On the probability that all eigenvalues of random matrices lie within an interval

Marco Chiani

marco.chiani@unibo.it

University of Bologna, Italy

Dept. of Electrical, Electronic, and Information Eng. “Guglielmo Marconi” (DEI)

Workshop Complex ODEs: Asymptotics, Orthogonal Polynomials and Random Matrices
Pisa, May 14-18, 2018

Outline

- Three communication theory applications of RMT
- Probability that all eigenvalues of a random matrix $\boldsymbol{M}$ lie within an arbitrary interval $[a, b]$,

$$\psi(a, b) \triangleq \Pr\{a \leq \lambda_{\min}(\boldsymbol{M}), \lambda_{\max}(\boldsymbol{M}) \leq b\}$$

  when $\boldsymbol{M}$ is a real or complex finite dimensional Wishart, double Wishart, GOE, GUE.
- Efficient recursive formulas for $\psi(a, b)$ for large matrices.
- Probability that all eigenvalues are within the limiting spectral support (Marčenko-Pastur or the semicircle laws) tends for large dimensions to the universal values 0.6921 and 0.9397 for the real and complex cases, respectively.

Three communication theory applications of RMT

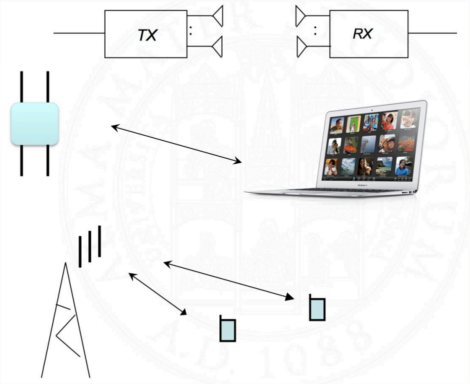
1. Wireless MIMO: tremendous $\Uparrow$ data with wireless devices $\Rightarrow$ need for increased [bits/s/Hs] $\Rightarrow$ use of multiple antenna systems (MIMO)

   Basic equations $\boldsymbol{y} = \boldsymbol{Hx} + \boldsymbol{n}$, all complex matrices/vectors.

   System depends on the eigenvalues/eigenvectors of $\boldsymbol{H}^\dagger \boldsymbol{H}$
2. Spectrum sensing: Tremendous $\Uparrow$ number of wireless sources $\Rightarrow$ need for spectrum estimation / source detection and enumeration methods.
   Typically complex matrices involved.
3. Compressive Sensing: Tremendous $\Uparrow$ of IoT data $\Rightarrow$ efficient acquisition and compression schemes.
   $\Rightarrow$ need for signal processing techniques for efficient data acquisition and communication.
   Typically real matrices involved.

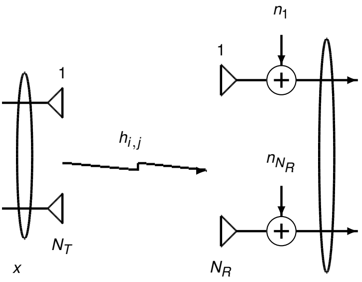
1) Wireless MIMO

MIMO is *the* technology for high data rate wireless



Now going towards a large number of antenna elements ("massive MIMO")

1) Wireless MIMO



$\mathbf{H} = \left\{ h_{i,j} \right\}$

$\mathbf{y} = \mathbf{H}x + \mathrm{n}$

Frequency-flat MIMO channel

$$\mathbf{y}(k) = \mathbf{H}\,\mathbf{x}(k) + \mathbf{n}(k) \tag{1}$$

Noise vector $\mathbf{n}(k) \in \mathbb{C}^{N_R}$ are independent, identically distributed (i.i.d.) with circularly-symmetric complex Gaussian (CSCG) entries and covariance matrix $\mathbf{R}_n = \mathbb{E}\left\{\mathbf{n}(k)\mathbf{n}^H(k)\right\} = \sigma^2\mathbf{I}$.

1) Wireless MIMO

Total transmitted power $\mathbb{E}\left\{\mathbf{x}^H(k)\mathbf{x}(k)\right\} \leq P_t$.

- If the channel is known at the receiver only the capacity is obtained with $\mathbf{x}$ Gaussian

$$
\begin{aligned}
C(\mathbf{H}) &= \log_2 \det\left(\mathbf{I} + \frac{P_t}{N_T\sigma^2}\mathbf{H}\mathbf{H}^H\right) \qquad \text{[info bits/s/Hz]} \qquad (2)\\
&= \sum_{\ell=1}^{n} \log_2\left(1 + \frac{P_t}{N_T\sigma^2}\lambda_\ell\right) \qquad \text{[info bits/s/Hz]}
\end{aligned}
$$

  where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ are the nonzero eigenvalues of $\mathbf{H}\mathbf{H}^H$.

- In the presence of multipath electromagnetic propagation the elements of $\mathbf{H}$ can be modeled as Gaussian r.v.s. $\Rightarrow$ $\mathbf{H}\mathbf{H}^H$ is Wishart.

Gaussian Orthogonal Ensemble (GOE) and Gaussian Unitary Ensemble (GUE)

Studied in physics

- Gaussian Orthogonal Ensemble (GOE): Symmetric matrices with Gaussian entries:

$$\underline{\underline{G}} = \underline{\underline{X}} + \underline{\underline{X}}^T$$

  where the entries of $\underline{\underline{X}}$ are i.i.d. zero mean Gaussian (e.g. $\mathcal{N}(0, 1/4)$.)

- Gaussian Unitary Ensemble (GUE): Hermitian matrices with Gaussian entries:

$$\underline{\underline{G}} = \underline{\underline{X}} + \underline{\underline{X}}^H$$

  where the entries of $\underline{\underline{X}}$ are i.i.d. zero mean complex Gaussian (e.g. $\mathcal{CN}(0, 1/4)$)

M. L. Mehta, Random Matrices, 1991

- *Central Wishart matrix with one-sided correlation*: the $n \times n$ random symmetric real matrix

$$\underline{\underline{W}} = \underline{\underline{X}}\,\underline{\underline{X}}^T \sim \mathcal{W}_n(m, \underline{\underline{\Sigma}})$$

$\underline{\underline{X}}$: a Gaussian $n \times m$ matrix with i.i.d. columns $\underline{X}_j \sim \mathcal{N}(\underline{0}, \underline{\underline{\Sigma}})$.

  - n is the dimension of the random Gaussian vector
  - m is the number of samples (dimension of the sample, degrees of freedom)
  - $\underline{\underline{\Sigma}}$ gives the correlation within each column

- Complex case: replace $()^T$, and $\mathcal{N}()$ with $()^H$, and $\mathcal{CN}()$.
- The sample covariance matrix (SCM) for samples of a Gaussian vector is Wishart distributed.
- Studied in multivariate statistics (real case) since Wishart

J. Wishart "The generalised product moment distribution in samples from a normal multivariate population", Biometrika, 1928

T. W. Anderson, An Introduction to Multivariate Statistical Analysis, 2003

R. J. Muirhead, Aspects of Multivariate Statistical Theory, 1982

Some known results

- The joint distribution of the elements of $\underline{\underline{X}}$ with Gaussian i.i.d. columns (real or complex)
- The joint distribution of the elements of $\underline{\underline{W}} = \underline{\underline{X}}\,\underline{\underline{X}}^H \sim \mathcal{W}_n(m, \underline{\underline{\Sigma}})$ (real or complex)
- The joint distribution of the eigenvalues for GOE $\underline{\underline{G}} = \underline{\underline{X}} + \underline{\underline{X}}^T$ (no correlation)
- The joint distribution of the eigenvalues for GUE $\underline{\underline{G}} = \underline{\underline{X}} + \underline{\underline{X}}^H$ (no correlation)
- The joint distribution of the eigenvalues for $\mathcal{W}_n(m, \underline{\underline{I}})$ (real or complex)
- Asymptotic distribution of a randomly picked eigenvalue for large dimensional matrices (Gaussian and Wishart)
- Asymptotic distribution of the largest eigenvalue for large dimensional matrices (Gaussian and Wishart)

Example: joint p.d.f. of the eigenvalues, uncorrelated Wishart $\mathbf{W} \sim \mathcal{W}_n(m, \underline{I})$, complex

The joint p.d.f. of the ordered eigenvalues $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_{n_{\min}}$ of the complex Wishart $\mathbf{W} \sim \mathcal{W}_n(m, \underline{I})$ is

$$f(x_1, \ldots, x_{n_{\min}}) = K \prod_{i=1}^{n_{\min}} e^{-x_i} x_i^{n_{\max}-n_{\min}} \cdot \prod_{i<j}^{n_{\min}} (x_i - x_j)^2$$

where K is a normalizing constant and $n_{\max} = \max\{n, m\}$.
Vandermonde matrix $\mathbf{V}_1(\mathbf{x})$

$$\mathbf{V}_1(\mathbf{x}) \triangleq \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_{n_{\min}} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n_{\min}-1} & x_2^{n_{\min}-1} & \cdots & x_{n_{\min}}^{n_{\min}-1} \end{bmatrix}.$$

Since $|\mathbf{V}_1(\mathbf{x})|^2 = \prod_{i<j}^{n_{\min}} (x_i - x_j)^2$

$$f(x_1, \ldots, x_{n_{\min}}) = K \, |\mathbf{V}_1(\mathbf{x})|^2 \prod_{i=1}^{n_{\min}} e^{-x_i} x_i^{n_{\max}-n_{\min}}$$

Example: large dimensional matrices (asymptotic) results

- The asymptotic distribution of an unordered eigenvalue for GOE and GUE $\underline{\underline{G}} = \underline{\underline{X}} + \underline{\underline{X}}^H$ (the Wigner semi-circle law)
- The asymptotic distribution of an unordered eigenvalue for Wishart $\underline{\underline{W}} = \underline{\underline{X}}\,\underline{\underline{X}}^H \sim \mathcal{W}_n(m, \underline{\underline{I}})$ (the Marčenko-Pastur law)

The limiting Wigner and Marčenko-Pastur laws apply not just when $\underline{\underline{X}}$ is Gaussian, but for a wide class of distributions for the entries of $\underline{\underline{X}}$ (assumed always i.i.d.)

Bai & Silverstein, Spectral Analysis of Large Dimensional Random Matrices, Science Press, 2006.

$$\mathbf{y}(k) = \mathbf{H}\,\mathbf{x}(k) + \mathbf{n}(k) \qquad C(\lambda_1, \ldots, \lambda_{n_{\min}}) = \sum_{\ell=1}^{n_{\min}} \log_2\left(1 + \frac{P_t}{n_T \sigma^2}\lambda_\ell\right)$$

We have to study for instance:

- $\mathrm{E}\left\{C(\lambda_1, \ldots, \lambda_{n_{\min}})\right\} = \int \ldots \int_{\mathcal{D}_{\text{ord}}(0,\infty)} f(x_1, \ldots, \mathbf{x}_{n_{\min}}) \sum_{\ell=1}^{n_{\min}} \log_2\left(1 + \frac{P_t}{n_T \sigma^2} x_\ell\right) d\mathbf{x}$
- $\Pr\left[C(\lambda_1, \ldots, \lambda_{n_{\min}}) \le x\right]$
- $\Pr(\lambda_1 \le x) = \int \ldots \int_{\mathcal{D}_{\text{ord}}(0,x)} f(x_1, \ldots, x_{n_{\min}}) d\mathbf{x}$

where $f(x_1, \ldots, x_{n_{\min}})$ is the joint p.d.f. of the ordered eigenvalues of $\mathbf{H}\mathbf{H}^H$,
$\mathcal{D}_{\text{ord}}(a, b) = \left\{b > x_1 > x_2 > \cdots > x_{n_{\min}} > a\right\}$, $d\mathbf{x} = dx_1\, dx_2 \cdots dx_{n_{\min}}$

