

Teoria macierzy przypadkowych jako rachunek prawdopodobieństwa XXI wieku dla realnych systemów złożonych

Maciej A. Nowak

Centrum Badań Systemów Złożonych im. Marka Kaca,
Instytut Fizyki Teoretycznej
Uniwersytet Jagielloński

Wydział Fizyki i Informatyki Stosowanej AGH
24 kwietnia 2020

NCN: Maestro Grant (2012-2017), Interdisciplinary applications of the random matrix models

FNP: TeamNet (2019-2022), Bio-inspired artificial neural networks

Czym jest teoria macierzy przypadkowych?

- "Random Matrix Theory is the theory which studies the matrix whose elements are random variables (equivalent to random variables which take values in the space of matrices). As such, it is a sort of the probability theory".
Terrence Tao, Random Matrix Theory, 2011
- "Suppose we had a theory that could explain everything. Not just atoms and quarks but aspects of our everyday lives too. Sound impossible? Perhaps not. It's all part of the recent explosion of work in the area of physics known as a random matrix theory...
"Enter the Matrix", Mark Buchanan "cover story", New Scientist, 4/10/2010, Vol. 205, Issue 2755, p 28-31.
- "-Neo: What is Matrix?...
-Morpheus: The Matrix is everywhere. It is all around us. Even now, in this very room. You can see it when you look out your window or when you turn on your television. You can feel it when you go to work... when you go to church...when you pay your taxes. "
film Matrix (1999), reż. rodzeństwo Wachowski

- Odrobina historii
- O korzyściach mechaniki kwantowej
- Nietypowe zastosowania macierzy przypadkowych
- Szersza perspektywa

Rozkłady chi-kwadrat (χ^2) i Wisharta

- Rozważmy x_i ze standardowych rozkładów Gaussa, i utwórzmy $y = \frac{1}{T} \sum_{i=1}^T x_i^2$. Gęstość prawdopodobieństwa takiego rozkładu $p(y) \sim y^{T/2-1} e^{-Ty/2}$
Kluczowy rozkład przy analizie wariancji testowaniu hipotez.
- Rozważmy wektor \vec{x}_i ze standardowego, rzeczywistego lub zespolonego rozkładu Gaussa, i utwórzmy macierz X

$$X = \begin{pmatrix} x_{11} & \dots & x_{1T} \\ \vdots & \vdots & \vdots \\ x_{N1} & \dots & x_{NT} \end{pmatrix}$$

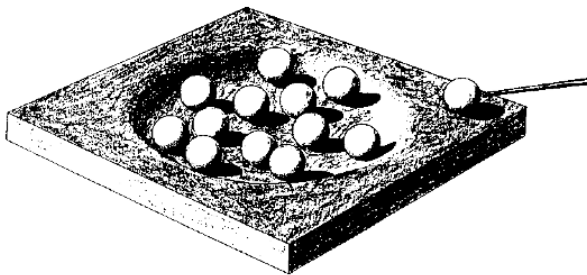
Następnie utwórzmy macierz korelacji $M = \frac{1}{T} X X^\dagger$.

Rozkład Wisharta (przy $N \leq T$)

$$P(M) \sim \det M^{(T-N-1)/2} e^{-T/2 \operatorname{tr} M}$$

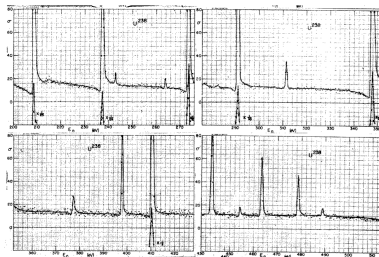
Przypadek zespolony $P(M) \sim \det M^{T-N} e^{-T \operatorname{tr} M}$

Drewniana zabawka Nielsa Bohra

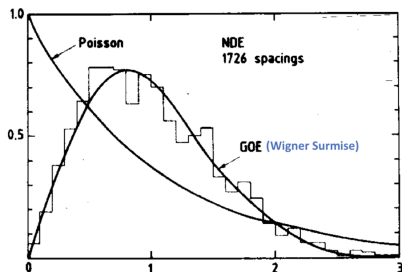


Macierze przypadkowe i Zimna Wojna

- 6.08.1945, 8.15, Hiroshima
- Zrozumienie oddziaływań powolnych neutronów z ciężkimi jądrami staje się problemem strategicznym



