

Calcolo efficiente di misure di centralità per Complex Networks

Luca Ferragina

Indice

1	Introduzione	2
2	Definizioni di base	4
2.1	Grafi e loro proprietà	4
3	Misure di centralità	6
3.1	Indice di Bonacich	6
3.2	Centralità esponenziale	7
4	Calcolo della misura di centralità esponenziale	10
4.1	Polinomi ortogonali	11
4.2	Calcolo di J_n	17
4.3	Formule di quadratura di Gauss	17
4.4	Stime a priori della centralità esponenziale	21
4.4.1	Stima con la regola di Gauss	22
4.4.2	Stime con la regola di Gauss-Radau	23
4.4.3	Stima con la regola di Gauss-Lobatto	24
5	Sperimentazione numerica	26
5.1	Grafi random	26
5.2	Analisi algoritmica	27

Capitolo 1

Introduzione

Con il termine *Complex Network* si intende un modello di una situazione reale effettuato attraverso una struttura di grafo, tipicamente composto da un grande numero di nodi e da poche connessioni. I campi di applicazione e gli esempi pratici sono innumerevoli e spaziano dai sistemi biologici a quelli fisici, dalle reti di trasporto alle reti sociali, dallo studio della diffusione delle epidemie all'analisi delle pagine web ([5], [6]). Questa grande duttilità di applicazione ha portato recentemente ad un veloce sviluppo della disciplina della *Network Analysis*, il cui scopo è la descrizione della struttura, della funzione e dell'evoluzione dei Complex Networks.

Uno dei problemi fondamentali in Network Analysis è determinare il "più importante" elemento in un data rete, dove l'interpretazione di cosa si intenda per "importante" cambia a seconda dell'applicazione. In questo ambito ricoprono un ruolo di primo piano le cosiddette *misure di centralità* che possiamo definire informalmente come misure di importanza dei nodi. Proprio queste misure di centralità sono l'argomento principale di questo elaborato in cui, dopo un breve riepilogo delle nozioni fondamentali riguardanti i grafi, e dopo una piccola esposizione di esempi reali di Complex Networks, andremo a formalizzare il concetto stesso di misura di centralità, esponendo all'inizio gli esempi di misura più ovvi, e andando successivamente a concentrarci su due misure: l'*indice di Bonacich* e, con maggior enfasi, la *centralità esponenziale*.

Quello che ci interessa è il ranking che le misure di centralità producono, ovvero come queste "classificano" i nodi in base alla loro importanza, pertanto generalmente i valori delle misure di centralità vengono approssimati con precisione non elevata. Il calcolo numerico dell'indice di Bonacich consisterà nella stima dell'autovettore relativo all'autovalore dominante della matrice di adiacenza del grafo ([2]), questo ci porterà ad applicare il metodo delle potenze, la cui convergenza è assicurata, sotto opportune ipotesi, dal teorema di Perron-Frobenius. Per stimare la centralità esponenziale sfrutteremo le relazioni fra le formule di quadratura di Gauss, il metodo di Lanczos per

la tridiagonalizzazione di matrici simmetriche e i metodi di costruzione di sequenze di polinomi ortogonali rispetto a prodotti scalari discreti e continui. In particolare, poche iterazioni dell'algoritmo costruito sono sufficienti a formulare delle stime a priori della centralità esponenziale ([4], [9], [1], [3]). Per concludere andremo ad implementare praticamente gli algoritmi di calcolo delle misure al fine di confrontarle e di valutarne l'efficacia.

Capitolo 2

Definizioni di base

2.1 Grafi e loro proprietà

In questa sezione forniremo una rapida panoramica dei concetti di base, delle definizioni e della terminologia di teoria dei grafi che saranno necessari in seguito.

Un *grafo* G è definito come una coppia $G = (V, E)$, dove V è un insieme finito i cui elementi sono detti *nodi* e E è un sottinsieme di $V \times V$ i cui elementi sono detti *archi*. Il numero $n = |V|$ di nodi è detto *ordine* del grafo, mentre $m = |E|$ indica il numero di archi. Noi considereremo solo grafi *non orientati* cioè tali che $\forall i, j \in V, (i, j) \in E$ se e solo se $(j, i) \in E$. In questo caso solitamente viene adottata la convenzione di indicare, con un piccolo abuso di notazione, con (u, v) entrambi gli archi che collegano u a v e v a u . In generale vale $0 \leq m \leq \binom{n}{2}$ poichè $\binom{n}{2} = O(n^2)$ è il numero di tutte le possibili coppie di nodi e quindi rappresenta il numero massimo di archi in un grafo: diremo che il grafo è *sparso* se $m = O(n)$ ed è *denso* se $m = O(n^2)$. Dato un nodo u , il *grado* di u è pari al numero di nodi ad esso collegati tramite un arco.

Successivamente sarà fondamentale il concetto di *cammino* da un nodo u a un nodo v , definito come una sequenza di nodi x_0, x_1, \dots, x_k tale che $x_0 = u$, $x_k = v$ e $(x_i, x_{i+1}) \in E$ per ogni $0 \leq i < k$. L'intero k è detto *lunghezza* del cammino. Un grafo si dice *connesso* se per ogni coppia di nodi esiste un cammino che li congiunge. Un *ciclo* è un cammino per cui vale $x_0 = x_k$, ovvero un cammino che alla fine torna nel nodo di partenza. Un cammino (o un ciclo) si dice *semplice* se non attraversa alcun nodo più di una volta, ossia se non esiste alcun ciclo annidato al suo interno. Un *cammino minimo* da u a v è caratterizzato dall'aver lunghezza minima tra tutti i cammini da u a v , la *distanza* tra due nodi è pari alla lunghezza di un cammino minimo che li congiunge, oppure a $+\infty$ nel caso in cui i due nodi siano disgiunti. Il *diametro* di un grafo è la distanza massima al variare delle coppie di nodi. Ad ogni grafo è associata in maniera naturale una matrice $A \in \mathbb{R}^{n \times n}$, detta

matrice di adiacenza, che verifica la proprietà

$$A = (a_{ij}); \quad a_{ij} = \begin{cases} 1 & \text{se } (i, j) \in E \\ 0 & \text{altrimenti} \end{cases} \quad i, j = 1, \dots, n.$$

dove poniamo per convenzione che $(i, i) \notin E$ e quindi $a_{ii} = 0$ per ogni i .

Si può osservare facilmente come la matrice A sia simmetrica se e solo se il grafo è non orientato. Un'osservazione fondamentale per i nostri scopi riguarda le potenze della matrice di adiacenza A : dato un intero k , l'elemento $[A^k]_{ij}$ rappresenta il numero di cammini nel grafo dal nodo i al nodo j di lunghezza esattamente k , in particolare l'elemento $[A^2]_{ii}$ indica i cicli di lunghezza 2 passanti per i e corrisponde esattamente al grado d_i del nodo i .

Capitolo 3

Misure di centralità

Come già detto le misure di centralità hanno il compito di classificare i nodi di un grafo in base alla loro importanza; in questa sezione andremo a presentare alcune di queste misure, non prima però di aver formalizzato il concetto stesso di misura di centralità.

Definizione 3.1. Sia \mathcal{G}_n l'insieme di tutti i grafi composti da n nodi e siano $G, G' \in \mathcal{G}_n$ tali che si possa ottenere uno dall'altro tramite una permutazione di indici, cioè tali che le corrispondenti matrici di adiacenza A_G e $A_{G'}$ soddisfino $A_{G'} = \Pi A_G \Pi^T$, con Π matrice di permutazione.

- Un *invariante* è una funzione $f : \mathcal{G}_n \rightarrow \mathbb{R}$ tale che $f(G) = f(G')$.
- Una *misura di centralità* è una funzione $c : \mathcal{G}_n \rightarrow \mathbb{R}^n$ tale che $\Pi c(G) = c(G')$.

Il più semplice esempio di misura di centralità di un grafo G con matrice di adiacenza A è dato dal vettore dei gradi $\mathbf{d} = (d_i)_{i=1, \dots, n}$, dove $d_i = (A\mathbf{1})_i$ e $\mathbf{1} = (1, \dots, 1)^T$. Come sappiamo dal celebre handshaking lemma di Eulero, si ha $\sum_{i=1}^n d_i = 2m$, questo valore, detto *volume* del grafo, è un invariante. La misura grado in pratica classifica i nodi in base al numero di nodi ad esso vicini, si tratta quindi di una misura prettamente locale che in alcune applicazioni può risultare inespressiva. Occorre in questi casi ricorrere a misure che in qualche modo tengano conto non solo dei legami fra nodi dati dagli archi, ma anche da quelli dati da cammini più lunghi.

3.1 Indice di Bonacich

L'idea alla base di questa misura, ideata dal sociologo americano Bonacich, è che un nodo è importante se è connesso ad altri nodi importanti. Questa definizione può essere formalizzata ponendo l'indice di ogni nodo i proporzionale alla somma degli indici di tutti i nodi j connessi ad i tramite un

arco:

$$\lambda b_i = \sum_{j=1}^n A_{ji} b_j \quad (3.1)$$

Quindi il vettore \mathbf{b} degli indici di Bonacich soddisfa la relazione $A^T \mathbf{b} = \lambda \mathbf{b}$, cioè è un autovettore di A^T .

