

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE
Corso di Laurea in Matematica

Reti complesse e centralità dei nodi

Tesi di Laurea in Matematica Computazionale

Relatore:
Chiar.mo Prof.
Valeria Simoncini

Presentata da:
Leonardo Lambertini

Anno Accademico 2022/2023

A tutte le persone a me care.

Introduzione

Il concetto di rete è abbastanza intuitivo per tutti nella società moderna. Non appena vediamo questa parola ci viene alla mente un gruppo di oggetti interconnessi. Come riporta Estrada nel suo libro [Est12], la parola “rete” è apparsa per la prima volta nel 1560, nella Bibbia di Ginevra, “E farai per esso una grata come una rete di ottone.” In questo caso si riferisce ad una disposizione simile ad una rete di fili. In questi secoli, tuttavia, è stata utilizzata per descrivere sempre più concetti. Attualmente, l’OED (Oxford English Dictionary) definisce una rete come una “disposizione di linee orizzontali e verticali che si intersecano” o come “un gruppo o sistema di persone o cose interconnesse”. Il motivo principale per il quale le reti vengono studiate consiste nel fatto che le reti complesse rappresentano lo scheletro dei sistemi complessi. I sistemi complessi sono sistemi costituiti da diverse parti interconnesse che presentano proprietà non evidenti dalle proprietà delle singole parti. Uno dei concetti più intuitivi che possiamo pensare quando affrontiamo l’analisi di una rete è l’identificazione dei nodi “più importanti”. La quantificazione del concetto di importanza degli attori nelle reti sociali è stata una delle prime applicazioni della teoria delle reti, risalente ai primi anni Cinquanta. Questo concetto è formalizzato nella teoria delle reti mediante le misure di centralità. In questa tesi ci concentreremo sull’analisi di alcune misure di centralità, studiando ed analizzando le differenze fra esse tramite esempi pratici su vari dati molto diversi fra loro. Nell’ultimo capitolo introdurremo un’analisi più approfondita della misura di centralità dell’eigenvector centrality, analizzata nel secondo capitolo, prendendo in considerazione anche le interazioni del secondo ordine, ossia attribuendo l’importanza ad un nodo non solo osservando i cammini di lunghezza uno, ma arrivando anche ad analizzare i cammini di lunghezza due.

Indice

Introduzione	i
1 Concetti preliminari	1
1.1 Funzioni di matrici	1
1.1.1 Forma canonica di Jordan	3
1.2 SVD e Pseudoinversa	5
1.3 Metodo delle potenze	7
2 Reti complesse	9
2.1 Misure di centralità	11
2.1.1 Grado	11
2.1.2 Centralità di Katz	14
2.1.3 Eigenvector centrality	15
2.1.4 Subgraph centrality	17
2.1.5 PageRank	18
2.2 Betweenness centrality	21
3 Eigenvector centrality del secondo ordine	23
3.1 Coefficiente di clustering di Watts-Strogatz	23
3.2 Modello generale dell'eigenvector centrality del secondo ordine	25
3.2.1 Scelte di M e di T	25
3.2.2 Scelte di α e di p	28
3.2.3 Esistenza e unicità	31
3.2.4 Alcuni esempi e considerazioni associate	33
Bibliografia	39

Capitolo 1

Concetti preliminari

In questa tesi useremo le seguenti notazioni:

- $\|x\|_2 = (\sum_{i=1}^n \|x_i\|^2)^{1/2}$, $x \in \mathbb{C}^n$ per indicare la norma Euclidea vettoriale;
- $\|A\|_F = (\sum_{i,j} \|a_{i,j}\|^2)^{1/2}$, $A \in \mathbb{C}^{n \times m}$ per indicare la norma di Frobenius matriciale;
- $\mathbf{1}$ per indicare un vettore composto da soli 1;
- $A > 0$ per indicare una matrice in cui tutti gli elementi sono positivi, e $A \geq 0$ una matrice in cui tutti gli elementi sono non negativi.
- \otimes per indicare il prodotto di Kronecker, ossia per $A \in \mathbb{R}^{n_A \times m_A}$ e $B \in \mathbb{R}^{n_B \times m_B}$

$$A \otimes B = \begin{bmatrix} a_{1,1}B & a_{1,2}B & \dots & a_{1,m_A}B \\ a_{2,1}B & a_{2,2}B & \dots & a_{2,m_A}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{n_A,1}B & a_{n_A,2}B & \dots & a_{n_A,m_A}B \end{bmatrix} \in \mathbb{R}^{n_A n_B \times m_A m_B}.$$

- \circ per indicare il prodotto di Hadamard, ossia $(A \circ B)_{ij} = (A)_{ij}(B)_{ij}$ con $A, B \in \mathbb{R}^{n \times m}$.

1.1 Funzioni di matrici

Il termine "funzione di matrice" può avere diversi significati, qui siamo interessati ad una definizione che prende una funzione scalare f , una matrice $A \in \mathbb{C}^{n \times n}$ e restituisce una funzione $f(A)$ delle stesse dimensioni di A . Prima di dare una definizione formale però, facciamo delle osservazioni. Quando $f(t)$ è una funzione polinomiale o razionale

con coefficienti e argomento t scalari, è naturale definire $f(A)$ sostituendo t con A , rimpiazzando la divisione con l'inversa della matrice (purchè le matrici siano invertibili), e scambiando i vari 1 con la matrice identità. Quindi, per esempio

$$f(t) = \frac{1+t^2}{1-t} \Rightarrow f(A) = (I-A)^{-1}(I+A^2) \text{ se } 1 \notin \Lambda(A),$$

dove, $\Lambda(A)$ rappresenta l'insieme degli autovalori di A (lo spettro di A). Notiamo che le funzioni razionali di una stessa matrice commutano, quindi non importa se scriviamo $(I-A)^{-1}(I+A^2) \circ (I+A^2)(I-A)^{-1}$. Se f ha una rappresentazione in serie di potenze convergente, come

$$\log(t+1) = t - \frac{t^2}{2} + \frac{t^3}{3} - \frac{t^4}{4} + \dots, \text{ se } |t| < 1,$$

possiamo semplicemente sostituire t con A

$$\log(I+A) = A - \frac{A^2}{2} + \frac{A^3}{3} - \frac{A^4}{4} + \dots, \text{ se } \rho(A) < 1.$$

Dove $\rho(A)$ denota il raggio spettrale di A , e la condizione $\rho(A) < 1$ assicura la convergenza della serie di matrici, infatti il teorema seguente dice proprio questo.

Teorema 1.1 (Convergenza serie di Taylor per matrici). *Sia f una funzione sviluppabile in serie di Taylor*

$$f(z) = \sum_{k=0}^{\infty} a_k (z - \alpha)^k \quad \left(a_k = \frac{f^{(k)}(\alpha)}{k!} \right)$$

con raggio di convergenza r . Se $A \in \mathbb{C}^{n \times n}$ allora è definita $f(A)$ come

$$f(A) = \sum_{k=0}^{\infty} a_k (A - \alpha I)^k$$

se e solo se ogni autovalore distinto $\lambda_1, \dots, \lambda_s$ di A soddisfa una delle due condizioni

i) $|\lambda_i - \alpha| < r,$

ii) $|\lambda_i - \alpha| = r$ e la serie per $f^{(n_i-1)}(\lambda)$ (dove n_i è l'indice di λ_i , Definizione 1.3) converge a $\lambda = \lambda_i$, $i = 1, \dots, s$.

La dimostrazione del Teorema 1.1 si trova in [Hig08].

Un esempio molto utilizzato è l'esponenziale di una matrice, che si definisce come

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}. \quad (1.1)$$

Alcune proprietà importanti dell'esponenziale sono

- $e^{A+B} = e^A e^B$ se e solo se $AB = BA$.
- $e^A e^B = e^B e^A$ se e solo se $AB = BA$.
- $e^A = e^{(A/\sigma)\sigma}$, per ogni $\sigma \in \mathbb{C}$.

Quest'ultima proprietà è molto utile per poter approssimare la funzione esponenziale, infatti usando l'approssimazione di Padè, otteniamo

$$e^A \approx \frac{p_k(A)}{q_m(A)}, \text{ con } p_k, q_m \text{ polinomi di grado } k \text{ ed } m.$$

Questa approssimazione è computazionalmente non efficiente a meno che $\|A\| < 1$, ed è in questo contesto che facciamo uso della terza proprietà. Scegliamo $\sigma = 2^s$ tale che $\|A/\sigma\| \approx 1$, successivamente $e^{A/\sigma} \approx r(A/\sigma)$ con r funzione razionale di Padè, ed infine facciamo $(r(A/\sigma))^\sigma \approx e^A$.

Utilizzando la definizione descritta sopra si possono definire molte funzioni di matrici, tale approccio tuttavia potrebbe essere sconveniente:

- Non ci fornisce la definizione di $f(A)$ per f funzione arbitraria.
- Una particolare formula potrebbe essere applicata solo ad un insieme ristretto di matrici, e se definiamo $f(A)$ con tale formula (piuttosto che usare una definizione generale), allora dobbiamo verificare che sia in linea con le altre definizioni di tale funzione.
- Per una funzione polidroma, come potrebbe essere la funzione logaritmica o la radice quadrata, è necessario classificare tutte le possibili $f(A)$ che si potrebbero ottenere utilizzando le differenti diramazioni della funzione e identificare ogni possibile valore.

Per questi motivi vedremo delle definizioni generali di funzioni di matrici.

1.1.1 Forma canonica di Jordan

Teorema 1.2 (Forma canonica di Jordan). *Ogni matrice quadrata è simile ad una matrice di Jordan, ovvero $A \in \mathbb{C}^{n \times n}$ può essere scritta come*

$$A = ZJZ^{-1} \Rightarrow Z^{-1}AZ = J = \text{diag}(J_1, J_2, \dots, J_p),$$

$$J_k = J_k(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_k & 1 \end{bmatrix} \in \mathbb{C}^{m_k \times m_k},$$

dove Z è una matrice non singolare, $\lambda_1, \dots, \lambda_s$ gli autovalori distinti di A e m_k la dimensione del blocco di Jordan J_k , con $m_1 + m_2 + \dots + m_p = n$.

Definizione 1.3. L'indice di λ_i è l'ordine del blocco di Jordan più grande nel quale λ_i compare.

Definizione 1.4. Posto n_i l'indice di λ_i , la funzione f si dice definita sullo spettro di A , se i valori

$$f^{(j)}(\lambda_i), \quad j = 1, \dots, n_i - 1, \quad i = 1, \dots, s$$

esistono. Questi sono chiamati i valori della funzione f sullo spettro di A .

In molti casi di particolare interesse f è data da una formula, ad esempio $f(t) = e^t$. Tuttavia, la definizione che daremo di $f(A)$ richiederà solo i valori di f sullo spettro di A , e non servirà nessun'altra informazione riguardante f .

Definizione 1.5 (Funzione di matrice tramite la forma canonica di Jordan). Sia f definita sullo spettro di $A \in \mathbb{C}^{n \times n}$, una matrice che possiede la forma canonica di Jordan, allora

$$f(A) := Z f(J) Z^{-1} = Z \operatorname{diag}(f(J_k)) Z^{-1}, \quad (1.2)$$

dove

$$f(J_k) = \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \dots & \frac{f^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \\ & f(\lambda_k) & \ddots & \vdots \\ & & \ddots & f'(\lambda_k) \\ & & & f(\lambda_k) \end{bmatrix}. \quad (1.3)$$

La Definizione 1.5 richiede che la funzione e le sue derivate siano ben definite sui valori dello spettro di A . Vediamo un piccolo esempio per capire meglio la definizione appena data. Per il blocco di Jordan $J = \begin{bmatrix} 1/2 & 1 \\ 0 & 1/2 \end{bmatrix}$ e $f(x) = x^3$, (1.3) fornisce

$$f(J) = \begin{bmatrix} f(1/2) & f'(1/2) \\ 0 & f(1/2) \end{bmatrix} = \begin{bmatrix} 1/8 & 3/4 \\ 0 & 1/8 \end{bmatrix},$$

e si può facilmente verificare che coincide con J^3 . Facciamo ora delle osservazioni sulla definizione appena data.

Osservazione 1.6. La Definizione 1.5 fornisce una $f(A)$ che è indipendente dalla particolare forma canonica di Jordan utilizzata.

Dimostrazione. Consideriamo due forme canoniche di Jordan

$$A = ZJZ^{-1} = WJW^{-1} \quad (1.4)$$

(possiamo assumere senza perdere di generalità che la matrice J è la stessa in entrambi i casi). La Definizione 1.5 fornisce $f_1(A) = Zf(J)Z^{-1}$, $f_2(A) = Wf(J)W^{-1}$ e dobbiamo mostrare che $f_1(A) = f_2(A)$, ovvero $W^{-1}Zf(J)Z^{-1}W = f(J)$, o $X^{-1}f(J)X = f(J)$ con $Z^{-1}W = X$. Ora da (1.4) abbiamo $X^{-1}JX = J$, che implica $f(J) = f(X^{-1}JX) = X^{-1}f(J)X$. Abbiamo quindi ottenuto che $f_1(J) = f_2(J)$. \square

Notiamo poi che se A è diagonalizzabile, allora la forma canonica di Jordan si riduce a $A = ZDZ^{-1}$ con $D = \text{diag}(\lambda_i)$, e sulle colonne di Z ci sono gli autovettori di A , quindi la definizione sopra diventa $f(A) = Zf(D)Z^{-1} = Z\text{diag}(f(\lambda_i))Z^{-1}$. Quindi se A è diagonalizzabile notiamo che $f(A)$ ha gli stessi autovettori di A , mentre gli autovalori sono ottenuti applicando f a quelli di A .

