

No estudo de Matrizes Aleatórias e Simulação de Gases de Coulomb

PIMENTA, J. V. A.

30 de abril de 2023

Resumo

Resumo dos estudos de Iniciação Científica em Matrizes Aleatórias e Simulação de Gases de Coulomb.

Conteúdo

1	Introdução	2
1.1	Porquê exponencial?	3
1.1.1	Shannon-Quem?	3
1.2	Independência ou Morte	3
1.3	Uma medida à Hermitiana	4
2	Movimento Browniano	7
2.1	Emsemble Biortogonal	7
2.2	Karlin-McGregor	7
2.2.1	O teorema	7
2.2.2	Consequências	9
3	Chapter Two Title	10
4	Chapter Three Title	11
5	Chapter Four Title	12
6	Conclusion	13
A	Appendix Title	15

Capítulo 1

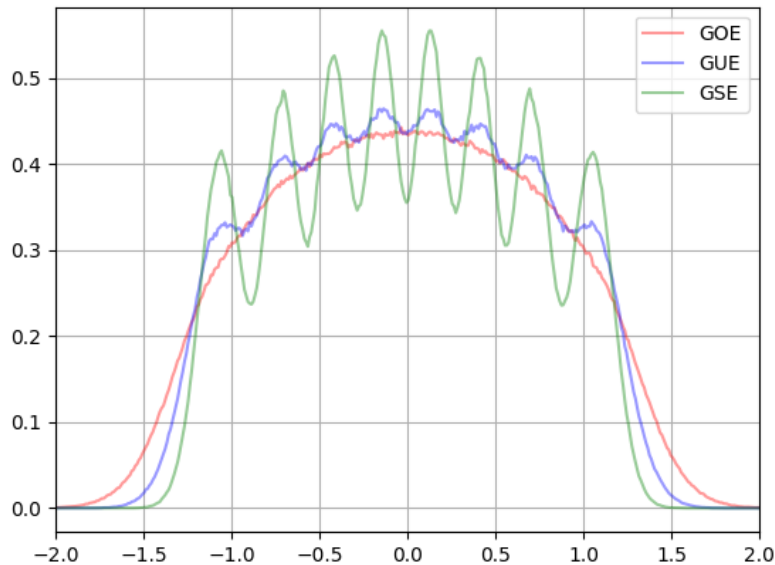
Introdução

Um esquema útil para organização é a classificação de Layman. Denomina-se

1. **Entradas Independentes:** Adicionada a necessidade de simetria, matrizes deste tipo são chamadas matrizes de Wigner.
2. **Invariantes por rotação:** Quaisquer duas matrizes que são relacionadas por uma transformação $\hat{H}' = \hat{U}\hat{H}\hat{U}^{-1}$ ocorrerão com a mesma probabilidade.

Só existe um tipo especial de matriz que se encontra na intersecção e são as classes Gaussianas.

Em geral no mundo das matrizes aleatórias três classes são muito importantes e serão atores centrais no nosso estudo. São as três classes associadas às matrizes que tem entradas gaussianas. Estas são: O Ensemble Gaussiano Ortogonal (GOE), O Ensemble Gaussiano Unitário (GUE) e o Ensemble Gaussiano Simplético (GSE). Em suma, eles tratam, em ordem, de matrizes com entradas gaussianas reais, complexas e quaterniônicas. Seus nomes estão relacionados com a matriz necessária para a transformação de diagonalização das matrizes. Unitária para o caso complexo, por exemplo. Seus autovalores possuem distribuições distintas (ao menos para a escala usada) e é ilustrada abaixo na Figura 1.



Mais desenvolvimento sobre a forma desta distribuição e as diferenças será feito posteriormente. Usaremos a referência [3] para a maior parte dos desenvolvimentos. Algum material interessante pode ser consultado em um livro de física em [4].

