Asymptotic expansions for GUE and Wishart random matrices

12. april 2011

ESI, Vienna

Uffe Haagerup, University of Copenhagen
Steen Thorbjørnsen, University of Aarhus

Stein's method (rough idea)

Let μ_0 be a probability measure on $\mathbb{R}$.

Part I: Find a class of functions $\mathcal{C} \subseteq \mathcal{M}_b(\mathbb{R})$ and an operator $A\colon \mathcal{C} \to \mathcal{M}_b(\mathbb{R})$ such that

$$\int_{\mathbb{R}} Af(t)\,\mu(\mathrm{d}t) = 0 \text{ for all } f \text{ in } \mathcal{C} \iff \mu = \mu_0.$$

or in terms of random variables:

$$\mathbb{E}\{Af(X)\} = 0 \text{ for all } f \text{ in } \mathcal{C} \iff X \sim \mu_0.$$

Part II: Establish that

$$\lim_{n\to\infty} \mathbb{E}\{Af(X_n)\} = 0 \text{ for all } f \text{ in } \mathcal{C} \implies X_n \to \mu_0, \quad \text{as } n \to \infty.$$

Stein's Method for the Gaussian distribution

Note that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{-t^2/2} = -t\mathrm{e}^{-t^2/2}.$$

Hence if f in $C_b^1(\mathbb{R})$, we have by partial integration:

$$\begin{aligned} 0 &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{-t^2/2} + t\mathrm{e}^{-t^2/2} \right) f(t)\,\mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (f'(t) - tf(t))\mathrm{e}^{-t^2/2}\,\mathrm{d}t \\ &= \mathbb{E}\{f'(X) - Xf(X)\}, \end{aligned}$$

if $X \sim N(0,1)$.

Stein's Method for the Gaussian distribution (con't)

Assume conversely that

$$\mathbb{E}\{f'(X) - Xf(X)\} = 0 \text{ for all } f \text{ in } C_b^1(\mathbb{R}).$$

Putting $f(t) = \mathrm{e}^{\mathrm{i}ut}$, it then follows that

$$\phi_X'(u) + u\phi_X(u) = 0, \text{ for all } u \text{ in } \mathbb{R}.$$

This implies that $X \sim N(0,1)$.

Stein's Method for the semi-circle distribution

Consider the SCD: $h(t)\,\mathrm{d}t$, where $h(t) = \frac{1}{2\pi}\sqrt{4-t^2}1_{[-2,2]}(t)\,\mathrm{d}t$.

Note that for t in $(-2,2)$,

$$h'(t) = -\frac{1}{2\pi}\frac{t}{\sqrt{4-t^2}},$$

so that

$$(4-t^2)h'(t) + th(t) = \frac{1}{2\pi}\big(-t\sqrt{4-t^2} + t\sqrt{4-t^2}\big) = 0.$$

If $f \in C_b^1(\mathbb{R})$ and $X \sim h(t)\,\mathrm{d}t$ we thus have that

$$0 = \frac{1}{2\pi}\int_{-2}^{2}\big((4-t^2)h'(t) + th(t)\big)f(t)\,\mathrm{d}t$$

$$= \frac{1}{2\pi}\int_{-2}^{2}\big((t^2-4)f'(t) + 3tf(t)\big)h(t)\,\mathrm{d}t$$

$$= \mathbb{E}\{(X^2-4)f'(X) + 3Xf(X)\}.$$

Solving the Stein diff. equation for the SCD

For any C^∞-function $g\colon \mathbb{R} \to \mathbb{C}$, the differential equation:

$$(t^2-4)f'(t)+3tf(t)=g(t)-\frac{1}{2\pi}\int_{-2}^{2} g(t)\sqrt{4-t^2}\,\mathrm{d}t, \quad (t\in\mathbb{R}), \tag{1}$$

has a unique C^∞-solution on $\mathbb{R}$.

If $g \in C_b^\infty(\mathbb{R})$ then so is f.

Stein's equation for the SCD

For a random variable X note now that

$$\mathbb{E}\{(X^2-4)f'(X)+3Xf(X)\}=0 \ \forall f\in C_b^\infty(\mathbb{R})$$

$$\iff \mathbb{E}\{g(X)-\tfrac{1}{2\pi}\textstyle\int_{-2}^{2} g(t)\sqrt{4-t^2}\,\mathrm{d}t\}=0 \quad \forall g\in C_b^\infty(\mathbb{R})$$

$$\iff \mathbb{E}\{g(X)\}=\frac{1}{2\pi}\int_{-2}^{2} g(t)\sqrt{4-t^2}\,\mathrm{d}t \quad \forall g\in C_b^\infty(\mathbb{R})$$

$$\iff X\sim h(t)\,\mathrm{d}t.$$

Weak convergence to the SCD

For a sequence (X_n) of random variables X_n we note similarly that

$$\lim_{n\to\infty} \mathbb{E}\{(X_n^2 - 4)f'(X_n) + 3Xf(X_n)\} = 0 \quad \forall f \in C_b^\infty(\mathbb{R})$$

