

# No estudo de Matrizes Aleatórias e Simulação de Gases de Coulomb

PIMENTA, J. V. A.

7 de maio de 2023

## **Resumo**

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

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>2</b>
1.1	Física Estatística . . . . .	3
1.1.1	Uma abordagem mais física . . . . .	3
1.1.2	A entropia de Shannon . . . . .	4
1.1.3	Conexão Micro-Macro . . . . .	5
1.2	Porquê exponencial? . . . . .	6
1.2.1	Shannon-Quem? . . . . .	6
1.3	Independência ou Morte . . . . .	7
1.4	Uma medida à Hermitiana . . . . .	7
<b>2</b>	<b>Movimento Browniano</b>	<b>10</b>
2.1	Processo Pontual . . . . .	10
2.1.1	Poisson & fries . . . . .	10
2.1.2	Função Correlação . . . . .	11
2.1.3	Pontual Determinante . . . . .	11
2.2	Emsemble Biortogonal . . . . .	12
2.3	Karlin-McGregor . . . . .	12
2.3.1	O teorema . . . . .	12
2.3.2	Consequências . . . . .	14
<b>A</b>	<b>Transformação Legendre</b>	<b>18</b>
A.1	Tangentes . . . . .	18
A.2	Otimização . . . . .	18
<b>B</b>	<b>Soma Assintóticas</b>	<b>20</b>

# Capítulo 1

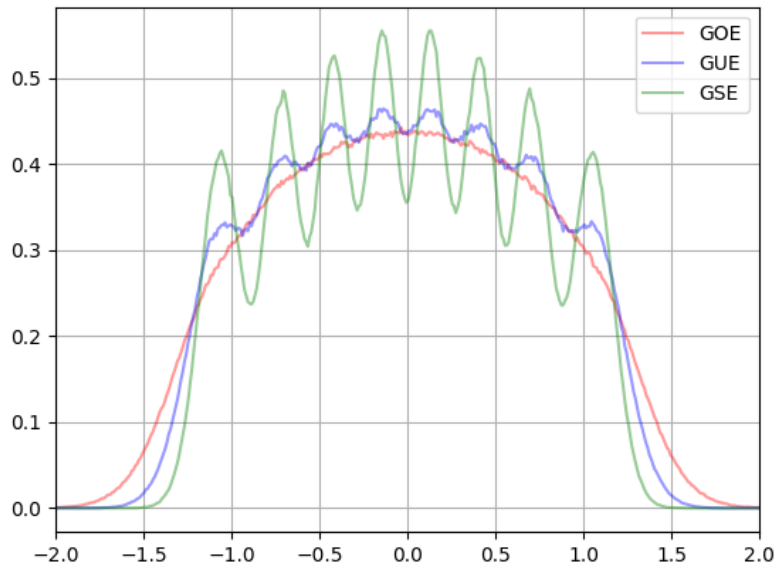
## Introdução

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).



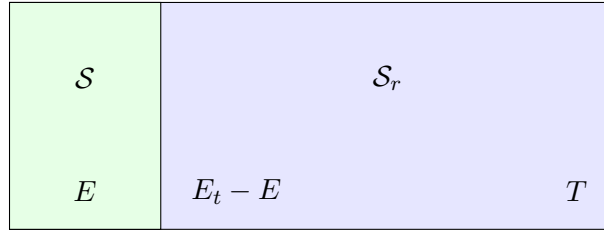
Mais desenvolvimento sobre a forma desta distribuição e as diferenças será feito posteriormente. Usaremos a referência [3] para a maior parte dos desenvolvimentos. Algum material interessante pode ser consultado em um livro de física em [4].

## 1.1 Física Estatística

Estaremos lidando o tempo todo no nosso estudo com funções partições, de alguma forma, compartilhada com a física. Entenderemos um pouco mais sobre os desenvolvimentos do ensemble canônico na termodinâmica.

