



Daniel Byron Souza Pereira de Andrade

Matrizes Aleatórias e a Lei do Semicírculo

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Matemática, do Departamento de Matemática da PUC-Rio .

Orientador : Prof. Simon Griffiths
Co-orientador: Prof. Thiago Barbosa dos Santos Guerreiro

Rio de Janeiro
Abril de 2022



Daniel Byron Souza Pereira de Andrade

Matrizes Aleatórias e a Lei do Semicírculo

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Matemática da PUC-Rio . Aprovada pela Comissão Examinadora abaixo:

Prof. Simon Griffiths

Orientador

Departamento de Matemática – PUC-Rio

Prof. Thiago Barbosa dos Santos Guerreiro

Co-orientador

Departamento de Física – PUC-Rio

Prof. Nicolau Corção Saldanha

Departamento de Matemática – PUC-Rio

Prof. Gregorio Malajovich Munoz

Departamento de Matemática Aplicada – UFRJ

Prof. Roberto Imbuzeiro Moraes Felinto de Oliveira

Instituto de Matemática Pura e Aplicada – IMPA

Rio de Janeiro, 28 de abril de 2022

Todos os direitos reservados. A reprodução, total ou parcial do trabalho, é proibida sem a autorização da universidade, do autor e do orientador.

Daniel Byron Souza Pereira de Andrade

Bacharel em Matemática e Física pela PUC-Rio.

Ficha Catalográfica

Andrade, Byron

Matrizes Aleatórias e a Lei do Semicírculo / Daniel Byron Souza Pereira de Andrade; orientador: Simon Griffiths; co-orientador: Thiago Barbosa dos Santos Guerreiro. – 2022.

51 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Matemática, 2022.

Inclui bibliografia

1. Matemática – Teses. 2. Matrizes Aleatórias. 3. Teorema de Wigner. 4. Lei do semicírculo. I. Griffiths, Simon. II. Guerreiro, Thiago. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Matemática. IV. Título.

CDD: 510

Agradecimentos

Aos meus orientadores Professor Simon Griffiths e o Professor Thiago Guerreiro pelo estímulo e parceria para a realização deste trabalho.

Ao CNPq e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

À matemática Luíze D'Urso, que contribuiu com o total e irrestrito apoio.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Andrade, Byron; Griffiths, Simon; Guerreiro, Thiago. **Matrizes Aleatórias e a Lei do Semicírculo**. Rio de Janeiro, 2022. 51p. Dissertação de Mestrado – Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

Nessa dissertação vamos abordar a famosa "lei do Semicírculo de Wigner", que dá uma descrição do comportamento do espectro de autovalores de matrizes aleatórias simétricas. A demonstração combina ideias e técnicas de Combinatória e Probabilidade, incluindo uma análise cautelosa dos momentos da distribuição de autovalores.

Palavras-chave

Matrizes Aleatórias; Teorema de Wigner; Lei do semicírculo.

Abstract

Andrade, Byron; Griffiths, Simon (Advisor); Guerreiro, Thiago (Co-Advisor). **Random Matrices and the Semicircle Law**. Rio de Janeiro, 2022. 51p. Dissertação de Mestrado – Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

In this dissertation we will approach the famous "Wigner's Semicircle Law", which gives a description of the behavior of the eigenvalue spectrum of symmetric random matrices. The proof combines ideas and techniques from Combinatorics and Probability, including a careful analysis of the moments of the eigenvalue distribution.

Keywords

Random Matrices; Wigner's Theorem; Semicircle law.

Sumário

1	Introdução	10
2	O Teorema de Wigner em três versões	15
2.1	Definindo a família de matrizes de Wigner	15
2.2	Enunciados do Teorema	17
2.3	Passos da prova	19
3	Prova da Afirmação 1	20
3.1	Abrindo o Traço	20
3.2	Interpretando o índice I como um caminho em um grafo	22
3.3	Calculando grafos	27
3.4	Caminhos de Dyck	30
4	Prova da Afirmação 2	32
4.1	A variância em termos de grafos	32
4.2	Cotando a variância	36
5	Conclusões	40
5.1	Lei empírica dos autovalores	40
5.2	Prova do Teorema 2.2.2	41
5.3	Prova do Teorema 2.2.1	43
6	Apêndice	45
6.1	Norma de Frobenius	45
6.2	Sequência de Catalan	47
6.3	Código em Python	49
7	Referências bibliográficas	51

*Pure mathematics is, in its way, the poetry
of logical ideas.*

Albert Einstein, *Obituary for Emmy Noether.*

1

Introdução

A ideia original de usar a teoria das Matrizes Aleatórias, que trata das matrizes cujas entradas são variáveis aleatórias, vem de um famoso problema de Física (WIGNER, 1955). Eugene Wigner utilizou da teoria de Matrizes Aleatórias como um modelo estatístico e reinterpretou o modelo no contexto da Física Nuclear. Isso historicamente foi um marco para a área da Física Matemática (TAO, 2010; AKEMANN; BAIK; FRANCESCO, 2011), a partir de então seria possível usar Matrizes Aleatórias para ajudar como ferramenta estatística nos mais variados modelos físicos, desde que interpretados como elementos de uma sequência de matrizes de Wigner cujas dimensões tendem a infinito.

Empiricamente foi constatado que uma sequência de matrizes aleatórias com certas condições bastante razoáveis (cf. Definição 2.1.1) tem para cada elemento uma distribuição de autovalores, e que quanto mais avançado for o termo na sequência, mais sua distribuição se assemelhará à função de distribuição de um semicírculo. A formalização dessa constatação é o que entendemos por "lei do semicírculo de Wigner" ou Teorema de Wigner.

O objetivo desta dissertação é estudar uma demonstração para o Teorema, da área de Matrizes Aleatórias. O Teorema em questão é não só um dos mais característicos, mas também um dos que possui resultado mais significativo, prevendo o comportamento da distribuição dos autovalores das matrizes aleatórias.

Este Teorema é clássico na literatura de Matrizes Aleatórias e tem algumas demonstrações. Nessa dissertação teremos como objeto de estudo uma demonstração longa e ao mesmo tempo elementar. Uma prova essencialmente combinatorial, isso é, ferramentas que envolvem grafos e combinatória básica serão encontradas no texto, que visa dissertar e familiarizar o leitor com o resultado do Teorema. A demonstração apresentada envolve teoremas simples e desigualdades como as de Markov e Chebyshev e chegaremos a uma contagem de grafos.

Esse trabalho busca um entendimento bem básico, um ponto de partida, para que em algum momento o leitor possa se sentir confortável em ler trabalhos que se utilizem das matrizes de Wigner como em modelos físicos ou em qualquer outro lugar (TAO, 2010).

A importância do Teorema é evidente em várias áreas da Matemática e da Física, ao tentar estudar matrizes cujas entradas expressam funções complexas

demais para uma abordagem analítica. Nesse tipo de problema, em que existe uma dada matriz \mathbf{X} complicada, o ferramental de Matrizes Aleatórias oferece matrizes, que podem ter alguns comportamentos estatísticos semelhantes aos da matriz \mathbf{X} , mas que era demasiadamente difícil para uma análise direta.

A seguir, vamos assumir a existência de uma sequência \mathbf{X}_n de matrizes aleatórias para que possamos analisar alguns histogramas de autovalores (gráficos da quantidade de autovalores por intervalo da reta).

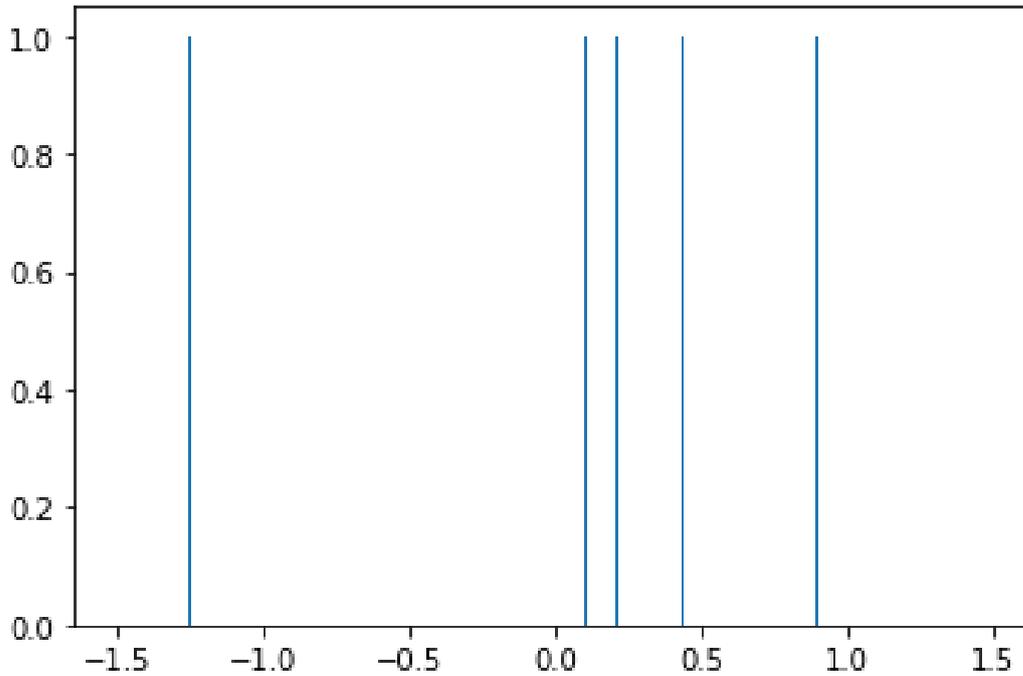


Figura 1.1: Matriz com distribuição Gaussiana, $\mathbf{X}_{n=5}$, intervalos de tamanho 0.01 da reta. Gerado com Colab Google (cf, Seção 6.3)

Conforme avançamos na sequência, vemos que os histogramas parecem convergir para uma distribuição do semicírculo.

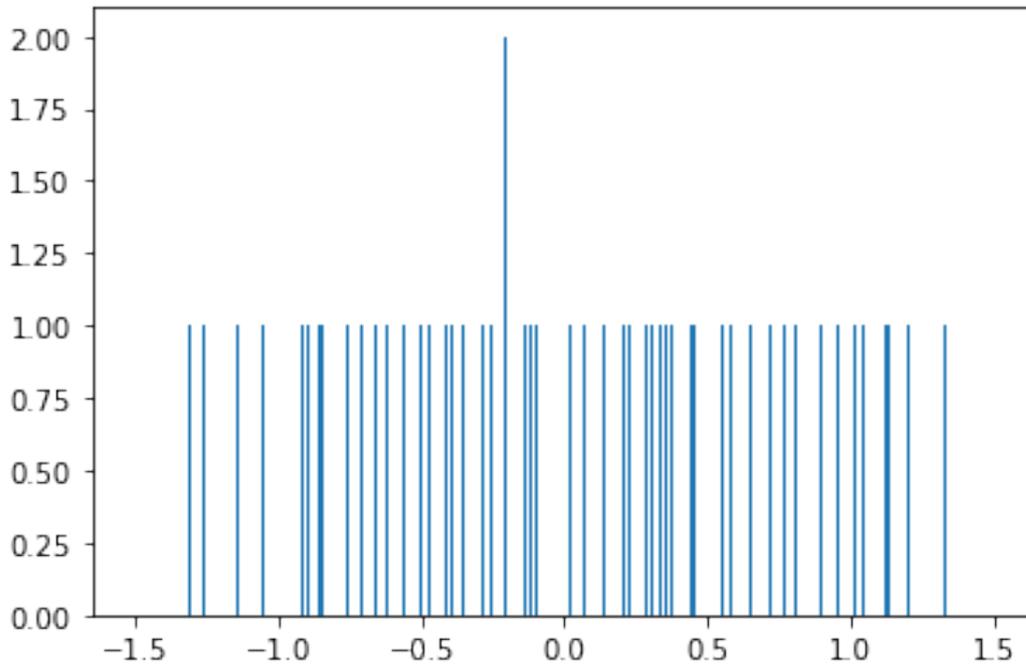


Figura 1.2: Matrizes com distribuição Gaussiana, $\mathbf{X}_{n=50}$, intervalos de tamanho 0.01 da reta. Gerado com Colab Google (cf, Seção 6.3)

PUC-Rio - Certificação Digital Nº 1920825/CA

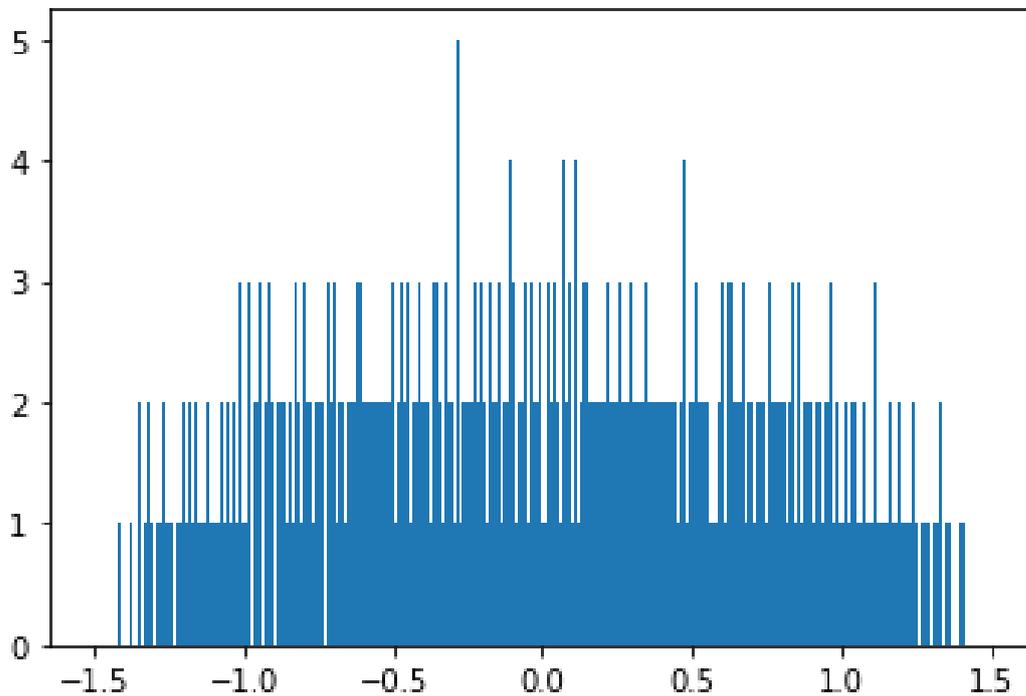


Figura 1.3: Matrizes com distribuição Gaussiana, $\mathbf{X}_{n=500}$, intervalos de tamanho 0.01 da reta. Gerado com Colab Google (cf, Seção 6.3)

