

Appunti del corso

Teoria dei Grafi e dei Giochi

Paolo Serafini, a.a. 2015-16

revisione: 5 maggio 2016

Teoria dei Grafi

1 Definizione di grafo non orientato

Un grafo (non orientato) è un insieme $G = (V, E)$ dove V è un insieme finito i cui elementi vengono detti *vertici* o *nodi*, e E è un sottoinsieme di coppie (non ordinate) di V dette *archi*, *spigoli* o *lati*.

È prassi abbastanza diffusa usare la notazione $n = |V|$ e $m = |E|$. Un arco può essere indicato sia con una notazione generica $e \in E$, che con una più specifica (i, j) , con $i, j \in V$. Dato l'arco (i, j) i vertici i e j vengono anche indicati come *estremi* dell'arco. È consuetudine visualizzare un grafo come un insieme di punti più o meno grandi (i vertici) e di segmenti che li connettono (gli archi), a volte curvilinei per esigenze di disegno.

Se esiste l'arco (i, j) , i vertici i e j vengono detti *adiacenti*, mentre il vertice i (e anche il vertice j) e l'arco (i, j) vengono detti *incidenti*. Anche due archi con un vertice in comune vengono detti incidenti.

Questa definizione di grafo non ammette il caso di molteplicità di archi per la medesima coppia di nodi. Se è necessario considerare questo caso, bisogna modificare la definizione di E in questo modo: E è un insieme finito ed è definita una funzione d'arco $E \rightarrow V \times V$ che associa ad ogni arco $e \in E$ una coppia (i, j) di vertici distinti. Se esistono due archi diversi e e e' associati alla stessa coppia di vertici si parla di *multigrafo*. Questa nuova definizione può considerare anche il caso di archi i cui estremi coincidano, i cosiddetti *cappi*, se si ammettono coppie di vertici uguali. Se sono presenti dei cappi si parla di *pseudografo*.

I grafi rivestono un'importanza fondamentale in molti campi, anche diversi fra loro, in quanto esprimono, in modo astratto, l'esistenza o meno di una relazione binaria fra elementi di un insieme. Si tratta evidentemente di una situazione quanto mai pervasiva. Come esempi possiamo citare: reti stradali (gli incroci sono i vertici e gli archi sono i tratti di strada fra due incroci); reti ferroviarie (analoghe alle reti stradali); tornei sportivi (le squadre sono i vertici e gli archi rappresentano gli incontri fra le squadre); reti elettriche (le connessioni sono i vertici e gli archi sono il conduttore fra le connessioni); internet (ogni homepage è un vertice e ogni link è un arco – in questo caso l'arco è orientato); facebook (ogni utente è un vertice e ogni legame di amicizia è un arco); molecole (gli atomi sono i vertici e i legami chimici sono gli archi).

Come si è visto nell'esempio di internet l'ordine della coppia di elemento può essere rilevante. In questi casi si parla di grafi orientati. La definizione di grafo orientato verrà data nella Sez. 12. In queste prime sezioni si tratta unicamente di grafi non orientati.

Gli esempi potrebbero continuare a lungo. Rappresentiamo in Fig. 1 il grafo dato dalla mappa della metropolitana di Londra, in Fig. 2-1 il grafo di una molecola di sulfon-ftaleina e in Fig. 2-2 un grafo più astratto in cui i vertici sono i numeri da 1 a 12 e c'è un arco se i numeri non sono primi fra loro. In Fig. 3 viene visualizzato un grande grafo che rappresenta la mappa dell'interazione proteica della *Drosophila melanogaster*.

Quindi fenomeni molto diversi fra loro, ma che abbiano la stessa struttura astratta di grafo, possono essere studiati ed analizzati direttamente dalle proprietà generali dei grafi. In questo corso verranno appunto

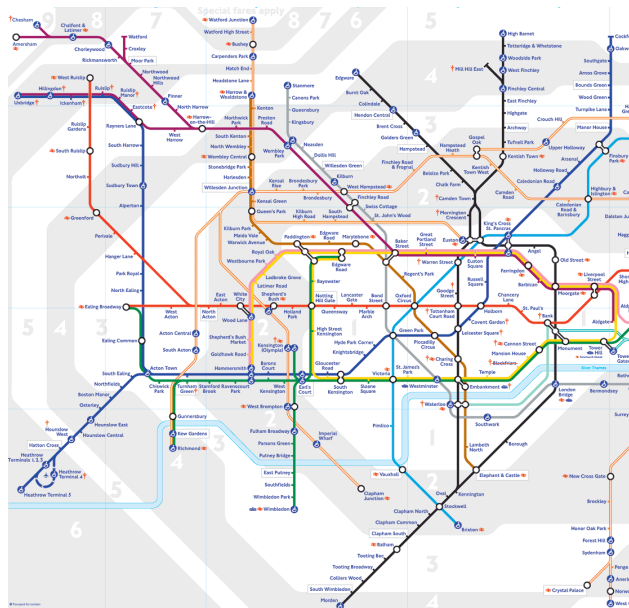
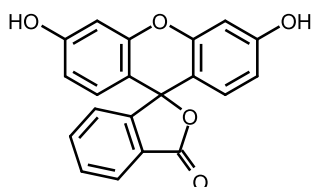
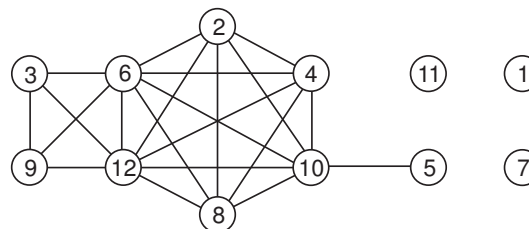


Figura 1. Metropolitana di Londra



1 – Molecola di Sulfon-Ftaleina



2 – Numeri non primi fra loro

Figura 2.

esposte le più importanti caratterizzazioni e proprietà dei grafi. La trattazione non può che essere parziale. Una trattazione esaustiva richiederebbe diversi volumi.

2 Cammini, circuiti e connessione

Un *cammino* in un grafo viene definito stabilendo un vertice di partenza detto *sorgente*, identificando un arco incidente nel vertice di partenza, selezionando l'altro vertice dell'arco e procedendo ricorsivamente fino ad arrivare in un vertice prefissato di arrivo detto *destinazione*. Se sorgente e destinazione coincidono ed esiste almeno un arco il cammino viene detto *circuito*. Bisogna considerare anche il caso degenero di un cammino

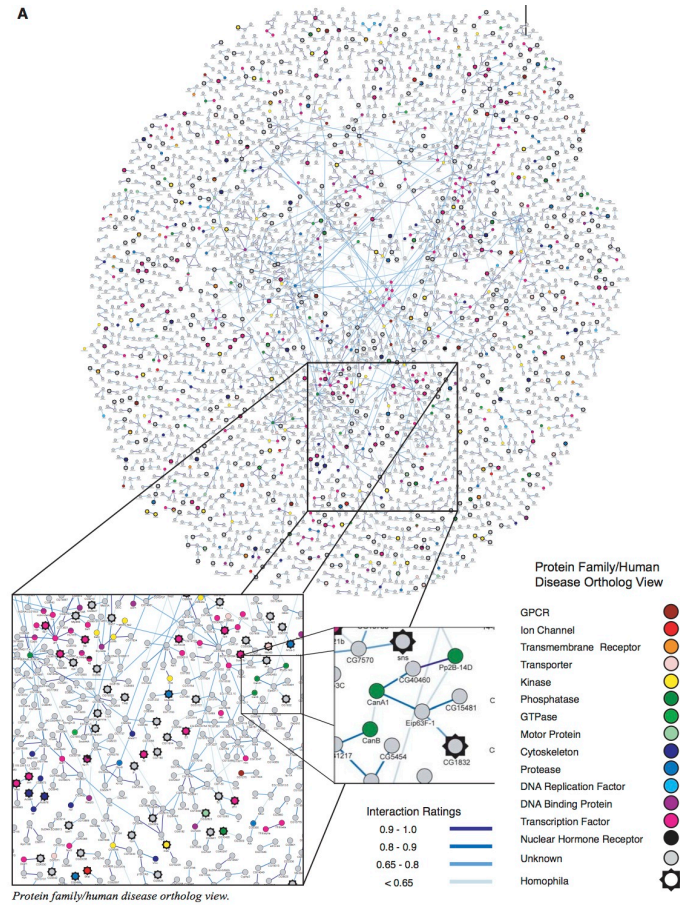


Figura 3. Mappa dell'interazione proteica della *Drosophila melanogaster*

che consiste solo di un vertice. In questo caso sorgente e destinazione coincidono ma non si tratta di un circuito.

Se gli archi di un cammino sono tutti diversi, il cammino viene detto *non molteplice*. Se anche i nodi del cammino sono tutti diversi il cammino viene detto *elementare* o *semplice*. Simili definizioni valgono anche per i circuiti. Si noti che un circuito semplice deve avere almeno tre archi.

Se in un grafo esiste un cammino da ogni vertice ad ogni altro vertice il grafo è detto *connesso*. Un grafo in cui $E = \emptyset$ viene detto *totalmente sconnesso*. Se un grafo non è connesso i sottoinsiemi di vertici connessi fra loro inducono dei sottografi detti *componenti connesse*.

Un circuito semplice che contiene tutti i vertici del grafo viene detto *hamiltoniano*. Un circuito che contiene tutti gli archi esattamente una volta viene detto *euleriano*. Le stesse definizioni valgono per i cammini hamiltoniani ed euleriani rispettivamente. Si può dimostrare che esiste un circuito euleriano in un

grafo connesso se e solo se tutti i nodi hanno grado pari (si veda la sezione seguente per la definizione di grado) ed esiste un cammino euleriano se e solo se esattamente due nodi hanno grado dispari (e il cammino deve necessariamente partire ed arrivare in questi nodi). Invece, a tutt'oggi non esiste una caratterizzazione così semplice per determinare l'esistenza di un circuito hamiltoniano in un grafo. Si tratta infatti di un problema **NP**-completo.

Se non intervengono valori numerici associati agli archi, la *lunghezza* di un cammino è il numero di archi del cammino. Un cammino è pari o dispari se la sua lunghezza è pari o dispari.

3 Grado di un nodo e regolarità

Si definisce come *grado* di un nodo il numero degli archi incidenti nel nodo. Non esiste una notazione univoca per indicare il grado. Qui usiamo la notazione g_i per indicare il grado del nodo i . Si vede facilmente che $\sum_{i \in V} g_i = 2m$ e quindi la somma dei gradi è pari. Da questo fatto discende che il numero di nodi con grado dispari è pari.

Grafi con lo stesso grado in ogni nodo vengono detti *regolari*, oppure *k-regolari* con k il grado di ogni nodo. Un grafo si dice *fortemente regolare*, se è regolare e inoltre ogni coppia di vertici adiacenti ha in comune lo stesso numero r di vertici adiacenti e ogni coppia di vertici non adiacenti ha in comune lo stesso numero s di vertici adiacenti (si trovano usualmente in letteratura i simboli λ e μ al posto di r e s ; siccome λ e μ avranno altri significati in seguito, si preferisce usare in questa sede una notazione meno ambigua).

I grafi completi (si veda la Sez. 5 per la definizione) vengono spesso inclusi fra i grafi fortemente regolari anche se il numero s non è definito e viene posto convenzionalmente a zero. Analogamente vengono anche inclusi i grafi totalmente sconnessi per i quali non è definito r .

I grafi fortemente regolari vengono contrassegnati dai quattro numeri (n, k, r, s) . In Fig. 4 sono rappresentati tutti i grafi fortemente regolari connessi e non completi con $n \leq 9$ vertici. Viene anche raffigurato il famoso grafo di Petersen con 10 vertici che è fortemente regolare (non è l'unico fortemente regolare connesso e non completo con 10 vertici).

Non tutti i valori di n , k , r and s possono essere ammissibili. Deve valere

$$(k - r - 1)k = (n - k - 1)s$$

La dimostrazione di questa relazione si basa su un argomento di conteggio. Si prenda un vertice (chiamiamolo v_0) e i suoi k vertici adiacenti (insieme A). Gli altri $n - k - 1$ vertici costituiscano l'insieme B . Per costruzione i vertici di B non sono adiacenti a v_0 . Si contino gli archi con un estremo in A e l'altro in B . Ognuno dei vertici di B ha in comune con v_0 s vertici, che appartengono necessariamente ad A . Quindi il numero di archi fra A e B è $s(n - k - 1)$. Un altro modo di contare gli archi tiene conto del fatto che ognuno dei vertici in A deve avere r vertici in comune, necessariamente in A , con v_0 . Quindi per ogni vertice di A ci sono $(k - r - 1)$ archi incidenti in B , per cui si ottiene il numero di archi $k(k - r - 1)$ che deve essere uguale alla precedente espressione.

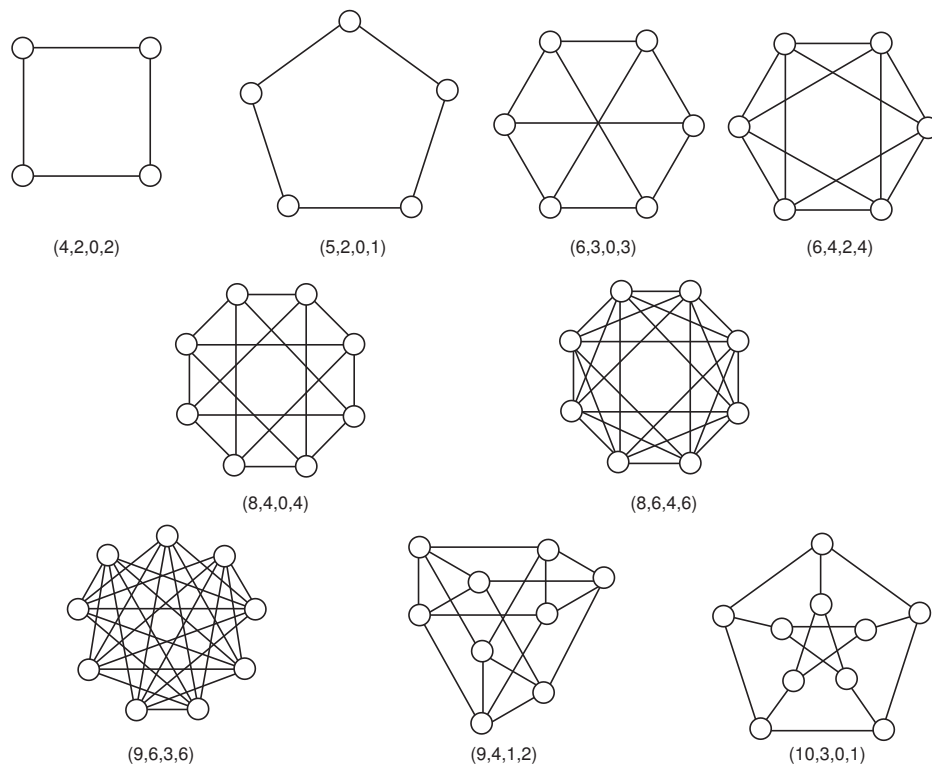


Figura 4. Grafi fortemente regolari

Il vettore g_i viene detto sequenza dei gradi. Non tutti i vettori di n componenti con valori compresi fra 0 e $n - 1$ possono essere la sequenza di un grafo. Inoltre grafi diversi possono avere la stessa sequenza dei gradi (Fig. 5). Quindi la funzione che assegna ad ogni grafo la sua sequenza dei gradi non è né suriettiva, né iniettiva. Una sequenza di n interi non negativi viene detta *grafica* se è la sequenza dei gradi di un grafo. Vale la seguente importante proprietà che permette di verificare se un vettore sia o no grafico.

Siano dati n numeri. Questi numeri vengano ordinati come $a_1 \geq a_2 \geq \dots \geq a_n$. Tali numeri sono una sequenza grafica se e solo se b_1, \dots, b_{n-1} è una sequenza grafica con

$$\begin{cases} b_i = a_{i+1} - 1 & \text{se } 1 \leq i \leq a_1 \\ b_i = a_{i+1} & \text{se } i > a_1 \end{cases}$$

(in altre parole b è ottenuto da a , eliminando a_1 e sottraendo 1 dai successivi a_1 numeri).

Ad esempio si consideri $a = (5, 4, 4, 3, 2, 2, 2)$. Quindi potrebbe trattarsi di un grafo con 7 nodi. Si assegni il grado 5 al nodo 1 e si creino gli archi $(1, 2), (1, 3), (1, 4), (1, 5), (1, 6)$. Adesso bisogna considerare il grafo dopo aver “tolto” il nodo 1. Quindi a diventa $(3, 3, 2, 1, 1, 2)$. Questo vettore va riordinato (tenendo conto di quali sono i nodi). Quello che si ottiene è (in ogni matrice la riga di sopra contiene i gradi e quella di sotto i

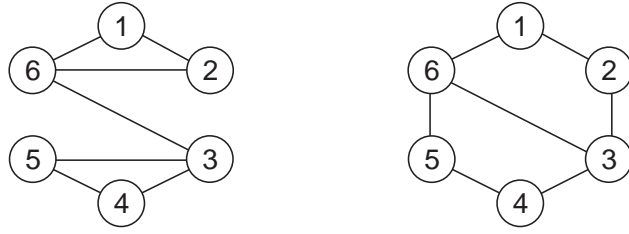


Figura 5. Grafi diversi (e non isomorfi) con la stessa sequenza dei gradi

nodi)

$$\begin{pmatrix} 5 & 4 & 4 & 3 & 2 & 2 & 2 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{pmatrix} \xrightarrow{(a)} \begin{pmatrix} 3 & 3 & 2 & 1 & 1 & 2 \\ 2 & 3 & 4 & 5 & 6 & 7 \end{pmatrix} \xrightarrow{(b)} \begin{pmatrix} 3 & 3 & 2 & 2 & 1 & 1 \\ 2 & 3 & 4 & 7 & 5 & 6 \end{pmatrix} \xrightarrow{(c)} \begin{pmatrix} 2 & 1 & 1 & 1 & 1 \\ 3 & 4 & 7 & 5 & 6 \end{pmatrix} \xrightarrow{(d)} \begin{pmatrix} 0 & 0 & 1 & 1 \\ 4 & 7 & 5 & 6 \end{pmatrix} \xrightarrow{(e)} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 5 & 6 & 4 & 7 \end{pmatrix} \xrightarrow{(f)} \begin{pmatrix} 0 & 0 & 0 \\ 6 & 4 & 7 \end{pmatrix}$$

dove in (a) si generano gli archi (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), in (b) gli archi (2, 3), (2, 4), (2, 7), in (c) gli archi (3, 4), (3, 7) e in (d) l'arco (5, 6).

Se invece la sequenza fosse $a = (5, 5, 5, 5, 3, 3)$ avremmo

$$\begin{pmatrix} 5 & 5 & 5 & 5 & 3 & 3 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix} \xrightarrow{(a)} \begin{pmatrix} 4 & 4 & 4 & 2 & 2 \\ 2 & 3 & 4 & 5 & 6 \end{pmatrix} \xrightarrow{(b)} \begin{pmatrix} 3 & 3 & 1 & 1 \\ 3 & 4 & 5 & 6 \end{pmatrix} \xrightarrow{(c)} \begin{pmatrix} 2 & 0 & 0 \\ 4 & 5 & 6 \end{pmatrix}$$

e a questo punto si vede che la sequenza (2, 0, 0) non può essere la sequenza di nodi di nessun grafo (a meno di ammettere cappi) e quindi la sequenza (5, 5, 5, 5, 3, 3) non è grafica.

Questa procedura è particolarmente importante in quei casi in cui si vuole generare un grafo casuale con distribuzione fissata dei gradi. Di questo si parlerà nella Sez.46 a pag. 116.

4 Isomorfismo

Due grafi sono uguali se gli insiemi dei nodi e degli archi coincidono. Questa definizione di uguaglianza tiene conto di quali sono i vertici del grafo. Molto spesso però si è interessati alla struttura astratta del grafo indipendentemente da chi siano i vertici.

Si introduce allora il concetto di *isomorfismo*. Due grafi $G_1 = (V_1, E_1)$ e $G_2 = (V_2, E_2)$ si dicono isomorfi se esiste una corrispondenza biettiva $\pi : V_1 \rightarrow V_2$ tale che $(i, j) \in E_1$ se e solo se $(\pi(i), \pi(j)) \in E_2$.

Se $V_1 = V_2 = V$ la corrispondenza biettiva è una permutazione su V . Se avviene che $\pi(E) = E$ (cioè l'elenco degli archi è il medesimo anche dopo avere rietichettato tutti i nodi), π viene detto *automorfismo*. Se definiamo equivalenti due nodi i e j tali che $\pi(i) = j$ e π è un automorfismo, le classi di equivalenza vengono dette *orbite*. Ovviamente tutti i nodi della stessa orbita devono avere lo stesso grado. In Fig. 6 si vedono tre

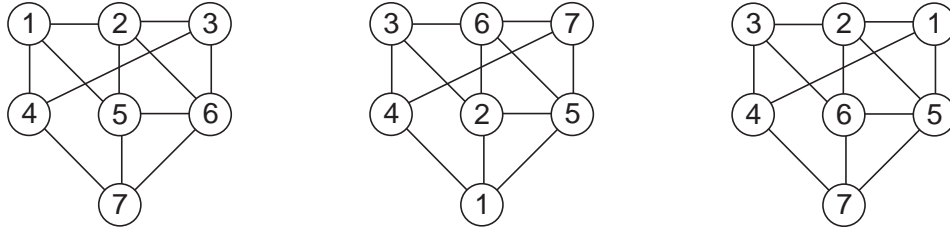


Figura 6. Tre automorfismi

automorfismi dello stesso grafo (ce ne sono altri, quali?). L'elenco degli archi è il medesimo per ognuno dei tre grafi. I nodi vengono ripartiti in tre orbite, date dai nodi rispettivamente $\{1, 3, 7\}$, $\{2, 5, 6\}$ e $\{4\}$ (c'è un modo migliore di 'disegnare' il grafo che rende più evidente la sua struttura, quale?).

Nodi nella stessa orbita sono, per così dire, indistinguibili uno dall'altro. Si immagini di voler identificare un nodo particolare, senza sapere chi sia, solo esplorando il grafo e conoscendo la 'mappa' del grafo (ovvero l'elenco degli archi). Il fatto che esiste una permutazione dei nodi che rende invariati gli archi impedisce di sapere di quale nodo della stessa orbita si tratti.

A tutt'oggi (2016) non è stato trovato un algoritmo polinomiale per determinare se due grafi sono isomorfi né è stato dimostrato che il problema sia **NP**-completo. Che sia **NP**-completo è molto improbabile perché ciò implicherebbe dei risultati poco attesi della teoria della **NP**-completezza. Dal lato della polinomialità va citato un importante risultato recente di Laszlo Babai (12/11/2015) che ha trovato un algoritmo quasi-polinomiale, cioè con complessità $2^{O(\log^c n)}$. In ogni caso dal punto di vista pratico vi sono degli algoritmi che, senza essere polinomiali, permettono abbastanza velocemente di determinare se i grafi sono isomorfi.

Si noti che, banalmente, due grafi devono avere lo stesso numero di archi e di nodi per essere isomorfi. Questa condizione necessaria si può rafforzare in modo iterativo, notando che i grafi in esame devono avere lo stesso numero di nodi di grado 0, di grado 1, eccetera. Se questo avviene allora il problema si divide in tanti sottoproblemi. Ma si può dire di più. Sia $\mathcal{V} = \{V_0, V_1, \dots, V_{n-1}\}$ la partizione di V indotta dai gradi dei nodi di uno dei due grafi. Ogni nodo in V_i ha come nodi adiacenti nodi appartenenti a certi sottoinsiemi V_k . Se due nodi nello stesso sottoinsieme V_i hanno una diversa distribuzione di nodi adiacenti non possono essere permutati fra di loro. In base a questa distinzione l'insieme V_i viene a sua volta partizionato. Si prosegue ricorsivamente finché non è più possibile raffinare la partizione. A questo punto se le due partizioni non hanno gli stessi numeri di elementi i grafi non sono isomorfi, altrimenti si cercano le permutazioni per ognuno dei sottoinsiemi.

Ad esempio siano dati i grafi

$$E_1 = \{(1, 2), (1, 4), (2, 3), (2, 4), (2, 5), (2, 6), (3, 6), (4, 5), (5, 6)\}$$

$$E_2 = \{(1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 3), (2, 4), (3, 4), (5, 6)\}$$

Per E_1 si ottiene la partizione $\mathcal{V}_1 = \{\{1, 3\}, \{4, 5, 6\}, \{2\}\}$ (nodi di grado 2, 3 e 5 rispettivamente) e per E_2 si ottiene la partizione $\mathcal{V}_2 = \{\{5, 6\}, \{2, 3, 4\}, \{1\}\}$. Quindi, fino a questo punto, i grafi potrebbero essere isomorfi in quanto gli insiemi della partizione hanno la stessa cardinalità. Raffinando \mathcal{V}_1 si esaminano i nodi 1 e 3. Il nodo 1 è adiacente ad un nodo in $\{2\}$ e ad un nodo in $\{4, 5, 6\}$. Altrettanto succede per il nodo 3. Quindi non è possibile dividere l'insieme $\{1, 3\}$. Invece per l'insieme $\{4, 5, 6\}$ si vede che, mentre 4 e 6 hanno la stessa struttura d'intorno, il nodo 5 ha una diversa struttura. Quindi la nuova partizione è $\mathcal{V}_1 = \{\{1, 3\}, \{4, 6\}, \{5\}, \{2\}\}$. La partizione \mathcal{V}_2 non è ulteriormente raffinata quindi i grafi non sono isomorfi.

Ovviamente questo metodo non permette di ridurre i calcoli se due grafi sono regolari. In particolare il problema di determinare se due grafi sono isomorfi è difficile se i grafi sono fortemente regolari.

5 Grafi particolari

Sono state definite molte classi di grafi a seconda della particolare struttura dell'insieme di archi. L'elenco sarebbe lunghissimo e ci limitiamo a presentare solo alcune classi, quelle che si incontrano più frequentemente.

Un grafo che abbia come archi tutte le possibili coppie non ordinate di nodi viene detto *completo* e viene indicato con K_n . Tutti i grafi completi di n nodi sono isomorfi fra loro.

Un grafo i cui nodi si possano ripartire in due sottoinsiemi V_1 e V_2 tali che $(i, j) \in E$ solo se $i \in V_1$ e $j \in V_2$ (o viceversa) viene detto *bipartito*. Per evidenziare la bipartizione di un grafo si usa la notazione $(V_1, V_2; E)$. I grafi bipartiti intervengono quando ad esempio bisogna modellare una relazione fra elementi di due insiemi diversi. Determinare se un grafo è bipartito è facile. Un grafo è bipartito se e solo se tutti i suoi circuiti sono pari (facile dimostrazione lasciata al lettore).

Un grafo bipartito viene detto *completo* se per ogni $i \in V_1$ e ogni $j \in V_2$ esiste l'arco (i, j) (quindi un grafo bipartito completo non è completo nel senso generale!) e viene indicato con $K_{|N_1|, |N_2|}$. I grafi bipartiti completi $K_{n,n}$ sono fortemente regolari (tre di questi sono inclusi in Fig. 4. Quali sono?)

Più in generale un grafo in cui i nodi possano essere ripartiti in sottoinsiemi V_1, \dots, V_k tali che $(i, j) \in E$ solo se $i \in V_h$ e $j \in V_{h'}$ con $h \neq h'$, viene detto *k-partito* oppure *multi-partito*. Un grafo *k-partito* si dice *completo* se per ogni coppia (i, j) esiste un arco se e solo se i e j appartengono a diversi sottoinsiemi. Si noti che un grafo *k-partito* è anche banalmente $(k + 1)$ -partito (se naturalmente ha almeno $k + 1$ vertici). I sottoinsiemi della partizione sono per definizione insiemi stabili (si veda la successiva Sez. 7) e il minimo valore di k per cui il grafo sia *k-partito* è il numero cromatico del grafo, il cui calcolo è NP-difficile.

Un grafo si dice *a livelli* se i nodi possono essere ripartiti in sottoinsiemi V_1, \dots, V_k tali che $(i, j) \in E$ solo se $i \in V_h$ e $j \in V_{h+1}$. Un grafo con k livelli è anche un grafo *k-partito*. Inoltre è anche un grafo con $(k - 1)$ livelli e quindi è anche $(k - 1)$ -partito. Quindi un grafo a livelli è anche bipartito. Viceversa un grafo bipartito (con almeno tre vertici) è anche un grafo con tre livelli (in diversi modi alternativi). Affinché un grafo bipartito abbia quattro livelli devono essere verificate delle condizioni. Quali? Quale grafo bipartito non può mai avere quattro livelli?

Una *stella* è un grafo in cui un nodo, detto centro, è adiacente a tutti gli altri e questi sono adiacenti solo al centro. Una stella con $n + 1$ nodi viene normalmente indicata con S_n . La stella S_n non è altro che il grafo bipartito completo $K_{1,n}$. Se i nodi esterni di una stella vengono resi adiacenti l'uno all'altro in sequenza circolare si ottiene una *ruota*, indicata come W_n (con n numero di nodi esterni, anche se si trova in letteratura il simbolo W_{n+1} per la ruota con $n + 1$ nodi).

Un *cubo* è un grafo con 2^n nodi, identificabili con vettori di n componenti di zeri e uni. C'è un arco se i vettori differiscono per esattamente una componente. Un cubo viene indicato con Q_n oppure anche con 2^n . Ovviamente $Q_1 = K_2$, Q_2 è il quadrato e Q_3 è il cubo ordinariamente detto. Un cubo è un particolare grafo bipartito (V_1 è dato dai vettori con un numero pari di uni e V_2 da un numero dispari).

Una *griglia*, denotata anche G_{mn} se bidimensionale e $G_{m_1 m_2 \dots m_n}$ se n -dimensionale (ma la notazione non è standard), ha vertici associati alle n -uple (i_1, i_2, \dots, i_n) , $i_1 = 1, \dots, m_1$, $i_2 = 1, \dots, m_2, \dots, i_n = 1, \dots, m_n$, e archi fra due n -uple che differiscono in una sola componente e la differenza vale 1. Anche una griglia è un particolare grafo bipartito (V_1 è dato dai nodi con $\sum_k i_k$ pari e V_2 con somma dispari). Un cubo è un particolare grafo a griglia (con $m_k = 2$ per ogni k).

Grafi molto importanti sono gli *alberi*, che verranno definiti e trattati più diffusamente nella Sez. 10.

6 Grafi ottenuti da grafi

Dato un grafo si possono ottenere altri grafi tramite varie costruzioni. Il grafo $\bar{G} = (V, \bar{E})$ ottenuto dal grafo $G = (V, E)$ tramite la relazione $(i, j) \in \bar{E} \iff (i, j) \notin E$ viene detto *grafo complementare* di E . Il complementare di un grafo fortemente regolare è anch'esso fortemente regolare.

Un'interessante proprietà che lega un grafo al suo complementare è che non possono essere entrambi sconnessi. Sia G sconnesso e si prendano due nodi i e j . Se l'arco (i, j) non è presente in G allora esiste nel complementare e i e j sono connessi. Se invece l'arco (i, j) è in G , necessariamente i e j stanno nella stessa componente connessa ed esiste un nodo k in un'altra componente connessa, e quindi non esistono gli archi (i, k) e (k, j) in G . Ma questo implica che esiste un cammino di due archi nel grafo complementare e quindi, dati due nodi arbitrari nel complementare questi sono sempre connessi.

Data una partizione V_1, \dots, V_k dei nodi di un grafo G , il grafo G' ottenuto identificando il sottoinsieme V_i con il nodo i di G' , detto anche *pseudonodo*, e definendo il seguente insieme di archi per G'

$$E' := \{(i, j) : \exists h \in V_i, k \in V_j, (h, k) \in E \text{ con } i \neq j\}$$

viene detto grafo *contratto*, oppure ottenuto per *contrazione* della partizione (si usano anche i termini *collassato* e *collassamento*).

Il grafo $G' = (V', E')$ è un *sottografo* di $G = (V, E)$ se $V' \subset V$ e $E' \subset E$. Si noti che V' e E' non possono essere sottoinsiemi qualsiasi. Siccome G' deve essere a sua volta un grafo, è richiesto che tutti gli archi in E' abbiano estremi in V' . Il grafo $G' = (V, E')$ è un *grafo parziale* o, alternativamente, un *sottografo di*

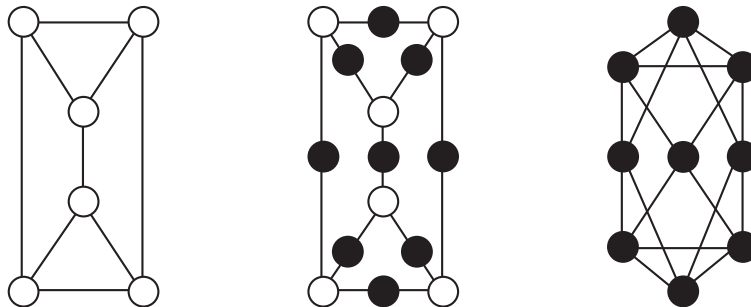


Figura 7. Grafo G , nodi di $L(G)$ (in nero), $L(G)$

supporto di $G = (V, E)$ se $E' \subset E$. Dato $V' \subset V$ il sottografo $G' = (V', E(V'))$ *indotto* da V' è il sottografo di $G = (V, E)$ che contiene tutti gli archi di G con entrambi gli estremi in V' , insieme indicato come $E(V')$.

Dato un grafo G si definisce come suo *grafo di linea* $L(G)$ il grafo in cui i nodi sono associati agli archi di G ed esiste un arco (i, j) se e solo se gli archi i e j di G sono incidenti. Si noti che necessariamente il grafo $L(G)$ contiene, per ogni nodo i di G una *cricca* di d_i nodi (per la definizione di ‘*cricca*’ si veda l’inizio della Sez. 7). Si veda in Fig. 7 un grafo e il suo grafo di linea.

Il grafo di linea di un grafo k -regolare è un grafo $(2(k-1))$ -regolare. Quindi un circuito (vedi dopo per la definizione) è isomorfo al suo grafo di linea, ed è l’unico tipo di grafi per cui questo succeda. Non tutti i grafi sono grafi di linea di un altro grafo. Affinché ciò sia vero devono esistere nel grafo delle cricche tali che ogni arco sia coperto da una cricca e ogni nodo appartenga al massimo a due cricche. Ovviamente grafi di linea di grafi isomorfi sono isomorfi. Esistono anche due grafi non isomorfi che producono due grafi di linea isomorfi e sono K_3 e la stella S_3 il cui grafo di linea è ancora K_3 . Tuttavia, a parte questo unico caso, due grafi di linea isomorfi derivano da grafi isomorfi.

Dati due grafi $G_1 = (V_1, E_1)$ e $G_2 = (V_2, E_2)$ si possono definire varie forme di *prodotto di grafi*. Il grafo prodotto $G = (V, E)$ ha come insieme di nodi il prodotto cartesiano $V = V_1 \times V_2$. Quindi ogni nodo può essere etichettato come $[i_1, i_2]$ con $i_1 \in V_1$ e $i_2 \in V_2$ e il numero di nodi del grafo prodotto è $n_1 n_2$. La definizione dell’insieme di archi invece varia e dà luogo a diversi tipi di prodotto. Il *prodotto cartesiano* $G_1 \square G_2$ ha come archi il seguente insieme

$$([i_1, i_2], [j_1, j_2]) \in E \iff (i_1 = j_1 \wedge (i_2, j_2) \in E_2) \vee ((i_1, j_1) \in E_1 \wedge i_2 = j_2)$$

con numero di archi $n_1 m_2 + n_2 m_1$. Come si vede una parte di archi ‘proviene’ dal grafo G_1 e ognuno di questi archi viene replicato tante volte quante sono i nodi di G_2 e l’altra parte di archi proviene da G_2 , replicati tante volte quante sono i nodi di G_1 . Il grado del nodo $[i_1, i_2]$ è, come evidente dalla definizione di arco, $g(i_1) + g(i_2)$ e il grado medio del grafo prodotto è la somma dei gradi medi dei due grafi. Si vedano in Fig. 8 due grafi e il loro prodotto cartesiano (sono evidenziati in tratto grosso gli archi provenienti da G_2). Intuitivamente si può pensare di ottenere il grafo prodotto sostituendo ogni nodo di G_1 con una copia del

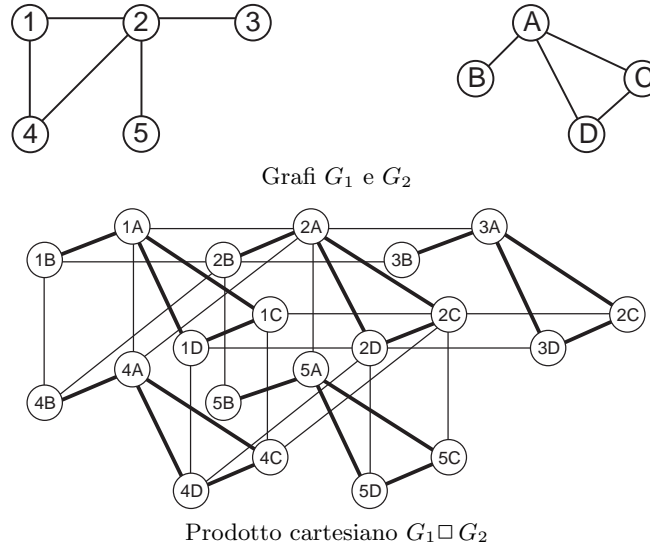


Figura 8.

grafo G_2 e poi collegando i nodi corrispondenti delle varie copie con archi di G_1 . Si ottiene ovviamente lo stesso risultato se si invertono i ruoli di G_1 e G_2 .

Il prodotto cartesiano di n cammini dà luogo ad una griglia n -dimensionale.

Nella definizione di *prodotto tensoriale* $G_1 \times G_2$ (detto qualche volta anche *prodotto di Kronecker* con simbolo $G_1 \otimes G_2$) gli archi sono definiti come

$$([i_1, i_2], [j_1, j_2]) \in E \iff ((i_1, j_1) \in E_1 \wedge (i_2, j_2) \in E_2)$$

con numero di archi $2m_1m_2$. Infatti ogni coppia di archi $e_1 = (i_1, j_1) \in E_1$, $e_2 = (i_2, j_2) \in E_2$ dà luogo alla coppia di archi nel grafo prodotto $([i_1, i_2], [j_1, j_2])$ e $([i_1, j_2], [j_1, i_2])$. In questo caso il grado del nodo $[i_1, i_2]$ è $g(i_1) \cdot g(i_2)$ e il grado medio del grafo prodotto è il prodotto dei gradi medi dei due grafi. Si noti che gli archi del prodotto cartesiano e del prodotto tensoriale formano due insiemi disgiunti. Si vedano in Fig. 9 due grafi e il loro prodotto tensoriale.

Se G_1 e G_2 sono bipartiti, anche il loro prodotto cartesiano e quello tensoriale sono bipartiti. Infatti si prenda un circuito in $G_1 \square G_2$. Ogni arco del circuito può essere messo in corrispondenza con un arco di G_1 oppure di G_2 (ma non di entrambi), con due archi successivi in G_1 incidenti fra loro ed altrettanto per quelli di G_2 . Quindi ogni circuito in $G_1 \square G_2$ dà luogo ad un circuito in G_1 e ad un altro in G_2 (non necessariamente semplici) la cui somma del numero di archi è uguale al numero di archi del circuito in $G_1 \square G_2$. Siccome G_1 e G_2 sono bipartiti i due circuiti sono pari e la loro somma deve essere pari.

Se invece consideriamo un circuito in $G_1 \times G_2$ ogni arco del circuito è in corrispondenza con un arco di G_1 e uno di G_2 . Quindi ogni circuito in $G_1 \times G_2$ dà luogo ad un circuito in G_1 e ad un altro in G_2

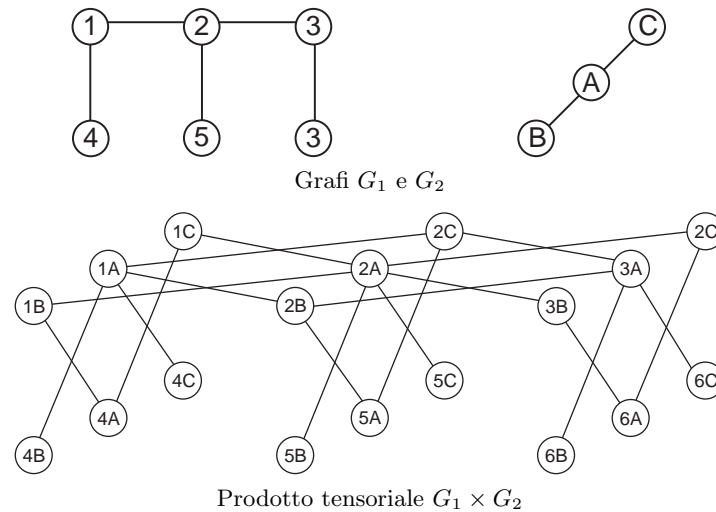


Figura 9.

(non necessariamente semplici) con lo stesso numero di archi del circuito in $G_1 \times G_2$. Siccome G_1 e G_2 sono bipartiti, ogni circuito in $G_1 \times G_2$ deve essere pari.

L'affermazione inversa vale invece solo per il prodotto cartesiano. Se $G_1 \square G_2$ è bipartito non esistono circuiti dispari in $G_1 \square G_2$. Siccome G_1 e G_2 sono contenuti come sottografi in $G_1 \square G_2$ non possono esistere circuiti dispari nemmeno per G_1 e G_2 . Come controesempio al prodotto tensoriale, si noti che $K_2 \times K_3$ è formato da un unico circuito di 6 nodi e quindi è bipartito, ma ovviamente K_3 non lo è.

Nella definizione di *prodotto forte* $G_1 \boxtimes G_2$ gli archi sono definiti come l'unione degli archi del prodotto cartesiano e del prodotto tensoriale. Quindi il grado del nodo $[i_1, i_2]$ è $g(i_1) + g(i_2) + g(i_1) \cdot g(i_2)$ e il numero di archi $n_1 m_2 + n_2 m_1 + 2 m_1 m_2$. Si vedano in Fig. 10 i grafi K_2 , $K_2 \square K_2$, $K_2 \times K_2$, $K_2 \boxtimes K_2$, $K_2 \square K_2 \square K_2$, $K_2 \times K_2 \times K_2$ e $K_2 \boxtimes K_2 \boxtimes K_2$. Il prodotto cartesiano di p termini $K_2 \square K_2 \square \dots \square K_2 = K_2^{\square p}$ dà luogo al cubo Q_p . In generale si ha $Q_m \square Q_n = Q_{m+n}$. Nel caso di prodotto tensoriale c'è un arco se i vettori differiscono in ogni componente e quindi gli archi di $K_2^{\boxtimes p}$ sono un accoppiamento perfetto (vedi sezione seguente). Nel caso di prodotto forte si ottiene il grafo completo $K_2^{\boxtimes p} = K_{2^p}$. In generale $K_m \boxtimes K_n = K_{mn}$.

Si notino i grafi $K_2 \square K_2$, $K_2 \times K_2$, $K_2 \boxtimes K_2$ e i simboli usati per indicare il prodotto. La scelta del simbolo è dovuta proprio all'aspetto dei rispettivi grafi.

7 Cricche, coperture e colorazioni

Un sottoinsieme K di nodi tale che $(K, E(K))$ è completo viene detto *cricca* (*clique*). Una cricca viene definita *massimale* se, aggiungendo un qualsiasi altro nodo alla cricca, il nuovo insieme non è più una cricca. Più in generale l'aggettivo 'massimale' viene usato quando è definito un insieme base V (ad esempio nodi di un

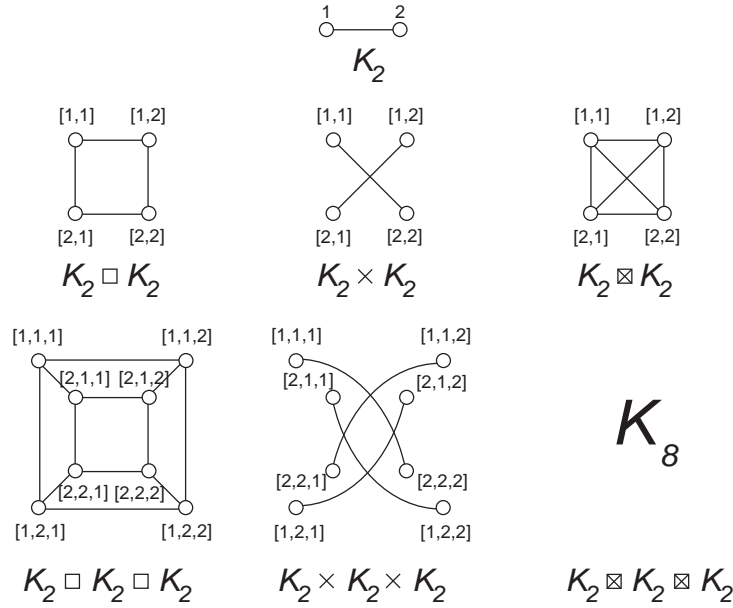


Figura 10.

grafo) ed è definita una famiglia $\mathcal{F} \subset 2^V$ di sottoinsiemi di V (ad esempio le cricche). Un insieme $S \in \mathcal{F}$ viene detto massimale se aggiungendo un qualsiasi elemento a S si ottiene un insieme non appartenente alla famiglia, cioè $S \cup v \notin \mathcal{F}$ per ogni $v \in V \setminus S$. Identificare un insieme massimale qualsiasi è quindi facile perché basta aggiungere un elemento alla volta finché questo non è più possibile.

Il concetto di ‘massimale’ non va confuso con il concetto di ‘massimo’. Un insieme della famiglia \mathcal{F} è massimo se la sua cardinalità è massima fra tutti gli insiemi della famiglia. Ovviamente un insieme massimo è massimale, ma il viceversa non è vero in generale. Trovare la cricca massima in un grafo è **NP**-difficile. Si usa la notazione $\omega(G)$ per indicare la cardinalità della cricca massima del grafo G . Si consideri il grafo di Fig. 2-2, per un valore generico n di nodi. Come sono fatte le cricche di tale grafo?

Un sottoinsieme K di nodi tale che nessun nodo di K è adiacente ad un altro nodo di K viene detto *insieme stabile* o *indipendente*. Si noti che un insieme stabile è una cricca nel grafo complementare. Quindi anche trovare il massimo insieme stabile è **NP**-difficile. Si usa la notazione $\alpha(G)$ per indicare la cardinalità del massimo insieme stabile del grafo G . Quali sono i vertici del massimo insieme stabile del grafo di Fig. 2-2?

Un sottoinsieme di nodi K tale che ogni arco è incidente ad almeno un nodo di K viene detto *copertura di nodi*. Un insieme di nodi è una copertura se e solo se l’insieme complementare è stabile (dimostrazione facile lasciata al lettore). Quindi anche trovare la minima copertura di nodi è **NP**-difficile.

Un sottoinsieme F di archi tale che ogni nodo è incidente in almeno un arco di F viene detto *copertura di archi*. Un sottoinsieme M di archi tale che ogni nodo è incidente in al più un arco di M viene detto

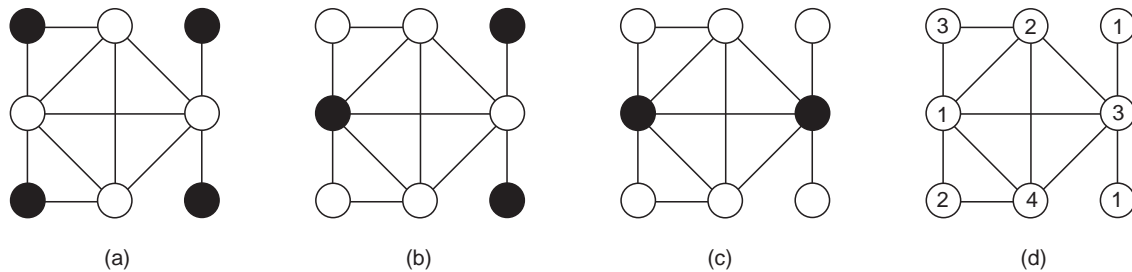


Figura 11. (a) Massimo insieme indipendente (nodi neri) e minima copertura (bianchi); (b) insieme indipendente massimale e insieme dominante minimale; (c) minimo insieme dominante; (d) minima colorazione.

accoppiamento. Se il sottoinsieme M è sia un accoppiamento che una copertura d'archi allora viene detto *accoppiamento perfetto*. Trovare coperture minime oppure accoppiamenti massimi si può fare in tempo polinomiale.

Un insieme di nodi tale che ogni nodo del grafo è adiacente ad almeno un nodo dell'insieme (intendendo che ogni nodo è anche adiacente a se stesso) prende il nome di *insieme dominante*. Trovare il minimo insieme dominante è **NP**-difficile. Un insieme indipendente massimale è necessariamente un insieme dominante minimale. Il complemento di un insieme dominante minimale è dominante. Però un insieme dominante non è necessariamente un insieme indipendente. Si vedano in Fig. 11(a), (b) e (c) alcuni esempi.

Il concetto di insieme stabile si estende decomponendo l'insieme dei nodi del grafo in tanti insiemi stabili. Una tale partizione viene detta *colorazione* del grafo. Il nome deriva dall'idea di colorare i nodi di un grafo con la regola che nodi adiacenti siano di colore diverso e quindi nodi con lo stesso colore devono formare un insieme stabile. Il minimo numero di colori con cui si può colorare il grafo viene detto *numero cromatico* $\chi(G)$. Il calcolo del numero cromatico è in generale **NP**-difficile. Un grafo G è bipartito se e solo se $\chi(G) = 2$. Inoltre $\chi(K_n) = n$ e in generale $\chi(G) \geq \omega(G)$. Si veda in Fig. 11(d) una colorazione di una grafo con 4 colori (in questo caso il numero cromatico dato che avviene $\chi(G) = \omega(G) = 4$).

Una partizione in insiemi stabili diventa nel grafo complementare una partizione in cricche. Il minimo numero di cricche in cui si può decomporre un grafo viene indicato con $\theta(G)$ e ovviamente $\theta(G) \geq \alpha(G)$ (per ogni cricca si può avere al più un nodo di un insieme stabile).

Un grafo G , per cui vale sempre $\chi(G') = \omega(G')$ per ogni sottografo indotto $G' \subset G$, prende il nome di *grafo perfetto*. Un celebre teorema di Lovász afferma che un grafo è perfetto se e solo se il suo complementare è perfetto. Quindi in un grafo perfetto anche $\theta(G') = \alpha(G')$ per ogni sottografo indotto G' . Una tale uguaglianza implica l'esistenza di un certificato sia per le istanze di tipo sì che di tipo no per i vari problemi di insieme stabile, copertura di nodi, massima cricca, numero cromatico e quindi si tratta di problemi che stanno nella classe **NP** \cap **co-NP**. Infatti esistono algoritmi polinomiali che risolvono questi problemi per i grafi perfetti.

Sono ad esempio perfetti i grafi bipartiti, i grafi d'intervallo e i grafi di comparabilità. I grafi d'intervallo sono definiti a partire da un insieme di intervalli I_1, \dots, I_k sulla retta reale. Ad ogni intervallo viene associato un vertice del grafo ed un arco esiste se i corrispondenti intervalli hanno intersezione non vuota. Un grafo di comparabilità viene definito a partire da una relazione d'ordine \prec definita su un insieme V . Esiste un arco fra $i \in V$ e $j \in V$ se $i \prec j$ oppure $j \prec i$. Si può vedere che un grafo d'intervallo è complementare di un grafo di comparabilità.

8 Lunghezza di un cammino e di un circuito

Come detto nella Sez. 2, se non intervengono valori numerici associati agli archi, la lunghezza di un cammino è il numero di archi del cammino. Se invece ad ogni arco (i, j) è associata una *lunghezza* d_{ij} , la lunghezza $d(P)$ del cammino P è la somma di questi valori lungo il cammino, cioè $d(P) = \sum_{(ij) \in P} d_{ij}$.

Dati due nodi s e t la distanza $D(s, t)$ fra i due nodi è la minima fra le lunghezze di tutti i cammini da s a t . La distanza fra due nodi è ben definita se il minimo esiste. Se le lunghezze degli archi sono non negative e il grafo è connesso, certamente il minimo esiste ed esiste quindi anche la distanza fra ogni coppia di nodi. Se il grafo non è connesso la distanza fra due nodi in diverse componenti connesse viene convenzionalmente posta uguale a $+\infty$. Se è presente un arco di lunghezza negativa esistono cammini che percorrono avanti e indietro tale arco un numero qualsiasi di volte e quindi non può esistere il minimo. In questo caso il minimo viene convenzionalmente posto uguale a $-\infty$.

Se si escludono cammini che percorrono avanti e indietro un arco il minimo non esiste se esiste un ciclo (di almeno tre archi) di lunghezza negativa. In questo caso il ciclo può essere percorso un numero qualsiasi di volte. Se si ammettono solo cammini semplici o non molteplici, questi sono in numero finito e quindi il minimo e la distanza esistono sempre.

Dal punto di vista della complessità computazionale la situazione è la seguente: se le lunghezze sono non negative, trovare il cammino minimo (necessariamente semplice) è polinomiale; se le lunghezze sono negative, determinare se esiste un circuito semplice negativo è polinomiale; se tale circuito non esiste, determinare un cammino minimo non multiplice (e necessariamente semplice) è polinomiale; se tale circuito esiste, determinare un cammino minimo semplice è **NP**-difficile.

Nella maggior parte dei casi le lunghezze degli archi sono non negative, per cui il concetto di distanza fra due nodi è ben definito (ammettendo pure il caso di distanza infinita se il grafo non è connesso). D'ora in poi supponiamo che le lunghezze siano non negative.

Per le distanze vale la *diseguaglianza triangolare*:

$$D(i, j) \leq D(i, k) + D(k, j)$$

Si definisce *eccentricità* ε_i di un nodo i la quantità

$$\varepsilon_i := \max_{j \in V} D(i, j)$$

e si definiscono *raggio* $\rho(G)$ e *diametro* $\Delta(G)$ del grafo G le quantità

$$\rho(G) := \min_{i \in V} \varepsilon_i = \min_{i \in V} \max_{j \in V} D(i, j), \quad \Delta(G) := \max_{i \in V} \varepsilon_i = \max_{i \in V} \max_{j \in V} D(i, j)$$

e si definisce *centro* $c(G)$ del grafo G il nodo (o anche i nodi) tale che

$$\rho(G) = \varepsilon_c$$

Si noti che, in base alla disuguaglianza triangolare, con s e t i nodi che determinano il diametro:

$$\Delta(G) = D(s, t) \leq D(s, i) + D(i, t) \quad \forall i \in V \quad \implies \quad \Delta(G) \leq D(s, c) + D(c, t) \leq \varepsilon_c + \varepsilon_c = 2\rho(G)$$

Inoltre siccome

$$\max_{i \in V} \max_{j \in V} D(i, j) \geq \min_{i \in V} \max_{j \in V} D(i, j)$$

si ha

$$1 \leq \frac{\Delta(G)}{\rho(G)} \leq 2$$

Il diametro è il doppio del raggio se il cammino che dà luogo al diametro contiene il centro. Il diametro è invece uguale al raggio tutte le volte in cui c'è simmetria fra tutti i nodi, ovvero per ogni coppia ordinata i e j di nodi esiste un automorfismo (cioè una permutazione delle etichette dei nodi che dà luogo al medesimo elenco di archi) che porta il nodo i in j . In questo caso ogni nodo “vede” il medesimo grafo e le eccentricità sono tutte uguali (esempi banali: un grafo completo, un circuito).

Le definizioni date di raggio e diametro sono molto generali e si possono applicare a qualsiasi insieme sul quale sia definita una distanza. Se tale insieme è un cerchio si ritrovano proprio il diametro e il raggio secondo la nota definizione geometrica. Però possiamo pensare di definire diametro e raggio anche per altri figure geometriche. Ad esempio per un triangolo equilatero di lato 1, il raggio vale $\sqrt{3}/3$ e il diametro è dato da un lato e vale 1. Il rapporto è $1 < \sqrt{3} < 2$. Per esercizio si può verificare che in un poligono regolare con n numero di lati dispari, se il raggio vale 1 allora il diametro vale $\sin(\pi/n)/\sin(\pi/(2n))$ (per n pari il diametro vale 2).

Si definisce *sequenza eccentrica* la sequenza delle eccentricità dei nodi. Una sequenza di interi non negativi si definisce eccentrica se è la sequenza eccentrica di un grafo. È un problema aperto caratterizzare in modo costruttivo una sequenza eccentrica, ovvero trovare una procedura che permetta di dire se una sequenza è eccentrica e, in caso affermativo, costruirne il grafo.

Si definisce *stato* s_i di un nodo i la somma delle distanze di tutti i nodi da i , ovvero $s_i = \sum_{j \in V} D(i, j)$. Sia l'eccentricità sia lo stato sono misure di centralità dei nodi. Tanto minori sono l'eccentricità e lo stato, tanto più centrale è il nodo rispetto al grafo.

La somma delle distanze fra tutte le coppie di nodi prende il nome di *indice di Wiener* del grafo, con notazione $W(G)$ (non si tratta di Norbert Wiener ma del chimico Harry Wiener), cioè (il fattore $1/2$ è dovuto

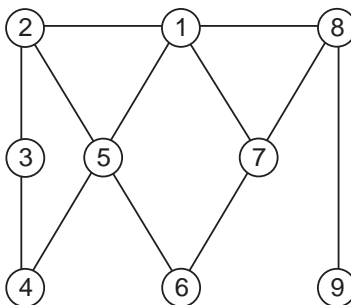


Figura 12.

al fatto che nella sommatoria ogni coppia viene contata due volte)

$$W(G) = \frac{1}{2} \sum_{i,j \in V} D(i,j) = \frac{1}{2} \sum_{i \in V} s_i$$

L'indice di Wiener ha una particolare importanza in chimica e viene usato per caratterizzare le molecole (viste come grafi in cui gli atomi sono i nodi e i legami fra gli atomi gli archi). Si vede facilmente che con lunghezze unitarie un cammino di n nodi ha indice di Wiener pari a

$$\frac{1}{2} \sum_{i=1}^n \frac{(n-i)(n-i+1)}{2} + \frac{i(i-1)}{2} = \frac{1}{2} \sum_{i=1}^n i(i-1) = \frac{n(2n+1)(n+1)}{12} - \frac{n(n+1)}{4} = \frac{n(n^2-1)}{6}$$

mentra una stella con $n+1$ nodi ha indice n^2 .

Ad esempio si consideri il grafo di Fig. 12 per il quale si hanno i seguenti valori di distanze ed eccentricità. Il centro è il nodo 1 (in generale il centro non è unico)

$$D(i,j) = \begin{pmatrix} 0 & 1 & 2 & 2 & 1 & 2 & 1 & 1 & 2 \\ 1 & 0 & 1 & 2 & 1 & 2 & 2 & 2 & 3 \\ 2 & 1 & 0 & 1 & 2 & 3 & 3 & 3 & 4 \\ 2 & 2 & 1 & 0 & 1 & 2 & 3 & 3 & 4 \\ 1 & 1 & 2 & 1 & 0 & 1 & 2 & 2 & 3 \\ 2 & 2 & 3 & 2 & 1 & 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 3 & 2 & 1 & 0 & 1 & 2 \\ 1 & 2 & 3 & 3 & 2 & 2 & 1 & 0 & 1 \\ 2 & 3 & 4 & 4 & 3 & 3 & 2 & 1 & 0 \end{pmatrix}, \quad s = \begin{pmatrix} 12 \\ 14 \\ 19 \\ 18 \\ 13 \\ 16 \\ 15 \\ 15 \\ 22 \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 4 \\ 3 \\ 3 \\ 3 \\ 3 \\ 4 \end{pmatrix}, \quad \rho = 2, \quad \Delta = 4$$

$$W(G) = 134$$

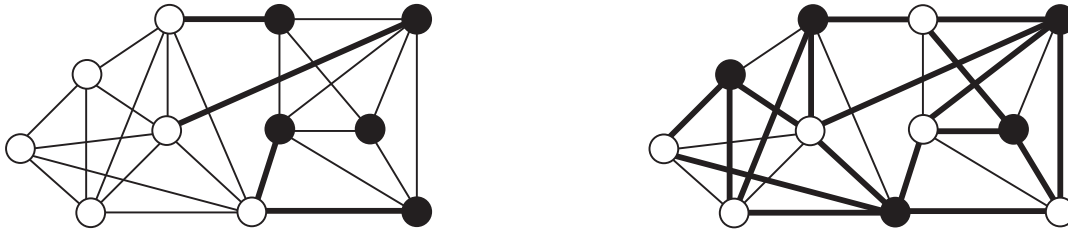


Figura 13.

9 Tagli

Il concetto di *taglio* di un grafo è simmetrico a quello di cammino. Dato un sottoinsieme proprio $S \subset V$ il taglio indotto da S è il sottoinsieme di archi

$$\delta(S) := \{(i, j) \in E : i \in S, j \notin S \text{ o viceversa}\}$$

La rimozione di questo insieme di archi sconnette il grafo, da cui il nome di taglio.

Un grafo è connesso se e solo se nessun taglio è vuoto. Inoltre un taglio ed un circuito hanno in comune un numero pari di archi. Di particolare interesse sono i problemi di trovare il taglio minimo (cioè con il minimo numero di archi) e quello di trovare il taglio massimo. Il primo è polinomiale e il secondo **NP**-difficile. I problemi verranno trattati in profondità a pag. 75 e seguenti. Si veda in Fig. 13 un grafo con il taglio minimo (non è l'unico) e quello massimo. Nelle figure S è dato dai vertici bianchi e il suo complemento dai vertici neri (o viceversa). Gli archi del taglio sono in tratto grosso.

Anche per i tagli, se sono definiti valori numerici c_e sugli archi, il valore del taglio (in questo caso si parla normalmente di *capacità del taglio*) è dato dalla somma di questi valori su $\delta(S)$ e il problema del taglio minimo o massimo consiste nel trovare il taglio di valore minimo o massimo. Ad esempio il grafo potrebbe rappresentare una rete di trasmissione dati e il valore di ogni arco potrebbe rappresentare la capacità dell'arco, cioè il numero di byte/sec che possono al massimo transitare sull'arco. Allora il minimo taglio dà informazione sulla massima quantità di informazione che può essere trasmessa da un nodo ad un altro nodo della rete.

La cardinalità del taglio minimo viene anche detta *connettività di arco*. Un concetto collegato è quello di *connettività di vertice* di un grafo G denotato $\kappa(G)$. Si tratta del minimo numero di vertici la cui rimozione (insieme con gli archi incidenti) rende sconnessa la parte rimanente del grafo. Questa definizione non si applica a grafi completi in quanto, anche rimuovendo vertici, la parte restante rimane un grafo completo e quindi non si riesce mai a sconnettere la parte rimanente. Esiste anche una definizione equivalente, che però si può applicare a grafi completi: $\kappa(G) < n$ è il più grande numero tale che, rimuovendo $k < \kappa(G)$ vertici, la parte rimanente non è sconnessa. In base a questa seconda definizione $\kappa(K_n) = n - 1$. Siccome per ogni grafo la rimozione di tutti i vertici incidenti in un dato nodo rende sconnesso il grafo (a meno che il nodo abbia grado $n - 1$), si ha $\kappa(G) \leq \min_i g_i(G)$. Il valore $\kappa(G)$ si può calcolare in tempo polinomiale.

10 Alberi e foreste

Un grafo connesso senza circuiti viene detto *albero*. Se si rimuove l'ipotesi di connessione e si mantiene quella di non esistenza di circuiti allora il grafo viene detto *foresta*. Quindi una foresta è un grafo le cui componenti connesse sono alberi. Si può dimostrare che per un albero vale la relazione $m = n - 1$, mentre per una foresta vale $m = n - k$ con k il numero di componenti connesse.

Un albero è inoltre contraddistinto dalle seguenti importanti proprietà: per ogni coppia di nodi esiste un unico cammino; l'eliminazione di un qualsiasi arco disconnette l'albero e identifica un taglio; l'aggiunta di un qualsiasi arco identifica un circuito.

Se un sottografo di supporto è una foresta o un albero, viene detto *foresta di supporto* oppure *albero di supporto* (*spanning tree*). Gli archi di G appartenenti ad un albero di supporto vengono detti *rami*, mentre quelli non appartenenti vengono detti *corde*.

Se agli archi sono associati dei pesi, è molto importante il problema di trovare un albero di supporto di peso minimo. Si tratta di un problema polinomiale per il quale esistono vari algoritmi. L'algoritmo detto di *Kruskal* esegue una scansione 'greedy' degli archi dal meno pesante al più pesante. Gli archi vengono inseriti uno alla volta, scartando quelli che generano un circuito. Dopo $n - 1$ inserzioni l'algoritmo termina.

Qualche volta interessa connettere fra loro solo alcuni vertici fissati del grafo, mentre gli altri possono anche non essere coperti dall'albero. In questo caso si parla di *alberi di Steiner*. Trovare un minimo albero di Steiner è **NP**-difficile. Si veda in Fig. 14 un grafo con un albero di supporto (a sinistra) e un albero di Steiner (a destra). I vertici neri sono quelli obbligati. Si tratta di un albero minimo (non è l'unico).

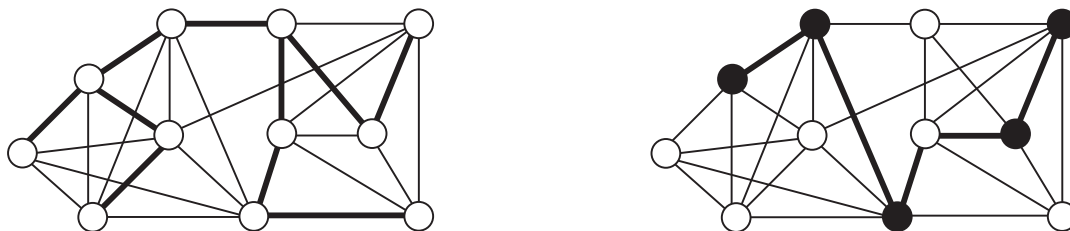


Figura 14.

11 Grafi planari

Normalmente i grafi vengono disegnati nel piano, associando i nodi a punti del piano e connettendo i nodi con linee a rappresentare gli archi. Anche se il disegno di un grafo non è intrinseco al grafo stesso, tuttavia vi sono alcune proprietà del grafo che si riflettono sulla rappresentazione del grafo. La domanda più importante che si può fare riguarda la possibilità di disegnare un grafo in modo che le linee dei suoi archi non si intersechino

(ovviamente al di fuori dei nodi). Se la cosa è possibile il grafo viene detto *planare*. Si può dimostrare che se gli archi, disegnati come linee qualsiasi di un grafo non si intersecano, allora esiste una rappresentazione planare in cui gli archi sono segmenti che non si intersecano. Il celebre teorema di Kuratowski afferma che un grafo è planare se e solo se non contiene un sottografo che non sia una sottodivisione di K_5 oppure di $K_{3,3}$. Per *sottodivisione* si intende un grafo ottenuto da un altro grafo inserendo nodi in un arco, ovvero sostituendo un arco (i, j) con due archi (i, k) , (k, j) ed eventualmente ripetendo questa operazione.

In un grafo planare il numero di archi è limitato. Vale la formula $m \leq 3n - 6$. Se, in particolare non ci sono circuiti di lunghezza 3 (ad esempio nei grafi bipartiti), vale la formula più restrittiva $m \leq 2n - 4$. Queste due formule possono costituire una rapida verifica di non planarità, se non soddisfatte. Determinare in generale se un grafo è planare si può eseguire in tempo lineare.

Si noti che, in base alla prima formula, in un grafo planare si ha $m \in O(n)$, mentre in generale $m \in O(n^2)$. Questo fatto permette di accelerare notevolmente alcuni algoritmi se applicati a grafi planari.

12 Grafi orientati

Un grafo si dice orientato quando le coppie di nodi sono ordinate. Quindi possono essere presenti sia l'arco (i, j) che l'arco (j, i) e si tratta di archi diversi. Una tale coppia di archi viene detta *antiparallela*. Nei grafi orientati è qualche volta utile poter disporre di più di una coppia (i, j) per gli stessi nodi (archi *paralleli* o *multipli*). Quindi si ha bisogno di una definizione di multigrafo come nei grafi non orientati.

I concetti di adiacenza di nodi e di incidenza nodi–archi sono indotti dal grafo (non orientato) associato. Il concetto di isomorfismo tra grafi viene naturalmente esteso tenendo conto anche dell'orientazione. Oltre al grado, definito come nel grafo associato, servono anche le definizioni di *grado esterno* g_i^+ , come il numero di archi uscenti dal nodo i , e *grado interno* g_i^- , come il numero di archi entranti nel nodo. Ovviamente il grado è la somma di grado interno e grado esterno.

Un grafo orientato *completo* ha come archi tutte le possibili coppie orientate di nodi. Quindi sia (i, j) che (j, i) sono presenti in un grafo orientato completo.

Un cammino ed un circuito vengono spesso definiti senza tener conto dell'orientazione degli archi. Ci sono casi in cui un cammino può percorrere gli archi indipendentemente dalla loro orientazione rispetto al cammino. In altri casi invece il cammino può percorrere gli archi solo secondo l'orientazione. Questa differenza va ovviamente evidenziata nel momento in cui si costruisce il modello del particolare problema. Nel secondo caso si usa il termine di *cammino orientato*, mentre per un circuito si usa il termine *ciclo*. Cammini e circuiti vengono definiti semplici o elementari come per i grafi non orientati. Un grafo orientato senza cicli viene detto *aciclico*.

Due nodi fra i quali esista un cammino vengono detti connessi e il grafo viene detto connesso se fra tutte le coppie di nodi esistono cammini (non necessariamente orientati). Quindi la proprietà di connessione è ereditata dal grafo associato. Nei grafi orientati c'è un ulteriore concetto. Due nodi fra i quali esista un ciclo vengono detti *fortemente connessi* e il grafo si dice fortemente connesso se tutte le coppie di nodi

sono fortemente connesse. Si noti che un grafo può essere connesso ma non fortemente connesso (come il semplice grafo dato da due nodi ed un arco). Definendo ogni nodo fortemente connesso con se stesso, anche la proprietà di connessione forte è una relazione di equivalenza che induce una partizione dei nodi del grafo in componenti fortemente connesse. Il grafo ottenuto per contrazione delle componenti fortemente connesse è aciclico.

Il concetto di taglio si particolarizza in un grafo orientato tenendo conto dell'orientazione. La notazione $\delta(S)$ si riferisce in ogni caso all'insieme di archi con esattamente un estremo in S , senza tener conto della loro orientazione, mentre si usano le notazioni

$$\delta^+(S) := \{(i, j) : i \in S, j \notin S\}, \quad \delta^-(S) := \{(i, j) : i \notin S, j \in S\}$$

per distinguere gli archi che vanno da S al suo complementare e quelli contrari. Ovviamente $\delta(S) = \delta^+(S) \cup \delta^-(S)$.

Anche i concetti di *albero* e *foresta* vengono definiti per un grafo orientato facendo riferimento alla struttura di grafo non orientato. Inoltre si definisce come *arborescenza* un albero in cui esista un nodo dal quale partono cammini orientati verso tutti gli altri nodi. In modo equivalente un'arborescenza è un grafo orientato in cui tutti i nodi hanno grado interno uguale ad 1 tranne un nodo in cui il grado interno vale 0.

13 Cosa è una matrice?

Prima di definire alcune matrici caratteristiche di un grafo, è forse utile ricordare cosa è una matrice. Si immagini una funzione che associa punti di uno spazio R^n a punti di uno spazio R^m . In questi casi si usa più frequentemente il termine *operatore*. In altri termini un vettore a n componenti viene “trasformato” dall'operatore in un vettore a m componenti. Indichiamo con $f : R^n \rightarrow R^m$ tale operatore. Operatori particolarmente semplici e di uso molto frequente nelle applicazioni sono quelli detti *lineari*. Un operatore è lineare se soddisfa le due seguenti proprietà

$$f(\alpha x) = \alpha f(x), \quad f(x + y) = f(x) + f(y) \tag{1}$$

dove α è un coefficiente che moltiplica tutte le componenti del vettore. Si noti che necessariamente $f(\mathbf{0}) = \mathbf{0}$ per ogni operatore lineare (verrà usato il simbolo $\mathbf{0}$ per indicare un vettore nullo, o una matrice nulla, e, analogamente il simbolo $\mathbf{1}$ per indicare un vettore con tutte le componenti uguali a 1). Si veda in Fig. 15 un possibile operatore lineare che trasforma il quadrato di sinistra nel parallelogramma di destra. Si noti che, in base alla linearità, un segmento che congiunge due punti A e B viene trasformato nel segmento che congiunge i punti $f(A)$ e $f(B)$. Infatti il segmento che congiunge i punti di coordinate (a_1, a_2) e (b_1, b_2) è dato dai punti di coordinate

$$\alpha \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + (1 - \alpha) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad 0 \leq \alpha \leq 1$$

e, applicando (1), si ha

$$f\left(\alpha \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + (1 - \alpha) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right) = f\left(\alpha \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}\right) + f\left((1 - \alpha) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right) = \alpha f\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}\right) + (1 - \alpha) f\left(\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right)$$

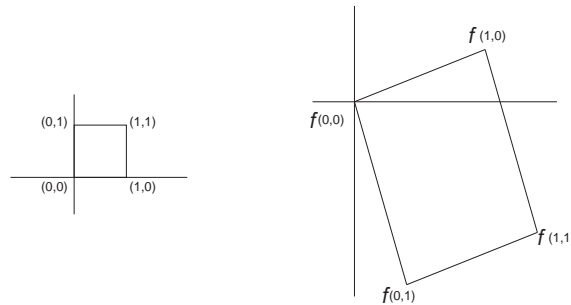


Figura 15. Funzione lineare: trasformazione di un quadrato

cioè ancora un segmento.

Quale tipo d'informazione è necessaria per identificare un operatore lineare? I punti di R^n e quelli di R^m vengono identificati con vettori, basati sulla scelta di assi cartesiani (non necessariamente ortogonali) nei rispettivi spazi. Ogni punto di R^n viene rappresentato, a partire da n vettori arbitrari (purché linearmente indipendenti) $\xi_1, \xi_2, \dots, \xi_n$, detti *base*, come loro combinazione lineare, cioè

$$x = x_1 \xi_1 + x_2 \xi_2 + \dots + x_n \xi_n$$

e quindi possiamo identificare x con i coefficienti della combinazione e scrivere

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$$

Tutto quello che serve conoscere per identificare un operatore lineare è come questo trasforma i vettori della base, cioè sapere chi sono i seguenti vettori in R^m

$$f(\xi_1), f(\xi_2), \dots, f(\xi_n)$$

Anche i punti in R^m sono identificati dalla combinazione lineare di m vettori $\zeta_1, \zeta_2, \dots, \zeta_m$, base di R^m e quindi dovrà essere

$$f(\xi_j) = a_{1j} \zeta_1 + a_{2j} \zeta_2 + \dots + a_{mj} \zeta_m$$

cioè

$$f(\xi_j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{pmatrix}$$

e siccome $x = x_1 \xi_1 + x_2 \xi_2 + \dots + x_n \xi_n$ si ha

$$f(x) = f(x_1 \xi_1 + x_2 \xi_2 + \dots + x_n \xi_n) = x_1 f(\xi_1) + x_2 f(\xi_2) + \dots + x_n f(\xi_n) =$$

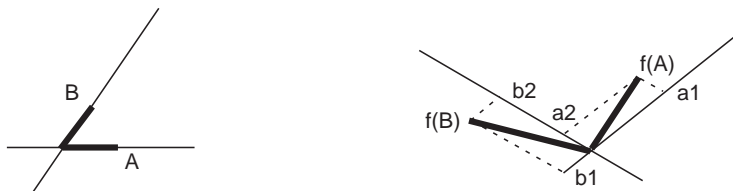


Figura 16. Operatore lineare: costruzione della matrice

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix}$$

Quindi tutta l'informazione sull'operatore lineare è data dalla 'matrice' di numeri

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

e il calcolo di $f(x)$ viene effettuato con la cosiddetta moltiplicazione righe per colonne della matrice per il vettore che identifica x , ovvero

$$f(x) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = Ax, \quad f(x)_i = (Ax)_i = \sum_j a_{ij} x_j$$

Si veda in Fig. 16 come costruire una possibile funzione. La matrice è data da

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

Allora la matrice che rappresenta l'operatore lineare della Fig. 15 è (vedi Fig. 17), supponendo che il quadrato in R^n sia il quadrato unitario,

$$\begin{pmatrix} 5 & 2 \\ 2 & -7 \end{pmatrix} \quad (2)$$

Se sono presenti due operatori lineari $f : R^n \rightarrow R^m$ e $g : R^m \rightarrow R^p$ allora la loro composizione $g \cdot f : R^n \rightarrow R^p$ è l'operatore $g(f(x)) = (g \cdot f)(x)$. Se f è rappresentato dalla matrice A e g dalla matrice B , si ha

$$(g \cdot f)(x)_i = g(f(x))_i = \sum_h b_{ih} f(x)_h = \sum_h b_{ih} \sum_j a_{hj} x_j = \sum_j \left(\sum_h b_{ih} a_{hj} \right) x_j$$

per cui l'operatore composto $g \cdot f$ è rappresentato dalla matrice BA dove il prodotto si ottiene come prodotto righe per colonne:

$$(BA)_{ij} = \sum_h b_{ih} a_{hj} \quad (3)$$

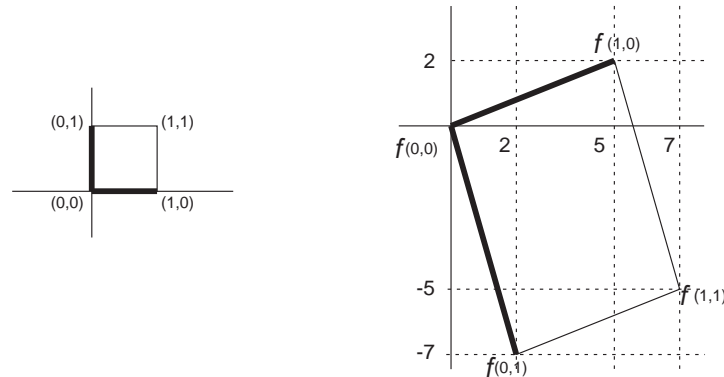


Figura 17. Operatore lineare: costruzione della matrice della Fig. 15

La matrice che si ottiene dalla matrice A scambiando le righe con le colonne, viene detta *matrice trasposta* e viene normalmente indicata con A^\top . Ovviamente $(A^\top)^\top = A$. Una matrice per cui $A = A^\top$ (necessariamente quadrata) viene detta *simmetrica*. Se i valori di A sono complessi, oltre all'operazione di trasposizione bisogna anche coniugare tutti i valori della matrice. La matrice così ottenuta viene detta *aggiunta* e si indica con la notazione A^\dagger . Una matrice per cui $A = A^\dagger$ viene detta *hermitiana*.