### 1.1.1 Uma abordagem mais física

Para o desenvolvimento da função partição estamos atentos aos requisitos do ensemble canônico. Um sistema termalizado por um banho térmico. Nosso sistema completo pode ser representado



Considere que os sistemas estão em contato térmico e o sistema completo, junto com o banho é um sistema isolado de energia  $E_t$ . Sabemos que a probabilidade de uma energia no nosso sistema termalizado ( $\rho_E$ ) é expressa:

$$\rho_E = \sum_{\sigma} \rho_{\sigma} = \Omega(E) \rho_{\sigma}$$

Onde note,  $\rho_{\sigma}$  é a probabilidade de, dada uma temperatura, um microestado possível. Note por outro lado que podemos expressar:

$$\rho_E = \frac{\Omega(E) \Omega_r(E_t - E)}{\sum_i \Omega(E_i) \Omega_r(E_t - E)}$$

Note que isso é nada mais do que dizer que os microestados são equiprováveis e basta uma contagem (normalizada) para definir a probabilidade. Neste sentido podemos também escrever:

$$\rho_{\sigma} \propto \Omega_r(E_t - E)$$

Isso é, quanto mais formas o reservatório possa se organizar para determinada energia, mais provável é o microestado associado à energia de  $\mathcal{S}$ .

$$\begin{aligned} \rho_{\sigma} &\propto e^{\beta T K_b \log(\Omega_r(E_t - E))} \\ &\propto e^{\beta T (\mathcal{S}_r(E_t) - \frac{E}{T})} \\ &\propto e^{-\beta E} e^{\beta T \mathcal{S}_r(E_t)} \\ &\propto e^{-\beta E} \end{aligned}$$

Onde fizemos a expansão de  $k_b \log(\Omega_r(E_t - E))$  (entropia) em Taylor (apesar de que, em realidade, apenas  $\frac{E}{E_t}$  é pequeno, não necessariamente  $E$ ) e obtemos

$$\mathcal{S}_r(E_t - E) \approx \mathcal{S}_r(E_t) + \left( \frac{\partial \mathcal{S}_r}{\partial E} \right)_{E=E_t} (-E)$$

Onde

$$\left( \frac{\partial \mathcal{S}_r}{\partial E} \right)_{E=E_t} = \frac{1}{T}$$

### 1.1.2 A entropia de Shannon

Podemos fazer algo um pouco mais matemático. Definiremos a entropia de Shannon

$$\mathcal{S} = -k_b \sum_i p_i \log p_i$$

#### Um vínculo

Se impormos um vínculo do tipo

$$\sum_i p_i = 1$$

Podemos usar os multiplicadores de Lagrange para realizar a maximização da nossa função. Pela definição, escrevemos:

$$S(p_i, \lambda) \equiv k_b \sum_{i=1}^N p_i \log(p_i) - \lambda \left( \sum_{i=1}^N p_i - 1 \right)$$

Com diferencial

$$\delta S = -k_b \sum_{i=1}^N \left( p_i \log p_i + \frac{p_i}{p_i} \delta p_i \right) - \lambda \sum_{i=1}^N \delta p_i$$

Resolveremos agora o sistema

$$\begin{cases} \delta S = 0 \\ \sum_i p_i = 1 \end{cases}$$

ou seja,

$$\begin{cases} -k_b(\log p_i + 1) - \lambda \\ \sum_i p_i = 1 \end{cases}$$

Note que  $p_i = cte \forall i$ . Teremos então  $p_i = \frac{1}{N}$ . A entropia neste caso será

$$S(p_i^*) = -k_b \sum_{i=1}^N \left( \frac{1}{N} \log \frac{1}{N} \right) = k_b \log N$$

### Dois vínculos

Faremos agora a implicação de um novo vínculo

$$\sum_{\sigma} p_{\sigma} E_{\sigma} = U$$

Desenvolvendo a equação

$$S(p_i, \lambda) \equiv k_b \sum_{i=1}^N p_i \log(p_i) - \lambda_1 \left( \sum_{i=1}^N p_i - 1 \right) - \lambda_2 \left( \sum_{\sigma} p_{\sigma} E_{\sigma} - U \right)$$

Chegaremos no sistema

$$\begin{cases} -k_b(\log p_i + 1) - \lambda_1 - \lambda_2 E_{\sigma} \\ \sum_{\sigma} p_i = 1 \\ \sum_{\sigma} p_{\sigma} E_{\sigma} = U \end{cases}$$

E finalmente, da primeira equação tiramos

$$p_{\sigma} = e^{AE_{\sigma}+B} = M e^{-\beta E_{\sigma}}$$

De forma que podemos reescrever

$$1 = \sum_{\sigma} M e^{-\beta E_{\sigma}} = M \sum_{\sigma} e^{-\beta E_{\sigma}}$$

