Comunicação e Redes – 2020.QS

Lista 2 – Universidade Federal do ABC

Aritanan Gruber

aritanan.gruber@ufabc.edu.br

http://professor.ufabc.edu.br/∼aritanan.gruber

Exercícios marcados com ▶ devem ser entregues via Tidia até 17/12/2020 às 23h00m.

Notação. Exercícios marcados com: ♦ são operacionais, e envolvem a execução de cálculos simples ou o fornecimento de (contra) exemplos; ♥ são sobre teoria de redes, e geralmente requerem provas matemáticas; ♣ também são sobre redes, mas requerem a construção de um algoritmo; ♠ envolvem matemática de maneira mais geral, não limitada a grafos e variações; ★ são exercícios ligeiramente mais difíceis; finalmente, ▶ são exercícios que devem ser entregues via Tidia.

1. ▶ ♥ Seja $G = (V, E)$ um grafo conexo com n vértices e defina o **comprimento médio de caminhos** em G como

$$\widehat{\ell}(G) := \frac{1}{n(n-1)} \sum_{\substack{u,v \in V \\ u \neq v}} \mathsf{dist}(u, v).$$

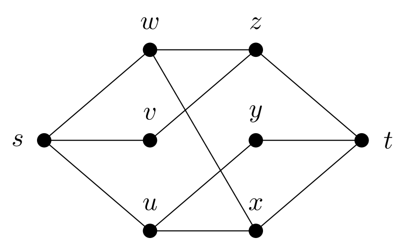
Lembre-se que o **diâmetro** de G é dado por: $\mathsf{diam}(G) := \max_{s \in V}\{\mathsf{dist}(s, v) : v \in V\}$. Para muitos grafos, os valores de comprimento médio de caminhos e diâmetro são próximos, mas para outros eles podem ser muito diferentes.

   (a) Forneça um grafo G tal que $3\,\widehat{\ell}(G) < \mathsf{diam}(G)$.

   (b) Descreva como estender sua construção acima para que, dada uma constante $C > 1$, seja possível obter um grafo G tal que $C\,\widehat{\ell}(G) < \mathsf{diam}(G)$.

2. ▶ ♦ Dados um grafo $G = (V, E)$ conexo e vértices distintos $s, t \in V$, um subconjunto $\emptyset \neq X \subseteq V \setminus \{s, t\}$ é um **(s, t)-separador** para G se: (i) $\kappa(G - X) > \kappa(G)$, isto é, se após a remoção dos vértices em X e das arestas que neles incidem, o subgrafo resultante tem mais componentes que o grafo original; (ii) os vértices s e t pertencem a componentes diferentes de $G - X$. Um (s, t)-separador X^* é **mínimo** se $|X^*| \leq |X|$ para todo (s, t)-separador X.

Dois (s, t)-caminhos $P_1 = \langle s = u_0, u_1, \ldots, u_i = t\rangle$ e $P_2 = \langle s = v_0, v_1, \ldots, v_j = t\rangle$ são **internamente disjuntos** se $V(P_1) \cap V(P_2) = \{s, t\}$. Considere o grafo G abaixo.



   (a) Determine um conjunto separador mínimo X^* para G. Seja k a cardinalidade de X^*.

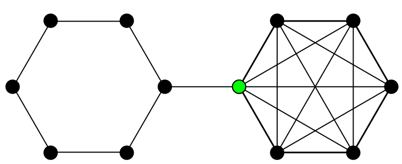
   (b) Determine ℓ (s, t)-caminhos internamente disjuntos em G, com ℓ o maior possível.

   (c) Qual a relação entre k e ℓ? Ela prova que X^* como determinado em (a) é mínimo? Por que?

3. ▶ ♣ Dado um grafo $G = (V, E)$, a **centralidade harmônica** de um vértice $u \in V$ é dada por

$$H_u := \frac{1}{n-1} \sum_{v \in V - u} \frac{1}{\mathsf{dist}(u, v)}.$$

Escreva um algoritmo (não exaustivo) que recebe G e devolve, para cada vértice $u \in V$, a centralidade harmônica H_u de u. Para crédito extra, determine o vértice de maior centralidade harmônica no grafo abaixo.



