UNIVERSITÀ DI PISA

Facoltà di Scienze Matematiche, Fisiche e Naturali

Corso di Laurea Triennale in Matematica

Tesi di Laurea

Grafi Expanders: Proprietà Spettrali e Applicazioni

Relatore:
Andrea Dario Bini

Candidato:
Barbarino Giovanni

ANNO ACCADEMICO 2013-2014

Introduzione

Lo scopo principale di questo documento è fare un excursus di alcune proprietà ed applicazioni dei Grafi Expanders, concentrandosi sulle caratteristiche che rendono questi grafi adatti a risolvere particolari tipologie di problemi. Non pretendiamo di coprire tutti gli utilizzi di questo potente mezzo, ma diamo un'idea delle sue applicazioni in molti campi della matematica, dall'algebra alla geometria e alla probabilità.

Nel Capitolo 1, introduciamo brevemente il concetto di Grafo, assieme alle sue matrici caratteristiche, ossia la Matrice di Adiacenza, la Matrice Laplaciana e la Matrice di Transizione, e studiamo gli spettri delle prime due, collegandole a propietà combinatorie del grafo.

Nel capitolo successivo, definiamo il Numero Isoperimetrico di un grafo, ed introduciamo la famiglia dei Grafi Expanders in relazione ad esso. Dimostrando la Diseguaglianza di Cheeger, mostriamo come le proprietà spettrali del grafo diano una misura del numero sopra definto, che in generale è molto complessa da calcolare.

I Capitoli 3 e 4 danno degli esempi di applicazioni.

Studiando infatti la Matrice di Transizione di un grafo, possiamo dedurre proprietà riguardanti la convergenza veloce di Catene di Markov, e dunque come una Camminata Aleatoria operata su un grafo possa essere assimilabile ad una Campionatura (o Sampling) casuale. Mostriamo quindi un utilizzo di queste caratteristiche nella Derandomizzazione di un algoritmo progettato da Micheal O. Rabin per un Test di Primalità.

Passando invece alla Teoria dei Codici, illustriamo il problema della correzione degli errori nella trasmissione di messaggi. Utilizziamo dunque particolari Grafi Expanders, chiamati Grafi Magici, che permettano di generare dei buoni Dizionari, oltre che ad efficienti algoritmi di decodifica dei messaggi, basati sulla Belief Propagation.

Nel quinto capitolo dimostriamo infine l'esistenza dei Grafi Magici, e mostriamo esplicitamente qualche esempio di Grafo Expander, illustrandone le relazioni con grafi regolari generati casualmente.

Indice

Capitolo 1

Grafi e Proprietà Spettrali

Dato V un insieme di elementi, e E un sottoinsieme di $V \times V$, definiamo $G = (V, E)$ un grafo, e chiameremo V l'insieme dei suoi vertici (o nodi), e $E \subseteq V \times V$ l'insieme dei suoi archi. Graficamente, un grafo è rappresentabile come un insieme di vertici e di archi che li collegano.

Per tutto il seminario, se non specificato, supporremo che i grafi utilizzati siano finiti, connessi, non diretti, e semplici, ossia

finito il numero dei vertici sia finito.

connesso per ogni coppia di nodi distinti u, v esistono degli archi in E del tipo $(u, x_1), (x_1, x_2), \ldots, (x_{n-1}, x_n), (x_n, v)$.

non diretto gli archi sono rappresentati da coppie non ordinate di elementi.

semplice non ci sono archi da un nodo in sè stesso(loop), e tra due nodi ci può essere al massimo un arco.

Inoltre, porremo $n = |V|$ il numero di nodi, $m = |E|$ il numero di archi, e indichiamo con d_v il grado del nodo v, ossia il numero di archi che hanno v come uno dei due estremi. In formule

$$d_v = |\{w \in V \ : \{v, w\} \in E\}|.$$

Infine numeriamo i vertici da 1 a n, in modo che i d_i siano ordinati in maniera non crescente.

Un modo di esprimere un grafo, è attraverso la sua *matrice di adiacenza*, ossia una matrice simmetrica A_G di dimensione $n \times n$ ad entrate binarie:

$$(A_G)_{ij} = \begin{cases} 1 & \text{se } \{i, j\} \in E \\ 0 & \text{se } \{i, j\} \notin E \end{cases}$$

Un'altra matrice importante è il *Laplaciano*, o *Matrice Laplaciana* L_G del grafo. Anche lei è una matrice $n \times n$, definita da

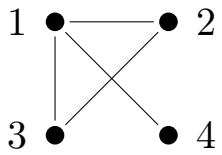
$$(L_G)_{ij} = \begin{cases} -1 & \text{se } \{i, j\} \in E \\ 0 & \text{se } \{i, j\} \notin E \text{ e } i \neq j \\ d_i & \text{se } i = j \end{cases}$$

Entrambe queste matrici sono diagonalizzabili nei reali, in quanto simmetriche.

Se invece dividiamo ogni riga della matrice di adiacenza per il grado del nodo corrispondente, otteniamo una matrice, che denoteremo con P_G e chiameremo *Matrice di Transizione*, fondamentale per quanto riguarda lo studio di Passeggiate Aleatorie sul grafo.

$$(P_G)_{ij} = (A_G)_{ij}/d_i$$

Diamo ora un esempio pratico:



Rappresentazione Grafica

$$L_G = \begin{pmatrix} 3 & -1 & -1 & -1 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 2 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

Matrice Laplaciana

$$A_G = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Matrice di Adiacenza

$$P_G = \begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 1/2 & 0 & 1/2 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Matrice di Transizione

Se D_G è la matrice diagonale che ha per elementi i gradi dei nodi del grafo, allora

$$P_G = D_G^{-1} A_G = D_G^{-1/2} D_G^{-1/2} A_G D_G^{-1/2} D_G^{1/2},$$

ma $D_G^{-1/2} A_G D_G^{-1/2}$ è ancora simmetrica poiché

$$(D_G^{-1/2} A_G D_G^{-1/2})^t = (D_G^t)^{-1/2} A_G^t (D_G^t)^{-1/2} = D_G^{-1/2} A_G D_G^{-1/2}.$$

Ne segue che P_G è simile ad una matrice simmetrica, e dunque diagonalizzabile.

Nelle prossime sezioni accenniamo alcune proprietà spettrali dei grafi, ossia informazioni che possiamo trarre dagli autovalori e autovettori delle matrici associate ai grafi, per dare un'idea dell'importanza dello studio degli spettri.

1.1 Proprietà del Laplaciano

Essendo una matrice simmetrica, il Laplaciano determina un prodotto scalare su $\mathbb{R}^n$, dato da

$$x^t L_G x = \sum_{1 \le i,j \le n} x_i x_j (L_G)_{ij} = \sum_{1 \le i \le n} x_i^2 d_i - 2 \sum_{\{i,j\} \in E} x_i x_j = \sum_{\{i,j\} \in E} (x_i - x_j)^2.$$

Questo è chiaramente semidefinito positivo, dunque gli autovalori del Laplaciano sono nonnegativi, ed inoltre, chiamato e il vettore $(1, \dots, 1) \in \mathbb{R}^n$, ci accorgiamo subito che è un autovettore relativo all'autovalore zero del Laplaciano. Ciò ci dice che il Laplaciano di un grafo è sempre singolare.

Più in generale, vedremo tra poco che la dimensione del nucleo dell'applicazione lineare associata al Laplaciano è un'informazione importante.

Definiamo adesso un *sottografo* di G relativo ad un sottoinsieme di nodi, e le *componenti connesse* di G.

Definizione 1.1.1 (Sottografo)**.**
Dato G un grafo, e $S \subseteq V$ un sottoinsieme dei nodi, allora il *sottografo* di G determinato da S è il grafo G' che ha per nodi $V' = S$, e per archi E', che sono gli archi di G che collegano tra loro i nodi in S. In formule

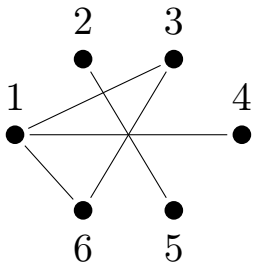
$$E' = \{\{v, w\} \in E \mid v, w \in S\}.$$

Definizione 1.1.2 (Componente Connessa)**.**
Dato G un grafo, $C \subseteq V$ è una *componente connessa* di G, se il sottografo determinato da C è connesso, e non vi sono archi che partano da C e finiscano in V/C.

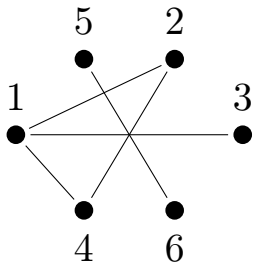
Preso ora G, possiamo partizionare i nodi nelle sue $c_1, c_2, \ldots, c_k$ componenti connesse, dove

$$\bigcup_{i=1}^{k} c_i = V \qquad \text{e} \qquad i \neq j \implies c_i \cap c_j = \emptyset.$$

É possibile rinumerare i nodi in modo che la matrice laplaciana del grafo (come anche quella di adiacenza e quella di transizione) sia diagonale a blocchi.
Un esempio è dato qui sotto:



Grafo G



Grafo G permutato

$$L_G = \begin{pmatrix} 3 & 0 & -1 & -1 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 2 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 2 \end{pmatrix} \qquad L_G = \begin{pmatrix} 3 & -1 & -1 & -1 & 0 & 0 \\ -1 & 1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

Laplaciano Laplaciano permutato

I blocchi in cui viene diviso il Laplaciano dopo la permutazione dei vertici, saranno le matrici laplaciane corrispondenti ai sottografi determinati dalle relative componenti connesse.

L'operazione di permutazione dei vertici viene inoltre letta a livello matriciale come una permutazione di righe e colonne, operata tramite un cambio base ortogonale. La matrice risultante, dunque, sarà ancora simile alla matrice di partenza, e chiamando *Polinomio Laplaciano* il polinomio caratteristico della matrice laplaciana associata ad un grafo, avremo che

Lemma 1.1.1. *Dato G un grafo, e $c_1, c_2, \ldots, c_k$ le sue componenti connesse, allora il polinomio laplaciano di G sarà il prodotto dei polinomi laplaciani dei sottografi determinati dai c_i.*

Questo, insieme a poche altre osservazioni, porta al seguente importante risultato:

Teorema 1.1.1. *la dimensione del nucleo dell'applicazione relativa alla matrice laplaciana di un grafo è pari al numero di componenti connesse del grafo.*

Dimostrazione.

Poniamo che il grafo G sia connesso. Sappiamo già che il Laplaciano è singolare, quindi dobbiamo dimostrare che l'autovalore zero è semplice, ossia ha molteplicità uno.

Dato x autovettore del Laplaciano di G relativo all'autovettore 0, avremo che

$$x^t L_G x = \sum_{\{i,j\} \in E} (x_i - x_j)^2 = 0 \implies \forall i, j \in E, x_i = x_j.$$

Sappiamo, però, che dati due nodi, riusciamo sempre ad arrivare dall'uno all'altro tramite un percorso, poiché il grafo è connesso, dunque gli x_i devono essere tutti uguali, e ciò vuol dire che x è un multiplo del vettore $e = (1, 1, \ldots, 1)$.

Questo prova che la dimensione del nucleo di L_G è uno nel caso di grafo connesso.

Preso ora un grafo G con $c_1, c_2, \ldots, c_k$ componenti connesse, la dimensione del nucleo del Laplaciano è pari alla molteplicità dell'autovalore zero, ma grazie al lemma 1.1.1 gli autovalori di G sono l'unione degli autovalori dei laplaciani delle singole componenti connesse, contati con molteplicità.

Dato che ogni sottografo relativo ai c_i è connesso, avremo che ogni c_i contribuisce al nucleo con esattamente un autovalore zero, e dunque la dimensione del nucleo sarà k.

□

Questo risultato ci dice quindi che un grafo è connesso sse la matrice laplaciana ha un singolo autovalore zero; riprenderemo questo fatto più avanti, ma per ora analizziamo alte proprietà relative al polinomio laplaciano.

Definizione 1.1.3 (Albero)**.**

Chiamiamo *albero* un grafo connesso con n nodi e $n - 1$ archi.

Definizione 1.1.4 (Albero di Copertura)**.**

dato un grafo G, un suo *albero di copertura* è un albero che abbia gli stessi nodi del grafo, e tale che i suoi archi siano un sottoinsieme degli archi del grafo.

Lemma 1.1.2. *Se a_1 è il coefficiente di grado 1 del polinomio laplaciano, e $\tau(G)$ è il numero di alberi di copertura del grafo, allora $a_1 = n \cdot \tau(G)$.*

Questo risultato è dimostrabile con semplici argomenti combinatori, o anche come caso particolare del *Matrix Tree Theorem*[19]. Per enunciarlo, abbiamo bisogno della seguente definizione:

Definizione 1.1.5 (Foresta di Copertura)**.**
Dato un grafo, una *foresta di copertura* è un insieme di alberi senza nodi in comune, tali che ogni albero sia un sottografo di G, e che l'unione di essi copra tutti i nodi di G.

Da notare che anche un nodo singolo, per definizione, è un albero.

Teorema 1.1.2 (Matrix Tree Theorem)**.**
Sia a_k il coefficiente di grado k nel polinomio caratteristico del laplaciano. Data $F = \{T_1, \ldots, T_r\}$ una foresta di copertura del grafo, sia $P(F)$ il prodotto del numero di nodi di ogni albero T_i in F. Detto inoltre S_k l'insieme delle foreste di copertura di G con esattamente k archi, allora

$$|a_{n-k}| = \sum_{F \in S_k} P(F).$$

In generale, una foresta di copertura ha k alberi sse ha $n-k$ archi, poiché ogni albero è di copertura per il sottografo dei nodi che comprende, ed un albero di copertura ha esattamente un numero di archi pari a quello dei nodi meno uno. Inoltre se un grafo ha k componenti connesse, esistono solo foreste di coperture con almeno k alberi.
Questo dice che il numero di componenti connesse coincide con il più piccolo indice i per cui $a_i \neq 0$, e dunque coincide anche con la molteplicità di zero come radice del polinomio caratteristico, dimostrando il lemma 1.1.2.

Le proprietà del polinomio laplaciano sono molte e variegate, ma qui ne riporteremo solo un'altra, tratta da [3].

Definizione 1.1.6 (Distanza tra Nodi)**.**
Dati due nodi in un grafo, la loro *distanza* è la lunghezza del cammino più corto che va da uno all'altro. La indichiamo come $d_G(v, w)$.