Facciamo vedere ora che (1.3) può essere ottenuta da considerazioni sulle serie di Taylor. In generale J_k si può riscrivere come $J_k = \lambda_k I + N_k \in \mathbb{C}^{m_k \times m_k}$, dove N_k è una matrice di tutti zero, tranne sulla diagonale superiore, nella quale ci sono solo uno. Se per esempio $m_k = 3$, abbiamo che

$$N_k = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad N_k^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad N_k^3 = 0.$$

In generale elevare alla potenza N_k , equivale a spostare la diagonale superiore il numero di volte per il quale stiamo elevando alla potenza fino a che non diventa zero, quindi abbiamo che N_k è nilpotente. Assumendo che f è sviluppabile in serie di Taylor convergente

$$f(t) = f(\lambda_k) + f'(\lambda_k)(t - \lambda_k) + \frac{f''(\lambda_k)(t - \lambda_k)^2}{2} + \dots + \frac{f^{(j)}(\lambda_k)(t - \lambda_k)^j}{j!} + \dots$$

Sostituendo $J_k \in \mathbb{C}^{m_k \times m_k}$ al posto di t , si ottiene

$$f(J_k) = f(\lambda_k)I + f'(\lambda_k)N_k + \dots + \frac{f^{(m_k-1)}(\lambda_k)N_k^{m_k-1}}{(m_k-1)!}, \quad (1.5)$$

poichè tutte le potenze di N_k dalla m_k -esima in poi fanno zero. Questa espressione è facilmente riconducibile a (1.3).

1.2 SVD e Pseudoinversa

Teorema 1.7 (Decomposizione in valori singolari). *Sia $A \in \mathbb{C}^{m \times n}$ e $q = \min\{n, m\}$. Allora esiste una matrice $\Sigma = (\sigma_1, \sigma_2, \dots, \sigma_q, 0, \dots, 0) \in \mathbb{R}^{m \times n}$ con $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_q$ e due matrici unitarie $U \in \mathbb{C}^{m \times m}$, $V \in \mathbb{C}^{n \times n}$ tali che $A = U\Sigma V^*$.*

Dimostrazione. Per fissare le idee, A è $m \times n$ con $m \geq n$ (non è restrittivo), e vogliamo

$$A = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^*.$$

Sia $x = \arg \max_{\|x\|_2=1} \|Ax\|_2$, e sia $Ax = \sigma_1 y$ con $\|y\|_2 = 1$, dove $\sigma_1 = \|A\|_2$. Definiamo quindi $X_1 = [x, \tilde{X}_2] \in \mathbb{R}^{n \times n}$ e $Y_1 = [y, \tilde{Y}_2] \in \mathbb{R}^{m \times m}$ in modo che siano entrambe unitarie. Quindi

$$A_1 := Y_1^T A X_1 = \begin{bmatrix} \sigma_1 & d^T \\ 0 & B \end{bmatrix}$$

(dato che $y^T Ax = \sigma_1$ e $\tilde{Y}_2^T Ax = \sigma_1 \tilde{Y}_2^T y = 0$). Inoltre, $d = 0$, infatti, se definiamo $z = \begin{bmatrix} \sigma_1 \\ d \end{bmatrix}$:

$$\begin{aligned} \sigma_1^2 &\geq \frac{\|A_1 z\|_2^2}{\|z\|_2^2} = \frac{\left\| \begin{bmatrix} \sigma_1 & d^T \\ 0 & B \end{bmatrix} \begin{bmatrix} \sigma_1 \\ d \end{bmatrix} \right\|_2^2}{\sigma_1^2 + d^T d} = \frac{\left\| \begin{bmatrix} \sigma_1^2 + d^T d \\ Bd \end{bmatrix} \right\|_2^2}{\sigma_1^2 + d^T d} \\ &= \frac{(\sigma_1^2 + d^T d)^2 + \|Bd\|_2^2}{\sigma_1^2 + d^T d} \geq \sigma_1^2 + d^T d. \end{aligned}$$

Queste disuguaglianze sono vere se e solo se $d = 0$. Quindi sia la prima riga che la prima colonna di A_1 sono zero, eccetto l'elemento diagonale. La procedura prosegue in modo iterativo con B , e $U = Y_1, \dots, Y_{n-1}$, $V = X_1, \dots, X_{m-1}$. \square

Ricordiamo che la decomposizione SVD fornisce una soluzione del problema dei minimi quadrati, infatti fornisce un x_{LS} e la norma del residuo minimo $\rho_{LS} = \|Ax_{LS} - b\|_2$.

Teorema 1.8. Sia $A \in \mathbb{R}^{m \times n}$, $r = \text{rank}(A)$ e $U\Sigma V^T$ la sua decomposizione tramite SVD. Siano $U = [u_1, \dots, u_m]$ e $V = [v_1, \dots, v_n]$, dove u_i e v_j rappresentano le colonne rispettivamente di U e V , e sia $b \in \mathbb{R}^m$, allora

$$x_{LS} = \sum_{i=1}^r \frac{u_i^T b}{\sigma_i} v_i \quad (1.6)$$

minimizza $\|Ax - b\|_2$. Inoltre

$$\rho_{LS}^2 = \|Ax_{LS} - b\|_2^2 = \sum_{i=r+1}^m (u_i^T b)^2. \quad (1.7)$$

Dimostrazione. Per ogni $x \in \mathbb{R}^n$, si ha

$$\begin{aligned} \|Ax - b\|_2^2 &= \|(U^T A V)(V^T x) - U^T b\|_2^2 = \|\Sigma \alpha - U^T b\|_2^2 \\ &= \sum_{i=1}^r (\sigma_i \alpha_i - u_i^T b)^2 + \sum_{i=r+1}^m (u_i^T b)^2 \end{aligned} \quad (1.8)$$

dove $\alpha = V^T x$. Chiaramente se x risolve il problema dei minimi quadrati, allora $\alpha_i = (u_i^T b / \sigma_i)$ per $i = 1, \dots, r$. Se poniamo $\alpha_i = 0$ per $i = r + 1, \dots, n$ allora x minimizza la norma. \square

Definiamo ora la pseudoinversa. Sia $A^\dagger \in \mathbb{R}^{n \times m}$ definita come $A^\dagger = V \Sigma^\dagger U^T$, con

$$\Sigma^\dagger = \text{diag} \left(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}, 0, \dots, 0 \right) \in \mathbb{R}^{n \times m}, \quad r = \text{rank}(A)$$

A^\dagger si chiama pseudoinversa di A , ed è l'unica soluzione del problema

$$\min_{X \in \mathbb{R}^{n \times m}} \|AX - I_m\|_F.$$

Osserviamo che se il rango di A fosse n , allora $A^\dagger = (A^T A)^{-1} A^T$, mentre se $m = n = \text{rank}(A)$, allora $A^\dagger = A^{-1}$. Normalmente, A^\dagger è definita per essere l'unica matrice $X \in \mathbb{R}^{n \times m}$ che verifichi le condizioni di Moore-Penrose;

- i) $AXA = A$
- ii) $XAX = X$
- iii) $(AX)^T = AX$
- iv) $(XA)^T = XA$

Notiamo che dall'equazione (1.6) e dalla definizione di pseudoinversa possiamo scrivere $x_{LS} = A^\dagger b$ e $\rho_{LS} = \|(I - AA^\dagger)b\|$.

1.3 Metodo delle potenze

Il metodo delle potenze è un metodo iterativo, particolarmente adatto per il calcolo dell'autovalore dominante di una matrice quadrata, ovvero quello di modulo massimo, λ_1 , ed il relativo autovettore associato. Sia quindi $A \in \mathbb{R}^{n \times n}$ e, fissato un vettore iniziale $x_0 \in \mathbb{R}^n$ di norma unitaria, l'iterazione del metodo delle potenze consiste in

Per $i = 0, 1, 2, \dots$, fino a convergenza,

$$\begin{aligned} y &= Ax^{(i)} \\ x^{(i+1)} &= \frac{y}{\|y\|} \\ \lambda^{(i+1)} &= x^{(i+1)T} Ax^{(i+1)} \end{aligned}$$

End

dove $\lambda^{(i+1)}$ è calcolato applicando il quoziente di Rayleigh in cui non è presente il denominatore, in quanto $x^{(i+1)}$ ha norma unitaria. Indicata con (λ_1, x_1) l'autocoppia dominante di A , il metodo delle potenze costruisce due successioni

$$\{x^{(k)}\}_{k \in \mathbb{N}} \rightarrow x_1, \quad \{\lambda^{(k)}\}_{k \in \mathbb{N}} \rightarrow \lambda_1, \quad \text{per } k \rightarrow \infty.$$

Affinchè le due successioni convergano alle quantità desiderate, λ_1 deve essere un autovalore semplice in modulo. D'altra parte, se λ_1 è autovalore con molteplicità algebrica (e geometrica) maggiore di 1, e non ci sono altri autovalori con lo stesso modulo ma distinti da λ_1 , allora la convergenza dell'autovalore si ha ancora, e l'iterazione converge a qualsiasi autovettore dell'autospazio associato.

Nel caso in cui A sia diagonalizzabile, cioè sia $A = X\Lambda X^{-1}$ con $X = [x_1, \dots, x_n]$ matrice invertibile avente come colonne gli autovettori di A e $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ con $|\lambda_i|$ in ordine decrescente. Allora è possibile scrivere l' i -esimo iterato $x^{(i)}$ del metodo delle potenze, opportunamente scalato, come combinazione lineare degli autovettori come segue:

$$\begin{aligned} \frac{x^{(i)}}{\lambda_1^i} &= \frac{A^i x^{(0)}}{\lambda_1^i} = \frac{1}{\lambda_1^i} X \begin{bmatrix} \lambda_1^i & & \\ & \ddots & \\ & & \lambda_n^i \end{bmatrix} \underbrace{X^{-1} x^{(0)}}_{=: \xi} = X \begin{bmatrix} 1 & & \\ & \lambda_2^i / \lambda_1^i & \\ & & \ddots \\ & & & \lambda_n^i / \lambda_1^i \end{bmatrix} \xi \\ &= (\xi)_1 x_1 + \sum_{j=2}^n \frac{\lambda_j^i}{\lambda_1^i} (\xi)_j x_j \rightarrow \xi_1 x_1 \quad \text{per } i \rightarrow \infty, \end{aligned}$$

dove si è supposto che $(\xi)_1 \neq 0$. Tale ipotesi è importante ai fini della convergenza. Abbiamo in pratica dimostrato il seguente risultato.

Teorema 1.9. *Siano $|\lambda_i|, i = 1, \dots, n$ ordinati in modo decrescente, e $|\lambda_1| > |\lambda_2|$, con $|\lambda_1|$ semplice e sia x_1 l'autovettore corrispondente. Sia $x^{(0)} = X\xi$. Se $(\xi)_1 \neq 0$, allora esiste una costante $C > 0$ tale che*

$$\|x^{(i)} - x_1\|_2 \leq C \left| \frac{\lambda_2}{\lambda_1} \right|^i, \quad i \geq 1,$$

dove $x^{(i)}$ è normalizzato in modo opportuno.

Dimostrazione. Sia $v^{(i)} := A^i x^{(0)} / ((\xi)_1 \lambda_1^i)$, dove $v^{(i)}$ differisce da $x^{(i)}$ per un fattore di normalizzazione. Allora, sfruttando $A = X\Lambda X^{-1}$ e le espressioni scritte sopra, si ha $v^{(i)} = x_1 + \sum_{j=2}^n ((\xi)_j / (\xi)_1) (\lambda_j / \lambda_1)^i x_j$. Quindi

$$\|v^{(i)} - x_1\| \leq \sum_{j=2}^n \left| \frac{(\xi)_j}{(\xi)_1} \right| \left| \frac{\lambda_j}{\lambda_1} \right|^i \|x_j\| \leq \left| \frac{\lambda_2}{\lambda_1} \right|^i \sum_{j=2}^n \left| \frac{(\xi)_j}{(\xi)_1} \right| \|x_j\| = C \left| \frac{\lambda_2}{\lambda_1} \right|^i.$$

□

Capitolo 2

Reti complesse

Le reti possono apparire in varie forme, ma possono essere sempre pensate come una collezione di oggetti, e di collegamenti fra loro. Per analizzare le reti, dobbiamo trasformare questa definizione vaga in una definizione formale. Per far questo, dobbiamo introdurre alcune notazioni. Consideriamo un insieme finito $V = \{v_1, \dots, v_n\}$, e sia $V \times V$ l'insieme di tutte le coppie ordinate (v_i, v_j) degli elementi di V . Una relazione sull'insieme V è qualsiasi sottoinsieme $E \subseteq V \times V$. Diamo ora la seguente definizione per le reti semplici.

Definizione 2.1. *Una rete (o grafo) semplice è una coppia $G = (V, E)$, dove V è un insieme finito di nodi e E è una relazione su V simmetrica e antiriflessiva.*

La definizione appena data non include collegamenti multipli e self-loops. Di conseguenza, introduciamo una nuova definizione più generale di rete.

Definizione 2.2 (Rete). *Una rete (o grafo) è una tripletta $G = (V, E, f)$, dove V è un insieme finito ed è detto insieme dei nodi, $E \subseteq V \times V = \{e_1, \dots, e_m\}$ è un insieme di connessioni ed è detto insieme degli archi, f è una mappa che associa ad ogni arco e in E due vertici di V .*

D'ora in poi i termini rete e grafo saranno intercambiabili.

