i \partial_a f \\ -\partial_a \partial_i f & -\partial_a \partial_b f \end{pmatrix} && (6) \\
&= \sum_{\text{eigenvectors}} \mathbf{1}(\sigma \ge x) \operatorname{sgn} \det \begin{pmatrix} \sigma I_m & -GBH^T \\ -HB^T G^T & \sigma I_n \end{pmatrix} && (7) \\
&= \sum_{i=1}^m \mathbf{1}(\sigma^{(i)} \ge x) \operatorname{sgn} {\sigma^{(i)}}^{n-m} {\sigma^{(i)}}^2 \det \left({\sigma^{(i)}}^2 I_{m-1} - B^{(i)} {B^{(i)}}^T \right), && (8)
\end{aligned}$$

where $\sigma^{(i)}$ is the i-th singular value of A, $g^{(i)}$ and $h^{(i)}$ are left and right eigenvectors, and $B^{(i)} = G^T(g^{(i)})AH(h^{(i)})$. The equality (6) is the Morse theorem for manifolds with boundaries. The equality of (6) and (7) can be shown as follows.

We have the relation $g_i^T g = 0$. By differentiating by u_j, we have $g_{ij}^T g + g_i^T g_j = 0$. Let us evaluate $\partial_i \partial_j f$. It is equal to, by the expression $A = \sigma g h^T + GBH^T$,

$$\begin{aligned} & \partial_i \partial_j f \\ = \; & g_{ij}^T A h \\ = \; & g_{ij}^T \sigma g h^T h + g_{ij}^T GBH^T h \\ = \; & -\sigma g_i^T g_j \quad \text{by } H^T h = 0. \end{aligned}$$

Next, we evaluate $\partial_i \partial_a f$.

$$\begin{aligned} & \partial_i \partial_a f \\ = \; & g_i^T A h_a \\ = \; & g_i^T g \sigma h^T h_a + g_i^T GBH^T h_a \\ = \; & g_i^T GBH^T h_a \quad \text{by } g_i^T g = h^T h_a = 0. \end{aligned}$$

Thirdly, we evaluate $\partial_a \partial_b f$.

$$\begin{aligned} & \partial_a \partial_b f \\ = \; & g^T A h_{ab} \\ = \; & g^T g \sigma h^T h_{ab} + g^T GBH^T h_{ab} \\ = \; & -\sigma h_a^T h_b \quad \text{by } g^T G = 0. \end{aligned}$$

Summarizing these calculation, we have that the Hessian is equal to

$$\begin{aligned} & \begin{pmatrix} -\partial_i \partial_j f & -\partial_i \partial_a f \\ -\partial_i \partial_a f & -\partial_a \partial_b \end{pmatrix} \\ = \; & \begin{pmatrix} \sigma g_i^T g_j & -g_i^T GBH^T h_a \\ -h_a^T HB^T G^T g_i & \sigma h_a^T h_b \end{pmatrix} \\ = \; & \begin{pmatrix} g_1 \cdots g_{n-1} & 0 \\ 0 & h_1 \cdots h_{m-1} \end{pmatrix}^T \begin{pmatrix} \sigma I_m & -GBH^T \\ -HB^T G^T & \sigma I_n \end{pmatrix} \begin{pmatrix} g_1 \cdots g_{n-1} & 0 \\ 0 & h_1 \cdots h_{m-1} \end{pmatrix}. \end{aligned}$$

Since $\det(PP^T) = \det(P)^2$, the sign of the determinant of the Hessian is equal to the sign of the determinant of the middle of the above 3 matrices.

Let us show the equality of (7) and (8). We fix i and omit the superscript (i) in the following discussion. We consider the product of the following two matrices.

$$\begin{pmatrix} \sigma I_m & -GBH^T \\ -HB^T G^T & \sigma I_n \end{pmatrix} \begin{pmatrix} \sigma I_m & 0 \\ HB^T G^T & \sigma^{-1} I_n \end{pmatrix}.$$

It is equal to

$$\begin{pmatrix} \sigma^2 I_m - GBB^T G^T & -\sigma^{-1} GBH^T \\ -\sigma HB^T G^T + \sigma HB^T G^T & I_n \end{pmatrix}.$$

Since the left-bottom block is 0, the determinant of this matrix is $\det(\sigma^2 I_m - GBB^T G)$. Put $C = BB^T$ and $\tilde{G} = (g|G)$. We have

$$\sigma^2 I_m - GCG^T = \sigma^2 I_m - \tilde{G}\left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & C \end{array}\right)\tilde{G}^T.$$

Since $\tilde{G}\tilde{G}^T = E$, the determinant of the matrix above is equal to $\sigma^2 \det(\sigma^2 I_{m-1} - C)$. In summary, we obtain the equality of (7) and (8).

Let us take the expectation of the Euler characteristic number. Exchanging the sum and the integral, we have

$$\begin{aligned} & E[\chi(M_x)] \\ = & \sum_{i=1}^{m} \int dA p(A) \mathbf{1}(\sigma^{(i)} \geq x) \operatorname{sgn} {\sigma^{(i)}}^{n-m} {\sigma^{(i)}}^2 \det\left({\sigma^{(i)}}^2 I_{m-1} - B^{(i)} {B^{(i)}}^T\right) \end{aligned}$$

To evaluate the expectation of the Euler characteristic number, we needs the Jacobian of (3). According to standard arguments of multivariate analysis (see, e.g., [23, (3.19)]), we have

$$dA = |\det(\sigma^2 I_{m-1} - BB^T)|\, d\sigma G^T dg H^T dh dB. \tag{9}$$

Then we have

$$\begin{aligned} & E[\chi(M_x)] \\ = & \frac{1}{2}\sum_{i=1}^{m} \int_x^\infty \sigma^{n-m} d\sigma \int_{B\in\mathcal{B}(i,\sigma^{(i)})} dB \int_{S^{m-1}} G^T dg \int_{S^{n-1}} H^T dh \det(\sigma I_{m-1} - BB^T) p(A). \end{aligned} \tag{10}$$

Note that the factor 1/2 comes from the fact that the multiplicity $(g,h) \mapsto gh^T$ is 2. Put $\mathcal{B}^{(i)} = \mathcal{B}(i, \sigma^{(i)})$. Since $\mathcal{B}^{(i)} \cap \mathcal{B}^{(j)}$, $i \neq j$ is a measure 0 set and $\mathbb{R}^{(m-1)(n-1)} \setminus \sum_{i=1}^m \mathcal{B}^{(i)}$ is also a measure 0 set, we may sum up integral domains for B into one domain as

$$\begin{aligned} & \sum_{i=1}^{m} \int_{B\in\mathcal{B}(i,\sigma^{(i)})} \det(\sigma I_{m-1} - BB^T) p(A) \\ = & \int_{B\in M(m-1,n-1)} \det(\sigma I_{m-1} - BB^T) p(A). \end{aligned}$$

Thus, we derive the conclusion. //

We note that the integral (5) does not depend on a choice of $G(g)$ nor $H(h)$. The column vectors of the matrix $G = G(g)$ have the length 1 and are orthogonal to the vector g. Let $\tilde{G}$ be a matrix which has the same property. In other words, we assume $(g, \tilde{G}) \in SO(m)$. Then, there exists an $(m-1) \times (m-1)$ orthogonal matrix P such that $\tilde{G} = GP$ and $|P| = 1$ hold. Taking the exterior product of elements of $\tilde{G}^T dg = PG^T dg$, we have

$$\wedge_{i=1}^m \tilde{g}_i^T dg = |P| \wedge_{i=1}^m g_i^T dg = \wedge_{i=1}^m g_i^T dg.$$

The case for H can be shown analogously.