Fra tutte le possibili soluzioni dell'equazione (3.1), la scelta più comoda è quella di prendere l'autovettore relativo all'autovalore di modulo massimo per sfruttare le proprietà garantite dal seguente teorema:

Teorema 3.1 (di Perron-Frobenius). *Sia $A \in \mathbb{R}^{n \times n}$ irriducibile e a valori non negativi. Allora:*

1. *il raggio spettrale $\rho(A)$ è autovalore di A*
2. *esiste $\mathbf{x} \in \mathbb{R}^n$ a valori positivi tale che $A\mathbf{x} = \rho(A)\mathbf{x}$*
3. *$\rho(A)$ è un autovalore semplice.*

La teoria di Perron-Frobenius ci assicura che il vettore \mathbf{b} sia definito in modo unico, a meno di prodotto per uno scalare, e che ad ogni nodo sia associato un indice b_i strettamente positivo. In questo caso è ben noto che l'autovettore \mathbf{b} può essere calcolato tramite il metodo delle potenze e soddisfa la relazione

$$\mathbf{b} = \lim_{k \rightarrow \infty} \frac{(A^T)^k \mathbf{1}}{\|(A^T)^k \mathbf{1}\|} \quad (3.2)$$

In particolare si può osservare che, se consideriamo il rapporto $\frac{b_i}{b_j}$ tra due indici del vettore \mathbf{b} , questo sarà pari a

$$\frac{b_i}{b_j} = \lim_{k \rightarrow \infty} \frac{[(A^T)^k \mathbf{1}]_i}{[(A^T)^k \mathbf{1}]_j}$$

In questa relazione l'espressione a destra esprime il rapporto fra il numero di cammini nel grafo di lunghezza k che partono dal nodo i e il numero di cammini di lunghezza k che partono dal nodo j . Se, per k che tende all'infinito, tale rapporto tende a un numero maggiore di uno, vuol dire quindi che dal nodo i hanno origine più cammini "lunghi" che dal nodo j . Questa osservazione ci fa capire come l'indice di Bonacich dia più peso ai cammini di lunghezza maggiore, si tratta quindi di una misura globale che si contrappone fortemente alla misura grado.

3.2 Centralità esponenziale

La misura di *centralità esponenziale* rappresenta la parte più sostanziale di questo elaborato e si pone a metà strada, in un certo senso che verrà successivamente chiarito, tra le altre due misure appena citate, catturando

aspetti sia globali che locali. Per definirla rigorosamente abbiamo bisogno di un po' di lavoro preliminare.

Definizione 3.2 (Funzioni di matrici diagonalizzabili). Sia $A \in \mathbb{R}^{n \times n}$ una matrice diagonalizzabile $A = Q\Lambda Q^{-1}$, con $\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ e sia $f(X)$ una funzione a valori reali definita sullo spettro di A . Allora:

$$f(A) := Q \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} Q^{-1}$$

La definizione precedente può essere estesa al caso generale in cui la matrice non è necessariamente diagonalizzabile usando la sua forma di Jordan, tuttavia a noi basta questo caso particolare in quanto avremo a che fare solo con matrici simmetriche.

Definizione 3.3 (Centralità esponenziale). Il vettore $\mathbf{c}(\beta) = (c_i(\beta))_{i=1, \dots, n}$ della misura di *centralità esponenziale* è dato da $c_i(\beta) = [e^{\beta A}]_{ii}$ con β numero reale positivo.

Osservazione 3.1. L'indice $c_i(\beta)$ conta il numero di cammini che iniziano e finiscono nel nodo i e ogni cammino di lunghezza k viene pesato per un fattore di $\frac{\beta^k}{k!}$ in modo che i cammini più lunghi siano penalizzati. Infatti:

$$\begin{aligned} c_i(\beta) - 1 &= \left[I + \beta A + \frac{\beta^2}{2} A^2 + \frac{\beta^3}{3!} A^3 + \dots \right] (i, i) - 1 \\ &= \left[1 + \beta A(i, i) + \frac{\beta^2}{2} A^2(i, i) + \frac{\beta^3}{3!} A^3(i, i) + \dots \right] - 1 \quad (3.3) \\ &= \frac{\beta^2}{2} A^2(i, i) + \frac{\beta^3}{3!} A^3(i, i) + \dots \end{aligned}$$

Un aspetto interessante di quest'ultima misura è il suo rapporto con le altre due precedentemente presentate: il prossimo teorema mostra infatti quanto la centralità esponenziale possa essere resa simile sia alla misura grado che all'indice di Bonacich, modificando il valore di β , in termini di ranking, ovvero di nodi ordinati secondo la loro misura.

Teorema 3.2. *Sia G un grafo connesso non orientato e sia $\mathbf{c}(\beta)$ il vettore della centralità esponenziale. Allora:*

- (i) *per $\beta \rightarrow 0$, il ranking prodotto da $\mathbf{c}(\beta)$ tende a quello prodotto da \mathbf{d} , il vettore della misura grado.*

(ii) per $\beta \rightarrow \infty$, il ranking prodotto da $\mathbf{c}(\beta)$ tende a quello prodotto da \mathbf{b} , il vettore dell'indice di Bonacich.

Dimostrazione. Per provare (i), consideriamo l'espansione di Taylor di $c_i(\beta)$:

$$c_i(\beta) = [e^{\beta A}]_{ii} = 1 + 0 + \frac{\beta^2}{2} d_i + \frac{\beta^3}{3!} [A^3]_{ii} + \dots$$

Sia $\Phi(\beta) = \frac{2}{\beta^2}(\mathbf{c} - \mathbf{1})$. Il ranking prodotto da $\Phi(\beta)$ sarà lo stesso di quello prodotto da $\mathbf{c}(\beta)$, in quanto l'indice di ogni nodo è stato shiftato e scalato nello stesso modo. La componente i -esima di $\Phi(\beta)$ è data da:

$$\Phi_i(\beta) = \frac{2}{\beta^2}(c_i(\beta) - 1) = d_i + \frac{2\beta}{3!} [A^3]_{ii} + \frac{2\beta^2}{4!} [A^4]_{ii} + \dots$$

che tende a d_i per $\beta \rightarrow 0$. Quindi, per $\beta \rightarrow 0$, il ranking prodotto dalla centralità esponenziale tende a quello prodotto dalla misura grado.

Per provare (ii), consideriamo l'espansione di $c_i(\beta)$ in termini di autovettori e autovalori di A :

$$c_i(\beta) = \sum_{k=1}^n e^{\beta \lambda_k} q_k(i)^2 = e^{\beta \lambda_1} q_1(i)^2 + \sum_{k=2}^n e^{\beta \lambda_k} q_k(i)^2$$

Sia $\Psi(\beta) = \frac{1}{e^{\beta \lambda_1}} \mathbf{c}(\beta)$. Come prima, il ranking prodotto da $\Psi(\beta)$ è lo stesso di quello prodotto da $\mathbf{c}(\beta)$, poichè gli indici di tutti i nodi sono stati scalati allo stesso modo. La componente i -esima di $\Psi(\beta)$ è data da:

$$\Psi_i(\beta) = q_1(i)^2 + \sum_{k=2}^n e^{\beta(\lambda_k - \lambda_1)} q_k(i)^2$$

Poichè $\lambda_k < \lambda_1$ per $2 \leq k \leq n$, per $\beta \rightarrow \infty$, $\Psi_i(\beta) \rightarrow q_1(i)^2$. Dal teorema di Perron-Frobenius si ha che q_1 ha elementi strettamente positivi, quindi il ranking prodotto da $q_1(i)^2$ è lo stesso di quello prodotto da $q_1(i) = \mathbf{b}$. Per concludere abbiamo ottenuto che per $\beta \rightarrow \infty$, il ranking prodotto da $\mathbf{c}(\beta)$ coincide con quello prodotto da \mathbf{b} . □

Come abbiamo visto tutte e tre le misure esaminate si basano sul concetto di cammino nel grafo: la misura grado esaminava i cammini di lunghezza 1 (gli archi) e aveva un significato locale, l'indice di Bonacich esprimeva proprietà più globali. Il teorema (3.2) ci mostra come la misura esponenziale inglobi entrambi gli aspetti: per β che tende a 0 i pesi corrispondenti alle lunghezze k maggiori tendono a 0 più velocemente e quindi i cammini più corti diventano più importanti nella misura, per β che tende all'infinito invece i pesi dei cammini di lunghezza maggiore crescono più rapidamente e quindi i cammini corti assumono un ruolo secondario. Grazie a questa proprietà la misura di centralità esponenziale può essere adattata, modificando il solo parametro β , alle varie applicazioni pratiche in modo da renderla più sensibile alla struttura globale o locale del grafo.

Capitolo 4

Calcolo della misura di centralità esponenziale

In questa sezione andremo a sviluppare un algoritmo che ci permetta di calcolare in maniera efficiente la centralità esponenziale. I grafi che derivano da situazioni reali, infatti, sono spesso di dimensioni così grandi che il calcolo diretto dell'esponenziale della matrice di adiacenza risulta troppo lento. Poiché il nostro scopo è velocizzare il calcolo della misura piuttosto che valutarne l'efficacia descrittiva, porremo d'ora in avanti il parametro $\beta = 1$, in modo da facilitare i calcoli; ovviamente tutti i passaggi che faremo valgono per un qualsiasi valore di β . Generalizzando il problema, quello che dobbiamo calcolare è un'espressione della forma:

$$u^T f(A)u$$

dove u è un vettore fissato e f è una funzione a valori reali definita sullo spettro di A . Osserviamo come la misura di centralità esponenziale $c_i(1) = c_i$ si ottenga ponendo $u = e_i$ e $f(x) = e^x$.