Os histogramas, são referentes as matrizes com dimensões $n = 5$, $n = 50$, $n = 500$ e por fim, onde já é praticamente evidente a forma do semi-circulo

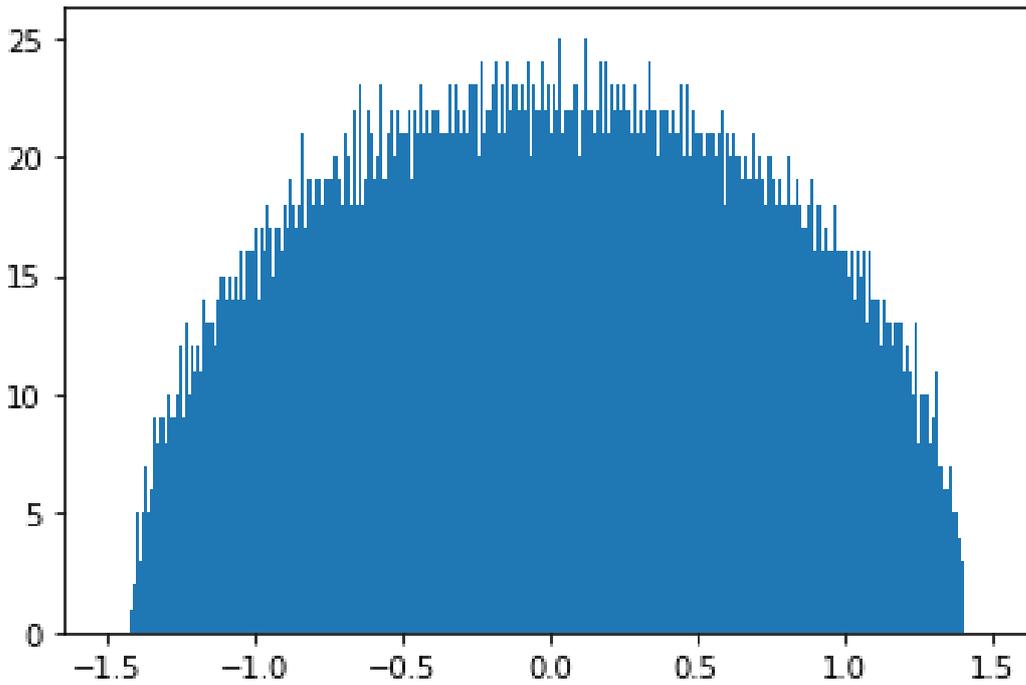


Figura 1.4: Matrizes com distribuição Gaussiana, $\mathbf{X}_{n=5000}$, intervalos de tamanho 0.01 da reta. Gerado com Colab Google (cf, Seção 6.3)

quando olhamos para $n = 5000$.

Antes de seguirmos para a descrição dos capítulos do texto, vamos dar uma breve ideia da alma da prova para o teorema 2.2.3 que vamos estudar:

Imagine que temos uma matriz aleatória simétrica $n \times n$. Suponha que as variáveis aleatórias $\{Y_{ij}\}_{1 \leq i < j \leq n}$ são i.i.d. e $\{Y_{ii}\}_{1 \leq i \leq n}$ também são i.i.d., todas elas independentes entre si. Por simetria, quando $i > j$ temos $Y_{ij} = Y_{ji}$. Em determinado momento, estaremos interessados em calcular a esperança

$$\mathbb{E}(Y_{i_1 i_2} Y_{i_2 i_3} \cdots Y_{i_k i_1})$$

para qualquer k -upla $(i_1, i_2, \dots, i_k) \in \{1, 2, \dots, n\}^k$. Note que para isso queremos identificar as variáveis que aparecem repetidamente e separar no produto de esperanças da forma $\mathbb{E}(Y_{ij}^m)$, já que variáveis diferentes são independentes.

Vamos ver que todas as informações que precisamos para calcular a esperança estão codificadas por um par (G, w) , onde $G = G(i_1, i_2, \dots, i_k)$ é um grafo, e $w = w(i_1, i_2, \dots, i_k)$ um circuito nesse grafo que percorre todas as suas arestas. A grosso modo, G consiste dos vértices i_l , e w é o circuito que percorre $i_1 \mapsto i_2 \mapsto \dots \mapsto i_k \mapsto i_1$.

Ou seja, duas variáveis aleatórias Y_{ab} e Y_{cd} são iguais (i.e. $\{a, b\} = \{c, d\}$ como conjuntos) quando suas correspondentes arestas em G são a mesma

aresta. Mais ainda, a quantidade de vezes que esta variável aparece na esperança (que corresponderia ao expoente m do termo $\mathbb{E}(Y_{ij}^m)$) é determinada pelo número de vezes que o circuito w passa pela sua correspondente aresta.

Além disso, precisamos diferenciar os termos Y_{ij} em que $i = j$ e os em que $i \neq j$, já que temos duas distribuições distintas. E nosso grafo diferencia tais arestas como laços e arestas conectoras, respectivamente.

A partir dessa dualidade, nosso problema se transforma em uma constante interpretação de $\mathbb{E}(Y_{i_1 i_2} Y_{i_2 i_3} \cdots Y_{i_k i_1})$ em função das duplas (G, w) e vamos resolvê-lo contando grafos e caminhos com as propriedades que desejamos.

A seguir está uma descrição dos seguintes capítulos desta dissertação:

Esse trabalho começa no Capítulo 2 pela definição das matrizes de Wigner e segue para uma explicação e algumas diferentes versões do Teorema de Wigner. Nesse Capítulo vamos reduzir a prova do Teorema a duas afirmações, que serão provadas nos dos capítulos seguintes.

No Capítulo 3 veremos a demonstração da Afirmação 1, que consiste na convergência da Esperança do traço de matrizes de Wigner.

No Capítulo 4 veremos a demonstração da Afirmação 2, que trata da convergência da Variância da mesma variável aleatória.

Finalmente, no Capítulo 5 veremos por que as diferentes versões do Teorema são equivalentes.

Esta dissertação também conta com um Apêndice para esclarecer assuntos que não convém ao corpo do texto.

2

O Teorema de Wigner em três versões

2.1

Definindo a família de matrizes de Wigner

Vamos falar de uma família de matrizes aleatórias, em particular a família que estamos interessados é a família das matrizes de Wigner. Para falarmos o que é isso vamos tentar uma construção intuitiva.

Comece com uma família infinita de variáveis aleatórias Y_{ij} com $j \geq i \geq 1$. Com essas variáveis, vamos construir as matrizes simétricas \mathbf{Y}_n . Onde as entradas dessa matriz são exatamente as variáveis acima (Y_{ij}).

$$\begin{array}{cccc} Y_{11} & Y_{12} & Y_{13} & \cdots \\ & Y_{22} & Y_{23} & \cdots \\ & & Y_{33} & \cdots \\ & & & \ddots \end{array}$$

$$\mathbf{Y}_1$$

$$\begin{array}{cccc} [Y_{11}] & Y_{12} & Y_{13} & \cdots \\ & Y_{22} & Y_{23} & \cdots \\ & & Y_{33} & \cdots \\ & & & \ddots \end{array}$$

$$\mathbf{Y}_2$$

$$\begin{array}{ccc} \left[\begin{array}{cc} Y_{11} & Y_{12} \\ Y_{12} & Y_{22} \end{array} \right] & \begin{array}{c} Y_{13} \\ Y_{23} \\ Y_{33} \end{array} & \begin{array}{c} \cdots \\ \cdots \\ \cdots \\ \ddots \end{array} \end{array}$$

$$\mathbf{Y}_3$$

$$\left[\begin{array}{ccc} Y_{11} & Y_{12} & Y_{13} \\ Y_{12} & Y_{22} & Y_{23} \\ Y_{13} & Y_{23} & Y_{33} \end{array} \right] \cdots$$

$$\ddots$$

$$\mathbf{Y}_n$$

$$\begin{bmatrix} Y_{11} & & Y_{ji} \\ & \ddots & \\ Y_{ij} & & Y_{nn} \end{bmatrix}.$$

Como todas essas matrizes são simétricas por construção, seus autovalores são reais e nesse caso, podemos ordená-los para cada matriz \mathbf{Y}_n .

Antes de definir as matrizes de Wigner, vamos mencionar algumas propriedades que estamos interessados:

1. Vamos assumir que as variáveis aleatórias que usamos nas entradas da matriz são independentes da diagonal para cima.
2. Vamos assumir que temos duas distribuições somente, uma para as variáveis aleatórias na diagonal, e uma para as acima dela. Ambas com média 0.
3. Vamos assumir que o segundo momento de qualquer uma das entradas é finito e igual. Isto é $r_2 := \max \{ \mathbb{E} (Y_{11}^2), \mathbb{E} (Y_{12}^2) \} < \infty$.

Existe uma versão mais forte da última condição:

4. Vamos assumir que o k -ésimo momento de qualquer uma das entradas é finito, para todo k . Isto é $r_k := \max \{ \mathbb{E} (Y_{11}^k), \mathbb{E} (Y_{12}^k) \} < \infty$.

Vai ser útil definir matrizes de Wigner já re-escaladas de tal maneira que os autovalores estão na escala apropriada. Em algum sentido, precisamos ir re-escalando os autovalores e \mathbf{Y}_n (para que cada X_n seja uma matriz de Wigner 2.1.1 e em particular, com auto valores "controlados"). Para isso vamos utilizar a norma de Frobenius (vide Seção 6.1).

Considere $\mathbf{X}_n := n^{-\frac{1}{2}} \mathbf{Y}_n$; onde \mathbf{Y}_n é a família que acabamos de construir que podem ser entendidas como as matrizes de Wigner ainda não normalizadas.

Observação 1 *Os autovalores de \mathbf{Y}_n , de acordo com as características que pedimos dela (a matriz), podem ir para o infinito.*

Isso fará sentido no futuro pois provaremos que os autovalores convergem para uma função com valores comportados (distribuição do semicírculo). Então começar controlando os auto valores e não deixando eles irem ao infinito, será importante.

Definição 2.1.1 (Matriz de Wigner) *Seja \mathbf{Y}_n uma matriz simétrica que satisfaz as condições 1,2 e 3 com entradas $\{Y_{ij}\}_{1 \leq i,j \leq n}$, e suponha que $\mathbb{E} (Y_{12}^2) > 0$. Definimos assim $\mathbf{X}_n = n^{-\frac{1}{2}} \mathbf{Y}_n$ como as matrizes de Wigner.*

Observação 2 Pedimos que o segundo momento fora da diagonal seja maior que zero para evitar uma situação degenerada, que seria uma matriz diagonal.

Para as provas será útil usar a condição mais forte em relação aos momentos das entradas. Por isso, acrescentamos a seguinte definição.

Definição 2.1.2 (Matriz de Wigner forte) *Seja \mathbf{Y}_n uma matriz simétrica que satisfaz condições 1, 2 e 4 com entradas $\{Y_{ij}\}_{1 \leq i, j \leq n}$, e suponha que $\mathbb{E}(Y_{12}^2) > 0$. Definimos assim $\mathbf{X}_n = n^{-\frac{1}{2}}\mathbf{Y}_n$ como as matrizes de Wigner forte.*

2.2

Enunciados do Teorema

Nessa seção introduziremos o resultado principal – a lei do semicírculo de Wigner. O foco da dissertação estará no Teorema 2.2.3 e uma demonstração combinatorial para o tal.

Para criar alguma intuição a respeito da origem do nosso objeto de estudo, vamos ver *três* versões do teorema. A primeira vai considerar a proporção de autovalores num intervalo. A segunda expressa a mesma ideia mas no contexto de funções integradas em relação a uma medida (que representa o conjunto de autovalores). A terceira (que terá nosso foco) corresponde à uma variação da segunda em que só funções polinomiais estão consideradas, e isso corresponde ao traço de potências da matriz \mathbf{X}_n .

Teorema 2.2.1 (Lei do semicírculo de Wigner) *Seja $\mathbf{X}_n = n^{-\frac{1}{2}}\mathbf{Y}_n$ uma sequencia de matrizes de Wigner cujas entradas satisfazem $\mathbb{E}(Y_{ij}) = 0 \forall i, j$ e $\mathbb{E}(Y_{12}^2) = t$. Seja $I \subset \mathbb{R}$ um intervalo. Defina as variáveis aleatórias*

$$E_n(I) = \frac{\#\left(\{\lambda_1(\mathbf{X}_n) \dots \lambda_n(\mathbf{X}_n)\} \cap I\right)}{n}.$$

Então $E_n(I) \rightarrow \sigma_t(I)$ em probabilidade, quando $n \rightarrow \infty$, onde

$$\sigma_t(dx) = \frac{1}{2\pi t} \sqrt{(4t - x^2)_+} dx.$$

Precisamente $\forall I \subset \mathbb{R}$ intervalo e todo $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(|E_n(I) - \sigma_t(I)| > \epsilon) = 0.$$

Observe que a ideia por trás de $E_n(I)$ é mensurar a quantidade normalizada de autovalores que vamos ter no intervalo I . Em algum sentido, fixando n e um tamanho pequeno ("binsize" no código da Seção 6.3) para os intervalos I , $E_n(I)$ gera um histograma ao tomar os intervalos consecutivamente.

O teorema é enunciado dessa forma para que fique claro que a quantidade de autovalores de X_n no intervalo I , converge para a medida $\sigma_t(I)$, que pode ser entendida como a medida da área abaixo do tal semicírculo, na parte que compete ao intervalo.

Agora vamos simplificar esse teorema, no sentido de tentar tornar a prova dele mais compreensível ou intuitiva.