One of the most important examples is that A is distributed as a Gaussian distribution $N_{m\times n}(M,\Sigma\otimes I_n)$:

$$p(A)dA=\frac{1}{(2\pi)^{mn/2}\mathrm{det}(\Sigma)^{n/2}}\exp\Big\{-\frac{1}{2}\mathrm{tr}(A-M)^T\Sigma^{-1}(A-M)\Big\}dA.$$

Then the largest singular value of A is the square root of the largest eigenvalue of a non-central Wishart matrix $W_m(n,\Sigma,\Sigma^{-1}MM^T)$. Substituting (5),

$$\begin{aligned}E[\chi(M_x)]=&\frac{1}{2}\int_x^\infty\sigma^{n-m}d\sigma\int_{\mathbb{R}^{(m-1)\times(n-1)}}dB\int_{\mathbb{S}^{m-1}}G^Tdg\int_{\mathbb{S}^{n-1}}H^Tdh\mathrm{det}\left(\sigma^2I_{m-1}-BB^T\right)\\&\times\frac{1}{(2\pi)^{nm/2}\mathrm{det}(\Sigma)^{n/2}}\exp\Big\{-\frac{1}{2}\mathrm{tr}(\sigma hg^T+HB^TG^T-M^T)\Sigma^{-1}(\sigma gh^T+GBH^T-M)\Big\}.\end{aligned}\tag{11}$$

In this expression, number of parameters is $m(m+1)/2+mn$, but it is over-parametrized. Note that

$$A=\Sigma^{1/2}V+M,\quad V=(v_{ij})_{m\times n},\quad v_{ij}\sim N(0,1)\ \text{ i.i.d.}$$

Let $\Sigma^{1/2}=P^TDP$, $D=\mathrm{diag}(d_i)$, is a spectral decomposition. Then, $PA=DPV+PM$. Let $PM=NQ$ be a QR decomposition, where N is $m\times n$ lower triangle matrix with nonnegative diagonal elements and $Q\in O(n)$. Then $PAQ^T=DV+N$. Since the largest eigenvalues of A and PAQ^T are the same, we can assume that Σ is a diagonal matrix, and M is a lower triangle with nonnegative diagonal elements without loss of generality. That is,

$$\Sigma^{-1}=\begin{pmatrix}s_1&&0\\&\ddots&\\0&&s_m\end{pmatrix},\quad s_i>0,\quad M=\begin{pmatrix}m_{11}&&0&0&\cdots&0\\\vdots&\ddots&&\vdots&&\vdots\\m_{mn}&\cdots&m_{mm}&0&\cdots&0\end{pmatrix},\quad m_{ii}\ge0.\tag{12}$$

When Σ has multiple roots, *i.e.*,

$$\Sigma^{-1}=\begin{pmatrix}s_1I_{n_1}&&0\\&\ddots&\\0&&s_rI_{n_r}\end{pmatrix},\quad \sum_{i=1}^r n_i=m,\tag{13}$$

by multiplying $\mathrm{diag}(P_1,\dots,P_r)\in O(n_1)\times\cdots\times O(n_r)$ and its transpose from left and right, we can assume

$$M=\begin{pmatrix}m_1I_{n_1}&&&&&0&\cdots&0\\M_{21}&m_2I_{n_2}&&&&0&\cdots&0\\\vdots&&\ddots&&&\vdots&&\vdots\\M_{r-1,1}&M_{r-1,2}&&m_{r-1}I_{n_{r-1}}&&0&\cdots&0\\M_{r1}&M_{r2}&\cdots&M_{r,r-1}&m_rI_{n_r}&0&\cdots&0\end{pmatrix},\quad m_i\ge0,\quad M_{ij}\in\mathbb{R}^{n_j\times n_i}.\tag{14}$$

Therefore, our problem is formalized as follow: To evaluate (11) with parameters (12) (or (13) and (14)).

In the following sections, we will evaluate the integral representation of the expectation of the Euler characteristic number given in the Theorem 1 for some interesting special cases. We can obtain approximate values of the probability of the first eigenvalue of random matrices by virtue of them. The Euler characteristic heuristic is

$$P\Big(\max_{g\in\mathbb{S}^{m-1},\, h\in\mathbb{S}^{n-1}} g^T A h \geq x\Big) = P\Big(\max_{U\in M} f(U) \geq x\Big) \approx E\big[\chi(M_x)\big].$$

The condition that $f(U)$ is a Morse function with probability one holds if A has distinct and non-zero m singular values with probability one.

3 The case of $m = n = 2$

We derive Theorem 1 in the special case of $m = n = 2$ with taking explicit coordinates. We have referred this derivation to find a proof for the general case discussed in the previous section. The case $m = n = 2$ will be studied numerically in the last section with the holonomic gradient method (HGM).

Fix two unit vectors

$$g = (\cos\theta, \sin\theta)^T, h = (\cos\phi, \sin\phi)^T \in S^1$$

for $0 \leq \theta, \phi < 2\pi$. Define

$$G = \left(\cos\left(\theta + \frac{\pi}{2}\right), \sin\left(\theta + \frac{\pi}{2}\right)\right)^T = (-\sin\theta, \cos\theta)^T,$$

which satisfies

$$(g, G) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in SO(2).$$

Similarly, we define $H = \left(\cos\left(\phi + \frac{\pi}{2}\right), \sin\left(\phi + \frac{\pi}{2}\right)\right)^T = (-\sin\phi, \cos\phi)^T$. Here, both $\theta + \frac{\pi}{2}$ and $\phi + \frac{\pi}{2}$ should be treated as mod 2π, in case the sum is greater than 2π. Now, any 2×2 matrix, say A, can be recovered by

$$A = \sigma g h^T + b G H^T$$

still with 4 variables $(\sigma, \theta, \phi, b)$, instead of $a_{11}, a_{12}, a_{21}, a_{22}$. We may further assume that $\sigma \in \mathbb{R}_{\geq 0}$, $b \in \mathbb{R}$ and $\phi, \theta \in [0, 2\pi)$.

Now, if we fix $\sigma_0, b_0, \theta_0, \phi_0$ and let

$$A_0 = \sigma_0 g(\theta_0) h(\phi_0)^T + b_0 G(\theta_0) H(\phi_0)^T,$$

by allowing σ, b vary in $\mathbb{R}$ and ϕ, θ vary in $[0, 2\pi)$, we will recover A_0 four times:

$$\begin{cases} A_0 &= \sigma_0 g(\theta_0) h(\phi_0)^T + b_0 G(\theta_0) H(\phi_0)^T \\ A_0 &= \sigma_0 g(-\theta_0) h(-\phi_0)^T + b_0 G(-\theta_0) H(-\phi_0)^T \\ A_0 &= b_0 g\left(\theta_0 + \frac{\pi}{2}\right) h\left(\phi_0 + \frac{\pi}{2}\right)^T + \sigma_0 G\left(\theta_0 + \frac{\pi}{2}\right) H\left(\phi_0 + \frac{\pi}{2}\right)^T \\ A_0 &= b_0 g\left(-\theta_0 + \frac{\pi}{2}\right) h\left(-\phi_0 + \frac{\pi}{2}\right)^T + \sigma_0 G\left(-\theta_0 + \frac{\pi}{2}\right) H\left(-\phi_0 + \frac{\pi}{2}\right)^T \end{cases}$$

- Here the first two are easily seen from the symmetry of the manifold M (shown below) that $(h, g) \sim (-h, -g)$.
- The second symmetry is given $(\sigma', b') = (b_0, \sigma_0)$, i.e., interchanging σ and b, there also exists $(\theta', \phi') = \left(\theta_0 + \frac{\pi}{2}, \phi_0 + \frac{\pi}{2}\right)$ recovering A_0, if noting $G(\theta) = g\left(\theta + \frac{\pi}{2}\right)$ and $H(\phi) = h\left(\phi + \frac{\pi}{2}\right)$.