Sfruttando il fatto che la nostra matrice è simmetrica, scriviamo $A = Q\Lambda Q^T$ e, come abbiamo già visto vale $f(A) = Qf(\Lambda)Q^T$. Da quest'ultima relazione segue che

$$u^T f(A)u = u^T Qf(\Lambda)Q^T u = v^T f(\Lambda)v = \sum_{i=1}^n f(\lambda_i)v_i^2$$

dove $v = Q^T u$.

Questa somma può essere vista come un integrale di Riemann-Stieltjes

$$I[f] = u^T f(A)u = \int_a^b f(\lambda)d\mu(\lambda)$$

dove la misura μ è una funzione crescente, positiva, costante a tratti, definita da

$$\mu(\lambda) = \begin{cases} 0 & \text{se } \lambda < a = \lambda_1 \\ \sum_{j=1}^i v_j^2 & \text{se } \lambda_i \leq \lambda \leq \lambda_{i+1} \\ \sum_{j=1}^n v_j^2 & \text{se } b = \lambda_n \leq \lambda \end{cases}$$

4.1 Polinomi ortogonali

Data una misura μ , con $\mu(x) : [a, b] \rightarrow \mathbb{R}$ funzione a valori positivi tale che per ogni polinomio $f(x)$ esista e sia finito $\int_a^b f(x)d\mu(x)$, si può definire un prodotto scalare sullo spazio dei polinomi tramite la seguente funzione:

$$\langle f, g \rangle := \int_a^b f(x)g(x)d\mu(x)$$

Analogamente la seguente funzione definisce una norma:

$$\|f(x)\| := \sqrt{\int_a^b f(x)^2 d\mu(x)}$$

Osservazione 4.1. Il prodotto scalare appena definito verifica la proprietà non banale $\langle xf(x), g(x) \rangle = \langle f(x), xg(x) \rangle$.

Definizione 4.1. Data una successione di polinomi $(p_i(x))_{i=0,1,\dots}$ questi ultimi si dicono *ortogonali* se $\deg(p_i) = i$ e $\langle p_i, p_j \rangle = 0$ se $i \neq j$. I polinomi $\tilde{p}_i := \frac{p_i}{\sqrt{h_i}}$, con $h_i = \langle p_i, p_i \rangle$ si dicono *ortonormali*.

I polinomi ortogonali $p_0(x), \dots, p_n(x)$ sono linearmente indipendenti, quindi costituiscono una base dello spazio dei polinomi di grado al più n , da questo segue che se $q(x)$ è un polinomio di grado k , allora per ogni $i > k$ vale $\langle p_i, q \rangle = 0$. Questa proprietà ci permette di dimostrare i due seguenti risultati fondamentali.

Teorema 4.1. *Le radici dei polinomi ortogonali rispetto al prodotto scalare $\langle f, g \rangle = \int_a^b f(x)g(x)d\mu(x)$ sono reali, semplici e contenute in (a, b) .*

Dimostrazione. Per prima cosa osserviamo che, per $n \geq 1$, si ha

$$0 = \langle p_n(x), p_0(x) \rangle = p_0 \int_a^b p_n(x)d\mu(x)$$

dove p_0 è una costante non nulla.

Questo ci dice che $p_n(x)$ sicuramente cambia segno almeno una volta in (a, b) , quindi avrà almeno una radice di molteplicità dispari. Siano t_1, \dots, t_r

le radici distinte di molteplicità dispari di $p_n(x)$ in (a, b) . Allora possiamo scrivere

$$p_n(x) = \prod_{k=1}^r (x - t_k)^{m_k} q(x)$$

dove $q(x)$ è un polinomio di grado $n - \sum_{k=1}^r m_k$ che non cambia segno in (a, b) . Allora anche il polinomio $h(x)$ definito da

$$h(x) = p_n(x) \prod_{k=1}^r (x - t_k) = \prod_{k=1}^r (x - t_k)^{m_k+1} q(x)$$

non cambia segno in (a, b) . Se ora supponiamo per assurdo $r < n$, per le proprietà di $p_n(x)$, otteniamo

$$0 = \langle p_n(x), \prod_{k=1}^r (x - t_k) \rangle = \int_a^b h(x) d\mu(x)$$

che contraddice quanto visto precedentemente riguardo $h(x)$. Perciò deve necessariamente valere $r = n$. \square

Teorema 4.2 (Ricorrenza a tre termini). *Siano $(p_i)_{i=0,1,\dots}$ polinomi ortogonali rispetto ad un prodotto scalare. Denotiamo con a_i e b_i i coefficienti di grado rispettivamente x^i e x^{i-1} di $p_i(x)$. Allora vale che:*

$$p_{i+1} = (xA_{i+1} - B_{i+1})p_i(x) - C_i p_{i-1}(x) \quad \text{per } i \geq 1$$

Dove, posto $h_i = \langle p_i, p_i \rangle$, i coefficienti A_i, B_i, C_i sono dati da:

- $A_{i+1} = \frac{h_{i+1}}{\langle xp_i, p_{i+1} \rangle} = \frac{a_{i+1}}{a_i}$
- $B_{i+1} = A_{i+1} \frac{\langle xp_i, p_i \rangle}{h_i} = -\frac{a_{i+1}}{a_i} \left(\frac{b_{i+1}}{a_{i+1}} - \frac{b_i}{a_i} \right)$
- $C_i = A_{i+1} \frac{\langle xp_i, p_{i-1} \rangle}{h_{i-1}} = \frac{A_{i+1}}{A_i} \frac{h_i}{h_{i-1}} = \frac{a_{i+1} a_{i-1} h_i}{a_i^2 h_{i-1}}$

Dimostrazione. Il polinomio $xp_i(x)$ è di grado $i + 1$ per cui p_0, \dots, p_i, xp_i sono linearmente indipendenti. Esistono allora degli scalari $\alpha_i \in \mathbb{R}$ tali che

$$p_{i+1} = \alpha_0 p_0 + \dots + \alpha_i p_i + \alpha_{i+1} xp_i \tag{4.1}$$

Moltiplicando scalarmente entrambi i membri della precedente espressione per $p_j(x)$ con $j \neq i + 1$, si ottiene:

$$0 = \langle p_j, p_{i+1} \rangle = \alpha_j \langle p_j, p_j \rangle + \alpha_{i+1} \langle p_j, xp_i \rangle = \alpha_j \langle p_j, p_j \rangle + \alpha_{i+1} \langle xp_j, p_i \rangle$$

Se $j < i - 1$, allora il polinomio xp_j ha grado minore di i e per quanto visto in precedenza vale $\langle xp_j, p_i \rangle = 0$ e di conseguenza $\alpha_j = 0$ per $j = 0, \dots, i - 2$. L'espressione (4.1) in questo modo diventa

$$p_{i+1} = \alpha_{i+1} xp_i + \alpha_i p_i + \alpha_{i-1} p_{i-1} \tag{4.2}$$

Vale quindi il risultato del teorema con $A_{i+1} = \alpha_{i+1}$, $B_{i+1} = -\alpha_i$ e $C_i = -\alpha_{i-1}$.

Per ricavare espressioni esplicite per le costanti A_i , B_i e C_i moltiplichiamo scalarmente entrambi i membri di (4.2) rispettivamente per p_{i+1} , p_i e p_{i-1} , ottenendo

$$\langle p_{i+1}, p_{i+1} \rangle = A_{i+1} \langle xp_i, p_{i+1} \rangle \quad (4.3)$$

$$A_{i+1} \langle xp_i, p_i \rangle - B_{i+1} \langle p_i, p_i \rangle = 0 \quad (4.4)$$

$$A_{i+1} \langle p_{i-1}, xp_i \rangle - C_i \langle p_{i-1}, p_{i-1} \rangle = 0 \quad (4.5)$$

Sfruttando il fatto che il prodotto scalare del polinomio p_i per tutti i polinomi di grado minore di i è nullo, possiamo sviluppare i prodotti scalari presenti nelle tre equazioni precedenti.

$$\langle xp_i, p_{i+1} \rangle = a_i \langle x^{i+1}, p_{i+1} \rangle = \frac{a_i}{a_{i+1}} \langle a_{i+1} x^{i+1}, p_{i+1} \rangle = \frac{a_i}{a_{i+1}} h_{i+1}$$

$$\langle p_{i-1}, xp_i \rangle = \langle xp_{i-1}, p_i \rangle = \frac{a_{i-1}}{a_i} h_i$$

$$\begin{aligned} \langle xp_i, p_i \rangle &= a_i \langle x^{i+1}, p_i \rangle + b_i \langle x^i, p_i \rangle \\ &= \frac{a_i}{a_{i+1}} \langle a_{i+1} x^{i+1}, p_i \rangle + \frac{b_i}{a_i} h_i \\ &= \frac{a_i}{a_{i+1}} \langle p_{i+1}, p_i \rangle - \frac{a_i}{a_{i+1}} \langle b_{i+1} x^i, p_i \rangle + \frac{b_i}{a_i} h_i \\ &= \frac{b_i}{a_i} h_i - \frac{b_{i+1}}{a_{i+1}} \langle a_i x^i, p_i \rangle \\ &= h_i \left(\frac{b_i}{a_i} - \frac{b_{i+1}}{a_{i+1}} \right) \end{aligned}$$

Otteniamo le relazioni scritte nella tesi sostituendo le espressioni appena trovate nelle (4.3) (4.4) (4.5). \square

Nel teorema appena visto i polinomi sono definiti a meno di una costante moltiplicativa, possiamo quindi imporre che questi siano monici, cioè $a_i = 1$ per ogni $i \geq 1$. In questo modo i coefficienti della ricorrenza diventano:

- $A_{i+1} = 1$
- $B_{i+1} = \frac{\langle xp_i, p_i \rangle}{h_i}$
- $C_i = \frac{h_i}{h_{i-1}} > 0$

Se i polinomi $p_i(x)$ sono monici, allora possiamo normalizzarli dividendoli per la loro norma, cioè possiamo considerare i polinomi ortonormali $\tilde{p}_i(x) = \frac{p_i(x)}{\sqrt{h_i}}$. Vorremmo trovare una relazione di ricorrenza a tre termini che generi questi polinomi ortonormali, a tale scopo scriviamo $p_i(x) = \sqrt{h_i}\tilde{p}_i(x)$ nella ricorrenza a tre termini per polinomi monici e otteniamo

$$\sqrt{h_i}\tilde{p}_i(x) = (x - B_i)\sqrt{h_{i-1}}\tilde{p}_{i-1}(x) - C_{i-1}\sqrt{h_{i-2}}\tilde{p}_{i-2}(x)$$

Ricordando che $C_i = \frac{h_i}{h_{i-1}}$ e dividendo per $\sqrt{h_{i-1}}$ otteniamo quanto voluto

$$\sqrt{C_i}\tilde{p}_i(x) = (x - B_i)\tilde{p}_{i-1}(x) - \sqrt{C_{i-1}}\tilde{p}_{i-2}(x) \quad (4.6)$$

Posto per semplicità di notazione $B_1 = -b_1$, sia T_n la matrice tridiagonale definita da

$$T_n := \begin{pmatrix} B_1 & 1 & & & \\ C_1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & C_{n-1} & 1 \\ & & & & B_n \end{pmatrix}$$

Indicando con $\mathbf{p}(x) = (p_0(x), \dots, p_{n-1}(x))^T$ il vettore che ha per componenti i primi n polinomi ortogonali monici, si può esprimere la relazione di ricorrenza a tre termini tramite la matrice T_n come

$$(xI_n - T_n)\mathbf{p}(x) = \mathbf{e}_n p_n(x).$$

Da questo possiamo concludere che gli autovalori di T_n sono esattamente le radici t_1, \dots, t_n di $p_n(x)$ e gli autovettori hanno per componenti i valori dei polinomi ortogonali calcolati negli zeri di $p_n(x)$.

Siano $\tilde{\mathbf{p}}(x) = (\tilde{p}_0(x), \dots, \tilde{p}_{n-1}(x))^T$ il vettore dei polinomi ortonormali e D la seguente matrice diagonale

$$D := \begin{pmatrix} d_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}$$

dove $d_i = \sqrt{h_{i-1}}$ vale che:

$$J_n := D^{-1}T_n D = \begin{pmatrix} B_1 & \sqrt{C_1} & & & \\ \sqrt{C_1} & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & \sqrt{C_{n-1}} & \\ & & & & B_n \end{pmatrix}$$

Con queste notazioni possiamo scrivere in termini matriciali anche la (4.6) come

$$(xI - J_n)\tilde{\mathbf{p}}(x) = \sqrt{C_n}\tilde{p}_n\mathbf{e}_n$$

Gli autovettori di J_n sono della forma $D^{-1}\mathbf{p}(x)$, dalla loro ortogonalità segue che:

- $\sum_{k=0}^{n-1} \frac{p_k(t_i)p_k(t_j)}{h_k} = 0$ se $i \neq j$
- $\sum_{k=0}^{n-1} \frac{p_k(t_i)^2}{h_k} \neq 0$

Grazie a questa proprietà, se consideriamo la matrice $V = (v_{ij})_{i,j=1,\dots,n}$ definita da $v_{ij} = \frac{p_{i-1}(t_j)}{\sqrt{h_{i-1}}}$ si vede immediatamente che $V^T V$ è una matrice diagonale che è possibile rendere uguale alla matrice identità tramite una semplice scalatura delle colonne di V . Più precisamente, se chiamiamo

$$w_j = \frac{1}{\sum_{i=0}^{n-1} \frac{p_i(t_j)}{h_i}} \quad \tilde{V} = V \begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_n \end{pmatrix}$$

si ha che la matrice \tilde{V} verifica la proprietà $\tilde{V}^T \tilde{V} = I$ cioè è ortogonale e vale anche $\tilde{V} \tilde{V}^T = I$. Questa ultima relazione può essere scritta per esteso come:

$$\sum_{k=1}^n \tilde{v}_{ik} \tilde{v}_{jk} = \sum_{k=1}^n \frac{p_i(t_k)}{\sqrt{h_i}} \frac{p_j(t_k)}{\sqrt{h_j}} w_k = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

In pratica abbiamo appena dimostrato come i polinomi $\frac{p_0(x)}{\sqrt{h_0}}, \dots, \frac{p_{n-1}(x)}{\sqrt{h_{n-1}}}$ siano ortonormali rispetto al prodotto scalare

$$\langle f(x), g(x) \rangle' := \sum_{k=1}^n f(t_k) g(t_k) w_k$$

Allora si ha che i due prodotti scalari $\langle \cdot, \cdot \rangle$ e $\langle \cdot, \cdot \rangle'$ coincidono sulla base dello spazio dei polinomi di grado minore di n , costituita da tali polinomi e quindi coincidono su tutto lo spazio, cioè, se $\deg f, \deg g < n$, vale che

$$\int_a^b f(x) g(x) d\mu(x) = \sum_{k=1}^n f(t_k) g(t_k) w_k \quad (4.7)$$

Grazie a questo risultato possiamo dimostrare il seguente importante teorema.

Teorema 4.3. *Sia $h(x)$ un polinomio tale che $\deg h < 2n$. Allora vale che*

$$\int_a^b h(x) d\mu(x) = \sum_{k=1}^n h(t_k) w_k$$

Dimostrazione. Applicando la divisione euclidea a $h(x)$ e $p_n(x)$ possiamo scrivere

$$h(x) = p_n(x)q(x) + r(x)$$

dove sia il quoziente $q(x)$ che il resto $r(x)$ hanno grado minore di n e vale inoltre che $h(t_j) = r(t_j)$, poichè $p_n(t_j) = 0$. Integrando entrambi i membri dell'equazione precedente si ottiene

$$\int_a^b h(x) d\mu(x) = \langle p_n(x), q(x) \rangle + \int_a^b r(x) d\mu(x)$$

Il primo addendo a destra è nullo dal momento che q è di grado minore di n e quindi ortogonale a $p_n(x)$, per il secondo addendo possiamo applicare l'equazione (4.7) con $f(x) = 1$ e $g(x) = r(x)$ e otteniamo

$$\int_a^b h(x) d\mu(x) = \sum_{k=1}^n r(t_k) w_k = \sum_{k=1}^n h(t_k) w_k$$

□

Il seguente lemma, di immediata dimostrazione, ci fornisce una comoda caratterizzazione dei coefficienti w_k che sarà molto utile successivamente.

Lemma 4.1. *I coefficienti w_k sono i quadrati delle prime componenti degli autovettori normalizzati di J_n .*

Dimostrazione. Come abbiamo visto $\tilde{\mathbf{p}}(t_j)$ è autovettore di J_n relativo a t_j . Sia ora Σ la matrice degli autovettori normalizzati di J_n , e sia $\sigma^{(j)}$ la sua j -esima colonna, cioè $\|\sigma^{(j)}\| = 1$, allora esiste uno scalare s tale che $\tilde{\mathbf{p}}(x_j) = s\sigma^{(j)}$. Vale che

$$1 = w_j \sum_{i=0}^{n-1} \frac{p_i(t_j)}{h_i} = w_j \tilde{\mathbf{p}}(t_j)^T \tilde{\mathbf{p}}(t_j) = w_j s^2 \|\sigma^{(j)}\|^2 = w_j s^2$$

Per calcolare s confronto le prime componenti di $\sigma^{(j)}$ e $\tilde{\mathbf{p}}(t_j)$:

$$s = \frac{p_0(t_j)}{\sqrt{h_0} \sigma_{1j}} = \frac{1}{\sqrt{h_0} \sigma_{1j}}$$

poichè $p_0(x) = 1$. Sostituendo nella relazione precedente

$$w_j = \frac{1}{s^2} = (\sigma_{1j})^2 h_0 = (\sigma_{1j})^2 \int_a^b d\mu(x)$$

Da questo concludiamo osservando che nel nostro caso l'integrale della misura è esattamente 1. □

4.2 Calcolo di J_n

In questa sezione presenteremo il metodo di Lanczos per la tridiagonalizzazione di matrici simmetriche e dimostreremo che questo genera i coefficienti della matrice J_n .

Definiamo $\mathbf{v}^{(-1)} = 0$ e $\mathbf{v}^{(0)} = \mathbf{u}$, ricordando che nel nostro caso $\mathbf{u} = \mathbf{e}_i$, il metodo di Lanczos è dato dalle seguenti relazioni

$$\begin{cases} B_i = (\mathbf{v}^{(i-1)})^T A \mathbf{v}^{(i-1)} \\ \sqrt{C_i} \mathbf{v}^{(i)} = \mathbf{r}^{(i)} = (A - B_i I) \mathbf{v}^{(i-1)} - \sqrt{C_{i-1}} \mathbf{v}^{(i-2)} \\ \sqrt{C_i} = \|\mathbf{r}^{(i)}\| \end{cases}$$

Teorema 4.4. *Il vettore $\mathbf{v}^{(i)}$ è dato da*

$$\mathbf{v}^{(i)} = \tilde{p}_i(A) \mathbf{v}^{(0)}$$

dove \tilde{p}_i è il polinomio ortonormale di grado i definito dalla relazione di ricorrenza a tre termini presentata nella sezione precedente.

