

Prova 1 - Exercícios práticos

Redes Complexas para Ciência da Computação

Luben Miguel, Rodrigo Lassance

Índice

Exercício 1	1
Item (a)	2
Item (b)	3
Item (c)	3
Item (d)	3
Exercício 2	4
Item (a)	4
Item (b)	6
Item (c)	9
Exercício 3	11
Exercício 4	11

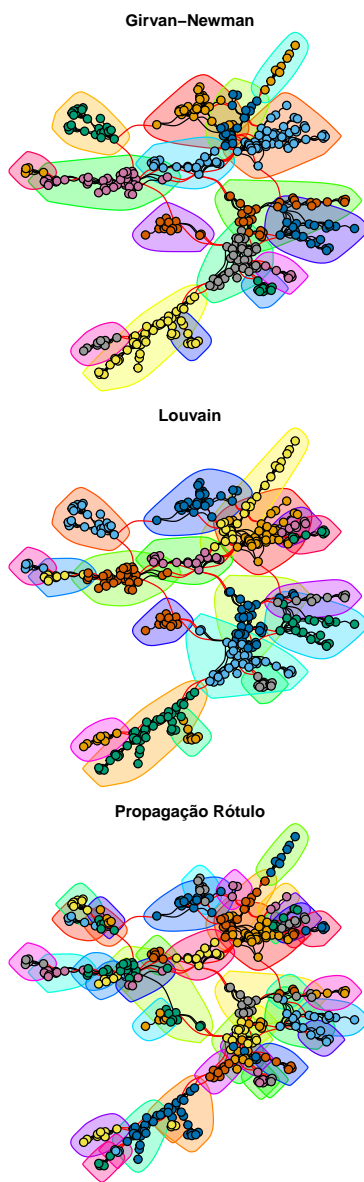
Exercício 1

Para esse exercício utilizamos as redes de conexão entre as fibras do cérebro de ratos [bn-mouse_visual-cortex_2](#), rede de co-autoria de artigos da área de redes complexas [ca-netscience](#) e a rede de emails mandados entre funcionários da empresa de energia Enron que deixou de existir em 2007 [email-enron-only](#).

Selecionada essas redes utilizamos como métodos de detecção de comunidades o método de Louven, Girvan-Newman e propagação de rótulos.

Item (a)

Selecionamos a rede de coautoria para visualização de suas respectivas comunidades estimadas em cada método como pode ser visualizado pela Figura abaixo:



Item (b)

A tabela a seguir sumariza o número de comunidades encontrados para cada rede (nas linhas) por cada método (nas colunas):

Tabela 1: Número de comunidades para cada rede usando cada método

Redes	Girvan-Newman	Louvain	Propagação de Rótulo
Rede do cortex dos ratos	10	9	10
Rede de emails	9	7	4
Rede de co-autoria	17	19	46

Item (c)

A seguir, assim como no item (b), construímos uma tabela com os valores calculados de modularidade de cada método (nas colunas) para cada rede (nas linhas).

Tabela 2: Modularidade para cada rede usando cada método

Redes	Girvan-Newman	Louvain	Propagação de Rótulo
Rede do cortex dos ratos	0.7499236	0.7523037	0.7385689
Rede de emails	0.5345388	0.5663787	0.4123642
Rede de co-autoria	0.8421850	0.8472006	0.7742825

Conforme visto em aula, a modularidade tende a ser maior em redes maiores. Ou seja, não sabemos se a modularidade da rede de co-autoria, com 379 vértices, é maior que a rede de ratos, com 193 vértices, por estar melhor separada ou por simplesmente ser uma rede maior.

Item (d)

Para cada rede utilizada, podemos calcular a similaridade entre as partições encontradas por cada algoritmo através do NMI. Assim, computando o NMI entre os diferentes métodos de detecção de comunidades obtemos os resultados mostrados pela Tabela abaixo:

Nota-se primeiramente que para a rede de emails, o valor do NMI calculado na comparação entre os métodos Girvan-Newman e Louvain versus a propagação de rótulos é razoavelmente baixo. Isso pode se dever ao fato de que no método de propagação de rótulos, o número resultante de comunidades é pequeno comparado aos outros dois métodos. Já, para a rede de co-autoria, esperavamos que o valor do NMI fosse maior nessas mesmas comparações devido à quantidade grande de partições obtidas pela propagação de rótulo, mas não foi isso o observado, vide os valores modestos de NMI sumarizados na tabela.