Teorema 2.2.2 (Lei do Semicírculo de Wigner) *Seja $\mathbf{X}_n = n^{-\frac{1}{2}}\mathbf{Y}_n$ uma sequência de matrizes de Wigner com entradas satisfazendo $\mathbb{E}(Y_{ij}) = 0 \forall i, j$ e $\mathbb{E}(Y_{12}^2) = t$.*

Então $\forall f \in C_b(\mathbb{R})$ (função contínua e limitada) e todo $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{1}{n} \sum_{i=1}^n f(\lambda_i(\mathbf{X}_n)) - \int f d\sigma_t \right| > \epsilon \right) = 0.$$

Observação 3 *Se $f = \mathbb{I}_I$, obtemos o teorema apresentado da forma anterior.*

Queremos provar para funções contínuas, mas bastaria trocar essas funções contínuas por polinômios e provar uma versão ‘Mais fraca’. Mais simplificada, seria suficiente provarmos para monômios, e a versão para polinômios se estende usando linearidade.

Com isso, podemos escrever uma terceira versão do teorema:

Teorema 2.2.3 (Lei de Wigner para os momentos) *Seja $\mathbf{X}_n = n^{-\frac{1}{2}}\mathbf{Y}_n$ uma sequência de Wigner com entradas satisfazendo $\mathbb{E}(Y_{ij}) = 0 \forall i, j$ e $\mathbb{E}(Y_{12}^2) = t$. Então para um fixado $k \in \mathbb{N}$ e $\epsilon > 0$.*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{1}{n} \text{Tr}(\mathbf{X}_n^k) - \int x^k \sigma_t(dx) \right| > \epsilon \right) = 0.$$

Observação 4 *Esse é um caso particular da segunda versão, tomando $f = x^k$:*

$$\frac{1}{n} \sum_{i=1}^n f(\lambda_i(\mathbf{X}_n)) = \frac{1}{n} \sum_{i=1}^n (\lambda_i(\mathbf{X}_n))^k = \frac{1}{n} \text{Tr}(\mathbf{X}_n^k).$$

Observação 5 *Podemos calcular a integral à direita:*

$$\int x^k \sigma_t(dx) = \begin{cases} 0 & k \text{ é ímpar} \\ t^{k/2} C_{k/2} & k \text{ é par} \end{cases}$$

onde C_m é a sequência de Catalan.

Para k ímpar o resultado é trivial, dado que se trata da integral de uma função ímpar. Para provar o caso k par, confira o Teorema 3 de (QI; GUO, 2017).

2.3

Passos da prova

O resto da dissertação vai focar no Teorema 2.2.3.

Primeiro observamos que o problema em provar o Teorema 2.2.3 pode ser reduzido ao problema de controlar a esperança e variância de $\text{Tr}(\mathbf{X}_n^k)$. Vamos usar o seguinte Lema, que é um Corolário da desigualdade de Chebychev:

Lema 2.3.1 *Seja W_n uma sequencia de variáveis aleatórias com $\lim_{n \rightarrow \infty} \mathbb{E}(W_n) = m$, e $\lim_{n \rightarrow \infty} \text{Var}(W_n) = 0$. Então $W_n \rightarrow m$ em probabilidade:*

$$\forall \epsilon > 0, \quad \lim_{n \rightarrow \infty} \mathbb{P}(|W_n - m| > \epsilon) = 0.$$

Como queremos provar uma convergência em probabilidade, basta mostrar que a esperança converge e que a variância converge a zero. Ou seja, para mostrar Teorema 2.2.3 basta mostrar as seguintes duas afirmações:

Afirmção 1

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \text{Tr}(\mathbf{X}_n^k) = \begin{cases} 0 & \text{se } k \text{ ímpar} \\ t^{\frac{k}{2}} C_{\frac{k}{2}} & \text{se } k \text{ par} \end{cases}$$

Afirmção 2 *Para cada k , temos que $\text{Var}\left(\frac{1}{n} \text{Tr}(\mathbf{X}_n^k)\right) \rightarrow 0$ quando $n \rightarrow \infty$.*

Considerando $W_n = \frac{1}{n} \text{Tr}(\mathbf{X}_n^k)$ é claro que Teorema 2.2.3 segue das afirmações usando Lema 2.3.1 e Observação 5.

3

Prova da Afirmação 1

O objetivo do capítulo é provar a primeira afirmação. Em resumo, queremos mostrar que uma dada esperança converge para 0 (se k for ímpar) ou $C_{\frac{k}{2}}$ (se k for par).

Veja a seguinte proposição (mostrar a Afirmação 1 corresponde a provar essa proposição). Vamos dar uma versão da prova que usa a condição mais forte sobre os mementos. Ou seja, consideramos $\mathbf{X}_n = n^{-\frac{1}{2}}\mathbf{Y}_n$ uma matriz de Wigner forte.

Proposição 3.0.1 *Seja $\mathbf{X}_n = n^{-\frac{1}{2}}\mathbf{Y}_n$ uma matriz de Wigner, e suponha que $\mathbb{E}(Y_{ij}) = 0 \forall i, j$ e $\mathbb{E}(Y_{12}^2) = t$. Então*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \operatorname{Tr}(\mathbf{X}_n^k) = \begin{cases} t^{\frac{k}{2}} C_{\frac{k}{2}}, & k \text{ par} \\ 0, & k \text{ ímpar} \end{cases}$$

3.1

Abrindo o Traço

Vamos começar a provar expandindo os termos do valor esperado do traço:

$$\frac{1}{n} \mathbb{E} \operatorname{Tr}(\mathbf{X}_n^k) = \frac{1}{n} \mathbb{E} \operatorname{Tr} \left[\left(n^{-\frac{1}{2}} \mathbf{Y}_n \right)^k \right] = n^{-\frac{k}{2}-1} \mathbb{E} \operatorname{Tr}(\mathbf{Y}_n^k).$$

Chegamos a isso só fazendo substituições e usando o fato de que a esperança é linear. Queremos o traço da matriz \mathbf{Y}_n elevada a k . Precisamos então identificar os elementos da diagonal da matriz \mathbf{Y}_n^k .

Vamos entender melhor a multiplicação, por exemplo, de \mathbf{Y}_3 por \mathbf{Y}_3 , matriz 3×3 :

$$\begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{21} & Y_{22} & Y_{23} \\ Y_{31} & Y_{32} & Y_{33} \end{bmatrix} \times \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{21} & Y_{22} & Y_{23} \\ Y_{31} & Y_{32} & Y_{33} \end{bmatrix}$$

Assim, temos que a soma que expressa o elemento $(2, 3)$ (por exemplo) da matriz $\mathbf{Y}_3 \times \mathbf{Y}_3$ é dada por:

$$\left[\mathbf{Y}_3^2 \right]_{23} = Y_{23}Y_{13} + Y_{22}Y_{23} + Y_{23}Y_{33} = \sum_{i=1}^3 Y_{2i}Y_{i3}$$

Se quisermos $\mathbf{Y}_3 \times \mathbf{Y}_3 \times \mathbf{Y}_3$, vamos ter algo como:

$$[\mathbf{Y}_3^3]_{23} = \sum_{i=1, j=1}^3 Y_{2i} Y_{ij} Y_{j3}$$

De forma mais geral, vamos ter para \mathbf{Y}_n^3 :

$$[\mathbf{Y}_n^3]_{23} = \sum_{i=1, j=1}^n Y_{2i} Y_{ij} Y_{j3}$$

Por fim, \mathbf{Y}_n é $n \times n$, e vamos fazer $\mathbf{Y}_n \times \mathbf{Y}_n \cdots \times \mathbf{Y}_n$, k vezes:

$$[\mathbf{Y}_n^k]_{23} = \sum_{i_2=1, \dots, i_k=1}^n Y_{2i_2} Y_{i_2 i_3} \cdots Y_{i_{k-1} i_k} Y_{i_k 3}$$

Onde o termo genérico, isto é $2 = i$ e $3 = j$, é exatamente o termo que vimos há pouco.

Finalmente, teremos $\forall 1 \leq i, j \leq n$

$$[\mathbf{Y}_n^k]_{ij} = \sum_{1 \leq i_2, \dots, i_k \leq n} Y_{i i_2} Y_{i_2 i_3} \cdots Y_{i_{k-1} i_k} Y_{i_k j} \quad (3-1)$$

Porém só estamos interessados nos termos da diagonal, então $i = j$. Note que o valor esperado do traço surge ao pegar a soma das esperanças das entradas da diagonal, isso justifica a primeira igualdade na equação a seguir:

$$\mathbb{E} \text{Tr}(\mathbf{Y}_n^k) = \sum_{i_1=1}^n \mathbb{E}([\mathbf{Y}_n^k]_{i_1 i_1}) = \sum_{1 \leq i_1, i_2, \dots, i_k \leq n} \mathbb{E}(Y_{i_1 i_2} Y_{i_2 i_3} \cdots Y_{i_k i_1}). \quad (3-2)$$

Na segunda igualdade, estamos “abrindo” o termo i_1, i_1 . Como a esperança é linear, na segunda igualdade estamos juntando o somatório das esperanças com o do símbolo $[\mathbf{Y}_n^k]_{i_1 i_1}$.

Para facilitar, a partir de agora, seja $[n] := \{1, \dots, n\}$ e $I := (i_1, \dots, i_k)$, define-se

$$Y_I = Y_{i_1 i_2} \cdots Y_{i_k i_1}$$

E a equação acima fica

$$\mathbb{E} \text{Tr}(\mathbf{Y}_n^k) = \sum_{1 \leq i_1, i_2, \dots, i_k \leq n} \mathbb{E}(Y_{i_1 i_2} Y_{i_2 i_3} \cdots Y_{i_k i_1}) \equiv \sum_{I \in [n]^k} \mathbb{E}(Y_I) \quad (3-3)$$

Observação 6 O índice I pertence a $[n]^k$ e pode ser interpretado como sendo uma “palavra” de “ k letras”, cada “letra” com n opções, isto é, I é um k -ordenado. Assim temos que I pode, por exemplo, ser: $I = 222 \cdots 222$, de k letras onde repete 2 k vezes. Nesse caso $Y_I = Y_{22} Y_{22} \cdots Y_{22}$. Ou ainda podemos ter outro exemplo como: $I = 1234 \cdots (k-1)k$, uma “palavra” de k letras, que possui os números em sequencia de 1 até k . Nesse caso $Y_I =$

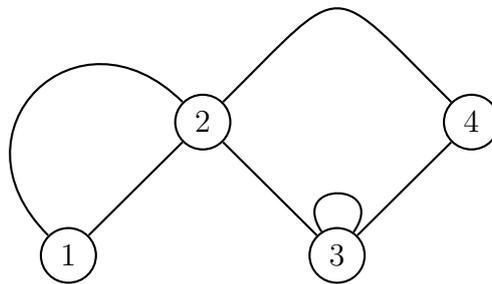
$Y_{12}Y_{23}\cdots Y_{k-1k}Y_{k1}$. Um último exemplo pode ser: $I = nn\cdots nnn$, mais uma palavra de k letras onde n se repete k vezes. Nesse caso $Y_I = Y_{nn}Y_{nn}\cdots Y_{nn}$.

3.2

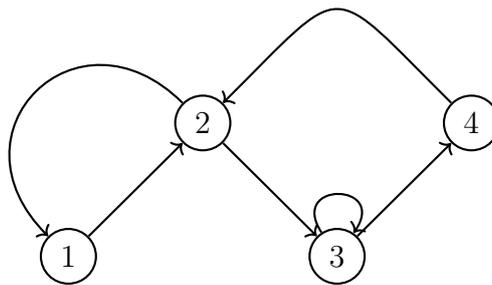
Interpretando o índice I como um caminho em um grafo

Vamos agora dar um salto nas interpretações e mudar um pouco a forma de ver esse índice I .

Por exemplo, se $I = (1, 2, 3, 3, 4, 2)$, imagine um grafo com conjunto de vértices $V_I = \{1, 2, 3, 4\}$ e $E_I = \{\{1, 2\}, \{2, 3\}, \{3, 3\}, \{3, 4\}, \{4, 2\}, \{2, 1\}\}$ (confira figura 3.2).



Como o grafo não possui uma orientação também será útil considerar o circuito no grafo gerado por I . Um *circuito* num grafo é simplesmente uma lista de vértices visitados, lista essa que começa e acaba no mesmo vértice. É exigido que vértices consecutivos da lista sejam ligados por uma aresta no grafo. Ou seja, também podemos ver I como um circuito:



Mais formalmente,

Definição 3.2.1 *Seja $I \in [n]^k$ uma k -upla, $I = (i_1, i_2 \dots i_k)$. Defina o grafo G_I da seguinte forma: seus vértices V_I são os elementos distintos de $\{i_1, i_2, \dots, i_k\}$, e suas arestas E_I são os pares distintos entre $\{i_1, i_2\}, \{i_2, i_3\}, \dots, \{i_{k-1}, i_k\}, \{i_k, i_1\}$. O circuito w_I é a sequencia*

$$w_I = (\{i_1, i_2\}, \{i_2, i_3\}, \dots, \{i_{k-1}, i_k\}, \{i_k, i_1\})$$

de arestas.

Podemos pensar em G como uma função que vai das k -uplas para os grafos.

O índice I que define o grafo, agora também define um circuito w_I nesse mesmo grafo.

Por definição, o circuito w_I visita todas as arestas de G_I (incluindo qualquer laço presente), começando e terminando no mesmo vértice. Em particular isso significa que o grafo G_I é conexo.

No nosso exemplo:

$$Y_I = Y_{12}Y_{23}Y_{33}Y_{34}Y_{42}Y_{21}$$

O ciclo w_I codifica uma rotulagem das arestas: o número de vezes que cada aresta foi atravessada. Agora vamos usar o índice I , e seguir extraindo informações dos grafos em questão a partir de mais uma definição : $w_I(e)$. Já vimos que w_I pode ser visto como um caminho fechado, um ciclo, agora podemos definir...

Definição 3.2.2 Dizemos que $w_I(e)$ é a quantidade de vezes que a aresta “ e ” aparece em w_I . Se e não está em w_I , definimos $w_I(e) = 0$.