Definizione 1.1.7 (Indice di Wiener)**.**
Dato G connesso, l'*indice di Wiener* del grafo è

$$W(G) = \sum_{v,w \in V} d_G(v, w).$$

Teorema 1.1.3. *Dato T un albero, a_2 il coefficiente di secondo grado del suo polinomio laplaciano, e μ_i gli autovalori del laplaciano, allora*

$$a_2 = W(T) = \sum_i \frac{n}{\mu_i}.$$

1.2 Autovalori del Laplaciano

Come abbiamo visto, L_G è una matrice simmetrica singolare e semidefinita positiva.

Indichiamo dunque con $\mu_1 \geq \mu_2 \geq .. \geq \mu_{n-1} \geq \mu_n = 0$ gli autovalori di L_G, e con $d_1 \geq d_2 \geq \cdots \geq d_n$ la sequenza dei gradi del grafo.

Inoltre definiamo

Definizione 1.2.1 (Diametro)**.**

Dato G grafo, il suo *diametro* è

$$l = \max_{v,w \in V} d_G(v, w).$$

Definizione 1.2.2 (Automorfismo)**.**

Dato G grafo, un *automorfismo* del grafo è una matrice di permutazione P tale che

$$PL_G P^t = L_G.$$

Applicare un automorfismo ad un grafo, vuol dire mantenere la stessa struttura grafica permutandone i nodi, quindi ci permette di scoprirne le simmetrie. Dato che gli automorfismi formano un gruppo, possiamo infatti farlo agire sui nodi, e ottenere una partizione dei nodi in orbite. Come mostrato in [3] e in [1], queste sono strettamente legate alle molteplicità autovalori:

Lemma 1.2.1. *Dato un automorfismo di G, con s cicli dispari e t cicli pari, allora il laplaciano ha al massimo $s + 2t$ autovalori semplici.*

Lemma 1.2.2. *Se il laplaciano ha tutti gli autovalori distinti, allora il gruppo degli automorfismi è abeliano, e tutti gli automorfismi hanno ordine al più 2.*

Questo influenza anche le proprietà del grafo, come mostrato dal seguente risultato.

Lemma 1.2.3. *In un grafo G, il diametro del grafo è sempre strettamente minore del numero degli autovalori distinti di L_G.*

Le proprietà più importanti, invece, le ricaviamo andando a studiare singolarmente gli autovalori.

Definizione 1.2.3 (Connettività Algebrica)**.**

Dato G grafo, definiamo la sua *connettività algebrica* come il secondo autovalore più piccolo del laplaciano, ossia

$$a(G) = \mu_{n-1}.$$

Come abbiamo già avuto modo di vedere, $a(G)$ è diverso da zero sse il grafo è connesso. Inoltre il suo valore contribuisce a determinare una quantità importante chiamata **Gap Spettrale**, che definiremo nel capitolo sugli Expanders.

Questo viene usato in quanto dà una misura di connessione del grafo, come mostrano i seguenti risultati, tratti principalmente da [4].

Lemma 1.2.4. *Dato l il diametro del grafo G, allora*

- $2m/(n-1) \geq a(G) \geq 4/ln$.
- $d_n \geq a(G) \geq d_{n-1} - n + 3$.
- *i, j non adiacenti* $\implies a(G) \leq (d_i + d_j)/2$.

Un altro risultato notevole riguarda invece la cardinalità dei *tagli* di G, ossia il numero di archi che partono da un insieme $S \subseteq V$ e finiscono in $S^c = V/S$. Più in generale, definiamo

Definizione 1.2.4 (Coboundary, Taglio e Neighbourhood)**.**
Dati X e Y due sottoinsiemi di V, chiamiamo il *coboundary di* (X, Y) l'insieme degli archi che partono in X e finiscono in Y, e lo indichiamo come $E(X, Y)$.
Quando $Y = X^c$, viene anche detto *taglio di* X, e si indica con $E(X)$. In formule

- $E(X,Y) := \{\{x,y\} \in E \mid x \in X,\ y \in Y\}. \qquad E(X) := E(X, X^c).$
- $e(X,Y) := |E(X,Y)|. \qquad e(X) := |E(X)|.$

Definizione 1.2.5 (Neighbourhood)**.**
Se invece vogliamo indicare i nodi non in X che sono collegati ad uno dei nodi in X, scriviamo $N(X)$, ossia il *neighborhood di* X. Se X è composto da un solo nodo v, scriviamo $N(v)$. In formule

$$N(X) = \{v \in V \mid \exists w \in X : \{v, w\} \in E\}.$$

Lemma 1.2.5. *Dato un qualsiasi $S \subseteq V$, vale*

$$\mu_{n-1}\frac{|S||S^c|}{n} \leq e(S) \leq \mu_1\frac{|S||S^c|}{n}.$$

Questo lemma ([6] e [5]) ci dice che se gli autovalori del laplaciano sono circa uguali, allora tutti i sottoinsiemi di vertici con la stessa cardinalità hanno un numero di archi uscenti quasi uguale.

Infine citiamo un risultato importante per quanto riguarda la costruzione di Grafi. ([3], [13] e [16])

Definizione 1.2.6 (Threshold Graph)**.**
Un grafo G è un *threshold graph* se può essere costruito usando solo le seguenti operazioni, partendo dal grafo vuoto:

- aggiungere un nodo isolato al grafo.
- aggiungere un nodo al grafo e collegarlo a tutti quelli già presenti.

Questi tipi di grafi hanno particolari proprietà che ne rendono quasi immediato il calcolo dello spettro:

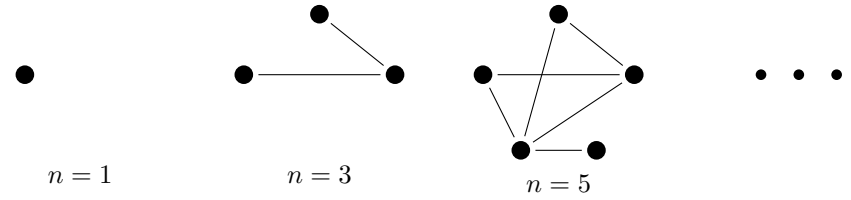
Definizione 1.2.7 (Sequenza Duale)**.**
Data $a = (a_1, a_2, \ldots, a_n)$ una sequenza di numeri naturali non crescente, allora la sua *sequenza duale* è data da

$$a^* = (b_1, b_2, \ldots, b_{a_1}), \text{ dove } b_i = |\{a_j : a_j \geq i\}|.$$

Teorema 1.2.1. *Se G è un threshold graph, con d la sequenza dei gradi dei suoi nodi, allora $\mu = d^*$, dove μ è la sequenza degli autovalori non nulli del laplaciano.*

Un esempio di grafo Threshold è un grafo i cui gradi dei nodi comprendono tutti i numeri da 1 a $n-1$, infatti si possono ricostruire le mosse usate per costruirlo, al contrario: se si rimuove il nodo con grado $n-1$ insieme agli archi che partono da esso, tutti i gradi dei nodi calano di uno, pertanto il nodo di grado 1 diviene di grado zero, e si può rimuovere, ottenendo un grafo con $n-2$ nodi e tutti i gradi da 1 a $n-3$. Proseguendo, otteniamo un grafo con 1 o 2 nodi, che sono in ogni caso grafi Threshold.

Ecco come appaiono i grafi appena descritti:



Data $a = (a_1, a_2, \ldots, a_n)$ una sequenza di numeri naturali, allora

Definizione 1.2.8 (Sequenza Grafica)**.**

Si dice che a è una *sequenze grafica* se esiste un grafo con n nodi tale che abbia proprio a come sequenza di gradi dei nodi.

Poniamo una relazione d'ordine parziale sulle sequenze grafiche lunghe n:

$$a = (a_1, a_2, \ldots, a_n) \lhd (b_1, b_2, \ldots, b_n) = b$$
$$\Updownarrow$$
$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i, \qquad \sum_{i=1}^{k} a_i \leq \sum_{i=1}^{k} b_i \quad \forall k < n.$$

Diciamo che un grafo è un *Maximal Degree Graph* se il suo vettore dei gradi è maggiore di qualsiasi sequenza grafica a lui confrontabile.

Tutto questo porta al seguente teorema:

Teorema 1.2.2. *G è un grafo threshold sse è un Maximal Degree Graph.*

1.3 Spettro della Matrice di Adiacenza

La matrice di adiacenza è quella che più comunemente viene usata per rappresentare un grafo, in quanto è a entrate binarie, e dunque meglio gestibile da un calcolatore.

In generale, data una matrice A simmetrica $n \times n$, il *grafo associato* è un grafo su n nodi, in cui i nodi i, j sono collegati sse $A_{ij} \neq 0$. Se inoltre il grafo generato è connesso, allora la matrice di partenza si dice *irriducibile*.

Di conseguenza, se prendiamo un grafo G connesso, la matrice di adiacenza (come anche il laplaciano e la matrice di transizione) è irriducibile, e ciò la rende soggetta al teorema di Perron-Frobenius[30].

Teorema 1.3.1 (Perron-Frobenius)**.**
Se A è una matrice irriducibile reale, con raggio spettrale $\rho(A) = \max\{|\lambda| : \lambda$ autovalore di $A\}$, e con tutte le entrate nonnegative, allora

- *$\rho(A)$ è un autovalore semplice.*
- *esiste un autovettore v relativo a $\rho(A)$ con tutte le componenti positive.*
- *gli unici autovettori con tutte le componenti positive sono relativi a $\rho(A)$.*

La struttura del grafo influisce sullo spettro in vari modi. Chiamiamo $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ gli autovalori di A_G, e diciamo che un grafo è

bipartito se $L \coprod R = V$, con L e R insiemi di nodi tali che tutti gli archi partano da L e finiscano in R.

completo se $E = \{\{i, j\} : 1 \leq i, j \leq n\}$, ossia se ogni coppia di nodi è collegata da un arco.

Queste sono proprietà del grafo che è possibile dedurre dallo spettro di A_G, in quanto

Teorema 1.3.2. *[1]Dato G grafo, sono equivalenti:*

- *G è un grafo bipartito.*
- *$\lambda_n = -\lambda_1$.*
- *lo spettro di A_G è simmetrico, ossia $\lambda_k = -\lambda_{n-k+1} \quad \forall k$.*

Lemma 1.3.1. *A_G ha esattamente due autovalori distinti sse G è completo, ed in questo caso avremo*

$$\lambda_1 = n - 1 \qquad \lambda_i = -1 \quad \forall i > 1.$$

Tr gli altri utilizzi dello spettro di A_G annoveriamo anche contributi all'approssimazione di quantità difficili da calcolare, quali la massima stazza di una *cricca*, ossia un sottografo completo di G, o anche la cardinalità del suo taglio massimo.

Lemma 1.3.2. *il numero di nodi di un sottografo completo di G è al massimo $\lambda_1 + 1$.*

Lemma 1.3.3. *il taglio massimo del grafo è inferiore a $m/2 - (n/4)\lambda_n$.*

Passiamo ora all'argomento principale che affronteremo: i grafi expanders.

Capitolo 2

Expanders

Una particolare classe di grafi, detti expanders, giocano un ruolo importante in molte applicazioni della teoria dei grafi, in particolare in informatica. Essi sono studiati da diverse comunità di matematici, e sono utili a diversi scopi, per esempio

- studiare algoritmi atti a trovare il minimo taglio in un grafo, che sono fondamentali nell'area degli algoritmi di approssimazione, e sono impiegati in vari campi, come la segmentazione di immagini, o nelle procedure di divide et impera.
- le costruzioni esplicite di buoni grafi expander hanno applicazioni nella crittografia, nella derandomizzazione di algoritmi, e in definire strutture di dati efficienti; molte di queste costruzioni sono algebriche, e sono collegate ad aspetti profondi di teoria dei gruppi.
- stabilire la velocità di convergenza di metodi probabilistici collegati alle Catene di Markov e agli algoritmi basati sul modello Monte-Carlo.

Dopo aver definito e studiato alcune proprietà di questi grafi, torneremo a discutere più approfonditamente di alcune delle loro applicazioni.

2.1 Definizione

Definiamo ora il parametro che ci accompagnerà da qui alla fine:

Definizione 2.1.1 (Numero Isoperimetrico o Edge Expansion Ratio)**.**
Dato G grafo con $|V| = n$, allora definiamo il *numero isoperimetrico di* G come

$$i(G) = \min_{0<|X|\leq n/2} \frac{e(X)}{|X|}.$$

Come dice il nome *Edge Expansion Ratio*, $i(G)$ dà informazioni su quanto velocemente riusciamo a muoverci sul grafo, in quanto per ogni insieme S di cardinalità piccola avremo $e(S) \geq i(G) \cdot |S|$. Ciò vuol dire che più è grande questa costante, più riusciamo a raggiungere i nodi del grafo in pochi passi. (Per un'analisi più approfondita, vedere il paragrafo Camminata Aleatoria).

Questo parametro, per esempio, viene usato nelle reti di network per testare la loro sicurezza e robustezza, o anche come indice di integrità della rete se alcuni collegamenti venissero compromessi, come mostra il seguente lemma:

Lemma 2.1.1. *Dato un grafo G, con $i(G) = r$, e rimuoviamo k archi, il grafo che rimane contiene una componente connessa con almeno $n - \frac{2k}{r}$ nodi.*

Dimostrazione.

Siano $c_1, c_2, \ldots, c_p$ le componenti connesse di G a cui sono stati tolti i k archi, con $|c_1| \geq |c_2| \geq \cdots \geq |c_p|$. Dato che originariamente G era connesso, allora abbiamo tolto almeno gli archi tra i c_i, e dunque

$$k \geq \sum_{p \geq i > j > 0} e(c_i, c_j) = \frac{1}{2} \sum_{p \geq i > 0} e(c_i).$$

Se ora $p = 1$, allora il lemma è banalmente vero. Se invece $p > 1$, allora $\forall i > 1 \quad |c_i| \leq n/2$, e dunque

$$k \geq \frac{1}{2} \sum_{p \geq i > 0} e(c_i) \geq \frac{1}{2} \sum_{p \geq i > 1} e(c_i) > \frac{r}{2} \sum_{p \geq i > 1} |c_i| = \frac{r}{2}(n - |c_1|) \implies |c_1| > n - \frac{2k}{r}.$$

□

Un altro utilizzo lo troviamo nell'analisi di reti come quelle dei social network, in cui è importante identificare gruppi di persone che condividono informazioni comuni, quali interessi, amicizie o anche solo luoghi di nascita.