1.1 Porquê exponencial?

1.1.1 Shannon-Quem?

O uso da p.d.f gaussiana pode ser justificado de algumas formas. Uma primeira abordagem a ser explorada é a de maximização entrópica ou minimização de informação similar aos trabalhos de Shannon-Kinchin. Definiremos uma grandeza $\mathcal{I}[\mathcal{P}(\hat{H})]$ associada à uma p.d.f tal que:

$$\mathcal{I}[\mathcal{P}(\hat{H})] = - \int d\mu(\hat{H}) \mathcal{P}(\hat{H}) \ln \mathcal{P}(\hat{H})$$

Que é uma extensão natural da definição discreta de informação $-\sum_{l=1}^m p_m \ln p_m$. Agora argumentaremos algo parecido com os argumentos usados em termodinâmica de maximização de entropia. Diremos que a incerteza sobre as matrizes será máxima, ou seja, teremos a maior aleatoriedade das matrizes quando a entropia for maximizada e a informação, minimizada. Assim como na entropia física impomos um vínculo de energia constante, aqui faremos algo do tipo $E(\text{Tr } \hat{H}) = b$ e $E((\text{Tr } \hat{H})^2) = a > 0$. Vamos introduzir esses vínculos como multiplicadores de lagrange com multiplicadores v_1 e v_2 .

$$\mathcal{I}[\mathcal{P}(\hat{H})] = - \int d\mu(\hat{H}) \mathcal{P}(\hat{H}) \left(\ln \mathcal{P}(\hat{H}) - v_1 \text{Tr } \hat{H} - v_2 \text{Tr } \hat{H}^2 \right)$$

que tem diferencial

$$\delta \mathcal{I}[\mathcal{P}(\hat{H})] = - \int d\mu(\hat{H}) \delta \mathcal{P}(\hat{H}) \left(1 + \ln \mathcal{P}(\hat{H}) - v_1 \text{Tr } \hat{H} - v_2 \text{Tr } \hat{H}^2 \right) = 0$$

Que só vai ser mínimo se

$$\mathcal{P}(\hat{H}) \propto e^{-v_1 \text{Tr } \hat{H} - v_2 \text{Tr } \hat{H}^2}$$

Onde os multiplicadores são unicamente definidos pelas constantes do vínculo. Esse fato motiva de alguma forma o estudo da p.d.f gaussiana para as matrizes.

1.2 Independência ou Morte

Consideremos matrizes com entradas independentes. Qual a função densidade de probabilidade (F.P.D.) da matriz simétrica \hat{H}_s ? Devemos fazer separadamente a diagonal da seção triangular que formos usar e teremos

$$\rho((\hat{H}_s)_{11}, \dots, (\hat{H}_s)_{NN}) = \prod_{i=1}^N \left[\frac{e^{-\frac{(H_s)_{ii}^2}{2}}}{2\pi} \right] \prod_{i < j} \left[\frac{e^{-(H_s)_{ij}^2}}{\sqrt{\pi}} \right]$$

Podemos também definir a distribuição para os autovalores de uma matriz Gaussiana de dimensão N como ¹

$$\rho(x_1, \dots, x_N) = \frac{1}{\mathcal{Z}_{N,\beta}} e^{-\frac{1}{2} \sum_{i=1}^N x_i^2} \prod_{j < k} |x_j - x_k|^\beta \quad (1.1)$$

¹Esse resultado é não óbvio e deve ser discutido em breve.

onde a constante de normalização é dada por

$$\mathcal{Z}_{N,\beta} = (2\pi)^{\frac{N}{2}} \prod_{j=1}^N \frac{\Gamma(1 + j\frac{\beta}{2})}{\Gamma(1 + \frac{\beta}{2})}$$

Para ressaltar um pouco do jargão, β é denominado *Dyson Index* que em suma se refere à "dimensão" das suas entradas na matriz. 1 para GOE, 2 para GUE e 4 para GSE.

Algumas observações sobre essa expressão são interessantes. Note que o fator exponencial deve matar qualquer chance de uma matriz com autovalor alto. Ao mesmo tempo o fator de dependência deve matar qualquer configuração com autovalores muito próximos entre si. Existe um efeito de repelência entre autovalores na expressão.

1.3 Uma medida à Hermitiana

Consideraremos no nosso estudo para referencia matrizes quadradas de entradas complexas com dimensão N . Nosso objetivo é afinal ter uma forma de mensurar a distribuição de autovalores e para isso, faremos os seguintes desenvolvimentos.