Therefore, to recover A, we could always assume that $\sigma \geq b$, and let $\theta, \phi \in [0, 2\pi)$. See Lemma 1 for a general claim.

Again, we consider the manifold

$$M = \left\{ts^T \mid s = (\cos\alpha, \sin\alpha), t = (\cos\beta, \sin\beta) \in S^1, 0 \leq \alpha, \beta < 2\pi\right\}$$

and function f on M such that

$$f\left(ts^T\right) = s^T A t = s^T \left(\sigma g h^T + b G H^T\right) t.$$

Apparently, A only has two pairs of eigenvectors, which can be easily verified easily by the following computations:

$$\begin{cases} Ah = \sigma g h^T h + b G H^T h & = \sigma g; \\ g^T A = \sigma g^T g h^T + b g^T G H^T & = \sigma h^T; \\ AH = \sigma g h^T H + b G H^T H & = bG; \\ G^T A = \sigma G^T g h^T + b G^T G H^T & = bH^T. \end{cases}$$

Namely, function f has two critical points on M, which are at

- point $P = hg^T \in M \Leftrightarrow (\alpha, \beta) = (\theta, \phi)$;
- and point $Q = HG^T \in M \Leftrightarrow (\alpha, \beta) = \left(\theta + \frac{\pi}{2}, \phi + \frac{\pi}{2}\right)$.

Further computation shows the following 4 facts.

- $f(P) = g^T A h = \sigma$ and $f(Q) = G^T A H = b$.
- From

$$\begin{aligned} \mathrm{Hess} f &= \begin{pmatrix} \frac{\partial^2}{\partial\alpha^2} f & \frac{\partial^2}{\partial\alpha\partial\beta} f \\ \frac{\partial^2}{\partial\beta\partial\alpha} f & \frac{\partial^2}{\partial\beta^2} f \end{pmatrix} \\ &= \begin{pmatrix} \frac{(b-\sigma)\cos(\alpha+\beta-\theta-\phi)-(b+\sigma)\cos(\alpha-\beta-\theta+\phi)}{2} & \frac{(b-\sigma)\cos(\alpha+\beta-\theta-\phi)+(b+\sigma)\cos(\alpha-\beta-\theta+\phi)}{2} \\ \frac{(b-\sigma)\cos(\alpha+\beta-\theta-\phi)+(b+\sigma)\cos(\alpha-\beta-\theta+\phi)}{2} & \frac{(b-\sigma)\cos(\alpha+\beta-\theta-\phi)-(b+\sigma)\cos(\alpha-\beta-\theta+\phi)}{2} \end{pmatrix}, \end{aligned}$$

  it follows that $\det\left(\mathrm{Hess}_P f\right) = \sigma^2 - b^2$ and $\det\left(\mathrm{Hess}_Q f\right) = b^2 - \sigma^2$. Therefore, we see

  1. if $x > \sigma \geq b$, M_x does not contain any critical points, so $\chi(M_x) = 0$;

2. if $x < b \leq \sigma$, M_x contains both critical points, so

$$\chi\left(M_x\right) = \operatorname{sgn}\left(\sigma^2 - b^2\right) + \operatorname{sgn}\left(b^2 - \sigma^2\right) = 0;$$

3. the only nontrivial case is $\sigma \geq x \geq b$, then

$$\chi\left(M_x\right) = \mathbf{1}\left(\sigma \geq x \geq b\right) \operatorname{sgn}\left(\sigma^2 - b^2\right).$$

- Since

$$A = \sigma g h^T + b G H^T = \begin{pmatrix} b\sin\theta\sin\phi + \sigma\cos\theta\cos\phi & \sigma\cos\theta\sin\phi - b\sin\theta\cos\phi \\ \sigma\sin\theta\cos\phi - b\cos\theta\sin\phi & b\cos\theta\cos\phi + \sigma\sin\theta\sin\phi \end{pmatrix},$$

we have

$$\begin{aligned}(\mathrm{d}A) = {} & \mathrm{d}b\sin\theta\sin\phi + \sigma\cos\theta\cos\phi \wedge \mathrm{d}\left(\sigma\cos\theta\sin\phi - b\sin\theta\cos\phi\right) \\ & \wedge \mathrm{d}\left(\sigma\sin\theta\cos\phi - b\cos\theta\sin\phi\right) \wedge \mathrm{d}\left(b\cos\theta\cos\phi + \sigma\sin\theta\sin\phi\right) \\ = {} & \left(b^2 - \sigma^2\right)\mathrm{d}\sigma\mathrm{d}b\mathrm{d}\theta\mathrm{d}\phi.\end{aligned}$$

- Let $M = \begin{pmatrix} m_{11} & 0 \\ m_{21} & m_{22} \end{pmatrix}$ and $\Sigma = \begin{pmatrix} 1/s_1 & 0 \\ 0 & 1/s_2 \end{pmatrix}$ such that

$$A = \sqrt{\Sigma}V + M, \text{ where } V = \left(v_{ij}\right), \ v_{ij} \sim \mathcal{N}\left(0,1\right) \text{ i.i.d.}$$

Then

$$P\left(A\right) = \frac{s_1 s_2}{\left(2\pi\right)^2} e^{-\frac{R}{2}},$$

where

$$\begin{aligned}R = {} & s_1\left(b\sin\theta\sin\phi + \sigma\cos\theta\cos\phi - m_{11}\right)^2 + s_2\left(\sigma\sin\theta\cos\phi - b\cos\theta\sin\phi - m_{21}\right)^2 \\ & + s_1\left(\sigma\cos\theta\sin\phi - b\sin\theta\cos\phi\right)^2 + s_2\left(b\cos\theta\cos\phi + \sigma\sin\theta\sin\phi - m_{22}\right)^2.\end{aligned}$$

Hence, we have

$$\begin{aligned}E\left(\chi\left(M_x\right)\right) & = \frac{1}{2}\underline{\int_{-\infty}^{\infty}}\mathrm{d}\sigma\underline{\int_{-\infty}^{\infty}}\mathrm{d}b\int_0^{2\pi}\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\left(\underline{\mathbf{1}\left(\sigma \geq x \geq b\right)}\overbrace{\operatorname{sgn}\left(\sigma^2 - b^2\right)}\right)\frac{s_1 s_2}{\left(2\pi\right)^2}e^{-\frac{R}{2}}\overbrace{\left|\left(b^2 - \sigma^2\right)\right|} \\ & = \frac{1}{2}\int_x^{\infty}\mathrm{d}\sigma\int_{-\infty}^{x}\mathrm{d}b\int_0^{2\pi}\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\,\overbrace{\left(\sigma^2 - b^2\right)}\frac{s_1 s_2}{\left(2\pi\right)^2}e^{-\frac{R}{2}}.\end{aligned}$$

Note that we have $\int_{-\infty}^{\infty} db \cdots = \int_{-\infty}^{x} db \cdots$ by an anti-symmetry of σ and b in this case. In other words, integrals over $\sigma > x > 0, b > x, \sigma > b$ and $\sigma > x > 0, b > x, \sigma < b$ are canceled. Thus, we have obtained Theorem 1 in the case that A is distributed as a Gaussian distribution.

Let us give a numerical example.