Definizione 2.3. *Un grafo si dice diretto se la relazione E è non simmetrica, ovvero se un nodo v_i è collegato a v_j allora non è detto che v_j sia necessariamente collegato a v_i . Viceversa un grafo si dice non diretto se la relazione E è simmetrica.*

Definizione 2.4 (Nodi adiacenti e incidenti). *Sia $G = (V, E)$ un grafo, diciamo che due nodi v_i e v_j sono adiacenti se sono collegati da un arco $e = (v_i, v_j)$. I nodi v_i e v_j sono incidenti con l'arco e , e l'arco e è incidente con i nodi v_i e v_j .*

Definizione 2.5 (Grafo pesato). Sia G grafo non diretto, con vertici $V = \{v_1, \dots, v_n\}$ e lati in E . Diremo che G è pesato, se per ogni lato fra due vertici v_i e v_j è associato un peso $w_{ij} \geq 0$. Inoltre, la matrice che ha come coefficienti i pesi del grafo, $W = (w_{ij})_{i,j=1,\dots,n}$ si chiama matrice di adiacenza.

Osservazione 2.6. Notiamo che se $w_{ij} = 0$ significa che il nodo v_i non è collegato al nodo v_j . Inoltre, se la matrice W di adiacenza è una matrice simmetrica, allora il grafo associato alla matrice sarà un grafo non diretto.

Definizione 2.7. Un sottoinsieme $A \subset V$ di un grafo si dice connesso, se ogni coppia di vertici in A è collegata da un percorso avente tutti i vertici in A . Inoltre, se A è connesso e non ha connessioni con il suo complementare, allora A si chiamerà componente connessa.

Definizione 2.8. Un grafo si dice fortemente connesso, se per ogni coppia di vertici esiste un cammino che li collega.

Definizione 2.9. Una matrice $A \in \mathbb{R}^{n \times n}$ si dice riducibile se esiste una matrice di permutazione T tale che

$$TAT^{-1} = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix}$$

con X e Z quadrate. Se la matrice T non esiste allora A si dice irriducibile.

Teorema 2.10. Una matrice $A \in \mathbb{R}^{n \times n}$ è irriducibile se e solo se il suo grafo associato è fortemente connesso. Equivalentemente una matrice è riducibile se il suo grafo associato non è fortemente connesso.

Dimostrazione. Dimostriamo che una matrice è riducibile se e solo se il grafo associato non è fortemente connesso.

Assumiamo quindi che A sia riducibile, e quindi che esista una matrice di permutazione T tale per cui

$$TAT^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

Notiamo che il grafo associato non varia se permuti gli elementi della matrice. Supponiamo che A_{11} abbia dimensione $r \times r$. Allora abbiamo che i nodi del grafo da $r + 1$ fino a n non saranno connessi con quelli da 1 a r , e quindi il grafo non è fortemente connesso. Viceversa, assumiamo che il grafo sia non fortemente connesso. In particolare, esiste un nodo p dal quale non posso raggiungere un nodo q . Ora definisco i due insiemi seguenti: P l'insieme dei nodi raggiungibili da p e Q l'insieme dei nodi non raggiungibili da p .

Notiamo che tutti i nodi di Q non sono raggiungibili dai nodi di P , allora dispongo la matrice in modo tale che su riga e colonna tutti gli indici di Q precedano quelli di P , ed ottengo una matrice ridotta nella forma desiderata. \square

2.1 Misure di centralità

Uno dei concetti più intuitivi a cui possiamo pensare quando affrontiamo l'analisi di una rete è l'identificazione dei nodi "più importanti". Questo concetto è formalizzato nella teoria delle reti mediante le misure di centralità.

Negli ultimi anni sono state fatte alcune ipotesi sul come calcolare l'importanza dei vari nodi in una rete, noi andremo ad analizzare alcune di queste ipotesi, cercando di evidenziare i vantaggi e gli svantaggi.

Per poter determinare quando un nodo è più importante di altri dobbiamo sicuramente tenere a mente almeno questi tre fattori:

- L'abilità del nodo di comunicare direttamente con altri nodi.
- La vicinanza a molti altri nodi.
- La quantità di coppie di nodi che hanno bisogno di quel nodo come intermediario.

2.1.1 Grado

Una prima ipotesi di centralità dei nodi è stata proposta da Freeman.

Definizione 2.11 (Grado di un nodo). *Sia G un grafo non orientato e non pesato, e A la sua matrice di adiacenza, allora definiamo il grado dell' i -esimo nodo, come*

$$d_i = \sum_{j=1}^n a_{ij} = (A\mathbf{1})_i.$$

Nel caso di grafi orientati il grado si distinguerà in indegree e outdegree

$$d_i^{in} = (\mathbf{1}^T A)_i, \quad d_i^{out} = (A\mathbf{1})_i, \quad (2.1)$$

in e *out* rappresentano rispettivamente il numero di archi incidenti al nodo i e il numero di archi incidenti dal nodo i .

L'idea di Freeman è stata quella di associare l'importanza dei nodi in base al grado. "Un nodo è più centrale (più influente) di un altro se il suo grado è più grande". Il grado

di un nodo conta i cammini di lunghezza uno di quel nodo dato, o equivalentemente, conta i cammini di lunghezza due che partono (e finiscono) da quel nodo:

$$d_i = (A^2)_{ii}.$$

Questo significa che il grado tiene conto degli effetti immediati che hanno luogo in una rete. Per esempio, se consideriamo un processo dove le “informazioni” sono passate da un nodo all’altro nella rete, la possibilità di ricevere tale informazione per un nodo è proporzionale al suo grado. Questo è il caso del rischio di infezione in una rete di persone, nel quale alcuni nodi sono persone infette, e si presuppone che il collegamento ad essi possa implicare un’infezione.

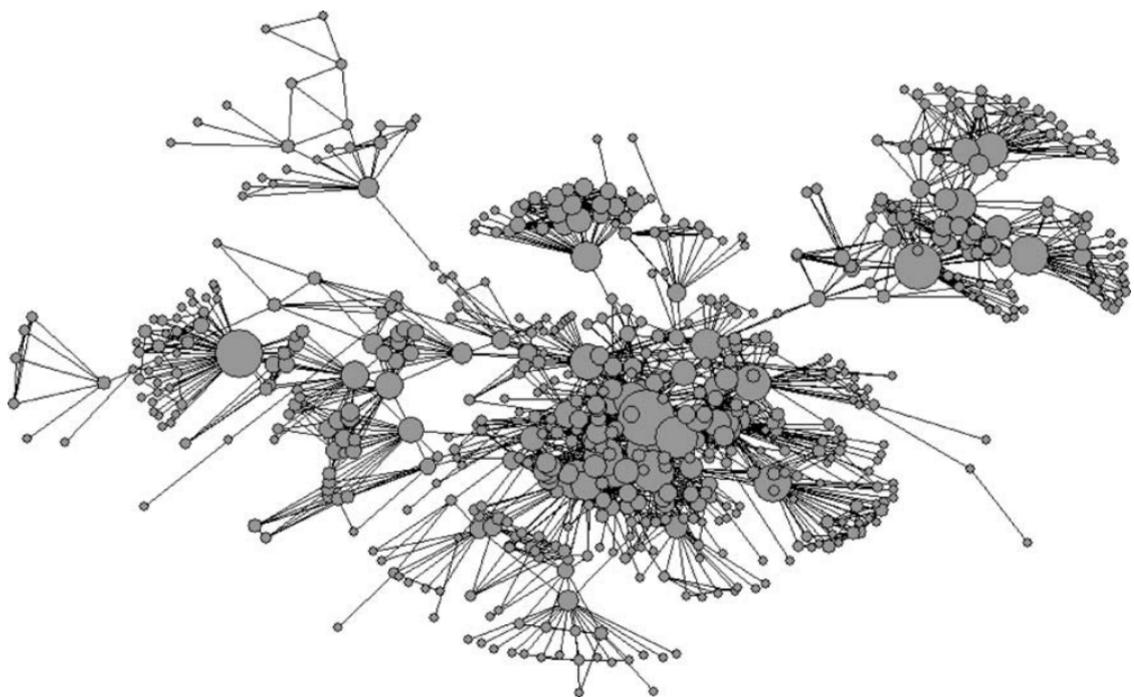


Figura 2.1: *Connessioni principali della rete sociale di consumatori di droghe iniettabili a Colorado Springs, USA. I nodi hanno il raggio proporzionale al loro grado. Figura presa da [Est12].*

La Figura 2.1 mostra una rete sociale di consumatori di droga a Colorado Springs, USA, dove gli archi indicano se due persone si sono scambiati degli aghi negli ultimi tre mesi. Il grado di ogni nodo è mostrato tramite la grandezza del nodo. Per gli individui con un grado alto e non ancora infetti dall’HIV, il rischio di diventare infetti è molto più alto rispetto agli individui non infetti che hanno un grado molto più basso.

Quando analizziamo grafi diretti invece, è necessaria l’analisi sia degli indegree che degli outdegree (2.1).

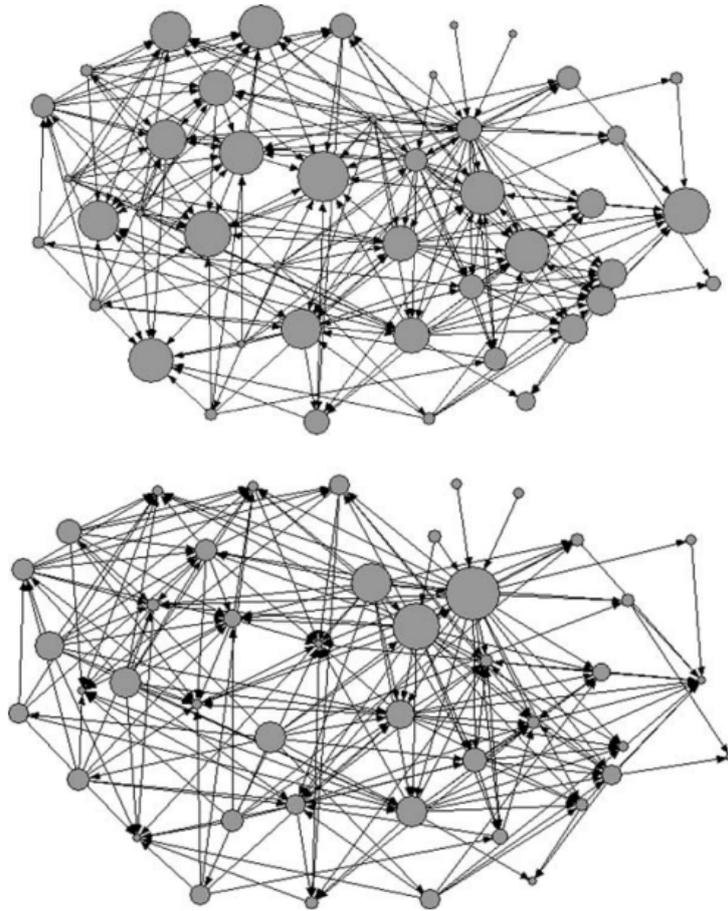


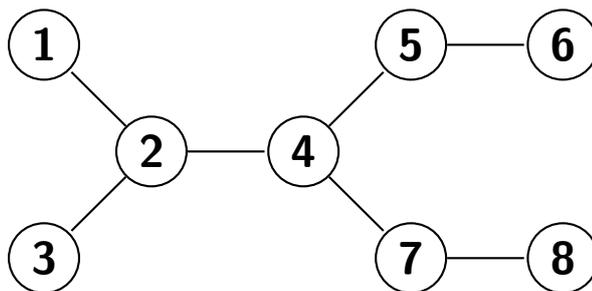
Figura 2.2: *I nodi rappresentano le specie e gli archi rappresentano i rapporti trofici. I nodi hanno raggio proporzionale all'indegree (sopra) e outdegree (sotto). Figura presa da [Est12].*

Una rappresentazione della catena alimentare animale dell'isola di St Martin è mostrata in Figura 2.2, dove il raggio dei nodi è proporzionale all'indegree (sopra) e outdegree (sotto). I nodi con un alto indegree rappresentano specie che sono prede di molti predatori in questo ecosistema, mentre i nodi con un alto outdegree rappresentano predatori con una grande varietà di prede. In generale, grandi predatori non sono delle prede, e per questo possiedono un grande outdegree, ma basso indegree, questo vale anche per il viceversa. Per lo stesso motivo, in una catena alimentare, ci aspettiamo una correlazione opposta fra indegree ed outdegree. Questa correlazione fra indegree ed outdegree può

essere misurata tramite il coefficiente di correlazione di Pearson:

$$\Omega = \frac{\sum_{i=1}^n (d_i^{in} - \overline{d^{in}}) (d_i^{out} - \overline{d^{out}})}{\sqrt{\sum_{i=1}^n (d_i^{in} - \overline{d^{in}})^2} \sqrt{\sum_{i=1}^n (d_i^{out} - \overline{d^{out}})^2}} \quad (2.2)$$

dove $\overline{d^{in}}$ e $\overline{d^{out}}$ sono l'indegree e outdegree medi della rete.



Questo metodo potrebbe però non andare sempre bene, infatti nel grafo riportato sopra, si ha che secondo questa teoria il nodo due e il nodo quattro hanno la stessa importanza, però è chiaro che il nodo quattro ha un posizionamento migliore rispetto al nodo due.

2.1.2 Centralità di Katz

Per parlare ora di altre misure di centralità, dobbiamo mostrare alcune operazioni matriciali.