Consideremos inicialmente um espaço de matrizes com entradas complexas $2N^2$ dimensional. Contido neste espaço temos um espaço de maior interesse correspondente ao espaço das matrizes *hermitianas* de dimensão N^2 . A escolha do subespaço está relacionada com o fato que matrizes hermitianas são diagonalizáveis e a distribuição de seus autovalores estará diretamente relacionada (com uma mudança de base) à distribuição do traço da matriz diagonalizada. Note que para a matriz diagonal ter a mesma medida que nossa matriz inicial, nossa medida deve ser invariável por rotação.

Mais detalhadamente podemos escrever nossa matriz hermitiana \hat{H} como

$$\hat{H} = \hat{U} \hat{\Lambda} \hat{U}^{-1}, \quad \hat{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N), \quad \hat{U} \cdot \hat{U}^* = I$$

onde, claro, $\hat{\Lambda}$ é diagonal de autovalores e \hat{U} é unitária e com colunas equivalentes aos autovetores de \hat{H} . Em geral, o conjunto de matrizes degeneradas tem medida nula e não é uma preocupação. Um cuidado deve ser tomado. A correspondência $\hat{H} \implies (\hat{U}, U(N), \hat{\Lambda})$ não é injetora, podemos tomar $\hat{U}_1 \hat{\Lambda} \hat{U}_1^{-1} = \hat{U}_2 \hat{\Lambda} \hat{U}_2^{-1}$ se $\hat{U}_1^{-1} \hat{U}_2 = \text{diag}(e^{i\phi_1}, \dots, e^{i\phi_N})$ para qualquer escolha de fases (ϕ_1, \dots, ϕ_N) . Para restringir nosso problema e tornar a função injetiva será necessário considerar as matrizes unitárias ao espaço de coset $U(N)/U(1) \times \dots \times U(1)$ ². Uma outra restrição necessárias é ordenar os autovalores, ou seja, $\lambda_1 < \dots < \lambda_N$. Temos que reescrever agora a medida $d\mu(\hat{H})$ em função de autovalores e da \hat{U} de autovetores.

Para resumir o desenvolvimento, alguns resultados serão diretamente enunciados. Essa seção pode ser encontrada no relatório [2]. Em especial recuperaremos o elemento de distância e volume no subespaço que vamos tratar

$$(ds)^2 = \text{Tr} d\hat{H} d\hat{H}^* = \sum_i (dx_{ii})^2 + 2 \sum_{i < j} [(dx_{ij})^2 + (dy_{ij})^2] \quad (1.2)$$

$$d\mu(\hat{H}) = 2^{\frac{N(N-1)}{2}} \prod_i dx_{ii} \prod_{i < j} dx_{ij} dy_{ij} \quad (1.3)$$

Ambos vem de um desenvolvimento da métrica do espaço discutido. Note que nossa medida de comprimento é invariante em respeito à automorfismos. Especificamente, se tomarmos os elementos (1.2) e (1.3) na decomposição espectral, obteremos

²Não tenho muita ideia de espaços de Coset. Pelo que entendo, existe um espaço onde toda \hat{U} pode ser representada por $\hat{U}_c \hat{U}_d$, onde \hat{U}_c compõe o espaço de coset e \hat{U}_d é uma matriz diagonal unitária. Dessa forma matrizes equivalentes são aquelas que multiplicadas por \hat{U}_d tem um mesmo resultado.

$$(ds)^2 = \sum_i (d\lambda)^2 + \sum_{i < j} (\lambda_i - \lambda_j)^2 \overline{\delta U_{ij}} \delta U_{ij} \quad (1.4)$$

e

$$d\mu(\hat{H}) = \prod_{i < j} (\lambda_i - \lambda_j)^2 \prod_i d\lambda_i \times d\mu(\hat{U}) \quad (1.5)$$

Tendo a medida de integração pronta, podemos definir uma F.D.P $\mathcal{P}(\hat{H})$ neste espaço de matizes hermitianas tal que $\mathcal{P}(\hat{H})d\mu(\hat{H})$ é a probabilidade da matriz \hat{H} estar no volume $d\mu(\hat{H})$. Queremos que nossa função seja invariante à rotação, ou seja, $\mathcal{P}(\hat{H}) = \mathcal{P}(\hat{U}^* \hat{H} \hat{U})$.

