Universidade Federal do Ceará
Departamento de Computação
Mestrado e Doutorado em Ciência da Computação

Sequências Convergentes e Testabilidade

Antonio Josefran de Oliveira Bastos

Fortaleza – Ceará
Fevereiro 2012

Universidade Federal do Ceará
Centro de Ciências
Departamento de Computação
Mestrado e Doutorado em Ciência da Computação

Sequências Convergentes e Testabilidade

Autor
Antonio Josefran de Oliveira Bastos

Orientador
Prof. Dr. Rudini Menezes Sampaio
Coorientador
Prof. Dr. Carlos Eduardo Fish de Brito

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Ciência da Computação da Universidade Federal do Ceará como parte dos requisitos para obtenção do título de Mestre em Ciência da Computação.

Fortaleza – Ceará
Fevereiro 2012

Resumo

Neste trabalho, estudamos a teoria recente de convergência de sequências de grafos e suas extensões para permutações e ordens parciais de dimensão fixa. Conjecturamos um lema de regularidade fraca de grafos em intervalos que, se for verdadeira, nos possibilita estender essa teoria para grafos ordenados, que são grafos tais que existe uma ordem total entre os vértices. Mostramos algumas relações interessantes de permutações e ordens parciais com grafos ordenados. Com isso, conseguimos uma prova alternativa para a existência de objetos limites de qualquer sequência convergente de permutações. Provamos também que toda propriedade hereditária de permutações ou grafos ordenados é testável.

PALAVRAS-CHAVE: Testabilidade, Sequências Convergentes, Propriedade Testável

Abstract

In this work, we studied the recent theory of convergent graph sequences and its extensions to permutation and partially ordered sets with fix dimension. We've conjectured a lemma of weak regularity on intervals that, if this conjecture is true, we can extend this theory to ordered graphs, which are graphs such that there is a total order on its vertices. We show some interesting relations on permutation and partially ordered sets with ordered graphs. Then, we obtain another proof to the existence of limit objects for all convergent permutation sequences. We also proved that all hereditary property of either permutation or ordered graph is testable.

KEY-WORDS: Testability, Convergence Sequences, Properties Testable

Agradecimentos

Seria difícil conseguir agradecer a todos que, de alguma forma, me ajudaram a conseguir realizar esse trabalho sem ter que escrever um trabalho a parte apenas para isso. Porém existem algumas pessoas nas quais eu direciono um agradecimento especial. São elas:

Em primeiro lugar, gostaria de agradecer a Deus por tudo.

A minha família. Aos meus pais, Bastos e Neide, que me criaram e educaram, que me ensinaram o valor do trabalho e me serviram de exemplos para a vida e que sacrificaram uma boa parte da vida deles para minha formação e dos meus irmãos sem esperar nada em troca. Com toda certeza, este trabalho não seria possível sem eles. A minha tia que é como uma segunda mãe para mim, Marineide, que sempre me incentivou a seguir adiante e nunca desistir. Aos meus irmãos, Patrícia e Jofran, que me deram apoio e pegaram bastante no meu pé e mesmo com alguns desentendimentos, sempre me ajudaram e foram muito pacientes comigo... De fato, verdadeiros irmãos.

A meu orientador, Rudini Sampaio, e o meu coorientador, Carlos Brito, pelo tempo dedicado, pela confiança que em mim foi depositada, pelos conselhos e pela amizade. Nosso trabalho passou, em muito, os laços mestre e aluno e as lições que me foram ensinadas durante este trabalho irão, com certeza, perpendurar pelas muitas fases que ainda irei passar.

Ao grupo de pesquisa, ou devo dizer família, ParGO, que me acolheu ao longo desses dois anos e que me fizeram sentir parte da família. Em especial ao professor Manoel Campelo que me apresentou a família e sempre foi muito paciente comigo.

A minha namorada, Herbia Castro, que em momentos de grande aperto estava ao meu lado e que varias vezes se sacrificou para me ajudar, e também pela compreensão de vários fins de semana sacrificados para estudar.

Ao Carlos Filho(mingau), Markos, Yuri, Pablo Mayckon, Jonas e o pessoal do NerdProud Futebol Clube, pelos incontáveis conselhos e almoços filosóficos (com filósofo mingau e as filosofias que não se repetem) e por vários outros momentos que proporcionaram boas gargalhadas.

Finalmente, a FUNCAP - Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico - pelo financiamento da bolsa de estudos para a realização do presente trabalho.

Sumário

Lista de Figuras

Capítulo 1

Introdução

1.1 Sequências de Grafos

L. Lovász e B. Szegedy [19] definiram o conceito de sequência convergente de grafos. Dados grafos simples F e G, seja $hom(F,G)$ o número de homomorfismos de F em G, ou seja, o número de mapeamentos $\Phi : V(F) \to V(G)$ que preservam arestas: Se $xy \in E(F)$ então $\Phi(x)\Phi(y) \in E(G)$. A densidade de homomorfismos de F em G é definida como:

$$t(F,G) = \frac{hom(F,G)}{|V(G)|^{|V(F)|}}.$$

Definição 1. *Uma sequência de grafos (G_n) é convergente se $|V(G_n)| \to \infty$ e, para todo grafo F, a sequência de valores reais $t(F,G_n)$ converge.*

Seja $inj(F,G)$ o número de homomorfismos injetivos de F em G. O número de subgrafos F em G é igual a $inj(F,G)$ dividido o número de automorfismos de F. Seja $(n)_k = n(n-1)\ldots(n-k+1)$. A densidade de homomorfismos injetivos é definida como

$$t_{inj}(F,G) = \frac{inj(F,G)}{(|V(G)|)_{|V(F)|}}.$$

Seja $ind(F,G)$ o número de homeomorfismos injetivos de F em G que também preservam não-arestas, ou seja, se $xy \notin E(F)$ então $\Phi(x)\Phi(y) \notin E(G)$. Definimos $t_{ind}(F,G)$ como

$$t_{ind}(F,G) = \frac{ind(F,G)}{(|V(G)|)_{|V(F)|}}.$$

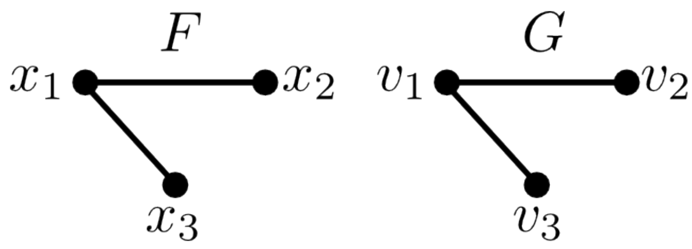

Figura 1.1: F e G

Ex. 1. *Para ilustrar a diferença entre as densidades descritas acima, observe os seguintes grafos:*

Note que, se quisermos apenas preservar as vizinhanças, devemos escolher $\phi : V(F) \to V(G)$ *tal que* $(\phi(x_1), \phi(x_2), \phi(x_3)) \in \{(v_2, v_1, v_1), (v_1, v_2, v_2), (v_3, v_2, v_2), (v_2, v_3, v_3), (v_3, v_1, v_1), (v_1, v_3, v_3), (v_1, v_2, v_3), (v_1, v_3, v_2), (v_2, v_1, v_3), (v_2, v_3, v_1), (v_3, v_1, v_2), (v_3, v_2, v_1)\}$. *Para preservar a vizinhança e a injetividade devemos escolher* $\phi_{inj} : V(F) \to V(G)$ *tal que* $(\phi_{inj}(x_1), \phi_{inj}(x_2), \phi_{inj}(x_3)) \in \{(v_1, v_2, v_3), (v_1, v_3, v_2), (v_2, v_1, v_3), (v_2, v_3, v_1), (v_3, v_1, v_2), (v_3, v_2, v_1)\}$. *Finalmente, observe que o grafo* G *não possui nenhum subgrafo induzido isomorfo a* F. *Desta forma temos*

$$t(F, G) = \frac{hom(F, G)}{|V(G)|^{|V(F)|}} = \frac{12}{3^3} = \frac{4}{9},$$

$$t_{inj}(F, G) = \frac{inj(F, G)}{(|V(G)|)_{|V(F)|}} = \frac{6}{3 \cdot 2 \cdot 1} = 1,$$

$$t_{ind}(F, G) = \frac{ind(F, G)}{(|V(G)|)_{|V(F)|}} = \frac{0}{3 \cdot 2 \cdot 1} = 0.$$

L. Lovász e B. Szegedy [19] observaram que as densidades acima estão relacionadas da seguinte forma: Para quaisquer grafos F e G temos:

$$|t_{inj}(F, G) - t(F, G)| \leq \binom{|V(F)|}{2} \frac{1}{|V(G)|},$$

$$t_{ind}(F, G) = \sum_{\substack{V(F')=V(F) \\ E(F')\supseteq E(F)}} (-1)^{|E(F')\setminus E(F)|} t_{inj}(F, G).$$

Portanto, para toda sequência convergente (G_n), temos que $t(F, G_n)$, $t_{inj}(F, G_n)$ e $t_{ind}(F, G_n)$ convergem para todo grafo F.

Ex. 2. *A sequência (H_n) de grafos, tal que $E(H_n) = \emptyset$ e $|V(H_n)| \to \infty$, converge, pois para todo grafo F, e n suficientemente grande, temos*

$$t(F, H_n) = t_{inj}(F, H_n) = t_{ind}(F, H_n) = \begin{cases} 0 & \textit{se } E(F) \neq \emptyset \\ 1 & \textit{caso contrário} \end{cases}$$

Ex. 3. *A sequência (K_n) de grafos completos, onde $|V(K_n)| \to \infty$, converge, pois para todo grafo F, e n suficientemente grande, temos*

$$t(F, K_n) = t_{inj}(F, K_n) = 1,$$

e

$$t_{ind}(F, K_n) = \begin{cases} 1 & \textit{se } F \textit{ é completo} \\ 0 & \textit{caso contrário} \end{cases}$$

Ex. 4. *Dado $0 \leq p \leq 1$ fixo, a sequência de grafos baseada no modelo de Erdős e Rényi, $(G_{n,p})$, converge com probabilidade 1, pois para todo grafo F temos que*

$$t(F, G_{n,p}) \xrightarrow[n \to infty]{} p^{|E(F)|}.$$

É intuitivo imaginar que se a sequência converge então ela converge para algum lugar. Porém olhar apenas para os grafos não nos fornece uma objeto limite para o qual seja viável trabalhar. Lovász e Szegedy provaram a existência de um "objeto limite" para qualquer sequência convergente de grafos.

Definição 2. *Um* graphon *é uma função simétrica mensurável $W : [0,1]^2 \to [0,1]$.*

Graphons podem ser visualizados como grafos ponderados infinitos (ou matrizes de adjacências infinitas) com vértices no intervalo $[0,1]$. Dado $n > 0$, seja $[n] = \{1, \ldots, n\}$. Para todo grafo simples G de ordem n, podemos construir um *graphon* W_G da seguinte maneira:

Definição 3. *Dado um grafo G com vértices em $[n]$, seja W_G o* graphon *o grafo a seguir: Para todo $x, y \in (0,1]$ temos*

$$W_G(x,y) = \begin{cases} 1 & \textit{se } ij \textit{ é aresta de } G, \textit{ onde } i = \lceil nx \rceil \textit{ e } j = \lceil ny \rceil, \\ 0 & \textit{caso contrário} \end{cases}$$

Se $x = 0$, respectivamente $y = 0$, tome $i = 1$, respectivamente $j = 1$, na igualdade anterior.

Ex. 5. *A figura abaixo mostra o exemplo de um grafo associado ao seu* graphon.

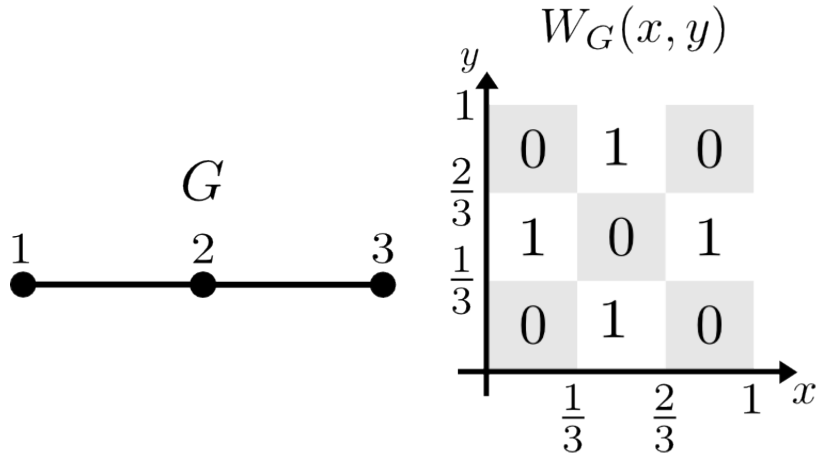

Figura 1.2: Grafo G e *Graphon* W_G

Assim como é possível gerar um *graphon* a partir de um grafo, podemos gerar um grafo, com qualquer número de vértices, a partir de um *graphon*.

Definição 4. *Dado um* graphon W*, definimos o grafo* W*-aleatório* $G(n,W)$ *de ordem* n *da seguinte forma:*

1. *Geramos independentemente* n *valores* $X_1, X_2, \ldots, X_n$ *segundo a distribuição uniforme em* $[0,1]$*;*

2. *Ligamos os vértices* $i, j \in [n]$ *com probabilidade* $W(X_i, X_j)$*.*

Note que o item *2* da definição acima pode ser substituido pelos seguintes passos:

2^* Para cada par de vértice $i, j \in [n]$ sorteamos uniformemente um valor $Y_{ij} \in [0,1]$;

3^* Ligamos os vértices $i, j \in [n]$ se $Y_{ij} \leq W(X_i, X_j)$.

O modelo $G(n,W)$ proposto por L. Lovász e B. Szegedy [19] é uma generalização do modelo $G(n,p)$ proposto por Erdős e Rényi [12]. Para isso, basta tomar o *graphon* W tal que $W(x,y) = p$ para todo $(x,y) \in [0,1]^2$.

Para todo graphon W e todo grafo simples F com k vértices, estendemos a definição de densidade de homomorfismo para *graphons*

$$t(F, W) = \int_{[0,1]^k} \Big(\prod_{ij \in E(F)} W(x_i, x_j) \Big) dx_1 \dots dx_k.$$

Com isso L. Lovász e B. Szegedy [19] provaram que toda sequência convergente possui um objeto limite para o qual as sequências de densidades convergem. Mais formalmente temos:

Teorema 5 ([19])**.** *Para toda sequência convergente* (G_n) *de grafos simples, existe um* graphon $W : [0,1]^2 \to [0,1]$ *tal que, para todo grafo simples* F*, temos que*

$$\lim_{n \to \infty} t(F, G_n) = t(F, W).$$

Provou-se ainda que todo *graphon* é limite de alguma sequência convergente de grafos.