Onde nomearemos  $M = \frac{1}{\sum_{\sigma} e^{-\beta E_{\sigma}}} = \frac{1}{Z}$  **função partição**. Faltam apenas definir  $\beta$  em

$$p_{\sigma} = \frac{e^{-\beta E_{\sigma}}}{Z}$$

Para  $\beta$  podemos aplicar o último vínculo (da temperatura térmica)

$$\begin{aligned} U &= \sum_{\sigma} p_{\sigma} E_{\sigma} = \sum_{\sigma} \left( \frac{1}{Z} e^{-\beta E_{\sigma}} \right) E_{\sigma} = \frac{1}{Z} \sum_{\sigma} E_{\sigma} e^{-\beta E_{\sigma}} \\ &= -\frac{1}{Z} \frac{\partial}{\partial \beta} \left( \sum_{\sigma} e^{-\beta E_{\sigma}} \right) = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial(\log Z)}{\partial \beta} \end{aligned}$$

De alguma forma esta equação transcendental nos define  $\beta$ .

#### 1.1.3 Conexão Micro-Macro

Quando tratamos destes sistemas no ensemble canônico a energia interna não mais será minimizada no nosso sistema termalizado. Introduzimos uma nova grandeza chamada Energia Livre de Helmholtz ( $F$ ), definida como a transformada de Legendre (discutida no Apêndice A) da energia interna em relação à entropia, ou seja

$$U(S, V, N) \mapsto F(T, V, N) \tag{1.1}$$

ou ainda mais especificamente

$$F(T, V, N) = U(S(T, V, N), V, N) - TS(T, V, N)$$

Note ainda as diferenciais

$$\begin{aligned}dU &= TdS - PdV + \mu dN \\dF &= -SdT - PdV + \mu dN\end{aligned}$$

Um sistema termalizado vai querer minimizar essa nova grandeza da energia livre de Helmholtz. Em todo caso iniciaremos com a expressão já deduzida da equação de partição

$$\mathcal{Z} = \sum_{\sigma} e^{-\beta E_{\sigma}}$$

Que pode ser reescrito em termos de uma soma na energia

$$\begin{aligned}\mathcal{Z} &= \sum_E e^{-\beta E} \Omega(E) \\ \mathcal{Z} &= \sum_E e^{-\beta T K_b \log(\Omega(E))} \Omega(E)\end{aligned}$$

Podemos argumentar que  $K_b \log(\Omega(E)) = S(E)$  e usaremos o logaritmo de somas assintóticas (discutido no Apêndice B) para terminar o desenvolvimento. Note

$$\begin{aligned}\log(\mathcal{Z}) &\approx \log\left(\max_x [e^{-\beta(E-TS(E))}]\right) \\ \log(\mathcal{Z}) &\approx \log\left(e^{-\beta \min_E ((E-TS(E)))}\right)\end{aligned}$$

Onde podemos reconhecer pela transformada de Legendre o termo referente à Energia livre de Helmholtz. Tendo assim

$$\log(\mathcal{Z}) \approx -\beta F$$

$$F = -K_b T \log(\mathcal{Z}) \tag{1.2}$$

## 1.2 Porquê exponencial?

### 1.2.1 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.

### 1.3 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 <sup>1</sup>

$$\rho(x_1, \dots, x_N) = \frac{1}{\mathcal{Z}_{N,\beta}} e^{-\frac{1}{2} \sum_{i=1}^N x_i^2} \prod_{j < k} |x_j - x_k|^\beta \quad (1.3)$$

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,  $\beta$  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.

### 1.4 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

<sup>1</sup>Esse resultado é não óbvio e deve ser discutido em breve.

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

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  $\hat{H} \implies (\hat{U} \in U(N), \hat{\Lambda})$  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  $(\phi_1, \dots, \phi_N)$ . Para restringir nosso problema e tornar a função injetiva será necessário considerar as matrizes unitárias ao espaço de coset  $U(N)/U(1) \times \dots \times U(1)$ <sup>2</sup>. Uma outra restrição necessária 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 [1]. Em especial recuperaremos o elemento de distância e volume no subespaço que vamos tratar

$$(ds)^2 = \text{Tr } d\hat{H} d\hat{H}^* = \sum_i (dx_{ii})^2 + 2 \sum_{i < j} [(dx_{ij})^2 + (dy_{ij})^2] \quad (1.4)$$

$$d\mu(\hat{H}) = 2^{\frac{N(N-1)}{2}} \prod_i dx_{ii} \prod_{i < j} dx_{ij} dy_{ij} \quad (1.5)$$