La trasposizione e il prodotto di matrici sono legati dalla relazione, facilmente deducibile da (3)

$$(AB)^\top = B^\top A^\top$$

da cui si vede che $A^\top A$ è necessariamente simmetrica.

Si definisce *spazio nullo* (o anche *nucleo*) di A , denotato anche $\mathcal{N}(A)$ l'insieme dei vettori x tali che $Ax = \mathbf{0}$. Si tratta cioè di tutti i vettori che vengono mandati nell'origine dall'operatore rappresentato da A . Necessariamente tale insieme è un sottospazio (cioè somme di elementi dell'insieme appartengono all'insieme come anche moltiplicazione per uno scalare). L'immagine $\mathcal{R}(A)$ (in inglese *range*) di A è un sottospazio la cui dimensione è un'importante caratteristica della matrice che viene detta *rango*, $r(A)$. Si può dimostrare che $r(A) = r(A^\top)$. Per una matrice $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ vale la relazione fondamentale di ortogonalità

$$\mathcal{N}(A) \perp \mathcal{R}(A^\top)$$

che implica $\dim \mathcal{N}(A) + r(A) = n$, cioè la dimensione dello spazio nullo è data dalla dimensione dello spazio stesso del dominio meno il rango.

Sono particolarmente interessanti quegli operatori lineari $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$, in cui dominio e codominio coincidono. In questo caso le matrici sono quadrate.

L'insieme dei valori a_{ii} di una matrice quadrata viene detto la *diagonale*. Una matrice quadrata in cui sono nulli tutti i termini non diagonali viene detta *matrice diagonale*. Una matrice diagonale corrisponde ad un operatore lineare particolarmente semplice in cui ogni componente agisce in modo indipendente dalle altre

componenti e di fatto la matrice è un operatore formato componendo assieme n operatori unidimensionali indipendenti. Una matrice diagonale può essere vista come derivata da un vettore $a \in R^n$ formando la diagonale con gli elementi di a (nell'ordine). In questo caso si usa spesso la notazione $\text{diag}(a)$.

Una matrice diagonale di grande importanza è la *matrice identica* in cui gli elementi diagonali sono tutti uni. La matrice identica viene denotata come I ($= \text{diag}(\mathbf{1})$). Quindi $Ix = x$ per ogni $x \in R^n$ e si tratta dell'unico operatore che lascia invariati tutti gli elementi del dominio.

Per una matrice $A : R^n \rightarrow R^n$, se $\mathcal{R}(A) = R^n$ allora anche $\mathcal{N}(A) = \mathbf{0}$ ed è quindi definito l'operatore inverso A^{-1} tale che $Ax = y$ se e solo se $x = A^{-1}y$, ovvero $A^{-1}A = I$. La matrice A^{-1} viene detta *matrice inversa* (o semplicemente *inversa*) di A . Prodotto di matrici e inversione sono legati dalla relazione (di facile dimostrazione)

$$(AB)^{-1} = B^{-1}A^{-1}$$

Una matrice quadrata in cui sono nulli tutti i termini a_{ij} con $j > i$, viene detta *triangolare inferiore*. Una matrice quadrata in cui sono nulli tutti i termini a_{ij} con $j < i$, viene detta *triangolare superiore*. Quindi una matrice diagonale è sia triangolare superiore che inferiore.

Spesso conviene pensare ad una matrice costituita non da numeri, ma da altre matrici. Una tale matrice viene detta *a blocchi*. Più formalmente sia data una partizione degli indici di riga $\{1, \dots, m\}$ in p insiemi I_1, \dots, I_p e una partizione degli indici di colonna $\{1, \dots, n\}$ in q insiemi J_1, \dots, J_q . Gli elementi con indice $i \in I_h$ e $j \in J_k$ formano una sottomatrice, altrimenti detta blocco. Interessano soprattutto matrici in cui alcuni blocchi sono matrici nulle. Ad esempio la matrice

$$\begin{pmatrix} 1 & 0 & 1 & 2 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 2 & 0 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \begin{matrix} A_{11} = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{pmatrix} & A_{12} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ A_{21} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} & A_{22} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \\ 2 & 0 \end{pmatrix} \end{matrix}$$

è una matrice a blocchi. Gli insiemi della partizione sono $I_1 = \{1, 2, 3, 4\}$, $I_2 = \{5, 6\}$, $J_1 = \{1, 2\}$, $J_2 = \{3, 4, 5\}$ e i blocchi sono le matrici A_{11} , A_{12} , A_{21} e A_{22} . Gli insiemi non devono essere costituiti necessariamente da interi consecutivi, ma in questo caso la struttura a blocchi è più evidente. Il prodotto di due matrici a blocchi segue le normali regole della moltiplicazione di due matrici applicate ai blocchi. Quindi

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} & A_{11}B_{13} + A_{12}B_{23} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} & A_{21}B_{13} + A_{22}B_{23} \end{pmatrix}$$

Se $p = q$ (quindi la matrice è quadrata se vista a blocchi) ed inoltre $a_{ij} = 0$ se $i \in I_h$, $j \in J_k$ e $h \neq k$, allora una matrice a blocchi rappresenta un operatore costituito da p operatori indipendenti fra loro. Se la matrice è quadrata e ogni blocco sulla diagonale è a sua volta una matrice quadrata allora la matrice viene

detta *diagonale a blocchi*, come ad esempio la seguente matrice

$$\begin{pmatrix} 1 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 3 \end{pmatrix} = \begin{pmatrix} A_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_3 \end{pmatrix} \quad \begin{aligned} A_1 &= \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 2 & 3 & 1 \end{pmatrix} \\ A_2 &= \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 0 & 1 & 1 \end{pmatrix} \\ A_3 &= \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \end{aligned}$$

Per matrici quadrate due parametri fondamentali sono il determinante e la traccia. Questi parametri sono intrinseci all'operatore e non dipendono dalla base tramite la quale l'operatore viene rappresentato come matrice. Il *determinante* viene definito in modo combinatorio attraverso la seguente sommatoria:

$$\det A = \sum_{\pi \in P_n} \sigma(\pi) \prod_{i=1}^n a_{i,\pi(i)} \quad (4)$$

dove π è una permutazione di n elementi e P_n è l'insieme di tutte le permutazioni di n elementi e $\sigma(\pi)$ è il segno della permutazione (pari o dispari a seconda del numero di scambi di due elementi adiacenti per passare dalla permutazione identica a π). Da (4) si vede che $\det A = \det A^T$.

Il calcolo del determinante tramite (4) ha una complessità esponenziale per cui è praticabile solo per bassi valori di n . Normalmente il calcolo numerico del determinante si effettua triangolando la matrice, attraverso un numero polinomiale di trasformazioni di righe. Per matrici triangolari la formula (4) si riduce a $\prod_{i=1}^n a_{ii}$ in quanto per ogni permutazione diversa dall'identica c'è almeno un $a_{i,\pi(i)} = 0$. Ovviamente anche per matrici diagonali il determinante è semplicemente dato da $\prod_{i=1}^n a_{ii}$. Analogamente il determinante di una matrice diagonale a blocchi è uguale al prodotto dei determinanti dei singoli blocchi.

Il determinante è nullo se e solo se le colonne (o equivalentemente le righe) sono linearmente dipendenti. Se il determinante è nullo la matrice si dice *singolare*. Allora esiste la matrice inversa se e solo se la matrice è non singolare.

Per matrici 2×2 la formula (4) diventa

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} a_{22} - a_{12} a_{21}$$

e per matrici 3×3

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{32} a_{21} - a_{31} a_{22} a_{13} - a_{32} a_{23} a_{11} - a_{33} a_{21} a_{12}$$

che può essere ricordata attraverso il seguente trucco, scrivendo la matrice aggiungendo a destra la prima e la seconda colonna

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{pmatrix}$$

e sommando il prodotto delle tre diagonali dall'alto in basso e sottraendo il prodotto delle tre diagonali dal basso verso l'alto (ma attenzione: il trucco vale solo per $n = 3$).

Il determinante ha un importante significato geometrico legato all'operatore che definisce la matrice. Il valore assoluto del determinante è uguale al volume del parallelepipedo che si ottiene trasformando il cubo unitario. Si consideri la Fig. 17. Il quadrato unitario viene trasformato nel parallelepipedo di destra, la cui area vale 39 (si facciano i conti per esercizio) cioè proprio il valore assoluto del determinante della matrice (2). Si noti che i vertici del parallelepipedo sono orientati in modo opposto a quelle del quadrato di sinistra. È come se la figura di destra fosse vista 'dal basso'. Il segno negativo del determinante indica proprio questa trasformazione dell'orientazione del quadrato. Se la matrice è singolare il quadrato viene deformato in un segmento, appunto ad area nulla.

Vale l'importante proprietà

$$\det(AB) = \det(A) \det(B) = \det(BA) \implies \det(A^k) = \det(A)^k$$

da cui abbiamo anche che $\det(A) \det(A^{-1}) = 1$. Inoltre vale la proprietà che permutando le righe o le colonne di una matrice il determinante rimane invariato in valore assoluto e cambia di segno se la permutazione è dispari. Quindi se si permutano nello stesso modo sia le righe che le colonne il determinante rimane invariato.

La *traccia* di una matrice viene definita da

$$\text{Tr}(A) = \sum_{i=1}^n a_{ii}$$

La traccia è lineare e quindi valgono le proprietà (di immediata verifica) $\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B)$, $\text{Tr}(\alpha A) = \alpha \text{Tr}(A)$, $\text{Tr}(A) = \text{Tr}(A^T)$. Inoltre $\text{Tr}(AB) = \text{Tr}(BA)$.

È opportuno notare che, date due matrici quadrate qualsiasi, mentre vale sempre $\det(AB) = \det(BA)$ e $\text{Tr}(AB) = \text{Tr}(BA)$, in generale si ha $AB \neq BA$, o, come si dice tecnicamente, le matrici non commutano. Quindi, nel calcolare $(A + B)^2$ il risultato è $A^2 + AB + BA + B^2$ e non $A^2 + 2AB + B^2$ a differenza dell'analogo calcolo per numeri.

Sono di grandissima importanza quei vettori non nulli che, una volta trasformati dalla funzione, mantengono inalterata la direzione (eventualmente cambiando verso). Questi vettori rappresentano dei valori invarianti per la trasformazione lineare e prendono il nome di *autovettori* e il coefficiente per cui il vettore viene moltiplicato prende il nome di *autovalore*. Quindi per un autovettore non nullo u e il corrispondente autovalore λ deve valere

$$Au = \lambda u \tag{5}$$

Si noti che ogni $u \in \mathcal{N}(A)$ è un autovettore con autovalore $\lambda = 0$. In generale gli autovalori e gli autovettori hanno valori complessi. Autovalori ed autovettori reali possono non esistere. Ad esempio la seguente matrice

$$\begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \quad (6)$$

ruota ogni vettore del piano in senso orario dell'angolo φ . Quindi, a meno che φ sia un multiplo di π , non può esistere un vettore, che, ruotato, abbia la medesima direzione. Gli autovalori di (6) sono $\lambda_1 = \cos \varphi - i \sin \varphi = e^{-i\varphi}$, $\lambda_2 = \cos \varphi + i \sin \varphi = e^{i\varphi}$ con autovettori $u^1 = (i, 1)$, $u^2 = (-i, 1)$ (cosa succede per φ multiplo di π ?).

Il polinomio $\det(\lambda I - A)$ (con variabile λ) viene detto *polinomio caratteristico* della matrice e gli autovalori si ottengono come radici del polinomio caratteristico, cioè

$$\det(\lambda I - A) = 0$$

Infatti (5) può essere scritta come $(\lambda I - A)u = 0$ e, per avere soluzioni non nulle, la matrice $(\lambda I - A)$ deve essere singolare. In base al teorema fondamentale dell'algebra esistono sempre n radici complesse (eventualmente con molteplicità) di un polinomio di grado n . Ad ogni autovalore distinto corrisponde sempre un autovettore distinto definito (a meno di una costante moltiplicativa) dalla soluzione di $(\lambda I - A)u = 0$.

Se un autovalore ha molteplicità k possono corrispondere h , con $1 \leq h \leq k$, autovettori linearmente indipendenti e ogni combinazione lineare di questi autovettori è a sua volta un autovettore con il medesimo autovalore. Quindi ad un autovalore con molteplicità k corrisponde un sottospazio invariante di dimensione h . Ad esempio la matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix}$$

ha l'autovalore 1 con molteplicità 3, ma l'unico autovettore $(0, 0, 1)$ se a e c sono entrambi diversi da 0. Se almeno uno fra a e c è nullo ma non sono nulli a , b e c , allora c'è un secondo autovettore. Se ad esempio $a = 0$ un secondo autovettore è $(-c, b, 0)$ ma lo è anche ogni vettore del sottospazio invariante dato da ogni combinazione lineare dei due vettori

$$\alpha \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} -c \\ b \\ 0 \end{pmatrix} = \begin{pmatrix} -\beta c \\ \beta b \\ \alpha \end{pmatrix}$$

Solo nel caso $a = b = c = 0$ (matrice identica) ci sono tre autovettori e il sottospazio invariante è in realtà tutto R^3 .

Per una matrice triangolare, e quindi anche per una matrice diagonale, gli autovalori sono uguali agli elementi della diagonale. Questo fatto discende dal fatto che il determinante di una matrice triangolare è uguale al prodotto dei termini sulla diagonale.

Se λ_i è autovalore di A , λ_i^k è autovalore di A^k , con gli stessi autovettori. Siccome $\det(\lambda I - A) = \det(\lambda I - A^T)$ la matrice trasposta ha gli stessi autovalori della matrice originale. Gli autovettori sono però

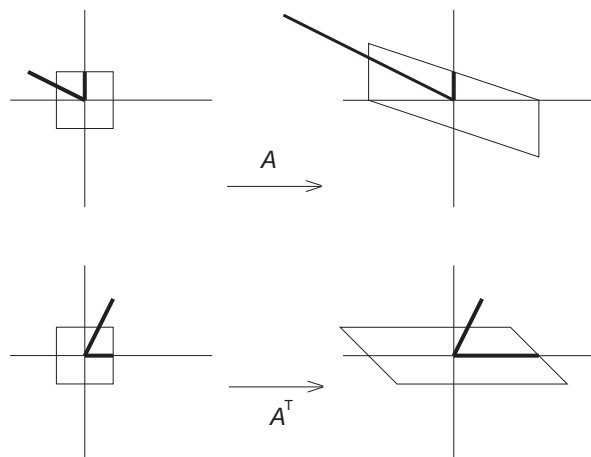


Figura 18.

in generale diversi. Un autovettore u di A e un autovettore v di A^T corrispondenti ad autovalori diversi sono ortogonali. Questa proprietà si vede facilmente da

$$v^T A u = v^T \lambda u = \mu v^T u \implies (\lambda - \mu) v^T u = 0$$

dove una prima volta si è sfruttato il fatto che u è autovettore di A con autovalore λ e una seconda volta si è sfruttato il fatto che v è autovettore di A^T (o autovettore sinistro di A come anche si dice) con autovalore μ . Se $\lambda \neq \mu$ deve necessariamente essere $v^T u = 0$. Si veda in Fig. 18 un esempio riferito alla matrice

$$A = \begin{pmatrix} 3 & 0 \\ -1 & 1 \end{pmatrix}, \quad \lambda_1 = 3, \quad \lambda_2 = 1, \quad u^1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad u^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad v^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v^2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad (7)$$

In Fig. 18 sono indicati in tratto grosso gli autovettori, destri nella parte alta della figura e sinistri in quella bassa.

Una matrice simmetrica ammette sempre autovalori reali ed n autovettori linearmente indipendenti ed ortogonali. Per dimostrare che gli autovalori sono reali si prenda una matrice hermitiana A con autovalore λ e corrispondente autovettore u (u e A possono avere valori complessi). Allora

$$u^\dagger A u = \lambda u^\dagger u = \lambda \|u\|^2$$

Siccome

$$(u^\dagger A u)^\dagger = u^\dagger A^\dagger u = u^\dagger A u$$

abbiamo che $u^\dagger A u$ è reale e quindi λ , dato dal rapporto di due numeri reali, è reale.

L'ortogonalità deriva dalla proprietà appena vista sugli autovettori di una matrice e della sua trasposta che coincidono in una matrice simmetrica. In più se un autovalore di una matrice simmetrica ha molteplicità

k , è presente un sottospazio di dimensione k in cui ogni elemento è un autovettore corrispondente a quell'autovalore. Se una matrice è simmetrica e positiva semidefinita (cioè $x^\top A x \geq 0$ per ogni x) allora tutti i suoi autovalori sono non negativi. Se è positiva definita (cioè $x^\top A x > 0$ per ogni $x \neq 0$) allora tutti i suoi autovalori sono positivi. Si noti che $A^\top A$ è necessariamente positiva semidefinita (oltre che simmetrica) ed è positiva definita se le colonne di A sono linearmente indipendenti.

L'insieme degli autovalori $\{\lambda_1, \dots, \lambda_n\}$ (contati con le loro molteplicità) prende il nome di *spettro della matrice*. Il valore $\max_i |\lambda_i|$ prende il nome di *raggio spettrale* della matrice e viene spesso indicato con $\rho(A)$. È una quantità importante perché le potenze A^k convergono alla matrice nulla se e solo se $\rho(A) < 1$, e divergono se $\rho(A) > 1$. Se $\rho(A) = 1$ le potenze di A possono rimanere limitate (senza tendere a zero) oppure divergere (questo dipende dall'esistenza o meno di autovettori distinti per autovalori multipli con valore assoluto uguale ad uno).

Se esistono n autovettori linearmente indipendenti lo spettro definisce intrinsecamente (ovvero indipendentemente dalla base dello spazio) la funzione lineare rappresentata dalla matrice. Non è difficile vedere che (da $\det(\lambda I - A) = \prod_i (\lambda - \lambda_i)$)

$$\text{Tr}(A) = \sum_i \lambda_i, \quad \det(A) = \prod_i \lambda_i$$

Limitandoci al caso di matrice simmetrica, per la quale esistono sempre n autovettori ortonormali (cioè ortogonali e di lunghezza unitaria) la relazione (5) si può scrivere

$$A u^1 = \lambda_1 u^1, \quad A u^2 = \lambda_2 u^2, \quad \dots, \quad A u^n = \lambda_n u^n,$$

che si può riscrivere in modo compatto, impaccando tutti i vettori colonna u^i nella matrice U ,

$$A U = U A$$

dove A è una matrice diagonale con gli autovalori sulla diagonale. Siccome U è formata da vettori ortonormali, $U^\top U = I$ e $U U^\top = I$, e quindi

$$A = U A U^\top \quad \text{oppure} \quad A = U^\top A U \quad (8)$$

La relazione (8) dice che, se come base dello spazio fossero stati scelti gli autovettori, allora, la funzione lineare sarebbe rappresentata dalla semplice matrice diagonale A . Siccome la scelta di una base non è una proprietà intrinseca della funzione lineare, si vede che ciò che caratterizza intrinsecamente la funzione lineare è il suo spettro.

Dati due vettori $a, b \in R^n$ si definisce *matrice diadica* la matrice M ottenuta come *prodotto esterno* dei due vettori $M_{ij} = a_i b_j$ rappresentabile anche come prodotto di un vettore colonna con uno riga:

$$M = a b^\top$$

Siccome esistono $n-1$ vettori linearmente indipendenti ed ortogonali ad un vettore dato, una matrice diadica ha sempre l'autovalore 0 presente con molteplicità $n-1$. Inoltre, siccome $b^\top u$ è uno scalare,

$$a b^\top u = \lambda u \quad \implies \quad u = a, \quad \lambda = b^\top a \quad (9)$$

Una forma diadica importante è data dalla matrice tutta di uni che possiamo scrivere come $\mathbf{1}\mathbf{1}^\top$ con autovalore $\mathbf{1}^\top\mathbf{1} = n$ e autovettore corrispondente $\mathbf{1}$ (spesso si trova il simbolo J per denotare la matrice $\mathbf{1}\mathbf{1}^\top$). Ogni altro vettore, purché sia ortogonale al vettore $\mathbf{1}$, è un autovettore con autovalore 0.

Lo spettro di una matrice diagonale a blocchi è semplicemente l'unione (tenendo conto delle molteplicità) degli spettri dei singoli blocchi e gli autovettori corrispondenti sono quelli dei singoli blocchi aggiungendo componenti uguali a zero in corrispondenza degli altri blocchi. Si veda ad esempio la matrice a blocchi (è la stessa matrice di (7) aggiungendo un blocco formato da una singola componente)

$$A = \begin{pmatrix} 3 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{matrix} \lambda_1 = 3 \\ \lambda_2 = 1, \\ \lambda_3 = 2 \end{matrix}, \quad u^1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \quad u^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Date due matrici A e B con i relativi spettri, che relazione c'è fra questi spettri e quello di $A + B$ oppure di AB ? In generale non c'è relazione. Solo se i sottospazi invarianti di A e di B coincidono allora anche i sottospazi invarianti di $A + B$ e AB coincidono con quelli di A e B e gli autovalori sono rispettivamente la somma e il prodotto. Infatti sia u appartenente ad un sottospazio invariante per il quale l'autovalore di A sia λ e l'autovalore di B sia μ . Allora

$$(A + B)u = Au + Bu = \lambda u + \mu u = (\lambda + \mu)u, \quad (AB)u = ABu = \mu Au = \mu \lambda u$$

Se esistono n autovettori indipendenti sia per A che per B e gli spazi invarianti coincidono allora le matrici commutano. Infatti per ogni autovettore u si ha

$$ABu = \mu \lambda u = BAu \implies (AB - BA)u = \mathbf{0}$$

Siccome ogni punto di R^n può essere espresso come combinazione lineare degli autovettori, ne consegue che $(AB - BA)$ è la matrice nulla, ovvero A e B commutano.

14 Matrici d'incidenza di grafi non orientati

Dato un grafo $G = (V, E)$, ogni sottoinsieme $E' \subset E$ oppure $V' \subset V$ può essere rappresentato tramite vettori d'incidenza

$$x(E')_e := \begin{cases} 1 & \text{se } e \in E' \\ 0 & \text{se } e \notin E' \end{cases} \quad y(V')_i := \begin{cases} 1 & \text{se } i \in V' \\ 0 & \text{se } i \notin V' \end{cases}$$

La matrice d'incidenza nodi-archi B di un grafo non orientato viene definita come

$$B_i^e := \begin{cases} 1 & \text{se } i \text{ ed } e \text{ sono incidenti} \\ 0 & \text{altrimenti} \end{cases}$$

Si noti che si assume un insieme di archi non vuoto. Un grafo totalmente sconnesso non è quindi rappresentabile in quanto si avrebbe un insieme vuoto di colonne. Dalla definizione ogni colonna di B ha solo due componenti diverse da 0 ed uguali ad 1. Come esempio la matrice d'incidenza del grafo in Fig. 19 è

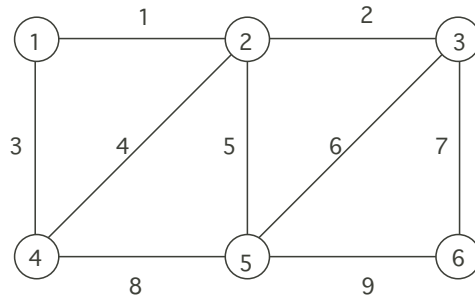


Figura 19.

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

Se il grafo non è connesso la matrice è a blocchi e ogni blocco non nullo si riferisce ad una componente connessa, con la particolarità però che nodi isolati dovrebbero formare un blocco di una riga e *nessuna* colonna (vedi sopra).

Si noti che la somma degli elementi di ogni riga è il grado del nodo corrispondente. La somma degli elementi di una riga, fatta riga per riga, si può anche scrivere moltiplicando la matrice d'incidenza per un vettore tutto di uni, ovvero dal vettore d'incidenza di E , cioè $Bx(E)$. Allora se $x(E')$ è il vettore d'incidenza di un sottoinsieme E' di archi, il grado di $G' = (V, E')$ può essere espresso da $Bx(E')$.

Se $y(V')$ è il vettore d'incidenza di un sottoinsieme V' di nodi, il numero di estremi contenuti in V' per ogni arco può essere espresso da $y(V')B$. Si può vedere che un sottoinsieme E' di archi è un taglio se e solo se esiste un vettore y di 1 e 0, per cui il vettore d'incidenza $x(E')$ soddisfa $yB = x(E') \pmod{2}$.

Inoltre è facile verificare che V' è una copertura di nodi se e solo se $y(V')B \geq \mathbf{1}$ e V' è un insieme stabile se e solo se $y(V')B \leq \mathbf{1}$.

Si consideri una qualsiasi sottomatrice quadrata B' di B ottenuta scegliendo m archi E' da E . Si può dimostrare che B' è non singolare se e solo se (V, E') ha esattamente un circuito dispari per ogni componente connessa. Ad esempio, se $E' = \{1, 3, 4, 6, 7, 9\}$ (grafo di Fig. 19) la sottomatrice B' è

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

che sappiamo essere non singolare perché E' ha due componenti connesse e ognuna ha un circuito dispari. Invece se $E' = \{1, 3, 5, 7, 8, 9\}$ la sottomatrice B' è

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

che sappiamo essere singolare perché E' ha un'unica componente connessa con un circuito pari. Come conseguenza di questo fatto ogni sottomatrice quadrata di un grafo bipartito deve essere singolare, perché non esistono circuiti dispari in un grafo bipartito, e quindi le righe della matrice d'incidenza di un grafo bipartito sono sempre linearmente dipendenti. Viceversa, dato un grafo non bipartito connesso, basta scegliere un sottoinsieme d'archi dato da un circuito dispari (che deve esistere in un grafo non bipartito) e aggiungere archi senza generare circuiti (basta pensare di collassare i nodi del circuito dispari e generare un albero di supporto sul grafo collassato). La sottomatrice quadrata individuata da questi archi non è singolare e quindi le righe della matrice sono linearmente indipendenti. Quindi, dato un grafo connesso, le righe di una matrice d'incidenza sono linearmente dipendenti se e solo se il grafo è bipartito. Se il grafo non è connesso si richiede anche che ogni componente connessa non sia bipartita per avere le righe linearmente indipendenti.

15 Matrici d'incidenza di grafi orientati

La matrice d'incidenza nodi-archi B di un grafo orientato viene definita come

$$B_i^e := \begin{cases} 1 & \text{se l'arco } e \text{ è uscente dal nodo } i \\ -1 & \text{se l'arco } e \text{ è entrante nel nodo } i \\ 0 & \text{altrimenti} \end{cases}$$

Anche in questo caso si richiede che l'insieme degli archi non sia vuoto. La matrice è a blocchi se il grafo è sconnesso (con la particolarità già vista per i nodi isolati).

Dalla definizione ogni colonna di B ha solo due componenti diverse da 0 che valgono 1 e -1 . Si vede immediatamente che le righe di B sono linearmente dipendenti perché sommandole si ottiene 0. Ad esempio la matrice d'incidenza del grafo di Fig. 20 è

$$B = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

Se aggiungiamo ad B una colonna tutta di zeri con un unico 1 in una posizione qualsiasi, ad esempio in prima posizione in corrispondenza del nodo 1, e otteniamo la matrice B' , questa matrice ha un'importante

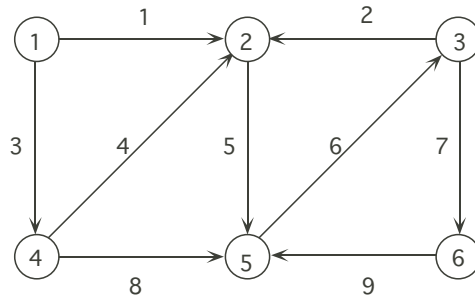


Figura 20.

proprietà: ogni sottomatrice C di B' , con n righe e n colonne, è non singolare se e solo se C contiene la colonna aggiunta e le altre $n - 1$ colonne si riferiscono ad un albero di supporto del grafo. In questo caso $|\det(C)| = 1$.

$$B' = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

Se C è formata con la colonna aggiunta e con le colonne relative agli archi (1, 2), (3, 2), (1, 4), (2, 5) e (3, 6) che formano un albero di supporto si ha

$$C = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

che è sopra triangolare ed ha determinante 1. Si noti che, proprio perché gli archi scelti formano un albero di supporto, le colonne possono essere ordinate in modo da aver una matrice C sopra triangolare. Basta 'costruire' l'albero partendo dal nodo 1 e 'attaccando' gli archi uno alla volta.

Se invece sostituiamo l'arco (3, 6) con l'arco (5, 3) (e quindi non si ha un albero) si ottiene la matrice singolare

$$C = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Si noti che in presenza di un circuito basta scegliere le colonne degli archi del circuito e si hanno delle colonne linearmente dipendenti (in ogni riga c'è esattamente un 1 e un -1) e quindi la matrice è singolare.

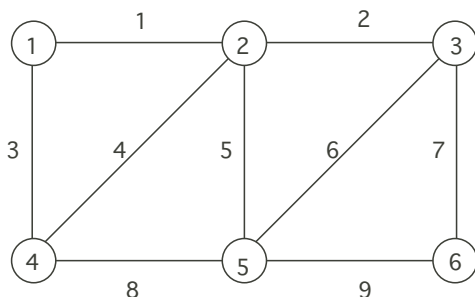


Figura 21.

Nell'esempio la singolarità della matrice si vede immediatamente dall'ultima riga nulla, ma in generale la singolarità si deduce dalla presenza di un circuito che nell'esempio corrisponde alle colonne 3, 5 e 6 di C .

Allora un grafo orientato connesso ha una matrice d'incidenza con righe sempre linearmente dipendenti e il suo spazio nullo sinistro ha dimensione 1, cioè esiste un unico vettore (a meno di moltiplicazione per uno scalare), ovvero il vettore $\mathbf{1}$, tale che $\mathbf{1}^\top A = \mathbf{0}$. Se il grafo non è connesso ogni componente connessa dà luogo ad un vettore tutto di uni sui nodi della componente e di zeri altrove che appartiene allo spazio nullo sinistro della matrice. Ne consegue che il numero delle componenti connesse è uguale alla dimensione dello spazio nullo sinistro.

Se $x(E')$ è il vettore d'incidenza di un sottoinsieme E' di archi, si ha per $G' = (V, E')$,

$$g^+(G') - g^-(G') = B x(E')$$

Se $y(V')$ è il vettore d'incidenza di un sottoinsieme V' di nodi, il vettore $y(V') B$ è la riga della matrice d'incidenza relativa allo pseudonodo V' del grafo contratto. Quindi se si definisce la seguente matrice d'incidenza partizione-nodi:

$$P_S^i := \begin{cases} 1 & \text{se il nodo } i \text{ appartiene al sottoinsieme } S \text{ della partizione} \\ 0 & \text{altrimenti} \end{cases}$$

la matrice d'incidenza \bar{B} del grafo contratto è data da $\bar{B} = P B$, a meno di alcune colonne tutte nulle.

16 Matrici d'adiacenza di un grafo non orientato

La matrice d'adiacenza A di un grafo non orientato viene definita come

$$A_{ij} := \begin{cases} 1 & \text{se i nodi } i \text{ e } j \text{ sono adiacenti} \\ 0 & \text{altrimenti} \end{cases}$$

Il grafo è sconnesso se e solo se la matrice è diagonale a blocchi. La matrice d'adiacenza del grafo di Fig. 19, qui ripetuto come Fig. 21, è

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

Dalla definizione si ha che la matrice d'adiacenza di un grafo non orientato è simmetrica. È interessante l'interpretazione delle potenze di A . Infatti si ha

$$(A^2)_{ij} = \sum_k A_{ik} A_{kj} \quad (10)$$

che, per il grafo di Fig. 21 diventa

$$A^2 = \begin{pmatrix} 2 & 1 & 1 & 1 & 2 & 0 \\ 1 & 4 & 1 & 2 & 2 & 2 \\ 1 & 1 & 3 & 2 & 2 & 1 \\ 1 & 2 & 2 & 3 & 1 & 1 \\ 2 & 2 & 2 & 1 & 4 & 1 \\ 0 & 2 & 1 & 1 & 1 & 2 \end{pmatrix}$$

Gli unici termini diversi da zero nella sommatoria (10) sono quelli per cui sia $A_{ik} = 1$ che $A_{kj} = 1$, cioè i e k sono adiacenti ed anche k e j sono adiacenti. Quindi $A_{ik} A_{kj} = 1$ se e solo se esiste un cammino di due archi da i a j passante per k . Sommando sui diversi possibili nodi intermedi, si ha che la sommatoria conta il numero di cammini di due archi fra i e j . Quindi $(A^k)_{ij}$ è il numero di cammini distinti di lunghezza k fra i nodi i e j . Inoltre $(A^k)_{ii}$ è il numero di circuiti distinti di lunghezza k passanti per il nodo i , contati tenendo conto del verso di percorrenza. Quindi, siccome il grafo non è orientato e anche i circuiti vengono definiti senza tener conto dell'orientazione, il numero di circuiti distinti passanti per i non è in realtà $(A^k)_{ii}$ e non è nemmeno $(A^k)_{ii}/2$, come forse si potrebbe pensare. Infatti, essendo ogni circuito percorso nei due sensi, sembrerebbe che basti dividere per due il numero totale. Questo è vero se k è dispari. Purtroppo per k pari ci sono circuiti palindromi che vengono contati solo una volta e quindi per conoscere il numero corretto di circuiti distinti di lunghezza $2k$ pari bisogna eseguire il seguente calcolo

$$\frac{(A^{2k})_{ii} + \sum_{h=1}^n (A^k)_{ih}}{2} \quad (11)$$

dove, appunto vengono aggiunti a $(A^{2k})_{ii}$ tutti i circuiti palindromi (termine $\sum_{h=1}^n (A^k)_{ih}$) in modo da avere tutti i circuiti contati due volte.

Si vede immediatamente che $(A^2)_{ii} = g(i)$ e quindi $\text{Tr}(A^2) = 2m$. I circuiti di lunghezza 3 sono necessariamente semplici e quindi $(A^3)_{ii}/2$ è il numero di triangoli che contengono il nodo i e $\text{Tr}(A^3)/6$ è il numero di triangoli del grafo.

Per il grafo di Fig. 19 si ha

$$A^3 = \begin{pmatrix} 2 & 6 & 3 & 5 & 3 & 3 \\ 6 & 6 & 8 & 7 & 9 & 3 \\ 3 & 8 & 4 & 4 & 7 & 5 \\ 5 & 7 & 4 & 4 & 8 & 3 \\ 3 & 9 & 7 & 8 & 6 & 6 \\ 3 & 3 & 5 & 3 & 6 & 2 \end{pmatrix}$$

Ad esempio i 6 cammini da 1 a 2 sono: $1 \rightarrow 2 \rightarrow 1 \rightarrow 2$, $1 \rightarrow 4 \rightarrow 1 \rightarrow 2$, $1 \rightarrow 2 \rightarrow 3 \rightarrow 2$, $1 \rightarrow 2 \rightarrow 4 \rightarrow 2$, $1 \rightarrow 2 \rightarrow 5 \rightarrow 2$, $1 \rightarrow 4 \rightarrow 5 \rightarrow 2$. Si noti che si contano cammini non semplici. Saper contare i cammini semplici di lunghezza k fra i e j in tempo polinomiale implicherebbe saper risolvere il problema del circuito hamiltoniano in tempo polinomiale. Si ha $\text{Tr}(A^3) = 24$ e quindi vi sono 4 triangoli. Per lo stesso grafo si ha

$$A^4 = \begin{pmatrix} 11 & 13 & 12 & 11 & 17 & 6 \\ 13 & 30 & 18 & 21 & 24 & 17 \\ 12 & 18 & 20 & 18 & 21 & 11 \\ 11 & 21 & 18 & 20 & 18 & 12 \\ 17 & 24 & 21 & 18 & 30 & 13 \\ 6 & 17 & 11 & 12 & 13 & 11 \end{pmatrix}$$

Gli 11 circuiti ‘contati’ da A_{11}^4 sono

$$(1, 2, 1, 2, 1), \quad (1, 4, 1, 2, 1), \quad (1, 2, 3, 2, 1), \quad (1, 2, 4, 2, 1), \quad (1, 2, 5, 2, 1), \quad (1, 4, 5, 2, 1), \\ (1, 2, 1, 4, 1), \quad (1, 4, 1, 4, 1), \quad (1, 4, 2, 4, 1), \quad (1, 2, 5, 4, 1), \quad (1, 4, 5, 4, 1)$$

però il n. 2 è uguale al n. 7 (con percorrenze in versi opposti), il n. 6 al n. 10. Gli altri 7 circuiti sono palindromi. Quindi in realtà il numero di circuiti diversi passanti per il vertice 1 è 9. Questo valore si trova anche applicando la formula (11) che dà $(11 + 7)/2 = 9$.

Queste proprietà del quadrato della matrice di adiacenza possono essere usate per caratterizzare i grafi fortemente regolari (si riveda la definizione a pag. 5) direttamente dalla matrice di adiacenza. Infatti $(A^2)_{ij} = r$ se $(i, j) \in E$, cioè $A_{ij} = 1$, $(A^2)_{ij} = s$ se $(i, j) \notin E$, cioè $A_{ij} = 0$ e $(A^2)_{ii} = k$, da cui

$$A^2 + (s - r)A + (s - k)I = s\mathbf{1}\mathbf{1}^\top \quad (12)$$

Il vettore $A\mathbf{1}$ è il vettore dei gradi. Se $x(S)$ è il vettore d’incidenza di un sottoinsieme $S \subset V$, il vettore $Ax(S)$ conta, per ogni nodo $i \in V$ quanti nodi di S sono adiacenti a i . Quindi

$$(\mathbf{1} - x(S))^\top Ax(S) \quad (13)$$

è la cardinalità del taglio indotto da S . Un altro modo utile di esprimere la cardinalità consiste nell’usare vettori che, anziché valere 0 o 1, valgono -1 o $+1$. Più esattamente, dato un sottoinsieme S si definisca

$$y_i(S) = \begin{cases} 1 & \text{se } i \in S \\ -1 & \text{se } i \notin S \end{cases}$$

Si vede immediatamente che la relazione che lega il vettore d'incidenza $x(S)$ al vettore $y(S)$ è data da

$$x(S) = \frac{1}{2}(\mathbf{1} + y(S)), \quad \text{cioè anche} \quad \mathbf{1} - x(S) = \frac{1}{2}(\mathbf{1} - y(S))$$

per cui (13) diventa (si ricordi che A è simmetrica)

$$\frac{1}{4}(\mathbf{1} - y(S))^\top A (\mathbf{1} + y(S)) = \frac{1}{4}(\mathbf{1}^\top A \mathbf{1} - y(S)^\top A y(S)) = \frac{m}{2} - \frac{1}{4}y(S)^\top A y(S) \quad (14)$$

Quindi la cardinalità del taglio indotto da S è legata alla forma quadratica $y(S)^\top A y(S)$ che in dettaglio è

$$\sum_{i,j \in V} a_{ij} y_i y_j$$

Nella sommatoria i termini diversi da zero sono solo quelli per nodi i e j adiacenti e allora si sommano (due volte) tutti gli archi che non sono nel taglio e si sottraggono (sempre due volte) gli archi del taglio da cui (14).

Il fatto importante è che il taglio di cardinalità minima o massima è associato al massimo o al minimo della forma quadratica $y(S)^\top A y(S)$.

Qualche volta, quando agli archi sono associati dei pesi w_{ij} può essere utile definire una *matrice d'adiacenza pesata*

$$(\tilde{A})_{ij} = \begin{cases} w_{ij} & \text{se } (i, j) \in E \\ 0 & \text{altrimenti} \end{cases}$$

In modo del tutto analogo la capacità del taglio indotto da S si può esprimere nei due modi alternativi come

$$w(\delta(S)) = (\mathbf{1} - x(S))^\top \tilde{A} x(S), \quad \frac{W}{2} - \frac{1}{4}y(S)^\top \tilde{A} y(S) \quad (15)$$

dove $W = \sum_{e \in E} w_e$.

17 Matrici d'adiacenza di grafi prodotto

Si considerino i grafi di Fig. 8 le cui matrici d'adiacenza sono

$$A_1 = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Si ricorda che i nodi del grafo prodotto sono le coppie $[i, h]$, $i \in V_1$, $h \in V_2$. Nel grafo prodotto cartesiano $G_1 \square G_2$ un arco è presente fra $[i, h]$ e $[j, k]$ se $i = j$ e $(h, k) \in E_2$ oppure $h = k$ e $(i, j) \in E_1$. Quindi se i nodi del grafo prodotto vengono numerati come $[1, A]$, $[1, B]$, $[1, C]$, $[1, D]$, $[2, A]$, $[2, B]$, \dots (usando la notazione

usata precedentemente), la matrice d'adiacenza del grafo prodotto è

$$\begin{pmatrix} A_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & A_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & A_2 \end{pmatrix} + \begin{pmatrix} \mathbf{0} & I & \mathbf{0} & I & \mathbf{0} \\ I & \mathbf{0} & I & I & I \\ \mathbf{0} & I & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ I & I & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

La prima matrice tiene conto degli archi provenienti da $i = j$ e $(h, k) \in E_2$ e la seconda (ottenuta da A_1 mettendo una matrice quadrata $n_2 \times n_2$ di zeri se in A_1 c'è uno zero oppure la matrice identica I di dimensione n_2 se l'elemento è un uno) tiene conto degli archi provenienti da $h = k$ e $(i, j) \in E_1$. Matrici ottenute in questo modo sono il risultato di un'operazione binaria detta *prodotto di Kronecker*. In generale il prodotto di Kronecker $B \otimes C$ è la matrice a blocchi

$$\begin{pmatrix} b_{11}C & b_{12}C & \dots & b_{1n}C \\ b_{21}C & b_{22}C & \dots & b_{2n}C \\ \dots & \dots & \dots & \dots \\ b_{m1}C & b_{m2}C & \dots & b_{mn}C \end{pmatrix}$$

per cui la matrice d'adiacenza del grafo prodotto cartesiano si può scrivere anche usando il prodotto di Kronecker come

$$I \otimes A_2 + A_1 \otimes I$$

Per il prodotto tensoriale la matrice d'adiacenza è data dalla matrice A_1 fatta di blocchi anziché di singoli elementi e ogni blocco della nuova matrice è un blocco di zeri se in A_1 c'è uno zero oppure la matrice A_2 se l'elemento è un uno. Quindi

$$\begin{pmatrix} \mathbf{0} & A_2 & \mathbf{0} & A_2 & \mathbf{0} \\ A_2 & \mathbf{0} & A_2 & A_2 & A_2 \\ \mathbf{0} & A_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_2 & A_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} = A_1 \otimes A_2$$

e la matrice d'adiacenza del prodotto forte è quindi

$$A_1 \otimes A_2 + I \otimes A_2 + A_1 \otimes I$$

18 Matrici d'adiacenza di un grafo orientato

La *matrice d'adiacenza* A di un grafo orientato viene definita come

$$A_{ij} := \text{numero di archi } (i, j) \text{ presenti nel grafo}$$

Ad esempio la matrice d'adiacenza del grafo di Fig. 20 è

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

Si trova in letteratura anche la definizione opposta, cioè A_{ij} è il numero di archi da j a i . La matrice di una definizione è semplicemente la trasposta dell'altra e quindi non c'è nessuna differenza sostanziale. Qui ci atteniamo alla definizione data che sembra più coerente con l'ordine della coppia (i, j) . La matrice d'adiacenza di un grafo orientato è simmetrica se e solo se ogni arco è presente con il suo arco antiparallelo.

Anche in questo caso è interessante l'interpretazione delle potenze di A . Infatti si ha che $(A^k)_{ij}$ = numero di cammini orientati distinti di lunghezza k fra i nodi i e j . Per il grafo di Fig. 20 si ha

$$A^4 = \begin{pmatrix} 0 & 2 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \end{pmatrix}$$

I due cammini di lunghezza 4 da 1 a 2 sono $1 \rightarrow 2 \rightarrow 5 \rightarrow 3 \rightarrow 2$ e $1 \rightarrow 4 \rightarrow 5 \rightarrow 4 \rightarrow 2$. Se un grafo è aciclico deve esistere $k \leq n$ tale che $A^h = 0$ per ogni $h \geq k$. A differenza dei grafi non orientati, i valori $(A^k)_{ii}$ contano i cicli distinti passanti per i e $\text{Tr}(A^k)$ conta tutti i cicli di G di lunghezza k , tenendo presente però che un ciclo viene contato tante volte quante sono i suoi nodi. Non si creda che dividendo $\text{Tr}(A^k)$ per k si ottenga il numero dei cicli. Non essendo i cicli semplici qualche ciclo di k archi contiene nodi ripetuti e quindi meno di k . Questo implica che la divisione per k produce risultati errati. Anche in questo caso il calcolo esatto del numero di cicli al netto delle ripetizioni dei suoi nodi è molto complesso.

Se il grafo è aciclico si possono sempre permutare righe e colonne in modo che la matrice sia triangolare superiore. Per un grafo aciclico i cammini sono necessariamente semplici e quindi A^k conta i cammini semplici di lunghezza k . Non potendo essere presenti cammini di lunghezza n si deduce che $A^n = 0$. La matrice

$$C = A + A^2 + A^3 + \dots + A^{n-1}$$

conta tutti i cammini (per ogni possibile lunghezza). Possiamo anche esprimere C in altro modo. Siccome

$$(I - A)(I + C) = I - A + C - AC = I$$

abbiamo anche

$$C = (I - A)^{-1} - I$$

Se ad esempio cambiano l'orientazione dell'arco $(5, 3)$ in $(3, 5)$ per il grafo Fig. 20 abbiamo la matrice d'adiacenza

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

per la quale si ha già $A^4 = \mathbf{0}$. Il calcolo di C porta a

$$C = (I - A)^{-1} - I = \begin{pmatrix} 1 & -1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 & -1 & -1 \\ 0 & -1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}^{-1} - I = \begin{pmatrix} 0 & 2 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 3 & 1 \\ 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

19 Matrice Laplaciana

Dato il grafo non orientato $G = (V, E)$ si orientino arbitrariamente gli archi e sia B la matrice d'incidenza del grafo orientato risultante. La matrice $L := BB^T$ viene detta *matrice Laplaciana* (o più brevemente *Laplaciano*) oppure *matrice di grado* del grafo G . Si vede facilmente che $L = \text{diag}(g) - A$, dove g e la matrice d'adiacenza A si riferiscono al grafo non orientato G .

Siccome le righe di B sono linearmente dipendenti, L è singolare. Inoltre, siccome $L := BB^T$, la matrice è positive semidefinita.

Il Laplaciano interviene in modo naturale quando si debba calcolare

$$\sum_{(ij) \in E} (x_i - x_j)^2 \quad (16)$$

dove i valori x_i sono numeri associati ai nodi. Infatti si ha

$$\sum_{(ij) \in E} (x_i - x_j)^2 = (B^T x)^T (B^T x) = x^T B B^T x = x^T L x$$

Usando questa relazione la cardinalità di un taglio può essere espressa in modo più sintetico dal Laplaciano che non dalla matrice d'adiacenza come nella formula (14). Infatti se si introducono le variabili y_i , tali che $y_i = 1$ se $i \in S$ e $y_i = -1$ se $i \notin S$, i termini che non si annullano in (16) corrispondono agli archi del taglio e per questi si ha $(x_i - x_j)^2 = 4$. Quindi

$$|\delta(S)| = \frac{1}{4} y(S)^T L y(S) \quad (17)$$

Se agli archi del grafo sono associati valori di capacità w_{ij} si può definire un Laplaciano pesato come

$$\tilde{L}_{ij} = \begin{cases} -w_{ij} & \text{se } (i, j) \in E \\ \sum_{k \in \delta(i)} w_{ik} & \text{se } i = j \\ 0 & \text{altrimenti} \end{cases} \quad (18)$$

o anche come

$$\tilde{L} = B \operatorname{diag}(w) B^\top$$

e si ha

$$\sum_{(ij) \in E} w_{ij} (x_i - x_j)^2 = (B^\top x)^\top \operatorname{diag}(w) (B^\top x) = x^\top B \operatorname{diag}(w) B^\top x = x^\top \tilde{L} x$$

per cui la capacità di un taglio $\delta(S)$ può anche in questo caso più generale essere espressa come

$$c(\delta(S)) = \frac{1}{4} y(S)^\top \tilde{L} y(S) \quad (19)$$

Il Laplaciano è strettamente collegato agli alberi di supporto del grafo. La cosa non deve meravigliare dato che $L = B B^\top$ e le sottomatrici quadrate di B' (B' è ottenuta aggiungendo a B la colonna $(1, 0, 0, \dots, 0)$, si riveda a pag. 35) sono singolari se e solo se le colonne sono associate ad un albero di supporto. Se definiamo $L' = B' B'^\top$ si vede facilmente che L' coincide con L tranne che per l'elemento $L'_{1,1} := L_{1,1} + 1$. Dal Teorema di Cauchy-Binet si ottiene che $|\det L'|$ è uguale al numero di alberi di supporto distinti di G . Analogo risultato si ottiene per $\det(L'')$ dove L'' è ottenuta da L sopprimendo la prima riga e la prima colonna. Per il grafo di Fig. 19 si ha

$$L = \begin{pmatrix} 2 & -1 & 0 & -1 & 0 & 0 \\ -1 & 4 & -1 & -1 & -1 & 0 \\ 0 & -1 & 3 & 0 & -1 & -1 \\ -1 & -1 & 0 & 3 & -1 & 0 \\ 0 & -1 & -1 & -1 & 4 & -1 \\ 0 & 0 & -1 & 0 & -1 & 2 \end{pmatrix}$$

e $\det(L') = 55$. Il numero 55 rappresenta appunto il numero di alberi di supporto del grafo (con un po' di pazienza si potrebbero elencare tutti quanti).

Per quel che riguarda i grafi prodotto si può notare che la matrice diagonale dei gradi del grafo prodotto cartesiano ha gli elementi diagonali uguali a $g_1^1 + g_1^2, g_1^1 + g_2^2$, ecc., dove g^1 è il vettore dei gradi di G_1 e dove g^2 è il vettore dei gradi di G_2 . La matrice diagonale dei gradi può essere scritta, sfruttando il prodotto di Kronecker

$$\operatorname{diag}(g^1) \otimes I + I \otimes \operatorname{diag}(g^2)$$

e allora il Laplaciano è

$$\operatorname{diag}(g^1) \otimes I + I \otimes \operatorname{diag}(g^2) - (I \otimes A_2 + A_1 \otimes I)$$

Il Laplaciano del grafo prodotto tensoriale è

$$\operatorname{diag}(g_1) \otimes \operatorname{diag}(g_2) - A_1 \otimes A_2$$

Un'altra matrice che si trova descritta in letteratura, anche se di minore importanza rispetto al Laplaciano, è il *Laplaciano non segnato* Q (*signless Laplacian*), definito come

$$Q := \text{diag}(g) + A$$

Questa matrice è anche uguale al prodotto $B B^\top$, dove B è la matrice d'incidenza nodi-archi del grafo (questa volta però del grafo stesso non orientato, senza cioè assegnare un'orientazione arbitraria agli archi come nel caso del Laplaciano). Il Laplaciano non segnato non è necessariamente singolare. In base alle osservazioni alla fine della Sez. 14, il Laplaciano non segnato è singolare se e solo se il grafo ha una componente bipartita.

Viene anche definito in letteratura il *Laplaciano normalizzato* \mathcal{L} definito come

$$\mathcal{L} := \text{diag}(g)^{-1/2} L \text{diag}(g)^{-1/2} \quad \text{dove} \quad \text{diag}(g)^{-1/2} = \begin{pmatrix} \frac{1}{\sqrt{g_1}} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{g_2}} & 0 & \dots & 0 \\ 0 & 0 & \frac{1}{\sqrt{g_3}} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \frac{1}{\sqrt{g_n}} \end{pmatrix}$$

e se $g_i = 0$ si pone 0 per convenzione. Quindi nel Laplaciano normalizzato si ha

$$\mathcal{L}_{ij} = \begin{cases} 1 & \text{se } i = j \\ -1/\sqrt{g_i g_j} & \text{se } (i, j) \in E \\ 0 & \text{se } (i, j) \notin E \end{cases}$$

Citiamo ancora la *matrice di Seidel* S definita come

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

o anche come $S = \mathbf{1}\mathbf{1}^\top - I - 2A$.

20 Spettro di un grafo

Si definisce come *spettro (ordinario) di un grafo* lo spettro $\{\lambda_1, \dots, \lambda_n\}$ della matrice d'adiacenza. Si definisce come *spettro Laplaciano di un grafo* lo spettro $\{\lambda_1^L, \dots, \lambda_n^L\}$ della matrice Laplaciana. In generale non c'è relazione fra lo spettro ordinario e quello Laplaciano tranne che nei grafi regolari, come si vedrà fra poco. Anche le altre matrici definite nella sezione precedente hanno uno spettro che caratterizza ulteriormente il grafo. Questi spettri sono però meno importanti e non verranno descritti in questa sede.

Se un grafo non è connesso, lo spettro (sia ordinario che Laplaciano) è l'unione (mantenendo eventuali molteplicità degli autovalori) degli spettri delle singole componenti connesse.

Siccome gli autovalori non dipendono da come righe e colonne sono state permutate (righe e colonne però devono essere permutate allo stesso modo), lo spettro si riferisce a proprietà di un grafo indipendentemente da come sono stati etichettati i nodi. Pertanto due grafi isomorfi hanno lo stesso spettro. Purtroppo non è

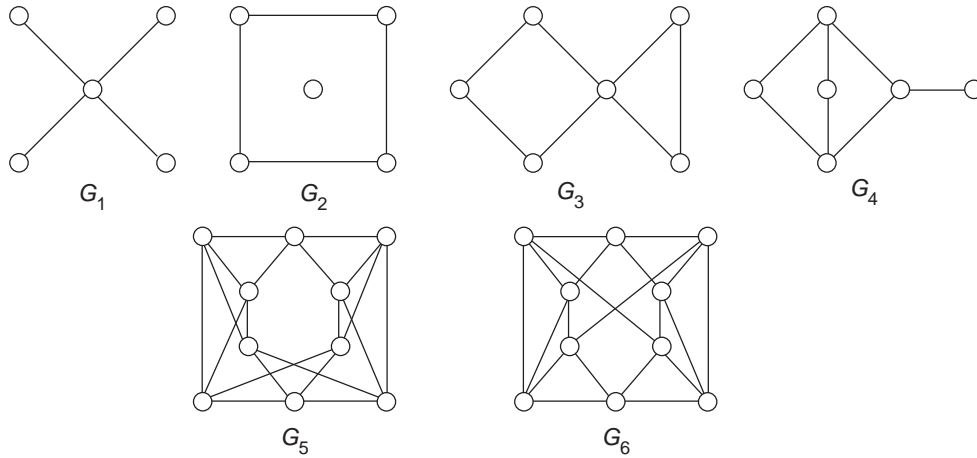


Figura 22.

vero il viceversa. Ci possono essere grafi non isomorfi con lo stesso spettro. Altrimenti sarebbe stato possibile decidere dell'isomorfismo di due grafi verificando l'uguaglianza degli spettri. Si vedano ad esempio in Fig. 22 i grafi G_1 e G_2 con lo stesso spettro ordinario ma differenti spettri Laplaciani, i grafi G_3 e G_4 con lo stesso spettro Laplaciano ma differenti spettri ordinari e i grafi regolari non isomorfi G_5 e G_6 con lo stesso spettro sia ordinario che Laplaciano. Quindi se dall'esame degli spettri risulta che gli spettri sono diversi, allora i grafi non sono isomorfi, ma se gli spettri sono uguali non si può decidere. Gli spettri dei sei grafi sono:

$$\begin{aligned} \lambda(G_1) &= \lambda(G_2) = (-2 \ 0 \ 0 \ 0 \ 2) \\ \lambda^L(G_1) &= (0 \ 1 \ 1 \ 1 \ 5) \\ \lambda^L(G_2) &= (0 \ 0 \ 2 \ 2 \ 4) \\ \lambda(G_3) &= (-2.1912 \ -1 \ -0.5767 \ 0 \ 1.2643 \ 2.5035) \\ \lambda(G_4) &= (-2.5243 \ -0.7923 \ 0 \ 0 \ 0.7923 \ 2.5243) \\ \lambda^L(G_3) &= \lambda^L(G_4) = (0 \ 3 - \sqrt{5} \ 2 \ 3 \ 3 \ 3 + \sqrt{5}) \\ \lambda(G_5) &= \lambda(G_6) = (-\frac{1}{2}(1 + \sqrt{17}) \ -\sqrt{5} \ -1 \ -1 \ -1 \ -1 \ 1 \ \sqrt{5} \ \frac{1}{2}(-1 + \sqrt{17}) \ \sqrt{5} \ 4) \\ \lambda^L(G_5) &= \lambda^L(G_6) = (0 \ 4 - \sqrt{5} \ \frac{1}{2}(9 - \sqrt{17}) \ 3 \ 5 \ 5 \ 5 \ 5 \ 4 + \sqrt{5} \ \frac{1}{2}(9 + \sqrt{17})) \end{aligned}$$

Siccome la matrice d'adiacenza è simmetrica ha n autovalori reali con n autovettori ortogonali. Dato che $\text{Tr}(A) = 0$ e $\text{Tr}(A) = \sum_i \lambda_i$, devono essere presenti sia autovalori positivi che negativi. In base alla teoria di Perron-Frobenius si può dimostrare che in un grafo connesso esiste un unico autovalore positivo di valore assoluto non inferiore al valore assoluto degli altri autovalori ed a tale autovalore è associato un autovettore le cui componenti hanno tutte lo stesso segno e ad esempio possono essere prese come positive. Siccome gli autovettori sono ortogonali, solo un autovettore può avere tutte le componenti non negative. Se il grafo non

è connesso si perde la garanzia dell'unicità dell'autovalore massimo e delle componenti strettamente positive dell'autovettore associato (fatto abbastanza evidente dato che in un grafo sconnesso lo spettro è l'unione delle singole componenti).

Da

$$A u = \lambda u \implies \sum_{j \in \delta(i)} u_j = \lambda u_i$$

si vede che un autovettore ordinario può essere interpretato come un peso u_i (in generale anche negativo) da assegnare ad ogni nodo in modo che la somma dei pesi sui nodi adiacenti sia uguale ad un multiplo (lo stesso per tutti i nodi) del peso del nodo.

Se vogliamo considerare pesi non negativi e pensiamo di attribuire un peso ad un nodo in base ai pesi dei nodi adiacenti e consideriamo tale peso come una misura di centralità, dobbiamo prendere in esame l'autovettore associato all'autovalore massimo. Ad esempio per il grafo di Fig. 12 si ottiene il seguente autovettore in corrispondenza dell'autovalore massimo $\lambda_1 = 3.02142$:

$$u^1 = \{0.501, 0.380, 0.197, 0.215, 0.452, 0.269, 0.361, 0.320, 0.106\}$$

Si confrontino questi valori con quelli di eccentricità e di stato elencati a pag. 18. In questo caso valori elevati corrispondono ad una maggior centralità del nodo.

Il raggio spettrale della matrice d'adiacenza (si veda a pag. 31) che possiamo chiamare anche raggio spettrale del grafo non può essere più grande del massimo grado nel grafo $\bar{g} := \max_i g_i$. Per dimostrarlo sia u un autovettore e λ l'autovalore corrispondente. Sia k tale che $|u_k| \geq |u_i|$, $i = 1, \dots, n$. Possiamo sempre supporre $u_k = 1$. Allora abbiamo

$$\lambda = \lambda u_k = \sum_{j \in \delta(k)} u_j \leq \sum_{j \in \delta(k)} |u_j| \leq \bar{g}$$

e anche

$$\lambda = \lambda u_k = \sum_{j \in \delta(k)} u_j \geq \sum_{j \in \delta(k)} -|u_j| \geq -\bar{g}$$

da cui $|\lambda| \leq \bar{g}$ per ogni autovalore λ . Dalla dimostrazione si vede anche che se esiste un autovalore $\lambda = \bar{g}$, allora deve essere $u_j = u_k = 1$ per ogni $j \in \delta(k)$ e che $g_k = \bar{g}$. A questo punto si possono ripetere gli argomenti per ogni $j \in \delta(k)$ e pervenire alla conclusione che $u_j = 1$ e $g_j = \bar{g}$ su una componente connessa del grafo. In conclusione se il raggio spettrale è uguale al massimo grado il grafo ha una componente connessa regolare con grado uguale al massimo grado.

Siccome $\text{Tr}(A^k)$ conta i circuiti di lunghezza k e siccome $\text{Tr}(A^k) = \sum_i \lambda_i^k$, lo spettro ordinario fornisce informazione sul numero di circuiti. Gli autovalori del grafo di Fig. 21 sono

$$(3.18194 \quad -1.80194 \quad -1.58836 \quad 1.24698 \quad -0.593579 \quad -0.445042)$$

con rispettivi autovettori

$$u^1 = \begin{pmatrix} 0.286 \\ 0.503 \\ 0.406 \\ 0.406 \\ 0.503 \\ 0.286 \end{pmatrix}, u^2 = \begin{pmatrix} 0.418 \\ -0.521 \\ 0.232 \\ -0.232 \\ 0.521 \\ -0.418 \end{pmatrix}, u^3 = \begin{pmatrix} -0.108 \\ -0.401 \\ 0.573 \\ 0.573 \\ -0.401 \\ -0.108 \end{pmatrix}, u^4 = \begin{pmatrix} 0.521 \\ 0.232 \\ -0.418 \\ 0.418 \\ -0.232 \\ -0.521 \end{pmatrix}, u^5 = \begin{pmatrix} 0.638 \\ -0.294 \\ -0.085 \\ -0.085 \\ -0.294 \\ 0.638 \end{pmatrix}, u^6 = \begin{pmatrix} 0.232 \\ 0.418 \\ 0.521 \\ -0.521 \\ -0.418 \\ -0.232 \end{pmatrix}$$

e i loro cubi sono

$$(32.2163 \quad -5.85088 \quad -4.00725 \quad 1.939 \quad -0.209139 \quad -0.0881461)$$

la cui somma fa 24, e questo numero, diviso per 6, fornisce i 4 triangoli.

Il Laplaciano, oltre ad essere simmetrico è anche positivo semidefinito, e quindi ha n autovalori reali non negativi con n autovettori ortogonali. Siccome L è singolare, almeno un autovalore deve essere nullo. L'autovettore corrispondente è $\mathbf{1}$. La molteplicità dell'autovalore nullo è uguale al numero di componenti connesse del grafo. Per dimostrare questo fatto si noti che L è diagonale a blocchi se il grafo è sconnesso e per ogni blocco deve essere presente un autovalore nullo da cui si ha che la molteplicità dell'autovettore nullo è almeno uguale al numero di componenti connesse. Resta da dimostrare che per ogni componente connessa c'è solo un autovettore nullo. Infatti

$$Lu = \mathbf{0} \implies BB^T u = \mathbf{0} \implies u^T B B^T u = 0 \implies u^T B = \mathbf{0}$$

cioè $u_i = u_j$ per ogni $(i, j) \in E$, quanto a dire che si hanno vettori costanti (ad esempio uguale ad uno) su una componente connessa e zero sulle altre componenti. Quindi il numero di autovettori associati ad autovalori nulli è esattamente uguale al numero di componenti connesse. Se ordiniamo gli autovalori in ordine crescente $0 = \lambda_1^L \leq \lambda_2^L \leq \dots \leq \lambda_n^L$, abbiamo allora che un grafo è connesso se e solo se $\lambda_2^L > 0$. Quindi lo spettro Laplaciano fornisce informazione sul numero di componenti connesse di un grafo, mentre lo spettro ordinario non dà questa informazione. Si vedano ad esempio i grafi G_1 e G_2 in Fig. 22.

Il fatto di avere un valore grande per λ_2^L rende sempre più forte la connettività di un grafo. Per questo motivo il secondo autovalore del Laplaciano viene anche chiamato *connettività algebrica*. Si può dimostrare che vale la seguente diseuguaglianza fra la connettività algebrica e quella di vertice (pag. 19)

$$\kappa(G) \geq \lambda_2^L(G)$$

Si noti però che la diseuguaglianza si basa sulla prima definizione di connettività di vertice e quindi non si applica ad un grafo completo (per il quale, come si vedrà in seguito, $\lambda_2^L = n$ e $\kappa(K_n) = n - 1$, dalla seconda definizione).

Ad esempio per il grafo di Fig. 19 si ha $\lambda_2^L = 1.18639$ e $\kappa = 2$. Siccome per ogni grafo $\kappa(G) \leq \min_i g_i(G)$ si ha anche

$$\min_i g_i(G) \geq \kappa(G) \geq \lambda_2^L(G)$$

Questo implica che per ogni albero vale $\lambda_2^L \leq 1$.

Si era fatto notare (pag. 43) come il numero di alberi di supporto di un grafo sia ottenibile dal Laplaciano calcolando il determinante del Laplaciano a cui sia stata tolta una qualsiasi colonna e la corrispondente riga. Lo stesso numero è anche ottenibile come

$$\frac{1}{n} \prod_{i=2}^n \lambda_i^L \quad (20)$$

Per dimostrare questo fatto si noti che il polinomio caratteristico del Laplaciano si può scrivere come

$$p(\lambda) = \det(\lambda I - L) = \lambda (\lambda - \lambda_2^L) (\lambda - \lambda_3^L) \dots (\lambda - \lambda_n^L)$$

e quindi si ha

$$\left. \frac{dp(\lambda)}{d\lambda} \right|_{\lambda=0} = (-1)^{n-2} \prod_{i=2}^n \lambda_i^L$$

Il calcolo della derivata si può anche fare direttamente da $\det(\lambda I - A)$ con il metodo di considerare, di volta in volta, variabile solo uno degli n termini λ e considerare costanti tutti gli altri e poi sommare le n derivate così ottenute. Questo porta a sommare i determinanti delle n matrici che si ottengono togliendo di volta in volta una riga e la corrispondente colonna. Così facendo si ottiene n volte il numero di alberi di supporto, da cui la formula sopra scritta.

21 Spettro di un grafo regolare

Se un grafo è k -regolare

$$A \mathbf{1} = k \mathbf{1}$$

Quindi esiste un autovalore pari al grado dei nodi. In base al precedente risultato un grafo connesso è k -regolare se e solo se il più grande autovalore vale k . Un grafo non connesso potrebbe avere una componente k -regolare, il che implica un autovalore del grafo uguale a k , ma le altre componenti potrebbero non essere regolari oppure regolari con un valore diverso del grado.

Quando un grafo è regolare lo spettro ordinario e quello Laplaciano sono legati da una semplice relazione. Infatti in questo caso si ha (supponendo il grafo k -regolare) $L = kI - A$ e quindi, se λ è un autovalore di A con autovettore u , si ha

$$L u = (kI - A) u = k u - \lambda u = (k - \lambda) u$$

cioè u è anche autovettore di L con autovalore $k - \lambda$, cioè

$$\lambda_i^L + \lambda_i = k, \quad i = 1 \dots, n \quad (21)$$

Si vedano ad esempio i grafi regolari G_5 e G_6 in Fig. 22. Come si vede, l'autovalore ordinario k è associato all'autovalore Laplaciano nullo. Quindi per un grafo k -regolare la molteplicità dell'autovalore k è uguale al numero di componenti connesse del grafo. Si noti che, mentre la proprietà della molteplicità dell'autovalore 0

della matrice Laplaciana vale per un grafo qualsiasi, la stessa proprietà riferita alla molteplicità dell'autovalore più grande della matrice d'adiacenza, vale solo per grafi regolari.

Si noti ancora che la proprietà (21) vale anche se il grafo non è k -regolare ma ha diverse componenti connesse ognuna delle quali regolare con gradi eventualmente diversi fra diverse componenti. In questo caso la relazione (21) vale componente per componente con i rispettivi valori di grado.

Se il grafo è fortemente regolare ed è connesso gli autovalori (sia ordinari che Laplaciani) possono assumere al più tre valori diversi. Infatti si riconsideri la relazione (12) e sia u un autovettore ordinario, diverso dall'autovettore $\mathbf{1}$, con autovalore λ . Allora si ha

$$A^2 u + (s - r) A u + (s - k) I u = s \mathbf{1} \mathbf{1}^\top u$$

che diventa

$$\lambda^2 u + (s - r) \lambda u + (s - k) u = \mathbf{0}$$

in quanto u è ortogonale a $\mathbf{1}$ (il grafo è connesso). Essendo $u \neq \mathbf{0}$ deve essere

$$\lambda^2 + (s - r) \lambda + (s - k) = 0$$

che ammette solo le due radici

$$\frac{(r - s) \pm \sqrt{(r - s)^2 + 4(k - s)}}{2} \quad (22)$$

Quindi gli unici autovalori possibili sono k e queste due radici. La molteplicità di queste radici si calcola facilmente ricordando che l'autovalore k ha molteplicità 1 e che la somma di tutti gli autovalori deve essere nulla.

Sia m_1 la molteplicità di una delle due radici e m_2 la molteplicità dell'altra. Quindi si deve avere

$$k + m_1 \frac{(r - s) + \sqrt{(r - s)^2 + 4(k - s)}}{2} + m_2 \frac{(r - s) - \sqrt{(r - s)^2 + 4(k - s)}}{2} = 0$$

che diventa (si ricordi che deve anche essere $m_1 + m_2 = n - 1$)

$$2k + (n - 1)(r - s) = (m_2 - m_1) \sqrt{(r - s)^2 + 4(k - s)}.$$

Questa relazione implica che l'espressione sotto radice deve essere un quadrato perfetto a meno che $m_1 = m_2$. Quindi i grafi fortemente regolari con un numero pari di nodi (per cui quindi $m_1 \neq m_2$) e quelli con un numero dispari di nodi ma con $m_1 \neq m_2$, devono avere valori di k , r e s tali che $(r - s)^2 + 4(k - s)$ sia un quadrato perfetto e quindi hanno autovalori interi. Al sito <http://www.win.tue.nl/~aeb/graphs/srg/srgtab.html> si possono trovare i parametri di moltissimi grafi fortemente regolari.

Per esercizio si calcolino gli spettri ordinario e Laplaciano dei grafi di Fig. 4 con le molteplicità degli autovalori e, tramite la formula (20) si calcoli il numero degli alberi di supporto dei vari grafi.

22 Spettro di un grafo completo

Gli autovalori del grafo completo K_n (che è anche $(n-1)$ -regolare) sono

$$\lambda_1 = n-1, \quad \lambda_i = -1, \quad i = 2, \dots, n, \quad (23)$$

e i corrispondenti autovettori sono

$$u^1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \dots \\ 1 \\ 1 \end{pmatrix}, \quad u^2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \dots \\ 1 \\ -n+1 \end{pmatrix}, \quad u^3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \dots \\ -n+2 \\ 0 \end{pmatrix}, \quad u^{n-1} = \begin{pmatrix} 1 \\ 1 \\ -2 \\ \dots \\ 0 \\ 0 \end{pmatrix}, \quad u^n = \begin{pmatrix} 1 \\ -1 \\ 0 \\ \dots \\ 0 \\ 0 \end{pmatrix}$$

Il calcolo si può effettuare tenendo conto che in questo caso $A = \mathbf{1}\mathbf{1}^\top - I$ e quindi

$$Au = \lambda u \implies (\mathbf{1}\mathbf{1}^\top - I)u = \lambda u \implies \mathbf{1}\mathbf{1}^\top u = (1 + \lambda)u$$

Allora gli autovettori di A sono gli stessi della matrice diadica $\mathbf{1}\mathbf{1}^\top$ e gli autovalori delle due matrici differiscono di 1. Da (9) gli autovalori di $\mathbf{1}\mathbf{1}^\top$ sono n e 0 (con molteplicità $n-1$) e quindi lo spettro è come in (23). Da notare che gli autovettori corrispondenti all'autovalore -1 non sono univocamente determinati, in quanto basta prendere un qualsiasi insieme di vettori linearmente indipendenti nel sottospazio ortogonale all'autovettore $(1, 1, \dots, 1)$. Gli autovalori del Laplaciano per il grafo completo K_n sono (da (23) e da (21))

$$\lambda_1^L = 0, \quad \lambda_i^L = n, \quad i = 2, \dots, n$$

Il numero di triangoli in un grafo completo è naturalmente $\binom{n}{3}$. Si noti che

$$\sum_{i=1}^n \lambda_i^3 = (n-1)^3 - (n-1) = n(n-1)(n-2) = 6 \binom{n}{3}$$

23 Spettro di un grafo bipartito

Gli autovalori di un grafo bipartito $G = (V_1, V_2; E)$ sono sempre presenti a coppie λ_i e $-\lambda_i$. Se il numero di nodi è dispari almeno un autovalore deve essere nullo. Si vedano ad esempio i grafi bipartiti G_1 , G_2 e G_4 in Fig. 22. Infatti la matrice d'adiacenza di un grafo bipartito può essere scritta come

$$A = \begin{pmatrix} \mathbf{0}_1 & D^\top \\ D & \mathbf{0}_2 \end{pmatrix} \quad (24)$$

dove le matrici $\mathbf{0}_1$ e $\mathbf{0}_2$ sono quadrate con componenti tutte uguali a zero e di dimensione n_1 e n_2 rispettivamente. Allora si deve avere

$$\begin{pmatrix} \mathbf{0}_1 & D^\top \\ D & \mathbf{0}_2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \lambda \begin{pmatrix} u \\ v \end{pmatrix} \implies \begin{matrix} D^\top v = \lambda u \\ Du = \lambda v \end{matrix} \implies \begin{matrix} D^\top(-v) = (-\lambda)u \\ Du = (-\lambda)(-v) \end{matrix}$$

e quindi anche $-\lambda$ è autovalore con autovettore $(u, -v)$. Inoltre si ricava

$$\begin{aligned} Du = \lambda v &\implies D^\top Du = \lambda D^\top v = \lambda^2 u \\ D^\top v = \lambda u &\implies DD^\top v = \lambda Du = \lambda^2 v \end{aligned} \quad (25)$$

cioè u è autovettore di $D^\top D$ con autovalore λ^2 e v è autovettore di DD^\top con lo stesso autovalore λ^2 . La matrice $D^\top D$ ha dimensione $n_1 \times n_1$ e la matrice DD^\top ha dimensione $n_2 \times n_2$. Se $n_1 \neq n_2$ hanno un numero diverso di autovalori. Se, ad esempio $n_1 > n_2$, gli $(n_1 - n_2)$ autovalori in più di $D^\top D$ sono tutti nulli e corrispondono ad autovettori u tali che $Du = \mathbf{0}$. Per questi autovalori nulli l'autovettore della matrice d'adiacenza è $(u, \mathbf{0})$ (cioè con $v = \mathbf{0}$).

L'elemento (i, j) (con $i, j \in V_1$) della matrice $D^\top D$ conta quanti vertici in V_2 sono adiacenti sia a i che a k e quindi l'elemento diagonale (i, i) è semplicemente il grado di i . Analogamente l'elemento (i, j) (con $i, j \in V_2$) della matrice DD^\top conta quanti vertici in V_1 sono adiacenti sia a i che a k e l'elemento diagonale (j, j) è il grado di j . Le matrici $D^\top D$ e DD^\top rivestono un ruolo importante, come si vedrà più avanti.

Si noti che per ottenere l'autovettore (u, v) dagli autovettori u e v rispettivamente di $D^\top D$ e DD^\top bisogna fare attenzione a normalizzare in modo opportuno u rispetto a v , cioè si deve avere $Du = \lambda v$ e $D^\top v = \lambda u$.

È vero anche il viceversa, cioè che se lo spettro di un grafo ha autovalori a coppie λ e $-\lambda$, più eventualmente qualche autovalore nullo, il grafo deve essere bipartito. Per dimostrarlo si sfrutta la proprietà che gli elementi diagonali di A^k contano il numero di circuiti di lunghezza k . Allora per k dispari

$$\text{Tr}(A^k) = \sum_i (A^k)_{ii} = \sum_i \lambda_i^k = 0 \implies (A^k)_{ii} = 0$$

in quanto ogni $(A^k)_{ii}$ è non negativo e se la somma è nulla devono essere tutti nulli. Se ogni $(A^k)_{ii}$ è nullo non esistono circuiti dispari e il grafo deve essere bipartito.