Definizione 2.12. *Un cammino di lunghezza l è un elenco di nodi $P_1, P_2, \dots, P_l, P_{l+1}$ non necessariamente distinti (si può ripassare dallo stesso nodo e si può avere $i = j$), tale che per ogni $i = 1, \dots, l$ c'è un collegamento da P_i a P_{i+1} .*

Proposizione 2.13. *La quantità $(A^l)_{ij}$ conta il numero di cammini distinti ($i \neq j$) o cammini chiusi ($i = j$) di lunghezza l fra P_i e P_j .*

Una misura di centralità che tiene conto di cammini di lunghezza anche maggiore di uno è proposta da Katz, il quale definisce l'importanza del nodo i come

$$\begin{aligned} k_i &= ((\alpha^0 A^0 + \dots + \alpha^k A^k + \dots) \mathbf{1})_i \\ &= \sum_{j=1}^n \sum_{k=0}^{\infty} \alpha^k (A^k)_{ij} = (((I - \alpha A)^{-1} - I) \mathbf{1})_i \end{aligned} \quad (2.3)$$

(con $\alpha < \lambda_{\max}(A)^{-1}$ per avere convergenza).

Un esempio per questa misura è dato dalla matrice di adiacenza relativa al Grafo 2.1.1

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Calcolando gli autovalori otteniamo che $\lambda_{\max} = 2.1010$, allora se prendiamo $\alpha = 0.4$ otteniamo che $k_2 = 6.84$ e $k_4 = 7.83$, quindi a differenza della misura di centralità precedente abbiamo che i due nodi hanno valori diversi.

2.1.3 Eigenvector centrality

Un'ulteriore misura di centralità che analizzeremo sarà la eigenvector centrality. Questa misura di centralità calcola l'importanza dei nodi utilizzando autovalori ed autovettori. Per definirla dobbiamo prima analizzare alcuni risultati importanti della teoria di Perron-Frobenius.

Definizione 2.14. *Definiamo il raggio spettrale della matrice A come*

$$\rho(A) = \max_{\lambda \in \text{Spec}(A)} |\lambda|,$$

dove $\text{Spec}(A)$ è l'insieme degli autovalori di A .

Teorema 2.15 (Teorema di Perron-Frobenius). *Sia $A \in \mathbb{R}^{n \times n}$, $A > 0$, allora:*

- i) $\rho(A) > 0$.*
- ii) $\rho(A)$ è un autovalore di A .*
- iii) La molteplicità algebrica di $\rho(A)$ è 1 (cioè $\rho(A)$ è un autovalore semplice).*
- iv) Esiste un vettore $x > 0$ tale che $Ax = \rho(A)x$.*

Teorema 2.16 (Generalizzazione teorema di Perron-Frobenius). *Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$, e A irriducibile, allora:*

- i) $\rho(A) > 0$.*

ii) $\rho(A)$ è un autovalore di A .

iii) La molteplicità algebrica di $\rho(A)$ è 1 (cioè $\rho(A)$ è un autovalore semplice).

iv) Esiste un vettore $x > 0$ tale che $Ax = \rho(A)x$.

Definizione 2.17. Data una matrice $A \in \mathbb{R}^{n \times n}$, l'unico vettore $x \in \mathbb{R}^n$, $x > 0$ tale che $Ax = \rho(A)x$ si chiama vettore di Perron-Frobenius.

La eigenvector centrality calcola l'importanza di un nodo in base all'importanza dei nodi ai quali è collegato, e quindi si avrà

$$b_i = \left(\frac{1}{\lambda_{\max}} Av \right)_i$$

dove (λ_{\max}, v) è l'autocoppia di Perron-Frobenius di A . Nel caso dei grafi non diretti, ossia grafi con matrice associata simmetrica, se ne avrà solo uno poichè per le matrici simmetriche si hanno solo autovettori destri. Per i grafi diretti invece, ci sarà eigenvector centrality “destra” o “sinistra”, dove l'eigenvector centrality destra valuta “l'importanza” del nodo in base all’ “importanza” dei nodi che punta, mentre la sinistra valuta “l'importanza” del nodo in base all’ “importanza” dei nodi che lo puntano. Per fare un esempio, se volessimo calcolare l'importanza dei nodi del Grafo 2.1.1 tramite la eigenvector centrality, basterebbe prendere la matrice di adiacenza A e calcolare le autocoppie, una volta calcolate le autocoppie prendiamo l'autocoppia di Perron-Frobenius e andiamo a vedere le entrate dell'autovalore v , che in questo caso valuta $b_2 = 0.5000$ e $b_4 = 0.5745$. Per un esempio più pratico, consideriamo la rete polbooks, nella quale i nodi rappresentano libri sulla politica degli stati uniti venduti su Amazon, gli archi invece rappresentano frequenti acquisti di libri da parte degli stessi acquirenti. I risultati che si hanno prima applicando la misura di centralità che tiene conto solo del grado e successivamente la eigenvector centrality sono abbastanza diversi, e sono riportati nella tabella sottostante.

Importanza	Eigenvector centrality	Grado
1	Bushwhacked	A National Party No More
2	Big Lies	Off with Their Heads
3	American Dynasty	Losing Bin Laden
4	The Price of Loyalty	Bushwhacked
5	The Great Unraveling	American Dynasty
6	The Lies of George W. Bush	The Great Unraveling
7	Worse Than Watergate	Big Lies
8	Plan of Attack	The Price of Loyalty
9	Lies and the Lying Liars Who Tell Them	Legacy
10	Against All Enemies	Deliver Us from Evil

2.1.4 Subgraph centrality

La subgraph centrality invece ha come idea di fondo il calcolo dell'importanza di un nodo considerando tutte le sue partecipazioni in cammini chiusi che cominciano (e finiscono) nel nodo stesso. Ricordiamo che la Proposizione 2.13 afferma che $(A^l)_{ii}$ conta i cammini chiusi di lunghezza l che partono dal nodo i . Di conseguenza, la subgraph centrality è definita con la seguente funzione

$$f_i(A) = \left(\sum_{l=0}^{\infty} c_l A^l \right)_{ii} \quad (2.4)$$

dove i coefficienti c_l sono selezionati in modo tale che la serie converga. Il numero di cammini chiusi di lunghezza zero che coinvolgono un certo nodo è sempre uno. In ogni rete, come abbiamo già visto, i cammini chiusi di lunghezza $l = 2$ che partono dal nodo i sono proprio il grado del nodo i . Inoltre, nel caso in cui un grafo non presenti self loops, il numero di cammini chiusi di lunghezza $l = 1$ equivale a zero, allora possiamo riscrivere la funzione in (2.4) come segue:

$$f_i(A) - 1 = c_2 d_i + \sum_{l=3}^{\infty} (c_l A^l)_{ii}. \quad (2.5)$$

Possiamo poi scegliere i coefficienti c_l in modo tale da dare più importanza ai cammini chiusi corti rispetto ai cammini chiusi lunghi. Ci sono ovviamente varie scelte degli indici che tengono conto di quello che abbiamo detto. Una scelta in particolare è usata per la definizione dell'Estrada index

$$EE(i) = \left(\sum_{l=0}^{\infty} \frac{A^l}{l!} \right)_{ii} \stackrel{(1.1)}{=} (e^A)_{ii}. \quad (2.6)$$

L'Estrada index è usato anche per calcolare la facilità di passaggio di una informazione tra nodi distinti:

$$\left(\sum_{l=0}^{\infty} \frac{A^l}{l!} \right)_{ij} \stackrel{(1,1)}{=} (e^A)_{ij}. \quad (2.7)$$

In Figura 2.3 viene mostrata la subgraph centrality dei nodi nel social network dei consumatori di droga a Colorado Springs analizzato precedentemente. In questo caso però, si nota chiaramente che gli individui più influenti sono posizionati al centro della rete, indicando un grande scambio di sangue nelle loro relazioni.

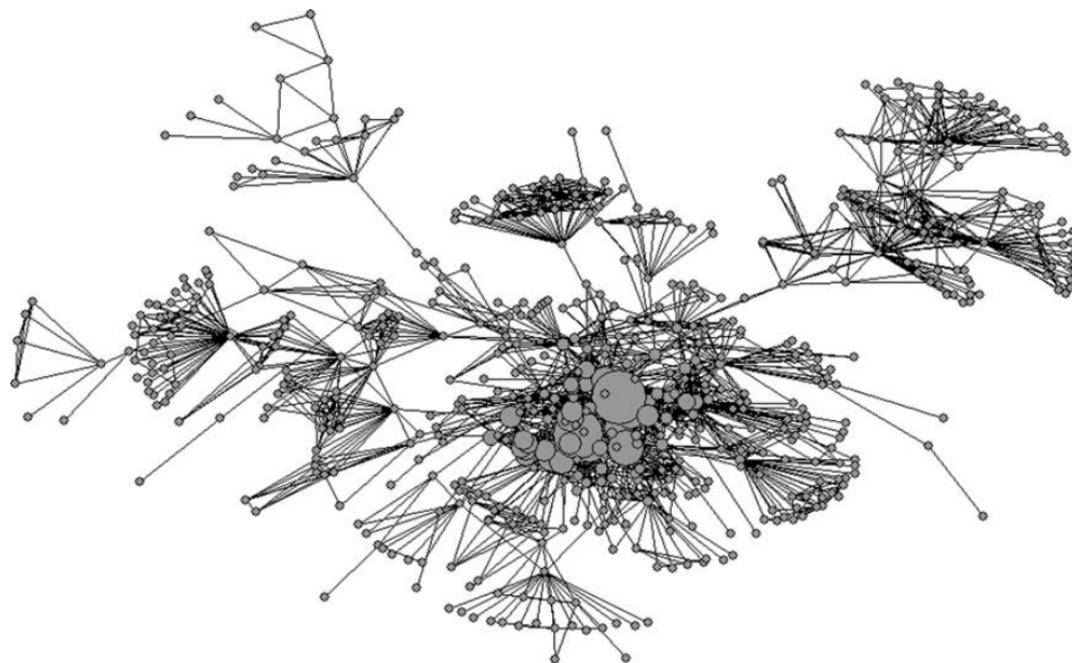


Figura 2.3: *Connessioni principali della rete sociale di consumatori di droghe iniettabili a Colorado Springs, USA. I nodi hanno il raggio proporzionale alla loro subgraph centrality. Figura presa da [Est12].*

2.1.5 PageRank

Un esempio classico di eigenvector centrality è il PageRank. Nel problema del PageRank dobbiamo capire l'ordine di importanza delle pagine web. La misura dell'importanza di una pagina web si basa sul numero di link da (outlinks) e verso (inlinks) la pagina. Se si utilizza un'analisi di importanza basata sul grado, Definizione 2.11, ovvero una pagina risulta importante se la puntano tante altre, si rischia una forte manipolazione. Un metodo più sicuro che si può adottare è definire l'importanza di una pagina

osservando l'importanza delle pagine che la puntano. Infatti, se definiamo P_j la j -esima pagina web, $\mathcal{I}(P_j)$ l'importanza della pagina P_j e l_j il numero di links della pagina P_j (outlinks), l'idea chiave diventa che se P_j punta P_k , allora P_j cede a P_k una frazione della sua importanza, uguale a $1/l_j$. Di conseguenza, l'importanza della pagina P_k sarà definita dalla somma di tutti i contributi ottenuti da altre pagine che puntano P_k

$$\mathcal{I}(P_k) = \frac{1}{l_1}\mathcal{I}(P_1) + \frac{1}{l_2}\mathcal{I}(P_2) + \frac{1}{l_3}\mathcal{I}(P_3) + \dots$$

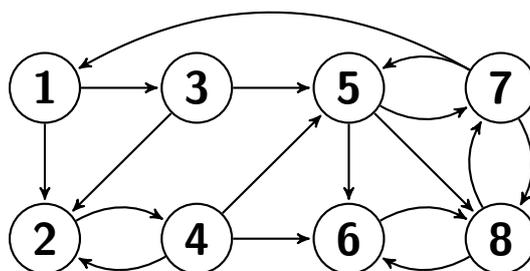
Vediamo il problema in termini di matrici, creiamo la matrice di hyperlink $H = (h_{ij})_{i,j=1,\dots,n}$ dove

$$h_{ij} = \begin{cases} \frac{1}{l_j} & \text{se } P_j \text{ punta } P_i, \\ 0 & \text{altrimenti.} \end{cases} \quad (2.8)$$

Per esempio

$$H = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & 0 & \frac{1}{3} \end{bmatrix} \Rightarrow \begin{bmatrix} \mathcal{I}(P_1) \\ \mathcal{I}(P_2) \\ \mathcal{I}(P_3) \\ \mathcal{I}(P_4) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \mathcal{I}(P_1) \\ \mathcal{I}(P_2) \\ \mathcal{I}(P_3) \\ \mathcal{I}(P_4) \end{bmatrix}$$

dove la riga i -esima rappresenta gli inlinks di P_i , mentre la colonna i -esima rappresenta gli outlinks della pagina P_i . La matrice di hyperlinks H è una matrice stocastica per colonna, ovvero tutti gli elementi sono non negativi, ed inoltre, la somma degli elementi sulle colonne è uno (a meno che una pagina non abbia outlinks). Se ora definiamo $\mathbf{I} = (\mathcal{I}(P_i))_{i=1,\dots,n}$ allora, il problema corrisponde a $\mathbf{I} = H\mathbf{I}$. Vediamo ora un esempio. Prendiamo un grafo della forma



Allora la matrice di adiacenza ed autovettore corrispondente saranno dati da

$$H = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & 0 \end{bmatrix}, \quad \mathbf{I} = \begin{bmatrix} 0.1400 \\ 0.1576 \\ 0.0700 \\ 0.1576 \\ 0.2276 \\ 0.4727 \\ 0.4201 \\ 0.6886 \end{bmatrix}$$

Si possono presentare però delle problematiche, non è detto che H abbia autovalore uno, ed inoltre se H ha delle colonne di zeri non è più stocastica per colonna. Per ovviare a questi problemi modifichiamo H , definendo

$$d_j = \begin{cases} 1 & \text{se } l_j = 0, \\ 0 & \text{altrimenti,} \end{cases}$$

e poniamo

$$P = H + \frac{1}{n} \mathbf{1} d^t.$$

Abbiamo quindi ottenuto ora una matrice stocastica per colonna (ogni colonna di zeri che aveva H è stata ora cambiata con un colonna di $1/n$) e una matrice con autovalore 1. Abbiamo ancora un problema però da risolvere, ci chiediamo se $\lambda = 1$ è un autovettore semplice, in modo tale da avere \mathbf{I} come autovettore unico.

