Random Matrix Theory and its applications

to Statistics and Wireless Communications

Eigenvalues and Singular Values of Random Matrices: A Tutorial Introduction

Sergio Verdú Princeton University

National University of Singapore
Institute for Mathematical Sciences

February 27, 2006

Applications of Random Matrices

- Condensed Matter Physics
- Statistical Physics
- String Theory
- Quantum Chaos
- Disordered Systems
- Number Theory

Applications of Random Matrices

- Riemann Hypothesis
- von Neumann and C*-algebra theory
- Multivariate Statistics
- Stochastic Differential Equations
- Numerical Linear Algebra
- Economics

Engineering Applications of Random Matrices

- Information Theory
- Wireless Communications
- Signal Processing
- Neural Networks
- Small-World Networks

Typical Random Matrix Questions

- Distribution of $\lambda(\mathbf{H})$
- Distribution of $\lambda(\mathbf{H}^\dagger\mathbf{H})$
- Distribution of $\lambda_{\max}(\mathbf{H})$
- $\mathbb{E}\left[\det\left(\mathbf{H}^k\right)\right]$
- $\mathbb{E}[\det(\mathbf{I} + \gamma\mathbf{W})]$
- Joint distribution of $\lambda_1(\mathbf{H}) \ldots \lambda_N(\mathbf{H})$
- Distribution of the spacings between adjacent eigenvalues
- Distribution of $\mathbf{H}^\dagger\mathbf{H}$
- Distribution of the matrix of eigenvectors of $\mathbf{H}^\dagger\mathbf{H}$

Our Motivation

Vector noisy channels of the form

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1}$$

where

- $\mathbf{x}$ is the K-dimensional input vector,
- $\mathbf{y}$ is the N-dimensional output vector,
- $\mathbf{n}$ is the circularly symmetric N-dimensional vector Gaussian noise.

Why Singular Values?

$$\text{Shannon Capacity} = \frac{1}{N}\sum_{i=1}^{N} \log\left(1 + \textsc{snr}\,\lambda_i(\mathbf{H}\mathbf{H}^\dagger)\right)$$

Why Singular Values?

$$\text{Minimum Mean Square Error} = \frac{1}{K}\sum_{i=1}^{K}\frac{1}{1+\text{SNR}\,\lambda_i(\mathbf{H}^\dagger\mathbf{H})}$$

Asymptotics

- $K \longrightarrow \infty$
- $N \longrightarrow \infty$
- $\frac{K}{N} \longrightarrow \beta$

Asymptotics

$$\text{Shannon Capacity} = \frac{1}{N}\sum_{i=1}^{N} \log\left(1 + \text{SNR}\,\lambda_i(\mathbf{HH}^\dagger)\right)$$

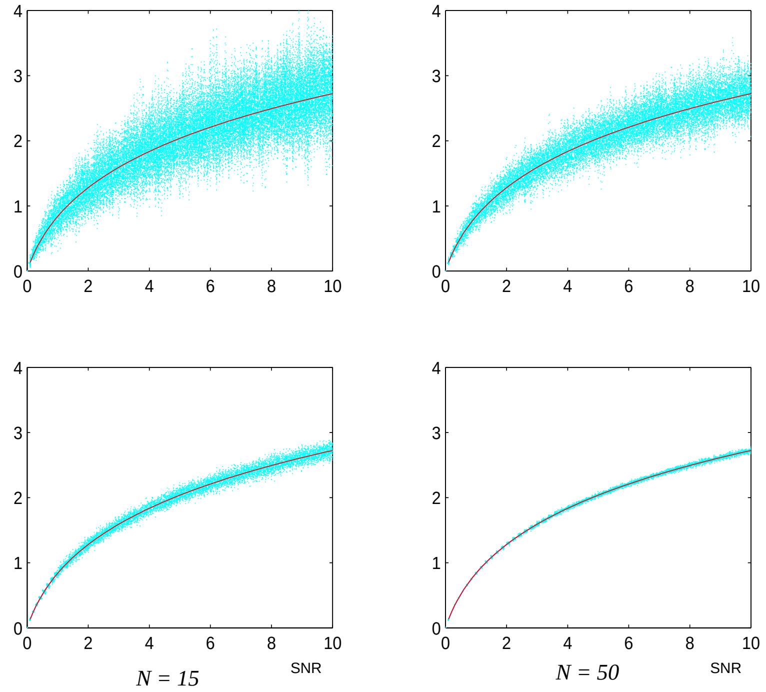

Figure 1: $\beta = 1$ for sizes: $N = 3, 5, 15, 50$

The Birth of (Nonasymptotic) Random Matrix Theory: (Wishart, 1928)

J. Wishart, “The generalized product moment distribution in samples from a normal multivariate population,” *Biometrika*, vol. 20 A, pp. 32–52, 1928.

Probability density function of the matrix:

$$\mathbf{v}_1\mathbf{v}_1^\dagger + \ldots + \mathbf{v}_n\mathbf{v}_n^\dagger$$

where $\mathbf{v}_i$ are iid Gaussian vectors.

Wishart Matrices

Definition 1. *The* $m \times m$ *random matrix* $\mathbf{A} = \mathbf{H}\mathbf{H}^\dagger$ *is a (central)* real/complex Wishart *matrix with* n *degrees of freedom and covariance matrix* $\mathbf{\Sigma}$*, (*$\mathbf{A} \sim \mathcal{W}_m(n, \mathbf{\Sigma})$*), if the columns of the* $m \times n$ *matrix* $\mathbf{H}$ *are zero-mean independent real/complex Gaussian vectors with covariance matrix* $\mathbf{\Sigma}$*.*[1]

The p.d.f. of a complex Wishart matrix $\mathbf{A} \sim \mathcal{W}_m(n, \mathbf{\Sigma})$ for $n \geq m$ is

$$f_{\mathbf{A}}(\mathbf{B}) = \frac{\pi^{-m(m-1)/2}}{\det \mathbf{\Sigma}^n \prod_{i=1}^{m} (n-i)!} \exp\left[-\mathrm{tr}\left\{\mathbf{\Sigma}^{-1}\mathbf{B}\right\}\right] \det \mathbf{B}^{n-m}. \qquad (2)$$

[1] If the entries of $\mathbf{H}$ have nonzero mean, $\mathbf{H}\mathbf{H}^\dagger$ is a non-central Wishart matrix.