Example 1. When $m = n = 2$, by letting $g = (\cos\theta, \sin\theta)^T$, $G = (-\sin\theta, \cos\theta)^T$, $h = (\cos\phi, \sin\phi)^T$, $H = (-\sin\phi, \cos\phi)^T$, $B = (b)_{1\times 1}$,

$$\Sigma^{-1} = \begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix}, \qquad M = \begin{pmatrix} m_{11} & 0 \\ m_{21} & m_{22} \end{pmatrix},$$

we have

$$\begin{aligned} E[\chi(M_x)] =& F(s_1, s_2, m_{11}, m_{21}, m_{22}; x) \\ =& \frac{1}{2}\int_x^\infty d\sigma \int_{-\infty}^\infty db \int_0^{2\pi} d\theta \int_0^{2\pi} d\phi (\sigma^2 - b^2)\frac{s_1 s_2}{(2\pi)^2} \exp\Big\{-\frac{1}{2}R\Big\}, \end{aligned}$$

where

$$\begin{aligned} R =& s_1(b\sin\phi\sin\theta + \sigma\cos\phi\cos\theta + m_{11})^2 + s_2(-b\sin\phi\cos\theta + \sigma\cos\phi\sin\theta + m_{21})^2 \\ &+ s_2(b\cos\phi\cos\theta + \sigma\sin\phi\sin\theta + m_{22})^2 + s_1(-b\cos\phi\sin\theta + \sigma\sin\phi\cos\theta)^2. \end{aligned}$$

Moreover, by letting $s_1 = 2$, $s_2 = 1$, $m_{11} = 1$, $m_{21} = -1$, $m_{22} = 1$, we have the following table:

x	0	1	2	3	4	5	6
$E[\chi(M_x)]$	-5.92828×10^{-8}	0.745833	0.567728	0.144874	0.0146727	0.000582529	8.75458×10^{-6}
$P(\sigma > x)$	1.	0.957375	0.576156	0.145001	0.0146561	0.000584400	0.0000105

Here, the probability $P(\sigma > x)$ is estimated by a Monte Carlo study with 10,000,000 iterations and the expectation of the Euler characteristic is evaluated by a numerical integration function `NIntegrate` on Mathematica [20]. As expected, $E[\chi(M_x)] \approx P(\sigma > x)$ when x is large.

4 The central case with a scalar covariance: Selberg type integral and Laguerre polynomials

In this section, we assume that $M = 0$ (central) and Σ in (12) is a scalar matrix and study this case by special functions. Under these assumptions, we will show that the expectation of the Euler characteristic can be expressed in terms of a Selberg type integral, which equals to a Laguerre polynomial by virtue of the works by K.Aomoto [2] and J.Kaneko [11]

Theorem 2. *Put*

$$M_x = \{hg^T \mid g^T A h \geq x, h, g \in S^{m-1}\}.$$

The distribution of $m \times m$ random matrices A is the Gaussian distribution with average 0 *and the covariance E_m/s. In other words, we have*

$$p(A) \sim \exp\left(-\frac{1}{2}\mathrm{tr}\,(sA^T A)\right).$$

Then we have

$$E[\chi(M_x(s))] = \prod_{i=1}^{5} c_i \int_x^{+\infty} \exp\left(-\frac{s}{2}\sigma^2\right) {}_1F_1(-(m-1), 1; s\sigma^2) d\sigma \tag{15}$$

where c_i, $i = 1, \ldots, 5$ are given in (16), (17), (20), (23), (26) respectively.

Proof. We put

$$\tilde{G} = \left(\, g \mid G \, \right) \in O(m), \quad g \text{ is a column vector,}$$

$$\tilde{H} = \left(\, h \mid H \, \right) \in O(n), \quad h \text{ is a column vector.}$$

Then the $m \times m$ matrix A can be written as

$$A = \tilde{G} \left(\begin{array}{c|c} \sigma & 0 \\ \hline 0 & B \end{array} \right) \tilde{H}^T.$$

We denote by $\tilde{B}$ the middle matrix in the expression above.

We denote $\exp(\mathrm{tr}(X))$ by $\mathrm{etr}(X)$. Put $S = \Sigma^{-1}$. We consider the central case $M = 0$ in (11). Since $\mathrm{tr}(PQ) = \mathrm{tr}(QP)$ and $\tilde{H}^T\tilde{H} = E$, we have

$$\begin{aligned} & \mathrm{etr}(-\frac{1}{2}A^T S A) \\ = \ & \mathrm{etr}(-\frac{1}{2}\tilde{H}\tilde{B}^T\tilde{G}^T S\tilde{G}\tilde{B}\tilde{H}^T) \\ = \ & \mathrm{etr}(-\frac{1}{2}S\tilde{G}\tilde{B}\tilde{H}^T\tilde{H}\tilde{B}^T\tilde{G}^T) \\ = \ & \mathrm{etr}(-\frac{1}{2}S\tilde{G}(\tilde{B}\tilde{B}^T)\tilde{G}^T). \end{aligned}$$

It follows from the Theorem 1 with $p(A)$ being the normal distribution that

$$\begin{aligned} E[\chi(M_x)] \ = \ & c_1(S) \int_x^{\infty} \sigma^{n-m} d\sigma \int_{\mathbb{R}^{(m-1)(n-1)}} dB \int_{S^{m-1}} G^T dg \int_{S^{n-1}} H^T dh \\ & \det(\sigma^2 I_{m-1} - BB^T)\mathrm{etr}(-\frac{1}{2}S\tilde{G}(\tilde{B}\tilde{B}^T)\tilde{G}^T), \end{aligned}$$

where

$$c_1(S) = \frac{1}{2} \cdot \frac{1}{(2\pi)^{nm/2}\det(S^{-1})^{n/2}}. \tag{16}$$

We denote by G_i the i-th column vector of G and by dg the column vector of the differential forms dg_i. Define

$$G^T dg = \wedge_{i=1}^{m-1} G_i^T \cdot dg.$$

It is an invariant measure for the rotations on S^{m-1} [10, Theorem 4.2]. We may define $H^T dh$ analogously.

Moreover, we assume that $S = \Sigma^{-1}$ is a scalar matrix $\underline{S = sE \text{ and } m = n}$. Then we have

$$\mathrm{etr}(-\frac{1}{2}S\tilde{G}(\tilde{B}\tilde{B}^T)\tilde{G}^T) = \mathrm{etr}(-\frac{s}{2}\tilde{B}\tilde{B}^T).$$

Since there is no G, H in the etr, we can separate the following integral

$$c_2(m) = \int_{S^{m-1}} G^T dg \int_{S^{m-1}} H^T dh = \left(\frac{2\pi^{m/2}}{\Gamma(m/2)}\right)^2. \tag{17}$$

Therefore, we may evaluate the integral

$$\int_{\mathbb{R}^{(m-1)^2}} dB \det(\sigma^2 I_{m-1} - BB^T)\mathrm{etr}\left(-\frac{s}{2}\tilde{B}\tilde{B}^T\right). \tag{18}$$

We denote the integral above by $q(s;\sigma)$. In terms of $q(s;\sigma)$, we have

$$E[\chi(M_x)] = c_1(S)c_2(m)\int_x^\infty q(s;\sigma)d\sigma.$$

We make the singular value decomposition of the matrix B as $B = PLQ^T$, where the matrices $P, Q \in O(m-1)$, $L = \mathrm{diag}(\ell_1, \ldots, \ell_{m-1})$ (see, e.g., [10] ([23, (3.1)]). It follows from [23, (3.1)] that

$$dB = \prod_{1\le i<j\le m-1}(\ell_i^2 - \ell_j^2)\left(\prod_{i=1}^{m-1} d\ell_i\right)\wedge\omega,$$

$$\omega = \wedge_{1\le i\le m-1, i<j\le m-1} P_j^T dP_i \ \wedge_{1\le i\le m-1, i<j\le m-1} Q_j^T dQ_i,$$

when $\ell_1 \ge \ell_2 \ge \cdots \ge \ell_{m-1}$. Here, P_i and Q_i are i-th column vectors, respectively. Since

$$\det(\sigma^2 I_{m-1} - PLQ^TQL^TP^T) = \det(P(\sigma^2 I_{m-1} - LL^T)P^T) = \det(\sigma^2 I_{m-1} - LL^T),$$

and

$$\begin{aligned}
& \mathrm{etr}\left(-\frac{s}{2}\tilde{B}\tilde{B}^T\right) \\
= \ & \exp\left(-\frac{s}{2}\sigma^2\right)\mathrm{etr}\left(-\frac{s}{2}BB^T\right) \\
= \ & \exp\left(-\frac{s}{2}\sigma^2\right)\mathrm{etr}\left(-\frac{s}{2}PLQ^TQL^TP^T\right) \\
= \ & \exp\left(-\frac{s}{2}\sigma^2\right)\exp\left(-\frac{s}{2}LL^T\right),
\end{aligned}$$

we have

$$q(s;\sigma) = c_3(m,\sigma)\int_{L\in\mathbb{R}^{m-1}}\prod_{1\le i<j\le m-1}|\ell_i^2 - \ell_j^2|\prod_{i=1}^{m-1}(\sigma^2 - \ell_i^2)\exp\left(-\frac{s}{2}\sum \ell_i^2\right)\prod_{i=1}^{m-1} d\ell_i. \tag{19}$$