4. ▶ ♥ Seja G um grafo com n vértices tal que $\delta(G) > \frac{1}{2}(n-2)$, em que $\delta(G) := \min\{d(v) : v \in V(G)\}$ é o grau mínimo dos vértices de G. Mostre que o **diâmetro** de G é no máximo 2.

5. ▶ ♥ Considere o modelo de grafos aleatórios definido a seguir. Sejam $n \geq 0$ inteiro, $p = p(n) \in [0,1]$ e $q = q(n) \in [0,1]$ duas medidas de probabilidade. Defina o espaço de probabilidade $\mathscr{G}(n,p,q)$ em que cada $G = ([2n], E) \sim \mathscr{G}(n,p,q)$ é um grafo aleatório que possui $2n$ vértices, sendo n vermelhos (R) e n azuis (B), e tal que para cada potencial aresta $\{x,y\} \in \binom{[2n]}{2}$,

$$\mathbb{P}\mathrm{r}[\{x,y\} \in E] = \begin{cases} p & \text{se } x \text{ e } y \text{ tiverem a mesma cor,} \\ q & \text{se } x \text{ e } y \text{ tiverem cores diferentes.} \end{cases}$$

Um grafo $G \sim \mathscr{G}(n,p,q)$ é **esnobe** se $p > q$, capturando a tendência a vértices de uma mesma cor a se conectarem. Se $q = 0$, o grafo possui dois componentes conexos disjuntos: um para cada cor de vértices.

   (a) Determine a distribuição de graus para os vértices de $G \sim \mathscr{G}(n,p,q)$.

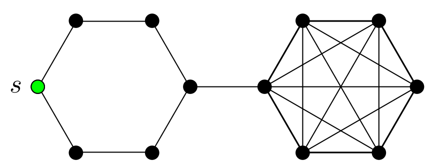
   (b) Calcule o grau médio de G e o grau médio do subgrafo $B \subseteq G$ induzido pelos vértices azuis – isto é, $B = G\big[\{v \in [2n] : v \text{ é azul}\}\big]$.

   (c) Sejam M e N duas variáveis aleatórias que contam os números de triângulos monocromáticos e não monocromáticos (com relação às cores dos vértices) em G, respectivamente. Determine $\mathbb{E}[M]$ e $\mathbb{E}[N]$.

   (d) Para $p = 1/\sqrt{2n}$, mostre que G tem triângulos monocromáticos com alta probabilidade. Sugestão: determine $\mathrm{Var}(M)$ e utilize a desigualdade de Chebyshev.

6. ▶ ♦ ♣ Nesta questão vamos considerar um modelo epidemiológico simples. Seja $G = (V, E)$ um grafo conexo com $n \geq 1$ vértices e $p \in [0,1]$. Inicialmente, todos os vértices em V são sadios. Um vértice $s \in V$ é então infectado com um patógeno para o qual não existe cura. A infecção se propaga em G da seguinte forma: se num dado momento, ao menos uma fração p dos vizinhos de um vértice $u \in V$ estão infectados, u contrai o patógeno e torna-se infectado. Como G é finito, claramente o processo de infecção termina num quadro que denominamos **equilíbrio**.

   (a) Determine o quadro de equilíbrio da infecção no grafo abaixo para $p = 1/2$.



   (b) Modifique o algoritmo de busca em largura para determinar, dado (G, s, p), o quadro de equilíbrio da infecção. Devolva a condição da cada vértice neste quadro e a fração de vértices infectados.

   (c) Suponha que s, o primeiro vértice infectado, tenha centralidade por grau máxima em G. O que pode ser dito sobre o quadro de equilíbrio da infecção neste caso? Sua resposta deve, naturalmente, depender de p.

Nota: um exemplo de tal patógeno é a pressão do grupo de "amigos" em adolescentes e jovens adultos (*peer pressure*).

7. ▶ ♦ Seja C_n um ciclo com $n \geq 3$ vértices. Determine o autovalor principal λ_1 de C_n. Naturalmente, sua resposta deve ser dada em função de n. O conceito de centralidade de vértices faz sentido para C_n? Justifique.

8. ▶ ♥ Seja $G \sim \mathscr{G}(n,p)$ um grafo aleatório Erdős-Rényi.