Teorema 6 ([19])**.** *Dado um* graphon W*, a sequência* $(G(n, W))$*, para* $n \to \infty$*, converge para* W *com probabilidade 1.*

Considerando a noção de convergência aqui definida. O Teorema 6 mostra que todo *graphon* é um ponto de acumulação do espaço de grafos.

Definição 7. *Seja* G *e* G' *grafos sobre o mesmo conjunto de vértice de ordem* n*. Definimos a distância retangular como sendo:*

$$d_\square(G, G') = \frac{1}{n^2} \max_{S,T \subseteq V} |e_G(S,T) - e_{G'}(S,T)|$$

onde $e_G(S,T) = \#\{(ij \in E(G); i \in S \text{ e } j \in T\}$*.*

A distância acima satisfaz, trivialmente, as condições de não negatividade e de simetria. Para provar a desigualdade triângular basta notar que dado um grafo F com vértices em V temos que

$$\max_{S,T \subseteq V} |e_G(S,T) - e_{G'}(S,T)| = \max_{S,T \subseteq V} |(e_G(S,T) - e_F(S,T)) + (e_F(S,T) - e_{G'}(S,T)|$$

$$\le \max_{S,T \subseteq V} \Big(|(e_G(S,T) - e_F(S,T))| + |(e_F(S,T) - e_{G'}(S,T)| \Big)$$

$$\le \max_{S,T \subseteq V} |(e_G(S,T) - e_F(S,T))| + \max_{S,T \subseteq V} |(e_F(S,T) - e_{G'}(S,T)|.$$

Note, a grosso modo, que a distância retangular mede o quanto o grafo G se difere do grafo G' em volume de arestas. A definição de distância retangular pode ser estendida para *graphons*. Dados *graphons* W_1 e W_2, a distância retangular entre eles é:

$$d_{\Box}(W_1, W_2) = \sup_{A,B \subseteq [0,1]} \Big| \int_A \int_B \Big(W_1(x,y) - W_2(x,y)\Big) dxdy \Big|.$$

Observe que, de maneira análoga ao caso de distância retangular entre grafos, verificamos diretamente as condições para ser uma distância.

L. Lovász e B. Szegedy [19] relacionaram a distância retangular com a densidade de homomorfismos através do seguinte lema:

Lema 8 ([19]). *Sejam U e W dois* graphons. *Então para todo grafo simples F temos:*

$$|t(F,U) - t(F,W)| \leq |E(F)| d_{\Box}(U,W).$$

Sejam G e G' grafos com mesmo conjunto de vértices de ordem n. E assim, para todo grafo F, temos que

$$|t(F,G) - t(F,G')| \leq |E(F)| d_{\Box}(G,G').$$

Dado um grafo simples G de ordem n, definimos por $\mathcal{K}_G$ o conjunto de todos os $n!$ grafos obtidos por rotulações distintas de G. Podemos estender a distância retangular para dois grafos quaisquer através da seguinte distância. Dados W_1 e W_2 *graphons* e G e H grafos simples, em que $|V(G)| \geq |V(H)|$, definimos a distância retangular $\delta_{\Box}$ da seguinte forma:

1. Entre grafos:
$$\delta_{\Box}(G,H) = \min_{\substack{H' \in \mathcal{K}_H \\ G' \in \mathcal{K}_G}} d_{\Box}(W_{H'}, W_{G'});$$

2. Entre grafo e *graphons*:
$$\delta_{\Box}(G,W) = \min_{G' \in \mathcal{K}_G} d_{\Box}(W, W_{G'});$$

3. Entre *graphons*:
$$\delta_{\Box}(W_1, W_2) = d_{\Box}(W_1, W_2).$$

Seja $\mathcal{G}$ o conjunto formado por todos os infinitos grafos possíveis e $\mathcal{W}$ o conjunto formado por todos os infinitos *graphons* possíveis. Em [6], mostrou-se que toda sequência convergente é de Cauchy com relação a distância $\delta_{\Box}$. Desta forma eles provaram que o espaço métrico $(\mathcal{G} \cup \mathcal{W}, \delta_{\Box})$ é completo.

1.2 Testabilidade

Na seção anterior apresentamos o conceito de sequências convergente de grafos. Nesta seção iremos apresentar uma aplicação para a teoria de sequências convergentes.

Definição 9. *Uma propriedade $\mathcal{P}$ de grafos é dita hereditária se, para qualquer grafo G que possua a propriedade $\mathcal{P}$, todo subgrafo de G também possui a propriedade $\mathcal{P}$.*

Definição 10. *Dizemos que um grafo G é ε-distante de uma propriedade $\mathcal{P}$ se não for possível obter um grafo $G' \in \mathcal{P}$ adicionando ou removendo εn^2 arestas de G.*

Ex. 6. *Observe os seguintes grafos:*

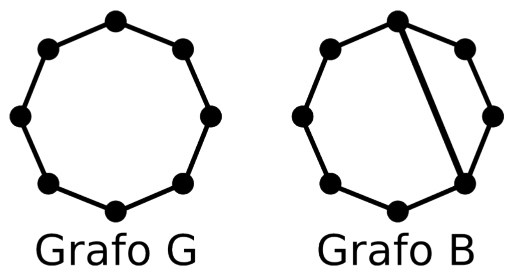

Figura 1.3:

Note que o grafo G é $\frac{1}{64}$-far do grafo B, pois neste caso basta adicionar uma aresta ao grafo G para que eles se tornem isomorpho ao grafo B.

Com isso, podemos definir a condição para uma propriedade ser testável, segundo a definição apresentada por Goldreich et al. [14]:

Definição 11. *Seja $\mathcal{P}$ uma propriedade de grafos. Dizemos que $\mathcal{P}$ é testável se, para todo $\varepsilon > 0$, existe um inteiro positivo $k(\varepsilon)$ e um algoritmo aleatório $\mathcal{T}$, chamado de* testador*, tal que dado um grafo $G(V, E)$, com $|V| > k$, as seguintes propriedades são satisfeitas:*

1. *O número de consultas feitas pelo algoritmo $\mathcal{T}$ depende apenas de k e não do tamanho da entrada.*

2. *Se G satisfaz $\mathcal{P}$, então o* testador *identifica isso com probabilidade maior que $1 - \varepsilon$.*

3. *Se G é "ε-distante" de satisfazer $\mathcal{P}$, então o testador confirma que G não satisfaz $\mathcal{P}$ com probabilidade maior que $1 - \varepsilon$.*

Observe que afirmar que uma propriedade de grafos $\mathcal{P}$ é testável significa que podemos verificar em *tempo constante* e com alta probabilidade se um dado grafo G, de ordem arbitrária, satisfaz $\mathcal{P}$.

Seja f um parâmetro de grafos. Assumimos f normalizado, ou seja,

$$0 \leq f \leq 1.$$

Definição 12. *Dizemos que f é testável se, para todo $\varepsilon > 0$, existe um inteiro positivo $k(\varepsilon)$ tal que dado um grafo G com $|V(G)| > k$, se selecionarmos aleatoriamente um subgrafo induzido G' com pelo menos k vértices temos que:*

$$\mathbb{P}(|f(G) - f(G')| > \varepsilon) < \varepsilon$$

Antes de apresentarmos alguns exemplos de propriedades testáveis iremos apresentar uma desigualdade muito poderosa que será utilizada nos exemplos e em algumas demonstrações.

Definição 13. *Dado uma constante $c \in \mathbb{R}$, dizemos que uma função $f(x_1, \ldots, x_n)$ é c-Lipschitz se $|f(x) - f(x')| \leq c$ para todos vetores x e x' que diferem em apenas uma única coordenada (ou seja, $x = (x_1, \ldots, x_n)$, $x' = (x'_1, \ldots, x'_n)$ e $x_i = x'_i$ para todo $i \in [n] - \{j\}$ para algum $j \in [n]$).*

Teorema 14 (Método das Diferenças Limitadas [11])**.** *Para toda função c-Lipschitz com n parâmetros $f(x_1, \ldots, x_n)$, temos que, se $X_1, \ldots, X_n$ são escolhidos aleatória e independentemente, então $\forall t > 0$*

$$\mathbb{P}\Big(\big|f(X_1, \ldots, X_n) - \mathbb{E}(f(X_1, \ldots, X_n))\big| \leq t\Big) \;\leq\; 2\exp\left\{\frac{-2t^2}{c^2 n}\right\}$$

Ex. 7. *Seja G um grafo com n vértices e sejam H e H' dois subgrafos induzidos de G com k vértices que diferem em apenas um vértice (ou seja, H e H' tem $k-1$ vértices de G em comum). A diferença entre o número de arestas de H e H' é no máximo de $k - 1$ (pois o vértice diferente pode formar arestas com no máximo todos os demais $k - 1$ vértices do subgrafo). Portanto, o número de arestas $e(H)$ é $(k-1)$-Lipschitz. Se H é um subgrafo aleatório de G com k vértices, então*

$$\mathbb{E}(e(H)) = \frac{e(G)}{\binom{n}{2}} \cdot \binom{k}{2}$$

.

Portanto, pelo método das diferenças limitadas

$$\mathbb{P}\Big(\big|e(H) - \frac{e(G)}{\binom{n}{2}} \cdot \binom{k}{2}\big| \leq \varepsilon\binom{k}{2}\Big) \;\leq\; 2\exp\left\{\frac{-2\varepsilon^2\binom{k}{2}^2}{(k-1)^2 \cdot k}\right\}$$

Portanto, tomando $k = 6/\varepsilon^3$ e $n > k$, temos

$$\mathbb{P}\Big(\big|de(H) - de(G)\big| \leq \varepsilon\Big) \;\leq\; 2\exp\Big\{-\frac{1}{2}\varepsilon^2 k\Big\} \;\leq\; \varepsilon$$

onde $de(G) = e(G)/\binom{n}{2}$ (resp. $de(H) = e(H)/\binom{k}{2}$) é a densidade de arestas de G (resp. H).

Portanto, por definição, a densidade de arestas é um parâmetro testável de grafos.

Ex. 8. *Como feito no exemplo anterior, seja G um grafo com n vértices e sejam H e H' dois subgrafos induzidos de G com k vértices que diferem em apenas um vértice. A diferença entre o número de triângulos de H e H' é de no máximo $\binom{k-1}{2}$ (pois o vértice diferente pode formar triângulos com no máximo todos os demais $\binom{k-1}{2}$ pares de vértices do subgrafo). Portanto, o número de triângulos $tr(H)$ é $\binom{k-1}{2}$-Lipschitz. Se H é um subgrafo aleatório de G com k vértices, então*

$$\mathbb{E}(tr(H)) = \frac{tr(G)}{\binom{n}{3}} \cdot \binom{k}{3}.$$

Portanto, pelo método das diferenças limitadas

$$\mathbb{P}\Big(\big|tr(H) - \frac{tr(G)}{\binom{n}{3}} \cdot \binom{k}{3}\big| \leq \varepsilon\binom{k}{3}\Big) \;\leq\; 2\exp\left\{\frac{-2\varepsilon^2\binom{k}{3}^2}{\binom{k-1}{2}^2 \cdot k}\right\}$$

Portanto, tomando $k = 6/\varepsilon^3$ e $n > k$, temos

$$\mathbb{P}\Big(\big|dtr(H) - dtr(G)\big| \leq \varepsilon\Big) \;\leq\; 2\exp\Big\{-\frac{2}{9}\varepsilon^2 k\Big\} \;\leq\; \varepsilon,$$

onde $dtr(G) = tr(G)/\binom{n}{3}$ (resp. $dtr(H) = tr(H)/\binom{k}{3}$) é a densidade de triângulos de G (resp. H).

Portanto, por definição, a densidade de triângulos é um parâmetro testável de grafos.

Do mesmo modo, com argumento similar aos feitos nos exemplos anteriores, as propriedades de ser um grafo vazio, ou de ser um grafo completo ou de ser um grafo livre de triângulos são propriedades testáveis de grafos.

Definição 15. *Para dois grafos G e G' com o mesmo conjunto de nós, definimos por*

$$d_1(G, G') := \frac{|E(G)\Delta E(G')|}{|V(G)|^2}.$$

a diferença de arestas entre G e G' normalizado por $|V(G)|^2$.

Note que a distância acima não é muito similar a $d_\square$, pois ela verifica a diferença entre os dois grafos aresta a aresta. Podemos estender a definição da distância d_1 para as propriedades. Para todo grafo G e um propriedade de grafos $\mathcal{P}$, denotamos por $d_1(G, \mathcal{P}) = \min_{G' \in \mathcal{P}} d_1(G, G')$. Note que, para todo grafo F fixo, podemos olhar para $d_1(F, G)$ como um parâmetro de um grafo G. Com isso, em [2], Alon e Shapira provaram o seguinte teorema:

Teorema 16 ([2])**.** *A distância d_1 para uma propriedade hereditária de grafos é testável.*

E consequentemente provaram o teorema principal de [2]:

Teorema 17 ([2])**.** *Toda propriedade hereditária de grafos é testável.*

L. Lovász e B. Szegedy [19] provaram os teoremas acima de maneira mais simples e curta utilizando o conceito de sequências convergentes.

Outros resultados de testabilidade foram obtidos para outras estruturas. N. Alon et al. [2] mostraram que toda propriedade hereditária de grafos é testável. N. Alon e A. Shapira [3] provaram que toda propriedade monótona de grafos é testável. T. Austin e T. Tao [4] provaram que toda propriedade hereditária para hipergrafos, que não são mistos, é testável. A. J. O. Bastos et al. [5] mostramos que toda propriedade hereditária de permutações é testável.

Neste trabalho, apresentaremos o conceito de sequências convergentes para permutações e ordens parciais, assim como os resultados obtidos em testabilidade. Estendemos o conceito de sequência convergente para grafos ordenados e provamos a existência de um objeto limite baseado em uma conjectura de regularidade fraca para grafos em intervalos. Com isso, mostramos uma outra prova para a existência de um objeto limite para permutações e posets. Provamos ainda a unicidade do objeto limite. Por fim, mostramos que toda propriedade hereditária de grafos ordenados é fortemente testável.

Capítulo 2

Permutações

Os resultados mostrados para grafos através de sequência convergentes, serviram de motivação para mostrar resultados análogos para permutações. Neste capítulo, apresentaremos o conceito de sequências convergentes de permutações e testabilidade. Na Seção 2.1, apresentaremos a definição e notações utilizadas neste capítulo. Na Seção 2.2, apresentaremos a noção de sequência de permutações assim como a condição para convergência, introduziremos um conceito de distância retangular entre permutações e apresentaremos o teorema de Cauchy para convergência. Na Seção 2.3, apresentaremos o conceito de testabilidade fraca. Na Seção 2.4, apresentaremos o conceito de testabilidade forte e alguns resultados. Na Seção 2.5, apresentaremos o conceito de quase-aleatoridade para grafos e para permutações.