Wishart Matrices: Eigenvectors

Theorem 1. *The matrix of eigenvectors of Wishart matrices is uniformly distributed on the manifold of unitary matrices (* Haar measure*)*

The Birth of Asymptotic Random Matrix Theory

E. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," *The Annals of Mathematics*, vol. 62, pp. 546–564, 1955.

$$\frac{1}{\sqrt{n}}\begin{bmatrix} 0 & +1 & +1 & -1 & -1 & +1 \\ +1 & 0 & -1 & -1 & +1 & +1 \\ +1 & -1 & 0 & +1 & +1 & -1 \\ -1 & -1 & +1 & 0 & +1 & +1 \\ -1 & +1 & +1 & +1 & 0 & -1 \\ +1 & +1 & -1 & +1 & -1 & 0 \end{bmatrix}$$

As the matrix dimension $n \to \infty$, the histogram of the eigenvalues converges to...?

Motivation: bypass the Schrödinger equation and explain the statistics of experimentally measured atomic energy levels in terms of the limiting spectrum of those random matrices.

Wigner Matrices

Definition 2. *An* $n \times n$ *Hermitian matrix* $\mathbf{W}$ *is a* Wigner *matrix if its upper-triangular entries are independent zero-mean random variables with identical variance. If the variance is* $\frac{1}{n}$, *then* $\mathbf{W}$ *is a standard Wigner matrix.*

Wigner Matrices: The Semicircle Law

E. Wigner, "On the distribution of roots of certain symmetric matrices," *The Annals of Mathematics*, vol. 67, pp. 325–327, 1958.

Theorem 2. *Consider an* $N \times N$ *standard Wigner matrix* $\mathbf{W}$ *such that, for some constant* κ*, and sufficiently large* N

$$\max_{1 \leq i \leq j \leq N} \mathbb{E}\left[|\mathsf{W}_{i,j}|^4\right] \leq \frac{\kappa}{N^2} \tag{3}$$

Then, the empirical distribution of $\mathbf{W}$ *converges almost surely to the semicircle law whose density is*

$$w(x) = \frac{1}{2\pi}\sqrt{4 - x^2} \tag{4}$$

with $|x| \leq 2$*.*

Wigner Matrices: The Semicircle Law

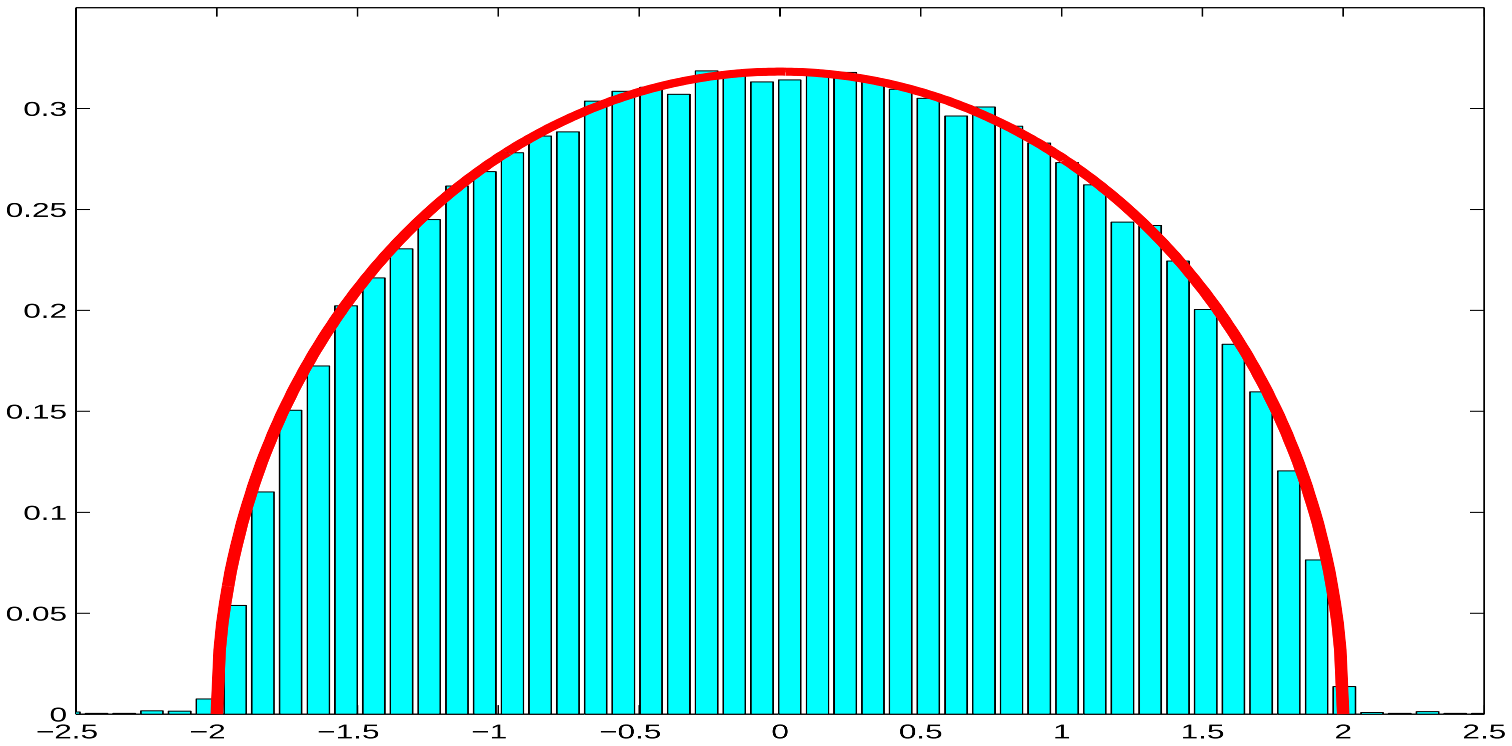

Figure 2: The semicircle law density function (4) compared with the histogram of the average of 100 empirical density functions for a Wigner matrix of size $n = 100$.

