Largest Eigenvalue Distributions & Their Applications

Craig A. Tracy
UC Davis

Interface Fluctuations and KPZ Universality Class
YITP Workshop
August 20–23, 2014

Outline

- Random matrix models

Outline

- Random matrix models
- Limiting distribution of largest eigenvalue, $F_\beta(x)$, for $\beta = 1, 2, 4$.
  - Fredholm determinant & Painlevé representations
  - Tail behavior
  - Next-largest, etc. eigenvalue distributions

Outline

- Random matrix models
- Limiting distribution of largest eigenvalue, $F_\beta(x)$, for $\beta = 1, 2, 4$.
  - Fredholm determinant & Painlevé representations
  - Tail behavior
  - Next-largest, etc. eigenvalue distributions
- Beyond determinantal class
  - Wigner matrices & universality theorems
  - F_β for $\beta > 0$
  - ASEP

Outline

- Random matrix models
- Limiting distribution of largest eigenvalue, $F_\beta(x)$, for $\beta = 1, 2, 4$.
  - Fredholm determinant & Painlevé representations
  - Tail behavior
  - Next-largest, etc. eigenvalue distributions
- Beyond determinantal class
  - Wigner matrices & universality theorems
  - F_β for $\beta > 0$
  - ASEP
- Role of F_β in data analysis

Random Matrix Models (RMM)

A RMM is a probability space $(\Omega, \mathbb{P}, \mathcal{F})$ where the sample space Ω is a set of matrices.

Random Matrix Models (RMM)

A RMM is a probability space $(\Omega, \mathbb{P}, \mathcal{F})$ where the sample space Ω is a set of matrices.

Three classic models: GOE, GUE and GSE where the $N \times N$ matrices are real symmetric, Hermitian or symplectic Hermitian and have probability density proportional to

$$\exp\left(-\mathrm{Tr}(A^2)\right)$$

Random Matrix Models (RMM)

A RMM is a probability space $(\Omega, \mathbb{P}, \mathcal{F})$ where the sample space Ω is a set of matrices.

Three classic models: GOE, GUE and GSE where the $N \times N$ matrices are real symmetric, Hermitian or symplectic Hermitian and have probability density proportional to

$$\exp\left(-\mathrm{Tr}(A^2)\right)$$

This measure $\mathbb{P}_{\beta,N}$ induces a probability density on the eigenvalues which has the form

$$P_{\beta,N}(x_1,\ldots,x_N) = C_{N,\beta} \prod_{i<j} |x_i - x_j|^\beta \prod_j e^{-\beta x_j^2/2}, \quad \beta = 1, 2, 4.$$

Largest Eigenvalue Distributions

Let $x_{\max}$ denote the largest (random) eigenvalue. Then

$$\begin{aligned} F_{\beta,N}(x) : &= \mathbb{P}_{\beta,N}(x_{\max} < x) = \mathbb{P}_{\beta,N}(x_1 < x, \ldots, x_N < x) \\ &= \int_{-\infty}^{x} \cdots \int_{-\infty}^{x} P_{\beta,N}(x_1, \ldots, x_N)\, dx_1 \cdots dx_N \end{aligned}$$

[1] Actually, Gaudin looked a different statistic—level spacing distribution—but his proof carries over to the largest eigenvalue distribution.

Largest Eigenvalue Distributions

Let $x_{\max}$ denote the largest (random) eigenvalue. Then

$$\begin{aligned} F_{\beta,N}(x) : &= \mathbb{P}_{\beta,N}(x_{\max} < x) = \mathbb{P}_{\beta,N}(x_1 < x, \ldots, x_N < x) \\ &= \int_{-\infty}^{x} \cdots \int_{-\infty}^{x} P_{\beta,N}(x_1, \ldots, x_N)\, dx_1 \cdots dx_N \end{aligned}$$

For $\beta = 2$ (unitary case) it was M. Gaudin who in 1961 showed that $F_{2,N}$ is a *Fredholm determinant*[1] whose kernel is

$$\begin{aligned} K_N(x,y) &= \left(\frac{N}{2}\right)^{\frac{1}{2}} \frac{\phi_N(x)\phi_{N-1}(y) - \phi_{N-1}(x)\phi_N(y)}{x-y}, \\ & \phi_n(x) = \text{harmonic oscillator wave fns.} \end{aligned}$$

[1] Actually, Gaudin looked a different statistic—level spacing distribution—but his proof carries over to the largest eigenvalue distribution.

Largest Eigenvalue Distributions

Let $x_{\max}$ denote the largest (random) eigenvalue. Then

$$
\begin{aligned}
F_{\beta,N}(x) : &= \mathbb{P}_{\beta,N}(x_{\max} < x) = \mathbb{P}_{\beta,N}(x_1 < x, \ldots, x_N < x) \\
&= \int_{-\infty}^{x} \cdots \int_{-\infty}^{x} P_{\beta,N}(x_1, \ldots, x_N)\, dx_1 \cdots dx_N
\end{aligned}
$$

For $\beta = 2$ (unitary case) it was M. Gaudin who in 1961 showed that $F_{2,N}$ is a *Fredholm determinant*[1] whose kernel is

$$
\begin{aligned}
K_N(x,y) &= \left(\frac{N}{2}\right)^{\frac{1}{2}} \frac{\phi_N(x)\phi_{N-1}(y) - \phi_{N-1}(x)\phi_N(y)}{x-y}, \\
& \quad \phi_n(x) = \text{harmonic oscillator wave fns.}
\end{aligned}
$$

We are interested in a limit theorem as $N \to \infty$.