Tabela 3: NMI associado a cada combinação dois a dois dos métodos para cada rede

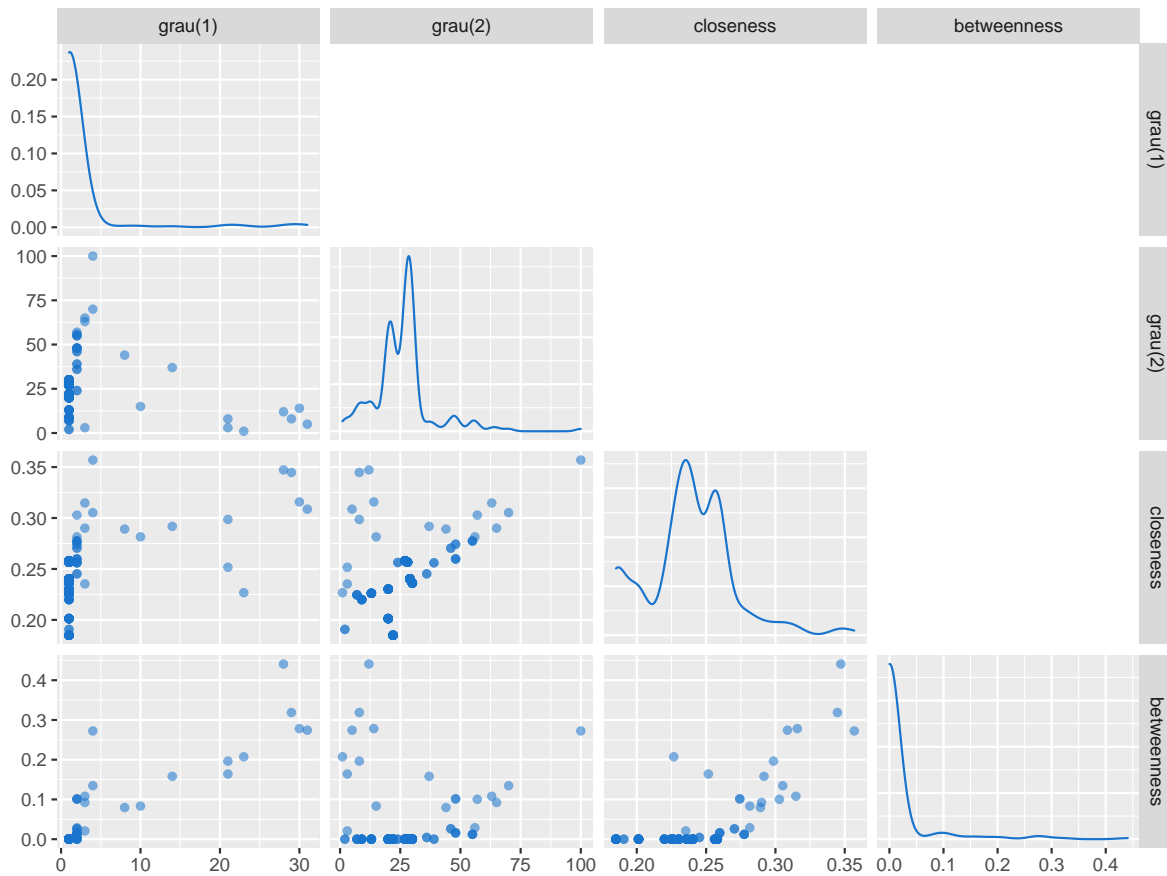
Comparação	Rede de ratos	Rede de emails	Rede de coautoria
Givarn-Newman vs Louvain	0.9488052	0.7754347	0.8683370
Girvan-Newman vs Label Prop	0.8762523	0.4213453	0.6952310
Louvain vs Label Prop	0.8782549	0.4315381	0.7260035