Wigner Matrices: The Semicircle Law

Wigner's original proof of the convergence to the semicircle law:

the empirical moments $\frac{1}{N}\mathrm{tr}\left\{\mathbf{W}^{2k}\right\} \to$ the Catalan numbers:

$$\begin{aligned}\lim_{N\to\infty}\frac{1}{N}\mathrm{tr}\left\{\mathbf{W}^{2k}\right\} &= \int_{-2}^{2} x^{2k} w(x)\,dx \\ &= \frac{1}{k+1}\binom{2k}{k}. \qquad (5)\end{aligned}$$

Asymptotics

Distribution Insensitivity: The asymptotic eigenvalue distribution does not depend on the distribution with which the independent matrix coefficients are generated.

“Ergodicity”: The eigenvalue histogram of one matrix realization converges almost surely to the asymptotic eigenvalue distribution.

Speed of Convergence: $8 = \infty$.

Gaussian case: Nonasymptotic joint distribution of eigenvalues known.

Gaussian Wigner Matrices: Nonasymptotic

M. L. Mehta and M. Gaudin, “On the density of the eigenvalues of a random matrix,” *Nuclear Physics*, vol. 18, pp. 420–427, 1960.

Theorem 3. *Let* $\mathbf{W}$ *be an* $n \times n$ *Wigner matrix whose entries are i.i.d. zero-mean Gaussian with unit variance. Then, its p.d.f. is*

$$2^{-n/2}\pi^{-n^2/2}\exp\left[-\frac{\mathrm{tr}\{\mathbf{W}^2\}}{2}\right] \tag{6}$$

while the joint p.d.f. of its ordered eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$ *is*

$$\frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\sum_{i=1}^{n}\lambda_i^2} \prod_{i=1}^{n-1}\frac{1}{i!}\prod_{i<j}^{n}(\lambda_i-\lambda_j)^2. \tag{7}$$

Gaussian Wigner Matrices: Nonasymptotic

E. Wigner, "Distribution laws for the roots of a random Hermitian matrix," in *Statistical Theories of Spectra: Fluctuations*, (C. E. Porter, ed.), New York: Academic, 1965.

Theorem 4. *Let* $\mathbf{W}$ *be an* $n \times n$ *Gaussian* Wigner *matrix. The marginal p.d.f. of the unordered eigenvalues is*

$$\frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{2^i\, i!\,\sqrt{2\pi}}\left(e^{-\frac{x^2}{4}}H_i(x)\right)^2 \tag{8}$$

with $H_i(\cdot)$ *the* i*th Hermite polynomial.*

Square matrix of iid coefficients

$$\frac{1}{\sqrt{n}}\begin{bmatrix} -1 & +1 & +1 & -1 & -1 & +1 \\ -1 & +1 & -1 & -1 & +1 & +1 \\ +1 & +1 & +1 & +1 & +1 & -1 \\ +1 & -1 & +1 & -1 & +1 & +1 \\ -1 & -1 & +1 & -1 & -1 & -1 \\ -1 & +1 & +1 & +1 & +1 & -1 \end{bmatrix}$$

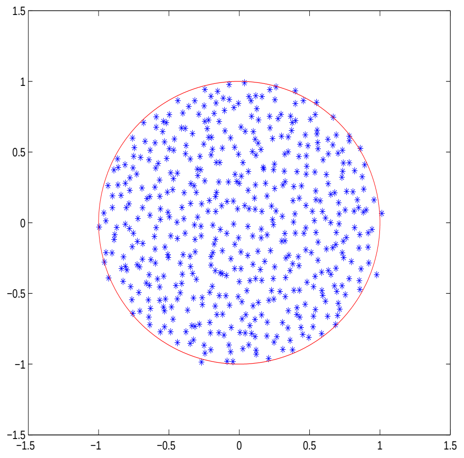

Figure 3: The full-circle law and the eigenvalues of a realization of a 500×500 matrix

Full Circle Law

V. L. Girko, "Circular law," *Theory Prob. Appl.*, vol. 29, pp. 694–706, 1984.
Z. D. Bai, "The circle law," *The Annals of Probability*, pp. 494–529, 1997.

Theorem 5. *Let $\mathbf{H}$ be an $N \times N$ complex random matrix whose entries are independent random variables with identical mean and variance and finite kth moments for $k \geq 4$. Assume that the joint distributions of the real and imaginary parts of the entries have uniformly bounded densities. Then, the asymptotic spectrum of $\mathbf{H}$ converges almost surely to the circular law, namely the uniform distribution over the unit disk on the complex plane $\{\zeta \in \mathbb{C} : |\zeta| \leq 1\}$ whose density is given by*

$$\mathsf{f}_c(\zeta) = \frac{1}{\pi} \qquad\qquad |\zeta| \leq 1 \tag{9}$$

(also holds for real matrices replacing the assumption on the joint distribution of real and imaginary parts with the one-dimensional distribution of the real-valued entries.)