Un esempio è dato dalle mappe terrestri, in cui la maggioranza delle connessioni stradali si registrano all'interno delle città, mentre vi sono pochi collegamenti tra un centro urbano ed un altro. Ciò vuol dire che data una mappa, se volessimo identificare le aree urbane, potremmo concentrarci nel cercare sottoinsiemi dei nodi con un rapporto tra il taglio relativo e la cardinalità del sottoinsieme molto basso.

In generale, questa tecnica permette di riconoscere sottografi ad alta densità di archi dentro a grafi sparsi.

Questo indice, inoltre, ci permette di definire cos'è un grafo *c-expander*:

Definizione 2.1.2 (Expander)**.**

Dato un reale $c > 0$, un grafo G è detto *c-expander* se $i(G) \geq c$.

Più il numero isoperimetrica di un grafo è alta, più viene detto che è un buon expander. Molti sono i risultati che riguardano le costruzioni di buoni expanders, sia esatte che randomizzate.

Spesso, però, è più utile costruire una famiglia di expander, ossia dei grafi sempre più grandi che mantengano buone proprietà di espansione, e richiediamo ai grafi di essere anche *d-regolari*, ossia che tutti i nodi abbiano grado d.

Da qui nasce la seguente definizione:

Definizione 2.1.3 (Famiglia di grafi (d, c)-Expanders)**.**

Una *Famiglia di grafi (d, c)-Expanders* è una successione di grafi c-expanders e d-regolari $\{G_n\}_{n \in \mathbb{N}}$ con $\lim_{n \to +\infty} |V_n| = +\infty$.

Non è facile costruire una tale famiglia di grafi, ed è ancora più complicato farlo in modo che siano computabili in poco tempo. Ciò è dovuto anche al fatto che calcolare la costante isoperimetrica di un grafo qualunque è un problema NP-hard, ma al contempo anche molto importante, dunque nel corso degli anni si sono cercate diversi metodi per stimarla o algoritmi per approssimarla.

Tutte le definizioni date in questa sezione sono puramente combinatorie, quindi nel prossimo paragrafo studieremo come queste si relazionano con le proprietà spettrali dei grafi.

2.2 Gap Spettrale e Cheeger

La condizione di d-regolarità fa sì che le relazioni tra le tre matrici descritte sopra diventino molto semplici:

$$L_G = dI - A_G, \qquad P_G = \frac{1}{d} A_G.$$

Quindi possiamo studiare indistintamente lo spettro di una delle tre, e siamo in grado di ricavarne lo spettro delle altre due immediatamente.

Scegliamo quindi di concentrarci sullo spettro di A_G, chiamando i suoi autovalori, come sopra, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

Notiamo innanzitutto che $\lambda_1 = d$, corrispondente all'unico autovalore zero del laplaciano, e all'autovettore $e = (1, 1, \ldots, 1)$. Ricordandoci che $a(G)$ è l'autovalore non nullo più piccolo del laplaciano, avremo $\lambda_2 = d - a(G)$, o anche $a(G) = \lambda_1 - \lambda_2$.

Definizione 2.2.1 (Gap Spettrale)**.**
Dato G grafo, e A_G la sua matrice di adiacenza, allora $\lambda_1 - \lambda_2$ è il *gap spettrale* del grafo.

In questo particolare caso, gap spettrale e connettività algebrica coincidono.

Nota: più in generale, la definizione di gap spettrale di una matrice diagonalizzabile è la differenza tra l'autovalore più grande e il secondo. Se A è una matrice con autovalori $a_1 \geq a_2 \geq \cdots \geq a_n$, allora la matrice $a_1 I - A$ avrà tutti gli autovalori nonnegativi, con il più piccolo autovalore non nullo pari a $a_1 - a_2$. Per questo, quando la matrice ha solo autovalori non negativi, si dice che il gap spettrale è il più piccolo autovalore non nullo. In questo caso possiamo scegliere una delle due definizioni indistintamente.

Ricordiamo che stiamo lavorando solo con grafi connessi, quindi $a(G) \neq 0$, e di conseguenza anche il gap spettrale non è nullo.

Il laplaciano di un grafo è anche interpretabile come la discretizzazione dell'operatore laplaciano su varietà differenziabili. I risultati in teoria dei grafi e in geometria spettrale Riemanniana sono dunque collegati, e permettono di ottenere nuovi risultati in entrambi i campi.

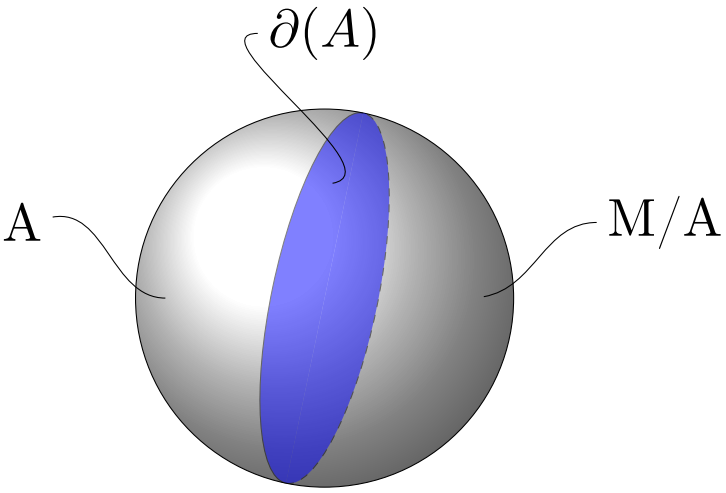
Il concetto di costante isoperimetrica, per esempio, è presente anche in geometria riemanniana, ed è legato al più piccolo autovalore non nullo dell'operatore laplaciano tramite la Disuguaglianza di Cheeger.

Per questo motivo, la costante isoperimetrica di una varietà è anche conosciuta col nome di *Costante di Cheeger*.

Definizione 2.2.2 (Costante di Cheeger)**.**
Data M una varietà compatta Riemanniana n-dimensionale, la sua costante di Cheeger è

$$h(M) := \inf_A \mu_{n-1}(\partial A)/\min(\mu_n(A), \mu_n(M-A)),$$

dove A varia tra tutti i sottoinsiemi aperti di M, e ∂A è la frontiera di A. μ_n e μ_{n-1} sono le misure di Lebesgue n e $n-1$ dimensionali.



L'idea qui è partizionare M in due parti, e considerare il rapporto tra la misura della frontiera di A e il minimo tra i due volumi in cui M è stato diviso.

Il laplaciano è un operatore lineare differenziale definito da funzioni reali $f : M \to \mathbb{R}$, come $\Delta(f) = div(grad(f))$. Si può mostrare che tutti gli autovalori di Δ sono nonnegativi, e che il minimo è zero, corrispondente alla funzione costante.

Con questi presupposti, possiamo finalmente citare il Teorema di Cheeger[31]:

Teorema 2.2.1 (Teorema di Cheeger)**.** *Data M una varietà Riemanniana compatta, e λ il più piccolo autovalore positivo del suo laplaciano, allora*

$$\lambda \geq h(M)^2/4.$$

Il risultato più importante che lega il gap spettrale con la definizione di expander, tratto da [8], è una versione discreta di questo teorema. Infatti possiamo discretizzare l'operatore laplaciano della varietà riemanniana, rappresentando lo spazio come una griglia di nodi.

In questo contesto, il volume della varietà si trasforma nel numero di nodi inclusi in essa, mentre la superficie della frontiera sarà la cardinalità del taglio determinato da quei nodi, ottenendo così la seguente diseguaglianza:

Teorema 2.2.2 (Disuguaglianza Discreta di Cheeger)**.**
Dato G grafo d-regolare, sia $i(G)$ la sua costante isoperimetrica, e $a(G)$ il suo gap spettrale. Se $n>3$, allora

$$\frac{a(G)}{2} \leq i(G) \leq \sqrt{a(G)(2d - a(G))}.$$

Dividiamo la dimostrazione di questo teorema in due parti:

Parte facile

$$\frac{a(G)}{2} \leq i(G).$$

Per dimostrare questa parte, usiamo un risultato di algebra lineare:

Lemma 2.2.1 (Courant-Fisher)**.**
Sia A una matrice simmetrica, e v l'autovettore relativo all'autovalore più grande di A. Se λ è il secondo autovalore più grande di A, allora

$$\lambda = \max\{x^t A x : \|x\|_2 = 1, x \perp v\}.$$

Cerchiamo quindi un vettore ortogonale a e, tale che il suo *Quoziente di Rayleigh* sia

$$R_{A_G}(x) = \frac{x^t A_G x}{\|x\|_2^2} \geq d - 2i(G),$$

così abbiamo automaticamente dimostrato che $\lambda_2 \geq d - 2i(G)$, e dunque questa parte del teorema.

Sia $S \subseteq V$ tale che $i(G) = e(S)/|S|$, e poniamo che $k = |S|$. Se 1_S rappresenta il vettore che ha 1 sui nodi di S, zero altrimenti, consideriamo $x = (n-k)1_S - k1_{S^c}$. Sicuramente $x^t e = k(n-k) - k(n-k) = 0$. valutiamo quindi il suo quoziente di Rayleigh.

$$\|x\|_2^2 = k^2(n-k) + (n-k)^2 k = nk(n-k).$$
$$x^t A_G x = 2e(S,S)(n-k)^2 + 2e(S^c,S^c)n^2 + 2n(n-k)e(S).$$

Dato che il grafo è d regolare, inoltre

$$2e(S,S) = dk - e(S) \qquad 2e(S^c,S^c) = d(n-k) - e(S).$$

Quindi avremo

$$R_{A_G}(x) = \frac{x^t A_G x}{\|x\|_2^2} = \frac{ndk(n-k) - n^2 e(S)}{nk(n-k)} = d - \frac{n}{n-k}\frac{e(S)}{k} \geq d - 2i(G).$$

Parte difficile

$$i(G) \leq \sqrt{a(G)(2d - a(G))}.$$

Per prima cosa, poniamo che G sia un grafo completo. Ciò vuol dire che $d = n-1$, ed è facile trovare che tutti gli autovalori non nulli del laplaciano sono n, con autovettori associati $e_i - e_{i+1}$, dove gli e_i sono la base canonica di $\mathbb{R}^n$. Inoltre dato che $|S| = k \implies e(S) = k(n-k)$, allora $e(S)/S = n-k$.

$$i(G) = \min_{0<|S|\leq n/2} n - |S| = \left\lceil \frac{n}{2} \right\rceil \leq \sqrt{n(n-2)} = \sqrt{a(G)(2d - a(G))}.$$

Da notare che la diseguaglianza sopra vale solo se $n \geq 4$.

Prima di procedere con il caso non completo, premettiamo un altro risultato di algebra lineare:

Teorema 2.2.3 (Monotonicity Theorem)**.** *[32]*

Data A matrice reale simmetrica di autovalori $\{\lambda_i\}_{i\leq n}$ ordinati in maniera non crescente, e sia B un minore simmetrico di autovalori $\{\mu_i\}_{i\leq n-k}$ ordinati anch'esso in maniera non crescente. Allora vale la seguente relazione:

$$\lambda_i \geq \mu_i \geq \lambda_{i+k} \quad \forall i \leq n-k.$$

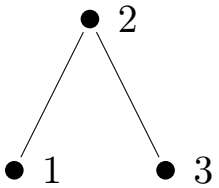
Preso adesso G, avremo che

Lemma 2.2.2. *Dato G connesso e non completo, con almeno 3 nodi, allora G contiene sicuramente un sottografo composto da 3 nodi e 2 archi.*

Dimostrazione.

Dato che G non è completo, allora esistono v e w nodi la cui distanza è maggiore di 1, ma dato che il grafo è connesso, possiamo prendere $\{v, a_1, \ldots, a_k, w\}$ un cammino di nodi minimale, ossia con $k+1$ pari alla distanza tra v e w, e $k > 0$. Considerato a_2, con $a_2 = w$ se $k = 1$, avremo che v e a_2 non possono essere collegati, poichè altrimenti il cammino $v, a_2, \ldots, a_k, w$ sarebbe più corto di quello originale. Ciò vuol dire che $\{v, a_1, a_2\}$, è un sottografo con 3 nodi e 2 archi, in quanto $\{v, a_1\}$ e $\{a_1, a_2\}$ sono archi. □

Questo sottografo G' identifica un minore simmetrico nella matrice di adiacenza, che ha per autovalori $\sqrt{2}, 0, -\sqrt{2}$, in quanto avremo



grafo G'

$$A_{G'} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$p_{A_{G'}}(t) = -t^3 + 2t$$

Per il Monotonicity Theorem, dunque, $\lambda_2 \geq 0$, o equivalentemente $a(G) \leq d$. Sia ora v autovettore relativo a $a(G)$ nella matrice laplaciana, in modo che

$$|W| = |\{i \in V : v_i > 0\}| \leq \frac{n}{2}.$$

Notiamo che è sempre possibile trovare un tale autovettore poiché dato x qualunque o x o $-x$ ha questa proprietà. Inoltre vale

$$Lv = a(G)v \implies (Lv)_i = a(G)v_i = dv_i - \sum_{j\in E(i)} v_j \implies \sum_{j\in E(i)} v_j = (d - a(G))v_i.$$

Se definiamo un nuovo vettore $g = v1_W$ e denotiamo con $EW = E(W, W)$, il risultato sopra ci dice che

$$\begin{aligned} a(G)\sum_{i\in W} v_i^2 &= \sum_{i\in W}\left(dv_i - \sum_{j\in E(i)} v_j\right) v_i \\ &= \sum_{i\in W}\sum_{\{i,j\}\in E} (v_i - v_j)v_i \\ &= \sum_{\{i,j\}\in EW} [(v_i - v_j)v_i + (v_j - v_i)v_j] + \sum_{\{i,j\}\in E(W)} (v_i - v_j)v_i \\ &= \sum_{\{i,j\}\in E} (g_i - g_j)^2 - \sum_{\{i,j\}\in E(W)} v_i v_j. \end{aligned}$$

Similmente, otteniamo

$$\begin{aligned} (2d - a(G))\sum_{i\in W} v_i^2 &= \sum_{i\in W}\left(dv_i + \sum_{j\in E(i)} v_j\right) v_i \\ &= \sum_{i\in W}\sum_{\{i,j\}\in E} (v_i + v_j)v_i \\ &= \sum_{\{i,j\}\in EW} [(v_i + v_j)v_i + (v_j + v_i)v_j] + \sum_{\{i,j\}\in E(W)} (v_i + v_j)v_i \\ &= \sum_{\{i,j\}\in E} (g_i + g_j)^2 + \sum_{\{i,j\}\in E(W)} v_i v_j. \end{aligned}$$

