

Redes Complexas - CPS765 - 2020.2

Primeira Lista de Exercícios

Pedro Maciel Xavier

116023847

7 de novembro de 2020

Questão 1.: Matriz de Adjacência

Podemos supor, neste caso, que as matrizes em questão vivem em um espaço vetorial construído sobre o semianel $(\mathcal{B}, \vee, \wedge)$ em vez de $(\mathbb{R}, +, \cdot)$, onde $\mathcal{B} = \{0, 1\}$. Assim, as entradas das matrizes serão sempre 0 ou 1 e as operações usuais de soma e multiplicação são substituídas pela disjunção e pela conjunção lógica, respectivamente.

a) A fim de obter uma expressão para a alcançabilidade em k passos do vértice i ao j , dado pela entrada $\mathbf{B}_{i,j}^{(k)}$ vamos empregar um raciocínio indutivo. É claro que a alcançabilidade em 0 passos é dada pela matriz identidade \mathbf{I} , uma vez que só é possível chegar ao vértice em que já encontramos. O caso para um único passo é dado pela matriz de adjacências \mathbf{A} , trivialmente. Logo, $\mathbf{B}^{(0)} = \mathbf{I}$ e $\mathbf{B}^{(1)} = \mathbf{A}$. Vamos supor, por hipótese de indução, que a matriz $\mathbf{B}^{(k)} \in \mathcal{B}^{n \times n}$ representa a alcançabilidade em exatamente k passos, isto é, se existe um caminho de comprimento k ligando o vértice i ao vértice j , então $\mathbf{B}_{i,j}^{(k)} = 1$. Caso contrário, $\mathbf{B}_{i,j}^{(k)} = 0$. Para saber se existe um caminho de tamanho $k + 1$ entre os vértices i e j é preciso que exista um caminho de tamanho k entre i e algum vértice ξ assim como ξ deve ser incidente em j . Portanto,

$$\mathbf{B}_{i,j}^{(k+1)} = \bigvee_{\xi=1}^n \mathbf{B}_{i,\xi}^{(k)} \wedge \mathbf{A}_{\xi,j}$$

de onde concluímos quem, para todo $k \geq 1$, $\mathbf{B}^{(k+1)} = \mathbf{B}^{(k)} \mathbf{A}$. O resultado é dado pelo produto usual de matrizes induzido pelo semianel booleano. Logo, escrevemos $\mathbf{B}^{(k)} = \mathbf{A}^k$.

b) Seguindo raciocínio semelhante, dizemos que i alcança j em k ou menos passos se $\mathbf{B}_{i,j}^{(\xi)} = 1$ para algum $0 \leq \xi \leq k$. Isto é,

$$\mathbf{C}_{i,j}^{(k)} = \mathbf{B}_{i,j}^{(0)} \vee \mathbf{B}_{i,j}^{(1)} \vee \mathbf{B}_{i,j}^{(2)} \dots \vee \mathbf{B}_{i,j}^{(k)} = \bigvee_{\xi=0}^k \mathbf{B}_{i,j}^{(\xi)}$$

resultado que, por conta do espaço onde as matrizes se encontram, é caracterizado pela soma usual. Ou seja, $\mathbf{C}^{(k)} = \sum_{\xi=0}^k \mathbf{B}^{(\xi)}$.

c) Análise da complexidade:

$\mathbf{B}^{(k)}$ - A multiplicação usual de matrizes tem custo $O(n^3)$. Como temos de calcular este produto $k - 1$ vezes, temos uma complexidade assintótica total de ordem $O(n^3 k)$.

$\mathbf{C}^{(k)}$ - A soma de matrizes possui complexidade $O(n^2)$. Contando as $k - 1$ somas temos um total de $O(n^2 k)$ para esta etapa. Se recalculamos $\mathbf{B}^{(k)}$ a cada passo, a complexidade das multiplicações segue uma progressão aritmética em k , totalizando $O(n^3 k^2)$. Se aproveitamos a matriz anterior a cada soma, podemos realizar este processo em tempo $O(n^3 k)$. O termo quadrático em n é de ordem inferior e pode ser omitido em ambos os casos.

d) Seguindo o conselho de multiplicar diferentemente, apresento duas abordagens para reduzir a complexidade do cálculo de $\mathbf{B}^{(k)}$ e $\mathbf{C}^{(k)}$. A primeira, se aplica a um grafo qualquer e se baseia

na seguinte relação:

$$\mathbf{A}^k = \begin{cases} \mathbf{I} & \text{para } k = 0 \\ \left(\mathbf{A}^{\frac{k}{2}}\right)^2 & \text{para } k \text{ par} \\ \left(\mathbf{A}^{\frac{k-1}{2}}\right)^2 * \mathbf{A} & \text{para } k \text{ ímpar} \end{cases}$$

para $k \geq 0$. Em geral, esta relação vale para qualquer operação $*$ associativa e, portanto, utilizaremos para o cálculo das potências de matrizes. Isso nos traz complexidade $O(\log k)$ nesta tarefa. Com este aprimoramento, somos capazes de calcular $\mathbf{B}^{(k)}$ em tempo $O(n^3 \log k)$ enquanto $\mathbf{C}^{(k)}$ sai por $O(n^3 \log k!) = O(n^3 k \log k)$. Apesar do ganho no cálculo de $\mathbf{B}^{(k)}$, simplesmente aplicar este método requer calcular cada ξ -ésima potência de \mathbf{A} . É melhor, portanto, multiplicar por \mathbf{A} e somar ao resultado iterativamente, com custo $O(n^3 k + n^2 k) = O(n^3 k)$.