Singular Values: Fisher-Hsu-Girshick-Roy

- R. A. Fisher, "The sampling distribution of some statistics obtained from non-linear equations," *The Annals of Eugenics*, vol. 9, pp. 238–249, 1939.
- M. A. Girshick, "On the sampling theory of roots of determinantal equations," *The Annals of Math. Statistics*, vol. 10, pp. 203–204, 1939.
- P. L. Hsu, "On the distribution of roots of certain determinantal equations," *The Annals of Eugenics*, vol. 9, pp. 250–258, 1939.
- S. N. Roy, "p-statistics or some generalizations in the analysis of variance appropriate to multivariate problems," *Sankhya*, vol. 4, pp. 381–396, 1939.

The joint p.d.f. of the ordered strictly positive eigenvalues of the Wishart matrix $\mathbf{H}\mathbf{H}^{\dagger}$, where the entries of $\mathbf{H}$ are i.i.d. complex Gaussian with zero mean and unit variance.

Singular Values^2: Fisher-Hsu-Girshick-Roy

Theorem 6. *Let the entries of* $\mathbf{H}$ *be i.i.d. complex Gaussian with zero mean and unit variance. The joint p.d.f. of the ordered strictly positive eigenvalues of the Wishart matrix* $\mathbf{H}\mathbf{H}^\dagger$, $\lambda_1 \geq \ldots \geq \lambda_t$, *equals*

$$e^{-\sum_{i=1}^{t} \lambda_i} \prod_{i=1}^{t} \frac{\lambda_i^{r-t}}{(t-i)!\,(r-i)!} \prod_{i<j}^{t} (\lambda_i - \lambda_j)^2 \tag{10}$$

where t *and* r *are the minimum and maximum of the dimensions of* $\mathbf{H}$. *The marginal p.d.f. of the unordered eigenvalues is*

$$\mathrm{g}_{r,t}(\lambda) = \frac{1}{t} \sum_{k=0}^{t-1} \frac{k!}{(k+r-t)!} \left[L_k^{r-t}(\lambda)\right]^2 \lambda^{r-t} e^{-\lambda} \tag{11}$$

where the Laguerre polynomials are

$$L_k^n(\lambda) = \frac{e^{\lambda}}{k!\lambda^n} \frac{d^k}{d\lambda^k} \left(e^{-\lambda} \lambda^{n+k}\right). \tag{12}$$

Singular Values^2: Fisher-Hsu-Girshick-Roy

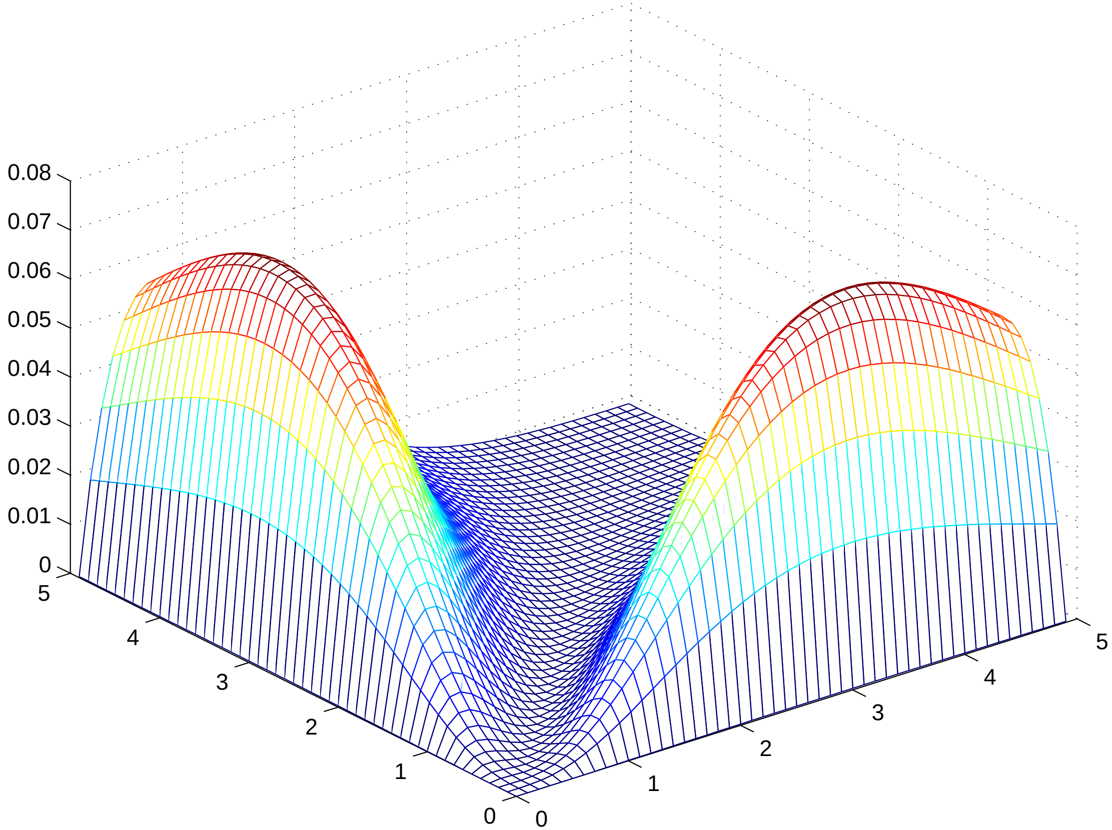

Figure 4: Joint p.d.f. of the unordered positive eigenvalues of the Wishart matrix $\mathbf{HH}^\dagger$ with $n = 3$ and $m = 2$. (Scaled version of (10).)