Dimostrazione. Dimostriamo l'enunciato per induzione su i . Poichè $\mathbf{v}^{(-1)} = 0$ e $\sqrt{C_1} \tilde{p}_1(x) = x - B_1$ vale

$$\sqrt{C_1} \mathbf{v}^{(1)} = (A - B_1 I) \mathbf{v}^{(0)} = \sqrt{C_1} \tilde{p}_1(A) \mathbf{v}^{(0)}$$

Usando l'ipotesi induttiva per i e $i - 1$ e considerando la ricorrenza a tre termini per polinomi ortonormali, otteniamo

$$\begin{aligned} \sqrt{C_{i+1}} \mathbf{v}^{(i+1)} &= (A - B_{i+1} I) \mathbf{v}^{(i)} - \sqrt{C_i} \mathbf{v}^{(i-1)} \\ &= (A - B_{i+1} I) \tilde{p}_i(A) \mathbf{v}^{(0)} - \sqrt{C_i} \tilde{p}_{i-1}(A) \mathbf{v}^{(0)} \\ &= [(A - B_{i+1} I) \tilde{p}_i(A) - \sqrt{C_i} \tilde{p}_{i-1}(A)] \mathbf{v}^{(0)} \\ &= \sqrt{C_{i+1}} \tilde{p}_{i+1}(A) \mathbf{v}^{(0)} \end{aligned}$$

Da cui segue la tesi dividendo per $\sqrt{C_{i+1}}$ □

Da questo teorema segue immediatamente che i coefficienti B_i e C_i del metodo di Lanczos sono gli stessi della matrice J_n .

4.3 Formule di quadratura di Gauss

La formula generale per approssimare un integrale di Riemann-Stieltjes è data da

$$I[f] = \int_a^b f(x) d\mu(x) = \sum_{j=1}^n f(t_j) w_j + \sum_{k=1}^m f(z_k) v_k + R[f] \quad (4.8)$$

dove i pesi w_j e v_k e i nodi t_j sono inizialmente ignoti mentre i nodi z_k sono stabiliti in partenza.

In precedenza abbiamo visto che se nella formula di quadratura poniamo $m = 0$ e prendiamo come t_j le radici del polinomio di grado n ortogonale rispetto alla misura μ , ovvero gli autovalori della matrice J_n , e come w_j i quadrati delle prime componenti dei corrispondenti autovalori, otteniamo che la formula è esatta, cioè il resto è nullo, per i polinomi di grado minore di $2n$.

In questo caso l'approssimazione $\int_a^b f(x)d\mu(x) = \sum_{j=1}^n f(t_j)w_j + R_G[f]$ prende il nome di *formula di quadratura di Gauss*

Teorema 4.5. *Supponiamo che f sia tale che $f^{(2n)}(x) > 0$ per ogni n e per ogni $x \in (a, b)$ e indichiamo con $L_G[f] = \sum_{j=1}^n f(t_j)w_j$. Allora:*

(i) $L_G[f]$ è una stima dal basso di $I[f]$

(ii) esiste $\eta \in (a, b)$ tale che $R_G[f] = \frac{f^{(2n)}(\eta)}{(2n)!}$

Dimostrazione. Sia $h(x)$ il polinomio di Hermite per f , ovvero il polinomio di grado minore di $2n$ tale che

$$h(t_i) = f(t_i) \quad h'(t_i) = f'(t_i) \quad i = 1, \dots, n$$

Le proprietà di tale polinomio ci permettono di scrivere il resto $r(x) = f(x) - h(x)$ come

$$r(x) = \prod_{i=1}^n (x - t_i)^2 \frac{f^{(2n)}(\theta)}{(2n)!} = \tilde{p}_n^2(x) \frac{f^{(2n)}(\theta)}{(2n)!}$$

con $\theta = \theta(x) \in (a, b)$

Poichè la formula di Gauss è esatta per i polinomi di grado minore di $2n$, si ha che

$$\int_a^b h(x)d\mu(x) = \sum_{i=1}^n w_i h(t_i) = \sum_{i=1}^n w_i f(t_i)$$

Quindi

$$\int_a^b f(x)d\mu(x) = \int_a^b h(x)d\mu(x) + \int_a^b r(x)d\mu(x) = \sum_{i=1}^n w_i f(t_i) + \int_a^b r(x)d\mu(x)$$

Da queste relazioni possiamo scrivere $R_G[f]$ come

$$R_G[f] = \int_a^b r(x)d\mu(x) = \frac{1}{(2n)!} \int_a^b \tilde{p}_n^2(x) f^{(2n)}(\theta) d\mu(x)$$

Poichè $\tilde{p}_n^2(x)$ non cambia segno e $f^{(2n)}(\theta)$ è una funzione continua nella x , per il teorema del valor medio del calcolo integrale esiste $\eta \in (a, b)$ tale che

$$R_G[f] = \frac{f^{(2n)}(\eta)}{(2n)!} \int_a^b \tilde{p}_n^2(x) d\mu(x) = \frac{f^{(2n)}(\eta)}{(2n)!} \langle \tilde{p}_n, \tilde{p}_n \rangle = \frac{f^{(2n)}(\eta)}{(2n)!}$$

Il punto (i) segue facilmente dall'osservazione del fatto che il resto $R_G[f]$ è sempre positivo. \square

Ponendo $m = 1$ in (4.8) si ottiene la *formula di Gauss-Radau*, che può assumere due forme, a seconda che si scelga $z_1 = a$ oppure $z_1 = b$. Poichè in questo modo abbiamo aggiunto un nodo alla formula di quadratura, abbiamo perso le proprietà relative alla matrice J_n , vorremmo quindi costruire una nuova matrice J'_{n+1} che si comporti in modo analogo e un polinomio \tilde{p}_{n+1} che abbia z_1 come radice. Dalla relazione di ricorrenza per i polinomi ortonormali, dobbiamo imporre

$$0 = \sqrt{C_{n+1}}\tilde{p}_{n+1}(z_1) = (z_1 - B_{n+1})\tilde{p}_n(z_1) - \sqrt{C_n}\tilde{p}_{n-1}(z_1)$$

da cui

$$B_{n+1} = z_1 - \sqrt{C_n} \frac{\tilde{p}_{n-1}(z_1)}{\tilde{p}_n(z_1)}$$

Al fine di calcolare effettivamente il coefficiente B_{n+1} , osserviamo che abbiamo imposto anche

$$(J_n - z_1 I)\tilde{\mathbf{p}}(z_1) = -\sqrt{C_n}\tilde{p}_n(z_1)\mathbf{e}_n$$

quindi definendo $\delta = (\delta_1, \dots, \delta_n)^T$ con $\delta_i = -\sqrt{C_n} \frac{\tilde{p}_{i-1}(z_1)}{\tilde{p}_n(z_1)}$, otteniamo

$$(J_n - z_1 I)\delta = C_n \mathbf{e}_n \quad (4.9)$$

Quanto visto ci permette di calcolare B_{n+1} tramite i seguenti passi:

- generiamo $\sqrt{C_n}$ col metodo di Lanczos
- risolviamo il sistema tridiagonale di incognita δ definito in (4.9)
- calcoliamo $B_{n+1} = z_1 + \delta_n$

In questo modo la matrice J'_{n+1} definita come

$$J'_{n+1} = \begin{pmatrix} J_n & \sqrt{C_n}\mathbf{e}_n \\ \sqrt{C_n}\mathbf{e}_n^T & B_{n+1} \end{pmatrix}$$

avrà z_1 come autovalore e produrrà i nodi e i pesi della formula di Gauss-Radau. Il seguente teorema, analogo a quello visto per la formula di Gauss, ci fornisce informazioni sul resto della formula di Gauss-Radau.

Teorema 4.6. *Supponiamo che f sia tale che $f^{(2n+1)}(x) < 0$ per ogni n e per ogni $x \in (a, b)$ e definiamo*

$$U_{GR}[f] = \sum_{j=1}^n f(t_j)w_j + v_a f(a)$$

$$L_{GR}[f] = \sum_{j=1}^n f(t_j)w_j + v_b f(b)$$

Allora $L_{GR}[f]$ è una stima inferiore di $I[f]$, mentre $U_{GR}[f]$ è una stima superiore, cioè:

$$L_{GR}[f] \leq I[f] \leq U_{GR}[f]$$

Inoltre la formula è esatta per i polinomi di grado minore di $2n + 1$.