Combinando queste due, otteniamo

$$\begin{aligned} a(G)(2d - a(G))\left(\sum_{i\in W} v_i^2\right)^2 &= \sum_{\{i,j\}\in E} (g_i + g_j)^2 \sum_{\{i,j\}\in E} (g_i - g_j)^2 \\ &\quad - \left(\sum_{\{i,j\}\in E(W)} v_i v_j\right)\left(4\sum_{\{i,j\}\in EW} v_i v_j + \sum_{\{i,j\}\in E(W)} v_i v_j\right). \end{aligned}$$

Notiamo innanzitutto che $\sum_{\{i,j\}\in E(W)} v_i v_j \leq 0$ poiché uno tra i e j in ogni addendo non sta in W. Inoltre,

$$\begin{aligned} 4\sum_{\{i,j\}\in EW} v_i v_j + \sum_{\{i,j\}\in E(W)} v_i v_j = \\ 2\sum_{\{i,j\}\in EW} v_i v_j + \sum_{i\in W} y_i \sum_{j\in E(i)} v_j = \\ 2\sum_{\{i,j\}\in EW} v_i v_j + \sum_{i\in W} (d - a(G))v_i^2 \geq 0, \end{aligned}$$

dove l'ultima disuguaglianza vale grazie a $d \geq a(G)$. Per concludere, ci conviene ora definire $0 = t_0 < t_1 < \cdots < t_m$ i diversi valori del vettore g, con $m \leq n/2$, e $V_k = \{i \in V | g_i \geq t_k\}$. Usando Cauchy-Schwartz, otteniamo quindi

$$\begin{aligned}
&\sqrt{a(G)(2d - a(G))} \sum_{i \in W} v_i^2 \geq \sqrt{\sum_{\{i,j\} \in E} (g_i + g_j)^2 \sum_{\{i,j\} \in E} (g_i - g_j)^2} \\
&\geq \sum_{\substack{\{i,j\} \in E \\ g_i > g_j}} (g_i^2 - g_j^2) = \sum_{k=1}^{m} e(V_k)(t_k^2 - t_{k-1}^2) \geq i(G) \sum_{k=1}^{m} |V_k|(t_k^2 - t_{k-1}^2) \\
&= i(G) \sum_{k=0}^{m} t_k^2 (|V_k| - |V_{k+1}|) = i(G) \sum_{i \in V} g_i^2 = i(G) \sum_{i \in W} v_i^2,
\end{aligned}$$

che conclude la dimostrazione.

□

Nei prossimi capitoli vedremo qualche applicazione e costruzione classica di questi grafi expanders.

Capitolo 3

Random Walk

Nella prima sezione abbiamo definito cos'è la matrice di transizione P_G associata ad un grafo. In probabilità, questa assume un ruolo molto importante, perché rappresenta la distribuzione di probabilità associata ad una *Passeggiata Aleatoria* sul grafo.

Immaginiamo infatti di partire da un nodo del grafo, e ad ogni passo scegliamo di muoverci lungo uno degli archi, scelto con probabilità uniforme tra quelli uscenti dal nodo. Avremo che $(P_G)_{ij}$ è esattamente la probabilità che partendo dal nodo i arriviamo al nodo j in un passo. Se inoltre calcoliamo le potenze di P_G otteniamo più in generale che

Lemma 3.0.3. *Dato un grafo G, avremo che la componente i, j della matrice P_G^k rappresenta la probabilità che in una passeggiata aleatoria sul grafo, se partiamo dal nodo i arriviamo nel nodo j con esattamente k passi.*

Dimostrazione.

Dimostriamolo per induzione.

Per $k = 1$ è vero, poiché la probabilità che dal nodo i arriviamo al nodo j con un passo è zero se i, j non sono collegati da un arco, mentre è $1/d_i$ se lo sono, e questo è esattamente $(P_G)_{ij}$.

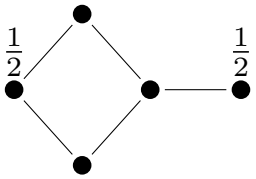
Chiamiamo ora $P_k(i, j)$ la probabilità che in una passeggiata aleatoria sul grafo, se partiamo dal nodo i arriviamo nel nodo j con esattamente k passi. Per ipotesi induttiva, avremo che $(P_G^k)_{ij} = P_k(i, j)$. Preso il nodo j, per calcolare la probabilità che dal nodo i arriviamo al nodo j con esattamente $k + 1$ passi, sommiamo le probabilità dei cammini di k passi che partono da i, e che poi con un ulteriore passo terminano in j. Avremo

$$P_{k+1}(i, j) = \sum_{r \in V} P_k(i, r) P_1(r, j) = \sum_{r \in V} (P_G^k)_{ir} (P_G)_{rj} = (P_G^{k+1})_{ij},$$

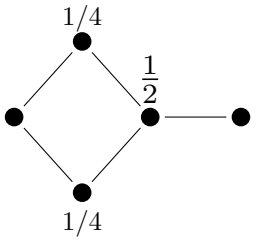
dunque il passo induttivo è vero. □

Sia ora π_0 un vettore ad entrate non negative in $\mathbb{R}^n$ tale che la somma delle sue componenti sia 1. Questo rappresenta una distribuzione di probabilità associata ai nodi del grafo. Se per esempio $\pi_0 = e_1$ vettore della base canonica, vuol dire che ci troviamo quasi certamente sul nodo etichettato da 1.

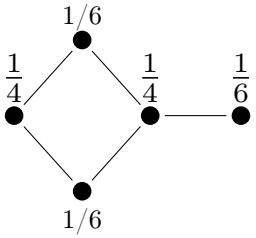
Applicare P_G a π_0 rappresenta il fare un passo sul grafo. Se $\pi_k^t = \pi_0^t P_G^k$, allora la componente i-esima di π_k rappresenta la probabilità di trovarsi nel nodo i dopo esattamente k passi, partendo dalla configurazione iniziale π_0.



$$\pi_0 = \begin{pmatrix} \frac{1}{2} \\ 0 \\ 0 \\ 0 \\ \frac{1}{2} \end{pmatrix}$$



$$\pi_1 = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 1 \\ \frac{1}{2} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ 0 \\ 0 \\ 0 \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \\ 0 \end{pmatrix}$$



$$\pi_2 = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 1 \\ \frac{1}{2} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{6} \\ \frac{1}{4} \\ \frac{1}{6} \\ \frac{1}{6} \end{pmatrix}$$

Tra le distribuzioni, avremo la *distribuzione uniforme* $\pi = (1/n)e$, ossia quella in cui è equiprobabile trovarsi in ogni nodo, e, se esistono, le *distribuzioni stazionarie*, ossia quelle in cui facendo un passo della camminata aleatoria, la probabilità di trovarsi in ogni singolo nodo non varia. In formule π è una distribuzione stazionaria sse $\pi^t P_G = \pi^t$. Visto che dunque una distribuzione è stazionaria sse è un autovettore sinistro di P_G relativo all'autovalore 1, e che P_G è simile alla sua trasposta, allora concentriamoci sullo studio del suo spettro.

3.1 Proprietà di P_G

Richiamiamo un risultato di algebra lineare, dovuto a Gerschgorin:

Teorema 3.1.1 (Cerchi di Gerschgorin)**.**
Data A una matrice complessa $n \times n$, allora siano

$$a_i = A_{ii} \quad r_i = \sum_{1 \le j \le n}^{j \neq i} |A_{ij}| \quad \forall i : 1 \le i \le n.$$

Definiamo i ***Cerchi di Gerschgorin*** *relativi alla matrice A come*

$$D_i = D(a_i, r_i) = \{y \in \mathbb{C} : |y - a_i| \leq r_i\}.$$

Avremo dunque che tutti gli autovalori di A appartengono ad almeno uno dei cerchi D_i*.*

Applicando il teorema a P_G, notiamo che il suo raggio spettrale è 1, poiché i centri dei suoi cerchi di Gerschorin sono tutti l'origine del piano complesso, e il raggio è pari ad 1. In particolare, 1 è sempre un autovalore e la matrice è stocastica, ossia abbiamo $e = (1, 1, \ldots, 1)$ autovettore relativo all'autovalore uno.

Dato G un grafo connesso, P_G è irriducibile e nonnegativa, dunque anche per lei vale il teorema 1.3.1 di Perron-Frobenius, pertanto l'autovalore 1 sarà semplice in questo caso. Ne segue che esiste ed è unica la distribuzione stazionaria.

Un'altra questione importante riguardo lo spettro della matrice di transizione, è la presenza o meno dell'autovalore -1. Se infatti questo non sarà presente, avremo che ogni autovalore diverso da 1 avrà modulo minore di uno, e questo avrà grande importanza nel prossimo capitolo.

La presenza o meno di questo autovalore, determina se il grafo sia bipartito o meno, ed infatti avremo che

Lemma 3.1.1. *Dato G grafo connesso, G è bipartito sse* -1 *è un autovalore di* P_G*.*

Dimostrazione.

Poniamo dapprima che G sia bipartito in L e R. Denotiamo con 1_L e 1_R i vettori binari le cui componenti sono 1 solo su L e su R rispettivamente. Avremo che

$$P_G(1_L - 1_R) = P_G 1_L - P_G 1_R = 1_R - 1_L = -(1_L - 1_R),$$

dunque $1_L - 1_R$ è un autovettore relativo all'autovalore -1.

Poniamo ora che esista l'autovalore -1, e che v sia un suo autovettore. Si avrebbe

$$P_G v = -v \implies -v_k = (P_G v)_k = \sum_{1 \geq j \geq n} (P_G)_{kj} v_j = \frac{1}{d_k} \sum_{\{k,j\} \in E} v_j.$$

$$\sum_{1 \leq k \leq n} -d_k v_k = \sum_{1 \leq k \leq n} \sum_{\{k,j\} \in E} v_j = \sum_{1 \leq k \leq n} d_k v_k \implies \sum_{1 \leq k \leq n} d_k v_k = 0.$$

Da ciò ricaviamo che v ha almeno una componente positiva ed una negativa. Possiamo quindi partizionare i nodi in I^+, I^- e I^0, secondo la regola che se v_i è positivo, allora $i \in I^+$, se è negativo $i \in I^-$, e se è nullo $i \in I^0$. Otteniamo

$$\sum_{k \in I^+} d_k v_k = -\sum_{k \in I^-} d_k v_k = \sum_{k \in I^-} \sum_{\{k,j\} \in E} v_j = \sum_{1 \leq k \leq n} v_k |N(k) \cap I^-|$$

$$\implies \sum_{k \in I^+} d_k v_k \leq \sum_{k \in I^+} v_k |N(k) \cap I^-| \implies \sum_{k \in I^+} (d_k - |N(k) \cap I^-|) v_k \leq 0.$$

Dato che i v_k con $k \in I^+$ sono strettamente positivi, e che $d_k \geq |N(k) \cap I^-|$, allora $d_k = |N(k) \cap I^-| \quad \forall k \in I^+$, ovvero tutti i vicini dei nodi in I^+ stanno in I^-.

Con un ragionamento simmetrico, riusciamo a dire anche che tutti i vicini dei nodi in I^- stanno in I^+, e visto che il grafo è connesso, allora I^0 deve essere vuoto. Visto che invece I^+ e I^- non sono vuoti, abbiamo partizionato V in due sottoinsiemi senza archi interni, e quindi abbiamo dimostrato che G è bipartito. □

Ora vedremo come queste considerazioni si relazionano con la costante isoperimetrica dei grafi.

3.2 Expander e Velocità di Convergenza

Consideriamo per un attimo un grafo G *orientato*, ossia in cui per ogni arco sono definiti un nodo di partenza ed uno di arrivo. Poniamo che sia anche connesso, ovvero che per ogni coppia di nodi distinti u, v esistono degli archi orientati in E del tipo $(u, x_1), (x_1, x_2), \ldots, (x_{n-1}, x_n), (x_n, v)$.

Sotto queste ipotesi, diremo che

Definizione 3.2.1 (Ciclico)**.**

G è *ciclico* se esiste un partizionamento dei nodi $S_0, S_1, \ldots, S_{k-1}$ con $k > 0$, tali che ogni arco inizi in un S_i e finisca in S_{i+1}, con indici contati modulo k.

Diremo inoltre che è *aciclico* se non è ciclico.

Un noto teorema nella teoria delle Catene di Markov stabilisce che dato un grafo connesso orientato e aperiodico, ogni distribuzione di probabilità tende, tramite una camminata aleatoria, ad una distribuzione stazionaria.

É facile convincersi che un grafo ciclico non può avere questa proprietà: se come distribuzione di probabilità iniziale scegliamo $1_{S_0}/|S_0|$, avremo che si trasforma in un passo in $1_{S_1}/|S_1|$, quindi in $1_{S_2}/|S_2|$, e così via, fino a tornare al $1_{S_0}/|S_0|$ iniziale, per poi continuare ciclicamente, senza mai convergere a nessun vettore.

Nel nostro caso, il grafo è non diretto, dunque l'unico modo in cui possa essere ciclico nel senso sopra, è che ci siano 2 insiemi di nodi in cui V è ripartito senza collegamenti interni, ossia il grafo è bipartito.

Teorema 3.2.1. *Data π_0 una distribuzione di probabilità sui nodi di un grafo connesso non bipartito, allora $\pi_k^t = \pi_0^t P_G^k$ tende al limite alla distribuzione stazionaria.*

Dimostrazione.

Dato che il grafo è connesso e non bipartito, sappiamo già che lo spettro di P_G^t è uguale a quello di P_G, ed è $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ dove $1 = \sigma_1 > |\sigma_i| \quad \forall i > 1$. Scegliamo dunque una base di autovettori sinistri $\{v_1, v_2, \ldots, v_n\}$, relativi ai rispettivi autovalori.