2.1 Introdução

Definição 18. *Uma permutação é uma função bijetiva $\sigma : A \to A$ onde A é um conjunto enumerável.*

Como todo conjunto enumerável possui uma bijeção com os números naturais, podemos, por simplicidade, definir uma permutação com n elementos como sendo uma bijeção $\sigma : [n] \to [n]$. Podemos representar uma permutação apenas mostrando a ordem dos elementos da imagem. Por exemplo, a permutação $\sigma : [3] \to [3]$ tal que $\sigma(1) = 2, \sigma(2) = 3, \sigma(3) = 1$ pode ser representáda através da ordem dos elementos da imagem, ou seja, $\sigma = (2, 3, 1)$.

Definição 19. *Dado um conjunto $A = \{a_1, a_2, \dots, a_k\} \subset [n]$ crescente, e uma permutação σ em $[n]$ definimos a permutação de σ restrita a A como sendo $\sigma^A : [|A|] \to [|A|]$ onde para todo $i, j \in [k]$ temos que $\sigma^A(i) < \sigma^A(j)$ se e somente se $\sigma(a_i) < \sigma(a_j)$.*

Ex. 9. *Considere a permutação $\sigma = (1, 5, 3, 2, 4, 6)$ e considere o conjunto $A = \{2, 3, 4\}$. A permutação σ restrita ao conjunto A é a permutação $\sigma^A = (3, 2, 1)$*

Desta maneira podemos definir o conceito de subpermutação:

Definição 20. *Dadas a permutações τ em $[k]$ e σ em $[n]$, com $k < n$, dizemos que τ é uma subpermutação de σ se existir um subconjunto $A \subset [n]$, com $|A| = k$, tal que a permutação restrita σ^A é igual a τ.*

Ex. 10. *Considere a permutação $\sigma = (1, 3, 2, 5, 4)$. A permutação $\tau = (1, 2, 3)$ é subpermutação de σ, pois $\sigma^A = \tau = (1, 2, 3)$ para $A = \{1, 3, 5\}$.*

Dizemos que um grafo $G(V, E)$ é ordenado se V é munido de uma ordem total sobre seus elementos. Seja $|V| = n$. Como todo ordem total em V possui uma relação biunívoca com $[n]$ então, por simplicidade, iremos assumir $V = [n]$.

Podemos associar permutações com grafos ordenados. Dada uma permutação σ em $[n]$, seja G_σ o grafo ordenado tal que, para todo $i < j \in [n]$, temos que $ij \in E(G_\sigma)$ se só se $\sigma(i) < \sigma(j)$.

Ex. 11. *Para a permutação $\sigma = (3, 2, 5, 4, 1)$ temos o seguinte grafo:*

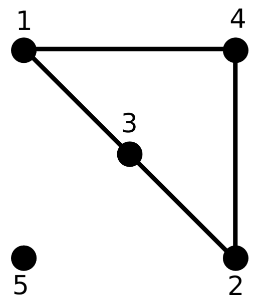

Figura 2.1: Grafo G_σ

Ex. 12. *Para a permutação $\eta = (7, 6, 1, 2, 5, 8, 9, 4, 3)$ temos o seguinte grafo:*

A permutação η é 3-simétrica, ou seja, para qualquer permutação τ em $[3]$ temos que $t(\tau, \eta) = \frac{1}{3!}$.

Observe que toda permutação pode ser relacionada a um grafo ordenado, mas nem todo grafo ordenado representa uma permutação, por causa das transitividades. Nosso lema abaixo prova uma caracterização de quais grafos ordenados representam permutações:

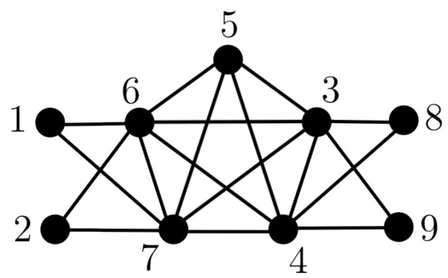

Figura 2.2: Grafo G_η

Lema 21. *Seja G um grafo ordenado em $[n]$. Então, para todo $i < j < k \in [n]$, temos que*

$$\begin{cases} ij, jk \in E(G) \Longrightarrow ik \in E(G) \\ ij, jk \notin E(G) \Longrightarrow ik \notin E(G) \end{cases}$$

se e só se existe uma permutação σ tal que $G = G_\sigma$.

Demonstração. Seja σ uma permutação em $[n]$. Para todo $i < j < k \in [n]$, se $ij, jk \in E(G_\sigma)$, temos que $\sigma(i) < \sigma(j)$ e $\sigma(j) < \sigma(k)$. Portanto $\sigma(i) < \sigma(k)$ e $ik \in E(G_\sigma)$. Da mesma forma temos que se $ij, jk \notin E(G_\sigma)$ temos que $\sigma(i) > \sigma(j)$ e $\sigma(j) > \sigma(k)$. Portanto $\sigma(i) > \sigma(k)$ e consequentemente $ik \notin E(G_\sigma)$.

Seja agora G um grafo ordenado em $[n]$ que satisfaz as condições do lema. Seja $\preceq_{\mathcal{P}}$ definida da seguinte forma: $i \preceq_{\mathcal{P}} j$ se e somente se $i < j \Leftrightarrow ij \in E(G)$. Afirmamos que $\preceq_{\mathcal{P}}$ é uma ordem total. É fácil ver que $\preceq_{\mathcal{P}}$ é reflexiva e antissimétrica. Para ver que $\preceq_{\mathcal{P}}$ é transitiva, note que, se $a \preceq_{\mathcal{P}} b$ e $b \preceq_{\mathcal{P}} c$, então

$$a < b \Leftrightarrow ab \in E(G),$$

$$b < c \Leftrightarrow bc \in E(G),$$

o que implica que

$$a < b < c \Leftrightarrow ab, bc \in E(G)$$

e, das restrições sobre G,

$$a < c \Leftrightarrow ac \in E(G)$$

e consequentemente $a \preceq_{\mathcal{P}} c$.

Note ainda que, para todo $i, j \in [n]$, $i \preceq_{\mathcal{P}} j$ ou $j \preceq_{\mathcal{P}} i$. Portanto, como $\preceq_{\mathcal{P}}$ é uma ordem total, podemos determinar qual é a ordem relativa de cada vértice do grafo em relação aos outros, ou seja, podemos determinar, para cada vértice, a posição dele em relação aos outros através dessa ordem total. Assim, geramos a nossa permutação σ colocando, para todo $i \in [n]$, $\sigma(i)$ como sendo a posição do vértice i na ordem total $\preceq_{\mathcal{P}}$. □

2.2 Sequências Convergentes

Dadas as permutações τ em $[k]$ e σ em $[n]$, com $k \leq n$, seja $\Lambda_\sigma(\tau) = \#\{A \subset [n]; \tau = \sigma^A\}$ o número de subpermutações τ em σ. A densidade de subpermutações τ em σ é definida como sendo

$$t(\tau, \sigma) = \frac{\Lambda_\sigma(\tau)}{\binom{n}{k}}.$$

De maneira análoga ao que se foi feito para grafos, temos a seguinte definição:

Definição 22. *Dizemos que uma sequência de permutações (σ_n) é convergente se, para toda permutação τ, a sequência de valores reais $t(\tau, \sigma_n)$ converge.*

Ex. 13. *A sequência de permutações (μ_n) tal que $\mu_n = (1, \ldots, n)$ converge, pois para toda permutação τ, temos que $t(\tau, \mu_n) = 1$ se τ é uma permutação identidade $\tau = (1, \ldots, k)$, e $t(\tau, \mu_n) = 0$ caso contrário.*

Ex. 14. *A sequência de permutações (λ_n) tal que λ_n é escolhida uniformemente entre as $n!$ permutações possíveis para $[n]$ é convergente, pois, para qualquer permutação τ em $[k]$, a sequência de valores reais $t(\tau, \lambda_n)$ converge para a probabilidade $\frac{1}{k!}$*

C. Hoppen et al. [15] provaram que toda sequência convergente de permutações possui um objeto limite definido abaixo.

Definição 23. *Uma permutação limite é um par de distribuição conjunta $Z = (X, Y)$ de variáveis aleatórias uniformes X e Y em $[0, 1]$ (não necessariamente independentes). Toda permutação limite $Z = (X, Y)$ pode ser identificada como a função de distribuição condicional de Y dado X, dada pela função mensurável $Z : [0, 1]^2 \to [0, 1]$ tal que $Z(x, y) = \mathbb{P}(Y \leq y \mid X = x)$. Não é difícil ver que:*

1. *Para todo $x_0 \in [0, 1]$, a função $Z(x_0, y)$ é uma função de distribuição acumulada (fda);*

2. *Para todo* $y_0 \in [0,1]$, *temos que*

$$\int_0^1 Z(x, y_0)dx = y_0.$$

Definição 24. *Dada uma permutação* σ *em* $[n]$, *a permutação limite* Z_σ *é obtida da função* f_σ *de densidade de probabilidade conjunta tal que, para cada* $(x, y) \in [0,1]^2$,

$$f_\sigma(x, y) \;=\; n \;\cdot\; \mathbb{1}\Big[\sigma(\lceil n \;\cdot\; x\rceil) = \lceil n \;\cdot\; y\rceil\Big]$$

De maneira similar ao que foi feito em grafos, podemos gerar permutações aleatórias a partir de uma permutações limite Z.

Definição 25. *Dada uma permutação limite* $Z = (X, Y)$, *a permutação* Z*-aleatória* $\sigma(n, Z)$ *em* $[n]$ *é gerada da seguinte forma:*

1. *Geramos* n *pares* $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ *de acordo com a distribuição de* $Z = (X, Y)$

2. *Ordenamos os* n *pares em ordem crescente em relação aos* x_i*'s:* $(x'_1, y'_1), \ldots, (x'_n, y'_n)$. *A permutação* $\sigma(n, Z)$ *é obtida da ordem relativa de* $(y'_1, \ldots, y'_n)$.

Ex. 15. *Suponha que os pares gerados sejam* $(0.88, 0.15)$, $(0.16, 0.44)$, $(0.35, 0.17)$ *e* $(0.56, 0.43)$. *Ordenando na primeira coordenada, obtemos* $(0.16, 0.44)$, $(0.35, 0.17)$, $(0.56, 0.43)$ *e* $(0.88, 0.15)$. *Nossa permutação* $\sigma(4, Z)$ *seria* $(4, 2, 3, 1)$, *que é a ordem relativa de* $(0.44, 0.17, 0.43, 0.15)$.

O modelo de permutação Z-aleatória estende o modelo clássico de permutação aleatória. Para isso, basta tomar $Z = (X, Y)$ tal que X e Y são independentes.

C. Hoppen et al. [16] provaram o seguinte teorema:

Teorema 26 ([16]). *Para toda sequência convergente de permutações* (σ_n), *existe uma permutação limite* Z, *tal que, para toda permutação* τ *em* $[k]$, *temos*

$$\lim_{n\to\infty} t(\tau, \sigma_n) = t(\tau, Z) := \mathbb{P}(\sigma(k, Z) = \tau)$$

Dada uma permutação limite Z, *a sequência de permutações* $\sigma(n, Z)$, $n \to \infty$, *é convergente com probabilidade 1 e seu limite é* Z.

Provou-se ainda uma relação entre as densidades de σ e as densidades de Z_σ.

Lema 27 ([16]). *Seja σ uma permutação em $[n]$. Então para toda permutação τ em $[k]$, com $k < n$, temos que:*

$$|t(\tau,\sigma) - t(\tau, Z_\sigma)| \leq \frac{1}{n}\binom{k}{2}$$

Abaixo define-se a distância retangular entre permutações, que é bastante útil para medir a convergência das sequências.

Definição 28. *Seja $I[n]$ o conjunto de todos os intervalos de $[n]$. Dadas duas permutações σ_1 e σ_2 em $[n]$, $n > 1$, definimos*

$$d_\square(\sigma_1,\sigma_2) = \frac{1}{n}\max_{S,T\in I[n]} ||\sigma_1^S \cap T| - |\sigma_2^S \cap T||,$$

onde $\sigma^S = \{\sigma(x)\forall x \in S\}$ é a

Essa distância foi baseada em uma das propriedades utilizadas por J. Copper [9] para definir a noção de quase-aleatoriedade.

Desta forma, introduzimos uma noção eficiente de distância, pois no caso das permutações aleatórias não é difícil ver que essa distância tende para zero. C. Hoppen et al. [15] provaram uma relação entre distância retangular para permutações e densidade de subpermutações através do seguinte teorema:

Teorema 29 ([15]). *Seja τ uma permutação em $[k]$ e sejam σ_1 e σ_2 permutações em $[n]$, com $k \leq n$. Então*

$$|t(\tau,\sigma_1) - t(\tau,\sigma_2)| \leq 2k^2 \cdot d_\square(\sigma_1,\sigma_2)$$

Podemos estender o conceito de distância retangular para objetos limite da seguinte maneira.

Definição 30. *Dadas permutações limite $Z_1 = (X_1, Y_1)$ e $Z_2 = (X_2, Y_2)$, definimos a distância retangular entre Z_1 e Z_2 como*

$$\begin{aligned} d_\square(Z_1,Z_2) &= \sup_{\substack{x_1<x_2\in[0,1]\\ y_1<y_2\in[0,1]}} \Big|\mathbb{P}(x_1 \leq X_1 \leq x_2, y_1 \leq Y_1 \leq y_2) \\ &\qquad - \mathbb{P}(x_1 \leq X_2 \leq x_2, y_1 \leq Y_2 \leq y_2)\Big| \end{aligned}$$

ou

$$d_\square(Z_1,Z_2) = \sup_{\substack{x_1<x_2\in[0,1]\\ y_1<y_2\in[0,1]}} \left| \int_{x_1}^{x_2} (Z_1(x,y_2) - Z_1(x,y_1))dx - \int_{x_1}^{x_2} (Z_2(x,y_2) - Z_2(x,y_1))dx \right|.$$

O teorema abaixo mostra uma equivalência entre sequências convergentes de permutações e sequências Cauchy na distância retangular.

Teorema 31 ([15]). *Uma sequência (σ_n) é convergente se, e somente se, (σ_n) é uma sequência de Cauchy na distância retangular*

Seja $\mathcal{Q}$ o conjunto de todas as permutações e seja $\mathcal{Z}$ o conjunto de todas as permutações limite. O Teorema 26 mostra que todo elemento do conjunto $\mathcal{Z}$ é um ponto de acumulação do conjunto $\mathcal{Q}$. O Teorema 31 prova que o espaço métrico $(\mathcal{Q} \cup \mathcal{Z}, d_\square)$ é completo.

2.3 Testabilidade Fraca

O trabalho apresentado nesta seção foi realizado por C. Hoppen et al. [16].

