Gaussian fluctuations for random matrices with correlated entries

Hermann Schulz-Baldes, Jeffrey Schenker

Erlangen, Michigan

Gaussian fluctuations = linear statistics

Ensemble $X_n = \frac{1}{\sqrt{n}}(a_n(p,q))_{p,q=1\dots n}$ hermitian
with $a_n(p,q) \in \mathbb{C}$ centered, unit variance, moment bound

If i.i.d., Wigner (1955):

$$\lim_{n\to\infty} \frac{1}{n}\,\mathbb{E}\,\mathrm{Tr}(X_n^k) = \int_{-2}^{2} \frac{dx}{2\pi}\sqrt{4-x^2}\,x^k$$

L. Arnold (1969): almost sure convergence of empirical distribution

Jonsson (1982), Khorunzhy, Khoruzhenko, Pastur (1996), Johansson, Sinai + Soshnikov, ...

$$\mathrm{Tr}\, f(X_n) - \mathbb{E}\,\mathrm{Tr}\, f(X_n) \quad \longrightarrow \quad \text{Gaussian process}$$

with f polynomial or analytic

Diagonalizing the covariance matrix

Chebyshev polynomials of the first kind

$$T_m(2\cos(\theta)) = 2\cos(m\theta)\,.$$

Johansson (1998): for i.i.d. Gaussian entries (GUE)

$$\{\mathrm{Tr}(T_m(X_n)) - \mathbb{E}\,\mathrm{Tr}(T_m(X_n))\}_{m\geq 1}$$

is an independent family of Gaussians with variance σ_m

Extended by Cabanal-Duvillard (2001): stochastic analysis

Further by Kusalik, Mingo, Speicher (2007): genus expansion

Anderson, Zeitouni (2006): i.i.d. non-Gaussian entries

Schenker, Schulz-Baldes (2007): combinatoriel proof, correlations

Definition of ensembles

For $n \in \mathbb{N}$, equivalence relation $\sim_n$ on pairs
$P = (p, q) \in [n]^2 \;=\; \{1, \dots, n\}^{\times 2}$ with $(p, q) \sim_n (q, p)$

Entries $a_n(P_1), \dots, a_n(P_j)$ independent
$\iff P_1, \dots, P_j$ in j distinct classes of $\sim_n$.

$a(P)$ centered, unit variance, moment bound

Control on size of equivalence classes: for all $\epsilon > 0$

$$\max_{p,q \in [n]^2} \#\{(p', q') \in [n]^2 \,|\, (p, q) \;\sim_n\; (p', q')\} \;=\; \mathcal{O}(n^{\epsilon}) \qquad \text{(C1)}$$

Sufficiently many independent rows: for some $\delta > 0$

$$\#\{(p, q, p') \in [n]^3 \,|\, (p, q) \;\sim_n\; (p', q) \;\&\; p \neq p'\} \;=\; \mathcal{O}(n^{1-\delta}) \qquad \text{(C2)}$$

Extending Wigner's result

Theorem *If* (C1) *and* (C2) *hold, then DOS* $\to$ *semicircle*

Proof: With sum over *consistent* pairs

$$\mathbb{E}\,\frac{1}{n}\mathrm{Tr}((X_n)^k) \;=\; \frac{1}{n^{1+\frac{k}{2}}}\sum_{P_1,\dots,P_k\in[n]^2}\mathbb{E}\,a(P_1)\cdots a(P_k)$$

Reorder sum: $P_1,\dots,P_k$ gives partition $\pi\in\mathcal{P}_k$ of $[k]$ by

$$l\sim_\pi m \qquad\Longleftrightarrow\qquad P_l\sim_n P_m$$

$$\mathbb{E}\,\frac{1}{n}\mathrm{Tr}((X_n)^k) \;=\; \frac{1}{n^{1+\frac{k}{2}}}\sum_{\pi\in\mathcal{P}_k}\sum_{(P_1,\dots,P_k)\cong\pi}\mathbb{E}\,a(P_1)\cdots a(P_k)$$

$\#\pi>\frac{k}{2}$ singleton, $\#\pi<\frac{k}{2}$ vanishes $\Longrightarrow$ pair partitions

crossings vanish $\Longrightarrow$ non-crossing pair partitions $\Longrightarrow$ Catalan $\#$