Si può affermare una proprietà ancora più forte: se il grafo è connesso e ci sono due autovalori opposti uguali al raggio spettrale, allora il grafo è bipartito (e di conseguenza tutti gli autovalori sono presenti a coppie opposte). Per dimostrare questo fatto sia $\hat{\lambda} > 0$ l'autovalore unico uguale al raggio spettrale e sia $u > \mathbf{0}$ l'autovettore associato (l'esistenza di un tale autovettore e l'unicità del autovalore $\hat{\lambda}$ discendono dal Teorema di Perron-Frobenius applicato alla matrice d'adiacenza di un grafo connesso). Sia $-\hat{\lambda}$ l'altro autovalore con autovettore v . Allora si ha

$$\hat{\lambda} |v_k| = |-\hat{\lambda} v_k| = \left| \sum_{j \in \delta(k)} v_j \right| \leq \sum_{j \in \delta(k)} |v_j| \quad k = 1, \dots, n \quad (26)$$

dove la seconda uguaglianza è dovuta al fatto che v è autovettore con autovalore $-\hat{\lambda}$. Moltiplicando per u_k il primo e l'ultimo termine di (26), e sommando su tutti i nodi k , si ha

$$\hat{\lambda} \sum_k u_k |v_k| \leq \sum_k \sum_{j \in \delta(k)} u_k |v_j| = \sum_j \sum_{k \in \delta(j)} u_k |v_j| = \sum_j |v_j| \sum_{k \in \delta(j)} u_k = \hat{\lambda} \sum_j |v_j| u_j$$

dove la prima uguaglianza deriva dal fatto che entrambe le sommatorie sommano (due volte) rispetto a tutti gli archi e l'ultima deriva dal fatto che u è autovettore con autovalore $\hat{\lambda}$. Siccome il primo e l'ultimo termine sono uguali anche i primi due termini devono essere uguali, cioè

$$\sum_k u_k \left(\sum_{j \in \delta(k)} |v_j| - \hat{\lambda} |v_k| \right) = 0$$

Siccome $u_k > 0$ e $\sum_{j \in \delta(k)} |v_j| - \hat{\lambda} |v_k| \geq 0$ in base a (26), si deve avere in realtà

$$\sum_{j \in \delta(k)} |v_j| - \hat{\lambda} |v_k| = 0 \quad \implies \quad \hat{\lambda} |v_k| = \sum_{j \in \delta(k)} |v_j| \quad k = 1, \dots, n \quad (27)$$

Quindi $|v|$ è autovettore con autovalore $\hat{\lambda}$. Data l'unicità dell'autovalore si deve concludere che $u = |v|$. In altri termini l'autovettore corrispondente all'autovalore $-\hat{\lambda}$ ha le stesse componenti dell'autovettore corrispondente all'autovalore $\hat{\lambda}$, ma alcune di segno opposto. Ora si confrontino (26) e (27), da cui si vede che

$$\left| \sum_{j \in \delta(k)} v_j \right| = \sum_{j \in \delta(k)} |v_j| \quad k = 1, \dots, n$$

Questa eguaglianza implica che i valori v_j con $j \in \delta(k)$ sono tutti dello stesso segno e siccome $-\hat{\lambda} v_k = \sum_{j \in \delta(k)} v_j$, il segno di v_k deve essere opposto a quello di v_j sui nodi adiacenti a v_j . Questo implica che il segno di v_k induce una bipartizione del grafo.

Si è quindi visto che lo spettro ordinario dà informazione sul fatto che un grafo sia bipartito, mentre lo spettro Laplaciano non dà in generale questa informazione. Si confrontino i grafi G_3 , non bipartito, e G_4 , bipartito, di Fig. 22, che hanno lo stesso Laplaciano. Se il grafo è k -regolare ovviamente anche lo spettro Laplaciano può dire se il grafo è bipartito. Dato che la coppia di autovalori $-k$ e k è sufficiente a garantire che un grafo connesso sia bipartito, allora la sola presenza dell'autovalore $\lambda^L = 2k$ è sufficiente a garantire che un grafo connesso k -regolare sia bipartito.

Lo spettro ordinario e quello Laplaciano sono comunque legati da una interessante relazione se il grafo bipartito, pur non essendo regolare, ha i nodi di ognuno dei due insiemi con lo stesso grado, k_1 per un insieme e k_2 per l'altro insieme. Si noti che deve essere $k_1 n_1 = k_2 n_2$ e in particolare un grafo bipartito completo è di questo tipo con $k_1 = n_2$ e $k_2 = n_1$. Allora si ha

$$L = \begin{pmatrix} k_1 I & -D^\top \\ -D & k_2 I \end{pmatrix}, \quad L \begin{pmatrix} u \\ v \end{pmatrix} = \lambda^L \begin{pmatrix} u \\ v \end{pmatrix} \quad \implies \quad \begin{aligned} k_1 u - D^\top v &= \lambda^L u \\ k_2 v - D u &= \lambda^L v \end{aligned} \quad \implies$$

$$\begin{aligned} D^\top v &= (k_1 - \lambda^L) u \\ D u &= (k_2 - \lambda^L) v \end{aligned} \quad (28)$$

$$\begin{aligned} D^\top D u &= (k_2 - \lambda^L) D^\top v = (k_1 - \lambda^L) (k_2 - \lambda^L) u \\ D D^\top v &= (k_1 - \lambda^L) D u = (k_1 - \lambda^L) (k_2 - \lambda^L) v \end{aligned} \quad (29)$$

Confrontando (25) con (29) deve essere

$$\lambda^2 = (k_1 - \lambda^L)(k_2 - \lambda^L) \quad (30)$$

Quindi ad ogni coppia di autovalori ordinari di segno opposto, corrispondono i due autovalori Laplaciani soluzioni di (30) (uno minore di $\min\{k_1, k_2\}$ e l'altro maggiore di $\max\{k_1, k_2\}$, come si vede immediatamente dal grafico della parabola $(k_1 - \lambda^L)(k_2 - \lambda^L)$). In particolare agli autovalori ordinari $\lambda = \pm\sqrt{k_1 k_2}$ sono associati gli autovalori Laplaciani $\lambda^L = 0$ e $\lambda^L = k_1 + k_2$.

Ad autovalori ordinari nulli corrispondono autovalori Laplaciani uguali a k_1 oppure a k_2 . Per capire esattamente a quale dei due autovalori Laplaciani corrisponde un autovalore ordinario nullo si consideri (28) e si supponga $\lambda^L = k_1$. Questo implica $D^\top v = \mathbf{0}$, cioè $v \in \mathcal{N}(D^\top)$ (si rivedano le definizioni a pag. 25). Dalla seconda relazione di (28) abbiamo anche $v \in \mathcal{R}(D)$ (stiamo necessariamente supponendo $k_1 \neq k_2$ e quindi $k_2 - \lambda^L \neq 0$) e siccome $\mathcal{N}(D^\top) \perp \mathcal{R}(D)$ si ha $v = \mathbf{0}$. Allora $u \in \mathcal{N}(D)$. A questo punto si vede che il numero di autovettori distinti corrispondenti all'autovalore $\lambda^L = k_1$ è uguale alla dimensione di $\mathcal{N}(D)$ e cioè a $n_1 - r$ dove r è il rango sia di D che di D^\top . Con ragionamento analogo si trova che la molteplicità di $\lambda^L = k_2$ è $n_2 - r$.

Se il grafo bipartito è completo il calcolo di tutti gli autovalori è semplice. Siano n_1 e n_2 i nodi dei due insiemi. L'autovalore massimo deve essere $\sqrt{n_1 n_2}$ in base ai risultati precedenti. Per il calcolo degli altri autovalori si noti che se il grafo è completo la matrice D in (24) può essere scritta come

$$D = \mathbf{1}_2 \mathbf{1}_1^\top$$

dove i vettori $\mathbf{1}_1$ e $\mathbf{1}_2$ sono vettori tutti di uni di dimensione n_1 e n_2 rispettivamente. In questo caso il rango di D è 1. Allora (24) è

$$\begin{pmatrix} \mathbf{0}_1 & \mathbf{1}_1 \mathbf{1}_2^\top \\ \mathbf{1}_2 \mathbf{1}_1^\top & \mathbf{0}_2 \end{pmatrix}$$

Questo implica che il rango della matrice d'adiacenza sia uguale a 2 (ci sono $n_1 - 1$ vettori ortogonali a $\mathbf{1}_1$ e $n_2 - 1$ vettori ortogonali a $\mathbf{1}_2$) il che implica che $n_1 + n_2 - 2$ autovalori sono nulli. Per i due autovalori non nulli si ha

$$\begin{aligned} \mathbf{1}_1 \mathbf{1}_2^\top v = \lambda u & \implies u = \frac{\mathbf{1}_2^\top v}{\lambda} \mathbf{1}_1 & \implies \mathbf{1}_1^\top u = \frac{\mathbf{1}_2^\top v}{\lambda} n_1 & \implies \lambda = \pm \sqrt{n_1 n_2} \\ \mathbf{1}_2 \mathbf{1}_1^\top u = \lambda v & \implies v = \frac{\mathbf{1}_1^\top u}{\lambda} \mathbf{1}_2 & \implies \mathbf{1}_2^\top v = \frac{\mathbf{1}_1^\top u}{\lambda} n_2 & \implies \lambda = \pm \sqrt{n_1 n_2} \end{aligned}$$

Dai risultati precedenti abbiamo che gli autovalori Laplaciani sono $\lambda^L = 0$, $\lambda^L = n_1 + n_2$, $\lambda^L = n_1$ con molteplicità $n_2 - 1$ e $\lambda^L = n_2$ con molteplicità $n_1 - 1$.

Se $n_1 = n_2$ il grafo bipartito completo è anche fortemente regolare. Per esercizio si ritrovino gli autovalori sfruttando la relazione (22).

24 Spettro di un circuito

Gli autovalori di un circuito di n nodi (che è un grafo 2-regolare ed inoltre bipartito se n è pari) derivano dalle radici di $\lambda^n = 1$. Infatti, supponendo che i nodi del circuito siano numerati come $1, 2, \dots, n$, dobbiamo avere

$$u_{j-1} + u_{j+1} = \lambda u_j \quad (31)$$

dove intendiamo che $j-1 = n$ se $j = 1$ e $j+1 = 1$ se $j = n$. Essendo il grafo 2-regolare, $\lambda = 2$ è il massimo autovalore con autovettore $\mathbf{1}$. Per il calcolo degli altri autovalori si sfrutta la simmetria dei nodi e si può vedere se ogni componente u_j è uguale ad un multiplo α della precedente, cioè

$$u_j = \alpha u_{j-1} \quad (32)$$

Questa posizione porta a

$$u_{j-1} + u_{j+1} = \lambda u_j \implies u_{j-1} + \alpha^2 u_{j-1} = \lambda \alpha u_{j-1} \implies \lambda = \alpha^{-1} + \alpha$$

Bisogna anche pensare che applicando ciclicamente $u_j = \alpha u_{j-1}$ si deve avere $u_1 = \alpha^n u_1$ e quindi α deve essere una radice dell'unità, ovvero

$$\alpha_k = e^{i2\pi k j/n} \quad k \in \{1, \dots, n\}$$

(attenzione! i è l'unità immaginaria, non l'indice di un nodo). Allora

$$\lambda_k = e^{-i2\pi k/n} + e^{i2\pi k/n} = 2 \cos \frac{2\pi k}{n} \quad k \in \{1, \dots, n\}$$

Si noti che $\lambda_n = 2$ (è il valore trovato precedentemente) e $\lambda_k = \lambda_{n-k}$. Ponendo $u_1^k = 1$ (una componente di un autovettore è arbitraria) e applicando (32) si ha

$$u_{j+1}^k = \alpha_k^j u_1^k = e^{i2\pi k j/n}$$

ed anche (sempre ponendo $u_1^{n-k} = 1$)

$$u_{j+1}^{n-k} = \alpha_{n-k}^j u_1^{n-k} = e^{i2\pi(n-k)j/n} = e^{i2\pi j} e^{-i2\pi k j/n} = e^{-i2\pi k j/n}$$

Quindi l'autovettore u^{n-k} è coniugato di u^k e siccome si riferiscono allo stesso autovalore, ogni loro combinazione lineare è un autovettore. Quindi possiamo scegliere delle combinazioni lineari che diano luogo ad autovettori v^k reali, ad esempio:

$$v_j^k = u_j^k + u_j^{n-k} = 2 \cos \frac{2\pi k j}{n}, \quad v_j^{n-k} = i u_j^{n-k} - i u_j^k = 2 \sin \frac{2\pi k j}{n}$$

Gi autovalori del Laplaciano per un circuito di n nodi sono (da (23))

$$\lambda_k^L = 2 \left(1 - \cos \frac{2\pi k}{n}\right) \quad k = 0, \dots, n-1$$

25 Spettro di un cammino

Lo spettro di un cammino (che è un grafo bipartito) si può dedurre da quello di un circuito e si trovano gli autovalori

$$\lambda_k = 2 \cos \frac{k \pi}{n+1}, \quad k \in \{1, \dots, n\} \quad (33)$$

con autovettori corrispondenti

$$u_j^k = \sin \frac{k j \pi}{n+1}, \quad j \in \{1, \dots, n\}, \quad k \in \{1, \dots, n\} \quad (34)$$

Si può effettuare la verifica sulla relazione che lega autovalori e autovettori che per un cammino di n nodi è

$$u_2^k = \lambda_k u_1^k, \quad u_{j-1}^k + u_{j+1}^k = \lambda_k u_j^k, \quad u_{n-1}^k = \lambda_k u_n^k \quad (35)$$

Si noti che si può estendere la formula (34) agli indici $j = 0$ e $j = n+1$ per i quali si ha $u_0^k = u_{n+1}^k = 0$. Con questa estensione la relazione (35) si può semplicemente scrivere

$$u_{j-1}^k + u_{j+1}^k = \lambda_k u_j^k$$

che, applicando (33) e (34), diventa

$$\sin \frac{k(j-1)\pi}{n+1} + \sin \frac{k(j+1)\pi}{n+1} = 2 \cos \frac{k\pi}{n+1} \sin \frac{kj\pi}{n+1}$$

la cui verifica si effettua tramite le formule di prostaferesi applicate al termine di sinistra.

Ritroviamo gli autovalori a coppie dei grafi bipartiti (per k e $n-k+1$ in (33)). L'autovalore più grande è λ_1 il cui autovettore ha componenti positive $\sin j\pi/(n+1)$, con valore più grande nel centro del cammino.

26 Spettro del grafo complementare

Dato un grafo G , lo spettro Laplaciano del grafo complementare \bar{G} è legato in modo semplice allo spettro Laplaciano di G . Si indichi con \bar{L} il Laplaciano e con \bar{A} la matrice d'adiacenza di \bar{G} . Si ha

$$\bar{A} = \mathbf{1}\mathbf{1}^\top - I - A$$

da cui

$$\bar{L} = \text{diag}(n-1-g) - \bar{A} = (n-1)I - \text{diag}(g) - \mathbf{1}\mathbf{1}^\top + I + A = nI - \mathbf{1}\mathbf{1}^\top - L$$

Ovviamente $\bar{L}\mathbf{1} = \mathbf{0}$ e quindi $\mathbf{1}$ è autovettore anche del grafo complementare con autovalore nullo (come del resto per ogni grafo). Per ogni altro autovettore u di L si noti che u è ortogonale a $\mathbf{1}$ e quindi si ha $\mathbf{1}\mathbf{1}^\top u = \mathbf{0}$, e allora

$$\bar{L}u = nu - Lu = nu - \lambda^L u = (n - \lambda^L)u$$

Quindi u è anche autovettore di \bar{L} con autovalore $(n - \lambda^L)$. Posiamo concludere che, tranne che per l'autovalore $\lambda_1^L = \bar{\lambda}_1^L = 0$, vale la relazione

$$\lambda_i^L + \bar{\lambda}_i^L = n, \quad i = 2, \dots, n \quad (36)$$

Questa relazione implica che se $\lambda_i^L = n$, allora $\bar{\lambda}_i^L = 0$, cioè il grafo complementare è sconnesso e in particolare il numero di componenti connesse di \bar{G} è uguale alla molteplicità di $\lambda_i^L = n$ aumentata di uno.

Se il grafo è k -regolare (e quindi il grafo complementare è $(n - k - 1)$ -regolare), anche gli spettri ordinari sono legati fra loro in base a (36) e (21), ovvero

$$k - \lambda_i + (n - k - 1) - \bar{\lambda}_i = n \implies \lambda_i + \bar{\lambda}_i = -1$$

Questa relazione non è valida per gli autovalori massimi $\lambda_i = k$ e $\bar{\lambda}_i = n - k - 1$ (corrispondenti agli autovalori Laplaciani nulli).

27 Spettro del grafo prodotto

Si era visto nella Sez. 17 che la matrice d'adiacenza del grafo prodotto cartesiano di due grafi con matrici d'adiacenza A_1 e A_2 è

$$I \otimes A_2 + A_1 \otimes I$$

Sia allora u un autovettore di A_1 con autovalore λ e v un autovettore di A_2 con autovalore μ , e si prenda il vettore di componenti $u_i v_h$ che possiamo scrivere come $u \otimes v$ (ordinati come (u_1, v_1) , (u_1, v_2) , ecc.; il vettore $v \otimes u$ corrisponde all'ordinamento (u_1, v_1) , (u_2, v_1)), ecc). Allora abbiamo

$$\begin{aligned} & (I \otimes A_2 + A_1 \otimes I)(u \otimes v) = \\ & \begin{pmatrix} A_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & A_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & A_2 \end{pmatrix} \begin{pmatrix} u_1 v \\ u_2 v \\ u_3 v \\ u_4 v \\ u_5 v \end{pmatrix} + \begin{pmatrix} \mathbf{0} & I & \mathbf{0} & I & \mathbf{0} \\ I & \mathbf{0} & I & I & I \\ \mathbf{0} & I & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ I & I & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} u_1 v \\ u_2 v \\ u_3 v \\ u_4 v \\ u_5 v \end{pmatrix} = \\ & \mu \begin{pmatrix} u_1 v \\ u_2 v \\ u_3 v \\ u_4 v \\ u_5 v \end{pmatrix} + \begin{pmatrix} (u_2 + u_4) v \\ (u_1 + u_3 + u_4 + u_5) v \\ (u_2) v \\ (u_1 + u_2) v \\ (u_2) v \end{pmatrix} = \mu \begin{pmatrix} u_1 v \\ u_2 v \\ u_3 v \\ u_4 v \\ u_5 v \end{pmatrix} + \lambda \begin{pmatrix} u_1 v \\ u_2 v \\ u_3 v \\ u_4 v \\ u_5 v \end{pmatrix} = (\mu + \lambda)(u \otimes v) \end{aligned}$$

Quindi gli autovalori del grafo prodotto cartesiano sono la somma degli autovalori delle due matrici e gli autovettori sono $u \otimes v$. Usando l'algebra del prodotto di Kronecker per la quale si ha $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$, si ottiene lo stesso risultato più rapidamente

$$(I \otimes A_2 + A_1 \otimes I)(u \otimes v) = (I \otimes A_2)(u \otimes v) + (A_1 \otimes I)(u \otimes v) =$$

$$u \otimes (A_2 v) + (A_1 u) \otimes v = \mu u \otimes v + \lambda u \otimes v = (\mu + \lambda) (u \otimes v)$$

Il Laplaciano è (si veda la Sez. 19)

$$\text{diag}(g^1) \otimes I + I \otimes \text{diag}(g^2) - (I \otimes A_2 + A_1 \otimes I)$$

e, se u è un autovettore di L_1 con autovalore λ^L e v un autovettore di L_2 con autovalore μ^L , possiamo scrivere

$$\begin{aligned} & (\text{diag}(g^1) \otimes I + I \otimes \text{diag}(g^2) - I \otimes A_2 - A_1 \otimes I) (u \otimes v) = \\ & \text{diag}(g^1) u \otimes v + u \otimes \text{diag}(g^2) v - u \otimes A_2 v - A_1 u \otimes v = \\ & (\text{diag}(g^1) - A_1) u \otimes v + u \otimes (\text{diag}(g^2) - A_2) v = \\ & L_1 u \otimes v + u \otimes L_2 v = \lambda^L u \otimes v + \mu^L u \otimes v = (\lambda^L + \mu^L) (u \otimes v) \end{aligned}$$

Si è visto che la molteplicità dell'autovalore nullo Laplaciano è uguale al numero di componenti connesse. Se G_1 ha k componenti connesse e G_2 ha h componenti connesse, si ottengono kh autovalori nulli del Laplaciano del grafo prodotto sommando in tutti i modi possibili i k autovalori nulli di L_1 e gli h autovalori nulli di L_2 . Quindi il grafo prodotto cartesiano ha kh componenti connesse.

Per il prodotto tensoriale la matrice d'adiacenza è data da $A_1 \otimes A_2$ e quindi si ha

$$(A_1 \otimes A_2) (u \otimes v) = (A_1 u) \otimes (A_2 v) = (\lambda \mu) (u \otimes v)$$

e quindi lo spettro ordinario del grafo prodotto tensoriale si ottiene come prodotto degli spettri dei singoli grafi. Da notare che gli autovettori sono gli stessi, sia per il prodotto cartesiano che per quello tensoriale. Quindi, essendo la matrice d'adiacenza del grafo prodotto forte la somma delle matrici d'adiacenza del prodotto cartesiano e del prodotto tensoriale, risulta che $u \otimes v$ sono autovettori anche del grafo prodotto forte con autovalori $\lambda \mu + \lambda + \mu$. Si verifichi questo risultato per la relazione $K_{mn} = K_n \boxtimes K_m$.

Non c'è in generale un'analogia semplice relazione che lega lo spettro Laplaciano del grafo prodotto tensoriale con quello dei singoli grafi. Il Laplaciano del grafo prodotto è

$$\text{diag}(g_1) \otimes \text{diag}(g_2) - A_1 \otimes A_2$$

Se i grafi sono regolari, rispettivamente k_1 -regolare e k_2 -regolare, allora la relazione è semplice, come è ovvio data la semplice relazione che lega lo spettro Laplaciano con quello ordinario nel caso di grafi regolari. Non è difficile trovare che lo spettro Laplaciano del grafo prodotto tensoriale (nel caso di grafi regolari) è dato da

$$k_2 \lambda^L + k_1 \mu^L - \lambda^L \mu^L = k_1 k_2 - \lambda \mu$$

dove λ e μ sono autovalori ordinari. Da quest'ultima relazione si vede che il numero di componenti connesse del grafo prodotto è dato dal prodotto delle componenti connesse di G_1 e G_2 (si sottolinea che la relazione è valida solo per grafi regolari).

28 Laplaciano e processi dinamici su reti

Il Laplaciano interviene in molte applicazioni. Si immagini ad esempio che per ogni nodo sia definito un ‘potenziale’ φ_i e che il potenziale in ogni nodo vari in modo proporzionale alla differenza rispetto ai potenziali nei nodi incidenti secondo la seguente formula

$$\frac{d\varphi_i(t)}{dt} = \sum_{j \in \delta(i)} (\varphi_j(t) - \varphi_i(t)), \quad i = 1, \dots, n \quad (37)$$

che in forma matriciale si può scrivere come

$$\frac{d\varphi(t)}{dt} = -L\varphi(t) \quad (38)$$

L’equazione (38) prende il nome di *equazione di diffusione* e permette di studiare come un fenomeno, misurato da $\varphi(t)$, si diffonde attraverso la rete. Qualsiasi soluzione $\varphi(t) = K\mathbf{1}$, con K costante qualsiasi, è una soluzione costante, invariante nel tempo, di (38), dato che $L\mathbf{1} = 0$. Si noti ancora che la quantità $\mathbf{1}^\top \varphi(t) = \sum_i \varphi_i(t)$ è costante nel tempo. Infatti

$$\frac{d(\mathbf{1}^\top \varphi(t))}{dt} = \mathbf{1}^\top \frac{d\varphi(t)}{dt} = -\mathbf{1}^\top L\varphi(t) = 0$$

Quindi, se la soluzione inizialmente è φ_i^0 , l’equazione (38) fa evolvere nel tempo $\varphi(t)$, ma facendo rimanere invariata la quantità $\sum_i \varphi_i(t)$ e il potenziale tenderà ad assumere il valore $\sum_i \varphi_i^0/n$ su tutti i nodi. La velocità con cui si raggiunge l’equilibrio dipende dal più piccolo autovalore positivo di L .

Se il valore $\varphi(t)$ rappresenta una quantità che fluisce nella rete passando da un nodo all’altro, allora l’equazione (38) permette di studiare come questa quantità si diffonde nella rete. La risoluzione di (38) è data da

$$\varphi(t) = e^{-Lt} \varphi^0 \quad (39)$$

dove l’esponenziale della matrice e^{-Lt} è definito come

$$e^{-Lt} = I - LT + \frac{L^2 t^2}{2} - \frac{L^3 t^3}{3!} + \frac{L^4 t^4}{4!} - \dots$$

e si può anche calcolare come

$$e^{-Lt} = U e^{-\Lambda t} U^{-1}$$

dove U è la matrice degli autovettori, Λ è la matrice diagonale degli autovalori e $e^{-\Lambda t}$ è una matrice diagonale con elementi $e^{-\lambda_i t}$ sulla diagonale e quindi

$$\varphi(t) = U e^{-\Lambda t} U^{-1} \varphi^0 \quad (40)$$

Un altro modo di arrivare a (40) consiste nello scrivere $\varphi(t)$ come combinazione lineare con coefficienti variabili nel tempo della base formata dagli autovettori di L , cioè

$$\varphi(t) = \sum_i a_i(t) u^i$$

per cui

$$\frac{d\varphi(t)}{dt} = \sum_i \frac{da_i(t)}{dt} u^i = -L \varphi(t) = -L \sum_i a_i(t) u^i = - \sum_i a_i(t) L u^i = - \sum_i a_i(t) \lambda_i u^i$$

e moltiplicando scalarmente questa espressione per u^j , essendo $u^{j\top} u^i = 0$ per $i \neq j$, e $u^{j\top} u^j = \|u^j\|^2$ si ha

$$\frac{da_i(t)}{dt} \|u^i\|^2 = -a_i(t) \lambda_i \|u^i\|^2 \implies a_i(t) = e^{-\lambda_i t} a_i^0$$

Quindi

$$\varphi(t) = \sum_i e^{-\lambda_i t} a_i^0 u^i \quad (41)$$

I coefficienti a_i^0 vanno calcolati risolvendo il sistema lineare $\varphi^0 = \sum_i a_i^0 u^i$ per cui (41) è esattamente (40).

Per il grafo completo il calcolo di (40) porta alla seguente espressione

$$\varphi_i(t) = e^{-nt} \varphi_i^0 + \frac{1}{n} (1 - e^{-nt}) \sum_j \varphi_j^0$$

Ad esempio se $\varphi_1^0 = 1$ e $\varphi_i^0 = 0$ per $i > 1$ (cioè tutta la quantità è concentrata nel nodo 1 all'istante iniziale) la diffusione è

$$\varphi_1(t) = \frac{1}{n} + \frac{n-1}{n} e^{-nt}, \quad \varphi_i(t) = \frac{1}{n} (1 - e^{-nt}), \quad i > 1$$

Quindi la diffusione è rapida, come del resto ci si aspetta, dato il coefficiente $-n$ ad esponente. In un circuito invece il tempo per raggiungere l'equilibrio viene dato dall'autovalore a valore positivo più piccolo, cioè $2(1 - \cos \frac{2\pi}{n})$, che, per n elevato, può essere approssimato da

$$2(1 - \cos \frac{2\pi}{n}) \approx 2(1 - 1 + \frac{2\pi^2}{n^2}) = \frac{4\pi^2}{n^2}$$

quindi con valori molto bassi. Ad esempio la diffusione nei nodi 1, 2, 3 e 4 per un circuito di 6 nodi è data dal grafico (i nodi 6 e 5 sono uguali a 2 e 3 rispettivamente) in Fig. 23. Si veda il confronto della diffusione nel nodo 4 rispetto al grafo completo (Fig. 24).

Spesso agli archi del grafo sono associati valori numerici che indicano il diverso grado di diffusione fra un arco e l'altro. Siano w_{ij} questi valori. In questo caso l'equazione (37) ha la forma più generale

$$\frac{d\varphi_i(t)}{dt} = \sum_{j \in \delta(i)} w_{ij} (\varphi_j(t) - \varphi_i(t)), \quad i = 1, \dots, n$$

e, usando il Laplaciano pesato, come definito in (18) l'equazione (38) rimane valida.

29 Reti elettriche

Una rete elettrica si modella naturalmente come un grafo. Indichiamo con u_i il potenziale elettrico nel nodo i , con x_{ij} la corrente che circola nell'arco (i, j) , con $v_{ij} = u_i - u_j$ la tensione (differenza di potenziale) nell'arco

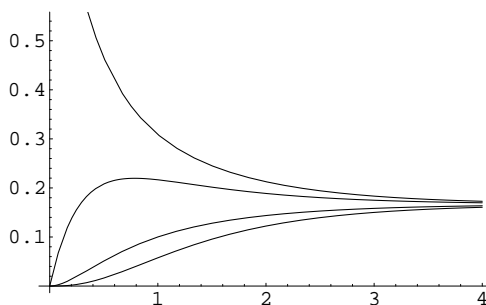


Figura 23.

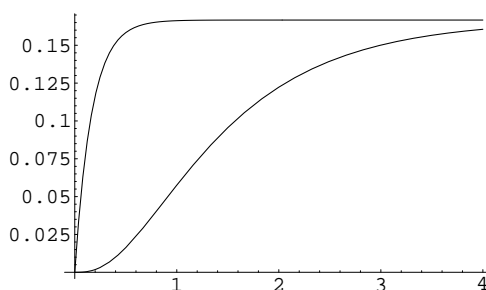


Figura 24.

(i, j) e con r_{ij} la resistenza elettrica nell'arco (i, j) . La corrente può attraversare un arco in entrambi i sensi. Conviene orientare arbitrariamente gli archi in modo da assegnare convenzionalmente un segno alla corrente: se x_{ij} è positivo nell'arco orientato (i, j) allora significa che la corrente fluisce da i a j . Se invece è negativo la corrente fluisce da j a i . Analoga convenzione vale per la tensione che è positiva se $u_i > u_j$ e negativa nel caso contrario.

Si supponga inoltre che la rete sia collegata con la rete esterna tramite due nodi, che identifichiamo come il nodo 1 e il nodo n . Dal nodo 1 entra una corrente \mathcal{I} (nota o variabile a seconda di come il problema è posto) e la stessa corrente esce dal nodo n .

Sia A la matrice d'incidenza nodi-archi del grafo orientato. Indichiamo con b il vettore $b_1 = 1$, $b_n = -1$, $b_i = 0$ altrimenti. Si indichi con C la matrice diagonale delle conduttanze (la conduttanza è l'inverso della resistenza). Allora, in base alle leggi della fisica, valgono le seguenti relazioni (scritte in modo matriciale)

$$\begin{aligned} C v &= x && \text{legge di Ohm} \\ A x &= \mathcal{I} b && \text{legge di Kirchhoff} \\ A^\top u &= v && \text{definizione di tensione} \end{aligned}$$

In dettaglio, arco per arco, la legge di Ohm è data da $v_{ij} = r_{ij} x_{ij}$. La legge di Kirchhoff esprime il fatto che la somma algebrica delle correnti che confluiscono in un nodo è sempre nulla (nei nodi non si crea né si

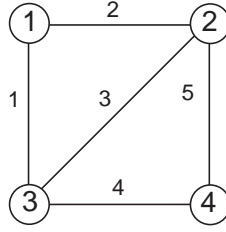


Figura 25.

distrugge corrente) tranne che nei nodi connessi con la rete esterna dove vale \mathcal{I} o $-\mathcal{I}$. Si ottiene allora

$$C A^\top u = C v = x \implies A C A^\top u = A x = \mathcal{I} b$$

La matrice $A C A^\top$ non è altro che il Laplaciano pesato e quindi si ha

$$L u = \mathcal{I} b$$

Il Laplaciano è singolare ($L \mathbf{1} = \mathbf{0}$), ma siccome anche $b^\top \mathbf{1} = 0$, il sistema lineare è risolubile con infinite soluzioni del tipo $\bar{u} + \alpha \mathbf{1}$ dove \bar{u} è una soluzione particolare. Che esistano infinite soluzioni non deve meravigliare perché i potenziali sono sempre noti a meno di una costante additiva arbitraria. Se imponiamo $u_n = 0$ allora le soluzioni u_1, \dots, u_{n-1} si calcolano definendo il *Laplaciano ridotto* L' ottenuto togliendo l'ultima riga e l'ultima colonna dal Laplaciano. Indicando con u' il vettore u senza u_n e b' il vettore b senza b_n , dobbiamo risolvere il sistema lineare non singolare

$$L' u' = \mathcal{I} b' \quad (42)$$

Se \mathcal{I} è noto e u_1 è variabile allora si risolve direttamente (42). Se invece fosse noto il potenziale u_1 ma non la corrente \mathcal{I} allora (42) rimane valida, con l'avvertenza che \mathcal{I} è ora variabile e u_1 dato. La potenza elettrica dissipata nella rete è data da

$$x^\top v = x^\top A^\top u = \mathcal{I} b^\top u = \mathcal{I} u_1 = \mathcal{I} (u_1 - u_n)$$

Quindi la potenza dissipata arco per arco ($\sum_{ij} v_{ij} x_{ij}$) è uguale al prodotto della corrente \mathcal{I} che entra nella rete per la differenza di potenziale agli estremi dei nodi di connessione. Inoltre

$$x^\top v = v^\top C v = u^\top A C A^\top u = u^\top L u$$

dove la potenza dissipata viene direttamente espressa dal Laplaciano. Si noti che con potenziale uniforme non c'è potenza dissipata (perché naturalmente non c'è corrente). Vista dall'esterno la rete si presenta come una resistenza di valore

$$\frac{u_1 - u_n}{\mathcal{I}} = \frac{x^\top v}{\mathcal{I}^2} = \frac{u^\top L u}{\mathcal{I}^2}$$

Si consideri la rete elettrica di Figura 25 dove i valori di resistenza in Ohm sono indicati vicino agli archi. Il Laplaciano è

$$L = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} & -1 & 0 \\ -\frac{1}{2} & \frac{31}{30} & -\frac{1}{3} & -\frac{1}{5} \\ -1 & -\frac{1}{3} & \frac{19}{12} & -\frac{1}{4} \\ 0 & -\frac{1}{5} & -\frac{1}{4} & \frac{9}{20} \end{pmatrix} \implies L' = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} & -1 \\ -\frac{1}{2} & \frac{31}{30} & -\frac{1}{3} \\ -1 & -\frac{1}{3} & \frac{19}{12} \end{pmatrix}$$

Risolviendo il sistema lineare (42) si ottiene

$$u_1 = \frac{61}{21} \mathcal{I}, \quad u_2 = \frac{45}{21} \mathcal{I}, \quad u_3 = \frac{48}{21} \mathcal{I}$$

Gli archi siano orientati da i verso j con $i < j$. Allora si hanno le tensioni (da $A^\top u = v$ e tenendo conto che $u_4 = 0$)

$$v_{12} = \frac{16}{21} \mathcal{I}, \quad v_{13} = \frac{13}{21} \mathcal{I}, \quad v_{23} = -\frac{3}{21} \mathcal{I}, \quad v_{24} = \frac{45}{21} \mathcal{I}, \quad v_{34} = \frac{48}{21} \mathcal{I},$$

da cui, applicando la legge di Ohm

$$x_{12} = \frac{8}{21} \mathcal{I}, \quad x_{13} = \frac{13}{21} \mathcal{I}, \quad x_{23} = -\frac{1}{21} \mathcal{I}, \quad x_{24} = \frac{9}{21} \mathcal{I}, \quad v_{34} = \frac{12}{21} \mathcal{I},$$

La potenza dissipata è $61/21 \mathcal{I}^2$ e la rete ha una resistenza equivalente di $61/21$ Ohm.

30 Catene e processi di Markov

Nelle *catene di Markov* vi è un insieme finito S di *stati*, il tempo è discretizzato e ad ogni istante di tempo la catena si muove dallo stato i allo stato j con probabilità p_{ij} . Se la transizione da i a j non è ammessa $p_{ij} = 0$. Inoltre sono ammesse transizioni verso lo stesso stato, che possiamo chiamare *auto-transizioni*. I valori p_{ij} costituiscono una matrice quadrata che viene detta *matrice delle probabilità di transizione* e che ovviamente deve soddisfare i vincoli

$$\sum_{j \in S} p_{ij} = 1, \quad p_i \geq 0, \quad i \in S$$

Ad una catena di Markov si associa in modo naturale un grafo orientato (con cappi) in cui vi è un arco (i, j) se $p_{ij} > 0$. Di particolare importanza sono le catene per le quali il grafo associato è fortemente connesso, dette *catene irriducibili*. Se indichiamo con π_i^t la probabilità che la catena si trovi nello stato i al tempo t , le probabilità π_i^{t+1} sono date da

$$\pi_j^{t+1} = \sum_{i \in S} p_{ij} \pi_i^t$$

dove si esprime il fatto che la catena si trova nello stato j essendoci arrivata da uno stato qualsiasi i , condizionatamente al fatto che si trovi nello stato i con probabilità π_i^t e che avvenga la transizione verso j con probabilità p_{ij} . La precedente espressione può essere scritta sinteticamente in modo matriciale come

$$\pi^{t+1} = \pi^t P$$

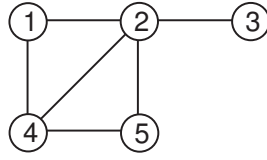


Figura 26.

da cui si vede anche che

$$\pi^t = \pi^0 P^t$$

Se avviene che

$$\lim_{t \rightarrow \infty} \pi^t = \bar{\pi}$$

la probabilità $\bar{\pi}$ viene detta *probabilità limite* e rappresenta la probabilità, indipendente dallo stato di partenza, di trovare la catena in un particolare stato dopo un tempo sufficientemente lungo. Per una catena irriducibile con un numero finito di stati la probabilità limite esiste sempre e soddisfa l'equazione

$$\bar{\pi} = \bar{\pi} P$$

quanto a dire che $\bar{\pi}$ è un autovettore sinistro di P con autovalore 1.

Consideriamo il caso particolare in cui è dato un grafo connesso non orientato $G = (V, E)$ in cui i nodi rappresentano gli stati e gli archi le transizioni ammesse. Per ogni arco (i, j) sono ammesse entrambe le transizioni $i \rightarrow j$ e $j \rightarrow i$. Quindi la catena è irriducibile. Le probabilità di transizione sono

$$p_{ii} = 0, \quad i \in V, \quad p_{ij} = \frac{1}{g_i}, \quad (i, j) \in E$$

Quindi la catena effettua sempre una transizione verso un nodo adiacente con probabilità uniforme. La matrice P è data da

$$P = A D^{-1}$$

dove A è la matrice d'adiacenza di un grafo non orientato e $D = \text{diag}(g)$. L'evoluzione della catena è descritta allora dalle potenze di $(A D^{-1})$. Ad esempio per il grafo di Fig. 26 si ha

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}, \quad D^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad A D^{-1} = \begin{pmatrix} 0 & \frac{1}{4} & 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & 1 & \frac{1}{3} & \frac{1}{2} \\ 0 & \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{4} & 0 & \frac{1}{3} & 0 \end{pmatrix},$$

$$(A D^{-1})^5 = \begin{pmatrix} 0.125 & 0.203 & 0.118 & 0.190 & 0.125 \\ 0.406 & 0.250 & 0.451 & 0.308 & 0.406 \\ 0.059 & 0.113 & 0.042 & 0.090 & 0.059 \\ 0.285 & 0.231 & 0.271 & 0.222 & 0.285 \\ 0.125 & 0.203 & 0.118 & 0.190 & 0.125 \end{pmatrix} \quad (43)$$

Il significato della matrice è il seguente: se la catena inizia nel nodo 1, questo è equivalente ad una probabilità iniziale $\pi_1^0 = 1$ e $\pi_i^0 = 0$ per $i \neq 1$. Il prodotto $(AD^{-1})^5 \pi^0$ è dato dalla prima colonna di (43) (in generale se parte dallo stato k bisogna guardare la colonna k -ma). Gli elementi di tale colonna danno la probabilità di trovare la catena nei rispettivi stati alla quinta iterazione. Quindi c'è una probabilità 0.406 di trovare la catena nello stato 2 e solo 0.059 di trovarla nello stato 3.

$$(AD^{-1})^{20} = \begin{pmatrix} 0.167 & 0.166 & 0.167 & 0.167 & 0.167 \\ 0.333 & 0.334 & 0.332 & 0.334 & 0.333 \\ 0.084 & 0.083 & 0.084 & 0.083 & 0.084 \\ 0.250 & 0.250 & 0.250 & 0.250 & 0.250 \\ 0.167 & 0.166 & 0.167 & 0.167 & 0.167 \end{pmatrix}$$

Dopo 20 iterazioni la probabilità di trovare la catena in un determinato stato è quasi indipendente dallo stato iniziale. Infatti le colonne differiscono di pochissimo fra loro. Le colonne sono appunto uguali alla probabilità limite per la quale

$$\bar{\pi} = AD^{-1} \bar{\pi} \implies (I - AD^{-1}) \bar{\pi} = 0 \implies (D - A) D^{-1} \bar{\pi} = 0 \implies L D^{-1} \bar{\pi} = 0$$

quanto a dire che $D^{-1} \bar{\pi}$ è l'autovettore del Laplaciano corrispondente all'autovalore 0, che è $\mathbf{1}$. Allora (con α costante di normalizzazione)

$$D^{-1} \bar{\pi} = \alpha \mathbf{1} \implies \bar{\pi} = \alpha D \mathbf{1} \implies \bar{\pi}_i = \frac{g_i}{2m} \quad (44)$$

con probabilità quindi proporzionale al grado. La cosa è anche intuitiva perché un nodo con grado elevato ha maggiori possibilità di essere raggiunto. Per il grafo di Fig. 26 si ha

$$\bar{\pi} = \left(\frac{1}{6} \quad \frac{1}{3} \quad \frac{1}{12} \quad \frac{1}{4} \quad \frac{1}{6} \right)$$

Comunque all'espressione (44) si può pervenire più agevolmente usando le *equazioni dettagliate di bilancio*, per le quali se esiste un vettore $\bar{\pi}$ tale che

$$\bar{\pi}_i p_{ij} = \bar{\pi}_j p_{ji}$$

per ogni i e j , allora tale vettore è una probabilità limite. In questo caso, essendo $p_{ij} = 1/g(i)$, si vede immediatamente che basta porre π_i proporzionale a $g(i)$.

Un modello particolare di catena di Markov si ha se ogni transizione fuori dallo stato ha la stessa probabilità p per ogni stato, cioè $p_{ij} = p$ per ogni $i \neq j$. Se p è tale per cui la probabilità di una transizione fuori dallo stato è inferiore ad 1, cioè $\sum_{j \neq i} p_{ij} < 1$, tale differenza viene assunta da una auto-transizione, cioè $p_{ii} = 1 - \sum_{j \neq i} p_{ij}$. Si noti che non necessariamente gli stati hanno il medesimo numero di transizioni possibili fuori dallo stato. Si supponga inoltre che se esiste la transizione da i a j esiste anche la transizione opposta da j a i (e necessariamente con la stessa probabilità per le ipotesi fatte). Allora le equazioni dettagliate di bilancio danno luogo ad una probabilità limite uniforme.

Può sorprendere che la probabilità limite sia uniforme anche quando gli stati hanno un numero diverso di transizioni fuori dallo stato. Si è infatti notato per il precedente modello di catena che si arriva con maggior

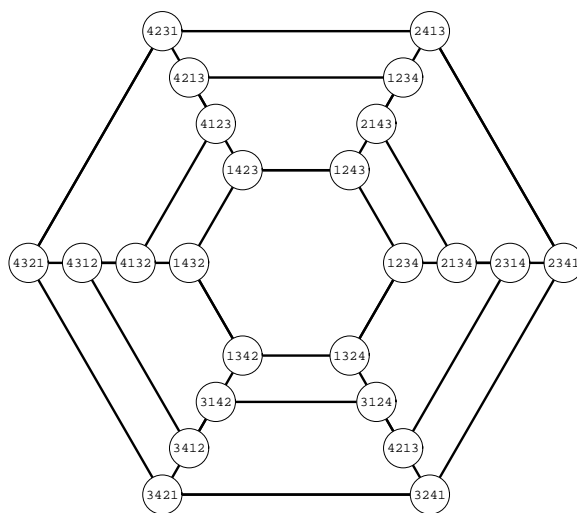


Figura 27.

frequenza ad uno stato se vi si può arrivare da un numero maggiore di stati. Tuttavia in questo secondo modello sono presenti anche le auto-transizioni e la probabilità di un'auto-transizione è maggiore se il grado è minore. I due fatti si compensano esattamente: si arriva più frequentemente ad uno stato con grado elevato ma si abbandona lo stato anche con maggior frequenza, mentre si arriva più raramente in uno stato con basso grado ma poi si resta più a lungo nello stato.

Avere una probabilità limite uniforme è molto utile quando si debba campionare un elemento casuale all'interno di un insieme V e sia possibile trasformare un elemento in un altro secondo opportune regole. Se la trasformazione è possibile questo corrisponde a percorrere un arco del grafo. Si suppone naturalmente che anche la trasformazione inversa sia possibile.

Ad esempio si immagini di dover generare una permutazione casuale di n elementi. Un modo molto efficiente di farlo consiste nello scegliere casualmente un intero fra 1 e n , poi uno fra i rimanenti e così di seguito. Si immagini però di voler seguire un'altra strada, generando una permutazione da un'altra semplicemente scambiando due elementi adiacenti. Il grafo risultante è in Fig. 27.

Si tratta di un grafo 3-regolare e quindi la probabilità stazionaria è uniforme. Si tratta però anche di un grafo bipartito (ogni permutazione è pari o dispari) quindi metà permutazioni si raggiungono in un numero pari di transizioni e l'altra metà in un numero dispari di transizioni. Tecnicamente si parla di catena *periodica* ed una probabilità limite non esiste. Però se si sceglie *casualmente* il numero di iterazioni si raggiunge una qualsiasi permutazione e dopo un numero elevato di iterazioni ogni permutazione viene raggiunta con probabilità stazionaria.

È interessante sapere con che velocità si raggiunge la probabilità stazionaria. Questa dipende dal più grande autovalore di (AD^{-1}) strettamente minore di 1 in valore assoluto. Per l'esempio in questione tale autovalore vale $\tilde{\lambda} = 0.804738$. Affinché le probabilità stazionarie degli stati differiscano al più di ε , il numero k di iterazioni deve essere tale che $\tilde{\lambda}^k \leq \varepsilon$. Ponendo ad esempio $\varepsilon = 0.01$ si ottiene $k \geq 22$. Quindi per avere un campionamento abbastanza uniforme (cioè con differenza di probabilità tra gli stati inferiore a 0.01) bisognerebbe effettuare almeno 22 transizioni e poi con probabilità 1/2 effettuarne ancora una.

L'esempio della permutazione casuale è meno astratto di quel che sembri. Un mazzo di carte da gioco è rappresentabile da tutte le permutazioni delle sue 40 o 52 carte e quando si mescolano le carte si cerca di ottenere una qualsiasi permutazione in modo casuale. In questo caso la generazione di una permutazione non si effettua con nessuno dei due metodi delineati, in quanto praticamente 'lenti'. Si effettua tuttavia con una catena di Markov, anche se diversa e più efficiente di quella indicata. La consueta tecnica di mescolare le carte consiste nello spezzare la permutazione in due parti e fare una fusione delle due parti in modo da mantenere l'ordine dell'una e dell'altra parte. Con quattro carte significa passare dalla permutazione $(1, 2, 3, 4)$ ad una delle possibili 11 permutazioni

2	2	2	1	1	3	3	3	1	1	4
1	3	3	3	3	1	1	4	2	4	1
3	1	4	2	4	2	4	1	4	2	2
4	4	1	4	2	4	2	2	3	3	3

In questo caso non si tratta di un grafo simmetrico. Infatti da $(3, 1, 4, 2)$ non si può tornare indietro a $(1, 2, 3, 4)$. È stato dimostrato che con questa catena di Markov bastano 7 smazzate per raggiungere una probabilità abbastanza uniforme.

Anche nei *processi Markoviani* vi è un numero finito di *stati* però il tempo è continuo. La transizione da uno stato all'altro (ora necessariamente diverso) avviene nel seguente modo: per ogni transizione possibile è definito un tasso di transizione q_{ij} come il numero medio di transizioni per unità di tempo dallo stato i allo stato j . Quando il sistema si trova nello stato i la transizione avviene secondo la funzione di distribuzione esponenziale

$$1 - e^{-\sum_j q_{ij} t}$$

che dà la probabilità che la transizione avvenga entro il tempo t . Quando avviene la transizione dallo stato i , questa avviene verso uno stato per cui esiste l'arco (i, j) con probabilità proporzionale a q_{ij} . Si può dimostrare che la probabilità p_i di trovare il sistema nello stato i obbedisce alla equazione

$$\frac{dp_i(t)}{dt} = \sum_{j \in \delta^+(i)} q_{ij} (p_j(t) - p_i(t))$$

che è del tutto simile a (38), con la notevole differenza però, che in questo caso la matrice non è simmetrica, perché il grafo è orientato. Se le transizioni possono avvenire da i a j e anche da j a i con lo stesso tasso $q_{ij} = q_{ji}$ allora la matrice Q è l'opposto del Laplaciano. Va detto che una matrice simmetrica non è molto interessante nei processi Markoviani, in quanto in questo caso si conosce a priori la probabilità stazionaria, che risulta uniforme su tutti gli stati. La matrice Q prende il nome di *generatore infinitesimale*

del processo Markoviano. Possiamo tuttavia considerare la controparte continua della catena di Markov vista precedentemente. Se per ogni nodo il tasso di transizione verso i nodi adiacenti è costante e assumiamo che in ogni nodo il soggiorno nel nodo stesso obbedisca alla stessa distribuzione $1 - e^{-t}$, allora $q_{ij} = 1/g(i)$.

31 Il Page Rank di Google

Consideriamo il grafo orientato in cui i nodi sono le pagine web ed è presente un arco orientato se c'è un link da una pagina ad un'altra pagina. In questo grafo possono essere presenti anche dei cappi, se una pagina ha un link verso se stessa.

Il Page Rank $r(i)$ della pagina i usato da Google è dato dalla formula

$$r(i) = \frac{1 - \alpha}{n} + \alpha \sum_{j:(j,i) \in E} \frac{r(j)}{g_j} \quad (45)$$

Nella sommatoria sono presenti le pagine j che hanno un link verso la pagina i . Per ognuna di queste pagine si considera il suo Page Rank normalizzato rispetto al suo grado. Il senso di questa normalizzazione è che è importante essere puntati da un'altra pagina, ma la cosa è tanto meno importante quanti più link ha questa pagina. Il valore dato dalla sommatoria viene poi combinato in modo convesso con il numero $1/n$. Se α fosse uguale a zero, ovviamente il Page Rank sarebbe uguale per tutte le pagine. Si noti ancora che

$$\sum_i r(i) = (1 - \alpha) + \alpha \sum_i \sum_{j:(j,i) \in E} \frac{r(j)}{g_j} = 1 - \alpha + \alpha \sum_j \frac{r(j)}{g_j} \sum_{i:(j,i) \in E} 1 = 1 - \alpha + \alpha \sum_j r(j)$$

da cui $\sum_i r(i) = 1$. Allora il Page Rank può essere visto come una distribuzione di probabilità sulle pagine. Infatti possiamo fare le seguenti considerazioni. Sia A la matrice d'incidenza del grafo, con eventualmente degli uni sulla diagonale se ci sono dei cappi e sia $D = \text{diag}(g)$. La matrice $D^{-1}A$ è la matrice d'adiacenza con le righe divise per il grado dei rispettivi nodi. Si consideri la matrice

$$P = \frac{1 - \alpha}{n} \mathbf{1} \mathbf{1}^\top + \alpha D^{-1}A \quad (46)$$

Si noti che

$$P \mathbf{1} = \frac{1 - \alpha}{n} \mathbf{1} \mathbf{1}^\top \mathbf{1} + \alpha D^{-1}A \mathbf{1} = (1 - \alpha) \mathbf{1} + \alpha D^{-1}g = \mathbf{1}$$

Inoltre $P > 0$. Quindi P può essere vista come la matrice di transizione di una catena di Markov sulle pagine web. La probabilità stazionaria è data dall'autovettore sinistro $r P = r$. Essendo $P > 0$, r è positivo. Quindi abbiamo (si ricordi che $\sum_i r(i) = 1$, ovvero $r \mathbf{1} = 1$)

$$r P = r \implies \frac{1 - \alpha}{n} r \mathbf{1} \mathbf{1}^\top + \alpha r D^{-1}A = \frac{1 - \alpha}{n} \mathbf{1}^\top + \alpha r D^{-1}A = r$$

cioè la formula (45). Il calcolo di r si può effettuare iterando $r^k = M r^{k-1}$. Anche per il grafo reale del web con milioni di nodi, essendo il grafo molto sparso, il calcolo richiede poche ore. Google suggerisce di usare $\alpha = 0.85$.

Esaminiamo in dettaglio la matrice di transizione (46). Questa matrice può essere interpretata come: con probabilità $(1 - \alpha)$ si sceglie una pagina web a caso (primo termine della somma in (46)) e con probabilità α si sceglie a caso uno dei link presenti nella pagina (secondo termine della somma).

Quindi il Page Rank non è altro che la probabilità, a lungo termine, di trovarsi su una pagina. Se fosse $\alpha = 1$, ci sarebbero solo transizioni dovute ai link, ma questo non garantirebbe l'irriducibilità della catena. Con valori di α di poco inferiori ad 1, la catena è irriducibile e quindi esiste la probabilità limite. Però il valore assoluto del secondo autovalore di P , che determina la velocità di convergenza della catena, decresce con α e quindi un basso valore di α sarebbe consigliabile per avere una convergenza veloce. Però bassi valori di α rendono i Page Rank sempre più uguali e quindi fanno perdere l'informazione che vogliamo ricavare dal grafo. Il valore di $\alpha = 0.85$ suggerito da Google, cerca di mediare fra le opposte esigenze.

Si consideri come esempio il grafo di Fig. 20. Il grafo non è fortemente connesso e quindi la catena, limitata agli archi, non è irriducibile. Infatti allo stato 1 non si può arrivare e quindi si tratta di uno stato transiente con probabilità stazionaria 0. Anche lo stato 4, a cui si accede solo da 1, è transiente. Rispetto agli altri stati la catena è periodica con periodo 3, perché la catena passa con regolarità attraverso la successione di stati $5 \rightarrow 3 \rightarrow \{2, 6\} \rightarrow 5$. Quindi per $\alpha = 1$ non si avrebbe la probabilità limite. Nella tabellina sono riportati i valori di Page Rank per diversi valori di α . Come si vede il valore di α non modifica l'ordine relativo del rank anche se, abbassando α comprime i valori verso $1/n$ (i valori di r_i della tabella non sommano necessariamente ad 1 perché sono stati arrotondati alla terza cifra decimale).

α	r_1	r_2	r_3	r_4	r_5	r_6
0.9	0.017	0.173	0.314	0.017	0.322	0.158
0.7	0.050	0.181	0.275	0.050	0.297	0.146
0.5	0.083	0.185	0.238	0.083	0.268	0.143
0.3	0.117	0.182	0.204	0.117	0.233	0.147
0.1	0.150	0.174	0.177	0.150	0.191	0.159

32 Famiglie di sottoinsiemi

Sono molto frequenti situazioni in cui è assegnato un insieme K di m elementi e una famiglia

$$\mathbf{F} = \{A^1, A^2, \dots, A^n\}$$

di sottoinsiemi $A^j \subset K$. Un modo frequente di rappresentare tale situazione fa uso di una matrice d'incidenza $m \times n$

$$D_{ij} = \begin{cases} 1 & \text{se } i \in A^j \\ 0 & \text{se } i \notin A^j \end{cases} \quad (47)$$

Se ogni insieme A^j avesse due elementi allora (K, \mathbf{F}) sarebbe un grafo e la matrice d'incidenza sarebbe la matrice d'incidenza nodi-archi di un grafo non orientato. È stata quindi introdotta in letteratura la definizione di *ipergrafo* ad indicare il caso generico in cui gli insiemi hanno un numero generico di elementi e ogni insieme

viene definito come *iperarco*. Tuttavia pochissimi risultati della teoria dei grafi si estendono agli ipergrafi, per cui il modello di ipergrafo viene usato raramente.

Piuttosto è più conveniente rappresentare la situazione come un grafo bipartito in cui K e \mathbf{F} sono i due insiemi di vertici e c'è un arco fra $i \in K$ e $A^j \in \mathbf{F}$ se $i \in A^j$. La matrice d'adiacenza è allora, elencando prima i vertici di K e poi quelli di \mathbf{F} ,

$$A = \begin{pmatrix} \mathbf{0}_1 & D \\ D^\top & \mathbf{0}_2 \end{pmatrix}$$

Mentre A contiene tutta l'informazione riguardante l'interazione fra K e \mathbf{F} , un'informazione parziale si può ottenere semplicemente guardando se due insiemi A^j e A^h hanno almeno un elemento in comune, ovvero, in termini di grafo bipartito, se sono adiacenti allo stesso nodo. Si ottiene in questo modo un ulteriore grafo G_1 con vertici $\{A^1, A^2, \dots, A^n\}$ e un arco fra A^j e A^h se e solo se $A^j \cap A^h \neq \emptyset$. Questo grafo viene chiamato *proiezione unimodale*. Si noti che G_1 è anche l'unione di m cricche, una per ogni elemento $i \in K$. La cricca associata ad i è individuata dagli insiemi A^j che contengono i .

Analogamente si può creare un altro grafo G_2 con K insieme di vertici e un arco fra i e h se e solo se esiste un insieme A^j che li contiene entrambi. Anche G_2 costituisce una proiezione unimodale ed è l'unione di cricche una per ogni sottoinsieme della famiglia. Nel caso particolare in cui (K, \mathbf{F}) è un grafo, G_2 è proprio il grafo stesso, mentre G_1 è il grafo di linea.

Se ad esempio l'insieme K è costituito dagli attori di cinema e ogni insieme A^j è l'elenco degli attori presenti nel film j , il grafo G_1 ha come nodi i film e c'è un arco fra due film se c'è almeno un attore che ha recitato in entrambi. Il grafo G_2 è invece basato sugli attori e c'è un arco fra due attori se hanno recitato assieme in un film.

Ad esempio sia $K = \{1, 2, \dots, 6\}$, e

$$\mathbf{F} = \{\{1, 2, 5\}, \{3, 6\}, \{1, 3, 4, 6\}, \{3\}, \{4, 5, 6\}, \{1, 6\}, \{2, 3, 5\}, \{2, 4\}\}$$

Allora

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

Il grafo bipartito, dove i sottoinsiemi della famiglia sono indicati con le lettere da A a H, e il grafo G_1 sono rappresentati in Fig. 28. Il grafo G_2 è “quasi” il grafo completo di 6 nodi, manca solo l'arco (2, 6) e quindi non viene disegnato. Se si esegue il prodotto $\tilde{B} = D D^\top$ si ottiene una matrice quadrata con righe e colonne associate all'insieme K . L'elemento (i, j) di \tilde{B} conta a quanti sottoinsiemi appartengono sia i che j . Quindi la matrice B ottenuta da \tilde{B} semplicemente ponendo $B_{ij} = 1$ se $\tilde{B}_{ij} > 0$ e $i \neq j$, e $B_{ij} = 0$ altrimenti, è esattamente la matrice d'adiacenza del grafo G_2 .

Analogamente si possono definire le matrici $\tilde{C} = D^\top D$ e $C_{ij} = 1$ se $\tilde{C}_{ij} > 0$ e $i \neq j$, e $C_{ij} = 0$ altrimenti. La matrice C è la matrice d'adiacenza del grafo G_1 . Per l'esempio si ottiene:

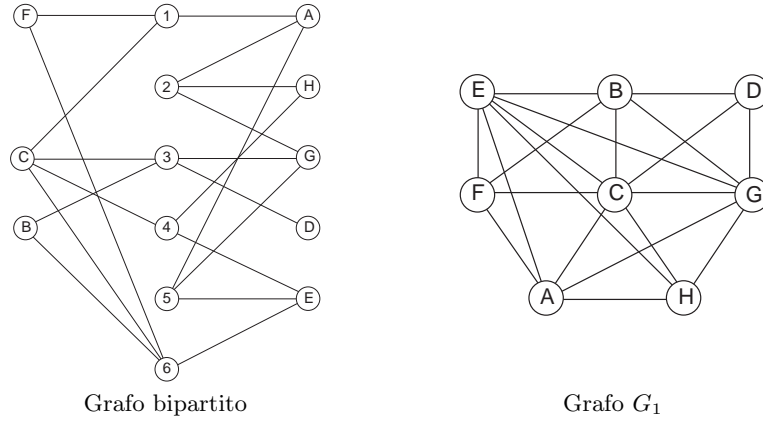


Figura 28.

$$\tilde{C} = \begin{pmatrix} 3 & 0 & 1 & 0 & 1 & 1 & 2 & 1 \\ 0 & 2 & 2 & 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 3 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 & 1 & 2 & 0 & 0 \\ 2 & 1 & 1 & 1 & 1 & 0 & 3 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 2 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} 3 & 1 & 1 & 1 & 1 & 2 \\ 1 & 3 & 1 & 1 & 2 & 0 \\ 1 & 1 & 4 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 & 1 & 2 \\ 1 & 2 & 1 & 1 & 3 & 1 \\ 2 & 0 & 2 & 2 & 1 & 4 \end{pmatrix}$$

L'informazione contenuta in G_1 e G_2 è parziale perché non è possibile ricostruire (K, \mathbf{F}) a partire da G_1 o da G_2 . Dato G_1 ogni famiglia \mathbf{F} ottenuta come copertura di cricche di G_1 ha come proiezione unimodale G_1 . Non è possibile ricostruire (K, \mathbf{F}) neppure a partire dalla conoscenza sia di G_1 che di G_2 . Si vedano in Figura 29 due grafi G_1 e G_2 che sono entrambi le proiezioni unimodali delle due diverse famiglie di sottoinsiemi:

	A	B	C	D		A	B	C	D
1	1	0	0	0	1	1	0	0	0
2	1	0	0	0	2	1	0	0	0
3	1	1	1	0	3	1	1	1	0
4	0	1	0	0	4	0	1	0	1
5	0	1	0	1	5	0	1	1	0
6	0	0	1	1	6	0	0	1	1

(48)

Tuttavia in molti casi è possibile, anche se non facile, ricostruire (K, \mathbf{F}) a partire dalle matrici \tilde{B} e \tilde{C} . In generale, siano date due matrici positive semidefinite A e B che sappiamo essere uguali a $A = DD^T$ e $B = D^T D$, con D matrice $m \times n$ ($m \leq n$) da trovare. Si vede facilmente che B ha gli stessi autovalori λ_i , $i = 1, \dots, m$, di A più $n - m$ volte l'autovalore 0. Supponiamo che gli autovalori di A siano distinti. Chiamiamo regolari gli autovettori di B in corrispondenza degli m autovalori di A . Siano u_1, \dots, u_m gli autovettori di A

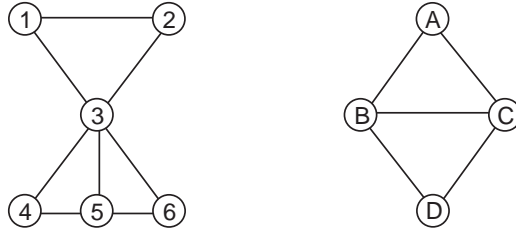


Figura 29. Proiezioni unimodali delle famiglie in (48)

e v_1, \dots, v_m gli autovettori regolari di B . Sia U la matrice formata dagli autovettori u_1, \dots, u_m sulle colonne e normalizzati $\|u_i\| = 1$ (quindi $UU^\top = U^\top U = I$). Sia V la matrice formata dagli autovettori v_1, \dots, v_m sulle colonne e normalizzati $\|v_i\| = \sqrt{\lambda_i}$ (quindi $V^\top V = \Lambda$). Da $AU = DD^\top U = U\Lambda$ (per definizione di autovettore di A), e $BV = V\Lambda$ (per definizione di autovettore di B) si ha

$$D^\top AU = D^\top DD^\top U = D^\top U\Lambda = BD^\top U$$

da cui si vede che le colonne di $D^\top U$ sono autovettori di B . Siccome gli autovettori sono definiti a meno di una costante moltiplicativa (abbiamo supposto autovalori distinti e quindi i sottospazi invarianti hanno dimensione uno) possiamo dire che esiste una matrice diagonale P tale che

$$VP = D^\top U \quad (49)$$

Analogamente abbiamo

$$DBV = DD^\top DV = DV\Lambda = ADV$$

da cui si vede che le colonne di DV sono autovettori di A . Quindi esiste una matrice diagonale Q tale che

$$UQ = DV \quad (50)$$

Da (49) e (50) ricaviamo

$$VPQ = BV \implies PQ = \Lambda$$

Inoltre da (49)

$$PV^\top VP = U^\top DD^\top U \implies P\Lambda P = U^\top AU = \Lambda$$

da cui si vede che gli elementi diagonali di P devono valere ± 1 . Da (50) abbiamo

$$QU^\top UQ = V^\top D^\top DV \implies Q^2 = V^\top BV = \Lambda V^{-1}BV = \Lambda^2$$

da cui si vede che gli elementi diagonali di Q devono valere $\pm \lambda_i$. Siccome $PQ = \Lambda$ si deve avere $P_{ii} = 1$ e $Q_{ii} = \lambda_i$ oppure $P_{ii} = -1$ e $Q_{ii} = -\lambda_i$. Applicando (49) otteniamo

$$D^\top = VP U^\top \implies D = UPV^\top \quad (51)$$

Da (51) e dalla possibilità di scelta del segno in P risulta che ci sono 2^m possibili matrici D che danno luogo alle stesse matrici A e B . Tuttavia, se sappiamo a priori che D deve essere non negativa le possibili matrici possono essere molte di meno, se non uniche. Se m è piccolo si possono provare tutte i 2^m casi. Ma con m grande tale approccio non è ipotizzabile. Risulta che determinare il segno di P per ottenere $D \geq 0$ è un problema NP-difficile. Lo si può dimostrare trasformando il problema di partizione in questo problema.

Nel problema di partizione sono dati m numeri interi $w_k \geq 0$ e ci si chiede se esiste un sottoinsieme $J \subset \{1, \dots, m\}$ tale che $\sum_{k \in J} w_k = \sum_{k \notin J} w_k$. Alternativamente ci si può chiedere se esiste un vettore x con m componenti uguali a ± 1 tale che $\sum_{k=1}^m w_k x_k = 0$. Questo problema è un caso particolare del precedente. Infatti da (51) abbiamo

$$D_{ij} = \sum_{k=1}^m U_{ik} P_{kk} V_{jk} = \sum_{k=1}^m w_k^{ij} x_k$$

dove $w_k^{ij} = U_{ik} V_{jk}$ e $x_k = P_{kk}$. Se avviene che esistono indici (i, j) e (i', j') tali che $w_k^{ij} = -w_k^{i'j'}$ allora chiedere che $D_{ij} \geq 0$ e $D_{i'j'} \geq 0$ è equivalente a chiedere se $D_{ij} = 0$, cioè se esiste una partizione dei numeri w_k^{ij} .

Come piccolo esempio sia

$$D = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix} \implies A = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

e si voglia ricostruire D a partire da A e B . Si ottiene

$$\lambda_1 = 7, \quad \lambda_2 = 3, \quad U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad V = \frac{1}{\sqrt{2}} \begin{pmatrix} 3 & 1 \\ 2 & -2 \\ 1 & 1 \end{pmatrix}$$

$$D = U P V^T = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 1 & -2 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & 4 & 0 \\ 4 & 0 & 2 \end{pmatrix}$$

Si ottengono le quattro matrici

$$D = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 2 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} -2 & 0 & -1 \\ -1 & -2 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} -1 & -2 & 0 \\ -2 & 0 & -1 \end{pmatrix}$$

viene ricostruita a meno del segno (infatti anche $-D$ dà luogo a A e B). Due delle quattro matrici sono non positive e possono essere escluse (se sappiamo che la matrice originaria non può essere negativa). Tuttavia le due matrici positive possono essere entrambe soluzione e quindi non sempre la soluzione può essere univoca. In particolare se vi sono autovalori multipli vi sono sottospazi invarianti e quindi le matrici P e Q non sono semplicemente diagonali.

33 Cocitazione e accoppiamento bibliografico

Sia dato un grafo orientato $G = (V, E)$. Per ogni coppia i e j definiamo la seguente quantità

$$C_{ij} = |\{k : (k, i) \in E, (k, j) \in E\}|$$

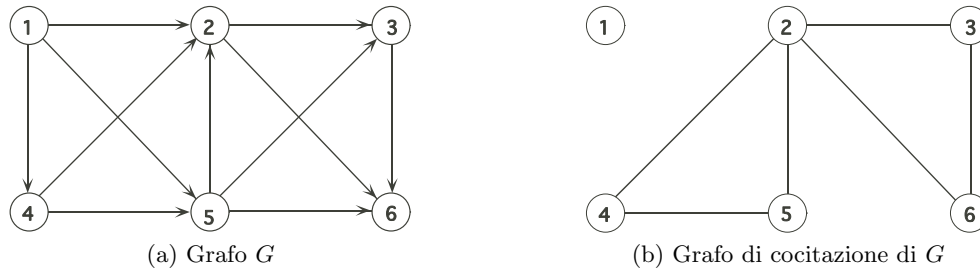


Figura 30.

Quindi C_{ij} conta quanti vertici hanno un arco orientato sia verso i che verso j . Per definizione la matrice C è simmetrica con elementi sulla diagonale uguali al grado interno dei nodi. La matrice C si può anche ottenere direttamente dalla matrice d'adiacenza A del grafo orientato:

$$C = A^T A$$

dato che $C_{ij} = \sum_k A_{ki} A_{kj}$. A partire da C si può definire un *grafo di cocitazione* dove i vertici sono ancora V ed è presente l'arco (i, j) , con $i \neq j$, se $C_{ij} > 0$.

Per il grafo in Fig. 30-(a) si ha la matrice di cocitazione

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

con il conseguente grafo di cocitazione in Fig. 30-(b)

Si noti che ogni vertice $j \in V$ del grafo orientato G può essere pensato come un sottoinsieme $A^j = \{i : (i, j) \in E\}$ di vertici, cioè il sottoinsieme dei vertici i , per i quali esiste l'arco (i, j) . Quindi è come se fossero definiti un insieme $K = V$ ed una famiglia di n sottoinsiemi $\mathbf{F} = \{A^1, \dots, A^n\}$, ciascuno associato ad un nodo. La matrice d'adiacenza del grafo orientato (V, E) è esattamente la matrice d'incidenza (47) per il sistema (V, \mathbf{F}) e quindi il grafo di cocitazione non è altro che la proiezione unimodale di questa particolare famiglia di sottoinsiemi. Il grafo bipartito derivato dalla famiglia è quello di Fig. 31 (dove ogni nodo è stato sdoppiato, quello tondo corrisponde al vertice visto come sottoinsieme e quello quadrato al vertice, visto come elemento, quindi ogni arco orientato è diventato un arco non orientato da un vertice quadrato ad uno tondo). Si noti che anche il grafo di cocitazione, come del resto ogni proiezione unimodale, è l'unione di cricche, in particolare, per ogni vertice i , la cricca formata dai successori di i .

Il termine 'cocitazione' deriva dall'interpretazione del grafo orientato come grafo delle citazioni nella letteratura scientifica. Ogni vertice rappresenta un articolo e un arco (i, j) sta ad indicare che l'articolo i cita

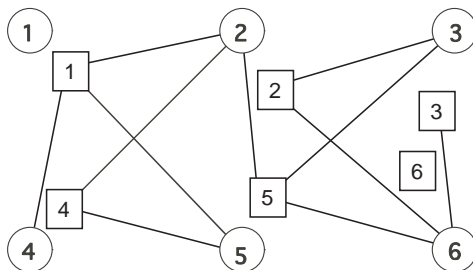


Figura 31. Grafo bipartito della famiglia (V, \mathbf{F})

l'articolo j . Data una coppia di articoli i e j la matrice di cocitazione indica quanti sono gli articoli che citano sia i che j . L'esistenza di un arco fra i e j sta quindi ad indicare una similarità di argomenti fra i due articoli. È verosimile infatti che articoli su argomenti molto diversi difficilmente siano citati da uno stesso articolo. Inoltre il valore numerico dato dalla matrice di cocitazione, che conta quante volte una coppia di articoli compare assieme nella bibliografia di un articolo, indica la maggiore o minore similarità fra i due articoli, oltre che l'importanza e l'influenza di questi articoli sugli articoli successivi. Gli elementi diagonali della matrice danno semplicemente il numero delle citazioni di un determinato articolo. Ovviamente articoli recenti che, per motivi temporali, non hanno ancora potuto ricevere citazioni appaiono valutati "scarsamente" dalla matrice di cocitazione. Ad esempio, interpretando il grafo dell'esempio come grafo delle citazioni, l'articolo 1 potrebbe sia essere un articolo recente, non ancora citato, come anche un articolo 'stagionato' ma per nulla importante.

In modo analogo si può definire la famiglia di sottoinsiemi (V, \mathbf{F}) dove ogni sottoinsieme $A^i \subset K$ è l'insieme dei vertici per cui esiste l'arco (i, j) . Viene quindi ribaltato l'ordine dell'orientazione e si può definire la quantità

$$B_{ij} = |\{k : (i, k) \in E, (j, k) \in E\}|$$

dove B_{ij} conta quanti vertici hanno un arco orientato sia da i che da j . Anche la matrice B è simmetrica e gli elementi diagonali danno il grado esterno dei nodi. La matrice B si può anche ottenere direttamente dalla matrice d'adiacenza A del grafo orientato:

$$B = A A^\top$$

A partire da B si può definire un *grafo di accoppiamento bibliografico* dove i vertici sono ancora V ed è presente l'arco (i, j) , con $i \neq j$, se $B_{ij} > 0$. Per il grafo di Fig. 30-(a) si ha la matrice

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

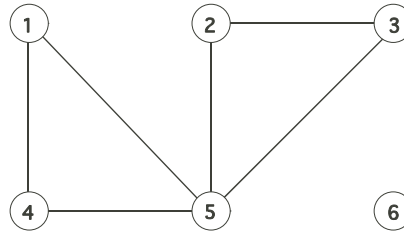


Figura 32. Grafo di accoppiamento bibliografico di G

Il grafo bipartito associato a (V, \mathbf{F}) è lo stesso di Fig. 31 con i ruoli dei vertici tondi e quadri ribaltati, cioè quello quadrato corrisponde al vertice visto come sottoinsieme e quello tondo al vertice, visto come elemento. Il grafo di accoppiamento bibliografico è in Fig. 32.

Rispetto al grafo di cocitazione il grafo di accoppiamento bibliografico rivela maggiormente una similarità fra argomenti senza invece rivelare una maggiore o minore importanza degli articoli. Infatti gli elementi diagonali danno semplicemente la lunghezza della bibliografia di un articolo, mentre gli elementi non diagonali danno il numero di citazioni comuni di due articoli. Inoltre non varia nel tempo ed è quindi più indicato per segnalare similarità di argomenti.

34 Minimi tagli e massimi flussi

Se un grafo rappresenta un insieme di possibili collegamenti, è importante sapere se vi sono degli archi critici nel grafo, la cui rimozione provoca la sconnessione del grafo. Il più critico di tali insiemi è quello costituito dal minor numero di archi. Un tale insieme prende il nome di *minimo taglio*.

Molto spesso ad ogni arco sono associate delle grandezze numeriche rappresentanti la massima quantità che può transitare sull'arco, quantità che può essere a seconda dell'applicazione, numero di bit al secondo in una rete di calcolatori, numeri di veicoli al minuto in una rete stradale, numero di litri al secondo in una rete idraulica, ecc. La quantità in transito sull'arco viene normalmente indicata come *flusso* mentre la massima quantità di flusso transitabile sull'arco è indicata come *capacità*. In questo caso un insieme critico di archi è un taglio di minima capacità, dove la capacità di un taglio è la somma delle capacità degli archi del taglio.

È allora importante determinare in un grafo un taglio di capacità minima. Sono stati costruiti vari algoritmi polinomiali per trovare un taglio minimo. Alcuni di questi si basano sui flussi. È opportuno definire i flussi su un grafo orientato. Se il grafo in origine non è orientato, nel senso che il flusso può scorrere in entrambi i sensi sull'arco, allora si orienta arbitrariamente l'arco e il vincolo fra il flusso x_e e la capacità c_e è

$$-c_e \leq x_e \leq c_e$$

Se il flusso è positivo allora significa che scorre nello stesso verso dell'arco. Se è negativo allora scorre in verso opposto all'arco. Rispetto al flusso i nodi possono comportarsi da *sorgenti* o *pozzi* o *transiti*. Nelle sorgenti il

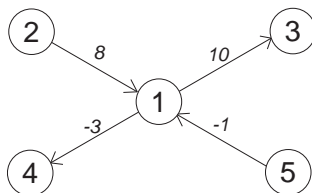


Figura 33.

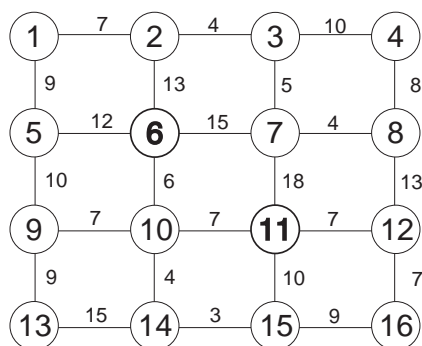


Figura 34.

flusso entrante è minore di quello uscente; nei pozzi quello entrante è maggiore di quello uscente e nei nodi di transito quello che entra nel nodo è esattamente uguale a quello che esce. In questo caso si parla anche di *conservazione del flusso*.

In base alla convenzione sul segno del flusso, un flusso è entrante in un nodo se è positivo e l'arco è orientato verso il nodo oppure se il flusso è negativo e l'arco è orientato dal nodo. Il flusso è entrante nel caso opposto. Quindi nell'esempio di Fig. 33 i flussi $x_{21} = 8$ e $x_{14} = -3$ sono entranti nel nodo 1, per un totale di 11 unità di flusso, mentre i flussi x_{13} e x_{51} sono uscenti, ancora per un totale di 11 unità di flusso, e quindi il nodo è di transito.

Un classico problema consiste nel calcolare il massimo flusso che può uscire da un nodo fissato (sorgente), attraversare la rete ed entrare in un nodo fissato (pozzo), compatibilmente con le capacità degli archi. Questo problema prende il nome di *Massimo flusso*. Si consideri ad esempio la rete di Fig. 34. Si vuole trovare il massimo flusso uscente dal nodo sorgente 6 ed entrante nel nodo pozzo 11.

Un semplice algoritmo per tale calcolo si basa sul calcolo di *cammini aumentanti*. Un cammino aumentante è un cammino dalla sorgente al pozzo sul quale il flusso può essere aumentato rispetto al flusso corrente. Per poter essere aumentato su un arco l'arco non deve essere positivamente saturo (cioè $x_e = c_e$) se l'arco è orientato come il cammino e non deve essere negativamente saturo (cioè $x_e = -c_e$) se l'arco è orientato contrariamente al cammino. L'algoritmo ricerca un cammino aumentante. Se lo trova aumenta il flusso della minima quantità aumentabile lungo il cammino. Se non lo trova, l'algoritmo termina con la soluzione corrente

corrispondente al massimo flusso. Che tale quantità corrisponda al massimo flusso verrà ora dimostrato e la dimostrazione prova anche l'uguaglianza fra la capacità di un minimo taglio e il massimo flusso.