Asymptotic Distribution of Singular Values: Quarter circle law

Consider an $N \times N$ matrix $\mathbf{H}$ whose entries are independent zero-mean complex (or real) random variables with variance $\frac{1}{N}$, the asymptotic distribution of the singular values converges to

$$q(x) = \frac{1}{\pi}\sqrt{4 - x^2}, \quad 0 \leq x \leq 2 \tag{13}$$

Asymptotic Distribution of Singular Values: Quarter circle law

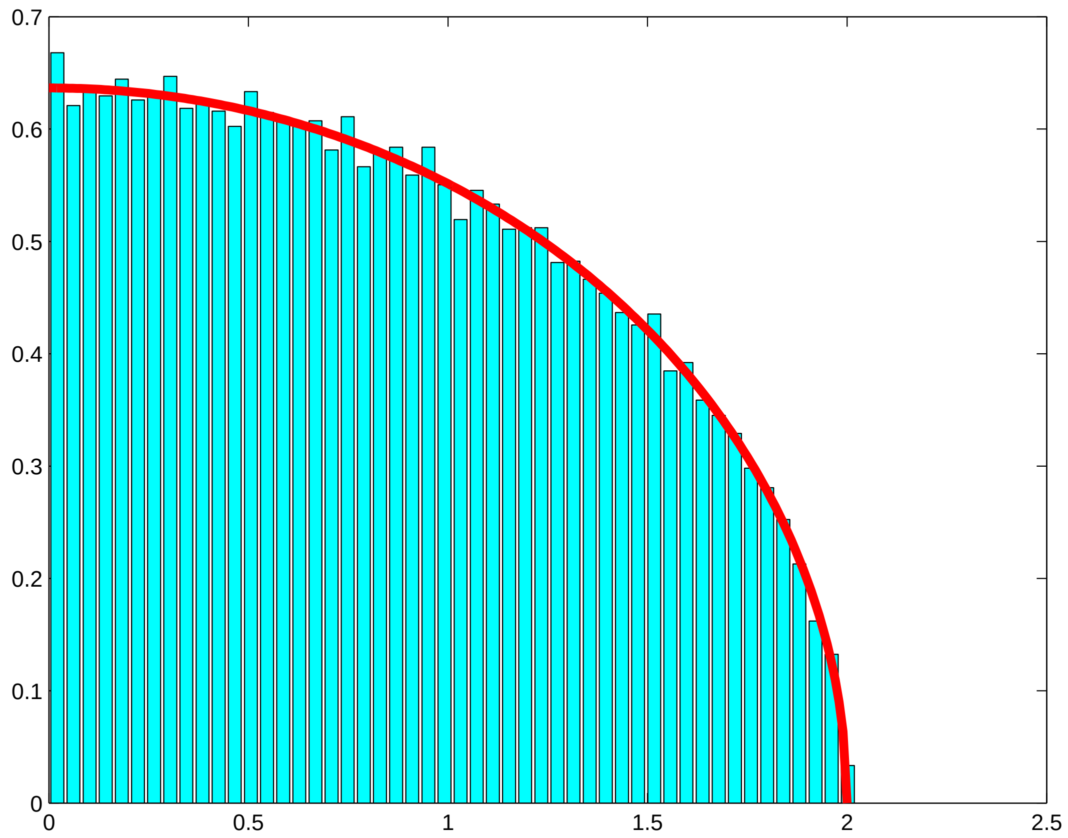

Figure 5: The quarter circle law compared a histogram of the average of 100 empirical singular value density functions of a matrix of size 100×100.

Minimum Singular Value of Gaussian Matrix

- A. Edelman, *Eigenvalues and condition number of random matrices*. PhD thesis, Dept. Mathematics, MIT, Cambridge, MA, 1989.

- J. Shen, "On the singular values of Gaussian random matrices," *Linear Algebra and its Applications*, vol. 326, no. 1-3, pp. 1–14, 2001.

Theorem 7. *The minimum singular value of an* $N \times N$ *standard complex Gaussian matrix* $\mathbf{H}$ *satisfies*

$$\lim_{N \to \infty} P[N\sigma_{\min} \geq x] = e^{-x-x^2/2}. \tag{14}$$

The Marčenko-Pastur Law

V. A. Marčenko and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices," *Math USSR-Sbornik*, vol. 1, pp. 457–483, 1967.

Theorem 8. *Consider an $N \times K$ matrix $\mathbf{H}$ whose entries are independent zero-mean complex (or real) random variables with variance $\frac{1}{N}$ and fourth moments of order $O(\frac{1}{N^2})$. As $K, N \to \infty$ with $\frac{K}{N} \to \beta$, the empirical distribution of $\mathbf{H}^\dagger \mathbf{H}$ converges almost surely to a nonrandom limiting distribution with density*

$$\mathsf{f}_\beta(x) = \left(1 - \frac{1}{\beta}\right)^+ \delta(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi\beta x} \tag{15}$$