Data una qualsiasi distribuzione di probabilità iniziale, possiamo scriverla in base, $\pi_0 = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$, ed avremo che

$$\pi_k^t = \pi_0^t P_G^k = (a_1 v_1 + a_2 v_2 + \cdots + a_n v_n)^t P_G^k = a_1 \sigma_1^k v_1^t + a_2 \sigma_2^k v_2^t + \cdots + a_n \sigma_n^k v_n^t,$$

ma $|\sigma_i| < 1 \implies \lim_{k\to\infty} \sigma_i^k \to 0$, dunque tutti gli autovalori minori di uno spariscono[1]. L'unico addendo che rimane è quello relativo all'autovalore 1, e si ha

$$\pi_k^t = \pi_0^t P_G^k = a_1\sigma_1^k v_1^t + a_2\sigma_2^k v_2^t + \cdots + a_n\sigma_n^k v_n^t \to a_1 v_1^t.$$

Osserviamo ora che se $i \neq 1$, allora $\lambda_i \neq 1$, e

$$\lambda_i v_i^t \cdot e = (v_i^t P_G)e = v_i^t \cdot e \implies v_i^t \cdot e = 0 \implies \pi_0^t \cdot e = a_1 v_1^t \cdot e = 1.$$

Inoltre, dato che $\pi_k \to a_i v_i$, allora $a_i v_i$ ha tutte le componenti nonnegative, ed in conclusione è la distribuzione di probabilità stazionaria, poiché è autovettore sinistro relativo all'autovalore 1.

□

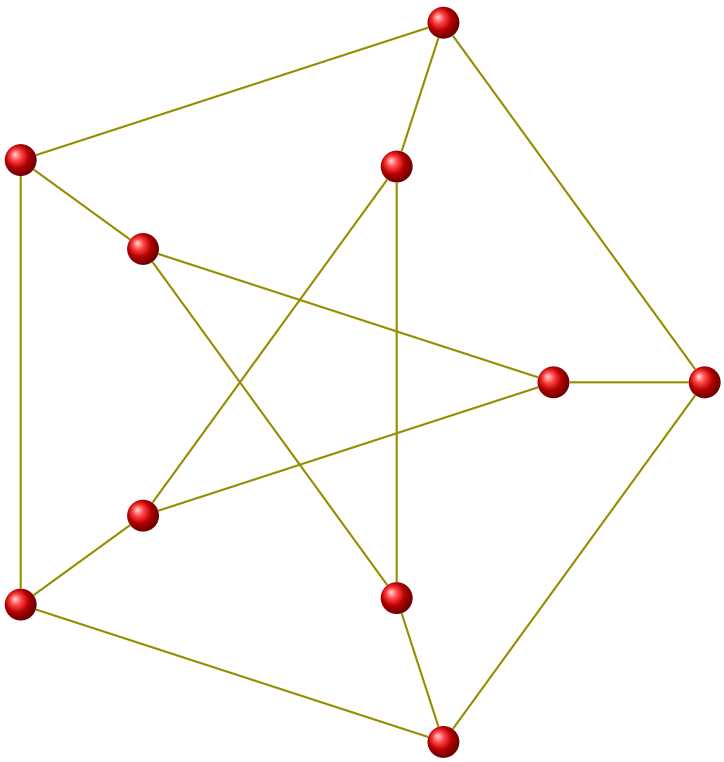
Questo risultato, non fornisce una stima della velocità di convergenza del metodo.

Definiamo quindi un nuovo tipo di grafo:

Definizione 3.2.2 ((n, d, α)-Grafo)**.**
G è un *(n, d, α)-Grafo* se è d-regolare, connesso, ha n nodi, e presi λ_i autovalori di A_G, si ha $|\lambda_2|, |\lambda_n| \leq \alpha d$, con $\alpha > 0$.

Per esempio, il grafo di Petersen raffigurato qui sotto è 3-regolare, ha 10 nodi, e la sua sequenza di autovalori è $\{3, 1, 1, 1, 1, 1, -2, -2, -2, -2\}$, dunque $\alpha \geq 2/3$.

Pertanto questo è un $(10, 3, 2/3)$-grafo:



[1] Questa procedura per trovare gli autovettori è comunemente detta *Metodo delle Potenze*

Come visto nella dimostrazione della disuguaglianza di Cheeger, nel caso di grafi d-regolari abbiamo $d - \lambda_2 = a(G)$, e se il grafo ha più di 2 nodi e non è completo, allora $\lambda_2 \geq 0$. Quindi un (n, d, α)-Grafo non completo è tale che $a(G) \geq d(1-\alpha)$, ed in virtù della disuguaglianza di Cheeger, $i(G) \geq d(1-\alpha)/2$, ovvero questi tipi di grafo sono dei buoni expanders quando la costante α è piccola.

Nel caso G sia completo, avremo comunque

$$i(G) = \left\lceil \frac{n}{2} \right\rceil \geq \frac{(n-1)(1-\alpha)}{2} = \frac{d(1-\alpha)}{2}.$$

La regolarità del grafo comporta anche che P_G sia una matrice simmetrica e doppiamente stocastica, ossia $e^t P_G = e^t$, $P_G e = e$, dunque la distribuzione stazionaria coincide con quella uniforme, ed è $(1/n)e$. Inoltre $P_G = (1/d)A_G$ comporta che il gap spettrale di P_G (differenza tra i due autovalori più grandi) sia $(d - \lambda_2)/d \geq 1 - \alpha$.

Notiamo anche che per calcolare la distribuzione di probabilità sul grafo, non c'è più bisogno di trasporre, in quanto

$$\pi_k^t = \pi_0^t P_G^k \iff \pi_k = P_G^k \pi_0.$$

Tutto questo porta ai risultati seguenti [7]:

Teorema 3.2.2. *Dato G un (n, d, α)-grafo, allora $\forall k$ intero e $\forall \pi_0$ distribuzione iniziale di probabilità , si ha*

$$\|P_G^k \pi_0 - \frac{1}{n} e\|_2 \leq \alpha^k \|\pi_0 - \frac{1}{n} e\|_2 \leq \alpha^k$$

$$\|P_G^k \pi_0 - \frac{1}{n} e\|_1 \leq \sqrt{n} \alpha^k.$$

Dimostrazione.

Procediamo per induzione. Per $k = 0$, avremo

$$\|P_G^0 \pi_0 - \frac{1}{n} e\|_2^2 = \|\pi_0 - \frac{1}{n} e\|_2^2 = \|\pi_0\|_2^2 + \|\frac{1}{n} e\|_2^2 - \frac{2}{n} \pi_0^t \cdot e = \|\pi_0\|_2^2 - \frac{1}{n} < 1,$$

dove l'ultima disuguaglianza è data dal fatto che π_0 è una distribuzione di probabilità, dunque tutte le sue componenti sono positive e al massimo 1, da cui $1 = \|\pi_0\|_1 \geq \|\pi_0\|_2^2$.

Per il passo induttivo, chiamiamo $\pi_k = P_G^k \pi_0$. Avremo

$$\|P_G^{k+1} \pi_0 - \frac{1}{n} e\|_2 = \|P_G \pi_k - \frac{1}{n} e\|_2 = \|P_G(\pi_k - \frac{1}{n} e)\|_2.$$

Ora che P_G è simmetrica, possiamo prendere una base di autovettori ortonormali ma, dalla dimostrazione procedente, sappiamo che se scriviamo una distribuzione di probabilità nella base di autovettori ortogonali, avremo che la componente relativa all'autovalore 1 dovrà essere la distribuzione stazionaria, dunque

$$\pi_k = \frac{1}{n} e + a_2 v_2 + \cdots + a_n v_n \implies \pi_k - \frac{1}{n} e = a_2 v_2 + \cdots + a_n v_n.$$

Dato che abbiamo preso v_i ortogonali tra loro, avremo

$$\|P_G(\pi_k - \frac{1}{n}e)\|_2^2 = \|P_G(a_2v_2 + \cdots + a_nv_n)\|_2^2 = \|a_2\sigma_2v_2 + \cdots + a_n\sigma_nv_n\|_2^2 =$$

$$= \|a_2\sigma_2v_2\|_2^2 + \cdots + \|a_n\sigma_nv_n\|_2^2 = |\sigma_2|^2\|a_2v_2\|_2^2 + \cdots + |\sigma_n|^2\|a_nv_n\|_2^2 \leq$$

$$\leq \alpha^2(\|a_2v_2\|_2^2 + \cdots + \|a_nv_n\|_2^2) = \alpha^2\|a_2v_2 + \cdots + a_nv_n\|_2^2 = \alpha^2\|\pi_k - \frac{1}{n}e\|_2^2,$$

il che prova la tesi:

$$\|P_G^{k+1}\pi_0 - \frac{1}{n}e\|_2 = \|P_G(\pi_k - \frac{1}{n}e)\|_2 \leq \alpha\|\pi_k - \frac{1}{n}e\|_2 \leq \alpha^{k+1}\|\pi_0 - \frac{1}{n}e\|_2$$

Infine, dato che per ogni vettore vale $\|v\|_1 \leq \sqrt{n}\|v\|_2$ usando Cauchy-Schwarz, otteniamo anche la seconda parte del teorema.

□

Un'osservazione da fare è che se $\alpha \geq 1$, allora non riusciamo a dire nulla sulla convergenza del metodo. Per esempio, se G fosse bipartito, sappiamo che il metodo non converge, ed infatti $|\lambda_n| = d$ in virtù del Teorema 1.3.2, e quindi $\alpha \geq 1$.

Questo risultato ci dice che a d fissato, più il grafo è un buon expander, più velocemente il metodo converge. L'idea di fondo è semplice: un grafo è un buon expander se riusciamo a raggiungere velocemente ogni nodo del grafo, ma allora ogni camminata aleatoria diviene presto in grado di raggiungere tutti i nodi con uguale probabilità.

Un'altra proprietà di questi particolari grafi è espressa dal seguente teorema:

Teorema 3.2.3 (Expander Mixing Lemma)**.**
Dato G un (n, d, α)-grafo, allora per ogni $S, T \subseteq V$, si ha

$$\left|\frac{d|S||T|}{n} - e(S,T)\right| \leq \alpha d\sqrt{|S||T|} \leq \alpha dn.$$

Dimostrazione.

Consideriamo $\{v_1, \ldots, v_n\}$ base di autovettori ortonormali di A_G, relativi a $d = \lambda_1 \geq \cdots \geq \lambda_n$, con $v_1 = (1/\sqrt{n})e$.

Siano 1_S e 1_T i vettori caratteristici dei sottoinsiemi S e T, e poniamo che le loro scritture in base siano $1_S = \sum_{i=1}^n a_iv_i$ e $1_T = \sum_{i=1}^n b_iv_i$. Avremo

$$e(S,T) = 1_S^t A_G 1_T = (\sum_{i=1}^n a_iv_i)^t \cdot (\sum_{i=1}^n b_i\lambda_iv_i) = \sum_{i=1}^n a_ib_i\lambda_i.$$

Sappiamo che

$$a_1 = 1_S^t \cdot \frac{1}{\sqrt{n}}e = \frac{|S|}{\sqrt{n}}, \qquad b_1 = 1_T^t \cdot \frac{1}{\sqrt{n}}e = \frac{|T|}{\sqrt{n}}$$

$$\implies e(S,T) = \frac{d|S||T|}{n} + \sum_{i=2}^n a_ib_i\lambda_i$$

$$\implies \left|\frac{d|S||T|}{n} - e(S,T)\right| = \left|\sum_{i=2}^{n} a_i b_i \lambda_i\right| \leq \alpha d \sum_{i=2}^{n} |a_i b_i|,$$

dove l'ultima diseguaglianza è data dalla definizione di (n,d,α)-grafo, secondo cui $|\lambda_2|, |\lambda_n| \leq \alpha d$. Usando Cauchy-Schwarz, otteniamo infine

$$\implies \left|\frac{d|S||T|}{n} - e(S,T)\right| \leq \alpha d \sum_{i=2}^{n} |a_i b_i| \leq \alpha d \sqrt{(\sum_{i=2}^{n} a_i^2)(\sum_{i=2}^{n} b_i^2)} \leq$$

$$\leq \alpha d \|1_S\|_2 \|1_T\|_2 = \alpha d \sqrt{|S||T|} \leq \alpha d n.$$

□

Notiamo che in un grafo d-regolare casuale, $d|S||T|/n$ è il numero medio di archi in $E(S,T)$, quindi questo simboleggia che più è piccolo α, più il grafo è vicino ad un grafo casuale. Questo è in accordo con altri risultati che vedremo in seguito, che mostreranno come la maggior parte dei grafi costruiti in maniera random siano effettivamente dei buoni expander.

Ma come si utilizza questo fatto?

3.3 Sampling

Uno dei problemi più comuni, e che trova varie applicazioni in diversi campi, è il problema del Sampling, ossia la campionatura. Enunciamolo, specializzandolo al caso dei grafi.

Problema Dato un grafo, sappiamo che esiste un set $B \subseteq V$ di nodi cattivo (bad), con $|B| = \beta n$. Esiste un metodo per essere sicuri di trovare un nodo al di fuori di B?

Prima di addentrarci nel metodo di risoluzione effettivo, utilizziamo il Mixing Lemma per dare un'intuizione della soluzione. Riscriviamolo quindi in questa forma:

$$\left|\frac{|S||T|}{n^2} - \frac{e(S,T)}{dn}\right| \leq \alpha$$

Consideriamo il seguente esperimento: scegliamo due nodi con qualche criterio, e calcoliamo la probabilità che il primo stia in S e il secondo in T.

- Se scegliamo i, j a caso tra tutti i nodi, allora la probabilità di successo è $|S||T|/n^2$.
- Se scegliamo i, j a caso tra gli elementi di E, ossia in modo che siano estremi di un arco, allora la probabilità di successo è $e(S,T)/nd$.

Il mixing lemma ci dice esattamente che queste due quantità differiscono di una costante, che possiamo porre piccola. Ciò vuol dire che dato un buon grafo expander, scegliere un cammino aleatorio di lunghezza 1, o due nodi a caso, è circa equivalente. Il risultato principale che presentiamo, generalizza questa intuizione a cammini di lunghezza arbitraria:

Teorema 3.3.1. *Dato un cammino aleatorio di lunghezza k su G (n, d, α)-grafo, sia (B, k) l'evento in cui tutti i nodi del cammino stiano in B, con $|B| = \beta n$. Allora*

$$Prob[(B, k)] \leq (\beta + \alpha)^k$$

. Se inoltre $\beta > 6\alpha$, allora

$$\beta(\beta - 2\alpha)^k \leq Prob[(B, k)] \leq \beta(\beta + 2\alpha)^k$$

.

Dimostrazione.
Dimostriamo solo la prima parte del teorema, $Prob[(B, k)] \leq (\beta + \alpha)^k$.

Sia $P = P_B$ matrice di proiezione sui nodi di B, ossia $P_{ij} = 1$ se $i = j \in B$, zero altrimenti.

Notiamo che PP_GP ritaglia da P_G il minore relativo a B, e come prima, si dimostra che $P_k = (PP_GP)^k$ è una matrice simmetrica tale che, presi due nodi in B i, j, si ha che $(P_k)_{ij}$ è la probabilità di arrivare da i a j in k passi, attraversando solo nodi in B.