Definição 32. *Uma propriedade $\mathcal{P}$ é hereditária se, dada uma permutação σ que possua a propriedade $\mathcal{P}$, então toda subpermutação de σ também possui a propriedade $\mathcal{P}$.*

Definição 33. *Uma propriedade $\mathcal{P}$ de permutações é fracamente testável por subpermutações, ou apenas fracamente testável, se, para todo $\varepsilon > 0$, existe um inteiro positivo k tal que, se σ é um permutação em $[n]$, com $n > k$, então as duas afirmações abaixo são garantidas para uma subpermutação aleatória $sub(k, \sigma)$ de σ de tamanho k:*

1. *$sub(k, \sigma)$ satisfaz $\mathcal{P}$ com probabilidade pelo menos $1 - \varepsilon$ quando σ satisfaz $\mathcal{P}$;*
2. *$sub(k, \sigma)$ não satisfaz $\mathcal{P}$ com probabilidade pelo menos $1 - \varepsilon$ quando $d_\square(\sigma, \mathcal{P}) \geq \varepsilon$ onde:*

$$d_\square(\sigma, \mathcal{P}) \ := \ \min\{d_\square(\sigma, \pi) : |\pi| = n \ e \ \pi \ satisfaz \ \mathcal{P}\}.$$

Seja f um parâmetro de permutações, ou seja, uma função de valores reais definida sobre permutações e invariante por isomorfismos. Assumimos que o parâmetro é normalizado, ou seja, $0 \leq f \leq 1$. Um parâmetro de permutação f é testável se, para todo $\varepsilon > 0$, existe um $k > 0$ tal que, se σ é um permutação de tamanho $n > k$, então

$$\mathbb{P}\Big(\Big|f(\sigma) - f(sub(k, \sigma))\Big| > \varepsilon\Big) \leq \varepsilon.$$

Lema 34 ([16]). *Um parâmetro f de permutações é testável se e somente se, para toda sequência convergente de permutações (σ_n), a sequência $f(\sigma_n)$ converge.*

Com isso, o resultado principal de Hoppen et al. é o seguinte:

Teorema 35 ([16]). *Toda propriedade hereditária é fracamente testável.*

2.4 Testabilidade Forte

Nesta seção, nós resumimos o nosso resultado principal publicado no *Eurocomb 2011* [5], que é o resultado análogo de Alon e Shapira sobre a testabilidade de propriedades hereditárias para permutações. Na seção anterior, mostramos o resultado principal de [16], a saber, toda propriedade hereditária é fracamente testável através de subpermutações. O termo *fracamente* testável vem do fato de utilizarmos a noção de distância retangular para definir testabilidade, ao invés de uma distância mais intuitiva entre permutações como a distância d_1 de edição (*edit distance* ou *Kendall tau distance*).

Definição 36. *Sejam σ e π duas permutações em $[n]$. A distância de edição, definida como*

$$d_1(\sigma,\pi) = \frac{1}{\binom{n}{2}}\left|\left\{\{i,j\} \in \binom{[n]}{2} : (\sigma(i) > \sigma(j)) \Leftrightarrow (\pi(i) < \pi(j))\right\}\right|,$$

mede a densidade de pares de σ e π que tem ordem relativa diferente.

Essa distância está relacionada também com o algoritmo bubble sort: mede a densidade de inversões de pares de elementos adjacentes para transformar a permutação σ na permutação π.

Uma pergunta natural é a seguinte: "Se utilizarmos a distância de edição para definir testabilidade, o Teorema 35 continuária válido?".

É fácil ver que, dadas permutações σ e π em $[n]$, temos que, para todo $\varepsilon > 0$, existe um $\delta > 0$ tal que se $d_1(\sigma,\pi) < \delta$ então $d_\square(\sigma,\pi) < \varepsilon$.

Desta forma, podemos definir o conceito de testabilidade forte através de subpermutações.

Definição 37. *Uma propriedade $\mathcal{P}$ de permutações é fortemente testável por subpermutações, ou apenas testável, se, para todo $\varepsilon > 0$, existe um inteiro positivo k tal que, se σ é um permutação em $[n]$, com $n > k$, então as duas afirmações abaixo são garantidas para uma subpermutação aleatória $sub(k,\sigma)$ de σ de tamanho k:*

1. *$sub(k,\sigma)$ satisfaz $\mathcal{P}$ com probabilidade pelo menos $1-\varepsilon$ quando σ satisfaz $\mathcal{P}$;*

2. *$sub(k, \sigma)$ não satisfaz $\mathcal{P}$ com probabilidade pelo menos $1 - \varepsilon$ quando $d_1(\sigma, \mathcal{P}) \geq \varepsilon$ onde:*

$$d_1(\sigma, \mathcal{P}) \ := \ \min\{d_1(\sigma, \pi) : |\pi| = n \text{ e } \pi \text{ satisfaz } \mathcal{P}\}.$$

O teorema abaixo é o resultado principal.

Teorema 38. *Seja $\mathcal{P}$ uma propriedade hereditária de permutações. Dada uma permutação σ, então, $d_1(\sigma, \mathcal{P})$ é um parâmetro testável. Consequentemente, toda propriedade hereditária de permutações é fortemente testável.*

A prova do Teorema 38 esta relacionada ao Teorema 71, visto que, na Seção 2.1, mostramos que uma permutação pode ser representada através de um grafo ordenado. Omitimos a prova, pois consideramos mais relevante nos concentrarmos na prova de que toda propriedade hereditária de grafos ordenados é testável.

2.5 Quase-Aleatoriedade

2.5.1 Grafos

Uma sequência de estruturas é dita quase-aleatória quando, mesmo ela sendo gerada através de um algoritmo determinístico, ela satisfaz as mesmas propriedades que uma estrutra aleatória satisfaz. Este conceito foi desenvolvido inicialmente para auxiliar métodos de demonstração baseados em estruturas aleatórias, pois uma vez que uma demonstração baseada em estruturas aleatórias nos fornece uma prova da existência de uma determinada estrutura com uma determinada propriedade, o conceito de quase-aleatoriedade nos fornece uma maneira, relativamente fácil, de mostrar uma construção explícita de uma estrutura com a mesma propriedade.

F. R. K. Chung et al. [8] introduziram o conceito de quase-aleatoriedade para grafos. Para isso, eles determinaram um conjunto de propriedades que são satisfeitas por grafos aleatórios e definiram grafos quase-aleatórios como sendo aqueles grafos que satisfazem pelo menos uma dessas propriedades.

O principal e surpreendente resultado de F. R. K. Chung et al. [8], foi provar a equivalência das propriedades quase-aleatórias de grafos. Ou seja, se um grafo tem uma dessas propriedades, então também possui as demais. Abaixo, listamos algumas delas:

1. O número de arestas de G é $\frac{(1+o(1))n^2}{4}$ e o número de C_4 em G é $(1 + o(1))(\frac{n}{2})^2$

2. Para todo grafo H com $s \geq 4$ vértices, o número de cópias rotuladas de H em G é $(1 + o(1))n^s 2^{-\binom{s}{2}}$

3. Para todo subconjunto $S \subset V(G)$ o número de arestas de G com as duas extremidades em S é $\frac{1}{4}|S|^2 + o(n^2)$

4. $\sum_{u,v \in V(G)} ||N(u) \cap N(v)| - \frac{n}{4}| = o(n^3)$. Em que $N(v)$ representa a vizinhança de v em G.

Com o conceito de convergência de sequência, podemos definir quase-aleatoriedade de outra forma.

Definição 39. *Uma sequência (G_n) é quase-aleatória se, para qualquer grafo F, temos*

$$\lim_{n \to \infty} t(F, G_n) = t(F, W).$$

Onde W é o graphon *tal que $W(x, y) = \frac{1}{2}$ para todo $(x, y) \in [0, 1]^2$.*

Observe que esta definição está de acordo com o trabalho feito por Chung et al. Em outras palavras, podemos utilizar essa definição para demonstrar as propriedades citadas por Chung et al.

2.5.2 Permutações

Em 2004, J. N. Cooper [9] apresentou conceito de quase-aleatoriedade para permutações e também provou-se que as propriedades quase-aleatórias são equivalentes entre si. Para isso, Cooper enfraqueceu a noção de Chung-Graham de conjunto quase-aleatório de inteiros apresentado por F. R. K. Chung e R. L. Graham [7] usando conceitos de discrepância de sequências de números reais, que ele irá chamar de conjuntos ε-balanceados.

Formalmente, dados dois subconjuntos $S, T \subset \mathbb{Z}_n$ definimos discrepância de S em T como sendo

$$D_T(S) = ||S \cap T| - \frac{|S||T|}{n}|.$$

A discrepância máxima de S é definida por

$$D(S) = \max_{T \subset I[n]} D_T(S).$$

Dizemos então que, um subconjunto $S \subset \mathbb{Z}_n$ é ε-balanceado se $D(S) < \varepsilon n$. É possível comparar e observar uma relação entre quase-aleatoriedade de Chung e Graham para conjuntos de $\mathbb{Z}_n$ e conjuntos ε-balanceados de Cooper.

Formalmente, definimos discrepância de uma permutações σ em $[n]$ como sendo

$$D(\sigma) = \max_{S \in I[n]} D(\sigma(S))$$

Cooper define uma sequência de permutações (σ_n) como sendo quase-aleatória se

$$\lim_{n\to\infty} \frac{D(\sigma_n)}{|\sigma_n|} = 0.$$

Com o auxílio de sequências convergentes de permutações, Sampaio et al. [22] provaram outra caracterização para permutações quase-aleatórias.

Teorema 40. *Uma sequência (σ_n) de permutações é quase-aleatória se e só se, para qualquer permutação τ, temos*

$$\lim_{n\to\infty} t(\tau, \sigma_n) = t(\tau, Z_{ind}) = \frac{1}{|\tau|!},$$

onde $Z_{ind} = (X, Y)$ é a permutação limite tal que X e Y são independentes.

Capítulo 3

Ordens Parciais

Neste capítulo, iremos apresentar a teoria de sequências convergentes para ordens parcias. Na Seção 3.1, definiremos os conceitos básicos utilizados neste capítulo. Na Seção 3.2, apresentaremos o conceito de sequência convergentes de ordens parciais e alguns resultados que já foram desenvolvidos.

3.1 Introdução

Definição 41. *Um conjunto parcialmente ordenado, ou poset, com n elementos é um par $P = ([n], \preceq_P)$ no qual $\preceq_P$ é uma relação binária reflexiva, antissimétrica e transitiva em $[n]$.*

Ex. 16. *O conjunto $M = ([n], \preceq_I)$, no qual dado $i, j \in [n]$ temos que $i \preceq_I j$ se, e somente se, $i = j$ é um poset.*

Ex. 17. *O conjunto $N = ([n], \leq)$ é um poset.*

Note que, no exemplo anterior, a relação binária é *total*, ou seja, para todo $i, j \in [n]$ temos que $i \leq j$ ou $j \leq i$.

Definição 42. *Sejam $P = ([n], \preceq_P)$ e $Q = ([k], \preceq_Q)$ dois posets, com $k < n$. Dizemos que Q é um subposet de P, se existir uma função injetiva $\phi : [k] \to [n]$ tal que, para todo $i, j \in [k]$, temos que, se, $i \preceq_Q j$, então, $\phi(i) \preceq_P \phi(j)$.*

Definição 43. *Sejam $P = ([n], \preceq_P)$ e $Q = ([k], \preceq_Q)$ dois posets, com $k < n$. Dizemos que Q é um subposet induzido de P se, existir uma função injetiva $\phi : [k] \to [n]$ tal que, para todo $i, j \in [k]$, temos que $i \preceq_Q j$ se, e somente se, $\phi(i) \preceq_P \phi(j)$*

Note que toda permutação σ em $[n]$ pode ser vista como uma ordem total $([n], \preceq_\sigma)$ tal que, para todo $i, j \in [n]$, $i \preceq_\sigma j$ se e só se $\sigma(i) \leq \sigma(j)$. Por outro lado, a interseção de um conjunto finito de permutações $\sigma_1, \sigma_2, \ldots, \sigma_k$ em $[n]$ gera um poset $P = ([n], \preceq_P)$, tal que, para todo $i, j \in [n]$, $i \preceq_P j$ se e somente se $\sigma_\ell(i) \leq \sigma_\ell(j)$ para todo $\ell \in [k]$.

Ex. 18. *Seja σ uma permutação em $[6]$ tal que $\sigma = (5, 3, 1, 2, 4, 6)$. Então temos a ordem total $3 \preceq_\sigma 4 \preceq_\sigma 2 \preceq_\sigma 5 \preceq_\sigma 1 \preceq_\sigma 6$.*

Ex. 19. *Sejam σ_1 e σ_2 permutações em $[6]$ tais que $\sigma_1 = (5, 3, 1, 2, 4, 6)$ e $\sigma_2 = (5, 2, 1, 3, 4, 6)$. Então temos o poset $\{(1, 6), (2, 1), (2, 5), (2, 6), (3, 1), (3, 2), (3, 4), (3, 5), (3, 6), (4, 1), (4, 5), (4, 6), (5, 1), (5, 6)\}$.*

Desta forma, temos:

Definição 44. *Seja $P = ([n], \preceq_P)$ um poset e sejam $\sigma_1, \ldots, \sigma_k$ permutações em $[n]$. Dizemos que as permutações $\sigma_1, \ldots, \sigma_k$ formam um k-realizador de P se P é gerado pela interseção das k permutações.*

Definição 45. *Seja $P = ([n], \preceq_P)$ um poset. A dimensão de um poset é o menor inteiro k tal que P possui um k-realizador.*

Como mostrado no Capítulo 2, podemos também representar um poset como um grafo ordenado da seguinte maneira: Dado um poset $P = ([n], \preceq_P)$ tal que, se $i \preceq_P j$ então $i < j$, seja $G_P = ([n], E)$ o grafo ordenado tal que, para todo $i < j \in [n]$, $ij \in E(G_P)$ se e somente se $i \preceq_P j$. É fácil ver que, dado qualquer poset P', podemos obter um poset P isomorfo a P' tal que, se $i \preceq_P j$ então $i < j$.

Ex. 20. *Seja $P = ([5], \preceq_P)$ um poset, onde $\{(1, 4), (1, 3), (2, 3), (2, 4)\}$. Logo o grafo ordenado abaixo representa P.*

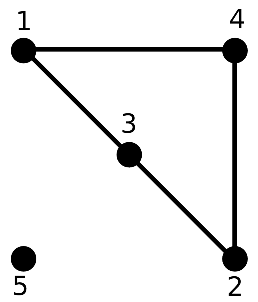

Figura 3.1: Grafo G_P

3.2 Sequências Convergentes

Seguindo a mesma ideia de grafos, S. Janson [17] definiu o critério para convergência de sequência de posets. Dados dois posets $P = ([k], \preceq_P)$ e $B = ([n], \preceq_B)$, com $k < n$, seja $\Gamma(P, B)$ o número mapeamentos injetivos de $\Phi : [k] \to [n]$ que preservam a ordem: Se $x \preceq_P y$ então $\Phi(x) \preceq_B \Phi(y)$. O número de subposets P em B é igual à $\Gamma(P, B)$ dividido pelo número de automorfismos de P. Observe que o número de mapeamentos injetivos de P em B é igual à $(n)_k$. Logo podemos definir o conceito de densidade fazendo

$$t(P, B) = \frac{\Gamma(P, B)}{(n)_k}.$$

Definição 46. *Dizemos que uma sequência (B_n) de posets é convergente se, para todo poset $P = ([k], \preceq_P)$, a sequência de valores reais $t(P, B_n)$ converge.*

R. C. Corrêa [10] provaram que toda sequência convergente de posets de dimensão k, para qualquer valor fixo de k, possui um objeto limite, o qual chamaram de *k-kernel*.