$$\iff \lim_{n\to\infty} \mathbb{E}\{g(X_n) - \tfrac{1}{2\pi}\textstyle\int_{-2}^{2} g(t)\sqrt{4-t^2}\,\mathrm{d}t\} = 0 \quad \forall g \in C_b^\infty(\mathbb{R})$$

$$\iff \lim_{n\to\infty} \mathbb{E}\{g(X_n)\} = \frac{1}{2\pi}\int_{-2}^{2} g(t)\sqrt{4-t^2}\,\mathrm{d}t \quad \forall g \in C_b^\infty(\mathbb{R})$$

$$\iff X_n \xrightarrow{\mathrm{w}} h(t)\,\mathrm{d}t.$$

The spectral distribution of a selfadjoint random matrix

Let T be a selfadjoint random $n \times n$ matrix with eigenvalues

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

Then the spectral distribution of T is defined at any Borel set B by

$$L_T(B) = \frac{1}{n} \int_\Omega \Big(\sum_{j=1}^n \delta_{\lambda_j(\omega)}(B) \Big) \, P(\mathrm{d}\omega)$$

$$= \frac{1}{n} \sum_{j=1}^n \int_\Omega 1_B(\lambda_j(\omega)) \, P(\mathrm{d}\omega).$$

By an extension argument it follows that for any bounded Borel function f,

$$\int_{\mathbb{R}} f(t) \, L_T(\mathrm{d}t) = \mathbb{E}\big[\mathrm{tr}_n(f(T))\big].$$

Stein's method for random matrices

Let (T_n) be a sequence of selfadjoint random matrices, such that

$$\lim_{n\to\infty} \mathbb{E}\{\mathrm{tr}_n((T_n^2-4)f'(T_n)+3T_nf(T_n)\}=0, \quad \forall f\in C_b^\infty(\mathbb{R}).$$

Then

$$L_{T_n} \xrightarrow{\mathrm{w}} \frac{1}{2\pi}\sqrt{4-t^2}1_{[-2,2]}(t)\,\mathrm{d}t.$$

In other words

$$\lim_{n\to\infty} \mathbb{E}\{\mathrm{tr}_n(f(T_n))\} = \frac{1}{2\pi}\int_{-2}^{2} f(t)\sqrt{4-t^2}1_{[-2,2]}(t)\,\mathrm{d}t \quad \forall f\in C_b(\mathbb{R}).$$

The Gaussian Unitary Ensemble (GUE)

Definition. By $\mathrm{GUE}(n, \sigma^2)$ we denote the class of $n \times n$ random matrices W satisfying that

- $\forall i \geq j$: $w_{ij} = \overline{w_{ji}}$.
- the random variables w_{ij}, $1 \leq i \leq j \leq n$ are independent.
- $\forall i < j$: $\mathrm{Re}(w_{ij}), \mathrm{Im}(w_{ij}) \sim \mathrm{i.i.d.}\ N(0, \frac{1}{2}\sigma^2)$.
- $\forall i$: $w_{ii} \sim N(0, \sigma^2)$.

The spectral distribution of $\mathrm{GUE}(n, \frac{1}{n})$

Let W_n be a random matrix in $\mathrm{GUE}(n, \frac{1}{n})$.

Then $L_{W_n}(\mathrm{d}x) = h_n(x)\,\mathrm{d}x$, where

- $h_n(x) = \dfrac{1}{\sqrt{2n}} \displaystyle\sum_{j=0}^{n-1} \varphi_j\big(\sqrt{\tfrac{n}{2}}x\big)^2.$

- $\varphi_0, \varphi_1, \varphi_2, \dots$ are the Hermite functions:

$$\varphi_k(x) = \tfrac{1}{(2^k k! \sqrt{\pi})^{1/2}} H_k(x) \exp(-\tfrac{x^2}{2}), \qquad (k \in \mathbb{N}_0),$$

- $H_0, H_1, H_2, \dots,$ are the Hermite polynomials:

$$H_k(x) = (-1)^k \exp(x^2) \cdot \Big(\frac{d^k}{dx^k} \exp(-x^2)\Big).$$

Stein's method for SCD in the GUE case

The spectral density h_n for the $\mathrm{GUE}(n, \frac{1}{n})$ satisfies the third order differential equation:

$$\frac{1}{n^2} h_n'''(t) + (4 - t^2) h_n'(t) + t h_n(t) = 0, (t \in \mathbb{R}).$$

Hence for any function f in $C_b^\infty(\mathbb{R})$ we find that

$$\begin{aligned} 0 &= \int_{\mathbb{R}} f(t) \big[n^{-2} h_n'''(t) + (4 - t^2) h_n'(t) + t h_n(t) \big] \, \mathrm{d}t \\ &= \int_{\mathbb{R}} \big[-n^{-2} f'''(t) - (4 - t^2) f'(t) + 3 t f(t) \big] h_n(t) \, \mathrm{d}t. \end{aligned}$$

Thus, if $W_n \sim \mathrm{GUE}(n, \frac{1}{n})$ we have that

$$\mathbb{E}\big\{ \mathrm{tr}_n \big[(W_n^2 - 4) f'(W_n) + 3 W_n f(W_n) \big] \big\} = \frac{1}{n^2} \mathbb{E}\big\{ \mathrm{tr}_n \big[f'''(W_n) \big] \big\}.$$