No exemplo acima, $w_I(e)$ é igual a 2 na aresta “ e ” = $\{1,2\}$, e igual a 1 para todas as outras arestas, incluindo o laço $\{3,3\}$ (o “caminhante” só passaria por esse laço uma única vez). Esse número está evidente na expansão Y_I : usando o fato de que $Y_{ij} = Y_{ji}$, nós teremos

$$Y_I = Y_{12}^2 Y_{23} Y_{33} Y_{34} Y_{24}$$

Isso é porque $Y_{12} = Y_{21}$, pois pedimos simetria, isso é, por definição: primeiro são criadas as variáveis que definem o triângulo superior da matriz, em baixo é só uma cópia.

Em cada caso o expoente de Y_{ij} é igual a $w_I(\{i,j\})$. Então,

$$Y_I = \prod_{1 \leq i \leq j \leq n} Y_{ij}^{w_I(\{i,j\})}$$

Ou seja, agora estamos escrevendo aquele termo da soma 3-3, em função de grafos. Aqui podemos desapegar dos índices, e nos concentrar nos caminhos. Como temos vários índices que podem ser interpretados como os mesmos caminhos, teremos uma soma mais simples.

Agora, sabemos que trabalhar com esperanças de variáveis independentes é melhor, então vamos juntar os termos dependentes para podermos usar propriedades da esperança. Isto é, quando A e B são variáveis independentes,

$$\mathbb{E}(AB) = \mathbb{E}(A)\mathbb{E}(B).$$

Logo é importante igualar as variáveis aleatórias Y_{ij} e Y_{ji} e representá-las com uma “potência”, que também significa a quantidade de vezes que o caminho passa pela tal aresta.

Como a diagonal e os termos fora dela têm diferentes distribuições, precisamos fazer alguma distinção entre os termos referentes a laços $E_I^s = \{\{i, i\} \in E_I\}$ e as arestas conectoras $E_I^c = \{\{i, j\} \in E_I : i \neq j\}$. Então nós teremos

$$\mathbb{E}(Y_I) = \prod_{1 \leq i \leq j \leq n} \mathbb{E}\left(Y_{ij}^{w_I(\{i,j\})}\right) = \prod_{e_s \in E_I^s} \mathbb{E}\left(Y_{11}^{w_I(e_s)}\right) \cdot \prod_{e_c \in E_I^c} \mathbb{E}\left(Y_{12}^{w_I(e_c)}\right). \quad (3-4)$$

Aqui separamos os termos da diagonal, dos fora dela, pois eles têm distribuições diferentes. Todos na diagonal tem a distribuição igual a de Y_{11} , e os fora a mesma que a de Y_{12} .

Observação 7 *Note que a equação mostra que a esperança só depende de G_I e w_I .*

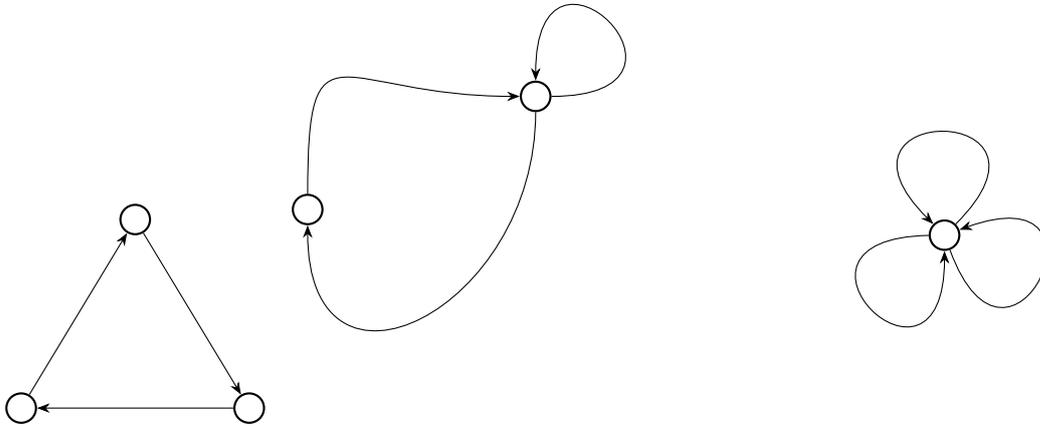
Para qualquer k -índice I e pela definição do G_I , sabemos que esse tem no máximo k vértices, além de ser conexo e orientado. Vemos que a quantidade de passos no ciclo é a mesma que a quantidade de entradas em I . Podemos concluir então que existem tantas entradas em I , quanto arestas percorridas:

$$|w_I| \equiv \sum_{e \in E_I} w_I(e) = k.$$

Nessa equação estamos somando quantas vezes cada aresta pertence ao caminho. Vamos agora usar uma notação mais cômoda:

Definição 3.2.3 *Seja \mathcal{G}_k o conjunto de todos os pares G e w , onde G é um grafo conexo com até k vértices, e w um caminho fechado de G de tamanho k , que passa por todas as arestas de G .*

Para ilustrar, podemos ver como se parecem os pares de \mathcal{G}_3 :



Observação 8 Todos os pares de \mathcal{G}_3 quando desenhados se parecem com um dos exemplos acima. A rigor, pares (G, w) diferentes que produzem o mesmo desenho, ainda são elementos diferentes de \mathcal{G}_3 . Por exemplo, na figura central, faz diferença se w começa pelo laço ou não.

Com essa nova definição, podemos enfatizar a Observação 7 com a seguinte Definição.

Definição 3.2.4 Dado um par $(G, w) \in \mathcal{G}_k$ como acima, definimos

$$\Pi(G, w) = \prod_{e_s \in E^s} \mathbb{E} \left(Y_{11}^{w(e_s)} \right) \cdot \prod_{e_c \in E^c} \mathbb{E} \left(Y_{12}^{w(e_c)} \right),$$

onde E^s e E^c são os conjuntos de laços e arestas conectoras de G , respectivamente.

Observação 9 Pela definição acima e pela equação (3-4), dada uma k -upla I ,

$$\Pi(G_I, w_I) = \mathbb{E}(Y_I).$$

Assim, a equação (3-3) torna-se:

$$\mathbb{E} \text{Tr} \left(\mathbf{Y}_n^k \right) = \sum_{I \in [n]^k} \Pi(G_I, w_I) = \sum_{(G, w) \in \mathcal{G}_k} \sum_{\substack{I \in [n]^k \\ (G_I, w_I) \cong (G, w)}} \Pi(G, w). \quad (3-5)$$

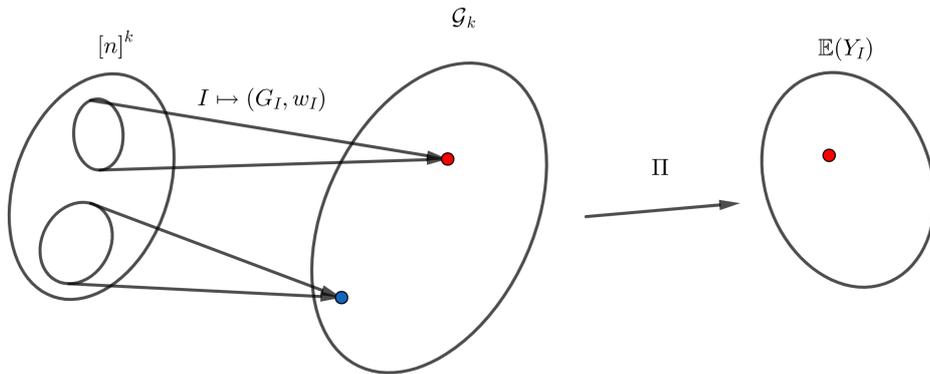
Como cada termo do somatório depende somente do par (G_I, w_I) , podemos calcular a soma de forma a juntar os termos que possuem mesmo par (G_I, w_I) .

Ao invertermos a ordem que percorreremos os índices das somas, olhamos primeiro para os pares (G, w) e depois os índices I que podem gerar aquele par. Temos então que cada I , que gera aquele par, contribui com o mesmo termo $\Pi(G_I, w_I)$.

Logo, uma soma com vários termos iguais, torna-se um produto, melhorando nossa equação:

$$\mathbb{E} \operatorname{Tr} \left(\mathbf{Y}_n^k \right) = \sum_{(G,w) \in \mathcal{G}_k} \Pi(G, w) \cdot \# \left\{ I \in [n]^k : (G_I, w_I) \cong (G, w) \right\}. \quad (3-6)$$

O ponto é que existem vários I que vão para o mesmo par (G, w) , e temos como descobrir quantos.



Podemos interpretar a função $I \mapsto \mathbb{E}(Y_I)$ como composição das funções $I \mapsto (G_I, w_I)$ e $(G, w) \mapsto \Pi(G, w)$. Para calcular a soma $\sum \mathbb{E}(Y_I)$ basta calcular a soma $\sum \Pi(G, w)$ com pesos, de acordo com quantos I estão na pré-imagem de cada (G, w) .

Queríamos calcular $\mathbb{E} \operatorname{Tr} (\mathbf{X}_n^k)$ e então chegamos a uma soma de termos $\mathbb{E}(Y_I)$. Com isso obtivemos a equação 3-6. Juntando as informações até agora, obtemos:

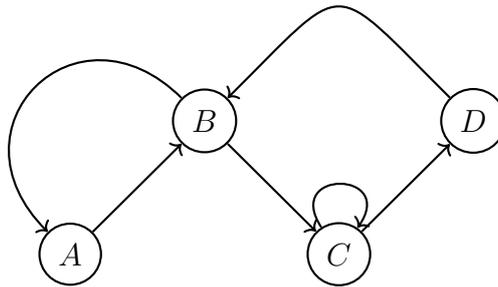
$$\frac{1}{n} \mathbb{E} \operatorname{Tr} \left(\mathbf{X}_n^k \right) = \sum_{(G,w) \in \mathcal{G}_k} \Pi(G, w) \cdot \frac{\# \left\{ I \in [n]^k : (G_I, w_I) \cong (G, w) \right\}}{n^{\frac{k}{2}+1}} \quad (3-7)$$

Vamos agora resolver a 3-7. Queremos o traço, mas vimos que podemos achar esse valor calculando o termo $\Pi(G, w)$ e $\#\{I \in [n]^k : (G_I, w_I) \cong (G, w)\}$.

3.3

Calculando grafos

Vamos começar calculando $\#\{I \in [n]^k : (G_I, w_I) \cong (G, w)\}$. Perceba que o caminho w contém informação de como vamos distribuir as k entradas. Começaremos com um par (G, w) de exemplo e vamos tentar construir um índice I compatível com o par. Podemos até conseguir mais de um.



Nomeamos os vértices apenas para descrever o ciclo w . Neste exemplo em particular, pedimos que o nosso ciclo percorra A, B, C, C, D, B, A nesta ordem.

Nesse caso, poderíamos ter $A = 1, B = 2, C = 3$ e $D = 4$ ou $A = 2, B = 3, C = 4$ e $D = 10$ ou $A = 8, B = 6, C = 3$ e $D = 1$. Os possíveis valores para cada letra dependem de n . Mais precisamente, eles podem ser qualquer valor inteiro entre 1 e n , tomando o cuidado de não ter letras diferentes com o mesmo valor. Se quisermos contar quantos I 's acarretam nessa dupla (G, w) , podemos dar n escolhas para a letra A , $n - 1$ escolhas para a letra B e assim por diante.

Mas geralmente, pela argumentação acima é claro que temos o seguinte lema:

Lema 3.3.1 Dado $(G, w) \in \mathcal{G}_k$ com $|G|$ o número de vértices de G , então

$$\#\{I \in [n]^k : (G_I, w_I) \cong (G, w)\} = n(n - 1) \cdots (n - |G| + 1).$$

Dessa forma, 3-7 pode ser reescrita como:

$$\frac{1}{n} \mathbb{E} \text{Tr}(\mathbf{X}_n^k) = \sum_{(G, w) \in \mathcal{G}_k} \Pi(G, w) \cdot \frac{n(n - 1) \cdots (n - |G| + 1)}{n^{\frac{k}{2} + 1}}. \quad (3-8)$$

Observação 10 Note que k é fixo mesmo que n vá para infinito, logo a soma é finita. Teremos então que $n(n - 1) \cdots (n - |G| + 1)$ é assintoticamente $n^{|G|}$.

De fato, $|G| \leq k$, e com n bem grande, $(n-1) \approx n$, $(n-c) \approx n$ para qualquer constante c .

Para eliminar muitos termos da nossa soma, observe que se tivermos alguma aresta que é percorrida apenas uma vez, isto é, $w(e) = 1$, o que significa que a equação vai ter pelo menos um termo:

$$\mathbb{E}(Y_{ij}^{w(e)}) = \mathbb{E}(Y_{ij}) = 0$$

por hipótese (pedimos que a distribuição de probabilidade Y_{ij} fosse centrada no zero), então a esperança desse termo é nula.

Ou seja, $\Pi(G, w) = 0$, toda vez que w passar apenas uma vez por alguma aresta. Podemos então nos concentrar apenas nos w que tenham todas as arestas com $w(e) \geq 2$. Vamos denotar tais ciclos como $w \geq 2$:

$$\frac{1}{n} \mathbb{E} \operatorname{Tr}(\mathbf{X}_n^k) = \sum_{(G,w) \in \mathcal{G}_k, w \geq 2} \Pi(G, w) \cdot \frac{n(n-1) \cdots (n-|G|+1)}{n^{\frac{k}{2}+1}}. \quad (3-9)$$

Como só estamos lidando com $w \geq 2$, estamos deixando vários possíveis grafos de fora. Ficando com aqueles em que as arestas são percorridas ao menos 2 vezes, assim o número de arestas é $\leq \frac{k}{2}$.

Lema 3.3.2 *Seja $G = (V, E)$ um grafo finito e conexo. Então $|G| := \#V \leq \#E + 1$, e a igualdade ocorre se e somente se G é uma árvore.*

Pelo Lema 3.3.2, teremos que $|G| \leq \#E + 1 \leq \frac{k}{2} + 1$, pois $\#E \leq \frac{k}{2}$ já que todas as arestas são atravessadas ao menos 2 vezes pelo ciclo de tamanho k .

