Introduzione ai grafi

Giandomenico Mastroeni

TFA A048
Anno Accademico 2012-13

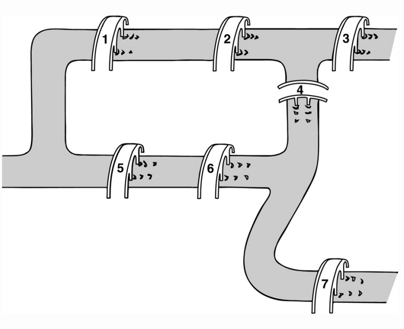
Outline

- Cenni storici sui grafi
- Nozioni introduttive: cammini, connessione, alberi, cicli
- Cammini di costo minimo

Origini storiche

La nascita della teoria dei grafi risale al 1736, anno in cui il matematico svizzero Eulero cominciò il proprio studio discutendo il cosiddetto *Problema dei ponti di Königsberg*.

Fu posto ad Eulero il problema di stabilire se fosse possibile attraversare una e una sola volta tutti i ponti, che collegano le zone della città, e tornare al punto di partenza.



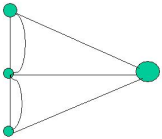
Origini storiche

Se definiamo un "cammino" come una successione di archi che connette una coppia di nodi, nel linguaggio della teoria dei grafi, la domanda si presenta in questa forma:

È possibile, dato un grafo connesso, trovare un cammino che attraversi tutti gli spigoli una e una sola volta, tornando al punto di partenza?

Tale cammino, se esiste, è detto *cammino di Eulero*.

I ponti di Konisberg



Grafi: definizione formale

Definizione. Un grafo $G = (N, A)$ e' una coppia di insiemi $N := \{1, 2, 3, ..., n\}$, ed A sottoinsieme del prodotto cartesiano $N \times N$, ossia un insieme di coppie (i, j), $i, j \in N$.
Se le coppie si considerano ordinate, il grafo si dice orientato, altrimenti, non orientato.
Il grafo si dice *rete* se ad ogni arco e' associato un numero reale detto peso o costo.

- N e' detto l'insieme dei nodi;
- A l'insieme degli archi, $|A| = m$;
- c_{ij} e' il costo dell'arco (i, j);
- Un arco (i, j) si dice incidente nei nodi i e j, in tal caso i nodi i e j sono detti adiacenti.
- $FS(i) := \{(i, j) \in A, \ j \in N\}$;
- $BS(i) := \{(j, i) \in A, \ j \in N\}$;

Alcuni classici problemi sui grafi

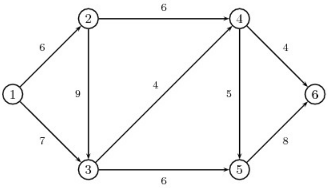
- La determinazione delle componenti connesse di un grafo;
- Il problema del cammino di costo minimo;
- Il problema del massimo flusso;
- Il problema dell'orientazione;

Nelle successive definizioni ci riferiremo ad un grafo orientato tale che $A \neq \emptyset$ e supporremo, inoltre, che non esistano archi che connettono lo stesso nodo, cioe' A non contenga coppie della forma $(\bar{n}, \bar{n})$, $\bar{n} \in N$.

Grado

Definizione. Si dice grado del nodo n, il numero degli archi $g(n)$ incidenti in tale nodo.

Rete: grafo orientato pesato.



Cammino

Definizione

Sia dato un grafo $G = (N, A)$. **(i)** Un *cammino orientato* dal nodo n_1 al nodo n_k, $k \geq 2$, e' un sottografo di G definito da una coppia di successioni $n_1, ..., n_k$ ed $a_1, ..., a_{k-1}$ tali che $n_i \in N$, $i = 1, ..., k$ e $a_i = (n_i, n_{i+1}) \in A$, per $i = 1, ..., k-1$.

(ii) Un *cammino non orientato* dal nodo n_1 al nodo n_k, $k \geq 2$, e' un sottografo di G definito da una coppia di successioni $n_1, ..., n_k$ ed $a_1, ..., a_{k-1}$ tali che $n_i \in N$, $i = 1, ..., k$ e $a_i = (n_i, n_{i+1}) \in A$ oppure $a_i = (n_{i+1}, n_i) \in A$, per $i = 1, ..., k-1$.