Un'ulteriore formula di quadratura, detta *formula di Gauss-Lobatto*, si ha prendendo $m = 2$ nodi prescritti, con $z_1 = a$ e $z_2 = b$. Come prima dobbiamo modificare la matrice J_n , quindi poniamo

$$\tilde{p}_{n+1}(a) = \tilde{p}_{n+1}(b) = 0$$

Usando la ricorrenza a tre termini, possiamo scrivere B_{n+1} e $\sqrt{C_n}$ come soluzione del seguente sistema lineare:

$$\begin{pmatrix} \tilde{p}_n(a) & \tilde{p}_{n-1}(a) \\ \tilde{p}_n(b) & \tilde{p}_{n-1}(b) \end{pmatrix} \begin{pmatrix} B_{n+1} \\ \sqrt{C_n} \end{pmatrix} = \begin{pmatrix} a\tilde{p}_{n+1}(a) \\ b\tilde{p}_{n+1}(b) \end{pmatrix}$$

Siano ora $\delta = (\delta_1, \dots, \delta_n)^T$ e $\gamma = (\gamma_1, \dots, \gamma_n)^T$ definiti da

$$\delta_i = -\frac{\tilde{p}_{i-1}(a)}{\sqrt{C_n}\tilde{p}_n(a)} \quad \gamma_i = -\frac{\tilde{p}_{i-1}(b)}{\sqrt{C_n}\tilde{p}_n(b)}$$

questi vettori sono soluzioni dei due sistemi tridiagonali

$$(J_n - aI)\delta = \mathbf{e}_n \quad (J_n - bI)\gamma = \mathbf{e}_n \quad (4.10)$$

Il sistema di ordine 2 si può scrivere come

$$\begin{pmatrix} 1 & -\delta_n \\ 1 & -\gamma_n \end{pmatrix} \begin{pmatrix} B_{n+1} \\ C_n \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \quad (4.11)$$

I coefficienti B_{n+1} e $\sqrt{C_n}$ possono quindi essere ottenuti risolvendo prima i due sistemi tridiagonalidi ordine n in (4.10) e poi quello di ordine 2 in (4.11), in questo modo possiamo costruire la matrice

$$J''_{n+1} = \begin{pmatrix} J_n & \sqrt{C_n}\mathbf{e}_n \\ \sqrt{C_n}\mathbf{e}_n^T & B_{n+1} \end{pmatrix}$$

che fornisce nodi e pesi della formula di quadratura in questione. Come nei casi precedenti abbiamo il seguente teorema sul resto.

Teorema 4.7. *Supponiamo che f sia tale che $f^{(2n)}(x) > 0$ per ogni n e per ogni $x \in (a, b)$ e definiamo*

$$U_{GL}[f] = \sum_{j=1}^n f(t_j)w_j + v_a f(a) + v_b f(b)$$

Allora $U_{GL}[f]$ è una stima superiore di $I[f]$, cioè:

$$I[f] \leq U_{GR}[f]$$

Inoltre la formula è esatta per polinomi di grado minore di $2n + 2$.

Concludiamo la discussione sulle formule di quadratura con il seguente teorema che ci sarà di grossa utilità nella pratica in quanto ci dice che per il calcolo dell'approssimazione dell'integrale possiamo fare a meno di passare dal calcolo di autovalori e autovettori della matrice J_n .

Teorema 4.8 (Golub-Meurant).

$$\sum_{i=0}^n w_i f(t_i) = \mathbf{e}_1^T f(J_n) \mathbf{e}_1 = [f(J_n)]_{11}$$

Dimostrazione.

$$\begin{aligned} \sum_{i=1}^n w_i f(t_i) &= \sum_{i=0}^n (\sigma_1^{(i)})^2 f(t_i) \\ &= \sum_{i=1}^n \mathbf{e}_1^T \sigma^{(i)} f(t_i) (\sigma^{(i)})^T \mathbf{e}_1 \\ &= \mathbf{e}_1^T \left(\sum_{i=1}^n \sigma^{(i)} f(t_i) (\sigma^{(i)})^T \right) \mathbf{e}_1 \\ &= \mathbf{e}_1^T \Sigma \begin{pmatrix} f(t_1) & & \\ & \ddots & \\ & & f(t_n) \end{pmatrix} \Sigma^T \mathbf{e}_1 \\ &= \mathbf{e}_1^T f(J_n) \mathbf{e}_1. \end{aligned}$$

□

Per come sono state costruite le matrici J'_n e J''_n , il teorema precedente può essere applicato anche ad esse per il calcolo delle formule di Gauss-Radau e Gauss-Lobatto.

4.4 Stime a priori della centralità esponenziale

Il teorema conclusivo della sezione precedente ci fornisce un algoritmo veloce di approssimazione della misura di centralità esponenziale, grazie ad esso inoltre possiamo stabilire limitazioni a priori, simulando un solo passo del metodo di Lanczos e applicando le tre diverse formule di quadratura introdotte. Per fare questo abbiamo bisogno di introdurre una sorta di estensione matriciale della formula di interpolazione di Lagrange.

Consideriamo una funzione $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ e una matrice $G \in \mathbb{R}^{n \times n}$ diagonalizzabile e con gli autovalori $\alpha_1, \dots, \alpha_n$ tutti distinti fra loro. Costruisco ora il polinomio di interpolazione di Lagrange relativo ai punti $(\alpha_1, f(\alpha_1)), \dots, (\alpha_n, f(\alpha_n))$:

$$r(x) = \sum_{k=1}^n \frac{\prod_{h \neq k} (x - \alpha_h)}{\prod_{j \neq k} (\alpha_k - \alpha_j)} f(\alpha_k)$$

Per le proprietà del polinomio r e della matrice $G = PLP^T$,

dove $L = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix}$, vale la seguente formula:

$$f(G) = Pf(L)P^T = Pr(L)P^T = r(G) = \sum_{k=1}^n \frac{\prod_{h \neq k} (G - \alpha_h I)}{\prod_{j \neq k} (\alpha_k - \alpha_j)} f(\alpha_k) \quad (4.12)$$

4.4.1 Stima con la regola di Gauss

Vorremmo applicare la formula (4.12) alla matrice J_2 ottenuta dopo due passi del metodo di Lanczos

$$J_2 = \begin{pmatrix} B_1 & \sqrt{C_1} \\ \sqrt{C_1} & B_2 \end{pmatrix}$$

Ricordiamo che i coefficienti di tale matrice sono dati dal metodo di Lanczos con $\mathbf{v}^{(-1)} = \mathbf{0}$ e $\mathbf{v}^{(0)} = \mathbf{e}_i$. Quindi si ha che:

- $B_1 = (\mathbf{v}^{(0)})^T A \mathbf{v}^{(0)} = \mathbf{e}_i^T A \mathbf{e}_i = 0$ poichè tutti gli elementi diagonali di A sono nulli.
- $\mathbf{r}^{(1)} = A \mathbf{v}^{(0)} = A \mathbf{e}_i$.
- $\sqrt{C_1} = \|\mathbf{r}^{(1)}\| = \|A \mathbf{e}_i\| = \sqrt{d_i}$ dove d_i indica il grado del nodo i .
- $\mathbf{v}^{(1)} = \frac{\mathbf{r}^{(1)}}{\sqrt{d_i}}$.
- $B_2 = (\mathbf{v}^{(1)})^T A \mathbf{v}^{(1)} = \frac{(\mathbf{r}^{(1)})^T A \mathbf{r}^{(1)}}{d_i} = \frac{\mathbf{e}_i^T A^3 \mathbf{e}_i}{d_i} = \frac{2T_i}{d_i}$ dove T_i indica il numero di triangoli del grafo contenenti il nodo i .

Otteniamo la matrice

$$J_2 = \begin{pmatrix} 0 & \sqrt{d_i} \\ \sqrt{d_i} & \frac{2T_i}{d_i} \end{pmatrix}$$

che ha per autovalori

$$x = \frac{T_i + S}{d_i} \quad \text{e} \quad y = \frac{T_i - S}{d_i} \quad \text{dove} \quad S = \sqrt{T_i^2 + d_i^3}$$

Applicando la formula (4.12) alla matrice J_2 e alla funzione esponenziale, si ha che

$$\exp(J_2) = \frac{(J_2 - yI)e^x}{(x - y)} + \frac{(J_2 - xI)e^y}{(y - x)} = \frac{J_2(e^x - e^y) + I(xe^y - ye^x)}{(x - y)}$$

Poichè siamo interessati solamente all'elemento in alto a sinistra, scriviamo

$$\begin{aligned} e_1^T \exp(J_2) e_1 &= \frac{xe^y - ye^x}{x - y} = \frac{\frac{T_i+S}{d_i} \exp\left(\frac{T_i-S}{d_i}\right) - \frac{T_i-S}{d_i} \exp\left(\frac{T_i+S}{d_i}\right)}{\frac{2S}{d_i}} = \\ &= \frac{[(T_i + S)(\sinh(-\frac{S}{d_i}) + \cosh(-\frac{S}{d_i})) - (T_i - S)(\sinh(\frac{S}{d_i}) + \cosh(\frac{S}{d_i}))]}{\frac{2S}{\exp(\frac{T_i}{d_i})}} = \\ &= \frac{e^{\frac{T_i}{d_i}}}{2S} \left[S \cosh\left(\frac{S}{d_i}\right) - T_i \sinh\left(-\frac{S}{d_i}\right) \right] \end{aligned}$$

Grazie a questa relazione, al teorema di Golub-Meurant e al fatto che il metodo di quadratura di Gauss è un'approssimazione dal basso, otteniamo la seguente stima inferiore della misura di centralità esponenziale

$$[e^A]_{ii} \geq \frac{e^{\frac{T_i}{d_i}}}{\sqrt{T_i^2 + d_i^3}} \left[\sqrt{T_i^2 + d_i^3} \cosh\left(\frac{\sqrt{T_i^2 + d_i^3}}{d_i}\right) - T_i \sinh\left(\frac{\sqrt{T_i^2 + d_i^3}}{d_i}\right) \right]$$