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 (1.4) e (1.5) na decomposição espectral, obteremos

$$(ds)^2 = \sum_i (d\lambda)^2 + \sum_{i < j} (\lambda_i - \lambda_j)^2 \overline{\delta U_{ij}} \delta U_{ij} \quad (1.6)$$

e

$$d\mu(\hat{H}) = \prod_{i < j} (\lambda_i - \lambda_j)^2 \prod_i d\lambda_i \times d\mu(\hat{U}) \quad (1.7)$$

Tendo a medida de integração pronta, podemos definir uma F.D.P  $\mathcal{P}(\hat{H})$  neste espaço de matrizes 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

$$\mathcal{P}(\hat{H}) = C e^{-\text{Tr } Q(\hat{H})} \quad (1.8)$$

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

---

<sup>2</sup>Não 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.

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

Nossa medida tomará a forma

$$\mathcal{P}(\hat{H}) = e^{-a[\sum_i x_{ii}^2 + 2\sum_{i<j} [x_{ij}^2 + y_{ij}^2]]} e^{-b\sum_i x_{ii}} e^{-cN} \quad (1.9)$$

$$= e^{-cN} \prod_{i=1}^N \left( e^{-ax_{ii}^2 - bx_{ii}} \right) \prod_{i<j} e^{-2ax_{ij}^2} \prod_{i<j} e^{-2ay_{ij}^2} \quad (1.10)$$

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  $\mathcal{P} = Ce^{-(a \text{Tr } \hat{H}^2 + b \text{Tr } \hat{H} + cN)}$  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

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

Dado  $\lambda_j$  os autovalores

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

e finalmente podemos escrever

$$E[f] = \int_{\mathcal{H}_n} f(\hat{M}) e^{-\text{Tr } (Q(\hat{M}))} d\hat{M} \quad (1.11)$$

$$= \frac{1}{\tilde{\mathcal{Z}}} \int \cdots \int f(\lambda_1, \dots, \lambda_n) \prod_{i<j} (\lambda_i - \lambda_j)^2 \prod_{j=1}^n e^{-Q(\lambda_j)} d\lambda_1 \dots d\lambda_n \quad (1.12)$$

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

$$\frac{1}{\tilde{\mathcal{Z}}_n} \prod_{i<j} (\lambda_i - \lambda_j)^2 \prod_{j=1}^n e^{Q(\lambda_j)} \quad (1.13)$$

Alguns resultados foram resgatadas da nota do autor em [2].

## Capítulo 2

# Movimento Browniano

### 2.1 Processo Pontual

Um processo pontual pode ser interpretado como um conjunto aleatório de pontos ou como a medida de probabilidade associada a esse conjunto. Um processo pontual possui  $n$  pontos se

$$\mathcal{P}(\#X = n) = 1$$

Onde  $X$  é um conjunto enumerável de  $\mathcal{X}$  ( $\mathbb{R}$ ,  $\mathbb{Z}$  ou um subconjunto destes). O conjunto de todas configurações possíveis é denominado  $Conf(\mathcal{X})$ . Se  $P(x_1, \dots, x_n)$  é uma função de densidade de probabilidade em  $\mathbb{R}^n$  invariante por permutações

$$\mathbb{R}^n \rightarrow Conf(\mathbb{R})$$

$$(x_1, \dots, x_n) \mapsto X = x_1, \dots, x_n$$

define naturalmente um processo pontual com  $n$  pontos.

#### 2.1.1 Poisson & fries

Tome  $\{N(t)\}$  o número de eventos no intervalo de tempo  $]0, t]$ .  $\{N(t)\}$  é um processo estocástico (de contagem). Se o processo de Poisson possui  $\lambda > 0$ , para um elemento fixo do espaço amostral a variável aleatória  $N$  assume valor  $k$  no tempo  $t$  com probabilidade

$$\mathcal{P}[N(t) = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \quad (2.1)$$

Onde  $\lambda$  é o número esperado de chegadas por unidade de tempo. Agora, como um processo pontual, a probabilidade de  $n$  eventos no intervalo  $]a, b]$  é

$$\mathcal{P}(N]a, b] = n) = \frac{(\lambda(b-a))^n e^{-\lambda(b-a)}}{n!}$$

Podemos usar a independência de cada evento de Poisson em intervalos disjuntos para escrever

$$\mathcal{P}(N]a_1, b_1] = n_1, \dots, N]a_k, b_k] = n_k) = \prod_{i=1}^k \frac{(\lambda(b_i - a_i))^{n_i} e^{-\lambda(b_i - a_i)}}{n_i!}$$

Podemos escrever para uma função  $f$  mensurável em  $\mathbb{R}$

$$\sum_{x_i \in \mathcal{X}} f(x_i) = \int f(x) dN(x)$$

Onde a medida  $dN$  é

$$dN(x) = \sum_{x_i \in \mathcal{X}} \delta_{x_i}(x)$$

Onde notamos que podemos interpretar tanto quanto uma soma de um processo pontual quanto uma medida de probabilidade.