Osservazione

Un cammino orientato risulta univocamente determinato dalla successione dei nodi $n_1, ..., n_k$. E' facile riconoscere che cio' non accade per un cammino non orientato.

Connessione

Definizione

(i) Un grafo $G = (N, A)$ si dice *fortemente connesso* se, per ogni coppia di nodi n_i,$n_j \in N$, esiste un cammino orientato da n_i a n_j.

(ii) Un grafo $G = (N, A)$ si dice *connesso* se, per ogni coppia di nodi n_i,$n_j \in N$, esiste un cammino non orientato da n_i a n_j.

Componenti connesse di un grafo

Se un grafo non e' connesso e' di interesse determinare quali siano i sottografi di G che risultano connessi.

Una componente connessa G' di un grafo G e' un sottografo connesso massimale di G.

Un grafo puo' essere sempre partizionato nelle sue componenti connesse.

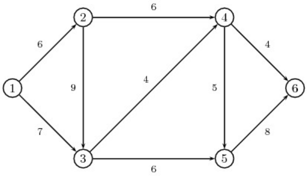
Cicli

Definizione

(i) Un *ciclo orientato (semplice)* e' un cammino orientato definito dalla successione di nodi $n_1, ..., n_k$, $k \geq 3$, tale che $n_1 = n_k$ e non esiste nessun altra coppia di nodi uguali.

(ii) Un *ciclo non orientato (semplice)* e' un cammino non orientato definito dalle successioni $n_1, ..., n_k$, $a_1, ..., a_{k-1}$, $k \geq 3$, tale che $n_1 = n_k$, non esiste nessun altra coppia di nodi uguali ed, inoltre gli archi a_i, $i = 1, ..., k-1$, sono distinti.

Esempio. (2,4) (3,4) (2,3) e' un ciclo non orientato.



Alberi

Definizione

Un grafo G connesso e privo di cicli (non orientati) e' detto *albero*.

Definizione

Un *albero di copertura* di un grafo $G = (N, A)$ e' un albero $G'(N', A')$ tale che $N' = N$ e $A' \subseteq A$.

Lemma 1

Un grafo G e' connesso se e solo se possiede almeno un albero di copertura.

Dimostrazione

Se G e' connesso ma non e' un albero, allora contiene almeno un ciclo non orientato. E' possibile cancellare da tale ciclo un arco mantenendo la connessione del grafo G. Iterando tale procedimento si arriva ad ottenere un albero di copertura di G. L'altra implicazione e' immediata.

Lemma

Se $G = (N, A)$ e' un grafo orientato tale che ogni nodo $\bar{n} \in N$ ha grado $g(\bar{n}) \geq 2$, allora G contiene almeno un ciclo non orientato.

Dimostrazione

Si scelga un qualsiasi nodo $i \in N$ e ci si sposti su un nodo adiacente utilizzando un arco (i, j) o (j, i) incidente in i. Siccome ogni nodo $\bar{n} \in N$ ha grado $g(\bar{n}) \geq 2$, e' possibile spostarsi da un nodo al successivo senza utilizzare l'arco selezionato al passo precedente: essendo il numero dei nodi finito, si dovra' necessariamente tornare su un nodo gia' considerato, generando cosi' un ciclo non orientato.

Corollario

Un albero contiene almeno un nodo $\bar{n}$ tale che $g(\bar{n}) = 1$.

Teorema

Sia $G = (N, A)$ un grafo connesso tale che $|N| = n$. Allora G e' un albero se e solo se $|A| = n - 1$.