Conhecer os N primeiros traços ($\text{Tr } \hat{H}^n$) de \hat{H} define unicamente o polinômio característico e junto com ele, os autovalores. Especificamente tomaremos

$$\mathcal{P}(\hat{H}) = C e^{-\text{Tr } Q(\hat{H})} \quad (1.6)$$

Onde Q deve ser um polinômio de até ordem $2j \leq N$ suficiente para garantir a convergência de

$$\mathcal{Z}_n = \int_{\mathcal{H}_n} e^{-\text{Tr } Q(\hat{M})} d\hat{M}$$

Comumente uma condição suficiente é

$$\lim_{x \rightarrow \pm\infty} \frac{Q(x)}{\ln(1+x^2)} = \infty$$

Mas em especial, se tomarmos

$$Q(x) = ax^2 + bx + c$$

Nossa medida tomará a forma

$$\mathcal{P}(\hat{H}) = e^{-a[\sum_i x_{ii}^2 + 2\sum_{i < j} [x_{ij}^2 + y_{ij}^2]]} e^{-b\sum_i x_{ii}} e^{-cN} \quad (1.7)$$

$$= e^{-cN} \prod_{i=1}^N \left(e^{-ax_{ii}^2 - bx_{ii}} \right) \prod_{i < j} e^{-2ax_{ij}^2} \prod_{i < j} e^{-2ay_{ij}^2} \quad (1.8)$$

Onde podemos notar que a distribuição de probabilidade da matriz \hat{H} pode ser representados por fatores independentes, cada um de forma gaussiana. Para este potencial, temos uma conexão entre as matrizes de entrada independentes e as matrizes invariáveis por rotação. Lembre-se que para as variáveis serem independentes \mathcal{P} deve ter a forma $\mathcal{P} = C e^{-(a \text{Tr } \hat{H}^2 + b \text{Tr } \hat{H} + cN)}$ para constantes $a > 0, b, c$. Em nota, sabemos então

$$e^{\text{Tr } V(\hat{H})} d\mu(\hat{H}) = e^{-\sum_j V(\lambda_j)} \prod_{i < j} (\lambda_i - \lambda_j)^2 d\mu(\lambda) d\mu(\hat{U})$$

ou mais geralmente para o ensemble com

$$\frac{1}{\tilde{\mathcal{Z}}_n} e^{\text{Tr } (V(\hat{M}))} d\hat{M}$$

Dado λ_j os autovalores

$$\text{Tr } (V(\hat{M})) = -\sum_{j=1}^n V(\lambda_j)$$

e finalmente podemos escrever

$$E[f] = \int_{\mathcal{H}_n} f(\hat{M}) e^{-\text{Tr}(Q(\hat{M}))} d\hat{M} \quad (1.9)$$

$$= \frac{1}{\mathcal{Z}} \int \cdots \int f(\lambda_1, \dots, \lambda_n) \prod_{i < j} (\lambda_i - \lambda_j)^2 \prod_{j=1}^n e^{-Q(\lambda_j)} d\lambda_1 \dots d\lambda_n \quad (1.10)$$

Assim, a probabilidade conjunta nas matrizes induz uma densidade de probabilidade de autovalores

$$\frac{1}{\mathcal{Z}_n} \prod_{i < j} (\lambda_i - \lambda_j)^2 \prod_{j=1}^n e^{-Q(\lambda_j)}$$

Alguns resultados foram resgatadas da nota do autor em [1].

Capítulo 2

Movimento Browniano

2.1 Emsemble Biortogonal

Um n-ponto processo é um ensemble biortogonal se existem duas sequências f_1, \dots, f_n e g_1, \dots, g_n em $L^2(R)$ e uma constante $Z_n \neq 0$ tais que:

$$\mathcal{P}(x_1, \dots, x_n) = \frac{1}{Z_n} \det [f_i(x_j)]_{i,j=1}^n \cdot \det [g_i(x_j)]_{i,j=1}^n$$

Onde todo f_i e g_i é independente nos i 's. Pode-se mostrar que se

$$\phi_j \in \text{span}(f_1, \dots, f_n) \quad \psi_j \in \text{span}(g_1, \dots, g_n)$$

tais que

$$\int_{-\infty}^{\infty} \phi_k(c) \psi_j(x) dx = \delta_{jk}$$

então

$$K_n(x, y) = \sum_{j=1}^n \phi_j(x) \psi_j(y)$$

onde $K_n(x, y)$ é um *Kernel* tal que

$$\mathcal{P}(x_1, \dots, x_n) = \frac{1}{n!} \det [K_n(x_i, x_j)]_{i,j=1}^n$$

O processo é determinado e K_n é o kernel de correlação.