Definição 47. *Um k-*kernel *$Z = (X_1, \ldots, X_k)$ é uma k-upla de variáveis aleatórias uniformes, não necessariamente independentes, $X_1, \ldots, X_k$ com valores no intervalo $[0, 1]$.*

Um *k-kernel* $Z = (X_1, \ldots, X_k)$ pode ser visto como uma função de distribuição $Z : [0, 1]^k \to [0, 1]$ dada por $Z(a_1, \ldots, a_k) = \mathbb{P}(X_1 \leq a_1, \ldots, X_k \leq a_k)$, e vice-versa.

Ex. 21. *Um k-*kernel *$Z = (X_1, \ldots, X_k)$ onde as variáveis aleatórias $X_i's$ são independentes corresponde à função de distribuição $Z(x_1, \ldots, x_k) = \prod_{i=1}^{k} x_i$*

Similar a grafos, dado um *k-kernel* Z, podemos gerar um poset $P(n, Z)$.

Definição 48. *Dado um k-*kernel *$Z = (X_1, \ldots, X_k)$, o poset k-dimensional aleatório $P(n, Z)$ é gerado da seguinte forma:*

1. *$Y_1, \ldots, Y_n$ são gerados no k-cubo $[0, 1]^k$ de acordo com Z;*
2. *Dado $i, j \in [n]$ dizemos que $i \preceq_P j$ se e somente se, para todo $h \in [k]$, a h-ésima coordenada de Y_i é menor ou igual a h-ésima coordenada de Y_j.*

Não é difícil ver que a dimensão de $P(n, Z)$ é no máximo k, para isso basta observar que as k permutações formadas pela ordem relativa a cada coordenada dos Y_i's é um k-realizador de $P(n, Z)$. Podemos generalizar o conceito de densidade de subposets para densidade em um k-*kernel* Z. Dado um poset $P = ([k], \preceq_P)$, definimos

$$t(P, Z) = \mathbb{P}\Big(P(k, Z) = P\Big).$$

R. C. Corrêa et al. [10] mostraram que, no critério de convergência, os teoremas a seguir são válidos.

Teorema 49 ([10]). *Seja k um inteiro fixo. Para toda sequência convergente (B_n) de posets k dimensionais, existe um k-*kernel *$Z = (X_1, \ldots, X_k)$ tal que, para todo poset P, temos*

$$\lim_{n\to\infty} t(P, B_n) = t(P, Z)$$

Teorema 50 ([10]). *Seja $Z = (X_1, \ldots, X_k)$ um k-*kernel*. A sequência $(P(n, Z))$, para $n \to \infty$, converge para Z com probabilidade 1.*

Capítulo 4

Grafos Ordenados

Motivado pelos resultados obtidos anteriormente, decidimos desenvolver a noção de sequência convergente para grafos ordenados, pois podemos visualizar permutações, posets e outras estruturas como casos particulares de grafos ordenados. Uma vez que mostramos a existência de um objeto limite para grafos ordenados, ganhamos uma demonstração existencial de objeto limite para uma estrutura que pode ser representada como grafo ordenado.

Na Seção 4.1, apresentaremos algumas definições básicas sobre grafos ordenados que utilizaremos no restante deste trabalho. Na Seção 4.2, definiremos o critério para convergência de sequência de grafos ordenados. Na Seção 4.3, apresentaremos o nosso objeto limite para tais sequências. Na Seção 4.4, apresentaremos a noção de distância retangular para o nosso espaço de grafos ordenados e provaremos o Teorema de Cauchy para sequência de grafos ordenados e a unicidade do objeto limite. Na Seção 4.5, apresentaremos alguns resultados que são herdados diretamente de grafos e enunciamos uma conjectura de regularidade fraca em intervalos que, se for verdadeira, prova a exostência do objeto limite que esta descrito na Seção 4.6. Na Seção 4.7, iremos apresentar o conceito de testabilidade forte para grafos ordenados e iremos provar que toda propriedade hereditária de grafos ordenados é fortemente testável.

4.1 Grafos Ordenados

No Capítulo 2, apresentamos o conceito de grafos ordenados, que nada mais é do que um grafo simples rotulado munido de uma ordem total sobre seus rótulos. Sejam F e G grafos ordenados com conjuntos de vértices $[k]$ e $[n]$, respectivamente, com $k \leq n$.

Definição 51. *Dizemos que F é um subgrafo ordenado de G se existe uma função injetiva $\phi : [k] \to [n]$ tal que, para todo $i, j \in [k]$,*

1. $i < j \iff \phi(i) < \phi(j)$.
2. $ij \in E(F) \implies \phi(i)\phi(j) \in E(G)$

Definição 52. *Dizemos que F é um subgrafo ordenado induzido de G se existe uma função injetiva $\phi : [k] \to [n]$ tal que, para todo $i, j \in [k]$,*

1. $i < j \iff \phi(i) < \phi(j)$.
2. $ij \in E(F) \iff \phi(i)\phi(j) \in E(G)$

Ex. 22. *Considere os grafos a seguir:*

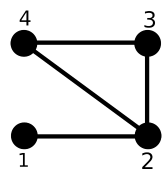

Figura 4.1: Um grafo ordenado F com 4 vértices

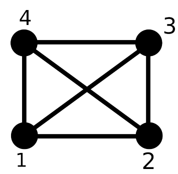

Figura 4.2: Um grafo ordenado G com 4 vértices

Nas figuras acima, temos que F é um subgrafo ordenado de G, mas não é um subgrafo ordenado induzido, enquanto que F e G são subgrafos ordenado induzido de G_1.

4.2 Sequências Convergentes

Dados grafos ordenado F e G com conjuntos de vértices $[k]$ e $[n]$, respectivamente, com $k \leq n$, seja $inj'(F, G)$ o número de funções injetivas $\phi : [k] \to [n]$ que satisfazem as condições da Definição 51. Observe que podemos ter no máximo $\binom{n}{k}$ funções injetivas ϕ distintas satisfazendo a Definição

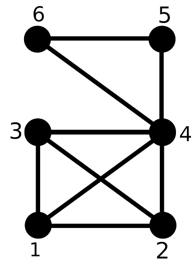

Figura 4.3: Um grafo ordenado G_1 com 6 vértices

51. Desta forma, definimos a densidade de subgrafos ordenados F em G como sendo

$$t'(F,G) = \frac{inj'(F,G)}{\binom{n}{k}}.$$

Observe que $t'(F,G)$ é a probabilidade $sub'(k,G) \supseteq F$, onde $sub'(k,G)$ é um subgrafo ordenado aleatório com k vértices de G.

De maneira análoga ao que fizemos até agora, estendemos o conceito de convergência para grafos ordenados.

Definição 53. *Dizemos que uma sequência (G_n) de grafos ordenados, com $|V(G_n)| \to \infty$, é convergente se, para todo grafo ordenado F, a sequência de densidades $t'(F,G_n)$ converge.*

Ex. 23. *A sequência (K_n) de grafos ordenados completos converge, pois, para todo grafo ordenado F, temos que $t'(F,K_n) = 1$ para todo n.*

Ex. 24. *A sequência $(\overline{K_n})$ de grafos ordenados vazios converge, pois, para todo grafo F, temos que*

$$t'(F,\overline{K_n}) = \begin{cases} 1 & se\ |E(F)| = 0 \\ 0 & caso\ contrário \end{cases}$$

Seja $ind'(F,G)$ o número de funções injetivas $\phi : [k] \to [n]$ que satisfazem as condições da Definição 52. Observe que podemos ter no máximo $\binom{n}{k}$ funções ϕ distintas. Desta forma, definimos a densidade de subgrafos ordenados induzidos F em G como sendo

$$t'_{ind}(F,G) = \frac{ind'(F,G)}{\binom{n}{k}}.$$

Apesar não parecerem equivalentes, o lema a seguir nos garante que a convergência em t' é equivalente à convergência em t'_{ind}.

Lema 54. *Seja (G_n) uma sequência de grafos ordenados. Temos que (G_n) é convergente se e somente se $t'_{ind}(F, G_n)$ converge para todo grafo ordenado F.*

Demonstração. Seja F um grafo ordenado e seja (G_n) uma sequência de grafos ordenados. Utilizando o princípio da inclusão e exclusão, temos que

$$ind'(F, G_n) = \sum_{F' \supset F} (-1)^{|E(F') \setminus E(F)|} inj'(F', G_n)$$

$$\Downarrow$$

$$t'_{ind}(F, G_n) = \sum_{F' \supset F} (-1)^{|E(F') \setminus E(F)|} t'(F', G_n)$$

onde F' é um grafo ordenado com o mesmo número de vértices de F e que contém F. Logo, se (G_n) é convergente, temos que $t'(F', G_n)$ converge qualquer que seja F' e, consequentemente, temos que $t'_{ind}(F, G_n)$ converge para todo F.

Além disso, é fácil ver que

$$inj'(F, G_n) = \sum_{F' \supset F} ind'(F', G_n)$$

$$\Downarrow$$

$$t'(F, G_n) = \sum_{F' \supset F} t'_{ind}(F', G_n).$$

Portanto se $t_{ind}(F', G_n)$ converge para qualquer F', então $t'(F, G_n)$ converge e, consequentemente, (G_n) é convergente. □

4.3 Objeto Limite

Nesta seção, iremos provar que o objeto limite para sequências convergentes de grafos ordenados também são *graphons*, como descritos no Capítulo 1.

Definição 55 (Densidade de subgrafos ordenados em um *graphon*)**.** *Seja W um* graphon *e seja F um grafo ordenado em $[k]$. A densidade de subgrafos ordenados F no* graphon *W é definido da seguinte forma:*

$$t'(F, W) = k! \int_{[0,1]^k_<} \prod_{ij \in E(F)} W(x_i, x_j) dx_1, \ldots, dx_k,$$

onde $[0, 1]^k_<$ é o conjunto de pontos $x = (x_1, \ldots, x_k) \in [0, 1]^k$ tais que $x_i < x_j$ para todo $i < j \in [k]$.

Podemos obter um *graphon* a partir de um grafo ordenado. Para isso basta notar que na Definição 3 o grafo G que utilizamos para construir um *graphon* W_G é um grafo ordenado. Ainda mais, cada grafo ordenado gera unicamente um *graphon*, enquanto que cada grafo gera uma família de *graphons*.

Note que o *graphon* herda da matriz de adjacência de G a simetria e a mensurabilidade. Nosso teorema abaixo é uma generalização do Teorema 5 de [19].

Teorema 56. *Seja (G_n) uma sequência convergente de grafos ordenados. Então existe uma função mensurável simétrica $W : [0,1]^2 \to [0,1]$ tal que, para todo grafo ordenado F temos que:*

$$\lim_{n\to\infty} t'(F, G_n) = t'(F, W)$$

O teorema acima será demonstrado mais a frente.

Dado um *graphon* $W : [0,1]^2 \to [0,1]$, podemos também definir um modelo $G'(n, W)$ de grafos ordenados aleatórios, que é similar ao $G(n, W)$ já apresentado no Capítulo 1.

Definição 57. *Dado um* graphon *W, definimos $G'(n, W)$ como sendo o grafo ordenado aleatório, com conjunto de vértices $[n]$, gerado da seguinte forma:*

1. *Geramos de maneira uniforme e independente n valores $X_1 \leq X_2 \leq \ldots \leq X_n$ em $[0,1]$;*

2. *Para cada par de vértices $i < j \in [n]$, colocamos a aresta ij com probabilidade $W(X_i, X_j)$.*

Note que o item *2* da definição acima pode ser substituido pelos seguintes passos:

2^* Para cada par de vértices $i, j \in [n]$ sorteamos uniformemente um valor $Y_{ij} \in [0,1]$;

3^* Ligamos os vértices $i, j \in [n]$ se $Y_{ij} \leq W(X_i, X_j)$.

Nosso teorema abaixo é uma generalização do Teorema 6 de [19].

Teorema 58. *A sequência $G'(n, W)$ converge para W com probabilidade* 1.

Demonstração. Seja $\mathcal{G}_n$ o conjunto de todos os grafos ordenados com vértices em $[n]$. Seja F um grafo ordenado em $[m]$. Seja $G' = G'(n, W)$. Então,

$$\begin{aligned}
\mathbb{E}\Big(t'(F, G'(n, W))\Big) &= \sum_{B \in \mathcal{G}_n} \mathbb{P}\Big(G' = B\Big) \cdot t'(F, B) \\
&= \sum_{B \in \mathcal{G}_n} \mathbb{P}\Big(G' = B\Big) \cdot \mathbb{P}\Big(sub'(m, B) \supseteq F\Big) \\
&= \sum_{B \in \mathcal{G}_n} \mathbb{P}\Big(G' = B\Big) \cdot \mathbb{P}\Big(sub'(m, G') \supseteq F | G' = B\Big) \\
&= \sum_{B \in \mathcal{G}_n} \mathbb{P}\Big(G' = B, sub'(m, G') \supseteq F\Big) \\
&= \mathbb{P}\Big(sub'(m, G') \supseteq F\Big) \\
&= \mathbb{P}\Big(G' \ \supseteq \ F\Big) \\
&= t'(F, W).
\end{aligned}$$

A primeira igualdade é a definição de esperança. A segunda e a última igualdade vem da definição de densidade de subgrafos ordenados.

Desejamos mostrar que $t'(F, G'(n, W))$ é uma função concentrada na média. Pela Definição 57, geramos independentemente os pontos $X_1 < \ldots < X_n$ em $[0, 1]$ e variáveis $Y_{i,j}$ em $[0, 1]$ para todo $i < j \in [n]$. O grafo $G'(n, W)$ é o grafo ordenado em $[n]$ tal que ij é uma aresta, para $i < j$, se e somente se $Y_{i,j} \leq W(X_i, X_j)$.

Em outras palavras, podemos expressar $G'(n, W)$ a partir de n variáveis aleatórias $Z_1, \ldots, Z_n$, onde $Z_i = (X_i, Y_{1,i}, Y_{2,i}, \ldots, Y_{i-1,i})$ satisfazendo $X_1 < \ldots < X_n$.