### 2.1.2 Função Correlação

Definimos uma variável aleatória  $N$  anteriormente. Naturalmente, poderíamos estar interessados em sua esperança. Mais especificamente, podemos procurar a esperança do número de pontos de uma configuração dentro de um intervalo  $A \subset \mathbb{R}$ .

$$A \mapsto \mathbb{E}[N(A)] = \mathbb{E}[\#(A \cap X)]$$

Que pode ser interpretada como uma medida com densidade  $p_1$

$$\mathbb{E}[\#(A \cap X)] = \int_A p_1(x) dx \quad (2.2)$$

A equação 2.2 é conhecida como *função de correlação de 1 ponto*. Em grosso modo,  $p_1(x)$  é a probabilidade de haver um ponto da configuração entre  $x$  e  $x + dx$ . Seja um configuração simples  $X = \{x_1, \dots, x_n\}$  e intervalos disjuntos na reta  $A_1, A_2, \dots, A_n$ ,

$$\int_A \dots \int_A \rho_n(x_1, \dots, x_n) dx_1, \dots, dx_n = \mathbb{E} \left( \prod_{j=1}^k \#(X \cap A_j) \right)$$

é o número esperado de  $n$ -uplas  $(x_1, x_2, \dots, x_n) \in A_1 \times \dots \times A_n$  tais que  $x_i \in A_i, i = 1, \dots, n$ . Seja  $\mathbb{P}(x_1, \dots, x_n)$  uma densidade de probabilidade em  $\mathbb{R}^n$ , então o processo pontual de  $n$  pontos gerado possui funções de correlação dadas por

$$p_k(x_1, \dots, x_k) = \frac{n!}{(n-k)!} \int \dots \int \mathbb{P}(x_1, \dots, x_n) dx_{k+1} \dots dx_n$$

### 2.1.3 Pontual Determinante

Um processo pontual vai ser chamado determinante se dada uma função de correlação  $\rho_n$ , existe um núcleo  $K(x, y)$  conhecido como núcleo de correlação tal que

$$\rho_n(x_1, \dots, x_n) = \det[K(x_i, x_j)]_{i,j=1}^n \quad (2.3)$$

onde

$$[K(x_i, x_j)]_{i,j=1}^n = \begin{bmatrix} K(x_1, x_1) & K(x_1, x_2) & \dots & K(x_1, x_n) \\ K(x_2, x_1) & K(x_2, x_2) & \dots & K(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ K(x_n, x_1) & K(x_n, x_2) & \dots & K(x_n, x_n) \end{bmatrix}$$

Nos casos de baixa dimensão

$$p_1(x_1) = K(x_1, x_1), \quad p_2(x_1, x_2) = \begin{vmatrix} K(x_1, x_1) & K(x_1, x_2) \\ K(x_2, x_1) & K(x_2, x_2) \end{vmatrix}$$

Para que o núcleo satisfaça a equação 2.3 enunciaremos um resultado de interesse.

**Teorema 1** *Seja  $K$  um núcleo tal que*

- (a)  $\int K(x, x) = n \in \mathbb{N}$ ,
- (b) Para todo  $x_1, x_2, \dots, x_n \in \mathbb{R}$ , o determinante é não negativo
- (c)  $K$  possui a propriedade de **núcleo reprodutor**, isto é;

$$K(x, y) = \int_{-\infty}^{\infty} K(x, s)K(s, y)ds$$

Então

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

será uma densidade de probabilidade em  $\mathbb{R}$  cujo processo de  $n$  pontos associado é determinante.

## 2.2 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

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

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.