Here, we put

$$\begin{aligned}
c_3(m;\sigma) &= \frac{1}{(m-1)!2^{m-1}2^{m-1}}\exp\left(-\frac{s}{2}\sigma^2\right)\int_{O(m-1)}\int_{O(m-1)}\omega \qquad (20) \\
&= \frac{1}{(m-1)!2^{m-1}}\exp\left(-\frac{s}{2}\sigma^2\right)\left(2^{m-2}\prod_{k=2}^{m-1}\frac{\pi^{k/2}}{\Gamma(k/2)}\right)^2.
\end{aligned}$$

We divide the integral by $(m-1)!2^{m-1}2^{m-1}$, because $(m-1)!2^{m-1}$ copies of the domain $\ell_1 \geq \ell_2 \geq \ldots \geq \ell_{m-1} \geq 0$ cover $\mathbb{R}^{m-1}$ and the correspondence of the B coordinates and the coordinates of the singular value decomposition is $1 : 2^{m-1}$. Note that the volume of $O(m-1)$ is two times of that of $SO(m-1)$.

In (19), we make a change of variables as $\ell'_i = \ell_i^2$. Then we have $d\ell'_i = 2\ell_i d\ell_i$, and

$$d\ell_i = \frac{1}{2\sqrt{\ell'_i}} d\ell'_i.$$

With this change of variables, we have the expression

$$\begin{aligned} q(s;\sigma) &= c_3(m;\sigma) \\ &\quad \int_{L' \in \mathbb{R}^{m-1}_{\geq 0}} \prod \ell_i'^{-1/2} \prod_{1 \leq i < j \leq m-1} |\ell'_i - \ell'_j| \prod_{i=1}^{m-1} (\sigma^2 - \ell'_i) \\ &\quad \exp\Big(-\frac{s}{2}\sum \ell'_i\Big) \prod_{i=1}^{m-1} d\ell'_i. \end{aligned} \tag{21}$$

Put $\ell'_i = \frac{2}{s}\ell''_i$ and factor out $s > 0$. Then it follows from $d\ell'_i = \frac{2}{s}d\ell''_i$ that

$$q(s;\sigma) = c_3(m;\sigma)c_4(m,s)\tilde{q}(s;\sigma),$$

where

$$\tilde{q}(s;\sigma) = \int_{L'' \in \mathbb{R}^{m-1}_{\geq 0}} \prod \ell_i''^{-1/2} \prod_{1 \leq i < j \leq m-1} |\ell''_i - \ell''_j| \prod_{i=1}^{m-1} (\frac{\sigma^2 s}{2} - \ell''_i) \exp\Big(-\sum \ell''_i\Big) \prod_{i=1}^{m-1} d\ell''_i, \tag{22}$$

and

$$c_4(m,s) = (s/2)^{(m-1)/2}(s/2)^{-\frac{1}{2}(m-1)(m-2)}(s/2)^{-(m-1)}(s/2)^{-(m-1)} = (s/2)^{-\frac{1}{2}(m^2-1)}. \tag{23}$$

This integral can be expressed as a polynomial in σ. Let us derive differential equations for this integral and express it in terms of a special polynomial. We utilize the result by Aomoto [2] and its generalization [11] by Kaneko. In [11], a system of differential equations, special values, and an expansion in terms of Jack polynomials are given for the integral

$$\int_{[0,1]^{m-1}} \prod_{1 \leq i \leq m-1, 1 \leq k \leq r} (\ell_i - \sigma_k)^{\mu} D(\ell_1, \ldots, \ell_{m-1}) d\ell_1 \cdots d\ell_{m-1}, \tag{24}$$

$$D = \prod_{i=1}^{m-1} \ell_i^{\lambda_1}(1-\ell_i)^{\lambda_2} \prod_{1 \leq i < j \leq m-1} |\ell_i - \ell_j|^{\lambda},$$

when $\mu = 1$ or $\mu = -\lambda/2$. Let us make the coordinate change $\ell_i = y_i/N$, $\lambda_2 = N$, $\sigma_i = \tau_i/N$. Then we have $d\ell_i = dy_i/N$, $(1-\ell_i)^{\lambda} = (1 - y_i/N)^N$,

$$(1 - y_i/N)^N \to \exp(-y_i), \quad N \to \infty.$$

The integral (24) changes to

$$c_N \int_{[0,N]^{m-1}} \prod_{1\le i\le m-1, 1\le k\le r} (y_i-\tau_k)^\mu D(y_1,\ldots,y_{m-1})dy_1\cdots dy_{m-1},$$

$$D=\prod_{i=1}^{m-1} y_i^{\lambda_1}(1-y_i/N)^N \prod_{1\le i<j\le m-1} |y_i-y_j|^\lambda, c_N=N^{-r(m-1)-(m-1)-\lambda_1(m-1)-\lambda(m-1)(m-2)/2}.$$

When $N\to\infty$, this integral divided by c_N converges to

$$\int_{\mathbb{R}_{\ge 0}^{m-1}} \prod_{1\le i\le m-1, 1\le k\le r} (y_i-\tau_k)^\mu D(y_1,\ldots,y_{m-1})dy_1\cdots dy_{m-1},$$

$$D=\prod_{i=1}^{m-1} y_i^{\lambda_1}\exp(-\sum_{i=1}^{m-1} y_i) \prod_{1\le i<j\le m-1} |y_i-y_j|^\lambda.$$

Let us apply this limiting procedure to the corresponding differential equation. When $r=1$, $\mu=1$, the differential equation for the integral (24) is

$$\sigma(1-\sigma)\partial_\sigma^2+(c-(a+b+1)\sigma)\partial_\sigma-ab$$

where $a=-(m-1)$, $b=\frac{2}{\lambda}(\lambda_1+\lambda_2+2)+(m-1)+1$, $c=\frac{2}{\lambda}(\lambda_1+1)$. This is the Gauss hypergeometric equation. Putting $\lambda_2=N$, $\sigma=\frac{z}{N}$, we can find the limit of this equation when $N\to\infty$. In fact, it can be performed as follows. Put $\theta_z=z\partial_z$. Note that it is invariant by the scalar multiplication of z. Then the limit of

$$\theta_z(\theta_z+\frac{2}{\lambda}(\lambda_1+1)-1)-\frac{z}{N}(\theta_z-(m-1))(\theta_z+\frac{2}{\lambda}(N+\lambda_1+2)+(m-1)+1)$$

when $N\to\infty$ is

$$\theta_z(\theta_z+\frac{2}{\lambda}(\lambda_1+1)-1)-\frac{2}{\lambda}z(\theta_z-(m-1)).$$