Se esiste un cammino aumentante, la soluzione corrente non può essere il massimo flusso perché evidentemente il flusso può essere aumentato. La semplice non esistenza di un cammino aumentante di per sé non prova che la soluzione corrente debba essere il massimo flusso. Bisogna prima notare che un qualsiasi flusso dalla sorgente al pozzo deve attraversare ogni taglio che separi la sorgente dal pozzo e quindi si trova limitato dalla capacità di ogni taglio. Quindi si ha la diseuguaglianza

$$\text{massimo flusso} \leq \text{minimo taglio}$$

Ora si immagini un algoritmo che cerca un cammino aumentante tramite una cosiddetta ricerca su un grafo. Partendo dalla sorgente la ricerca va a marcare nodi raggiungibili tramite archi aumentanti. Da ogni nodo marcato si va a marcare altri nodi adiacenti tramite archi aumentanti. Se si riesce a marcare il pozzo, allora esiste un cammino aumentante. Se invece il pozzo non è stato raggiunto e l'insieme marcato di nodi è un certo insieme S , allora ogni arco (i, j) , $i \in S$ e $j \notin S$ deve essere positivamente saturo, cioè $x_e = c_e$ e ogni arco (i, j) , $i \notin S$ e $j \in S$ deve essere negativamente saturo, cioè $x_e = -c_e$. Il flusso corrente che esce dalla sorgente è tutto quello che transita sul taglio, cioè la somma dei flussi x_e per gli archi orientati da S a $V \setminus S$ e dei flussi $-x_e$ per gli archi orientati da $V \setminus S$ a V , ovvero è esattamente uguale alla somma delle capacità c_e degli archi del taglio, cioè è uguale alla capacità del taglio.

Allora, in base alla precedente diseuguaglianza, il flusso corrente non può che essere quello massimo e il taglio quello minimo e vale l'importante risultato

$$\text{massimo flusso} = \text{minimo taglio}$$

Tale algoritmo per la rete di Fig. 34 produce i passi raffigurati nelle Figure 35-(a)-(f). Il primo flusso corrente vale 0 su tutti gli archi. Il primo cammino scelto (il cammino è arbitrario purché sia aumentante) è il cammino $6 \rightarrow 7 \rightarrow 11$. Il flusso può essere aumentato di 15 e di 18 rispettivamente sui due archi. Quindi il massimo valore di cui si può aumentare il flusso è 15 (Fig. 35-(a)). L'arco $(6, 7)$ è positivamente saturo (indicato in grosso) e in tale direzione il flusso non è ulteriormente aumentabile.

Il secondo cammino è $6 \rightarrow 2 \rightarrow 1 \rightarrow 5 \rightarrow 9 \rightarrow 10 \rightarrow 11$ con valore minimo aumentabile 7 dovuto agli archi $(2, 1)$, $(9, 10)$ e $(10, 11)$ che diventano positivamente saturi. Il flusso corrente ora vale 22 ed è indicato in Fig. 35-(b).

Il terzo cammino è $6 \rightarrow 2 \rightarrow 3 \rightarrow 7 \rightarrow 11$. Sull'arco $(6, 2)$ il flusso è ancora aumentabile di 6, sull'arco $(2, 3)$ di 4, sull'arco $(3, 7)$ di 5 ma sull'arco $(7, 11)$ che ha già 15 unità di flusso e capacità 18, è aumentabile di solo 3. Quindi 3 è il flusso che si può aumentare, rendendo l'arco $(7, 11)$ positivamente saturo. La nuova soluzione di valore 25 è in Fig. 35-(c).

Il quarto cammino è $6 \rightarrow 5 \rightarrow 9 \rightarrow 13 \rightarrow 14 \rightarrow 15 \rightarrow 11$ con massimo flusso aumentabile 3 (Fig. 35-(d)) e il quinto cammino è $6 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 8 \rightarrow 12 \rightarrow 11$ con massimo valore aumentabile 1 (Fig. 35-(e)).

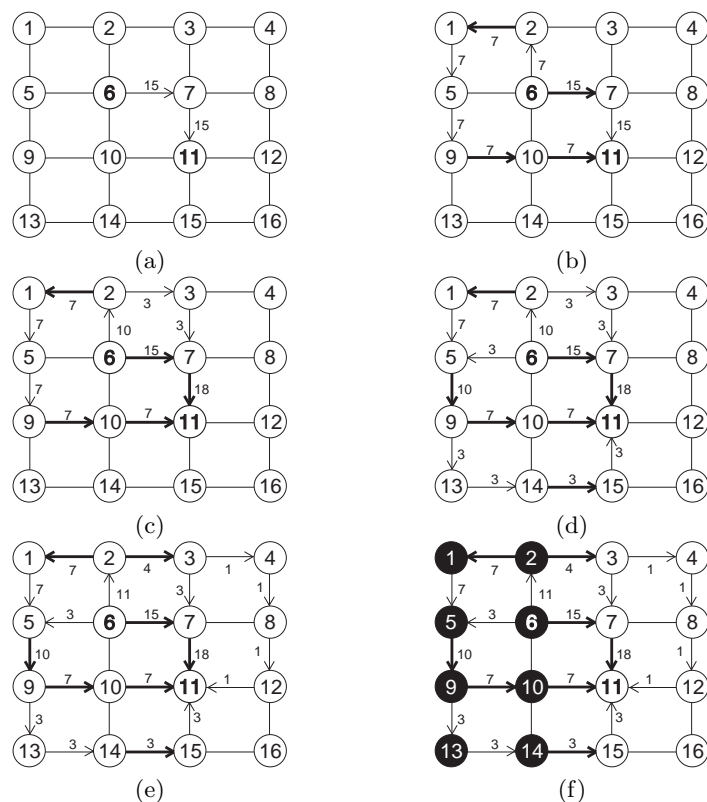


Figura 35.

A questo punto l'algoritmo non riesce a trovare cammini aumentanti. Si raggiungono solo i nodi in nero di Fig. 35-(f) (si noti che il nodo 9 viene raggiunto dal nodo 10 e non dal nodo 5 che è positivamente saturo; anche l'arco (9, 10) è positivamente saturo ma viene percorso in senso contrario (che corrisponderebbe ad una diminuzione del flusso) e quindi potrebbe essere aumentabile. Il taglio minimo è allora quello individuato dall'insieme $S = \{1, 2, 5, 6, 9, 10, 13, 14\}$ con taglio di valore 29 esattamente uguale al flusso.

35 Cammini disgiunti e tagli minimi

Quando le capacità sono uguali a 1, il flusso che si ottiene si può decomporre in un numero finito di cammini dalla sorgente al pozzo. Dati i valori di capacità i cammini non possono condividere archi, ma possono condividere nodi. Il teorema del massimo flusso e minima capacità di taglio si particularizza nel risultato che il numero di cammini distinti fra sorgente e pozzo è uguale al minimo taglio di cardinalità fra sorgente e pozzo. Si consideri la rete della Fig. 36. Vogliamo calcolare il massimo numero di cammini disgiunti negli archi fra 6 e 11. Procedendo con il medesimo algoritmo dei cammini aumentanti si perviene ai cinque cammini

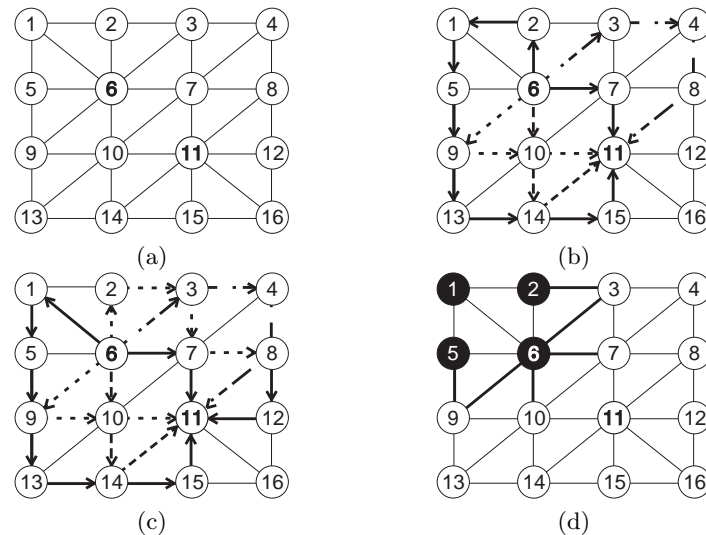


Figura 36.

indicati in Fig. 36-(b). A questo punto la ricerca di un cammino aumentante va a marcare il nodo 1. Il nodo 5 non è raggiungibile da 1 perché l'arco è saturo positivamente. Ma il nodo è raggiungibile perché l'arco viene percorso in senso contrario al flusso. Di fatto corrisponde ad una cancellazione di flusso nell'arco (1,2). Una volta raggiunto il nodo 2, l'algoritmo prosegue raggiungendo i nodi 3, 7, 8, 12 e 11, quindi trovando ancora un cammino aumentante. I sei cammini sono evidenziati in Fig. 36-(c). Da questa soluzione sono raggiungibili da 6 solo i nodi 1, 2 e 5. Questi identificano un taglio la cui capacità è 6. Quindi si tratta dell'ottimo.

Se, oltre alla non condivisione degli archi, si vuole che i cammini non condividano neppure i nodi, bisogna trasformare la rete in modo che l'attraversamento di un nodo diventi l'attraversamento di un arco. A questo fine si crea la rete $\mathcal{G} = (V' \cup V'', E')$ a partire dalla rete $G = (V, E)$ nel seguente modo: V' e V'' sono copie di V ($i' \in V'$ e $i'' \in V''$ sono copie di $i \in V$); gli archi E' sono: per ogni $i \in V$ si crea un arco (i', i'') in \mathcal{G} e per ogni arco $(i, j) \in E$ si creano due archi (i'', j') e (j'', i') in \mathcal{G} . Tutti gli archi di \mathcal{G} possono essere percorsi solo nel verso dell'orientamento dell'arco e quindi il vincolo di capacità per ogni arco è $0 \leq x_e \leq 1$. Quindi l'attraversamento dell'arco (i', i'') è equivalente all'attraversamento del nodo i nel grafo G . Tutti i cammini che passano per il nodo i in G devono passare in \mathcal{G} per l'arco (i', i'') . Si veda in Fig. 37-(a) il grafo G e in Fig.37-(b) il grafo \mathcal{G} con i tre cammini dal nodo in alto a sinistra a quello in basso a destra.

36 Tagli minimi in un grafo

Spesso viene richiesto di trovare in un grafo non orientato un taglio di capacità minima senza necessariamente specificare due nodi che vengano separati dal taglio. Oltre al problema di trovare in una rete di trasporto o di comunicazione quali collegamenti siano più critici perché ai fini della connessione del grafo, in altri tipi di

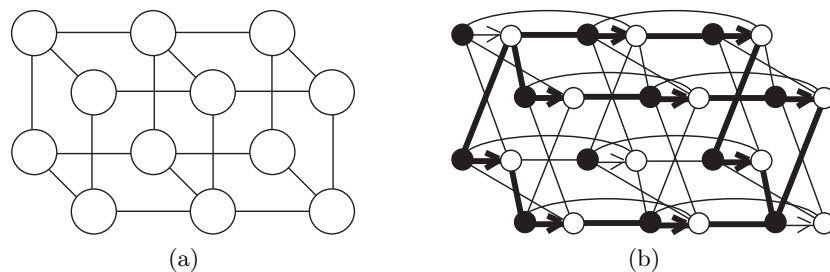


Figura 37.

problemi, dove gli archi rappresentano similarità fra oggetti, può essere invece interessante capire se si può separare il grafo in due componenti che presentino il minor numero di connessioni fra loro, in modo che le due componenti del grafo indotte dal taglio rappresentino invece comunità simili.

Un modo per calcolare il taglio di capacità minima consiste nel risolvere ripetutamente un problema di massimo flusso. Si sceglie arbitrariamente un nodo come sorgente s e poi si risolvono $(n - 1)$ problemi di massimo flusso in cui ogni volta si prende come destinazione uno degli altri nodi. Alla fine si considera il minimo dei tagli di minima capacità trovati. La complessità di questo metodo dipende dall'algoritmo di massimo flusso usato.

Vi sono anche algoritmi diretti che non usano concetti di flusso. Presentiamo due algoritmi, il secondo dei quali è stocastico e permette di trovare la soluzione solo con probabilità prefissata.

L'idea del primo algoritmo si basa sull'osservazione che, dati due nodi qualsiasi, o il taglio minimo separa i due nodi oppure non li separa. In questo secondo caso i due nodi possono essere fusi in uno fondendo eventualmente archi incidenti in entrambi i nodi e sommandone le capacità, e il taglio minimo del grafo ridotto è uguale a quello del grafo originario. Quindi, se si è in grado di trovare due nodi di cui si conosce il minimo taglio separatore, basta confrontare questo valore con quello del minimo taglio del grafo ridotto. Ricorsivamente nel grafo ridotto si determinano due nodi di cui si conosce il minimo taglio separatore e si prosegue finché il grafo è ridotto a due soli nodi. A questo punto basta confrontare tutti i tagli generati e prendere il migliore.

La parte difficile consiste ovviamente nel determinare due nodi e il minimo taglio che li separa. Naturalmente non è pensabile definire a priori i due nodi, altrimenti dovremmo risolvere un problema di massimo flusso per calcolare il minimo taglio. La procedura per generare due nodi di cui si conosca anche il minimo taglio separatore è abbastanza semplice, anche se la dimostrazione di correttezza non è banale e viene qui tralasciata. Per descrivere l'algoritmo è necessaria la seguente definizione: dati due insiemi A e B sia

$$c(A : B) := \sum_{\substack{i \in A \\ j \in B}} c_{ij}$$

L'algoritmo procede nel seguente modo: si sceglie arbitrariamente un nodo come nodo iniziale v_1 e si pone $S := \{v_1\}$. Ad esempio sia $v_1 := 1$. Poi si sceglie un nodo v_2 come

$$v_2 := \operatorname{argmax}_{k \notin S} c(S : \{k\})$$

e si aggiorna $S := S \cup \{v_2\}$. La procedura viene ripetuta fino a definire il nodo v_n . A questo punto si calcola la capacità del taglio indotto da v_n , cioè $c(\{v_n\}) =: C_n$.

Terminata questa fase, i nodi v_{n-1} e v_n vengono fusi generando il grafo G' . La procedura viene ripetuta sul grafo G' generando una successione di nodi (possibilmente diversa dalla precedente) v'_1, \dots, v'_{n-1} e una capacità di taglio $C_{n-1} := c(\{v'_{n-1}\})$.

La procedura viene ripetuta fino ad avere un grafo di due nodi in cui il valore C_2 è la capacità dell'unico arco del grafo e che corrisponde al taglio indotto da $v_1 = 1$ sul grafo originale (per costruzione il nodo 1 non viene mai fuso con altri nodi). Infine si sceglie il minimo fra i valori C_i . Questo individua il taglio di minima capacità. La complessità computazionale di questo algoritmo è $O(mn \log n)$.

Il secondo algoritmo è probabilistico e permette di trovare il taglio di capacità minima con probabilità arbitrariamente bassa.

Prima di descrivere l'algoritmo è utile dire cosa sia un algoritmo probabilistico o randomizzato o stocastico. In un algoritmo probabilistico alcune scelte sono lasciate al caso secondo un preciso meccanismo probabilistico. Questo fa sì che la soluzione desiderata sia ottenuta solo con una certa probabilità p . Se si ripete l'algoritmo la probabilità che in nessuna delle due iterazioni si sia trovata la soluzione desiderata è $(1-p)^2$. In generale la probabilità di non ottenere mai la soluzione in k iterazioni è $(1-p)^k$. Se si fissa a priori un valore ε e si vuole che $(1-p)^k \leq \varepsilon$, il numero di iterazioni deve essere almeno $\ln \varepsilon / \ln(1-p)$.

Nella maggior parte dei casi il valore di p decresce con l'aumentare dei dati del problema. Il modo come p decresce è cruciale. Ad esempio un algoritmo stocastico ingenuo potrebbe semplicemente generare a caso un sottoinsieme S . Il sottoinsieme potrebbe essere generato decidendo, elemento per elemento, se appartiene o no a S con probabilità $1/2$. In questo modo ogni sottoinsieme viene generato con probabilità 2^{-n} . Quindi in base al ragionamento appena esposto il numero di iterazioni deve essere almeno $\ln \varepsilon / \ln(1-2^{-n}) \approx 2^n \ln \varepsilon^{-1}$, dato che $(1-2^{-n}) \approx e^{-2^{-n}}$.

In questo caso l'algoritmo ha un tempo di calcolo che cresce esponenzialmente con il numero di nodi del grafo e non è quindi praticabile. Affinché l'algoritmo abbia un tempo di calcolo accettabile, la probabilità deve decrescere in modo polinomiale rispetto ai dati del problema.

L'algoritmo si basa sulla semplice osservazione che gli archi di capacità maggiore non sono probabilmente presenti nel taglio di capacità minima. Se tale affermazione fosse vera, basterebbe contrarre tali archi fondendo i relativi nodi e procedere ricorsivamente fino a rimanere con un grafo di due nodi rappresentativo del taglio di capacità minima. Naturalmente l'affermazione precedente non è vera in generale e quindi un algoritmo deterministico che procedesse nel modo prima indicato fallirebbe in molti casi. Tuttavia, se anziché scegliere sistematicamente gli archi di capacità maggiore, li scegliamo casualmente ma con probabilità proporzionale al valore di capacità, potremmo forse trovare il taglio di capacità minima. Se non lo si trova in un'iterazione,

si può sperare di trovarlo in una seconda iterazione con una scelta diversa degli archi da contrarre e così di seguito.

L'algoritmo allora consiste dei seguenti passi: un arco $\hat{e} = (u, v)$ viene scelto con probabilità proporzionale a $c_{\hat{e}}$, i nodi u e v vengono fusi, \hat{e} viene rimosso e gli archi eventualmente incidenti in u e in v vengono fusi con somma delle capacità. Se \hat{e} non appartiene al taglio minimo, il nuovo grafo ha il medesimo taglio minimo con lo stesso valore. Infatti ogni taglio del grafo originario che non contiene \hat{e} è presente anche nel grafo contratto, mentre ogni taglio del grafo originario che contiene \hat{e} non esiste nel grafo contratto. Inoltre i tagli corrispondenti hanno le medesime capacità di taglio. La procedura continua finché rimangono solo due nodi che individuano il taglio finale. Per analizzare l'algoritmo si definisca preliminarmente

$$C^* = \text{capacità del minimo taglio}, \quad C_{\Sigma} = \sum_{e \in E} c_e, \quad C_i = \sum_{e \in \delta(i)} c_e$$

Ovviamente si ha $C_i \geq C^*$ e sommando su tutti i nodi si ottiene $\sum_i C_i \geq n C^*$. Siccome $\sum_i C_i = 2 C_{\Sigma}$ si ottiene

$$\frac{C^*}{C_{\Sigma}} \leq \frac{2}{n}$$

Quindi la probabilità di non scegliere un arco del taglio minimo è almeno $1 - 2/n$. Nel grafo successivo con $(n - 1)$ nodi la probabilità diventa $1 - 2/(n - 1)$. L'ultimo grafo ha 3 nodi e la probabilità è allora $1 - 2/3$. Quindi la probabilità di non scegliere mai un arco del taglio minimo (e pertanto fornire come soluzione proprio il taglio minimo) è almeno

$$\left(1 - \frac{2}{n}\right) \left(1 - \frac{2}{n-1}\right) \left(1 - \frac{2}{n-2}\right) \dots \left(1 - \frac{2}{5}\right) \left(1 - \frac{2}{4}\right) \left(1 - \frac{2}{3}\right) = \frac{2}{n(n-1)} \geq \frac{2}{n^2}$$

e la probabilità di non trovare il taglio minimo è al più $1 - 2/n^2$. Dopo k tentativi la probabilità di non trovare mai il taglio minimo è al più

$$\left(1 - \frac{2}{n^2}\right)^k \approx e^{-2k/n^2}$$

Allora

$$e^{-2k/n^2} \leq \varepsilon \implies -\frac{2k}{n^2} \leq \ln \varepsilon \implies k \geq \frac{n^2 \ln \varepsilon^{-1}}{2}$$

Per valutare la complessità computazionale dell'algoritmo bisogna considerare quanto costa ogni calcolo di un taglio. Anche se la descrizione dell'algoritmo fa uso dell'idea di contrarre il grafo ad ogni scelta di un arco, sarebbe troppo costoso in pratica riscrivere il grafo. Il grafo viene allora mantenuto e ogni fusione di nodi viene realizzata con una struttura Union-Find con complessità $O(\log n)$.

La scelta casuale di un arco richiede una ricerca binaria fra gli archi ancora da scegliere. In generale la scelta casuale di un elemento $e \in E$ con probabilità proporzionale a c_e e sua eventuale rimozione da E si può effettuare efficientemente usando un albero binario. Inizialmente si crea un albero binario (pieno e bilanciato con l'ultima riga piena a 'sinistra', come uno heap) in cui le foglie sono in corrispondenza con gli elementi

di E . Ad ogni foglia si associa il valore c_e . Ad ogni altro nodo dell'albero si associa la somma dei valori c_e delle foglie del sottoalbero relativo al nodo. La scelta di una foglia viene effettuata partendo dal nodo radice e scegliendo il ramo di sinistra o di destra dell'albero con probabilità proporzionale ai valori associati ai due figli. Procedendo ricorsivamente si perviene alla foglia. Per rimuovere la foglia e basta porre $c_e = 0$ e diminuire di c_e tutti i valori associati ai nodi intermedi risalendo nell'albero fino alla radice. Questo calcolo ha complessità $O(\log m) = O(\log n)$.

Si noti che si potrebbe anche scegliere un arco non più esistente nel grafo contratto. In questo caso si riconosce che l'arco 'non esiste' più per il fatto che i due nodi appartengono alla stessa componente connessa della struttura Union-Find. Se un arco nel grafo contratto proviene da più archi e quindi la sua capacità sarebbe la somma delle capacità originarie, in ogni caso la scelta casuale di questo pseudoarco è equivalente alla scelta di uno qualsiasi degli archi originali.

Il numero di iterazioni per la ricerca di un taglio è al più pari a $m - 1$ e quindi un taglio si trova con complessità $O(m \log n)$. Complessivamente, per avere il taglio minimo con probabilità $1 - \varepsilon$, la complessità è $O(m n^2 \log n \log \varepsilon^{-1})$.

37 Tagli e autovalori

Dato un sottoinsieme di nodi $S \subset V$, si definisca

$$x_i := \begin{cases} 1 & \text{se } i \in S \\ -1 & \text{se } i \notin S \end{cases}$$

Si è visto che la cardinalità del taglio indotto da S si può esprimere come in (17), che qui si ripete,

$$|\delta(S)| = \frac{1}{4} \sum_{i,j} L_{ij} x_i x_j \quad (52)$$

dove L è la matrice Laplaciana. Si aggiunga ora il vincolo di avere un numero fissato di nodi n_1 nell'insieme S e n_2 in $V \setminus S$. Allora il calcolo del taglio di minima cardinalità con il vincolo aggiuntivo può essere espresso come

$$\begin{aligned} \min \quad & \sum_{i,j} L_{ij} x_i x_j \\ & \sum_i x_i = n_1 - n_2 \\ & x_i \in \{-1, +1\} \end{aligned} \quad (53)$$

mentre per il taglio di massima cardinalità si ha lo stesso problema con l'obiettivo da massimizzare invece che da minimizzare. Il calcolo di (53) non è facile (si tratta di problemi NP-difficili, anche il taglio minimo con il vincolo sui nodi è NP-difficile). Allora si opera un rilassamento del vincolo $x_i \in \{-1, +1\}$ e si pone $\sum_i x_i^2 = n$, cioè basta che il vettore x stia sulla sfera di raggio \sqrt{n} , alla quale comunque appartengono anche

tutti i punti di coordinate $+1$ o -1 . Allora il rilassamento di (53) è

$$\begin{aligned} \min \quad & \sum_{i,j} L_{ij} x_i x_j \\ & \sum_i x_i^2 = n \\ & \sum_i x_i = n_1 - n_2 \end{aligned} \quad (54)$$

Il minimo (o il massimo) è uno dei punti di stazionarietà e il calcolo dei punti di stazionarietà della funzione obiettivo vincolata di (54) può essere effettuato tramite la funzione lagrangiana:

$$\sum_{i,j} L_{ij} x_i x_j + \lambda (n - \sum_i x_i^2) + 2\mu (n_1 - n_2 - \sum_i x_i)$$

e ponendo a zero la sua derivata

$$\frac{\partial}{\partial x_k} \sum_{i,j} L_{ij} x_i x_j + \lambda (n - \sum_i x_i^2) + 2\mu (n_1 - n_2 - \sum_i x_i) = 2 \sum_j L_{kj} x_j - 2\lambda x_k - 2\mu = 0$$

che dà

$$\sum_j L_{kj} x_j = \lambda x_k + \mu \implies Lx = \lambda x + \mu \mathbf{1}$$

Siccome $L\mathbf{1} = 0$ possiamo scrivere

$$L(x + \frac{\mu}{\lambda} \mathbf{1}) = \lambda(x + \frac{\mu}{\lambda} \mathbf{1}) \quad (55)$$

Quindi

$$u = x + \frac{\mu}{\lambda} \mathbf{1}$$

è un autovettore di L con autovalore λ . Questa proprietà permette di calcolare u e λ , anche se rimane da definire quale autovalore e quale autovettore scegliere e come normalizzare u (gli autovettori sono definiti a meno di una costante moltiplicativa).

Bisogna precisare che se il grafo è connesso c'è un unico autovalore nullo con autovettore $\mathbf{1}$. Però se il grafo consiste di più componenti connesse l'autovalore nullo è presente con molteplicità pari al numero di componenti connesse e per ognuno di questi autovalori ci sono autovettori diversi associati con tutti i valori nulli tranne quelli relativi ad una particolare componente connessa dove sono 1.

Se il grafo è connesso si può escludere l'autovettore $\mathbf{1}$ al quale è associato l'autovalore nullo. Infatti se si avesse $x = \alpha \mathbf{1}$ (cioè tutte le componenti uguali a α), avremmo, da $x^\top x = n$

$$\alpha^2 \mathbf{1}^\top \mathbf{1} = n \implies \alpha^2 = \pm 1$$

e da $\mathbf{1}^T x = n_1 - n_2$

$$\mathbf{1}^T x = \alpha n = (n_1 - n_2) \implies \alpha = \frac{n_1 - n_2}{n}$$

e i due risultati sono incompatibili. Per il calcolo di μ si noti che $\mathbf{1}^\top u = 0$ (autovettori diversi sono ortogonali essendo L simmetrica) e quindi, sfruttando anche il vincolo $\sum_i x_i = n_1 - n_2$,

$$0 = \mathbf{1}^\top u = \mathbf{1}^\top \left(x + \frac{\mu}{\lambda} \mathbf{1} \right) = \sum_i x_i + \frac{\mu}{\lambda} n = n_1 - n_2 + \frac{\mu}{\lambda} n$$

da cui

$$\mu = -\frac{n_1 - n_2}{n} \lambda$$

e allora

$$u = x - \frac{n_1 - n_2}{n} \mathbf{1}$$

Per trovare la corretta normalizzazione di u si sfrutta il vincolo $x^\top x = n$:

$$\begin{aligned} u^\top u &= \left(x - \frac{n_1 - n_2}{n} \mathbf{1} \right)^\top \left(x - \frac{n_1 - n_2}{n} \mathbf{1} \right) = x^\top x - 2 \frac{n_1 - n_2}{n} x^\top \mathbf{1} + \frac{(n_1 - n_2)^2}{n^2} \mathbf{1}^\top \mathbf{1} = \\ &= n - 2 \frac{(n_1 - n_2)^2}{n} + \frac{(n_1 - n_2)^2}{n^2} n = \frac{n^2 - (n_1 - n_2)^2}{n} = \\ &= \frac{(n - n_1 + n_2)(n + n_1 - n_2)}{n} = 4 \frac{n_1 n_2}{n} \end{aligned}$$

La funzione obiettivo di (53) si può allora esprimere come

$$x^\top L x = \left(u + \frac{n_1 - n_2}{n} \mathbf{1} \right)^\top L \left(u + \frac{n_1 - n_2}{n} \mathbf{1} \right) = u^\top L u = \lambda u^\top u = 4 \lambda \frac{n_1 n_2}{n}$$

per cui per ottenere il taglio minimo bisogna scegliere l'autovalore minimo (escludendo il caso $\lambda_1 = 0$ che corrisponde all'autovalore $\mathbf{1}$) e per ottenere il taglio massimo bisogna scegliere l'autovalore massimo. Essendo (54) un rilassamento di (53) e indicando con $M(n_1, n_2)$ il valore del massimo taglio con $|S| = n_1$ e con $m(n_1, n_2)$ il valore del minimo taglio:

$$M(n_1, n_2) \leq \lambda_n^L \frac{n_1 n_2}{n}, \quad m(n_1, n_2) \geq \lambda_2^L \frac{n_1 n_2}{n} \quad (56)$$

Riassumendo, x punto stazionario viene calcolato da un autovettore u di L , escludendo però l'autovettore $\mathbf{1}$ perché non ammissibile per il vincolo $\sum_i x_i = n_1 - n_2$. L'autovettore u viene scelto in base al più piccolo autovalore positivo in caso di minimizzazione e al più grande autovalore in caso di massimizzazione e viene poi normalizzato in modo che $u^\top u = 4 n_1 n_2 / n$. A questo punto $x = u + (n_1 - n_2) / n \mathbf{1}$.

Per ottenere il taglio del grafo bisogna però ottenere una soluzione $\hat{x}_i \in \{-1, 1\}$, tale che $\sum_i \hat{x}_i = n_1 - n_2$, a partire da x . La cosa più semplice è trovare quel punto \hat{x} di coordinate ± 1 che minimizza la distanza da x :

$$\min_{\substack{\hat{x}_i \in \{-1, 1\} \\ \sum \hat{x}_i = n_1 - n_2}} \sum_i (\hat{x}_i - x_i)^2$$

Il minimo si ottiene ordinando in modo decrescente i valori x_i e ponendo $\hat{x}_i = 1$ per i primi n_1 valori. Si noti che sia u che $-u$ sono normalizzazioni ammissibili per l'autovettore. Quindi è presente anche una

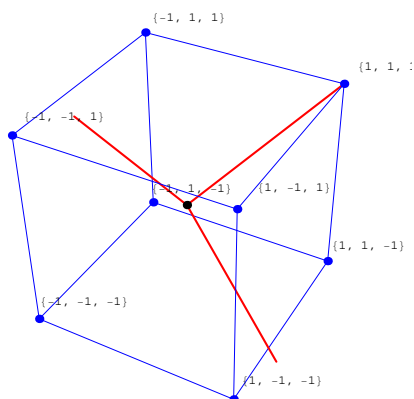


Figura 38.

soluzione alternativa che corrisponde all'ordine opposto per i valori x_i . Delle due soluzioni si prende la migliore. Siccome l'ordine degli elementi di x è lo stesso dell'autovettore u , si può semplicemente partire da u senza effettivamente calcolare x .

Questo metodo di calcolo può essere applicato anche quando siano definite capacità d'arco e si voglia calcolare il taglio di minima o massima capacità. In questo caso si usa la formula (19) di pag. 43 con il Laplaciano esteso. Ad esempio si consideri il grafo completo di 3 nodi con capacità $c_{12} = 8$, $c_{13} = 5$, $c_{23} = 2$. Il Laplaciano è

$$L = \begin{pmatrix} 13 & -8 & -5 \\ -8 & 10 & -2 \\ -5 & -2 & 7 \end{pmatrix}$$

con autovettori e autovalori

$$\lambda_1^L = 0, \quad u^1 = (1, 1, 1), \quad \lambda_2^L = 9.80, \quad u^2 = (-0.37, -1., 1.36), \quad \lambda_3^L = 20.19, \quad u^3 = (1.36, -1., -0.37)$$

Si vedano in Fig. 38 i tre autovettori.

Quindi l'autovalore che interessa per il minimo è $\lambda_2^L = 9.80$ e per il massimo $\lambda_3^L = 20.19$. Il grafico della funzione $x^\top L x$ sulla superficie della sfera viene rappresentato in Fig. 39-(a) dove sono state adottate come coordinate della superficie serica la longitudine (ascisse, da 0 a 180 gradi in quanto da 180 a 360 la funzione si ripete) e la latitudine (ordinate, da -90 a +90 gradi), dopo aver scelto come asse della sfera l'autovettore $(1, 1, 1)$. I poli degenerano nei lati alto e basso della figura. Per la scelta fatta sull'asse della sfera il minimo (non vincolato) di $x^\top L x$ si ha ai poli (i valori sono crescenti dal nero al bianco) e il massimo (non vincolato) sull'equatore nel punto (in rosso) di longitudine 75 gradi. Questo punto corrisponde al terzo autovettore. Il punto corrispondente al secondo autovalore è il punto sull'equatore (in rosso) di longitudine 165 gradi. I vertici del cubo (tranne i due vertici $\pm(1, 1, 1)$) sono indicati in blu.

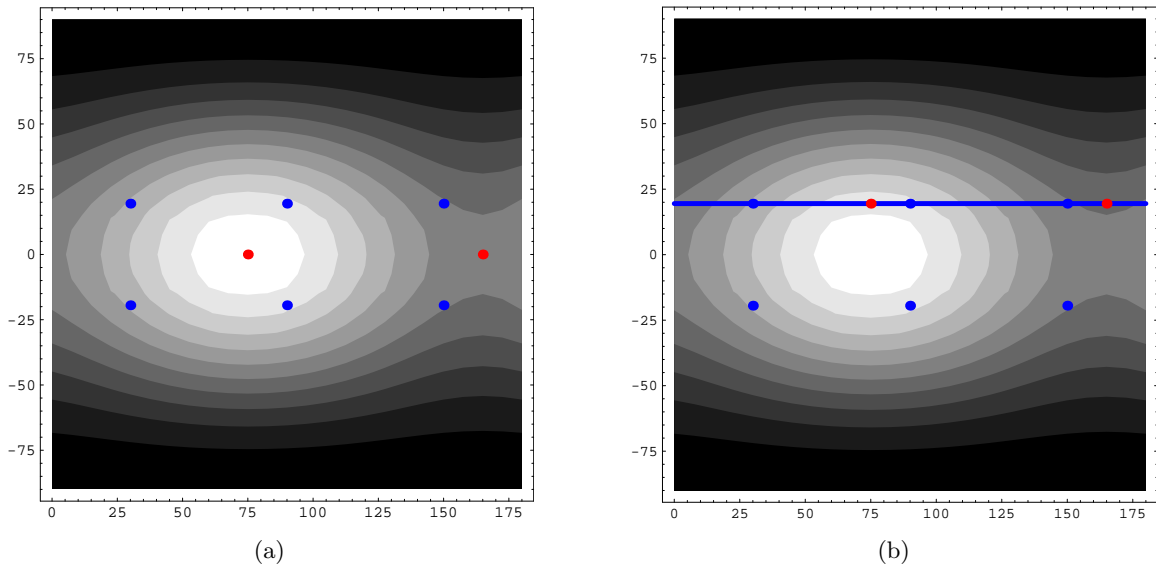


Figura 39.

Se ora aggiungiamo il vincolo $\sum_i x_i = n_1 - n_2$, i punti ammissibili sono diventati una circonferenza sulla superficie sferica. In particolare con la scelta dell'asse con il vettore $(1, 1, 1)$ tale circonferenza è un parallelo. Se nell'esempio poniamo $n_1 = 2$ e $n_2 = 1$, l'equazione $x_1 + x_2 + x_3 = 2 - 1 = 1$ diventa il parallelo di latitudine 19.47 gradi (si veda la Fig. 39-(b)). I due autovettori normalizzati sono

$$u^2 = (-0.34, -0.94, 1.29), \quad u^3 = (1.29, -0.94, -0.34)$$

da cui

$$x^2 = (-0.01, -0.61, 1.62), \quad x^3 = (1.62, -0.61, -0.01)$$

I due punti x^2 e x^3 devono stare sul parallelo.

Si noti che (54) è effettivamente un rilassamento di (53). Infatti si ottiene

$$\lambda_2^L \frac{n_1 n_2}{n} = 6.5359 < 7, \quad \lambda_3^L \frac{n_1 n_2}{n} = 13.4641 > 13$$

Come esempio più significativo si consideri il grafo in Fig. 40 di cui vogliamo conoscere il minimo taglio bilanciato (cioè $n_1 = n_2 = 6$). Il Laplaciano del grafo è

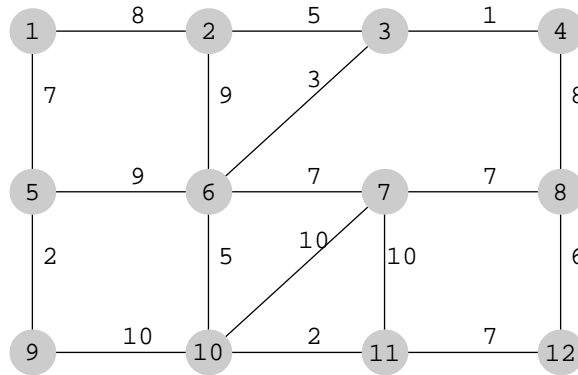


Figura 40.

$$L = \begin{pmatrix} 15 & -8 & 0 & 0 & -7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -8 & 22 & -5 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -5 & 9 & -1 & 0 & -3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 9 & 0 & 0 & 0 & -8 & 0 & 0 & 0 & 0 \\ -7 & 0 & 0 & 0 & 0 & 18 & -9 & 0 & 0 & -2 & 0 & 0 \\ 0 & -9 & -3 & 0 & -9 & 33 & -7 & 0 & 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -7 & 34 & -7 & 0 & -10 & -10 & 0 \\ 0 & 0 & 0 & -8 & 0 & 0 & -7 & 21 & 0 & 0 & 0 & -6 \\ 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 12 & -10 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -5 & -10 & 0 & -10 & 27 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -10 & 0 & 0 & -2 & 19 & -7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -6 & 0 & 0 & -7 & 13 \end{pmatrix}$$

con autovettore u^2 corrispondente al più piccolo autovalore positivo $\lambda_2^L = 2.899$

$$u^2 = (0.388, 0.325, 0.288, -0.428, 0.298, 0.186, -0.120, -0.362, 0.057, -0.007, -0.242, -0.383)$$

il cui ordinamento degli indici è $(1, 2, 5, 3, 6, 9, 10, 7, 11, 8, 12, 4)$. In base alla regola $S = \{1, 2, 5, 3, 6, 9\}$ il cui taglio ha valore 23. Non si tratta dell'ottimo che è invece $S = \{1, 2, 3, 4, 5, 6\}$ di valore 22. L'autovettore u^{12} corrispondente al massimo autovalore $\lambda_{12}^L = 47.6816$ è

$$u^{12} = (-0.082, 0.192, 0.011, 0.046, 0.161, -0.482, 0.718, -0.222, 0.056, -0.2323, -0.256, 0.090)$$

con ordinamento degli indici $(7, 2, 5, 12, 9, 4, 3, 1, 8, 10, 11, 6)$.

Se si deve risolvere il problema del massimo taglio senza imposizioni sul numero dei nodi, si può pensare di adottare il metodo precedentemente descritto per tutti i possibili valori di n_1 da 1 a $n - 1$. Algoritmicamente questo significa provare tutti gli $n - 1$ tagli che si ottengono dall'ordinamento dei nodi indotto dall'ordine dell'autovettore u . Nell'esempio si trova il taglio indotto da $S = \{7, 2, 5, 12, 9, 4\}$ di valore 104.

Da (56) si ha per il massimo taglio M

$$M = \max_{n_1, n_2} M(n_1, n_2) \leq \max_{n_1, n_2} \lambda_n^L \frac{n_1 n_2}{n} =$$

$$\begin{aligned}
(\text{se } n \text{ pari}) &= \lambda_n^L \frac{\frac{n}{2} \frac{n}{2}}{n} = \lambda_n^L \frac{n}{4} \\
(\text{se } n \text{ dispari}) &= \lambda_n^L \frac{\frac{n-1}{2} \frac{n+1}{2}}{n} = \lambda_n^L \frac{(n-1)^2}{4n}
\end{aligned}$$

Si noti che la limitazione è esatta per un circuito pari (in cui $\lambda_n^L = 4$) e per un grafo completo (in cui $\lambda_n^L = n$). Analogamente si ha da (56) per il taglio minimo m

$$m = \min_{n_1, n_2} m(n_1, n_2) \geq \min_{n_1, n_2} \lambda_2^L \frac{n_1 n_2}{n} = \lambda_2^L \frac{n-1}{n}$$

che fornisce la relazione fra la connettività algebrica e quella di arco di un grafo (si confronti con la connettività di vertice a pag. 47)

$$\lambda_2^L \leq \frac{n}{n-1} m$$

relazione valida anche per un grafo completo, mentre si può dimostrare che, per grafi non completi, vale il più forte risultato

$$\lambda_2^L \leq \kappa \leq m.$$

38 Partizione dei vertici in sottoinsiemi (clustering): introduzione

Una delle operazioni più frequenti su insiemi di oggetti consiste nel dare una definizione di similarità fra gli oggetti e poi nel ripartirli in classi di oggetti (clusters) il più possibile simili fra loro.

Se la definizione di similarità è puramente binaria (due oggetti sono simili oppure non lo sono) allora l'insieme è un grafo e si tratta di ripartire i nodi in sottoinsiemi il più possibile densi.

È naturalmente necessario definire delle misure che permettano di valutare una partizione. Varie misure sono state proposte in letteratura. Dato un sottoinsieme $S \subset V$ una possibile misura di *densità* potrebbe essere data dal rapporto fra gli archi presenti nel cluster diviso il massimo numero possibile di archi del cluster (per un nodo isolato la densità viene convenzionalmente considerata uguale a zero)

$$\frac{2|E(S)|}{|S|(|S|-1)}$$

Quindi, data una partizione (S_1, \dots, S_q) , questa può essere misurata valutando la densità media dei cluster

$$\frac{1}{q} \sum_{i=1}^q \frac{2|E(S_i)|}{|S_i|(|S_i|-1)}$$

Si noti però che per un sottoinsieme S la massima densità possibile è uguale ad uno e si ottiene quando il sottoinsieme è una cricca. Quindi si ottiene una densità media uguale ad uno (e massima possibile) semplicemente prendendo un accoppiamento perfetto nel grafo (se ne esiste uno). Non è quindi molto significativa

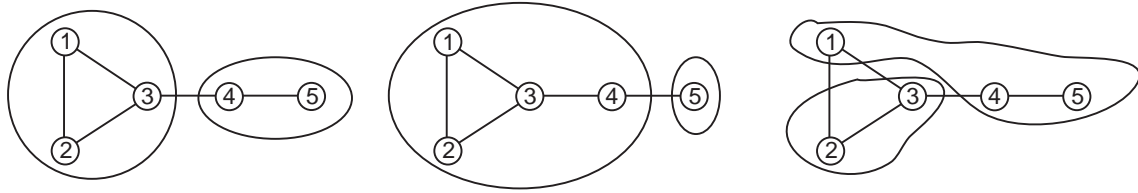


Figura 41.

una misura basata sulla densità come definita sopra. Una definizione alternativa di densità si basa invece sul rapporto fra archi e nodi dell'insieme S , cioè $|E(S)|/|S|$, per cui la densità media è

$$\frac{1}{q} \sum_{i=1}^q \frac{|E(S_i)|}{|S_i|}$$

Questa definizione valuta in modo diverso cricche con diversa cardinalità. La densità di K_n risulta essere $(n-1)/2$.

Si considerino le tre partizioni del grafo in Fig. 41. Con la densità basata sul numero di archi le tre partizioni hanno valore rispettivamente: 1 (si noti che è una partizione in cricche), $1/2(4/6 + 0) = 1/3$ e $1/2(1/3 + 1) = 2/3$. Con la densità basata sul numero di nodi si ottiene: $1/2(3/3 + 1/2) = 3/4$, $1/2(4/4 + 0) = 1/2$ e $1/2(1/3 + 1/2) = 5/12$. La prima partizione è la migliore in entrambi i casi, ma le altre due sono valutate in modo diverso.

Il criterio di similarità può essere più complesso di una semplice relazione binaria. Ad esempio per ogni coppia di oggetti (i, j) potrebbe essere definito un valore di similarità $0 \leq w_{ij} \leq 1$ con l'idea che il valore 1 rappresenta similarità piena e 0 nessuna similarità. Si noti che normalmente si ha $w_{ii} = 1$ e quindi le misure di densità includono anche questi termini, che precedentemente, con il semplice conto del numero di archi, non intervenivano. La misura di densità di un insieme S diventa allora

$$\frac{\sum_{i \in S} \sum_{j \in S} w_{ij}}{2|S|}$$

I valori di similarità possono dipendere dalla particolare situazione reale che viene modellata, ma possono anche essere derivati dal grafo stesso in modo più complicato della semplice relazione di adiacenza. Se la distanza fra due nodi è d_{ij} (misurata ad esempio come numero di archi) si può porre $w_{ij} = 1/d_{ij}$. Nodi adiacenti hanno similarità $w_{ij} = 1$ e per convenzione possiamo porre $w_{ii} = 1$. Per il grafo di Fig. 41 si ha la seguente tabella di valori di similarità

$$w = \begin{pmatrix} 1 & 1 & 1 & \frac{1}{2} & \frac{1}{3} \\ 1 & 1 & 1 & \frac{1}{2} & \frac{1}{3} \\ 1 & 1 & 1 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 & 1 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{2} & 1 & 1 \end{pmatrix}$$

che porta alla seguente valutazione delle tre partizioni:

$$\frac{1}{2} \left(\frac{9}{3} + \frac{4}{2} \right) = \frac{5}{2} = 2.5, \quad \frac{1}{2} \left(\frac{14}{4} + \frac{1}{1} \right) = \frac{9}{4} = 2.25, \quad \frac{1}{2} \left(\frac{6+2/3}{3} + \frac{4}{2} \right) = \frac{19}{9} = 2.111\dots$$

Un'altra misura di similarità proposta si basa sull'indice di Jaccard. Per ogni nodo i sia $N(i)$ l'insieme dei nodi adiacenti a i . Allora il valore di similarità è dato dall'indice di Jaccard dei due insiemi $N(i)$ e $N(j)$, ovvero

$$w_{ij} = \frac{|N(i) \cap N(j)|}{|N(i) \cup N(j)|}$$

che porta alla seguente tabella di valori di similarità per il grafo di Fig. 41

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

Con questa misura le tre partizioni sono valutate rispettivamente

$$\frac{23}{18} = 1.278, \quad \frac{11}{8} = 1.375, \quad \frac{89}{72} = 1.236$$

La seconda partizione è migliore della prima in questo caso. Con questa definizione si è trascurata l'informazione data dall'adiacenza di due nodi. Volendo recuperare questa informazione si può ridefinire l'intorno di un nodo i aggiungendo anche il nodo i stesso. Operando in questo modo si ottiene una diversa tabella di similarità

$$w = \begin{pmatrix} 1 & 1 & \frac{3}{4} & \frac{1}{5} & 0 \\ 1 & 1 & \frac{3}{4} & \frac{1}{5} & 0 \\ \frac{3}{4} & \frac{3}{4} & 1 & \frac{2}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{2}{5} & 1 & \frac{2}{3} \\ 0 & 0 & \frac{1}{5} & \frac{2}{3} & 1 \end{pmatrix}$$

(si noti che i nodi 1 e 2 che stanno sulla medesima orbita di automorfismo, adesso risultano più simili fra loro che con la precedente tabella). Con questi valori le tre partizioni hanno valore

$$\frac{13}{6} = 2.167, \quad \frac{73}{40} = 1.825, \quad \frac{599}{360} = 1.664$$

Un'interessante visualizzazione della partizione si ha dalla matrice di adiacenza una volta che si siano riordinati i nodi secondo la partizione. Si vedano le successive Fig. 47.

39 Modularità con numero di archi fissato – Partizione in due insiemi

Una soddisfacente misura di una partizione è la *modularità*. La misura è simile alla densità basata sul numero di archi però è differente in quanto non si confronta il numero di archi esistenti con il massimo

numero possibile ma lo si confronta con il numero di archi che sarebbero presenti se il grafo fosse casuale. Inoltre si misura una differenza anziché un rapporto. Quindi si tratta di valutare se un insieme di nodi ha un numero di archi superiore a quello che sarebbe normale in una situazione generica. Questo tipo di misura prende il nome di modularità.

Quando si parla di ‘numero atteso di archi’, bisogna sapere quali ipotesi sono state fatte sull’insieme di grafi casuali. Possiamo ad esempio assumere che sia costante il numero totale di archi al variare del grafo in esame. Se è m il numero totale di archi, la probabilità di avere i nodi i e j connessi da un arco è

$$p = \frac{2m}{n(n-1)}$$

Se assumiamo che un grafo casuale venga generato generando in modo stocasticamente indipendente gli archi (si tratta di un’ipotesi forte che può valere solo per certe classi di grafi casuali), allora, dato un insieme $S \subset V$ il numero di archi atteso in S è

$$p \frac{|S|(|S|-1)}{2}$$

mentre il numero di archi effettivamente presente in S è $|E(S)|$. Analogamente si ha per l’insieme complementare $\bar{S} = V \setminus S$. Allora la differenza fra gli archi presenti e quelli attesi per la partizione (S, \bar{S}) è data da

$$\begin{aligned} Q &= |E(S)| - p \frac{|S|(|S|-1)}{2} + |E(\bar{S})| - p \frac{|\bar{S}|(|\bar{S}|-1)}{2} & (57) \\ |E(S)| + |E(\bar{S})| - \frac{p}{2} (|S|^2 + |\bar{S}|^2 - |S| - |\bar{S}|) &= \\ m - |\delta(S)| - \frac{p}{2} (|S|^2 + |\bar{S}|^2 - n) &= \\ m + \frac{pn}{2} - (|\delta(S)| + \frac{p}{2} (|S|^2 + |\bar{S}|^2)) & \end{aligned}$$

Tale valore normalizzato al numero totale di archi è la modularità Q della partizione (S, \bar{S}) .

$$Q = \frac{n}{n-1} - \frac{1}{m} (|\delta(S)| + \frac{p}{2} (|S|^2 + |\bar{S}|^2)) \quad (58)$$

Si consideri nuovamente il semplice grafo di Fig. 41 dove $m = 5$, $n = 5$ e $p = 1/2$. Data la struttura del grafo i nodi 1, 2 e 3 formano un cluster ben definito (sono addirittura una clique) e altrettanto per i due nodi 4 e 5. Se si considera l’insieme $S = \{1, 2, 3\}$ si contano 3 archi effettivi mentre ci sono $3p = 1.5$ archi attesi. Quindi la differenza è 1.5. Per l’insieme complementare $\bar{S} = \{4, 5\}$ si ha un arco effettivo meno 0.5 archi attesi. Sommando sui due insiemi si ottiene 2, che, normalizzato, fornisce il valore di modularità $2/5 = 0.4$. Applicando la formula (58) si ottiene

$$Q = \frac{5}{4} - \frac{1}{5} (1 + \frac{1}{4} (9 + 4)) = \frac{5}{4} - \frac{1}{5} \frac{17}{4} = \frac{2}{5}$$

Se si considera la seconda partizione del grafo in Fig. 41 dove $S = \{1, 2, 3, 4\}$ e $\bar{S} = \{5\}$, si hanno in S 4 archi effettivi contro $6 \cdot 1/2 = 3$ archi attesi, mentre in \bar{S} si hanno zero archi effettivi e zero archi attesi.

La modularità è quindi $1/5 = 0.2$. Anche se il minimo taglio è il medesimo, tuttavia la prima partizione è migliore ed effettivamente il valore di modularità è più alto nel primo caso.

Infine si consideri la terza partizione in Fig. 41 con $S = \{1, 4, 5\}$, $\bar{S} = \{2, 3\}$. Si ottiene $Q = 0$. Quindi la definizione data di modularità fornisce proprio la misura che ci aspettiamo dal grafo.

Siccome vogliamo trovare la partizione che massimizza la modularità, massimizzare (57) è equivalente a minimizzare

$$|\delta(S)| + \frac{p}{2} (|S|^2 + |\bar{S}|^2) \quad (59)$$

Abbiamo visto che si ha

$$|\delta(S)| = \frac{1}{4} x^\top L x$$

dove, al solito, L è il Laplaciano di G e

$$x_i = \begin{cases} +1 & i \in S \\ -1 & i \notin S \end{cases}$$

Inoltre

$$|S|^2 = \left(\frac{\mathbf{1}^\top (x + \mathbf{1})}{2} \right)^2 = \frac{1}{4} (x + \mathbf{1})^\top \mathbf{1} \mathbf{1}^\top (x + \mathbf{1}) = \frac{1}{4} (x^\top \mathbf{1} \mathbf{1}^\top x + 2n \mathbf{1}^\top x + n^2)$$

e analogamente

$$|\bar{S}|^2 = \frac{1}{4} (x^\top \mathbf{1} \mathbf{1}^\top x - 2n \mathbf{1}^\top x + n^2)$$

per cui

$$|S|^2 + |\bar{S}|^2 = \frac{1}{2} (x^\top \mathbf{1} \mathbf{1}^\top x + n^2)$$

Allora minimizzare (59) è equivalente a minimizzare

$$x^\top L x + p x^\top \mathbf{1} \mathbf{1}^\top x = x^\top (L + p \mathbf{1} \mathbf{1}^\top) x \quad (60)$$

Si indichi

$$\tilde{L} = L + p \mathbf{1} \mathbf{1}^\top$$

Si noti che

$$\tilde{L} \mathbf{1} = L \mathbf{1} + p \mathbf{1} \mathbf{1}^\top \mathbf{1} = \mathbf{0} + p n \mathbf{1}$$

Quindi $\mathbf{1}$ è autovettore di \tilde{L} con autovalore $p n$. Se u è un altro autovettore di L con autovalore λ si ha (si ricordi che $\mathbf{1}^\top u = 0$)

$$\tilde{L} u = L u + p \mathbf{1} \mathbf{1}^\top u = \lambda u$$

Quindi \tilde{L} eredita da L tutti gli altri autovalori ed autovettori.

Per minimizzare (60) si procede come nella Sez. 37. Si risolve il problema rilassato

$$\begin{aligned} \min \quad & x^\top \tilde{L} x \\ & x^\top x = n \end{aligned}$$

Dal Lagrangiano

$$x^\top \tilde{L} x + \mu (n - x^\top x)$$

si ottiene, derivando e uguagliando a zero,

$$\tilde{L}x = \mu x$$

Quindi x deve essere un autovettore e μ deve essere un autovalore. Siccome

$$x^\top \tilde{L}x = \mu x^\top x = \mu n$$

μ deve essere uguale al più piccolo autovalore di \tilde{L} . Siccome

$$pn = \frac{2m}{n-1} > \frac{2m}{n} \geq \min_i g_i \geq \kappa \geq \lambda_2^L$$

(dove $2m/n$ è il grado medio e l'ultima diseuguaglianza è la nota relazione fra connettività algebrica e connettività di vertice (si riveda a pag. 19 e 47)), il più piccolo autovalore è λ_2^L e quindi

$$\mu = \lambda_2^L$$

Si noti che si ottiene lo stesso risultato come se si minimizzasse il taglio minimo. La differenza è che qui non si è imposto a priori il vincolo sulla cardinalità di S , ma questa viene invece automaticamente determinata dal numero di componenti di u con lo stesso segno.

Si vede, dai calcoli fatti, che il valore di modularità del problema rilassato è

$$\frac{n}{n-1} - \frac{n}{4m} (\lambda_2^L + pn)$$

Si riconsideri il grafo di Fig. 41. Si trova

$$\lambda_2^L = 0.519, \quad x = \{-0.938, -0.938, -0.451, 0.756, 1.571\}$$

da cui

$$\frac{n}{n-1} - \frac{n}{4m} (\lambda_2^L + pn) = 0.495299$$

ovviamente superiore al valore 0.4 dato che si è risolto un problema rilassato.

40 Modularità con gradi fissati – Partizione in due insiemi

Se si vuole operare un'analisi del numero di archi attesi in un insieme di nodi fissando anche i gradi nei nodi bisogna tener conto di come può essere generato un grafo casuale con una sequenza dei gradi fissata. Un metodo, basato sulle catene di Markov, per generare correttamente un grafo con probabilità uniforme e sequenza dei gradi fissata verrà esposto più avanti. Ora vorremmo ottenere una formula, il più possibile semplice, che ci dia la probabilità di avere l'arco (i, j) in un grafo casuale, conoscendo i gradi g_i e g_j ,

Se il grafo è k -regolare è abbastanza facile trovare la formula. In questo caso $g_i = k$ per ogni nodo ed $m = nk/2$. Immaginiamo di generare tutti i possibili grafi k -regolari e sia N il loro numero. Vogliamo valutare in quanti grafi è presente l'arco (i, j) . Questo numero diviso per N dà la probabilità che l'arco sia

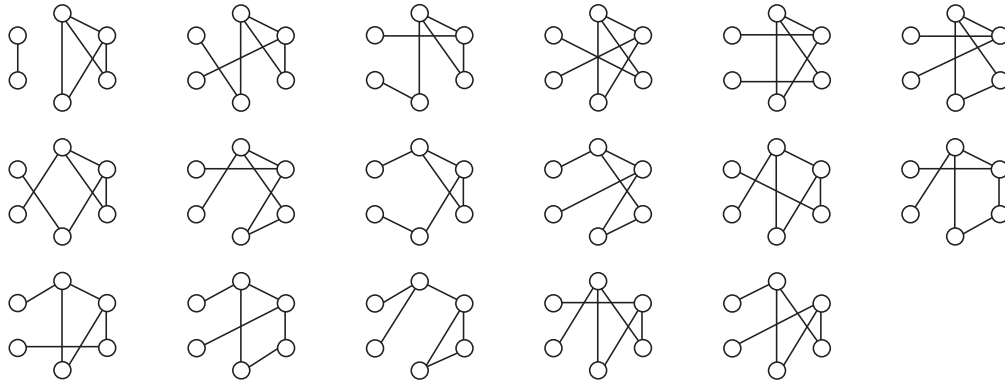


Figura 42.

presente (considerando, come è naturale, che tutti i grafi siano equiprobabili). La simmetria di un grafo regolare permette di calcolare facilmente quanti sono i grafi in cui compare un fissato arco (i, j) . Ogni grafo ha $m = nk/2$ archi e quindi su tutti gli N grafi si contano $Nnk/2$ archi. Per la simmetria ogni possibile arco deve comparire lo stesso numero di volte, cioè

$$\frac{Nnk}{n(n-1)} = \frac{Nk}{n-1}$$

Siccome un arco compare al più una volta in ogni grafo, la frazione di grafi in cui compare l'arco, ovvero la probabilità di trovare un qualsiasi arco fissato è

$$p = \frac{k}{n-1}$$

Si ottiene lo stesso valore della sezione precedente in cui era noto soltanto il numero totale m di archi. In questo caso $m = kn/2$.

Se il grafo invece non è regolare, i gradi dei vari nodi interagiscono tutti assieme nel generare la probabilità per ogni arco fissato. Ad esempio per $n = 6$ ed una sequenza di gradi $g = (3, 3, 2, 2, 1, 1)$ vi sono $N = 17$ grafi diversi (raffigurati in Fig. 42) e ogni arco compare in un numero di grafi come indicato nella seguente tabella:

$$\begin{aligned} (1, 2) &\rightarrow 15, & (1, 3) &\rightarrow 12, & (1, 4) &\rightarrow 12, & (1, 5) &\rightarrow 6, & (1, 6) &\rightarrow 6, \\ (2, 3) &\rightarrow 12, & (2, 4) &\rightarrow 12, & (2, 5) &\rightarrow 6, & (2, 6) &\rightarrow 6, & (3, 4) &\rightarrow 6, \\ (3, 5) &\rightarrow 2, & (3, 6) &\rightarrow 2, & (4, 5) &\rightarrow 2, & (4, 6) &\rightarrow 2, & (5, 6) &\rightarrow 1 \end{aligned} \tag{61}$$

e la probabilità di ogni arco è quindi tale numero diviso per 17. Come ulteriore esempio si consideri un grafo a stella, ovvero con una sequenza di gradi $(n-1, 1, 1, \dots, 1)$. Esiste solo un grafo con questi gradi e quindi la probabilità di un arco $(1, j)$ è 1 e la probabilità di un arco (i, j) , $i, j \neq 1$, è 0.

Per ottenere una formula approssimata e di facile uso si può ragionare nel seguente modo, abbastanza valido quando il numero dei nodi sia molto elevato. Il grafo viene generato a caso generando ognuno degli m

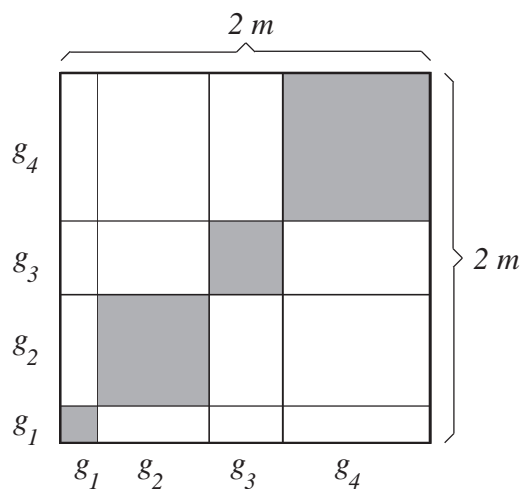


Figura 43.

archi nel seguente modo: si sceglie un primo nodo i con probabilità proporzionale a g_i , si sceglie un secondo nodo nello stesso modo e l'arco che connette questi due nodi viene generato. Si consideri la Fig. 43 dove si è raffigurato il caso di quattro nodi di gradi g_1, \dots, g_4 . Il quadrato ha lati di lunghezza $2m$ divisi in segmenti di lunghezza g_i . In base a quanto detto la coppia (i, j) viene scelta generando un punto all'interno del quadrato con probabilità uniforme escludendo però i casi in cui $i = j$, raffigurati dai quadrati in grigio. Quindi la probabilità che venga generato un particolare arco (i, j) (ovvero che il punto cada all'interno del rettangolo $g_i \times g_j$ oppure $g_j \times g_i$) è data da

$$\frac{2g_i g_j}{4m^2 - \sum_i g_i^2} \quad (62)$$

cioè, l'area data dai due rettangoli diviso l'area del quadrato senza i quadrati in grigio. La formula (62) non è comunque esatta. Si consideri il precedente esempio con sequenza di gradi $g = (3, 3, 2, 2, 1, 1)$. Applicando (62) per $(i, j) = (1, 2)$ si ottiene $18/116 = 0.155$, mentre dalla tabella (61) l'arco $(1, 2)$ compare 15 volte su un totale di 102 ($= 17 \cdot 6$, cioè tutti gli archi sommati sui 17 grafi), ovvero la probabilità di scegliere a caso l'arco $(1, 2)$ è $15/102 = 0.147$.

Il motivo di tale inesattezza è dovuto al fatto che i gradi dei nodi implicitamente determinano una distribuzione congiunta degli archi, per cui non si può isolare un singolo arco da tutto il contesto determinato dai vari gradi. Ad esempio il grafo con sequenza $(3, 2, 2, 1)$ (e 4 archi) è unico, quindi la scelta di un arco a caso non può che essere $1/4$ per ciascun arco $(1, 2)$, $(1, 3)$, $(1, 4)$ e $(2, 3)$ mentre è zero per gli archi $(2, 4)$ e $(3, 4)$. Ovviamente l'applicazione di (62) porta a risultati diversi. Tuttavia, anche se la formula (62) non è esatta, tende a dare il risultato esatto asintoticamente.

Il meccanismo delineato di generare un grafo generando ognuno degli m archi nel modo descritto potrebbe dare luogo ad incongruenze, in quanto potrebbe avvenire di generare lo stesso arco due o più volte. Però

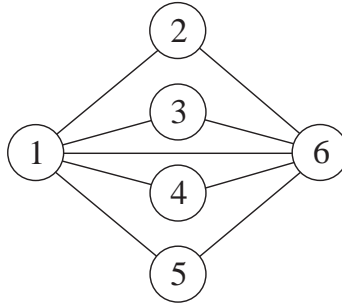


Figura 44.

con un numero elevato di nodi la probabilità che due punti cadano all'interno dello stesso rettangolo tende a zero.

Si noti ancora che la probabilità di generare un arco incidente in un nodo k è data dall'area interessata al nodo k divisa per l'area totale, cioè

$$\frac{g_k 2m + 2m g_k - 2g_k^2}{4m^2 - \sum_i g_i^2} = \frac{4m g_k - 2g_k^2}{4m^2 - \sum_i g_i^2}$$

Per n che tende ad infinito i termini dei quadrati diagonali tendono a zero e quindi

$$\frac{4m g_k - 2g_k^2}{4m^2 - \sum_i g_i^2} \xrightarrow{n \rightarrow \infty} \frac{g_k}{m}$$

per cui, generando m archi si ottiene che il numero medio di archi incidenti nel nodo k è proprio g_k .

Per valutare la probabilità di trovare l'arco (i, j) avendo generato m archi dobbiamo moltiplicare (62) per m e, tenendo conto della semplificazione valida asintoticamente per cui si trascura il termine $\sum_i g_i^2$ a denominatore, si ottiene

$$\frac{g_i g_j}{2m} \tag{63}$$

Si ribadisce che la formula (63) è un surrogato della probabilità di trovare l'arco (i, j) . In certi casi il suo valore può addirittura superare l'unità. Si considerino grafi del tipo (Fig. 44)

$$\{(1, i) : i = 2, \dots, n-1\} \cup \{(i, n) : i = 2, \dots, n-1\} \cup \{(1, n)\}$$

La formula (63), per $i = 1$ e $j = n$, è uguale a $(n-1)^2/(4n-6)$, che è maggiore di 1 per $n \geq 5$. La seguente tabella fornisce i valori di probabilità a confronto con i valori forniti dalla formula (63) per la sequenza di gradi $g = (3, 3, 2, 2, 1, 1)$, vista precedentemente. Come si vede in qualche caso il valore è inferiore e in qualche altro superiore.

arco	probabilità	formula (63)
1, 2	0.882	0.750
1, 3	0.706	0.500
1, 4	0.706	0.500
1, 5	0.353	0.250
1, 6	0.353	0.250
2, 3	0.706	0.500
2, 4	0.706	0.500
2, 5	0.353	0.250
2, 6	0.353	0.250
3, 4	0.353	0.333
3, 5	0.118	0.167
3, 6	0.118	0.167
4, 5	0.118	0.167
4, 6	0.118	0.167
5, 6	0.059	0.083

Si noti ancora che

$$\sum_{(i,j):i<j} \frac{g_i g_j}{2m^2} = \frac{1}{2} \sum_{(i,j):i \neq j} \frac{g_i g_j}{2m^2} = \frac{1}{4m^2} \sum_i g_i (\sum_j g_j - g_i) = 1 - \sum_i \frac{g_i^2}{4m^2} < 1$$

mentre la somma dei valori di probabilità deve essere uguale a 1. Quindi mediamente il valore (63) è una sottostima della probabilità (anche se per qualche termine il valore potrebbe essere superiore all'unità). Tuttavia se la somma viene estesa a tutte le coppie (i, j) , incluso il caso $i = j$ allora

$$\sum_{(i,j)} \frac{g_i g_j}{4m^2} = \frac{\sum_i g_i \sum_j g_j}{4m^2} = \frac{4m^2}{4m^2} = 1$$

Allora, facendo uso della formula (63), se l'insieme dei nodi è partizionato come S_1, \dots, S_p , il numero di archi atteso all'interno dei vari insiemi S_r è

$$\sum_{i,j} \frac{g_i g_j}{4m} [i \in S_q, j \in S_r, q = r] \quad (64)$$

dove $[i \in S_q, j \in S_r, q = r] = 1$ se i e j stanno nello stesso sottoinsieme, altrimenti vale 0. Si ribadisce che la somma viene valutata anche per i valori $i = j$. Questa quantità è costante e vale $\sum_i g_i^2 / (4m)$ e quindi è indipendente dalla particolare partizione. La somma (64) è strettamente positiva a meno che tutti i nodi abbiano grado nullo e in questo caso degenerare la divisione non è nemmeno definita.

Gli archi effettivamente presenti sono forniti dalla matrice d'adiacenza A e quindi la modularità di una partizione S_1, \dots, S_p viene valutata tramite la formula

$$\frac{1}{2} \sum_{i,j} (A_{ij} - \frac{g_i g_j}{2m}) [i \in S_q, j \in S_r, q = r]$$

A sua volta questa formula viene normalizzata rispetto al numero totale di archi e si ottiene

$$Q := \frac{1}{2m} \sum_{i,j} (A_{ij} - \frac{g_i g_j}{2m}) [i \in S_q, j \in S_r, q = r] \quad (65)$$

dove Q è il valore di modularità. Anche se normalizzato, Q non può mai essere uguale a 1, per l'osservazione precedente sulla stretta positività di (64). Ad esempio per un grafo che sia l'unione di k grafi regolari ed uguali, si ottiene $Q = 1 - 1/k$.

Il calcolo della partizione (in due o più sottoinsiemi) che massimizza la modularità è NP-difficile in generale, per cui si adottano euristiche. Un'euristica particolarmente semplice ed efficace per la partizione in due sottoinsiemi consiste nel provare a spostare un nodo da una parte all'altra e vedere se c'è un miglioramento nella modularità. Il calcolo del cambiamento di Q è rapido perché si tratta solo di valutare gli archi incidenti nel nodo che viene spostato.

Si indichi

$$B_{ij} = A_{ij} - \frac{g_i g_j}{2m}$$

La matrice B viene detta *matrice di modularità*. Si noti che

$$\sum_j B_{ij} = \sum_j A_{ij} - \frac{g_i g_j}{2m} = g_i - \frac{g_i}{2m} \sum_j g_j = g_i - \frac{g_i}{2m} 2m = 0$$

Data una partizione S_1, \dots, S_q , se $i \in S_1$, spostarlo in S_2 comporta la variazione di modularità pari a

$$\frac{1}{2m} \left(\sum_{j \in S_2} B_{ij} - \sum_{j \in S_1 \setminus i} B_{ij} \right)$$

Questo calcolo effettuato su tutti i nodi i ha un costo computazionale $O(m)$, tramite opportune strutture dati.

In forma compatta la matrice di modularità può anche essere espressa come

$$B = A - \frac{1}{2m} g g^\top$$

e siccome $g = A \mathbf{1}$, si può anche scrivere

$$B = A - \frac{1}{2m} A \mathbf{1} \mathbf{1}^\top A$$

Se il grafo è regolare c'è un legame molto stretto fra gli spettri della matrice d'adiacenza e della matrice di modularità. Siano λ e u autovalore e autovettore corrispondente di A . Allora, per un grafo k -regolare (per il quale $g = k \mathbf{1}$ e $2m = kn$)

$$B = A - \frac{k}{n} \mathbf{1} \mathbf{1}^\top$$

e

$$B u = A u - \frac{k}{n} \mathbf{1} \mathbf{1}^\top u$$

Siccome gli autovettori di A sono ortogonali e $\mathbf{1}$ è un autovettore (con autovalore k), per tutti gli altri autovettori $\mathbf{1}^\top u = 0$ e quindi

$$B u = A u = \lambda u$$

Allora per grafi regolari B ha gli autovalori e gli autovettori di A , tranne l'autovettore $\mathbf{1}$ per il quale A ha autovalore k e B ha autovalore 0.

In base al Teorema di Weyl, applicato alle matrici A e B con rispettivi autovalori $\lambda_1 \leq \dots \leq \lambda_n$ e $\lambda_1^B \leq \dots \leq \lambda_n^B$ si ha

$$\lambda_i^B \leq \lambda_i \leq \lambda_{i+1}^B, \quad i = 1, \dots, n-1$$

Inoltre $\lambda_n^B < \lambda_n$. Infatti per ogni matrice simmetrica A valgono le relazioni

$$\lambda_1 = \min_x \frac{x^\top A x}{x^\top x}, \quad \lambda_n = \max_x \frac{x^\top A x}{x^\top x}$$

e il minimo e il massimo si ottengono in corrispondenza degli autovettori relativi. Allora possiamo scrivere

$$\begin{aligned} \lambda_n^B &= \max_x \frac{x^\top B x}{x^\top x} = \frac{\hat{x}^\top B \hat{x}}{\hat{x}^\top x} = \frac{\hat{x}^\top A \hat{x}}{\hat{x}^\top x} - \frac{1}{2m} \frac{\hat{x}^\top g g^\top \hat{x}}{\hat{x}^\top \hat{x}} = \\ &= \left(\frac{\hat{x}^\top A \hat{x}}{\hat{x}^\top \hat{x}} - \frac{1}{2m} \frac{(g^\top \hat{x})^2}{\hat{x}^\top \hat{x}} \right) \leq \frac{\hat{x}^\top A \hat{x}}{\hat{x}^\top \hat{x}} \leq \max_x \frac{x^\top A x}{x^\top x} = \frac{\bar{x}^\top A \bar{x}}{\bar{x}^\top \bar{x}} = \lambda_n \end{aligned}$$

Tuttavia almeno una delle due disequazioni è stretta. Infatti il vettore \bar{x} che massimizza l'ultima espressione ha tutte le componenti dello stesso segno (in base alla teoria di Perron-Frobenius) e non può essere ortogonale a g . Quindi, se $\hat{x} = \bar{x}$ si ha $g^\top \hat{x} > 0$ e se $\hat{x} \neq \bar{x}$ è stretta l'ultima disequazione. Allora si ha

$$\lambda_n^B < \lambda_n \leq \max_i g_i$$

Non sempre λ_n^B è positivo. È stato dimostrato che $\lambda_n^B = 0$ se e solo se il grafo è multi-partito completo (si riveda la definizione a pag. 9), che include anche il caso di grafo completo. Il grafo di Fig. 44 è tripartito completo e infatti si ha $\lambda_n^B = 0$.

Come nel caso del taglio di capacità minima si può definire un vettore $x_i \in \{-1, 1\}$ a seconda se $i \in S$ oppure no. Pertanto la modularità di una partizione $(S, V \setminus S)$ è

$$\frac{1}{2m} \sum_{i,j} \frac{1+x_i x_j}{2} B_{ij} = \frac{1}{4m} \sum_{i,j} x_i x_j B_{ij} = \frac{1}{4m} x^\top B x$$

Per massimizzare si possono a questo punto usare le stesse tecniche spettrali usate per massimizzare il taglio. Basta calcolare l'autovettore di B associato all'autovalore più grande. In questo caso non è presente il vincolo che uno dei due insiemi abbia cardinalità fissata. Il problema da risolvere è quindi

$$\begin{aligned} \max \quad & x^\top B x \\ & x_i \in \{-1, 1\} \end{aligned}$$

che viene rilassato in

$$\max x^\top B x$$

$$\sum_i x_i^2 = n$$

per il quale viene costruita la funzione lagrangiana

$$x^\top B x + \lambda (n - \sum_i x_i^2)$$

che, derivando ed eguagliando a zero, porta a

$$B x = \lambda x$$

Quindi anche in questo caso la soluzione è data da un autovettore. Siccome $x^\top x = n$, si ha

$$Q = \frac{1}{4m} x^\top B x = \frac{n}{4m} \lambda$$

e si tratta allora di prendere l'autovalore λ_n^B più grande. Si noti che quando $\lambda_n^B = 0$, allora la modularità massima si ha in corrispondenza del grafo stesso senza partizioni. Come si è detto, questo fatto avviene per i grafi multi-partiti completi che quindi vengono detti *indivisibili*.

Si consideri ad esempio il grafo in Fig. 45-(a). Adottando la prima euristica e partendo dall'insieme $S = \{1, 2, 3, 4, 5\}$ si ottengono i grafi in Fig. 45-(b) con i valori di modularità indicati sotto ciascun grafo. Calcolando l'autovettore associato all'autovalore più grande si ottiene

$$(0.324, 0.367, -0.140, -0.357, 0.386, 0.245, -0.415, -0.380, 0.193, -0.221)$$

da cui si vede che l'insieme S è costituito dai nodi $\{1, 2, 5, 6, 9\}$.

Cosa succede se il grafo è già sconnesso? Come si riflette questo fatto nella modularità? Supponiamo che il grafo sia l'unione $G \cup G$ di due grafi uguali entrambi con matrice d'incidenza A , m archi, vettore di gradi g e matrice di modularità B . Allora la matrice di modularità del grafo unione è

$$\mathcal{B} = \begin{pmatrix} A - \frac{g g^\top}{4m} & -\frac{g g^\top}{4m} \\ -\frac{g g^\top}{4m} & A - \frac{g g^\top}{4m} \end{pmatrix} = \begin{pmatrix} B + \frac{g g^\top}{4m} & -\frac{g g^\top}{4m} \\ -\frac{g g^\top}{4m} & B + \frac{g g^\top}{4m} \end{pmatrix}$$

Allora si vede che, se u è un autovettore di A con autovalore λ si ha

$$\mathcal{B} \begin{pmatrix} u \\ -u \end{pmatrix} = \lambda \begin{pmatrix} u \\ -u \end{pmatrix}$$

mentre se v è un autovettore di B con autovalore λ^B si ha

$$\mathcal{B} \begin{pmatrix} v \\ v \end{pmatrix} = \lambda^B \begin{pmatrix} v \\ v \end{pmatrix}$$

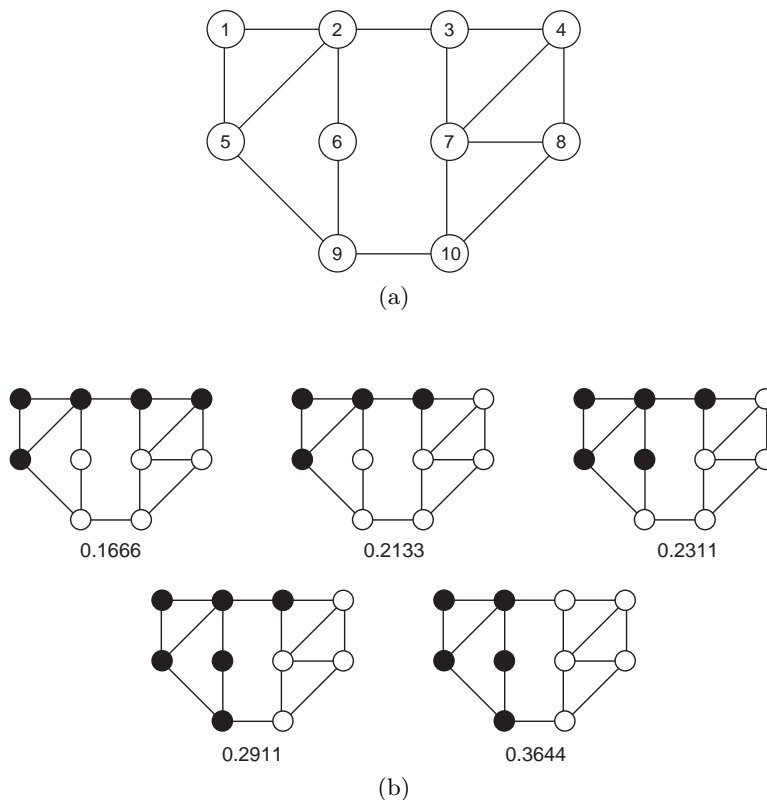


Figura 45.