- Nowy paradygmat - przypadkowe Hamiltoniany



- Domysł Wignera, oparty o przypadkowe macierze 2 na 2:
$$P(s) = \frac{\pi s}{2} e^{-\pi s^2/4}$$
 gdzie $s = S/D$
- Uniwersalność - triumfujący pochodź teorii macierzy przypadkowych, od fizyki jądrowej, poprzez atomową, teorię chaosu, układy mezoskopowe....

Genialny pomysł Gaudina-Mehty

- $P(\lambda_1, \dots, \lambda_N) \sim \prod \lambda_i^{T-N} e^{-T \sum \lambda_i} \Delta(\Lambda)^2$
- Vandermonian!

$$\Delta = \prod_{i < j} (\lambda_i - \lambda_j) \sim \det \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_{N-1} & \lambda_N \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_{N-1}^2 & \lambda_N^2 \\ \dots & \dots & \dots & \dots & \dots \\ \lambda_1^{N-1} & \lambda_2^{N-1} & \dots & \lambda_{N-1}^{N-1} & \lambda_N^{N-1} \end{pmatrix}$$

- Własności wyznaczników po raz pierwszy

$$\Delta(\lambda_1, \dots, \lambda_N) \sim \det \begin{pmatrix} P_0(\lambda_1) & \dots & P_0(\lambda_N) \\ \dots & \dots & \dots \\ P_{N-1}(\lambda_1) & \dots & P_{N-1}(\lambda_N) \end{pmatrix}$$

- Wyznacznik Slatera

$$P_N(\lambda_1, \dots, \lambda_N) = \frac{1}{N!} \left[\det \psi_{j-1}^{(N)}(\lambda^k) \Big|_{j,k=1}^N \right]^2 = \frac{1}{N!} [\det K_N(\lambda_i, \lambda_j)]$$

gdzie jądro

$$K_N(\lambda, \mu) = \sum_{l=0}^{N-1} \psi_l^{(N)}(\lambda) \psi_l^{(N)}(\mu)$$

zaś $\psi_l^{(N)}(\lambda) = e^{-T\lambda/2} \lambda^{(T-N)/2} P_l^{(N)}(\lambda)$ jest funkcją falową

- Mechanika kwantowa I
Radialne równanie Schroedingera dla atomu wodoru (w jednostkach $2\mu = 1$). Całkowicie rozwiązywalny problem dla dowolnego $N, T!$

The same equations have the same solutions!

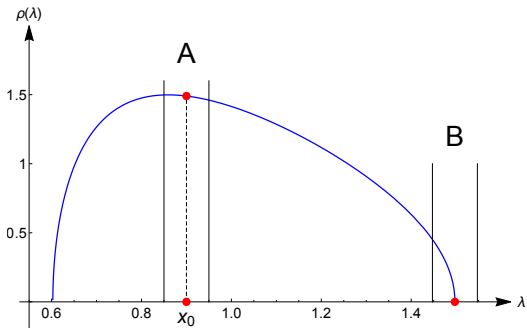
- $\hbar \sim \frac{1}{T}$
- Wzór Bohra-Sommerfelda $\oint p(r)dr = (N + \frac{1}{2})2\pi$
- Semi-klasycznie, $p^2 - \frac{1}{r} + \frac{l(l+1)}{r^2} = E$.
W granicy $N, T \rightarrow \infty$, $N/T = c$ ustalone, relacja Bohra-Sommerfelda daje wartości własnych dla zespołu Wisharta (tzw. wzór Marczenki-Pastura).
-

$$\int_{r_-}^{r_+} \rho(x) dx = 1$$

$$\rho(x) = \frac{1}{2\pi c x} \sqrt{(r_+ - x)(x - r_-)}$$

zaś $r_{\pm} = (1 \pm \sqrt{c})^2$ są klasycznymi punktami powrotu.

