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

Yi Zhang

Department of Mathematical Sciences
University of Texas at Dallas, USA

Joint work with Nobuki Takayama, Lin Jiu and Satoshi Kuriki

Largest Eigenvalue of Real Wishart Matrix

Let $\xi_i \in \mathbb{R}^m$ be distributed as $N_m(\mu_i, \Sigma)$.

The Wishart distribution $W_m(n, \Sigma; \Omega)$ is induced by the random matrix

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

where $\Omega = \Sigma^{-1} \sum_{i=1}^n \mu_i \mu_i^\top$ is the parameter matrix.

We call $W_m(n, \Sigma; \Omega)$ non-central if $\Omega \neq 0$.

Let $\lambda_1(W)$ be the largest eigenvalue of W. The distribution of $\lambda_1(W)$ is of particular interest in testing hypothesis.

Motivation and Previous works

Let $W_m(n, \Sigma; \Omega)$ be non-central.

Goal: Efficient evaluation of $\Pr(\lambda_1(W) \geq x)$ for many x.

Motivation and Previous works

Let $W_m(n, \Sigma; \Omega)$ be non-central.

Goal: Efficient evaluation of $\Pr(\lambda_1(W) \geq x)$ for many x.

- (James *et al.*, 1954) When $\Omega = 0$, express $\Pr(\lambda_1(W) \geq x)$ as a hypergeometric function ${}_1F_1$
- (Hashiguchi *et al.*, 2013) Efficient evaluation of ${}_1F_1$ using holonomic gradient method
- (Danufane *et al.*, 2017) In MIMO problem, evaluation of $\Pr(\lambda_1(W) \geq x)$ if W is a complex matrix and $\Omega \neq 0$.

Motivation and Previous works

Let $W_m(n, \Sigma; \Omega)$ be non-central.

Goal: Efficient evaluation of $\Pr(\lambda_1(W) \geq x)$ for many x.

- (James *et al.*, 1954) When $\Omega = 0$, express $\Pr(\lambda_1(W) \geq x)$ as a hypergeometric function ${}_1F_1$
- (Hashiguchi *et al.*, 2013) Efficient evaluation of ${}_1F_1$ using holonomic gradient method
- (Danufane *et al.*, 2017) In MIMO problem, evaluation of $\Pr(\lambda_1(W) \geq x)$ if W is a complex matrix and $\Omega \neq 0$.

Our contribution: Efficient evaluation of $\Pr(\lambda_1(W) \geq x)$ if W is a real matrix and $\Omega \neq 0$.

Euler Characteristic Method

Let $W_m(n, \Sigma; \Omega)$ be non-central and W be a real matrix.

Difficulty: No explicit formula for $\Pr(\lambda_1(W) \geq x)$.

Euler Characteristic Method

Let $W_m(n, \Sigma; \Omega)$ be non-central and W be a real matrix.

Difficulty: No explicit formula for $\Pr(\lambda_1(W) \geq x)$.

Adler, Tayler and Takemura (2000, 2005), Kuriki and Takemura (2001, 2008, 2009): Use Euler characteristic heuristic to approximate probabilities of random fields.

Fact: $\lambda_1(W)^{1/2}$ is the maximum of a Gaussian field

$$\left\{ u^\top \Xi v \mid \|u\|_{\mathbb{R}^m} = \|v\|_{\mathbb{R}^n} = 1 \right\}.$$

Idea: Approximation by the expected Euler characteristic heuristic:

$$\Pr(\lambda_1(W) \geq x) \approx E\left[\chi(M_x)\right] \quad \text{when } x \text{ is large,}$$

where M_x is a manifold induced by W and x.

Outline

- Explicit formula for the expectation of th Euler characteristic number of a manifold related to a random matrix

- Numerical evaluation for the integral formula by holonomic gradient method

Manifold of a Random Matrix

Let A be a real 2×2 random matrix. Define a manifold

$$M = \{hg^T \,|\, g \in S, h \in S\}.$$

Set

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

and

$$M_x = \{U \in M \,|\, f(U) \geq x\},$$

which is a manifold induced by A and x.

Euler Characteristic Number

Let A be a real 2×2 random matrix and M_x be the related manifold.

Recall: The Euler characteristic is defined for the surfaces of polyhedra by

$$\chi = V - E + F.$$

For convex polyhedron's surface, $\chi = 2$.

We can also define the Euler characteristic for M_x and denote it by $\chi(M_x)$.

Expectation of the Euler Characteristic Number

Let A be a real 2×2 random matrix and M_x be the related manifold.

Recall: $f(U) = \mathrm{tr}(UA), \quad U \in M_x.$

Let hg^T be a critical point of f. Take $(g, G) \in SO(2)$ and $(h, H) \in SO(2)$. Set

$$\sigma = g^T Ah, \; b = G^T AH,$$

which are singular values of A.