   (a) Para $k \geq 3$ inteiro fixado, seja $\mathscr{C}_k$ o evento em que um k-ciclo surge em G. Determine $\mathbb{P}\mathrm{r}[\mathscr{C}_k]$. Em outras palavras, calcule a probabilidade de G ter (ao menos) um k-ciclo.

   (b) Seja C_k uma variável aleatória que conta o número de k-ciclos em G. Determine $\mathbb{E}[C_k]$. Defina agora $C = \sum_{k=3}^{n} C_k$ e determine $\mathbb{E}[C]$. O que $\mathbb{E}[C]$ representa?

   (c) Suponha que $p = p(n) = 1/n^2$. Mostre que, com alta probabilidade, G é uma floresta. *Dica:* utilize a desigualdade de Markov.

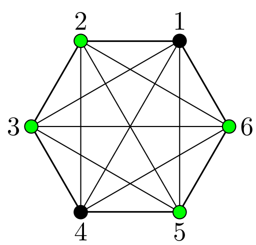
   (d) Suponha que $p = p(n) = 1/\sqrt{n}$. Mostre que, com alta probabilidade, G tem ciclos. *Dica:* ciclos em G podem não ser independentes; calcule $\mathbb{V}\mathrm{ar}[C]$ e utilize a desigualdade de Chebyshev.

9. ▶ ♥ ♦ Seja $K_n = (V, E)$ o grafo completo com $n \geq 0$ vértices e $r : E \to \mathbb{R}_{\geq 0}$ uma função para as arestas $e \in E$ que denota um coeficiente de **rejeição** entre as pontas de e. Para $S \subseteq V$, defina o **índice de rejeição** entre S e $V \setminus S$ por $r(\partial(\emptyset)) = r(\partial(V)) := 0$ e

$$r(\partial(S)) := \sum_{e \in \partial(S)} r(e),$$

em que $\partial(S) := \big\{\{u, v\} \in E : u \in S,\ v \in V \setminus S\big\}$ é o conjunto de arestas com uma ponta em S e a outra em $V \setminus S$. Podemos interpretar S e $S \setminus V$ como duas comunidades: os vértices em uma mesma comunidade tem rejeição menor entre si do que para com vértices da outra comunidade.

(a) No grafo abaixo ($n = 6$), com vértices em S pretos/sólidos e em $V \setminus S$ verdes/ocos, determine o valor $r(\partial(S))$.



$$r = \left(\begin{array}{c|cccccc} u/v & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & 0 & 1 & 1 & 3 & 1 & 5 \\ 2 & 1 & 0 & 1 & 2 & 0 & 5 \\ 3 & 1 & 1 & 0 & 2 & 1 & 5 \\ 4 & 3 & 2 & 2 & 0 & 2 & 5 \\ 5 & 1 & 0 & 1 & 2 & 0 & 5 \\ 6 & 5 & 5 & 5 & 5 & 5 & 0 \end{array}\right)$$

(b) Ainda para o grafo acima, determine $S^* \subseteq V$ tal que $r(\partial(S^*))$ tem valor máximo: $S^* = \operatorname{argmax}_{S \subseteq V} r(\partial(S))$. Qual a interpretação para $\partial(S^*)$?

(c) Agora para K_n e r quaisquer, considere o seguinte algoritmo probabilístico: inicialmente, $S = \emptyset$; cada vértice $v \in V$ é adicionado a S, de forma independente e uniforme, com probabilidade $1/2$; ao final, S é devolvido. Mostre que $\mathbb{E}[r(\partial(S))] \geq r(\partial(S^*))/2$, em que $r(\partial(S^*))$ tem valor máximo (sobre K_n e r).
Dica: defina uma variável aleatória $I_e = r(e)$ se $e \in \partial(S)$ e $I_e = 0$ em caso contrário, para cada aresta $e \in E$; calcule $\mathbb{E}[I_e]$; use a linearidade da esperança.

10. ▶ ♦ Seja $G = (V, E)$ um grafo em que os vértices representam indivíduos e as arestas relações de amizade. Um subconjunto $S \subseteq V$ é uma **super anti-comunidade** se os indivíduos em S não possuem amizades entre si. Escreva um algoritmo que recebe G e devolve uma super anti-comunidade maximal. Forneça um grafo de exemplo em que uma super anti-comunidade maximal não é máxima.