Exercício 2

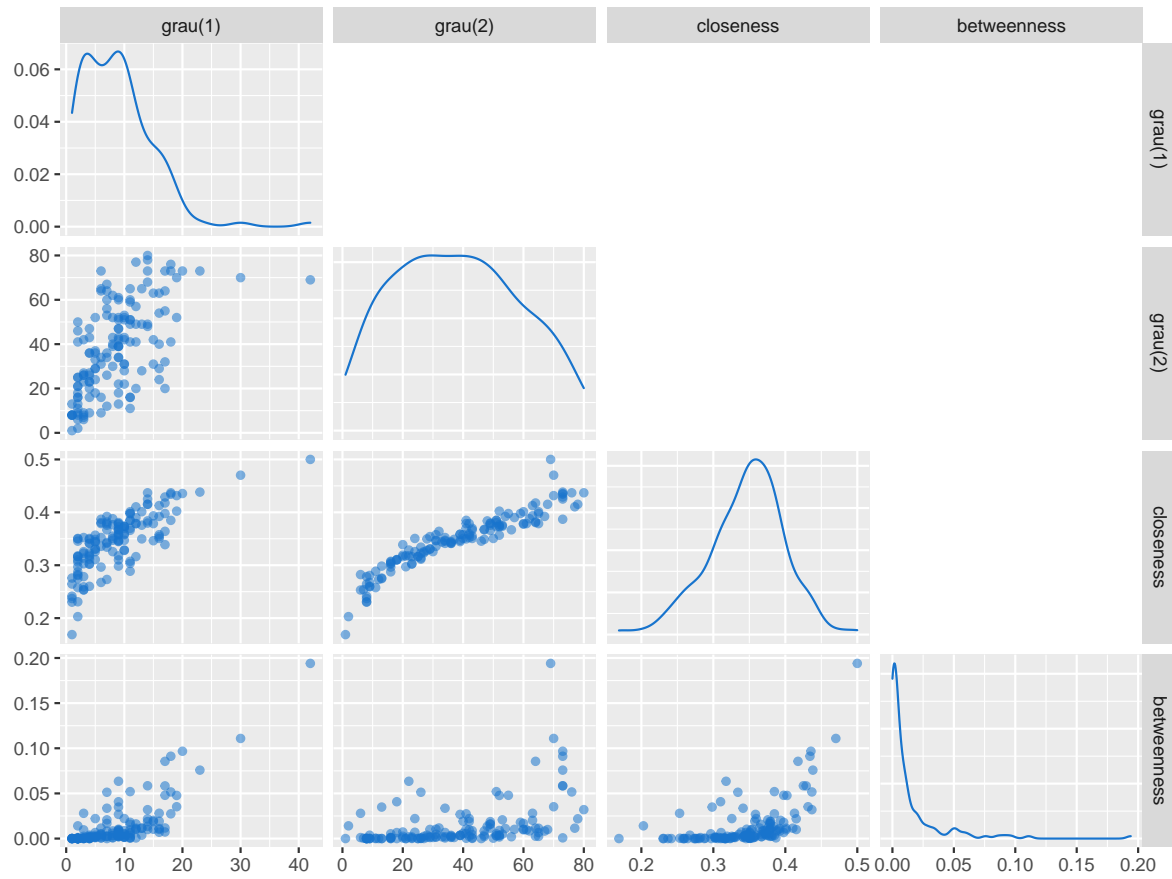
Item (a)

A seguir, podemos visualizar o scatter-plot entre cada par de medida para cada rede a seguir:

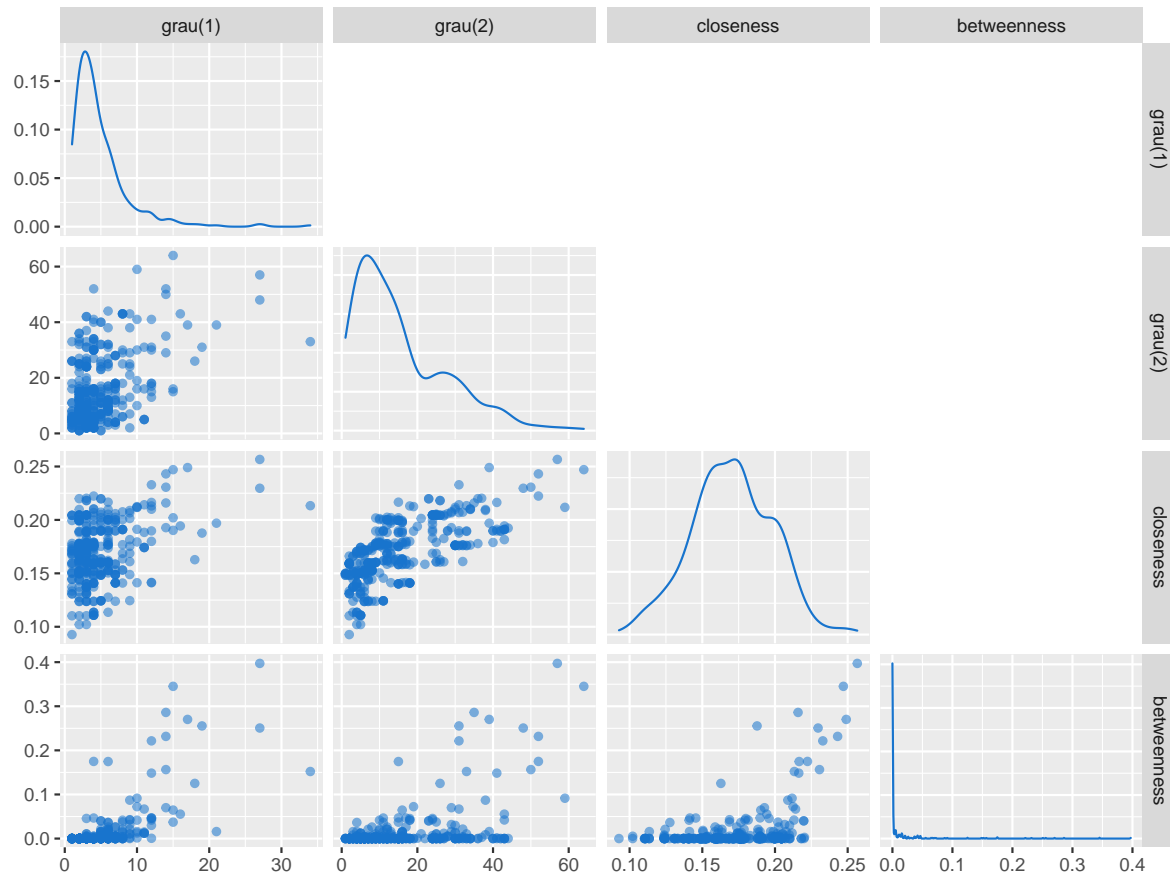
- Rede de conexão de fibras do cortex de ratos