Theorem 1: Assume $x > 0$ and $f(U)$ is a Morse function for almost all A's. Then $E[\chi(M_x)]$ is equal to

$$\frac{1}{2} \int_x^{\infty} d\sigma \int_{-\infty}^{\infty} db \int_S G^T dg \int_S H^T dh (\sigma^2 - b^2) p(A).$$

Expectation of the Euler Characteristic Number

Recall: Approximation by the expected Euler characteristic heuristic:

$$\Pr(\lambda_1(W) \geq x) \approx E\big[\chi(M_x)\big] \quad \text{when} \quad x \quad \text{is large,}$$

where M_x is a manifold induced by $W = \Xi\Xi^T$ and x.

Goal: Efficient evaluation of the integral in Theorem 1 when W is a non-central Wishart matrix and x is large.

Expectation of the Euler Characteristic Number

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

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

Then the integral in Theorem 1 becomes

$$\int_x^{\infty} d\sigma \int_{-\infty}^{\infty} db \int_{-\infty}^{\infty} ds \int_{-\infty}^{\infty} f(\sigma, b, s, t) dt, \tag{1}$$

where

$$f = \frac{s_1 s_2 (\sigma^2 - b^2)}{(1+s^2)(1+t^2)} \exp\left\{ -\frac{1}{2} R \right\}, \quad R \in \mathbb{Q}(\sigma, b, s, t)$$

We denote (1) by $F(M, \Sigma; x)$.

Challenge for Evaluation

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

- $F(M, \Sigma; x)$ contains parameters M, Σ.
- Numerical integration for $F(M, \Sigma; x)$ is time-consuming and not reliable for many x.

Observation: the integrand of $F(M, \Sigma; x)$ is holonomic (D-finite).

Idea: Use holonomic gradient method to evaluate $F(M, \Sigma; x)$.

Holonomic Gradient Method

$f(\theta, t)$: unnormalized probability distribution function w.r.t. $t = (t_1, \ldots, t_n)$, where $\theta = (\theta_1, \ldots, \theta_m)$ is a parameter vector.

$$z(\theta) = \int_{\Omega} f(\theta, t) dt$$

is the normalizing constant. $f(t, \theta)/z(\theta)$ is a probability distribution function on Ω. Evaluation of $z(\theta)$ is a fundamental problem in statistics.

Example: $f(\theta, t) = \exp\left(\frac{-t^2}{2\theta^2}\right)$ with $\Omega = (-\infty, +\infty)$. Then

$$z(\theta) = \sqrt{2\pi\theta^2}.$$

Holonomic Gradient Method

An analytic function $f(x)$ is called holonomic or D-finite when it satisfies n linear ODE's (holonomic system)

$$\sum_{j=0}^{r_i} a_{ij} \left(\frac{\partial}{\partial x_i} \right)^j f = 0, \quad a_{ij}(x) \in \mathrm{C}[x_1, \ldots, x_n], \quad i = 1, \ldots, n.$$

Theorem (Zeilberger, 1990): If $f(x)$ is holonomic, then the integral $\int_\Omega f(x) dx_n$ is holonomic in $(x_1, \ldots, x_{n-1})$ (under some conditions on Ω).

Holonomic Gradient Method (Nakayama *et al.*, 2011): When $f(\theta, t)$ is holonomic, the normalizing constant $z(\theta)$ satisfies a system of linear PDEs, which can be constructed by Gröbner bases. Evaluate $z(\theta)$ and its derivatives by the system with methods in numerical analysis.

3 Steps of Holonomic Gradient Method

1. Construct a Pfaffian system for $z(\theta)$.
2. Evaluate numerically $z(\theta)$ and its derivatives at $\theta = \theta_0$.
3. Apply numerical analysis methods for the Pfaffian system.

Example:

$$z(\theta) = \int_\Omega \exp(\theta t) t^{1/2}(1-t)^{1/2} dt, \quad \Omega = [0,1]$$

By creative telescoping,

$$\left(\theta\partial_\theta^2 + (3-\theta)\partial_\theta - 3/2\right) z = 0, \quad \partial_\theta = \frac{\partial}{\partial\theta}$$

Then $\frac{\partial}{\partial\theta} Z = PZ$, where

$$Z = \begin{pmatrix} z \\ \frac{\partial}{\partial\theta} z \end{pmatrix}, \quad P = \begin{pmatrix} 0 & 1 \\ \frac{3}{2\theta} & -\frac{3-\theta}{\theta} \end{pmatrix}$$