In particular, when $\lambda=1$ and $\lambda_1=-1/2$, it is

$$\theta_z^2-2z(\theta_z-(m-1)).$$

A polynomial solution of this can be written as

$$c_5(m)\cdot {}_1F_1(-(m-1),1;2z)$$

with a constant $c_5(m)$. Therefore, we have

$$\begin{aligned} q(s;\sigma) &= c_3(m;\sigma)c_4(m,s)c_5(m)\cdot {}_1F_1(-(m-1),1;\sigma^2 s) \\ &= c_3(m;\sigma)c_4(m,s)c_5(m)\big(1+\frac{-(m-1)}{1}(\sigma^2 s)+\frac{(m-1)(m-2)}{(2!)^2}(\sigma^2 s)^2 \\ &\quad +\frac{-(m-1)(m-2)(m-3)}{(3!)^2}(\sigma^2 s)^3+\cdots+\frac{(-1)^{m-1}(m-1)!}{((m-1)!)^2}(\sigma^2 s)^{m-1}\big) \end{aligned} \tag{25}$$

and

$$c_5(m)=(\text{the expression (22)})_{|\sigma=0}=\prod_{i=1}^{m-1}\frac{\Gamma\left(1+\frac{i}{2}\right)\Gamma\left(\frac{3}{2}+\frac{i-1}{2}\right)}{\Gamma\left(\frac{3}{2}\right)} \tag{26}$$

by taking a limit of the Selberg integral formula [22]. //

Let us make a numerical evaluation by utilizing Theorem 2 when $m = 3$. When $m = 3$, we have

$$c_1 c_2 c_3 c_4 c_5 = 2\sqrt{2/\pi}\sqrt{s}\exp(-\sigma^2 s/2).$$

Since

$$\begin{aligned} u(s,k,x) &= \int_x^{+\infty} \exp(-\sigma^2 s/2)\sigma^{2k} d\sigma \\ &= \Gamma(k+1/2)\left(\frac{2}{s}\right)^{k+1/2} \frac{1}{2}\int_{x^2}^{+\infty} \frac{y^{k+1/2-1}\exp(-y/(2/s))dy}{\Gamma(k+2)(2/s)^{k+1/2}}, \end{aligned}$$

where the integral of the second line is equal to the upper tail probability of the Gamma distribution with the scale $2/s$ and the shape $k + 1/2$, it follows from Theorem 2 that the expectation $E[\chi(M_x)]$ is equal to

$$2\sqrt{2/\pi}\sqrt{s}\left(u(s,0,x) - 2su(s,1,x) + \frac{s^2}{2}u(s,2,x)\right). \tag{27}$$

An R code of evaluating $E[\chi(M_x)]$ is as follows.

```
ug2<-function(s,k,x) {
  return(pgamma(x^2, scale=2/s, shape=k+1/2, lower = FALSE)*
        gamma(k+1/2)*(2/s)^(k+1/2)/2);
}
ec3<-function(x,s) {
 cc<- 2*(2/pi)^(1/2)*s^(1/2);
 c5<-1;
 return(cc*c5*
   (ug2(s,0,x)-2*s*ug2(s,1,x)+(1/2)*s^2*ug2(s,2,x)));
}
## Draw a graph
curve(ec3(x,1),from=1,to=10)
```

When $s = 1$, some values are as follows:

x	$E[\chi(M_x)]$	simulation (with 100000 tries)
3	0.215428520	0.217072
4	0.016122970	0.016195
5	0.000357368	0.000386

5 Computer algebra and the expectation for small m and n

In this section, we will study the non-central case $M \neq 0$ with the help of computer algebra. When $m = n = 2$, we can perform a general method of the holonomic gradient method (HGM) [6] to evaluate the integral (5).

By Theorem 1, we know that

$$\begin{aligned} E[\chi(M_x)] =& F(s_1, s_2, m_{11}, m_{21}, m_{22}; x) \\ =& \frac{1}{2}\int_x^{\infty} d\sigma \int_{-\infty}^{\infty} db \int_0^{2\pi} d\theta \int_0^{2\pi} d\phi (\sigma^2 - b^2)\frac{s_1 s_2}{(2\pi)^2}\exp\Big\{-\frac{1}{2}R\Big\}, \end{aligned} \tag{28}$$

where R is specified in Example 1, s_1, s_2, m_{11}, m_{21} and m_{22} are parameters.

In (28), we set

$$\sin\theta = \frac{2s}{1+s^2}, \quad \cos\theta = \frac{1-s^2}{1+s^2}, \quad \sin\phi = \frac{2t}{1+t^2}, \quad \cos\phi = \frac{1-t^2}{1+t^2}.$$

Then we have that

$$\begin{aligned} E[\chi(M_x)] =& F(s_1, s_2, m_{11}, m_{21}, m_{22}; x) \\ =& \frac{1}{2\pi^2}\int_x^\infty d\sigma \int_{-\infty}^\infty db \int_{-\infty}^\infty ds \int_{-\infty}^\infty dt \frac{s_1 s_2(\sigma^2 - b^2)}{(1+s^2)(1+t^2)} \exp\Big\{-\frac{1}{2}\tilde{R}\Big\}, \end{aligned} \tag{29}$$

where $\tilde{R}$ is a rational function in σ, b, s, t. Since the integrand is a holonomic function in σ, b, s, t, we can apply the creative telescoping method [25] to derive holonomic systems for the integrals. It is straightforward to do that for the inner single integral of $E[\chi(M_x)]$ by the classic methods [13] (such as Zeilberger's algorithm, Takayama's algorithm and Chyzak's algorithm). Below is an example:

Example 2. *Consider the inner single integral of* (29)*:*

$$f_1(\sigma, b, s) = \int_{-\infty}^\infty \frac{s_1 s_2(\sigma^2 - b^2)}{(1+s^2)(1+t^2)} \exp\Big\{-\frac{1}{2}\tilde{R}\Big\} dt,$$

where $\tilde{R}$ is a rational function in σ, b, s, t. Since the integrand of f_1 is a holonomic function, we can compute a holonomic system ann of it by using the Mathematica package `HolonomicFunctions` *[14]. Using ann and Chyzak's algorithm, we can then derive a holonomic system of f_1, which is of holonomic rank* 2. *The detailed calculation can be found in [18].*

In the above example, we use Chyzak's algorithm to derive a holonomic system of the inner single integral of $E[\chi(M_x)]$. It can be done within 5 seconds in a Linux computer with 15.10 GB RAM. However, experiments show that it is not efficient enough to derive a holonomic system for the inner double integral in the same way within reasonable computational time because of the complexity of this algorithm. In order to speed up the computation, our idea is to utilize Stafford theorem [9, 19] empirically. Let us first recall the theorem. Assume that $\mathbb{K}$ is a field of characteristic 0 and n is a positive integer. Let $R_n = \mathbb{K}(x_1, \ldots, x_n)[\partial_1, \ldots, \partial_n]$ and $D_n = \mathbb{K}[x_1, \ldots, x_n][\partial_1, \ldots, \partial_n]$ be the ring of differential operators with rational coefficients and the Weyl algebra in n variables, respectively.

Theorem 3. *Every left ideal in R_n or D_n can be generated by two elements.*

Assume that I is a left ideal in R_n or D_n. We observe from experiments that for any two random operators $a, b \in I$, it is of high probability that $I = \langle a, b\rangle$. This suggests the following heuristic method for computing a holonomic system for the inner double integral of $E[\chi(M_x)]$. As a matter of notation, we set

$$T_{n-1} = \{\partial_1^{i_1}\partial_2^{i_2}\cdots\partial_{n-1}^{i_{n-1}} \mid (i_1, \ldots, i_{n-1}) \in \mathbb{N}^{n-1}\}.$$

Recall that a D-finte system [4] in R_n is a finite set of generators of a zero-dimensional ideal in R_n. The relation between D-finite systems and holonomic systems is illustrated in [8, Section 6.9]. For the application of the holonomic gradient method, D-finite systems are alternative to holonomic systems. Here, we use D-finite systems because they are more efficient for computation.