Teorema 2.18 (Variante teorema Perron-Frobenius). *Sia A irriducibile e stocastica per colonna. Allora l'autovalore dominante ($|\lambda_1| = \max |\lambda|$) è uguale ad 1. Inoltre, esiste un unico autovettore associato, r , con componenti tutte positive, e questo è l'unico autovalore con componenti tutte non negative. Infine, se A ha tutti elementi positivi, $|\lambda_i|$, $i = 2, 3, \dots, n$.*

Normalmente però H è riducibile, di conseguenza la procedura di PageRank non è ben definita (l'autovalore non è semplice). Per assicurarci l'irriducibilità della matrice, prendiamo

$$P = \alpha H + (1 - \alpha) \frac{1}{n} \mathbf{1} \mathbf{1}^T, \quad \alpha \in [0, 1].$$

Teorema 2.19. *Sia $\text{Spec}(H) = \{1, \lambda_2, \dots, \lambda_n\}$. Allora $\text{Spec}(P) = \{1, \alpha \lambda_2, \dots, \alpha \lambda_n\}$ (anche se H ha autovalore 1 multiplo, P ha autovalore 1 semplice).*

Dimostrazione. Sia $\bar{\mathbf{1}} = \frac{1}{\sqrt{n}} \mathbf{1}$, allora $P^T \bar{\mathbf{1}} = \alpha H \bar{\mathbf{1}} + (1 - \alpha) \frac{1}{n} \mathbf{1} \mathbf{1}^T \bar{\mathbf{1}} = \bar{\mathbf{1}}$. Sia ora $U = [\bar{\mathbf{1}}, U_0]$ unitaria, ovvero $\bar{\mathbf{1}}^T U_0 = 0$, allora

$$U^* H U = \begin{bmatrix} \bar{\mathbf{1}}^T \\ U_0^* \end{bmatrix} H \begin{bmatrix} \bar{\mathbf{1}} & U_0 \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{1}}^T H \bar{\mathbf{1}} & \bar{\mathbf{1}}^T H U_0 \\ U_0^* H \bar{\mathbf{1}} & U_0^* H U_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ * & T \end{bmatrix}.$$

Per cui

$$U^*PU = \alpha \begin{bmatrix} 1 & 0 \\ * & T \end{bmatrix} + (1 - \alpha) \frac{1}{n} ne_1 \bar{\mathbf{1}}^T U = \begin{bmatrix} 1 & 0 \\ * & \alpha T \end{bmatrix}.$$

□

2.2 Betweenness centrality

In questa sezione considereremo l'importanza relativa di un nodo nella comunicazione tra altre coppie di nodi. Queste misure di centralità rappresentano la proporzione di informazioni che passano tramite un determinato nodo nella comunicazione con altre coppie di nodi nella rete, e sono genericamente chiamate *betweenness centrality*. Il modo più intuitivo per definire la *betweenness centrality* sfrutta il fatto che l'informazione che va da un nodo all'altro viaggia solo attraverso i percorsi più brevi che collegano quei nodi. Quindi, se $\rho(i, j)$ è il numero dei cammini più brevi che vanno dal nodo i al nodo j e $\rho(i, k, j)$ è il numero di cammini più brevi che partono dal nodo i e arrivano al nodo j passando per il nodo k , la *betweenness centrality* del nodo k è data da:

$$BC(k) = \sum_i \sum_j \frac{\rho(i, k, j)}{\rho(i, j)}, \quad i \neq j \neq k. \quad (2.9)$$

Ovviamente la comunicazione in una rete complessa non avviene sempre attraverso i percorsi più brevi che collegano coppie di nodi. In molte situazioni del mondo reale tale comunicazione avviene utilizzando alcuni o addirittura tutti i canali disponibili per spostarsi da un punto all'altro della rete. Sono state proposte diverse misure per tenere conto della *betweenness* di un nodo quando la comunicazione avviene utilizzando non solo i percorsi più brevi. Uno di questi esempi è la *random walk betweenness centrality*, la quale si basa sul numero di volte che una *random walk* fra i e j passa per k . Questa misura è definita come segue:

$$BC_{RW}(r) = \sum_{p < q} I_r^{pq}, \quad (2.10)$$

dove $I_p^{pq} = I_q^{pq} = 1$ e

$$I_r^{pq} = \frac{1}{2} \sum_j A_{rj} |T_{rp} - T_{rq} - T_{jp} + T_{jq}|, \quad r \neq p, q. \quad (2.11)$$

La matrice T è ottenuta rimuovendo la k -esima riga e colonna dalla matrice A , ottenendo una matrice L_k di dimensione $(n - 1) \times (n - 1)$, successivamente invertiamo L_k ed aggiungiamo una riga e colonna di tutti zeri dove abbiamo rimosso la riga e la

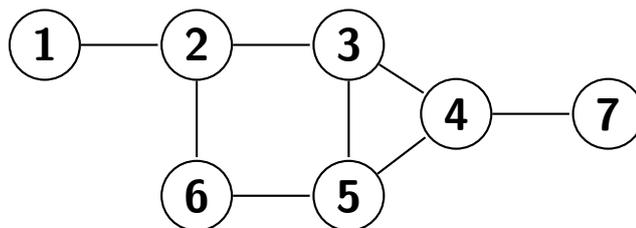
colonna prima. L'ultima misura che vedremo è la communicability betweenness centrality, la quale fornisce l'importanza all' r -esimo nodo togliendolo dalla rete e calcolando la comunicabilità senza di esso. In questo caso la communicability betweenness centrality del nodo r è data da:

$$BC_r = \frac{1}{C} \sum_i \sum_j \frac{(f(A)_{ij} - f(A - E(r))_{ij})}{f(A)_{ij}} \in (0, 1), \quad i \neq j, i \neq r, j \neq r. \quad (2.12)$$

dove $C = (n-1)^2 - (n-1)$, $A - E(r)$ equivale alla matrice di adiacenza dove la colonna e la riga r -esima sono zero. Se come funzione prendiamo l'esponenziale, allora otterremo

$$BC_r = \frac{1}{C} \sum_i \sum_j \frac{(e^A)_{ij} - (e^{A-E(r)})_{ij}}{(e^A)_{ij}}, \quad i \neq j, i \neq r, j \neq r. \quad (2.13)$$

Vediamo ora le differenze fra queste tre misure di centralità analizzando il grafo sottostante.



k	$BC(k)$	$BC_{RW}(k)$	$BC_C(k)$
1	0.0	0.286	0.029
2	5.5	0.619	0.451
3	5.0	0.576	0.505
4	5.0	0.584	0.446
5	2.5	0.545	0.397
6	1.0	0.450	0.193
7	0.0	0.286	0.027

Come possiamo vedere, la betweenness centrality che conta i cammini più brevi identifica il nodo 2 come più centrale. Tuttavia, assegna la stessa “importanza” ai nodi 3 e 4. Quando non solo i cammini più brevi vengono presi in considerazione, i risultati cambiano. In particolare la random walk betweenness centrality tiene come nodo più centrale il secondo, ma assegna al nodo 4 un valore maggiore al nodo 3. La communicability betweenness invece valuta come nodo più “importante” il nodo 3 al posto del nodo 2, questo quindi significa che il nodo 3 gioca un ruolo più centrale nel distribuire informazioni agli altri nodi quando tutti i tipi di percorsi sono usati come percorsi per le informazioni.

Capitolo 3

Eigenvector centrality del secondo ordine

In questo capitolo vedremo un'analisi di centralità dei nodi fino al secondo ordine, utilizzando nozioni di base sui tensori. D'ora in poi assumeremo che i grafi siano non diretti, cosicchè per ogni $(i, j) \in E$ si ha anche $(j, i) \in E$; non pesati, cosicchè tutte le connessioni della rete abbiano la stessa “forza”; e connessi, cosicchè si possa raggiungere ogni nodo del grafo da qualsiasi altro nodo. Assumeremo anche per semplicità che i grafi non abbiano self-loops, ovvero, un nodo non potrà puntarsi da solo. Sotto le ipotesi che abbiamo appena fatto allora abbiamo che la matrice di adiacenza A del grafo è sicuramente simmetrica, irriducibile e binaria.

3.1 Coefficiente di clustering di Watts-Strogatz

Molte reti reali sono caratterizzate dalla presenza di un grosso numero di triangoli. Questa caratteristica è generalmente conseguenza della forte transitività delle reti. Per esempio, in una rete sociale è fortemente probabile che, se Marco e Filippo sono entrambi amici di Giorgio, alla fine verranno presentati tra loro da Giorgio, chiudendo una relazione transitiva, o equivalentemente, formando un triangolo. Watts e Strogatz hanno proposto un coefficiente per analizzare i triangoli nelle reti. Più precisamente, per definire il coefficiente, useremo $\Delta(i) = (A^3)_{ii}/2$, che denota il numero di triangoli non orientati che coinvolgono il nodo i . Notiamo che il nodo i è coinvolto in esattamente $d_i(d_i - 1)/2$ spigoli centrati in i (dove d_i rappresenta il grado del nodo i , Definizione 2.11), i quali sono percorsi della forma hij dove h, i, j sono tutti distinti, inoltre $d_i(d_i - 1)$ rappresenta anche il numero massimo di triangoli non orientati nei quali il nodo i può essere coinvolto. Il coefficiente di clustering di Watts-Strogatz del nodo i è definito come il numero di

triangoli nei quali il nodo i è coinvolto fratto il numero di tutti i triangoli possibili nei quali potrebbe essere coinvolto, quindi

$$c_i = \begin{cases} \frac{2\Delta(i)}{d_i(d_i-1)} & \text{se } d_i \geq 2, \\ 0 & \text{altrimenti.} \end{cases} \quad (3.1)$$

Si nota facilmente che $c_i \in [0, 1]$ con $c_i = 0$ se il nodo i non partecipa in nessun triangolo, mentre $c_i = 1$ se non lascia nessuno spigolo aperto. Collegati a questa misura di transitività ci sono due versioni a livello di rete:

- Il coefficiente di clustering medio

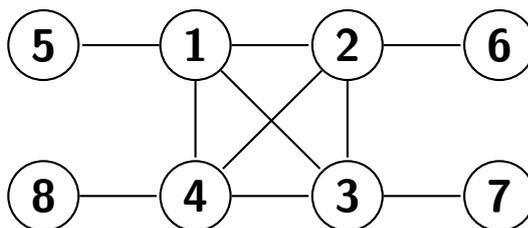
$$\bar{C} = \frac{1}{n} \sum_{i=1}^n c_i = \frac{2}{n} \sum_{i:d_i \geq 2} \frac{\Delta(i)}{d_i(d_i-1)}.$$

- Il coefficiente di clustering globale

$$\hat{C} = \frac{6|K_3|}{\sum_i d_i(d_i-1)},$$

dove $|K_3|$ è il numero di triangoli non orientati della rete.

Il coefficiente di clustering medio rappresenta la media dei coefficienti di clustering di tutti i nodi, invece il coefficiente di clustering globale rappresenta il numero di triangoli di tutta la rete diviso tutti i possibili triangoli che si potrebbero formare. Vediamo ora un esempio semplice.



I nodi 1 e 3 sono uguali, entrambi sono coinvolti in due triangoli non orientati, quindi avremo

$$c_1 = c_3 = \frac{2 \cdot 2}{4 \cdot 3} = \frac{1}{3},$$

Allo stesso modo otteniamo $c_2 = c_4 = 1/3$, mentre $c_5 = c_6 = c_7 = c_8 = 0$ poichè i nodi 5, 6, 7, 8 non sono coinvolti in alcun triangolo. Con questi valori otteniamo

$$\bar{C} = \frac{14}{8 \cdot 3} = \frac{1}{6}, \quad \hat{C} = \frac{6 \cdot 8}{48} = 1.$$

3.2 Modello generale dell'eigenvector centrality del secondo ordine

Per incorporare informazioni del secondo ordine, dato un tensore $T \in \mathbb{R}^{n \times n \times n}$ ed un parametro $p \in \mathbb{R}$, definiamo una funzione $T_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ in questo modo

$$T_p(x)_i = \sum_{j,k=1}^n T_{ijk} \mu_p(x_j, x_k), \quad (3.2)$$

dove $\mu_p(a, b)$ è la media di potenze

$$\mu_p(a, b) = \left(\frac{|a|^p + |b|^p}{2} \right)^{1/p}. \quad (3.3)$$

Riportiamo brevemente alcune proprietà della media di potenze

- i) $\lim_{p \rightarrow 0} \mu_p(a, b) = \sqrt{|ab|}$ è la media geometrica.
- ii) $\mu_{-1}(a, b) = 2(|a|^{-1} + |b|^{-1})^{-1}$ è la media armonica.
- iii) $\lim_{p \rightarrow \infty} \mu_p = \max\{|a|, |b|\}$, mentre al contrario $\lim_{p \rightarrow -\infty} \mu_p = \min\{|a|, |b|\}$.