Evaluation of the Expected Euler Characteristic

Recall: $E[\chi(M_x)] = F(M, \Sigma; x)$ is equal to

$$\int_x^{\infty} d\sigma \int_{-\infty}^{\infty} db \int_{-\infty}^{\infty} ds \int_{-\infty}^{\infty} f(\sigma, b, s, t) dt,$$

where f is hyperexponential over $\mathbb{Q}(\sigma, b, s, t)$. Thus, $-F'(M, \Sigma; x)$ is equal to

$$\int_{-\infty}^{\infty} db \int_{-\infty}^{\infty} ds \int_{-\infty}^{\infty} f(x, b, s, t) dt,$$

Idea: Use creative telescoping method to derive an ODE for $F'(M, \Sigma; x)$

Creative Telescoping Method

Given a holonomic function $f(\theta, t)$ with annihilator

$$\mathrm{ann}(f) \subset \mathrm{C}(\theta, t)[\partial_\theta, \partial_t].$$

Find nontrivial

$$P(\theta, \partial_\theta) + \partial_t Q(\theta, t, \partial_\theta, \partial_t) \in \mathrm{ann}(f)$$

Then $z(\theta) = \int_\Omega f(\theta, t)dt$ satisfies $P(z) = 0$ (under some conditions on Ω). We call P a telescoper for $ann(f)$.

Creative Telescoping Method

- (Zeilberger, 1990): Sylvester's dialytic elimination for multiple integrals
- (Takayama, 1992; Oaku, 1997): D-module theoretical algorithms for multiple integrals
- (Chyzak, 2000): a generalization of Gosper's algorithm for single integrals of multivariate holonomic functions
- (Koutschan, 2010): rational ansatz approach for multiple integrals
- (Bostan *et al.*, 2010, 2013; Chen *et al.*, 2015, 2016): reduction-based algorithms for single integrals of bivariate holonomic functions

Chyzak's algorithm

Given a holonomic function $f(\theta, t)$ with annihilator

$$\mathrm{ann}(f) \subset R = \mathsf{C}(\theta, t)[\partial_\theta, \partial_t].$$

We call $\dim_{\mathsf{C}}(R/\mathrm{ann}(f))$ the (holonomic) rank of $\mathrm{ann}(f)$.

Goal: Drive an ODE for

$$G(M, \Sigma; x) = \int_{-\infty}^{\infty} db \int_{-\infty}^{\infty} ds \int_{-\infty}^{\infty} f(x, b, s, t) dt,$$

where f is hyperexponential over $\mathbb{Q}(\sigma, b, s, t)$.

Using Chyzak's algorithm, find a holonomic system of rank 2 for

$$f_1(x, b, s) = \int_{-\infty}^{\infty} f(x, b, s, t) dt$$

in 5 seconds using a Linux computer with 15.10 GB RAM.

Chyzak's algorithm

Goal: Drive an ODE for

$$G(M,\Sigma;x) = \int_{-\infty}^{\infty} db \int_{-\infty}^{\infty} f_1(x,b,x)ds,$$

where $\mathrm{ann}(f_1)$ has holonomic rank 2.

Using Chyzak's algorithm, find a holonomic system of rank 6 for

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

in 16 mins by specifying M and Σ.

Question: Is it possible to compute a holonomic system for f_2 without specifying M and Σ?

Stafford Heuristic

Consider

$$R_n = \mathbb{K}(x_1, \ldots, x_n)[\partial_1, \ldots, \partial_n],$$
$$T_n = \{\partial_1^{i_1} \cdots \partial_n^{i_n} \mid (i_1, \ldots, i_n) \in \mathbb{N}^n\}.$$

Heuristic: Given a holonomic system H in R_n, compute new holonomic system H_1 in R_{n-1} s.t. $H_1 \subset (R_n \cdot H + \partial_n R_n) \cap R_{n-1}$.

1. Pick $S_1, S_2 \in T_{n-1}$.
2. Using rational ansatz method, check existence of telescoper P_i of H with support S_i, $i = 1, 2$. If P_i exists, go to step 3. Otherwise, go to step 1.
3. Compute Gröbner basis H_1 of $\{P_1, P_2\}$. If H_1 is holonomic, then output G_1. Otherwise, go to step 1.

Stafford Theorem: Every left ideal in R_n can be generated by 2 elements.

Stafford Heuristic

Goal: Drive an ODE for

$$G(M,\Sigma;x)=\int_{-\infty}^{\infty}db\int_{-\infty}^{\infty}f_1(x,b,x)ds,$$

where $\operatorname{ann}(f_1)=\langle H\rangle$ has holonomic rank 2.

1. Pick

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

2. Using rational ansatz method, find telescoper P_i of H with support S_i, $i=1,2$.

3. Compute Gröbner basis H_1 of $\{P_1,P_2\}$. We find that H_1 has holonomic rank 6.

Chyzak's algorithm vs Stafford Heuristic

Goal: Drive an ODE for

$$G(M, \Sigma; x) = \int_{-\infty}^{\infty} db \int_{-\infty}^{\infty} f_1(x, b, x) ds,$$

where $\mathrm{ann}(f_1)$ has holonomic rank 2.