Dato che un cammino aleatorio parte da ogni nodo con uguale probabilità, avremo che

$$Prob[(B, k)] = \left\| P_k \frac{1}{n} e \right\|_1 = \left\| (PP_GP)^k \frac{1}{n} e \right\|_1 .$$

Sia ora v un qualsiasi vettore a componenti nonnegative, e dimostriamo che

$$\|PP_GPv\|_2 \leq (\beta + \alpha)\|v\|_2.$$

Dato che $P^2 = P$, se lo dimostriamo per Pv abbiamo finito, poiché allora

$$\|PP_GPv\|_2 = \|PP_GP(Pv)\|_2 \leq (\beta + \alpha)\|Pv\|_2 \leq (\beta + \alpha)\|v\|_2,$$

quindi poniamo v con componenti non nulle solo su B. Inoltre, possiamo prenderlo normalizzato per linearità del problema $\|v\|_1 = 1$.

Sotto queste ipotesi, v è una distribuzione di probabilità, e dunque si scrive come $v = (1/n)e + z$, con z ortogonale a e. Avremo dunque

$$\|PP_GPv\|_2 = \|PP_Gv\|_2 \leq \|P\frac{1}{n}e\|_2 + \|PP_Gz\|_2.$$

Usando Cauchy-Schwarz otteniamo

$$1 = \|v\|_1 \leq \sqrt{|B|}\|v\|_2 \implies \|P\frac{1}{n}e\|_2 = \frac{1}{n}\sqrt{|B|} \leq \frac{|B|}{n}\|v\|_2 \leq \beta\|v\|_2,$$

ed inoltre, dato che z si scompone come somma di autovettori relativi ad autovalori diversi da 1,

$$\|PP_Gz\|_2 \leq \|P_Gz\|_2 \leq \alpha\|z\|_2 = \alpha\left(\|v\|_2 - \left\|\frac{1}{n}e\right\|_2\right) \leq \alpha\|v\|_2,$$

il che dimostra ciò che vogliamo.

Ci accorgiamo quindi che $(PP_GP)^r(1/n)e$ è un vettore ad entrate nonnegative, e dunque possiamo applicare il risultato appena dimostrato

$$Prob[(B,k)] = \left\| (PP_GP)^k \frac{1}{n} e \right\|_1 \leq \sqrt{n} \left\| (PP_GP)^k \frac{1}{n} e \right\|_2 \leq$$

$$\leq \sqrt{n}(\alpha+\beta)^k \left\| \frac{1}{n} e \right\|_2 = (\alpha+\beta)^k,$$

che conclude il teorema.

□

Notiamo che scegliendo invece k nodi casuali, avremmo ottenuto una probabilità di β^k, ma come vedremo nel prossimo capitolo, considerare una camminata aleatoria, al posto di dati totalmente casuali, permette di risparmiare sul numero di bit casuali generati per definire il cammino.

3.4 Derandomizzazione

Consideriamo ora un algoritmo di teoria dei numeri presentato da Micheal O. Rabin nel 1977 in [23]. Si tratta di un test di primalità dei numeri, ad oggi superato, ma che ci permette di dare un esempio di metodo di derandomizzazione di algoritmi.

Questo algoritmo si basa sui due seguenti risultati:

Lemma 3.4.1. *Dato n naturale dispari, sia $2^s d = n-1$ in modo che d sia dispari. Allora n è primo sse per ogni $0 < a < n$ vale una delle seguenti condizioni:*

- $a^d \equiv 1 \mod n$.
- $\exists\, 0 \leq r < s : a^{d2^r} \equiv -1 \mod n$.

Dimostrazione.

Ponendo n primo, avremo che $\mathbb{Z}/n\mathbb{Z}$ ammette un generatore g, dunque, preso un qualsiasi a classe di resto non zero, avremo che $a = g^m$ per un qualche $0 \leq m < n-1$. Scomponiamo $m = 2^t d'$, ed avremo

$$t \geq s \implies a^d = g^{dm} = g^{(n-1)2^{s-t}d'} \equiv 1 \pmod{n},$$

$$t < s \implies a^{d2^{s-t-1}} = g^{dm2^{s-t-1}} = g^{d2^{n-2}d'} \equiv (-1)^{d'} = -1 \pmod{n},$$

dove $0 \leq s-t-1 < s$.

Viceversa, poniamo che sia vera la proprietà per ogni classe di resto non zero, ma che n non sia primo. Allora esiste q divisore di n diverso da 1 e n, che sia anche un numero primo.

Sicuramente q non rispetta la proprietà sopra, poiché la classe di resto di una potenza di q deve comunque essere un multiplo di q, e 1 e -1 non hanno divisori non banali, assurdo.

Dunque n è primo.

□

Dato n composito, chiamiamo *testimoni* tutti i naturali $0 < a < n$ tali che la proprietà sopra non venga rispettata.

Teorema 3.4.1. *Dato $n > 4$ composito, e $T(n)$ l'insieme dei suoi testimoni, allora*

$$\frac{3(n-1)}{4} \leq |T(n)|.$$

Questo teorema ci dice che se scegliamo a caso una classe di resto modulo n, allora c'è al massimo una probabilità $1/4$ di *non* prendere un testimone. Se rappresentiamo le classi di resto come nodi di un grafo, allora nel formalismo del paragrafo precedente, $B(n) = T(n)^c$ e $\beta \leq 1/4$.

Passiamo quindi ad illustrare l'algoritmo di Rabin-Miller:

Input L'algoritmo riceve come input il numero n dispari di cui testare la primalità, e un numero $0 < k < n$.

Parte Random L'algoritmo sceglie in maniera casuale k classi di resto modulo n distinte $b_1, \ldots, b_k$.

Esecuzione Partendo da b_1, l'algoritmo testa se è un testimone di n. Se lo è si ferma, e ritorna l'output. Altrimenti ripete il test su b_2, etc, sino a che non si ferma o non arriva a b_k.

Output Se uno dei b_i era un testimone, ritorna che n è composito, altrimenti, ritorna che n è primo.

Se l'output dell'algoritmo è che n è composito, allora in base al lemma, n è davvero composito; se invece il risultato è che n è primo, vi è la possibilità che non sia corretto. In particolare il risultato non è corretto sse tutti i b_i scelti casualmente non fossero testimoni, e quindi con una probabilità minore di $1/4^k$.

Tramite implementazioni rapide di moltiplicazione e passaggio al modulo, l'algoritmo sopra ha un costo computazionale di $k(2\log_2 n + s\log_2 n)$ passi al massimo, dove $n = 2^s d$, con d dispari.

Il basso costo e la convergenza esponenziale dell'errore rendono questo un buon test di primalità randomizzato, ma la generazione di troppi bit casuali $(b_1, \ldots, b_k)$ può essere un problema.

Ed è qui che entrano in gioco i grafi: preso il grafo descritto sopra, scegliamo il primo nodo a caso, e costruiamo una camminata aleatoria che parte dal nodo scelto. Se il grafo è d-regolare, allora ad ogni passo dobbiamo generare $\log_2(d)$ bit casuali, in confronto ai $\log_2(n)$ bit di b_i. Ciò vuol dire che l'algoritmo normale usa $k\log_2(n)$ bit casuali, contro i $\log_2(n) + (k-1)\log_2(d)$ dell'algoritmo derandomizzato.

Per quanto riguarda la correttezza dell'algoritmo derandomizzato, in virtù dei teoremi sopra, avremo che la probabilità di fallimento aumenta a $(1/4 + \alpha)^k$ utilizzando un (n, d, α)-grafo.

Un altro metodo per derandomizzare algoritmi comporta l'utilizzo dei *Grafi Magici*, che introdurremo nel prossimo paragrafo; non tratteremo questa alternativa, in quanto il numero di bit casuali generati scende a $\log_2 n$, ma la probabilità di fallimento aumenta a $O(1/d)$.

Capitolo 4

Error Correcting Codes

In questa sezione, introduciamo un problema fondamentale di teoria dei codici: la correzione di errori nella trasmissione dei messaggi. Vediamo le specifiche.

Problema

> Alice e Bob vogliono scambiarsi un messaggio, che consiste in una stringa di k bit. Nella trasmissione, però, un rumore compromette al massimo una percentuale p del messaggio. Come possono fare per ricostruire il messaggio originale?

Soluzione Generale

> Si sceglie un *dizionario* o *codice* $C \subseteq \{0,1\}^n$ di stringhe binarie con $n \geq k$, ed una funzione biunivoca dal dizionario ai messaggi possibili. Quando si riceve una stringa di lunghezza n, si cerca la stringa in C più vicina al messaggio ricevuto.

Uno degli strumenti più usati in questo caso è il *test di parità*, ossia aggiungere un bit ad ogni messaggio, che indichi la parità del numero degli 1 nella stringa iniziale. Se infatti il codice venisse modificato in solo una posizione, questo metodo permetterebbe di certificare che il messaggio è stato compromesso, anche se non consentirebbe di ricostruire il messaggio originale.

Varie generalizzazioni del principio sopra sono state proposte, fino a giungere alla teoria dei *codici lineari*, e ai LDPC code (*Low Density Parity Check*), analizzate in [27], [7]. L'utilizzo dei grafi per risolvere tali problemi fu alla base di queste scoperte, insieme alla risoluzione del problema (NP-Hard) di rimozione dell'errore.

Nei prossimi paragrafi, cerchiamo di fare un rapido excursus di questi argomenti.

4.1 Bound Combinatorici

La soluzione generale implica che bisogna trovare un dizionario accettabile, ed una funzione distanza tra le stringhe.

Per la seconda si usa quella più ovvia:

Definizione 4.1.1 (Distanza di Hamming)**.**
Date due stringhe x e y della stessa lunghezza, la loro *distanza di Hamming* è il numero di bit di differenza tra le due. In formule, $d_H(x, y) = sum(xor(x, y))$.

Notiamo che la distanza di Hamming è a tutti gli effetti una distanza per lo spazio vettoriale $\mathbb{F}_2^n$, dove n è la lunghezza delle stringhe, in quanto indotta da una norma:

$$d_H(x, y) = sum(xor(x, y)) = \sum_{i=1}^{n}(x_i - y_i) = \|x - y\|_1.$$

Se chiamiamo x il messaggio originale di n bit, e $\tilde{x}$ il messaggio corrotto, sappiamo che il numero di bit cambiati sarà al massimo $pn \geq d_H(x, \tilde{x})$. Ciò vuol dire che se il destinatario del messaggio deve essere in grado di ricostruire i dati originali, allora tutti gli elementi del dizionario diversi da x devono avere distanza da $\tilde{x}$ maggiore di pn.
Ciò è possibile se, per esempio, **tutte le coppie di stringhe del dizionario avessero distanza maggiore di *2pn*** , poiché allora avremmo

$$\forall y \neq x \quad d_H(y, \tilde{x}) \geq d_H(y, x) - d_H(x, \tilde{x}) > 2pn - pn = pn,$$

grazie alla diseguaglianza triangolare data dalla distanza.

Poniamo per esempio che Alberto debba trasmettere un messaggio di lunghezza 5 tra 8 disponibili, e sa che la percentuale del messaggio corrotta sarà non superiore a 1/5, ossia al massimo un bit verrà modificato.
Per quanto detto sopra, bisognerebbe trovare un codice di 8 stringhe, con distanza almeno $2pn > 2$, quindi almeno 3. Come vedremo in seguito, è impossibile trovarne uno con $n = 5$, mentre ne esiste uno con $n = 6$.

	Originale	Codice
s_1	10101	000000
s_2	10001	111000
s_3	10111	100110
s_4	10010	010101
s_5	00101	001011
s_6	00000	110011
s_7	11011	011110
s_8	01001	101101

Messaggio Originale
$s_4 = 10010$
↓
Messaggio Codificato
$s'_4 = 010101$
↓
Messaggio Compromesso
$\tilde{s}'_4 = 011101$

Alberto, dunque, trasmette il messaggio codificato s'_4, ma Barbara lo riceve compromesso, quindi cerca la stringa del codice con distanza minore di 2, e lo decodifica:

Corrotto	Codice	Distanza
	000000	4
	111000	3
	100110	5
011101	010101	1
	001011	3
	110011	4
	011110	2
	101101	2

Messaggio Compromesso
$\tilde{s}'_4 = 011101$
↓
Messaggio Codificato
$s'_4 = 010101$
↓
Messaggio Originale
$s_4 = 10010$

In termini di costi, vorremmo che il nostro dizionario sia il più grande possibile, così da decodificare più messaggi, ma contemporaneamente che n sia vicino a k, per non far salire troppo il numero di bit trasmessi.

Un indice della bontà del dizionario è dato dal suo *Rate*:

Definizione 4.1.2 (Rate)**.**
Se $C \subseteq \{0,1\}^n$ è il nostro dizionario, allora definiamo il suo *rate* come

$$R(C) = \frac{\log_2 |C|}{n}.$$

Notiamo che $R(C) \leq k/n$ poichè $|C| \leq 2^k$ in quanto è in corrispondenza biunivoca con alcune stringhe da k bit. Quindi, più grande è il rate del dizionario, meno costoso e più espressivo sarà.

Per quanto riguarda invece la correttezza del metodo, essa sarà descritta dalla *distanza* del dizionario:

Definizione 4.1.3 (Distanza)**.**
Dato C un dizionario, allora definiamo la sua *distanza* come

$$D(C) = \frac{\min_{c_1 \neq c_2}^{c_1, c_2 \in C} d_H(c_1, c_2)}{n}.$$

Da quanto osservato prima, avremmo bisogno quindi che $D(C) > 2p$. Tutto ciò ci porta ad una riformulazione del problema:

Specifiche É possibile trovare una sequenza C_s di dizionari con $|C_s| = 2^s$, le cui rispettive distanze siano maggiori di $\delta > 0$ e i cui rate siano maggiori di $R > 0$?

Un altro modo di vedere il problema, coinvolge l'impacchettamento di sfere n-dimensionali, definite tramite la distanza di Hamming: Se una sfera con centro x e raggio r è l'insieme delle stringhe a distanza di Hamming minore o uguale a r da x, allora trovare il massimo numero possibile di elementi a distanza maggiore di $2pn$ diventa equivalente a trovare il numero massimo di sfere di raggio pn che si possono inserire in $\mathbb{F}_2^n$ senza che si intersechino.