Ainda podemos fazer melhor em alguns casos, dadas algumas condições. Vamos retornar ao espaço euclidiano usual $\mathbb{R}^{n \times n}$, pagando um custo $O(n^2)$ ao final do cálculo para definir cada entrada de $\mathbf{B}^{(k)}$ e $\mathbf{C}^{(k)}$ como 0 ou 1 verificando se cada elemento da matriz é ou não nulo, respectivamente. Se $\det(\mathbf{I} - \mathbf{A}) = (-1)^n \det(\mathbf{A} - \mathbf{I}) \neq 0$, isto é, para cada autovalor λ de \mathbf{A} temos que $\lambda \neq 1$, podemos dizer que

$$\sum_{\xi=0}^k \mathbf{A}^\xi = (\mathbf{I} - \mathbf{A})^{-1} (\mathbf{I} - \mathbf{A}^{k+1})$$

A inversão da matriz $\mathbf{I} - \mathbf{A}$ pode ser feita em tempo $O(n^3)$ pela eliminação de *Gauss-Jordan*. Isso nos permite calcular $\mathbf{C}^{(k)}$ em tempo $O(n^3) + O(n^3 \log k)$, ou seja, $O(n^3 \log k)$.

Temos ainda um caso ainda mais específico, para grafos não-direcionados. Neste caso, a estrutura confere $\mathbf{A} = \mathbf{A}^T$. O Teorema Espectral nos garante, portanto, que a matriz \mathbf{A} é diagonalizável e, além disso, a simetria permite encontrar a forma $\mathbf{A} = \mathbf{S} \mathbf{\Lambda} \mathbf{S}^{-1}$ em tempo $O(n^3)$ através da transformação de *Householder* seguida da aplicação do algoritmo *QR*. Em seguida, calculamos $\mathbf{B}^{(k)} \sim \mathbf{A}^k = \mathbf{S} \mathbf{\Lambda}^k \mathbf{S}^{-1}$ em tempo $O(n^3 + n \log k)$, uma vez que basta calcular a potência de cada uma das n entradas da diagonal principal de $\mathbf{\Lambda}$, o que possui complexidade $O(\log k)$ segundo o método visto acima. Por fim, isso também se aplica ao cálculo de $\mathbf{C}^{(k)}$, que pelo método da soma geométrica de matrizes descrito acima pode ser feito em tempo $O(n^3 + n \log k)$. Isso se verifica também por outra propriedade da forma diagonal de \mathbf{A} , uma vez que

$$\sum_{\xi=0}^k \mathbf{A}^\xi = \mathbf{S} \left[\sum_{\xi=0}^k \mathbf{\Lambda}^\xi \right] \mathbf{S}^{-1} = \mathbf{S} \begin{bmatrix} \sum_{\xi=0}^k \lambda_1^\xi & & \\ & \ddots & \\ & & \sum_{\xi=0}^k \lambda_n^\xi \end{bmatrix} \mathbf{S}^{-1} = \mathbf{S} \begin{bmatrix} \frac{1 - \lambda_1^{k+1}}{1 - \lambda_1} & & \\ & \ddots & \\ & & \frac{1 - \lambda_n^{k+1}}{1 - \lambda_n} \end{bmatrix} \mathbf{S}^{-1}$$

onde fica clara a condição de que $\lambda \neq 1$.

Questão 2.: Grau médio e densidade

Para analisar o comportamento das duas propriedades (grau médio \bar{g} e densidade ρ) vamos escrever cada uma em função da outra. Partimos das expressões

$$\bar{g} = \frac{2m}{n} \text{ e } \rho = \frac{2m}{n(n-1)}$$

de onde definimos as funções

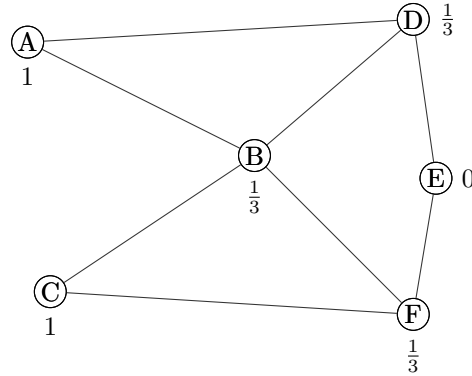
$$\bar{g}(\rho) = \rho(n-1) \text{ e } \rho(\bar{g}) = \frac{\bar{g}}{(n-1)}$$

O grau médio é, portanto, uma dilatação da densidade por um fator $(n-1)$ para $n \in \mathbb{N}$ qualquer. Possui comportamento monotônico crescente como função de ρ . Desta forma, para um grafo

qualquer de n vértices, independentemente de sua configuração de arestas, sabemos que o grau médio e a densidade estão relacionados de maneira linear. Mais do que isso, dizemos que $\bar{g} \propto \rho$. Isso garante que altas densidades levam a um alto grau médio da rede, e vice-versa.