Quindi gli autovalori di \mathcal{B} sono l'unione degli autovalori di A e di B . In particolare l'autovalore massimo è fra quelli di A . Siccome $\lambda_n^B < \lambda_n$ la massima modularità si ottiene per l'autovettore $(u, -u)$ corrispondente a λ_n , per il quale u ha componenti tutte dello stesso segno e quindi l'autovettore $(u, -u)$ determina correttamente la partizione del grafo. Si noti che, se G è k -regolare, allora $u = \mathbf{1}$ e $\lambda = k$ e in questo caso il problema rilassato fornisce proprio la soluzione intera.

41 Modularità con gradi fissati – Calcolo dell'autovalore più grande

Il calcolo dell'autovalore più grande in valore assoluto di una matrice C si calcola facilmente eseguendo la seguente iterazione a partire da un vettore arbitrario u^0 purché diverso da un autovettore (se scelto casualmente la probabilità che sia uguale ad un autovettore è nulla)

$$v^k = C u^k, \quad u^{k+1} = \frac{v^k}{\|v^k\|}, \quad k = 0, 1, 2, \dots \quad (66)$$

La successione $\|v^k\|$ tende all'autovalore più grande in valore assoluto e la successione u^k tende all'autovettore corrispondente. La velocità di convergenza è lineare con tasso uguale al rapporto fra il più grande autovalore in valore assoluto e quello immediatamente successivo. Tanto più elevato è questo rapporto e tanto più velocemente la successione tende ai valori limite.

Si noti che la matrice $C + \alpha I$ ha gli stessi autovettori di C e tutti i suoi autovalori si ottengono da quelli di C sommando α . Quindi se si prende $\alpha = \rho(C)$ (il raggio spettrale della matrice C) si ha la garanzia che l'autovalore più grande in valore assoluto di $C + \alpha I$ corrisponde all'autovalore più grande di C . Prendere α troppo grande non conviene perché si diminuirebbe la velocità di convergenza. La scelta ideale per ottenere la massima velocità di convergenza sarebbe

$$\alpha = -\frac{\lambda_{n-1} + \lambda_1}{2}$$

che corrisponde a 'spostare' gli autovalori in modo che l'origine sia a metà strada fra $\lambda_1 + \alpha$ e $\lambda_{n-1} + \alpha$, i due autovalori che possono entrare in gioco come secondo autovalore in valore assoluto. Ma ovviamente, per usare tale valore di α sarebbe richiesta la conoscenza degli autovalori e quindi la formula non è praticabile. Siccome vale la relazione $\rho(C) \leq \|C\|$ per qualsiasi norma $\|C\|$, è più semplice usare

$$\alpha = \|C\|$$

e una norma di facile calcolo è

$$\|C\|_1 = \max_i \sum_j |C_{ij}|$$

Applicando questo risultato alla matrice di modularità B si ha

$$\|B\|_1 = \max_i \sum_j |A_{ij} - \frac{1}{2m} g_i g_j|$$

e quindi bisogna applicare l'iterazione (66) a $B + \|B\|_1 I$.

42 Modularità con gradi fissati – Partizione in più insiemi

Una delle tecniche più usate per massimizzare la modularità partizionando in più di due sottoinsiemi il grafo, consiste nella partizione in due sottoinsiemi dei vari sottoinsiemi generati. Naturalmente la partizione viene effettuata solo se la modularità aumenta, altrimenti l'insieme non viene ulteriormente suddiviso. Si tratta allora di capire quando conviene suddividere un sottoinsieme di nodi S .

Il contributo dell'insieme S alla modularità di tutto il grafo è dato dall'espressione (65)

$$\frac{1}{2m} \sum_{i,j \in S} B_{ij} \tag{67}$$

Se invece S viene suddiviso secondo la partizione S_1 e S_2 il contributo è

$$\frac{1}{2m} \sum_{i,j} B_{ij} [i \in S_1, j \in S_1] + \frac{1}{2m} \sum_{i,j} B_{ij} [i \in S_2, j \in S_2] \quad (68)$$

che possiamo esprimere con il vettore

$$x_i = \begin{cases} 1 & \text{se } i \in S_1 \\ -1 & \text{se } i \in S_2 \end{cases} \quad i \in S$$

(si noti che il vettore è definito solo su S) come

$$\frac{1}{2m} \sum_{i,j \in S} B_{ij} \frac{1 + x_i x_j}{2}$$

La differenza ΔQ fra (68) e (67) è quindi

$$\Delta Q = \frac{1}{2m} \left(\sum_{i,j \in S} B_{ij} \frac{1 + x_i x_j}{2} - \sum_{i,j \in S} B_{ij} \right) = \frac{1}{4m} \left(\sum_{i,j \in S} B_{ij} x_i x_j - \sum_{i,j \in S} B_{ij} \right) \quad (69)$$

Si noti che in questo caso non si può sfruttare la relazione $\sum_{i,j} B_{ij} = 0$, in quanto gli indici i e j sono ristretti all'insieme S . Per poter scrivere (69) come una forma quadratica si può notare che

$$\begin{aligned} \sum_{i \in S} \sum_{j \in S} B_{ij} &= \sum_{i \in S} \sum_{k \in S} B_{ik} = \sum_{i \in S} \sum_{k \in S} B_{ik} x_i^2 = \sum_{i \in S} \sum_{k \in S} \sum_{j \in S} B_{ik} x_i x_j [i = j] = \\ &= \sum_{i \in S} \sum_{j \in S} \sum_{k \in S} B_{ik} x_i x_j [i = j] = \sum_{i \in S} \sum_{j \in S} x_i x_j [i = j] \sum_{k \in S} B_{ik} \end{aligned}$$

Sia D_S la matrice diagonale in cui $d_{ii} = \sum_{k \in S} B_{ik}$. Allora (69) è data dalla seguente forma quadratica

$$\Delta Q = \frac{1}{4m} (x^\top (B_S - D_S) x)$$

dove B_S è la matrice di modularità ristretta all'insieme S . Si noti che anche la matrice simmetrica $B_S - D_S$ ha autovalore nullo con autovettore associato $\mathbf{1}$. Infatti

$$\sum_{j \in S} (B_{ij} - [i = j] \sum_{k \in S} B_{ik}) = \sum_{j \in S} B_{ij} - \sum_{k \in S} B_{ik} = 0$$

Ora si tratta semplicemente di massimizzare ΔQ . Se il valore massimo è positivo la suddivisione, come indicato dal vettore soluzione x , viene effettuata, altrimenti l'insieme non viene ulteriormente suddiviso.

Come esempio si consideri il grafo in Fig. 46-(a) con $n = 80$ che è stato generato a caso nel seguente modo: se $\lfloor (i-1)/16 \rfloor = \lfloor (j-1)/16 \rfloor$ allora l'arco (i, j) viene generato con probabilità 0.6, altrimenti viene generato con probabilità 0.02. In questo modo si generano 5 gruppi di nodi (da 1 a 16, da 17 a 31 ecc.) con molti archi al loro interno, mentre il numero di archi fra i gruppi di nodi è molto minore. Si veda la

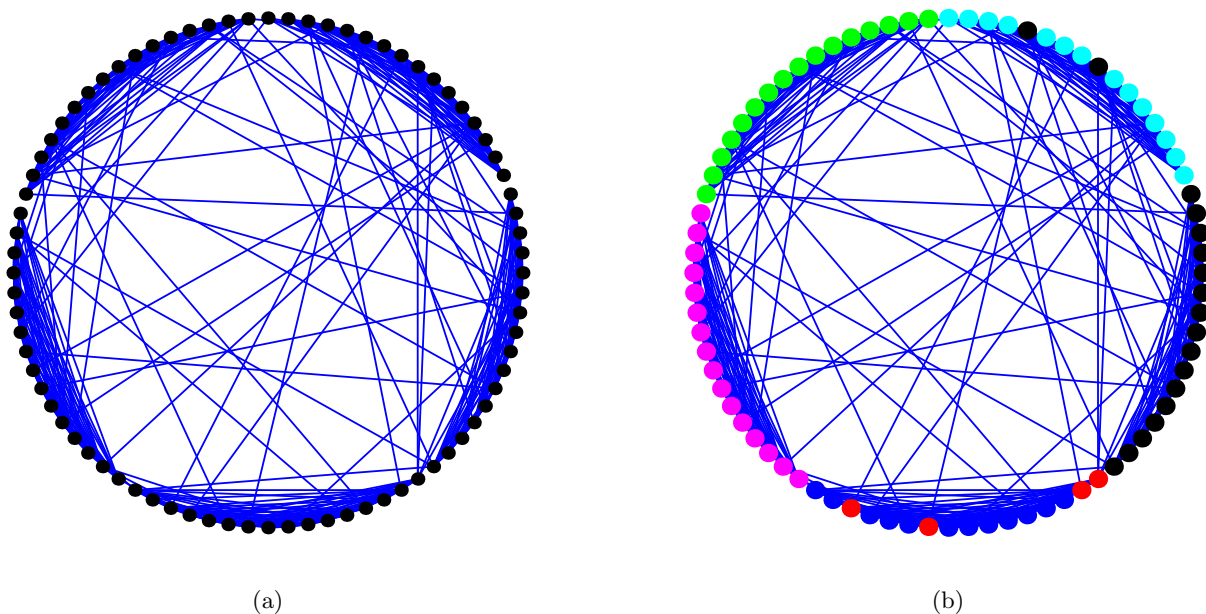


Figura 46.

figura dove i nodi sono disposti lungo una circonferenza numerati in senso orario a partire dall'alto. La prima partizione che si trova con valore di modularità 0.401 è

$$S_1 = \{5, 9, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 65, \\ 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80\}$$

e $S_2 = V \setminus S_1$. Massimizzando ΔQ per S_1 si ottiene $\Delta Q = 0.065$ con ulteriore partizione di S_1 in

$$S_3 = \{5, 9, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32\}$$

e $S_4 = S_1 \setminus S_3$. Massimizzando ΔQ per S_3 si ottiene un valore negativo e quindi S_3 rimane inalterato. Massimizzando ΔQ per S_2 si ottiene $\Delta Q = 0.069$ con partizione di S_2 data da:

$$S_5 = \{1, 2, 3, 4, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 33, 34, 42, 46\}$$

e $S_6 = S_2 \setminus S_5$. Massimizzando ΔQ per S_5 si ottiene $\Delta Q = 0.011$ con ulteriore partizione di S_5 in

$$S_7 = \{33, 34, 42, 46\}$$

e $S_8 = S_5 \setminus S_7$. S_4 e S_7 non sono ulteriormente partizionabili. Invece per S_6 si ottiene $\Delta Q = 0.045$ con partizione

$$S_9 = \{35, 36, 37, 38, 39, 40, 41, 43, 44, 45, 47, 48\}$$

e $S_{10} = S_6 \setminus S_9$. Infine S_8 , S_9 e S_{10} non sono partizionabili. Allora il grafo è stato suddiviso nei 6 insiemi

$$\begin{aligned} &\{1, 2, 3, 4, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16\} \\ &\{5, 9, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32\} \\ &\{33, 34, 42, 46\} \\ &\{35, 36, 37, 38, 39, 40, 41, 43, 44, 45, 47, 48\} \\ &\{49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64\} \\ &\{65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80\} \end{aligned}$$

raffigurati con colori diversi in Fig. 46-(b). Il risultato è un po' diverso da quanto ci si attendeva. I nodi 5 e 9 sono stati messi in un gruppo diverso da quello immaginato e i nodi $\{33, 34, 42, 46\}$ formano un gruppo a sé staccato dagli altri nodi a cui dovrebbero per costruzione appartenere. Se calcoliamo la modularità per la partizione con cui è stato generato il grafo si ottiene il valore 0.678, migliore quindi di quello trovato. Se nella prima fase si fosse scelto un valore di soglia pari a 0.1 per accettare gli elementi di S si sarebbe ottenuta esattamente la partizione cercata.

Alternativamente si possono usare algoritmi di ricerca locale, che, nel caso di partizione di grafi, risultano abbastanza semplici concettualmente. Partendo da una partizione totale si prova a spostare ogni nodo in un altro insieme. Lo spostamento che produce il migliore risultato viene mantenuto e la partizione viene modificata. L'algoritmo termina quando non ci sono più miglioramenti. Nell'esempio del grafo di 80 nodi, dopo 77 iterazioni si ottiene proprio la partizione ottima.

Applicando la ricerca locale al grafo di Fig. 45-(a) si ottiene la partizione (dopo 8 iterazioni)

$$\{\{1, 2, 5\}, \{6, 9\}, \{3, 4, 7, 8, 10\}\}$$

con modularità 0.33111, inferiore al valore 0.3644 ottenuto partizionando in due soli sottoinsiemi. Evidentemente spostando solo il nodo 6 o solo il nodo 9 nel primo insieme non si migliora la modularità. Questa viene migliorata spostando insieme entrambi i nodi. Naturalmente se la ricerca locale dovesse prendere in esame coppie di nodi, il suo costo computazionale diventerebbe eccessivo.

Come secondo esempio si consideri il grafo in Fig. 47-(1). Il grafo è stato generato casualmente, generando prima 300 punti in modo uniforme nel quadrato unitario. Poi per ogni coppia di punti x_i, x_j , il corrispondente arco (i, j) è stato generato con probabilità e^{-d} con $d = \alpha (x_i - x_j)^\top (x_i - x_j)$ e $\alpha = .0035$, cioè più vicini sono i punti tanto maggiore è la probabilità che ci sia un arco. Calibrando opportunamente il valore di α si ottengono grafi più o meno densi. La matrice d'adiacenza con i nodi ordinati casualmente si vede in Fig. 47-(1a). Come si vede gli uni della matrice d'adiacenza (nel disegno rappresentati come punti) sono distribuiti casualmente.

Si è applicata la separazione con metodi spettrali seguita da una ricerca locale eseguita con l'euristica di scambio. La prima divisione si vede in Fig. 47-(2) con la matrice d'adiacenza in Fig. 47-(2a), dove già si può notare l'effetto della suddivisione. Nelle Fig. 47-(3) e 47-(3a) si vede la partizione dopo la suddivisione del primo insieme in due. Sotto le figure sono riportati i valori di modularità che si ottengono dopo la

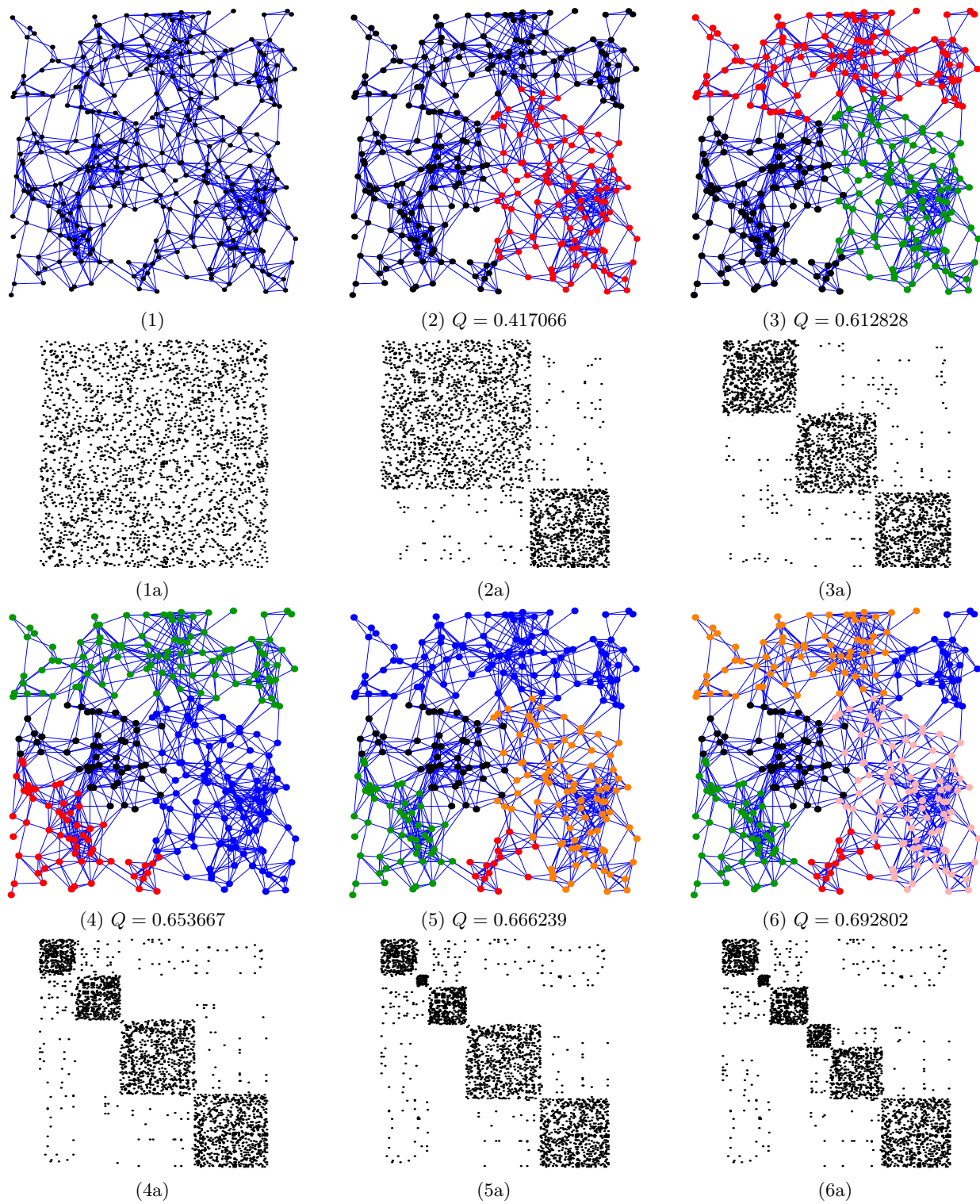


Figura 47.

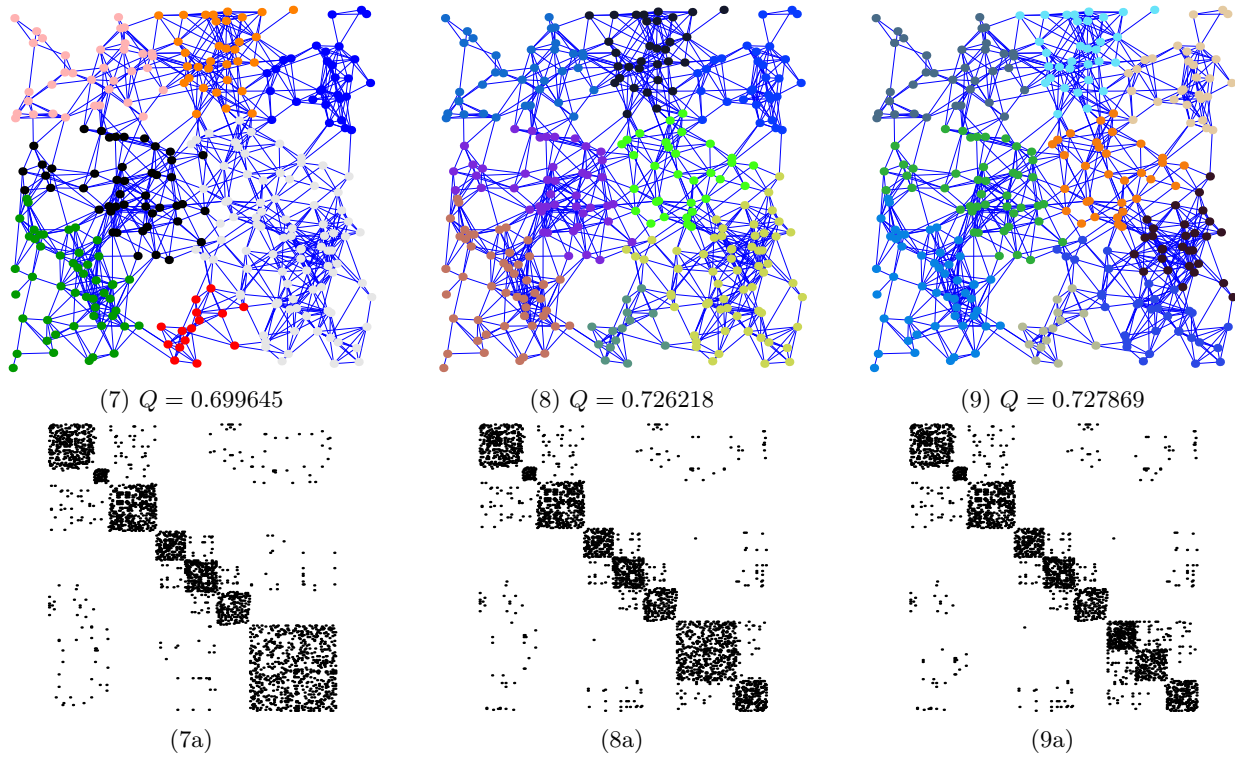


Figura 48.

suddivisione. Il primo insieme viene ulteriormente suddiviso in due ancora nei due passi successivi (Fig. 47-(4), 47-(4a), 47-(5) e 47-(5a)). Poi viene suddiviso il quarto insieme (Fig. 47-(6) e 47-(6a)), successivamente il quinto (Fig. 48-(7) e 47-(7a)). Infine il sesto viene suddiviso due volte (Fig. 48-(8), 48-(8a), 48-(9) e 48-(9a)).

A questo punto si sono ottenuti nove sottoinsiemi S_1, \dots, S_9 con un valore di modularità $Q = 0.727869$. Tuttavia è importante anche poter capire se i sottoinsiemi possono essere raggruppati fra loro in modo da evidenziare una struttura meno fine a più alto livello. Siccome i sottoinsiemi sono stati ottenuti non semplicemente per suddivisione, ma si è anche operato uno spostamento di nodi da un insieme all'altro con la procedura di scambio, tale gerarchia di inclusione si è persa e deve essere recuperata in altro modo.

A ben guardare, si tratta di ripetere per i sottoinsiemi la procedura finora fatta per i nodi del grafo. Cioè bisogna raggruppare gli insiemi usando una tecnica di modularità. Possiamo pensare ad ogni sottoinsieme come ad un nodo di un nuovo grafo. Questo grafo è completo e ad ogni arco (h, k) viene assegnato un peso w_{hk} uguale al numero di archi con un estremo in S_h e l'altro in S_k . Sia \tilde{A} una matrice d'adiacenza pesata data dai valori w_{hk} (ponendo $w_{hh} = 0$) e sia $\tilde{g}_h = \sum_k w_{hk}$ il 'grado pesato' dell'insieme h e sia $\tilde{m} = \sum_h \tilde{g}_h / 2$.

Si può allora definire una matrice di modularità pesata

$$\tilde{B} = \tilde{A} - \frac{\tilde{g}\tilde{g}^\top}{2\tilde{m}}$$

La matrice di adiacenza è per l'esempio

$$\tilde{A} = \begin{pmatrix} 0 & 1 & 30 & 0 & 2 & 15 & 0 & 5 & 12 \\ 1 & 0 & 3 & 0 & 0 & 0 & 0 & 10 & 0 \\ 30 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 14 & 0 & 1 & 0 & 7 \\ 2 & 0 & 0 & 14 & 0 & 19 & 0 & 0 & 17 \\ 15 & 0 & 0 & 0 & 19 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 30 & 18 \\ 5 & 10 & 0 & 0 & 0 & 0 & 30 & 0 & 6 \\ 12 & 0 & 0 & 7 & 17 & 0 & 18 & 6 & 0 \end{pmatrix}$$

Usando la tecnica di modularità si ottiene una prima suddivisione negli insiemi

$$\{\{S_1, S_3, S_4, S_5, S_6\}, \{S_2, S_7, S_8, S_9\}\}$$

che, ripetuta, produce la seguente suddivisione finale

$$\{\{\{S_1, S_3\}, \{S_4, S_5, S_6\}\}, \{S_2, S_7, S_8, S_9\}\}$$

In Figura 49(a) si vede la matrice di adiacenza con le righe permutate secondo la suddivisione finale con la suddivisione evidenziata secondo i blocchi diagonali. In Figura 49(b) si vede invece la matrice di adiacenza del grafo originario con le righe e colonne permutate secondo l'ultima suddivisione. Infine in Fig. 50 si vede la partizione sul grafo stesso.

43 Grafi casuali

Dato un insieme finito \mathcal{G} , un elemento casuale (o stocastico o aleatorio) è un elemento $G \in \mathcal{G}$ campionato con probabilità p_G (con ovviamente $\sum_{G \in \mathcal{G}} p_G = 1$). Se la probabilità non viene specificata, si intende che la probabilità è uniforme, cioè $p_G = 1/|\mathcal{G}|$ per ogni elemento G . Nel caso di grafi l'insieme \mathcal{G} viene definito normalmente dal numero di nodi n del grafo G ed eventualmente anche dal numero di archi m oppure, in modo più dettagliato, specificando i gradi dei nodi.

Ad esempio \mathcal{G} potrebbe essere l'insieme dei grafi con 3 nodi rappresentato in Fig. 51. Siccome ogni arco può essere o non essere presente il numero di grafi con n nodi è 2^M , dove si è indicato con $M = n(n-1)/2$ il numero di tutte le coppie non ordinate di nodi. In Fig. 52 è raffigurato l'insieme dei grafi con $n = 4$ e $m = 4$. Il numero di grafi con n nodi e m archi è $\binom{M}{m}$.

Se la specifica riguarda i gradi dei nodi si può ad esempio richiedere che il grado medio \bar{g} sia un certo valore prefissato. Per grado medio s'intende che la media è rispetto ai nodi del grafo generato, non rispetto a

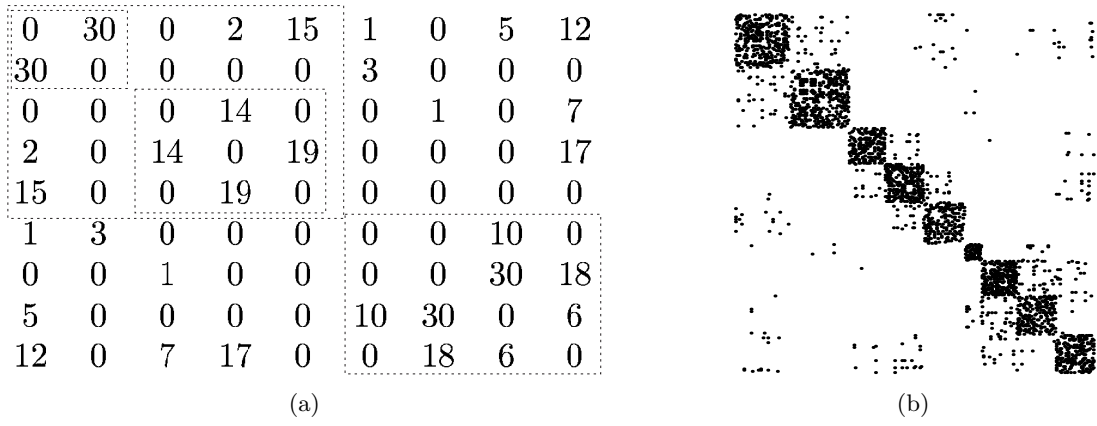


Figura 49.

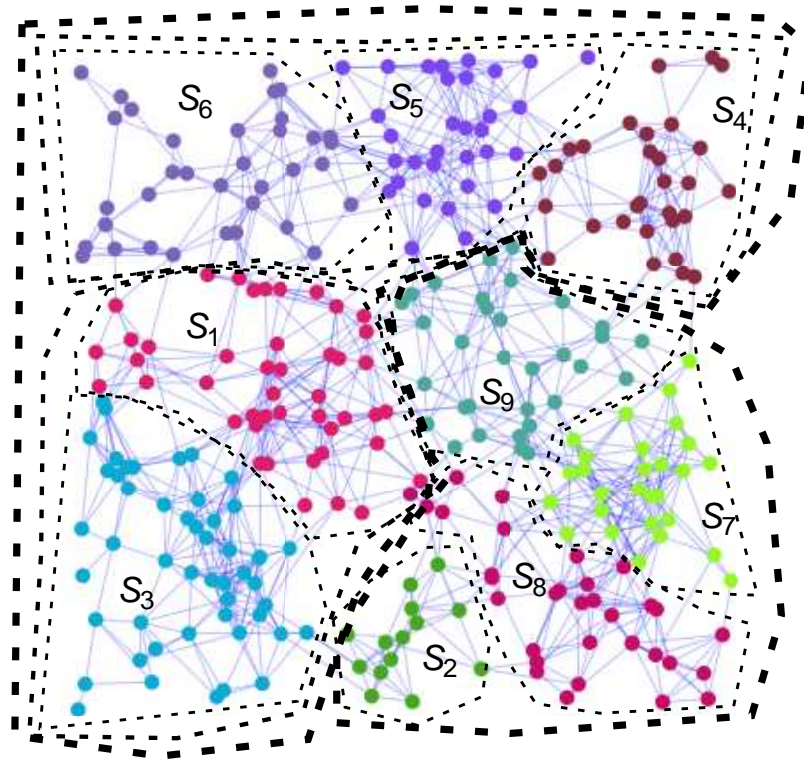


Figura 50.

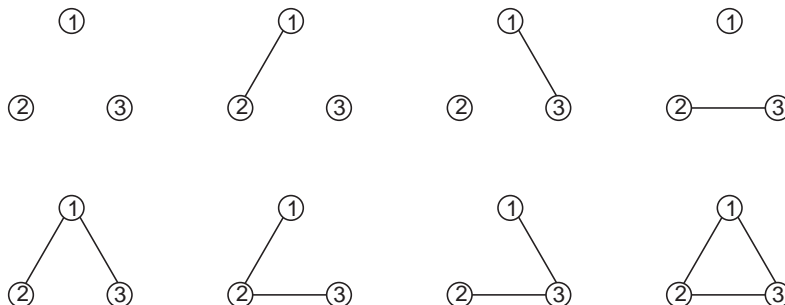


Figura 51. Insieme dei grafi con 3 nodi

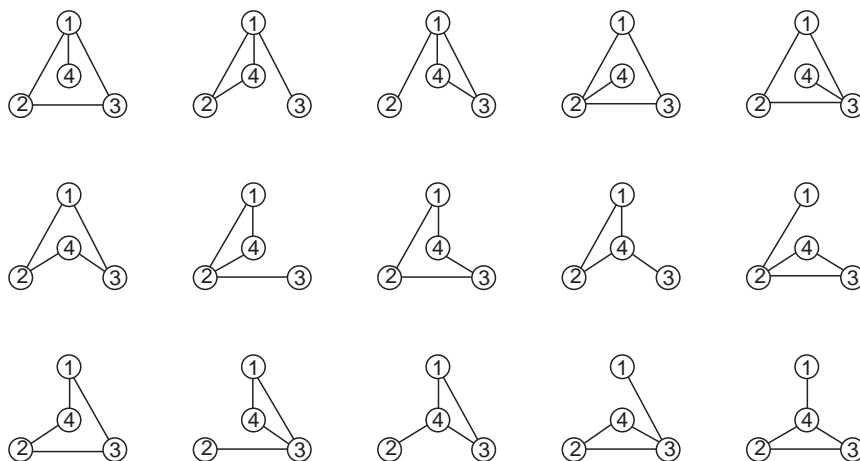


Figura 52. Insieme dei grafi con 4 nodi e 4 archi

tutte le possibili realizzazioni del grafo. Siccome la somma dei gradi dei nodi è $2m$, specificare il grado medio è equivalente a specificare il valore $2m/n$ e quindi, con n fissato, a specificare m . In modo più dettagliato si può specificare una particolare sequenza di gradi per i nodi (che necessariamente determina anche il numero di archi). Se ad esempio si volesse la sequenza $(3, 3, 2, 1, 1)$ nell'ordine per i nodi $(1, 2, 3, 4, 5)$ l'insieme dei grafi sarebbe l'insieme dei due grafi in Fig. 53. La sequenza può riferirsi a nodi arbitrari. In questo caso l'insieme dell'esempio contiene 60 grafi. Si noti che questo valore è inferiore a tutte le permutazioni di 5 nodi per la presenza di automorfismi.

L'aspetto critico per la scelta di un grafo particolare fra quelli specificati consiste nel fatto che non è possibile generare esplicitamente l'insieme \mathcal{G} e poi scegliere un grafo all'interno di \mathcal{G} . La cardinalità di \mathcal{G} è normalmente esponenziale in n e quindi bisogna adottare metodi diversi di generazione del grafo casuale.

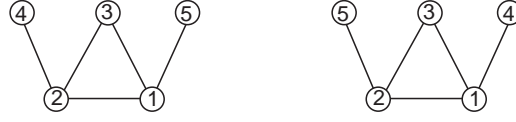


Figura 53. Insieme dei grafi con sequenza di gradi (3, 3, 2, 1, 1)

44 Modello $G_{n,p}$

La generazione di un grafo casuale con n nodi con probabilità uniforme si effettua generando ognuno degli $M = n(n-1)/2$ archi con probabilità $1/2$. In questo modo ognuno dei grafi di Fig. 51 può venir generato con la stessa probabilità $1/8$. In generale la probabilità di campionamento di ogni grafo è $p_G = 2^{-M}$ per ogni $G \in \mathcal{G}$. I grafi campionati in questo modo non hanno ovviamente tutti lo stesso numero di archi.

Questo modo di generare un grafo può essere generalizzato pensando di generare ogni arco con probabilità p . Tale modello di grafo viene indicato con la notazione $G_{n,p}$. Se $p \neq 1/2$ il campionamento non avviene più in modo uniforme all'interno di \mathcal{G} . Uno specifico grafo con m archi viene campionato con probabilità

$$p^m (1-p)^{M-m}$$

e un generico grafo con m archi con probabilità

$$\binom{M}{m} p^m (1-p)^{M-m} \quad (70)$$

Indicando con X_{ij} una variabile aleatoria che vale 1 se l'arco (i, j) viene generato e 0 altrimenti, il numero medio \bar{m} di archi è

$$\bar{m} = E[m] = E\left[\sum_{ij} X_{ij}\right] = \sum_{ij} E[X_{ij}] = \sum_{ij} p = pM$$

e la varianza di m è

$$E[(m - \bar{m})^2] = \sum_{ij} E[(X_{ij} - p)^2] = \sum_{ij} p(1-p) = p(1-p)M$$

La probabilità che un nodo generico abbia grado k è

$$q_k = \binom{n-1}{k} p^k (1-p)^{n-1-k} \quad (71)$$

Rispetto ad ogni possibile realizzazione il grado medio di ogni nodo è

$$\bar{g}_i = E[g_i] = \sum_{j \neq i} E[X_{ij}] = \sum_{j \neq i} p = p(n-1) \quad (72)$$

costante per ogni nodo e con varianza $p(1-p)(n-1)$. La relazione (72) suggerisce di scegliere la probabilità p dal valore desiderato di grado medio \bar{g} , cioè $p = \bar{g}/(n-1)$, per cui (71) diventa

$$q_k = \binom{n-1}{k} \left(\frac{\bar{g}}{n-1}\right)^k \left(1 - \frac{\bar{g}}{n-1}\right)^{n-1-k} \quad (73)$$

Il grado più probabile si ottiene trovando il massimo valore di q_k rispetto a k . Siccome q_k è prima crescente e poi decrescente, conviene valutare quale è il minimo valore k per cui $q_k > q_{k+1}$. Quindi

$$\binom{n-1}{k} p^k (1-p)^{n-1-k} > \binom{n-1}{k+1} p^{k+1} (1-p)^{n-1-k-1}$$

cioè

$$\frac{1-p}{p} > \frac{n-k-1}{k+1} \implies \frac{1}{p} > \frac{n}{k+1} \implies k > np - 1 = \frac{n}{n-1} \bar{g} - 1$$

Quindi il valore massimo di ha per $k = \lceil n/(n-1) \bar{g} \rceil - 1$. Se \bar{g} è intero si ha $k = \bar{g}$.

Il valore di probabilità (70) è praticamente trascurabile per valori di m abbastanza diversi da \bar{m} . Si noti ancora che la varianza è molto bassa per valori bassi di p . Siccome normalmente il grado medio scelto rimane costante anche al crescere di n , si deduce che p è normalmente un valore molto piccolo e quindi un grafo del modello $G_{n,p}$ tende ad avere un numero di archi molto concentrato intorno a \bar{m} con grado anche molto concentrato attorno a \bar{g} . Il grafico di (70) in funzione di m per $p = 0.25$ con $n = 10$ e $n = 100$ è riportato nelle Figure 54-(a)-(b).

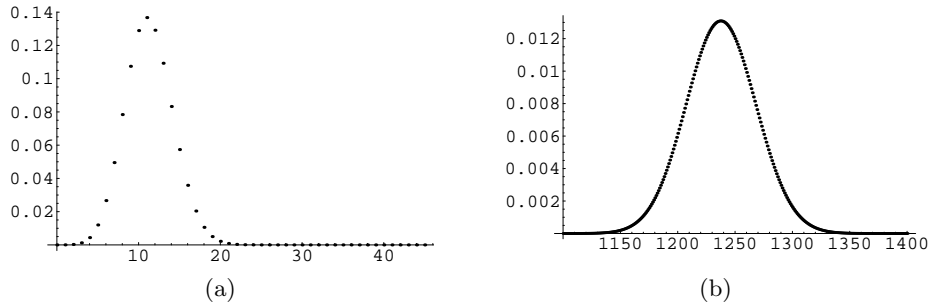


Figura 54.

È interessante capire come siano distribuiti i gradi nel grafo generato. Sia $\bar{q} = (\bar{q}_0, \bar{q}_1, \dots, \bar{q}_{n-1})$, dove \bar{q}_k è la frazione di nodi di grado k . Ad esempio per $n = 2$ sono possibili solo due grafi, quello di due nodi isolati senza arco con probabilità $1-p$ e quello con l'arco con probabilità p . Quindi con probabilità p si ha $\bar{q} = (0, 1)$ e con probabilità $1-p$ si ha $\bar{q} = (1, 0)$. Moltiplicando questi due casi per le rispettive probabilità e sommando si ottiene $q = ((1-p), p)$, cioè (71). Per $n = 3$ si ottengono i seguenti quattro casi (si guardi la Fig. 51) con le probabilità indicate

$$\bar{q} = (1, 0, 0) \rightarrow (1-p)^3; \quad \bar{q} = \left(\frac{1}{3}, \frac{2}{3}, 0\right) \rightarrow 3p(1-p)^2; \quad \bar{q} = \left(0, \frac{2}{3}, \frac{1}{3}\right) \rightarrow 3p^2(1-p); \quad \bar{q} = (0, 0, 1) \rightarrow p^3$$

Anche in questo caso se i diversi valori di \bar{q} vengono moltiplicati per le rispettive probabilità e poi sommati si ottiene

$$q = ((1-p)^2, 2p(1-p), p^3)$$

che è esattamente (71) per $n = 3$. Quello che ci si chiede in questi casi è se, quando n è molto grande, la distribuzione dei gradi \bar{q}_k sui nodi per il *singolo* grafo campionato rifletta la distribuzione dei gradi q_k su tutte le realizzazioni data da (71). Questa proprietà di tipo ergodico effettivamente vale per n che tende ad infinito. Si vedano nelle Figure 55 la probabilità q_k dei gradi data da (71) (Figure (a), (c) e (e)) e la frazione \bar{q}_k ottenuta rispetto ai nodi per una particolare realizzazione (Figure (b), (d) e (f)) nei tre casi, $n = 100, 1000, 10000$.

Può essere utile approssimare l'espressione (73) per $n \gg k$ nel seguente modo:

$$\begin{aligned} \binom{n-1}{k} \left(\frac{\bar{g}}{n-1}\right)^k \left(1 - \frac{\bar{g}}{n-1}\right)^{n-1-k} &\approx \frac{(n-1)^k}{k!} \left(\frac{\bar{g}}{n-1}\right)^k \left(1 - \frac{\bar{g}}{n-1}\right)^{n-1-k} = \\ &\frac{\bar{g}^k}{k!} \left(1 - \frac{\bar{g}}{n-1}\right)^{n-1-k} = \frac{\bar{g}^k}{k!} \left(\left(1 - \frac{\bar{g}}{n-1}\right)^{(n-1-k)/\bar{g}}\right)^{\bar{g}} \approx \frac{\bar{g}^k}{k!} e^{-\bar{g}} \end{aligned} \quad (74)$$

A motivo di questa espressione i grafi casuali $G_{n,p}$ prendono anche il nome di *grafi di Poisson*. In Fig. 56 si vedono sovrapposti i grafici di (73) (linea solida) e (74) (linea tratteggiata) in funzione di k per $n = 100$ (figura (a)) e $n = 1000$ (figura (b)) con $\bar{g} = 10$. Per $n = 1000$ i due grafici coincidono quasi esattamente.

45 Grafi casuali con numero d'archi prefissato

In questo caso l'insieme \mathcal{G} consiste di tutti i grafi con n nodi ed m archi. Vi sono

$$\binom{M}{m}$$

grafi in \mathcal{G} . Per generare uno di questi grafi in modo uniforme, la cosa più semplice è scegliere casualmente m archi fra le M possibili coppie di nodi. Un generico nodo è presente in $(n-1)$ coppie. Affinché un nodo generico abbia grado k devono essere scelte k di queste $n-1$ coppie dal totale di M . Allora la probabilità che un nodo generico abbia grado k è

$$q_k = \frac{\binom{n-1}{k} \binom{M-n+1}{m-k}}{\binom{M}{m}} \quad (75)$$

(la formula si giustifica in questo modo: al denominatore c'è il numero di tutti i sottoinsiemi di m elementi da un insieme di M elementi; al numeratore ci sono tutti i modi possibili di prendere k elementi all'interno di $n-1$ elementi e i rimanenti $m-k$ elementi all'interno dei rimanenti $M-n+1$ elementi). Per confrontare (75) con (71), bisogna porre $p = \bar{g}/(n-1)$ e siccome $n\bar{g} = 2m$ si ha $p = m/M$. La formula (71) calcolata per questo valore di p è quasi uguale a (75) per i valori di q_k non trascurabili.

Quindi si ottiene quasi lo stesso risultato se si genera ogni arco con probabilità m/M oppure si scelgono a caso m archi. Infatti, come già detto precedentemente il numero di archi, nel primo caso, si discosta pochissimo dal valore medio.

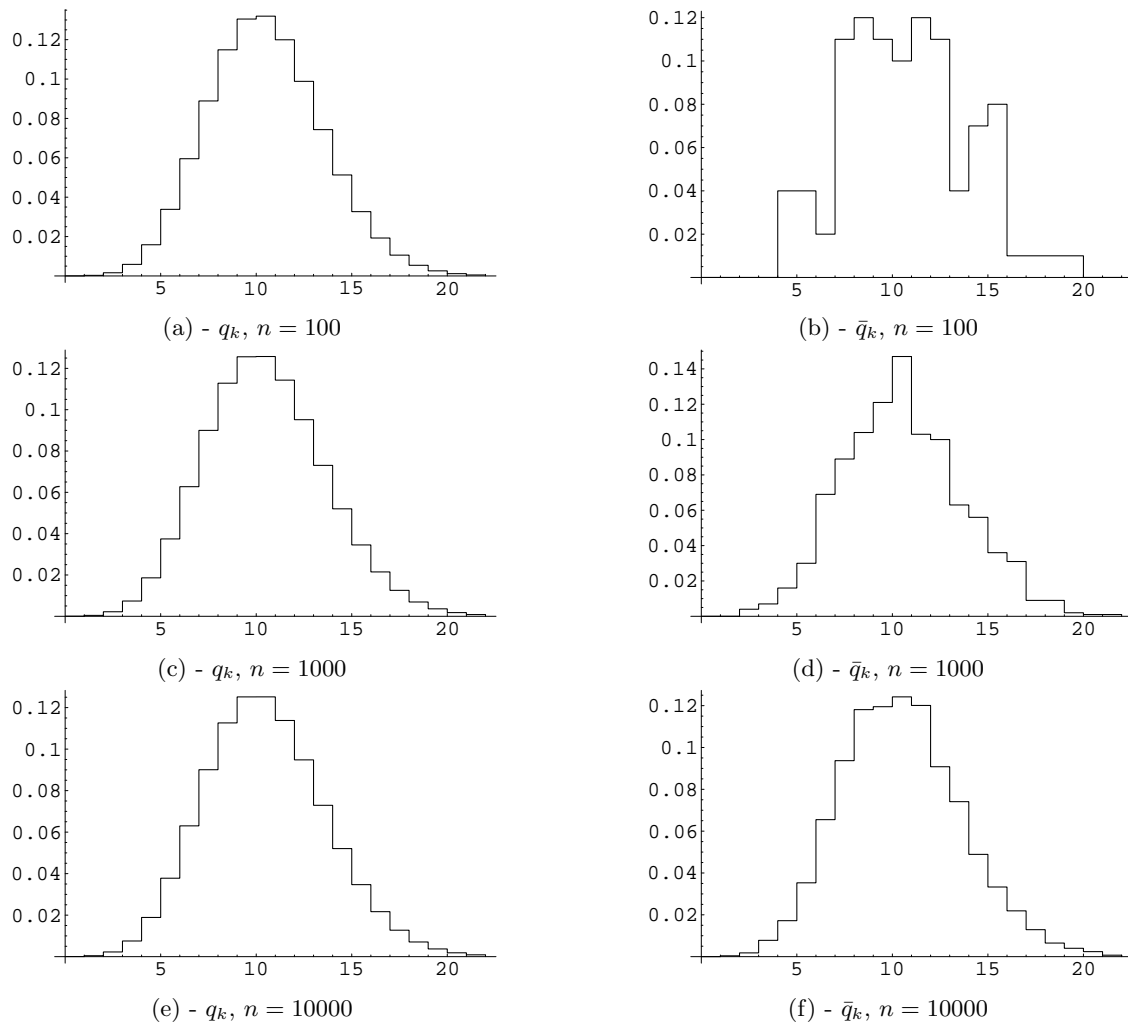


Figura 55.

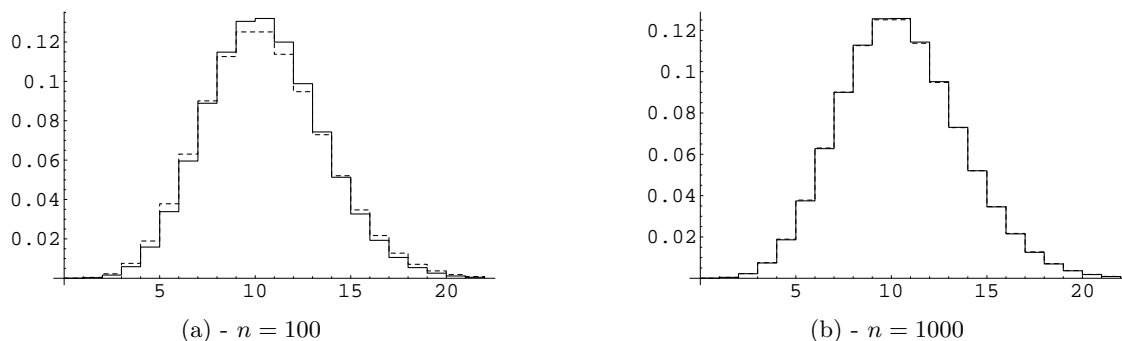


Figura 56.

46 Grafi casuali con gradi fissati

Se si vuole avere una distribuzione dei gradi sui nodi diversa da quella binomiale rappresentata da (71), bisogna vincolare i gradi in modo diverso. Il modo più drastico consiste nel definire l'insieme \mathcal{G} consistente di tutti i grafi i cui nodi hanno una sequenza di gradi prefissata. Questo corrisponde anche a fissare il numero degli archi. Allora si tratta di inserire un certo numero di archi in modo che la sequenza dei gradi sia rispettata. Come già visto a pag. 6 è disponibile una procedura deterministica che, in base alla sequenza dei gradi, genera un grafo (se la sequenza è ammissibile). Per ottenere un grafo casuale è necessario allora modificare il grafo ottenuto in modo casuale senza però alterare i gradi nei nodi. Il modo per generare un grafo casuale corrisponde allora ad una catena di Markov. In generale il grafo che si ottiene non è connesso. Se si vuole che il grafo sia connesso bisogna aggiungere dei passi nella procedura che verranno illustrati più avanti. Se il grafo non deve essere necessariamente connesso si opera nel seguente modo:

- 1) si genera un grafo G_0 con la sequenza di gradi desiderata; sia $r = 0$
- 2) da G_r si genera casualmente un altro grafo G_{r+1} con la stessa sequenza di gradi di G_r , si pone $r := r + 1$ e si ripete questo passo un numero sufficientemente elevato di volte.

Il passo 1) viene effettuato con la procedura di pag. 6. Nel passo 2) si scelgono a caso due archi non incidenti (i, h) e (j, k) di G_r , si rimuovono e al loro posto si mettono gli archi (i, h) e (j, k) oppure gli archi (i, k) e (j, h) , con scelta equiprobabile fra i due casi. Sia G' questo grafo. Se G' non contiene archi multipli, allora G' viene accettato e si pone $G_{r+1} = G'$, altrimenti il nuovo grafo non viene accettato e si pone semplicemente $G_{r+1} = G_r$.

Se invece si vuole ottenere un grafo connesso, si controlla preliminarmente che nessun nodo abbia grado zero e che il numero di archi sia almeno $n - 1$. Se queste due condizioni non sono verificate il grafo non può essere connesso. Poi fra il passo 1) e il passo 2) si inserisce un passo intermedio in cui si scelgono due archi da due componenti connesse diverse. Una delle due componenti non deve essere un albero e l'arco va scelto fra le corde di un generico albero di supporto. Siano (i, j) e (h, k) i due archi. I due archi vengono rimossi e al loro posto si introducono gli archi (i, h) e (j, k) . La scelta di uno dei due archi fra le corde impedisce che una

delle due componenti venga sconnessa e quindi c'è la garanzia che il numero delle componenti connesse cali di uno dopo l'operazione. Inoltre si è anche sicuri che non vengono creati archi multipli, dato che gli archi appartenevano a componenti connesse diverse. Si noti che deve esistere almeno una componente connessa che non sia un albero, altrimenti sarebbe $m < n - 1$, condizione esclusa preliminarmente. Questo passo viene ripetuto tante volte quante sono le componenti connesse meno 1 (ad ogni passo si fondono due componenti connesse). Poi si esegue il passo 2) che viene modificato aggiungendo anche una verifica di connessione sul grafo G' . Se G' non è connesso si pone $G_{r+1} = G_r$.

Il meccanismo delineato è pertanto, in entrambi i casi, una catena di Markov su \mathcal{G} e ogni grafo $G \in \mathcal{G}$ è uno stato. Se esiste una transizione da un grafo G' ad un grafo G'' , cioè è possibile ottenere G'' da G' tramite uno scambio, si vede subito che tale scambio deve essere unico. C'è la transizione se G' e G'' differiscono solo per due archi e questi due archi sono incidenti negli stessi quattro nodi sia per G' che per G'' . La transizione è simmetrica perché si può riottenere G' da G'' tramite lo stesso scambio. La probabilità di tale transizione è uguale a $1/(m(m-1))$. Infatti vi sono in totale $m(m-1)/2$ coppie di archi da scegliere per effettuare lo scambio e per ogni coppia sono possibili due scambi (da (i, j) e (h, k) si può ottenere sia (i, h) e (j, k) come anche (i, k) e (j, h)). Quindi in totale gli scambi sarebbero $m(m-1)$. Di questi un certo numero non è valido. Per ogni scelta di uno scambio non valido si ha una transizione $G' \rightarrow G'$. Quindi la probabilità di una autotransizione è data dal numero di scambi non validi diviso per $m(m-1)$, mentre la probabilità della transizione $G' \rightarrow G''$ è data da $1/(m(m-1))$, trattandosi di un'unica possibilità di scambio.

La matrice di transizione è pertanto simmetrica, cosa che implica l'esistenza di una probabilità stazionaria uniforme. Inoltre la presenza di autotransizioni con probabilità strettamente positiva implica la non periodicità della catena. Che esistano in ogni caso autotransizioni deriva dalla presenza di almeno uno scambio non valido in presenza degli archi (i, j) e (j, k) , ed una tale terna deve esistere in un grafo con almeno un nodo di grado maggiore di 1. Se invece tutti i nodi hanno grado 1 o 0, si vede che le tre coppie di archi $((i, j), (k, h))$, $((i, h), (k, j))$ e $((i, k), (j, h))$ formano una cricca di transizioni possibili e quindi la catena non può essere periodica. È stato inoltre dimostrato che la catena è irriducibile, cioè ogni grafo si riesce ad ottenere da ogni altro grafo tramite operazioni di scambio. Come conseguenza la probabilità stazionaria uniforme è anche la probabilità limite uniforme. Ogni grafo in \mathcal{G} viene quindi campionato con probabilità uniforme.

Rimane da valutare con che velocità si raggiunge la probabilità limite, fatto che determina il numero di ripetizione del passo 2). Non sono al momento note delle limitazioni teoriche polinomiali. L'evidenza computazionale è comunque buona.

Come esempio si consideri la sequenza di gradi $(3, 3, 2, 2, 2)$ per la quale $m = 6$. Vi sono in totale 7 grafi (necessariamente connessi) con questa sequenza, evidenziati in Fig. 57 come stati di una catena di Markov. Le transizioni sono indicate, ma per semplicità non sono indicate le autotransizioni. Ogni transizione ha probabilità $1/30$. Il secondo autovalore vale 0.8666 (presente con molteplicità 4) e quindi la convergenza avviene in tempi ragionevoli. Nell'esempio sono richieste 40 iterazioni per avere una probabilità uniforme alla terza cifra decimale.

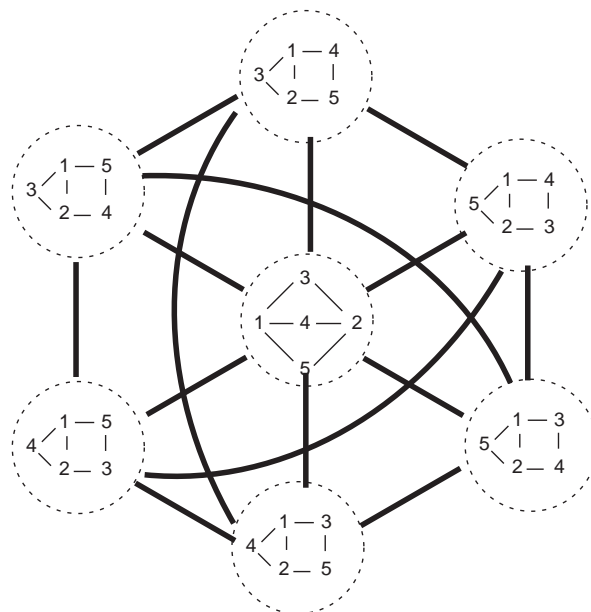


Figura 57. Catena markoviana per la generazione di un grafo casuale con gradi (3, 3, 2, 2, 2)

Per la sequenza di gradi (3, 4, 2, 2, 3, 4) vi sono 17 grafi diversi (anche in questo caso necessariamente connessi). Eseguendo l'algoritmo con 100 transizioni markoviane prima di generare il grafo, e ripetendo 1000 volte l'algoritmo si ottengono i 17 grafi secondo le seguenti frequenze

$$\{50, 57, 57, 68, 51, 57, 68, 61, 57, 57, 47, 67, 58, 54, 57, 67, 67\}$$

quindi in modo abbastanza uniforme. La media teorica è $1000/17 = 58.82$.

47 Funzioni generatrici in generale

Un altro modo di specificare il tipo di grafi da generare riguarda la probabilità con cui i gradi sono distribuiti nel grafo. A questo scopo è opportuno introdurre il concetto di funzione generatrice. Data una successione di valori p_k , $k = 0, \dots, \infty$ (la successione può comunque anche essere finita) si definisce *funzione generatrice* della successione la funzione:

$$F(x) = \sum_{k \geq 0} p_k x^k$$

Nel seguito tratteremo il caso in cui i valori p_k sono probabilità, ma la definizione è del tutto generale e i numeri p_k possono essere qualsiasi. In campo combinatorio le funzioni generatrici intervengono in modo molto efficiente per tutti i problemi in cui bisogna "contare" oggetti e in questo caso p_k rappresenta quanti sono gli oggetti definiti da un particolare parametro k . Per fare un esempio semplice p_k potrebbe essere il

numero di sottoinsiemi di un insieme di cardinalità k . Sappiamo che in questo caso $p_k = 2^k$. Spesso però non è così elementare trovare una formula per p_k . Le funzioni generatrici intervengono proprio per trovare una formula chiusa nonché altre proprietà interessanti della successione.

Anche se la definizione presuppone la convergenza della serie nel caso di infiniti termini, spesso non è importante che la serie converga. Le operazioni formali che vengono effettuate prescindono dalla proprietà di convergenza. Inizialmente conviene mettere in evidenza alcune proprietà che legano operazioni sulla funzione generatrice ad operazioni sui coefficienti. Indichiamo con

$$\langle p_k \rangle \leftrightarrow F(x) \quad \text{se} \quad F(x) = \sum_{k \geq 0} p_k x^k$$

Sia $G(x) = \sum_{k \geq 0} q_k x^k$. Allora

$$\langle \alpha p_k + \beta q_k \rangle \leftrightarrow \alpha F(x) + \beta G(x)$$

Se $G(x) = x F(x)$, allora

$$G(x) = x \sum_{k \geq 0} p_k x^k = \sum_{k \geq 0} p_k x^{k+1} = \sum_{k \geq 1} p_{k-1} x^k$$

quindi

$$\langle p_k \rangle \leftrightarrow F(x) \quad \text{e} \quad \langle 0, p_0, p_1, \dots \rangle \leftrightarrow x F(x)$$

cioè moltiplicare per x una funzione generatrice corrisponde a spostare a destra di una posizione i coefficienti, mettendo 0 per il coefficiente di grado 0. Se $p_0 = 0$ allora

$$\langle p_1, p_2, \dots \rangle \leftrightarrow \frac{F(x)}{x}$$

e se $p_0 \neq 0$ si ha

$$\langle p_1, p_2, \dots \rangle \leftrightarrow \frac{F(x) - p_0}{x}$$

Si vede subito che vale

$$\langle p_k a^k \rangle \leftrightarrow F(ax) \tag{76}$$

Sia $F^{(n)}(x)$ la derivata n -ma di $F(x)$ calcolata in x (indichiamo come d'abitudine $F'(x) = F^{(1)}(x)$). Allora

$$F'(x) = \sum_{k \geq 0} k p_k x^{k-1} = \sum_{k \geq 1} k p_k x^{k-1} = \sum_{k \geq 0} (k+1) p_{k+1} x^k$$

cioè

$$\langle (k+1) p_{k+1} \rangle \leftrightarrow F'(x) \tag{77}$$

ed anche

$$\begin{aligned} \langle k p_k \rangle &\leftrightarrow x F'(x) \\ \langle k^2 p_k \rangle &\leftrightarrow x \frac{d(x F'(x))}{dx} = x F'(x) + x^2 F^{(2)}(x) \end{aligned} \tag{78}$$

Si noti che i coefficienti p_k sono calcolabili come

$$p_k = \frac{F^{(k)}(0)}{k!} \quad (79)$$

$$\int_0^x F(y) dy = \int_0^x \sum_{k \geq 0} p_k y^k dy = \sum_{k \geq 0} p_k \int_0^x y^k dy = \sum_{k \geq 0} \frac{p_k}{k+1} x^{k+1} = \sum_{k \geq 1} \frac{p_{k-1}}{k} x^k$$

quindi

$$\langle 0, \frac{p_{k-1}}{k} \rangle \leftrightarrow \int_0^x F(y) dy$$

Una proprietà particolarmente importante è la seguente

$$F(x)G(x) = \sum_{h \geq 0} p_h x^h \sum_{j \geq 0} q_j x^j = \sum_{h \geq 0, j \geq 0} p_h q_j x^{h+j}$$

e, raccogliendo tutti i termini per cui $h + j = k$

$$F(x)G(x) = \sum_{k \geq 0} \left(\sum_{0 \leq h \leq k} p_h q_{k-h} \right) x^k \quad (80)$$

cioè

$$\langle \sum_{0 \leq h \leq k} p_h q_{k-h} \rangle \leftrightarrow F(x)G(x) \quad (81)$$

La successione $\sum_{0 \leq h \leq k} p_h q_{k-h}$, $k = 0, \dots$ prende il nome di *convoluzione*. Quindi (81) indica che la convoluzione delle successioni corrisponde al prodotto delle funzioni generatrici. Vediamo ora alcune funzioni generatrici di uso frequente:

$$F(x) = \frac{1}{1-x} \implies F^{(k)}(x) = \frac{k!}{(1-x)^k}$$

da cui, in base a (79) $p_k = 1$, quindi

$$\langle 1, 1, \dots \rangle \leftrightarrow \frac{1}{1-x}$$

e, applicando (76) si ha

$$\langle a^k \rangle \leftrightarrow \frac{1}{1-ax}$$

Applicando (80) si ha

$$\frac{F(x)}{1-x} = \sum_{k \geq 0} \left(\sum_{0 \leq h \leq k} p_h \right) x^k \quad (82)$$

e, in particolare,

$$\frac{1}{(1-x)^2} = \sum_{k \geq 0} (k+1) x^k$$

ottenibile anche da (77) e (78). Siccome

$$\frac{d^k}{dx^k} \frac{1}{(1-x)^m} = m(m+1) \cdots (m+k-1) \frac{1}{(1-x)^{m+k}}$$

il coefficiente p_k è

$$p_k = \frac{m(m+1) \cdots (m+k-1)}{k!} = \binom{m+k-1}{k} = \binom{m+k-1}{m-1} \quad (83)$$

Presentiamo alcuni esempi per far vedere quale sia la potenza delle funzioni generatrici nel risolvere problemi combinatorici.

Esempio 1

Si voglia trovare una formula per $S(n, h) := \sum_{k=1}^n k^h$, ad esempio per $h = 2$ oppure $h = 3$. Il caso $h = 0$ è banale in quanto ovviamente $S(n, 0) = n$. Il caso $h = 1$ è ben noto e si ottiene facilmente $S(n, 1) = n(n+1)/2$. Usando (78) e (82) congiuntamente, possiamo ottenere $S(n, h)$ per ogni valore di h . Qui ci limitiamo ai casi $h = 2$ e $h = 3$. Per valori superiori bisogna eseguire derivate sempre più complicate. Da (78) abbiamo

$$\langle k p_k \rangle \leftrightarrow x F'(x), \quad \langle k^2 p_k \rangle \leftrightarrow x F'(x) + x^2 F^{(2)}(x)$$

e

$$\langle k^3 p_k \rangle \leftrightarrow x \frac{d}{dx} (x F'(x) + x^2 F^{(2)}(x)) = x F'(x) + 3x^2 F^{(2)}(x) + x^3 F^{(3)}(x)$$

In questo caso $p_k = 1$ e quindi $F(x) = 1/(1-x)$ e allora si ha

$$\langle k \rangle \leftrightarrow \frac{x}{(1-x)^2}, \quad \langle k^2 \rangle \leftrightarrow \frac{x}{(1-x)^2} + \frac{2x^2}{(1-x)^3}$$

e

$$\langle k^3 \rangle \leftrightarrow \frac{x}{(1-x)^2} + \frac{6x^2}{(1-x)^3} + \frac{6x^3}{(1-x)^4}$$

Dividendo queste funzioni generatrici per $(1-x)$ si ottiene una funzione generatrice in cui ogni coefficiente è la somma dei coefficienti originali. Allora il coefficiente n -mo di

$$\frac{x}{(1-x)^3}$$

deve essere $S(n, 1)$. Il coefficiente n -mo di $1/(1-x)^3$ è $\binom{3+n-1}{3-1}$ (formula (83)). Moltiplicare per x equivale a spostare a destra la successione, per cui

$$S(n, 1) = \binom{3+(n-1)-1}{3-1} = \binom{n+1}{2} = \frac{n(n+1)}{2}$$

A questo punto possiamo calcolare i casi più difficili $h = 2$ e $h = 3$. Per $h = 2$ dobbiamo valutare il coefficiente n -mo di

$$\frac{x}{(1-x)^3} + \frac{2x^2}{(1-x)^4}$$

Tabella 1.

Alcune funzioni generatrici di uso frequente

successione	funzione generatrice	forma chiusa
$\langle 1, 0, 0, 0, \dots \rangle$	$\sum_{k \geq 0} [k = 0] x^k$	1
$\langle 0, \dots, 0, 1, 0, \dots \rangle$	$\sum_{k \geq 0} [k = m] x^k$	x^m
$\langle 1, 1, \dots, 1, 0, \dots \rangle$	$\sum_{k \geq 0} [k \leq m] x^k$	$\frac{1 - x^{m+1}}{1 - x}$
$\langle 1, 1, 1, \dots \rangle$	$\sum_{k \geq 0} x^k$	$\frac{1}{1 - x}$
$\langle 1, -1, 1, -1, 1, -1, \dots \rangle$	$\sum_{k \geq 0} (-1)^k x^k$	$\frac{1}{1 + x}$
$\langle 1, 2, 4, 8, 16, \dots \rangle$	$\sum_{k \geq 0} 2^k x^k$	$\frac{1}{1 - 2x}$
$\langle 1, a, a^2, a^3, \dots \rangle$	$\sum_{k \geq 0} a^k x^k$	$\frac{1}{1 - ax}$
$\langle 1, -a, a^2, -a^3, \dots \rangle$	$\sum_{k \geq 0} (-1)^k a^k x^k$	$\frac{1}{1 + ax}$
$\langle 1, 0, 1, 0, 1, 0, \dots \rangle$	$\sum_{k \geq 0} [k \text{ pari}] x^k$	$\frac{1}{1 - x^2}$
$\langle 0, 1, 0, 1, 0, 1, \dots \rangle$	$\sum_{k \geq 0} [k \text{ dispari}] x^k$	$\frac{x}{1 - x^2}$
$\langle 1, 2, 3, 4, 5, 6, \dots \rangle$	$\sum_{k \geq 0} (k + 1) x^k$	$\frac{1}{(1 - x)^2}$
$\langle 1, \binom{m+1}{m}, \binom{m+2}{m}, \dots \rangle$	$\sum_{k \geq 0} \binom{m+k}{m} x^k$	$\frac{1}{(1 - x)^{m+1}}$
$\langle 1, m+1, \binom{m+2}{2}, \binom{m+3}{3}, \dots \rangle$	$\sum_{k \geq 0} \binom{m+k}{k} x^k$	$\frac{1}{(1 - x)^{m+1}}$
$\langle 1, 2, 1, 0, 0, \dots \rangle$	$\sum_{k \geq 0} \binom{2}{k} x^k$	$(1 + x)^2$
$\langle 1, m, \binom{m}{2}, \binom{m}{3}, \dots \rangle$	$\sum_{k \geq 0} \binom{m}{k} x^k$	$(1 + x)^m$
$\langle 1, 1, 2, 5, 14, 42, \dots \rangle$	$\sum_{k \geq 0} \frac{1}{k+1} \binom{2k}{k} x^k$	$\frac{1 - \sqrt{1 - 4x}}{2x}$
$\langle 0, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots \rangle$	$\sum_{k \geq 1} \frac{1}{k} x^k$	$\ln \frac{1}{1 - x}$
$\langle 0, 1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \dots \rangle$	$\sum_{k \geq 1} (-1)^{k+1} \frac{1}{k} x^k$	$\ln(1 + x)$
$\langle 1, 1, \frac{1}{2}, \frac{1}{6}, \frac{1}{24}, \dots \rangle$	$\sum_{k \geq 1} \frac{1}{k!} x^k$	e^x

che è

$$S(n, 2) = \binom{3 + (n-1) - 1}{3-1} + 2 \binom{4 + (n-2) - 1}{4-1} = \frac{n(n+1)}{2} + 2 \binom{n+1}{3} =$$

$$\frac{n(n+1)}{2} + \frac{n(n+1)(n-1)}{3} = \frac{n(n+1)(2n+1)}{6}$$

Applicando lo stesso principio per $S(n, 3)$ si deve valutare il coefficiente n -mo di

$$\frac{x}{(1-x)^3} + \frac{6x^2}{(1-x)^4} + \frac{6x^3}{(1-x)^5}$$

cioè

$$\frac{n(n+1)}{2} + 6 \binom{n+1}{3} + 6 \binom{5 + (n-3) - 1}{5-1} =$$

$$\frac{n(n+1)}{2} + n(n+1)(n-1) + \frac{n(n+1)(n-1)(n-2)}{4} = \frac{n^2(n+1)^2}{4}$$

Si noti che si è ottenuto anche il seguente sorprendente risultato

$$\sum_{k=1}^n k^3 = \left(\sum_{k=1}^n k \right)^2$$

■

Esempio 2

Come secondo esempio si supponga di avere una successione di cui si conosce la formula ricorsiva

$$p_n = 5p_{n-1} - 6p_{n-2}, \quad p_0 = 0, \quad p_1 = 1$$

e si voglia trovare una formula chiusa per p_n . Calcolando esplicitamente alcuni valori si ottiene

$$p_2 = 5, \quad p_3 = 19, \quad p_4 = 65, \quad p_5 = 211, \quad p_6 = 665$$

La successione non suggerisce nessuna formula da provare per tentativi. Conviene pensare alla successione definita per ogni n (anche negativo) e porre $p_n = 0$ per $n < 0$. La ricorsione è vera per ogni n tranne $n = 1$ (infatti la ricorsione darebbe $p_1 = 0$ mentre $p_1 = 1$). A questo punto si moltiplica ogni termine della successione per x^n ottenendo

$$p_n x^n = 5p_{n-1} x^n - 6p_{n-2} x^n \quad n \neq 1, \quad p_1 x^1 = 1 x^1$$

e si sommano su tutti gli n

$$\sum_n p_n x^n = \sum_{n \neq 1} 5p_{n-1} x^n - \sum_{n \neq 1} 6p_{n-2} x^n + x$$

Siccome $p_{1-1} = 0$ e $p_{1-2} = 0$ possiamo estendere le sommatorie a tutti gli n e quindi

$$\sum_n p_n x^n = 5 \sum_n p_{n-1} x^n - 6 \sum_n p_{n-2} x^n + x$$

Si indichi con $F(x) = \sum_n p_n x^n$ la funzione generatrice della successione in esame.

$$\begin{aligned} F(x) &= 5x F(x) - 6x^2 F(x) + x \\ F(x) &= \frac{x}{1 - 5x + 6x^2} = \frac{x}{(1 - 2x)(1 - 3x)} \end{aligned}$$

Adesso si cercano coefficienti α e β tali che

$$F(x) = \frac{x}{(1 - 3x)(1 - 2x)} = \frac{\alpha}{1 - 3x} + \frac{\beta}{1 - 2x}$$

da cui in base ai risultati precedenti

$$p_n = \alpha 3^n + \beta 2^n$$

Siccome $p_0 = 0$ e $p_1 = 1$ si ottiene $0 = \alpha + \beta$ e $1 = 3\alpha + 2\beta$ da cui $\alpha = 1$ e $\beta = -1$. Quindi

$$p_n = 3^n - 2^n$$

■

Come esercizio si calcolino i coefficienti della successione

$$p_n = p_{n-1} - p_{n-2}, \quad p_0 = 0, \quad p_1 = 1$$

Il calcolo immediato dalla formula ricorsiva porta alla successione $0, 1, 1, 0, -1, -1, 0, 1, 1, -1, -1, \dots$ che è periodica con periodo 6. La formula chiusa è data da $p_n = 2/\sqrt{3} \sin(n\pi/3)$. Si trovi questo risultato applicando il metodo visto.

Come ulteriore esercizio si calcoli la formula chiusa per i numeri di Fibonacci definiti da

$$F_n = F_{n-1} + F_{n-2}, \quad F_0 = 0, \quad F_1 = 1$$

Il risultato è

$$F_n = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^n$$

48 Funzioni generatrici di probabilità

Fra le funzioni generatrici vi sono alcune che hanno un significato particolare in quanto i valori p_k sono valori di probabilità. Sia X una variabile aleatoria che assume valori su $\{0, 1, \dots\}$ e sia $p_k = \Pr\{X = k\}$. La *funzione generatrice della variabile aleatoria* X è

$$F_X = \sum_{k \geq 0} p_k x^k$$

da cui

$$F_X(1) = \sum_{k \geq 0} p_k = 1$$

Inoltre

$$F'_X(x) = \sum_{k \geq 1} k p_k x^{k-1} \implies F'_X(1) = \sum_{k \geq 1} k p_k = E[X] = \bar{X}$$

e

$$F''_X(x) = \sum_{k \geq 2} k(k-1) p_k x^{k-2} = \sum_{k \geq 2} k^2 p_k x^{k-2} - \sum_{k \geq 2} k p_k x^{k-2} \implies$$

$$F''_X(1) = E[X^2] - \bar{X} = E[(X - \bar{X})^2] + \bar{X}^2 - \bar{X} \quad (84)$$

Se $Z = X + Y$ con $p_k = \Pr\{X = k\}$ e $q_k = \Pr\{Y = k\}$ e funzioni generatrici

$$F_X(x) = \sum_{k \geq 0} p_k x^k, \quad F_Y(x) = \sum_{k \geq 0} q_k x^k$$

la funzione generatrice della variabile Z si ottiene come

$$F_Z(x) = F_X(x) F_Y(x)$$

in quanto

$$\Pr\{Z = k\} = \sum_{0 \leq h \leq k} p_h q_{k-h}$$

In particolare una variabile deterministica che vale k con probabilità 1 ha la funzione generatrice $F(x) = x^k$. Quindi se $F(x)$ è la funzione generatrice della variabile casuale X , la funzione generatrice $x^k F(x)$ corrisponde alla variabile casuale $X + k$.

Diamo alcuni esempi dell'utilità delle funzioni generatrici per il calcolo delle probabilità.

Esempio 3

Come primo esempio si immagina di lanciare 5 dadi tutti assieme e di fare la somma dei numeri usciti. Qual è la probabilità p_k di ottenere la somma uguale a k ? Ovviamente $p_k = 0$ per $k < 5$. Il calcolo di p_5 non è difficile. Infatti la somma 5 si può ottenere in un solo modo, in cui esce l'1 su tutti i dadi, quindi $p_5 = (1/6)^5$. Il calcolo di p_6 è leggermente più complicato. Deve uscire 4 volte l'1 e una volta il 2. Ci sono 5 modi in cui questo può avvenire, quindi $p_6 = 5 (1/6)^5$. Il calcolo di p_7 comincia a presentare problemi. La somma 7 si ottiene se esce 3 volte l'1 e 2 volte il 2 e vi sono $\binom{5}{3}$ modi in cui questo può avvenire, ma anche si ottiene con l'uscita del 3 (una volta) e 4 volte l'uscita dell'1 e questo può avvenire in 5 modi diversi. Quindi

$$p_7 = \left(\binom{5}{3} + 5 \right) \left(\frac{1}{6} \right)^5 = 15 \left(\frac{1}{6} \right)^5$$

Proseguendo in questo modo il calcolo diventa sempre più complicato per i modi numerosi in cui una certa somma può ripartirsi sui cinque dadi. Adottiamo invece l'approccio tramite funzioni generatrici. La funzione generatrice $F(x)$ di un singolo lancio di dadi è

$$F(x) = \frac{1}{6}x + \frac{1}{6}x^2 + \frac{1}{6}x^3 + \frac{1}{6}x^4 + \frac{1}{6}x^5 + \frac{1}{6}x^6$$

e quindi la funzione generatrice $F_5(x)$ di cinque lanci simultanei è

$$F_5(x) = (F(x))^5 = \left(\frac{1}{6}\right)^5 x^5 (1 + x + x^2 + x^3 + x^4 + x^5)^5 \quad (85)$$

Si noti che il calcolo di come una somma si possa ripartire sui vari dadi *viene automaticamente fatto eseguendo la moltiplicazione*. Il coefficiente di x^k verrà calcolato appunto da come i coefficienti dei singoli termini di potenza x^h verranno moltiplicati e sommati fra di loro se la somma degli esponenti dà proprio k . Il calcolo di (85) può essere effettuato direttamente, ma è un po' noioso. Possiamo anche scrivere

$$F(x) = \frac{x}{6} (1 + x + x^2 + x^3 + x^4 + x^5) = \frac{x}{6} \frac{1 - x^6}{1 - x}$$

per cui

$$F_5(x) = \frac{x^5}{6^5} \frac{(1 - x^6)^5}{(1 - x)^5}$$

Usando la tabella delle funzioni generatrici abbiamo

$$\begin{aligned} F_5(x) &= \frac{x^5}{6^5} (1 - x^6)^5 \sum_{k \geq 0} \binom{5+k-1}{4} x^k = \frac{x^5}{6^5} \sum_{h=0}^5 (-1)^h \binom{5}{h} x^{6h} \sum_{k \geq 0} \binom{5+k-1}{4} x^k = \\ &= \frac{1}{6^5} \sum_{h=0}^5 \sum_{k \geq 0} (-1)^h \binom{5}{h} \binom{4+k}{4} x^{6h+k+5} \end{aligned}$$

che possiamo riscrivere come

$$\frac{1}{6^5} \sum_k \sum_{h=0}^5 [k \geq 0] (-1)^h \binom{5}{h} \binom{4+k}{4} x^{6h+k+5}$$

e ponendo $n = 6h + k + 5$, la stessa espressione è uguale a

$$\frac{1}{6^5} \sum_n \sum_{h=0}^5 [n - 6h - 5 \geq 0] (-1)^h \binom{5}{h} \binom{4+n-6h-5}{4} x^n$$

Quindi la probabilità p_n è data da

$$p_n = \frac{1}{6^5} \sum_{h=0}^{\min\{5, \lfloor (n-5)/6 \rfloor\}} (-1)^h \binom{5}{h} \binom{n-6h-1}{4}$$

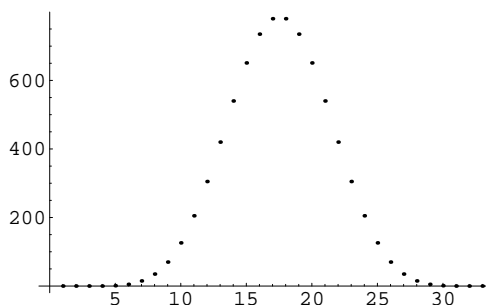


Figura 58.