[1] Actually, Gaudin looked a different statistic—level spacing distribution—but his proof carries over to the largest eigenvalue distribution.

Using the known asymptotics of Hermite polynomials in the transition region (*Plancherel-Rotach formulas*) it is an easy exercise to show that (at least formally) that the operator K_N converges to the *Airy kernel*[2]

$$K_{\mathrm{Airy}}(x, y) = \frac{\mathsf{Ai}(x)\,\mathsf{Ai}'(y) - \mathsf{Ai}'(x)\,\mathsf{Ai}(y)}{x - y}, \;\; \mathsf{Ai}(x) = \text{Airy function.}$$

[2] For a proof one must show trace-class convergence of the operator.

Using the known asymptotics of Hermite polynomials in the transition region (*Plancherel-Rotach formulas*) it is an easy exercise to show that (at least formally) that the operator K_N converges to the *Airy kernel*[2]

$$K_{\mathrm{Airy}}(x,y) = \frac{\mathsf{Ai}(x)\,\mathsf{Ai}'(y) - \mathsf{Ai}'(x)\,\mathsf{Ai}(y)}{x-y}, \;\; \mathsf{Ai}(x) = \text{Airy function}.$$

Thus in distribution

$$F_2(s) := \lim_{N\to\infty} \mathbb{P}_{2,N}\left[N^{2/3}\left(\frac{x_{\mathrm{max}}}{\sqrt{N}} - 2\right) \le s\right] = \det(I - K_{\mathrm{Airy}})$$

where K_{Airy} acts on $L^2((s,\infty))$.

[2] For a proof one must show trace-class convergence of the operator.

Using the known asymptotics of Hermite polynomials in the transition region (*Plancherel-Rotach formulas*) it is an easy exercise to show that (at least formally) that the operator K_N converges to the *Airy kernel*[2]

$$K_{\mathrm{Airy}}(x,y) = \frac{\mathsf{Ai}(x)\,\mathsf{Ai}'(y) - \mathsf{Ai}'(x)\,\mathsf{Ai}(y)}{x-y}, \;\; \mathsf{Ai}(x) = \text{Airy function}.$$

Thus in distribution

$$F_2(s) := \lim_{N\to\infty} \mathbb{P}_{2,N}\left[N^{2/3}\left(\frac{x_{\mathrm{max}}}{\sqrt{N}} - 2\right) \le s\right] = \det(I - K_{\mathrm{Airy}})$$

where K_{Airy} acts on $L^2((s,\infty))$.

For $\beta = 1,4$ the analysis is more involved.

[2] For a proof one must show trace-class convergence of the operator.

Using the known asymptotics of Hermite polynomials in the transition region (*Plancherel-Rotach formulas*) it is an easy exercise to show that (at least formally) that the operator K_N converges to the *Airy kernel*[2]

$$K_{\mathrm{Airy}}(x,y) = \frac{\mathsf{Ai}(x)\,\mathsf{Ai}'(y) - \mathsf{Ai}'(x)\,\mathsf{Ai}(y)}{x-y}, \;\; \mathsf{Ai}(x) = \text{Airy function.}$$

Thus in distribution

$$F_2(s) := \lim_{N\to\infty} \mathbb{P}_{2,N}\left[N^{2/3}\left(\frac{x_{\max}}{\sqrt{N}} - 2\right) \le s\right] = \det(I - K_{\mathrm{Airy}})$$

where K_{Airy} acts on $L^2((s,\infty))$.

For $\beta = 1,4$ the analysis is more involved.

First of all the resulting kernels are 2×2 matrix kernels; and secondly, for the orthogonal case ($\beta = 1$) the convergence is no longer trace-class and regularized determinants are involved.

[2]For a proof one must show trace-class convergence of the operator.

Using the known asymptotics of Hermite polynomials in the transition region (*Plancherel-Rotach formulas*) it is an easy exercise to show that (at least formally) that the operator K_N converges to the *Airy kernel*[2]

$$K_{\mathrm{Airy}}(x,y) = \frac{\mathsf{Ai}(x)\,\mathsf{Ai}'(y) - \mathsf{Ai}'(x)\,\mathsf{Ai}(y)}{x-y}, \ \ \mathsf{Ai}(x) = \text{Airy function.}$$

Thus in distribution

$$F_2(s) := \lim_{N\to\infty} \mathbb{P}_{2,N}\left[N^{2/3}\left(\frac{x_{\max}}{\sqrt{N}} - 2\right) \le s\right] = \det(I - K_{\mathrm{Airy}})$$

where K_{Airy} acts on $L^2((s,\infty))$.

For $\beta = 1,4$ the analysis is more involved.

First of all the resulting kernels are 2×2 matrix kernels; and secondly, for the orthogonal case ($\beta = 1$) the convergence is no longer trace-class and regularized determinants are involved.

