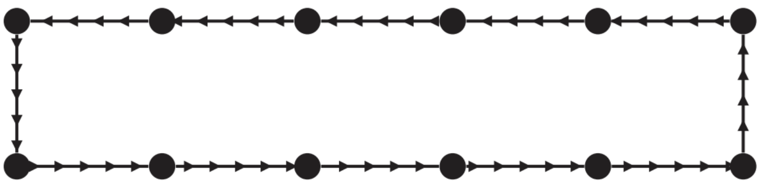
Esercizi sugli algoritmi approssimati

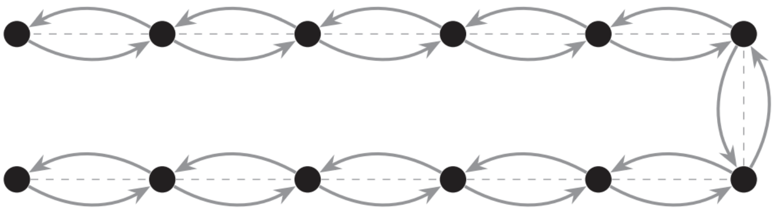
1. Si applichi l'algoritmo primale-duale del set covering all'esempio numerico illustrato nella slide 12 della lezione 2. Che valore di approssimazione ci si aspetta? E' possibile ottenere il valore tight di approssimazione su questa istanza?
   Soluzione:
   Inzialmente poniamo x=0, y=0.
   Siccome inizialmente nessun elemento è coperto, considero ad es. il primo ossia e=1: questo è presente in S2, S3, S4 e S5 che hanno costo 6, 10, 14 e 5 rispettivamente.
   Perciò poniamo y1=5 così S5 diventa un insieme tight (ossia $\sum_{e:e\in S_5} y_e = c(S_5)$). Di conseguenza S5 è scelto per il cover ossia x5=1. Perciò gli elementi sinora coperti sono 1, 4 e 5.
   Preleviamo il successivo elemento non coperto, ossia 2.
   Questo è presente nei set S3 e S4 che hanno costo 10 e 14.
   Per rendere tight S3 occorre $\sum_{e:e\in S_3} y_e = y_1 + y_2 + y_5 = 10 \rightarrow y_2 = 5$, mentre per rendere tight S4 occore che $\sum_{e:e\in S_4} y_e = y_1 + y_2 + y_4 = 14 \rightarrow y_2 = 9$.
   Perciò poniamo y2=5 in modo da rendere S3 tight. S3 è aggiunto al cover e perciò x3=1.
   L'unico elemento non ancora coperto è 3. Gli insiemi che lo contengono sono S1, S2 e S6.
   Per rendere tight S1 occorre che y3+y5=c(1)=4 → y3=4.
   Per rendere tight S2 occorre che y1+y3+y5=c(2)=6 → y3=6-5=1.
   Per rendere tight S6 occorre che y3+y4=c(6)=6 → y3=6.
   Quindi poniamo y3=1 e in questo modo S2 è tight e x2=1.
   Perciò la soluzione finale è il cover {S2, S3, S5} di costo 21.

   L'elemento più frequente è il 5 con f=4.
   Il valore ottimo è z*=16 dato da {S3, S6}.
   Il valore di approssimazione massimo atteso è f=4 ossia un valore ≤ 16*4=64 che è ampiamente rispettato avendo ottenuto una soluzione di valore 21.
   Su questa istanza non è possibile ottenere il valore di approssimazione tight perchè la somma dei costi di tutti i set è 45<f z*=64.

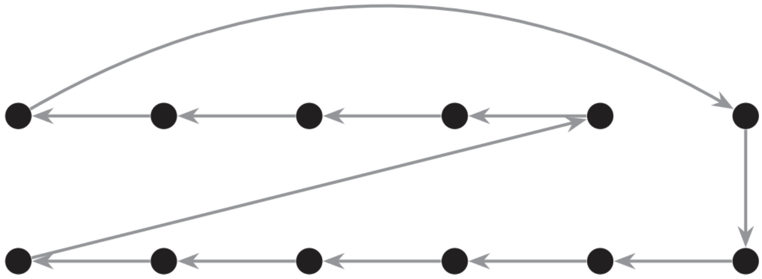
2. Si consideri l'algoritmo del doppio albero per il TSP. Si costruisca un'istanza tight ossia dove l'algoritmo del doppio albero fornisce una soluzione ammissibile che è il doppio del valore ottimo (come previsto nel caso peggiore).
   Soluzione:
   Si consideri una griglia n × 2-grid nel piano Euclideo con lati di lunghezza 1 e distanza Euclidea per tutti gli altri lati (vedere Fgure 3.1a). Un tour ottimo (come mostrato in figura) ha lunghezza 2n (dato che non è possibile un tour più corto che visita tutti i nodi).