Dimostrazione

Sia G un albero. Dimostriamo, per induzione su n, che $|A| = n - 1$. Se $n = 2$ la tesi e' vera. Supponiamo la tesi vera per $n = k > 2$ e consideriamo un grafo $G = (N, A)$ tale che $|N| = k + 1$. Per il Corollario precedente, se G e' un albero allora possiede almeno un nodo $\bar{n}$ tale $g(\bar{n}) = 1$. Sia $\bar{a} \in A$ l'unico arco incidente in $\bar{n}$ e consideriamo il grafo $G' := (N', A')$ ove $(N', A') := (N \setminus \{\bar{n}\}, A \setminus \{\bar{a}\})$: G' e' un albero tale che $|N'| = k$. Per ipotesi induttiva $|A'| = k - 1$ e, pertanto, $|A| = k$. Viceversa, supponiamo $G = (N, A)$ con $|A| = n - 1$. Essendo G connesso, per il Lemma 1, possiede almeno un albero di copertura $G' \subseteq G$, che, per l'implicazione appena dimostrata, ha $n - 1$ archi. Pertanto $G' = G$, da cui segue che G e' esso stesso un albero.

Proposizione

Un grafo $G(N, A)$ tale che $|N| = n$, $|A| = n - 1$ e privo di cicli e' connesso.

Dimostrazione

Detta componente connessa di G un sottografo connesso massimale di G, se per assurdo G non fosse connesso dovrebbe essere rappresentato come unione di $k \geq 2$ componenti connesse $G(N_i, A_i)$. Sia $n_i := |N_i| = n$, $i = 1, k$, il numero di nodi della i-esima componente connessa. Osserviamo che $\sum_{i=1}^{k} n_i = n$, e che, essendo G privo di cicli ciascuna componente connessa e' un albero. Per il Teorema precedente avremo che:

$$|A_i| = n_i - 1$$

da cui

$$\sum_{i=1}^{k} |A_i| = \sum_{i=1}^{k} (n_i - 1) < n - 1,$$

contro l'ipotesi.

Matrici associate a un grafo

Definizione

La matrice di adiacenza associata al grafo $G = (N, A)$ e' una matrice quadrata B di ordine n le cui componenti sono

$$b_{ij} := \begin{cases} 1 & \text{se } (i,j) \in A \\ 0 & \text{altrimenti} \end{cases}$$

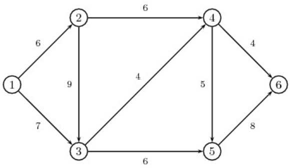
Definizione

La matrice di incidenza associata al grafo (orientato) $G = (N, A)$ e' una matrice E di ordine $n \times m$ le cui componenti sono

$$e_{ij} := \begin{cases} -1 & \text{se } i \text{ e' il nodo iniziale dell'arco } j \\ +1 & \text{se } i \text{ e' il nodo finale dell'arco } j \\ 0 & \text{altrimenti} \end{cases}$$

Nel caso in cui il grafo G non sia orientato si pone $+1$ in luogo di -1 essendo la distinzione tra nodo iniziale e nodo finale di un arco superflua.

Esempio



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

Il problema del flusso di costo minimo

Dato il grafo $G = (N, A)$ siano:

- E la matrice di incidenza di G;
- x il vettore (colonna) di componenti x_{ij}, $(i, j) \in A$, x_{ij} e' detto *flusso* sull'arco $(i, j) \in A$;
- $c := (c_{ij},\ (i, j) \in A)$ il vettore dei costi associati a ciascun arco $(i, j) \in A$;
- $u := (u_{ij},\ (i, j) \in A)$ il vettore delle capacita' superiori associate a ciascun arco $(i, j) \in A$;
- $b := (b_1, .., b_n)$ il vettore dei bilanci associati a ciascun nodo $i \in N$;

Si puo' supporre che l'insieme degli archi $(i, j) \in A$ e di conseguenza il vettore x sia ordinato mediante l'ordinamento lessicografico.

Il problema del flusso di costo minimo sul grafo G e' definito dal seguente problema di programmazione lineare:

$$\begin{cases} \min \ \langle c, x \rangle \\ Ex = b \\ 0 \leq x \leq u \end{cases}$$

Lo scalare b_i stabilisce quale sia la differenza tra il flusso entrante e il flusso uscente, nel nodo i, $i = 1, .., n$.
I nodi tali che $b_i < 0$ vengono detti *sorgenti* mentre quelli tali che $b_i > 0$ sono detti *pozzi*.
Come e' evidente dalla definizione della matrice E, condizione necessaria affinche' la regione ammissibile del problema sia non vuota e' che

$$\sum_{i=1}^{n} b_i = 0. \tag{1}$$

Osserviamo che la condizione (1) non e' restrittiva, nel senso che, nel caso non fosse verificata, e' sempre possibile formulare un problema equivalente a quello dato tale che (1) sia soddisfatta.