Grazie a questa interpretazione, possiamo trovare dei limiti imposti dalla combinatoria a Rate e Distanza di un codice. Chiamato $B(n,d)$ il massimo numero di stringhe di lunghezza n tale che la distanza tra ogni coppia di tali stringhe sia almeno d, sappiamo che ([25])

Teorema 4.1.1. *Valgono i seguenti bound:*

- $B(n,1) = 2^n$.
- $B(n-1, 2k-1) = B(n, 2k)$.
- $B(n, 2k-1) \leq \frac{2^n}{\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{k-1}}$.

Dimostrazione.

Prese tutte le stringhe, esse hanno distanza l'una dall'altra non nulla, dunque almeno 1. Si ha pertanto che $B(n,1) = 2^n$.

Consideriamo ora stringhe di lunghezza $n-1$ e distanza minima $2k-1$, e C un codice con cardinalità $B(n-1, 2k-1)$. Se ad ogni stringa di C aggiungiamo come n-esima coordinata un bit, che è la somma delle prime $n-1$ componenti, otteniamo un codice C' con stringhe di lunghezza n e distanza minima $2k-1$. Se però consideriamo una coppia di stringhe in C con distanza esattamente $2k-1$, otteniamo che il bit aggiunto sarà sicuramente diverso, poiché abbiamo scambiato un numero dispari di bit dall'una all'altra. Ne segue che le stringhe corrispondenti in C' avranno distanza $2k$, e di conseguenza tutte le stringhe in C' avranno distanza almeno $2k$ le une dalle altre: $B(n-1, 2k-1) \leq B(n, 2k)$.

Per il viceversa, prendiamo un codice C con stringhe di lunghezza n, distanza minima $2k$ e cardinalità $B(n, 2k)$. Creiamo dunque un codice C' eliminando l'ultima componente. Avremo che ogni coppia di stringhe ha distanza al minimo $2k-1$, ma C' ha la stessa cardinalità di C, e dunque $B(n-1, 2k-1) \geq B(n, 2k)$.

Questo porta dunque alla formula $B(n-1, 2k-1) = B(n, 2k)$.

Come detto sopra, $B(n, 2k-1)$ è anche il massimo numero di sfere di raggio $k-1$ che possiamo inserire in $\mathbb{F}_2^n$ senza che queste si intersechino.

Dato che tutte le sfere con lo stesso raggio contengono un numero uguale di stringhe (che chiamiamo *volume* della sfera), avremo che $B(n, 2k-1)$ sarà limitato da 2^n diviso il volume di ogni palla.

In particolare, possiamo considerare la sfera $S(0, k-1)$ centrata nella stringa composta da soli zeri, che chiameremo l'origine, e di raggio $k-1$. La distanza di una qualsiasi stringa dall'origine è il numero di componenti 1 che contiene, dunque le stringhe contenute in $S(0, k-1)$ saranno quelle con al massimo $k-1$ componenti 1.

Dato che il numero di stringe con esattamente r componenti 1 è $\binom{n}{r}$, allora

$$B(n, 2k-1) \leq \frac{2^n}{Vol(S(0, k-1))} = \frac{2^n}{\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{k-1}}.$$

□

Nell'esempio sopra, abbiamo detto che è impossibile trovare un codice con almeno 8 stringhe di lunghezza 5 e distanza 3; infatti, dai bound appena trovati, avremo che

$$B(5,3) \leq \frac{2^5}{\binom{5}{0} + \binom{5}{1}} < 8.$$

In generale, determinare $B(n,d)$ è un problema aperto, e i bound conosciuti sono molto larghi.[1]

4.2 Grafi Magici

Nella grande famiglia dei grafi expander annoveriamo anche alcuni grafi bipartiti che hanno buone proprietà di espansione ristrette ad uno dei due set di nodi del grafo.

Definizione 4.2.1 (Grafo Magico)**.**
Dato G tale che $L \coprod R = V$ rappresenta una bipartizione dei nodi, allora è detto un (n,m,d)-grafo magico se valgono le seguenti:

- $|L| = n$, $|R| = m$, e ogni nodo in L ha grado d.
- se $S \subseteq L$, con $|S| \leq n/10d$, allora $|N(S)| \geq 5d|S|/8$.
- se $S \subseteq L$, con $n/10d < |S| \leq n/2$, allora $|N(S)| \geq |S|$.

Dimostriamo un lemma che sarà utile tra poco:

Lemma 4.2.1. *Se G è un $(n, 3n/4, d)$-grafo magico, allora $\forall S \subseteq L$ non vuoto con $|S| \leq n/10d$, esiste u nodo in R tale che ha esattamente un vicino in S.*

Dimostrazione.
$e(S) = d|S|$, ma $|N(S)| \geq 5d|S|/8$, quindi ogni nodo in $N(S)$ ha in media

$$\frac{e(S)}{|N(S)|} \leq \frac{d|S|}{5d|S|/8} = \frac{8}{5}$$

archi che lo collegano a nodi in S.

Dato che però ogni nodo in $N(S)$ ha almeno un arco uscente che finisce in S, ne esisterà uno con meno di $8/5$ vicini in S, e di conseguenza con esattamente un vicino in S.

□

Adesso che li abbiamo definiti, usiamo questi grafi per generare buoni dizionari.

Preso un $(n, 3n/4, d)$-grafo magico, sia A la matrice $3n/4 \times n$ che ha per righe i nodi in R, per colonne i nodi in L, e vale 0 o 1 a seconda se la coppia di nodi sia collegata da un arco oppure no. Prendiamo quindi il codice su n bit dato da

$$C = \{x \in \mathbb{F}_2^n : Ax = 0\}.$$

Un modo per vederlo è questo: dato x vettore di n bit, associamo ogni sua componente ad un nodo in L; allora $x \in C$ sse per ogni nodo in R, la somma dei valori sui nodi collegati ad esso vale zero in $\mathbb{F}_2$.

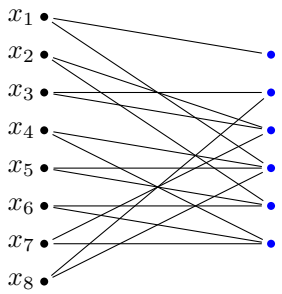


Figura 4.1:
$(8,6,2)$-*grafo magico*

[1] su http://www.win.tue.nl/∼aeb/codes/binary-1.html si trova una tavola aggiornata

In Figura 1, per esempio, avremo che

$$\begin{cases} x_1 = 0 \\ x_3 + x_8 = 0 \\ x_2 + x_3 + x_7 = 0 \\ x_1 + x_4 + x_5 + x_8 = 0 \\ x_2 + x_5 + x_6 = 0 \\ x_4 + x_6 + x_7 = 0 \end{cases} \rightarrow \quad A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

Come abbiamo già detto, $\mathbb{F}_2^n$ è uno spazio vettoriale, e imporre $Ax = 0$ vuol dire soddisfare $3n/4$ equazioni lineari, quindi C sarà un sottospazio vettoriale di dimensione almeno $n/4$.

Definizione 4.2.2 (Codice Lineare)**.**
Dato C codice in $\mathbb{F}_2^n$, è un codice lineare se è un suo sottospazio.

Lavorare con codici lineari è comodo poiché è relativamente facile stimarne la distanza:

Lemma 4.2.2. *Se C è un codice lineare, definiamo* peso *di un elemento di C il numero di cifre 1 in esso contenuto $w(x) = sum(x)$. Allora*

$$D(C) = \min_{x \in C} w(x)/n.$$

Dimostrazione.
Se x, y sono elementi di C, allora $x + y$ è ancora un elemento di C, poiché è lineare. Ma visto che $d_h(x, y) = d_H(0, x + y) = w(x + y)$, allora il minimo delle distanze è pari al minimo dei pesi.

□

Detto questo, adesso è facile vedere che il codice costruito in questa maniera ha buone proprietà:

Teorema 4.2.1. *Dato C costruito sopra, avremo $D(C) > 1/10d$ e $R(C) \geq 1/4$.*

Dimostrazione.
Consideriamo S in L di cardinalità minore di $n/10d$. Per il lemma sopra, esiste un nodo in R con un solo vicino in S, ma allora preso $x \in C$, è impossibile che $w(x) \leq n/10d$, poiché altrimenti scegliendo $S = \{i \in L : x_i = 1\}$ il nodo di R associato determina un equazione non soddisfatta. Quindi $w(x) > n/10d$, e quindi $D(C) > 1/10d$.
Come detto sopra, C è un sottospazio di dimensione maggiore di $n/4$, dunque $|C| \geq 2^{n/4}$ e $R(C) \geq 1/4$.

□

Grazie a questi grafi riusciamo dunque a trovare una famiglia di dizionari come richiesto dal problema.

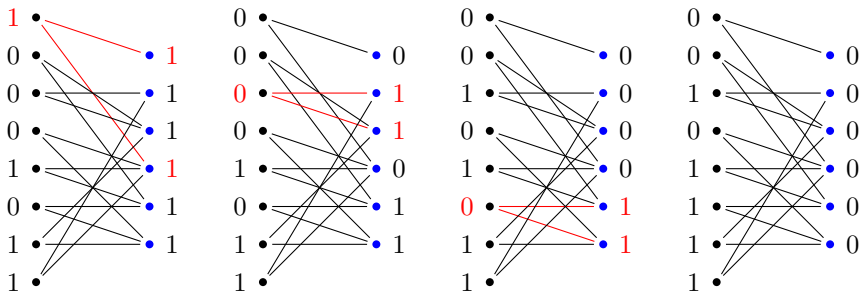
L'ultima questione aperta resta la decodifica del messaggio: confrontare la stringa ricevuta con tutte le stringhe del codice per trovare la distanza di hamming minima è un processo computazionalmente non efficace. Qui entra in campo il parametro $L(G, k)$.

Definizione 4.2.3 (left vertex expansion ratio)**.**
Dato G tale che $L \coprod R = V$ rappresenta una bipartizione dei nodi, allora

$$L(G,k) = \min_{\substack{0<|S|\leq k \\ S\subseteq L}} \frac{|N(S)|}{|S|}.$$

Chiamiamo *belief propagation* la seguente operazione: data una stringa x tale che $Ax \neq 0$, scegliamo un nodo i tale che la maggior parte dei controlli relativi ai vicini di i siano sbagliati (ossia $w(A(x+e_i)) < w(A(x))$), e cambiamo la corrispondente componente di x.

Prendiamo per esempio il grafo in Figura 1, e applichiamo più volte la belief propagation al vettore 10001011.



Come possiamo vedere, 3 iterazioni della belief propagation ci hanno permesso di trasformare il vettore iniziale in una stringa del codice.

Il seguente teorema ci assicura che questa procedura, applicata poche volte, porta alla correzione di una stringa corrotta e all'eliminazione dell'errore.

Teorema 4.2.2. *Dato G un grafo bipartito in cui i nodi di L abbiano tutti grado d, supponiamo che la percentuale di corruzione del messaggio sia al massimo p. Se $k = \lfloor pn \rfloor$ e $L(G, 2k) > 3d/4$, allora un numero di iterazioni lineare in n della belief propagation ripristina il messaggio originale.*

Dimostrazione.
Sia x il messaggio originale, e y quello corrotto, con $d_H(x,y) \leq k$.

Sia inoltre $y^{(i)}$ il vettore dopo i applicazioni della belief propagation, dove $y = y^{(0)}$, e $A_i = \{v : y_v^{(i)} \neq x_v\}$, ossia le componenti ancora errate di $y^{(i)}$.

Vogliamo dunque provare che $A_i = \emptyset$ per $i = O(n)$.

Fissato quindi i generico, e poniamo che $A = A_i$ non sia vuoto, ed abbia non più di $2k$ elementi. Consideriamo $N(A)$ e partizioniamolo in vicini soddisfatti S (ossia uguali a 0), e vicini insoddisfatti (ossia uguali a 1) U. Notiamo che U sono i vincoli insoddisfatti non solo in $N(A)$, ma in tutto R, poiché un vincolo insoddisfatto deve forzatamente essere un vicino di almeno una componente sbagliata.

Per la left expansion, si ha

$$|S| + |U| = |N(A)| > \frac{3}{4} d|A|.$$

Notiamo che un nodo sta in U sse ha un numero dispari di archi che lo collegano ad A, così come un nodo sta in S sse ne ha un numero pari, poiché i nodi

dovrebbero essere zero se collegato con x, dunque quando li colleghiamo con $y^{(i)}$ contano solo il numero dei bit cambiati che sono a loro vicini.

Da questo, deduciamo che da ogni nodo di U parte almeno un arco per A, mentre da ogni nodo in S ne partono almeno 2, e dato che i nodi che partono da A sono $d|A|$, avremo

$$|U| + 2|S| \leq d|A| \implies |U| = 2(|S| + |U|) - (|U| + 2|S|) > \frac{1}{2}d|A|.$$

Ricaviamo che c'è un nodo in A con più di $d/2$ vicini insoddisfatti, e dunque, finché $|A_i| \leq 2k$ possiamo continuare ad applicare la belief propagation.

Inoltre, ad ogni passo, la cardinalità di U decresce almeno di 1, poiché i vincoli insoddisfatti diminuiscono, e dunque, se $|A_i| \leq 2k$ per ogni i, in $|U_0|$ passi l'algoritmo termina e converge a x, ma $|U_0| \leq d|A_0| \leq dk = d\lfloor pn \rfloor$, dunque abbiamo bisogno di $O(n)$ iterazioni.

Infine, se esiste un j per cui $|A_j| > 2k$, allora esiste un i per cui $|A_i| = 2k$, dunque

$$dk \geq d|A_0| \geq |N(A_0)| \geq |U_0| \geq |U_i| > \frac{1}{2}d|A_i| = dk,$$

assurdo.

□

Vari risultati dimostrano che tali grafi esistono, e sono costruibili sia per via deterministiche che randomizzate.

Capitolo 5

Costruzione di Grafi Expander

Nelle sezioni precedenti, abbiamo dato risultati riguardanti l'utilizzo degli expanders, e visto che più i grafi hanno buone proprietà, più gli algoritmi diventano efficienti. Adesso affrontiamo il problema dell'esistenza di tali grafi, dando costruzioni esplicite di questi ultimi.

Ricordiamo che per far sì che gli algoritmi funzionino correttamente, il costo per scrivere e lavorare con i grafi deve essere basso, e allo stesso tempo, questi devono essere descritti esplicitamente.

5.1 Esistenza e Costruzioni

Partendo con i grafi magici, come osservato da Pinsker, con metodi probabilistici si può dimostrare l'esistenza della maggior parte di essi:

Lemma 5.1.1. *Esiste una costante n_0 tale che per ogni $d \geq 32$ e $n \geq n_0$, $m \geq 3n/4$, esiste un grafo (n, m, d)-magico.*

Dimostrazione.