F. Dyson (1970) was the first to recognize that matrix kernels arise in the orthogonal and symplectic cases. This was further developed by M. L. Mehta, TW and others.

[2] For a proof one must show trace-class convergence of the operator.

Integrable Structure: Painlevé Representations

The *sine kernel* (bulk scaling limit) and the *Airy kernel* are both of the form

$$K(x,y) = \frac{\varphi(x)\psi(y) - \psi(x)\varphi(y)}{x-y} \tag{1}$$

where φ, ψ satisfy DEs of the form

$$\frac{d}{dx}\begin{pmatrix} \varphi(x) \\ \psi(x) \end{pmatrix} = \Omega(x)\begin{pmatrix} \varphi(x) \\ \psi(x) \end{pmatrix}$$

where $\Omega(x)$ is a 2×2 traceless matrix whose elements are rational functions of x.

Integrable Structure: Painlevé Representations

The *sine kernel* (bulk scaling limit) and the *Airy kernel* are both of the form

$$K(x,y) = \frac{\varphi(x)\psi(y) - \psi(x)\varphi(y)}{x - y} \tag{1}$$

where φ, ψ satisfy DEs of the form

$$\frac{d}{dx}\begin{pmatrix} \varphi(x) \\ \psi(x) \end{pmatrix} = \Omega(x)\begin{pmatrix} \varphi(x) \\ \psi(x) \end{pmatrix}$$

where $\Omega(x)$ is a 2×2 traceless matrix whose elements are rational functions of x.

For example, for the Airy kernel

$$\Omega(x) = \begin{pmatrix} 0 & 1 \\ x & 0 \end{pmatrix}$$

Integrable Structure: Painlevé Representations

The *sine kernel* (bulk scaling limit) and the *Airy kernel* are both of the form

$$K(x,y) = \frac{\varphi(x)\psi(y) - \psi(x)\varphi(y)}{x-y} \tag{1}$$

where φ, ψ satisfy DEs of the form

$$\frac{d}{dx}\begin{pmatrix} \varphi(x) \\ \psi(x) \end{pmatrix} = \Omega(x)\begin{pmatrix} \varphi(x) \\ \psi(x) \end{pmatrix}$$

where $\Omega(x)$ is a 2×2 traceless matrix whose elements are rational functions of x.

For example, for the Airy kernel

$$\Omega(x) = \begin{pmatrix} 0 & 1 \\ x & 0 \end{pmatrix}$$

Let J be a union of open intervals: $J = \bigcup_{j=1}^{m}(a_{2j-1}, a_{2j})$, and consider the Fredholm determinant

$$\tau(a) := \det(I - K)$$

where K has kernel (1) and acts on $L^2(J)$.

It is a theorem (TW) that $\tau(a)$ can be expressed in terms of a solution to a total system of PDEs (independent variables the end-points a_j).

It is a theorem (TW) that $\tau(a)$ can be expressed in terms of a solution to a total system of PDEs (independent variables the end-points a_j).

For the Airy kernel and $J = (s, \infty)$, the PDEs reduce to a single ODE—the *Painlevé II* DE:

$$\begin{aligned} F_2(s) &= \exp\left(-\int_s^\infty (x-s)q^2(x)\,dx\right) := (F(s))^2 \\ \frac{d^2q}{dx^2} &= x\,q + 2q^3 \\ q(x) &\sim \mathrm{Ai}(x),\ \ x\to\infty,\ \ \text{Hastings-McLeod solution} \end{aligned}$$

It is a theorem (TW) that $\tau(a)$ can be expressed in terms of a solution to a total system of PDEs (independent variables the end-points a_j).

For the Airy kernel and $J = (s, \infty)$, the PDEs reduce to a single ODE—the *Painlevé II* DE:

$$\begin{aligned} F_2(s) &= \exp\left(-\int_s^\infty (x-s) q^2(x)\, dx\right) := (F(s))^2 \\ \frac{d^2 q}{dx^2} &= x\, q + 2q^3 \\ q(x) &\sim \mathrm{Ai}(x), \;\; x \to \infty, \;\; \text{Hastings-McLeod solution} \end{aligned}$$

Somewhat surprisingly, both F_1 and F_4 can be expressed in terms of the same Painlevé II function q, e.g.

$$F_1(s) = \exp\left(-\frac{1}{2}\int_s^\infty q(x)\, dx\right) (F_2(s))^{1/2} := E(x)F(x).$$

Historical remarks:

- The first connection between Fredholm deteminants/Toeplitz determinants and Painlevé functions was in the 2D Ising model where the 2-point massive scaling functions were shown to be expressible in terms of Painlevé III (Wu, McCoy, CT, Barouch, 1973–76).

Historical remarks:

- The first connection between Fredholm deteminants/Toeplitz determinants and Painlevé functions was in the 2D Ising model where the 2-point massive scaling functions were shown to be expressible in terms of Painlevé III (Wu, McCoy, CT, Barouch, 1973–76).
- Total systems of PDEs for Ising n-point scaling functions were derived by Sato, Miwa and Jimbo (1977–79)—isomondromy enters statistical physics.