Infatti, se fosse $\sum_{i=1}^{n} b_i \neq 0$, si puo' considerare il grafo G' ottenuto aggiungendo a G un nodo fittizio $n+1$ con

$$b_{n+1} := -\sum_{i=1}^{n} b_i$$

e gli archi che connettono il nodo $n+1$ con ogni pozzo, se $b_{n+1} < 0$ o che connettono ogni sorgente con il nodo $n+1$ se $b_{n+1} > 0$, a costo nullo e capacita' $+\infty$.

Il problema di flusso di costo minimo ottenuto definendo G' in luogo di G e' equivalente a quello dato e soddisfa la condizione (1).

Il problema del cammino di costo minimo

Supponiamo che

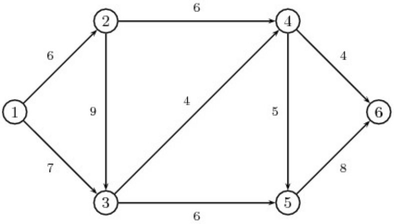
- c_{ij} sia il costo dell'arco (i, j);
- il costo di un cammino sia data dalla somma dei costi degli archi che lo compongono;

Il problema del cammino di costo minimo

Fissata una coppia di nodi i e j si vuole determinare il cammino di costo minimo che connette i e j;

Esempio

Siano $i = 1$, $j = 6$, i nodi tra i quali si vuole determinare il cammino di costo minimo.
$(1, 2), (2, 4), (4, 6)$ e' il cammino di costo minimo.



Supponiamo che

- I nodi 1 ed n siano il nodo iniziale e finale del cammino C cercato;
- Poniamo

$$x_{ij} := \begin{cases} 1 & se\ (i,j) \in C \\ 0 & altrimenti \end{cases}$$

Formulazione del problema

$$\min \sum_{(i,j)\in A} c_{ij} x_{ij}$$

con i vincoli

$$\sum_{(i,j)\in FS(1)} x_{ij} - \sum_{(j,i)\in BS(1)} x_{ji} = 1$$

$$\sum_{(i,j)\in FS(i)} x_{ij} - \sum_{(j,i)\in BS(i)} x_{ji} = 0, \quad i = 2, .., n-1$$

$$\sum_{(j,i)\in BS(n)} x_{ji} - \sum_{(i,j)\in FS(n)} x_{ij} = 1$$

$$x_{ij} \in \{0,1\} \quad (i,j) \in A$$

Formulazione mediante la matrice di incidenza

- E sia la matrice di incidenza del grafo $G = (N, A)$;
- x il vettore (colonna) di componenti x_{ij}, $(i, j) \in A$;
- c il vettore di componenti c_{ij}, $\quad (i, j) \in A$;
- $b = (-1, 0, ..., 0, 1)^T \in \mathbb{R}^n$;

$$\min c^T x$$

con i vincoli

$$\begin{cases} Ex = b \\ x \in \{0,1\}^m \end{cases}$$

Albero dei cammini di costo minimo

In modo analogo, il problema della ricerca dell'albero dei cammini di costo minimo di radice il nodo $r = 1$, ossia dell' insieme dei cammini di costo minimo dal nodo r ad un qualsiasi altro nodo del grafo G, puo' essere formulato mediante il problema

$$\min c^T x$$

con i vincoli

$$\begin{cases} Ex = b \\ x \geq 0 \end{cases}$$

ove $b := (-(n-1), 1, 1, .., 1)^T \in \mathbb{R}^n$.

Osservazione

Si puo' dimostrare che se il problema ammette ottimo finito, allora ammette una soluzione ottima $\bar{x}$ a componenti intere, in particolare $\bar{x}_{ij}$ coincide con il numero di cammini minimi passanti per l'arco (i, j) ed il valore ottimo coincide con la somma dei costi dei cammini minimi.