In general, for MIMO

... we need to evaluate expressions like

$$\int \cdots \int f(x_1, \ldots, x_{n_{\min}}) \prod_{\ell=1}^{n_{\min}} \xi(x_\ell) d\mathbf{x}$$

or

$$\int \cdots \int f(x_1, \ldots, x_{n_{\min}}) \sum_{\ell=1}^{n_{\min}} \xi(x_\ell) d\mathbf{x}$$

or

$$\int \cdots \int f(x_1, \ldots, x_{n_{\min}}) \prod_{\ell=1}^{n_{\min}} \xi_\ell(x_\ell) d\mathbf{x}$$

where $f(x_1, \ldots, x_{n_{\min}})$ is the joint p.d.f. of the ordered eigenvalues of a complex correlated Wishart matrix

Two steps

1. Find the joint p.d.f. of the eigenvalues of a Wishart matrix with correlation

$$f(x_1, \dots, x_{n_{\min}})$$

2. Find a method to evaluate the multiple integrals

$$\int \cdots \int f(x_1, \dots, x_{n_{\min}}) \prod_{\ell=1}^{n_{\min}} \xi(x_\ell) d\mathbf{x}$$

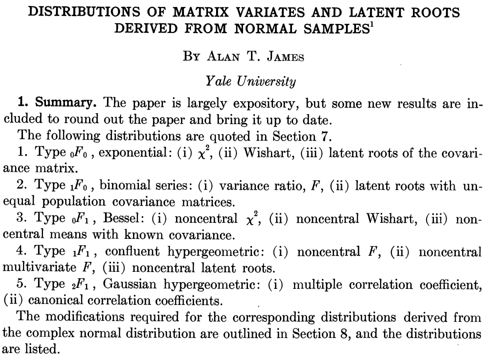
or

$$\int \cdots \int f(x_1, \dots, x_{n_{\min}}) \sum_{\ell=1}^{n_{\min}} \xi(x_\ell) d\mathbf{x}$$

or

$$\int \cdots \int f(x_1, \dots, x_{n_{\min}}) \prod_{\ell=1}^{n_{\min}} \xi_\ell(x_\ell) d\mathbf{x}$$

A. T. James, Annals Math. Stat., 1964



DISTRIBUTIONS OF MATRIX VARIATES AND LATENT ROOTS DERIVED FROM NORMAL SAMPLES[1]

By Alan T. James

Yale University

1. Summary. The paper is largely expository, but some new results are included to round out the paper and bring it up to date.

The following distributions are quoted in Section 7.

1. Type ${}_0F_0$, exponential: (i) χ^2, (ii) Wishart, (iii) latent roots of the covariance matrix.
2. Type ${}_1F_0$, binomial series: (i) variance ratio, F, (ii) latent roots with unequal population covariance matrices.
3. Type ${}_0F_1$, Bessel: (i) noncentral χ^2, (ii) noncentral Wishart, (iii) noncentral means with known covariance.
4. Type ${}_1F_1$, confluent hypergeometric: (i) noncentral F, (ii) noncentral multivariate F, (iii) noncentral latent roots.
5. Type ${}_2F_1$, Gaussian hypergeometric: (i) multiple correlation coefficient, (ii) canonical correlation coefficients.

The modifications required for the corresponding distributions derived from the complex normal distribution are outlined in Section 8, and the distributions are listed.

James [1964]: The joint p.d.f. of the (real) ordered eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{n_{\min}}$ of $\mathbf{W}$ is

$$f(x_1, \ldots, x_{n_{\min}}) = K \, |\mathbf{\Sigma}|^{-n_{\max}} \, {}_0\tilde{F}_0 \left(-\mathbf{\Sigma}^{-1}, \mathbf{W} \right) |\mathbf{W}|^{n_{\max} - n_{\min}} \cdot \prod_{i<j}^{n_{\min}} (x_i - x_j)^2$$

where ${}_0\tilde{F}_0\,(\mathbf{A}, \mathbf{B})$ is the hypergeometric function of Hermitian matrix arguments, defined as a series involving *zonal polynomials*. These polynomials are in general very difficult to manage and the equation above does not lend itself into a tractable form for further analysis.

James [1964]: The hypergeometric functions of two Hermitian $m \times m$ matrices $\mathbf{\Lambda}$ and $\mathbf{W}$ are

$$ {}_p\tilde{F}_q\left(a_1,\ldots,a_p;b_1,\ldots,b_q;\mathbf{\Lambda},\mathbf{W}\right) \triangleq \sum_{k=0}^{\infty}\sum_{\kappa}\frac{(a_1)_\kappa\cdots(a_p)_\kappa}{(b_1)_\kappa\cdots(b_q)_\kappa}\frac{C_\kappa(\mathbf{\Lambda})C_\kappa(\mathbf{W})}{k!C_\kappa(\mathbf{I}_m)} $$

where $C_\kappa(\cdot)$ is a symmetric homogeneous polynomial of degree k in the eigenvalues of its argument, called *zonal polynomial*, the sum $\sum_\kappa$ is over all partitions of k, i.e., $\kappa=(k_1,\ldots,k_m)$ with $k_1 \geq k_2 \geq \cdots \geq k_m \geq 0$, $k_1+k_2+\cdots+k_m=k$

$\Rightarrow$ the expression

$$ f(x_1,\ldots,x_{n_{\min}}) = K\left|\mathbf{\Sigma}\right|^{-n_{\max}}\,{}_0\tilde{F}_0\left(-\mathbf{\Sigma}^{-1},\mathbf{W}\right)\left|\mathbf{W}\right|^{n_{\max}-n_{\min}}\cdot\prod_{i<j}^{n_{\min}}(x_i-x_j)^2 $$

given in [James1964] is not useful at all!

An important forgotten result

Lemma (Khatri, 1970)

Let $\mathbf{\Lambda} = \lambda_1, \dots \lambda_m$ *and* $\mathbf{W} = w_1, \dots w_m$ *with* $\lambda_1 > \cdots > \lambda_m$ *and* $w_1 > \cdots > w_m$. *Then we have*

$$ {}_p\tilde{F}_q\left(a_1, \dots, a_p; b_1, \dots, b_q; \mathbf{\Lambda}, \mathbf{W}\right) = \Gamma_{(m)}(m) \frac{\psi_q^{(m)}(b)}{\psi_p^{(m)}(a)} \frac{|\mathbf{G}|}{\prod_{i<j}\left(\lambda_i - \lambda_j\right) \prod_{i<j}\left(w_i - w_j\right)} $$

where $\Gamma_{(m)}(n) \triangleq \prod_{i=1}^{m}(n-i)!$, $\psi_q^{(m)}(b) = \prod_{i=1}^{m}\prod_{j=1}^{q}(b_j - i + 1)^{i-1}$ *and the ij element of the* $(m \times m)$ *matrix* $\mathbf{G}$ *is defined in terms of hypergeometric functions of scalar arguments as follows*

$$ g_{i,j} = {}_pF_q\left(\tilde{a}_1, \dots, \tilde{a}_p; \tilde{b}_1, \dots, \tilde{b}_q; \lambda_i w_j\right) $$

with $\tilde{a}_i = a_i - m + 1$ *and* $\tilde{b}_i = b_i - m + 1$.

First result: joint p.d.f. of the eigenvalues for correlated complex Wishart

Lemma (M.C., M. Win, A. Zanella, IEEE Trans. on Information Theory, 2003)

Let $\mathbf{W} \sim \mathcal{W}_n(m, \underline{\underline{\Sigma}})$ *be a complex Wishart matrix. Denote* $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{n_{min}} \geq 0$ *the ordered eigenvalues of* $\mathbf{\Sigma}$*. Then, the joint p.d.f. of the ordered eigenvalues of* $\mathbf{W}$ *is*

$$f(x_1, \ldots, x_{n_{min}}) = K_{\mathbf{\Sigma}} \left| \mathbf{E}(\mathbf{x}, \boldsymbol{\sigma}) \right| \cdot \left| \mathbf{V}(\mathbf{x}) \right| \cdot \prod_{j=1}^{n_{min}} x_j^{n_{max}-n_{min}}$$

where $K_{\mathbf{\Sigma}}$ *is a normalizing constant,* $\mathbf{V}(\mathbf{x}) = \{x_j^{i-1}\}$ *is the* $(n_{min} \times n_{min})$ *Vandermonde matrix, and* $\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma})$ *is defined by*

$$\mathbf{E}(\mathbf{x}, \boldsymbol{\sigma}) \triangleq \begin{bmatrix} e^{-\frac{x_1}{\sigma_1}} & e^{-\frac{x_2}{\sigma_1}} & \cdots & e^{-\frac{x}{\sigma_1}} \\ e^{-\frac{x_1}{\sigma_2}} & e^{-\frac{x_2}{\sigma_2}} & \cdots & e^{-\frac{x}{\sigma_2}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-\frac{x_1}{\sigma_{n_{min}}}} & e^{-\frac{x_2}{\sigma_{n_{min}}}} & \cdots & e^{-\frac{x_{n_{min}}}{\sigma_{n_{min}}}} \end{bmatrix}.$$

Note that the new expression for joint p.d.f. involves a product of determinants

Lemma (M.C., M. Win, A. Zanella, IEEE Trans. on Information Theory, 2003)

Given two arbitrary $N \times N$ *matrices* $\mathbf{\Phi}(\mathbf{x})$, $\mathbf{\Psi}(\mathbf{x})$ *with ij elements* $\Phi_i(x_j)$ *and* $\Psi_i(x_j)$*, and an arbitrary function* $\xi(\cdot)$ *the following identity holds:*

$$\int \cdots \int_{\mathcal{D}} |\mathbf{\Phi}(\mathbf{x})| \cdot |\mathbf{\Psi}(\mathbf{x})| \prod_{i=1}^{N} \xi(x_i) d\mathbf{x} = N! \det \left(\left\{ \int_a^b \Phi_i(x) \Psi_j(x)\, \xi(x) dx \right\}_{i,j=1\ldots,N} \right)$$

where the multiple integral is over the domain $\mathcal{D} = \{a \le x_1 \le b, a \le x_2 \le b, \ldots, a \le x_N \le b\}$ *and* $d\mathbf{x} = dx_1\, dx_2 \cdots dx_N$.

Two steps:

- use the joint p.d.f. of the eigenvalues as product of determinants of matrices

$$f(x_1, \ldots, x_{n_{\min}}) = K_{\mathbf{\Sigma}} \left|\mathbf{E}\left(\mathbf{x}, \boldsymbol{\sigma}\right)\right| \cdot \left|\mathbf{V}(\mathbf{x})\right| \cdot \prod_{j=1}^{n_{\min}} x_j^{n_{\max} - n_{\min}}$$

- integrate the joint p.d.f. using the identity

$$\int \cdots \int_{\mathcal{D}} \left|\mathbf{\Phi}\left(\mathbf{x}\right)\right| \cdot \left|\mathbf{\Psi}\left(\mathbf{x}\right)\right| \prod_{i=1}^{N} \xi(x_i) d\mathbf{x} = N! \det\left(\left\{\int_a^b \Phi_i(x)\Psi_j(x)\, \xi(x) dx\right\}_{i,j=1\ldots,N}\right)$$

- joint p.d.f. of the eigenvalues for a correlated (one-sided) central Wishart matrix as product of determinants of matrices (2003)

  This was the first usable expression, since that of [James1964] is of no practical use
- distribution of the Shannon capacity for MIMO with correlation (expected value, characteristic function)
- marginal distribution of an arbitrary subset of the eigenvalues (largest, smallest, first two eigenvalues, ...)
- extension to the case where the covariance matrix $\underline{\underline{\Sigma}}$ has eigenvalues with arbitrary multiplicity

Lemma (M.C., M. Win, H. Shin, IEEE Trans. on Information Theory, 2010)