Seja G o grafo ordenado em $[n]$ obtido por $Z_1, \ldots, Z_n$. Seja G' outro grafo ordenado em $[n]$ obtido a partir de $Z_1, \ldots, Z_n$ alterando no máximo um Z_i por outro $Z'_i = (X'_i, Y'_{1,i}, Y'_{2,i}, \ldots, Y'_{i-1,i})$ satisfazendo $X_{i-1} < X'_i < X_{i+1}$. É fácil ver que

$$|t'(F, G) - t'(F, G')| \ \leq \ \frac{\binom{n-1}{m-1}}{\binom{n}{m}} \ = \ \frac{m}{n},$$

pois o número de subgrafos ordenados F contendo o vértice alterado é no máximo $\binom{n-1}{m-1}$, ou seja, $t'(F, G)$ é $\frac{m}{n}$-Lipschitz. Então, aplicando o método das diferenças limitadas (Teorema 14), temos que

$$\mathbb{P}\Big\{\Big|t'(F, G'(n, W)) - t'(F, W)\Big| > \varepsilon\Big\} \ \leq \ 2 \cdot \exp\{-2\varepsilon^2 n/m^2\}$$

para $\varepsilon > 0$ fixo. Seja A_n o evento ("$|t'(F, G'(n, W)) - t'(F, W)| > \varepsilon$"). A equação acima mostra que $\mathbb{P}(A_n) \leq 2\exp\{-2\varepsilon^2 n/m^2\}$. Isto garante que $\sum_{n=2}^{\infty} \mathbb{P}(A_n) < \infty$ e podemos então aplicar o Lema de Borel-Cantelli para provar que $t'(F, G'(n, W))$ converge para $t'(F, W)$ com probabilidade 1. Como o conjunto de todos os grafos ordenados é enumerável, então $t'(F, G'(n, W))$ converge para $t'(F, W)$ com probabilidade 1 para todo grafo ordenado F. $\square$

Seja $\mathcal{G}'$ o conjunto formado por todos os grafos ordenados. O Teorema 56 mostra que $\overline{\mathcal{G}'} \subseteq \mathcal{W}$, ou seja, que todo ponto de acumulação de $\mathcal{G}'$ é um *graphon*. Por outro lado o Teorema 58 mostra que $\overline{\mathcal{G}'} \subseteq \mathcal{W}$, ou seja, todo *graphon* é um ponto de acumulação de $\mathcal{G}'$. Com isso concluímos que o conjunto $\mathcal{G}' \cup \mathcal{W}$ é um conjunto fechado.

4.4 Distância

Uma vez tendo construído o conjunto fechado $\mathcal{G}' \cup \mathcal{W}$ estamos interessado agora em definir uma métrica para o nosso conjunto. Portanto iremos adotar a distância retangular utilizada para grafos, ver Definição 7. Uma definição alternativa para a distância retangular, dada por Lovász e Szegedy (ver equação 6 de [19]) é:

$$d_\square(U, W) = \sup_{\substack{0 \leq f, \\ g \leq 1}} \Big| \int_0^1 \int_0^1 f(x)g(y)(U(x,y) - W(x,y))dxdy \Big| \qquad (4.1)$$

O nosso lema auxiliar abaixo relaciona a distância retangular com a densidade de subgrafos ordenados e é fundamental para a demonstração do Teorema 56.

Lema 59. *Sejam U e W dois* graphons. *Então, para todo grafo ordenado F em $[n]$, temos que:*

$$|t'(F, U) - t'(F, W)| \leq n^2 \cdot |E(F)| \cdot d_\square(U, W)$$

Demonstração. Seja $E(F) = \{e_1, e_2, \ldots, e_{|E(F)|}\}$. Seja $e_t = i_t j_t$, onde $i_t < j_t$. Seja $\Delta = |t'(F, U) - t'(F, W)|$. Temos que:

$$\Delta = n! \Big| \int_{[0,1]^n_<} \Big(\prod_{ij \in E(F)} U(x_i, x_j) - \prod_{ij \in E(F)} W(x_i, x_j) \Big) dx_1 \ldots dx_n \Big| \qquad (4.2)$$

Seja $E_t(F) = \{e_1, e_2, \ldots, e_t\}$. Como feito em [19], podemos escrever:

$$\prod_{ij\in E(F)} U(x_i,x_j) - \prod_{ij\in E(F)} W(x_i,x_j) \;=\; \sum_{t=0}^{|E(F)|-1} X_t(x_1,\dots,x_n) \qquad (4.3)$$

onde

$$X_t(x_1,\dots,x_n) = \Big(\prod_{ij\in E_{t-1}(F)} W(x_i,x_j)\Big)\Big(\prod_{ij\in E(F)\setminus E_t} U(x_i,x_j)\Big)(U(x_{i_t},x_{j_t})-W(x_{i_t},x_{j_t}))$$

Assim, por (4.2) e (4.3), temos:

$$\begin{aligned}
\Delta &= n!\Big|\int_{[0,1]^n_<} \sum_{t=0}^{|E(F)|-1} X_t(x_1,\dots,x_n)dx_1\dots dx_n\Big| \\
&= n!\Big|\sum_{t=0}^{|E(F)|-1} \int_{[0,1]^n_<} X_t(x_1,\dots,x_n)dx_1\dots dx_n\Big| \\
&\leq n!\sum_{t=0}^{|E(F)|-1} \Big|\int_{[0,1]^n_<} X_t(x_1,\dots,x_n)dx_1\dots dx_n\Big| \\
&\leq n!\sum_{t=0}^{|E(F)|-1} \Big|\int_{[0,1]^{n-2}_<} \Big(\int_{x_{i_t-1}}^{x_{i_t+1}}\int_{x_{j_t-1}}^{x_{j_t+1}} X_t(x_1,\dots,x_n)dx_{i_t}dx_{j_t}\Big)dx_1\dots dx_n\Big| \\
&\leq n!\sum_{t=0}^{|E(F)|-1} \int_{[0,1]^{n-2}_<} \Big|\int_{x_{i_t-1}}^{x_{i_t+1}}\int_{x_{j_t-1}}^{x_{j_t+1}} X_t(x_1,\dots,x_n)dx_{i_t}dx_{j_t}\Big|dx_1\dots dx_n,
\end{aligned}$$

Portanto, de (4.1), temos que:

$$\begin{aligned}
\Delta &\leq n!\sum_{t=0}^{|E(F)|-1} \int_{[0,1]^{n-2}_<} d_\square(U,W)dx_1\dots dx_n \\
&\leq n!\; d_\square(U,W)\sum_{t=0}^{|E(F)|-1} \int_{[0,1]^{n-2}_<} dx_1\dots dx_n \\
&\leq n!\; d_\square(U,W)\sum_{t=0}^{|E(F)|-1} \frac{1}{(n-2)!},
\end{aligned}$$

Como desejado.

□

O lema a seguir mostra que uma sequência gerada através de $G'(n,W)$ converge para W, na distância retangular, com probabilidade 1.

Lema 60. *Seja W um* graphon. *Então*

$$\lim_{n\to\infty} d_\square(W, G'(n,W)) = 0$$

Demonstração. Seja $S = [x_1, x_2]$ e $T = [y_1, y_2]$ intervalos fixos em $[0,1]$. Gere $G := G'(n,W)$ da forma descrita na Definição 57. Seja $L(G)$ o número de arestas ij tais que $\lfloor x_1 n\rfloor \le i \le \lfloor x_2 n\rfloor$ e $\lfloor y_1 n\rfloor \le j \le \lfloor y_2 n\rfloor$. Temos:

$$\begin{aligned}
L(G) &= n^2 \int_{\frac{\lfloor x_1\rfloor}{n}}^{\frac{\lfloor x_2\rfloor}{n}} \int_{\frac{\lfloor y_1\rfloor}{n}}^{\frac{\lfloor y_2\rfloor}{n}} W_G(x,y)dxdy \\
&= n^2 \int_{x_1}^{x_2}\int_{y_1}^{y_2} W_G(x,y)dxdy \;\pm\; n(|S|+|T|-\frac{1}{n}) \\
&= n^2\Big(\int_{x_1}^{x_2}\int_{y_1}^{y_2} W_G(x,y)dxdy \;\pm\; \frac{2}{n}\Big) \qquad (4.4)
\end{aligned}$$

A segunda igualdade vem do fato que W_G contém apenas $0's$ e $1's$. Desta forma podemos aumentar o valor da integral em no máximo $n(|S|+|T|-\frac{1}{n})$.

Seja A_{ij} o evento tal que $\lfloor x_1 n\rfloor < i \le \lfloor x_2 n\rfloor$, $\lfloor y_1 n\rfloor < j \le \lfloor y_2 n\rfloor$ e ij é uma aresta. Seja I_{ij} a variável indicadora de A_{ij}. Logo

$$\mathbb{E}(L(G)) \le \sum_{i,j\in[n]} \mathbb{E}(I_{ij}) = n^2\mathbb{P}(A_{ij})$$

Seja E_{ij}^- o evento tal que $x_1^+ < X_i \le x_2^-$, $y_1^+ < X_j \le y_2^-$ e $Y_{i,j} \ge W(X_i,X_j)$, onde $x^\pm = x \pm n^{-\frac{1}{4}}$. Seja E_{ij}^+ o evento tal que $x_1^- < X_i \le x_2^+$, $y_1^- < X_j \le y_2^+$ e $Y_{i,j} \ge W(X_i,X_j)$. Observe que na definição do $G'(n,W)$ geramos n pontos uniformes $X_1 < ... < X_n$, esses pontos seguem a distribuição *Beta Chebyschev* e portanto, para $k \in n$, temos

$$\mathbb{P}\Big(|X_k - \frac{k}{n+1}| \ge n^{-\frac{1}{4}}\Big) \;\le\; n^{-\frac{1}{2}} \qquad (4.5)$$

Seja D_{ij} o evento tal que $X_i = \frac{i}{n+1} \pm n^{-\frac{1}{4}}$ e $X_j = \frac{j}{n+1} \pm n^{-\frac{1}{4}}$. Desta forma temos

$$\mathbb{P}(A_{ij}) \;=\; \mathbb{P}(A_{ij}|D_{ij})\mathbb{P}(D_{ij}) + \mathbb{P}(A_{ij}|\overline{D_{ij}})\mathbb{P}(\overline{D_{ij}}) \qquad (4.6)$$

Note que

$$\begin{aligned}\mathbb{P}(E_{ij}^-|D_{ij}) &= \frac{\mathbb{P}(E_{ij}^-, D_{ij})}{\mathbb{P}(D_{ij})}\\ &= \frac{\mathbb{P}(E_{ij}^-) - \mathbb{P}(E_{ij}^-, \overline{D_{ij}})}{\mathbb{P}(D_{ij})}\\ &\geq \frac{\mathbb{P}(E_{ij}^-) - \mathbb{P}(\overline{D_{ij}})}{\mathbb{P}(D_{ij})}\end{aligned}$$

E ainda

$$\begin{aligned}\mathbb{P}(E_{ij}^+|D_{ij}) &= \frac{\mathbb{P}(E_{ij}^+, D_{ij})}{\mathbb{P}(D_{ij})}\\ &\leq \frac{\mathbb{P}(E_{ij}^+)}{\mathbb{P}(D_{ij})}\end{aligned}$$

Para n suficientemente grande e de (4.5), temos que

$$\mathbb{P}(E_{ij}^{+/-}|D_{ij}) = \mathbb{P}(E_{ij}^{+/-}) \pm 2n^{-\frac{1}{2}}$$

Não é difícil ver que

$$\mathbb{P}(E_{ij}^-|D_{ij}) \leq \mathbb{P}(A_{ij}|D_{ij}) \leq \mathbb{P}(E_{ij}^+|D_{ij})$$

De maneira similar ao que fizemos no começo da demonstração, temos:

$$\begin{aligned}\mathbb{P}(E_{ij}^-) &= \int_{x_1^+}^{x_2^-}\int_{y_1^+}^{y_2^-} W(x,y)dxdy\\ &= \int_{x_1}^{x_2}\int_{y_1}^{y_2} W(x,y)dxdy \pm 4n^{-\frac{1}{4}}\end{aligned}$$

O mesmo vale para $\mathbb{P}(E_{ij}^+)$. Por outro lado, de (4.6) temos:

$$\mathbb{P}(A_{ij}) \leq \mathbb{P}(E_{ij}^+|D_{ij}) + n^{-\frac{1}{2}} = \mathbb{P}(E_{ij}^+) \pm 3n^{-\frac{1}{2}}$$

$$\mathbb{P}(A_{ij}) \geq \mathbb{P}(E_{ij}^-|D_{ij}) - n^{-\frac{1}{2}} = \mathbb{P}(E_{ij}^-) \pm 3n^{-\frac{1}{2}}$$

E assim temos

$$\mathbb{P}(A_{ij}) = \int_{x_1}^{x_2}\int_{y_1}^{y_2} W(x,y)dxdy \pm (3n^{-\frac{1}{2}} + 4n^{-\frac{1}{4}})$$

Portanto

$$\mathbb{E}(L(G)) = n^2\Big(\int_{x_1}^{x_2}\int_{y_1}^{y_2} W(x,y)dxdy \pm (3n^{-\frac{1}{2}} + 4n^{-\frac{1}{4}})\Big)$$

Note que $L(G)$ conta o número de arestas entre dois subconjuntos de vértices. Se alterarmos apenas um vértice em qualquer um dos dois subconjuntos de vértices então $L(G)$ aumenta em no máximo $|S| + |T|$. Portanto $L(G)$ é $2n$-Lipschitz. Usando o método das diferenças limitadas temos

$$\begin{aligned}
\mathbb{P}\Big(|L(G) - \mathbb{E}(L(G))| > \varepsilon n^2\Big) &\leq 2exp\Big(-\frac{2(\varepsilon n^2)^2}{4n^3}\Big)\\
&\leq 2exp\Big(-\frac{2\varepsilon^2 n^4}{4n^3}\Big)\\
&\leq 2exp\Big(-\frac{\varepsilon^2 n}{2n}\Big) \qquad (4.7)
\end{aligned}$$

Observe que para n suficientemente grande temos que $2n^{-1} + 3n^{-\frac{1}{2}} + 4n^{-\frac{1}{4}} \leq 5n^{-\frac{1}{4}}$. Desta forma

$$\mathbb{P}\Big(\Big|\int_{x_1}^{x_2}\int_{y_1}^{y_2} W_G(x,y) - W(x,y)dxdy\Big| > \varepsilon + 5n^{-\frac{1}{4}}\Big) \leq 2exp\Big(-\frac{\varepsilon^2 n}{2n}\Big)$$

Seja $V_n = \{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}$. Temos:

$$\mathbb{P}\Big(\exists x_1 < x_2, y_1 < y_2 \in V_n, \Big|\int_{x_1}^{x_2}\int_{y_1}^{y_2} W_G(x,y) - W(x,y)dxdy\Big| > \varepsilon + 5n^{-\frac{1}{4}}\Big)$$

$$\leq 2\binom{n+1}{2}^2 exp\Big(-\frac{\varepsilon^2 n}{2n}\Big)$$

Fixe $\varepsilon > 0$. Considere o evento

$$B_n = \Big(\exists x_1 < x_2, y_1 < y_2 \in [0,1], \Big|\int_{x_1}^{x_2}\int_{y_1}^{y_2} W_G(x,y) - W(x,y)dxdy\Big| > \varepsilon + 6n^{-\frac{1}{4}}\Big)$$

Observe que dado $x_1 < x_2, y_1 < y_2 \in [0,1]$ tomando $x_1', x_2' \in V_n$ tal que $x_1' \geq x_1$ e $x_1' - x_1 < \frac{1}{n}$ e $x_2' \leq x_2$ e $x_2' - x_2 < \frac{1}{n}$ e, de maneira análoga, tomando $y_1', y_2' \in V_n$ temos

$$\Big|\int_{x_1}^{x_2}\int_{y_1}^{y_2} W_G(x,y) - W(x,y)dxdy\Big| \leq \Big|\int_{x_1'}^{x_2'}\int_{y_1'}^{y_2'} W_G(x,y) - W(x,y)dxdy\Big| + \frac{4}{n}$$

Para n suficientemente grande temos que $\frac{4}{n} \leq n^{-\frac{1}{4}}$. Portanto

$$\mathbb{P}\Big(B_n\Big) \leq 2\binom{n+1}{2}^2 exp\Big(-\frac{\varepsilon^2 n}{2n}\Big) \leq exp\Big(-\frac{\varepsilon^2 n}{2n} + 4\ln(n)\Big)$$

Em outras palavras

$$\mathbb{P}\Big(d_\square(W, G'(n,W)) > \varepsilon + 6n^{-\frac{1}{4}}\Big) \leq exp\Big(-\frac{\varepsilon^2 n}{2n} + 4\ln(n)\Big)$$

Por Borel-Cantelli, temos que a probabilidade de que B_n,com $n = 1, \ldots$, ocorra para infinitos valores de n é nula. Com isso provamos o lema. $\square$

Dados W_1 e W_2 *graphons*, o lema a seguir mostra, que para m suficientemente grande, se W_1 e W_2 são próximos em relação às densidades então eles são próximos na norma retangular.