Algoritmo SPT

Descriviamo l'Algoritmo SPT (*shortest path tree*), dovuto a Dijkstra, per la ricerca dell'albero dei cammini minimi. Definiamo i seguenti elementi:

- $d \in \mathbb{R}^n$, il vettore delle etichette associate ai nodi;
- $p \in \mathbb{R}^n$, il vettore dei predecessori di ciascun nodo;
- $Q \subseteq N$, l'insieme dei nodi candidati ad essere considerati ad ogni passo dell'algoritmo.

Algoritmo SPT (Dijkstra)

1. Si ponga $d(r) = 0$, $d(i) = +\infty$, $\forall i \neq r$ e $Q = \{r\}$.
2. Si selezioni $i \in Q$ e si consideri l'insieme $U(i) := \{j \in N : (i, j) \in A\}$; si ponga $Q = Q \setminus \{i\}$ e, per ogni $j \in U(i)$, se
$$d(i) + c_{ij} < d(j),$$
allora si ponga $d(j) = d(i) + c_{ij}$, $p(j) = i$ e $Q = Q \cup \{j\}$ (potrebbe essere j gia' incluso in Q).
3. Se $Q = \emptyset$ allora l'algoritmo termina, altrimenti si vada al Passo 2.

Diciamo d un vettore di costi ammissibili se $d(r) = 0$ ed esiste un insieme di cammini $\mathcal{P}_{r,i}$, $i \in N, i \neq r$ che connette r con il nodo i tali che $d(i)$ sia uguale al costo del cammino $\mathcal{P}_{r,i}$, $i \in N, i \neq r$.
Vale il seguente risultato:

Teorema 1

$d \in \mathbb{R}^n$ e' il vettore dei minimi costi per un insieme di cammini orientati dal nodo r agli altri nodi del grafo G se e solo se d e' un vettore di costi ammissibili e risulta:

$$d(i) + c_{ij} \geq d(j), \quad \forall (i,j) \in A. \tag{2}$$

Dimostrazione

La condizione (2) e' ovviamente necessaria, infatti, se, per assurdo, fosse

$$d(i) + c_{ij} < d(j),$$

per un arco $(i,j) \in A$, allora il cammino $\mathcal{P}_{r,i}, (i,j)$ (ove $\mathcal{P}_{r,k}$ e' il cammino ottimale dal nodo r al nodo k fornito dall'algoritmo) risulterebbe di costo minore rispetto a $\mathcal{P}_{r,j}$, contro l'ipotesi.

Segue dimostrazione

Viceversa, supponiamo che la condizione (2) sia verificata e consideriamo un generico cammino orientato dal nodo r al nodo $j \neq r$. rappresentiamo tale cammino mediante la successione di nodi $n_1, ..., n_k$ ove $n_1 = r$ e $n_k = j$. Dalla (2) otteniamo le seguenti disuguaglianze:

$$d(n_1) + c_{n_1 n_2} \geq d(n_2)$$

$$d(n_2) + c_{n_2 n_3} \geq d(n_3)$$

$$\vdots$$

$$d(n_{k-1}) + c_{n_{k-1} n_k} \geq d(n_k)$$

Sommando membro a membro le precedenti disuguaglianze, ricordando che $d(n_1) = 0$, otteniamo

$$\sum_{i=1}^{k-1} c_{n_i n_{i+1}} \geq d(n_k) = d(j),$$

ossia il costo di un generico cammino da r a j e' maggiore o uguale di quello fornito dall'algoritmo SPT.

Esistenza della soluzione

Teorema 2

Se il grafo G non possiede cicli orientati di costo negativo ed esiste un vettore $d \in \mathbb{R}^n$ di costi ammissibili, allora il problema della determinazione dell'albero dei cammini minimi ammette soluzione.

Dimostrazione

Consideriamo l'insieme dei cammini orientati dal nodo r ad un fissato nodo $j \neq r$. Tale insieme puo' essere partizionato nei due sottoinsiemi S_1 ed S_2 formati, rispettivamente, dai cammini che non contengono cicli e da quelli che contengono cicli. Dato che ogni ciclo ha costo non negativo, il cammino di costo minimo puo' essere cercato nell'insieme S_1. Essendo S_1 un insieme finito, abbiamo che esiste un cammino di costo minimo che connette r e j. Ripetendo tale ragionamento per ogni $j \neq r$, si ha la tesi.