11. ★ ♥ *High chromatic number II.* Prove que para quaisquer inteiros $k, \ell \geq 0$, existe um grafo G com cintura $g(G) > \ell$ e número cromático $\chi(G) > k$.

12. ▶ ♥ Seja $G = (V, E)$ uma **árvore** – isto é, um grafo acíclico e conexo – com $n \geq 0$ vértices e seja $e = \{u, v\} \in E$ uma aresta. A remoção de e divide G em duas sub-árvores com n_u e n_v vértices cada. Mostre que as centralidades por proximidade C_u e C_v de u e v, respectivamente, satisfazem:

$$\frac{1}{C_u} + \frac{n_u}{n} = \frac{1}{C_v} + \frac{n_v}{n}.$$

Lembre-se que a centralidade de um vértice w é dada por

$$C_w = \frac{n}{\sum_{x \in V} \mathsf{dist}(w, x)},$$

em que $\mathsf{dist}(w, x)$ é a distância entre w e x.

Cheat Sheet

Algoritmo de Busca em Largura em um grafo G à partir de uma origem $s \in V(G)$.

```
BFS (G, s)
  p[u] ← u para cada u ∈ V(G)        # vetor de predecessores
  d[u] ← ∞ para cada u ∈ V(G) \ {s}  # vetor de distâncias
  d[s] ← 0                           # s é o vértice de origem

  Q ← fila ({s})
  enquanto Q ≠ ∅ faca
    u ← remova-frente (Q)
    para cada v ∈ N(u) faca
      se d[v] > d[u] + 1 entao
        d[v] ← d[u] + 1
        p[v] ← u
        insira-final (Q, v)
  devolva d, p
```

Abaixo, $n \geq 0$ inteiro e $p = p(n) \in [0,1]$.