Heuristic 1. *Given a D-finite system G in R_n, compute another D-finite system G_1 in R_{n-1} such that $G_1 \subset (R_n \cdot G + \partial_n R_n) \cap R_{n-1}$.*

1. *Choose two finite support set $S_1, S_2 \in T_{n-1}$.*
2. *Using the polynomial ansatz method [13, Section 3.4], check whether there exist telescopers $P_1, P_2 \in R_{n-1}$ of G with support sets S_1, S_2 or not. If P_1 and P_2 exist, then go to next step. Otherwise, go to step 1.*
3. *Compute the Gröbner basis G_1 of $\{P_1, P_2\}$ with respect to a term order in T_{n-1}. If G_1 is D-finite, then output G_1. Otherwise, go to step 1.*

In the above heuristic method, we need to find two finite support set $S_1, S_2 \in T_{n-1}$ through trial and error so that it will terminate and finish in a reasonable computational time. Next, we show how to use it to derive a D-finite system for the inner double integral of $E[\chi(M_x)]$.

Example 3. *Consider the inner double integral of* (29)*:*

$$f_2(\sigma, b) = \int_{-\infty}^{\infty} f_1(\sigma, b, s) ds \tag{30}$$

where $f_1(\sigma, b, s)$ is defined in Example 2.

Let G be a D-finite system of f_1, which is derived from Example 2. Using G and the polynomial ansatz method, we find two nonzero annihilators P_1 and P_2 for f_2 with support sets S_1 and S_2, respectively, where

$$S_1 = \{1, \partial_b, \partial_\sigma, \partial_b^2, \partial_b\partial_\sigma, \partial_\sigma^2, \partial_\sigma^3\},$$
$$S_2 = S_1 \cup \{\partial_b^2\partial_\sigma, \partial_b\partial_\sigma^2, \partial_b^3\}.$$

Then we compute the Gröbner basis G_1 of $\{P_1, P_2\}$ in $\mathbb{Q}(b, \sigma)[\partial_b, \partial_\sigma]$ with respect to a total degree lexicographic order. We find that G_1 is a D-finite system of holonomic rank 6*. The details of the calculation can be found in [18].*

In the above example, we specify the parameters in the integrand as that in Example 1. Using Heuristic 1, we can further compute a holonomic system for the inner double integral of $E[\chi(M_x)]$ without specifying those parameters (pars). It is much more efficient than Chyzak's algorithm. Below is a table for the comparison between Chyzak's algorithm (chyzak) and Heuristic 1 (heuristic) for the computational time (seconds).

# pars	0	1	2	3	4	5
chyzak	976	9.8323×10^4	-	-	-	-
heuristic	43.49	394.4	8527	4.3957×10^5	-	1.5519×10^6

Next, we use Heuristic 1 to derive a D-finite system of the inner triple integral of $E[\chi(M_x)]$ and then numerically solve the corresponding ordinary differential equation. Finally, we use numerical integration to evaluate $E[\chi(M_x)]$.

Example 4. *Consider*

$$E[\chi(M_x)] = \frac{1}{2\pi^2} \int_x^\infty d\sigma \int_{-\infty}^\infty db f_2(\sigma, b), \tag{31}$$

where $f_2(\sigma, b)$ is specified in (30).

By Example 3, we have derived a D-finite system for f_2. Using Heuristic 1, we derive a D-finite system for the inner first integral f_3 of (31) *of the following form:*

$$P = c_{10} \cdot \partial_\sigma^{10} + c_9 \cdot \partial_\sigma^9 + \cdots + c_0,$$

where $c_i \in \mathbb{Q}[\sigma], i = 0, \ldots, 10$.

Afterwards, we first numerically solve the ordinary differential equation $P(f_3) = 0$ to evaluate f_3, and then we evaluate $E[\chi(M_x)]$ by using numerical integration. Below are the results.

x	1	2	3	4	5	6
HGM	0.745835	0.567729	0.144879	0.0146728	0.000582526	8.79942×10^{-6}
mc	0.745802	0.567623	0.144986	0.0146901	0.0005933	9.6×10^{-6}

where mc is the result for a Monte Carlo study of $E[\chi(M_x)]$ by the following formula with 10,000,000 iterations:

$$E[\chi(M_x)] \approx \frac{\sum_{i=1}^n \chi(M_{x,i})}{n},$$

with

$$\chi(M_{x,i}) = \mathbf{1}(\sigma_i \geq x)(\sigma_i^2 - b_i^2) + \mathbf{1}(b_i \geq x)(b_i^2 - \sigma_i^2),$$

where σ_i and b_i are singular values of $M_{x,i}$, $i = 1, \ldots, n$.

As expected, the results of HGM are approximate to that of mc. The detailed computation can be found in [18].

Note that the evaluations of $E[\chi(M_x)]$ in the above example are also approximate to that in Example 1. The source codes for this section and a demo notebook are freely available as part of the supplementary electronic material [18].

Example 5. We consider the evaluation of (29) with parameters

$$m_{11} = 1, m_{21} = 2, m_{22} = 3, s_1 = 10^3, s_2 = 10^2.$$

As far as we have tried, it is hard to evaluate (29) for these relatively large parameters s_i by numerical integration (even the Monte Carlo integration). Thus, we take a different approach. Using Heuristic 1, we can compute a linear ODE for (29) of rank 11 with respect to the independent variable x. Then we construct series solutions for this differential equation and use them to extrapolate results by simulations.

Although this extrapolation method is well-known, we explain it in a subtle form with application in our evaluation problem. Consider an ODE with coefficients in $\mathbf{Q}(x)$ of rank r. Let $c \in \mathbf{Q}$ be a point in the x-space and we take r increasing numbers $y_j \in \mathbf{Q}$, where $j = 0, 1, \ldots, r-1$. We construct a series solution $f_i(x)$ as a series in $x - (c + y_i)$. We may further assume that $c + y_i$ is not a singular point of the ODE for each i. The initial value vector may be taken suitably so that the series is determined uniquely over $\mathbf{Q}$.

We assume that the vector $(f_i(x))$ converges in a segment I containing all $c+y_i$'s and it is a basis of the solution space. Once we construct such a basis of series solutions, we can construct the solution $f(x)$ which takes values b_j at $x = p_j \in \mathbf{Q} \cap I$, $j = 0, 1, \ldots, r-1$. To be specific, set

$$f(x) = \sum_{i=0}^{r-1} t_i f_i(x)$$

with unknown coefficients t_i's. Then we have

$$f(p_j) = \sum_{i=0}^{r-1} t_i f_i(p_j), \quad j = 0, 1, \ldots, r-1.$$

The unknown coefficients t_i's can be determined by solving the system of linear equations

$$b_j = \sum_{i=0}^{r-1} t_i f_i(p_j) \tag{32}$$

We call f the *extrapolation function by series solutions* of ODE. We call b_j the reference value of f at the reference point p_j.

Let us come back to our example. The linear ODE for (29) has rank $r = 11$. We set $c = 370/100 - 1/100$ and y_j's are $[0, 1/100, \ldots, 10/100]$. Then we have

$$c + y_0 = 3.69, c + y_1 = 3.70, \ldots, c + y_{10} = 3.79.$$

We construct an approximate series solution $f_i(x)$ by taking 20000 terms with the rational arithmetic.

We set the reference points $p_j = \frac{38}{100} + \frac{j}{1000}$, $p_0 = 3.8, \ldots, p_{10} = 3.81$ and construct a matrix related to (32). Numbers in the matrix are translated to approximate rational numbers to avoid the unstability problem of solving linear equations (32) with floating point numbers.