Mikroskopia kwantowa



"Hurt (Bulk)"

$$n_{bulk} \sim N \int_{x_0-s/2}^{x_0+s/2} \rho(x_0) dx \sim Ns\rho(x_0)$$

czyli mamy skalowanie $s \sim 1/(N\rho(x_0))$

"Miękki brzeg"

$$n_{soft} \sim N \int_{-s/2}^{s/2} \sqrt{x} dx \sim Ns^{3/2}$$

czyli mamy skalowanie $s \sim 1/N^{2/3}$.

"Twardy brzeg"

$$n_{hard} \sim N \int_0^s \frac{dx}{\sqrt{x}} \sim N\sqrt{s}$$

czyli mamy skalowanie $s \sim 1/N^2$.

Mechanika kwantowa II

$\hat{K}_N = \sum_i^N |\psi_i\rangle\langle\psi_i|$ jest operatorem rzutowym ($\hat{K}^2 = \hat{K}$)

Warunek spektralny $\langle\psi|\hat{H}_N|\psi\rangle \leq E_N$, gdzie $E_N = -1/4N^2$, połączony z odpowiednim skalowaniem mikroskopowym, pokaże jak zdeformuje się dziedzina operatora. Odczytanie tej deformacji daje rozwiązanie dla jądra K w każdym z trzech powyższych obszarach mikroskopowych.

MAN, Tarnowski, *Narain transform for spectral deformation of random matrix models*, Nucl. Phys. B (2020)

Hurt (Bulk)

- Równanie $\hat{H}_N \leq E_N$, przy skalowaniu $x/T = x_0 + \frac{s}{N\rho(x_0)}$, przechodzi (modulo zmiany zmiennych) w równanie $-\frac{d^2}{ds^2} \leq \pi^2$
- Mechanika kwantowa każe użyć fal płaskich, wtedy $(2\pi t)^2 \leq \frac{1}{4}$, czyli deformacja polega na ograniczeniu pędów do paska $[-1/2, 1/2]$.
- Operator identyczności $F(t') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{2\pi i t' s} e^{-2\pi i t s} F(t) dt ds$
deformuje się do $\mathbf{P}[F(t')] = \int_{-\infty}^{\infty} \left[\int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i t' s} e^{-2\pi i t s} ds \right] F(t) dt$
- Stąd uniwersalna funkcja Dysona

$$K_{Sine}(t, t') = \frac{\sin(\pi(t' - t))}{\pi(t' - t)}$$

Miękkie i twarde brzegi

Te samą metodę stosujemy do miękkich i twardych brzegów.

- Deformacja w przypadku miękkiego brzegu daje równanie $-\frac{d^2}{ds^2} + s \leq 0$ (trojkątny potencjał), co prowadzi do uniwersalnego jądra Airy'ego

$$K_{\text{Airy}}(t, t') = \frac{Ai(t')Ai'(t) - Ai'(t')Ai(t)}{t' - t}$$

- Deformacja w przypadku twardego brzegu daje równanie $-\frac{d^2}{ds^2} - \frac{1}{s} \frac{d}{ds} - \frac{\alpha^2}{s^2} \leq 1$ (Hankela), co prowadzi do uniwersalnego jądra Bessela

$$K_{\text{Bessel}}(x, y) = \frac{J_\alpha(\sqrt{x})J'_\alpha(\sqrt{y})\sqrt{y} - \sqrt{x}J'_\alpha(\sqrt{x})J_\alpha(\sqrt{y})}{2(x - y)}$$

gdzie $\alpha = T - N$.

- I. Uniwersalne mikroskopowe skalowanie typu "hurt" (Dysona, tzw. sine kernel)
- II. Uniwersalne mikroskopowe skalowanie typu "miękki brzeg" (Airy kernel)
- III. Uniwersalne mikroskopowe skalowanie typu "twardy brzeg" (Bessel kernel)
- IV. Makroskopowe centralne twierdzenie graniczne - rozkład Marczenko-Pastura

Przykład I - uniwersalna funkcja Dysona a hipoteza Riemanna

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{\text{prime}} \left(1 - \frac{1}{p^s}\right)^{-1}$$