Historical remarks:

- The first connection between Fredholm deteminants/Toeplitz determinants and Painlevé functions was in the 2D Ising model where the 2-point massive scaling functions were shown to be expressible in terms of Painlevé III (Wu, McCoy, CT, Barouch, 1973–76).
- Total systems of PDEs for Ising *n*-point scaling functions were derived by Sato, Miwa and Jimbo (1977–79)—isomondromy enters statistical physics.
- For the Fredholm determinant with sine kernel, Jimbo, Miwa, Môri, Sato (1980) showed the connection with Painlevé V.

Historical remarks:

- The first connection between Fredholm deteminants/Toeplitz determinants and Painlevé functions was in the 2D Ising model where the 2-point massive scaling functions were shown to be expressible in terms of Painlevé III (Wu, McCoy, CT, Barouch, 1973–76).
- Total systems of PDEs for Ising n-point scaling functions were derived by Sato, Miwa and Jimbo (1977–79)—isomondromy enters statistical physics.
- For the Fredholm determinant with sine kernel, Jimbo, Miwa, Môri, Sato (1980) showed the connection with Painlevé V.
- TW developed the general theory above with parallel developments by Adler, van Moerbeke and Shiota.

Historical remarks:

- The first connection between Fredholm deteminants/Toeplitz determinants and Painlevé functions was in the 2D Ising model where the 2-point massive scaling functions were shown to be expressible in terms of Painlevé III (Wu, McCoy, CT, Barouch, 1973–76).
- Total systems of PDEs for Ising n-point scaling functions were derived by Sato, Miwa and Jimbo (1977–79)—isomondromy enters statistical physics.
- For the Fredholm determinant with sine kernel, Jimbo, Miwa, Môri, Sato (1980) showed the connection with Painlevé V.
- TW developed the general theory above with parallel developments by Adler, van Moerbeke and Shiota.
- Riemann-Hilbert methods by Its, Deift and others further developed this area particularly with regards to the difficult questions of asymptotics (connection formulas).

Some numerics

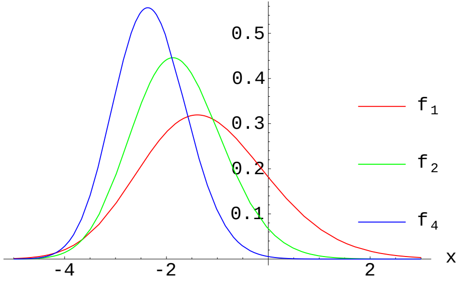


Figure : Largest eigenvalue densities $f_\beta(x) = dF_\beta/dx$, $\beta = 1, 2, 4$.

Some numerics

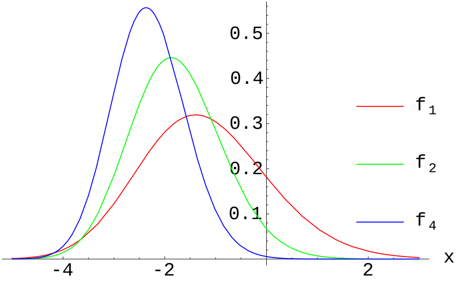
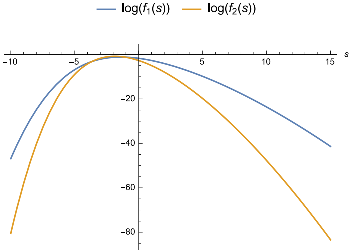


Figure : Largest eigenvalue densities $f_\beta(x) = dF_\beta/dx$, $\beta = 1, 2, 4$.

Table : The mean (μ_β), variance (σ_β^2), skewness (S_β) and kurtosis (K_β) of F_β.

β	μ_β	σ_β^2	S_β	K_β
1	-1.206 533 574	1.607 781 034	0.293 464 524	0.165 242 938
2	-1.771 086 807	0.813 194 792	0.224 084 203	0.093 448 087
4	-2.306 884 489	0.517 723 720	0.165 509 494	0.049 195 156

Tail behavior of F_β



As $x \to +\infty$

$$F(x) - 1 \sim -\frac{\exp(-\frac{4}{3}x^{3/2})}{32\pi x^{3/2}}, \quad E(x) - 1 \sim -\frac{\exp(-\frac{2}{3}x^{3/2})}{4\sqrt{\pi}x^{3/2}}$$

As $x \to -\infty$

$$F_1(x) \sim \tau_1 \frac{\exp(-\frac{1}{24}|x|^3 - \frac{1}{3\sqrt{2}}|x|^{3/2})}{|x|^{1/16}}, \quad F_2(x) \sim \tau_2 \frac{\exp(-\frac{1}{12}|x|^3)}{|x|^{1/8}}$$

$$\tau_1 = 2^{-11/48}\, e^{\zeta'(-1)}, \tau_2 = 2^{1/24}\, e^{\zeta'(-1)}$$

Next-largest, next-next-largest, etc. eigenvalue distributions

There are Painlevé II type representations. (TW, $\beta = 2$; M. DIENG, $\beta = 1, 4$.)