Sia G un grafo random bipartito con $|L| = n$, $|R| = m$, e d-regolare a sinistra. Sia $S \subseteq L$ con cardinalità $s = |S| \leq n/10d$, $T \subseteq R$ con $t = |T| < 5ds/8$, e $P_{S,T}$ la probabilità che tutti gli archi che partono da S finiscano in T.

Usando la disuguaglianza $\binom{n}{k} \leq (ne/k)^k$, otteniamo

$$\begin{aligned}
\sum_{S,T} P_{S,T} \leq \sum_{S,T} (t/m)^{sd} &\leq \sum_{s=1}^{n/10d} \binom{n}{s}\binom{m}{5ds/8}\left(\frac{5ds}{8m}\right)^{sd} \\
&\leq \sum_{s=1}^{n/10d} \left(\frac{ne}{s}\right)^{s}\left(\frac{8me}{5ds}\right)^{5sd/8}\left(\frac{5ds}{8m}\right)^{sd} \\
&= \sum_{s=1}^{n/10d} \left[\left(\frac{s}{n}\right)^{\frac{3}{8}d-1}\left(\frac{5dn}{8m}\right)^{\frac{3}{8}d} e^{\frac{5}{8}d+1}\right]^{s} \\
&\leq \sum_{s=1}^{n/10d} \left[\left(\frac{1}{10d}\right)^{\frac{3}{8}d}\left(\frac{5d}{6}\right)^{\frac{3}{8}d} 10de^{\frac{5}{8}d+1}\right]^{s} \\
&= \sum_{s=1}^{n/10d} \left[\left(\frac{e^5}{12^3}\right)^{\frac{1}{8}d} 10de\right]^{s} \leq \sum_{s=1}^{n/10d} \left[\left(\frac{e^5}{12^3}\right)^{4} 320e\right]^{s} \\
&\leq \sum_{s=1}^{n/10d} \left(\frac{1}{20}\right)^{s} \leq \sum_{s=1}^{\infty} \left(\frac{1}{20}\right)^{s} \leq \frac{1}{19},
\end{aligned}$$

dove la quarta disuguaglianza è data da $s \leq n/10d$ e $m \geq 3n/4$, mentre la quinta è vera poiché $d > 32$, ed ogni termine della sommatoria decresce all'aumentare di d.

Se invece $n/10d < s = |S| \leq n/2$, e $t = |T| < |S|$, come prima avremo

$$\begin{aligned}
\sum_{S,T} P_{S,T} \leq \sum_{S,T} (t/m)^{sd} &\leq \sum_{s=n/10d}^{n/2} \binom{n}{s}\binom{m}{s}\left(\frac{s}{m}\right)^{sd} \\
&\leq \sum_{s=n/10d}^{n/2} \left[\left(\frac{ne}{s}\right)\left(\frac{me}{s}\right)\left(\frac{s}{m}\right)^{d}\right]^{s} \\
&= \sum_{s=n/10d}^{n/2} \left[e^2\left(\frac{s}{n}\right)^{d-2}\left(\frac{n}{m}\right)^{d-1}\right]^{s} \\
&\leq \sum_{s=n/10d}^{n/2} \left[e^2\left(\frac{1}{2}\right)^{d-2}\left(\frac{4}{3}\right)^{d-1}\right]^{s} \\
&= \sum_{s=n/10d}^{n/2} \left[e^2\frac{4}{3}\left(\frac{2}{3}\right)^{d-2}\right]^{s} \\
&\leq \sum_{s=n/10d}^{n/2} \left[e^2\frac{4}{3}\left(\frac{2}{3}\right)^{30}\right]^{s} \\
&\leq \sum_{s=n/10d}^{n/2} \left(\frac{1}{10^4}\right)^{s} \leq \frac{1}{9999}.
\end{aligned}$$

Ciò vuol dire che con grande probabilità, il grafo generato sarà un (n, m, d)-grafo magico. □

Questo dimostra che riusciamo sempre a creare grafi magici, ma inoltre dice che dato un grafo generato casualmente, esso sarà con molte probabilità un grafo magico.

Ottenere grafi bipartiti regolari a sinistra, con una buona left expansion ratio, è possibile tramite costruzione diretta. In particolare, tramite il prodotto zig-zag di grafi bipartiti, riusciamo ad ottenere per ogni costante $\delta > 0$ e per un d abbastanza grande, un grafo bipartito con $L(G, k) > (1 - \delta)d$ per $k = \Omega(n)$. Non daremo qui la costruzione, poiché lunga, e richiede il concetto di Entropia di una distribuzione di probabilità su un grafo.

Poniamo adesso di voler costruire una famiglia di expanders G_i tali che il numero di nodi cresca non troppo velocemente ($n_{i+1} \leq n_i^2$). Questa famiglia è chiamata *Mildly Explicit* se vi è un algoritmo per generare G_i in tempo polinomiale in i, mentre è chiamata *Very Explicit* se dati i, k, v, computa il k-esimo vicino del nodo v nel grafo G_i in tempo polinomiale nel numero di bit della tripla (i, k, v).

Esempi di costruzioni esplicite di Expander Graph sono

Margulis Graph il grafo G_m di Margulis ha come nodi gli elementi di $\mathbb{Z}_m \times \mathbb{Z}_m$ ed è un grafo 8-regolare, in quanto colleghiamo (x, y) con $(x + y, y), (x - y, y), (x, y+x), (x, y-x), (x+y+1, y), (x-y+1, y), (x, y+x+1), (x, y-x+1)$, dove le operazioni sono modulo m. Questo grafo può avere loop e archi multipli, ma comunque si riesce a dimostrare (utilizzando Serie di Fourier e Teoria delle Rappresentazioni) che il gap spettrale è $8 - 5\sqrt{2}$ $\forall m$. Questa famiglia di grafi è Very Explicit.

Cubi Un esempio semplice sono i cubi d-dimensionali. Questi sono grafi d-regolari e bipartiti, e il gap spettrale è costante e pari a 2. L'unica pecca è che il numero di nodi cresce esponenzialmente, e che il grado aumenta.

Grafi di Caley Dato H un gruppo finito, e S un sottoinsieme di H, allora il grafo $C(H, S)$ ha per nodi gli elementi di H, e due nodi g e h sono collegati se esiste un elemento in S t.c. $gs = h$. Assumendo che S sia un insieme di generatori simmetrico (ossia $S = S^{-1}$), e che l'identità non vi appartenga, $C(H, S)$ risulta non diretto, senza loop, connesso, e $|S|$-regolare. In molti casi, il gap spettrale di questi grafi è facile da calcolare. Un risultato importante è che per ogni $n \geq 3$, $m \geq 1$, p primo, il gruppo $SL_n(\mathbb{F}_{p^m})$ ammette un set di generatori $S_{m,n,p}$ di cardinalità costante, tale che i grafi di Caley generati siano una famiglia di Expanders.

Mild Un altro esempio di grafi derivati da quelli di Caley sono costruiti su $\mathbb{Z}_p$ con p primo, tali che x sia collegato con $x + 1, x - 1$, e il suo inverso x^{-1}, dove poniamo $0^{-1} = 0$. Questi sono grafi 3-regolari, ma la loro particolarità è di essere solo Mildly Explicit, in quanto non siamo in grado di generare deterministicamente grandi primi.

Un'ultima nota va invece alle costruzioni randomizzate di grafi. Come preannunciato, i grafi regolari che sono buoni expanders, hanno caratteristiche simili a grafi regolari costruiti casualmente.

Il risultato di Friedman mostra esplicitamente questa dipendenza:

Teorema 5.1.1. *Presi G_n grafi casuali su n nodi d-regolari, con $\lambda(G) = \max\{|\lambda_2|, |\lambda_n|\}$ allora per ogni $\epsilon > 0$ vale*

$$Prob\left[\lambda(G) \leq 2\sqrt{d-1} + \epsilon\right] = 1 - o_n(1).$$

il che porta a sospettare che la maggior parte dei grafi d-regolari abbiano $\lambda(G) \leq 2\sqrt{d-1}$, (chiamati *Grafi di Ramanujan*).

Capitolo 6

Altri Utilizzi

Questo trattato non pretende di coprire tutti gli utilizzi di questo potente mezzo che sono i grafi expanders.

Al contrario, molti sono stati omessi, quali

- applicazioni nella teoria degli embedding in spazi metrici, e utilizzi in algoritmi per problemi di taglio.
- studio di gruppi tramite Grafi di Caley.
- applicazioni alla teoria della complessità e approssimazioni di algoritmi Np-hard o completi.
- studio di ricoprimenti universali di grafi e alberi infiniti.
- costruzione di superconcentratori e complessità relativa a matrici super regolari.

e molti altri, reperibili in [7].

Abbiamo comunque dato un'idea del fatto che questi grafi riescano ad avere applicazioni inattese in molti campi della matematica, dall'algebra alla geometria e alla probabilità.

Bibliografia

[1] László Lovász (2007), *Eigenvalues of graphs*, unpublished notes, available at http://www.cs.elte.hu/∼lovasz/eigenvals-x.pdf

[2] Ya-Hong Chen, Rong-Ying Pan, Xiao-Dong Zhang (2011), *The Laplacian Spectra of Graphs and Complex Networks*. CoRR abs/1111.2896 .

[3] Russel Merris (1994), *Laplacian Matrices of Graphs: A Survey*, Linear Algebra Appl. 197&198, 143-176.

[4] Xiao-Dong Zhang (2007), *The Laplacian eigenvalues of graphs: A Survey*, Gerald D. Ling (Ed.), Linear Algebra Research Advances, Nova Science Publishers Inc., pp. 201–228, Chapter 6

[5] Bojan Mohar (1997), *Some Applications of Laplace Eigenvalues of Graphs*, Graph Symmetry: Algebraic Methods and Approximations, pp.255-275.

[6] Bojan Mohar (1992), *Laplace eigenvalues of graphs*, Discrete Math. 109 171-183

[7] Shlomo Hoory, Nathan Linial, Avi Widgerson (2006), *Expander Graphs and their Application*, Bulletin of the American Mathematical Society, Volume 43, Number 4, October 2006, Pages 439–561.

[8] Bojan Mohar (1987), *Isoperimetric Numbers of Graphs*, Journal of Combinatorial Theory, pp. 274-291.

[9] P.J. Cameron, J.M. Goethals, J.J. Seidel, E.E. Shult (1975), *Line Graphs, Root Systems, and Elliptic Geometry*, J. Algebra, 43 (1976), pp. 305–327

[10] Miroslav Fiedler (1973), *Algebraic connectivity of graphs*, Czechoslovak Mathematical Journal 23(2):298–305

[11] Inderjit S. Dhillon (2001), *Co-clustering documents and words using Bipartite Spectral Graph Partitioning*, Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, p.269-274

[12] K.Ch. Das (2004), *The Laplacian Spectrum of a Graph*, Comput. Math. Appl. 48 715–724.

[13] Bit-Shun Tam, Shu-Hui Wu (2010), *On the reduced signless Laplacian spectrum of a degree maximal graph*, Linear Algebra and its Appl., 432, pp. 1734-1756.

[14] Satish B. Rao (1992), *Faster Algorithms for Finding Small Edge Cuts in Planar Graphs*, Proceedings of the 24th Annual ACM Symposium on the Theory of Computing, pp. 229–240

[15] Eric A. Carlen, Elliott H. Lieb (2012), *Short Proofs of Theorems of Mirsky and Horn on Diagonals and Eigenvalues of Matrices*, Linear Algebra and its Applications, Volume 437, Issue 10, Pages 2680–2682

[16] R.B. Bapat, V.S. Sunder (1985), *On Majorization and Schur Products*, Linear Algebra Appl. 72, 107-117.

[17] Xiao-Dong Zhang (2011), *Vertex Degrees and Doubly Stochastic Graph Matrices* , Journal of Graph Theory, 66:104-114

[18] Russel Merris (1997), *Doubly Stochastic Graph Matrices*, Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. 8 64-71

[19] S. Chaiken, D.J. Kleitman (1978), *Matrix Tree Theorem*, Journal of Combinatorial Theory, Series A Volume 24, Issue 3, Pages 377–381

[20] Mark Jerrum, Alistair sinclair, Eric Vigoda (2004), *A Polynomial-Time Approximation Algorithm for the Permanent of a Matrix with Nonnegative Entries*, Journal of the ACM (JACM), Volume 51 Issue 4, Pages 671-697

[21] Bojan Mohar, Svatopluk Poljak (1993), *Eigenvalues in Combinatorial Optimization*, Combinatorial and Graph-Theoretical Problems in Linear Algebra, IMA Volume 50

[22] Reza Olfati-Saber (2005), *Ultrafast Consensus in Small-World Networks*, Proceeding of the American Control Conference, pp. 2371–2378.

[23] Micheal O. Rabin (1977), *Probabilistic Algorithm for Testing Primality*, Journal of Number Theory 12, 128-138.

[24] N. Alon, V.D. Milman (1984), λ_1*, lsoperimetric Inequalities for Graphs, and Superconcentrators*, Journal of Combinatorial Theory, Series B 38, 73-88.

[25] R.W. Hamming (1950), *Error Detecting and Error Correcting Codes*, Bell Syst. Tech. J., vol. 29, pp.147 -160

[26] Leslie G. Valiant (1976), *Graph-Theoretic Properties in Computational Complexity*, Journal of Computer and System Sciences, Volume 13 Issue 3, December, Pages 278-285

[27] Michael Sipser, Daniel A. Spielman (1996), *Expander Codes*, IEEE Transactions on Information Theory, vol. 42, no. 6, pp. 1710–1722.

[28] F.J. MacWilliams, N.J.A. Sloane (1977), *The Theory of Error-Correcting Codes*, SIAM Rev., 22(4), 513–519.

[29] J. Friedman. *A proof of Alon's second eigenvalue conjecture*, Proceedings of the thirty-fifth annual ACM symposium on Theory of computing, Pages 720-724

[30] M.R.F. Smyth (2002), *A Spectral Theoretic Proof of Perron-Frobenius*, Mathematical Proceedings of the Royal Irish Academy, page 29–35. JSTOR.

[31] J. Cheeger (1970), *A lower bound for the smallest eigenvalue of the Laplacian*, in Problems in Analysis (R. C. Gunnig, Ed.), pp. 195-199, Princeton Univ. Press, Princeton, NJ.

[32] Willem H. Haemers (1995). *Interlacing eigenvalues and graphs*. Linear Algebra Appl. 226/228:593–616.