- HR: Wszystkie nietrywialne zera tej funkcji mają postać $x = \frac{1}{2} + it$, czyli leżą na linii krytycznej.
- Jeden z 7 problemów milenijnych Instytutu Matematycznego Claya
- Numeryczne metody - Riemann 1859 (kilka pierwszych zer), Turing 1952 (1054 zera), Odlyzko 1980-2000 (10^{23} zer) , Bober i Hiary 2016 (10^{32} zer).
- Hipoteza Dysona-Montgomery-Odlyzko
- Test Odlyzko (1980) - obliczenia odległości pomiędzy 176 milionami zer wokół zera nr 10^{20}
($x = 1/2 + i 1520244115920747268, 6290299$)
- Dowód poprzez pokazanie istnienia Hamiltonianu, którego wartościami własnymi są części urojone zer funkcji Riemanna?

Przykład I - uniwersalna funkcja Dysona a hipoteza Riemanna, cd

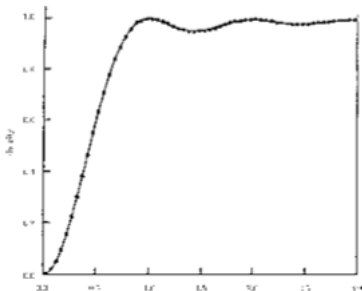
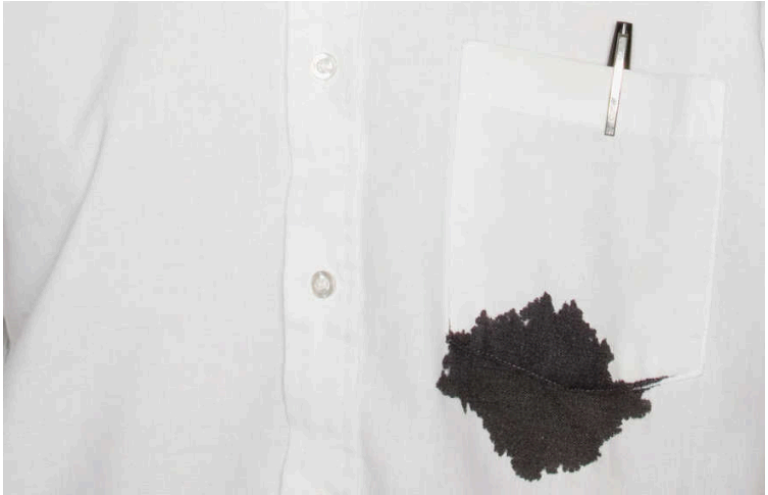


Figure 4: Testing Montgomery's conjecture using 70 million zeros near zero 10^{20} .

Przykład II - uniwersalna funkcja Airy'ego a granice międzyfazowe



Przykład II - uniwersalna funkcja Airy'ego

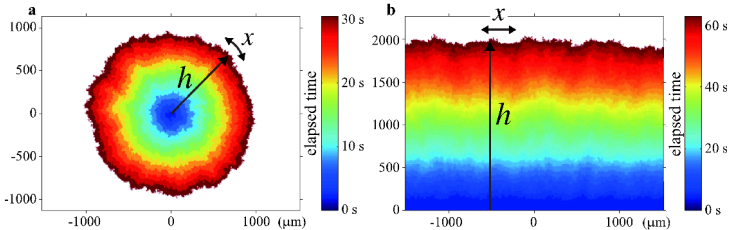


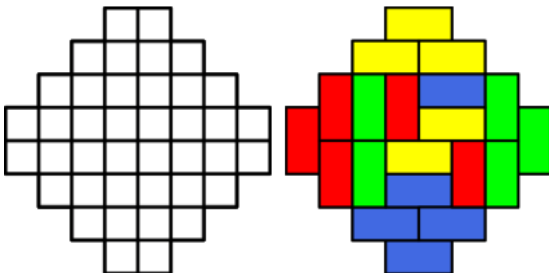
Figure 1 | Growing DSM2 cluster with a circular (a) and flat (b) interface. Binarised snapshots at successive times are shown with different colours. Indicated in the colour bar is the elapsed time after the laser emission. The local height $h(x, t)$ is defined in each case as a function of the lateral coordinate x along the mean profile of the interface (a circle for a and a horizontal line for b). See also Supplementary Movies 1 and 2.