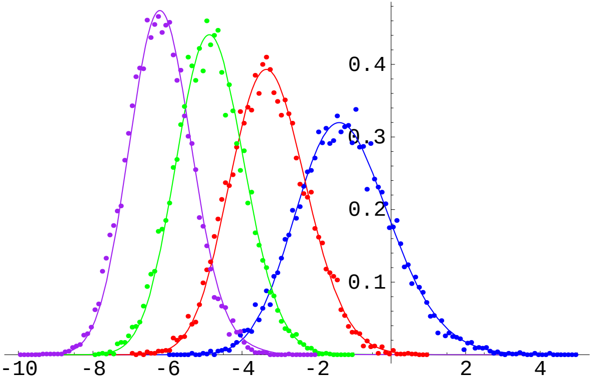
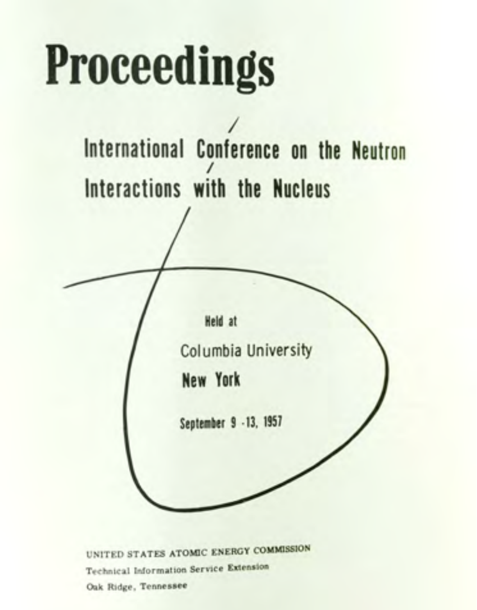


Figure : Histogram of the four largest (centered and normalized) eigenvalues for 10^4 realizations of 1000×1000 GOE matrices. Solid curves are F_1 and Dieng's limiting next-largest, etc. distributions. Figure courtesy of Dieng.

Beyond Determinantal Class

Some historical remarks concerning universality



SESSION IIB

INTERPRETATION OF LOW ENERGY NEUTRON SPECTROSCOPY

CHAIRMAN—W. W. Havens, Jr.

IIB1. DISTRIBUTION OF NEUTRON RESONANCE LEVEL SPACING.

E. P. WIGNER, *Princeton University*
Presented by E. P. Wigner

The problem of the spacing of levels is neither a terribly important one nor have I solved it. That is really the point which I want to make very definitely. As we go up in the energy scale it is evident that the detailed analyses which we have seen for low

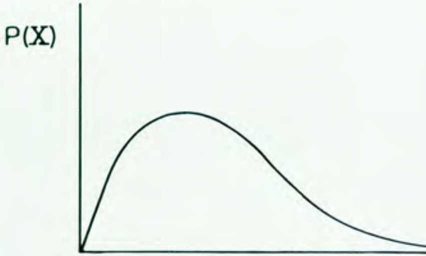


Fig. IIB1-1. Probability of a level spacing X.

Let me say only one more word. It is very likely that the curve in Figure I is a universal function. In other words, it doesn't depend on the details of the model with which you are working. There is one particular model in which the probability of the

Let me say only one more word. It is very likely that the curve in Figure I is a universal function. In other words, it doesn't depend on the details of the model with which you are working. There is one particular model in which the probability of the

This is the origin of the famous Wigner-Dyson-Mehta **bulk universality conjecture** (sine-kernel).

Let me say only one more word. It is very likely that the curve in Figure I is a universal function. In other words, it doesn't depend on the details of the model with which you are working. There is one particular model in which the probability of the

This is the origin of the famous Wigner-Dyson-Mehta **bulk universality conjecture** (sine-kernel).

First proof for Wigner matrices: Tau & Vu (2009) and independently Erdös, Péché, Ramírez, Schlein, Yau (2009). Wigner matrices

$$A = \frac{1}{\sqrt{N}} (a_{ij}) , \text{ algebraically indep. elements } a_{ij} \text{ are i.i.d. random variables}$$

Universality theorem for largest eigenvalue of Wigner matrices

A. Soshnikov (1999): Properly centered and normalized largest eigenvalue converges in distribution to the corresponding Gaussian result; namely, F_β. Since Soshnikov's work the moment conditions in his theorem have been relaxed (Tao, Vu, Erdös, Yau and coworkers)

Universality theorem for largest eigenvalue of Wigner matrices

A. Soshnikov (1999): Properly centered and normalized largest eigenvalue converges in distribution to the corresponding Gaussian result; namely, F_β. Since Soshnikov's work the moment conditions in his theorem have been relaxed (Tao, Vu, Erdös, Yau and coworkers)

> *The importance of universality for Wigner matrices is that these ensembles of matrices are representative of the "non-integrable" case; namely, "integrable techniques" such as Fredholm determinants, Riemann-Hilbert methods, Painlevé theory are not directly applicable to the Wigner case. Yet the limit theorems found by these integrable techniques continue to hold in a broader class of ensembles.*

The Beta Ensembles: F_β for $\beta > 0$

Recall

$$P_{\beta,N}(x_1, \ldots, x_N) = C_{N,\beta} \prod_{i<j} |x_i - x_j|^\beta \prod_j e^{-\beta x_j^2/2}, \qquad (2)$$

which makes sense for all real, positive β.

The Beta Ensembles: F_β for $\beta > 0$

Recall

$$P_{\beta,N}(x_1, \ldots, x_N) = C_{N,\beta} \prod_{i<j} |x_i - x_j|^\beta \prod_j e^{-\beta x_j^2/2}, \tag{2}$$

which makes sense for all real, positive β.