where

$$a = (1 - \sqrt{\beta})^2 \qquad b = (1 + \sqrt{\beta})^2.$$

The Marčenko-Pastur Law

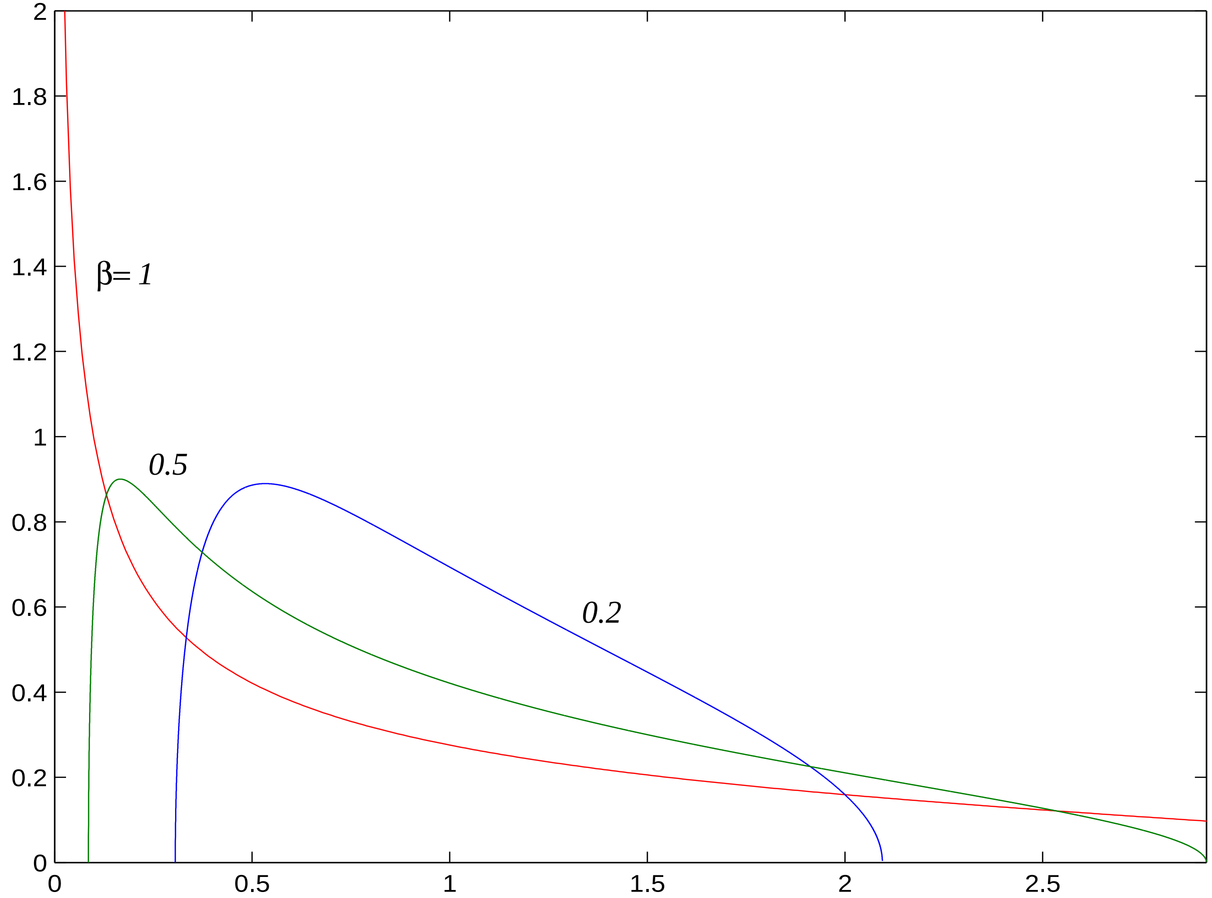

Figure 6: The Marčenko-Pastur density function for $\beta = 1, 0.5, 0.2$.

Rediscovering/Strenghtening the Marčenko-Pastur Law

- U. Grenander and J. W. Silverstein, “Spectral analysis of networks with random topologies,” *SIAM J. of Applied Mathematics*, vol. 32, pp. 449–519, 1977.
- K. W. Wachter, “The strong limits of random matrix spectra for sample matrices of independent elements,” *The Annals of Probability*, vol. 6, no. 1, pp. 1–18, 1978.
- J. W. Silverstein and Z. D. Bai, “On the empirical distribution of eigenvalues of a class of large dimensional random matrices,” *J. of Multivariate Analysis*, vol. 54, pp. 175–192, 1995.
- Y. L. Cun, I. Kanter, and S. A. Solla, “Eigenvalues of covariance matrices: Application to neural-network learning,” *Physical Review Letters*, vol. 66, pp. 2396–2399, 1991.

Generalizations needed!

- $\mathbf{W} = \mathbf{H}\mathbf{T}\mathbf{H}^\dagger$
- $\mathbf{W} = \mathbf{W}_0 + \mathbf{H}\mathbf{H}^\dagger$
- $\mathbf{W} = \mathbf{W}_0 + \mathbf{H}\mathbf{T}\mathbf{H}^\dagger$

Transforms

1. Stieltjes transform
2. η transform
3. Shannon transform
4. Mellin transform
5. R-transform
6. S-transform

Stieltjes Transform

Let X be a real-valued random variable with distribution $F_X(\cdot)$. Its Stieltjes transform is defined for complex arguments as

$$\mathcal{S}_X(z) = \mathbb{E}\left[\frac{1}{X - z}\right] = \int_{-\infty}^{\infty} \frac{1}{\lambda - z} \, dF_X(\lambda). \tag{16}$$

Stieltjes Transform and Moments

$$\mathcal{S}_X(z) = -\frac{1}{z}\sum_{k=0}^{\infty}\frac{\mathbb{E}[X^k]}{z^k}. \tag{17}$$

Inverse Stieltjes Transform

T. J. Stieltjes, “Recherches sur les fractions continues,” *Annales de la Faculte des Sciences de Toulouse*, vol. 8 (9), no. A (J), pp. 1–47 (1–122), 1894 (1895).

$$\mathsf{f}_X(\lambda) = \lim_{\omega \to 0^+} \frac{1}{\pi} \mathrm{Im}\Big[\mathcal{S}_X(\lambda + \mathsf{j}\,\omega)\Big]. \tag{18}$$

Stieltjes Transform of Semicircle law

$$w(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4 - x^2} & \text{if} \quad |x| \leq 2 \\ 0 & \text{if} \quad |x| > 2 \end{cases} \tag{19}$$

$$\mathcal{S}_w(z) = \frac{1}{2\pi}\int_{-2}^{2} \frac{\sqrt{4-\lambda^2}}{\lambda - z} d\lambda = \frac{1}{2}\left[-z \pm \sqrt{z^2 - 4}\right]. \tag{20}$$

Stieltjes Transform of Marčenko-Pastur law

$$\mathsf{f}_\beta(x) = \left(1 - \frac{1}{\beta}\right)^+ \delta(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi\beta x} \tag{21}$$

$$\begin{aligned} \mathcal{S}_{\mathsf{f}_\beta}(z) &= \int_a^b \frac{1}{\lambda - z} \mathsf{f}_\beta(\lambda)\, d\lambda \\ &= \frac{1 - \beta - z \pm \sqrt{z^2 - 2(\beta+1)z + (\beta-1)^2}}{2\beta z}. \end{aligned} \tag{22}$$

η transform

S. Verdú, "Large random matrices and wireless communications," *2002 MSRI Information Theory Workshop*, Feb 25–Mar 1, 2002.