Below is a table of time (seconds) for deriving holonomic systems of

$$f_2(x, b) = \int_{-\infty}^{\infty} f_1(x, b, s) ds.$$

# pars	0	1	2	3	4	5
Chyzak	976	9.8×10^4	-	-	-	-
Heuristic	43.49	394.4	8527	4.3957×10^5	-	1.5×10^6

Evaluation of the Expected Euler Characteristic

Goal: Drive an ODE for

$$G(M, \Sigma; x) = \int_{-\infty}^{\infty} f_2(x, b) db,$$

where $\mathrm{ann}(f_2)$ has rank 6 (Recall: $G(M, \Sigma; x) = -F'(M, \Sigma; x)$).

Example 1: Set

$$\Sigma^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \quad M = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Using Heuristic, find an 11-th order ODE $P(F) = 0$ of $F(M, \Sigma; x)$.
By numerical solving of $P(F) = 0$, get

x	1	2	3	4	5
HGM	0.745835	0.567729	0.144879	0.0146728	0.000582526
mc	0.745802	0.567623	0.144986	0.0146901	0.0005933

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

Evaluation of the Expected Euler Characteristic

Example 2: Set

$$\Sigma^{-1} = \begin{pmatrix} 10^3 & 0 \\ 0 & 10^2 \end{pmatrix} \quad M = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}.$$

Using Heuristic, find an 11-th order ODE $P(F) = 0$ of $F(M, \Sigma; x)$.

Difficulty:

- Initial value: numerical integration is time-consuming and not reliable.
- Numerical solving of ODEs: the Runge-Kutta method only works locally since $F(M, \Sigma; x)$ is not dominant among solutions of $P(F) = 0$.

Recall: Let $f_1, \ldots, f_n$ be a basis of solutions of a linear ODE $L(y) = 0$. A solution f of $L(y) = 0$ is dominant if

$$\lim_{x \to \infty} \frac{|f_i(x)|}{|f(x)|} < \infty, \quad i = 1, \ldots, n.$$

Evaluation of the Expected Euler Characteristic

Let $P(y) = 0$ be the r-th order linear ODE of $F(M, \Sigma; x)$.

Idea: Compute approximation series solutions of the linear ODE $P(y) = 0$ and use them to extrapolate results by simulations.

1. Construct approximation series solutions $f_1, \ldots, f_r$ of $P(y) = 0$ up to $20,000$ terms.
2. Make an ansatz $f(x) = \sum_{i=0}^{r-1} t_i f_i(x)$, where t_i's are unknown. Chose $x = p_j$ for $j = 0, \ldots, r-1$. Then

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

Evaluation of the Expected Euler Characteristic

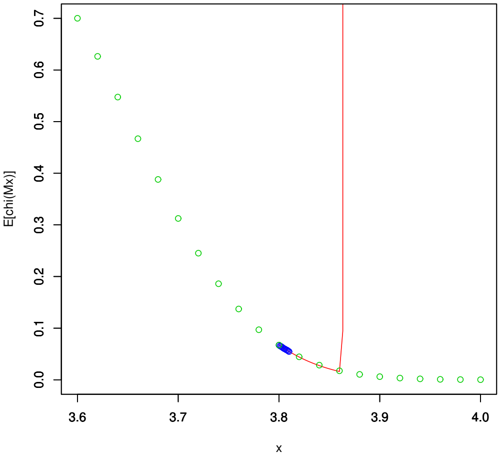
$$f(p_j) = \sum_{i=0}^{r-1} t_i f_i(p_j), \quad j = 0, 1, \ldots, r-1.$$

3. Compute $f(p_j)$ by Monte-Carlo simulation and then determine t_i's by solving linear equations.

4. Use $f(x) = \sum_{i=0}^{r-1} t_i f_i(x)$ to extrapolate $F(M, \Sigma; x)$ at target points.

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

Evaluation of the Expected Euler Characteristic



The extrapolation function $f(x)$ with 20,000 terms. Solid line is $f(x)$, which diverges when $x > 3.8633$. Dots are values by simulations.

Conclusion

Let $W_m(n, \Sigma; \Omega)$ be non-central and W be a real matrix.

- Approximate formula of $\Pr(\lambda_1(W) \geq x)$ by Euler characteristic method
- Numerical evaluation for the integral formula by holonomic gradient method

Conclusion

Let $W_m(n, \Sigma; \Omega)$ be non-central and W be a real matrix.

- Approximate formula of $\Pr(\lambda_1(W) \geq x)$ by Euler characteristic method
- Numerical evaluation for the integral formula by holonomic gradient method

Thanks!