I. Dumitriu and A. Edelman discovered a class of tridiagonal random matrices H_N^β for which (2) is the induced probability density for the eigenvalues.

The Beta Ensembles: F_β for $\beta > 0$

Recall

$$P_{\beta,N}(x_1,\ldots,x_N) = C_{N,\beta} \prod_{i<j} |x_i - x_j|^\beta \prod_j e^{-\beta x_j^2/2}, \tag{2}$$

which makes sense for all real, positive β.

I. Dumitriu and A. Edelman discovered a class of tridiagonal random matrices H_N^β for which (2) is the induced probability density for the eigenvalues.

B. Sutton and A. Edelman gave heuristic arguments that the rescaled operators

$$\tilde{H}_N^\beta = N^{1/6}\left(2\sqrt{N} - H_N^\beta\right)$$

converges to the differential operator

$$\mathcal{H}_\beta = -\frac{d^2}{dx^2} + x + \frac{2}{\sqrt{\beta}} b'_x$$

where b' is white noise.

Theorem (Ramírez, Rider, Virág): With probability one, for each $k \geq 0$ the set of eigenvalues of $\mathcal{H}_\beta$ has a well-defined $(k+1)$st lowest element Λ_k. Moreover, let $\lambda_1 \geq \lambda_2 \geq \cdots$ denote the eigenvalues of the Hermite β-ensemble H_N^β. Then the vector

$$\left(N^{1/6}(2\sqrt{N} - \lambda_{\beta,\ell})\right)_{\ell=1,\dots,k}$$

converges in distribution as $N \to \infty$ to $(\Lambda_0, \Lambda_1, \dots, \Lambda_{k-1})$. The distribution of $-\Lambda_0(\beta)$ defines F_β for all $\beta > 0$.

Theorem (Ramírez, Rider, Virág): With probability one, for each $k \geq 0$ the set of eigenvalues of $\mathcal{H}_\beta$ has a well-defined $(k+1)$st lowest element Λ_k. Moreover, let $\lambda_1 \geq \lambda_2 \geq \cdots$ denote the eigenvalues of the Hermite β-ensemble H_N^β. Then the vector

$$\left(N^{1/6}(2\sqrt{N} - \lambda_{\beta,\ell})\right)_{\ell=1,\dots,k}$$

converges in distribution as $N \to \infty$ to $(\Lambda_0, \Lambda_1, \dots, \Lambda_{k-1})$. The distribution of $-\Lambda_0(\beta)$ defines F_β for all $\beta > 0$.

As a corollary of their theorem

$$-\Lambda_0(\beta) = \sup_f \left\{ \frac{2}{\sqrt{\beta}} \int_0^\infty f^2(x)\, db(x) - \int_0^\infty \left[f'(x)^2 + xf^2(x)\right]\, dx \right\}$$

where the supremum is taken over the space of functions satisfying $f(0) = 0$ and $\int_0^\infty [(f')^2 + (1+x)f^2\, dx < \infty$ with norm one.

This last expression is a characterization of F_β independent of RMT.

ASEP

The *Asymmetric Simple Exclusion Process* (ASEP) is a stochastic interacting particle system on the lattice $\mathbb{Z}$ where particles hop one site to the right (left) with probability p ($q = 1 - p$) subject to the exclusion rule.

ASEP

The *Asymmetric Simple Exclusion Process* (ASEP) is a stochastic interacting particle system on the lattice $\mathbb{Z}$ where particles hop one site to the right (left) with probability p ($q = 1 - p$) subject to the exclusion rule.

ASEP is *not* a determinantal process but nevertheless we have, for step initial condition,

ASEP

The *Asymmetric Simple Exclusion Process* (ASEP) is a stochastic interacting particle system on the lattice $\mathbb{Z}$ where particles hop one site to the right (left) with probability p ($q = 1 - p$) subject to the exclusion rule.

ASEP is *not* a determinantal process but nevertheless we have, for step initial condition,

Theorem (TW) Let $m = [\sigma t]$, $\gamma = q - p > 0$ fixed, then

$$\lim_{t\to\infty} \mathbb{P}_{\mathbb{Z}^+}\left(x_m(t/\gamma) \le c_1(\sigma)t + c_2(\sigma)\, s\, t^{1/3}\right) = F_2(s)$$

uniformly for σ in compact subsets of $(0,1)$ where $c_1(\sigma) = -1 + 2\sqrt{\sigma}$, $c_2(\sigma) = \sigma^{-1/6}(1 - \sqrt{\sigma})^{2/3}$.

Here $x_m(t)$ is the position of the mth particle from the left.

Role of F_β in Data Analysis

The greatest root distribution is found everywhere in classical multivariate analysis. It describes the null hypothesis distribution for the union intersection test for any number of classical problems, including multiple response linear regression, MANOVA, canonical correlations, equality of covariance matrices and so on. However, the exact null distribution is difficult to calculate and work with, and so the use of extensive tables or special purpose software has always been necessary.