Definition 3. *The η-transform of a nonnegative random variable X is*

$$\eta_X(\gamma) = \mathbb{E}\left[\frac{1}{1+\gamma X}\right] \tag{23}$$

where $\gamma \geq 0$.

Note: $0 < \eta_X(\gamma) \leq 1$.

Why Singular Values?

$$\text{Minimum Mean Square Error} = \frac{1}{K}\sum_{i=1}^{K}\frac{1}{1+\text{SNR}\,\lambda_i(\mathbf{H}^\dagger\mathbf{H})}$$

η transform and Stieltjes transform

$$\eta_X(\gamma) = \frac{\mathcal{S}_X(-\frac{1}{\gamma})}{\gamma} \tag{24}$$

$$\eta_X(\gamma) = \sum_{k=0}^{\infty} (-\gamma)^k \mathbb{E}[X^k], \tag{25}$$

η-Transform of Marčenko-Pastur law

Example: The η-transform of the Marčenko-Pastur law is

$$\eta(\gamma) = 1 - \frac{\mathcal{F}(\gamma, \beta)}{4\,\beta\,\gamma} \tag{26}$$

with

$$\mathcal{F}(x, z) = \left(\sqrt{x(1+\sqrt{z})^2+1} - \sqrt{x(1-\sqrt{z})^2+1}\right)^2 . \tag{27}$$

Random Quadratic Forms

Z. D. Bai and J. W. Silverstein, “No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices,” *The Annals of Probability*, vol. 26, pp. 316–345, 1998.

Theorem 9. *Let the components of the N-dimensional vector $\mathbf{x}$ be zero-mean and independent with variance $\frac{1}{N}$. For any $N \times N$ nonnegative definite random matrix $\mathbf{B}$ independent of $\mathbf{x}$ whose spectrum converges almost surely,*

$$\lim_{N\to\infty} \mathbf{x}^{\dagger} \left(\mathbf{I} + \gamma \mathbf{B}\right)^{-1} \mathbf{x} = \eta_{\mathbf{B}}(\gamma) \quad \textit{a.s.} \tag{28}$$

$$\lim_{N\to\infty} \mathbf{x}^{\dagger} \left(\mathbf{B} - z\mathbf{I}\right)^{-1} \mathbf{x} = \mathcal{S}_{\mathbf{B}}(z) \quad \textit{a.s.} \tag{29}$$

Shannon transform

S. Verdú, “Random matrices in wireless communication, proposal to the National Science Foundation,” Feb. 1999.

Definition 4. *The Shannon transform of a nonnegative random variable* X *is defined as*

$$\mathcal{V}_X(\gamma) = \mathbb{E}[\log(1 + \gamma X)]. \tag{30}$$

where $\gamma \geq 0$

Why Singular Values?

$$\text{Shannon Capacity} = \frac{1}{N}\sum_{i=1}^{N} \log\left(1 + \text{SNR}\,\lambda_i(\mathbf{H}\mathbf{H}^\dagger)\right)$$

Stieltjes, Shannon and η

$$\frac{\gamma}{\log e}\frac{d}{d\gamma}\mathcal{V}_X(\gamma) = 1 - \frac{1}{\gamma}\mathcal{S}_X\left(-\frac{1}{\gamma}\right) = 1 - \eta_X(\gamma) \tag{31}$$

Shannon transform of Marčenko-Pastur law

Example: The Shannon transform of the Marčenko-Pastur law is

$$\begin{aligned}\mathcal{V}(\gamma) &= \log\left(1+\gamma-\frac{1}{4}\mathcal{F}\left(\gamma,\beta\right)\right) \\ &+ \frac{1}{\beta}\log\left(1+\gamma\beta-\frac{1}{4}\mathcal{F}\left(\gamma,\beta\right)\right) \\ &- \frac{\log e}{4\,\beta\,\gamma}\mathcal{F}\left(\gamma,\beta\right) \qquad (32)\end{aligned}$$

Mellin transform

Definition 5. *The Mellin transform of a positive random variable X is given by*

$$\mathcal{M}_X(z) = \mathbb{E}[X^{z-1}] \tag{33}$$

where z belongs to a strip of the complex plane where the expectation is finite.

Mellin and Shannon transforms

R. Janaswamy, “Analytical expressions for the ergodic capacities of certain MIMO systems by the Mellin transform,” *Proc. of IEEE Global Telecomm. Conf.*, vol. 1, pp. 287–291, Dec. 2003.

Theorem 10.

$$\mathcal{V}_X(\gamma) = \mathcal{M}_{\Upsilon}^{-1}(\gamma) \tag{34}$$

where $\mathcal{M}_{\Upsilon}^{-1}$ *is the inverse Mellin transform of*

$$\Upsilon(z) = z^{-1}\Gamma(z)\Gamma(1-z)\mathcal{M}_X(1-z) \tag{35}$$

R-transform

D. Voiculescu, "Addition of certain non-commuting random variables," *J. Funct. Analysis*, vol. 66, pp. 323–346, 1986.

Definition 6.

$$\mathsf{R}_X(z) = \mathcal{S}_X^{-1}(-z) - \frac{1}{z}. \tag{36}$$

R-transform and η-transform

$$\eta_X(\gamma) = \frac{1}{1 + \gamma\, \mathsf{R}_X(-\gamma\, \eta_X(\gamma))} \tag{37}$$

R-transform of the semicircle law

$$\mathrm{R}(z) = z. \tag{38}$$

R-transform of the Marčenko-Pastur law

$$\mathsf{R}(z) = \frac{1}{1 - \beta z}. \tag{39}$$

R-transform: Key theorem

D. Voiculescu, “Addition of certain non-commuting random variables,” *J. Funct. Analysis*, vol. 66, pp. 323–346, 1986.

Theorem 11. *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are asymptotically free random matrices, then the R-transform of their sum satisfies*

$$\mathsf{R}_{\mathbf{A}+\mathbf{B}}(z) = \mathsf{R}_{\mathbf{A}}(z) + \mathsf{R}_{\mathbf{B}}(z) \tag{40}$$

free analog of the log-moment generating function

S-transform

D. Voiculescu, “Multiplication of certain non-commuting random variables,” *J. Operator Theory*, vol. 18, pp. 223–235, 1987.