Cumulants

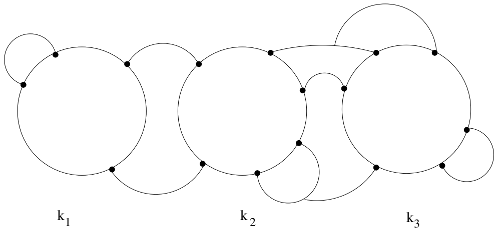
$$C_j(Y_1,\ldots,Y_j) \;=\; \sum_{\pi\in\mathcal{P}_j}(-1)^{\#\pi-1}(\#\pi-1)!\prod_{l=1}^{\#\pi}\mathbb{E}\prod_{i\in B_l}Y_i\ ,$$

where first product over blocks of partition $\pi=\{B_1,\ldots,B_{\#\pi}\}$

If all joint cumulants with $j\geq 3$ vanish, then Gaussian family

Theorem *If* (C1) *holds and* $j\geq 3$*, then*

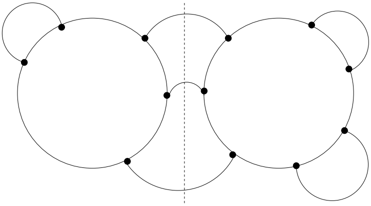
$$C_j\big(\mathrm{Tr}(X_n^{k_1}),\ldots,\mathrm{Tr}(X_n^{k_j})\big) \;=\; o(1)$$



Variances

Up to $o(1)$ errors: only non-crossing pair partitions, no twist in m connectors (dihedral group D_{2m}), $a_n(\mathbf{P}_i) = \prod_{\ell=1}^{k_i} a_n(P_{i,\ell})$

$$C_2(\mathrm{Tr} X_n^{k_1}, \mathrm{Tr} X_n^{k_2}) \approx \frac{1}{n^{\frac{k}{2}}} \sum_{m=1}^{\min\{k_1,k_2\}} \sum_{\pi \in \mathcal{DNPP}^m_{[k_1]\cup[k_2]}} \sum_{(\mathbf{P}_1,\mathbf{P}_2)\cong\pi} C_2(a(\mathbf{P}_1), a(\mathbf{P}_2))$$



$$\mathcal{DNPP}^m_{[k_1]\cup[k_2]} = \mathcal{NHPP}^m_{k_1} \times D_{2m} \times \mathcal{NHPP}^m_{k_2}$$

$\mathcal{NHPP}^m_k$ non-crossing half pair partitions of $[k]$, m connectors

Variances

Thus $C_2(\mathrm{Tr} X_n^{k_1}, \mathrm{Tr} X_n^{k_2})$ is up to $o(1)$

$$\frac{1}{n^{\frac{k}{2}}} \sum_{m=1}^{\min\{k_1,k_2\}} \sum_{g \in D_{2m}} \sum_{\pi_1 \in \mathcal{NHPP}^m_{[k_1]}} \sum_{\pi_2 \in \mathcal{NHPP}^m_{[k_2]}} \sum_{(\mathbf{P}_1,\mathbf{P}_2) \cong (\pi_g, \pi_1, \pi_2)} C_2(a(\mathbf{P}_1), a(\mathbf{P}_2))$$

$$\approx \sum_{m=1}^{\min\{k_1,k_2\}} t_{k_1,m}\, t_{k_2,m} \frac{1}{n^m} \sum_{g \in D_{2m}} \sum_{(\mathbf{P}_1,\mathbf{P}_2) \cong \pi_g} C_2(a(\mathbf{P}_1), a(\mathbf{P}_2))$$

where $t_{k,m} = \# \mathcal{NHPP}^m_k$

Theorem (Kusalik, Mingo, Speicher)

$$x^k = \sum_{m=0}^{k} t_{k,m}\, T_m(x)$$

Inverse of matrix $(t_{k,m})$ *is* $(T_{m,k})$ *where* $T_m(x) = \sum_{k=0}^m T_{m,k} x^k$

Variance: conclusion

Theorem *If* (C1) *and* (C2) *hold, then*

$$C_2\big(\mathrm{Tr}(T_m(X_1)), \mathrm{Tr}(T_k(X_j))\big) \;=\; \delta_{m=k}\, V_m \;+\; o(1)$$

Explicit formula for V_m given.

Theorem *For a generalized Wigner ensembles of matrices having T symmetries, on top of being hermitian. Then*

$$V_m \;=\; \begin{cases} T\,\mathbb{E}(d) & m=1 \\ 2T\,(\mathbb{E}(a^4)-1) & m=2 \\ 2mT & m\geq 3 \end{cases}$$