Role of F_β in Data Analysis

The greatest root distribution is found everywhere in classical multivariate analysis. It describes the null hypothesis distribution for the union intersection test for any number of classical problems, including multiple response linear regression, MANOVA, canonical correlations, equality of covariance matrices and so on. However, the exact null distribution is difficult to calculate and work with, and so the use of extensive tables or special purpose software has always been necessary.

Our main claim is that for many applied purposes, the Tracy–Widom approximation can often, if not quite always, substitute for the elaborate tables and computational procedures that have until now been needed.

Iain Johnstone, 2009

Review of multivariate statistical analysis

Let $x_1, x_2, \ldots, x_n$ be a random sample from $N_p(\mu, \Sigma)$. Form $n \times p$ data matrix X

$$X = \begin{pmatrix} \longleftarrow & x_1 & \longrightarrow \\ \longleftarrow & x_2 & \longrightarrow \\ & \vdots & \\ \longleftarrow & x_n & \longrightarrow \end{pmatrix}$$

then the $p \times p$ matrix

$$A = X^T X \sim W_p(n, \Sigma) = \text{Wishart distr.}$$

For $p = 1$, A is distributed as $\sigma^2 \chi^2_{(n)}$.

In the univariate case if X and Y are independent and χ^2-distributed with m and n degrees of freedom, respectively, then $X/(X+Y)$ has the beta distribution

$$\frac{1}{B(m/2, n/2)} t^{m/2-1}(1-t)^{-(m+n)/2}$$

Multivariate case: if $A \sim W_p(m, \Sigma)$ and $B \sim W_p(n, \Sigma)$, independent of A, then the matrix analogue of the beta distribution is

$$(A+B)^{-1}B$$

In the univariate case if X and Y are independent and χ^2-distributed with m and n degrees of freedom, respectively, then $X/(X+Y)$ has the beta distribution

$$\frac{1}{B(m/2, n/2)} t^{m/2-1}(1-t)^{-(m+n)/2}$$

Multivariate case: if $A \sim W_p(m, \Sigma)$ and $B \sim W_p(n, \Sigma)$, independent of A, then the matrix analogue of the beta distribution is

$$(A+B)^{-1}B$$

Definition: Let A and B be as above, $m \geq p$. The largest eigenvalue θ_1 of $(A+B)^{-1}B$ is called the *greatest root statistic*; and we denote a random variable having this distribution by $\theta_1(p, m, n)$.

In the univariate case if X and Y are independent and χ^2-distributed with m and n degrees of freedom, respectively, then $X/(X+Y)$ has the beta distribution

$$\frac{1}{B(m/2, n/2)} t^{m/2-1}(1-t)^{-(m+n)/2}$$

Multivariate case: if $A \sim W_p(m, \Sigma)$ and $B \sim W_p(n, \Sigma)$, independent of A, then the matrix analogue of the beta distribution is

$$(A+B)^{-1}B$$

Definition: Let A and B be as above, $m \geq p$. The largest eigenvalue θ_1 of $(A+B)^{-1}B$ is called the *greatest root statistic*; and we denote a random variable having this distribution by $\theta_1(p, m, n)$.

The eigenvalues of $(A+B)^{-1}B$ have density (Jacobi ensemble)

$$C \prod_{i=1}^{\min(n,p)} \theta_i^{(|n-p|-1)/2}(1-\theta_i)^{(m-p-1)/2} \prod_{i \neq j} |\theta_i - \theta_j|.$$

Let

$$W(p,m,n) := \log\left(\frac{\theta(p,m,n)}{1-\theta(p,m,n)}\right)$$

Theorem (Johnstone): Assume

$$\lim_{p\to\infty}\frac{\min(p,n)}{m+n} > 0, \quad \lim_{p\to\infty}\frac{p}{m} < 1$$

then

$$\frac{W(p,m,n)-\mu(p,m,n)}{\sigma(p,m,n)} \overset{\mathcal{D}}{\Rightarrow} F_1$$

where

$$\begin{aligned}
\mu(p,m,n) &= 2\log\tan(\frac{\varphi+\gamma}{2}) \\
\sigma^3(p,m,n) &= \frac{16}{(m+n-1)^2}\frac{1}{\sin^2(\varphi+\gamma)\sin\varphi\sin\gamma} \\
\sin^2(\gamma/2) &= \frac{\min(p,n)-1/2}{m+n-1} \\
\sin^2(\varphi/2) &= \frac{\max(p,n)-1/2}{m+n-1}
\end{aligned}$$

Note: "-1/2" and "-1" make convergence $\mathrm{O}(p^{-2/3})$.

Let f_α denote the α-percentile for $X \sim F_1$,

$$F_1(f_\alpha) = \alpha$$

and similarly θ_α the α-percentile of $\theta(p, m, n)$, then corollary of above is

$$\theta_\alpha \approx \frac{\exp(\mu + f_\alpha \sigma)}{1 + \exp(\mu + f_\alpha \sigma)}$$

Tables use variables

$$s = \min(n, p), \;\; \mathfrak{m} = (|n - p| - 1)/2, \;\; \mathfrak{n} = (m - p - 1)$$