Let $\mathbf{H}$ *be a complex Gaussian* $(p \times n)$ *random matrix with zero-mean, unit variance, independent, identically distributed (i.i.d.) entries and let* $\mathbf{\Phi}$ *be an* $(n \times n)$ *positive definite matrix. The joint probability density function (p.d.f.) of the (real) non-zero ordered eigenvalues* $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{n_{min}} \geq 0$ *of the* $(p \times p)$ *quadratic form* $\mathbf{W} = \mathbf{H}\mathbf{\Phi}\mathbf{H}^\dagger$ *is given by*

$$f(x_1, \ldots, x_{n_{min}}) = K \, |\mathbf{V}(\mathbf{x})| \, \left|\tilde{\mathbf{G}}(\mathbf{x}, \boldsymbol{\mu})\right| \prod_{i=1}^{n_{min}} x_i^{p-n_{min}}$$

where $\mathbf{V}(\mathbf{x}) = \{x_j^{i-1}\}$ *is the* $(n_{min} \times n_{min})$ *Vandermonde matrix,* K *is a constant and* $\mu_{(1)} > \mu_{(2)} \cdots > \mu_{(L)}$ *are the* L *distinct eigenvalues of* $\mathbf{\Phi}^{-1}$*, with multiplicities* $m_1, \ldots, m_L$ *such that* $\sum_{i=1}^{L} m_i = n$*.*
The $(n \times n)$ *matrix* $\tilde{\mathbf{G}}(\mathbf{x}, \boldsymbol{\mu})$ *has elements*

$$\tilde{g}_{i,j} = \begin{cases} (-x_j)^{d_i} \, e^{-\mu_{(e_i)} x_j} & j = 1, \ldots, n_{min} \\ [n-j]_{d_i} \, \mu_{(e_i)}^{n-j-d_i} & j = n_{min}+1, \ldots, n \end{cases}$$

Other results for correlated complex Wishart

- Expected value, moments and cumulative distribution of the i-th ordered eigenvalue
- Marginal joint distribution of L arbitrary unordered eigenvalues
- Probability that all eigenvalues are within a given interval
- Expected value, moments and cumulative distribution of L unordered eigenvalues joint moments, m.g.f.
- Marginal distribution of an eigenvalue as a mixture of gamma

Definition (M.C., M. Win, A. Zanella 2003)

Given a rank 3 tensor $\mathbf{A} = \left\{a_{i,j,k}\right\}_{i,j,k=1,\ldots,N}$, we define the operator $\mathcal{T}\left(\mathbf{A}\right)$ as

$$\mathcal{T}\left(\mathbf{A}\right) \triangleq \sum_{\mu}(\mu)\sum_{\alpha}(\alpha)\prod_{k=1}^{N} a_{\mu_k,\alpha_k,k} \ ,$$

where the sums are over all possible permutations, μ and α, of the integers $1, \ldots, N$.

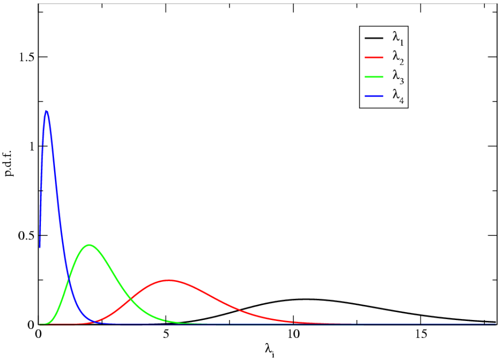
Theorem (M.C., M. Win, A. Zanella 2003)

Given two arbitrary $N \times N$ matrices $\boldsymbol{\Phi}\left(\mathbf{x}\right), \boldsymbol{\Psi}\left(\mathbf{x}\right)$ with (i,j) elements $\Phi_i(x_j)$ and $\Psi_i(x_j)$, and arbitrary functions $\xi_i(\cdot)$ the following identity holds:

$$\int \cdots \int_{\mathcal{D}} \left|\boldsymbol{\Phi}\left(\mathbf{x}\right)\right| \cdot \left|\boldsymbol{\Psi}\left(\mathbf{x}\right)\right| \prod_{k=1}^{N} \xi_k(x_k) d\mathbf{x} = \mathcal{T}\left(\left\{\int_a^b \Phi_i(x)\Psi_j(x)\,\xi_k(x) dx\right\}_{i,j,k=1,\ldots,N}\right),$$

where the multiple integral is over the domain $\mathcal{D} = \{a \le x_1 \le b, a \le x_2 \le b, \ldots, a \le x_N \le b\}$ and $d\mathbf{x} = dx_1\, dx_2 \cdots dx_N$.

Example: marginal p.d.f. of $\lambda_1, \lambda_2, \lambda_3$ and λ_4 for a central Wishart matrix with $p = 4$ and $n = 5$ (M.C., A. Zanella PIMRC 2008).



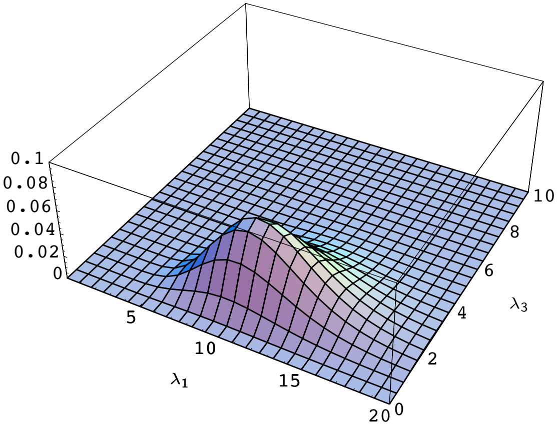


Figure: Joint probability density function of λ_1 and λ_3 of the uncorrelated central Wishart matrix with $p = 4$ and $n = 5$.

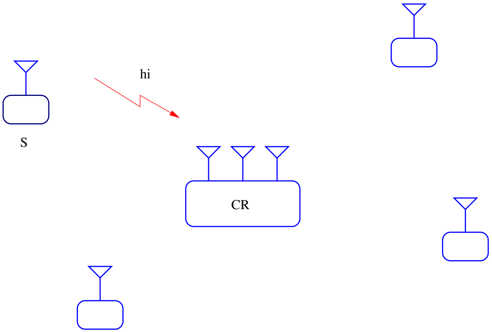
2) Spectrum sensing

General goal: find if someone is transmitting in a given area and frequency band.

- A detector can be used to distinguish between ”noise only” and ”signal(s) + noise”.
- Can we also estimate the <u>number</u> of signals?

2) Spectrum Sensing Scenario

Assuming the receiver is equipped with multiple antenna elements (sensors), spatial correlation can be exploited to discriminate signal sources from the spatially uncorrelated thermal noise.



2) Spectrum Sensing

Received N_R-dimensional vector

$$\mathbf{y}(t) = \sum_{i=1}^{N_T} \mathbf{h}_i x_i(t) + \mathbf{n}(t) = \mathbf{H}\,\mathbf{x}(t) + \mathbf{n}(t) \tag{3}$$

where

- N_T is the (unknown) number of signals (active transmitters)
- $\mathbf{n}(t) \in \mathbb{C}^{N_R}$ is the thermal noise, $\mathbf{n}(t) \sim \mathcal{CN}_{N_R}\left(\mathbf{0}, \sigma^2 \mathcal{E}_{N_R}\right)$
- $x_i(t) \in \mathbb{C}$ is the symbol transmitted by the $i^{\underline{\text{th}}}$ source at time t
- $\mathbf{h}_i \in \mathbb{C}^{N_R},\ i = 1, \dots, N_T$, are linearly independent vectors describing the gain of the radio channel between $i^{\underline{\text{th}}}$ transmitting source and the N_R receiving antennas.

Problem: determine the number of sources N_T by observing n_S samples $\mathbf{y}(t_1), \cdots, \mathbf{y}(t_{n_S})$ over which the number of sources and the channel do not change, assuming **nothing** is known at the receiver.

2) Equivalent $N_R \times N_T$ virtual MIMO formulation

$$\mathbf{y}(t) = \mathbf{H}\,\mathbf{x}(t) + \mathbf{n}(t) \tag{4}$$

where $\mathbf{x}(t) = (x_1(t), \ldots, x_{N_T}(t))^T$ and $\mathbf{H} = [\mathbf{h}_1 | \cdots | \mathbf{h}_{N_T}]$.

- The only information available is the structure of the model. ($\mathbf{H}$, N_T, σ^2, $\mathbf{x}(t)$ and $\mathbf{n}(t)$ are unknowns)
- Assume iid samples:
  $\mathbf{x}(t_i) \sim \mathcal{CN}_{N_T}(\mathbf{0}, \mathbf{R}_x)$, with $\mathbf{x}(t_1), \mathbf{x}(t_2)$, ... statistically independent.
  $\mathbf{n}(t_i) \sim \mathcal{CN}_{N_R}\left(\mathbf{0}, \sigma^2\mathcal{E}\right)$, with $\mathbf{n}(t_1), \mathbf{n}(t_2)$, ... statistically independent.
  Put the observed vectors in a $N_R \times n_S$ matrix
  $$\mathbf{Y} = \left[\mathbf{y}(t_1) | \cdots | \mathbf{y}(t_{n_S})\right] \tag{5}$$
- For a given $\mathbf{H}$ the observed vector has covariance matrix
  $$\begin{aligned} \mathbf{R} &= \mathbb{E}\left\{\mathbf{y}(t)\mathbf{y}^\dagger(t)\right\} = \mathbf{H}\,\mathbb{E}\left\{\mathbf{x}(t)x^\dagger(t)\right\}\,\mathbf{H}^\dagger + \sigma^2\mathcal{E} \\ &= \mathbf{H}\mathbf{R}_x\mathbf{H}^\dagger + \sigma^2\mathcal{E}. \end{aligned} \tag{6}$$
- The observed matrix is $\mathbf{Y} \sim \mathcal{CN}_{N_R, n_S}\left(\mathbf{0}, \mathbf{R}, \mathcal{E}_{n_S}\right)$, and the matrix $W = \mathbf{Y}\mathbf{Y}^\dagger \in \mathbb{C}^{N_R \times N_R}$ is Wishart.

2) The multiplicity of the smallest eigenvalue of $\mathbf{R}$

Ordered eigenvalues of $\mathbf{R}$: $\lambda_1 \geq \ldots \geq \lambda_{N_R}$.
If $N_T < N_R$, the smallest $N_R - N_T$ eigenvalues of $\mathbf{R}$ are all equal to σ^2.
Example: $N_R = 5$ antennas, $N_T = 2$ sources

$$\mathbf{R} = \mathbf{U} \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3 = \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & \lambda_4 = \sigma^2 & 0 \\ 0 & 0 & 0 & 0 & \lambda_5 = \sigma^2 \end{bmatrix} \mathbf{U}^\dagger$$

$\Rightarrow$ determining the number of source $\equiv$ determining of the multiplicity of the smallest eigenvalue of $\mathbf{R}$.

Unfortunately, $\mathbf{R}$ is not available at the receiver $\Rightarrow$ use its estimate based on the available n_S observations.

2) The problem seen as estimate of the multiplicity of the smallest eigenvalue of **R** based on the sample covariance matrix

The sample covariance matrix (SCM) is

$$\hat{\mathbf{R}} = \frac{1}{n_S} \sum_{i=1}^{n_S} \mathbf{y}(t_i)\mathbf{y}^\dagger(t_i) = \frac{1}{n_S}\mathbf{Y}\mathbf{Y}^\dagger = \frac{1}{n_S}\mathcal{T}\mathcal{W}. \tag{7}$$

This is an estimate of the covariance matrix **R**.

The problem of determining the number of sources is reformulated as:

find an estimate of the multiplicity
of the smallest eigenvalue of **R** based on $\hat{\mathbf{R}}$.

2) Estimate the multiplicity of the smallest eigenvalue of **R** based on the sample covariance matrix

The SCM is

$$\hat{\mathbf{R}} = \frac{1}{n_S}\sum_{i=1}^{n_S}\mathbf{y}(t_i)\mathbf{y}^{\dagger}(t_i) = \frac{1}{n_S}\mathbf{Y}\mathbf{Y}^{\dagger} = \frac{1}{n_S}W.$$

This is an estimate of the covariance matrix **R**.

- The problem of determining if there is only noise is reformulated as:

  test the hypothesis that all eigenvalues of **R** are equal, based on $\hat{\mathbf{R}}$ (sphericity test).