- Rede de emails mandados entre funcionários da empresa Enron

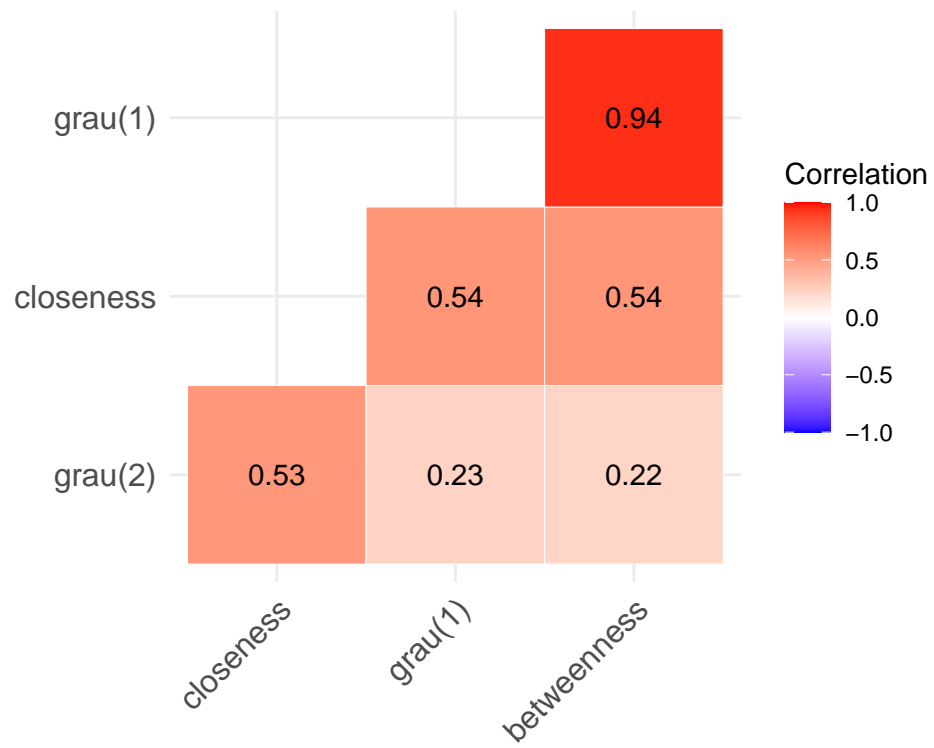


- Rede de co-autoria de artigos

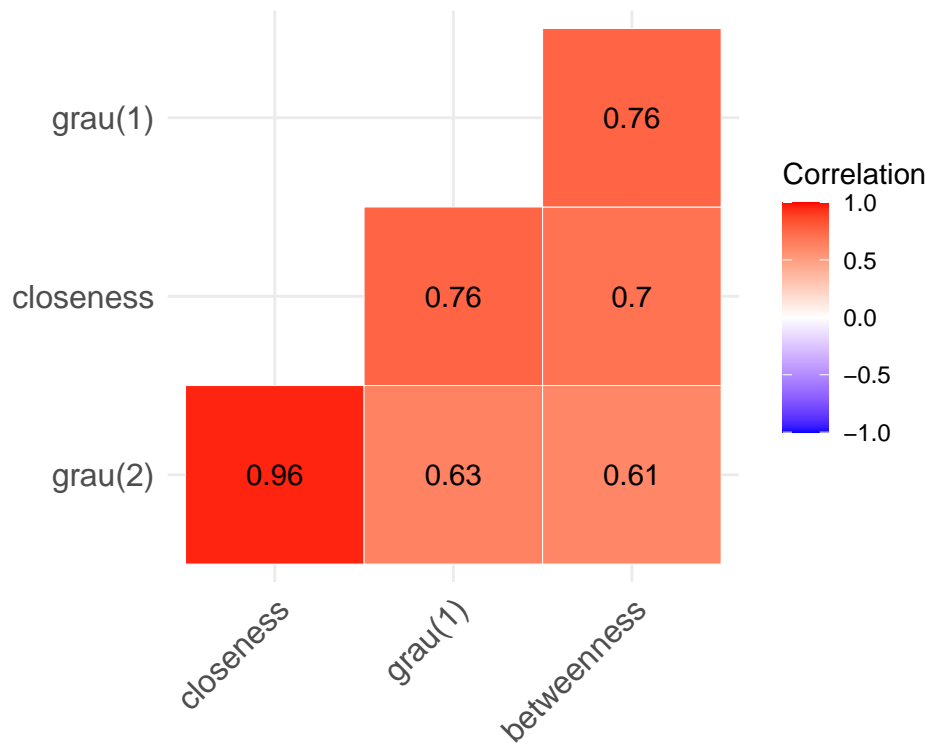


Item (b)