4.4.2 Stime con la regola di Gauss-Radau

Per prima cosa osserviamo che il teorema riguardo le stime di Gauss-Radau ha come ipotesi sulla funzione f che questa abbia tutte le derivate dispari negative. Ovviamente ciò non è soddisfatto dalla funzione esponenziale quindi in questo caso useremo il trucco di considerare $f(x) = e^{-x}$ applicando tutti i calcoli alla matrice $-A$. Osserviamo inoltre che se $a = \lambda_1 \leq \dots \leq \lambda_n = b$ sono gli autovalori di A , allora $-b = -\lambda_n \leq \dots \leq -\lambda_1 = -a$ sono gli autovalori di $-A$. Dal primo passo del metodo di Lanczos otteniamo le seguenti relazioni

- $B_1 = (\mathbf{v}^{(0)})^T (-A) \mathbf{v}^{(0)} = \mathbf{e}_i^T (-A) \mathbf{e}_i = 0$ poichè tutti gli elementi diagonali di $-A$ sono nulli.
- $\mathbf{r}^{(1)} = (-A) \mathbf{v}^{(0)} = -A \mathbf{e}_i$.
- $\sqrt{C_1} = \|\mathbf{r}^{(1)}\| = \|-A \mathbf{e}_i\| = \|A \mathbf{e}_i\| = \sqrt{d_i}$ dove d_i indica il grado del nodo i .

che ci permettono di costruire la matrice

$$J'_2 = \begin{pmatrix} 0 & \sqrt{d_i} \\ \sqrt{d_i} & B'_2 \end{pmatrix}$$

dove il coefficiente B'_2 non è ottenuto tramite il metodo di Lanczos, ma risolvendo un sistema tridiagonale di incognita δ che in questo caso è di taglia 1 e si riduce a

$$-z\delta = d_i$$

per ottenere infine $B'_2 = z + \delta$. Nelle espressioni precedenti z può prendere come valore una stima inferiore dell'autovalore massimo di $-A$, ossia $-a$, oppure una stima superiore dell'autovalore minimo, ossia $-b$. Per quanto visto in precedenza, nel primo caso, considerando $[e^{J'_2}]_{11}$, si ottiene una stima inferiore della misura di centralità, mentre nel secondo caso si ha una stima superiore. Consideriamo il primo caso, cioè $z = -a$. La matrice diventa

$$J'_2 = \begin{pmatrix} 0 & \sqrt{d_i} \\ \sqrt{d_i} & \frac{d_i}{a} - a \end{pmatrix}$$

e si verifica facilmente che ha autovalori $-a$ e $\frac{d_i}{a}$. Usando come prima la formula (4.12) otteniamo

$$e_1^T \exp(J'_2) e_1 = \frac{\frac{d_i}{a} e^a + a e^{-\frac{d_i}{a}}}{\frac{d_i}{a} + a} = \frac{d_i e^a + a^2 e^{-\frac{d_i}{a}}}{d_i + a^2}$$

Procedendo in maniera analoga per il caso $z = -b$ si ricavano le seguenti limitazioni per l'indice di centralità

$$\frac{d_i e^a + a^2 e^{-\frac{d_i}{a}}}{d_i + a^2} \leq [e^A]_{ii} \leq \frac{d_i e^b + b^2 e^{-\frac{d_i}{b}}}{d_i + b^2}$$

4.4.3 Stima con la regola di Gauss-Lobatto

Questa volta l'unica informazione che ricaviamo dal metodo di Lanczos è che $B_1 = 0$, quindi abbiamo una matrice della forma

$$J''_2 = \begin{pmatrix} 0 & \sqrt{C_1} \\ \sqrt{C_1} & B_2 \end{pmatrix}$$

dove, come abbiamo visto, per calcolare i coefficienti C_1 e B_2 dobbiamo prima risolvere due sistemi tridiagonali, di incognite δ e γ , che in questo caso hanno dimensione 1

$$-a\delta = 1 \quad -b\gamma = 1$$

e successivamente risolvere un altro sistema di dimensione 2

$$\begin{pmatrix} 1 & -\delta \\ 1 & -\gamma \end{pmatrix} \begin{pmatrix} B_2 \\ C_1 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

Da queste relazioni segue che $C_1 = -ab$, $B_2 = b$ e

$$J_2'' = \begin{pmatrix} 0 & \sqrt{-ab} \\ \sqrt{-ab} & b \end{pmatrix}$$

Si verifica facilmente che, come ci aspettavamo, la matrice J_2'' ha per autovalori a e b , quindi la stima data dalla formula di Gauss-Lobatto si scrive come

$$[e^A]_{ii} \leq \frac{ae^b + be^a}{a - b}$$

Capitolo 5

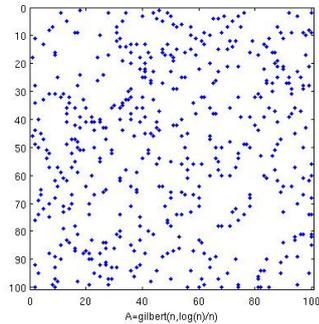
Sperimentazione numerica

5.1 Grafi random

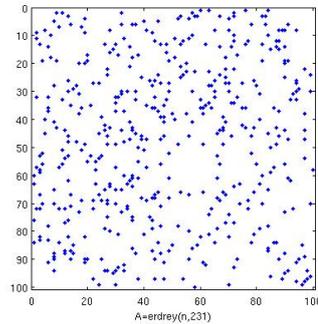
Per ottenere esempi di grafi su cui testare l'algoritmo che scriveremo adopereremo CONTEST ([7]), il pacchetto di Matlab creato da Taylor e Higham che sfrutta la casualità per creare matrici di adiacenza di grafi sparsi che possono modellizzare reti reali.

Qui di seguito elencheremo i tipi di grafo utilizzati nella sperimentazione che possono essere creati tramite le funzioni del pacchetto.

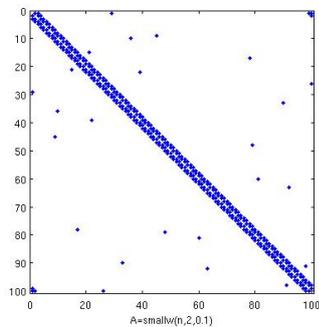
- *Grafi di Gilbert* $G(n, p)$. Grafo di taglia n in cui ogni coppia di nodi è connessa, in maniera indipendente, con probabilità p . La funzione del pacchetto che crea la matrice di adiacenza di tale grafo è `gilbert(n,p)`.
- *Grafi di Erdős-Rényi* $ER(n, m)$. Grafo scelto con probabilità uniforme fra tutti i nodi con n nodi e m archi. La funzione corrispondente è `erdrey(n,m)`.
- *Grafi di tipo "small world"* $S(n, k, p)$. Grafo creato a partire da un k -anello (cioè un grafo in cui i nodi i e j sono connessi se e solo se $|i - j| \leq k$ oppure $|n - |i - j|| \leq k$) a cui sono aggiunti altri archi con probabilità p . Dal punto di vista dell'algebra lineare, la matrice creata dalla funzione `smallw(n,k,p)` è simmetrica e di Toeplitz, con pochi elementi non nulli extra.
- *Grafi di tipo Renga* $R(n, \lambda, \alpha)$. Grafo di taglia n in cui inseriamo un arco fra il nodo i e il nodo j con probabilità $\alpha\lambda^{|j-i|-1}$, dove $\alpha > 0$ e $\lambda \in (0, 1)$. La scelta $\alpha = 1$ assicura che nodi adiacenti siano connessi, mentre il fattore geometrico $\lambda^{|j-i|-1}$ provoca il fatto che archi fra nodi vicini siano più comuni di archi fra nodi lontani. La funzione corrispondente è `renga(n,lambda,alpha)`.



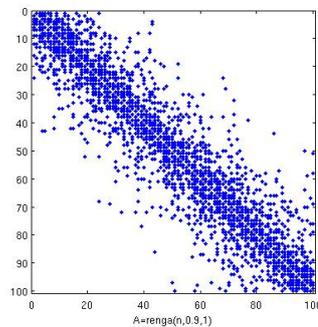
(a) Gilbert.



(b) Erdős-Rényi.



(c) Small world.



(d) Renga.

Le immagini precedenti mostrano un esempio per ogni modello di grafo presentato, per costruirle si è usato il comando `spy` di Matlab e si è posto $n = 100$.

5.2 Analisi algoritmica

Svilupperemo l'algoritmo per il calcolo della centralità esponenziale considerando la stima inferiore data dalla regola di Gauss e la stima superiore data dalla regola di Gauss-Radau con b come nodo prescritto, interrompendo il processo iterativo quando le due stime sono fra loro a una distanza minore di quella da noi desiderata. Come detto in precedenza non pretenderemo una precisione esagerata, ci basterà che la differenza fra le due stime sia minore di 10^{-3} .

L'algoritmo per il calcolo del singolo elemento i diagonale procede come descritto nel codice seguente, ovviamente per calcolare tutto il ranking del grafo dovremo inserire tale codice in un ciclo che lavori su tutti i nodi del grafo.