Proof of the differential equation for h_n (sketch)

- If Φ denotes the confluent hypergeometric function, we have

$$\psi(z) := \int \mathrm{e}^{zx} h_n(x)\,\mathrm{d}x = \exp(\tfrac{z^2}{2n}) \cdot \Phi(1-n, 2; -\tfrac{z^2}{n}).$$

- The function $w(x) = \Phi(a, b; x)$ satisfies the diff. eq.

$$x\frac{d^2 w}{dx^2} + (b-x)\frac{dw}{dx} - aw = 0.$$

- Setting $z = -\mathrm{i}y$ for y in $\mathbb{R}$, it follows that

$$n^2 \mathrm{i}y \widehat{h}_n''(y) + 3n^2 \mathrm{i}\widehat{h}_n'(y) + (4n^2 \mathrm{i}y - \mathrm{i}y^3)\widehat{h}_n(y) = 0.$$

- Fourier inversion!

Asymptotic expansion for the GUE

Let g be a function from $C_b^\infty(\mathbb{R})$, and let f be the solution to

$$(t^2-4)f'(t)+3tf(t)=g(t)-\frac{1}{2\pi}\int_{-2}^{2}g(t)\sqrt{4-t^2}\,\mathrm{d}t,\quad (t\in\mathbb{R}).$$

Then,

$$\begin{aligned}\mathbb{E}\big\{\mathrm{tr}_n\big[g(W_n)-\tfrac{1}{2\pi}\textstyle\int_{-2}^{2}g(t)\sqrt{4-t^2}\,\mathrm{d}t\big]\big\}\\ &=\mathbb{E}\big\{\mathrm{tr}_n\big[(W_n^2-4)f'(W_n)+3W_nf(W_n)\big]\big\}\\ &=\frac{1}{n^2}\mathbb{E}\big\{\mathrm{tr}_n\big[f'''(W_n)\big]\big\}.\end{aligned}$$

In other words,

$$\mathbb{E}\big\{\mathrm{tr}_n\big[g(W_n)\big]\big\}=\frac{1}{2\pi}\int_{-2}^{2}g(t)\sqrt{4-t^2}\,\mathrm{d}t+\frac{1}{n^2}\mathbb{E}\big\{\mathrm{tr}_n\big[f'''(W_n)\big]\big\}.$$

Asymptotic expansion for GUE random matrices

For any function g in $C_b^\infty(\mathbb{R})$ there exists a (unique) sequence $(\alpha_j(g))_{j\in\mathbb{N}}$ of complex numbers, such that for any k in $\mathbb{N}$

$$\mathbb{E}\{\mathrm{tr}_n(g(X_n))\} = \int_{\mathbb{R}} g(x)h_n(x)\,\mathrm{d}x =$$

$$\frac{1}{2\pi}\int_{-2}^{2} g(x)\sqrt{4-x^2}\,\mathrm{d}x + \frac{\alpha_1(g)}{n^2} + \frac{\alpha_2(g)}{n^4} + \cdots + \frac{\alpha_k(g)}{n^{2k}} + O(n^{-2k-2}).$$

More precisely

$$\alpha_j(g) = \frac{1}{2\pi}\int_{-2}^{2}[T^j g](x)\sqrt{4-x^2}\,\mathrm{d}x,$$

and

$$O(n^{-2k-2}) = \frac{1}{n^{2k+2}}\int_{\mathbb{R}}\left[T^{k+1}g\right](x)h_n(x)\,\mathrm{d}x.$$

Here $T\colon C_b^\infty(\mathbb{R}) \to C_b^\infty(\mathbb{R})$ is the linear operator given by $T(g) = f'''$.

Asymptotic expansion for the GUE (con't)

One may verify that

$$\alpha_j(g) = \frac{1}{\pi} \sum_{k=2j}^{3j-1} C_{j,k} \frac{k!}{(2k)!} \int_{-2}^{2} g^{(k)}(x) \frac{T_k(\frac{x}{2})}{\sqrt{4-x^2}} \, \mathrm{d}x,$$

where $C_{j,2j}, C_{j,2j+1}, \dots, C_{j,3j-1}$ are constants.

Moreover $T_0, T_1, T_2, \dots$ are the Chebychev polynomials given by

$$T_k(\cos\theta) = \cos(k\theta), \qquad (\theta \in [0,\pi], \ k \in \mathbb{N}_0).$$

A diff. equation for the Cauchy transform of h_n

The Cauchy transform of a $\mathrm{GUE}(n, \frac{1}{n})$ random matrix X_n is given by

$$G_n(\lambda) = \mathbb{E}\{\mathrm{tr}_n((\lambda \mathbf{1}_n - X_n)^{-1})\} = \int_{\mathbb{R}} \frac{1}{\lambda - x} h_n(x)\,\mathrm{d}x, \qquad (\lambda \in \mathbb{C}\setminus\mathbb{R}).$$

Set $$f_\lambda(x) = \frac{1}{\lambda - x}, \qquad (x \in \mathbb{R},\ \lambda \in \mathbb{C} \setminus \mathbb{R}).$$