We assume that the expectation of the Euler characteristic of M_x is almost equal to the probability $P(\ell_1 > x)$ of the first eigenvalue is larger than x. In fact, we have the Euler expectation $E[\chi(M_x)] = P(\ell_1 > x) - P(\ell_2 > x)$ in this case, where ℓ_i is the i-th eigenvalue. We have $P(\ell_2 > 3.8) = 0$ by a Monte-Carlo simulation with with $1,000,000$ tries. Then we may suppose that reference values $f(p_j)$ are estimated by a Monte-Carlo simulation for $P(\ell_1 > x)$. We construct a solution $f(x)$ with these reference values. Evaluation of $f(x)$ is done with big float.

The Figure 1 is the table of values of the extrapolation function $f(x)$ obtained by the above method with the big float of 380 digits and that by simulation with $1,000,000$ samples. One simulation takes about $573s$.*.

*R and the package `mnormt` on a machine with Intel Xeon CPU(2.70GHz) and 256G memory.

x	$f(x)$	simulation
3.8133	0.051146	0.051176
3.8166	0.047517	0.047695
3.82	0.044120	0.044515

Figure 1: Numerical evaluation by extrapolation series

The solid line in the Figure 2 is obtained by this extrapolation function. The line goes to a big value at $x = 3.866$ because this x is out of the domain of convergence of this approximate series. Dots are values obtained by simulation and that on the thick solid line are values used as reference values to obtain the extrapolation function.

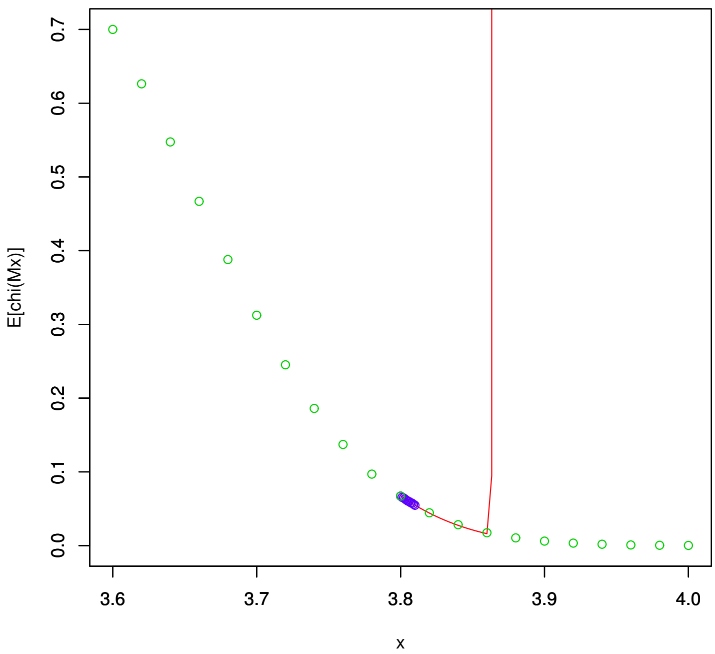

Figure 2: The extrapolation function with 20000 terms. Solid line is the extrapolation function, which diverges when $x > 3.8633$. Dots are values by simulations.

The time to obtain the series f_i with $20,000$ terms is $5661s^{\dagger}$. The time to evaluate the extrapolation function at 61 points is $14.03s$. On the other hand, if we want to obtain simulation values at 61 points, we need about $573 \times 61 = 34953s$. Thus, our extrapolation method has advantages when we want to evaluate the function $E[\chi(M_x)]$ for many x.

†Risa/Asir on a machine with Intel Xeon CPU(2.70GHz) and 256G memory.

Acknowledgement

We deeply thank Christoph Koutschan, who is the author of the package `HolonomicFunctions` used in this paper, for his helps and encouragements.

References

[1] R. J. Adler, J. E. Taylor, Random fields and geometry, Springer, 2007.

[2] K. Aomoto, Jacobi Polynomials associated with Selberg Integrals, SIAM Journal on Mathematical Analysis 18 (1987), 545–549.

[3] Y. Chikuse, Properties of Hermite and Laguerre polynomials in matrix argument and their applications, Linear Algebra and its Applications 176 (1992), 237–260.

[4] S. Chen, M. Kauers, Z. Li, and Y. Zhang, Apparent singularities of D-finite systems, Journal of Symbolic Computation, in press, (2019).

[5] F. H. Danufane, C. Siriteanu, K. Ohara, N. Takayama, Holonomic gradient method-based CDF evaluation for the largest eigenvalue of a complex noncentral Wishart matrix, arxiv:1707.02564.

[6] H. Hashiguchi, Y. Numata, N. Takayama, A. Takemura, Holonomic gradient method for the distribution function of the largest root of a Wishart matrix, Journal of Multivariate Analysis, 117 (2013), 296–312.

[7] T. Hayakawa, On the distribution of the latent roots of a positive definite random symmetric matrix I, Annals of the Institute of Statistical Mathematics 21 (1969), 1–21.

[8] T. Hibi et al., Gröbner Bases: Statistics and Software Systems, Springer, (2013).

[9] A. Hillebrand, W. Schmale, Towards a effective version of a theorem of Stafford, Journal of Symbolic Computation 32 (2001), 699–716.

[10] A. T. James, Normal Multivariate Analysis and the Orthogonal Group, The Annals of Mathematical Statistics (1954), 40–75.

[11] J. Kaneko, Selberg Integrals and Hypergeometric Functions associated with Jack Polynomials, SIAM Journal on Mathematical Analysis 24 (1993), 1086–1110.

[12] M. Kang, M. S. Alouini, Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems, IEEE Journal on Selected Areas in Communications 21 (2003), 418–426.

[13] C. Koutschan, Advanced applications of the holonomic systems approach, PhD thesis, Johannes Kepler University Linz, (2009).

[14] C. Koutschan, HolonomicFunctions user's guide, Technical Report 10-01, RISC Report Series, Johannes Kepler University Linz, Austria, (2010). http://www.risc.jku.at/publications/download/risc_3934/hf.pdf

[15] S. Kuriki, A. Takemura, Tail probabilities of the maxima of multilinear forms and their applications, The Annals of Statistics 29 (2001), 328–371.

[16] S. Kuriki, A. Takemura, Euler characteristic heuristic for approximating the distribution of the largest eigenvalue of an orthogonally invariant random matrix, Journal of Statistical Planning and Inference 138 (2008), 3357–3378.

[17] S. Kuriki, A. Takemura, Volume of tubes and the distribution of the maximum of a Gaussian random field, Selected Papers on Probability and Statistics, American Mathematical Society Translations Series 2, 227 (2009), 25–48.

[18] N. Takayama, L. Jiu, S. Kuriki, N. Takayama, Y. Zhang, Supplementary electronic material to the article *Euler Characteristic Method for the Largest Eigenvalue of a Random Matrix*. https://yzhang1616.github.io/ec1/ec1.html

[19] A. Leykin, Algorithmic proofs of two theorems of Stafford. Journal of Symbolic Computation 38 (2004), 1535–1550.

[20] Mathematica, Version 11.0, Wolfram Research, Inc., Champaign, IL, 2016.

[21] R. J. Muirhead, Aspects of multivariate statistical theory, Wiley, 2005.

[22] A. Selberg, Remarks on a multiple integral, Norsk Matematisk Tidsskrift 26 (1944), 71–78.

[23] A. Takemura, S. Kuriki, Shrinkage to smooth non-convex cone: principal component analysis as Stein estimation, Communications in Statistics: Theory and Methods 28 (1999), 651–669.

[24] M. Morse, S. S. Cairns, Critical point theory in global analysis and differential topology: an introduction, Academic Press, 1969

[25] D. Zeilberger, The method of creative telescoping, Journal of Symbolic Computation 11 (1991), 195–204.