

FGV EMAp

João Pedro Jerônimo

Ciência de Redes

Revisão para A2

Rio de Janeiro

2025

Conteúdo

1	Correlação de Graus	3
1.1	Descobrindo os Critérios	4
1.2	Average Next Neighbour Degree	6
1.3	Cutoffs	8
2	Robustez	10

Correlação de Graus

Nesse capítulo, temos o objetivo de entender como nós se relacionam dentro de uma rede. Imagine uma rede que os nós são pessoas e as arestas são direcionadas e apontam para nós que aquela pessoa conhece. Quando paramos para reparar, nós com muitos graus (Que muitas pessoas conhecem, logo, celebridades), tendem a apontar para outros nós com muitos graus de entrada (Celebridades conhecem muitas celebridades). Porém, muitas vezes esses critérios podem nem estar explícitos na rede em si, e queremos estudar todos os casos possíveis nesse capítulo

1.1 Descobrindo os Critérios

Dado uma rede $G(V, E)$, vamos supor que conhecemos a distribuição dos graus dessa rede ($\mathbb{P}(\delta(v_i) = k) = p(k)$ para simplificação de notação), queremos ver se o grau do nó tem alguma influência em como eles se ligam. Então temos que ver a probabilidade de que um nó de grau k' se ligue com um de grau k , ou seja:

$$\mathbb{P}(\{v_i, v_j\} \in E \mid \delta(v_i) = k, \delta(v_j) = k') = p(\{k, k'\} \in E) \quad (1)$$

Vamos escrever dessa forma para simplificar notação. Vamos ter que:

$$p(\{k, k'\} \in E) = \frac{kk'}{2|E|} \quad (2)$$

Onde $k'/2L$ é a probabilidade do meu nó com grau k se ligar com um nó de grau k' e eu multiplico por k pois meu nó possui k arestas ligadas a ele. Essa probabilidade indica que nós com graus maiores tem mais probabilidade de se ligar com outros nós de grau grande

Definição 1.1.1 (Matriz de Correlação por Grau): A matriz de correlação por grau é a matriz que a entrada e_{ij} é a probabilidade de encontrar nós de graus i e j nas pontas de uma aresta selecionada aleatoriamente

Essa matriz nos tráz informações diretas da relação **linear** entre os nós. Podemos dividir e classificar as redes de 3 formas de acordo com a relação entre seus graus

- **Rede Assortativa:** Nós com graus parecidos intelrigam entre si
- **Redes Neutras:** Os nós não apresentam um padrão para intelrigarem entre si
- **Redes Desassortativas:** Nós com graus maiores se ligam com graus menores e vice-versa

Podemos então avaliar a matriz de covariância ou a correlação entre os dados. Porém a relação entre esses pontos pode não ser linear, o que pode dar falsas impressões sobre a correlação