It follows then as before by partial integration that

$$\begin{aligned} 0 &= \int_{\mathbb{R}} f_\lambda(x)\big(n^{-2} h_n'''(x) + (4 - x^2) h_n'(x) + x h_n(x)\big)\,\mathrm{d}x \\ &= \int_{\mathbb{R}} \big[- n^{-2} f_\lambda'''(x) - (4 - x^2) f_\lambda'(x) + 3x f_\lambda(x)\big] h_n(x)\,\mathrm{d}x \\ &= n^{-2} G_n'''(\lambda) + (4 - \lambda^2) G_n'(\lambda) + \lambda G_n(\lambda) - 2. \end{aligned}$$

Asymptotic expansion for G_n

Consider the asymptotic expansion:

$$G_n(\lambda) = \int_{\mathbb{R}} \frac{1}{\lambda - x} h_n(x)\,\mathrm{d}x$$

$$= H_0(\lambda) + \frac{H_1(\lambda)}{n^2} + \cdots + \frac{H_k(\lambda)}{n^{2k}} + O(n^{-2k-2}).$$

Insert this into the differential equation:

$$n^{-2} G_n'''(\lambda) + (4 - \lambda^2) G_n'(\lambda) + \lambda G_n(\lambda) = 2.$$

Then we obtain a sequence of differential equations:

$$(\lambda^2 - 4) H_k'(\lambda) - \lambda H_k(\lambda) = H_{k-1}'''(\lambda), \qquad (k \in \mathbb{N}).$$

These can be solved successively using that

$$H_0(\lambda) = \frac{1}{2\pi} \int_{-2}^{2} \frac{1}{\lambda - x} \sqrt{4 - x^2}\,\mathrm{d}x = \frac{\lambda}{2} - \frac{1}{2}(\lambda^2 - 4)^{1/2}.$$

Explicit formulae

Using the technique described above we obtain

$$H_1(\lambda) = \frac{1}{(\lambda^2 - 4)^{5/2}},$$
$$H_2(\lambda) = \frac{21(\lambda^2 + 1)}{(\lambda^2 - 4)^{11/2}},$$
$$H_3(\lambda) = \frac{1738 + 6138\lambda + 1485\lambda^2}{(\lambda^2 - 4)^{17/2}}.$$

Explicit formulae

Generally we have that

$$H_j(\lambda) = \sum_{r=2j}^{3j-1} C_{j,r}(\lambda^2 - 4)^{-r-1/2}.$$

Here $C_{j,r}$, $2j \le r \le 3j-1$, are constants satisfying the recursion formulae

$$C_{j+1,r} = \frac{(2r-3)(2r-1)}{r+1}\big((r-1)C_{j,r-2} + (4r-10)C_{j,r-3}\big).$$

A formula of Pastur and Scherbina

Let X_n be a random matrix in $\mathrm{GUE}(n, \frac{1}{n})$.

Then for any functions f, g in $C_b^\infty(\mathbb{R})$ we have

$$\mathrm{Cov}\big\{\mathrm{Tr}_n(f(X_n)), \mathrm{Tr}_n(g(X_n))\big\} = \int_{\mathbb{R}^2} \Big(\frac{f(x) - f(y)}{x - y}\Big)\Big(\frac{g(x) - g(y)}{x - y}\Big)\rho_n(x, y)\,\mathrm{d}x\,\mathrm{d}y.$$

Here,

$$\rho_n(x, y) = \frac{n}{4}\Big(\varphi_n\big(\sqrt{\tfrac{n}{2}}x\big)\varphi_{n-1}\big(\sqrt{\tfrac{n}{2}}y\big) - \varphi_{n-1}\big(\sqrt{\tfrac{n}{2}}x\big)\varphi_n\big(\sqrt{\tfrac{n}{2}}y\big)\Big).$$

Weak convergence of $\rho_n(x,y)\,\lambda(\mathrm{d}x,\mathrm{d}y)$

Proposition. We have the following weak convergence:

$$\rho_n(x,y)\,\lambda_2(\mathrm{d}x,\mathrm{d}y) \xrightarrow{\mathrm{w}} \rho(x,y)\,\lambda_2(\mathrm{d}x,\mathrm{d}y) \quad \text{as } n \to \infty.$$

Here,

$$\rho(x,y) = \frac{1}{4\pi^2}\frac{4-xy}{\sqrt{4-x^2}\sqrt{4-y^2}}1_{[-2,2]}(x)1_{[-2,2]}(y),$$