## 2.3 Karlin-McGregor

### 2.3.1 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 [2]. Note que temos

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

$$= \int_{E_1} \dots \int_{E_n} \begin{vmatrix} p_t(a_1, x_1) & p_t(a_2, x_1) & \dots & p_t(a_{n-1}, x_1) & p_t(a_n, x_1) \\ p_t(a_1, x_2) & p_t(a_2, x_2) & \dots & p_t(a_{n-1}, x_2) & p_t(a_n, x_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_t(a_1, x_{n-1}) & p_t(a_2, x_{n-1}) & \dots & p_t(a_{n-1}, x_{n-1}) & p_t(a_n, x_{n-1}) \\ p_t(a_1, x_n) & p_t(a_2, x_n) & \dots & p_t(a_{n-1}, x_n) & p_t(a_n, x_n) \end{vmatrix} dx_1 \dots dx_n \quad (2.4)$$

$$= \sum_{\sigma} \text{sgn}(\sigma) \prod_{j=1}^n p_t(a_j, E_{\sigma(j)}) \quad (2.5)$$

$$= \sum_{\sigma} \text{sgn}(\sigma) \mathcal{P}(A_{\sigma}) \quad (2.6)$$

Onde denotamos

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

$\sigma$  é uma permutação de  $1, \dots, n$  e  $A_{\sigma}$  é o evento que  $X_j(t) \in E_{\sigma(j)}$  para todo  $j$ . Os caminhos devem ser independentes para (2.6).

De alguma forma o determinada 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

$$\begin{vmatrix} p_t(a_1, x_1) & p_t(a_2, x_1) & p_t(a_3, x_1) \\ p_t(a_1, x_2) & p_t(a_2, x_2) & p_t(a_3, x_2) \\ p_t(a_1, x_3) & p_t(a_2, x_3) & p_t(a_3, x_3) \end{vmatrix} = \quad (2.7)$$

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

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

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

$$- p_t(a_3, x_1) p_t(a_2, x_2) p_t(a_1, x_3) \quad (2.11)$$

$$- p_t(a_2, x_1) p_t(a_1, x_2) p_t(a_3, x_3) \quad (2.12)$$

$$- p_t(a_1, x_1) p_t(a_3, x_2) p_t(a_2, x_3) \quad (2.13)$$

Logo

$$\int_{E_1} \cdots \int_{E_n} \det [p_t(a_i, x_j)]_{i,j=1}^n dx_1 \dots dx_n = + p_t(a_1, E_1) p_t(a_2, E_2) p_t(a_3, E_3) \quad (2.14)$$

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

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

$$- p_t(a_3, E_1) p_t(a_2, E_2) p_t(a_1, E_3) \quad (2.17)$$

$$- p_t(a_2, E_1) p_t(a_1, E_2) p_t(a_3, E_3) \quad (2.18)$$

$$- p_t(a_1, E_1) p_t(a_3, E_2) p_t(a_2, E_3) \quad (2.19)$$

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

### 2.3.2 Consequências

Considere  $n$  cópias do processo de Markov condicionado para começar em  $t = 0$  nas determinadas posições  $a_1 < a_2 < \dots < a_n$ . 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  $a_i \rightarrow a$  e  $b_j \rightarrow b$  ficamos, por consequência que quando mais de um  $a_i$  ou  $b_j$  convergem ao mesmo ponto teremos  $\mathcal{Z} \rightarrow 0$  pela sua dependência destes termos (eles não serão mais 'ortogonais', pontos distintos). Aplicando L'Hôpital

$$\det \left[ \frac{\partial^{i-1}}{\partial a^{i-1}} p_t(a_i, x_j) \right]_{i,j=1}^n$$

Onde

$$\frac{\partial^{i-1}}{\partial a^{i-1}} p_t(a_i, x_j) = \left[ \left( \frac{x_j - a_i}{t} \right)^{i-1} + O(x^{i-2}) \right] p_t(a_i, x_j)$$



Que nos permite afirmar que o determinante é proporcional à

$$\det [x_j^{i-1}]_{i,j=1}^n \prod_{j=1}^n e^{-\frac{x_j^2 - 2ax_j}{2t}}$$

Analogamente para o outro determinante

$$\det [x_j^{i-1}]_{i,j=1}^n \prod_{j=1}^n e^{-\frac{x_j^2 - 2bx_j}{2(T-t)}}$$

Ou seja;

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

onde  $F$  e  $G$  são polinômios em  $x$  de grau  $j-1$ . Tomaremos agora o limite em que  $a, b \rightarrow 0$ . Reescrevemos a distribuição conjunta destes pontos