Se vogliamo calcolare p_{19} dobbiamo valutare la somma per $h = 0, 1, 2$. Quindi

$$p_{19} = \frac{1}{6^5} \left(\binom{5}{0} \binom{18}{4} - \binom{5}{1} \binom{12}{4} + \binom{5}{2} \binom{6}{4} \right) = \frac{735}{6^5} = 0.0945216$$

Si noti che se si applica la formula per valori superiori a 30 si ottiene 0, cioè i vari termini si cancellano. Ad esempio per $n = 35$ si ottengono nella sommatoria i seguenti termini

$$\{46376, -102375, 73150, -18200, 1050, -1\}$$

la cui somma dà 0. Il grafico delle probabilità è in Fig. 58 (i valori dell'asse delle ordinate vanno divisi per 6^5). ■

Esempio 4

In questo esempio si fa vedere l'utilità delle funzioni generatrici per trattare la ricorsività. Si supponga di effettuare il seguente gioco ai dadi: si tira un dado e si prendono i punti ottenuti. Inoltre, se esce 3 o 6, si può tirare ancora con la stessa regola, continuando ad accumulare i punti. Ci si chiede con che probabilità p_k si ottengono k punti. Intanto si può anticipare che $p_k = 0$ se k è multiplo di 3, perché in questi casi si deve tirare ancora e ci si fermerà soltanto quando l'ultimo punteggio non è un multiplo di 3. Inoltre $p_1 = p_2 = 1/6$ perché c'è un unico modo di ottenere 1 o 2. Per altri valori si può ragionare tenendo conto dei casi possibili. Ad esempio si può ottenere 4 alla prima uscita con probabilità $1/6$ oppure alla seconda avendo fatto 3 e 1 nei due tiri e questo avviene con probabilità $1/36$. Quindi $p_4 = 1/6 + 1/36 = 7/36$. Il punteggio 10 si può ottenere come 6+4 oppure come 6+3+1 oppure come 3+6+1 oppure come 3+3+4 oppure come 3+3+3+1, che porta ad una probabilità

$$\frac{1}{36} + \frac{1}{216} + \frac{1}{216} + \frac{1}{216} + \frac{1}{1296} = \frac{55}{1296}$$

Però se si dovesse procedere in questo modo per ogni possibile giocata sarebbe difficile tener conto di tutte le possibilità create dalla ricorsività del gioco. Con una funzione generatrice la ricorsione si esprime invece

in modo molto efficace. Se indichiamo con $F(x)$ la funzione generatrice dei p_k , deve valere l'uguaglianza

$$F(x) = \frac{1}{6}x + \frac{1}{6}x^2 + \frac{1}{6}x^3 F(x) + \frac{1}{6}x^4 + \frac{1}{6}x^5 + \frac{1}{6}x^6 F(x) \quad (86)$$

Il termine $1/6 x^3 F(x)$ sta ad indicare che se esce il 3 con probabilità $1/6$, intanto si accumulano 3 punti e di questo si tiene conto nel termine x^3 e poi è come se si ricominciasse nuovamente il gioco. Quindi ai 3 punti iniziali si sommeranno quelli che si otterranno dopo, corrispondenti ad un gioco che parte da zero. Questa somma corrisponde al prodotto fra x^3 e $F(x)$.

Prima di calcolare la funzione generatrice si può calcolare velocemente il valore atteso derivando (86).

$$F'(x) = \frac{1}{6} + \frac{2}{6}x + \frac{3}{6}x^2 F(x) + \frac{1}{6}x^3 F'(x) + \frac{4}{6}x^3 + \frac{5}{6}x^4 + \frac{6}{6}x^5 F(x) + \frac{1}{6}x^6 F'(x)$$

da cui, ricordando che il valor medio è $F'(1)$ e $F(1) = 1$,

$$F'(1) = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{1}{6}F'(1) + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} + \frac{1}{6}F'(1) \implies F'(1) = \frac{21}{4}$$

Per il calcolo della funzione generatrice si ha da (86)

$$F(x) = \frac{\frac{1}{6}x + \frac{1}{6}x^2 + 16x^4 + \frac{1}{6}x^5}{1 - \frac{1}{6}x^3 - \frac{1}{6}x^6} = \frac{1}{6} \frac{x + x^2 + x^4 + x^5}{(1 - \frac{1}{2}x^3)(1 + \frac{1}{3}x^3)}$$

Dobbiamo calcolare due costanti α e β tali che valga l'identità

$$\frac{1}{(1 - \frac{1}{2}x^3)(1 + \frac{1}{3}x^3)} = \frac{\alpha}{1 - \frac{1}{2}x^3} + \frac{\beta}{1 + \frac{1}{3}x^3} = \frac{\alpha + \frac{1}{3}\alpha x^3 + \beta - \frac{1}{2}\beta x^3}{(1 - \frac{1}{2}x^3)(1 + \frac{1}{3}x^3)}$$

Affinché valga l'identità deve essere

$$\alpha + \beta = 1, \quad \frac{1}{3}\alpha - \frac{1}{2}\beta = 0$$

da cui $\alpha = 3/5$ e $\beta = 2/5$. Allora

$$\begin{aligned} F(x) &= \frac{1}{6}(x + x^2 + x^4 + x^5) \left(\frac{\frac{3}{5}}{1 - \frac{1}{2}x^3} + \frac{\frac{2}{5}}{1 + \frac{1}{3}x^3} \right) \\ F(x) &= \frac{1}{6}(x + x^2 + x^4 + x^5) \left(\frac{3}{5} \sum_{k \geq 0} \left(\frac{1}{2}\right)^k x^{3k} + \frac{2}{5} \sum_{k \geq 0} \left(-\frac{1}{3}\right)^k x^{3k} \right) \\ F(x) &= \sum_{k \geq 0} \frac{1}{6} \left(\frac{3}{5} \left(\frac{1}{2}\right)^k + \frac{2}{5} \left(-\frac{1}{3}\right)^k \right) (x^{3k+1} + x^{3k+2} + x^{3k+4} + x^{3k+5}) \end{aligned}$$

Se vogliamo calcolare p_{10} dobbiamo valutare per quali valori di k si ha un termine x^{10} . Questo avviene quando $k = 3$ (da $3k + 1 = 10$) oppure quando $k = 2$ (da $3k + 4 = 10$) e quindi

$$p_{10} = \frac{1}{6} \left(\frac{3}{5} \left(\frac{1}{2}\right)^3 + \frac{2}{5} \left(-\frac{1}{3}\right)^3 + \frac{3}{5} \left(\frac{1}{2}\right)^2 + \frac{2}{5} \left(-\frac{1}{3}\right)^2 \right) = \frac{55}{1296} = 0.0424383$$

In generale, siccome per ogni k non multiplo di 3, entrano in gioco due termini, e cioè $3h + 1 = k$ e $3(h - 1) + 4 = k$ oppure $3h + 2 = k$ e $3(h - 1) + 5 = k$ e $h = \lfloor k/3 \rfloor$ si ottiene:

$$p_k = \frac{1}{6} \left(\frac{9}{5} \left(\frac{1}{2} \right)^{\lfloor k/3 \rfloor} - \frac{4}{5} \left(-\frac{1}{3} \right)^{\lfloor k/3 \rfloor} \right), \quad k > 3, \quad k \text{ non multiplo di } 3$$

I primi valori sono

$$\begin{aligned} p_1 = p_2 = 0.1667, & \quad p_4 = p_5 = 0.1944, & \quad p_7 = p_8 = 0.0602, \\ p_{10} = p_{11} = 0.0424, & \quad p_{13} = p_{14} = 0.0171, & \quad p_{16} = p_{17} = 0.0099, \\ p_{19} = p_{20} = 0.0045, & \quad p_{22} = p_{23} = 0.0024, & \quad p_{25} = p_{26} = 0.0011 \end{aligned}$$

Come esercizio si calcolino le probabilità di vincere k punti se, all'uscita del 3 o del 6, si guadagna un solo punto (e si tira ancora una volta). ■

Esempio 5

In questo terzo esempio la ricorsione entra in un modo più complesso. Si immagini il seguente gioco: si lancia un dado e se escono i numeri da 1 a 5 si vince 1, mentre se esce il 6 non si vince nulla però vengono offerti due lanci di dado, ognuno dei quali con la stessa regola. Durante il gioco quindi i lanci a disposizione si possono accumulare. Il gioco termina quando non si hanno più lanci a disposizione. Ad esempio: esce il 6 e si hanno a disposizione due lanci, nel primo esce di nuovo il 6 si hanno altri due lanci (quindi in questo momento si hanno 3 lanci a disposizione), poi esce il 4 e si guadagna 1, nel terzo esce il 5 e nel quarto esce il 2. Complessivamente si è vinto 3. La differenza sostanziale rispetto al caso precedente consiste nel fatto che, anziché semplicemente riprendere il gioco ad un'uscita particolare, si hanno a disposizione due giocate.

Si noti come il gioco possa anche essere simulato nel seguente modo: partendo da 1 si esegue un cammino casuale sui numeri interi non negativi partendo dal numero 1. Con probabilità $1/6$ si aumenta di 1 e con probabilità $5/6$ si diminuisce di 1. Ad ogni diminuzione di 1 si guadagna 1. Il gioco termina quando si raggiunge lo 0.

Sia $F(x)$ la funzione generatrice i cui coefficienti p_k rappresentano la probabilità di vincere k . Allora si ha

$$F(x) = \frac{5}{6}x + \frac{1}{6}F^2(x) \tag{87}$$

Il termine $F^2(x)$ sta ad indicare che con probabilità $1/6$ si effettueranno due lanci e quindi si sommeranno le vincite dei due lanci, ovvero si moltiplicano le corrispondenti funzioni generatrici. Da (87) si ha

$$F^2(x) - 6F(x) + 5x = 0$$

che è una equazione di secondo grado in $F(x)$. Risolvendo si ottiene

$$F(x) = 3 \pm \sqrt{9 - 5x}$$

Delle due radici, quella corretta è quella per cui $F(1) = 1$ e allora

$$F(x) = 3 - \sqrt{9 - 5x} = 3 \left(1 - \sqrt{1 - \frac{5}{9}x} \right) \quad (88)$$

Prima di calcolare i coefficienti di (88) conviene calcolare il valore atteso della vincita totale che si può fare facilmente in due modi alternativi. Un calcolo diretto di tipo ricorsivo porta a dire che il valore atteso \bar{v} è uguale a 1 con probabilità $5/6$ e con probabilità $1/6$ è $2\bar{v}$, dato che si hanno a disposizione due giocate fatte da capo. Allora

$$\bar{v} = \frac{5}{6} + \frac{1}{6} 2\bar{v}$$

che porta a $\bar{v} = 5/4$. Altrimenti si derivi (87)

$$F'(x) = \frac{5}{6} + 2 \frac{1}{6} F(x) F'(x)$$

Siccome $\bar{v} = F'(1)$ e $F(1) = 1$ si ottiene nuovamente $\bar{v} = 5/4$.

Il calcolo dei coefficienti di (88) invece non è del tutto semplice. Tuttavia funzioni del genere sono già state studiate e i coefficienti sono spesso disponibili. In questo caso il calcolo dei valori p_k non è proibitivo. Si tratta di calcolare le derivate di $F(x)$ sperando di trovare una formula chiusa per la derivata n -ma. A questo scopo calcoliamo il caso generico

$$G(x) = 1 - \sqrt{1 - ax}$$

si trova

$$\begin{aligned} G'(x) &= \frac{a}{2} (1 - ax)^{-\frac{1}{2}}, & G''(x) &= \left(\frac{a}{2}\right)^2 (1 - ax)^{-\frac{3}{2}}, \\ G^{(3)}(x) &= 3 \left(\frac{a}{2}\right)^3 (1 - ax)^{-\frac{5}{2}}, & G^{(4)}(x) &= 5 \cdot 3 \left(\frac{a}{2}\right)^4 (1 - ax)^{-\frac{7}{2}} \end{aligned}$$

e quindi

$$G^{(n)}(x) = (2n - 2)!_o \left(\frac{a}{2}\right)^n (1 - ax)^{-\frac{2n-1}{2}}$$

dove $n!_o$ è il prodotto dei numeri dispari minori o uguali a n . Da (79) abbiamo

$$p_n = \frac{G^{(n)}(0)}{n!} = \frac{(2n - 2)!_o}{n!} \left(\frac{a}{2}\right)^n, \quad n \geq 1$$

Ora, sia $n!_e$ il prodotto dei numeri pari minori o uguali a n . Si noti che vale $(2n)!_e = n! 2^n$. Quindi si ha

$$(2n)! = (2n)!_e \cdot (2n)!_o = n! 2^n (2n)!_o \implies (2n)!_o = \frac{(2n)!}{n! 2^n} \implies (2n - 2)!_o = \frac{(2n - 2)!}{(n - 1)! 2^{n-1}}$$

e allora

$$p_n = \frac{(2n - 2)!}{n! (n - 1)!} \frac{a^n}{2^{2n-1}} = \frac{1}{n} \binom{2(n - 1)}{n - 1} \frac{a^n}{2^{2n-1}}, \quad n \geq 1$$

che diventa nel nostro caso, con $a = 5/9$ e coefficiente moltiplicativo 3,

$$p_n = 3 \frac{1}{n} \binom{2(n-1)}{n-1} \frac{5^n}{9^n 2^{2n-1}} = \frac{1}{n} \binom{2(n-1)}{n-1} \frac{5^n}{6^{2n-1}}, \quad n \geq 1$$

I numeri

$$C_n = \frac{1}{n+1} \binom{2n}{n}$$

sono molto importanti nell'Analisi Combinatoria e vengono detti *numeri di Catalan*, dal matematico belga che li ha studiati (ma non è stato il primo, già Eulero li aveva trattati). I primi numeri della successione sono

$$C_0 = 1, \quad C_1 = 1, \quad C_2 = 2, \quad C_3 = 5, \quad C_4 = 14, \quad C_5 = 42, \quad C_6 = 132, \quad C_7 = 429, \quad C_8 = 1430$$

Allora possiamo scrivere

$$p_n = C_{n-1} \left(\frac{5}{6}\right)^n \left(\frac{1}{6}\right)^{n-1}, \quad n \geq 1$$

I primi 10 coefficienti sono

$$p_1 = 0.833333, \quad p_2 = 0.115741, \quad p_3 = 0.032150, \quad p_4 = 0.011163, \quad p_5 = 0.004341, \\ p_6 = 0.001809, \quad p_7 = 0.000790, \quad p_8 = 0.000356, \quad p_9 = 0.000166, \quad p_{10} = 0.000078$$

Può essere interessante considerare lo stesso esempio con probabilità p generica di ottenere due giocate e $(1-p)$ di vincere 1. In questo caso si ottiene

$$F(x) = \frac{1 - \sqrt{1 - 4p(1-p)x}}{2p} = \tag{89}$$

$$\frac{1}{2p} \sum_{n \geq 1} \frac{1}{n} \binom{2(n-1)}{n-1} \frac{(2^2 p(1-p))^n}{2^{2n-1}} x^n = \frac{1}{p} \sum_{n \geq 1} C_{n-1} p^n (1-p)^n x^n$$

Il calcolo di $F(1)$ dà luogo ad un risultato sorprendente per $p > 1/2$. Si ricordi che $\sqrt{a^2} = |a|$. Quindi $F(1) = (1-p)/p < 1$ per $p > 1/2$. Non si creda che per $p > 1/2$ si debba prendere l'altra radice (quella con il segno +) dell'equazione

$$F(x) = (1-p)x + pF^2(x)$$

perché il calcolo dei coefficienti porterebbe a valori negativi e quindi non potrebbero rappresentare valori di probabilità anche se la somma fosse uguale ad 1 (si ha $p_0 = 2$ e tutti i p_n , $n > 0$, sono l'opposto di quelli calcolati sopra). Quindi cosa succede per $p > 1/2$? In questo caso fa differenza se si viene pagati alla fine del gioco o dopo ogni giocata? ■

Si cambi il gioco assumendo una perdita di 1 quando esce il 6 (ovvero in generale con probabilità p). Si imponi il calcolo della funzione generatrice considerando un termine x^{-1} che tiene conto della perdita di 1. Si ottiene un risultato forse inaspettato. Lo si spieghi in modo puramente combinatorio. Anche in questo caso, cosa succede per $p > 1/2$?

49 Processi di diramazione

Prima di applicare le funzioni generatrici ai grafi casuali è utile considerare un particolare tipo di processo stocastico, detto *processo di diramazione* (*branching process*) o anche processo di Galton-Watson dai nomi dei due matematici che per primi l'hanno proposto. Galton era interessato a stabilire con che probabilità un cognome si poteva estinguere all'interno di una popolazione, ma il modello matematico ha molte applicazioni interessanti. In particolare ci sono delle forti connessioni con i grafi casuali.

In un processo di diramazione sono definiti dei valori di probabilità p_k , $k = 0, 1, \dots$, (con $\sum_k p_k = 1$) che rappresentano la probabilità per ogni individuo della popolazione di generare k figli.

Si vuole stabilire con che probabilità la popolazione si estingue se inizialmente consiste di un singolo individuo e, nei casi in cui si estingua, con che probabilità la popolazione finale consiste di h individui. Si suppone che nessuno muoia e per estinzione quindi si intende che la popolazione non può più crescere perché tutti gli individui hanno già generato i figli che dovevano generare, includendo ovviamente anche il caso di nessun figlio, altrimenti la popolazione cresce continuamente. Ovviamente, assumendo una vita finita per ogni individuo, una popolazione che non cresce più corrisponde ad una vera e propria estinzione.

Il processo si può rappresentare come un albero con radice e ad ogni vertice v si possono aggiungere altri k vertici adiacenti a v con probabilità p_k . Vogliamo conoscere con che probabilità l'albero ha un numero finito di vertici. L'equazione chiave che permette tale calcolo si basa sulla seguente considerazione: il nodo radice genererà un albero finito se e solo ogni ognuno dei suoi figli genererà un albero finito.

Sia u la probabilità che l'albero sia finito. Se il nodo radice ha k figli, la probabilità che tutti i figli generino alberi finiti è u^k . Tenendo conto del fatto che sono presenti k figli con probabilità p_k perveniamo alla relazione

$$u = \sum_k p_k u^k = F(u) \quad (90)$$

dove $F(u)$ è la funzione generatrice delle probabilità. Si noti che $u = 1$ è sempre soluzione di (90). Da questo fatto Galton e Watson avevano incorrettamente dedotto che la popolazione si estingue con probabilità uno. Tuttavia vi possono essere altre soluzioni con valore $u < 1$. Questo significa che con probabilità $1 - u$ la popolazione continua a crescere senza mai fermarsi.

Consideriamo alcuni semplici casi. Sia $p_k = 0$ per $k > 1$ e sia $p_1 = p$ e $p_0 = 1 - p$. È evidente che in questo caso ci sarà sempre estinzione non appena uno dei figli unici generati non avrà figli. L'equazione (90) diventa

$$u = (1 - p) + p u$$

che ha come unica soluzione $u = 1$. Consideriamo il caso più interessante in cui $p_2 = p$, $p_0 = 1 - p$ e $p_k = 0$, per $k \notin \{0, 2\}$. Allora (90) diventa

$$u = (1 - p) + p u^2 \quad (91)$$

In Fig. 59 sono raffigurate la retta u e la parabola $(1 - p) + p u^2$ per tre valori di p . La retta e la parabola si intersecano sempre per $u = 1$, come già osservato. L'altra radice è maggiore di uno se $p = 1/4$ e in questo

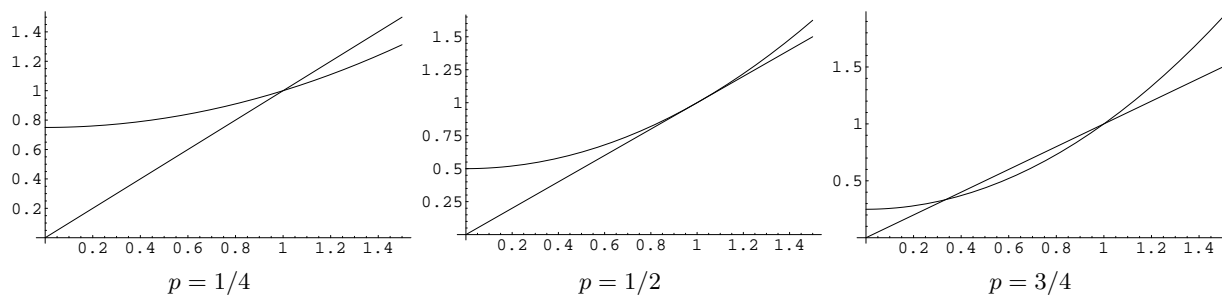


Figura 59.

caso la popolazione si estingue con probabilità uno. Il valore $p = 1/2$ è critico con le due radici coincidenti. Per valori superiori di probabilità si ha una soluzione $u < 1$, in particolare per $p = 3/4$ si ottiene $u = 1/3$, cioè la popolazione si estingue con probabilità $1/3$ e con probabilità $2/3$ cresce senza limiti.

Come esercizio si consideri il caso di probabilità $p_k = 0$ per $k > 2$ (quindi rispetto al caso precedente si può avere $p_1 > 0$) e si trovi che la probabilità critica si ha per $p_0 = p_2$ con estinzione se e solo se $p_2 > p_0$.

Consideriamo la seconda questione riguardante la numerosità della popolazione in caso di estinzione. Siano r_k , $k = 1, 2, \dots$, le probabilità che la popolazione consista di k individui (il caso $k = 0$ è escluso in quanto esiste sempre almeno il capostipite) e sia $H(x)$ la corrispondente funzione generatrice. Sfruttiamo le proprietà ricorsive delle funzioni generatrici. Se la radice ha k figli, ognuno di questi è a sua volta radice di un albero e quindi la discendenza totale della radice è data dalla somma dei discendenti di ognuno dei figli (inclusi i figli stessi), che si traduce per le funzioni generatrici nel loro prodotto, ovvero $H^k(x)$. Tenendo conto del fatto che sono presenti k figli con probabilità p_k perveniamo alla relazione

$$H(x) = x \sum_k p_k H^k(x) = x F(H(x)) \quad (92)$$

dove la moltiplicazione per x corrisponde ad aggiungere il capostipite alla sua discendenza. Valutiamo (92) per i due semplici casi visti. Nel primo caso abbiamo $p_k = 0$ per $k > 1$ e $p_1 = p$ e $p_0 = 1 - p$. Allora (92) diventa

$$H(x) = x(1 - p + p H(x)) \implies H(x) = \frac{(1 - p)x}{1 - px}$$

Dalla Tabella 1 di pagina 122 possiamo esprimere $H(x)$ come

$$H(x) = (1 - p)x \sum_{k \geq 0} p^k x^k = \sum_{k \geq 1} (1 - p) p^{k-1} x^k$$

da cui ricaviamo la probabilità r_k che l'albero (di fatto un cammino) abbia k vertici:

$$r_k = (1 - p) p^{k-1}$$

In questo caso avremmo potuto ottenere lo stesso risultato per via diretta. Ma nel secondo esempio le cose sono meno semplici. Quindi sia $p_2 = p$, $p_0 = 1 - p$ e $p_k = 0$, per $k \notin \{0, 2\}$. Allora (92) diventa

$$H(x) = x(1 - p + pH^2(x)) \implies pxH^2(x) - H(x) + (1 - p)x = 0$$

da cui

$$H(x) = \frac{1 \pm \sqrt{1 - 4p(1 - p)x^2}}{2px}$$

Si noti che non necessariamente $H(1) = 1$. Si tratta infatti di una variabile aleatoria difettiva, ovvero con integrale della densità non necessariamente uguale ad uno, in quanto l'evento 'albero finito con k vertici' potrebbe anche non avvenire. Però per valori bassi di p l'albero non può essere infinito e quindi possiamo basarci sul fatto che $H(1) = 1$ per bassi valori di p . Allora

$$H(x) = \frac{1 - \sqrt{1 - 4p(1 - p)x^2}}{2px}$$

Abbiamo già incontrato un'espressione simile. Si veda l'equazione (89) a pag. 131. Quindi abbiamo

$$H(x) = \sum_{k \geq 1} C_{k-1} p^{k-1} (1 - p)^k x^{2k-1}$$

e

$$r_{2k-1} = C_{k-1} p^{k-1} (1 - p)^k$$

Si noti che, correttamente, si ottengono valori positivi solo per un numero dispari di vertici, in quanto l'albero necessariamente ha un numero dispari di vertici. Ad esempio si ha (ponendo $k = 1$) $r_1 = 1 - p$, valore che si ottiene anche più facilmente per via diretta e (ponendo $k = 2$) $r_3 = p(1 - p)^2$, anche questo valore non difficile da calcolare per via diretta.

Se $p > 1/2$ si ottiene $\sum_k r_k < 1$, in particolare si ottiene esattamente $\sum_k r_k = H(1) = u$ come si vede sia risolvendo direttamente (91) ma anche, in modo del tutto generale, confrontando (90) con (92) da cui si vede che $H(1) = u$.

50 Funzioni generatrici dei gradi dei nodi

Immaginiamo di avere un grafo generato casualmente con p_k uguale alla probabilità che un generico nodo del grafo abbia grado k . Quindi la funzione generatrice è un polinomio di grado $n - 1$. Spesso è però conveniente pensare che il numero di nodi tenda ad infinito, in modo da poter ottenere dei risultati di tipo asintotico. Ad esempio la funzione generatrice dei gradi di un grafo di Poisson è

$$F(x) = e^{-\bar{g}} \sum_{k \geq 0} \frac{\bar{g}^k}{k!} x^k = e^{-\bar{g}(1-x)}$$

Il grado medio calcolato direttamente dalla funzione generatrice è,

$$F'(x)|_{x=1} = \bar{g} e^{-\bar{g}(1-x)}|_{x=1} = \bar{g}$$

Una quantità spesso utile da prendere in considerazione è il cosiddetto *grado in eccesso*. Introduciamo il concetto attraverso queste considerazioni. Supponiamo di osservare il grado di un nodo scelto a caso in tre modi alternativi. Nel primo modo si sceglie a caso direttamente il nodo e si osserva il suo grado. Questo corrisponde per definizione alla probabilità p_k descritta sopra. Nel secondo modo si sceglie a caso un arco e si osserva il grado di uno dei suoi due nodi, scelto ancora a caso. Nel terzo modo si sceglie a caso un nodo di grado almeno uno e poi si osserva il grado di un nodo adiacente scelto ancora a caso.

Nel primo modo la probabilità di osservare un grado k è, come già detto, p_k . Nel secondo modo si immagina di spezzare un arco in due ‘mezzi’ archi. Da ogni nodo di grado k escono k mezzi archi. Se il numero di nodi di grado k è $p_k n$, il numero di mezzi archi incidenti in nodi di grado k è $k p_k n$. Il numero totale di mezzi archi è $2m$. Quindi la probabilità p'_k di avere un mezzo arco incidente in un nodo di grado k è

$$p'_k := \frac{k p_k n}{2m} = \frac{k p_k}{\bar{g}} \quad (93)$$

Si noti che correttamente

$$\sum_k p'_k = \frac{1}{\bar{g}} \sum_k k p_g = \frac{\bar{g}}{\bar{g}} = 1$$

Nel terzo modo la scelta di un nodo j a caso fra quelli non isolati avviene con probabilità $1/(n(1-p_0))$, dato che il numero di nodi di grado diverso da zero è $n(1-p_0)$. Affinché si possa poi scegliere un particolare nodo i come adiacente al nodo j , il nodo j deve essere fra gli adiacenti al nodo i . Se avviene questo, il nodo i viene scelto con probabilità $1/g_j$. Quindi il nodo i viene scelto con probabilità

$$\frac{1}{n(1-p_0)} \sum_{j:(i,j) \in E} \frac{1}{g_j} \quad (94)$$

Si noti che se si somma su tutti gli i con $g_i > 0$ si ha correttamente

$$\begin{aligned} \frac{1}{n(1-p_0)} \sum_{i:g_i > 0} \sum_{j:(i,j) \in E} \frac{1}{g_j} &= \frac{1}{n(1-p_0)} \sum_i [g_i > 0] \sum_j [(i,j) \in E] [g_j > 0] \frac{1}{g_j} = \\ \frac{1}{n(1-p_0)} \sum_j [g_j > 0] \frac{1}{g_j} \sum_i [(i,j) \in E] [g_i > 0] &= \frac{1}{n(1-p_0)} \sum_j [g_j > 0] \frac{1}{g_j} g_j = 1 \end{aligned}$$

Da (94) si vede che la probabilità p''_k di scegliere un nodo di grado k è

$$\begin{aligned} p''_k &= \frac{1}{n(1-p_0)} \sum_i [g_i = k] \sum_j [(i,j) \in E] [g_j > 0] \frac{1}{g_j} = \\ \frac{1}{n(1-p_0)} \sum_j [g_j > 0] \frac{1}{g_j} \sum_i [(i,j) \in E] [g_i = k] &= \frac{1}{n(1-p_0)} \sum_j [g_j > 0] \frac{g_{jk}}{g_j} \end{aligned} \quad (95)$$

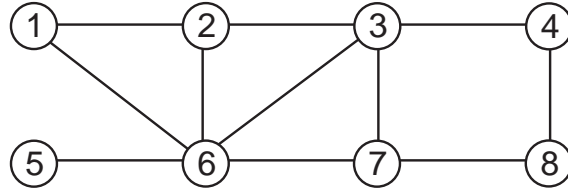


Figura 60.

dove g_{jk} è il numero di nodi adiacenti a j di grado k . L'espressione (95) è esatta e vale per qualsiasi grafo. La sua valutazione richiede però il calcolo dei valori g_{jk} . Asintoticamente il valore p_k'' si può approssimare nel seguente modo. Una volta scelto un nodo il nodo j , anziché considerare il valore corretto g_j , consideriamo il valore medio $\bar{g}/(1-p_0)$ di archi incidenti nel nodo j . Tale valore medio tiene conto del fatto che i nodi isolati non sono scelti a priori e quindi le probabilità p_k , condizionate al fatto che non ci sono nodi di grado 0, diventano $p_k/(1-p_0)$ il che spiega il valor medio $\bar{g}/(1-p_0)$. Allora (95) può essere approssimata da

$$\frac{1}{n(1-p_0)} \sum_j [g_j > 0] \frac{g_{jk}(1-p_0)}{\bar{g}} = \frac{1}{n\bar{g}} \sum_j [g_j > 0] g_{jk}$$

L'espressione $\sum_j [g_j > 0] g_{jk}$ conta il numero di mezzi archi incidenti in nodi di grado k . I nodi di grado k sono in numero np_k e quindi $\sum_j [g_j > 0] g_{jk} = nk p_k$ e la formula (95) diventa

$$p_k'' = \frac{1}{n(1-p_0)} \sum_j [g_j > 0] \frac{g_{jk}}{g_j} = \frac{1}{n(1-p_0)} \sum_j [g_j > 0] \frac{k p_k}{\bar{g}} = \frac{k p_k}{\bar{g}} \quad (96)$$

Quindi si ottiene che p_k'' è uguale a p_k' se consideriamo valida l'approssimazione. Però se il grafo è di Poisson allora si ha $p_k'' = p_k'$ in modo esatto. Questo risultato si ottiene con il seguente ragionamento. Si immagini di scegliere un nodo a caso e di generare solo adesso gli archi incidenti nel nodo. Supponiamo che almeno un arco venga generato. Ogni nodo del grafo può essere scelto come nodo adiacente con la stessa probabilità. Si scelga allora uno di questi nodi. Un arco per questo nodo è stato appena generato. Gli altri $(n-2)$ sono generati con probabilità $p = \bar{g}/(n-1)$. Quindi la probabilità che questo nodo abbia grado k è

$$p_k'' = \binom{n-2}{k-1} p^{k-1} (1-p)^{n-k-1}$$

e quindi, sfruttando l'approssimazione asintotica (74), ritroviamo (96).

Il confronto di p_k con p_k'' porta ad una considerazione che sembra paradossale. Secondo (96) si ottiene un nodo di grado più elevato se anziché campionare direttamente un nodo scegliamo un nodo e poi prendiamo un suo adiacente. Detto in altri termini: un mio amico ha probabilmente più amici di me!

Si vedano applicati al grafo in Fig. 60 i tre modi diversi di scegliere un nodo. Si noti che $\bar{g} = 22/8 = 2.75$. Nel primo caso ogni nodo viene scelto con la stessa probabilità e quindi le probabilità p_k sono

p_1	p_2	p_3	p_4	p_5
$\frac{1}{8}$	$\frac{3}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Se si scelgono gli archi con probabilità uniforme e per ogni arco si sceglie a caso uno dei due nodi, i nodi vengono scelti con probabilità

1	2	3	4	5	6	7	8
$\frac{2}{22}$	$\frac{3}{22}$	$\frac{4}{22}$	$\frac{2}{22}$	$\frac{1}{22}$	$\frac{5}{22}$	$\frac{3}{22}$	$\frac{2}{22}$

da cui

p'_1	p'_2	p'_3	p'_4	p'_5
$\frac{1}{22}$	$\frac{6}{22}$	$\frac{6}{22}$	$\frac{4}{22}$	$\frac{5}{22}$

Si noti la correttezza dell'espressione (93). Se invece si scelgono nodi a caso e poi i nodi adiacenti, ogni nodo viene scelto con probabilità

1	2	3	4
$\frac{1}{8}(\frac{1}{3} + \frac{1}{5})$	$\frac{1}{8}(\frac{1}{2} + \frac{1}{4} + \frac{1}{5})$	$\frac{1}{8}(\frac{1}{3} + \frac{1}{2} + \frac{1}{5} + \frac{1}{3})$	$\frac{1}{8}(\frac{1}{4} + \frac{1}{2})$

5	6	7	8
$\frac{1}{8}(\frac{1}{5})$	$\frac{1}{8}(\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + 1 + \frac{1}{3})$	$\frac{1}{8}(\frac{1}{4} + \frac{1}{5} + \frac{1}{2})$	$\frac{1}{8}(\frac{1}{2} + \frac{1}{3})$

cioè

1	2	3	4	5	6	7	8
$\frac{32}{480}$	$\frac{57}{480}$	$\frac{82}{480}$	$\frac{45}{480}$	$\frac{12}{480}$	$\frac{145}{480}$	$\frac{57}{480}$	$\frac{50}{480}$

e quindi

p''_1	p''_2	p''_3	p''_4	p''_5
$\frac{12}{480}$	$\frac{127}{480}$	$\frac{114}{480}$	$\frac{82}{480}$	$\frac{145}{480}$

Ovviamente c'è differenza in questo caso fra p' e p'' perché il grafo non è di Poisson. I valori di p , p' e p'' sono riassunti nel grafico di Fig. 61. Si noti invece in Fig. 62 il risultato di una simulazione su un grafo $G_{n,p}$ con $n = 100$, e grado medio $\bar{g} = 5$. Le distribuzioni di p' e p'' sono quasi identiche. All'aumentare di n la differenza diventa trascurabile.

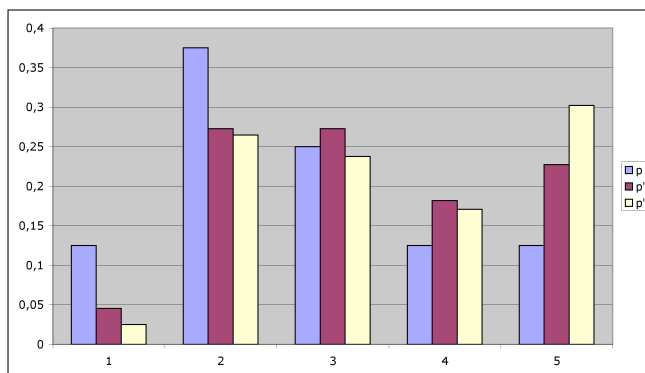


Figura 61.

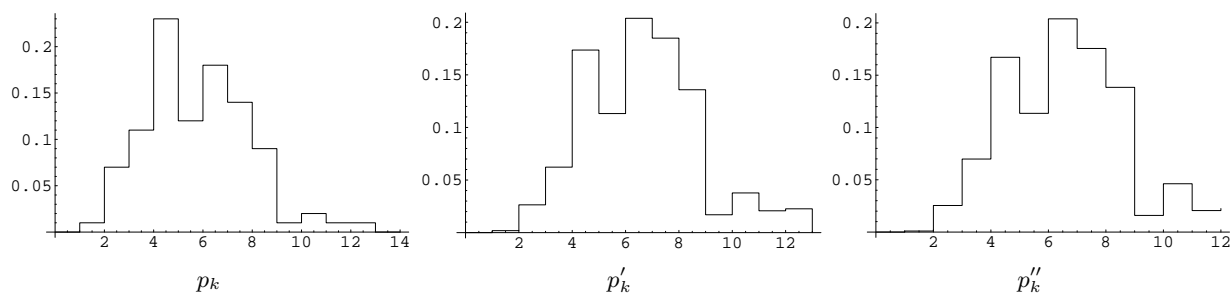


Figura 62.

Si definisce *grado in eccesso* il grado del nodo adiacente *meno uno*, cioè il numero di archi adiacenti nel nodo meno l'arco da cui si è arrivati. Sia q_k la probabilità che un nodo abbia grado in eccesso k . Allora

$$q_k = p''_{k+1} = \frac{(k+1)p_{k+1}}{\bar{g}}$$

Si noti che in quest'espressione si è fatto uso dell'approssimazione $p' = p''$. Definiamo le funzioni generatrici per i due tipi di gradi:

$$G_0(x) = \sum_k p_k x^k$$

$$G_1(x) = \sum_k q_k x^k = \sum_k \frac{(k+1)p_{k+1}}{\bar{g}} x^k = \frac{1}{\bar{g}} G'_0(x) = \frac{G'_0(x)}{G'_0(1)} \quad (97)$$

$$G'_1(x) = \frac{1}{\bar{g}} G''_0(x) \quad (98)$$

Si ribadisce che anche queste relazioni fra $G_0(x)$ e $G_1(x)$ dipendono dall'ipotesi $p' = p''$.

51 Componenti connesse piccole e componente gigante

Le considerazioni che seguono sono di tipo asintotico. Cioè si suppone che il numero di vertici tenda ad infinito. Il grafo non è connesso in generale. Si definisce *componente gigante* una componente connessa il cui numero di vertici tende ad infinito come i vertici del grafo stesso, mentre si definiscono *componenti piccole* quelle componenti che rimangono finite anche al tendere ad infinito del numero di vertici del grafo.

Siamo interessati a capire sotto quali condizioni possa esistere nel grafo una componente gigante insieme ad altre piccole componenti. A questo fine dobbiamo fare le seguenti ipotesi:

Ipotesi 1: Asintoticamente esiste al più una componente gigante.

Ipotesi 2: Asintoticamente le componenti piccole sono alberi.

L'Ipotesi 1 si giustifica pensando che la probabilità che non esista nemmeno un arco fra due componenti il cui numero di nodi tende ad infinito deve tendere a 0. Infatti, supponiamo che le due componenti giganti abbiano un numero di nodi $\alpha_1 n$ e $\alpha_2 n$ rispettivamente e immaginiamo di generare gli m archi un arco alla volta. Sia $M = n(n-1)/2$ il numero di tutte le coppie di nodi. Il numero di coppie di nodi fra le due componenti è $\alpha_1 \alpha_2 n^2$. Allora la probabilità che il primo arco scelto a caso non abbia un estremo in una componente e l'altro nell'altra componente è

$$\frac{M - \alpha_1 \alpha_2 n^2}{M} = 1 - \frac{\alpha_1 \alpha_2 n^2}{M}$$

e che la stessa cosa succeda per il secondo arco è

$$\frac{M - 1 - \alpha_1 \alpha_2 n^2}{M - 1} = 1 - \frac{\alpha_1 \alpha_2 n^2}{M - 1}$$

e che succeda per tutti gli m archi

$$\left(1 - \frac{\alpha_1 \alpha_2 n^2}{M}\right) \left(1 - \frac{\alpha_1 \alpha_2 n^2}{M - 1}\right) \cdots \left(1 - \frac{\alpha_1 \alpha_2 n^2}{M - m + 1}\right) \leq \left(1 - \frac{\alpha_1 \alpha_2 n^2}{M}\right)^m = \left(1 - \frac{2 \alpha_1 \alpha_2 n^2}{n(n-1)}\right)^m$$

Per n che tende ad infinito, il termine dentro parentesi tende ad un valore strettamente minore di 1, mentre m tende ad infinito. Quindi la probabilità che non esista nemmeno un arco fra due componenti giganti tende a 0 e quindi non può che esserci al più una sola componente gigante.

L'Ipotesi 2 si giustifica osservando che il numero di componenti piccole deve essere infinito (oppure nullo). Allora la probabilità che i due estremi dell'arco appartengano entrambi alla stessa componente piccola tende a zero e quindi le componenti sono alberi.

Questa analisi viene inizialmente condotta su una componente piccola sfruttando l'Ipotesi 2. Con questa ipotesi la valutazione della probabilità di avere una componente di dimensione data si può effettuare tramite le funzioni generatrici. Definiamo i seguenti eventi:

- D_k : un nodo scelto a caso ha grado k ;
- S : un nodo scelto a caso appartiene ad una componente piccola;

– S_k : un nodo scelto a caso fra i nodi delle componenti piccole appartiene ad una componente piccola di cardinalità k ;

– R : dato un nodo i , scelto a caso fra i nodi di grado almeno uno di tutto il grafo ed un nodo j , scelto a caso fra i nodi adiacenti ad i , il nodo j appartiene ad una componente piccola dopo aver rimosso l'arco (i, j) ;

– R_k : dato un nodo i , scelto a caso fra i nodi di grado almeno uno delle componenti piccole ed un nodo j , scelto a caso fra i nodi adiacenti ad i , il nodo j appartiene ad una componente piccola di cardinalità k , dopo aver rimosso l'arco (i, j) ;

– T_k : una componente piccola scelta a caso ha cardinalità k .

Per definizione $p_k = \Pr \{D_k\}$. Vogliamo valutare le seguenti probabilità

$$\begin{aligned} u &:= \Pr \{R\}, & v &:= \Pr \{S\}, & \hat{p}_k &:= \Pr \{D_k | S\}, \\ s_k &:= \Pr \{S_k\}, & r_k &:= \Pr \{R_k\}, & t_k &:= \Pr \{T_k\} \end{aligned}$$

Il valore u può essere calcolato tramite la funzione generatrice $G_1(x)$. La rimozione di un arco può creare una componente piccola se e solo se lo stesso fatto accade per tutti gli altri archi incidenti nel vertice. Allora se il vertice ha grado in eccesso k , la probabilità di creare una componente è u^k . Facendo la media su tutti i gradi si ottiene

$$u = \sum_{k \geq 0} u^k q_k = G_1(u) \quad (99)$$

L'equazione di punto fisso (99) è il punto di partenza fondamentale di tutti i calcoli successivi. Se non c'è la componente gigante allora $u = v = 1$ e $\hat{p}_k = p_k$. Se invece c'è la componente gigante si ha $u < 1$ e $v < 1$ e per il calcolo di \hat{p}_k si può sfruttare la regola di Bayes

$$\hat{p}_k = \Pr \{D_k | S\} = \frac{\Pr \{S | D_k\} p_k}{\Pr \{S\}}$$

Ovviamente $\Pr \{S | D_0\} = 1$ e di conseguenza

$$\hat{p}_0 = \frac{p_0}{v}$$

Se $k > 0$, $\Pr \{S | D_k\}$ è la probabilità che *tutti* i k nodi adiacenti appartengano ad una componente piccola dopo che sono stati rimossi gli archi corrispondenti. Quindi

$$\hat{p}_k := \frac{\Pr \{R\}^k p_k}{v} = \frac{u^k}{v} p_k \quad (100)$$

valida anche per $k = 0$. Quindi si ha

$$1 = \sum_k \hat{p}_k = \frac{G_0(u)}{v}$$

da cui

$$v = G_0(u) \quad (101)$$

Quindi da (99) si calcola u . Poi si calcola v da (101) e infine si calcola \hat{p}_k da (100). Definiamo le funzioni generatrici

$$\hat{G}_0(x) = \sum_k \hat{p}_k x^k, \quad \hat{G}_1(x) = \sum_k \hat{q}_k x^k,$$

dove $\hat{q}_k = (k+1)\hat{p}_{k+1}/\hat{g}$ e \hat{g} è il grado medio nelle componenti piccole. La relazione fra $G_0(x)$ e $\hat{G}_0(x)$ è data da

$$\hat{G}_0(x) = \sum_k \hat{p}_k x^k = \sum_k \frac{u^k}{v} p_k x^k = \frac{1}{v} G_0(ux) = \frac{G_0(ux)}{G_0(u)}$$

da cui

$$\hat{G}'_0(x) = \frac{u G'_0(ux)}{G_0(u)} = \frac{u}{v} G'_0(ux)$$

cosicché

$$\hat{g} = \hat{G}'_0(1) = \frac{u G'_0(u)}{G_0(u)} = \frac{u}{v} G'_0(u) = \frac{u}{v} G_1(u) \bar{g} = \frac{u^2}{v} \bar{g} \quad (102)$$

dove si è tenuto conto di (99). Per trovare la relazione fra $G_1(x)$ e $\hat{G}_1(x)$ si ha per definizione

$$\hat{G}_1(x) = \frac{\hat{G}'_1(x)}{\hat{G}'_1(1)} = \frac{G'_0(ux)}{G'_0(u)} = \frac{u}{v} \frac{G'_0(ux)}{\hat{g}} \quad (103)$$

Da $G_1(x) = G'_0(x)/G'_0(1)$ si ottiene usando una delle eguaglianze in (103),

$$G_1(ux) = \frac{G'_0(ux)}{G'_0(1)} = \hat{G}_1(x) \frac{G'_0(u)}{G'_0(1)} = \hat{G}_1(x) G_1(u)$$

cioè,

$$\hat{G}_1(x) = \frac{G_1(ux)}{G_1(u)} = \frac{G_1(ux)}{u}$$

da cui $\hat{G}'_1(x) = G'_1(ux)$, $\hat{G}'_1(1) = G'_1(u)$ e quindi si vede che u soddisfa la seguente relazione rispetto a \hat{G}_1

$$\hat{G}_1(u^{-1}) = u^{-1} \quad (104)$$

Il calcolo delle probabilità s_k , r_k e t_k è più complesso. A tale fine bisogna definire le funzioni generatrici delle probabilità e sfruttare le proprietà delle funzioni generatrici. Siano

$$H_0(x) = \sum_k s_k x^k, \quad H_1(x) = \sum_k r_k x^k, \quad K(x) = \sum_k t_k x^k$$

Si consideri la Fig. 63. Il nodo i è un nodo scelto a caso fra i nodi delle componenti piccole. La probabilità che la componente a cui appartiene abbia cardinalità k è, per definizione, s_k con funzione generatrice $H_0(x)$. Sia g_i il numero di nodi adiacenti a i (Fig. 63-a). Per “contare” i nodi della componente si operi in questo modo: per ogni nodo adiacente si rimuove l’arco corrispondente e così facendo il nodo adiacente appartiene ad una nuova componente piccola (Fig. 63-b). Allora il numero totale di nodi della componente piccola iniziale

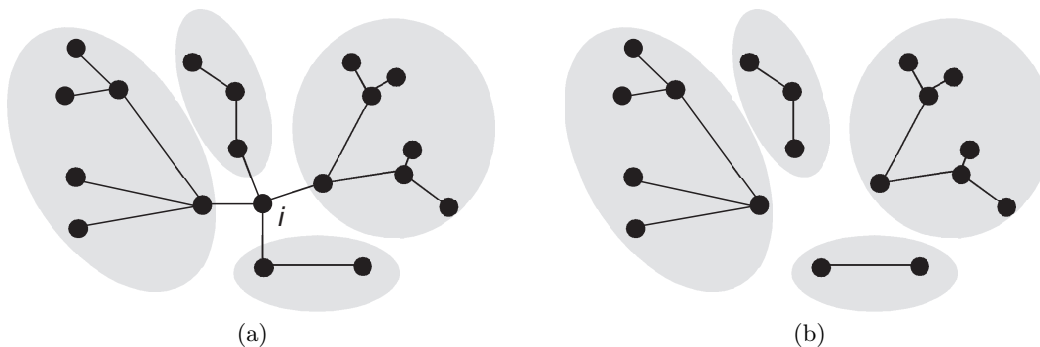


Figura 63.

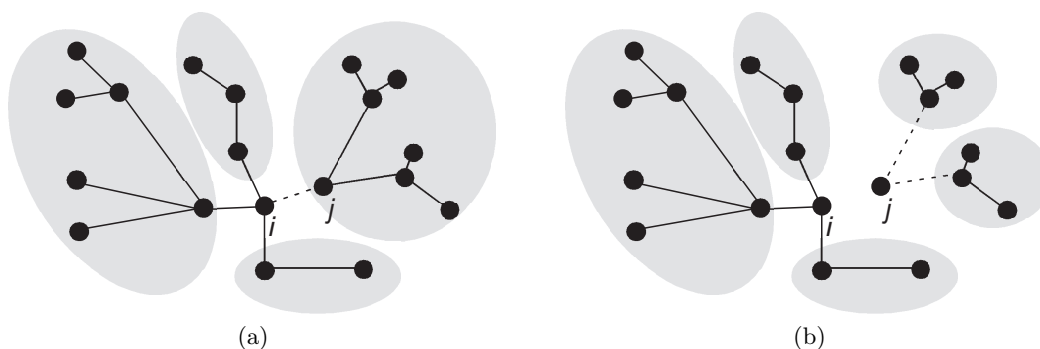


Figura 64.

è dato da 1 (il nodo i) più la somma dei nodi delle nuove componenti piccole. La probabilità di avere k nodi in ognuna delle componenti piccole è, per definizione, r_k , con funzione generatrice $H_1(x)$. Allora, se il grado di i è g_i il numero totale di nodi è

$$x H_1(x)^{g_i}$$

dove si è sfruttata la proprietà che la somma di variabili aleatorie indipendenti si traduce nel prodotto delle funzioni generatrici corrispondenti e il termine x sta per la somma del nodo i stesso. A questo punto la funzione generatrice $H_0(x)$ si trova come media probabilistica rispetto a tutti i possibili gradi g_k e quindi

$$H_0(x) = x \sum_k \hat{p}_k H_1(x)^k = x \hat{G}_0(H_1(x)) \quad (105)$$

In modo analogo si possono contare i nodi di una componente piccola a cui appartiene il nodo j dopo aver rimosso l'arco (i, j) (si veda la Fig. 64-a). Il numero di nodi della componente piccola a cui appartiene il nodo j è dato da 1 (nodo j stesso) più la somma dei nodi delle nuove componenti piccole ottenute dopo la rimozione degli archi corrispondenti (Fig. 64-b). Anche in questo caso la probabilità di avere k nodi in

ognuna delle componenti piccole è r_k . Il numero delle piccole componenti che si formano è dato dal grado in eccesso di j e allora si ha

$$H_1(x) = x \sum_k \hat{q}_k H_1(x)^k = x \hat{G}_1(H_1(x)) \quad (106)$$

Per trovare il legame fra $K(x)$ e le funzioni $H_0(x)$ e $H_1(x)$ si derivino entrambi i termini in (105)

$$H'_0(x) = \hat{G}_0(H_1(x)) + x \hat{G}'_0(H_1(x)) H'_1(x) \quad (107)$$

che possiamo rielaborare come

$$H'_0(x) = \frac{H_0(x)}{x} + x \hat{g} \hat{G}_1(H_1(x)) H'_1(x) = \frac{H_0(x)}{x} + \hat{g} H_1(x) H'_1(x) \quad (108)$$

Ora notiamo che

$$s_k = \frac{k t_k}{c} \quad (109)$$

dove $c = K'(1)$ è la cardinalità media di una componente piccola. Questa relazione si giustifica nel seguente modo: siano n_k il numero di componenti piccole di cardinalità k . Allora il numero totale di nodi delle componenti piccole è $\sum_h h n_h$. Allora si ha

$$s_k = \frac{k n_k}{\sum_h h n_h} = \frac{k n_k}{\sum_h n_h} \frac{\sum_h n_h}{\sum_h h n_h}$$

e la probabilità t_k di scegliere a caso una componente di cardinalità k è $n_k / \sum_h n_h$ e la cardinalità media c è data da $\sum_h h n_h / \sum_h n_h$. Sfruttando (109) si ha

$$K'(x) = \frac{d}{dx} \sum_{k \geq 0} t_k x^k = \sum_{k \geq 1} k t_k x^{k-1} = \sum_{k \geq 1} c s_k x^{k-1} = c \frac{H_0(x)}{x} = c (H'_0(x) - \hat{g} H_1(x) H'_1(x))$$

la cui integrazione porta a

$$K(x) = c (H_0(x) - \frac{\hat{g}}{2} H_1^2(x)) = \frac{2}{2 - \hat{g}} (H_0(x) - \frac{\hat{g}}{2} H_1^2(x)) \quad (110)$$

La relazione

$$c = \frac{2}{2 - \hat{g}} \quad (111)$$

deriva dal fatto che ogni componente piccola è un albero. Da (106) si ha ancora

$$H'_1(x) = \hat{G}_1(H_1(x)) + x \hat{G}'_1(H_1(x)) H'_1(x)$$

da cui

$$H'_1(1) = \hat{G}_1(H_1(1)) + \hat{G}'_1(H_1(1)) H'_1(1) = 1 + \hat{G}'_1(1) H'_1(1)$$

e allora, indicando con \bar{r} il valor medio della variabile aleatoria r_k , abbiamo

$$\bar{r} = H'(1) = \frac{1}{1 - \hat{G}'_1(1)}$$

Il momento in cui (variando i gradi che definiscono il grafo) si forma la componente gigante viene detto *transizione di fase*. In transizione di fase il numero medio di nodi di una componente piccola tende ad infinito e quindi si deve avere $\hat{G}'_1(1) = 1$. Il valor medio \bar{s} della variabile aleatoria s_k è dato da $H'_0(1)$ e quindi abbiamo da (107)

$$\bar{s} = H'_0(1) = 1 + \hat{G}'_0(1) H'_1(1) = 1 + \frac{\hat{G}'_0(1)}{1 - \hat{G}'_1(1)} = 1 + \frac{\hat{g}}{1 - \hat{G}'_1(1)} \quad (112)$$

I valori medi c , \bar{s} e \bar{r} richiedono solo la risoluzione dell'equazione $u = G_1(u)$ e non la risoluzione delle equazioni funzionali (105) e (106). In alcuni rari casi le funzioni $H_0(x)$ e $H_1(x)$ si possono ottenere in forma analitica. Quando ciò non sia possibile è però facilmente ottenibile una risoluzione numerica di (106).

Dall'equazione ricorsiva

$$H_1(x) = x G_1(H_1(x)) = x \sum_{k \geq 0} \hat{q}_k H_1(x)^k \quad (113)$$

e da

$$H_1(x) = \sum_{k \geq 1} r_k x^k$$

(necessariamente $r_0 = 0$) deriviamo

$$\begin{aligned} \sum_{k \geq 1} r_k x^k &= x \hat{q}_0 + x \sum_{h=1}^{n-1} \hat{q}_h \left(\sum_{j \geq 1} r_j x^j \right)^h \\ \sum_{k \geq 1} r_k x^k &= x \hat{q}_0 + \sum_{h=1}^{n-1} \hat{q}_h x^{h+1} \left(\sum_{j \geq 0} r_{j+1} x^j \right)^h = x \hat{q}_0 + \sum_{h=2}^n \hat{q}_{h-1} x^h \left(\sum_{j \geq 0} r_{j+1} x^j \right)^{h-1} \end{aligned} \quad (114)$$

Sia b_k^h il coefficiente di x^k in $(\sum_{j \geq 0} r_{j+1} x^j)^h$. Si noti che $b_k^1 = r_{k+1}$. Da (114)

$$r_k = \sum_{h=2}^k \hat{q}_{h-1} b_{k-h}^{h-1} \quad (115)$$

Quindi il calcolo di r_k richiede i coefficienti $b_{k-2}^1, b_{k-3}^2, \dots, b_0^{k-1}$. A sua volta il calcolo di b_k^h richiede i termini r_1, \dots, r_{k+1} e allora per calcolare r_k basta conoscere i valori r_1, \dots, r_{k-1} . La ricorsione è allora: inizialmente $r_1 = \hat{q}_0$, e poi

$$\left. \begin{aligned} r_k &= \sum_{h=2}^k \hat{q}_{h-1} b_{k-h}^{h-1}, \\ b_{k-1}^1 &= r_k \\ b_{k-1}^h &= \sum_{j=0}^{k-1} b_j^{h-1} b_{k-1-j}^1 = \sum_{j=0}^{k-1} b_j^{h-1} r_{k-j}, \quad h = 2, \dots, k-1 \\ b_h^k &= \sum_{j=0}^h b_j^{k-1} r_{h+1-j}, \quad h = 0, \dots, k \end{aligned} \right\} k = 2, \dots \quad (116)$$

Inoltre da

$$H_0(x) = x \hat{G}_0(H_1(x)), \quad H_0(x) = \sum_{k \geq 1} s_k x^k$$

si deriva l'espressione

$$\sum_{k \geq 1} s_k x^k = x \hat{p}_0 + \sum_{h=2}^n \hat{p}_{h-1} x^h \left(\sum_{j \geq 0} r_{j+1} x^j \right)^{h-1}$$

e allora

$$s_1 = \hat{p}_0, \quad s_k = \sum_{h=2}^k \hat{p}_{h-1} b_{k-h}^{h-1}, \quad k \geq 2 \quad (117)$$

In questo caso il calcolo è immediato perché coinvolge quantità già calcolate precedentemente. Ovviamente il calcolo va interrotto dopo aver ottenuto il numero desiderato n di termini r_k e s_k , ma questo fatto non ha conseguenze sull'accuratezza del calcolo in quanto n non è presente in (116) e (117).

52 Componenti connesse in grafi particolari

Prima di applicare i concetti visti a dei grafi casuali 'interessanti', può essere istruttivo 'mettere alla prova' tutte le relazioni viste su alcuni grafi molto semplici.

Esempio 6

Supponiamo di studiare il grafo casuale in cui sono possibili solo due gradi: 0 o 1, con probabilità rispettivamente $1 - p$ e p . L'analisi è comunque asintotica per un grafo il cui numero di vertici tende ad infinito. Date le condizioni sui gradi, il grafo è fatto di vertici isolati oppure di coppie di vertici connesse con un arco. Quindi non ci può essere una componente gigante, e tutte le componenti hanno dimensione 1 o 2. Applicando le definizioni si ha

$$G_0(x) = 1 - p + px, \quad G'_0(x) = p, \quad G'_0(1) = \bar{g} = p, \quad G_1(x) = \frac{G'_0(x)}{\bar{g}} = 1$$

Effettivamente se si sceglie a caso un nodo non isolato, esiste solo un arco incidente nel nodo e il grado dell'altro estremo deve essere 1 e quindi il grado in eccesso deve essere 0 con probabilità 1. Da $u = G_1(u) = 1$ si ricava $u = 1$ e quindi $p_k = \hat{p}_k$. In queste condizioni il calcolo delle funzioni $H_0(x)$ e $H_1(x)$ è molto semplice.

$$H_1(x) = x G_1(H_1(x)) \implies H_1(x) = x,$$

$$H_0(x) = x G_0(H_1(x)) = x(1 - p + p H_1(x)) = x(1 - p + px) = (1 - p)x + px^2$$

Il risultato è coerente con il calcolo diretto che in questo caso si può fare facilmente. Scegliendo un nodo non isolato si trova esattamente un vertice all'altro estremo e quindi il coefficiente di x^1 in $H_1(x)$ non può che

essere 1. Scegliendo un nodo a caso, con probabilità p si sceglie un nodo di grado 1 e quindi la componente connessa ha dimensione 2, mentre con probabilità $1 - p$ si sceglie un vertice isolato. I coefficienti di $H_0(x)$ riflettono questa situazione. Abbiamo inoltre

$$\bar{s} = H_0'(1) = 1 - p + 2p = 1 + p$$

Infatti scegliendo a caso un vertice si sceglie con probabilità p uno di grado 1 e quindi la dimensione della componente è 2. Allora il valor medio è proprio $1(1 - p) + 2p = 1 + p$. L'equazione (111) diventa

$$c = \frac{2}{2 - \hat{g}} = \frac{2}{2 - p}$$

Anche questa relazione è coerente con un calcolo diretto. Il numero di vertici di grado 1 è np , quindi il numero di componenti connesse con 2 nodi è $np/2$ e il numero di componenti con un solo vertice è $n(1 - p)$. Il numero totale di componenti è $np/2 + n(1 - p)$ e il valor medio è dunque

$$\frac{1 \cdot n(1 - p) + 2 \cdot np/2}{np/2 + n(1 - p)} = \frac{1}{1 - \frac{p}{2}}$$

Esempio 7

Come secondo esempio leggermente più complesso si consideri un grafo in cui il grado massimo è 2 e le probabilità dei vari gradi sono p_0 , p_1 e p_2 . Quindi

$$G_0(x) = p_0 + p_1 x + p_2 x^2$$

$$G_0'(x) = p_1 + 2p_2 x \implies G_0'(1) = \bar{g} = p_1 + 2p_2$$

$$G_1(x) = \frac{p_1 + 2p_2 x}{p_1 + 2p_2}$$

Indichiamo per semplicità

$$\pi = \frac{2p_2}{p_1 + 2p_2} \implies 1 - \pi = \frac{p_1}{p_1 + 2p_2}$$

Allora

$$G_1(x) = (1 - \pi) + \pi x$$

Si può vedere che da $u = G_1(u)$ si ottiene soltanto la soluzione $u = 1$ da cui $v = G_0(u) = 1$ e quindi non c'è la componente gigante. Per il calcolo di $H_1(x)$ e $H_0(x)$ si ottiene

$$H_1(x) = x(1 - \pi + \pi H_1(x))$$

Quindi

$$H_1(x) = \frac{(1 - \pi)x}{1 - \pi x} = (1 - \pi)x \sum_{k \geq 0} \pi^k x^k$$

Per $H_0(x)$ si ha

$$\begin{aligned} H_0(x) &= x G_0(H_1(x)) = x (p_0 + p_1 H_1(x) + p_2 H_1^2(x)) = \\ &= x \left(p_0 + p_1 \frac{(1-\pi)x}{1-\pi x} + p_2 \frac{(1-\pi)^2 x^2}{(1-\pi x)^2} \right) = \\ &= p_0 x + p_1 (1-\pi) x^2 \sum_{k \geq 0} \pi^k x^k + p_2 (1-\pi)^2 x^3 \sum_{k \geq 0} (k+1) \pi^k x^k \end{aligned}$$

da cui

$$s_1 = p_0, \quad s_2 = p_1 (1-\pi),$$

e per $k \geq 3$

$$s_k = p_1 (1-\pi) \pi^{k-2} + p_2 (1-\pi)^2 (k-2) \pi^{k-3} = (1-\pi) \pi^{k-3} (p_1 \pi + p_2 (1-\pi) (k-2))$$

I valori \bar{s} (equazione (112)) e c (equazione (111)) sono:

$$\begin{aligned} \bar{s} &= 1 + \frac{\hat{G}'_0(1)}{1 - \hat{G}'_1(1)} = 1 + \frac{p_1 + 2p_2}{1-\pi} = 1 + \frac{(p_1 + 2p_2)^2}{p_1} \\ c &= \frac{2}{2 - \bar{g}} = \frac{2}{2 - (p_1 + 2p_2)} \end{aligned}$$

Eseguendo il calcolo numerico per i valori di probabilità $p_0 = 0$, $p_1 = 2/3$, $p_2 = 1/3$ si ottiene $\pi = 1/2$ e

$$\begin{aligned} G_0(x) &= \frac{2}{3} x + \frac{1}{3} x^2, & G_1(x) &= \frac{1}{2} + \frac{1}{2} x \\ H_1(x) &= \frac{1}{2} \frac{x}{1 - \frac{x}{2}} = \sum_{k \geq 1} \left(\frac{1}{2}\right)^k x^k \\ H_0(x) &= \frac{1}{3} x^2 + \sum_{k \geq 3} \left(\frac{1}{2}\right)^{k-1} \frac{k}{3} x^k & (118) \\ \bar{s} &= \frac{11}{3} = 3.667, & c &= 3 \end{aligned}$$

È stata effettuata una simulazione per i valori $n = 1200$, $p_0 = 0$, $p_1 = 2/3$ e $p_2 = 1/3$ generando un grafo casuale in questo modo: ai nodi da 1 a $2n/3$ è stato assegnato grado 1 e ai rimanenti nodi grado 2. In questo modo le probabilità p_k sono quelle assegnate. Poi è stato generato un grafo con questi gradi con la procedura nota. Da questo grafo si sono generati casualmente altri grafi usando la procedura della catena di Markov, già vista, con la differenza che non si fa nessuna verifica affinché il grafo sia connesso. Il grafo campionato è quello ottenuto dopo 5000 iterazioni.

Sono stati generati 800 archi con 400 componenti connesse. Come si vede da questi numeri in nessuna componente connessa è presente un circuito (in una foresta il numero di archi è uguale al numero di nodi

meno il numero di componenti connesse). Le componenti connesse hanno i seguenti valori (con n_k eguale al numero di componenti connesse di k vertici)

k	2	3	4	5	6	7	8	9	10
n_k	197	104	49	22	15	9	2	1	1

dai quali si deducono i valori di probabilità indicati nella tabella, abbastanza consistenti con i valori teorici dati da (118).

	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
simulati	0.3283	0.2600	0.1633	0.0916	0.0750	0.0525	0.0133	0.0075	0.0083
teorici	0.3333	0.2500	0.1667	0.1042	0.0625	0.0365	0.0208	0.0117	0.0065

Inoltre si trova dalla simulazione $\bar{s} = 1087/300 = 3.62333$. Il valore di c della simulazione è necessariamente uguale al valore teorico (nell'ipotesi di assenza di circuiti). Infatti ogni componente deve essere un cammino e quindi il numero di componenti è uguale alla metà del numero di nodi di grado 1. Quindi $c = n/(np_1/2) = 3$.

Esempio 8

Come terzo esempio di complessità ancora crescente si consideri un grafo in cui il grado massimo è 3 e le probabilità dei vari gradi sono p_0 , p_1 , p_2 e p_3 . Quindi

$$G_0(x) = p_0 + p_1 x + p_2 x^2 + p_3 x^3$$

$$G'_0(x) = p_1 + 2p_2 x + 3p_3 x^2 \quad \implies \quad G'_0(1) = \bar{g} = p_1 + 2p_2 + 3p_3$$

$$G_1(x) = \frac{p_1 + 2p_2 x + 3p_3 x^2}{p_1 + 2p_2 + 3p_3}$$

$$u = G_1(u) \quad \implies \quad u = \frac{p_1 + 2p_2 u + 3p_3 u^2}{p_1 + 2p_2 + 3p_3} \quad \implies \quad 3p_3 u^2 - (p_1 + 3p_3)u + p_1 = 0$$

Siccome sappiamo che una radice vale 1, l'altra si trova immediatamente come $u = p_1/(3p_3)$. Allora

$$u = \begin{cases} 1 & \text{se } p_1 \geq 3p_3 \\ p_1/(3p_3) & \text{se } p_1 < 3p_3 \end{cases}$$

La componente gigante può essere presente se $p_1 < 3p_3$. Si noti il fatto interessante che l'esistenza della componente gigante dipende unicamente dai gradi 1 e 3, ma non dal grado 2. Si ha

$$v = p_0 + p_1 \frac{p_1}{3p_3} + p_2 \left(\frac{p_1}{3p_3} \right)^2 + p_3 \left(\frac{p_1}{3p_3} \right)^3$$

$$\hat{p}_k = \frac{p_k \left(\frac{p_1}{3p_3} \right)^k}{p_0 + p_1 \frac{p_1}{3p_3} + p_2 \left(\frac{p_1}{3p_3} \right)^2 + p_3 \left(\frac{p_1}{3p_3} \right)^3}$$

e

$$H_1(x) = x \frac{\hat{p}_1 + 2\hat{p}_2 H_1(x) + 3\hat{p}_3 H_1^2(x)}{\hat{p}_1 + 2\hat{p}_2 + 3\hat{p}_3}$$

Il calcolo di $H_1(x)$ in modo analitico per valori generici di probabilità si può fare, ma è molto laborioso. Calcoliamo alcune grandezze per i valori numerici $p_0 = 0$, $p_1 = 1/2$, $p_2 = 1/6$, $p_3 = 1/3$. Quindi la condizione per la componente gigante è soddisfatta e $u = 1/2$ e $\bar{g} = 11/6$. Allora si ha

$$G_0(x) = \frac{1}{2}x + \frac{1}{6}x^2 + \frac{1}{3}x^3, \quad G_1(x) = \frac{3}{11} + \frac{2}{11}x + \frac{6}{11}x^2$$

$$v = G_0(u) \implies v = \frac{1}{2} \frac{1}{2} + \frac{1}{6} \frac{1}{4} + \frac{1}{3} \frac{1}{8} = \frac{1}{3}$$

Quindi 1/3 del grafo è componente piccola e 2/3 componente gigante. Si ha

$$\hat{p}_k = \frac{u^k}{v} p_k \implies \hat{p}_1 = \frac{\frac{1}{2}}{\frac{1}{3}} \frac{1}{2} = \frac{3}{4}, \quad \hat{p}_2 = \frac{\frac{1}{4}}{\frac{1}{3}} \frac{1}{6} = \frac{1}{8}, \quad \hat{p}_3 = \frac{\frac{1}{8}}{\frac{1}{3}} \frac{1}{3} = \frac{1}{8}$$

da cui il grado medio nelle componenti piccole

$$\hat{g} = \frac{3}{4} + 2 \frac{1}{8} + 3 \frac{1}{8} = \frac{11}{8}$$

Inoltre

$$\hat{G}_0(x) = \frac{3}{4}x + \frac{1}{8}x^2 + \frac{1}{8}x^3, \quad \hat{G}_1(x) = \frac{6}{11} + \frac{2}{11}x + \frac{3}{11}x^2$$

$$\bar{s} = 1 + \frac{\hat{g}}{1 - \hat{G}'_1(1)} = 1 + \frac{121}{24} = 6.0416$$

$$c = \frac{2}{2 - \hat{g}} = \frac{2}{2 - \frac{11}{8}} = \frac{16}{5} = 3.2$$

Eseguendo una simulazione, e generando a caso un grafo di 6000 nodi con le percentuali di grado indicate si ottengono i seguenti risultati: è presente una componente gigante di 3927 nodi, che occupa quindi il 65.45% del grafo, in accordo con il risultato teorico di 66,67% (2/3). Le componenti piccole, contenenti i rimanenti 2073 nodi, sono 632. Si noti che $2073 - 632 = 1441$ che è anche il numero di archi nelle componenti piccole. Quindi ogni componente piccola è un albero. La più grande di queste ha 37 vertici ed è un albero. Si vedano in Fig. 65 una rappresentazione della componente gigante e delle componenti piccole con almeno 10 vertici.

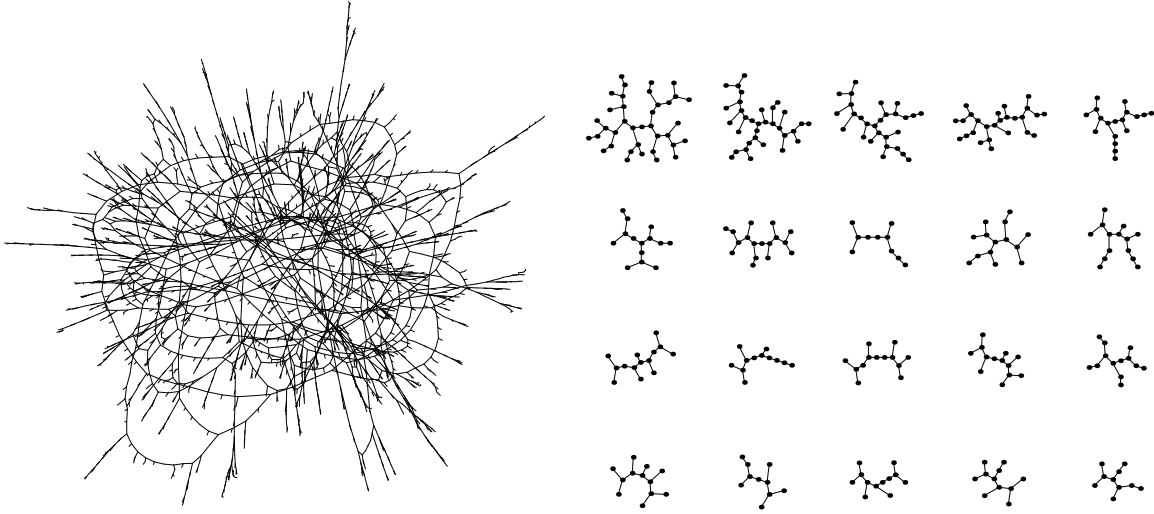


Figura 65. Componente gigante e alcune componenti piccole per un grafo con $p_0 = 0$, $p_1 = 1/2$, $p_2 = 1/6$, $p_3 = 1/3$

Il valor medio c delle componenti piccole è 3.157 abbastanza vicino al valore teorico $c = 3.2$. In Tabella 2 si riportano i valori teorici e sperimentali di alcune quantità. Il valore di u è calcolato in base alla sua definizione contando il rapporto fra i casi favorevoli e tutti i casi possibili. I valori di s_k , r_k e t_k sono vicini ai valori teorici per bassi valori di k . È chiaro che 6000 nodi sono ‘pochi’ per poter avere dei valori affidabili di simulazione per componenti con un discreto numero di vertici.

53 Grafi di Poisson

Per un grafo di Poisson si ha

$$G_0(x) = e^{-\bar{g}(1-x)}, \quad G'_0(x) = \bar{g} e^{-\bar{g}(1-x)}, \quad G_1(x) = \frac{1}{\bar{g}} G'_0(x) = G_0(x), \quad G'_1(x) = \bar{g} e^{-\bar{g}(1-x)}$$

Quindi $G_0(x) = G_1(x)$, cioè grado di un nodo e grado in eccesso sono identicamente distribuiti, ed anche $H_0(x) = H_1(x)$. L'equazione (99) diventa

$$u = e^{\bar{g}(u-1)} \tag{119}$$

che si può facilmente risolvere in modo numerico.

In Fig. 66 sono disegnate le funzioni u e $e^{\bar{g}(u-1)}$ nei due casi $\bar{g} = 0.5$ (figura (a)) e $\bar{g} = 2$ (figura (b)). L'equazione $u = e^{\bar{g}(u-1)}$ ha due soluzioni per $\bar{g} \neq 1$, di cui una è sempre $u = 1$. Per $\bar{g} > 1$ la seconda soluzione è minore di 1, mentre per $\bar{g} < 1$ la seconda soluzione è maggiore di 1. Per $\bar{g} = 1$ le due soluzioni coincidono in 1. Allora, siccome la soluzione non deve essere più grande di 1, trattandosi di una probabilità,

<i>teorico</i>	u	v	\hat{g}					
sperimentale	0.5080	0.4217	1.3903					
<i>teorico</i>	\hat{p}_1	\hat{p}_2	\hat{p}_3					
sperimentale	0.7347	0.1404	0.1249					
<i>teorico</i>	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
sperimentale	0	0.4091	0.1116	0.1082	0.0615	0.0526	0.0376	0.0314
		0.3936	0.1143	0.1100	0.0386	0.0463	0.0439	0.0502
<i>teorico</i>	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8
sperimentale	0.5455	0.0992	0.0992	0.0475	0.0408	0.0269	0.0223	0.0168
	0.5319	0.1055	0.0959	0.0355	0.0406	0.0301	0.035	0.0266
<i>teorico</i>	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
sperimentale	0	0.6545	0.1190	0.0866	0.0393	0.0281	0.0172	0.0126
	0	0.6456	0.1250	0.0902	0.0253	0.0253	0.0206	0.0206

Tabella 2. Valori teorici e simulati per il grafo con $p_0 = 0$, $p_1 = 1/2$, $p_2 = 1/6$, $p_3 = 1/3$

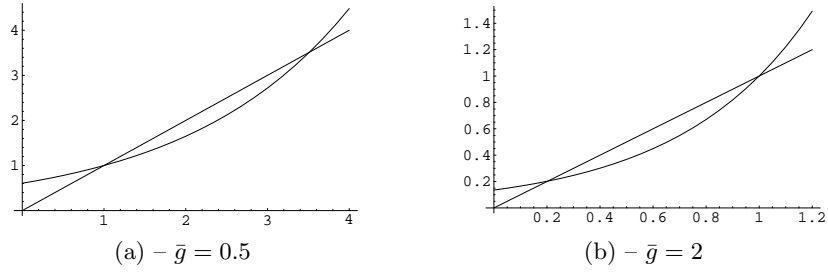


Figura 66.

la soluzione di (119) è (si veda il grafico in Fig. 67).

$$u = 1 \quad \text{se } \bar{g} \leq 1, \quad u < 1 \quad \text{se } \bar{g} > 1$$

La transizione di fase avviene quindi per il valore $\bar{g} = 1$. Quando il grado medio supera tale valore appare la componente gigante. La funzione $1 - v$, che dà la frazione di vertici nella componente gigante, rispetto al grado medio \bar{g} è quella di Fig. 67 ribaltata. Si veda il grafico in Fig. 68, dove in (a) il grafico è disegnato per i valori di grado $0 \leq \bar{g} \leq 4$ e in (b) per i valori di grado $4 \leq \bar{g} \leq 7$. Come si vede con un grado medio di 7 più del 99% del grafo è nella componente gigante. Quando $\bar{g} > 1$ abbiamo

$$v = G_0(u) = e^{\bar{g}(u-1)} = u, \quad \hat{g} = \frac{u G_0'(u)}{G_0(u)} = u \bar{g}$$

$$\hat{p}_k = \frac{u^k}{v} p_k = \frac{u^k}{e^{\bar{g}(u-1)}} \frac{\bar{g}^k}{k!} e^{-\bar{g}} = \frac{(u \bar{g})^k}{k!} e^{-u \bar{g}} = \frac{\hat{g}^k}{k!} e^{-\hat{g}}$$

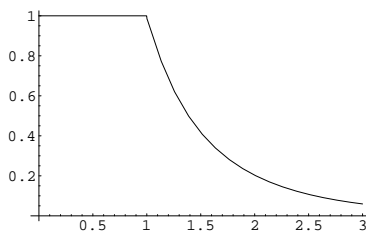


Figura 67. u in funzione del grado medio \bar{g}

	u	v	\hat{g}					
<i>teorico</i>	0.4175	0.4175	0.6260					
<i>sperimentale</i>	0.4178	0.4217	0.6398					
	\hat{p}_0	\hat{p}_1	\hat{p}_2	\hat{p}_3	\hat{p}_4	\hat{p}_5	\hat{p}_6	\hat{p}_6
<i>teorico</i>	0.5347	0.3347	0.1048	0.0219	0.0034	0.0004	0	0
<i>sperimentale</i>	0.5290	0.3332	0.1110	0.0232	0.0028	0.0007	0	0
	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
<i>teorico</i>	0.5347	0.1790	0.0899	0.0535	0.0350	0.0243	0.0176	0.0131
<i>sperimentale</i>	0.5290	0.1693	0.0960	0.0541	0.0403	0.0228	0.0183	0.0133
	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8
<i>teorico</i>	0.5347	0.1790	0.0899	0.0535	0.0350	0.0243	0.0176	0.0131
<i>sperimentale</i>	0.5176	0.1910	0.0885	0.0584	0.0321	0.0241	0.0177	0.0164
	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
<i>teorico</i>	0.7783	0.1303	0.0436	0.0195	0.0102	0.0059	0.0037	0.0024
<i>sperimentale</i>	0.7779	0.1245	0.0471	0.0199	0.0119	0.0056	0.0038	0.0024

Tabella 3. Valori teorici e simulati per un grafo di Poisson con $\bar{g} = 1.5$

Quindi la distribuzione dei gradi nelle componenti piccole è ancora di Poisson ma con un grado medio inferiore. Inoltre

$$\bar{s} = 1 + \frac{\bar{g}u}{1 - \bar{g}u} = \frac{1}{1 - \bar{g}u}$$

Il grafico di \bar{s} in funzione di \bar{g} è in Fig. 69(a). Si noti come il valore tenda ad infinito in prossimità della transizione di fase. Inoltre

$$c = \frac{1}{1 - \frac{1}{2}\bar{g}u}$$

Si veda in Fig. 69(b) il grafico di c in funzione del grado medio \bar{g} . Per un grafo di Poisson si può anche eseguire in modo simbolico la ricorsione (116) e si ottiene

$$s_k = r_k = \frac{k^{k-1} \hat{g}^{k-1} e^{-k\hat{g}}}{k!}, \quad t^k = \frac{k^{k-2} \hat{g}^{k-1} e^{-k\hat{g}}}{k!} c$$

La Tabella 3 riporta i dati ottenuti da una simulazione su un grafo di 10.000 nodi con grado medio $\bar{g} = 1.5$, per il quale esiste una componente gigante.