In particular, if $X_n \sim \mathrm{GUE}(n,\frac{1}{n})$ for all n, then

$$\lim_{n\to\infty} \mathrm{Cov}\{\mathrm{Tr}_n[f(X_n)],\mathrm{Tr}_n[g(X_n)]\}$$
$$= \int_{\mathbb{R}^2}\Big(\frac{f(x)-f(y)}{x-y}\Big)\Big(\frac{g(x)-g(y)}{x-y}\Big)\rho(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

for all suitably "nice" functions f, g

ρ_n in terms of h_n

Proposition. For any n in $\mathbb{N}$ we have that

$$\rho_n(x,y) = \frac{1}{4}\Big(\tilde{h}_n(x)\tilde{h}_n(y) - 4h_n'(x)h_n'(y) + \frac{1}{n^2}h_n''(x)h_n''(y)\Big),$$

where

$$\tilde{h}_n(x) = h_n(x) - xh_n'(x).$$

Consequently, for any function f in $C_b^\infty(\mathbb{R}^2)$,

$$\int_{\mathbb{R}^2} f(x,y)\rho_n(x,y)\,\mathrm{d}x\,\mathrm{d}y =$$

$$\frac{1}{4}\int_{\mathbb{R}^2} f(x,y)\tilde{h}_n(x)\tilde{h}_n(y)\,\mathrm{d}x\,\mathrm{d}y - \int_{\mathbb{R}^2} f(x,y)h_n'(x)h_n'(y)\,\mathrm{d}x\,\mathrm{d}y - \frac{1}{4n^2}\int_{\mathbb{R}^2} f(x,y)h_n''(x)h_n''(y)\,\mathrm{d}x\,\mathrm{d}y.$$

By Fubinis Theorem and partial integration we get that

$$\int_{\mathbb{R}^2} f(x,y) h_n'(x) h_n'(y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^2} \frac{\partial^2}{\partial x \partial y} f(x,y) h_n(x) h_n(y)\,\mathrm{d}x\,\mathrm{d}y,$$

$$\int_{\mathbb{R}^2} f(x,y) h_n''(x) h_n''(y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^2} \frac{\partial^4}{\partial x^2 \partial y^2} f(x,y) h_n(x) h_n(y)\,\mathrm{d}x\,\mathrm{d}y,$$

and

$$\int_{\mathbb{R}^2} f(x,y) \tilde{h}_n(x) \tilde{h}_n(y)\,\mathrm{d}x\,\mathrm{d}y =$$

$$\int_{\mathbb{R}^2} \left[4f(x,y) + 2x\tfrac{\partial}{\partial x} f(x,y) + 2y\tfrac{\partial}{\partial y} f(x,y) + xy\tfrac{\partial^2}{\partial x \partial y} f(x,y)\right] \cdot h_n(x) h_n(y)\,\mathrm{d}x\,\mathrm{d}y.$$

Asymptotic expansion for $h_n(x)h_n(y)\,\mathrm{d}x\,\mathrm{d}y$.

Proposition. For any function f in $C_b^\infty(\mathbb{R}^2)$ there exists a sequence $(\beta_n(f))_{n\in\mathbb{N}_0}$ of complex numbers such that

$$\int_{\mathbb{R}^2} f(x,y)h_n(x)h_n(y)\,\mathrm{d}x\,\mathrm{d}y = \sum_{j=0}^{k} \frac{\beta_j(f)}{n^{2j}} + O(n^{-2k-2})$$

for any k in $\mathbb{N}_0$.

Sketch of Proof.

For fixed x in $\mathbb{R}$, $f(x,\cdot) \in C_b^\infty(\mathbb{R})$ and hence

$$\int_{\mathbb{R}} f(x,y)h_n(y)\,\mathrm{d}y = \sum_{j=1}^{k} \frac{\eta_j(x)}{n^{2j}} + \frac{1}{n^{2k+2}} R_k(x),$$

where

$$\eta_j(x) = \frac{1}{2\pi}\int_{-2}^{2} \left[T^j f(x,\cdot)\right](t)\sqrt{4-t^2}\,\mathrm{d}t, \qquad (x \in \mathbb{R},\ j = 1,\dots,k),$$

and

$$R_k(x) = \int_{\mathbb{R}} \left[T^{k+1} f(x,\cdot)\right](t) h_n(t)\,\mathrm{d}t, \qquad (x \in \mathbb{R}).$$

By Fubinis Theorem,

$$\int_{\mathbb{R}^2} f(x,y)h_n(x)h_n(y)\,\mathrm{d}x\,\mathrm{d}y = \sum_{j=1}^{k} \int_{\mathbb{R}} \frac{\eta_j(x)h_n(x)}{n^{2j}}\,\mathrm{d}x + \int_{\mathbb{R}} \frac{R_k(x)h_n(x)}{n^{2k+2}}\,\mathrm{d}x.$$

Sketch of Proof (continued)

Using continuity properties of T it follows that

$$C_k^f := \sup_{x\in\mathbb{R}} |R_k(x)| < \infty, \quad \text{and that} \quad \eta_j \in C_b^\infty(\mathbb{R}).$$

Therefore,

$$\int_{\mathbb{R}} \frac{R_k(x)h_n(x)}{n^{2k+2}}\,\mathrm{d}x = O(n^{-2k-2}).$$

In addition, for each j we have an expansion:

$$\int_{\mathbb{R}} \eta_j(x)h_n(x)\,\mathrm{d}x = \sum_{l=0}^{k-j} \frac{\xi_l^j(f)}{n^{2l}} + O(n^{-2k+2j-2}),$$

for suitable complex numbers $\xi_0^j(f), \ldots, \xi_{k-j}^j(f)$.