Lema 61. *Sejam* W_1 *e* W_2 graphons *tais que* $|t'(F,W_1) - t'(F,W_2)| \leq \frac{1}{2^{\binom{m+1}{2}}}$, *para todo* $m > 0$ *e todo grafo ordenado* F *em* $[m]$*. Então*

$$d_\square(W_1, W_2) = 0$$

Demonstração. Fixe $\varepsilon > 0$. Tome $m > \log_2 \frac{3}{\varepsilon}$. Como $\mathbb{P}(G'(m, W_1) = F) = t'(F, W_1)$, para todo grafo ordenado F temos que

$$\sum_{F([m],E)} \Big|\mathbb{P}\big(G'(m,W_1) = F\big) - \mathbb{P}\big(G'(m,W_2) = F\big)\Big| \leq 2^{\binom{m}{2}} \frac{1}{2^{\binom{m+1}{2}}} \leq \frac{\varepsilon}{3}$$

De maneira idêntica a feita por Borgs et al. na prova do Teorema 3.7(b) em [6], podemos acoplar $G'(m, W_1)$ e $G'(m, W_2)$ de tal forma que $G'(m, W_1) \neq G'(m, W_2)$ com probabilidade menor que $\frac{\varepsilon}{3}$.

Para isso, podemos gerar $G'(m, W_1)$ e $G'(m, W_2)$ de forma que, para todo grafo ordenado F em $[m]$, $\mathbb{P}(G'(m, W_1) = F \;\&\; G'(m, W_2) = F)$ é igual ao menor valor entre $\mathbb{P}(G'(m, W_1) = F)$ e $\mathbb{P}(G'(m, W_2) = F)$.

Desta forma, pelo Lema 60,

$$\begin{aligned} d_\square(W_1, W_2) &\leq \mathbb{E}(d_\square(W_1, G'(m, W_1))) + \mathbb{E}(d_\square(G'(m, W_1), G'(m, W_2))) \\ &\quad + \mathbb{E}(d_\square(G'(m, W_2), W_2)) \\ &\leq \varepsilon, \end{aligned}$$

como queríamos mostrar. $\square$

A prova da unicidade segue diretamente dos Lemas 60 e 61.

Corolário 62. *$d_\square(W_1, W_2) = 0$ se e só se $t'(F, W_1) = t'(F, W_2)$ para todo grafo ordenado F. Além disso, seja (G_n) uma sequência convergente de grafos ordenados com $G_n \to W_1$. Se $G_n \to W_2$ então $d_\square(W_1, W_2) = 0$.*

E finalmente a prova do Teorema de Cauchy para grafos ordenados segue diretamente dos Lemas 59 e 61.

Corolário 63. *A sequência (G_n) é convergente se e somente se é de Cauchy na norma retangular.*

O Corolário 63 mostra que o espaço métrico $(\mathcal{G} \cup \mathcal{W}, d_\square)$ é um espaço completo.

4.5 Resultados Herdados de Grafos

Iremos apresentar alguns resultados que herdamos diretamente de grafos. Isso acontece devido ao fato de que esses resultados são relacionados à matriz de adjacência do grafo, que é a mesma para um grafo ordenado. Utilizaremos a definição abaixo de distância retangular entre matrizes.

Definição 64. *Dadas duas matrizes $n \times n$ Q_1 e Q_2 com valores em $[0, 1]$, definimos a distância retangular $d_\square(Q_1, Q_2)$ como*

$$d_\square(Q_1, Q_2) = \max_{A,B \subseteq [n]} \frac{1}{n^2} \Big| \sum_{i \in A} \sum_{j \in B} \Big(Q_1(i, j) - Q_2(i, j) \Big) \Big|$$

Observe que, dados grafos G e G' com n vértices, a distância retangular entre G e G' é igual à distância retangular entre as matrizes de adjacências A_G e $A_{G'}$ de G e G' respectivamente. Ou seja,

$$d_\square(G, G') = d_\square(A_G, A_{G'}).$$

O Lema 1 abaixo é a forma fraca do Lema de Szemerédi (ver [13] ou o Lema 4.2 de [19]) o qual mostra que todo grafo pode ser aproximado por um grafo ponderado.

Lema 65 (Forma fraca do Lema de Szemerédi [19]). *Para todo $\varepsilon > 0$, existe um inteiro $K > 0$ tal que, para todo grafo simples G com vértices em $[n]$, existe uma partição $\mathcal{P}$ de $[n]$ em $k \leq K$ classes $V_1, \ldots, V_k$, e uma matriz $k \times k$ simétrica Q com todos os valores em $[0, 1]$, tal que:*

$$||V_i| - |V_j|| \leq 1 \qquad (1 \leq i, j \leq k).$$

e

$$d_\square(A_G, K_{\mathcal{P},Q}) \leq \varepsilon,$$

onde A_G *é a matriz de adjacência de* G *e* $K_{\mathcal{P},Q}$ *é a matriz* $n \times n$ *tal que, para todo* $x, y \in [n]$, $K_{\mathcal{P},Q}(x, y) = Q_{ij}$ *onde* $x \in V_i$ *e* $y \in V_j$. *Dizemos que a partição* $\mathcal{P}$ *é* ε*-regular fraca.*

Abaixo conjecturamos que a partição obtida da regularidade fraca acima pode ser em intervalos.

Conjectura 1 (Regularidade fraca em intervalos)**.** *Para todo* $\varepsilon > 0$, *existe um inteiro* $K > 0$ *tal que, para todo grafo simples* G *com vértices em* $[n]$, *existe uma partição* $\mathcal{P}$ *de* $[n]$ *em* $k \leq K$ *intervalos* $V_1, \ldots, V_k$, *de* $[n]$ *e uma matriz* $k \times k$ *simétrica* Q *com todos os valores em* $[0, 1]$, *tal que:*

$$||V_i| - |V_j|| \leq 1 \qquad (1 \leq i, j \leq k).$$

e

$$d_\square(A_G, K_{\mathcal{P},Q}) \leq \varepsilon,$$

onde A_G *é a matriz de adjacência de* G *e* $K_{\mathcal{P},Q}$ *é a matriz* $n \times n$ *tal que, para todo* $x, y \in [n]$, $K_{\mathcal{P},Q}(x, y) = Q_{ij}$ *onde* $x \in V_i$ *e* $y \in V_j$. *Dizemos que a partição* $\mathcal{P}$ *é* ε*-regular fraca.*

No restante do texto iremos considerar que esta conjectura é verdadeira.

O lema a seguir de [19] mostra que qualquer sequência de grafos possui uma subsequência de grafos no qual as suas matrizes de adjacência são uma sequência de refinamento e os grafos são bem comportados em relação à forma fraca do Lema de Szeméredi.

Lema 66 ([19])**.** *Toda sequência de grafos ordenados* (G_n) *possui uma subsequência de grafos ordenados* (G'_m) *tal que existe uma sequência de inteiros* (k_m) *e uma sequência de matrizes* (Q_m) *com as seguintes propriedades:*

1. Q_m *é uma matriz* $k_m \times k_m$ *simétrica e todos os valores estão em* $[0, 1]$*;*

2. *Se* $i < j$ *então* $k_i | k_j$ *e a matriz* Q_i *é obtida da matriz* Q_j *pelo particionamento das suas linhas e colunas em* k_i *blocos consecutivos de tamanho* k_j/k_i *e substituindo cada bloco por um único valor que é a média do bloco inteiro;*

3. *Para todo* $j < m$. G'_m *possui uma partição* $(1/m)$*-regular fraca* $\mathcal{P}_{m,j}$ *com a matriz de densidade* $Q_{m,j}$ *tal que*

$$d_\square(Q_{m,j}, Q_j) < 1/j$$

e, para todo $1 \leq i < j \leq m$, $\mathcal{P}_{m,j}$ *é um refinamento de* $\mathcal{P}_{m,i}$.

O lema seguinte de [19] mostra que, para toda sequência de matrizes obtida do lema anterior, existe um *graphon* W para qual ela converge.

Lema 67 ([19]). *Seja* (k_m) *uma sequência de inteiros positivos e* (Q_m) *uma sequência de matrizes satisfazendo (i) e (ii) do lema anterior. Então existe um* graphon W *tal que:*

1. $W_{Q_m} \to W \quad (m \to \infty)$ *com probabilidade 1;*

2. *Para todo* m *e* $1 \leq i, j \leq k_m$ *temos*

$$Q_m(i,j) = k_m^2 \int_{(i-1)/k_m}^{i/k_m} \int_{(j-1)/k_m}^{j/k_m} W(x,y)dxdy$$

4.6 Existência do Objeto Limite

Antes de provarmos o Teorema 56 iremos necessitar de um lema auxiliar. O lema auxiliar abaixo mostra uma relação entre a densidade t' de grafos ordenados e a densidade de *graphons*.

Lema 68. *Sejam* F *e* G *grafos ordenados, com* $k < n$. *Então*

$$|t'(F,G) - t'(F,W_G)| \leq \frac{(k-1)k!n^{k-1}}{(n)_k}.$$

Consequentemente, se $n > 2k$, então

$$|t'(F,G) - t'(F,W_G)| \leq \frac{(k+1)!2^k}{n}.$$

Demonstração. Para todo $i, j \in [n]$, seja $A_{ij} = \{(x_1, x_2) \in [0,1]^2 : \frac{i-1}{n} < x_1 \leq \frac{i}{n} \text{ e } \frac{j-1}{n} < x_2 \leq \frac{j}{n}\}$. Seja $\Omega_{inj}(F,G)$ o conjunto de todas as funções injetivas $\phi : [k] \to [n]$ que satisfazem a Definição 51. Para simplificar usaremos apenas a notação Ω_{inj}. Dada uma função $\phi \in \Omega_{inj}$, seja $\Gamma_\phi = A_{\phi(1)\phi(2)} \times \ldots A_{\phi(i)\phi(j)} \times \ldots \times A_{\phi(k-1)\phi(k)}$, para todo $i < j$. Desta forma temos que

$$\begin{aligned} t'(F,W_G) &= k! \int_{[0,1]^k_<} \prod_{ij \in E(F)} W_G(x_i,x_j)dx_1 \ldots dx_k \\ &\geq k! \sum_{\phi \in \Omega_{inj}} \int_{\Gamma_\phi} \prod_{ij \in E(F)} W_G(x_i,x_j)dx_1 \ldots dx_k \end{aligned}$$

$$
\begin{aligned}
&\geq k! \sum_{\phi \,\in\, \Omega_{inj}} \int_{\Gamma_\phi} \prod_{ij \in E(F)} A_G(\phi(i), \phi(j)) dx_1 \ldots dx_k \\
&\geq k! \sum_{\phi \,\in\, \Omega_{inj}} \prod_{ij \in E(F)} A_G(\phi(i), \phi(j)) \int_{\Gamma_\phi} 1 dx_1 \ldots dx_k \\
&\geq k! \sum_{\phi \,\in\, \Omega_{inj}} \prod_{ij \in E(F)} A_G(\phi(i), \phi(j)) \frac{1}{n^k} \\
&\geq \frac{k!}{n^k} \sum_{\phi \,\in\, \Omega_{inj}} \prod_{ij \in E(F)} A_G(\phi(i), \phi(j)) \\
&\geq \frac{k!}{(n)_k} \frac{(n)_k}{n^k} \sum_{\phi \,\in\, \Omega_{inj}} \prod_{ij \in E(F)} A_G(\phi(i), \phi(j)) \\
&\geq \frac{(n)_k}{n^k} \binom{n}{k}^{-1} \sum_{\phi \,\in\, \Omega_{inj}} \prod_{ij \in E(F)} A_G(\phi(i), \phi(j)) \\
&\geq \frac{(n)_k}{n^k} t'(F, G) \\
&\geq \Big(\prod_{i=1}^{k-1} (1 - \frac{i}{n})\Big) t'(F, G) \\
&\geq \Big(1 - \binom{k}{2} \frac{1}{n}\Big) t'(F, G) \\
&\geq t'(F, G) - \binom{k}{2} \frac{1}{n} \\
&\geq t'(F, G) - \frac{(k-1)k! n^{k-1}}{(n)_k}
\end{aligned}
$$

Desta forma conseguimos o primeiro limitante

$$
t'(F, G) - t'(F, W_G) \quad \leq \quad \frac{(k-1)k! n^{k-1}}{(n)_k}
$$

Agora note que podemos reescrever a densidade da seguinte forma

$$
t'(F, G) = \binom{n}{k}^{-1} \sum_{X \in [n]^k_<} \prod_{ij \in E(F)} W_G\Big(\frac{X_i}{n}, \frac{X_j}{n}\Big),
$$

onde $[n]^k_<$ é o conjunto das k-uplas $(X_1 < X_2 < \ldots < X_k) \in [n]^k$. Note que, por definição, $W_G(x_i, x_j) = W_G(\frac{X_i}{n}, \frac{X_j}{n})$ para todo $x_i \in (\frac{X_i}{n}, \frac{X_i+1}{n}]$. Podemos então reescrever o somatório como uma integral múltipla compactando $[n]$ no intervalo $[0, 1]$.

$$t'(F,G) = \frac{n^k}{\binom{n}{k}} \int_{(0,1]} \int_{\left(\frac{1}{n}\lceil nx_1\rceil,1\right]} \cdots \int_{\left(\frac{1}{n}\lceil nx_{k-1}\rceil,1\right]} \prod_{ij\in E(F)} W_G(x_i,x_j)dx_1\dots dx_k$$

Note que a diferença dessa integral para a integral utilizada em $t'(F, W_G)$ é apenas nos casos em que algum dos $k-1$ pares $\{(x_1, x_2), (x_2, x_3), \dots, (x_{k-1}, x_k)\}$ caem em um mesmo intervalo. Dado $x \in [0, 1]^k_<$, seja $W_G^F(x) = \prod_{ij\in E(F)} W_G(x_i, x_j)$. Pelo princípio da inclusão e exclusão de conjunto, temos:

$$t'(F,G) = \frac{n^k}{\binom{n}{k}} \int_{(0,1]} \int_{\left(\frac{1}{n}\lceil nx_1\rceil,1\right]} \int_{\left(\frac{1}{n}\lceil nx_2\rceil,1\right]} \cdots \int_{\left(\frac{1}{n}\lceil nx_{k-1}\rceil,1\right]} W_G^F(x)dx$$

$$\geq \frac{n^k}{\binom{n}{k}} \Big(\int_{[0,1]^k_<} W_G^F(x)dx - \sum_{i=1}^{k-1} \int_0^1 \cdots \int_{x_{i-1}}^1 \int_{x_i}^{\frac{1}{n}\lceil nx_i\rceil} \int_{x_{i+1}}^1 \cdots \int_{x_{k-1}}^1 W_G^F(x)dx \Big)$$

Portanto,

$$\begin{aligned}
t'(F,G) &\geq \frac{n^k}{\binom{n}{k}} \Big(\int_{[0,1]^k_<} W_G^F(x)dx - \sum_{i=1}^{k-1} \frac{1}{n} \Big) \\
&\geq \frac{n^k}{\binom{n}{k}} \Big(\int_{[0,1]^k_<} W_G^F(x)dx - \frac{k-1}{n} \Big) \\
&\geq \frac{n^k}{\binom{n}{k}} \Big(\frac{t'(F,W_G)}{k!} - \frac{k-1}{n} \Big) \\
&\geq \frac{n^k}{(n)_k} t'(F,W_G) - \frac{(k-1)k!n^{k-1}}{(n)_k} \\
&\geq t'(F,W_G) - \frac{(k-1)k!n^{k-1}}{(n)_k}
\end{aligned}$$

Como queríamos mostrar.