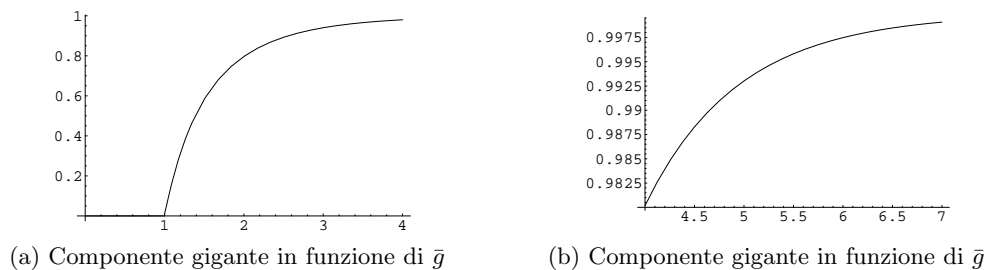


Figura 68.

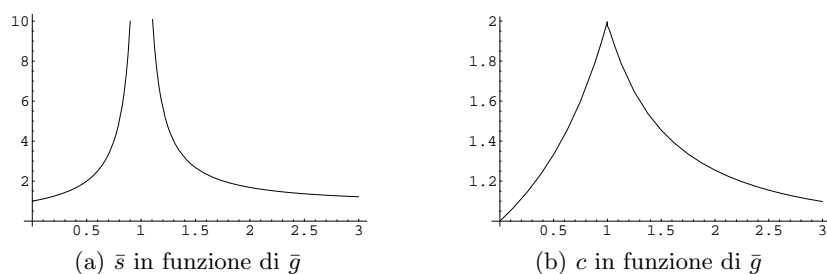


Figura 69.

54 Grafi con distribuzione esponenziale

Un grafo ha la distribuzione esponenziale se p_k è proporzionale a a^k con $0 < a < 1$ parametro. Sia allora

$$p_k = (1 - a) a^k, \quad k \geq 0$$

Allora si ha

$$G_0(x) = \sum_{k \geq 0} (1 - a) a^k x^k = \frac{1 - a}{1 - a x}$$

$$G'_0(x) = \frac{(1 - a) a}{(1 - a x)^2}, \quad G'_0(1) = \frac{a}{1 - a} = \bar{g}$$

$$G_1(x) = \frac{(1 - a)^2}{(1 - a x)^2}, \quad G'_1(x) = \frac{2 a (1 - a)^2}{(1 - a x)^3}$$

L'equazione $u = G_1(u)$ è

$$u = \frac{(1 - a)^2}{(1 - a u)^2}$$

Si tratta di un'equazione di terzo grado. Però una radice è nota ($u = 1$) e quindi basta risolvere un'equazione di secondo grado. Si ottiene

$$u_1 = 1, \quad u_2 = \frac{2 - a - \sqrt{a(4 - 3a)}}{2a}, \quad u_3 = \frac{2 - a + \sqrt{a(4 - 3a)}}{2a}$$

Per $a < 1/3$ si ha $u_3 > u_2 > 1$ e quindi $\hat{u} = 1$. Per tali valori non è presente quindi la componente gigante. Per $1/3 < a < 1$ si ha $u_3 > 1 > u_2 = \hat{u}$ e quindi c'è la componente gigante. La transizione di fase si ha per il valore $a = 1/3$. Riassumendo

$$\hat{u} = \begin{cases} 1 & 0 \leq a \leq 1/3 \\ \frac{2-a-\sqrt{4a-3a^2}}{2a} & 1/3 \leq a < 1 \end{cases} \quad (120)$$

Si può esprimere $a \geq 1/3$ in funzione di u da (120), ovvero,

$$a = \frac{1}{1 + \sqrt{u} + u} \quad (121)$$

Per $1/3 \leq a < 1$ c'è la componente gigante e si ha

$$v = G_0(u) = \frac{1-a}{1-au} = \frac{\sqrt{(4-3a)a}-a}{2a} = \sqrt{u}, \quad \hat{g} = \frac{u G_0'(u)}{G_0(u)} = \frac{au}{1-au} = \frac{u}{1+\sqrt{u}}$$

$$\hat{p}_k = \frac{u^k}{v} p_k = (1-au)(au)^k$$

Quindi anche le componenti piccole sono esponenziali ma con un parametro $au < 1/3$ come si vede da (121). Allora si ha

$$\hat{G}_0(x) = \frac{1-au}{1-aux}, \quad \hat{G}_1(x) = \frac{(1-au)^2}{(1-aux)^2}$$

Anche per i grafi esponenziali è possibile eseguire simbolicamente la ricorsione (116) e ottenere

$$s_k = \frac{1}{2k-1} \binom{3(k-1)}{k-1} (1-au)^{2k-1} (au)^{k-1}, \quad r_k = \frac{1}{k} \binom{3(k-1)+1}{k-1} (1-au)^{2k} (au)^{k-1}$$

e

$$t_k = \frac{1}{k(2k-1)} \binom{3(k-1)}{k-1} (1-au)^{2k-1} (au)^{k-1} c$$

È noto che

$$\sum_{k \geq 0} \frac{1}{2k+1} \binom{3k}{k} x^k = \frac{2}{\sqrt{3x}} \sin \left(\frac{1}{3} \arcsin \sqrt{\frac{27x}{4}} \right)$$

e quindi si ha

$$H_0(x) = \sum_{k \geq 1} s_k x^k = \frac{2}{au} \sqrt{\frac{aux}{3}} \sin \left(\frac{1}{3} \arcsin \left(\frac{3(1-au)}{2} \sqrt{3aux} \right) \right)$$

dove $au = u/(1 + \sqrt{u} + u)$ usando (121). Si ha ancora

$$\bar{s} = H_0'(1) = \frac{1-2au}{1-3au}, \quad c = \frac{1}{1-\hat{g}/2} = \frac{1-au}{1-3au/2}$$

che è uguale a $4/3$ all'apparire della componente gigante ($au = 1/3$).

La Tabella 4 riporta i dati ottenuti da una simulazione su un grafo di 10.000 nodi con $a = .5$, per il quale esiste una componente gigante.

<i>teorico</i>	u	v	\hat{g}					
sperimentale	0.3820 0.3837	0.6180 0.6197	0.2361 0.2346					
<i>teorico</i>	\hat{p}_0	\hat{p}_1	\hat{p}_2	\hat{p}_3	\hat{p}_4	\hat{p}_5	\hat{p}_6	\hat{p}_7
sperimentale	0.8090 0.8068	0.1545 0.1585	0.0295 0.0287	0.0056 0.0053	0.0011 0.0005	0.0002 0.0002	0 0	0 0
<i>teorico</i>	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
sperimentale	0.8090 0.8068	0.1011 0.1065	0.0379 0.0416	0.0190 0.0181	0.0109 0.0105	0.0067 0.0058	0.0044 0.0023	0.0030 0.0026
<i>teorico</i>	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8
sperimentale	0.6545 0.6743	0.1636 0.1712	0.0716 0.0680	0.0384 0.0354	0.0229 0.0180	0.0145 0.0078	0.0097 0.0053	0.0067 0.0058
<i>teorico</i>	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
sperimentale	0.9173 0.9141	0.0573 0.0603	0.0143 0.0157	0.0054 0.0051	0.0025 0.0024	0.0013 0.0011	0.0007 0.0004	0.0004 0.0004

Tabella 4. Valori teorici e simulati per un grafo esponenziale con $a = 0.5$

Teoria dei Giochi

55 Definizione normale di gioco

In un gioco sono definiti n giocatori. Ogni giocatore ha a sua disposizione un insieme S_i finito di *strategie*. Ogni giocatore i sceglie una strategia $x_i \in S_i$ e riceve un compenso definito da una funzione

$$u_i : S_1 \times S_2 \times \cdots \times S_n \rightarrow R$$

che valuta quindi l'esito del gioco per il giocatore i tenendo conto della strategia scelta da ogni giocatore.

Le funzioni u_i sono in generale *utilità*. Si ricorda che una funzione d'utilità u_i viene definita fissando due esiti fra tutti quelli possibili del gioco, uno il peggiore possibile e l'altro il migliore possibile per il giocatore i . Poi si valuta per quale probabilità p il giocatore i è indifferente fra ricevere in modo certo un determinato esito oppure ricevere con probabilità p l'esito migliore e con probabilità $(1 - p)$ l'esito peggiore. Tale valore di probabilità viene definito utilità di quel determinato esito. Ovviamente l'utilità è compresa fra 0 (utilità dell'esito peggiore) e 1 (esito migliore). Comunque nulla vieta di misurare gli esiti del gioco con altre unità di misura, ad esempio monetarie. Tuttavia si tenga presente che anche i valori monetari hanno una loro utilità che non necessariamente è lineare e può essere diversa da giocatore a giocatore.

L'obiettivo della teoria dei giochi è capire se il gioco ha una 'soluzione', ovvero se esistono strategie, una per ogni giocatore, verso le quali il gioco 'razionalmente' si indirizza e che rappresenta il 'migliore' risultato possibile per ogni giocatore.

Nella maggior parte dei casi non è univoca la definizione di miglior risultato. Innanzitutto vanno specificati alcuni elementi fondamentali del gioco, se ad esempio è possibile formare *coalizioni* e, se questo è possibile e il compenso è di tipo monetario, se sia possibile condividere il compenso fra i membri della coalizione. Se i giocatori sono soltanto due allora non sono possibili coalizioni e si parla più genericamente di *cooperazione* fra i giocatori.

Due concetti di soluzione sono molto importanti. Si definisce soluzione di *equilibrio* un insieme di strategie (x_1^*, \dots, x_n^*) tali che:

$$u_i(x_1^*, \dots, x_i^*, \dots, x_n^*) \geq u_i(x_1^*, \dots, x_i, \dots, x_n^*) \quad \forall x_i \in S_i \quad \forall i$$

In altre parole, ogni giocatore non ha interesse a cambiare unilateralmente la sua strategia x_i^* con un'altra strategia x_i . Quindi tutti i giocatori si trovano 'bloccati' nell'insieme di strategie (x_1^*, \dots, x_n^*) .

Si definisce soluzione di *salvaguardia* o di *sicurezza* per il giocatore i la strategia x_i^* calcolata nel seguente modo: per ogni fissata strategia x_i si trova il minimo rispetto a tutte le strategie di tutti gli altri giocatori

$$\psi_i(x_i) := \min_{x_j: j \neq i} u_i(x_1, \dots, x_i, \dots, x_n)$$

e poi si massimizza $\psi_i(x_i)$, cioè

$$\psi_i(x_i^*) = \max_{x_i \in S_i} \psi_i(x_i)$$

L'idea è quindi che il giocatore, con un atteggiamento di grande prudenza e pessimismo, valuta, per ogni sua strategia, quale sia il peggior esito del gioco causato dalle scelte degli altri giocatori e poi sceglie quella strategia che massimizza questo peggior esito. Per questo motivo questa strategia viene anche chiamata *maximin*. Se ogni giocatore sceglie una strategia di salvaguardia ha la garanzia di non ottenere meno di $\psi_i(x_i^*)$. Potrebbe invece, a causa delle scelte degli altri giocatori, non necessariamente motivati da 'malvolenza' nei confronti del giocatore stesso, ricevere di più di $\psi_i(x_i^*)$. Infatti si ha dalla definizione della funzione $\psi_i(x_i)$

$$\psi_i(x_i^*) \leq u_i(x_1^*, \dots, x_i^*, \dots, x_n^*), \quad \forall i$$

Se invece avviene che

$$\psi_i(x_i^*) = u_i(x_1^*, \dots, x_i^*, \dots, x_n^*), \quad \forall i$$

allora si dice che il gioco ha una *soluzione di salvaguardia*.

Un altro concetto importante è quello di *dominanza*. Per un giocatore k la strategia x_k domina la strategia y_k se $u_i(x_1, \dots, x_k, \dots, x_n) \geq u_i(x_1, \dots, y_k, \dots, x_n)$ per ogni possibile scelta $x_i, i \neq k$. Per un giocatore razionale non ha senso usare la strategia y_k . Si noti che una strategia dominata non può essere in una soluzione di equilibrio.

Un altro concetto apparentemente simile è quello di *dominanza secondo Pareto*. L'insieme di strategie (x_1, \dots, x_n) domina secondo Pareto l'insieme di strategie (y_1, \dots, y_n) se $u_i(x_1, \dots, x_n) \geq u_i(y_1, \dots, y_n)$ per ogni giocatore i con almeno una disuguaglianza stretta. Sembrerebbe naturale che la soluzione di un gioco non possa essere dominata secondo Pareto da qualche insieme di strategie. Vedremo che non è così in generale.

56 Giochi a due persone a somma zero

Il caso più semplice di gioco è quello in cui sono presenti solo due giocatori e $u_1(x_1, x_2) = -u_2(x_1, x_2)$. Quindi quello che un giocatore guadagna, l'altro perde. Di conseguenza non c'è spazio per la cooperazione e questo fatto rende semplice lo studio di questi giochi. Nei giochi a somma zero con due giocatori, siccome $u_1(x_1, x_2) = -u_2(x_1, x_2)$, si definisce solo una funzione $u(x_1, x_2)$, valida per entrambi i giocatori, con la convenzione che il primo giocatore è interessato a massimizzare u mentre il secondo giocatore è interessato a minimizzare u .

È consuetudine rappresentare il gioco con una tabella in cui le righe corrispondono alle strategie del giocatore A (che ad esempio massimizza) e le colonne a quelle del giocatore B (che invece minimizza). In ogni riquadro della tabella si indica l'ammontare che A riceve (e B paga) in base alle due scelte delle strategie. Ad esempio il gioco potrebbe essere descritto dalla seguente tabella dove $A1, A2, A3, A4$ sono le strategie di A e $B1, B2, B3, B4, B5$ quelle di B:

	$B1$	$B2$	$B3$	$B4$	$B5$
$A1$	-4	1	-3	-5	-1
$A2$	3	-8	-6	0	2
$A3$	5	3	8	2	3
$A4$	-4	-3	-1	1	-5

Vediamo se il gioco ha una soluzione di salvaguardia. Se il giocatore A sceglie la strategia $A1$, l'esito peggiore è $\psi_1(A1) = -5$, se sceglie la strategia $A2$ l'esito peggiore è $\psi_1(A2) = -8$. Proseguendo si trovano gli esiti $\psi_1(A3) = 2$ e $\psi_1(A4) = -5$. Il migliore di questi è $\psi_1(A3) = 2$. Quindi $A3$ è la strategia di salvaguardia per A. Analogamente, se il giocatore B sceglie la strategia $B1$ l'esito peggiore è $\psi_2(B1) = 5$ e proseguendo gli esiti peggiori sono $\psi_2(B2) = 3$, $\psi_2(B3) = 8$, $\psi_2(B4) = 2$, $\psi_2(B5) = 3$. Allora per B la strategia di salvaguardia è $B4$ con l'esito $\psi_2(B4) = 2$. Siccome

$$u_1(A3, B4) = \psi_1(A3), \quad u(A3, B4) = \psi_2(B4)$$

in questo caso il gioco ha una soluzione di salvaguardia che corrisponde all'esito 2 per la coppia di strategie $A3, B4$.

Si noti che la coppia di strategie $A3, B4$ è anche soluzione di equilibrio in quanto non è conveniente per nessuno dei due giocatori un cambiamento unilaterale di strategia. Questa coincidenza fra soluzione di equilibrio e soluzione di salvaguardia non è casuale. Dimostriamo adesso che deve essere così per i giochi a somma zero con due giocatori. Sia (x_1^*, x_2^*) una soluzione di salvaguardia. Allora in base alle scelte del giocatore B (che valuta l'esito peggiore per ogni sua strategia e in particolare per x_2^*)

$$u(x_1^*, x_2^*) \geq u(x_1, x_2^*) \quad \forall x_1$$

e in base alle scelte di A

$$u(x_1^*, x_2^*) \leq u(x_1^*, x_2) \quad \forall x_2$$

e dalle due relazioni si vede che (x_1^*, x_2^*) è anche una soluzione di equilibrio. Viceversa sia (x_1^*, x_2^*) una soluzione di equilibrio. Quindi

$$\forall x_1 \quad u(x_1, x_2^*) \leq u(x_1^*, x_2^*) \leq u(x_1^*, x_2) \quad \forall x_2 \quad (122)$$

Sia (\hat{x}_1, \hat{x}_2) una soluzione di salvaguardia con $u(\hat{x}_1, \hat{x}_2) \neq u(x_1^*, x_2^*)$. Allora deve essere

$$\forall x_1 \quad u(x_1, \hat{x}_2) \leq u(\hat{x}_1, \hat{x}_2) \leq u(\hat{x}_1, x_2) \quad \forall x_2$$

e in particolare

$$u(x_1^*, \hat{x}_2) \leq u(\hat{x}_1, \hat{x}_2) \leq u(\hat{x}_1, x_2^*) \quad (123)$$

Anche (122) si può particularizzare a

$$u(\hat{x}_1, x_2^*) \leq u(x_1^*, x_2^*) \leq u(x_1^*, \hat{x}_2) \quad (124)$$

Mettendo assieme (123) e (124) si ottiene

$$u(\hat{x}_1, x_2^*) \leq u(x_1^*, x_2^*) \leq u(x_1^*, \hat{x}_2) \leq u(\hat{x}_1, \hat{x}_2) \leq u(\hat{x}_1, x_2^*)$$

che ovviamente implica

$$u(\hat{x}_1, x_2^*) = u(x_1^*, x_2^*) = u(x_1^*, \hat{x}_2) = u(\hat{x}_1, \hat{x}_2) = u(\hat{x}_1, x_2^*) \quad (125)$$



Figura 70.

contraddicendo l'ipotesi $u(\hat{x}_1, \hat{x}_2) \neq u(x_1^*, x_2^*)$. Se invece si ammette $u(\hat{x}_1, \hat{x}_2) = u(x_1^*, x_2^*)$ allora da (125) si deduce che anche (x_1^*, \hat{x}_2) e (\hat{x}_1, x_2^*) sono soluzioni. In altre parole A può scegliere indifferentemente fra x_1^* e \hat{x}_1 e B può scegliere indifferentemente fra x_2^* e \hat{x}_2 ad avere comunque una strategia che è sia di salvaguardia che di equilibrio.

Consideriamo il seguente esempio tratto da un episodio della Seconda Guerra Mondiale e precisamente la Battaglia del Mare di Bismarck (si veda la Fig. 70). Per una conoscenza più approfondita di questo episodio, si veda http://en.wikipedia.org/wiki/Battle_of_the_Bismarck_Sea. Secondo le informazioni raccolte dai servizi segreti americani, un convoglio giapponese sarebbe partito da Rabal, all'estremo destro dell'isola di Nuova Britannia verso Lae nella Nuova Guinea. Non era tuttavia noto se il comandante giapponese avrebbe scelto la rotta a nord o a sud dell'isola. Il viaggio avrebbe comunque richiesto tre giorni. La rotta nord avrebbe comportato scarsa visibilità, a differenza della rotta sud. Il comandante americano Kenney poteva concentrare la sua aviazione o su una rotta o sull'altra. Una volta avvistato il convoglio, questo poteva essere bombardato fino all'arrivo a Lae.

Il numero di giorni di bombardamento possibili a seconda delle rispettive scelte può essere rappresentato dalla tabella, dove le righe corrispondono alla scelta di Kenney di inviare l'aviazione a nord (N) o a sud (S). Le colonne invece corrispondono alla scelta del comandante giapponese di inviare il convoglio a nord (N) o a sud (S).

	N	S
N	2	2
S	1	3

Se il comandante giapponese sceglie la rotta nord e Kenney indovina questa scelta, i tre giorni di viaggio, a causa della scarsa visibilità, diventano due giorni effettivi di bombardamento. Se invece entrambi scelgono la rotta sud, i tre giorni di viaggio corrispondono a tre giorni di bombardamento. Se invece Kenney sbaglia nella scelta della rotta, i giorni di bombardamento si riducono di uno a causa del riposizionamento dell'aviazione sull'altra rotta.

Kenney affrontò il problema secondo un'ottica di salvaguardia. Se avesse inviato l'aviazione a nord avrebbe avuto la garanzia di 2 giorni di bombardamento, mentre inviandola a sud avrebbe avuto la garanzia di solo un giorno. Quindi scelse di inviare l'aviazione a nord. Anche il comandante giapponese ragionò in termini di salvaguardia. Se avesse inviato il convoglio a nord avrebbe potuto incorrere in due giorni di bombardamento e inviandolo a sud in tre giorni di bombardamento. Quindi scelse di inviare il convoglio a nord. Si tratta evidentemente di una soluzione di equilibrio e anche di salvaguardia.

Anche se l'esito storico della Battaglia del Mare di Bismarck si risolse in una disfatta per i giapponesi a seguito dei due giorni di bombardamento, e dalla quale seguì la decisione di abbandonare la Nuova Guinea, il comandante giapponese non può però essere accusato di avere 'sbagliato' strategia. Data la situazione la sua decisione fu la migliore possibile.

Non è detto che esistano strategie di equilibrio e di salvaguardia. Ad esempio il gioco della morra cinese (carta, forbici e sasso) ha la seguente tabella

	C	F	S
C	0	-1	1
F	1	0	-1
S	-1	1	0

e certamente non presenta nessuna soluzione di equilibrio. Quando non esistono soluzioni di equilibrio (o di salvaguardia) si estende il concetto di strategia pensando di scegliere una particolare strategia secondo una distribuzione di probabilità e valutando il valore atteso del gioco. Con quest'estensione una strategia per il giocatore A è una distribuzione di probabilità p_i , $i \in S_1$ e per il giocatore B una distribuzione di probabilità q_j , $j \in S_2$ e l'esito di una particolare scelta di strategia è dato da

$$\sum_{i \in S_1} \sum_{j \in S_2} p_i q_j u(x_i, x_j)$$

In questo caso si parla di *strategie miste* e le strategie scelte in modo deterministico vengono quindi dette *pure*. Per determinare una strategia di salvaguardia A deve valutare

$$\max_p \min_q \sum_{i \in S_1} \sum_{j \in S_2} p_i q_j u(x_i, x_j) \quad (126)$$

Per ogni fissato p , il termine $\sum_{i \in S_1} p_i u(x_i, x_j)$ è costante e dipende solo da j . Allora possiamo scrivere (126) come

$$\max_p \min_q \sum_{j \in S_2} q_j \sum_{i \in S_1} p_i u(x_i, x_j) \quad (127)$$

Siccome le probabilità q_j devono avere somma 1, il minimo su q si ottiene assegnando il valore $q_j = 1$ a quell'indice j per cui $\sum_{i \in S_1} p_i u(x_i, x_j)$ è minimo. Allora (127) diventa

$$\max_p \min_j \sum_{i \in S_1} p_i u(x_i, x_j) \quad (128)$$

che si può a sua volta scrivere come

$$\begin{aligned}
 \max \quad & v \\
 v \leq \quad & \sum_{i \in S_1} p_i u(x_i, x_j) \quad j \in S_2 \\
 \sum_{i \in S_1} \quad & p_i = 1 \\
 p_i \geq 0 \quad & i \in S_1
 \end{aligned} \tag{129}$$

In modo del tutto simmetrico B deve valutare

$$\min_q \max_p \sum_{i \in S_1} \sum_{j \in S_2} p_i q_j u(x_i, x_j) \tag{130}$$

Per ogni fissato q , il termine $\sum_{j \in S_2} q_j u(x_i, x_j)$ è costante e dipende solo da i . Quindi (130) diventa

$$\min_q \max_i \sum_{j \in S_2} q_j u(x_i, x_j) \tag{131}$$

che si può a sua volta scrivere come

$$\begin{aligned}
 \min \quad & w \\
 w \geq \quad & \sum_{j \in S_2} q_j u(x_i, x_j) \quad i \in S_1 \\
 \sum_{j \in S_2} \quad & q_j = 1 \\
 q_j \geq 0 \quad & j \in S_2
 \end{aligned} \tag{132}$$

La coppia di problemi (129) e (132) sono di programmazione lineare e sono uno il duale dell'altro. In base al celebre teorema di dualità forte della programmazione lineare i due valori ottimi coincidono e quindi esistono strategie p^* e q^* di salvaguardia e d'equilibrio tali che

$$\max_p \min_q \sum_{i \in S_1} \sum_{j \in S_2} p_i q_j u(x_i, x_j) = \sum_{i \in S_1} \sum_{j \in S_2} p_i^* q_j^* u(x_i, x_j) = \min_q \max_p \sum_{i \in S_1} \sum_{j \in S_2} p_i q_j u(x_i, x_j)$$

Per giochi con solo due strategie per giocatore il calcolo si effettua per via grafica in modo molto semplice. Sia data ad esempio la tabella

	<i>B1</i>	<i>B2</i>
<i>A1</i>	-4	1
<i>A2</i>	5	-3

Sia p la probabilità per la strategia *A1* e $1 - p$ per *A2*. Il valore per ognuna delle due strategie di B in funzione di p è rappresentata in Fig. 71-(a). Siccome A valuta il peggiore esito deve prendere il minimo delle due funzioni (Fig. 71-(b)). Il massimo di questa funzione si ottiene per $p = 8/13$ e tale massimo vale $-7/13$

Sia invece q la probabilità per la strategia *B1* e $1 - q$ per *B2*. Il valore per ognuna delle due strategie di A in funzione di q è rappresentata in Fig. 72-(a). Per valutare il peggiore esito B deve prendere il massimo delle due funzioni (Fig. 72-(b)). Il minimo di questa funzione si ottiene per $q = 4/13$ e vale $-7/13$

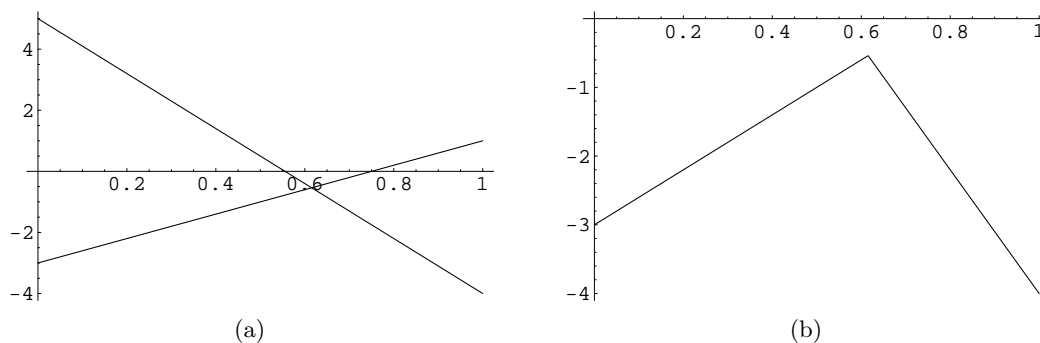


Figura 71.

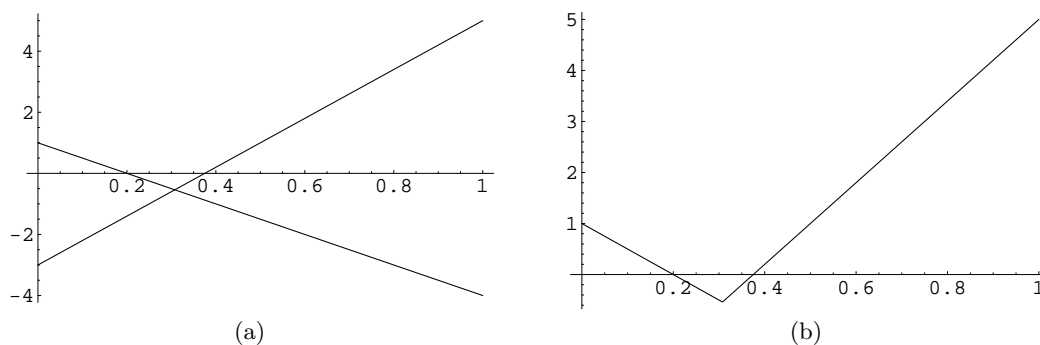


Figura 72.

57 Giochi a somma zero in forma estesa

Sono giochi a somma zero tutti i giochi normalmente definiti tali, fra cui ad esempio gli scacchi, il nim, la morra, i giochi a carte. Molti giochi si svolgono secondo una sequenza di mosse da parte dei due giocatori. Normalmente la sequenza è finita (il gioco degli scacchi potrebbe avere una sequenza infinita di mosse teoricamente, ad esempio spostando avanti e indietro sempre gli stessi pezzi, però tralasciando le mosse che riportano ad una precedente situazione si può pensare che anche il gioco degli scacchi abbia un numero finito di mosse). Dopo l'esecuzione dell'ultima mossa si decide il compenso per uno dei due giocatori (che per l'altro diventa un pagamento). Spesso il compenso è semplicemente il valore $+1$ o -1 ad indicare la vittoria o la sconfitta (può anche essere presente lo 0 , ad indicare la parità).

Se nel gioco l'informazione è completa (il giocatore conosce la situazione dell'avversario) e non sono presenti elementi casuali (nei giochi a carte, le carte che si pescano costituiscono una mossa della 'natura' che avviene con probabilità definita), è noto a priori il vincitore.

Un gioco ad informazione completa e deterministico può essere rappresentato secondo un albero. Tale rappresentazione viene chiamata *forma estesa*. La radice dell'albero corrisponde all'inizio del gioco. Dal nodo radice escono tanti archi quante sono le mosse che il giocatore che inizia può fare. Da questi nodi escono tanti archi quante sono le mosse che può fare il secondo giocatore e così di seguito sino ad arrivare alle foglie dell'albero dove il gioco termina. Le foglie dell'albero vengono etichettate con il valore del compenso (eventualmente negativo) per il giocatore A (che è un pagamento per il giocatore B). Se il giocatore che effettua l'ultima mossa è A sceglierà la mossa che massimizza l'esito del gioco e se è invece B quella che minimizza. A questo punto l'esito del gioco non è solo noto nelle foglie dell'albero ma anche nei nodi adiacenti alle foglie. Procedendo ricorsivamente verso la radice si arriva ad assegnare un esito del gioco alla radice e questo stabilisce chi vince e chi perde.

Può sembrare strano, ma il precedente risultato implica che teoricamente è noto a priori il vincitore di un incontro di scacchi (o se magari è sempre garantito un esito di parità). Siccome un tale albero ha un numero di nodi al di là di ogni possibile rappresentazione e computazione, la procedura ricorsiva delineata non è fattibile (né mai lo sarà probabilmente) e quindi il gioco degli scacchi continua a mantenere il suo interesse.

Il gioco del Nim è anche un gioco ad informazione completa e deterministica. Analizziamo il gioco del Nim eseguito con tre mucchi di due pezzi ciascuno. In Fig. 73-(a) è raffigurato l'albero delle mosse. A dire il vero non è un albero. Tuttavia, per risparmiare spazio, nodi equivalenti sono stati rappresentati da un unico nodo. Ricordiamo che nel gioco del Nim la mossa di ogni giocatore consiste nel togliere un numero arbitrario non nullo di pezzi da un mucchio. In una versione del gioco vince chi effettua l'ultima mossa, in un'altra versione perde chi effettua l'ultima mossa. In ogni nodo dell'albero è indicata la situazione dei tre mucchi dopo la mossa. Nei nodi tondi gioca A e in quelli quadrati gioca B. Nei nodi terminali è indicato il vincitore.

In Fig. 73-(b) sono indicati i vincitori ai vari nodi dopo avere effettuato la procedura ricorsiva. Se il giocatore ha almeno una mossa che lo porta in un nodo dove è vincitore, allora sceglierà quella mossa e il suo nodo può essere etichettato con la sua vittoria, altrimenti la vittoria è dell'avversario. Come si vede il vincitore è A.

Una strategia va vista come un manuale d'istruzioni che dica al giocatore cosa fare in ogni possibile circostanza. Ogni strategia è un diverso manuale. In questo esempio A ha a disposizione 144 strategie, che corrispondono al prodotto del numero di archi uscenti dai nodi dove gioca A. B ne ha invece 36. Quindi la rappresentazione di questo particolare gioco in forma normale richiederebbe una tabella con 144 righe e 36 colonne. In ogni riquadro ci sarebbe un +1 oppure un -1 a seconda di chi vince (scelta la strategia il gioco evolve in maniera automatica individuando un percorso dalla radice ad una foglia).

La strategia vincente (per A) corrisponde ad una riga di tutti 1. Chiaramente si tratta di una soluzione sia di equilibrio che di salvaguardia. Per il gioco del Nim il calcolo della strategia per vincere il gioco in generale è molto interessante e viene qui descritto nel riquadro.

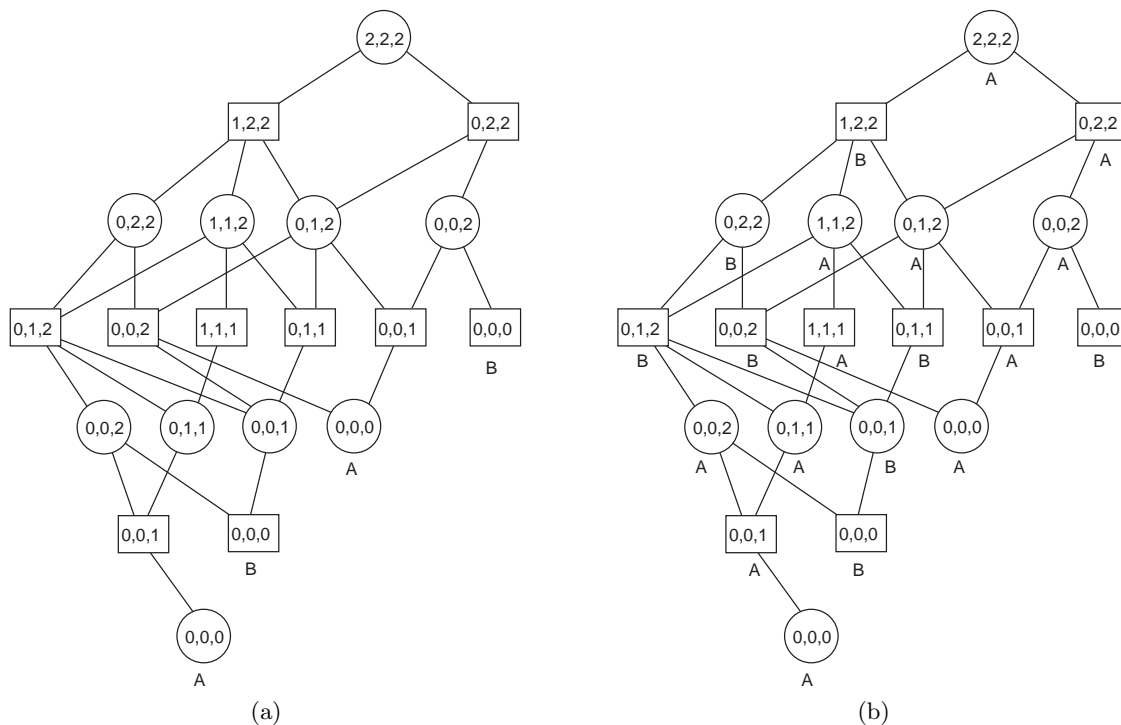


Figura 73.

Gioco del Nim

Sono dati n mucchi di pezzi con $M = (m_1, m_2, \dots, m_n)$ pezzi per mucchio. In ogni mossa un giocatore toglie un numero arbitrario di pezzi da un solo mucchio. I giocatori sono due. In una versione del gioco vince chi toglie l'ultimo pezzo. In un'altra versione perde chi toglie l'ultimo pezzo.

Siano (a_i^j) le cifre della rappresentazione binaria di m_i

$$m_i = a_i^0 2^0 + a_i^1 2^1 + a_i^2 2^2 + a_i^3 2^3 + \dots$$

siano

$$b^j := \sum_i a_i^j$$

Definiamo come 'canonica' una m -pla M per la quale tutti i b^j sono pari. Sia \mathcal{K} l'insieme delle m -ple canoniche. Si noti che $(0, 0, \dots, 0) \in \mathcal{K}$. Definiamo 'canonica' una mossa che porta in una configurazione canonica e come 'non canonica' una mossa che porta in una configurazione non canonica.

Se $M \in \mathcal{K}$ e si effettua una qualsiasi mossa $M \rightarrow M'$ allora $M' \notin \mathcal{K}$. Infatti se m_i e m'_i sono i due valori cambiati, almeno uno dei valori $a_i^j \in \{0, 1\}$ deve cambiare, altrimenti $m_i = m'_i$, e quindi almeno uno dei b^j deve diventare dispari. Allora, da una configurazione canonica si possono effettuare solo mosse non canoniche.

Se invece $M \notin \mathcal{K}$ esiste sempre una mossa $M \rightarrow M'$ tale che $M' \in \mathcal{K}$. Sia J l'insieme degli indici j per cui b^j è dispari e sia $k = \max\{j : j \in J\}$. Sia i tale che $a_i^k = 1$ (deve esistere altrimenti $a_i^k = 0, \forall i$, e $k \notin J$) e siano $J_0 := \{j : a_i^j = 0\}$ e $J_1 := \{j : a_i^j = 1\}$. Allora sia

$$\bar{a}_i^j = \begin{cases} 1 & \text{se } j \in J \cap J_0 \\ 0 & \text{se } j \in J \cap J_1 \\ a_i^j & \text{se } j \notin J \end{cases}$$

Quindi \bar{a}_i^j si ottiene 'flipando' i coefficienti in corrispondenza degli j con b^j dispari e quindi per $m'_i = \sum_j \bar{a}_i^j 2^j$ i valori b^j sono tutti pari. Inoltre la mossa è fattibile in quanto $m'_i = \sum_j \bar{a}_i^j 2^j < m_i$, per la scelta di k e di i . Quindi $M' = (m_1, m_2, \dots, m'_i, \dots, m_n) \in \mathcal{K}$. Allora, da una configurazione non canonica esiste sempre una mossa canonica, ma esistono anche mosse non canoniche.

Si noti che se l'avversario ha eseguito una mossa non canonica a partire da una configurazione canonica il coefficiente a_i^k del mucchio i da cui l'avversario ha tolto dei pezzi deve essere nullo, altrimenti ci sarebbero più pezzi di prima nel mucchio. Quindi una successiva mossa canonica non preleva mai pezzi dal mucchio da cui l'avversario ha appena tolto.

Esempio di mossa canonica da una configurazione non canonica:

$$\begin{aligned} 5 &= 1 \cdot 2^0 + 0 \cdot 2^1 + 1 \cdot 2^2 + 0 \cdot 2^3 \\ 3 &= 1 \cdot 2^0 + 1 \cdot 2^1 + 0 \cdot 2^2 + 0 \cdot 2^3 \\ 6 &= 0 \cdot 2^0 + 1 \cdot 2^1 + 1 \cdot 2^2 + 0 \cdot 2^3 \\ (5, 3, 6, 2, 1, 9, 11) &\implies 2 = 0 \cdot 2^0 + 1 \cdot 2^1 + 0 \cdot 2^2 + 0 \cdot 2^3 \\ &1 = 1 \cdot 2^0 + 0 \cdot 2^1 + 0 \cdot 2^2 + 0 \cdot 2^3 \\ &9 = 1 \cdot 2^0 + 0 \cdot 2^1 + 0 \cdot 2^2 + 1 \cdot 2^3 \\ &11 = 1 \cdot 2^0 + 0 \cdot 2^1 + 1 \cdot 2^2 + 1 \cdot 2^3 \end{aligned}$$

Allora $J = \{0, 1, 2\}$, $k = 2$ e le mosse canoniche sono tre a seconda se $i = 1$, $i = 3$ oppure $i = 7$. Se $i = 1$ (mucchio da 5) allora $J_0 = \{1\}$ e $J_1 = \{0, 2\}$ per cui $m'_1 = 0 \cdot 2^0 + 1 \cdot 2^1 + 0 \cdot 2^2 + 0 \cdot 2^3 = 2$. Se $i = 3$ (mucchio da 6) il nuovo numero è $m'_3 = 1 \cdot 2^0 + 0 \cdot 2^1 + 0 \cdot 2^2 + 0 \cdot 2^3 = 1$, e se $i = 7$ (mucchio da 11) il nuovo numero è $m'_7 = 0 \cdot 2^0 + 1 \cdot 2^1 + 0 \cdot 2^2 + 1 \cdot 2^3 = 9$.

Se chi toglie l'ultimo pezzo vince, la strategia ottima consiste nell'effettuare sempre una mossa canonica. Tale strategia è vincente anche nel caso di vittoria di chi non toglie l'ultimo pezzo però con una variante nelle ultime mosse. Chiamiamo 'configurazione speciale' una configurazione in cui è presente un solo mucchio con più di un pezzo e gli altri mucchi hanno zero o un pezzo. In questo caso la mossa canonica porterebbe ad un numero pari di mucchi da un pezzo con conseguente sconfitta. Bisogna invece prelevare dal mucchio più

grande tutti i pezzi del mucchio oppure tutti meno uno in modo che rimanga un numero dispari di mucchi da un pezzo assicurando così la vittoria.

Ci si deve chiedere se una configurazione speciale possa capitare anche all'avversario e se si tratta dell'unico caso in cui abbandonare la strategia di eseguire mosse canoniche. Una configurazione speciale non può essere canonica e quindi se l'avversario si trova sempre di fronte una configurazione canonica non può trovarsi di fronte una configurazione speciale. Prima di arrivare alla configurazione finale $(0, 0, \dots, 0)$ deve necessariamente presentarsi una configurazione speciale. Infatti ad un certo punto deve rimanere un solo mucchio. Se il mucchio ha più di un pezzo si tratta di una configurazione speciale. Se il mucchio rimasto ha un solo pezzo, nella mossa precedente aveva ancora un pezzo (altrimenti la configurazione speciale era già presente nella precedente mossa) e quindi c'era qualche altro mucchio con dei pezzi; ricorsivamente si arriva ad una configurazione speciale.

Se entrambi i giocatori giocano in modo ottimale è noto a priori il vincitore a seconda della configurazione iniziale ed indipendentemente dalla versione del gioco. Quindi (a meno che la configurazione iniziale non sia proprio una speciale) se la configurazione iniziale è canonica, necessariamente chi effettua la prima mossa esegue una mossa non canonica e quindi il secondo giocatore può effettuare sempre mosse canoniche e vincere; se invece la configurazione iniziale è non canonica, chi effettua la prima mossa può effettuare un mossa canonica e quindi vincere.

La difficoltà del gioco consiste nell'effettuare mosse canoniche senza l'aiuto di carta e matita. 'Vedere' i mucchi rappresentati in forma binaria richiede un buon allenamento. Però effettuare il gioco con carta e matita toglie ogni interesse perché diventa troppo facile eseguire la strategia ottimale.

Esempio: al sito http://www.archimedes-lab.org/game_nim/play_nim_game.html si può giocare online con un computer. La configurazione iniziale è $(1, 3, 5, 7)$ ed è canonica in quanto $3 = 2 + 1$, $5 = 4 + 1$, $7 = 4 + 2 + 1$, con quattro 1, due 2 e due 4. Quindi conviene far giocare per primo il computer (altrimenti si perderà sempre, il computer non sbaglia). In questa versione perde chi toglie l'ultimo pezzo. Ecco il resoconto di una giocata con il computer (C=computer, H=Human):

$$(1, 3, 5, 7) \rightarrow^C (1, 3, 4, 7) \rightarrow^H (1, 2, 4, 7) \rightarrow^C (1, 2, 3, 7) \rightarrow^H (1, 2, 3, 0) \rightarrow^C \\ (1, 2, 2, 0) \rightarrow^H (0, 2, 2, 0) \rightarrow^C (0, 1, 2, 0) \rightarrow^H (0, 1, 0, 0) \rightarrow^C (0, 0, 0, 0)$$

La terz'ultima configurazione è speciale e la mossa che si effettua non è più canonica (tutte le altre effettuate da H sono canoniche).

Come secondo esempio si consideri la seguente versione in scala ridotta del gioco della briscola: i semi sono solo due, ad esempio spade e denari, e le carte per seme sono solo tre, asso, due e tre. Indichiamo le sei carte con S1, S2, S3, D1, D2 e D3. I punti e il valore delle carte sono i medesimi che nella briscola ordinaria. Inoltre vengono date soltanto due carte ai giocatori, per cui rimangono sul tavolo solo due carte

di cui una determina la briscola. Per semplificare ancora assumiamo che il gioco sia a carte scoperte e quindi l'informazione è completa.

Descriviamo il gioco per una particolare smazzata (ce ne sono $\binom{6}{2}\binom{4}{2}2 = 180$ in tutto): il giocatore A riceve S1 e D2, B riceve S3 e D1 e sul tavolo rimangono S2 e D3, che determina la briscola di denari. L'albero è rappresentato in Fig. 74. Si noti che, date le regole del gioco, il livello dell'albero non determina automaticamente a chi spetta la mossa. Il simbolo indicato in ogni nodo denota il giocatore cui spetta la mossa. L'esito del gioco è indicato nelle foglie dell'albero con riferimento al giocatore A. In Fig. 75 si vede il calcolo delle strategie ottime. Questa smazzata vede B avvantaggiato in quanto riesce a garantirsi un guadagno di 2.

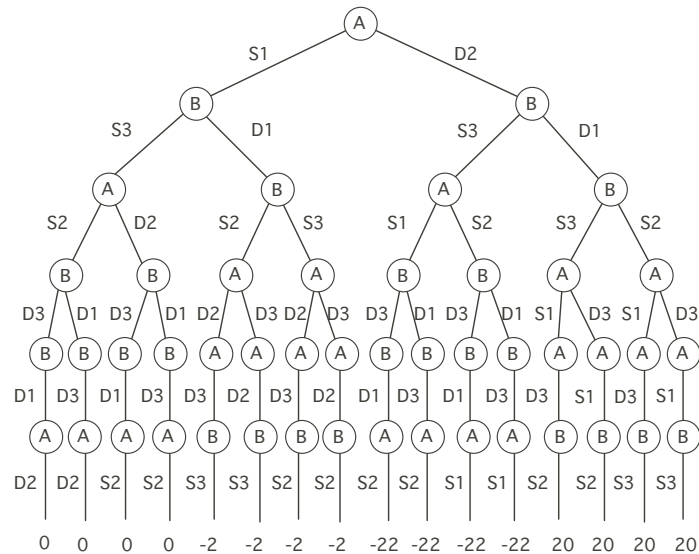


Figura 74.

In Fig. 76 sono evidenziati i nodi controllati dal giocatore A e dal giocatore B. Come si vede A ha a disposizione 16 strategie diverse e B ne ha a disposizione 36. Le strategie di A sono elencate qui sotto, indicando per ogni nodo controllato da A le scelte da effettuare. Le strategie di B sono indicate nella tabella del gioco in forma normale in Fig. 77. Dalla tabella si vede che il gioco ha due coppie di strategie di equilibrio (l'esistenza di strategie di equilibrio è garantita essendo il gioco ad informazione completa) corrispondenti a (1-D1; 2-S3; 3-S3; 7-D1; 8-S2;) per B e le strategie (1-D2; 3-S1; 6-S1; 7-S2) oppure (1-D2; 3-D3; 6-S1; 7-S2;) per A.

1 =1-S1; 2-D3; 4-D2; 5-D2;

2 =1-S1; 2-D3; 4-D2; 5-S2;

3 =1-S1; 2-D3; 4-S2; 5-D2;

4 =1-S1; 2-D3; 4-S2; 5-S2;

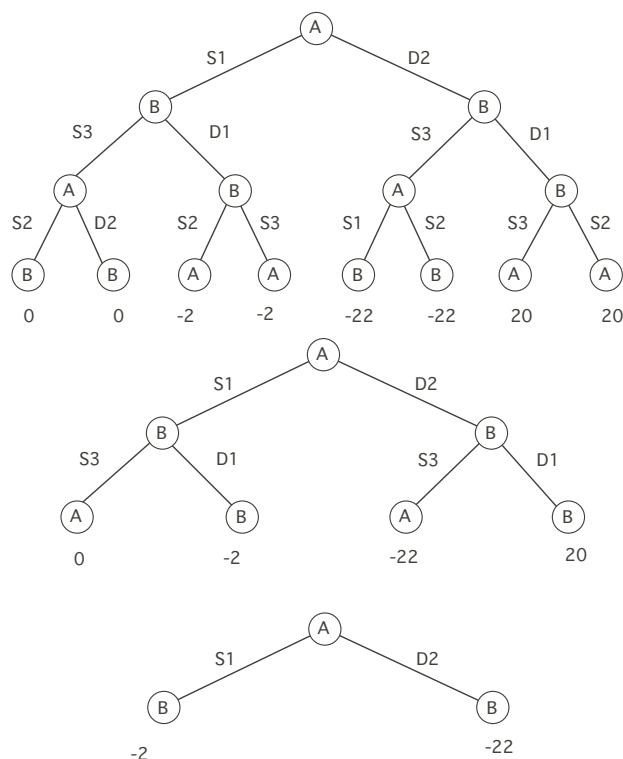


Figura 75.

5 =1-S1; 2-D2; 4-D2; 5-D2;

6 =1-S1; 2-D2; 4-D2; 5-S2;

7 =1-S1; 2-D2; 4-S2; 5-D2;

8 =1-S1; 2-D2; 4-S2; 5-S2;

9 =1-D2; 3-S1; 6-S1; 7-S1;

10 =1-D2; 3-S1; 6-S1; 7-S2;

11 =1-D2; 3-S1; 6-S2; 7-S1;

12 =1-D2; 3-S1; 6-S2; 7-S2;

13 =1-D2; 3-D3; 6-S1; 7-S1;

14 =1-D2; 3-D3; 6-S1; 7-S2;

15 =1-D2; 3-D3; 6-S2; 7-S1;

16 =1-D2; 3-D3; 6-S2; 7-S2;

Se l'informazione non è completa allora una rappresentazione del gioco secondo un albero va modificata. Siccome un giocatore non sa quale sia esattamente la situazione corrente del gioco, ma solo sa che la situazione fa parte di un insieme di situazioni possibili, i nodi corrispondenti a tale insieme formano un *insieme d'informazione* e le mosse del giocatore sono le stesse da tutti i nodi dell'insieme d'informazione.

Come esempio si consideri il seguente gioco: il giocatore A ha in mano tre carte, due re ed un asso. Toglie una carta, senza farla vedere, e depone le altre due sul tavolo a faccia in giù. Se depone asso e re questa mano vale più dell'altra di due re. Poi dichiara "Asso, re" oppure "Due re". Il giocatore B può passare oppure vedere. Se la dichiarazione di A è veritiera e B passa, A vince 1. Se A dichiara la mano migliore di quella

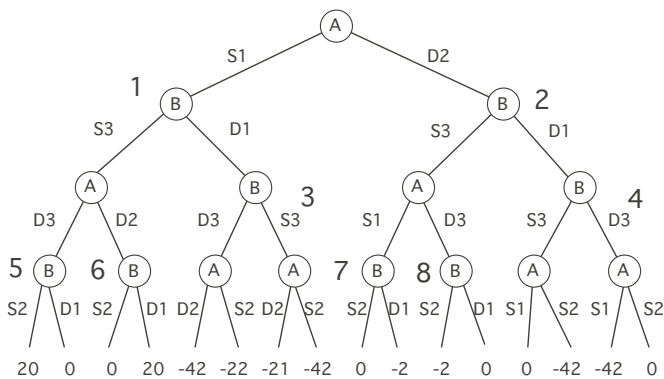
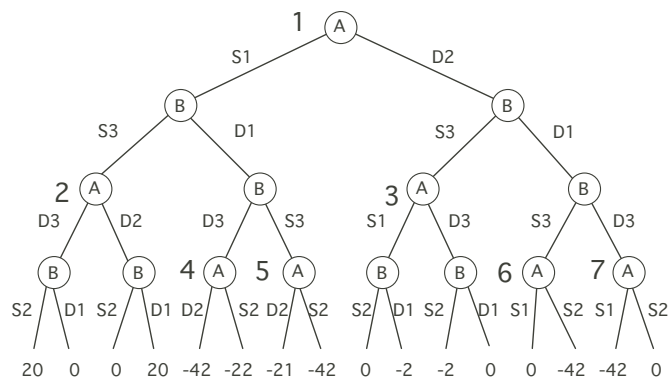


Figura 76.

che è e B passa, A vince 2. Se A dichiara la mano peggiore di quella che è e B passa, A perde 2. Se B vede, le precedenti vincite sono raddoppiate e invertite di segno.

Per rappresentare il gioco in forma estesa possiamo sempre pensare che la dichiarazione di A precede la deposizione delle carte sul tavolo. A questo punto al giocatore B è nota la dichiarazione di A ma non la scelta delle carte. Pertanto il gioco è ad informazione non completa e i nodi corrispondenti ad uno stesso insieme d'informazione sono raggruppati. Il diagramma è in Fig. 78 dove la radice dell'albero corrisponde alla dichiarazione "asso-re" (AR) oppure "due re" (RR). Il successivo livello corrisponde alla scelta delle carte deposte sul tavolo. A questo punto B può passare (P) oppure vedere (V).

Cercando di applicare la tecnica ricorsiva vista precedentemente, dobbiamo tener presente che B, quando decide, non sa quale sia la situazione e quindi deve valutare entrambi gli esiti possibili. Nell'insieme di informazione di sinistra se decide di passare può perdere 1 o 2, mentre se decide di vedere può guadagnare 2 o 4. In questo caso la scelta di vedere domina quella di passare, per cui la scelta migliore per B nell'insieme d'informazione di sinistra è di vedere. Con questa scelta di B i valori -2 e -4 delle foglie vengono riportati

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1-S3;2-S3;5-S2;6-S2;7-S2;8-S2;	20	20	20	20	0	0	0	0	0	0	0	0	-2	-2	-2	-2
1-S3;2-S3;5-S2;6-S2;7-S2;8-D1;	20	20	20	20	0	0	0	0	0	0	0	0	0	0	0	0
1-S3;2-S3;5-S2;6-S2;7-D1;8-S2;	20	20	20	20	0	0	0	0	-2	-2	-2	-2	-2	-2	-2	-2
1-S3;2-S3;5-S2;6-S2;7-D1;8-D1;	20	20	20	20	0	0	0	0	-2	-2	-2	-2	0	0	0	0
1-S3;2-S3;5-S2;6-D1;7-S2;8-S2;	20	20	20	20	20	20	20	20	0	0	0	0	-2	-2	-2	-2
1-S3;2-S3;5-S2;6-D1;7-S2;8-D1;	20	20	20	20	20	20	20	20	0	0	0	0	0	0	0	0
1-S3;2-S3;5-S2;6-D1;7-D1;8-S2;	20	20	20	20	20	20	20	20	-2	-2	-2	-2	-2	-2	-2	-2
1-S3;2-S3;5-S2;6-D1;7-D1;8-D1;	20	20	20	20	20	20	20	20	-2	-2	-2	-2	0	0	0	0
1-S3;2-S3;5-D1;6-S2;7-S2;8-S2;	0	0	0	0	0	0	0	0	0	0	0	0	-2	-2	-2	-2
1-S3;2-S3;5-D1;6-S2;7-S2;8-D1;	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1-S3;2-S3;5-D1;6-S2;7-D1;8-S2;	0	0	0	0	0	0	0	0	-2	-2	-2	-2	-2	-2	-2	-2
1-S3;2-S3;5-D1;6-S2;7-D1;8-D1;	0	0	0	0	0	0	0	0	-2	-2	-2	-2	0	0	0	0
1-S3;2-S3;5-D1;6-D1;7-S2;8-S2;	0	0	0	0	20	20	20	20	0	0	0	0	-2	-2	-2	-2
1-S3;2-S3;5-D1;6-D1;7-S2;8-D1;	0	0	0	0	20	20	20	20	0	0	0	0	0	0	0	0
1-S3;2-S3;5-D1;6-D1;7-D1;8-S2;	0	0	0	0	20	20	20	20	-2	-2	-2	-2	-2	-2	-2	-2
1-S3;2-S3;5-D1;6-D1;7-D1;8-D1;	0	0	0	0	20	20	20	20	-2	-2	-2	-2	0	0	0	0
1-S3;2-D1;5-S2;6-S2;4-S3;	20	20	20	20	0	0	0	0	0	0	-42	-42	0	0	-42	-42
1-S3;2-D1;5-S2;6-S2;4-D3;	20	20	20	20	0	0	0	0	-42	0	-42	0	-42	0	-42	0
1-S3;2-D1;5-S2;6-D1;4-S3;	20	20	20	20	20	20	20	20	0	0	-42	-42	0	0	-42	-42
1-S3;2-D1;5-S2;6-D1;4-D3;	20	20	20	20	20	20	20	20	-42	0	-42	0	-42	0	-42	0
1-S3;2-D1;5-D1;6-S2;4-S3;	0	0	0	0	0	0	0	0	0	0	-42	-42	0	0	-42	-42
1-S3;2-D1;5-D1;6-S2;4-D3;	0	0	0	0	0	0	0	0	-42	0	-42	0	-42	0	-42	0
1-S3;2-D1;5-D1;6-D1;4-S3;	0	0	0	0	20	20	20	20	0	0	-42	-42	0	0	-42	-42
1-S3;2-D1;5-D1;6-D1;4-D3;	0	0	0	0	20	20	20	20	-42	0	-42	0	-42	0	-42	0
1-D1;2-S3;3-D3;7-S2;8-S2;	-42	-42	-22	-22	-42	-42	-22	-22	0	0	0	0	-2	-2	-2	-2
1-D1;2-S3;3-D3;7-S2;8-D1;	-42	-42	-22	-22	-42	-42	-22	-22	0	0	0	0	0	0	0	0
1-D1;2-S3;3-D3;7-D1;8-S2;	-42	-42	-22	-22	-42	-42	-22	-22	-2	-2	-2	-2	-2	-2	-2	-2
1-D1;2-S3;3-D3;7-D1;8-D1;	-42	-42	-22	-22	-42	-42	-22	-22	-2	-2	-2	-2	0	0	0	0
1-D1;2-S3;3-S3;7-S2;8-S2;	-21	-42	-21	-42	-21	-42	-21	-42	0	0	0	0	-2	-2	-2	-2
1-D1;2-S3;3-S3;7-S2;8-D1;	-21	-42	-21	-42	-21	-42	-21	-42	0	0	0	0	0	0	0	0
1-D1;2-S3;3-S3;7-D1;8-S2;	-21	-42	-21	-42	-21	-42	-21	-42	-2	-2	-2	-2	-2	-2	-2	-2
1-D1;2-S3;3-S3;7-D1;8-D1;	-21	-42	-21	-42	-21	-42	-21	-42	-2	-2	-2	-2	0	0	0	0
1-D1;2-D1;3-D3;4-S3;	-42	-42	-22	-22	-42	-42	-22	-22	0	0	-42	-42	0	0	-42	-42
1-D1;2-D1;3-D3;4-D3;	-42	-42	-22	-22	-42	-42	-22	-22	-42	0	-42	0	-42	0	-42	0
1-D1;2-D1;3-S3;4-S3;	-21	-42	-21	-42	-21	-42	-21	-42	0	0	-42	-42	0	0	-42	-42
1-D1;2-D1;3-S3;4-D3;	-21	-42	-21	-42	-21	-42	-21	-42	-42	0	-42	0	-42	0	-42	0

Figura 77.

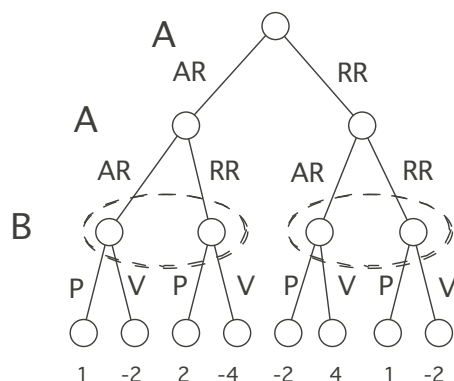


Figura 78.

al livello superiore (uno per ogni nodo, A conosce la situazione e quindi i valori sono esattamente legati ai rispettivi nodi). La scelta di A è di deporre sul tavolo asso-re che dà luogo ad una perdita minore (2 contro 4). L'esito -2 viene quindi riportato al livello superiore.

Nell'insieme d'informazione di destra la situazione è diversa. Se B passa può guadagnare 2 o perdere 1, se invece vede può perdere 4 o guadagnare 2. Nessuna scelta domina l'altra. Possiamo pensare che B adotti una strategia prudente di minimizzare il rischio e quindi sceglie dove la perdita è minore, cioè passa. Allora i due esiti -2 e 1 vengono portati al livello superiore e A sceglie di deporre re-re che dà luogo al guadagno di 1 contro la perdita di 2.

Infine in fase di dichiarazione A dichiara re-re perché comporta la vittoria di 1 contro la perdita di 2. Tuttavia la situazione non è di equilibrio. Una volta che B realizza che la strategia migliore per A è di dichiarare re-re e di deporre asso-re, può decidere di vedere nell'insieme d'informazione di destra ottenendo così un guadagno di 2. Ma a questo punto A potrebbe anticipare questa scelta di B e deporre invece asso-re guadagnando 4. Continuando a ragionare così si oscilla fra le varie scelte senza mai arrivare ad una scelta definitiva da parte dei giocatori.

Questo succede perché effettivamente il gioco non presenta una soluzione con strategie pure. Per valutare meglio il gioco è più conveniente riscriverlo in forma normale. A ha a disposizione 4 strategie a seconda delle combinazioni fra dichiarazione e disposizione effettiva delle carte. B ha 4 strategie perché può decidere di passare o vedere se si trova nel primo insieme d'informazione (corrispondente alla dichiarazione "asso-re") e analogamente nel secondo. Quindi il gioco in forma normale è dato dalla seguente tabellina 4×4 :

	P-P	P-V	V-P	V-V
AR-AR	1	1	-2	-2
AR-RR	2	2	-4	-4
RR-AR	-2	4	-2	4
RR-RR	1	-2	1	-2

Risolvendo il gioco si trova $x_A = (0, 0, 1/3, 2/3)$ e $y_B = (1/3, 1/3, 1/3, 0)$. Si noti che A, in ottimalità, non dovrebbe mai dichiarare “Asso-re”, come risulta anche evidente dalla precedente discussione sul gioco in forma estesa. Nel primo insieme d’informazione (corrispondente alla dichiarazione “Asso-re”) B possiede una strategia vincente pur non conoscendo le carte sul tavolo e quindi A non ha interesse a scegliere questa parte del gioco. A quindi dichiarerà sempre “Due re” (con il che il gioco diminuisce d’interesse ovviamente) e due volte su tre disporrà le carte in modo veritiero rispetto alla dichiarazione. Data questa strategia di A, la strategia di B si può sintetizzare in “passare due volte su tre”. Il valore atteso del gioco è zero, quindi il gioco è onesto.

58 Giochi non cooperativi a somma non zero a due persone

Anche se il gioco in questione non è strettamente competitivo, cioè non è detto che la vincita di un giocatore corrisponda alla perdita dell’altro, supponiamo che ai giocatori non sia concesso né di scambiarsi informazioni preventive sulle proprie strategie né tanto meno di cooperare. Il concetto di soluzione del gioco, come sviluppato da Nash, si basa sull’idea di equilibrio fra le strategie.

Consideriamo il seguente esempio dove nella stessa tabella sono indicati sia u_1 che u_2 (adesso ogni giocatore cerca di massimizzare la propria funzione):

	B_R	B_A
A_R	(6,10)	(0,0)
A_A	(4,4)	(10,6)

Questo gioco prende il nome, un po’ esagerato, di *battaglia dei sessi*, perché si può interpretare in questo modo: il giocatore A è un ragazzo e B è una ragazza. Entrambi vogliono andare al cinema e devono scegliere se andare a vedere un film romantico o un film d’azione. Indulgendo nei più scontati stereotipi supponiamo che il ragazzo preferisca vedere il film d’azione e la ragazza quello romantico. Entrambi preferiscono andare al cinema assieme piuttosto che andare da soli in sale separate. Le strategie per i giocatori sono quindi due: scegliere il film romantico (suffisso R) o quello d’azione (suffisso A). Se entrambi vanno a vedere il film romantico la ragazza ha il massimo esito, cioè 10, mentre il ragazzo ha un esito di 6 perché non è proprio contento del film. L’alto punteggio di entrambi è dovuto al piacere di vedere il film assieme. La situazione si ribalta se entrambi vanno a vedere il film d’azione. Gli altri due casi contemplanò il fatto che entrambi

vanno a vedere i film separatamente. Se vedono il film gradito, l'esito è 4 per tutti e due. Se invece vanno a vedere separatamente il film sgradito l'esito è 0.

Esaminiamo la coppia di strategie (A_R, B_R) . Presa la decisione di andare a vedere il film romantico, nessuno dei due ha interesse a cambiare idea da solo, perché la nuova situazione sarebbe peggiore. La stessa considerazione si applica alla coppia di strategie (A_A, B_A) .

Quindi si tratta di due soluzioni di equilibrio. Esaminiamo ora il gioco dal punto di vista delle strategie di salvaguardia. Ogni giocatore deve valutare quale sia il minimo esito che è in grado di ottenere in ogni caso, indipendentemente da cosa faccia l'altro giocatore. Il giocatore (supponiamo si tratti del giocatore A, con esiti u_1) deve allora assumere di stare giocando un gioco a somma zero in cui l'avversario B, invece di valutare le proprie azioni in base ai suoi esiti u_2 , cerchi di penalizzare al massimo A, cioè come se avesse degli esiti $-u_1$.

Anche se questo atteggiamento può sembrare troppo pessimista (in fin dei conti perché B dovrebbe essere più interessato a penalizzare A che a favorire se stesso?) pur tuttavia, se vogliamo valutare la strategia di salvaguardia, è l'unico atteggiamento corretto. Il calcolo della strategia di salvaguardia porta ad una strategia mista con probabilità $(1/2, 1/2)$ che garantisce un esito di 5. Infatti se B sceglie B_R si ottiene la media fra 6 e 4, cioè 5, e se B sceglie B_A si ottiene la media fra 0 e 10, cioè ancora 5. Data la simmetria del gioco, anche B ha una strategia di salvaguardia $(1/2, 1/2)$. Se entrambi adottano le loro strategie di salvaguardia il risultato è di avere ogni possibile esito della tabella con probabilità $1/4$. La diffidenza reciproca non ha portato ad un grande risultato, dominato dalle due strategie di equilibrio.

Un altro celebre (forse il più celebre paradigma di gioco non cooperativo) è il cosiddetto *dilemma del prigioniero* la cui tabellina è (i valori numerici della tabella possono variare, ciò che conta è la relazione d'ordine fra i valori)

	B_N	B_C
A_N	(-1,-1)	(-8,0)
A_C	(0,-8)	(-5,-5)

L'interpretazione del gioco, da cui deriva il nome, è la seguente: due criminali sospettati di aver commesso assieme un reato sono interrogati separatamente. Potrebbero entrambi negare (suffisso N) le circostanze e quindi, senza prove evidenti, la polizia potrebbe solo incriminarli per un reato minore, quale ad esempio il possesso indebito di armi, e quindi essere condannati ad un anno di carcere, con un esito $(-1, -1)$. La polizia però promette però ad ognuno di loro un beneficio di pena che porterebbe alla scarcerazione, se forniscono le evidenze probatorie che accuserebbero l'altro (suffisso C , confessare). Allora, se uno confessa e l'altro no, l'esito è di zero anni di carcere per chi confessa e di 8 anni invece per l'altro. Potrebbero però entrambi cedere alle promesse della polizia, il che porterebbe a 5 anni di carcere per entrambi, che corrisponde agli 8 anni, scontati per la collaborazione con la giustizia.

Questo paradigma si applica a molti casi di vita reale e quindi è di grandissima importanza. Ad esempio la corsa agli armamenti è un caso di dilemma del prigioniero. Ogni nazione può decidere di armarsi o di non armarsi (o quanto meno più concretamente di armarsi molto o poco). Se due nazioni non si armano hanno un esito (0,0) nel senso che non spendono in armamenti. Se una si arma e l'altra no, c'è la concreta possibilità che la nazione armata imponga la propria politica all'altra (se non addirittura la conquista) a causa della superiorità militare. Quindi l'esito è di 2 per la nazione armata (vantaggio politico di 4 meno la spesa per gli armamenti di 2) e di -4 per la nazione che non si è armata (indipendenza persa). Se entrambi si armano rimangono indipendenti però con un'elevata spesa militare, con esito allora (-2,-2). La tabellina è allora

	B_N	B_C
A_N	(0,0)	(-4,2)
A_C	(2,-4)	(-2,-2)

I valori sono diversi dalla tabella precedente ma la relazione d'ordine è la medesima e quindi si tratta dello stesso gioco. Un altro esempio potrebbe riguardare l'osservanza delle leggi. In questo caso un giocatore è un singolo cittadino e l'altro giocatore è la totalità degli altri cittadini. Ogni cittadino può rispettare le leggi oppure no. Non rispettare le leggi quando tutti gli altri lo fanno è certamente vantaggioso (tralasciando l'aspetto dell'eventuale punizione) per chi non le rispetta, ma se tutti si sottraggono al rispetto delle leggi si va incontro al caos e non c'è vantaggio per nessuno.

Analizziamo allora il dilemma del prigioniero e vediamo se il gioco ha una soluzione nel senso di Nash. Quella che sembrerebbe la strategia più desiderabile, cioè (A_N, B_N) , non è purtroppo stabile. I due giocatori hanno entrambi interesse a deviare dalla strategia N , il che li porta alla strategia (A_C, B_C) che risulta dominata da (A_N, B_N) , ma purtroppo è stabile. Si noti che la strategia (A_C, B_C) è anche di salvaguardia. Quindi è molto naturale che ogni situazione modellabile come il dilemma del prigioniero evolva naturalmente verso la strategia dominata (A_C, B_C) .

Data questa natura 'indesiderabile' della soluzione del gioco, sono stati fatti molti studi per capire come modificare il gioco per avere una soluzione che sia invece (A_N, B_N) . Ad esempio nel caso del rispetto delle leggi, la presenza di una pena (possibilmente certa) per il non rispetto, altera i valori della tabellina e rende stabile (A_N, B_N) . Nel caso della corsa agli armamenti non esiste una legge sovranazionale che possa imporre pene ai trasgressori di un accordo internazionale e quindi non è possibile modificare gli esiti del gioco in questo senso. È chiaro che un accordo di disarmo funziona finché c'è fiducia reciproca che l'avversario non abbandonerà unilateralmente l'accordo.

Si può ipotizzare che la fiducia nasce dall'esperienza e dalla conoscenza reciproca dei giocatori. Ad esempio ci si può chiedere cosa succede nel dilemma del prigioniero se il gioco viene iterato. Sembrerebbe che la consapevolezza di finire nell'esito $(u_1(A_C, B_C), u_2(A_C, B_C))$ porti i giocatori a mantenere le strategie (A_N, B_N) . Ma se è noto il numero di iterazioni, alla fine c'è inevitabile la defezione di entrambi i giocatori. Sono state proposte molte varianti del dilemma del prigioniero per ottenere una soluzione del gioco più

consona al desiderio di far coincidere l'interesse personale con quello generale, più o meno convincenti.

L'aspetto importante che emerge dal Dilemma del Prigioniero è che l'azione libera delle persone non produce necessariamente il bene di tutti. In altre parole la cosiddetta 'mano invisibile' che, secondo Adam Smith, produce il bene comune come somma delle libere azioni individuali che perseguono indipendentemente i propri obiettivi, può in realtà fallire. Purtroppo gli attori del Dilemma del Prigioniero sono due criminali e il mancato raggiungimento della soluzione migliore viene invece visto dal di fuori come 'cosa buona' perché porta all'incriminazione di due criminali. Forse per questo motivo la lezione negativa del Dilemma del Prigioniero non è ancora entrata profondamente nel 'sentire comune'.

Un altro paradigma, molto simile al dilemma del prigioniero, detto *dilemma del viaggiatore*, prevede il seguente scenario: il bagaglio di due viaggiatori è stato rovinato durante il viaggio e la compagnia aerea deve rifondere i viaggiatori. Si suppone che il bagaglio sia identico con il medesimo valore. La compagnia aerea prima di rifondere i viaggiatori (e sapendo che i bagagli erano uguali) chiede singolarmente ai viaggiatori una valutazione del valore, concedendo un rimborso massimo di 100 euro. Il rimborso è un valore intero compreso fra 2 e 100. Se i due viaggiatori dichiarano lo stesso valore, tale è anche il rimborso che ricevono. Se però dichiarano valori diversi, la compagnia prende per veritiera la dichiarazione inferiore e tale è il rimborso che entrambi ricevono, però, se non hanno dichiarato lo stesso valore, chi ha dichiarato di meno riceve un bonus di 2 euro che viene prelevato dal rimborso dell'altro. Quindi, ad esempio, se un viaggiatore dichiara 90 e l'altro dichiara 95, il primo riceve 92 e il secondo 88 e, se uno dichiara 2 e l'altro 3, il primo riceve 4 e il secondo 0.

Sembrerebbe ovvio che i viaggiatori dichiarino 100, però un viaggiatore potrebbe essere tentato di dichiarare 99 in modo da ottenere 101, grazie al bonus. Essendo la situazione simmetrica i viaggiatori ragionano allo stesso modo e quindi sono portati a dichiarare 99, ma poi, per sfruttare il bonus dichiarano 98 e così via fino a dichiarare alla fine 2 euro!

La tabella degli esiti è, per il dilemma del viaggiatore,

$$u_1(x, y) = \begin{cases} x + 2 & \text{se } x < y \\ x & \text{se } x = y \\ y - 2 & \text{se } x > y \end{cases} \quad u_2(x, y) = \begin{cases} x - 2 & \text{se } x < y \\ y & \text{se } x = y \\ y + 2 & \text{se } x > y \end{cases}$$

ed è facile da vedere che solo la strategia (2, 2) è un equilibrio di Nash. Rispetto al dilemma del prigioniero non si pone la possibilità di iterare il gioco e quindi l'aspetto di 'razionalità' della soluzione sembra molto problematico. Mentre l'equilibrio di Nash del dilemma del prigioniero corrisponde a quanto può accadere in realtà, è molto inverosimile che due giocatori reali del dilemma del viaggiatore dichiarino 2. In letteratura si possono trovare molte discussioni su questo aspetto. Una conclusione che si potrebbe trarre è che, in certe circostanze, diventa più utile cooperare.

Il gioco è del tutto simile al *gioco del millepiedi* esemplificato in forma estesa in Fig. 79 per un massimo rimborso di 10 (ma si può ovviamente estendere a qualsiasi valore). Il calcolo induttivo all'indietro per determinare la strategia di equilibrio porta sia il giocatore A che il giocatore B a scegliere la mossa corrispondente al tratto verticale e quindi la strategia finale porta al guadagno (2, 2).

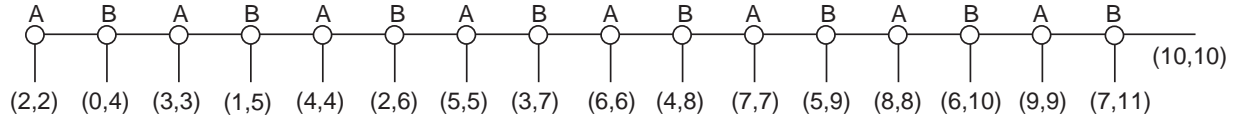


Figura 79.