- The problem of determining the number of sources is reformulated as:

  find an estimate of the multiplicity of the smallest eigenvalue of **R** based on $\hat{\mathbf{R}}$.

2) Some sphericity tests proposed for Spectrum Sensing in CR (Eigenvalue-based threshold tests)

$$T \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta$$

Possible test metrics T:

- GRLT-based or Arithmetic to geometric mean ratio (AGM) $T_{\mathrm{GLRT}} = \frac{\frac{1}{n_{\mathrm{R}}}\sum_{i=1}^{n_{\mathrm{R}}} l_i}{\left(\prod_i l_i\right)^{1/n_{\mathrm{R}}}}$
- GLRT-based algorithms (single user) $T_{\mathrm{GLRT}} = \frac{l_1}{\frac{1}{n_{\mathrm{R}}}\sum_{i=2}^{n_{\mathrm{R}}} l_i}$

where l_i are the ordered eigenvalues of the SCM.

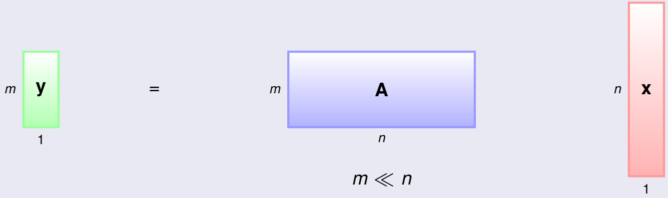
Issues: threshold determination, complexity for large matrices, ...

Other metrics proposed in some literature: Maximum to minimum eigenvalues ratio, Energy to minimum eigenvalue ratio.

3) Compressed Sensing

Compressed Sensing (CS)

CS considers recovering a signal $\mathbf{x} \in \mathbb{R}^n$ from $\mathbf{y} \in \mathbb{R}^m$



measurement vector | measurement matrix | sparse vector

- This requires solving an under-determined system of linear equations
- Possible only when at most $s \ll n$ elements of $\mathbf{x}$ are non-zero, i.e., $\mathbf{x}$ is sparse
- The CS model appears in many engineering fields, e.g., ADC, source coding, imaging, Radar, MRI, and WSN

3) Sparse Recovery Algorithms

The sparse vector $\mathbf{x}$ can be recovered from $\mathbf{y} = \mathbf{A}\mathbf{x}$ using:

- Optimization programs
- Greedy and thresholding algorithms

Optimization programs

- The s-sparse vector $\mathbf{x}$ can be reconstructed as
$$\hat{\mathbf{x}} = \arg\min \|\mathbf{x}\|_0 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}$$
If all possible submatrices consisting of $2s$ columns of $\mathbf{A}$ are maximum rank ($m > 2\,s$)
- However, the solution is computationally prohibitive
- A much easier problem is the ℓ_1-minimization solution
$$\hat{\mathbf{x}} = \arg\min \|\mathbf{x}\|_1 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}$$
under some **sufficient recovery conditions** on $\mathbf{A}$

3) Signal Recovery Guarantees: The RIC

The Restricted Isometry Constant (RIC) indicates how well a linear transformation preserves distances between sparse vectors.

RIC definition [Candes and Tao,'05]

The RIC of order s of $\mathbf{A}$, $\delta_s(\mathbf{A})$, is the smallest constant, larger than zero, that satisfies

$$1 - \delta_s(\mathbf{A}) \leq \frac{\|\mathbf{A}_S\,\mathbf{c}\|^2}{\|\mathbf{c}\|^2} \leq 1 + \delta_s(\mathbf{A})$$

for every $\mathbf{c} \in \mathbb{R}^s$ and every $m \times s$ submatrix $\mathbf{A}_S$ of $\mathbf{A}$ with columns indexed by $S \subset \Omega \triangleq \{1, 2, \ldots, n\}$ with $\mathrm{card}(S) = s$.

Why the RIC is important?

- Sufficient conditions for recovery via ℓ_1, greedy and thresholding algorithms are given in terms of the RIC (e.g., $\delta_s(\mathbf{A}) < 1/3$ [Cai *et al.*,'13])
- Bounding recovery error for non-sparse and/or noisy signals

3) RIC for Gaussian Matrices

Build $\mathbf{A}$ with entries $a_{i,j} \sim \mathcal{N}(0, 1/m)$.
What is the maximum s ensuring that $\mathbb{P}\{\delta_s(\mathbf{A}) < \delta\} \simeq 1$?

For a given submatrix $\mathbf{A}_S$, and the corresponding Wishart $\mathbf{W} = \mathbf{A}_S^T \mathbf{A}_S$

$$\lambda_{\min}(\mathbf{W}) \leq \frac{\|\mathbf{A}_S\,\mathbf{c}\|^2}{\|\mathbf{c}\|^2} \leq \lambda_{\max}(\mathbf{W})$$

$\mathbf{A}_S$ is well conditioned with probability

$$P_{sw}(\delta) \triangleq \mathbb{P}\{\lambda_{\min}(\mathbf{W}) \geq 1 - \delta, \lambda_{\max}(\mathbf{W}) \leq 1 + \delta\}$$

Using the union bound RIP is satisfied with probability

$$\mathbb{P}\{\delta_s(\mathbf{A}) < \delta\} \geq 1 - \binom{n}{s}[1 - P_{sw}(\delta)]$$

The statistics of the eigenvalues of Wishart matrices plays a crucial role to calculate $P_{sw}(\delta)$.

3) from the literature

[Candes and Tao,'05]

- Analyzing the RIC using the concentration of measure inequality bounds
- The concentration inequalities give quite loose bounds on the extreme eigenvalues
- Leading to a pessimistic over-estimation of the RIC (under-estimation of the recoverable sparsity)

[Blanchard, *et al.*,'11]

- Analyzing the RIC using the asymptotic behavior of some bounds on the eigenvalues distribution
- Provides better results in the asymptotic settings
- Can not be used for finite matrices, which are important for real applications

Application to compressed sensing

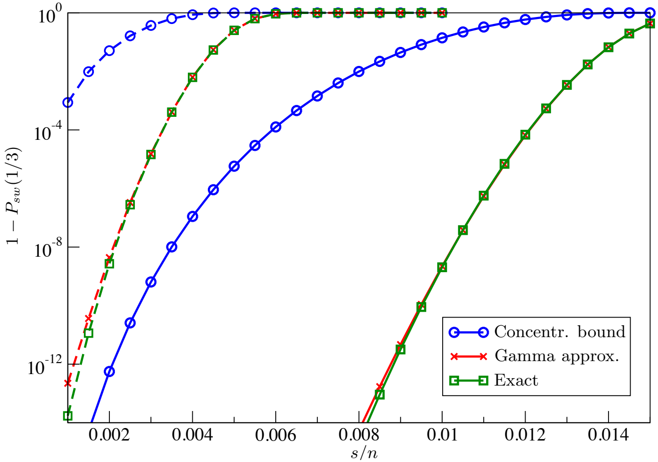


Figure: The probability that a measurement submatrix $\mathbf{A}_S$ is ill conditioned, $\delta = 1/3$, as a function of the sparsity ratio, s/n, for compression ratios $m/n = 0.2$ (dashed) and $m/n = 0.6$ (solid). The signal dimension is $n = 10^4$.

Computing the Distribution of Eigenvalues: Notations

$\gamma\,(a; x, y) = \int_x^y t^{a-1} e^{-t} dt$: the generalized incomplete gamma function

$P(a, x) = \frac{1}{\Gamma(a)} \gamma(a; 0, x)$: the regularized lower incomplete gamma function

$P(a; x, y) = \frac{1}{\Gamma(a)} \int_x^y t^{a-1} e^{-t} dt = P(a, y) - P(a, x)$: the generalized regularized incomplete gamma function

$\mathcal{B}\,(x, y; a, b) = \int_x^y t^{a-1} (1-t)^{b-1} dt$: the incomplete beta function

$$\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{i=1}^{m} \Gamma(a - (i-1)/2)$$

$$\tilde{\Gamma}_n(m) = \pi^{n(n-1)/2} \prod_{i=1}^{n} (m-i)!$$

Real Wishart matrices: joint eigenvalues p.d.f.

$\boldsymbol{X}$: Gaussian real $n \times N$ matrix $\boldsymbol{X}$ with i.i.d. columns, each with zero mean and covariance $\boldsymbol{\Sigma}$, and $N \geq n$.

$\boldsymbol{M} = \boldsymbol{X}\boldsymbol{X}^T$: real Wishart matrix, $W_n(N, \boldsymbol{\Sigma})$.

$\boldsymbol{\Sigma} \propto \boldsymbol{I}$: white or uncorrelated Wishart

The joint probability distribution function (p.d.f.) of the (real) ordered eigenvalues $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n \geq 0$ of the real Wishart matrix $\boldsymbol{M} \sim W_n(N, \boldsymbol{I})$ is

$$f(x_1, \ldots, x_n) = K \prod_{i=1}^{n} e^{-x_i/2} x_i^{\alpha} \prod_{i<j}^{n} (x_i - x_j) \qquad (13)$$

where $x_1 \geq x_2 \geq \cdots \geq x_n \geq 0$, $\alpha \triangleq (N - n - 1)/2$, and K is a normalizing constant

Real symmetric Gaussian matrices (GOE): joint eigenvalues p.d.f.

Gaussian Orthogonal Ensemble: real $n \times n$ symmetric matrices with i.i.d. Gaussian $\mathcal{N}(0, 1/2)$ on the upper-triangle, and i.i.d. $\mathcal{N}(0, 1)$ on the diagonal.

GOE: the joint p.d.f. of the eigenvalues is

$$f(x_1, \ldots, x_n) = K_{GOE} \prod_{i=1}^{n} e^{-x_i^2/2} \prod_{i<j}^{n} (x_i - x_j) \tag{14}$$

with $x_1 \geq x_2 \cdots \geq x_n$ and K_{GOE} is a normalization constant.

Real multivariate beta (double Wishart) matrices: joint eigenvalues p.d.f.

$\boldsymbol{X}$, $\boldsymbol{Y}$: independent real Gaussian $\mathrm{p} \times \mathrm{m}$ and $\mathrm{p} \times \mathrm{n}$ matrices with $\mathrm{m}, \mathrm{n} \geq \mathrm{p}$, zero mean i.i.d. columns with common covariance.

$\boldsymbol{A} = \boldsymbol{X}\boldsymbol{X}^T$ and $\boldsymbol{B} = \boldsymbol{Y}\boldsymbol{Y}^T$ are independent Wishart matrices

Multivariate analysis of variance (MANOVA): based on the eigenvalues of $(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}$ (beta matrix), simply related to the eigenvalues of $\mathbf{A}^{-1}\mathbf{B}$ (double Wishart or multivariate beta).

The joint distribution of the s non-null eigenvalues of the eigenvalues of $(\boldsymbol{A} + \boldsymbol{B})^{-1}\boldsymbol{B}$ is

$$f(x_1, \ldots, x_s) = K_{MB} \prod_{i=1}^{s} x_i^m (1 - x_i)^n \prod_{i<j}^{s} (x_i - x_j) \tag{15}$$

where $1 \geq x_1 \cdots \geq x_s \geq 0$, and K_{MB} is a normalizing constant,
$s = \mathrm{p}, \quad m = (\mathrm{n} - \mathrm{p} - 1)/2, \quad n = (\mathrm{m} - \mathrm{p} - 1)/2\,.$

Roy's largest root criterion: largest eigenvalue

The function $\psi(a, b)$ for real Wishart matrices

Theorem 4

The probability that all non-zero eigenvalues of the real Wishart matrix $\boldsymbol{M} \sim W_n(N, \boldsymbol{I})$ are within the interval $[a, b] \subset [0, \infty)$ is

$$\psi(a, b) = K' \sqrt{|\boldsymbol{A}(a, b)|} \tag{16}$$

with the constant $K' = K\, 2^{\alpha n + n(n+1)/2} \prod_{\ell=1}^{n} \Gamma(\alpha + \ell)$.