Sketch of Proof (continued)

Altogether,

$$\int_{\mathbb{R}^2} f(x,y) h_n(x) h_n(y) \, \mathrm{d}x \, \mathrm{d}y$$

$$= \sum_{j=1}^{k} \int_{\mathbb{R}} \frac{\eta_j(x) h_n(x)}{n^{2j}} \, \mathrm{d}x + \int_{\mathbb{R}} \frac{R_k(x) h_n(x)}{n^{2k+2}} \, \mathrm{d}x$$

$$= \sum_{j=0}^{k} \left(\sum_{l=0}^{k-j} \frac{\xi_l^j(f)}{n^{2(l+j)}} + O(n^{-2k-2}) \right) + O(n^{-2k-2})$$

$$= \sum_{r=0}^{k} n^{-2r} \left(\sum_{j=0}^{r} \xi_{r-j}^j(f) \right) + O(n^{-2k-2}).$$

Asymptotic expansion for covariances

For any functions f, g in $C_b^\infty(\mathbb{R})$, there exists a sequence $(\beta_n(f,g))_{n\in\mathbb{N}}$ of complex numbers, such that

$$\begin{aligned}
&\mathrm{Cov}\{\mathrm{Tr}_n[f(X_n)], \mathrm{Tr}_n[g(X_n)]\} \\
&\quad = \int_{\mathbb{R}^2} \Big(\frac{f(x)-f(y)}{x-y}\Big)\Big(\frac{g(x)-g(y)}{x-y}\Big)\rho_n(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&\quad = \sum_{j=0}^{k} \frac{\beta_j(f,g)}{n^{2j}} + O(n^{-2k-2}).
\end{aligned}$$

for any k in $\mathbb{N}_0$.

Proof of Asymptotic expansion for covariances

Use the previous Theorem on the function $\Delta f(x,y)\Delta g(x,y)$, where

$$\Delta f(x,y) = \begin{cases} \frac{f(x)-f(y)}{x-y}, & \text{if } x \neq y, \\ f'(x), & \text{if } x = y. \end{cases}$$

Note that

$$\Delta f(x,y) = \int_0^1 f'(sx+(1-s)y)\,\mathrm{d}s, \qquad ((x,y)\in\mathbb{R}^2),$$

so that

$$\Delta f \in C_b^\infty(\mathbb{R}^2).$$

The two-dimensional Cauchy transform

For λ, μ in $\mathbb{C} \setminus \mathbb{R}$, $\lambda \neq \mu$, we have

$$G_n(\lambda, \mu) := \mathrm{Cov}\big\{\mathrm{Tr}_n[(\lambda \mathbf{1}_n - X_n)^{-1}], \mathrm{Tr}_n[(\mu \mathbf{1}_n - X_n)^{-1}]\big\}$$

$$= \int_{\mathbb{R}^2} (\mu - \lambda)^{-2} \Big(\frac{1}{\lambda - x} - \frac{1}{\mu - x}\Big)\Big(\frac{1}{\lambda - y} - \frac{1}{\mu - y}\Big) \rho_n(x, y)\, \mathrm{d}x\, \mathrm{d}y.$$

Recall that

$$\rho_n(x, y) = \frac{1}{4}\Big(\tilde{h}_n(x)\tilde{h}_n(y) - 4h_n'(x)h_n'(y) + \frac{1}{n^2} h_n''(x)h_n''(y)\Big),$$

where

$$\tilde{h}_n(x) = h_n(x) - x h_n'(x).$$

The two-dimensional Cauchy transform (continued)

From the above formulae we obtain that

$G_n(\lambda, \mu)$

$$= -\frac{1}{2(\lambda-\mu)^2}\Big[\tilde{G}_n(\lambda)\tilde{G}_n(\mu) - \hat{G}_n(\lambda)\hat{G}_n(\mu) + 1 - \frac{1}{n^2} G_n''(\lambda) G_n''(\mu)\Big],$$

where

$$\tilde{G}_n(\lambda) = G_n(\lambda) - \lambda G_n'(\lambda)$$

$$\hat{G}_n(\lambda) = 2G_n'(\lambda) - 1.$$

By inserting the asymptotic expansions for $G_n(\lambda)$ and $G_n(\mu)$ we then obtain

$$G_n(\lambda, \mu) = G(\lambda, \mu) + \sum_{j=1}^{k} \frac{J_j(\lambda, \mu)}{n^{2j}} + O(n^{-2k-2}).$$

General Wishart distributions

Let α and γ be real numbers in $(-1, \infty)$ and $(0, \infty)$, respectively, and let n be a positive integer.

The Wishart distribution Wish(n, α, γ) with parameters (n, α, γ) is the probability measure on $M_n(\mathbb{C})_{\mathrm{sa}}$ with density

$$S \mapsto c_1 \cdot (\det S)^{\alpha} \cdot \exp(-\gamma^{-1}\mathrm{Tr}_n(S)) \cdot 1_{M_n(\mathbb{C})_+}(S),$$

with respect to Lebesgue measure $M_n(\mathbb{C})_{\mathrm{sa}}$.

A selfadjoint random $n \times n$ matrix with distribution Wish(n, α, γ) is called a Wishart matrix with parameters (n, α, γ).