Vejam os que acontece quando k é ímpar. Vamos nos aproveitar do fato de que $|G|$ é um número inteiro, e apertar a cota para o caso k ímpar:

$$|G| \leq \frac{(k-1)}{2} + 1 = \frac{k}{2} + \frac{1}{2}.$$

Com essa cota mais apertada que conseguimos, para o caso k ímpar, temos que o termo dependente de n na soma 3-9, é da forma $O(n^{-\frac{1}{2}})$. Como já observamos que a soma possui um número finito de termos, segue que, para k ímpar, temos

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \operatorname{Tr}(\mathbf{X}_n^k) = 0.$$

Vamos agora ver o caso par. Note que se você consegue melhorar a cota de $|G|$ ou seja, se $|G| \leq \frac{k}{2}$, caímos no caso analisado quando k ímpar, isso é, o numerador vai ser menor que o denominador na equação 3-9, e assim, os

termos da soma vão a zero, exatamente como visto antes. Vamos então olhar para o caso complementar aos descritos na Proposição 3.3.1.

Definição 3.3.1 *Seja $\mathcal{G}_k^{\frac{k}{2}+1} = \{(G, w) \in \mathcal{G}_k : w \geq 2, |G| = \frac{k}{2} + 1\}$.*

Pelo observado acima, podemos apenas considerar este subconjunto de pares. Para nos ajudar, vamos evocar a seguinte proposição:

Proposição 3.3.1 *Seja k um inteiro par e seja $(G, w) \in \mathcal{G}_k$ com $w \geq 2$.*

- (a) *Se existe laço em G , ou seja se $E_s \neq \emptyset$, então $|G| \leq \frac{k}{2}$.*
- (b) *Se existe uma aresta e em G tal que $w(e) \geq 3$, então $|G| \leq \frac{k}{2}$.*

Prova. (a) Como o grafo $G = (V, E)$ possui laços, em particular não é uma árvore; pelo que segue do lema 3.3.2 temos $\#V < \#E + 1$. Mas $w \geq 2$ implica que $\#E \leq \frac{k}{2}$, e então $\#V < \frac{k}{2} + 1$, logo $|G| = \#V \leq \frac{k}{2}$.

(b) A soma de w sobre todas as arestas E em G é k . Então, a soma de w sobre $E \setminus \{e\}$ é $\leq k - 3$. Como $w \geq 2$, isso significa que o número de arestas exceto e é $\leq \frac{(k-3)}{2}$. Então $\#E \leq \frac{(k-3)}{2} + 1 = \frac{(k-1)}{2}$. Como resultado do Lema 3.3.2 isso significa que $\#V \leq \frac{(k-1)}{2} + 1 = \frac{(k+1)}{2}$. Como k é par, segue que $|G| = \#V \leq \frac{k}{2}$. ■

Observe que se k é par, muitos termos vão para zero quando estivermos nos casos (G, w) e w for um ciclo como descrito acima. Isso não implica que $\mathbb{E}(X_n^k)$ vá para 0, pois ainda restam vários termos a serem somados: os pares (G, w) com $w(e) = 2 \forall e$. E isso vai nos assegurar o resultado da Proposição 3.0.1.

Observação 11 *Como consequência da Proposição 3.3.1, todos os elementos (G, w) de $\mathcal{G}_k^{\frac{k}{2}+1}$ são livres de laços, e w passa por todas as arestas exatamente 2 vezes, ou seja, $\#E = \frac{k}{2}$. Já pelo Lema 3.3.2, G é necessariamente uma árvore.*

Teremos então:

$$\frac{1}{n} \mathbb{E} \text{Tr}(\mathbf{X}_n^k) = \sum_{(G,w) \in \mathcal{G}_k^{\frac{k}{2}+1}} \Pi(G, w) \cdot \frac{n(n-1) \cdots (n - |G| + 1)}{n^{\frac{k}{2}+1}} + O_k(n^{-1}) \quad (3-10)$$

Observe que $O_k(n^{-1})$ é referente à parte que assintoticamente é zero, dos pares (G, w) , onde $|G| \leq \frac{k}{2}$. Isso é, esse termo é o termo que vai a zero e não vai contribuir na soma no limite $n \rightarrow \infty$.

Agora, se $|G| = \frac{k}{2} + 1$ e $n(n-1) \cdots (n - \frac{k}{2}) \approx n^{\frac{k}{2}+1}$, cada fração (termo que multiplica o $\Pi(G, w)$) vai para 1 quando n vai para infinito. Temos então:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \operatorname{Tr} (\mathbf{X}_n^k) = \sum_{(G,w) \in \mathcal{G}_k^{\frac{k}{2}+1}} \Pi(G, w) \quad (3-11)$$

Olhando cuidadosamente para um dos termos dessa soma, temos que:

$$\Pi(G, w) = \prod_{e_c \in E^c} \mathbb{E} (Y_{12}^{w(e_c)}) = \prod_{e_c \in E^c} \mathbb{E} (Y_{12}^2) = t^{\#E} = t^{\frac{k}{2}}. \quad (3-12)$$

Assim temos:

$$\lim_{n \rightarrow \infty} \mathbb{E} \operatorname{Tr} (\mathbf{X}_n^k) = t^{\frac{k}{2}} \cdot \#\mathcal{G}_k^{\frac{k}{2}+1} \quad (3-13)$$

Agora só falta mostrar que $\#\mathcal{G}_k^{\frac{k}{2}+1}$ é o número de Catalan $C_{\frac{k}{2}}$, vamos fazer isso associando cada par do $\mathcal{G}_k^{\frac{k}{2}+1}$ a um caminho de Dyck.

3.4 Caminhos de Dyck

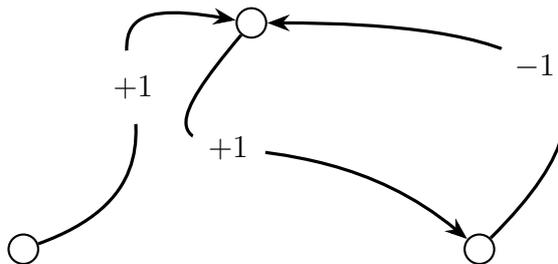
Definição 3.4.1 Para todo k par, conjunto de caminhos de Dyck de tamanho k é:

$$\mathcal{D}_k = \left\{ (d_1, \dots, d_k) \in \{\pm 1\}^k : \sum_{i=1}^j d_i \geq 0 \ \forall j \in \{1, \dots, k\} \text{ e } \sum_{i=1}^k d_i = 0 \right\}.$$

No Apêndice, Seção 6.2, está provado que $|\mathcal{D}_k| = C_{\frac{k}{2}}$. Para completar a prova da primeira afirmação 1, o objetivo desse capítulo, só falta provar o seguinte lema.

Lema 3.4.1 Se k é par, então existe uma bijeção $\mathcal{G}_k^{\frac{k}{2}+1} \rightarrow \mathcal{D}_k$.

Prova. Vamos construir uma aplicação $d : \mathcal{G}_k^{\frac{k}{2}+1} \rightarrow \mathcal{D}_k$ que se mostrará uma bijeção. Seja $d(G, w)$ a sequência ordenada dada pelos valores que ao caminhar pelo grafo G por w , em passos, indexamos com $+1$ quando passamos pela primeira vez a aresta e -1 ao passar pela segunda vez por uma aresta. Por exemplo:



Note que, no nosso caso, a soma dos termos de $d(G, w)$, é zero, pois temos a mesma quantidade de $+1$ e -1 . No entanto, o -1 , só aparece depois de um

+1, pois só aparece ao passar pela aresta pela segunda vez. Logo, a soma parcial das entradas de $d(G, w) \geq 0$. Isso mostra que d está bem definida.

Também é fácil de obter uma inversa $\mathcal{D}_k \rightarrow \mathcal{G}_k^{\frac{k}{2}+1}$. Dado um caminho de Dyck, vamos construir um grafo e circuito nele a começando com um único vértice v_0 . A cada +1 na sequência, se estamos no vértice v_i , podemos chamar o próximo vértice de v_{i+1} e caminhamos de v_i para v_{i+1} . A cada -1, se estamos no vértice v_i , voltamos para o vértice v_{i-1} . Não é difícil concluir que, pela construção, esta aplicação define unicamente uma inversa de d para cada caminho de Dyck. ■

4

Prova da Afirmação 2

Para concluirmos a demonstração da terceira versão do Teorema de Wigner, nos resta mostrar a Afirmação 2: que a variância de $W_n = \frac{1}{n} \text{Tr} \left([X_n]^k \right)$ vai para zero. Vamos mostrar algo um pouco mais forte, pois podemos estimar a velocidade com que isso ocorre.

Proposição 4.0.1 *Seja $\mathbf{X}_n = n^{-\frac{1}{2}} \mathbf{Y}_n$ uma matriz de Wigner forte, e suponha que $\mathbb{E}(Y_{ij}) = 0 \forall i, j$ e $\mathbb{E}(Y_{12}^2) = t$. Então*

$$\text{Var} \left(\frac{1}{n} \text{Tr} \left(\mathbf{X}_n^k \right) \right) = O_k \left(\frac{1}{n^2} \right).$$

Observação 12 *Para provar a Proposição 4.0.1 vamos usar o fato de serem matrizes de Wigner forte. Entretanto é possível relaxar a restrição sobre os momentos e ainda provar o teorema 2.2.3 (cf. cap 5 (KEMP, 2016))*

4.1

A variância em termos de grafos

Vamos então começar como um dos preparativos para a prova, abrir os termos da variância:

$$\text{Var} \left(\frac{1}{n} \text{Tr} \left(X_n^k \right) \right) = \mathbb{E} \left[\frac{1}{n} \cdot \frac{1}{n^{\frac{k}{2}}} \text{Tr} \left(\mathbf{Y}_n^k \right) \right]^2 - \left(\mathbb{E} \left[\frac{1}{n} \cdot \frac{1}{n^{\frac{k}{2}}} \text{Tr} \left(\mathbf{Y}_n^k \right) \right] \right)^2 \quad (4-1)$$

$$= \frac{1}{n^{k+2}} \left\{ \mathbb{E} \left[\text{Tr} \left(\mathbf{Y}_n^k \right) \right]^2 - \left(\mathbb{E} \text{Tr} \left(\mathbf{Y}_n^k \right) \right)^2 \right\}. \quad (4-2)$$

Pela definição de Esperança e de traço podemos re-escrever os termos:

$$\mathbb{E} \left[\text{Tr} \left(\mathbf{Y}_n^k \right) \right]^2 = \sum_{I, J \in [n]^k} \mathbb{E} (Y_I Y_J),$$

$$\left(\mathbb{E} \text{Tr} \left(\mathbf{Y}_n^k \right) \right)^2 = \sum_{I, J \in [n]^k} \mathbb{E} (Y_I) \mathbb{E} (Y_J).$$

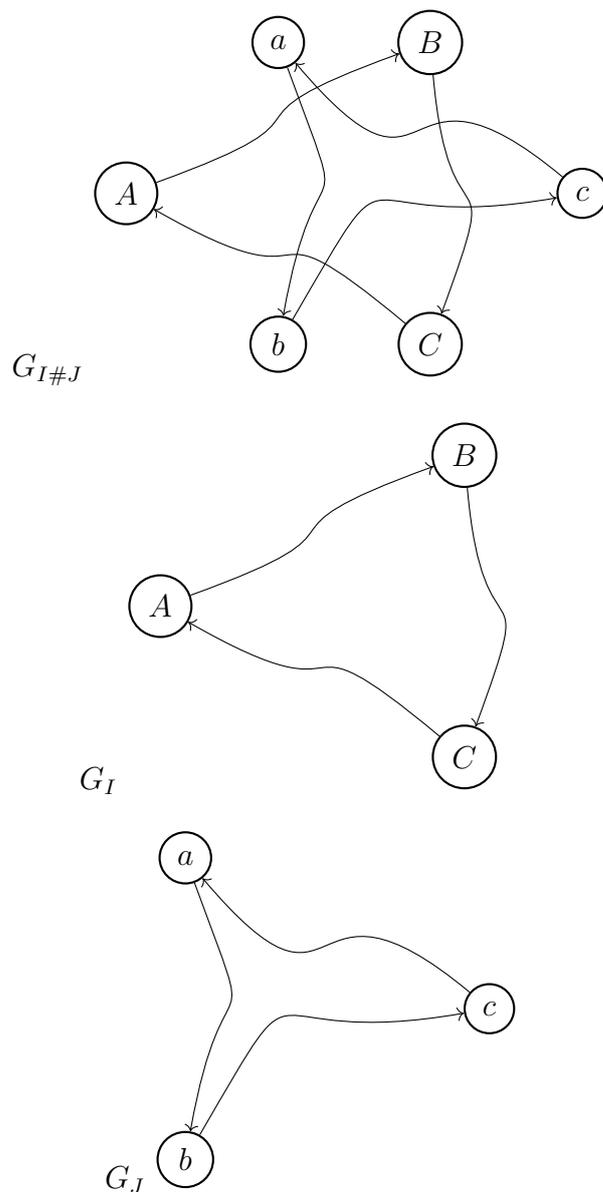
Dessa forma é possível obter:

$$\text{Var} \left(\frac{1}{n} \text{Tr} \left(\mathbf{X}_n^k \right) \right) = \frac{1}{n^{k+2}} \sum_{I, J \in [n]^k} [\mathbb{E} (Y_I Y_J) - \mathbb{E} (Y_I) \mathbb{E} (Y_J)].$$

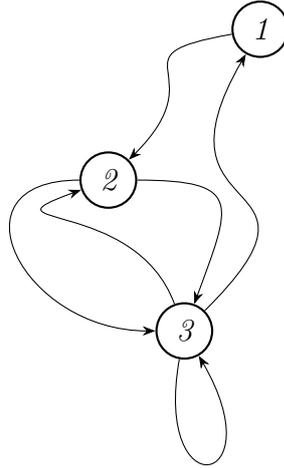
Para calcular essa soma, vamos repetir a ideia do capítulo anterior. Queremos, em vez de trabalhar com os índices I e J , achar uma interpretação deles como caminhos em certos grafos.

Inspirados no truque anterior, vamos reinterpretar as k -uplas I e J como pares (G_I, w_I) , (G_J, w_J) . Ainda melhor, será possível colar os grafos G_I e G_J de maneira intuitiva de forma que coordenadas iguais correspondam ao mesmo vértice.

Exemplo 4.1.1 *Por exemplo, para $I = (1, 2, 3)$ e $J = (4, 5, 6)$, como estes não compartilham nenhuma coordenada, gostaríamos de visualizar grafos disjuntos:*



Exemplo 4.1.2 *Porém, se $I = (1, 2, 3)$ e $J = (3, 3, 2)$, queremos que os vértices 2 e 3 colemb os grafos G_I e G_J :*



Mais precisamente, considere a seguinte definição.