When n is even the elements $a_{i,j}$ can be computed iteratively, without numerical integration, starting from $a_{i,i} = 0$ with the iteration

$$a_{i,j+1} = a_{i,j} + \frac{\Gamma(\alpha_i + \alpha_j)}{\Gamma(\alpha_j + 1)\Gamma(\alpha_i) 2^{\alpha_i + \alpha_j - 1}} P(\alpha_i + \alpha_j; a, b) - \frac{g(\alpha_j, a/2) + g(\alpha_j, b/2)}{\Gamma(\alpha_j + 1)} P(\alpha_i; a/2, b/2) \tag{17}$$

for $j = i, \ldots, n-1$, with $\alpha_\ell = \alpha + \ell = (N - n - 1)/2 + \ell$ and $g(a, x) = x^a e^{-x}$.

When n is odd, add a column/row as $a_{i,n+1} = P(\alpha_i; a/2, b/2)$, $a_{n+1,j} = -a_{j,n+1}$, $a_{n+1,n+1} = 0$.

The function $\psi(a, b)$ for real Wishart matrices

Proof (1/3)

We have to integrate the p.d.f. in (13). First we observe that an identity related to Vandermonde matrices gives $\det\left[\left\{y_i^{j-1}\right\}\right] = \prod_{i<j}(y_j - y_i)$. Then, for arbitrary constants $\gamma_\ell \neq 0$ we write

$$\int_{a \le x_1 < \cdots < x_n \le b} \cdots \int f(x_n, \ldots, x_1) d\boldsymbol{x} =$$

$$= K' \int_{a/2 \le y_1 < \cdots < y_n \le b/2} \cdots \int \prod_{i=1}^{n} \gamma_i \, y_i^{\alpha} e^{-y_i} \prod_{i<j}^{n} (y_j - y_i) d\boldsymbol{y}$$

$$= K' \int_{a/2 \le y_1 < \cdots < y_n \le b/2} \cdots \int \det\left[\left\{\Phi_i(y_j)\right\}\right] d\boldsymbol{y} \tag{18}$$

with $\Phi_i(y) = \gamma_i \, y^{\alpha+i-1} e^{-y}$ and $K' = K 2^{\alpha n + n(n+1)/2} \prod_{\ell=1}^{n} \gamma_\ell^{-1}$.

The function $\psi(a, b)$ for real Wishart matrices

Proof (2/3)

For a generic $m \times m$ matrix $\mathbf{\Phi}(\boldsymbol{w})$ with elements $\{\Phi_i(w_j)\}$ where the $\Phi_i(x)$, $i = 1, \ldots, m$ are generic functions, the following identity holds [Deb:55]

$$\int \cdots \int_{a \le w_1 < \ldots < w_m \le b} |\mathbf{\Phi}(\boldsymbol{w})| \, d\boldsymbol{w} = \mathrm{Pf}\,(\boldsymbol{A}) \tag{19}$$

where $\mathrm{Pf}\,(\boldsymbol{A})$ is the Pfaffian, $(\mathrm{Pf}\,(\boldsymbol{A}))^2 = |\boldsymbol{A}|$, and the skew-symmetric matrix $\boldsymbol{A}$ is $m \times m$ for m even, and $(m+1) \times (m+1)$ for m odd, with

$$a_{i,j} = \int_a^b \int_a^b \mathrm{sgn}(y - x)\Phi_i(x)\Phi_j(y)dxdy \qquad i, j = 1, \ldots, m. \tag{20}$$

For m odd the additional elements are $a_{i,m+1} = -a_{m+1,i} = \int_a^b \Phi_i(x)dx$, $i = 1, \ldots, m$, and $a_{m+1,m+1} = 0$.
We then note that, by writing $F_\ell(y) = \int_0^y \Phi_\ell(x)dx$ a primitive of $\Phi_\ell(x)$, we can rewrite (20) as

$$a_{i,j} = \left[F_j(b) + F_j(a)\right]\left[F_i(b) - F_i(a)\right] - 2\int_a^b \Phi_i(x)F_j(x)dx \,. \tag{21}$$

The function $\psi(a, b)$ for real Wishart matrices

Proof (3/3)

To avoid integration first we observe that for an integer n we have

$$P(a+n, x) = P(a, x) - e^{-x} \sum_{k=0}^{n-1} \frac{x^{a+k}}{\Gamma(a+k+1)}. \qquad (22)$$

Therefore, $P(a, x)$ and $P(a; x, y)$ can be written in closed form when a is integer or half-integer, starting from

$$P(0, x) = 1 \qquad P(1/2, x) = \mathrm{erf}\sqrt{x}$$

Using the relation

$$P(a+1, x) = P(a, x) - e^{-x} x^{a} / \Gamma(a+1)$$

in (21) gives, after simple manipulations, the iteration (17).

□

Algorithm for real Wishart matrices

Algorithm 1 $\psi(a,b)$ for real Wishart matrices

Input: n, N, a, b
Output: $\psi(a,b) = \Pr\{a \leq \lambda_{\min}(\boldsymbol{M}), \lambda_{\max}(\boldsymbol{M}) \leq b\}$

$\boldsymbol{A} = \boldsymbol{0}$
$\alpha_\ell = (N-n-1)/2 + \ell$
$g(\alpha_\ell, x) = x^{\alpha_\ell} e^{-x}$
for $i = 1 \to n-1$ **do**
  for $j = i \to n-1$ **do**

$$a_{i,j+1} = a_{i,j} + \frac{\Gamma(\alpha_i+\alpha_j)\, 2^{1-\alpha_i-\alpha_j}}{\Gamma(\alpha_j+1)\Gamma(\alpha_i)} P(\alpha_i+\alpha_j; a, b) - \frac{g(\alpha_j, a/2) + g(\alpha_j, b/2)}{\Gamma(\alpha_j+1)} P(\alpha_i; a/2, b/2)$$

  end for
end for
if n is odd **then**
  append to $\boldsymbol{A}$ one column according to Theorem 4 and a zero row
end if
$\boldsymbol{A} = \boldsymbol{A} - \boldsymbol{A}^T$
return $K' \sqrt{|\boldsymbol{A}|}$

In *Mathematica*®, we get the exact value $\psi(a,b)$ for $n = N = 500$ in few seconds.

The function $\psi(a, b)$ for real symmetric Gaussian matrices (GOE)

Theorem 5

The probability that all eigenvalues of the real GOE matrix $\boldsymbol{M}$ are within the interval $[a, b] \subset (-\infty, \infty)$ is

$$\psi(a, b) = K'_{GOE} \sqrt{|\boldsymbol{A}(a, b)|} \tag{23}$$

with the constant $K'_{GOE} = K_{GOE}\, 2^{n(n+1)/4} \prod_{\ell=1}^{n} \Gamma(\ell/2)$.
When n is even the elements $a_{i,j}$ can be computed iteratively, starting from

$$a_{2,1} = \frac{1}{4}\left\{\sqrt{2}\,[erfc(b) - erfc(a)] + \left(e^{-a^2/2} + e^{-b^2/2}\right)\left[erfc\left(\frac{a}{\sqrt{2}}\right) - erfc\left(\frac{b}{\sqrt{2}}\right)\right]\right\}$$

and using the antisymmetry $a_{j,i} = -a_{i,j}$, together with the iteration

$$a_{i,j+2} = a_{i,j} + \frac{\Gamma\left(\frac{i+j}{2}\right)2^{-(i+j)/2}}{\Gamma(i/2)\Gamma(j/2+1)} F_{i+j}(a, b) - \frac{q(j, a/\sqrt{2}) + q(j, b/\sqrt{2})}{2\Gamma(j/2+1)} F_i\left(\frac{a}{\sqrt{2}}, \frac{b}{\sqrt{2}}\right) \tag{24}$$

where $q(j, x) = x^j e^{-x^2}$. When n is odd, add a column/row as $a_{i,n+1} = F_i\left(\frac{a}{\sqrt{2}}, \frac{b}{\sqrt{2}}\right)$ where

$$F_j(y) = \frac{sgn^j(y)}{2} P\left(\frac{j}{2}, y^2\right) \tag{25}$$

and $F_j(x, y) \triangleq F_j(y) - F_j(x)$.

The function $\psi(a, b)$ for real symmetric Gaussian matrices (GOE)

Proof

Similar to Wishart.
To avoid integration, substituting (22) in (25) we have

$$P_{j+2}(y) = P_j(y) - y^j e^{-y^2} \frac{1}{2\,\Gamma(j/2+1)}.$$

Using the relation

$$\int_0^\beta x^{n-1} e^{-2x^2} dx = 2^{-1-n/2} \mathrm{sgn}(\beta)^n \, \Gamma\left(\frac{n}{2}\right) P\left(\frac{n}{2}, 2\beta^2\right)$$

in (**??**) gives, after some manipulations, the iteration (24). □

To build iteratively the upper half of the skew-symmetric matrix $\boldsymbol{A}(a, b)$ we just need (24) and the first two diagonals $a_{i,i}$ and $a_{i,i+1}$. The first diagonal is clearly identically zero due to skew-symmetry, giving $a_{i,i} = 0$. The odd first diagonal $a_{i,i+1}$ can be obtained by a zig-zag iteration

$$a_{1,2} \xrightarrow{(a)} a_{2,1} \xrightarrow{(b)} a_{2,3} \xrightarrow{(a)} a_{3,2} \xrightarrow{(b)} a_{3,4} \xrightarrow{(a)} a_{4,3} \cdots$$

where steps (a) use skew-symmetry, and steps (b) use (24). The element $a_{1,2}$ is directly obtained in closed form from (**??**).

Algorithm for GOE

Algorithm 2 $\psi(a,b)$ for real Gaussian matrices (GOE)

Input: n, N, a, b
Output: $\psi(a,b) = \Pr\{a \le \lambda_{\min}(\boldsymbol{M}), \lambda_{\max}(\boldsymbol{M}) \le b\}$

$\boldsymbol{A} = \boldsymbol{0}$

$$a_{1,2} = -\frac{1}{4}\left\{\sqrt{2}\,[\operatorname{erfc}(b) - \operatorname{erfc}(a)] + \left(e^{-a^2/2} + e^{-b^2/2}\right)\left[\operatorname{erfc}\left(\frac{a}{\sqrt{2}}\right) - \operatorname{erfc}\left(\frac{b}{\sqrt{2}}\right)\right]\right\}$$

for $i = 1 \to n-2$ **do**
  for $j = i \to n-2$ **do**
    derive $a_{i,j+2}$ from $a_{i,j}$ using (24)
  end for
  $a_{i+1,i} = -a_{i,i+1}$
  derive $a_{i+1,i+2}$ from $a_{i+1,i}$ using (24)
  $a_{i+1,i} = 0$
end for
if n is odd **then**
  append to $\boldsymbol{A}$ one column according to Theorem 5 and a zero row
end if
$\boldsymbol{A} = \boldsymbol{A} - \boldsymbol{A}^T$
return $K'_{GOE}\sqrt{|\boldsymbol{A}|}$

The algorithm can be used to evaluate numerically or symbolically $\psi(a,b)$. Evaluating numerically $\psi(a,b)$ for an arbitrary interval $[a,b]$ requires few seconds for matrices of dimensions $n = 500$.

Some examples of explicit expressions, GOE

With the Algorithm the exact expression of $\psi(a, b)$ can be also derived symbolically.