Gaussian Wishart matrices (Wishart $\sim$ 1928)

Suppose m, n are positive integers such that $m \geq n$, and put $\alpha = m - n$.

Let further $B = (b_{ij})$ be a random $m \times n$ matrix such that

$$\mathrm{Re}(b_{ij}), \mathrm{Im}(b_{ij}), \quad 1 \leq i \leq m, \; 1 \leq j \leq n,$$

are i.i.d. with distribution $N(0, \frac{1}{2})$.

Then $B^*B \sim \mathrm{Wish}(n, \alpha, 1)$.

The spectral distribution of a Wishart matrix

Let W be a Wishart matrix with parameters (n, α, γ).

Then for any bounded Borel function $f\colon [0,\infty) \to \mathbb{R}$ we have

$$\mathbb{E}\{\mathrm{Tr}_n(f(W))\} = \gamma^{-1} \int_0^\infty f(x)\sigma_{n,\alpha,\gamma}(x)\,\mathrm{d}x,$$

- $\sigma_{n,\alpha,\gamma}(x) = \sum_{j=0}^{n-1} \varphi_k^\alpha(\gamma^{-1}x)^2\,\mathrm{d}x,$
- $\varphi_k^\alpha(x) = \left[\frac{k!}{\Gamma(k+\alpha+1)} x^\alpha \exp(-x)\right]^{1/2} \cdot L_k^\alpha(x), \qquad (k \in \mathbb{N}_0).$
- $(L_k^\alpha)_{k\in\mathbb{N}_0}$ is the sequence of generalized Laguerre polynomials of order α:

$$L_k^\alpha(x) = (k!)^{-1} x^{-\alpha} \exp(x) \cdot \frac{d^k}{dx^k}\left(x^{k+\alpha}\exp(-x)\right), \qquad (k \in \mathbb{N}_0).$$

The Marchenko-Pastur Law

For each n in $\mathbb{N}$, let W_n be a Wishart matrix with parameters $(n, \alpha_n, \frac{1}{n})$, and put $c_n = 1 + \frac{\alpha_n}{n}$.

Assume that $c_n \to c \in [1, \infty)$ as $n \to \infty$.

Then for any function f in $C_b(\mathbb{R})$, we have that

$$\lim_{n\to\infty} \mathbb{E}\{\mathrm{tr}_n[f(W_n)]\} = \int_a^b f(x)\,\mu_c(x)\,\mathrm{d}x.$$

Here μ_c is the Marchenko-Pastur Law with parameter c, i.e.,

$$\mu_c(\mathrm{d}x) = \frac{\sqrt{(x-a)(b-x)}}{2\pi x} 1_{[a,b]}(x)\,\mathrm{d}x,$$

where $a = (c^{1/2} - 1)^2$ and $b = (c^{1/2} + 1)^2$.

Asymptotic expansion for Wishart matrices

Let c be a number in $[1,\infty)$, and for each n in $\mathbb{N}$, let W_n be a Wishart matrix with parameters $(n, n(c-1), \frac{1}{n})$.

For any function $f \in C^\infty_{\mathrm{pol}}(\mathbb{R})$ and for any k in $\mathbb{N}$ we have

$$\mathbb{E}\big\{\mathrm{tr}_n\big(f(W_n)\big)\big\}$$

$$= \int_0^\infty f(x)\sigma_{n,c}(x)\,\mathrm{d}x$$

$$= \int_a^b f(x)\,\mu_c(\mathrm{d}x) + \frac{\beta_1(f)}{n^2} + \frac{\beta_2(f)}{n^4} + \cdots + \frac{\beta_k(f)}{n^{2k}} + O(n^{-2k-2}),$$

for suitable constants $\beta_1(f), \ldots, \beta_k(f)$.

Outline of proof – Step I

For f in $C^{\infty}_{\mathrm{pol}}(\mathbb{R})$, put

$$\tilde{f}(x) = \begin{cases} \frac{f(x)-f(0)}{x}, & \text{if } x > 0, \\ f'(0), & \text{if } x = 0. \end{cases}$$

Then $\tilde{f} \in C^{\infty}_{\mathrm{pol}}(\mathbb{R})$, and

$$f(x) = f(0) + xf(x), \qquad (x \in \mathbb{R}).$$

Hence

$$\int_0^{\infty} f(x)\sigma_{n,c}(x)\,\mathrm{d}x = f(0) + \int_0^{\infty} f(x)x\sigma_{n,c}(x)\,\mathrm{d}x.$$

Therefore it suffices to establish asymptotic expansion for

$$\int_0^{\infty} f(x)x\sigma_{n,c}(x)\,\mathrm{d}x.$$

Note that

$$\tilde{\mu}_{n,c}(\mathrm{d}x) := c^{-1}x\sigma_{n,c}(x)\,\mathrm{d}x \xrightarrow{\mathrm{w}} (2\pi c)^{-1}\sqrt{(x-a)(b-x)}\,\mathrm{d}x =: \tilde{\mu}_c(\mathrm{d}x).$$