Questão 3.: clusterização

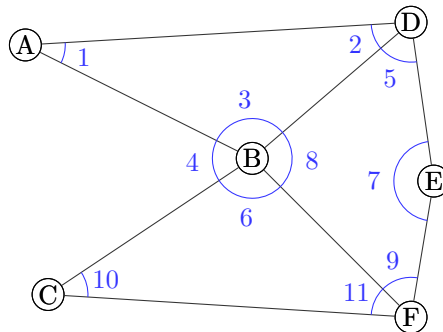
1 Calculando a clusterização local de cada vértice do grafo:



A clusterização média será, portanto,

$$\frac{1}{6} \sum_{i=A}^F c_i = \frac{1}{2} = 0.5$$

2 Contando as triplas:



Temos, então, 11 triplas e 2 triângulos ($\triangle ABD$ e $\triangle BCF$). Logo, a clusterização global é dada por

$$\frac{3 \times \text{n}^\circ \text{ de triângulos}}{\text{n}^\circ \text{ de triplas}} = \frac{6}{11} = 0.\overline{54}$$

3 A densidade da rede é dada por simplesmente $\rho = \frac{2m}{n(n-1)}$, ou seja, $\rho = \frac{16}{30} = 0.\overline{53}$. Observa-se que o valor da densidade se encontra entre o da clusterização média e o da clusterização global, estando mais próximo desta última.

Questão 4.: Closeness

Lembrando que o *closeness* de um vértice v é dado pela expressão

$$c_v = \frac{1}{n-1} \sum_{u \neq v \in V} d(u, v)$$

onde $d(u, v)$ é a distância entre u e v definimos as métricas globais:

- *Average Pairwise Distance* (APD):

$$\text{APD}(V, E) = \frac{1}{\binom{n}{2}} \sum_{u \in V} \sum_{v > u \in V} d(u, v)$$

- *Average Path Length* (APL):

$$\text{APL}(V, E) = \frac{1}{n(n-1)} \sum_{u \in V} \sum_{v \in V} d(u, v)$$

É importante notar que a métrica APD só se aplica a grafos não-direcionados, uma vez que os pares (u, v) do somatório sempre tem $v > u$. Aplicar sobre grafos direcionados geraria assimetria conforme a numeração dos vértices.

1 Reescrevendo APD:

$$\begin{aligned} \text{APD}(V, E) &= \frac{1}{\binom{n}{2}} \sum_{u \in V} \sum_{v > u \in V} d(u, v) \\ &= \frac{2}{n(n-1)} \sum_{u \in V} \sum_{v > u \in V} d(u, v) \\ &= \frac{1}{n(n-1)} \sum_{u \in V} \left[\sum_{v > u \in V} d(u, v) + \sum_{v < u \in V} d(u, v) \right] \\ &= \frac{1}{n(n-1)} \sum_{u \in V} \left[\sum_{v > u \in V} d(u, v) + \sum_{v < u \in V} d(u, v) + \sum_{v = u \in V} d(u, v) \right] \\ &= \frac{1}{n(n-1)} \sum_{u \in V} \sum_{v \in V} d(u, v) \\ &= \text{APL}(V, E) \end{aligned}$$

Aqui assumimos que $d(u, v) = d(v, u) \forall u, v \in V$ e $d(u, u) = 0 \forall u \in V$. Esta última suposição apenas explicita que *loops* não são levados em consideração, deixando assim a definição compatível com o termo $n(n-1)$ que divide o total.

Vamo agora reescrever a métrica APD em função do *closeness* dos vértices:

$$\begin{aligned} \text{APD}(V, E) &= \frac{1}{n(n-1)} \sum_{u \in V} \left[\sum_{v > u \in V} d(u, v) + \sum_{v < u \in V} d(u, v) \right] \\ &= \frac{1}{n(n-1)} \sum_{u \in V} \sum_{v \neq u \in V} d(u, v) \\ &= \frac{1}{n} \sum_{u \in V} c_u \end{aligned}$$

onde c_u é o *closeness* do vértice u . Vale notar que isso diz que APD e APL podem ser entendidos como o *closeness* médio.

2 Se o APL e o diâmetro de uma rede são iguais, então

$$\begin{aligned} \text{APL}(V, E) &= D(V, E) \\ \frac{1}{n(n-1)} \sum_{u \in V} \sum_{v \in V} d(u, v) &= \max_{u, v \in V} d(u, v) \\ \frac{1}{n(n-1)} \sum_{u, v \in V} |d(u, v)| &= \lim_{p \rightarrow \infty} \left[\sum_{u, v \in V} |d(u, v)|^p \right]^{\frac{1}{p}} \end{aligned}$$

$$\therefore \frac{1}{n(n-1)} =$$