Example: the probability that all eigenvalues are negative (or all positive, due to symmetry), obtained from Algorithm 2, are:

$$n = 1 \qquad \psi(-\infty, 0) = \frac{1}{2}$$

$$n = 2 \qquad \psi(-\infty, 0) = \frac{1}{4}\left(2 - \sqrt{2}\right)$$

$$n = 3 \qquad \psi(-\infty, 0) = \frac{\pi - 2\sqrt{2}}{4\pi}$$

$$n = 4 \qquad \psi(-\infty, 0) = \frac{\sqrt{\frac{1}{2}\left(9 - 4\sqrt{2}\right)}\left(-16 - 4\sqrt{2} + 7\pi\right)}{56\pi}$$

$$n = 5 \qquad \psi(-\infty, 0) = \frac{-8 - \sqrt{2} + 3\pi}{24\pi}$$

$$n = 10 \qquad \psi(-\infty, 0) = \frac{\sqrt{\frac{1}{2}\left(44217 - 27392\sqrt{2}\right)}}{183377510400\pi^2} \cdot \left[432799744 + 6251520\sqrt{2}\right.$$
$$\left. -\left(278413220 + 1989925\sqrt{2}\right)\pi + 44769900\pi^2\right].$$

The expressions for $n = 1, 2$ and 3 were already known [DeaMaj:06,DeaMaj:08].

Some asymptotics for GOE

Asymptotic bound [DeaMaj:06,DeaMaj:08]

$$\psi(-\infty, 0) \approx e^{-n^2 \ln(3)/4} . \tag{26}$$

By comparison with the exact value we found an improved approximation

$$\psi(-\infty, 0) \approx e^{-n^2 \ln(3)/4 - n \ln(10)/6} . \tag{27}$$

Table: Probability $\psi(-\infty, 0)$ that all eigenvalues are negative, GOE.

n	exact (Alg. 2)	approx. (26)	approx. [NadMaj:11]	approx. (27)
2	0.146	0.333	0.322	0.155
5	1.40E-4	1.04E-3	1.91E-3	1.53E-4
10	2.27E-14	1.18E-12	1.23E-12	2.54E-14
50	2.43E-307	6.30E-299	3.31E-304	2.92E-307
100	2.72E-1210	1.57E-1193	1.49E-1206	3.39E-1210
500	2.85E-29904	8.35E-29821	3.88E-29899	3.87E-29904

Some asymptotics for real Wishart matrices

[VivMaj:07] gives an approximation for $\psi(0, n)$ for Wishart matrices $\boldsymbol{M} \sim W_n(n, \boldsymbol{I})$.

Table: Probability $\psi(0, n)$, real Wishart, $N = n$.

n	exact (Alg. 1)	approx. [VivMaj:07]
2	0.315	0.491
5	3.71E-3	1.18E-2
10	1.90E-9	1.95E-8
50	1.70E-198	1.81E-193
100	10.2E-781	1.07E-771
500	7.33E-19325	6.22E-19275

The function $\psi(a, b)$ for real multivariate Beta (double Wishart) matrices

Theorem 6

The probability that all eigenvalues of a real multivariate beta matrix $\boldsymbol{M}$ *are within the interval* $[a, b] \subset [0, 1]$ *is*

$$\psi(a, b) = K'_{MB} \sqrt{|\boldsymbol{A}(a, b)|} \tag{28}$$

with $K'_{MB} = K_{MB} \prod_{\ell=1}^{s} \frac{\Gamma(m+\ell)}{\Gamma(m+\ell+n+1)}$.

When s *is even the elements* $a_{i,j}$ *can be computed iteratively, without numerical integration or infinite series expansion, starting from* $a_{i,i} = 0$ *with the iteration*

$$a_{i,j+1} = a_{i,j} - k_i \left[g_{j+1}(a) + g_{j+1}(b)\right] \mathcal{B}(a, b; m+i, n+1) + \frac{2k_i k_{j+1}}{m+n+j+1} \mathcal{B}(a, b; 2m+i+j, 2n+2)$$

for $j = i, \ldots, s-1$, *with* $g_\ell(x) = x^{m+\ell-1}(1-x)^{n+1} k_\ell/(m+n+\ell)$.

When s *is odd, add a column/row as* $a_{i,s+1} = k_i \mathcal{B}(a, b; m+i, n+1)$.

Algorithm for real multivariate Beta (double Wishart) matrices

Algorithm 3 $\psi(a, b)$ for real multivariate beta matrices

Input: n, m, n, a, b
Output: $\psi(a, b) = \Pr\{a \leq \lambda_{\min}(\boldsymbol{M}), \lambda_{\max}(\boldsymbol{M}) \leq b\}$

$\boldsymbol{A} = \boldsymbol{0}$
$g_\ell(x) = x^{m+\ell-1}(1-x)^{n+1}k_\ell/(m+n+\ell)$
$k_\ell = \Gamma(m+n+\ell+1)/\Gamma(m+\ell)$
for $i = 1 \to n-1$ **do**
  for $j = i \to n-1$ **do**
    $a_{i,j+1} = a_{i,j} - k_i\left[g_{j+1}(a) + g_{j+1}(b)\right]\mathcal{B}(a, b; m+i, n+1) + \dfrac{2k_ik_{j+1}}{m+n+j+1}\mathcal{B}(a, b; 2m + i + j, 2n+2)$
  end for
end for
if n is odd **then**
  append to $\boldsymbol{A}$ one column and a zero row according to Theorem 6
end if
$\boldsymbol{A} = \boldsymbol{A} - \boldsymbol{A}^T$
return $K'_{MB}\sqrt{|\boldsymbol{A}|}$

For example, we get the exact distribution of the largest eigenvalue in less than 0.1 seconds for all tables in [Pil:67], [And:03] and [Joh:09].

Complex matrices

The analysis for complex random matrices is easier than for the real case.

Important results are known since 1964 for complex multivariate Beta matrices and for uncorrelated complex Wishart [Khatri:1964].

Joint Eigenvalues p.d.f. Complex Wishart matrices, $\boldsymbol{\Sigma} = \boldsymbol{I}$

Assume a Gaussian complex $n \times N$ matrix $\boldsymbol{X}$ with i.i.d. columns, each circularly symmetric with covariance $\boldsymbol{\Sigma} = \boldsymbol{I}$, and $N \geq n$.

The joint p.d.f. of the (real) ordered eigenvalues $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_n \geq 0$ of the complex Wishart matrix $\boldsymbol{M} = \boldsymbol{X}\boldsymbol{X}^H \sim \mathcal{CW}_n(N, \boldsymbol{I})$ (identity covariance) is [Jam64; Ede:88; Joh:01]

$$f(x_1, \ldots, x_n) = K \prod_{i=1}^{n} e^{-x_i} x_i^{N-n} \prod_{i<j}^{n} (x_i - x_j)^2 \tag{29}$$

where $x_1 \geq x_2 \geq \cdots \geq x_n \geq 0$ and K is a normalizing constant

Joint Eigenvalues p.d.f. Complex Wishart matrices, arbitrary $\mathbf{\Sigma}$

Correlated case: Gaussian complex $n \times N$ matrix $\boldsymbol{X}$ with i.i.d. columns, each circularly symmetric with covariance $\mathbf{\Sigma}$, and $N \geq n$.

The joint distribution of the ordered eigenvalues of $\boldsymbol{M} = \boldsymbol{X}\boldsymbol{X}^H \sim CW_n(N, \mathbf{\Sigma})$ has been found in a usable form in [M.C., Win, Zanella, T-IT 2003]

Lemma 7 (M.C., Win, Zanella, IEEE Trans. on Information Theory, 2003)

Let $\boldsymbol{M} \sim \mathcal{CW}_n(N, \mathbf{\Sigma})$ be a complex Wishart matrix, $N \geq n$. Denote $\sigma_1 > \sigma_2 > \ldots > \sigma_n > 0$ the ordered eigenvalues of $\mathbf{\Sigma}$. Then, the joint p.d.f. of the ordered eigenvalues of $\boldsymbol{M}$ is

$$f(x_1, \ldots, x_n) = K_{\mathbf{\Sigma}} \left| \mathbf{E}(\boldsymbol{x}, \boldsymbol{\sigma}) \right| \cdot \prod_{i<j}^{n} (x_i - x_j) \cdot \prod_{j=1}^{n} x_j^{N-n} \tag{30}$$

where $\mathbf{E}(\boldsymbol{x}, \boldsymbol{\sigma}) = \left\{ e^{-x_i/\sigma_j} \right\}_{i,j=1}^{n}$ *and*

$$1/K_{\mathbf{\Sigma}} = \prod_{i<j}^{n} (\sigma_i - \sigma_j) \prod_{i=1}^{n} \sigma_i^{N-n+1} (N-i)!\,. \tag{31}$$

Joint Eigenvalues p.d.f. Complex Wishart matrices, arbitrary Σ

There are books where it is reported that the result was already in [James:1964] ... it was not ...

Theorem 2.18. Let $\mathbf{W}$ be a central complex Wishart matrix $\mathbf{W} \sim W_m(n, \boldsymbol{\Sigma})$ with $n \geq m$, where the eigenvalues of $\boldsymbol{\Sigma}$ are distinct and their ordered values are $a_1 > \ldots > a_m > 0$. The joint p.d.f. of the ordered positive eigenvalues of $\mathbf{W}$, $\lambda_1 \geq \ldots \geq \lambda_m$, equals [125]

$$\frac{\det(\{e^{-\lambda_j/a_i}\})}{\det \boldsymbol{\Sigma}^n} \prod_{\ell=1}^{m} \frac{\lambda_\ell^{n-m}}{(n-\ell)!} \prod_{k<\ell}^{m} \frac{\lambda_k - \lambda_\ell}{a_k - a_\ell} a_\ell \, a_k. \qquad (2.25)$$

Figure: Page 32 of the book by Tulino, Verdu, 2004, attributing the pdf with correlation to James 1964.

Theorem 2.5 (Section 8.7 of [James, 1964]). *Let the columns of $\mathbf{X} \in \mathbb{C}^{N \times n}$ be i.i.d. zero mean Gaussian with positive definite covariance $\mathbf{R}$, and $n \geq N$. The joint p.d.f. $P^{\geq}_{(\lambda_i)}$ of the ordered positive eigenvalues $\lambda_1 \geq \ldots \geq \lambda_N$ of the central Wishart matrix $\mathbf{X}\mathbf{X}^{\mathsf{H}}$ reads:*

$$P^{\geq}_{(\lambda_i)}(\lambda_1, \ldots, \lambda_N) = (-1)^{\frac{1}{2}N(N-1)} \frac{\det(\{e^{-\frac{\lambda_i}{r_j}}\}_{1 \leq i,j \leq N})}{\det \mathbf{R}^n} \frac{\Delta(\boldsymbol{\Lambda})}{\Delta(\mathbf{R}^{-1})} \prod_{j=1}^{N} \frac{\lambda_j^{n-N}}{(n-j)!}$$

with $r_1 > \ldots > r_N > 0$ the eigenvalues of $\mathbf{R}$ and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$.

Figure: Page 23 of the book by Couillet, Debbah, 2010, attributing the pdf with correlation to James 1964.

Generalization of the Cauchy-Binet formula [M.C., Win,Zanella:03]

Definition 8 (pseudo determinant)

Given a rank 3 tensor $\boldsymbol{A} = \{a_{i,j,k}\}_{i,j,k=1,\dots,N}$, define

$$\mathcal{T}(\boldsymbol{A}) \triangleq \sum_{\mu} \operatorname{sgn}(\mu) \sum_{\alpha} \operatorname{sgn}(\alpha) \prod_{k=1}^{N} a_{\mu_k,\alpha_k,k} \ , \tag{32}$$

where the sums are over all possible permutations, μ and α, of the integers $1, \dots, N$.

Note that when $a_{i,j,k}$ are independent of k, $a_{i,j,k} = a_{i,j,1}$, we have

$$\mathcal{T}(\boldsymbol{A}) = \sum_{\mu} \operatorname{sgn}(\mu) \sum_{\alpha} \operatorname{sgn}(\alpha) \prod_{k=1}^{N} a_{\mu_k,\alpha_k,k} = N! \det\left(\{a_{i,j,1}\}_{i,j=1\dots,N}\right) \ , \tag{33}$$

Generalization of the Cauchy-Binet formula [M.C., Win,Zanella:03]

Theorem 9

Given two arbitrary $N \times N$ matrices $\mathbf{\Phi}(\boldsymbol{x})$, $\mathbf{\Psi}(\boldsymbol{x})$ with (i,j) elements $\Phi_i(x_j)$ and $\Psi_i(x_j)$, and arbitrary functions $\xi_i(\cdot)$ the following identity holds:

$$\int \cdots \int_{\mathcal{D}} |\mathbf{\Phi}(\boldsymbol{x})| \cdot |\mathbf{\Psi}(\boldsymbol{x})| \prod_{k=1}^{N} \xi_k(x_k) d\boldsymbol{x} = \mathcal{T}\left(\left\{\int_a^b \Phi_i(x)\Psi_j(x)\,\xi_k(x)dx\right\}_{i,j,k=1,\dots,N}\right),$$

where the multiple integral is over the domain $\mathcal{D} = \{a \le x_1 \le b, a \le x_2 \le b, \dots, a \le x_N \le b\}$ and $d\mathbf{x} = dx_1\, dx_2 \cdots dx_N$.