Definição 4.1.1 *Seja $(I, J) \in [n]^k \times [n]^k$ um par de k -uplas, $I = (i_1, i_2 \dots i_k)$, $J = (j_1, j_2 \dots j_k)$. Defina o grafo $G_{I\#J}$ da seguinte forma: seus vértices $V_{I\#J}$ são os elementos distintos de $\{i_1, i_2 \dots, i_k, j_1, j_2 \dots j_k\}$, e suas arestas $E_{I\#J}$ são os pares distintos entre $\{i_1, i_2\}, \{i_2, i_3\}, \dots, \{i_{k-1}, i_k\}, \{i_k, i_1\}, \{j_1, j_2\}, \{j_2, j_3\}, \dots, \{j_{k-1}, j_k\}, \{j_k, j_1\}$.*

Também podemos definir $w_{I\#J} = w_I + w_J$ como a concatenação do circuito w_I com w_J , e para cada aresta $e \in E_{I\#J}$ definimos $w_{I\#J}(e) = w_I(e) + w_J(e)$, ou seja, é a quantidade de vezes que o circuito concatenado $w_{I\#J}$ passa pela aresta e .

Observação 13 *Note que $G_{I\#J}$ possui como subgrafos os grafos G_I e G_J . Em particular, os circuitos w_I e w_J definidos no capítulo anterior podem ser vistos como circuitos de $G_{I\#J}$. Entretanto, nesse grafo os circuitos w_I e w_J não precisam nem ter intersecção um com o outro, como no exemplo 4.1.1*

Mais uma vez, observe que os termos da soma $\mathbb{E}(Y_I Y_J) - \mathbb{E}(Y_I) \mathbb{E}(Y_J)$ dependem apenas da tripla $(G_{I\#J}, w_I, w_J)$. Para ilustrar, relembre o exemplo 4.1.2: Neste exemplo teríamos,

$$\mathbb{E}(Y_I) = \mathbb{E}(Y_{12} Y_{23} Y_{31}) = (\mathbb{E}(Y_{12}))^3,$$

$$\mathbb{E}(Y_J) = \mathbb{E}(Y_{33} Y_{32} Y_{23}) = \mathbb{E}(Y_{11}) \mathbb{E}(Y_{12}^2),$$

$$\mathbb{E}(Y_I Y_J) = \mathbb{E}(Y_{12} Y_{23} Y_{31} Y_{33} Y_{32} Y_{23}) = \mathbb{E}(Y_{11}) \mathbb{E}(Y_{12})^2 \mathbb{E}(Y_{12}^3).$$

Note que cada termo $\mathbb{E}(Y_{11})$, $\mathbb{E}(Y_{12})$, $\mathbb{E}(Y_{12})$ e $\mathbb{E}(Y_{12}^3)$ de $\mathbb{E}(Y_I Y_J)$ corresponde a uma aresta de $G_{I\#J}$. Se a aresta é um laço, usamos a distribuição da diagonal (Y_{11}) , e se for uma aresta conectora usamos a distribuição fora

dela (Y_{12}). Os expoentes dentro das esperanças correspondem à quantidade de vezes que a aresta é atravessada por $w_{I\#J}$, ou seja, $w_{I\#J}(e)$.

Mais geralmente, podemos concluir que

$$\begin{aligned} \mathbb{E}(Y_I Y_J) - \mathbb{E}(Y_I) \mathbb{E}(Y_J) &= \prod_{e_s \in E_{I\#J}^s} \mathbb{E}\left(Y_{11}^{w_{I\#J}(e_s)}\right) \cdot \prod_{e_c \in E_{I\#J}^c} \mathbb{E}\left(Y_{12}^{w_{I\#J}(e_c)}\right) \\ &- \prod_{e_s \in E_I^s} \mathbb{E}\left(Y_{11}^{w_I(e_s)}\right) \cdot \prod_{e_c \in E_J^c} \mathbb{E}\left(Y_{12}^{w_J(e_c)}\right) \cdot \prod_{e_s \in E_J^s} \mathbb{E}\left(Y_{11}^{w_J(e_s)}\right) \cdot \prod_{e_c \in E_I^c} \mathbb{E}\left(Y_{12}^{w_I(e_c)}\right) \end{aligned}$$

Vamos agora definir um conjunto que terá o papel de \mathcal{G}_k no capítulo anterior:

Definição 4.1.2 *Seja $\mathcal{G}_{k,k}$ o conjunto das triplas (G, w_1, w_2) onde G é um grafo com $|G| \leq 2k$ vértices e w_1, w_2 são dois circuitos de tamanho exatamente k cuja união percorre todas as arestas de G .*

Definição 4.1.3 *Dada uma tripla $(G, w_1, w_2) \in \mathcal{G}_{k,k}$ como acima, definimos*

$$\Pi'(G, w_1, w_2) = \prod_{e_s \in E^s} \mathbb{E}\left(Y_{11}^{(w_1+w_2)(e_s)}\right) \cdot \prod_{e_c \in E^c} \mathbb{E}\left(Y_{12}^{(w_1+w_2)(e_c)}\right) \quad (4-3)$$

$$- \prod_{e_s \in E^s} \mathbb{E}\left(Y_{11}^{w_1(e_s)}\right) \cdot \prod_{e_c \in E^c} \mathbb{E}\left(Y_{12}^{w_1(e_c)}\right) \cdot \prod_{e_s \in E^s} \mathbb{E}\left(Y_{11}^{w_2(e_s)}\right) \cdot \prod_{e_c \in E^c} \mathbb{E}\left(Y_{12}^{w_2(e_c)}\right) \quad (4-4)$$

onde E^s e E^c são os conjuntos de laços e arestas conectoras de $G_{I\#J}$, respectivamente.

Com esta definição, podemos facilmente reescrever o termo da variância como

$$\mathbb{E}(Y_I Y_J) - \mathbb{E}(Y_I) \mathbb{E}(Y_J) = \Pi'(G_{I\#J}, w_I, w_J).$$

Nosso objetivo é novamente ressaltar que o termo acima depende apenas dos elementos combinatoriais que definimos.

Seguindo a trilha que percorremos antes, poderemos reescrever a variância:

$$\text{Var}\left(\frac{1}{n} \text{Tr}(\mathbf{X}_n^k)\right) = \frac{1}{n^{k+2}} \sum_{I, J \in [n]^k} \Pi'(G_{I\#J}, w_I, w_J).$$

Este somatório terá vários termos iguais, pois várias duplas (I, J) podem gerar os mesmos trios como interpretações combinatoriais desses índices:

Podemos pensar em um mapa que vai dos pares (I, J) para os trios (G, w, w') . No entanto esse mapa não é injetivo, assim teremos mais de um par gerando o mesmo trio.

Podemos então reinterpretar a soma, no mesmo espírito de antes, isso é, inverter a ordem, primeiro pegar um trio e ver quantos pares geram

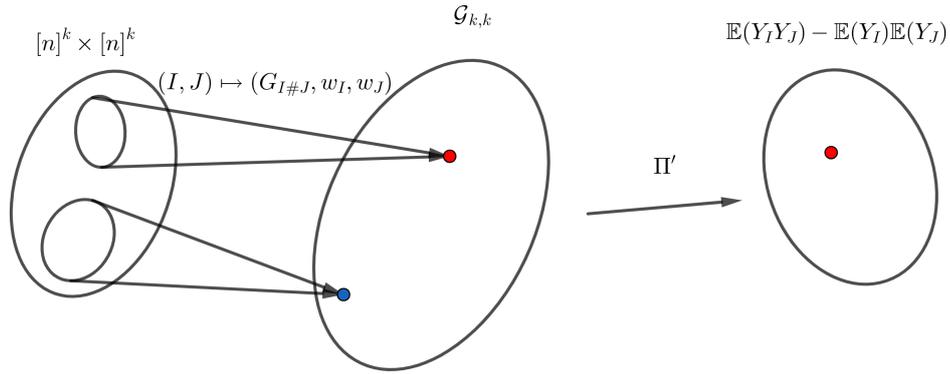


Figura 4.1: Diagrama

aquele mesmo trio, assim, na soma, teremos termos diferentes, mas que estão multiplicados pela quantidade de vezes que cada termo aparece. Então a soma se transforma em uma soma em $\mathcal{G}_{k,k}$:

$$\sum_{(G,w,w') \in \mathcal{G}_{k,k}} \Pi'(G, w, w') \cdot \frac{\#\{(I, J) \in ([n]^k)^2 : (G_{I\#J}, w_I, w_J) \cong (G, w, w')\}}{n^{k+2}}.$$

Note que antes era um somatório nos (I, J) e agora temos um somatório dos grafos.

Podemos generalizar o Lema 4.1.1 da seguinte forma:

Lema 4.1.1 *Dado $(G, w, w') \in \mathcal{G}_{k,k}$ com $|G|$ o número de vértices de G . Então*

$$\#\{(I, J) \in ([n]^k)^2 : (G_{I\#J}, w_I, w_J) \cong (G, w, w')\} = n(n-1) \cdots (n - |G| + 1).$$

A demonstração segue da mesma forma: podemos escolher o primeiro vértice de w de n formas, depois seguimos o caminho w cada vez com uma escolha a menos. Ao fim, recomeçamos pelo caminho w' .

4.2

Cotando a variância

Nosso objetivo agora será cotar os termos da variância. Vamos começar encontrando uma constante M_{2k} de forma que

$$|\Pi'(G_{I\#J}, w_I, w_J)| \leq 2M_{2k} \quad \forall I, J \in [n]^k.$$

Primeiro, lembre que individualmente cada momento $\mathbb{E}(Y_{11}^m)$ e $\mathbb{E}(Y_{12}^m)$ é cotada por r_m (que é por definição o máximo entre os dois momentos). Pela definição de Matriz de Wigner forte, cada r_m é finito. Então observe que:

$$\mathbb{E}(Y_I Y_J) = \prod_{e_s \in E_{I\#J}^s} \mathbb{E}(Y_{11}^{w_{I\#J}(e_s)}) \cdot \prod_{e_c \in E_{I\#J}^c} \mathbb{E}(Y_{12}^{w_{I\#J}(e_c)}) \leq \prod_{e \in E_{I\#J}} r_{w_{I\#J}(e)},$$

$$\mathbb{E}(Y_I) = \prod_{e_s \in E_I^s} \mathbb{E}(Y_{11}^{w_I(e_s)}) \cdot \prod_{e_c \in E_I^c} \mathbb{E}(Y_{12}^{w_I(e_c)}) \leq \prod_{e \in E_I} r_{w_I(e)},$$

$$\mathbb{E}(Y_J) = \prod_{e_s \in E_J^s} \mathbb{E}(Y_{11}^{w_J(e_s)}) \cdot \prod_{e_c \in E_J^c} \mathbb{E}(Y_{12}^{w_J(e_c)}) \leq \prod_{e \in E_J} r_{w_J(e)}.$$

Sendo assim, tanto $\mathbb{E}(Y_I Y_J)$ quanto $\mathbb{E}(Y_I) \mathbb{E}(Y_J)$ são cotados por um produto de r_m 's cuja soma dos índices m 's é fixa $2k$. De fato, os circuitos w_I e w_J têm tamanho k , exatamente. Logo não importando quantas vezes cada caminho use uma aresta, a soma da quantidade de vezes que o caminho passa por cada aresta, no fim, tem tamanho k :

$$\sum_{e \in G_{I\#J}} w_{I\#J}(e) = 2k = \sum_{e \in G_I} w_I(e) + \sum_{e \in G_J} w_J(e).$$

(Lembre-se que $w_I(e)$ é a quantidade de vezes que o caminho w passa pela aresta e no grafo G_I .)

Por um instante, esqueça os grafos e pense apenas nos índices como números, então perceba que existe uma configuração de índices de r , com a restrição da soma dos índices dar $2k$, que maximiza o produto pois são finitas combinações. Isso é, como os índices de r somam $2k$, de todas as formas diferentes de fazer essa soma, vamos pegar aquela que maximiza o valor do produto dos r . Vamos maximizar o produto para tornar óbvio que é cotado. E vamos chamar essa cota de M_{2k} .

Por exemplo, suponha que $k = 2$. Existem 5 formas de somar inteiros positivos para obter $2k = 4$: $1 + 1 + 1 + 1$, $1 + 1 + 2$, $1 + 3$, $2 + 2$, 4 . Então dentre os 5 valores $r_1 r_1 r_1 r_1$, $r_1 r_1 r_2$, $r_1 r_3$, $r_2 r_2$ e r_4 , definiremos $M_4 = M_{2k}$ como o maior desses valores.

Portanto teremos que $\forall I, J \in [n]^k$,

$$|\Pi'(G_{I\#J}, w_I, w_J)| \leq |\mathbb{E}(Y_I Y_J)| + |\mathbb{E}(Y_I) \mathbb{E}(Y_J)| \leq 2M_{2k}.$$

Agora vamos mostrar que muitos desses termos são nulos. Fixada uma aresta e , se $w_{I\#J}(e) = 1$ temos que a aresta aparece uma vez só na união dos caminhos, isso é ou $w_I(e) = 1$ ou $w_J(e) = 1$ o que implica que o termo $\Pi'(G_{I\#J}, w_I, w_J)$ é zero, pois ou teremos um termo $\mathbb{E}(Y_{11})$ ou $\mathbb{E}(Y_{12})$. Ignorando esses termos, a variância torna-se apenas

$$\sum_{(G, w, w') \in \mathcal{G}_{k,k}, w+w' \geq 2} \Pi'(G, w, w') \cdot \frac{n(n-1) \cdots (n-|G|+1)}{n^{k+2}}.$$

Como temos $2k$ passos na união dos caminhos, e cada aresta é passada no mínimo 2 vezes, no caso em que percorremos cada aresta exatamente 2 vezes, teremos k arestas diferentes. Isso é, no máximo temos k arestas diferentes. Como já vimos no Lema 3.3.2, temos no máximo $k+1$ vértices.