Takeuchi, Sano, Sasamoto, Spohn; Nature Scientific Report 2011

Dynamika opisywana przez równanie Kardara-Parisiego-Zhanga. Matematyczna teoria dla równania KPZ - medal

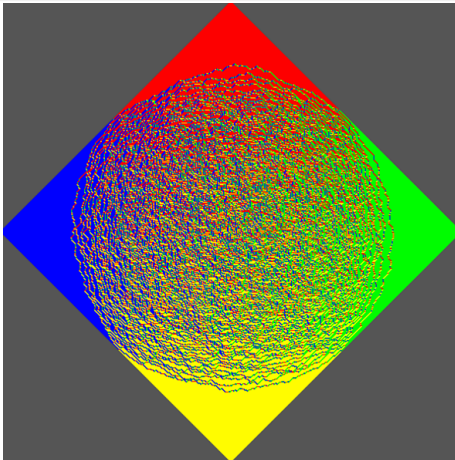
Fieldsa dla Martina Hairera 2016.

Przykład II - Diament Azteków, "ropuchy" i koło podbiegunowe



Ropucha (ang. TOAD = Tiling Of Aztec Diamont)

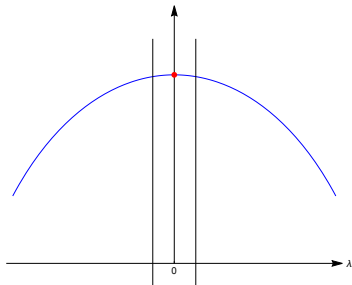
Przykład II - Diament Azteków i koło podbiegunowe, cd



"Koło podbiegunowe" oddzielające kryształ ("lód") od fazy nieuporządkowanej ("ciecz").

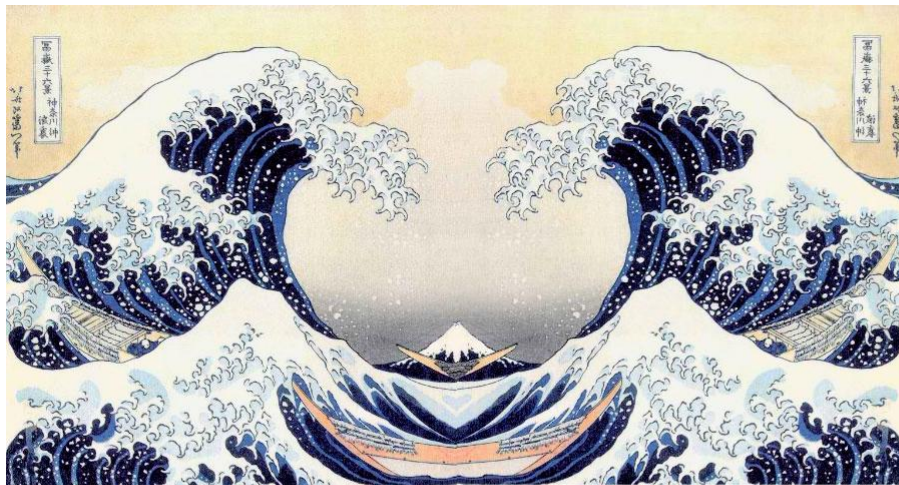
Przykład III - Chromodynamika Kwantowa i uniwersalna funkcja Bessela

- Spontaniczne łamanie symetrii
- Chiralność
- Kondensat kwarkowy jako magnetyzacja
- Wzór Banksa-Cashera $\langle \bar{\psi}\psi \rangle_{\rho(\lambda)} \sim \frac{\rho_D(0)}{V_4}$

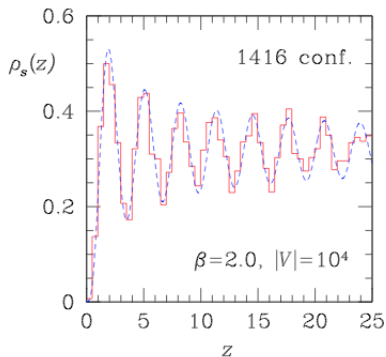


Przykład III - Chromodynamika Kwantowa i uniwersalna funkcja Bessela, cd

Chiralny GUE \equiv Wishart



Przykład III - Chromodynamika Kwantowa i uniwersalna funkcja Bessela, cd



Berbenni-Bitsch et al. (1998)