From Iain Johnstone

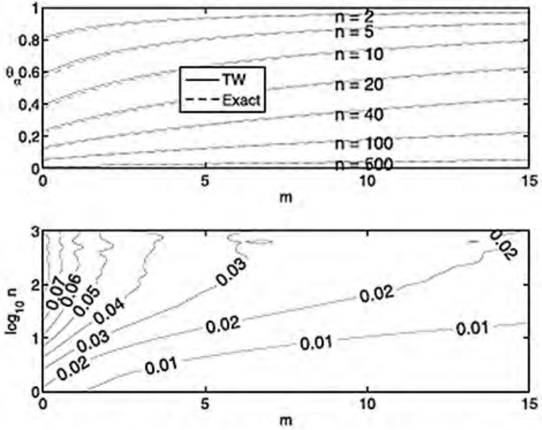


FIG. 2. *Comparison of exact and approximate 95th percentiles for* s $=2$. *Top panel: solid line is the Tracy–Widom approximation* $\theta_\alpha^{TW}(2, \mathsf{m}, \mathsf{n})$ *plotted as a function of* m *for values of* n *shown. Dashed lines are the exact percentiles* $\theta_\alpha(2, \mathsf{m}, \mathsf{n})$ *from Chen's tables. Bottom panel: Contour plots of relative error* $r = (\theta_\alpha^{TW}/\theta_\alpha) - 1$. *Horizontal axis is* m, *vertical axis is* $\log_{10}$ n, *thus covering the range from* n $= 1$ *to* 1000.

From Iain Johnstone

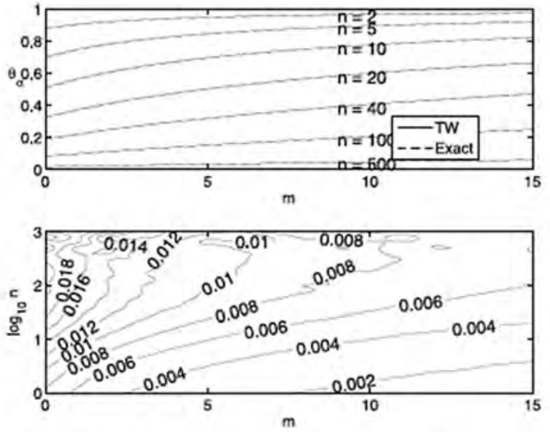


FIG. 3. *Comparison of exact and approximate 90th percentiles for* $s = 4$. *Top panel: solid line is the Tracy–Widom approximation* $\theta_\alpha^{TW}(4, m, n)$ *plotted as a function of* m *for values of* n *shown. Dashed lines are the exact percentiles* $\theta_\alpha(4, m, n)$ *from Chen's tables. Bottom panel: Contour plots of relative error* $r = (\theta_\alpha^{TW}/\theta_\alpha) - 1$. *Horizontal axis is* m, *vertical axis is* $\log_{10} n$, *thus covering the range from* $n = 1$ *to* 1000.

A Short Bibliography

- Books and review articles

1. *An Introduction to Random Matrices*, G. Anderson, A. Guionnet, O. Zeitouni, Cambridge University Press, 2010.
2. *Random Matrices: High Dimensional Phenomena*, G. Blower, Cambridge University Press, 2009.
3. *Random Matrix Theory: Invariant Ensembles and Universality*, P. Deift and D. Gioev, Amer. Math. Soc. 2009.
4. *Planar Ising Correlations*, J. Palmer, Birkhäuser, 2007.
5. *High dimensional statistical inference and random matrices*, I. Johnstone, Proceedings of the ICM, 2006; arXiv:math/0611589.
6. *Distribution functions for largest eigenvalues and their applications*, C. A. Tracy and H. Widom, Proceedings of the ICM, 2002; arXiv:math-ph/0210255.
7. *Two ways to solve ASEP*, I. Corwin, in *Topics in Percolative and Disordered Systems*, Springer, 2014; arXiv:1212.2267.

- Research papers
    1. *Universality at the edge of the spectrum in Wigner random matrices*, A. Soshnikov, Commun. Math. Phys. **207** (1999), 697–733.
    2. *Asymptotics of Tracy-Widom distributions and the total integral of a Painlevé II function*, J. Baik, R. Buckingham and J. DiFranco, Commun. Math. Phys. **280** (2008), 463–497; arXiv:0704.3636.
    3. *Higher order analogues of the Tracy-Widom distribution and the Painlevé II hierarchy*, T. Claeys, I. Krasovsky, A. Its, Commun. on Pure and Applied Math. **63** (2010), 362–412; arXiv:0901.2473.
    4. *Universal structure and universal equations (PDE) for unitary ensembles*, I. Rumanov, J. Math. Phys. **51** (2010), 083512.
    5. *Beta ensembles, stochastic Airy spectrum, and a diffusion*, J. Ramírez, B. Rider and B. Virág, J. Amer. Math. Soc. **24** (2011), 919–944; arXiv:math/0607331.
    6. *Edge universality of beta ensembles*, P. Bourgade, L. Erdös, H.-T. Yau, arXiv:1306.5728.
    7. *Tracy-Widom at high temperature*, R. Allez and L. Dumaz, arXiv:1312.1283
    8. *Approximate null distribution of the largest root in multivariate analysis*, I. Johnstone, Annals of Applied Statistics **3** (2009), 1616–1633; arXiv:1009.5854.
    9. Work by Tracy & Widom, see *arXiv.org/a/tracy_c_1*