2.2 Karlin-McGregor

2.2.1 O teorema

Exploraremos os caminhos não cruzantes providos por processos de Markov. Considere uma partícula de movendo com uma regra qualquer, vamos descrever esse movimento de forma que denotaremos $p_t(a; x)$ a densidade de probabilidade de transição; isto é, a chance uma partícula em a ir para x em um próximo momento. Um teorema clássico enuncia a probabilidade de um certo número de caminhos não se intersectarem passado um tempo t .

O teorema diz: Considere $X_1(t), \dots, X_n(t)$ cópias independentes de um processo forte de Markov com caminhos condicionados tais que

$$X_j(0) = a_j$$

onde $a_1 < a_2 < \dots < a_n$ são valores dados. Notamos novamente $p_t(x, y)$ ser a densidade do processo de transição. Vamos definir regiões E_1, E_2, \dots, E_n onde E 's vizinhos não se intersectam. Temos

$$\int_{E_1} \dots \int_{E_n} \det [p_t(a_i, x_j)]_{i,j=1}^n dx_1 \dots dx_n$$

vai ser a probabilidade de que os caminhos não tenham se intersectados no intervalo de tempo $[0, t]$ e $X_j(t)$ nos intervalos correspondentes. A demonstração está em [1]. Note que temos

$$\begin{aligned} & \int_{E_1} \dots \int_{E_n} \det [p_t(a_i, x_j)]_{i,j=1}^n dx_1 \dots dx_n \\ &= \int_{E_1} \dots \int_{E_n} \begin{vmatrix} p_t(a_1, x_1) & p_t(a_2, x_1) & \dots & p_t(a_{n-1}, x_1) & p_t(a_n, x_1) \\ p_t(a_1, x_2) & p_t(a_2, x_2) & \dots & p_t(a_{n-1}, x_2) & p_t(a_n, x_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_t(a_1, x_{n-1}) & p_t(a_2, x_{n-1}) & \dots & p_t(a_{n-1}, x_{n-1}) & p_t(a_n, x_{n-1}) \\ p_t(a_1, x_n) & p_t(a_2, x_n) & \dots & p_t(a_{n-1}, x_n) & p_t(a_n, x_n) \end{vmatrix} dx_1 \dots dx_n \end{aligned} \quad (2.1)$$

$$= \sum_{\sigma} \text{sgn}(\sigma) \prod_{j=1}^n p_t(a_j, E_{\sigma(j)}) \quad (2.2)$$

$$= \sum_{\sigma} \text{sgn}(\sigma) \mathcal{P}(A_{\sigma}) \quad (2.3)$$

Onde denotamos

$$p_t(a_j, E_{\sigma(j)}) = \int_{E_j} p_t(a_i, x_j) dx_j$$

σ é uma permutação de $1, \dots, n$ e A_{σ} é o evento que $X_j(t) \in E_{\sigma(j)}$ para todo j . Os caminhos devem ser independentes para 2.3.

De alguma forma o determinanda permuta os caminhos em todas ordens possíveis e calcula a probabilidade de todos se manterem nos intervalos adequados. Um exemplo de baixas dimensões pode mostrar que

$$\begin{vmatrix} p_t(a_1, x_1) & p_t(a_2, x_1) & p_t(a_3, x_1) \\ p_t(a_1, x_2) & p_t(a_2, x_2) & p_t(a_3, x_2) \\ p_t(a_1, x_3) & p_t(a_2, x_3) & p_t(a_3, x_3) \end{vmatrix} = \quad (2.4)$$