Definizione 3.1. Sia $\alpha \in \mathbb{R}$ tale che $\alpha \in [0, 1]$, $p \in \mathbb{R}$ e siano $M \in \mathbb{R}^{n \times n}$ e $T \in \mathbb{R}^{n \times n \times n}$ rispettivamente una matrice quadrata non negativa ed un tensore cubico non negativo associati alla rete. Definiamo $\mathcal{M} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ come

$$\mathcal{M}(x) = \alpha Mx + (1 - \alpha) T_p(x). \quad (3.4)$$

Allora i corrispettivi primo e secondo ordine eigenvector centrality del nodo i è dato da $x_i \geq 0$, dove $x \in \mathbb{R}^n$ risolve il problema agli autovalori non lineare

$$x \geq 0 \text{ tale che } \mathcal{M}(x) = \lambda x. \quad (3.5)$$

Se poniamo $\alpha = 1$ in (3.4) avremo che verranno considerate solo le interazioni del primo ordine, in questo modo torniamo all'eigenvector centrality classico analizzato nella Sezione 2.1.3. Allo stesso modo, se poniamo $\alpha = 0$ verranno considerate solo interazioni del secondo ordine. Affinchè la Definizione 3.1 sia ben posta, la (3.5) dovrà avere un'unica soluzione. Questo problema sarà discusso nella Sezione 3.2.3.

3.2.1 Scelte di M e di T

Nella Definizione 3.1, la matrice M possiede le informazioni riguardanti le interazioni del primo ordine, mentre il tensore T possiede le informazioni riguardanti le interazioni

del secondo ordine, ovvero tiene in considerazione i cammini di lunghezza due del grafo. Scelte utili per M potrebbero essere la matrice di adiacenza, o anche la matrice di PageRank. Ovviamente queste scelte devono essere fatte in base ai dati disponibili e a come vogliamo analizzarli.

Vediamo ora scelte del tensore T per analizzare le interazioni del secondo ordine.

Definizione 3.2. *Definiamo il tensore triangolare binario T_B come il tensore con coefficienti:*

$$(T_B)_{ijk} = \begin{cases} 1 & \text{se } i, j, k \text{ formano un triangolo,} \\ 0 & \text{altrimenti.} \end{cases} \quad (3.6)$$

Come discusso in [SW05] il più alto costo computazionale per costruire T_B è $O(n^3)$ o $O(m^{3/2})$, dove n è il numero di nodi della rete, mentre m è il numero di archi. Si è studiato però in [BGL15] che costruendo il tensore triangolare binario di quattro reti reali il numero di non zeri in T_B risulta $O(6m)$. Notiamo infine che questo tensore è strettamente collegato alla matrice $A \circ A^2$, dove \circ rappresenta il prodotto di Hadamard. Si verifica facilmente che, indifferentemente dalla scelta di p , $((T_B)_p(\mathbf{1}))_i = (A^3)_{ii} = 2\Delta(i) \forall i \in V$.

Definizione 3.3. *Definiamo il tensore triangolare del random walk T_W come il tensore con coefficienti:*

$$(T_W)_{ijk} = \begin{cases} \frac{1}{\Delta(j,k)} & \text{se } i, j, k \text{ formano un triangolo,} \\ 0 & \text{altrimenti,} \end{cases} \quad (3.7)$$

dove $\Delta(j, k) = (A \circ A^2)_{jk}$ è il numero di triangoli nel quale è coinvolto il lato (j, k) .

Definizione 3.4. *Definiamo il tensore triangolare dei coefficienti di clustering T_C come il tensore con coefficienti*

$$(T_C)_{ijk} = \begin{cases} \frac{1}{d_i(d_i-1)} & \text{se } i, j, k \text{ formano un triangolo,} \\ 0 & \text{altrimenti.} \end{cases} \quad (3.8)$$

Notiamo che se i, j, k formano un triangolo, allora $d_i \geq 2$ e quindi $(T_C)_{ijk}$ è ben definito. Questo tensore fornisce informazioni aggiuntive, le quali non sono date da (3.6) e (3.7), ovvero il numero di relazioni transitive (cioè di triangoli) nelle quali il nodo i potrebbe essere coinvolto.

Per definire il prossimo, ed ultimo tensore è necessaria una definizione preliminare.

Definizione 3.5. *Il coefficiente di chiusura locale del nodo i è definito come*

$$h_i = \frac{2\Delta(i)}{\omega(i)}, \quad (3.9)$$

dove

$$\omega(i) = \sum_{j \in N(i)} d_j - d_i = \sum_{j \in N(i)} (d_j - 1) \quad (3.10)$$

è il numero di cammini di lunghezza due che partono dal nodo i , e $N(i)$ è l'insieme dei nodi vicini del nodo i . Si potrebbe riscrivere come $\omega = A\mathbf{d} - \mathbf{d} = A^2\mathbf{1} - A\mathbf{1}$.

Il risultato seguente segue immediatamente dalla definizione di $\omega(i)$, e mostra che in pratica è possibile assumere $\omega(i) \neq 0$ nel caso di reti reali.

Proposizione 3.6. *Sia $G = (V, E)$ un grafo non diretto, non pesato e connesso. Allora $\omega(i) = 0$ se e solo se tutti i nodi vicini a i hanno grado uguale ad uno. Inoltre, se $\omega(i) = 0$ per qualche i , allora G sarà o un grafo con due nodi, o un grafo a stella con $n \geq 3$, avente il nodo i al centro.*

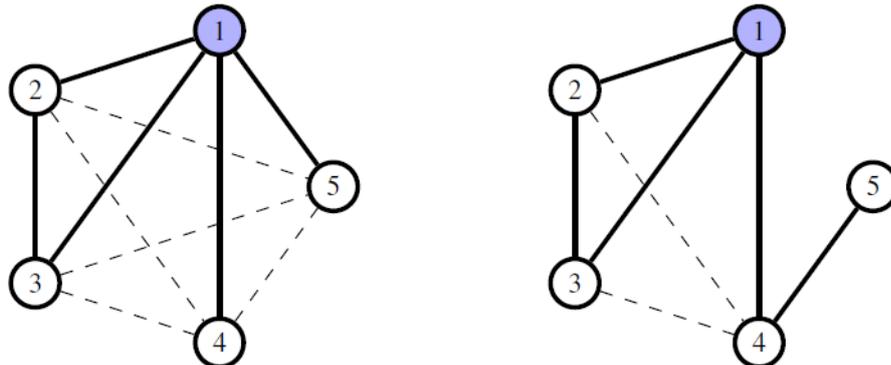
Definizione 3.7. *Definiamo il tensore triangolare di chiusura locale T_L come il tensore di coefficienti*

$$(T_L)_{ijk} = \begin{cases} \frac{1}{\omega(i)} & \text{se } i, j, k \text{ formano un triangolo,} \\ 0 & \text{altrimenti,} \end{cases} \quad (3.11)$$

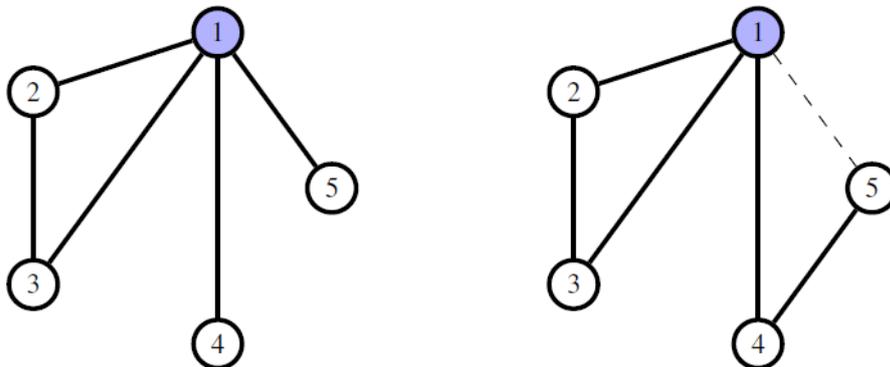
Si verifica facilmente che $(T_L)_p(\mathbf{1}) = h \forall p$. Questo tensore analizza informazioni che non sono analizzate dagli altri tensori, ovvero il potenziale di un nodo di essere coinvolto in un triangolo connettendo nodi a distanza due da esso.

Analizziamo brevemente ora le differenze dei quattro tensori appena introdotti. Il tensore triangolare binario (3.6) e il tensore triangolare del random walk (3.7) non contengono alcuna informazione che riguarda gli spigoli che coinvolgono ciascun nodo, e quindi il potenziale che avrebbe un nodo di formare triangoli. Quindi reti che hanno differenti strutture dal punto di vista delle potenziali ed effettive relazioni transitive sono trattati allo stesso modo.

a



b



Queste immagini rappresentano grafi con lo stesso numero di archi e di triangoli. Il grafo a sinistra di (a) ha il nodo 1 che potrebbe essere coinvolto in ulteriori 5 triangoli non orientati, e sono illustrati usando le linee tratteggiate, invece il grafo a destra di (a) ha il nodo 1 che potrebbe essere coinvolto in due ulteriori triangoli. Si fa un ragionamento analogo per (b).

Per vedere le differenze dei tensori notiamo che le due reti di (a) e anche quelle di (b) sono rappresentate dagli stessi tensori se usiamo il tensore triangolare binario (3.6) ed il tensore triangolare del random walk (3.7), ma dal momento che dal punto di vista delle relazioni transitive i due grafi sono molto diversi, se prendiamo il tensore triangolare dei coefficienti di clustering (3.8) per (a), il quale conta anche i potenziali triangoli che si potrebbero formare, otteniamo che $(T_C)_{123} = (T_C)_{132} = 1/12$ per il grafo sulla sinistra e $(T_C)_{123} = (T_C)_{132} = 1/6$ per il grafo sulla destra. Notiamo che il valore $1/12$ viene dal fatto che stiamo contando i triangoli orientati in questo caso, e vale lo stesso per $1/6$. Analizziamo invece il grafo (b) con il tensore triangolare di chiusura locale (3.11), il quale ci fornisce informazioni sul potenziale di un nodo di essere coinvolto in un triangolo collegando nodi a distanza due da esso: Il grafo a sinistra sarà $(T_L)_{123} = (T_L)_{132} = 1/2$, mentre il grafo a destra sarà $(T_L)_{123} = (T_L)_{132} = 1/3$.

3.2.2 Scelte di α e di p

La mappa \mathcal{M} definita in (3.4) diventa una mappa lineare per alcune scelte di α o p . Il caso più semplice avviene quando $\alpha = 1$, che corrisponde al semplice prodotto matrice vettore, $\mathcal{M}x = Mx$, e quindi (3.5) si riduce ad un generico problema agli autovalori lineare. In questo caso le varie scelte di M viste nella Sezione 3.2.1 daranno risultati in base a cosa stiamo cercando, quindi ad esempio il PageRank rientra in questo sottocaso.

Sia ora $\alpha \in [0, 1)$ e $p = 1$. La mappa $T_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ diventa anche in questo caso lineare, poichè la media delle potenze diventa media aritmetica

$$T_1(x)_i = \frac{1}{2} \sum_{j,k=1}^n T_{ijk}x_k + T_{ijk}x_j = \frac{1}{2} \left\{ \sum_{j=1}^n \left(\sum_{k=1}^n T_{ikj} \right) x_j + \sum_{j=1}^n \left(\sum_{k=1}^n T_{ijk} \right) x_j \right\}, \quad (3.12)$$

e $T_1(x)$ si riduce al prodotto tra il vettore x e la matrice con entrate $\frac{1}{2}(\sum_k T_{ijk} + T_{ikj})$. In particolare se il tensore T è simmetrico rispetto al secondo e terzo fattore ($T_{ijk} = T_{ikj}$) per ogni j, k , si ha

$$T_1(x)_i = \sum_{j=1}^n \left(\sum_{k=1}^n T_{ijk} \right) x_j. \quad (3.13)$$

Questo è il caso di tutti i tensori definiti nella Sezione 3.2.1.

È interessante soffermarsi sul comportamento di alcuni tensori definiti nella Sezione 3.2.1 per $p = 1$. Se $T = T_B$, si ha che

$$\sum_{k=1}^n (T_B)_{ijk} = (A \circ A^2)_{ij} \quad (3.14)$$

e quindi

$$(T_B)_1(x) = (A \circ A^2)x. \quad (3.15)$$

Allora in questo caso la mappa \mathcal{M} sarà

$$\mathcal{M}(x) = \alpha Mx + (T_B)_1(x) = (\alpha A + (1 - \alpha)(A \circ A^2))x, \quad (3.16)$$

e quindi la soluzione del problema degli autovalori sarà l'autovalore di Perron-Frobenius della matrice $\alpha A + (1 - \alpha)(A \circ A^2)$.

Se invece $T = T_C$, avremo che

$$\sum_{k=1}^n (T_C)_{ijk} = \begin{cases} \frac{(A \circ A^2)_{ij}}{d_i(d_i-1)} & \text{se } d_i \geq 2, \\ 0 & \text{altrimenti,} \end{cases} \quad (3.17)$$

allora (3.4) diventa

$$\mathcal{M}(x) = \alpha Mx + (1 - \alpha)(T_C)_1(x) = \left(\alpha A + (1 - \alpha)(D^2 - D)^\dagger (A \circ A^2) \right)x, \quad (3.18)$$

dove D è la matrice diagonale dei nodi $D_{ii} = d_i$, e \dagger rappresenta la pseudoinversa di Moore-Penrose, definita della Sezione 1.2. Se come ultimo caso si considera il tensore triangolare di chiusura locale si ottiene

$$\sum_{k=1}^n (T_L)_{ijk} = \begin{cases} \frac{(A \circ A^2)_{ij}}{w(i)} & \text{se } d_j \geq 2, \\ 0 & \text{altrimenti,} \end{cases} \quad (3.19)$$

con $\omega(i)$ definito dall'equazione (3.10). Notiamo in (3.19) che ci possono essere dei casi nei quali $\sum_{k=1}^n (T_L)_{ijk} = 0$ anche quando $d_j \geq 2$. Infatti, dalla Proposizione 3.6 segue che esistono dei casi nei quali $\omega(i) > 0$ ma i non forma nessun triangolo, e quindi $(A \circ A^2)_{ij} = 0$ per ogni j con $d_j \geq 2$.