Assume a (ordered) p.d.f. of the form

$$f_{\boldsymbol{\lambda}}(\boldsymbol{x}) = K\,|\mathbf{\Phi}(\boldsymbol{x})| \cdot |\mathbf{\Psi}(\boldsymbol{x})| \cdot \prod_{l=1}^{n} \xi(x_l), \tag{34}$$

where K is the normalizing constant, $\mathbf{\Phi}(\boldsymbol{x})$, $\mathbf{\Psi}(\boldsymbol{x})$ are two arbitrary $n \times n$ matrices with (i,j) elements $\Phi_i(x_j)$ and $\Psi_i(x_j)$, and $\xi(\cdot)$ is an arbitrary function.

Multiplying by $\prod_k \theta_k(x_k)$ and integrating we get $\mathbb{E}\left\{\prod_k \theta_k(\lambda_k)\right\}$.

Particular cases: marginals

Joint Eigenvalues p.d.f. Complex Wishart matrices, arbitrary $\mathbf{\Sigma}$

[M.C., Win, Zanella, T-IT 2003] extended to the case where $\mathbf{\Sigma}$ has eigenvalues of arbitrary multiplicity and to the marginal eigenvalues distribution in [M.C., Zanella 2008; Zanella, M.C., Win 2009; M.C., Win, Shin 2010].

In particular, when $\mathbf{\Sigma}$ is spiked with $\sigma_1 > \sigma_2 = \sigma_3 = \sigma_4 = \cdots = \sigma_n$, we have

Lemma 10 (M.C., Win, Shin, IEEE Trans. on Information Theory, 2010)

Let $\boldsymbol{M} \sim \mathcal{CW}_n(N, \mathbf{\Sigma})$ be a complex Wishart matrix, $N \geq n$. Denote $\sigma_1 > \sigma_2 = \ldots = \sigma_n > 0$ the ordered eigenvalues of $\mathbf{\Sigma}$ (spiked covariance matrix). Then, the joint p.d.f. of the ordered eigenvalues of $\boldsymbol{M}$ is

$$f(x_1, \ldots, x_n) = K_1 \left|\mathbf{E}\left(\boldsymbol{x}, \boldsymbol{\sigma}\right)\right| \cdot \prod_{i<j}^{n}(x_i - x_j) \cdot \prod_{j=1}^{n} x_j^{N-n} \tag{35}$$

where $\mathbf{E}\left(\boldsymbol{x}, \boldsymbol{\sigma}\right)$ has elements

$$e_{i,j} = \begin{cases} e^{-x_i/\sigma_1} & j = 1 \\ x_i^{n-j} e^{-x_i/\sigma_2} & j = 2, \ldots, n \end{cases}$$

and

$$\frac{1}{K_1} = \sigma_1^{N-n+1} \sigma_2^{(N-1)(n-1)} (\sigma_1 - \sigma_2)^{n-1} \prod_{i=1}^{n} (N-i)! \prod_{\ell=2}^{n-2} \ell! \quad .$$

$\psi(a,b)$ for Complex Wishart matrices, $\mathbf{\Sigma} = \boldsymbol{I}$

For complex Wishart matrices $\boldsymbol{M} \sim CW_n(N, \mathbf{\Sigma})$, $N \geq n$, the probability that all eigenvalues are within $[a, b] \subset [0, \infty)$ is given below, depending on the covariance $\mathbf{\Sigma}$.

Theorem 11

For the uncorrelated complex Wishart matrix $\boldsymbol{M} \sim CW_n(N, \boldsymbol{I})$:

$$\psi(a,b) = K\,|\boldsymbol{A}(a,b)|$$

where the elements of the $n \times n$ matrix $\boldsymbol{A}(a,b)$ are

$$a_{i,j} = \int_a^b t^{N+n-i-j} e^{-t} dt = \gamma\left(N+n-i-j+1; a, b\right)$$

and K is a normalizing constant.

$\psi(a, b)$ for Complex Wishart matrices, arbitrary $\mathbf{\Sigma}$

$\psi(a, b)$ can be obtained for the general case where $\mathbf{\Sigma}$ has eigenvalues of arbitrary multiplicity. Below some special cases.

Theorem 12

For the correlated complex Wishart matrix $\boldsymbol{M} \sim CW_n(N, \mathbf{\Sigma})$ where $\mathbf{\Sigma}$ has distinct eigenvalues $\sigma_1 > \sigma_2 > \cdots > \sigma_n$:

$$\psi(a, b) = K_{\mathbf{\Sigma}} \, |\boldsymbol{A}(a, b)|$$

where the elements of the $n \times n$ matrix $\boldsymbol{A}(a, b)$ are

$$a_{i,j} = \int_a^b t^{N-i} e^{-t/\sigma_j} dt = \sigma_j^{N-i+1} \gamma \left(N - i + 1; \frac{a}{\sigma_j}, \frac{b}{\sigma_j} \right)$$

and $K_{\mathbf{\Sigma}}$ is given in (31).

$\psi(a, b)$ for Complex Wishart matrices, spiked covariance

Theorem 13

For the correlated complex Wishart matrix $\boldsymbol{M} \sim CW_n(N, \boldsymbol{\Sigma})$ with a spiked covariance $\boldsymbol{\Sigma}$ having eigenvalues $\sigma_1 > \sigma_2 = \sigma_3 = \sigma_4 = \cdots = \sigma_n$:

$$\psi(a, b) = K_1 \left|\boldsymbol{A}(a, b)\right|$$

where the elements of the $n \times n$ matrix $\boldsymbol{A}(a, b)$ are

$$a_{i,1} = \int_a^b t^{N-i} e^{-t/\sigma_1} dt = \sigma_1^{N-i+1} \gamma\left(N - i + 1; \frac{a}{\sigma_1}, \frac{b}{\sigma_1}\right)$$

and, for $j = 2, \ldots, n$,

$$a_{i,j} = \int_a^b t^{N+n-i-j} e^{-t/\sigma_2} dt = \sigma_2^{N+n-i-j+1} \gamma\left(N + n - i - j + 1; \frac{a}{\sigma_2}, \frac{b}{\sigma_2}\right) .$$

The constant K_1 is given in Lemma 10.

$\psi(a, b)$ for Hermitian Gaussian matrices (GUE)

Gaussian Unitary Ensemble (GUE): complex Hermitian random matrices with i.i.d. $\mathcal{CN}(0, 1/2)$ entries on the upper-triangle, and $\mathcal{N}(0, 1/2)$ on the main diagonal.

Theorem 14

The probability that all eigenvalues of a $n \times n$ GUE matrix $\boldsymbol{M}$ are within the interval $[a, b] \subset (-\infty, \infty)$ is

$$\psi(a, b) = K_{GUE}\, |\boldsymbol{A}(a, b)| \tag{36}$$

where the elements of the $n \times n$ matrix $\boldsymbol{A}(a, b)$ are

$$\begin{aligned} a_{i,j} &= \int_a^b t^{i+j-2} e^{-t^2}\, dt \\ &= \frac{1}{2}\Gamma\left(\frac{i+j-1}{2}\right)\left[P\left(\frac{i+j-1}{2}, b^2\right) sgn(b)^{i+j-1} - P\left(\frac{i+j-1}{2}, a^2\right) sgn(a)^{i+j-1}\right] \end{aligned}$$

and $K_{GUE} = 2^{n(n-1)/2}(\pi^{n/2} \prod_{i=1}^{n} \Gamma[i])^{-1}$ *is a normalizing constant.*

$\psi(a, b)$ for Hermitian Gaussian matrices (GUE)

Proof.

As for the complex white Wishart, this theorem for GUE is easily derived from known results. In fact, the joint distribution of the ordered eigenvalues can be written as

$$f(x_1, \ldots, x_n) = K_{GUE} \prod_{i<j}^{n} (x_i - x_j)^2 \prod_{i=1}^{n} e^{-x_i^2} \,. \tag{37}$$

Then, use [ChiWinZan:J03, Corollary 2] with $\Psi_i(x_j) = \Phi_i(x_j) = x_j^{i-1}, \xi(x) = e^{-x^2}$. □

$\psi(a, b)$ for Complex multivariate beta (double Wishart) matrices

When $\boldsymbol{X}$, $\boldsymbol{Y}$ are two independent *complex* Gaussian, the analogous of (15) is the complex multivariate beta, where the joint distribution of the eigenvalues is [Kha:64]

$$f(x_1, \ldots, x_s) = K_{MB} \prod_{i=1}^{s} x_i^m (1 - x_i)^n \cdot \prod_{i<j}^{s} (x_i - x_j)^2 \tag{38}$$

with $1 > x_1 \geq x_2 \cdots \geq x_s > 0$, and K_{MB} a normalizing constant.

Theorem 15

For a complex multivariate Beta matrix $\boldsymbol{M}$

$$\psi(a, b) = K_{MB} \ |\boldsymbol{A}(a, b)| \tag{39}$$

where the elements of the $s \times s$ *matrix* $\boldsymbol{A}(a, b)$ *are*

$$a_{i,j} = \mathcal{B}(a, b; m + i + j - 1, n + 1)$$

for $i, j = 1, \ldots, s$.

Note that (39) can be seen as an extension of [eq. (3), Kha:64].

Asymptotics and approximations

1. For white Wishart matrices $\boldsymbol{M} \sim W_p(m, \boldsymbol{I})$ or $\boldsymbol{M} \sim CW_p(m, \boldsymbol{I})$, when $m, p \to \infty$ and $m/p \to \gamma \in [0, \infty]$
$$\frac{\lambda_{\max}(\boldsymbol{M}) - \mu_{mp}}{\sigma_{mp}} \xrightarrow{\mathcal{D}} \mathcal{TW}_\beta \tag{40}$$
where $\mathcal{TW}_\beta$ denotes the Tracy-Widom random variable of order β.
$\mu_{mp} = (\sqrt{m} + \sqrt{p})^2 \quad \sigma_{mp} = \sqrt{\mu_{mp}}\left(\frac{1}{\sqrt{p}} + \frac{1}{\sqrt{m}}\right)^{\frac{1}{3}}$. $\beta = 1$ and $\beta = 2$ for real and complex matrices, respectively.
2. For the smallest eigenvalue when $m, p \to \infty$ and $m/p \to \gamma \in (1, \infty)$,
$$-\frac{\lambda_{\min}(\boldsymbol{M}) - \mu_{mp}^-}{\sigma_{mp}^-} \xrightarrow{\mathcal{D}} \mathcal{TW}_\beta \tag{41}$$
with
$$\mu_{mp}^- = \left(\sqrt{m} - \sqrt{p}\right)^2 \quad \sigma_{mp}^- = \sqrt{\mu_{mp}^-}\left(\frac{1}{\sqrt{p}} - \frac{1}{\sqrt{m}}\right)^{\frac{1}{3}}. \tag{42}$$
The variance of the smallest eigenvalue is smaller than that of the largest eigenvalue.