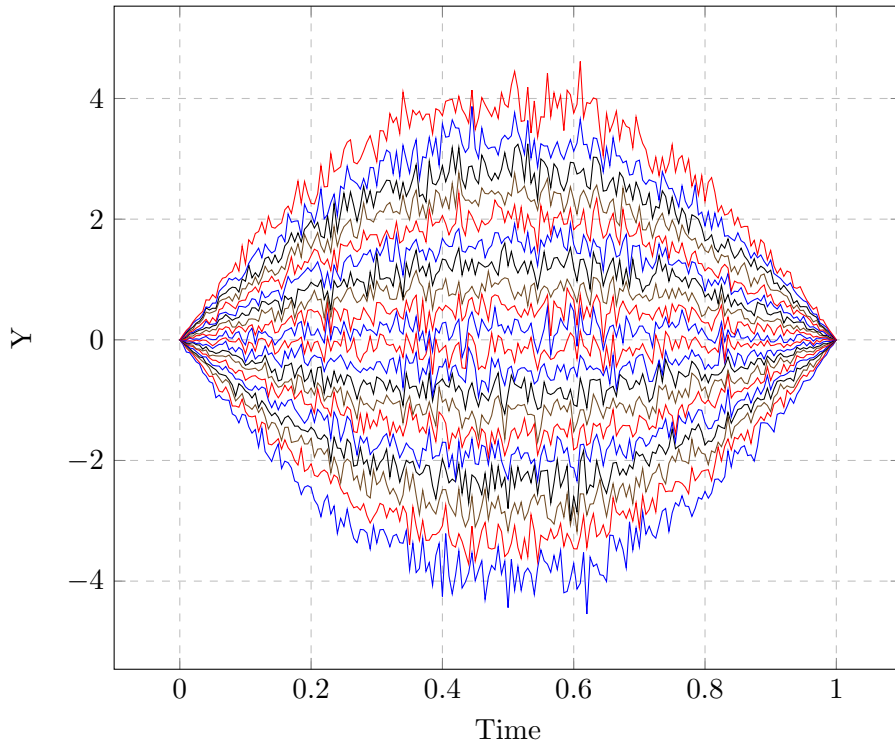
$$P(x_1, \dots, x_n) = \frac{1}{Z_n} \det [p_t(a_i, x_j)]_{i,j=1}^n \det [p_{T-t}(x_i, b_j)]_{i,j=1}^n \quad (2.20)$$

$$= \frac{1}{Z_n} \det [x_j^{i-1}]_{i,j=1}^n \prod_{j=1}^n e^{-\frac{x_j^2}{2t}} \det [x_j^{i-1}]_{i,j=1}^n \prod_{j=1}^n e^{-\frac{x_j^2}{2(T-t)}} \quad (2.21)$$

$$= \frac{1}{Z_n} \det [x_j^{i-1}]_{i,j=1}^n \det [x_j^{i-1}]_{i,j=1}^n \prod_{j=1}^n e^{-\frac{Tx_j^2}{2t(T-t)}} \quad (2.22)$$

Relembrando a equação (1.13), podemos interpretar esta distribuição como uma distribuição de autovalores em um espaço de matrizes hermitianas do GUE com entradas de variância  $\sigma^2 = \frac{2t(T-t)}{T}$  e média  $\mu = 0$ . De forma que poderemos simular a evolução destes movimentos brownianos com o sistema de matrizes descrito.

Dyson Bridge



---

```
1  import numpy as np
2
3  n_particles = 20
4  steps = 1000
5
6  def GUE(N, sigma = 1, mu = 0, escale = False):
7      beta = 2
8      A = (np.random.randn(N,N) * sigma + mu)+ 1j*(np.random.randn(N,N) * sigma + mu)
9      A = (A + A.T.conj())/2
10     eig = np.linalg.eigvalsh(A)
11     if escale:
12         return eig/(np.sqrt(N*beta))
13     return eig
14
15     eig_GUE = np.zeros(n_particles)
16
17     final_time = 1
18     dt = final_time/steps
19     Memory = np.zeros((steps+1,n_particles))
20
21     for step in range(steps+1):
22         partial_time = step * dt
23         sigma = 2*partial_time*(final_time-partial_time)/final_time
24         mu = 0
25         eig_GUE = GUE(n_particles, sigma = sigma, mu = mu, escale = False)
26         eig_order = np.argsort(eig_GUE)
27         Memory[step,:] = eig_GUE[eig_order]
28
29     np.savetxt('Memory.txt', Memory)
```

---

# Bibliografia

- [1] Yan V. Fyodorov. Introduction to the random matrix theory: Gaussian unitary ensemble and beyond, 2010.
- [2] Arno Kuijlaars. Lecture notes on riemann-hilbert problems and multiple orthogonal polynomials.
- [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

# Transformação Legendre

Transformações de Legendre tem ampla aplicação e interpretação. Faremos uma breve dissertação de duas possíveis visualizações do processo.