Algoritmo 5.1: Calcolo della misura di centralità del nodo i .

```

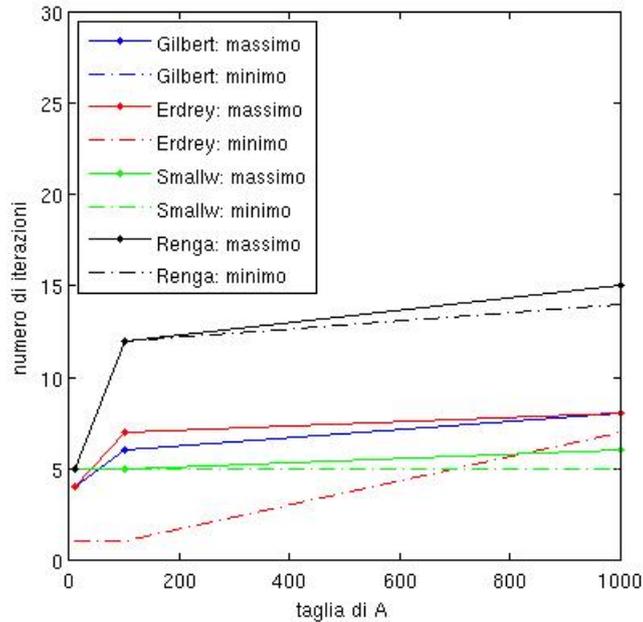
1  function [gauss,radau,k-1] = centrality(A,i)
2  A=sparse(A);
3  k = 1;
4  gauss = 4;
5  radau = 5;
6  x_0 = zeros(n,1);
7  x = zeros(n,1);
8  x(i) = 1;
9  x_0 = sparse(x_0);
10 x = sparse(x);
11 r = zeros(n,1);
12 gamma = 0;
13 while ( (abs(gauss-radau)) > 0.001 )
14     temp = A*x;
15     omega = x'*temp;
16     r = temp - omega*x - gamma*x_0;
17     gamma = norm(r);
18     x_0 = x;
19     x = r/gamma;
20     J(k,k) = omega;
21     J(k,k+1) = gamma;
22     J(k+1,k) = gamma;
23     Jp = J(1:k,1:k);
24     [V,D] = eig(Jp);
25     M = V*diag(exp(diag(D)))/V;
26     gauss = M(1,1);
27     b = norm(A,1);
28     L = Jp-b*eye(k);
29     f = zeros(k,1);
30     f(k) = gamma^2;
31     delta = L\f;
32     omeganew = b+delta(k);
33     f(k) = gamma;
34     JN = [Jp,f;f',omeganew];
35     [V,D] = eig(JN);
36     MN = V*diag(exp(diag(D)))/V;
37     radau = MN(1,1);
38     k = k+1;
39 end

```

Le righe di codice 2 – 12 sono di inizializzazione: vengono definiti i valori di partenza del metodo di Lanczos e vengono resi sparsi i vettori in uso e la matrice A al fine di velocizzare tutti i prodotti matrice per vettore. La prima parte del ciclo (righe 14 – 22) consiste esattamente nel passo j dell'iterazione di Lanczos, nella parte centrale (righe 23 – 26) invece si costruisce la matrice J_k della regola di Gauss e si calcola la stima inferiore. Infine, nelle righe 27 – 37, si calcola la stima superiore di Gauss-Radau dopo aver costruito la

matrice J'_k e aver risolto un sistema lineare di taglia k .

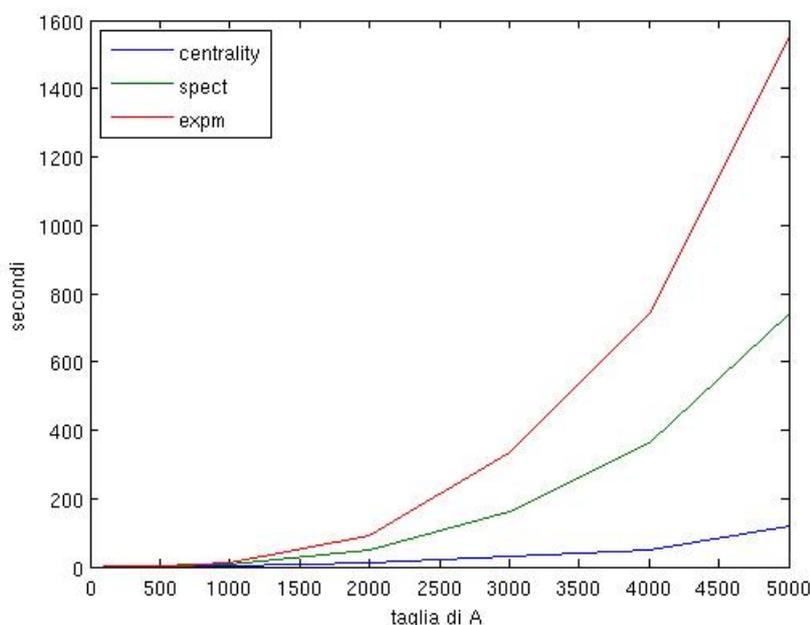
Il punto di forza di questo algoritmo sta nel fatto che le matrici J_k e J'_k sono generalmente di taglia molto più piccola rispetto ad A , quindi i calcoli delle due stime sono molto efficienti.



Infatti, come mostrato dal grafico in figura, per i vari tipi di grafo, il numero di iterazioni del metodo di Lanczos necessarie ad ottenere la precisione richiesta si mantiene grosso modo costante all'aumentare della dimensione del grafo. In particolare questo grafico mostra come le linee continue e le linee tratteggiate, che indicano il numero massimo e il numero minimo di iterazioni effettuate, siano generalmente vicine fra loro e addirittura coincidano per grafi di Gilbert. Questo fatto ci assicura che l'algoritmo abbia un costo lineare nel numero n di nodi, tuttavia il codice (5.1) può essere ulteriormente raffinato. Alla k -esima iterazione per trovare le stime di Gauss e Gauss-Radau dobbiamo effettuare due esponenziali di matrici e risolvere un sistema lineare di taglia k , questo vuol dire che se effettuiamo m iterazioni dobbiamo effettuare $2m$ esponenziali di matrici e risolvere m sistemi lineari di dimensione via via crescente. Per migliorare questo aspetto modifichiamo leggermente il codice (5.1) in modo che le due stime vengano calcolate non in tutte le iterazioni k , ma solamente nelle iterazioni in cui k è una potenza di due. In questo modo otteniamo che bisogna effettuare un numero sensibilmente minore di esponenziali e sistemi lineari al prezzo di aumentare il numero di iterazioni. Sostanzialmente il vantaggio che abbiamo acquisito è quello che fra tutte le iterazioni che effettuiamo, solo poche sono veramente costose dal punto di vista computazionale. Sperimentalmente si può verificare come l'algoritmo sia stato effettivamente velocizzato e come il guadagno

sia più sensibile su grafi, ad esempio, di tipo Renga in cui tipicamente servono un numero alto di iterazioni.

Verifichiamo ora come il metodo che abbiamo costruito sia davvero efficiente. Il seguente grafico confronta, per matrici di taglia crescente ottenute con la funzione `erdrey` del pacchetto CONTEST, tre diversi metodi di calcolo della centralità esponenziale: il primo consiste nell'esecuzione della funzione `expm` di Matlab, il secondo usa la decomposizione spettrale della matrice $A = P^{-1}DP$ e calcola $A = P^{-1} \exp(D)P$ sfruttando il fatto che A è simmetrica e quindi D diagonale, il terzo è l'algoritmo (5.1) modificato come detto in precedenza.



Un altro dato emerso dalla sperimentazione che rafforza la bontà del nostro algoritmo riguarda grafi con un numero di nodi ancora maggiore, per $n > 5000$ infatti i due metodi che utilizzano le funzioni predefinite di Matlab si interrompono per uso eccessivo della memoria a disposizione, mentre l'algoritmo (5.1) impiega, per esempio, meno di 6 minuti per un grafo con 8000 nodi e poco più di 20 minuti per un grafo con 16000 nodi. Questo fatto ci incoraggia ad applicare il nostro algoritmo su reti reali che tipicamente sono composti da un numero di nodi dell'ordine delle decine di migliaia.

Per concludere vorremmo renderci conto di come varia il ranking dei nodi a seconda della misura di centralità che si sceglie, verificando sperimentalmente il risultato del teorema (3.2). Per fare ciò utilizzeremo come esempio la rete *Power Grid* ([8]) costruita come modello della rete elettrica degli Stati Uniti occidentali in cui i nodi rappresentano generatori, trasformatori o stazioni succursali e gli archi rappresentano cavi di trasmissione ad alta

tensione.

<i>ranking</i>	d	c(0.1)	c(1)	c(2)	b
1°	2554	2554	4346	4346	4382
2°	4459	4459	4382	4382	4346
3°	4346	4346	4353	4353	4337
4°	3469	832	4385	4337	4333
5°	832	3469	4337	4385	4353

La tabella precedente mostra gli indici dei migliori 5 nodi del grafo secondo la misura grado **d**, l'indice di Bonacich **b** e la centralità esponenziale **c**(β), per vari valori di β .

Bibliografia

- [1] Michele Benzi and Paola Boito. Quadrature rule-based bounds for functions of adjacency matrices. *Linear Algebra and its Applications*, 433(3):637–652, 2010.
- [2] Dario Fasino. The structure and function of complex networks.
- [3] Feliks R Gantmacher. *The theory of matrices*, volume 1. Taylor & Francis, 1959.
- [4] Gene H Golub and Gérard Meurant. *Matrices, moments and quadrature with applications*. Princeton University Press, 2009.
- [5] Christine Klymko. *Centrality and Communicability Measures in Complex Networks: Analysis and Algorithms*. PhD thesis, Emory University, 2013.
- [6] Mark EJ Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.
- [7] Alan Taylor and Desmond J Higham. Contest: A controllable test matrix toolbox for matlab. *ACM Transactions on Mathematical Software (TOMS)*, 35(4):26, 2009.
- [8] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.
- [9] Herbert S Wilf. *Mathematics for the physical sciences*. Courier Corporation, 2013.