- Se $X : \Omega \to \mathbb{Z}$ é uma variável aleatória inteira, a **esperança** de X é $\mathbb{E}[X] := \sum_j j\,\mathbb{P}\mathrm{r}[X = j]$. A **variância** de X é $\mathbb{V}\mathrm{ar}[X] := \mathbb{E}\big[(X - \mathbb{E}[X])^2\big] = \mathbb{E}[X^2] - \mathbb{E}^2[X]$. ($\Omega$ é um **espaço amostral** e $\mathbb{P}\mathrm{r}$ uma medida de probabilidade.)
- Se X e Y são variáveis aleatórias e $\alpha \in \mathbb{R}$, então $\mathbb{E}[\alpha X + Y] = \alpha\mathbb{E}[X] + \mathbb{E}[Y]$, fato conhecido como **linearidade** da esperança. Se X e Y são **independentes**, isto é, $\mathbb{P}\mathrm{r}[X \wedge Y] = \mathbb{P}\mathrm{r}[X]\,\mathbb{P}\mathrm{r}[Y]$, então $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$.
- Se X e Y são variáveis aleatórias e $\alpha \in \mathbb{R}$, então $\mathbb{V}\mathrm{ar}[\alpha X + Y] = \alpha^2\mathbb{V}\mathrm{ar}[X] + \mathbb{V}\mathrm{ar}[Y] + 2\alpha\mathbb{C}\mathrm{ov}[X,Y]$, em que $\mathbb{C}\mathrm{ov}[X,Y] := \mathbb{E}\big[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\big] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$ é a **covariância** de X e Y. $\mathbb{C}\mathrm{ov}[X,Y] = 0$ se X e Y são independentes.
- Distribuição **Bernoulli**: se $X \sim \mathsf{Ber}(p)$, então $\mathbb{P}\mathrm{r}[X = 1] = p$ e $\mathbb{P}\mathrm{r}[X = 0] = 1 - p$; $\mathbb{E}[X] = 1 \cdot p + 0 \cdot (1-p) = p$; $\mathbb{V}\mathrm{ar}(X) = p(1-p)$. $\mathbb{E}[X] \geq \mathbb{V}\mathrm{ar}[X]$.
- Distribuição **Binomial**: se $X \sim \mathsf{Bin}(n,p)$, então $\mathbb{P}\mathrm{r}[X = k] = \binom{n}{k}p^k(1-p)^{n-k}$. Além disso, $X = \sum_{j=1}^{n} Y_j$ em que cada Y_j é independente e identicamente distribuída como $Y_j \sim \mathsf{Ber}(p)$. Logo, $\mathbb{E}[X] = \mathbb{E}\big[\sum_{j=1}^{n} Y_j\big] = \sum_{j=1}^{n}\mathbb{E}[Y_j] = np$; $\mathbb{V}\mathrm{ar}(X) = np(1-p)$. $\mathbb{E}[X] \geq \mathbb{V}\mathrm{ar}[X]$.
- Distribuição **Poisson**: se $X \sim \mathsf{Poi}(\lambda)$, com $\lambda \in \mathbb{R}$, então $\mathbb{P}\mathrm{r}[X = k] = e^{-\lambda}\lambda^k/k!$; $\mathbb{E}[X] = \mathbb{V}\mathrm{ar}[X] = \lambda$.
- Desigualdade de **Markov**: Se X é uma variável aleatória não negativa e $t \in \mathbb{R}_{>0}$, então $\mathbb{P}\mathrm{r}[X \geq t] \leq \mathbb{E}[X]/t$. De outra forma, $\mathbb{P}\mathrm{r}\big[X \geq t\mathbb{E}[X]\big] \leq 1/t$.
- **Primeiro momento** probabilístico: se X é uma variável aleatória e $\mathbb{E}[X] = k$, então $\mathbb{P}\mathrm{r}[X \geq k] > 0$ e $\mathbb{P}\mathrm{r}[X \leq k] > 0$.
- Desigualdade de **Chebyshev**: Se X é uma variável aleatória e $t \in \mathbb{R}_{>}$, então $\mathbb{P}\mathrm{r}\big[|X - \mathbb{E}[X]| \geq t\big] \leq \mathbb{V}\mathrm{ar}(X)/t^2$. De outra forma, $\mathbb{P}\mathrm{r}[X = 0] \leq \mathbb{P}\mathrm{r}\big[|X - \mathbb{E}[X]| \geq \mathbb{E}[X]\big] \leq \mathbb{V}\mathrm{ar}(X)/\mathbb{E}^2[X]$.
- **Segundo momento** probabilístico: Se X é uma variável aleatória não negativa, $X > 0$ com **alta probabilidade** se $\mathbb{P}\mathrm{r}[X = 0] = o(1)$, em que $o(1) = f(n) \to 0$ quando $n \to \infty$; $X = 0$ com alta probabilidade se $\mathbb{P}\mathrm{r}[X = 0] = 1 - o(1)$.
- **Grafo aleatório Binomial**: $\mathscr{G}(n,p)$ é um espaço de probabilidade em que cada elemento $G \sim \mathscr{G}(n,p)$ é um grafo aleatório com $V(G) = \{1, 2, \ldots, n\}$ e número médio de arestas $\mathbb{E}\big[|E(G)|\big] = p\binom{n}{2}$; isto é, cada possível aresta de $\binom{[n]}{2}$ está presente em G com probabilidade p, independente das demais.
- **Distribuição de graus**: Com $G \sim \mathscr{G}(n,p)$, para $v \in V(G)$ e $k \geq 0$ inteiro, $\mathbb{P}\mathrm{r}[d(v) = k] = \binom{n-1}{2}p^k(1-p)^{n-1-k} \approx \binom{n}{2}p^k(1-p)^{n-k}$; $\mathbb{E}[\mu(G)] = (n-1)p \approx np$; $\mathbb{V}\mathrm{ar}[\mu(G)] = (n-1)p(1-p) \approx np(1-p)$. Se $p = \lambda/n$ para $\lambda \in \mathbb{R}_{>0}$ constante, $d(v) \sim \mathsf{Poi}(\lambda)$ e $\mathbb{P}\mathrm{r}[d(v) = k] = e^{-\lambda}\lambda^k/k!$; $\mathbb{E}[\mu(G)] = \mathbb{V}\mathrm{ar}[\mu(G)] = \lambda$.