Então $n(n-1) \cdots (n-|G|+1) \leq n^{|G|} \leq n^{k+1}$ e assim temos que

$$\text{Var} \left(\frac{1}{n} \text{Tr}(\mathbf{X}_n^k) \right) \leq \sum_{(w, w') \in \mathcal{G}_{k,k}} \Pi'(G, w, w') \cdot \frac{n^{k+1}}{n^{k+2}} \leq \frac{1}{n} \cdot 2M_{2k} \cdot \#\mathcal{G}_{k,k}.$$

Note que, o número $2M_{2k} \cdot \#\mathcal{G}_{k,k}$ pode ser grande, no entanto não depende de n .

Ao provar que a variância vai com $O_k(\frac{1}{n})$ temos o suficiente para provar a convergência em probabilidade. Isso prova a Afirmação 2, mas não exatamente a Proposição 4.0.1.

Vamos além e mostrar que não há termos com $|G| = k+1$. Pelo Lema 3.3.2, sabemos que existem exatamente k arestas e isso é o mesmo que G ser uma árvore, isso é, $w + w' = 2$ (todas as arestas aparecem exatamente 2 vezes no caminho). Assim, podemos afirmar que G_I e G_J não têm arestas em comum.

E como w_I e w_J cobrem todas as arestas e retornam ao mesmo ponto inicial, o caminho passa por toda aresta uma quantidade de vezes par, dado que, por hipótese, não existem laços. Então, temos que o único $(G, w, w') \in \mathcal{G}_{k,k}$ com $|G| = k+1$ deve ter o conjunto de arestas coberto por w e w' distintos.

Portanto, se $|G_{I\#J}| = k+1$ tem-se que $\{\{i_1, i_2\}, \dots, \{i_k, i_1\}\} \cap \{\{j_1, j_2\}, \dots, \{j_k, j_1\}\} = \emptyset$. Com isso, podemos afirmar que Y_I e Y_J são independentes e que $\Pi'(G_{I\#J}, w_I, w_J) = \mathbb{E}(Y_I Y_J) - \mathbb{E}(Y_I) \mathbb{E}(Y_J) = 0$. Ou seja, nossa variância se resume a:

$$\sum_{\substack{(G,w,w') \in \mathcal{G}_{k,k} \\ w+w' \geq 2, |G| \leq k}} \Pi'(G, w, w') \cdot \frac{n(n-1) \cdots (n-|G|+1)}{n^{k+2}}.$$

Dessa forma pelos mesmos argumentos de contagem, podemos cotar a variância:

$$\text{Var} \left(\frac{1}{n} \text{Tr} \left(X_n^k \right) \right) \leq \sum_{\substack{(G,w,w') \in \mathcal{G}_{k,k} \\ w+w' \leq 2, |G| \leq k}} \Pi'(G, w, w') \cdot \frac{n^k}{n^{k+2}} \leq \frac{1}{n^2} \cdot 2M_{2k} \cdot \#\mathcal{G}_{k,k}.$$

5 Conclusões

Vamos começar este Capítulo com um breve resumo do que fizemos até o momento:

No Capítulo 2, estabelecemos 3 versões para o Teorema de Wigner. A primeira delas é o Teorema 2.2.1, que possui o enunciado mais intuitivo no sentido de obter uma convergência das distribuições dos autovalores para um semicírculo. A segunda versão é o Teorema 2.2.2, que tem um enunciado em termos de medidas. E finalmente a última versão é dada pelo Teorema 2.2.3, que é claramente um caso particular da versão anterior trocando funções contínuas quaisquer por monômios.

Para obtermos o resultado da última versão do Teorema, reduzimos a provar duas afirmações, que foram demonstradas nos dois seguintes capítulos. Com isso, obtivemos uma prova para o Teorema 2.2.3.

Nosso objetivo final será mostrar como chegaremos do resultado do Teorema 2.2.3 para o Teorema 2.2.2 e em seguida para o Teorema 2.2.1. Consideraremos a partir de agora $t = 1$.

5.1 Lei empírica dos autovalores

Definição 5.1.1 *Seja \mathbf{X}_n uma matriz de Wigner (forte). Sua “Lei empírica dos autovalores” $\mu_{\mathbf{X}_n}$ é uma medida de probabilidade*

$$\mu_{\mathbf{X}_n} = \frac{1}{n} \sum_{j=1}^n \delta_{\lambda_j(\mathbf{X}_n)}$$

Isso é, $\mu_{\mathbf{X}_n}$ é definido como a medida aleatória de probabilidade tal que, para qualquer função contínua $f \in C(\mathbb{R})$. A integral $\int f d\mu_{\mathbf{X}_n}$ é uma variável aleatória

$$\int f d\mu_{\mathbf{X}_n} = \frac{1}{n} \sum_{j=1}^n f(\lambda_j(\mathbf{X}_n)).$$

Observação 14 *Com esta definição, podemos reescrever o Teorema 2.2.2 como uma convergência entre medidas:*

Seja $\mathbf{X}_n = n^{-\frac{1}{2}} \mathbf{Y}_n$ uma sequência de matrizes de Wigner com entradas satisfazendo $\mathbb{E}(Y_{ij}) = 0 \forall i, j$ e $\mathbb{E}(Y_{12}^2) = t$.

Então $\forall f \in C_b(\mathbb{R})$ (função contínua e limitada) e todo $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \int f d\mu_{\mathbf{X}_n} - \int f d\sigma_t \right| > \epsilon \right) = 0.$$

Relembre que pela definição dada no Teorema 2.2.1 que:

$$E_n(I) = \frac{\#\{(\lambda_1(\mathbf{X}_n) \dots \lambda_n(\mathbf{X}_n)) \cap I\}}{n} = \int \mathbb{1}_I d\mu_{\mathbf{X}_n}.$$

Ou seja, ao trocar f pela função indicatriz de um intervalo $\mathbb{1}_I$, voltamos para a primeira versão do teorema: Teorema 2.2.1.

Como a função indicatriz de um intervalo pode ser aproximada arbitrariamente bem por funções contínuas, Teorema 2.2.1 seguirá facilmente desta versão do resultado.

5.2

Prova do Teorema 2.2.2

Lema 5.2.1 *Seja $k \in \mathbb{N}$ e $\epsilon > 0$. Então $\forall b > 4$,*

$$\limsup_{n \rightarrow \infty} \mathbb{P} \left(\int_{|x|>b} |x|^k \mu_{\mathbf{X}_n}(dx) > \epsilon \right) = 0.$$

Prova. Da desigualdade de Markov temos que:

$$\mathbb{P} \left(\int_{|x|>b} |x|^k \mu_{\mathbf{X}_n}(dx) > \epsilon \right) \leq \frac{1}{\epsilon} \mathbb{E} \left(\int_{|x|>b} |x|^k \mu_{\mathbf{X}_n}(dx) \right).$$

podemos então supor ν uma medida aleatória $\nu(dx) = |x|^k \mu_{\mathbf{X}_n}(dx)$. Então podemos usar a desigualdade de Markov pois podemos aplicar a todas as medidas positivas, não apenas a medidas de probabilidade:

$$\begin{aligned} \int_{|x|>b} |x|^k \mu_{\mathbf{X}_n}(dx) &= \nu\{x : |x| > b\} = \nu\{x : |x|^k > b^k\} \\ &\leq \frac{1}{b^k} \int |x|^k \nu(dx) = \frac{1}{b^k} \int x^{2k} \mu_{\mathbf{X}_n}(dx). \end{aligned}$$

Tomando os valores esperados temos então:

$$\mathbb{P} \left(\int_{|x|>b} |x|^k \mu_{\mathbf{X}_n}(dx) > \epsilon \right) \leq \frac{1}{\epsilon b^k} \mathbb{E} \left(\int x^{2k} \mu_{\mathbf{X}_n}(dx) \right) = \frac{1}{\epsilon b^k} \cdot \frac{1}{n} \mathbb{E} \text{Tr}(\mathbf{X}_n^k).$$

Assim, como já visto a parte direita converge para $\frac{C_k}{\epsilon b^k}$ onde C_k é um número de Catalan, logo é cotado por 4^k . Dessa fora segue que:

$$\limsup_{n \rightarrow \infty} \mathbb{P} \left(\int_{|x|>b} |x|^k \mu_{\mathbf{X}_n} > \epsilon \right) \leq \frac{1}{\epsilon} \left(\frac{4}{b} \right)^k.$$

Por outro lado temos que $|x| > b > 4 > 1$. Sabemos que a função $k \mapsto |x|^k$ é estritamente crescente, o que significa que a sequência é crescente. Como essa sequência cai exponencialmente pois $\frac{4}{b} < 1$. A única forma que a sequência de \limsup ser constante é ser 0. ■

Então podemos provar o Teorema 2.2.2

Prova. Fixe $f \in \mathcal{C}_b(\mathbb{R})$, $\epsilon > 0$, $b > 4$. Sabemos que pelo teorema de aproximação de Stone–Weierstrass (STONE, 1948) (ou qualquer livro texto de análise real)

$$\sup_{|x| \leq b} |f(x) - P_\epsilon(x)| < \frac{\epsilon}{6}$$

podemos então usar a desigualdade triangular para estimar

$$\begin{aligned} \left| \int f d\mu_{\mathbf{X}_n} - \int f d\sigma_1 \right| &\leq \left| \int f d\mu_{\mathbf{X}_n} - \int P_\epsilon d\mu_{\mathbf{X}_n} \right| + \left| \int P_\epsilon d\mu_{\mathbf{X}_n} - \int P_\epsilon d\sigma_1 \right| \\ &\quad + \left| \int P_\epsilon d\sigma_1 - \int f d\sigma_1 \right| \end{aligned}$$

Então, o evento $\left| \int f d\mu_{\mathbf{X}_n} - \int f d\sigma_1 \right| > \epsilon$ está contido na união dos três eventos onde cada um dos quatro acima é $> \frac{\epsilon}{3}$. Isso significa que

$$\begin{aligned} \mathbb{P} \left(\left| \int f d\mu_{\mathbf{X}_n} - \int f d\sigma_1 \right| > \epsilon \right) &\leq \mathbb{P} \left(\left| \int f d\mu_{\mathbf{X}_n} - \int P_\epsilon d\mu_{\mathbf{X}_n} \right| > \epsilon/3 \right) \\ &\quad + \mathbb{P} \left(\left| \int P_\epsilon d\mu_{\mathbf{X}_n} - \int P_\epsilon d\sigma_1 \right| > \epsilon/3 \right) \\ &\quad + \mathbb{P} \left(\left| \int P_\epsilon d\sigma_1 - \int f d\sigma_1 \right| > \epsilon/3 \right) \end{aligned}$$

Por construção, $|P_\epsilon - f| < \frac{\epsilon}{6}$ em $[-b, b]$, o que inclui o suporte $[-2, 2]$ de σ_1 assim, a última probabilidade é 0. Para o primeiro termo, dividimos a integral sobre $[-b, b]$ e seu complemento:

$$\begin{aligned} \left| \int (f - P_\epsilon) d\mu_{\mathbf{X}_n} \right| &\leq \int_{|x| \leq b} |f(x) - P_\epsilon(x)| \mu_{\mathbf{X}_n}(dx) \\ &\quad + \int_{|x| > b} |f(x) - P_\epsilon(x)| \mu_{\mathbf{X}_n}(dx). \end{aligned}$$

Pela mesma razão descrita acima, nos podemos fazer a seguinte estimativa

$$\mathbb{P}\left(\left|\int f d\mu_{\mathbf{X}_n} - \int P_\epsilon d\mu_{\mathbf{X}_n}\right| > \epsilon/3\right) \leq \mathbb{P}\left(\int |f - P_\epsilon| \mathbb{1}_{|x| \leq b} d\mu_{\mathbf{X}_n} > \epsilon/6\right) + \mathbb{P}\left(\int |f - P_\epsilon| \mathbb{1}_{|x| > b} d\mu_{\mathbf{X}_n} > \epsilon/6\right).$$

Novamente, por construção, $|f - P_\epsilon| < \frac{\epsilon}{6}$ em $[-b, b]$, e assim como $\mu_{\mathbf{X}_n}$ é uma medida de probabilidade, o primeiro termo é idêntico a 0. Temos, portanto, a estimativa

$$\mathbb{P}\left(\left|\int f d\mu_{\mathbf{X}_n} - \int f d\sigma_1\right| > \epsilon\right) \leq \mathbb{P}\left(\int |f - P_\epsilon| \mathbb{1}_{|x| > b} d\mu_{\mathbf{X}_n} > \epsilon/6\right) \quad (5-1)$$

$$+ \mathbb{P}\left(\left|\int P_\epsilon d\mu_{\mathbf{X}_n} - \int P_\epsilon d\sigma_1\right| > \epsilon/3\right). \quad (5-2)$$

Que esse ultimo termo 5-1 tende a 0 quando $n \rightarrow \infty$ segue imediatamente do Teorema 2.2.3 (já que convergência em probabilidade respeita adição e multiplicação escalar). Então nos resta apenas estimativa 5-2. Fazemos isso da seguinte forma. Seja $k = \deg P_\epsilon$. Como f é limitado, temos $|f(x) - P_\epsilon(x)| \leq \|f\|_\infty + |P_\epsilon(x)|$, e no conjunto $|x| > b$, isso é $\leq c|x|^k$ para alguma constante c (uma vez que $b > 0$). Isso significa que

$$\mathbb{P}\left(\int |f - P_\epsilon| \mathbb{1}_{|x| \geq b} d\mu_{\mathbf{X}_n} > \epsilon/6\right) \leq \mathbb{P}\left(\int c|x|^k \mathbb{1}_{|x| \geq b} d\mu_{\mathbf{X}_n}(dx) > \epsilon/6\right).$$

Assim, pelo lema 5.2.1, essa sequencia tem $\limsup_{n \rightarrow \infty} = 0$. ■

5.3

Prova do Teorema 2.2.1

Finalmente podemos então provar o teorema 2.2.1

Prova. A variável aleatória $E_n(I)$ no enunciado do teorema pode ser escrita como

$$E_n(I) = \frac{1}{n} \cdot \#\{j \in [n] : \lambda_j(\mathbf{X}_n) \in I\} = \mu_{\mathbf{X}_n}(I) \quad (5-3)$$

Ou seja, a conclusão desejada é que $\mu_{\mathbf{X}_n}(I) \rightarrow \sigma_t(I)$ em probabilidade para todos os intervalos I . Por um argumento de escala fácil, podemos tomar $t = 1$, caso em que provamos acima que $\mu_{\mathbf{X}_n} \rightarrow \sigma_1$ fracamente em probabilidade. Por um teorema de convergência padrão na teoria da probabilidade (c.f. Teorema 25.8 (BILLINGSLEY, 1995)), isso implica que $\mu_{\mathbf{X}_n}(A) \rightarrow \sigma_1(A)$ em probabilidade para cada σ_1 contínuo mensurável em A . Mas como σ_1 tem uma

densidade contínua, todo intervalo é um conjunto σ_1 contínuo. Isso prova o teorema.