Usando (3.18) otteniamo

$$(T_L)_1(x) = W^\dagger (A \circ A^2) x, \quad (3.20)$$

dove $W = \text{diag}(\omega(1), \dots, \omega(n))$. Quindi (3.4) diventa

$$\mathcal{M}(x) = \alpha Mx + (1 - \alpha) (T_L)_1(x) = (\alpha A + (1 - \alpha) W^\dagger (A \circ A^2)) x. \quad (3.21)$$

Mentre la scelta $\alpha = 1$ fornisce una mappa lineare e del primo ordine, la scelta $\alpha = 0$ corrisponde ad avere una mappa che considera solo le interazioni del secondo ordine tra i nodi. In particolare, questa scelta di α permette di definire la versione spettrale del coefficiente di clustering di Watts-Strogatz e il coefficiente di chiusura locale, dove il parametro p della media di potenze controlla come i coefficienti dei nodi vicini sono combinati.

Definizione 3.8. *Sia $T \in \mathbb{R}^{n \times n \times n}$ un tensore non negativo associato alla rete. Il coefficiente spettrale di clustering del nodo i è l' i -esimo elemento del vettore $x \geq 0$, il quale risolve il problema degli autovalori (3.5) con $\alpha = 0$ in (3.4), cioè*

$$T_P(x) = \lambda x. \quad (3.22)$$

La soluzione per $T = T_C$ sarà chiamata coefficiente spettrale di clustering di Watts-Strogatz, mentre la soluzione per $T = T_L$ sarà chiamata coefficiente spettrale di chiusura locale.

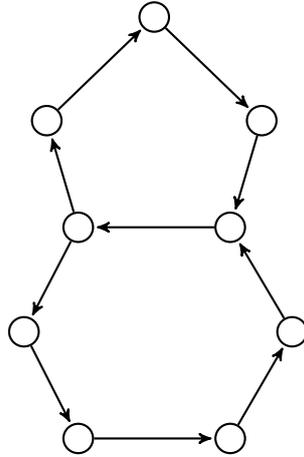
Se il nodo i non partecipa in nessun triangolo, allora la sommatoria che descrive la corrispondente entrata in $T_p(x)$ è vuota, e quindi il coefficiente spettrale di clustering per quel nodo è zero, come ci si aspettava. Ricordiamo però che poichè il coefficiente spettrale di clustering x è definito tramite un problema agli autovettori per T_p , segue che non può essere unico. Infatti, si ha che $T_p(\theta x) = \theta T_p(x)$ per ogni $\theta \geq 0$. Quindi quando $T = T_C$, a differenza del classico coefficiente di clustering di Watts-Strogatz, non è più vero che il coefficiente spettrale di clustering identifica i nodi che partecipano in tutti i possibili triangoli.

3.2.3 Esistenza e unicità

Fino ad ora abbiamo scelto grafi non orientati, le informazioni del secondo ordine però possono essere ovviamente incorporate anche nei casi di grafi diretti. Il problema nell'affrontare grafi diretti è che il range di possibilità da considerare aumenta notevolmente (ad esempio, tenendo conto di ogni triangolo orientato) rendendo le interpretazioni delle misure meno chiare. Nonostante questo però, in termini di studio di unicità, esistenza e problemi di calcolabilità non si ha quasi nessun svantaggio nel considerare anche i grafi diretti. Abbiamo discusso i casi lineari $\alpha = 1$ e $p = 1$, per i quali la mappa $\mathcal{M} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ diventa una matrice non negativa B . In questo caso, i risultati della teoria di Perron-Frobenius forniscono condizioni su \mathcal{M} che garantiscono l'esistenza di una soluzione di (3.5), la quale si può trovare tramite il metodo delle potenze classico.

Definizione 3.9. *Un grafo G si dice aperiodico se il massimo comune divisore della lunghezza di tutti i suoi cicli è 1.*

Il grafo sottostante è un classico esempio di grafo aperiodico, dal momento che i suoi cicli sono di lunghezza 6 e 5 e quindi il massimo comune divisore equivale ad 1.



Definizione 3.10. *Una matrice $B \in \mathbb{R}^{n \times n}$ si dice primitiva se e solo se esiste un intero $k \geq 1$ tale che $B^k > 0$.*

Proposizione 3.11. *Una matrice $B \geq 0$ è primitiva se e solo se il suo grafo associato G_B è aperiodico.*

Nel caso generale, faremo riferimento alla teoria di Perron-Frobenius non lineare, per mostrare che le proprietà di esistenza ed unicità della soluzione di (3.5) si preservano anche per il caso non lineare. Innanzitutto notiamo che per ogni $\alpha \in [0, 1]$, per ogni $p \in \mathbb{R}$ e per ogni $\theta > 0$ abbiamo

$$\mathcal{M}(\theta x) = \alpha \mathcal{M}(\theta x) + (1 - \alpha) T_p(\theta x) = \theta \mathcal{M}(x),$$

quindi se x risolve (3.5), allora ogni suo multiplo positivo lo risolve, allora come nel caso lineare, l'unicità è definita a meno di multipli scalari. Continuiamo introducendo il grafo di \mathcal{M} .

Definizione 3.12. *Data una matrice $M \in \mathbb{R}^{n \times n}$ ed un tensore cubico $T \in \mathbb{R}^{n \times n \times n}$, entrambi non negativi, definiamo il grafo di adiacenza $G_{\mathcal{M}}$ di \mathcal{M} (3.4) come una coppia $G_{\mathcal{M}} = (V, E_{\mathcal{M}})$, dove $V = \{1, \dots, n\}$ e, per ogni $i, j \in V$, $(i, j) \in E_{\mathcal{M}}$ se e solo se $(A_{\mathcal{M}})_{ij} = 1$, dove $A_{\mathcal{M}}$ è la matrice di adiacenza definita come*

$$(A_{\mathcal{M}})_{ij} = \begin{cases} 1 & \text{se } \alpha M_{ij} + (1 - \alpha) \sum_{k=1}^n (T_{ijk} + T_{ikj}) > 0, \\ 0 & \text{altrimenti.} \end{cases}$$

Ora possiamo enunciare il teorema che fornisce l'esistenza e l'unicità

Teorema 3.13. *Data una matrice non negativa $M \in \mathbb{R}^{n \times n}$ e un tensore non negativo $T \in \mathbb{R}^{n \times n \times n}$, sia \mathcal{M} definito come in (3.4) e sia $G_{\mathcal{M}}$ il suo grafo di adiacenza. Se $G_{\mathcal{M}}$ è fortemente connesso, allora*

- i) Esiste un unico (a meno di multipli) autovettore positivo di \mathcal{M} , equivalentemente, esiste un'unica soluzione di (3.5).*
- ii) L'autovettore positivo di \mathcal{M} è massimale, ovvero il suo autovalore è $\rho(\mathcal{M}) = \max\{|\lambda| : \mathcal{M}(x) = \lambda x\}$.*
- iii) Sia x un qualsiasi autovettore non negativo $\mathcal{M}(x) = \lambda x$ con almeno una entrata nulla, allora $\lambda < \rho(\mathcal{M})$.*

Se in aggiunta $G_{\mathcal{M}}$ è aperiodico, allora

- iv) Per ogni punto iniziale $x_0 > 0$, il metodo delle potenze non lineare*

$$\begin{cases} y_{k+1} = \alpha M x_k + (1 - \alpha) T_p(x_k) \\ x_{k+1} = \frac{y_{k+1}}{\|y_{k+1}\|} \end{cases}$$

converge all'autovettore positivo di \mathcal{M} .

La dimostrazione si trova nell'articolo [AHT20].

Notiamo che, poichè la mappa \mathcal{M} del Teorema 3.13 è definita per una matrice arbitraria non negativa M e un tensore non negativo T , il grafo $G_{\mathcal{M}}$ potrebbe essere un qualsiasi grafo diretto. Il prossimo lemma mostra che quando M e T sono definiti come nella Sezione 3.2.1 il grafo $G_{\mathcal{M}}$ coincide con le reti che abbiamo già visto. Quindi, per il caso non diretto e per ogni scelta della matrice e del tensore vista nella Sezione 3.2.1, il Teorema 3.13 si applica quando il grafo è connesso.

Lemma 3.14. *Sia $\alpha \neq 0$, e siano M e T una qualsiasi scelta fra quelle elencate nella Sezione 3.2.1. Allora M e $A_{\mathcal{M}}$ hanno la stessa sparsità, che equivale a $M_{ij} > 0$ se e solo se $(A_{\mathcal{M}})_{ij} = 1$.*

Dimostrazione. Se $(i, j) \in E$ è un arco nel grafo associato di M , ovvero $M_{ij} > 0$, allora $(A_{\mathcal{M}})_{ij} = 1$, dal momento che il tensore T ha tutte entrate non negative. Se invece $(i, j) \notin E$, allora dalle possibili definizioni del tensore T date nella Sezione 3.2.1, segue che $T_{ijk} = T_{ikj} = 0$ per ogni k . Quindi $(A_{\mathcal{M}})_{ij} = M_{ij} = 0$. Vice versa, se $(A_{\mathcal{M}})_{ij} = 0$, allora $\alpha M_{ij} + (1 - \alpha) \sum_{k=1}^n (T_{ijk} + T_{ikj}) = 0$. Dal momento che M e T sono non negative, avremo che entrambi i termini saranno zero, in particolare $M_{ij} = 0$. Se $(A_{\mathcal{M}})_{ij} = 1$, d'altra parte, abbiamo che $\alpha M_{ij} + (1 - \alpha) \sum_{k=1}^n (T_{ijk} + T_{ikj}) > 0$, e quindi almeno uno dei due termini dovrà essere positivo, ma dalle possibili definizioni di T date, si ha che T_{ijk} e T_{ikj} non può essere non zero a meno che $(i, j) \in E$, ovvero a meno che $M_{ij} > 0$. \square

3.2.4 Alcuni esempi e considerazioni associate

Lo scopo di questa sezione è mostrare come l'introdurre dell'eigenvector centrality del secondo ordine, in alcuni casi, può fare una notevole differenza nella classifica dell'importanza dei nodi di una rete. A tal fine consideriamo la rete riportata in Figura 3.1.

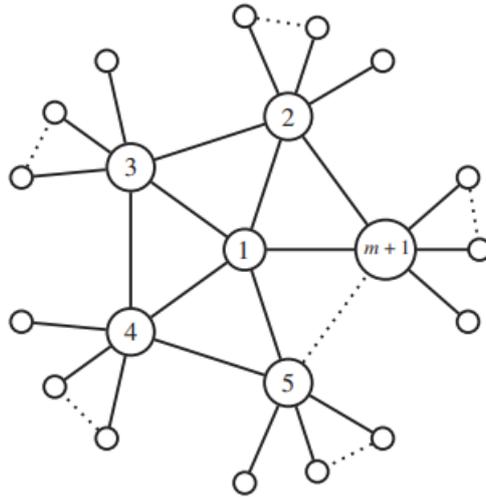


Figura 3.1: *Rappresentazione della rete analizzata. La rete è un grafo a ruota nel quale ciascuno degli m nodi è collegato a k foglie. Figura presa da [AHT20].*

Questo grafo ha tre tipi diversi di nodi: (i) Nodo 1, il centro della ruota, il quale ha grado m ed è connesso ad m nodi del secondo tipo. (ii) m nodi attaccati al nodo

1 e connessi ai loro vicini. Ogni nodo del secondo tipo è connesso a k nodi del terzo tipo. (iii) mk nodi foglia attaccati in gruppi di k agli m nodi del secondo tipo. La rete è costruita in modo tale che il nodo 1 è collegato a nodi importanti, e in aggiunta forma tanti triangoli. I nodi del secondo tipo, invece, sono coinvolti in solo due triangoli e sono collegati a nodi meno importanti. Se teniamo fisso m , ed aumentiamo il numero di nodi foglia k , ci aspettiamo che l'importanza dei nodi del secondo tipo prevalga quella del nodo 1. Mostriamo che questo accadrà per valori di k più grandi quando useremo l'analisi del secondo ordine. Più precisamente, posto $p = 1$, facciamo vedere come il nodo 1 è considerato più centrale per valori alti di k rispetto ai nodi del secondo tipo, quando usiamo l'analisi del secondo ordine, al posto del classico eigenvector centrality. La matrice di adiacenza A della rete, sarà della forma

$$A = \left[\begin{array}{c|cc} 0 & \mathbf{1}_m^T & 0 \dots 0 \\ \mathbf{1}_m & C & I_m \otimes \mathbf{1}_k^T \\ \hline 0 & & \\ \vdots & I_m \otimes \mathbf{1}_k & \\ 0 & & \end{array} \right] \quad \text{con } C = \begin{bmatrix} 0 & 1 & & 1 \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ 1 & & 1 & 0 \end{bmatrix} \in \mathbb{R}^{m \times m},$$

dove \otimes è il prodotto di Kronecker. L'autovettore $v = [x \ y \mathbf{1}_m^T \ z \mathbf{1}_{mk}^T]^T$ associato all'autovalore più grande $\lambda = 1 + \sqrt{1 + m + k}$ di A è tale che $\lambda x = my$, e si verifica che

$$x > y \text{ se e solo se } k < m(m - 3).$$

Ora ci spostiamo sull'analisi del secondo ordine. Prendiamo $T = T_B$, ovviamente avremo che $(T_B)_{ijk} = 0$ per ogni $i = m + 2, \dots, mk + m + 1$. Inoltre

$$(T_B)_{1jk} = \begin{cases} 1 & \text{se } j, k = 2, \dots, m + 1 \text{ sono tali che } (j, k) \in E, \\ 0 & \text{altrimenti,} \end{cases}$$

e per ogni $i = 2, \dots, m + 1$

$$(T_B)_{ijk} = \begin{cases} 1 & \text{se } j = 1 \text{ e } (i, k) \in E \text{ o } k = 1 \text{ e } (i, j) \in E, \\ 0 & \text{altrimenti.} \end{cases}$$