$$+ p_t(a_1, x_1)p_t(a_2, x_2)p_t(a_3, x_3) \quad (2.5)$$

$$+ p_t(a_2, x_1)p_t(a_3, x_2)p_t(a_1, x_3) \quad (2.6)$$

$$+ p_t(a_3, x_1)p_t(a_1, x_2)p_t(a_2, x_3) \quad (2.7)$$

$$- p_t(a_3, x_1)p_t(a_2, x_2)p_t(a_1, x_3) \quad (2.8)$$

$$- p_t(a_2, x_1)p_t(a_1, x_2)p_t(a_3, x_3) \quad (2.9)$$

$$- p_t(a_1, x_1)p_t(a_3, x_2)p_t(a_2, x_3) \quad (2.10)$$

Logo

$$\int_{E_1} \cdots \int_{E_n} \det [p_t(a_i, x_j)]_{i,j=1}^n dx_1 \dots dx_n = + p_t(a_1, E_1) p_t(a_2, E_2) p_t(a_3, E_3) \quad (2.11)$$

$$+ p_t(a_2, E_1) p_t(a_3, E_2) p_t(a_1, E_3) \quad (2.12)$$

$$+ p_t(a_3, E_1) p_t(a_1, E_2) p_t(a_2, E_3) \quad (2.13)$$

$$- p_t(a_3, E_1) p_t(a_2, E_2) p_t(a_1, E_3) \quad (2.14)$$

$$- p_t(a_2, E_1) p_t(a_1, E_2) p_t(a_3, E_3) \quad (2.15)$$

$$- p_t(a_1, E_1) p_t(a_3, E_2) p_t(a_2, E_3) \quad (2.16)$$

Onde somamos os casos onde as partículas se matém ordenadas e subtraímos os casos onde elas se cruzam.

2.2.2 Consequências

Considere n cópias do processo de Markov condicionado para começar em $t = 0$ nas determinadas posições $a_1 < a_2 < \cdots < a_n$. Se condicionarmos estes processos para não intersectar no intervalo $[0, t]$, o teorema vai nos dizer que os caminhos em um tempo t vão ter uma densidade de probabilidade conjunta

$$\frac{1}{\mathcal{Z}_n} \det [p_t(a_i, x_j)]_{i,j=1}^n$$

Mas este não pode ser considerado um processo pontual determinado. Não é expreso por um produto de determinantes. Isso pode ser ajeitado se considerarmos um tempo $T > t$ no nosso processo. Tomaremos b_1, b_2, \dots, b_n posições finais e condicionaremos os caminhos a não intersectar no intervalo $[0, T]$ com $X_j(0) = a_j$ e $X_j(T) = b_j$ para todos. É possível mostrar que a distribuição conjunta deles será

$$\frac{1}{\mathcal{Z}'_n} \det [p_t(a_i, x_j)]_{i,j=1}^n \det [p_{T-t}(x_i, b_j)]_{i,j=1}^n$$

Que será biortogonal com as funções

$$f_j = p_t(a_j, x) ; g_j = p_{T-t}(x, b_j)$$

E nosso caso de interesse é quando $a_j \rightarrow a$ e $b_j \rightarrow b$. Note que usando as duas funções podemos forçar que o movimento browniano se inicie em um ponto e encerre em outro determinado. Em uma, reverteremos o tempo e, nos limites 0 e T , forçaremos que apenas uma das funções seja predominante de forma que a posição inicial de cada uma prevaleça. Podemos impor a posição inicial e final do movimento. No caso browniano teremos

$$p_t(a, x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-a)^2}{2t}}$$

No caso dos limites de a e b ficamos com

$$f_j = F_{j-1}(x) e^{-\frac{(x-a)^2}{2t}} ; g_j = G_{j-1}(x) e^{-\frac{(x-b)^2}{2(T-t)}}$$

onde F e G são polinômios em x de grau $j - 1$. Este processo podemos escalar e transladar para uma versão do GUE $n \times n$.

Capítulo 3

Chapter Two Title

Capítulo 4

Chapter Three Title

Capítulo 5

Chapter Four Title

Capítulo 6

Conclusion

Bibliografia

- [1] Lecture notes on riemann-hilbert problems and multiple orthogonal polynomials.
- [2] Yan V. Fyodorov. Introduction to the random matrix theory: Gaussian unitary ensemble and beyond, 2010.
- [3] Giacomo Livan, Marcel Novaes, and Pierpaolo Vivo. *Introduction to Random Matrices*. Springer International Publishing, 2018.
- [4] M.L. MEHTA. Preface. In M.L. MEHTA, editor, *Random Matrices and the Statistical Theory of Energy Levels*. Academic Press, 1967.

Apêndice A

Appendix Title