□

Iremos provar agora o Teorema 56.

Prova do Teorema 56. Aplicando o Lema 66 obtemos uma subsequência de grafos ordenados (G'_n) da sequência original. Obtemos, também, uma sequência de inteiros e uma sequência de matrizes satisfazendo as condições (1) e (2). Utilizando o Lema 67, obtemos um *graphon* W com as propriedades (1) e (2) do Lema 67. Para concluir a demonstração, falta mostrar que, para todo grafo F, temos que $f(F) := \lim_{n\to\infty} t'(F, G'_n) = t'(F, W)$. Para $1 \leq j \leq m$, seja $G^*_{m,j} = K_{\mathcal{P}_{m,j},Q_{m,j}}$ e $G^{**}_{m,j} = K_{\mathcal{P}_{m,j},Q_j}$. Desta forma, temos pelo Lema 1 que

$$d_\square(G'_m, G^*_{m,j}) \quad \leq \quad \frac{1}{j} \tag{4.8}$$

E pela propriedade (3) do Lema 66 temos que

$$d_\square(G^*_m, G^{**}_{m,j}) \quad \leq \quad \frac{1}{j} \tag{4.9}$$

Seja $W_{m,j} = W_{G^{**}_{m,j}}$ e $W_j = W_{Q_j}$. Pelo Lema 68 temos que

$$|t'(F, W_{m,j}) - t'(F, G^{**}_{m,j})| \quad \leq \quad \frac{2(k+1)!}{n} \tag{4.10}$$

E pelo Lema 66 temos com probabilidade 1 que

$$W_{m,j} \quad \to \quad W_j \qquad (m \to \infty) \tag{4.11}$$

E finalmente, pelo Lema 67, temos que

$$W_j \quad \to \quad W \qquad (j \to \infty) \tag{4.12}$$

Para $\varepsilon > 0$ seja m_0 um inteiro positivo tal que para todo $m > m_0$ temos

$$|t'(F, G'_m) - f(F)| \quad < \quad \frac{\varepsilon}{4} \tag{4.13}$$

De (4.12), podemos escolher um inteiro positivo j tal que

$$|t'(F, W) - t'(F, W_j)| \quad < \quad \frac{\varepsilon}{4} \tag{4.14}$$

Podemos tomar $j > \frac{16}{k^2|E(F)|}$ e $j > m_0$. Assim, de (4.11), temos que

$$|t'(F, W_{m,j}) - t'(F, W_j)| \quad \leq \quad \frac{\varepsilon}{4} \tag{4.15}$$

E, pelo Lema 68 e pelas desigualdades (4.10), (4.9) e (4.8), temos que

$$\begin{aligned} |t'(F, G'_m) - t'(F, G^{**}_{m,j})| &\leq |E(F)|k^2 d_\square(G'_m, G^{**}_{m,j}) + \frac{\varepsilon}{8} \\ &\leq |E(F)|k^2 \frac{2}{j} + \frac{\varepsilon}{8} \\ &\leq \frac{\varepsilon}{4} \end{aligned} \tag{4.16}$$

Combinando as desigualdades (4.13), (4.14), (4.15) e (4.16), obtemos que

$$|f(F) - t(F, W)| \leq \varepsilon.$$

Como queríamos mostrar. □

4.7 Testabilidade Forte

Assim como foi definido para permutações, podemos definir uma noção de testabilidade forte para grafos ordenados.

Definição 69. *Uma propriedade $\mathcal{P}$ de grafos ordenas é fortemente testável por subgrafos, ou apenas testável, se, para todo $\varepsilon > 0$, existe um inteiro positivo k tal que, se G é um grafo em $[n]$, com $n > k$, então as duas afirmações abaixo são garantidas para um subgrafo aleatória $sub(k, G)$ de G de tamanho k:*

1. *$sub(k, G)$ satisfaz $\mathcal{P}$ com probabilidade pelo menos $1 - \varepsilon$ quando σ satisfaz $\mathcal{P}$;*

2. *$sub(k, G)$ não satisfaz $\mathcal{P}$ com probabilidade pelo menos $1 - \varepsilon$ quando $d_1(G, \mathcal{P}) \geq \varepsilon$ onde:*

$$d_1(G, \mathcal{P}) \ := \ \min\{d_1(G, F) : |V(F)| = n \text{ e } F \text{ satisfaz } \mathcal{P}\}.$$

Em [5], provou-se que toda propriedade hereditária para permutações é fortemente testável. No Lema 21 provamos que podemos visualizar toda permutação como um grafo ordenado. Nesta Seção, iremos mostrar uma demonstração alternativa ao resultado obtido por Bastos A. J. O. et al. em [5]. Para isso, iremos provar o resultado obtido em [5] para grafos ordenados, mais formalmente:

Teorema 70. *Toda propriedade hereditária para grafos ordenados é fortemente testável.*

Para isso, basta provar o seguinte resultado:

Teorema 71. *A distância d_1 de uma propriedade hereditária de grafos ordenados é testável.*

Seja $\mathcal{P}$ o conjunto formado por todas as funções mensuráveis simétricas $W : [0,1]^2 \to \mathbb{R}$. Seja $\mathcal{W}_0 \subset \mathcal{W}$ o conjunto dos *graphons*. Dizemos que um *graphon* W satisfaz $\mathcal{P}$ se existe um grafo ordenado G que satisfaz $\mathcal{P}$ tal que $W_G = W$. Em [21] provou-se os seguintes lemas, que serão fundamental para a demonstração do Teorema 71.

Lema 72 ([21])**.** *Seja $\mathcal{P}$ uma propriedade hereditária para grafos ordenados, e seja $\mathcal{H} \subseteq \mathcal{W}$ o conjunto formado por todos os* graphons *que satisfaz essa propriedade. Então $d_1(W, \mathcal{H})$ é uma função continua de W na norma $||.||_\square$.*

Lema 73 ([21])**.** *Se (G_n) é uma sequência convergente de grafos ordenados com a propriedade $\mathcal{P}$, e $G_n \to W$, então $W \in \mathcal{H}$.*

Lema 74 ([21])**.** *Para todo grafo ordenado G,*

$$d_1(G, \mathcal{P}) \leq d_1(W_G, \mathcal{H})$$

Podemos agora demonstrar o Teorema 71.

Demonstração do Teorema 71. Seja (G_n) una sequência convergente de grafos ordenados, e seja $W \in \mathcal{W}_0$ sei limite. Desejamos mostrar que $d_1(G_n, \mathcal{P})$ é convergente. Ou seja, temos que mostrar que

$$d_1(G_n, \mathcal{P}) \to d_1(W, \mathcal{H})$$

Pela unicidade, sabemos que $W_{G_n} \to W$, pelo Lema 72, $d_1(W_{G_n}) \to d_1(W, \mathcal{H})$. Pelo Lema 74, nos temos

$$d_1(G_n, \mathcal{P}) \leq d_1(W_{G_n}, \mathcal{H}) = d_1(W, \mathcal{H}) + o(1),$$

e então

$$\limsup_{n\to\infty} d_1(G_n, \mathcal{P}) \leq d_1(W, \mathcal{H}).$$

Para mostrar que

$$\liminf_{n\to\infty} d_1(G_n, \mathcal{P}) \geq d_1(W, \mathcal{H}).$$

suponha por absurdo que o limite inferior é estritamente menor. Então, tomando subsequências, temos que

$$d_1(G_n, \mathcal{P}) \to c < d_1(W, \mathcal{H}).$$

Para cada G_n, seja H_n um grafo ordenado que satisfaça a propriedade $\mathcal{P}$ e possui o mesmo conjunto de vértices de G_n tal que $d_1(G_n, H_n) = d_1(G_n, \mathcal{P})$. Podemos assumir que (H_n) é convergente, seja $U \in \mathcal{W}_0$ o seu limite. Pelo Lema 73, temos que $U \in \mathcal{H}$. Claramente

$$d_1(W_{G_n}, W_{H_n}) = d_1(G_n, H_n) = d_1(G_n, \mathcal{P}).$$

Consequentemente

$$d_1(W_{G_n}, \mathcal{H}) \leq d_1(W_{G_n}, W_{H_n}) + d_1(W_{H_n}, \mathcal{H}) = d_1(G_n, \mathcal{P}) + d_1(W_{H_n}, \mathcal{H})$$

Nesta desigualdade temos que $d_1(W_{G_n}, \mathcal{H}) \to d_1(W, \mathcal{H})$ e $d_1(W_{H_n}, \mathcal{H}) \to d_1(U, \mathcal{H}) = 0$ pelo Lema 72. Portanto $d_1(W, \mathcal{H}) \leq c$, que é um absurdo.

□

Capítulo 5

Considerações Finais

Neste trabalho apresentamos um estudo baseado em convergência de sequência de estruturas. Que consiste em construir e provar que determinados espaços de estruturas são completos. Apresentamos também as aplicações mais imediatas desta ferramenta (testabilidade e quase aleatoriedade). Em especial demos destaque para a testabilidade, pois serviu de incentivo inicial para desenvolver a teoria de sequências convergentes.

Uma propriedade, sobre um determinado tipo de estrutura, é dita testável se podemos decidir com alta probabilidade e em tempo constante se uma estrutura qualquer possui ou não essa propriedade. Apresentamos neste trabalho que toda propriedade hereditária para grafos e permutações é testável. Obtemos alguns resultados preliminares para gerar uma possível generalização de algumas classes de estruturas. Para isso, desenvolvemos a teoria de sequência convergente para grafos ordenados, pois como mostrado neste trabalho, estruturas que previamente já haviam sido desenvolvidos essa teoria podem ser vistas como um caso particular de grafos ordenados. Por fim, provamos que toda propriedade hereditária de grafos ordenados é fortemente testável.

Atualmente Lovász e Szegedy vêm desenvolvendo resultados relacionados à sequência convergente de grafos e *graphon*. Temos nos aprofundado mais nestes assuntos e esperamos conseguir criar resultados similares para permutações e grafos ordenados. Outra questão interessante seria mostrar se existe alguma relação entre os objetos limites mostrada neste trabalho.

Índice Remissivo

Referências Bibliográficas

[1] N. Alon, E. Fischer, I. Newman, and A. Shapira. A combinatorial characterization of the testable graph properties: it's all about regularity. *SIAM J. Comput.*, 39(1):143–167, 2009.

[2] N. Alon and A. Shapira. A characterization of the (natural) graph properties testable with one-sided error. *SIAM J. Comput.*, 37(6):1703–1727, 2008.

[3] N. Alon and A. Shapira. Every monotone graph property is testable. *SIAM J. Comput.*, 38(2):505–522, 2008.

[4] T. Austin and T. Tao. Testability and repair of hereditary hypergraph properties. *Random Structures Algorithms*, 36(4):373–463, 2010.

[5] A.J.O. Bastos, C. Hoppen, Y. Kohayakawa, C. G. Moreira, and R. M. Sampaio. Every hereditary permutation property is testable. In *Proc. of EuroComb'11*, volume 38 of *Electronic Notes in Discrete Mathematics*, 2011.

[6] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008.

[7] F. R. K. Chung and R. L. Graham. Quasi-random subsets of Z_n. *J. Combin. Theory Ser. A*, 61(1):64–86, 1992.

[8] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.

[9] J. N. Cooper. Quasirandom permutations. *J. Combin. Theory Ser. A*, 106(1):123–143, 2004.

[10] R. C. Corrêa, C. Hoppen, Y. Kohayakawa, and R. M. Sampaio. A note on random k-dimensional posets. In *Proc. VI Latin-American Algorithms,*

Graphs, and Optimization Symposium - LAGOS 2011, volume 37 of *Electronic Notes in Discrete Mathematics*, pages 51 – 56, 2011.

[11] D. P. Dubhashi and A. Panconesi. *Concentration of measure for the analysis of randomized algorithms.* Cambridge University Press, Cambridge, 2009.

[12] P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[13] A. Frieze and R. Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19(2):175–220, 1999.

[14] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45(4):653–750, 1998.

[15] C. Hoppen, Y. Kohayakawa, C. G. Moreira, and R. M. Sampaio. Limits of permutation sequences. *preprint*, 2011.

[16] C. Hoppen, Y. Kohayakawa, C. G. Moreira, and R. M. Sampaio. Testing permutation properties through subpermutations. *to appear in Theoretical Computer Science*, 2011.

[17] S. Janson. Poset limits and exchangeable random posets, 2009.

[18] L. Lovász and V. T. Sós. Generalized quasirandom graphs. *J. Combin. Theory Ser. B*, 98(1):146–163, 2008.

[19] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.

[20] L. Lovász and B. Szegedy. Testing properties of graphs and functions. *Israel J. Math.*, 178:113–156, 2010.

[21] L. Lovász and B. Szegedy. Graph limits and testing hereditary graph properties, 2005.

[22] R. M. Sampaio. Limits of permutation sequences. Tese de doutorado, IME-USP, 2008.