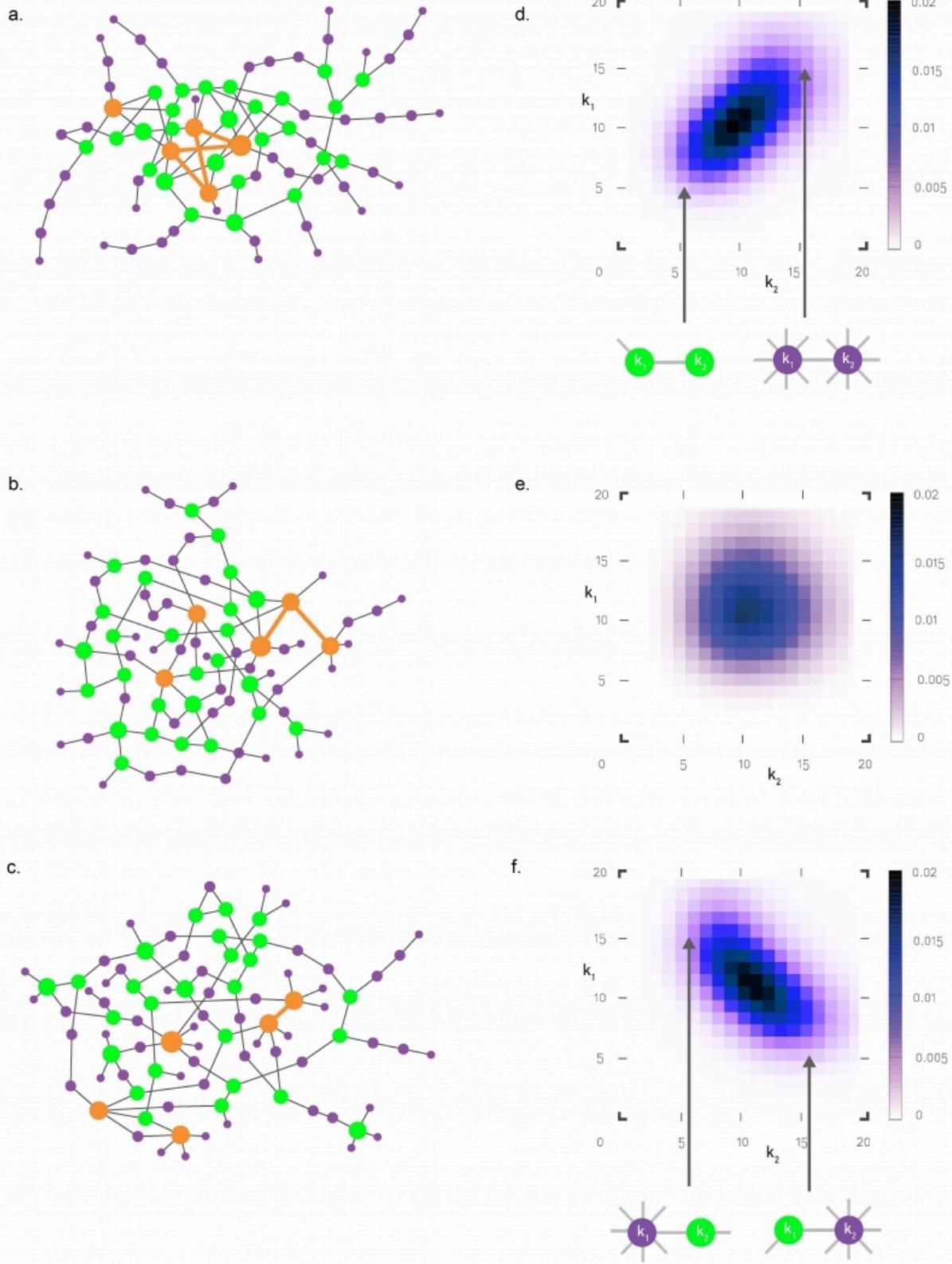


Figura 1: Matriz de covariância de graus com 3 redes. A primeira é assortativa, a segunda é neutra e a última é desassortativa

A matriz de correlação acaba por representar uma distribuição conjunta em i e j de forma que:

$$\sum_{i,j} e_{ij} = 1 \quad (3)$$

então definindo um pouco melhor essa probabilidade, defina K como a variável aleatória que representa o grau de um nó selecionado aleatoriamente. Seja também $\{v_i, v_j\} \in E$ uma aresta escolhida aleatoriamente dentre todas as do conjunto E . Para simplificar notações, também definimos $A = \delta(v_i)$ e $B = \delta(v_j)$. Então vamos ter que:

$$e_{ij} = \mathbb{P}(A = i, B = j) = \mathbb{P}(A = i|B = j) \cdot \mathbb{P}(B = j) = \mathbb{P}(B = j|A = i) \cdot \mathbb{P}(A = i) \quad (4)$$

vamos então calcular cada um desses termos (Até onde for possível). Primeiro vamos calcular $\mathbb{P}(A = i)$

$$\mathbb{P}(A = i) = \frac{iN_i}{2|E|} \times \frac{\frac{1}{N}}{\frac{1}{N}} = \frac{ip(i)}{\mathbb{E}[K]} \quad (5)$$

onde N_i é a quantidade de nós com grau i e eu multiplico por i pois eu posso escolher qualquer uma das i arestas presentes no nó que eu gostaria que fosse escolhido (O de grau i). E divido isso pela quantidade total de nós. Eu posso ainda, dividir o termo de cima e o de baixo por N , obtendo então a expressão na direita

Se os eventos A e B **não são independentes**, seria necessário saber a distribuição entre eles dois para calcular $\mathbb{P}(A = i|B = j)$. Porém, se A e B são independentes, então $\mathbb{P}(A = i|B = j) = \mathbb{P}(A = i) \cdot \mathbb{P}(B = j)$, então, no caso de A e B independentes, obtemos:

$$\mathbb{P}(A = i|B = j) = \mathbb{P}(A = i) \cdot \mathbb{P}(B = j) = \frac{ip(i)jp(j)}{(\mathbb{E}[K])^2} \quad (6)$$