Definition 7. *The S-transform of a nonnegative random variable X is*

$$\Sigma_X(x) = -\frac{x+1}{x}\,\eta_X^{-1}(1+x), \qquad (41)$$

which maps $(-1, 0)$ onto the positive real line.

S-transform of the Marčenko-Pastur law

$$\Sigma(x) = \frac{1}{1 + \beta x}. \tag{42}$$

S-transform: Key Theorem

D. Voiculescu, "Multiplication of certain non-commuting random variables," *J. Operator Theory*, vol. 18, pp. 223–235, 1987.

Theorem 12. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be nonnegative asymptotically free random matrices. The S-transform of their product satisfies*

$$\Sigma_{\mathbf{AB}}(x) = \Sigma_{\mathbf{A}}(x)\Sigma_{\mathbf{B}}(x) \tag{43}$$

Generalizations needed!

- $\mathbf{W} = \mathbf{H}\mathbf{T}\mathbf{H}^\dagger$
- $\mathbf{W} = \mathbf{W}_0 + \mathbf{H}\mathbf{H}^\dagger$
- $\mathbf{W} = \mathbf{W}_0 + \mathbf{H}\mathbf{T}\mathbf{H}^\dagger$

Asymptotic Spectrum of $\mathbf{HTH}^\dagger$

V. A. Marčenko and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices," *Math USSR-Sbornik*, vol. 1, pp. 457–483, 1967.

Theorem 13.

- *Let $\mathbf{H}$ be an $N \times K$ matrix whose entries are i.i.d. complex random variables with variance $\frac{1}{N}$.*

- *Let $\mathbf{T}$ be a $K \times K$ Hermitian nonnegative random matrix, independent of $\mathbf{H}$, whose empirical eigenvalue distribution converges almost surely to a nonrandom limit.*

The empirical eigenvalue distribution of $\mathbf{HTH}^\dagger$ converges almost surely, as $K, N \to \infty$ with $\frac{K}{N} \to \beta$, with $\eta_{\mathbf{HTH}^\dagger}(\gamma) = \eta$ the solution of:

$$\beta = \frac{1 - \eta}{1 - \eta_{\mathbf{T}}(\gamma\eta)} \tag{44}$$

Example

Further, if $\mathbf{T} = \mathbf{I}$, we have $\eta_T(\gamma) = \frac{1}{1+\gamma}$, and (44) becomes:

$$\eta = 1 - \beta + \frac{\beta}{1 + \gamma\eta}$$

Asymptotic Spectrum of $\mathbf{HTH}^\dagger$

Theorem 14.

- *Let* $\mathbf{H}$ *be an* $N \times K$ *matrix whose entries are i.i.d. complex random variables with variance* $\frac{1}{N}$*.*

- *Let* $\mathbf{T}$ *be a* $K \times K$ *Hermitian nonnegative random matrix, independent of* $\mathbf{H}$*, whose empirical eigenvalue distribution converges almost surely to a nonrandom limit.*

The empirical eigenvalue distribution of $\mathbf{HTH}^\dagger$ *converges almost surely, as* $K, N \to \infty$ *with* $\frac{K}{N} \to \beta$*, with Shannon transform*

$$\mathcal{V}_{\mathbf{HTH}^\dagger}(\gamma) = \beta \mathcal{V}_{\mathbf{T}}(\eta\gamma) + \log\frac{1}{\eta} + (\eta - 1)\log e \tag{45}$$

Asymptotic Spectrum of $\mathbf{W}_0 + \mathbf{HTH}^\dagger$

A. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications,* Now Publishers, 2004

Theorem 15.

- *Let $\mathbf{H}$ be an $N \times K$ matrix whose entries are i.i.d. complex random variables with zero-mean and variance $\frac{1}{N}$.*

- *Let $\mathbf{T}$ be a $K \times K$ positive definite random matrix whose empirical eigenvalue distribution converges almost surely to a nonrandom limit.*

- *Let $\mathbf{W}_0$ be an $N \times N$ nonnegative definite diagonal random matrix with empirical eigenvalue distribution converging almost surely to a nonrandom limit.*

- $\mathbf{H}$*,* $\mathbf{T}$*, and* $\mathbf{W}_0$ *are independent*

The empirical eigenvalue distribution of

$$\mathbf{W} = \mathbf{W}_0 + \mathbf{H}\mathbf{T}\mathbf{H}^\dagger \tag{46}$$

converges almost surely, as $K, N \to \infty$ with $\frac{K}{N} \to \beta$, to a nonrandom limiting distribution whose η-transform is the solution of the following pair of equations:

$$\gamma\,\eta = \varphi\,\eta_0\,(\varphi) \tag{47}$$

$$\eta = \eta_0\,(\varphi) - \beta\,(1 - \eta_{\mathbf{T}}(\gamma\,\eta)) \tag{48}$$

with η_0 and $\eta_{\mathbf{T}}$ the η-transforms of $\mathbf{W}_0$ and $\mathbf{T}$ respectively.