### A.1 Tangentes

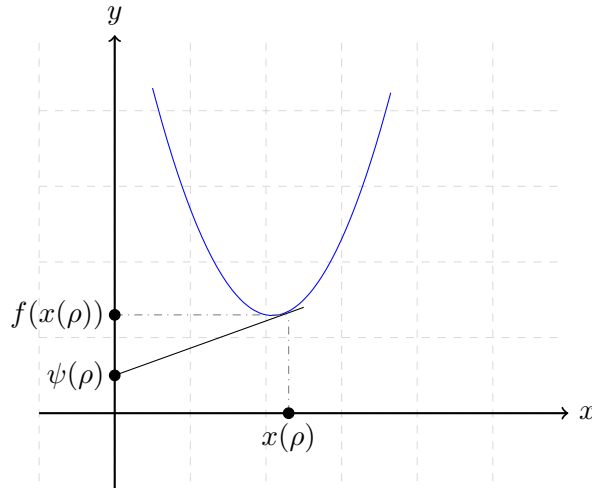
Uma primeira interpretação do processo representa uma mudança clara de variável a partir de tangentes de uma função de concavidade bem definida. Faremos

$$y = f(x) \mapsto \psi(\rho) = f(x(\rho)) - \rho x(\rho)$$

onde

$$\rho \equiv \frac{y - f(x(\rho))}{x - x(\rho)} = \frac{\psi(\rho) - f(x(\rho))}{0 - x(\rho)}$$

Que podemos visualizar



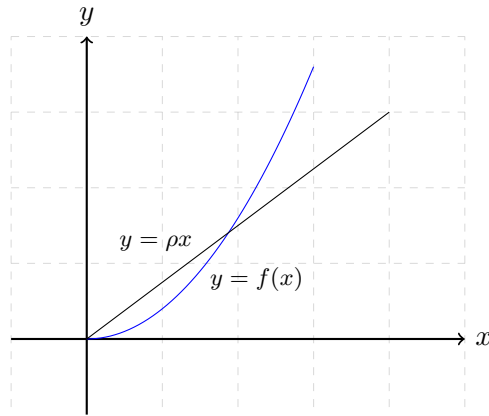
De forma que a transformada é criada pela projeção desta tangente no eixo  $x=0$ .

### A.2 Otimização

Outra forma de visualizar a transformada é por um problema mais conveniente de otimização. Definiremos

$$\psi(\rho) = \max_x [\rho x - f(x)]$$

Onde teremos



Se definirmos

$$g(x) = \rho x - f(x)$$

Minimizaremos  $g(x)$  e teremos a condição  $f'(x^*) = \rho$ . Ou seja

$$f'(x(\rho)) = \rho \implies \max_x [\rho x - f(x)] = \rho x(\rho) - f(x(\rho))$$

Que é equivalente à dizer

$$\psi(\rho) = \min_x [f(x) - \rho x] = f(x(\rho)) - \rho x(\rho)$$

## Apêndice B

### Soma Assintóticas

Existe uma aproximação a se fazer no logaritmo de somas assintóticas que pode ser de interesse. Assuma

$$S = \sum_{i=1}^M a_i$$

Se  $a_i(N) \sim e^{\phi_i N}$  ou  $\log(a_i) \approx \phi_i N$ , podemos afirmar

$$\log(S) \sim \log(a_{max})$$

Para demonstrar isso, notamos

$$\begin{aligned} a_{max} < S < M a_{max} \\ \frac{\log(a_{max})}{N} < \frac{\log S}{N} < \frac{\log(a_{max})}{N} + \frac{\log M}{N} \end{aligned}$$

Ou seja, desde que  $\frac{\log M}{N} \rightarrow 0$ . Isto ocorrerá desde que  $M$  seja sub-exponencial. Contudo **NOTE** que dizer

$$\log(n!) \sim n \log n - n$$

Não implica que

$$n! \sim \left(\frac{n}{e}\right)^n$$

Em algumas situações é possível afirmar contudo

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$