1.2 Average Next Neighbour Degree

É interessante fazer o gráfico de uma matriz de covariância, porém, por ela perder essas informações antes mencionadas, as vezes ela pode ser ineficiente ou ser de difícil leitura se houver muitos vértices, então criamos a seguinte definição:

Definição 1.2.1 (Average Next Neighbour Degree): Seja $G(V, E)$ uma rede, A a matriz de adjacência da mesma e $v_i \in V$, então:

$$k_{nn}(v_i) := \frac{1}{\delta(v_i)} \sum_{j=1}^{|V|} A_{ij} \delta(v_j) \quad (7)$$

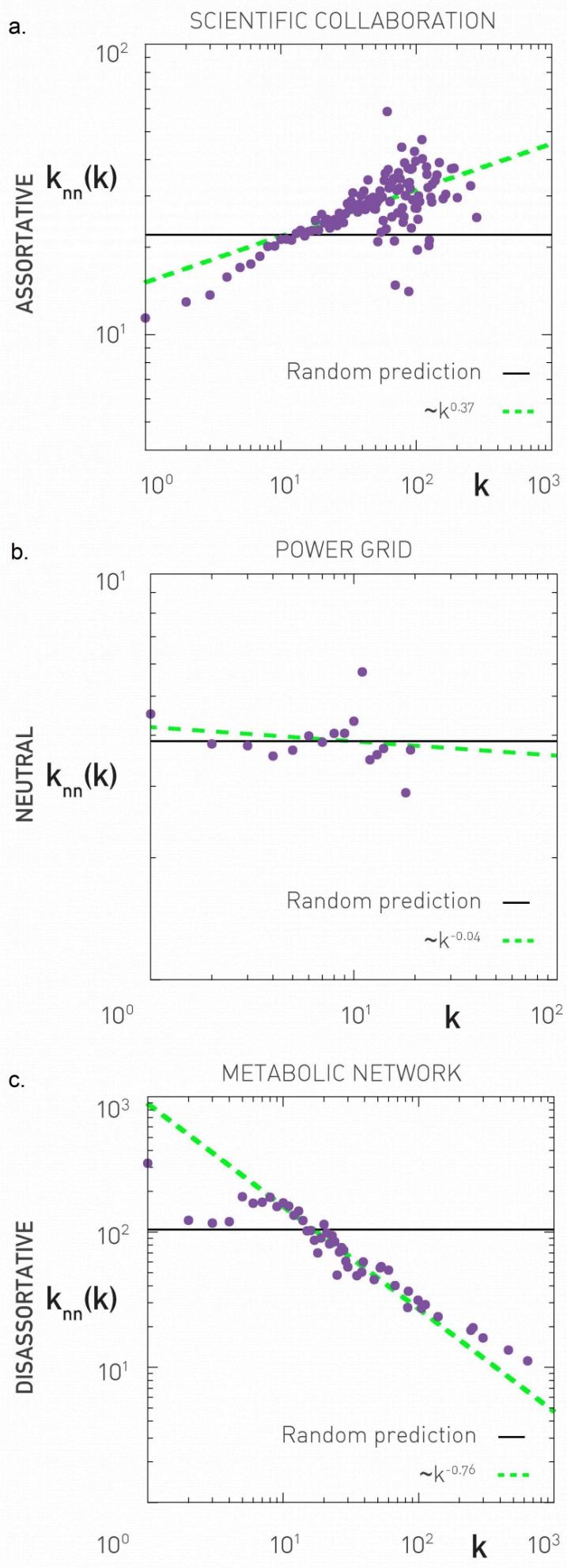
Essa definição é o k_{nn} para um vértice específico, mas e se quisermos o valor médio dessa medida para nós de grau k ?

$$k_{nn}(k) = \mathbb{E}[B|A = k] = \sum_{k'} k' \mathbb{P}(B = k'|A = k) \quad (8)$$

Podemos ver o que acontece com essa medida quando A e B são independentes, logo, analisar o caso das **redes neutras**

$$k_{nn} = \sum_{k'} k' \mathbb{P}(B = k') = \sum_{k'} k' \frac{k' p(k')}{\mathbb{E}[K]} = \frac{\mathbb{E}[K^2]}{\mathbb{E}[K]} \quad (9)$$

Perceba que, quando consideramos o caso das redes neutras, **a média dos vizinhos de um nó não depende em quantos nós ele tem ou de informações sobre como o nó se relaciona com os outros, mas única e exclusivamente de características globais da rede**. A mesma lógica vale para as redes assortativas e desassortativas:



- **Assortativas:** A medida tende a crescer conforme aumentamos os valores de k já que quanto maior o grau, maior a probabilidade de os graus se interligarem
- **Desassortativas:** Nós de grau menor tendem a interligar com nós de grau maior, o que acaba por fazer com que as probabilidades sejam altas em valores de k muito pequenos, isso faz com que conforme aumentamos o k , o valor da medida diminua

O livro aponta que, de acordo com o gráfico mostrado anteriormente, que podemos aproximar a medida do k_{nn} como:

$$k_{nn}(k) = \beta k^\mu \quad (10)$$

de forma que:

- $\mu > 0 \Rightarrow$ Rede assortativa
- $\mu \approx 0 \Rightarrow$ Rede neutra
- $\mu < 0 \Rightarrow$ Rede desassortativa

1.3 Cutoffs

A gente até agora ta assumindo que as redes analisadas são simples, ou seja, entre dois nós, só pode existir **no máximo** uma aresta. Porém, em algumas redes, quando calculamos o número esperado de arestas entre nós

$$\mathbb{E}_{kk'} = e_{kk'} \mathbb{E}[K]N \quad (11)$$

esse valor é **maior que 1**. Então como fazemos para achar o valor de k que faz com que esse valor seja > 1 ?

Primeiro, vamos definir

$$r_{kk'} = \frac{E_{kk'}}{m_{kk'}} \quad (12)$$

onde $E_{kk'}$ é o número de arestas entre todos os nós de grau k e de k' e $m_{kk'} = \min\{kN_k, k'N_{k'}, N_k N_{k'}\}$ é a maior quantidade **possível** de restas que ligam nós de grau k e k' . Vou explicar o porquê dessa fórmula para $m_{kk'}$

- Cada nó com grau k só pode se ligar a **no máximo** k nós de grau k' , então:

$$E_{kk'} \leq kN_k \quad (13)$$

- O mesmo argumento vale para os nós de grau k'

$$E_{kk'} \leq k'N_{k'} \quad (14)$$

- Eu também não consigo ligar mais que o número máximo disponível de pares entre os dois grupos de graus

$$E_{kk'} \leq N_k N_{k'} \quad (15)$$

Ou seja, temos que $r_{kk'} \leq 1 \forall k, k'$. Só que como mencionei anteriormente, por conta do valor esperado “não-esperado” de mais que uma aresta entre dois nós, pode ocorrer que $r_{kk'} > 1$, então queremos encontrar k_s tal que:

$$r_{k_s k_s} = 1 \quad (16)$$

Mas antes de achar k_s , vamos reescrever $r_{kk'}$ de outra forma e fazer uma análise. Perceba que, quando $k > Np(k')$ ou $k' > Np(k)$, os efeitos das arestas múltiplas aparecem, transformando a expressão em:

$$r_{kk'} = \frac{\mathbb{E}_{kk'}}{m_{kk'}} = \frac{\mathbb{E}[K]e_{kk'}}{Np(k)p(k')} \quad (17)$$

Para redes livre-de-escala, as condições são satisfeitas para $k, k' > (a |V|)^{\frac{1}{\gamma+1}}$ onde a depende de $p(k)$. Só que esse valor é abaixo do cutoff natural, de forma que temos CERTEZA de que, se k e k' são maiores que isso, não podemos garantir que o valor esperado dos links será menor que 1. Olhando o caso específico das redes neutras, sabemos que:

$$e_{kk'} = \frac{kk'p(k)p(k')}{\mathbb{E}[K]^2} \quad (18)$$

Logo $r_{kk'}$ vira:

$$r_{kk'} = \frac{kk'}{\mathbb{E}[K]N} \quad (19)$$

Ou seja, chegamos que $k_s \propto (\mathbb{E}[K]N)^{\frac{1}{2}}$

Tendo isso em mente, conseguimos dividir as **redes livre-de-escala** em dois regimes ao compararmos k_{\max} e k_s (Lembrando: $k_{\max} \propto N^{\frac{1}{\gamma-1}}$)

- **Sem cutoff estrutural:** Para redes aleatórias e livre-de-escala com $\gamma \geq 3$, o expoente de $k_{\max} \leq \frac{1}{2} \Rightarrow k_{\max} \leq k_s$, logo os valores esperados dos links entre nós será **sempre** menor ou igual a 1
- **Dissassortividade estrutural:** Para redes livre-de-escala com $\gamma < 3$, temos $-\frac{1}{1-\gamma} > \frac{1}{2}$, logo k_s pode ser menor que k_{\max} , logo, há nós entre k_s e k_{\max} que podem ter $E_{kk'} > 1$

Robustez