Non è detto che un gioco ammetta una soluzione in termini di strategie pure. Nash ha dimostrato che una soluzione usando strategie miste invece esiste sempre. Non è detto che sia unica. Possono esistere diverse soluzioni non equivalenti (si veda la battaglia dei sessi per esempio).

Per definizione allora una soluzione di equilibrio è data da due distribuzioni di probabilità (p_1^*, \dots, p_n^*) e (q_1^*, \dots, q_n^*) tali che

$$\sum_{i \in S_1} \sum_{j \in S_2} p_i^* q_j^* u_1(i, j) \geq \sum_{i \in S_1} \sum_{j \in S_2} p_i q_j^* u_1(i, j) \quad \forall p$$

e

$$\sum_{i \in S_1} \sum_{j \in S_2} p_i^* q_j^* u_2(i, j) \geq \sum_{i \in S_1} \sum_{j \in S_2} p_i^* q_j u_2(i, j) \quad \forall q$$

Le due relazioni si possono scrivere come

$$\sum_{i \in S_1} (p_i^* - p_i) \sum_{j \in S_2} q_j^* u_1(i, j) \geq 0 \quad \forall p \quad (133)$$

e

$$\sum_{j \in S_2} (q_j^* - q_j) \sum_{i \in S_1} p_i^* u_2(i, j) \geq 0 \quad \forall q \quad (134)$$

Siano h e k tali che $p_h^* > 0$ e $p_k^* > 0$. Se non esistono allora p^* è una strategia pura. Allora si prenda in (133) in particolare

$$p_i = \begin{cases} p_i^* & \text{se } i \neq h, i \neq k \\ p_h^* - \varepsilon & \text{se } i = h \\ p_k^* + \varepsilon & \text{se } i = k \end{cases} \quad (135)$$

si ha

$$\varepsilon \left(\sum_{j \in S_2} q_j^* u_1(h, j) - \sum_{j \in S_2} q_j^* u_1(k, j) \right) \geq 0$$

Siccome ε può essere sia positivo che negativo (dato che $p_h^* > 0$ e $p_k^* > 0$), deve essere

$$\sum_{j \in S_2} q_j^* u_1(h, j) = \sum_{j \in S_2} q_j^* u_1(k, j)$$

Allora deve esistere un valore v_1^* tale che

$$\sum_{j \in S_2} q_j^* u_1(h, j) = v_1^* \quad \text{se} \quad p_h^* > 0$$

Siano ora h e k tali che $p_h^* > 0$ e $p_k^* = 0$. Allora si prenda p_i come in (135) con la differenza che $\varepsilon > 0$ e si ha

$$\varepsilon \left(\sum_{j \in S_2} q_j^* u_1(h, j) - \sum_{j \in S_2} q_j^* u_1(k, j) \right) = \varepsilon (v_1^* - \sum_{j \in S_2} q_j^* u_1(k, j)) \geq 0$$

Siccome $\varepsilon > 0$ deve essere

$$v_1^* \geq \sum_{j \in S_2} q_j^* u_1(k, j) \quad \text{se} \quad p_k^* = 0$$

Si può fare un calcolo analogo per q^* e quindi si perviene alle relazioni

$$v_1^* \geq \sum_{j \in S_2} q_j^* u_1(h, j), \quad v_2^* \geq \sum_{i \in S_1} p_i^* u_2(i, k) \quad (136)$$

e inoltre

$$\begin{aligned} p_h^* > 0 &\implies v_1^* = \sum_{j \in S_2} q_j^* u_1(h, j), & v_1^* > \sum_{j \in S_2} q_j^* u_1(h, j) &\implies p_h^* = 0 \\ q_k^* > 0 &\implies v_2^* = \sum_{i \in S_1} p_i^* u_2(i, k), & v_2^* > \sum_{i \in S_1} p_i^* u_2(i, k) &\implies q_k^* = 0 \end{aligned} \quad (137)$$

Le relazioni (137) vengono dette di *complementarità lineare*. Se $u_1 = -u_2$ (cioè il gioco è a somma zero) sono esattamente le relazioni di complementarità della programmazione lineare. Il calcolo di p^* , q^* e v_1^* , v_2^* che soddisfano (136) e (137) non è del tutto banale in generale e richiede algoritmi ad hoc.

Riguardo alla battaglia dei sessi, oltre ai due equilibri con strategie pure, esiste un terzo equilibrio con strategie miste, dato da

$$p^* = \left(\frac{1}{6} \quad \frac{5}{6} \right), \quad q^* = \left(\frac{5}{6} \quad \frac{1}{6} \right)$$

e con esito (5, 5) (come con la strategia di salvaguardia, ma con differenti strategie miste). In Fig. 80-(a) si vede l'insieme ottenibile con strategie miste e i tre equilibri. Si noti che il terzo equilibrio è dominato dagli altri due.

59 Giochi cooperativi a somma non zero a due persone

Supponiamo ora che i giocatori possano cooperare. A questo scopo si possono fare diverse ipotesi sul funzionamento del gioco. A livello minimo i giocatori possono scambiarsi informazioni e comunicare all'altro le proprie intenzioni. La parola 'cooperazione' non deve però trarre in inganno. Anche se la possibilità di agire congiuntamente può arrecare vantaggio ad entrambi i giocatori, può anche avvenire che l'informazione sia in realtà una minaccia e quindi il giocatore che si trovi in condizioni di minacciare (con credibilità) può portare il gioco ad un esito che in realtà è di vantaggio per uno solo dei due giocatori. Vedremo più avanti un esempio di minaccia e delle sue conseguenze.

Se gli esiti del gioco rappresentano utilità per i giocatori, queste non sono confrontabili e quindi non ci possono essere scambi di utilità fra i giocatori. Ad un maggior livello di cooperazione le utilità sono

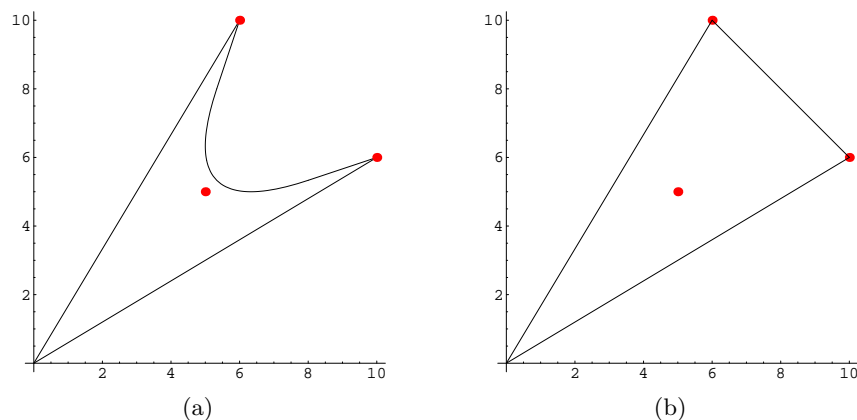


Figura 80.

confrontabili, ma non sono permessi scambi di utilità. Infine si può supporre che gli esiti del gioco siano monetari e come tali siano confrontabili. In più ci possono essere scambi monetari fra i giocatori.

Il semplice fatto di potersi mettere d'accordo e quindi di adottare congiuntamente strategie miste, aumenta gli esiti ottenibili. Ogni punto all'interno dell'involuppo convesso di tutti i punti $\bigcup_{x,y} (u_1(x,y), u_2(x,y))$ è ottenibile con un'opportuna strategia mista *congiunta*. Ad esempio per la battaglia dei sessi, l'insieme ottenibile diventa quello in Fig. 80-(b).

Il segmento che congiunge i punti $(10, 6)$ e $(6, 10)$ non sarebbe mai ottenibile con strategie miste indipendenti da parte dei giocatori. È invece ottenibile con una strategia mista congiunta in cui con probabilità p entrambi scelgono le strategie (A_R, B_R) e con probabilità $(1 - p)$ le strategie (A_A, B_A) .

Potendo cooperare i due giocatori necessariamente cercheranno una soluzione che non sia dominata. Nell'esempio le uniche strategie non dominate sono quelle che portano sul segmento congiungente $(10, 6)$ e $(6, 10)$.

Si consideri il seguente gioco

	B_1	B_2
A_1	$(2,3)$	$(-9,1)$
A_2	$(3,2)$	$(-9,1)$

i cui esiti ottenibili tramite strategie miste congiunte sono contenute nel triangolo di Fig. 81-(a)

Gli esiti su cui i giocatori possono mettersi d'accordo sono i punti del segmento congiungente $(3, 2)$ e $(2, 3)$. Su quale punto si metteranno d'accordo? Se il gioco fosse non cooperativo l'equilibrio di Nash è dato da (A_2, B_1) (punto verde in Fig. 81-(a)). Il fatto che i giocatori possano parlarsi introduce nuovi fattori nel gioco. Il giocatore B si trova in maggiore posizione di forza rispetto a A perché potrebbe minacciare di

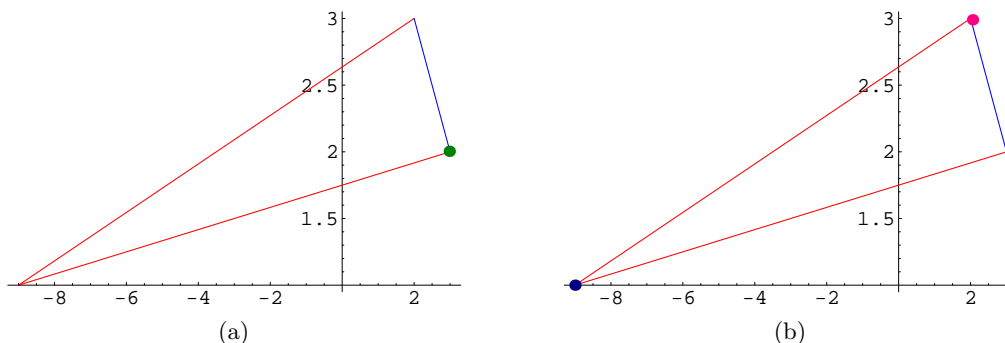


Figura 81.

adottare la strategia B_2 . È vero che ci rimette, ma molto meno di A, per il quale la scelta di B_2 da parte di B si configura quasi come un ricatto. Quindi B potrebbe obbligare A ad usare la strategia A_1 sotto la minaccia di usare la strategia B_2 .

Come formalizzare questa situazione? Il notevole contributo di Nash è stato quello di trovare una formalizzazione razionale al processo di negoziazione basandosi solo sugli aspetti strategici del gioco e non su quelli psicologici o etici. Il metodo adottato da Nash consiste nella formulazione di alcuni requisiti che uno schema di arbitrato fra i giocatori dovrebbe possedere e nel trovare uno schema che soddisfa i requisiti. In particolare Nash ha dimostrato che tale schema è unico.

I requisiti per uno schema di arbitrato, in cui il giocatore i riceve \hat{u}_i , sono:

- 1) Esistono due valori u_1^* e u_2^* che vengono definiti lo *status quo*, che corrispondono a quanto ogni giocatore avrebbe se agisse senza cooperare. Il punto (u_1^*, u_2^*) viene anche detto *punto di minaccia*.
- 2) la soluzione dovrebbe dare ad ogni giocatore almeno tanto quanto avrebbe se giocasse in condizioni di non cooperazione; quindi $\hat{u}_i \geq u_i^*$;
- 3) la soluzione deve essere non dominata, cioè non deve esistere una soluzione u_1, u_2 tale che $u_1 \geq \hat{u}_1$, $u_2 \geq \hat{u}_2$ con almeno una disuguaglianza stretta;
- 4) la soluzione non dovrebbe dipendere dall'unità di misura usata per le utilità dei giocatori (cioè se si moltiplicano per una costante gli esiti di un giocatore la soluzione dovrebbe essere la stessa; si noti che in figura le scale dei due assi sono differenti);
- 5) lo schema dovrebbe essere indipendente da alternative irrilevanti, ovvero, se il gioco viene cambiato togliendo alcuni esiti, ma lo status quo risulta inalterato e (\hat{u}_i, \hat{u}_2) è ancora ammissibile, allora (\hat{u}_i, \hat{u}_2) è soluzione anche del nuovo gioco.

Nash ha dimostrato che la soluzione (\hat{u}_1, \hat{u}_2) è quella che massimizza

$$(u_1 - u_1^*)(u_2 - u_2^*)$$

con (u_1, u_2) all'interno dell'involuppo convesso di tutti i punti $(u_1(x), u_2(y))$, $x \in S_1$, $y \in S_2$.

Prima di iniziare la negoziazione i giocatori dichiarano quale strategia adotterebbero se un accordo non venisse raggiunto. È naturale pensare che in assenza di accordo ogni giocatore può garantirsi solo il livello di salvaguardia. Quindi se x^* e y^* sono le due strategie di salvaguardia lo status quo potrebbe essere il punto

$$(u_1^*, u_2^*) = (\min_y u_1(x^*, y), \min_x u_2(x, y^*)) = (\max_x \min_y u_1(x, y), \max_y \min_x u_2(x, y))$$

che sembra più corretto del punto

$$(u_1^*, u_2^*) = (u_1(x^*, y^*), u_2(x^*, y^*))$$

dove l'esito u_i^* dipende anche dalla scelta dell'altro giocatore. Ma in qualche caso potrebbe essere più vantaggioso usare altre strategie.

Nell'esempio della battaglia dei sessi lo status quo, definito dalle strategie di salvaguardia è il punto $(u_1^*, u_2^*) = (5, 5)$. Bisogna pertanto massimizzare $(u_1 - 5)(u_2 - 5)$. Data la simmetria del gioco è evidente che la soluzione è il punto di mezzo del segmento cioè $(\hat{u}_1, \hat{u}_2) = (8, 8)$.

Più interessante è valutare lo schema di arbitrato per il precedente esempio con il giocatore in posizione di forza perché può esercitare delle minacce, che qui ripetiamo per comodità

	B_1	B_2
A_1	(2,3)	(-9,1)
A_2	(3,2)	(-9,1)

Se si considerano soltanto le strategie di salvaguardia, A sceglie A_2 : siccome per una scelta di B_2 da parte di B, c'è comunque una perdita di 9, allora è meglio optare per A_2 che garantisce un miglior risultato se B dovesse scegliere B_1 . In ogni caso si ha il valore di salvaguardia $u_1^* = -9$. B invece ha la strategia di salvaguardia B_1 con valore $u_2^* = 2$. Lo status quo quindi sarebbe $(-9, 2)$. Si noti che se si valutasse lo status quo come $(u_1(A_2, B_1), u_2(A_2, B_1)) = (3, 2)$, tale scelta non rivela adeguatamente la forza di B.

Infatti B potrebbe addirittura dichiarare di usare la strategia B_2 in caso di non accordo. Tutto sommato guadagnerebbe solo 1, ma non sarebbe una grave perdita rispetto ad un guadagno di 3 o di 2. Con una tale dichiarazione di B, lo status quo diventa $(u_1(A_2, B_2), u_2(A_2, B_1)) = (-9, 1)$ e quindi $(\hat{u}_1, \hat{u}_2) = (2, 3)$ (Fig. 81-(b)).

Ci si può chiedere se, data l'arbitrarietà della scelta del punto di status quo, i giocatori non trovino un modo di determinare la scelta dello status quo proprio per massimizzare il punto di arbitrato. Si consideri il seguente gioco

	B_1	B_2
A_1	(0,0)	(10,0)
A_2	(0,100)	(0,0)

che corrisponde ad una situazione in cui i due giocatori devono accordarsi su un valore, che supponiamo monetario, sul segmento $(0, 100) - (10, 0)$. Se non trovano un accordo non prendono nulla. La strategia (A_1, B_1) è un equilibrio di Nash ed è appunto la soluzione in caso di fallimento dell'accordo (l'altra strategia (A_2, B_2) non è di equilibrio).

Sia (\hat{u}_1, \hat{u}_2) il punto (per il momento non noto) di arbitrato da calcolare in base al punto (anch'esso non noto) di status quo. In base agli assiomi deve essere il punto che massimizza $(u - u^*)(v - v^*)$ sul segmento $(0, 100) - (10, 0)$. Sul segmento si ha $v = 100 - 10u$ e quindi bisogna massimizzare

$$(u - u^*)(100 - 10u - v^*)$$

Derivando e eguagliando a 0, si ha

$$(100 - 10u - v^*) - 10(u - u^*) = 0 \implies \\ \hat{u} = \frac{100 - v^* + 10u^*}{20}, \quad \hat{v} = 100 - 10 \frac{100 - v^* + 10u^*}{20} = \frac{100 + v^* - 10u^*}{2}$$

Allora, nella scelta di u^* il primo giocatore è interessato a massimizzare $10u^* - v^*$ mentre il secondo giocatore è interessato a *minimizzare* $10u^* - v^*$. Quindi è come se i giocatori si trovassero di fronte ad un gioco a somma zero con esiti dati da

$$10u(x, y) - v(x, y)$$

ovvero con tabella

	B_1	B_2
A_1	0	100
A_2	-100	0

La strategia (A_1, B_1) è evidentemente la soluzione di questo gioco a somma zero e quindi lo status quo è $(0, 0)$ e il punto di arbitrato è $(\hat{u}, \hat{v}) = (5, 50)$. Questo approccio di calcolare lo status quo da un gioco a somma zero si può estendere a qualsiasi gioco cooperativo. Questo verrà sviluppato dopo la successiva estensione al caso di trasferimenti di utilità.

L'ipotesi di utilità lineare ha prodotto un risultato abbastanza plausibile. Si immagini ora invece che le utilità non siano lineari e i due giocatori siano molto diversi, ad esempio A povero e B ricco e che le funzioni di utilità riflettano questo loro stato. Modifichiamo il gioco in modo che la somma da dividere sia 100 e che in caso di mancato accordo non ricevono niente. Il giocatore A che è povero ha la funzione di utilità $u_1(x)$ in Figura 82 (x è normalizzato da 0 a 1) e B ha la funzione $u_2(x)$. Le funzioni sono

$$u_1(x) = -\frac{3}{4}x^2 + \frac{7}{4}x, \quad u_2(x) = \frac{3}{4}x^2 + \frac{1}{4}x$$

Il massimo del prodotto $u_1(1-x)u_2(x)$ si ottiene (si veda la Figura 82) per $x = 2/3$. Certamente il punto $(0, 0)$ è lo status quo e quindi la soluzione che riflette la diversa posizione di forza dei due giocatori si vede

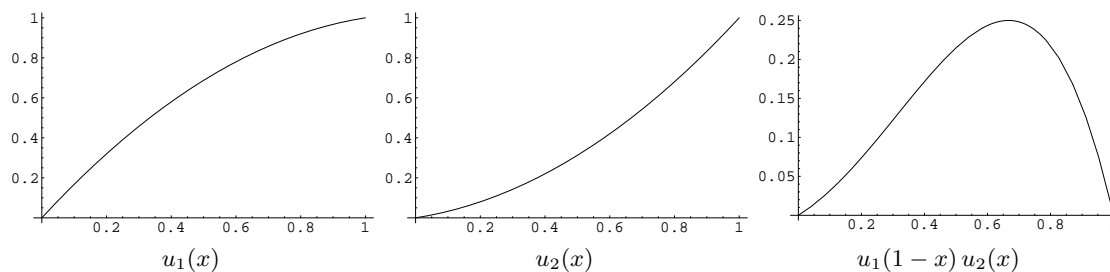


Figura 82.

nella scelta del punto di arbitrato. Il risultato ‘antipatico’, ma coerente con l’aspetto strategico, è che il ricco riceve più soldi. A pensarci bene non è affatto un risultato insolito.

Se sono ammessi dei pagamenti fra i giocatori e quindi stiamo assumendo che le utilità in realtà sono misure monetarie e i giocatori hanno utilità lineari, allora l’insieme dei possibili esiti si estende. Infatti, se i giocatori hanno interesse ad accordarsi, devono semplicemente decidere quale sia l’esito del gioco che complessivamente dà il più grande guadagno. I due giocatori si mettono d’accordo di adottare esattamente quelle strategie. Poi devono decidere come dividere fra loro l’ammontare globale. Dobbiamo immaginare che i giocatori non possano esimersi dal giocare. Se vogliamo dare alla possibilità ai giocatori di uscire dal gioco, possiamo sempre immaginare che esista una strategia di uscita dal gioco con valore $(0,0)$ per tutti i casi in cui almeno un giocatore esce.

Dato che i giocatori sono costretti a giocare e potrebbero anche perdere soldi, dobbiamo poter prevedere il caso in cui se il massimo esito del gioco è (a_k, b_h) (e quindi i giocatori devono dividersi la quantità $a_k + b_h$) il giocatore B, per esempio, si trovi in condizione di pagare ad A più di b_h . Sia allora

$$\sigma = \max_{x \in S_1} \max_{y \in S_2} u_1(x, y) + u_2(x, y)$$

la quantità che i giocatori devono dividersi, sempreché trovino un accordo. La soluzione sarà quindi un punto (\hat{u}_1, \hat{u}_2) tale che $\hat{u}_1 + \hat{u}_2 = \sigma$. Come determinare tale punto? La teoria dell’arbitrato permette tale calcolo noto lo status quo. In questo caso però i giocatori possono determinare il più conveniente status quo proprio sfruttando il fatto che la somma σ va divisa. Sia (u_1^*, u_2^*) lo status quo da determinare. Allora il punto che minimizza $(u_1 - u_1^*)(u_2 - u_2^*)$ sulla retta $u_1 + u_2 = \sigma$ è il punto

$$\frac{1}{2} (\sigma + u_1^* - u_2^*, \sigma - u_1^* + u_2^*) \quad (138)$$

Il primo giocatore è interessato a massimizzare $\sigma + u_1^* - u_2^*$ e il secondo a massimizzare $\sigma - u_1^* + u_2^*$, ovvero il primo giocatore è interessato a massimizzare $u_1^* - u_2^*$ e il secondo a minimizzare la stessa quantità. Quindi sceglieranno rispettivamente una strategia x^* e y^* tale che $u_1(x^*, y^*) - u_2(x^*, y^*)$ è massimizzato per il primo giocatore e minimizzato per il secondo giocatore. A ben vedere si tratta di un gioco a somma zero per una funzione di utilità $\tilde{u}(x, y) = u_1(x, y) - u_2(x, y)$. Quindi si tratta semplicemente di calcolare la soluzione di

questo gioco a somma zero. Sia δ il valore della soluzione. La soluzione di arbitrato è allora

$$\frac{1}{2}(\sigma + \delta, \sigma - \delta)$$

Si immagini il seguente gioco

	B_1	B_2
A_1	(2,2)	(-9,1)

Il giocatore A ha solo una strategia e quindi si trova in balia dell'altro giocatore, che potrebbe infliggergli la perdita di 9. Entrambi si accordano sulla strategia (A_1, B_1) (o meglio B concede ad A di adottare la strategia B_1). Però nel calcolo dello status quo si ha il gioco a somma zero

	B_1	B_2
A_1	0	-10

il cui valore è ovviamente $\delta = -10$. Quindi i due giocatori si dividono la somma di 4 nel seguente modo

$$\frac{1}{2}(\sigma + \delta, \sigma - \delta) = \frac{1}{2}(4 - 10, 4 + 10) = (-3, 7)$$

cioè B si prende tutta la somma di 4 e in più si fa pagare 3 da A. Suona ingiusto? Certamente non sembra etico, ma la teoria dei giochi cerca di modellare il comportamento razionale e non quello etico e nemmeno quello psicologico. A potrebbe rifiutarsi di dividere il guadagno, ma allora B adotterebbe la strategia B_2 , accontentandosi del guadagno di 1, contro i 7 della divisione e A dovrebbe perdere 9, contro i 3 della divisione. Entrambi guadagnano 6 dalla divisione e quindi conviene ad entrambi accordarsi. Il fatto che entrambi abbiano lo stesso guadagno non è casuale. Si abbia genericamente il payoff

	B_1	B_2
A_1	(a_1, b_1)	(a_2, b_2)

che porta al gioco a somma zero

	B_1	B_2
A_1	$a_1 - b_1$	$a_2 - b_2$

Supponiamo $a_2 - b_2 < a_1 - b_1$ e $a_2 + b_2 < a_1 + b_1$. Allora $\delta = a_2 - b_2$ e $\sigma = a_1 + b_1$ e

$$\frac{1}{2}(\sigma + \delta, \sigma - \delta) = \frac{1}{2}(a_1 + b_1 + a_2 - b_2, a_1 + b_1 + b_2 - a_2)$$

e

$$\frac{1}{2}(a_1 + b_1 + a_2 - b_2) - a_2 = \frac{1}{2}(a_1 + b_1 - a_2 - b_2)$$

$$\frac{1}{2}(a_1 + b_1 + b_2 - a_2) - b_2 = \frac{1}{2}(a_1 + b_1 - b_2 - a_2)$$

È interessante vedere cosa succede se $\sigma = a_2 + b_2 > a_1 + b_1$ (ad esempio si immagini di avere (-2,8) anziché (-9,1) nell'esempio precedente. Il gioco a somma zero rimane invariato quindi la divisione è

$$\frac{1}{2}(\sigma + \delta, \sigma - \delta) = \frac{1}{2}(a_2 + b_2 + a_2 - b_2, a_2 + b_2 - a_2 + b_2) = (a_2, b_2)$$

Di fatto è come se non ci fosse divisione. Siccome B sceglie B_2 'naturalmente', senza bisogno di interagire con A, la cooperazione non può aggiungere nulla al gioco. In generale, se $\max_{x \in S_1} \max_{y \in S_2} u_1(x, y) + u_2(x, y)$ e $\min_{x \in S_1} \min_{y \in S_2} u_1(x, y) - u_2(x, y)$ si ottengono per la medesima coppia di strategie ed anche il gioco a somma zero ha questa coppia come soluzione, allora i giocatori ottengono esattamente quello che otterrebbero senza cooperazione.

Naturalmente se diamo ad A la possibilità di non partecipare al gioco aggiungendo esplicitamente questa strategia (A_2) le cose cambiano. Adesso $\delta = 0$ e quindi ci si accorda su $(\hat{u}_1, \hat{u}_2) = (2, 2)$.

	B_1	B_2
A_1	(2,2)	(-9,1)
A_2	(0,0)	(0,0)

Il metodo di usare un trasferimento dei pagamenti può essere adattato al caso di trovare un punto d'arbitrato quando non sia permesso scambiare un valore monetario. Tale approccio prende il nome di λ -transfer. L'utilità del primo giocatore sia cambiata in $\lambda u_1(x, y)$ e il gioco diventi uno in cui si possano trasferire delle somme. Sia $\sigma(\lambda) = \max_{x,y} \lambda u_1(x, y) + u_2(x, y)$ e sia $\delta(\lambda)$ il valore del gioco a somma zero

$$\lambda u_1(x, y) - u_2(x, y)$$

Allora il punto

$$\frac{1}{2} \left(\frac{\sigma + \delta(\lambda)}{\lambda}, \sigma - \delta(\lambda) \right) \quad (139)$$

è il punto d'arbitrato se sta sulla frontiera di Pareto dell'involuppo convesso. Bisogna naturalmente trovare, ad esempio con ricerca binaria, il valore di λ per cui questo succede. Applicando questo metodo alla tabella precedente

	B_1	B_2
A_1	(2,3)	(-9,1)
A_2	(3,2)	(-9,1)

si ottiene per il valore $\lambda = 2/11$ la tabella

	B_1	B_2
A_1	$(4/11, 3)$	$(-18/11, 1)$
A_2	$(6/11, 2)$	$(-18/11, 1)$

con valore $\sigma(\lambda) = 37/11$. Il gioco a somma zero

	B_1	B_2
A_1	$-29/11$	$-29/11$
A_2	$-16/11$	$-29/11$

ha come valore $\delta(\lambda) = -29/11$ e quindi da (139)

$$\frac{1}{2} \left(\frac{\sigma + \delta(\lambda)}{\lambda}, \sigma - \delta(\lambda) \right) = \frac{1}{2} \left(\frac{\frac{37}{11} - \frac{29}{11}}{\frac{2}{11}}, \frac{37}{11} + \frac{29}{11} \right) = (2, 3)$$

60 Giochi a molti giocatori

Nei giochi a molti giocatori c'è la possibilità di formare coalizioni. Nell'idea di coalizione è implicita la possibilità di trasferimenti di utilità fra i giocatori. Allora è utile assumere fin dall'inizio che le utilità rappresentino in realtà valori monetari.

Sia $N = \{1, \dots, n\}$ l'insieme dei giocatori. Per ogni coalizione $R \subset N$ di giocatori si definisca una funzione d'utilità della coalizione come

$$u_R(x_1, x_2, \dots, x_n) = \sum_{i \in R} u_i(x_1, x_2, \dots, x_n)$$

Quindi per ogni possibile strategia, i membri della coalizione mettono in comune i loro guadagni. Immaginando un gioco a somma zero in cui da una parte c'è R (con insieme di strategia $\times_{i \in R} S_i$) e dall'altra l'anticoalizione $N \setminus R$ (con insieme di strategia $\times_{i \notin R} S_i$) con esiti del gioco $u_R(x_1, x_2, \dots, x_n)$, la coalizione calcola la soluzione del gioco (a somma zero, quindi sia di salvaguardia che di equilibrio) che viene definita come valore $v(R)$ della coalizione R . La funzione $v(R)$ prende il nome di *funzione caratteristica del gioco*. Si definisce $v(\emptyset) = 0$.

Si noti che, se R e T sono coalizioni disgiunte, deve valere la proprietà di superadditività

$$v(R) + v(T) \leq v(R \cup T)$$

Infatti per la coalizione R esiste una strategia $\times_{i \in R} x_i(R)$ che garantisce almeno $v(R)$, e analogamente per la coalizione T esiste una strategia $\times_{i \in T} x_i(T)$ che garantisce almeno $v(T)$. Essendo le coalizioni disgiunte la

strategia $\times_{i \in R} x_i(R) \times_{i \in T} x_i(T)$ è una strategia realizzabile per $R \cup T$ e siccome

$$u_R(x_1, \dots, x_n) + u_T(x_1, \dots, x_n) = u_{R \cup T}(x_1, \dots, x_n)$$

tale strategia realizza almeno $v(R) + v(T)$.

Spesso conviene riferirsi alla *funzione caratteristica normalizzata* di un gioco. Se ad esempio si moltiplicano tutte le $v(R)$ per una medesima costante c , l'aspetto strategico del gioco non viene certamente alterato. Quindi tutti i giochi le cui funzioni caratteristiche differiscono per una costante moltiplicativa sono da considerarsi equivalenti. Inoltre se il giocatore i viene obbligato a ricevere (o pagare nel caso di segno negativo) preliminarmente la somma a_i , tale quantità altera la funzione caratteristica secondo $v'(R) = v(R) + \sum_{i \in R} a_i$, ma non altera l'aspetto strategico del gioco e quindi anche $v'(R)$ deve essere considerata equivalente a $v(R)$.

Allora possiamo definire *equivalenti* due giochi con funzione caratteristiche v e v' tali che

$$v'(R) = c(v(R) + \sum_{i \in R} a_i)$$

e possiamo cercare, all'interno di una classe di equivalenza, quella funzione che ha la particolare forma

$$v'(\{i\}) = 0, \quad i = 1, \dots, n, \quad v(N) = 1$$

Questa funzione è la funzione caratteristica normalizzata. Per calcolarla da una funzione v basta definire

$$a_i = -v(\{i\}), \quad c = \frac{1}{v(N) + \sum_{i=1}^n a_i}$$

Se ad esempio ci sono tre giocatori con due strategie ciascuno potremmo avere i seguenti valori

C_1	B_1	B_2
A_1	(0,2,-1)	(3,2,0)
A_2	(-2,1,-3)	(0,0,2)

C_2	B_1	B_2
A_1	(3,-2,-1)	(-1,-1,3)
A_2	(0,3,0)	(2,2,1)

Se A e B si coalizzano hanno a disposizione quattro strategie e il loro gioco contro C è

	C_1	C_2
A_1, B_1	2	1
A_1, B_2	5	-2
A_2, B_1	-1	3
A_2, B_2	0	4

La strategia d'equilibrio è la strategia mista $(0, 4/11, 0, 7/11)$ che dà un valore di $20/11$. Quindi $v(\{A, B\}) = 20/11$. Proseguendo nello stesso modo per tutti i sottoinsiemi si trova

$$v(\{A\}) = -\frac{2}{5}; \quad v(\{B\}) = -1; \quad v(\{C\}) = -1;$$

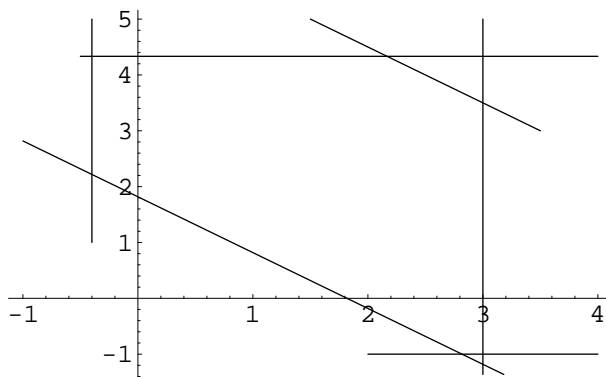


Figura 83.

$$v(\{A, B\}) = \frac{20}{11}; \quad v(\{A, C\}) = 2; \quad v(\{B, C\}) = 2; \quad v(\{A, B, C\}) = 5$$

La funzione normalizzata si trova calcolando

$$a_A = \frac{2}{5}, \quad a_B = 1, \quad a_C = 1, \quad c = \frac{5}{37}$$

da cui

$$\begin{aligned} v'(\{A\}) &= 0; & v'(\{B\}) &= 0; & v'(\{C\}) &= 0; & v'(\{A, B, C\}) &= 1 \\ v'(\{A, B\}) &= \frac{177}{407} = 0.434; & v'(\{A, C\}) &= \frac{17}{37} = 0.459; & v'(\{B, C\}) &= \frac{20}{37} = 0.541; \end{aligned}$$

Assumendo utilità monetarie trasferibili fra i giocatori, la quantità x_i che il giocatore i riceve dalla propria coalizione viene chiamata *imputazione*. Affinché si formi la grande coalizione in cui tutti i giocatori trovano conveniente cooperare, devono esistere imputazioni x_i tali che

$$\sum_{i \in R} x_i \geq v(R), \quad R \subset N, R \neq N; \quad \sum_{i \in N} x_i = v(N) \quad (140)$$

L'uguaglianza indica che le imputazioni provengono dal migliore guadagno del gioco. Le disuguaglianze indicano che nessun sottoinsieme trova vantaggioso staccarsi e formare una coalizione. L'insieme ammissibile di (140) prende il nome di *nucleo del gioco*.

In Fig. 83 si vede il nucleo (in funzione di x_1 e x_2 , dato che $x_3 = 5 - x_1 - x_2$) per l'esempio precedente. Qualsiasi imputazione all'interno del nucleo rende stabile la grande coalizione. Non è detto che il nucleo contenga punti. Nei casi in cui sia vuoto, e sia necessario avere una grande coalizione, un utile strumento è il *valore di Shapley*.

Per valutare il valore di Shapley si prenda in considerazione una particolare permutazione dei giocatori, ad esempio A, B e C. L'idea è quella di dare ad ogni giocatore un valore pari al suo contributo incrementale alla coalizione. Quindi per questa permutazione A riceve $-2/5$, quanto avrebbe se stesse da solo, B riceve la differenza fra $20/11$ (A insieme a B) e $-2/5$ (A da solo), quindi B riceve $122/55$. Infine C riceve $5 - 20/11 = 35/11$.

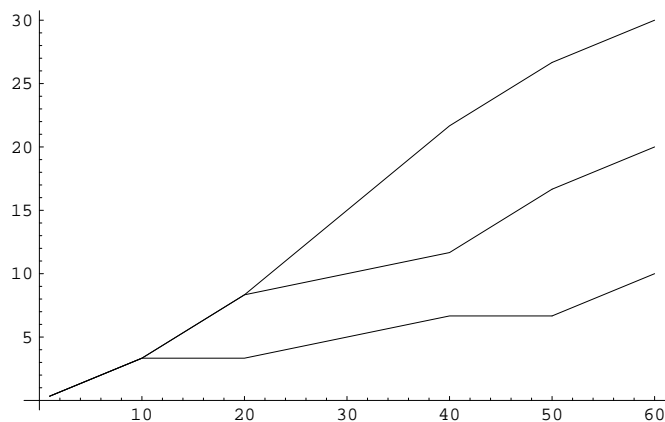


Figura 84.

Per uniformità si prendono in considerazione tutte le permutazioni e si effettua la media

$$\begin{pmatrix} -\frac{2}{5} & \frac{122}{55} & \frac{35}{11} \\ -\frac{2}{5} & \frac{13}{3} & \frac{16}{15} \\ \frac{31}{11} & -1 & \frac{35}{11} \\ 3 & -1 & 3 \\ \frac{13}{6} & \frac{13}{3} & -\frac{3}{2} \\ 3 & \frac{7}{2} & -\frac{3}{2} \end{pmatrix}$$

e si ottiene

$$x_A = 1.69747, \quad x_B = 2.06414, \quad x_C = 1.23838$$

Il valore di Shapley può entrare in gioco ad esempio anche nei casi in cui bisogna valutare come dividere una spesa o un guadagno. Forniamo due esempi: ci sono tre creditori A, B e C di una piccola ditta che vantano i tre crediti di 30.000, 20.000 e 10.000 euro. Il debitore fallisce e in cassa ci sono solo 18.000 euro per rifondere i creditori. Come dividere i 18.000 euro fra i creditori? Se, ad esempio si dividono i crediti in maniera proporzionale, cioè come 9, 6 e 3, si ha che il maggior creditore ha anche la maggior perdita e quindi tale idea non sembra molto equa. Affrontiamo allora il problema secondo la teoria dei giochi a n persone.

In questo caso non ci sono strategie a priori che i giocatori possono utilizzare. Immaginiamo allora che una coalizione R riceva tutto il denaro di cui è creditrice oppure la massima somma disponibile. Quindi possiamo definire (in migliaia di euro)

$$v(A) = 18, \quad v(B) = 18, \quad v(C) = 10, \quad v(AB) = 18, \quad v(AC) = 18, \quad v(BC) = 18, \quad v(ABC) = 18$$

Si vede subito che il nucleo è vuoto. Quindi si deve ricorrere al valore di Shapley. Le sei permutazioni danno i seguenti valori

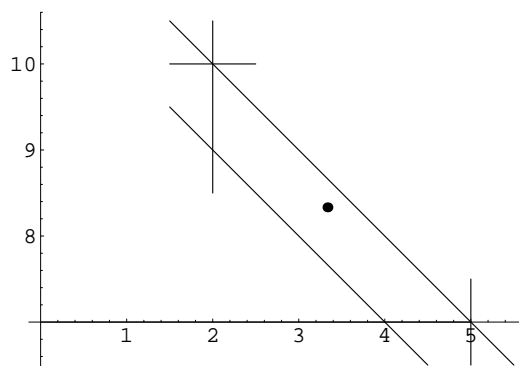


Figura 85.

	A	B	C
ABC	18	0	0
ACB	18	0	0
BAC	0	18	0
BCA	0	18	0
CAB	8	0	10
CBA	0	8	10
	44	44	20

Quindi A e B ricevono $22/3$ (la stessa somma!) e C riceve $10/3$. In Fig. 84 si vede la ripartizione della somma disponibile fra i tre creditori in funzione della stessa. Ovviamente per una somma di 10 o meno i creditori ricevono la stessa somma, mentre per una somma superiore a 60 ricevono in parti proporzionali. Per valori intermedi si hanno diverse possibilità come evidenziato in figura.

Come secondo esempio si immagini che tre persone devono prendere il taxi per andare in tre posti diversi partendo dallo stesso posto (ad esempio una stazione). Le corse singole costano $v(A) = 5$, $v(B) = 10$, $v(C) = 3$. Se il taxi porta con la stessa corsa due persone il prezzo della corsa è $v(AB) = 12$, $v(AC) = 7$, $v(BC) = 12$. Se infine viene effettuata una corsa singola il prezzo è $v(ABC) = 14$. Cosa conviene fare ai tre viaggiatori?

Correttamente i valori dovrebbero essere negativi perché si tratta di una spesa e non di un guadagno, ma per semplicità conviene pensare direttamente ai valori positivi e tener conto del fatto che si tratta di una spesa. Quindi il nucleo viene definito da $\sum_{i \in S} x_i \leq v(R)$. In Fig. 85 si vedano il nucleo e il valore di Shapley che è $(3 + 1/3, 8 + 1/3, 2 + 1/3)$ (calcolo lasciato come esercizio). Il nucleo non è vuoto quindi ai tre viaggiatori conviene effettuare un'unica corsa la cui spesa viene divisa secondo il valore di Shapley.

Formalmente il valore di Shapley può essere definito a partire da tutte le permutazioni dei giocatori (come abbiamo fatto negli esempi) come

$$u_i = \frac{1}{n!} \sum_{\pi \in \Pi} v(\{j : \pi_j \leq \pi_i\}) - v(\{j : \pi_j < \pi_i\}) \quad (141)$$

oppure, tenendo conto che in tutte le permutazioni che mantengono invariato l'insieme dei giocatori che precedono i (e quindi anche l'insieme di quelli che seguono) la differenza $v(\{j : \pi_j \leq \pi_i\}) - v(\{j : \pi_j < \pi_i\})$ è sempre uguale, dall'espressione

$$u_i = \sum_{R \subset N, i \notin R} \frac{|R|! (n - |R| - 1)!}{n!} (v(R \cup i) - v(R)) \quad (142)$$

Questa seconda espressione è molto più conveniente computazionalmente dato che 2^n cresce molto più lentamente di $n!$. Va detto che il valore di Shapley si calcola comunque per valori non elevati di n . Dall'espressione (141) si ha

$$\begin{aligned} \sum_i u_i &= \frac{1}{n!} \sum_i \sum_{\pi \in \Pi} v(\{j : \pi_j \leq \pi_i\}) - v(\{j : \pi_j < \pi_i\}) = \\ &= \frac{1}{n!} \sum_{\pi \in \Pi} \sum_i v(\{j : \pi_j \leq \pi_i\}) - v(\{j : \pi_j < \pi_i\}) = \frac{1}{n!} \sum_{\pi \in \Pi} v(N) = v(N) \end{aligned}$$

come deve essere trattandosi di un'imputazione.

Il valore di Shapley viene anche usato come indice di potere per capire il peso di un partito (o di un gruppo) nella formazione di coalizioni politiche. In questo caso $v(R) \in \{0, 1\}$ per ogni $R \subset N$, cioè ogni coalizione è vincente oppure perdente. Ad ogni giocatore è associato un valore numerico w_i (as esempio i voti o i seggi ottenuti) e sia $W = \sum_i w_i/2$. Allora si definiscono

$$v(R) = \begin{cases} 1 & \text{se } \sum_{i \in R} w_i > W \\ 0 & \text{altrimenti} \end{cases}$$

Ad esempio siano presenti quattro partiti 1, 2, 3 e 4 con percentuali rispettivamente 35, 30, 25 e 10. Calcolando il valore di Shapley usando (142) si vede che la differenza $v(R \cup i) - v(R)$ è diversa da 0 ed uguale a 1 solo quando l'aggiunta del giocatore i ad R rende la coalizione vincente. Per il partito 1 questo avviene solo nei casi $R = \{2\}$, $R = \{3\}$, $R = \{2, 4\}$, $R = \{3, 4\}$. Quindi

$$u_1 = \frac{|1|! (4 - 1 - 1)! + |1|! (4 - 1 - 1)! + |2|! (4 - 2 - 1)! + |2|! (4 - 2 - 1)!}{4!} = \frac{1}{3}$$

Per il partito 2 questo avviene solo nei casi $R = \{1\}$, $R = \{3\}$, $R = \{1, 4\}$, $R = \{3, 4\}$. Quindi si ha

$$u_2 = u_1 = \frac{1}{3}$$

Per il partito 3 questo avviene solo nei casi $R = \{1\}$, $R = \{2\}$, $R = \{1, 4\}$, $R = \{2, 4\}$. Quindi si ha

$$u_3 = u_2 = u_1 = \frac{1}{3}$$

Per il partito 4 invece si ottiene $u_4 = 0$. Infatti unito a singoli partiti non è in grado di far raggiungere la maggioranza mentre le coalizioni di due partiti hanno già la maggioranza. Quindi il partito 4 è ininfluente e il suo indice di potere è nullo. Per esercizio si possono calcolare gli indici di potere di tre partiti con percentuali 45, 45 e 10. I tre indici sono uguali!

Come ulteriore esempio si consideri il Consiglio di Sicurezza dell'ONU che è formato da 16 nazioni di cui 5 permanenti e con diritto di veto (indici 1,2,3,4 e 5). In questo caso $v(R) = 1$ se e solo se i membri permanenti sono tutti in R ed inoltre c'è la maggioranza. Per un membro non permanente $v(R \cup i) - v(R) = 1$ se e solo se $\{1, 2, 3, 4, 5\} \subset R$ e $|R| = 8$. Questo avviene in $\binom{10}{3}$ casi per cui la formula (142) diventa

$$u_i = \binom{10}{3} \frac{8! 7!}{16!} = \frac{1}{858} = 0.0011655$$

Quindi l'indice di potere di un membro permanente k è (per esercizio si calcoli direttamente questo valore)

$$u_k = \frac{1}{5} (1 - 11 u_i) = \frac{77}{390} = 0.197436$$

Si noti come il potere di veto dia ai membri permanenti un indice di potere complessivo pari a 0.987179.

61 Nucleolo e Talmud

Un altro importante concetto che entra in gioco quando il nucleo sia vuoto è il *nucleolo*, introdotto da Schmeidler nel 1969. Tuttavia si è visto che in realtà tale concetto è già presente nel Talmud babilonese. Il concetto di nucleolo è legato al concetto di *eccesso di un'imputazione*. Data un'imputazione x si definisce eccesso dell'imputazione rispetto ad una coalizione S la quantità

$$v(S) - \sum_{i \in S} x_i$$

che misura di quanto la coalizione S 'manchi' il guadagno $v(S)$ che le spetta se riceve l'imputazione x . Più grande è questo valore tanto più 'ingiusta' deve essere considerata l'imputazione. Dalla definizione di nucleo si vede che il nucleo esiste se e solo se esiste un'imputazione per la quale tutti gli eccessi siano nulli o negativi.

L'idea è allora quella di minimizzare la massima ingiustizia, ovvero risolvere il seguente problema di Programmazione lineare, dove \mathcal{S} è l'insieme di tutti i sottoinsiemi di N escluso l'insieme vuoto ed N stesso.

$$\begin{aligned} \min \quad & w \\ & w \geq v(S) - \sum_{i \in S} x_i \quad S \in \mathcal{S} \\ & \sum_{i \in N} x_i = v(N) \end{aligned}$$

La soluzione ottima che si ottiene ha un valore w^1 per cui alcune disequaglianze sono attive e le rimanenti non lo sono. Più esattamente si considerino le disequaglianze attive con valore positivo della variabile duale. Per

queste disequaglianze w non può essere ulteriormente diminuito. Sia \mathcal{S}^1 l'insieme di queste disequaglianze. Per poter diminuire ancora w per le rimanenti disequaglianze si deve risolvere il seguente problema di Programmazione lineare (dove w^1 è un valore costante)

$$\begin{aligned} \min \quad & w \\ & w^1 \geq v(S) - \sum_{i \in S} x_i \quad S \in \mathcal{S}^1 \\ & w \geq v(S) - \sum_{i \in S} x_i \quad S \in \mathcal{S} \setminus \mathcal{S}^1 \\ & \sum_{i \in N} x_i = v(N) \end{aligned}$$

Anche in questo caso alcune disequaglianze in $\mathcal{S} \setminus \mathcal{S}^1$ sono attive con variabile duale positiva. Siano \mathcal{S}^2 queste disequaglianze e sia w^2 il valore ottimo del problema. Allora si passa a risolvere

$$\begin{aligned} \min \quad & w \\ & w^1 \geq v(S) - \sum_{i \in S} x_i \quad S \in \mathcal{S}^1 \\ & w^2 \geq v(S) - \sum_{i \in S} x_i \quad S \in \mathcal{S}^2 \\ & w \geq v(S) - \sum_{i \in S} x_i \quad S \in \mathcal{S} \setminus (\mathcal{S}^1 \cup \mathcal{S}^2) \\ & \sum_{i \in N} x_i = v(N) \end{aligned}$$

Procedendo ricorsivamente si perviene alla soluzione finale che si può dimostrare essere unica.

Ottimi ottenuti con procedure di questo genere vengono detti *ottimi lessicografici non ordinati*. In generale un ordine lessicografico fra elementi di \mathbb{R}^m (in questo caso $m = 2^n - 2$, cioè il numero dei sottoinsiemi di N esclusi l'insieme vuoto e N stesso) viene definito dalla seguente relazione di preferenza \prec (in questo caso riferita a problemi di minimizzazione):

$$a \prec b \iff \exists k = 1, \dots, m : a_i = b_i, \quad i = 1, \dots, k-1, \quad a_k < b_k$$

L'ordine lessicografico è un ordine totale. Dati due elementi $a \neq b$ in \mathbb{R}^m , $a \prec b$ oppure $b \prec a$. Non vi sono altre alternative.

Gli ottimi lessicografici non ordinati sono definiti nel seguente modo: dato un elemento $a \in \mathbb{R}^m$ sia $\theta(a)$ il vettore ottenuto permutando le componenti di a in modo che le componenti di $\theta(a)$ siano in ordine non crescente (se vi sono componenti uguali possono essere ordinate in modo arbitrario). Allora dati due elementi a e b in \mathbb{R}^m , a è preferito a b se $\theta(a) \prec \theta(b)$ (in senso lessicografico) e a è ottimo lessicografico non ordinato se non esiste b preferito a a .

In alcuni casi semplici il nucleolo può essere calcolato direttamente con semplici ragionamenti. Nel caso dell'esempio del taxi si consideri inizialmente l'imputazione $x_A = 4$, $x_B = 6$, $x_C = 4$. Allora gli eccessi sono

$$\begin{array}{lll} 5 - x_A = 1 & 10 - x_B = 4 & 3 - x_C = -1 \\ 12 - x_A - x_B = 2 & 7 - x_A - x_C = -1 & 12 - x_B - x_C = 2 \end{array}$$

Il peggiore eccesso si ottiene per l'insieme $\{C\}$ e per l'insieme $\{A, C\}$. Se si abbassa x_C di una quantità Δ si deve alzare x_A oppure x_B oppure entrambi della medesima quantità. Supponiamo di aumentare x_B di Δ . Allora gli eccessi che cambiano sono quelli dell'insieme $\{B\}$ che diventa $10 - x_B - \Delta = 4 - \Delta$, dell'insieme $\{C\}$ che diventa $3 - x_C + \Delta = -1 + \Delta$, dell'insieme $\{A, B\}$ che diventa $12 - x_A - x_B - \Delta = 2 - \Delta$ e dell'insieme $\{A, C\}$ che $7 - x_A - x_C + \Delta = -1 + \Delta$. Il massimo valore di Δ per cui si ottiene il minimo peggior eccesso è dato da $2 - \Delta = -1 + \Delta$, cioè $\Delta = 3/2$.

La nuova imputazione è allora $x_A = 4$, $x_B = 7.5$, $x_C = 2.5$ con eccessi

$$\begin{array}{lll} 5 - x_A = 1 & 10 - x_B = 2.5 & 3 - x_C = 0.5 \\ 12 - x_A - x_B = 0.5 & 7 - x_A - x_C = 0.5 & 12 - x_B - x_C = 2 \end{array}$$

A questo punto l'eccesso dell'insieme $\{C\}$ non è ulteriormente modificabile e quindi proviamo ad aumentare x_B di Δ e diminuire x_A di Δ per cui si avrebbe:

$$\begin{array}{lll} 5 - x_A + \Delta = 1 + \Delta & 10 - x_B - \Delta = 2.5 - \Delta & 3 - x_C = 0.5 \\ 12 - x_A - x_B = 0.5 & 7 - x_A - x_C + \Delta = 0.5 + \Delta & 12 - x_B - x_C - \Delta = 2 - \Delta \end{array}$$

da cui si vede che si ottiene $\Delta = 0.75$ ponendo $0.5 + \Delta = 2 - \Delta$ e quindi il nucleolo è

$$x_A = 3.25, \quad x_B = 8.25, \quad x_C = 2.5$$

. Si noti che è un valore diverso dal valore di Shapley.

Il problema di ottenere indietro i soldi dati in prestito è ovviamente molto antico e come dividere le somme fra i creditori è stato oggetto anche del Talmud babilonese. In particolare si trova la seguente Mishna del quinto secolo a.C. che si trova nel capitolo Baba Metzia 2a (la parola ebraica Mishna può significare sia l'intero complesso dei testi da imparare a memoria - da cui la parola - sia singoli paragrafi o versi, come nel caso attuale):

Due persone hanno in mano un indumento, e uno di loro dice "è tutto mio" e l'altro dice "metà è mio". Allora il primo riceverà tre quarti e il secondo un quarto.

Chiamiamo MI (Mishna dell'indumento) questa regola. La logica della MI è questa: metà indumento è reclamata da entrambi e quindi viene divisa fra i due contendenti, l'altra metà è reclamata solo da uno e viene data tutta a questo. Quindi il primo riceve una metà e la metà dell'altra metà, cioè i tre quarti.

La MI può essere estesa nel seguente modo. Sia E la somma da dividere fra due creditori e siano c_1 e c_2 i crediti reclamati dai due creditori. Allora se $c_1 > E$ e $c_2 > E$ (e quindi entrambi reclamano tutto), ricevono entrambi la metà di E .

Se invece $c_1 < E$ e $c_2 > E$ (solo il secondo reclama tutto ed è il caso descritto dalla MI), allora la parte reclamata dal primo viene divisa fra i due e il secondo prende il resto. Quindi $x_1 = c_1/2$ e $x_2 = E - c_1/2$.

Se infine $c_1 < E$ e $c_2 < E$, allora $(c_1 + c_2 - E)$ è la parte reclamata da entrambi, che viene divisa in parti uguali, $E - c_2$ è la parte reclamata solo dal primo, che quindi gli viene data tutta e $E - c_1$ è la parte reclamata solo dal secondo, che gli viene data tutta. In questo caso il primo riceve $(c_1 + c_2 - E)/2 + E - c_2 = (E + c_1 - c_2)/2$ e il secondo $(E + c_2 - c_1)/2$. Si noti la somiglianza di queste formule con l'espressione (138) a pag. 182.

In un altro passo del Talmud (Kethubot 93 a) si trova scritta la seguente Mishna (che possiamo chiamare MC, Mishna dei creditori)

Tre creditori vantano crediti rispettivamente di 100, 200 e 300. Se il denaro disponibile in cassa è 100 i tre creditori ricevono parti uguali. Se è 200 ricevono 50, 75, 75 rispettivamente e se è 300 ricevono 50, 100 e 150 rispettivamente

L'analisi di questi passi e le loro connessioni con la moderna teoria dei giochi si trova nell'interessante articolo di R.J. Aumann e M. Maschler, Game theoretic analysis of a bankruptcy problem from the Talmud, J. of Economic Theory, vol.36, 195-213, 1985.

I due passi non sembrano legati fra loro e nemmeno è chiara a prima vista la logica del secondo passo. Un passo che sembra legare le due Mishna si trova nel Talmud di Gerusalemme, che è considerato meno autorevole di quello babilonese, però spesso viene usato come ulteriore fonte per dirimere una questione. Il passo in questione, che si riferisce a tre donne creditrici, dice:

Samuele dice, la Mishna assume che i creditori si uniscano; in particolare che la terza dia mandato alla seconda di negoziare per lei con la prima. Lei (la seconda) può dire all'altra (la prima): tu vuoi 100? Bene, prendi 50 e vattene.

In questo passo l'ordine (prima, seconda e terza) è dato da crediti crescenti. Esaminiamo alla luce di questo passo la MC prendendo in esame il caso della somma di 300 disponibile. Il terzo e il secondo creditore si uniscono vantando un credito complessivo di 500, ben superiore ai 300 disponibili. Il primo ha un credito di 100. Questi 100 sono reclamati quindi sia dal primo che dalla coalizione del secondo e del terzo. In base alla MI questi 100 sono divisi in parti uguali fra il primo da una parte e gli altri due, dall'altra parte. Il primo riceve 50 e se ne va. Adesso i restanti 250 devono essere divisi fra il secondo e il terzo. Il credito di 300 del terzo è superiore alla somma disponibile. Di questa somma 200 sono reclamati da entrambi e vanno divisi in due parti uguali secondo la MI. Quindi il secondo prende 100 e se ne va. Come si vede si ottiene esattamente la divisione indicata dalla MC.

L'analisi del caso con somma disponibile 200, fatta nello stesso modo, porta di nuovo alla divisione fornita dalla MC. Tuttavia l'applicazione al caso di somma disponibile 100, dà luogo ad un risultato diverso e contraddittorio. Infatti il minor creditore riceverebbe 50 e gli altri due 25 a testa!

La MI può però essere invocata in modo più sistematico. Sia E l'ammontare disponibile e siano (c_1, \dots, c_n) i crediti vantati, con $c_1 \leq c_2 \leq \dots \leq c_n$ e $E < c_1 + c_2 + \dots + c_n$. Una soluzione x è consistente rispetto alla MI se la divisione (x_i, x_j) per una somma disponibile $x_i + x_j$ rispetto alle quantità reclamate c_i e c_j è conforme

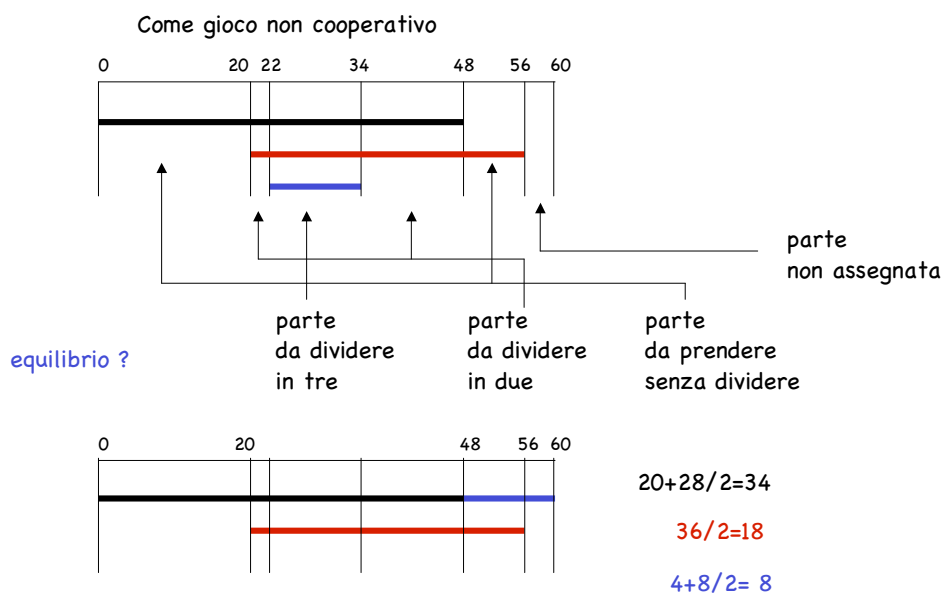


Figura 86.

alla MI, per ogni coppia (i, j) . Chiamiamo *problema della bancarotta* (*bankruptcy problem*), il problema di determinare le somme da versare ai creditori.

In base a questo requisito si vede che la MC è consistente rispetto alla MI. Si possono dimostrare le seguenti proprietà (si veda l'articolo citato dove vengono evidenziate anche altre proprietà che qui tralasciamo):

Esiste sempre ed è unica la soluzione al problema della bancarotta. Inoltre la soluzione corrisponde al nucleolo del gioco.

Esaminiamo quanto detto su un particolare problema di bancarotta. Tre creditori vantano crediti per $c_1 = 12$, $c_2 = 36$ e $c_3 = 48$, ma la somma disponibile è $E = 60$. Se si applica il valore di Shapley si ottiene $x_1 = 8$, $x_2 = 20$, $x_3 = 32$. Se si applica il nucleolo si ottiene $x_1 = 6$, $x_2 = 21$ e $x_3 = 33$.

Possiamo anche chiederci quali sono gli equilibri di Nash del problema. Per poter definire un equilibrio di Nash dobbiamo immaginare la somma disponibile come un intervallo I sulla retta reale di lunghezza E e il credito c_i come l'unione J_i di uno o più intervalli disgiunti di lunghezza complessiva c_i . Questi intervalli rappresentano i punti che il giocatore i reclama. Una "strategia" da parte del giocatore i è data dalla scelta degli intervalli. Per ogni punto $a \in I$ sia $N(a) := \{i : a \in J_i\}$, cioè l'insieme dei giocatori che reclamano il punto a . Suddividiamo I in un insieme di intervalli I_k definito come

$$I_k := \{a \in I : |N(a)| = k\}$$

cioè tutti i punti reclamati da k giocatori. Sia $L(I_k)$ la lunghezza complessiva degli intervalli in I_k . Si noti che $\sum_{k=0}^n L(I_k) = E$. La somma $L(I_k)$ corrispondente agli intervalli I_k viene suddivisa in parti uguali fra i

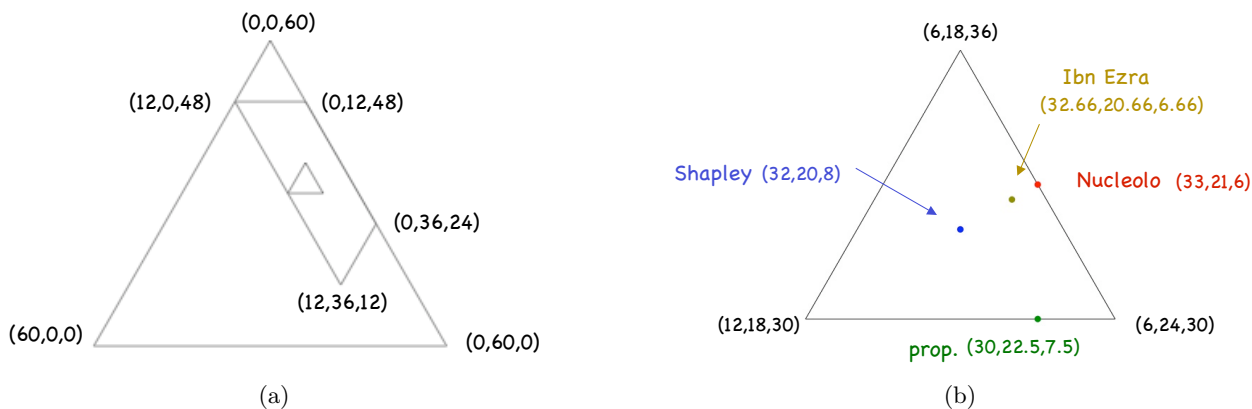


Figura 87.

k giocatori che la reclamano. La parte $L(I_0)$ non viene data a nessuno. Allora quello che riceve il giocatore i sarà dato da

$$x_i = \sum_{k=1}^n \frac{L(J_i \cap I_k)}{k}$$

Si veda in Fig. 86 una raffigurazione del metodo. In base a questa divisione e immaginando che ogni giocatore possa cambiare strategia, ovvero la scelta degli intervalli, si vede che una situazione di equilibrio è data da una scelta di intervalli in cui gli unici valori possibili di k , corrispondenti ad insiemi non vuoti I_k , sono dati da

$$\left[\frac{\sum_i c_i}{E} \right], \quad \left[\frac{\sum_i c_i}{E} \right]$$

Si veda in Fig. 87(a) una raffigurazione dell'insieme (simplex) di tutte le imputazioni possibili

$$\{x \geq 0 : x_1 + x_2 + x_3 = 60\}$$

(triangolo grande), del nucleo (trapezio all'interno del triangolo) e degli equilibri di Nash (triangolo all'interno del trapezio) e in Fig. 87(b) l'insieme degli equilibri di Nash ingrandito con l'indicazione delle varie imputazioni inclusa una divisione proporzionale ed una divisione suggerita dal rabbino Abraham Ibn Ezra (Toledo 1140 d.C.). A questo riguardo si veda anche l'articolo di Barry O'Neill, A problem of rights arbitration from the Talmud, *Mathematical Social Sciences*, vol. 2, 345-371, 1982.

62 Reti e Teoria dei giochi

Un aspetto molto interessante della Teoria dei Giochi riguarda il calcolo dei flussi ottimi su una rete di comunicazione. Prendiamo in esame il caso del traffico automobilistico su un insieme di strade. Ogni automobilista

tende a scegliere la strada più rapida per arrivare a destinazione. Modelliamo allora l'insieme di strade come una rete orientata, supponiamo che tutto il traffico vada da un nodo sorgente s ad un nodo destinazione t e che il tempo di percorrenza di un arco sia funzione del traffico stesso presente sull'arco, secondo una funzione non decrescente.

Il traffico viene modellato secondo un flusso in quantità data r che da s va a t distribuendosi sugli archi della rete. Il tempo di percorrenza di un arco viene chiamato *latenza* e viene definito come una funzione del flusso x_e sull'arco e

$$\ell_e(x_e) = a_e + b_e x_e$$

Vari tipi di funzioni possono essere usate. Qui valutiamo solo il caso di funzioni affini (termine costante più termine lineare). Quindi a_e è il tempo di percorrenza in assenza di traffico. La presenza di traffico crea un po' di congestione e quindi il tempo aumenta della quantità $b_e x_e$. Questa funzione non è del tutto realistica in quanto è verosimile che per valori elevati di x_e il tempo tenda ad infinito. Comunque, in prima approssimazione, una funzione affine è una valida misura di latenza.

Due sono gli scenari che possiamo immaginare. Da un lato possiamo pensare che ci sia un'autorità centrale che ha il potere di indirizzare il traffico e mira a ridurre i tempi di percorrenza per tutti gli automobilisti. Tale autorità vuole minimizzare la funzione obiettivo

$$\Phi(x) = \sum_{e \in E} \ell_e(x_e) x_e = \sum_{e \in E} (a_e + b_e x_e) x_e$$

dove, per ogni arco, si moltiplica la latenza per la quantità di traffico che effettivamente ha quel ritardo. Si può dimostrare che se $\ell_e(x_e) x_e$ è una funzione convessa, e $(a_e + b_e x_e) x_e$ lo è, e \bar{x} è il flusso che minimizza $\Phi(x)$ allora esiste una funzione $U_i(r)$ (con $U_s(r) = 0$) per cui

$$\begin{aligned} U_j(r) - U_i(r) &= \ell'_{ij}(\bar{x}_{ij}) \bar{x}_{ij} + \ell_{ij}(\bar{x}_{ij}) = a_{ij} + 2 b_{ij} \bar{x}_{ij} && \text{se } \bar{x}_{ij} > 0 \\ U_j(r) - U_i(r) &\leq \ell_{ij}(0) = a_{ij} && \text{se } \bar{x}_{ij} = 0 \end{aligned}$$

Se si sommano le uguaglianze lungo un cammino P sul quale ci sia del flusso positivo si ottiene

$$U_t(r) = \sum_{e \in P} (a_e + 2 b_e \bar{x}_e) = \sum_{e \in P} (a_e + b_e \bar{x}_e) + \sum_{e \in P} b_e \bar{x}_e$$

Siccome la durata del tragitto per chi deve viaggiare sul cammino P è data da $\sum_{e \in P} (a_e + b_e \bar{x}_e)$, si vede che la durata è uguale ad una quantità costante $U_t(r)$, uguale per tutti, meno una parte variabile $\sum_{e \in P} b_e \bar{x}_e$ che dipende dal cammino P .

Nel secondo scenario ogni automobilista decide per proprio conto dove andare minimizzando il proprio ritardo. Quindi il comportamento degli automobilisti è assimilabile ad un gioco non cooperativo e il traffico risultante sarà un equilibrio di Nash.

Se il flusso \hat{x} è un equilibrio di Nash significa che ogni cammino da s a t , sul quale ci sia del flusso, deve avere la stessa latenza, altrimenti ci sarebbe un travaso di flusso dal cammino più lungo a quello più corto. La condizione di equilibrio permette di definire una funzione $L_i(r)$ che rappresenta la latenza da s fino al

nodo i , indipendente dal cammino percorso, purché ci sia del flusso positivo lungo il cammino. La funzione $L_i(r)$ deve obbedire alle seguenti condizioni di equilibrio

$$\begin{aligned} L_j(r) - L_i(r) &= \ell_{ij}(\hat{x}_{ij}) = a_{ij} + b_{ij} \hat{x}_{ij} && \text{se } \hat{x}_{ij} > 0 \\ L_j(r) - L_i(r) &\leq \ell_{ij}(0) = a_{ij} && \text{se } \hat{x}_{ij} = 0 \end{aligned}$$

Il valore $L_t(r)$ è la latenza alla destinazione, cioè è la durata del tragitto per tutti gli automobilisti. Allora deve valere la relazione

$$r L_t(r) = \Phi(\hat{x})$$

La quantità $\Phi(\bar{x})$ è la minima latenza totale e quindi il flusso \bar{x} sarebbe il traffico ottimo per tutti. Tuttavia per ottenere questa quantità sarebbe richiesta (ma non sempre) un'autorità centrale che impone le rotte agli automobilisti. La quantità $\Phi(\hat{x})$ è invece la latenza totale che si ottiene in modo spontaneo senza regolazioni e lasciando agli automobilisti di seguire il proprio istinto 'egoistico'. Ovviamente, per definizione, $\Phi(\bar{x}) \leq \Phi(\hat{x})$ ma siamo interessati a valutare di quanto può degradare il traffico lasciato a se stesso. Il rapporto

$$\frac{\Phi(\hat{x})}{\Phi(\bar{x})}$$

viene definito *prezzo dell'anarchia*. Per capire meglio il problema si consideri il seguente esempio paradossale, che è stato proposto da Braess nel 1968. In Fig. 88(a) si vede una rete stradale. Sugli archi 1-3 e 2-4 il ritardo è costante e non dipende dalla quantità del traffico, mentre sugli archi 1-2 e 3-4 il ritardo dipende linearmente dal traffico. Si supponga che un flusso unitario deve andare da 1 a 4. Per questa rete $\bar{x} = \hat{x}$. In particolare il traffico si divide in due parti uguali e quindi le latenze dei due cammini sono

$$\frac{1}{2} + 1 = 1 + \frac{1}{2} = \frac{3}{2}$$

La latenza totale è

$$\frac{1}{2} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{2}$$

Per accelerare il traffico si decide di costruire una strada ad alta velocità fra 2 e 3 come in Fig. 88(b) dove la latenza viene considerata nulla. Se si minimizza il ritardo totale si ottiene che sul nuovo arco non circola traffico e quindi la migliore soluzione rimane quella precedente con valore totale $3/2$. Però l'equilibrio di Nash cambia. Essendo ora presente il cammino 1-2-3-4 parte del traffico trova conveniente usare questo cammino che ha latenza $1/2$, in assenza di traffico sull'arco 2-3. Si vede facilmente che il cammino più rapido risulta essere il cammino 1-2-3-4 e quindi tutto il traffico confluisce su questo cammino con una latenza

$$1 + 0 + 1 = 2$$

Il risultato di avere aperto una strada veloce è che ora ci si mette più tempo ad arrivare alla destinazione! La latenza totale $\Phi(\hat{x})$ dell'equilibrio di Nash è 2 per cui il prezzo dell'anarchia risulta essere $4/3$. Si può dimostrare che questo rapporto non può essere peggiore di $4/3$ in generale per latenze date da funzioni affini.

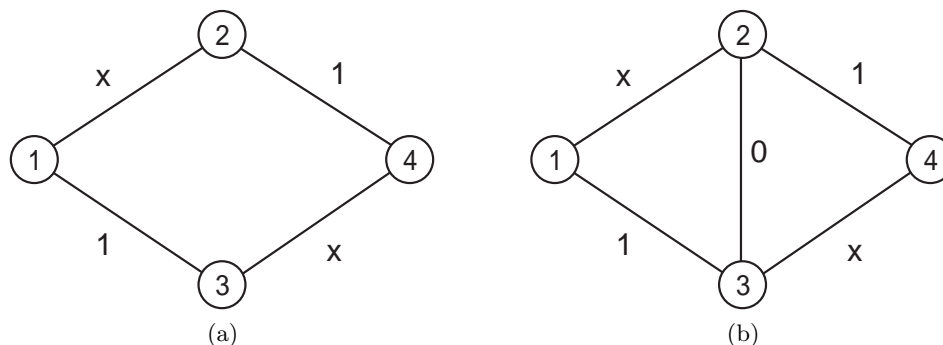


Figura 88.

Il paradosso rimane valido anche assumendo che sugli archi 1-2 e 3-4 la latenza sia una generica funzione crescente $\ell(x)$. Se $\ell(1) \leq 1$ il traffico percorrerà sempre il cammino 1-2-3-4 con latenza $2\ell(1)$ piuttosto che di uno dei due cammini 1-2-4 o 1-3-4 con latenza $1 + \ell(1)$. In questo caso il traffico che si divide fra i due cammini 1-2-4 e 1-3-4 ha latenza $1 + \ell(1/2) < 1 + \ell(1)$. Se inoltre $1 + \ell(1/2) < 2\ell(1)$ l'equilibrio di Nash non è l'ottimo totale.

Se invece $\ell(1) > 1$ con $\alpha > 1/2$ tale che $\ell(\alpha) = 1$ allora il traffico si divide come

$$\hat{x}_{12} = \alpha, \quad \hat{x}_{13} = 1 - \alpha, \quad x_{23} = 2\alpha - 1, \quad x_{24} = 1 - \alpha, \quad x_{34} = \alpha$$

con latenza 2, mentre il traffico \bar{x} ha latenza $1 + \ell(1/2) < 1 + \ell(\alpha) = 2$. Solo se $\ell(1) > 1$ con $\alpha < 1/2$ si ha che $\bar{x} = \hat{x}$ e il traffico si divide nei due cammini 1-2-4 e 1-3-4 senza usare la nuova strada.

Il risultato è però sorprendente in ogni caso: la nuova strada serve solo se $\ell(x)$ è particolarmente bassa e in quel caso rende inutili le strade 2-4 e 1-3. Altrimenti non serve o fa aumentare i tempi!

Riassumendo: gli automobilisti 'collaborativi' percorrerebbero i cammini 1-2-4 o 1-3-4, ma ci sono degli automobilisti 'furbi' che trovano conveniente approfittare della situazione data dagli automobilisti collaborativi per effettuare il più corto tragitto 1-2-3-4. Però tale tragitto rimane corto se gli archi 1-2 e 3-4 non hanno molto traffico e l'effetto dei furbi non può che indurre altri automobilisti ad abbandonare i cammini 1-2-4 e 1-3-4 in favore di 1-2-3-4, con i risultati visti. A ben vedere, è una variante del dilemma del prigioniero.

Si può dimostrare che non c'è differenza fra \bar{x} e \hat{x} se $\ell_e(x) = a_e$ (per tutti gli archi) oppure $\ell_e(x) = b_e x$ (per tutti gli archi). Nel primo caso tutto il flusso si riversa sul cammino minimo (rispetto ai costi a_e). Nel secondo caso il flusso si dispone come si disporrebbe la corrente elettrica su archi in cui la resistenza vale b_e per cui la latenza diventa la differenza di potenziale fra i due estremi dell'arco e la soluzione di minima latenza è la stessa di minima potenza dissipata.

Indice

1. Definizione di grafo non orientato.....	2
2. Cammini, circuiti e connessione.....	3
3. Grado di un nodo e regolarità.....	5
4. Isomorfismo.....	7
5. Grafi particolari.....	9
6. Grafi ottenuti da grafi.....	10
7. Cricche, coperture e colorazioni.....	13
8. Lunghezza di un cammino e di un circuito.....	16
9. Tagli.....	19
10. Alberi e foreste.....	20
11. Grafi planari.....	20
12. Grafi orientati.....	21
13. Cosa è una matrice?.....	22
14. Matrici d'incidenza di grafi non orientati.....	32
15. Matrici d'incidenza di grafi orientati.....	34
16. Matrici d'adiacenza di un grafo non orientato.....	36
17. Matrici d'adiacenza di grafi prodotto.....	39
18. Matrici d'adiacenza di un grafo orientato.....	40
19. Matrice Laplaciana.....	42
20. Spettro di un grafo.....	44
21. Spettro di un grafo regolare.....	48
22. Spettro di un grafo completo.....	50
23. Spettro di un grafo bipartito.....	50
24. Spettro di un circuito.....	54
25. Spettro di un cammino.....	55
26. Spettro del grafo complementare.....	55
27. Spettro del grafo prodotto.....	56
28. Laplaciano e processi dinamici su reti.....	58
29. Reti elettriche.....	59
30. Catene e processi di Markov.....	62
31. Il Page Rank di Google.....	67
32. Famiglie di sottoinsiemi.....	68
33. Cocitazione e accoppiamento bibliografico.....	72
34. Minimi tagli e massimi flussi.....	75
35. Cammini disgiunti e tagli minimi.....	78
36. Tagli minimi in un grafo.....	79
37. Tagli e autovalori.....	83
38. Partizione dei vertici in sottoinsiemi (clustering): introduzione.....	89

39. Modularità con numero di archi fissato – Partizione in due insiemi	91
40. Modularità con gradi fissati – Partizione in due insiemi	94
41. Modularità con gradi fissati – Calcolo dell’autovalore più grande	102
42. Modularità con gradi fissati – Partizione in più insiemi	103
43. Grafi casuali	109
44. Modello $G_{n,p}$	112
45. Grafi casuali con numero d’archi prefissato	114
46. Grafi casuali con gradi fissati	116
47. Funzioni generatrici in generale	118
48. Funzioni generatrici di probabilità	124
49. Processi di diramazione	132
50. Funzioni generatrici dei gradi dei nodi	134
51. Componenti connesse piccole e componente gigante	139
52. Componenti connesse in grafi particolari	145
53. Grafi di Poisson	150
54. Grafi con distribuzione esponenziale	153
55. Definizione normale di gioco	156
56. Giochi a due persone a somma zero	157
57. Giochi a somma zero in forma estesa	162
58. Giochi non cooperativi a somma non zero a due persone	172
59. Giochi cooperativi a somma non zero a due persone	177
60. Giochi a molti giocatori	185
61. Nucleolo e Talmud	191
62. Reti e Teoria dei giochi	196