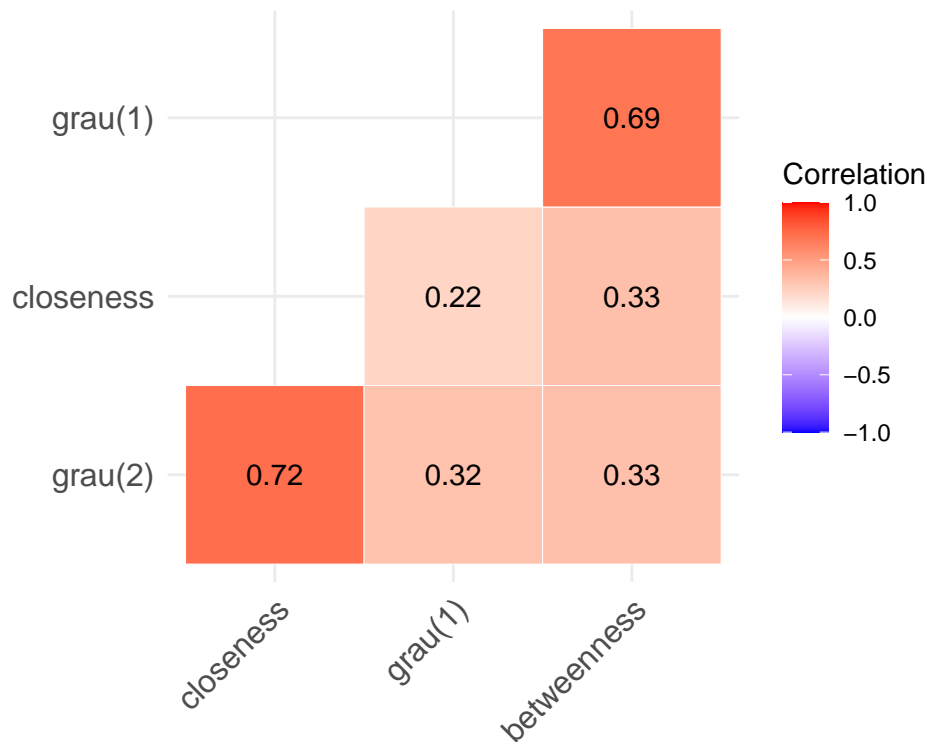
- Rede de conexão de fibras do cortex de ratos



- Rede de emails mandados entre funcionários da empresa Enron



- Rede de co-autoria de artigos



Item (c)

Tomando a rede de emails e as comunidades obtidas pelo método de Louvain, podemos realizar um ranqueamento dos vértices para as maiores comunidades. Primeiramente, a tabela abaixo nos dá o tamanho de cada comunidade:

Tabela 4: Tamanho de cada comunidade

Comunidade	Tamanho
1	29
2	18
3	21
4	26
5	22
6	15
7	12

Visualiza-se que as comunidades 1,3,4 e 5 tem mais que 20 vértices membros. Assim, tomando a comunidade 3, que possui um tamanho grande, podemos tomar o subgrafo correspondente

a essa comunidade e computar as medidas de centralidade locais como dada pela Tabela 5:

Tabela 5: Top 10 vértices com maior centralidade local para o subgrafo da comunidade 3

vertices	grau(1)	grau(2)	closeness	betweenness
75	14	6	0.769	0.277
47	11	6	0.625	0.054
9	10	10	0.667	0.035
13	10	10	0.667	0.033
32	10	10	0.667	0.067
49	10	10	0.667	0.028
56	10	10	0.667	0.040
66	10	7	0.606	0.032
74	10	10	0.667	0.079
137	9	11	0.645	0.072

Pela Tabela 5, destaca-se o vértice 75 como o mais central na comunidade 3, tendo maiores valores de grau, closeness e betweenness. Já, para a centralidade global, podemos observar os top 10 vértices mais centrais de forma geral pela Tabela 6:

Tabela 6: Top 10 vértices com maior centralidade global

vertices	grau(1)	grau(2)	closeness	betweenness
105	42	69	0.500	0.194
17	30	70	0.470	0.111
95	23	73	0.438	0.076
48	20	73	0.436	0.097
43	19	70	0.432	0.035
132	19	52	0.402	0.048
31	18	41	0.385	0.027
74	18	76	0.437	0.052
91	18	73	0.434	0.091
20	17	20	0.339	0.007

Vemos que o vértice 75 nem sequer consta entre os 10 vértices mais globalmente centrais. De fato, outros vértices mais relevantes para a comunidade 3 como o 47, 9, 13 e 32 também não aparecem entre os top 10 vértices mais centrais. Ou seja, pode-se afirmar que existem vértices altamente relevantes dentro da comunidade de interesse sem um grande destaque global.

Exercício 3

Exercício 4