6 Apêndice

6.1 Norma de Frobenius

Para a definição de Matrizes de Wigner, primeiramente precisamos definir uma sequência de matrizes aleatórias \mathbf{Y}_n , que ainda não estava normalizada. Note que os autovalores de \mathbf{Y}_n , de acordo com as características que pedimos, podem ir para o infinito. Gostaríamos de reescalonar os autovalores de \mathbf{Y}_n , para que no final nossa definição de Matriz de Wigner tenha autovalores uniformemente cotados para cada \mathbf{X}_n .

Faz sentido querermos isso, pois estamos visando que os autovalores convirjam para um função com valores finitos, então começar controlando os autovalores vai nos ser essencial. Vamos então entender como devemos definir α_n de forma apropriada para que os autovalores de uma matriz $\mathbf{X}_n = \alpha_n \mathbf{Y}_n$ sejam como desejamos.

Perceba que bastaria controlar o primeiro e o último autovalor, se os colocarmos em ordem crescente. Além disso, se $\lambda_1, \lambda_2, \dots, \lambda_n$ são os autovalores de \mathbf{X}_n , observe o seguinte:

$$(\lambda_1)^2 + \dots + (\lambda_n)^2 \leq n \cdot \max\{(\lambda_j)^2\} = n \cdot \max\{(\lambda_1)^2, (\lambda_n)^2\}. \quad (6-1)$$

É possível ver que, se λ_1 e λ_n convergem para algum valor real, essa soma é limitada. Essa soma está relacionada com a norma de Frobenius de uma matriz simétrica \mathbf{X}_n :

$$\|\mathbf{X}_n\|_2^2 = \frac{1}{n} \sum_{i=1}^n \lambda_i^2 = \frac{1}{n} \sum_{1 \leq i, j \leq n} X_{ij}^2, \quad (6-2)$$

onde X_{ij} são as entradas da matriz. Está implícita a seguinte

Afirmção 3 *Para matrizes simétricas, somar todos os autovalores ao quadrado equivale a que somar todas as suas entradas ao quadrado.*

Prova.

$$\sum_{i=1}^n \lambda_i^2 = \text{Tr}(\mathbf{X}_n^2) = \sum_{i=1}^n (\mathbf{X}_n^2)_{ii} = \sum_{i=1}^n \left(\sum_{j=1}^n X_{ij} X_{ji} \right) = \sum_{i=1}^n \sum_{j=1}^n X_{ij}^2.$$

■

Observação 15 *Isso facilita as coisas, já que poderemos tratar diretamente da soma do quadrado das entradas, e não da dos autovalores, que seria muito complicado de encontrar no caso de matrizes com entradas aleatórias.*

Devido às equações 6-1 e 6-2, nosso objetivo passa a ser simplesmente que para todo n a norma de Frobenius da matriz X_n seja uniformemente limitada. Queremos limitar os autovalores mas isso é equivalente a limitar a norma de Frobenius na matriz aleatória.

Como a norma de Frobenius será uma variável aleatória, vamos tentar limitar sua esperança. Então, para $\mathbf{X}_n = \alpha_n \mathbf{Y}_n$, teremos

$$\mathbb{E}(\mathbf{X}_n)_2^2 = \frac{1}{n} \alpha_n^2 \sum_{1 \leq i, j \leq n} \mathbb{E}(Y_{ij}^2).$$

Isso ocorre primeiro usando a definição da Norma, depois trocamos \mathbf{X}_n por $\alpha_n \mathbf{Y}_n$.

Por simetria, os elementos dos triângulos superior e inferior são iguais, e portanto

$$\frac{1}{n} \alpha_n^2 \sum_{1 \leq i, j \leq n} \mathbb{E}(Y_{ij}^2) = \frac{\alpha_n^2}{n} \sum_{i=1}^n \mathbb{E}(Y_{ii}^2) + \frac{2\alpha_n^2}{n} \sum_{1 \leq i < j \leq n} \mathbb{E}(Y_{ij}^2).$$

$$\frac{1}{n} \alpha_n^2 \sum_{1 \leq i, j \leq n} \mathbb{E}(Y_{ij}^2) = \frac{\alpha_n^2}{n} \sum_{i=1}^n \mathbb{E}(Y_{ii}^2) + \frac{2\alpha_n^2}{n} \sum_{1 \leq i < j \leq n} \mathbb{E}(Y_{ij}^2).$$

Mas como temos distribuições i.i.d, as duas somas são somas de termos iguais, portanto basta contar a quantidade de termos. Na primeira soma, temos n termos iguais, e na segunda temos $\frac{n(n-1)}{2}$ elementos no triângulo superior de uma matriz (sem contar a diagonal). Sendo assim, podemos concluir que

$$\mathbb{E}(\mathbf{X}_n)_2^2 = \alpha_n^2 \cdot \mathbb{E}(Y_{11}^2) + (n-1)\alpha_n^2 \cdot \mathbb{E}(Y_{12}^2).$$

Note, que se $\mathbb{E}(Y_{12}^2) = 0$, isso implica que estamos no caso degenerado em que $Y_{ij} = 0$ sempre que $i \neq j$, isto é, no caso de uma matriz diagonal. Vamos simplesmente assumir que $\mathbb{E}(Y_{12}^2) > 0$. Nesse caso, tomando $\alpha_n = n^{-\frac{1}{2}}$, vemos que $\mathbb{E} \|\mathbf{X}_n\|_2^2 \rightarrow \mathbb{E}(Y_{12}^2)$, e como isso é limitado pela condição 3 da Seção 2.1, temos o que desejávamos.

Observação 16 *Perceba que o fato*

$$\mathbb{E}(\mathbf{X}_n)_2^2 \rightarrow \mathbb{E}(Y_{12}^2)$$

explica o porquê de considerarmos apenas o segundo momento fora da diagonal no Teorema de Wigner 2.2.1, mas não com o segundo momento da diagonal.

6.2 Sequência de Catalan

A sequência de Catalan é uma sequência de números naturais que aparece frequentemente no contexto de Combinatória em problemas de contagem. Como referência para esta seção, confira o capítulo 6 de (STANLEY, 1999).

Definição 6.2.1 A sequência de Catalan é dada pela fórmula

$$C_n = \frac{1}{n+1} \binom{2n}{n} = \frac{(2n)!}{(n+1)!n!} = \prod_{k=2}^n \frac{n+k}{k} \quad \forall n \geq 0.$$

Os primeiros números de Catalan são

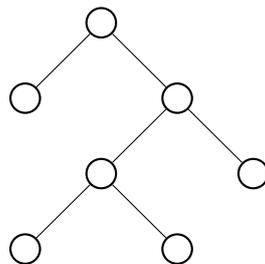
$$1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, \dots$$

Alguns exemplos de problemas de contagem onde os números de Catalan são a resposta incluem:

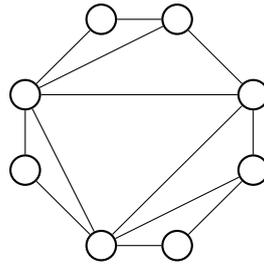
1. A quantidade de maneiras de distribuir n pares de parênteses, de forma que todo “)” pareie com um “(” anterior. Para $n = 3$, temos as 5 combinações

$$((())), (())(), (())(), ()(()), ()()().$$

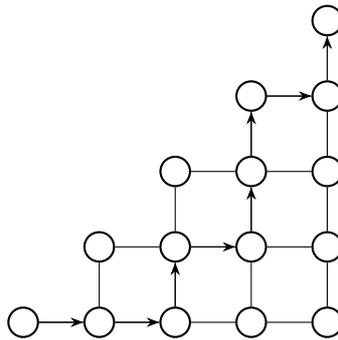
2. A quantidade de árvores binárias com exatamente $n + 1$ folhas.



3. A quantidade de triangulações de um polígono de $n + 2$ lados.



4. A quantidade de caminhos em um triângulo reticulado com vértices $(0, 0)$, $(n, 0)$ e (n, n) que começam em $(0, 0)$ e terminam em (n, n) e que cada passo consiste em andar na direção $(1, 0)$ ou na direção $(0, 1)$.



Para provar que os números de Catalan são a resposta a certo problema de contagem, é natural prová-lo de maneira recursiva. Mais precisamente, a sequência de Catalan também está unicamente determinada recursivamente, de forma que

$$C_0 = 1 \quad \text{e} \quad C_{n+1} = \sum_{i=0}^n C_i C_{n-i} \quad \forall n \geq 0,$$

ou também como

$$C_0 = 1 \quad \text{and} \quad C_{n+1} = \frac{2(2n+1)}{n+2} C_n.$$

Lema 6.2.1 *A quantidade de maneiras de distribuir n pares de parênteses é C_n*

Prova. Digamos que a quantidade de maneiras de distribuir n pares de parênteses é x_n . É trivial que $x_0 = 1$ e não é difícil perceber que satisfaz a mesma recorrência que C_n . De fato, para distribuir $n + 1$ pares de parênteses, precisamos começar com um parênteses abrindo, e este vai precisar se fechar em algum momento:

$$\left(\underbrace{\dots}_{i \text{ pares}} \right) \underbrace{\dots}_{n-i \text{ pares}}$$

Digamos que entre o primeiro parênteses abrindo e seu fechamento há i pares de parênteses, onde i pode ser qualquer valor entre 0 e n . Uma vez fixado o i , precisamos organizar de forma independente tanto os i pares de dentro (de x_i possíveis maneiras), quanto os $n - i$ pares de fora da primeira dupla de parênteses (de x_{n-i} possíveis maneiras). Então temos para cada i entre 0 e n , $x_i x_{n-i}$ formas de fazer isso.

Fica claro que

$$x_{n+1} = \sum_{i=0}^n x_i x_{n-i} \quad \forall n \geq 0.$$

Logo, por indução, isso prova que $x_n = C_n$ para todo $n \geq 0$. ■

Para os fins desta dissertação, estaremos especialmente interessados seguinte conjunto:

Definição 6.2.2 Para todo k par, conjunto de caminhos de Dyck de tamanho k é:

$$\mathcal{D}_k = \left\{ (d_1, \dots, d_k) \in \{\pm 1\} : \sum_{i=1}^j d_i \geq 0 \quad \forall j \in \{1, \dots, k\} \text{ e } \sum_{i=1}^k d_i = 0 \right\}.$$

Lema 6.2.2 Vale que $|\mathcal{D}_k| = C_{\frac{k}{2}}$

Prova. O problema de contar $|\mathcal{D}_k|$ é análogo ao de contar como distribuir $\frac{k}{2}$ pares de parênteses. Isso porque podemos associar cada caminho de Dyck a uma sequência de parênteses abrindo e fechando. Se $d_i = 1$, o substituímos pelo parêntese “(”, e se $d_i = -1$, por um parêntese “)”. A condição

$$\sum_{i=1}^j d_i \geq 0 \text{ para } 1 \leq j \leq k$$

garante que nunca um parênteses fecha sem ter um aberto correspondente, e a condição

$$\sum_{i=1}^k d_i = 0$$

garante que todos os parênteses se fecham. Por construção isto define uma bijeção. ■

6.3 Codigo em Python

```
1 import matplotlib.pyplot as plt
2 import numpy as np
```

```
3 from statistics import mean
4 from scipy.stats import unitary_group
5 from random import randint
6 from numpy.linalg import matrix_power
7
8 #dimension and amount of random matrices
9 n = 500
10 amount = 100
11 np.random.seed(42)
12
13 #Gaussian Orthogonal Ensemble - GOE
14
15 splits_GOE = []
16 eigenV_GOE = []
17
18 for r in range(amount):
19     A = np.random.normal(scale=n**-0.5, size=(n,n))
20
21     M = 0.5*(A + A.T)
22
23     w = np.linalg.eigvalsh(M)
24     eigenV_GOE.append(w)
25
26     for x in range(n - 1):
27         splits_GOE.append((w[x+1]-w[x])/mean(w))
28
29 eigenV_GOE_flat_list = list(np.concatenate(eigenV_GOE).flat)
30
31 binssize = 0.01
32 bins = np.arange(-1.5, 1.5, binssize)
33
34 plt.hist(eigenV_GOE_flat_list, bins = bins)
35 plt.show()
```

7

Referências bibliográficas

WIGNER, E. P. Characteristic vectors of bordered matrices with infinite dimensions. **Annals of Mathematics**, v. 62, n. 3, p. 548–564, 1955.

TAO, T. **Topics in Random Matrix Theory**. [S.l.]: American Mathematical Society, 2010. (Graduate Studies in Mathematics Volume 132). ISBN 978-0-8218-7430-1.

AKEMANN, G.; BAIK, J.; FRANCESCO, P. D. **The Oxford Handbook of Random Matrix Theory**. [S.l.]: Oxford University Press, 2011. ISBN 978-0-19-957400-1.

QI, F.; GUO, B.-N. Integral representations of the Catalan numbers and their applications. **Mathematics**, v. 5, n. 3: 40, 2017. (31 páginas).

KEMP, T. **Introduction to Random Matrix Theory**. 2016. Disponível em: <<https://mathweb.ucsd.edu/~tkemp/RM.LectureNotes.pdf>>.

STONE, M. H. The generalized Weierstrass approximation theorem. **Mathematics Magazine**, v. 21, n. 4, p. 167–184, 1948.

BILLINGSLEY, P. **Probability and Measure Third Edition**. [S.l.]: Wiley-Interscience, 1995. ISBN 0-471-00710-2.

STANLEY, R. P. **Enumerative Combinatorics Volume 2**. [S.l.]: Cambridge University Press, 1999. (Cambridge Studies in Advanced Mathematics 62). ISBN 0-521-56069-1.