Usando (3.2) otteniamo

$$((T_B)_p(v))_1 = 2my, \quad ((T_B)_p(v))_2 = 4\mu_p(x, y), \quad ((T_B)_p(v))_{m+2} = 0,$$

dove $v = [x \ y \mathbf{1}_m^T \ z \mathbf{1}_{mk}^T]^T$ come prima. Nel complesso abbiamo quindi che l'equazione 3.5 può essere scritta come

$$\begin{cases} \lambda x = (2 - \alpha)my \\ \lambda y = \alpha(x + 2y + kz) + 4(1 - \alpha)\mu_p(x, y) \\ \lambda z = \alpha y. \end{cases}$$

Per $p = 1$ e $\alpha \in (0, 1]$ si ha

$$x > y \iff k < \frac{(2 - \alpha)}{\alpha^2} ((2 - \alpha)m^2 + (\alpha - 4)m). \quad (3.23)$$

Se prendiamo ora $T = T_C$ si nota che

$$((T_C)_p(v))_1 = \frac{2y}{m-1}, \quad ((T_C)_p(v))_2 = \frac{4\mu_p(x, y)}{(k+3)(k+2)}, \quad ((T_C)_p(v))_{m+2} = 0,$$

e quindi la soluzione di (3.5) dovrà soddisfare

$$\begin{cases} \lambda x = \left(\alpha m + \frac{2(1-\alpha)}{m-1}\right) y \\ \lambda y = \alpha(x + 2y + kz) + \frac{4(1-\alpha)}{(k+3)(k+2)} \mu_p(x, y) \\ \lambda z = \alpha y. \end{cases}$$

Dopo vari calcoli si ottiene

$$x > y \iff \alpha m + \frac{2(1-\alpha)}{m-1} > \lambda. \quad (3.24)$$

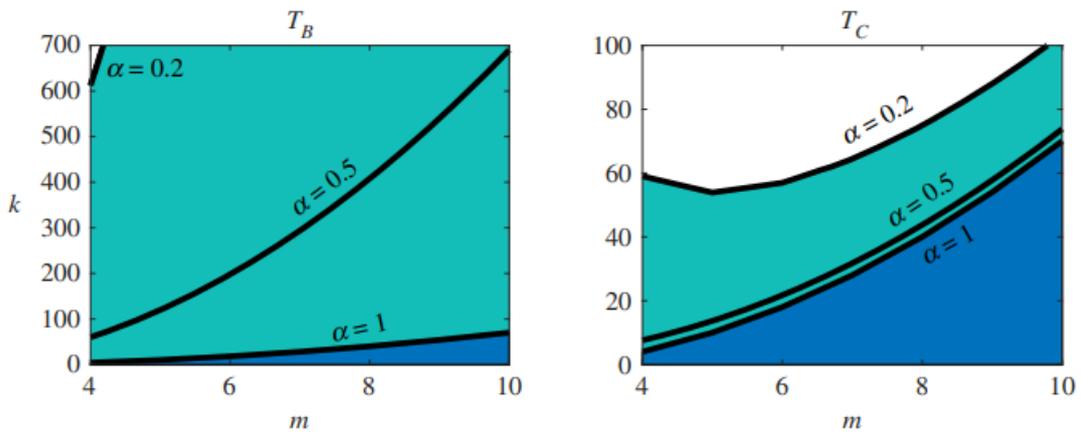


Figura 3.2: Valori di m e k , per i quali $(x > y)$ per differenti valori di α , $p = 1$ e tensori T_B e T_C . Figura presa da [AHT20].

Nella Figura 3.2 vediamo per $\alpha = 1$, $\alpha = 0,5$, $\alpha = 0,2$ e $p = 1$ come varia la centralità del nodo 1, rispetto ai nodi del secondo tipo. In particolare notiamo che per $\alpha = 1$ il

primo e secondo grafico sono ovviamente uguali perchè è un'analisi solo del primo ordine, per gli altri valori di α , in particolare per T_B , che come si vede dalla formula (3.23) la quale ha al denominatore α^2 , si deve aumentare di tantissimo k a parità di m per ottenere maggiore centralità per i nodi del secondo tipo.

Ora analizzeremo i risultati visti fino ad ora al variare di α per alcune reti reali. Consideriamo le seguenti reti

- KARATE: social network che rappresenta l'amicizia fra 34 membri di un club di karate all'US university.
- C.ELEGANS: rete neurale, qui useremo una versione non pesata e non diretta del network C.elegans fatto da Watts-Strogatz.
- ADJNOUN: rete basata sugli aggettivi e sostantivi più comuni nel libro David Copperfiel di Cahrls Dickens.
- CHESAPEAKE: rappresenta le interazioni nell'ecosistema della baia di Chesapeake.

Nomi	n	m	Δ
KARATE	34	78	45
C.ELEGANS	39	170	194
ADJNOUN	112	425	284
CHESAPEAKE	277	1918	2699

La tabella sopra ci fornisce il numero di nodi n , il numero di spigoli m , ed il numero di triangoli (non diretti) $\Delta = \text{traccia}(A^3)/6$ delle quattro reti che abbiamo preso in considerazione.

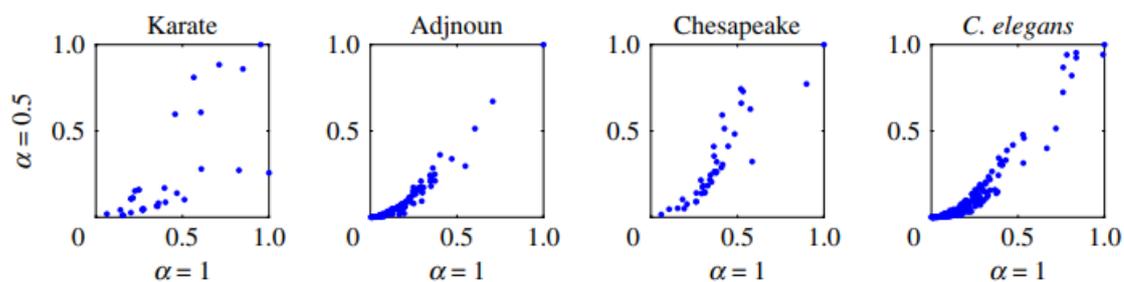


Figura 3.3: Grafico di dispersione per le soluzioni di (3.5) per $M = A$ e $T = T_B$. Importanza dei nodi per $\alpha = 1$ (ascisse) e per $\alpha = 0.5$ e $p = 1$ (ordinate). Figura presa da [AHT20].

La Figura 3.3 mostra che nel network KARATE c'è poca correlazione fra le due misu-
re. Nelle altre tre reti la differenza non è così evidente, in C.ELEGANS e CHESAPEAKE
i nodi più importanti differiscono leggermente per le due misure, mentre in ADJNOUN
i primi tre sono gli stessi.

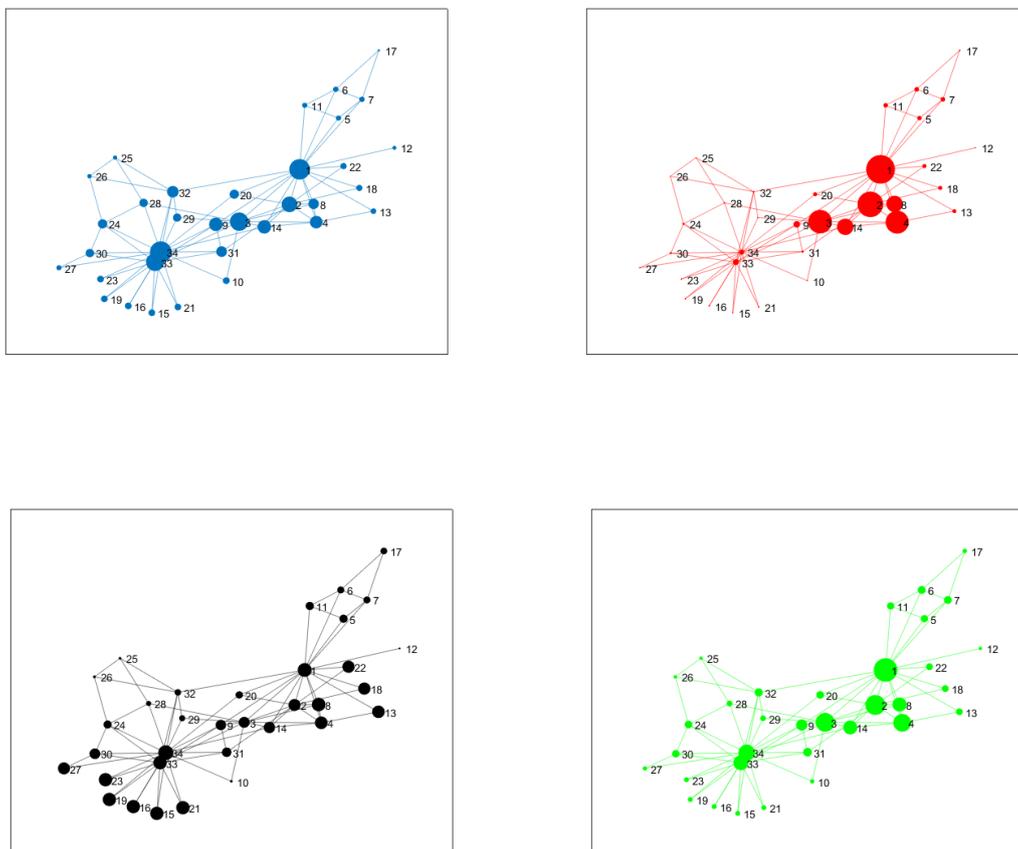


Figura 3.4: *Network KARATE: la grandezza dei nodi è proporzionale alla loro importanza. In alto a sinistra: analisi fatta per $\alpha = 1$; in alto a destra per T_B , in basso a sinistra per T_C , e in basso a destra per T_L , tutti e tre per $\alpha = 0.1$, $p = 1$.*

L'ultimo esempio che andremo a vedere consiste nell'analisi del network KARATE tramite Matlab. In particolare, analizzeremo le variazioni della centralità dei nodi usando il caso $\alpha = 1$ e il caso $\alpha = 0.1$, $p = 1$ per i tensori T_B , T_C e T_L . Come si può notare il caso del tensore T_L è molto simile al caso $\alpha = 1$, invece per i tensori T_B e T_C abbiamo risultati molto differenti. Infatti, dal grafo osserviamo che il tensore T_C in questo caso tende ad attribuire la stessa importanza a quasi tutti i nodi, anche a quelli ai quali era stata data più importanza nel caso $\alpha = 1$. Il tensore T_B al contrario attribuisce molta importanza

ai nodi che stanno sulla destra, mentre diminuisce drasticamente l'importanza dei nodi situati sulla sinistra.

Bibliografia

- [GV96] Gene H. Golub e Charles F. Van Loan. *Matrix computations*. Third. Johns Hopkins Studies in the Mathematical Sciences. Baltimore, MD: Johns Hopkins University Press, 1996, pp. xxx+698. ISBN: 0-8018-5413-X; 0-8018-5414-8.
- [SW05] Thomas Schank e Dorothea Wagner. “Finding, counting and listing all triangles in large graphs, an experimental study”. In: *International workshop on experimental and efficient algorithms*. Springer. 2005, pp. 606–609.
- [Hig08] Nicholas J. Higham. *Functions of Matrices - Theory and Computation*. Philadelphia, USA: SIAM, 2008.
- [Est12] Ernesto Estrada. *The structure of complex networks: theory and applications*. Oxford University Press, USA, 2012.
- [BGL15] Austin R Benson, David F Gleich e Jure Leskovec. “Tensor spectral clustering for partitioning higher-order network structures”. In: *Proceedings of the 2015 SIAM International Conference on Data Mining*. SIAM. 2015, pp. 118–126.
- [AHT20] Francesca Arrigo, Desmond J Higham e Francesco Tudisco. “A framework for second-order eigenvector centralities and clustering coefficients”. In: *Proceedings of the Royal Society A* 476.2236 (2020), p. 20190724.
- [Dav] Valeria Simoncini e Davide Palitta. *Dispense Calcolo Numerico*. URL: <https://www.dm.unibo.it/~simoncin/Mate/dispense051222.pdf>.
- [Sim] Valeria Simoncini. *Dispense DataMining*. URL: <http://www.dm.unibo.it/~simoncin/Mate/DataMining2223.pdf>.

Ringraziamenti

Giunto al termine di questa meravigliosa esperienza desidero ringraziare tutti coloro che hanno contribuito, direttamente o indirettamente, a renderla possibile e che mi sono stati vicino in questi tre anni.

Innanzitutto ringrazio la mia famiglia, in particolare i miei genitori, che mi hanno dato la possibilità di intraprendere questa esperienza, e mi sono stati vicini in momenti difficili. Ringrazio i miei compagni di corso con cui ho condiviso risate, sessioni e problemi. Ringrazio anche la mia ragazza e i miei amici che mi hanno aiutato senza dubbio in un modo o nell'altro in questi anni. Ringrazio infine la Prof.ssa Valeria Simoncini per aver accettato la mia richiesta di tesi.