Przykład IV: z (mojej) głowy



Figure: OchabLAB

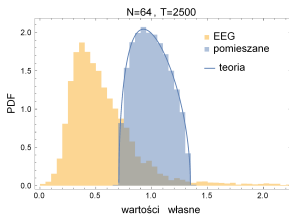
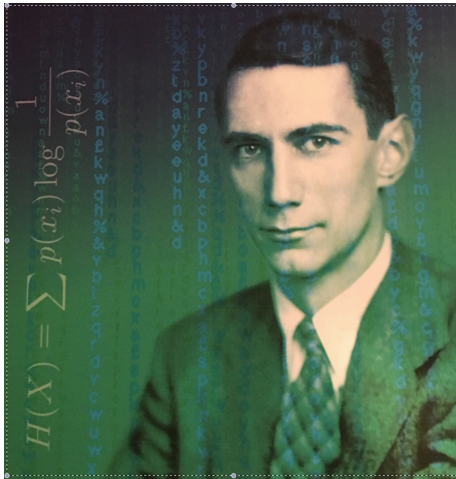


Figure: *Dolce far niente...*

- Poważniejsze zastosowanie w sztucznej inteligencji:
Dynamical Isometry is Achieved in Residual Networks in a Universal Way for any Activation Function
[W. Tarnowski, P. Warchoł, S.Jastrzębski, J. Tabor, MAN; AISTAT2019]

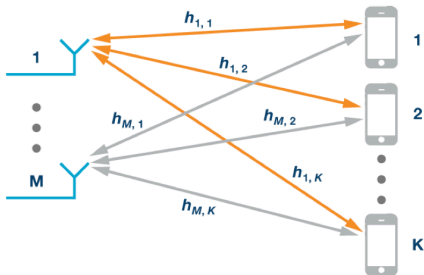
Przykład IV - Teoria informacji



Okładka książki "Information Theory - A Tutorial Introduction", J.V. Stone (2015).

Przykład IV - Marczenko-Pastur i Twój telefon

- Twierdzenie Shannona: Pojemność informacyjna kanału gaussowskiego wynosi $C = \frac{1}{2} \log(1 + \text{SINR})$ bitów gdzie SINR to Signal-to-Noise-Ratio.
- Jak zaspokoić rosące eksponencjalnie potrzeby użytkowników telefonów komórkowych?
- Foschini i Telatar (1996-1999) - Systemy MIMO, $C \sim \log \det(\mathbf{1} + \text{SINR} \mathbf{H}\mathbf{H}^\dagger)$.
- Rozkład Marchenko-Pastura maksymalizuje pojemność informacyjna MIMO.
- Dowód poprawności oraz praktyka: $\infty = 8$.



$$\mathbf{H} = \begin{bmatrix} h_{1,1} & \dots & h_{1,K} \\ \vdots & \ddots & \vdots \\ h_{M,1} & \dots & h_{M,K} \end{bmatrix}$$

Przykład IV - Marczenko-Pastur i Twój telefon, cd

Podziękowania dla Prof. Ralfa Muellera (Uniwersytet w Erlangen)



Wishart inside

- *If you asked me to describe the rising philosophy of the day, I'd say it is Data-ism, [David Brooks, New York Times, February 2013] relying on data can reduce cognitive biases and "illuminate patterns of behavior we haven't yet noticed".*
- Big Data is transforming society, using the term to describe the Big Data revolution [Steve Lohr, *Data-ism*, 2015]
- *Dataism declares that the universe consists of data flows, and the value of any phenomenon or entity is determined by its contribution to data processing, [Yuval N. Harari, *Homo Deus*, 2016]*

- Cztery ery rozwoju technologicznego: Era Pary, Era Elektryczności, Era Elektroniki, Era Informacji
- **Teoria wielkich macierzy losowych wydaje się być idealnym narzędziem poznawczym w czwartej erze technologicznej**