Outline of proof – Step II

The function $\tau_c(x) = \frac{1}{2\pi c}\sqrt{(x-a)(b-x)}$ satisfies the differential equation:

$$(x-a)(b-x)\tau_c'(x) + (x-(c+1))\tau_c(x) = 0.$$

Hence for any function f in $C^\infty_{\rm pol}(\mathbb{R})$ we get by partiel integration:

$$\int_a^b \left[(x-a)(x-b)f'(x) + 3(x-(c+1))f(x)\right]\tilde{\mu}_c(\mathrm{d}x) = 0.$$

Outline of proof – Step III

Proposition. For any function g from $C^\infty_{\mathrm{pol}}(\mathbb{R})$, there is a unique function f from $C^\infty_{\mathrm{pol}}(\mathbb{R})$ satisfying that

$$g(x) - \int_0^\infty g(x)\,\tilde{\mu}_c(\mathrm{d}x) = (x-b)(x-a)f'(x) + 3(x-(c+1))f(x).$$

Steins Method for $\tilde{\mu}_c$: For a sequence μ_n of probability measures it follows that

$$\mu_n \xrightarrow{\mathrm{w}} \tilde{\mu}_c$$

$$\Updownarrow$$

$$\forall f: \int_0^\infty \left[(x-b)(x-a)f'(x) + 3(x-(c+1))f(x)\right] \mu_n(\mathrm{d}x) \to 0.$$

Outline of proof – Step IV

Put $\tau_{n,c}(x) = x\sigma_{n,c}(x)$.

Then $\tau_{n,c}$ satisfies the differential equation

$$(x-a)(x-b)\tau'_{n,c}(x) - \big(x-(c+1)\big)\tau_{n,c}(x) = \frac{1}{n^2}\big(x^2\tau'''_{n,c}(x) + x\tau''_{n,c}(x)\big).$$

Hence by partiel integration (in the distribution sense) it follows that

$$\int_0^\infty \Big((x-b)(x-a)f'(x) + 3\big(x-(c+1)\big)f(x)\Big)\tau_{n,c}(x)\,\mathrm{d}x = \frac{1}{n^2}\int_0^\infty \Big(4f'(x) + 5xf''(x) + x^2 f'''(x)\Big)\tau_{n,c}(x)\,\mathrm{d}x.$$

Outline of proof – Step V

Moreover, for any function g in $C^{\infty}_{\mathrm{pol}}(\mathbb{R})$ we have that

$$\int_0^\infty g(x)\,\tau_{n,c}(x)\,\mathrm{d}x$$

$$= \int_a^b g(x)\,\tilde{\mu}_c(\mathrm{d}x) + \frac{1}{n^2}\int_0^\infty \Big(4f'(x) + 5xf''(x) + x^2 f'''(x)\Big)\tau_{n,c}(x)\,\mathrm{d}x,$$

where f is the solution to

$$g(x) - \int_0^\infty g(x)\,\tilde{\mu}_c(\mathrm{d}x) = (x-b)(x-a)f'(x) + 3(x-(c+1))f(x).$$

Asymptotic exp. for the Cauchy transform of σ_n

Consider for λ in $\mathbb{C} \setminus \mathbb{R}$ the Cauchy transform

$$S_{n,c}(\lambda) = \int_0^\infty \frac{1}{\lambda - x} \tau_{n,c}(x)\,\mathrm{d}x = -1 + \lambda \int_0^\infty \frac{1}{\lambda - x} \sigma_{n,c}(x)\,\mathrm{d}x.$$

Then we have the asymptotic expansion

$$S_{n,c}(\lambda) = l_0(\lambda) + \frac{l_1(\lambda)}{n^2} + \frac{l_2(\lambda)}{n^4} + \cdots + \frac{l_k(\lambda)}{n^{2k}} + O(n^{-2k-2}).$$

It follows from the differential equation for $\tau_{n,c}$ that

$$\frac{1}{n^2}\left[\lambda^2 S'''_{n,c}(\lambda) + \lambda S''_{n,c}(\lambda)\right] + (b-\lambda)(\lambda-a)S'_{n,c}(\lambda) + (\lambda-c-1)S_{n,c}(\lambda) = 2c.$$

Concrete formulae

As in the Wigner case, we get a sequence of differential equations:

$$(\lambda - b)(\lambda - a)I'_{k+1}(\lambda) - (\lambda - c - 1)I_{k+1}(\lambda) = \lambda^2 I'''_k(\lambda) + \lambda I''_k(\lambda),$$

which may be solved successively:

$$I_0(\lambda) = \frac{\lambda}{2} - \frac{c+1}{2} - \frac{1}{2}\big((\lambda - b)(\lambda - a)\big)^{1/2},$$

$$I_1(\lambda) = \frac{c\lambda^2}{(\lambda - a)^{5/2}(\lambda - b)^{5/2}},$$

$$I_2(\lambda) = \frac{c\lambda^2(1 + 6c^2 - 4c^3 + c^4 - 4c + 10(1 - c - c^2 + c^3)\lambda)}{(\lambda - a)^{11/2}(\lambda - b)^{11/2}} + \frac{c\lambda^2((-15 + 39c - 15c^2)\lambda^2 - 4(1 - c)\lambda^3)}{(\lambda - a)^{11/2}(\lambda - b)^{11/2}}.$$