Asymptotics and approximations

1. For the Gaussian ensemble, for $n \to \infty$ [TraWid:94,96]

$$\frac{\lambda_{\max}(\boldsymbol{M}) - \mu_n}{\sigma_n} \xrightarrow{\mathcal{D}} \mathcal{TW}_\beta \qquad -\frac{\lambda_{\min}(\boldsymbol{M}) - \mu_n^-}{\sigma_n^-} \xrightarrow{\mathcal{D}} \mathcal{TW}_\beta$$

   with

$$\mu_n = 2\sigma_0\sqrt{n} \quad \mu_n^- = -\mu_n \quad \sigma_n = \sigma_n^- = \sigma_0(n)^{-1/6}. \tag{43}$$

2. For $n \to \infty$ the largest and the smallest eigenvalues are asymptotically independent [E. Basor, Y. Chen, L. Zhang 2011].

$$\Pr\{a \le \lambda_{\min}(\boldsymbol{M}), \lambda_{\max}(\boldsymbol{M}) \le b\} \approx \Pr\{a \le \lambda_{\min}(\boldsymbol{M})\}\Pr\{\lambda_{\max}(\boldsymbol{M}) \le b\} \tag{44}$$

   which is like to say $\psi(a,b) \approx \psi(a,\infty)\psi(-\infty,b)$.

Approximating the Tracy-Widom with a gamma

$\Gamma(k, \theta)$: gamma r.v. with shape parameter k and scale parameter θ

Tracy-Widom approximated by scaled and shifted gamma (matching first three moments) [Chi:J14]

$$\mathcal{TW}_\beta \simeq \Gamma(k, \theta) - \alpha \tag{45}$$

Table: Parameters for approximating $\mathcal{TW}_\beta$ with $\Gamma[k, \theta] - \alpha$.

	$\mathcal{TW}_1$	$\mathcal{TW}_2$	$\mathcal{TW}_4$
k	46.446	79.6595	146.021
θ	0.186054	0.101037	0.0595445
α	9.84801	9.81961	11.0016

$\Longrightarrow$ The CDF of $\mathcal{TW}_\beta$ approximated by an incomplete gamma function:

$$\Pr\{\mathcal{TW}_\beta \leq x\} = F_\beta(x) \simeq P\left(k, \frac{(x+\alpha)^+}{\theta}\right) \tag{46}$$

Approximating the Tracy-Widom with a gamma

Comparison with pre-calculated p.d.f. values from [PraSpo:04].

Very good agreement for the right tail.

Left tail less precise but of small relative error for values of practical statistical uses.

Differently from the true distribution, the left tail is exactly zero for $x < -\alpha$.

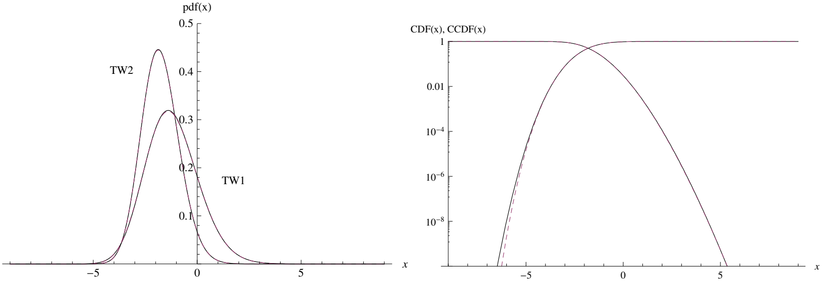


Figure: LEFT: the exact (solid line) and approximated (dashed) PDFs for the Tracy-Widom 1 and Tracy-Widom 2, linear scale. RIGHT: the exact (solid line) and approximated (dashed) CDF, CCDF, for Tracy-Widom $\mathcal{TW}_2$, log scale.

Asymptotics, Tracy-Widom, gamma

Asymptotic independence of the extremes + Tracy-Widom + gamma =

$$\psi(a,b) \simeq P\left(k, \left(\frac{\alpha}{\theta} + \frac{b-\mu}{\theta\sigma}\right)^{+}\right) P\left(k, \left(\frac{\alpha}{\theta} - \frac{a-\mu^{-}}{\theta\sigma^{-}}\right)^{+}\right) . \tag{47}$$

Probability the all eigenvalues are within the support of the limiting Marčenko-Pastur and Wigner spectral distribution

- $\boldsymbol{M} = \boldsymbol{X}\boldsymbol{X}^H$, $\boldsymbol{X}$ is $(p \times m)$ with i.i.d. entries with zero mean and variance $\sigma^2 = 1$
  Marčenko-Pastur law: asymptotic p.d.f. of an unordered eigenvalue $\lambda = \lambda(\boldsymbol{M})$ for large p, m with a fixed ratio $p/m \leq 1$ as
  $$f(\lambda) = \begin{cases} \frac{1}{2\pi p \lambda}\sqrt{(\tilde{b}-\lambda)(\lambda-\tilde{a})} & \tilde{a} \leq \lambda \leq \tilde{b} \\ 0 & \textit{otherwise} \end{cases}$$
  where $\tilde{a} = \left(\sqrt{m} - \sqrt{p}\right)^2$ and $\tilde{b} = \left(\sqrt{m} + \sqrt{p}\right)^2$

  Also, for increasing p, m $\lambda_{\min}(\boldsymbol{M}) \to \tilde{a}$ and $\lambda_{\max}(\boldsymbol{M}) \to \tilde{b}$ [BaiSil:06].
  Note that when the entries of $\boldsymbol{X}$ are Gaussian the matrix $\boldsymbol{M}$ is white Wishart.
- Wigner matrix $\boldsymbol{M} = \boldsymbol{X} + \boldsymbol{X}^H$ where $\boldsymbol{X}$ is $(p \times p)$ with i.i.d. entries with zero mean and variance $\sigma^2 = 1/4$, the Wigner semicircle law gives the asymptotic p.d.f. of an unordered eigenvalue $\lambda = \lambda(\boldsymbol{M})$ for large p as
  $$f(\lambda) = \begin{cases} \frac{1}{\pi p}\sqrt{2p - \lambda^2} & |\lambda| \leq \sqrt{2p} \\ 0 & \textit{otherwise} \end{cases}$$
  where $[-\sqrt{2p}, \sqrt{2p}]$ is the support of the semicircle law.
  Also $\lambda_{\min}(\boldsymbol{M}) \to -\sqrt{2p}$ and $\lambda_{\max}(\boldsymbol{M}) \to \sqrt{2p}$ [BaiSil:06].

Probability the all eigenvalues are within the support of the limiting MP and Wigner spectral distribution

We could be tempted to think that for increasing matrix sizes all eigenvalues are within the MP or semicircle supports with probability tending to one.
However, this is not the case: the prob. is $F_1^2(0) = 0.6921$ and $F_2^2(0) = 0.9397$ for the real and complex cases, respectively

Theorem 16

1. *Let $\boldsymbol{M} \sim W_n(N, \boldsymbol{I})$ be a real Wishart matrix with $N > n$. When $N, n \to \infty$ and $N/n \to \gamma \in (1, \infty)$, the probability that all eigenvalues are within the MP support is*

$$\psi\left((\sqrt{N} - \sqrt{n})^2, (\sqrt{N} + \sqrt{n})^2\right) \to F_1^2(0) = 0.6921\,.$$

2. *Let $\boldsymbol{M} \sim CW_n(N, \boldsymbol{I})$ be a complex Wishart matrix with $N > n$. When $N, n \to \infty$ and $N/n \to \gamma \in (1, \infty)$, the probability that all eigenvalues are within the MP support is*

$$\psi\left((\sqrt{N} - \sqrt{n})^2, (\sqrt{N} + \sqrt{n})^2\right) \to F_2^2(0) = 0.9397\,.$$

3. *Let $\boldsymbol{M}$ be a $(n \times n)$ real symmetric GOE matrix. When $n \to \infty$ the probability that all eigenvalues are within the semicircle support is*

$$\psi\left(-\sqrt{2n}, \sqrt{2n}\right) \to F_1^2(0) = 0.6921\,.$$

4. *Let $\boldsymbol{M}$ be a $(n \times n)$ complex symmetric GUE matrix. When $n \to \infty$ the probability that all eigenvalues are within the semicircle support is*

$$\psi\left(-\sqrt{2n}, \sqrt{2n}\right) \to F_2^2(0) = 0.9397\,.$$

Probability the all eigenvalues are within the support of the limiting Marčenko-Pastur and Wigner spectral distribution

Table: Probability $\psi(\tilde{a}, \tilde{b})$ that all eigenvalues of a real Wishart matrix are within the Marčenko-Pastur edges for $n/N = 2/3, 1/2, 1/5, 1/10$. Numerical values for finite n obtained by Algorithm 1, and for $n = \infty$ by Theorem 16.

	$\psi(\tilde{a}, \tilde{b})$			
n	$n/N = 2/3$	$n/N = 1/2$	$n/N = 1/5$	$n/N = 1/10$
10	0.7678	0.7645	0.7625	0.7624
20	0.7499	0.7483	0.7476	0.7477
50	0.7332	0.7326	0.7327	0.7329
100	0.7239	0.7238	0.7242	0.7244
200	0.7169	0.7171	0.7175	0.7177
500	0.7101	0.7103	0.7108	0.7109
∞	0.6921	0.6921	0.6921	0.6921

Conclusions

- Iterative algorithms are available for the exact value of the probability that all eigenvalues lie within an arbitrary interval $[a, b]$,
- Applicable for quite large (e.g. 500×500) real white Wishart, complex Wishart with arbitrary correlation, double Wishart, and Gaussian symmetric/Hermitian matrices.
- Simple approximations based on shifted incomplete gamma functions have also been proposed,
- For increasingly large matrices the probability that all eigenvalues are within the limiting support is 0.6921 for real white Wishart and GOE, and 0.9397 for complex white Wishart and GUE.

Acknowledgements

Thanks to

A. Elzanaty, A. Giorgetti, A. Mariani, P. Smith, A. Zanella, M. Win.

Research supported in part by the EC under the H2020 project EuroCPS

References

- M. C., “On the probability that all eigenvalues of Gaussian, Wishart, and double Wishart random matrices lie within an interval,” IEEE Trans. on Information Theory, Jul 2017, arxiv.
- M. Elzanaty, A. Giorgetti and M.C. “Weak RIC Analysis of Finite Gaussian Matrices for Joint Sparse Recovery,” IEEE Signal Processing Letters, 2017.
- M. C., “Distribution of the largest root of a matrix for Roy’s test in multivariate analysis of variance,” Journal of Multivariate Analysis, 2016, also in arxiv, 2014.
- M. C., “Distribution of the largest eigenvalue for real Wishart and Gaussian random matrices and a simple approximation for the Tracy-Widom distribution”, Journal of Multivariate Analysis, 2014, also arxiv 2012.
- M. C., M. Z. Win, and H. Shin, “MIMO networks: the effects of interference,” IEEE Trans. Inform. Theory, 2010.
- A. Zanella, M. C., and M. Z. Win, “On the marginal distribution of the eigenvalues of Wishart matrices,” IEEE Trans. Commun., 2009.
- H. Shin, M. Win, J. H. Lee, and M. C., “On the capacity of doubly correlated MIMO channels,” IEEE Trans. Wireless Commun., 2006.
- M. C., M. Z. Win, and A. Zanella, “On the capacity of spatially correlated MIMO Rayleigh fading channels,” IEEE Trans. Inform. Theory, 2003.

References

- M. C. and A. Zanella, “Joint distribution of an arbitrary subset of the ordered eigenvalues of Wishart matrices,” in Proc. IEEE PIMRC 2008.
- H. Shin, M. Win, J. H. Lee, and M. C., “Asymptotic statistics of mutual information for doubly correlated MIMO channels,” IEEE Trans. Wireless Commun., 2008.
- A. Zanella and M. C., “Reduced complexity power allocation strategies for MIMO systems with singular value decomposition,” IEEE Trans. Veh. Technol., 2012.
- A. Zanella, M. C., and M. Z. Win, “Statistical analysis of Steepest Descend and LMS detection algorithms for MIMO systems,” IEEE Trans. Veh. Technol., Nov. 2011.
- M. C., Moe Z. Win “Estimating the number of signals observed by multiple sensors,” in IEEE 2010 IAPR Workshop on Cognitive Information Processing (CIP2010), 2010.