(a) TSP instance and optimal solution



(b) minimum spanning tree and Euler tour



(c) solution provided by the double tree algorithm

D’altra parte, si consideri lo spanning tree e il corrispondente tour Euleriano rappresentato in Figura 3.1b. Partendo dal vertice sinistro inferiore e seguendo il tour Euleriano verso il basso si ottiene il ciclo Hamiltoniano di Figura 3.1c.

Per una griglia 2 ×n con distanze euclidee, questo fornisce un ciclo Hamiltoniano di lunghezza.

$$3n - 3 + \sqrt{1 + (n-2)^2} > 4n - 5.$$

Thus, the approximation ratio of this solution is

$$\frac{4n-5}{2n} = 2 - \frac{5}{2n} \to 2 \quad \text{for } n \to \infty.$$

3. Si scriva la formulazione duale del rilassamento lineare del problema del Minimum Weight Node Cover Problem (MWNCP) presentato nella slide 20: che cosa rappresenta tale formulazione? In che modo può consentire di costruire un algoritmo approssimato per il MWNCP?

   Soluzione:

   Il rilassamento lineare del MWNCP è:

$$\min z = \sum_{i=1}^{n} c_i x_i$$

$$x_i + x_j \geq 1 \ \forall [i,j] \in E$$

$$x_i \geq 0 \text{ for } i = 1,..,n$$

si noti che non occorre imporre xi<=1 perchè la f.o. è da minimizzare con coefficienti ci>=0 e i vincoli di copertura non richiedono alle variabili di assumere valori maggiori di 1.

Il duale del suo rilassamento è:

$$\max \sum_{[i,j]\in E} y_{ij}$$

$$\sum_{[i,j]\in E} y_{ij} \le c_i \qquad \forall i = 1, .., n \ \ (1)$$

$$y_{ij} \ge 0 \ \ \forall [i,j] \in E$$

Il significato del duale è di scegliere il maggior numero di lati (anche più volte) con il vincolo che per ogni nodo *i*, il numero di lati scelti (contati con la loro molteplicità) che sono incidenti il nodo *i* non deve superare il costo del nodo *i*.

Si noti che nel caso particolare in cui tutti i coefficienti ci=1 (o sono costanti, perché in tal caso i vincoli (1) si possono semplificare in modo da avere i termini noti unitari) il problema duale rappresenta il matching di cardinalità massima perché i vincoli (1) modellano la condizione che per ogni nodo ci sia al più un lato incidente.

Tornando al caso generale, attraverso il problema duale possiamo ottenere un algoritmo approssimato nel seguente modo. Partiamo dalla soluzione duale $y_{ij} = 0 \ \forall [i,j] \in E$ e definiamo inizialmente il cover C={}. Chiaramente questa soluzione è ammissibile, perchè i vincoli (1) sono soddisfatti. Scegliamo un lato qualunque $[u,v] \in E$ e incrementiamo il più possibile il valore di y_{uv} sinchè il vincolo (1) sarà attivo ossia soddisfatto con l'uguaglianza o per il nodo *u* o per il nodo *v*. Ossia $y_{uv} = y_{uv} + \min\{c_u - \sum_{[u,j]\in E} y_{uj}\, , c_v - \sum_{[v,j]\in E} y_{vj}\}$. Supponiamo che sia u il nodo che rende attivo il vincolo (1), allora aggiungiamo il nodo u al cover C ed eliminiamo $[u,v]$ dall'insieme E dei lati. Andiamo avanti così sinché non ci sono più lati.

Alla fine della procedura C è un vertex cover, perché un lato viene eliminato solo se uno dei suoi nodi estremi è stato inserito in C. Vogliamo ora dimostrare che il costo di C è al massimo il doppio di quello ottimo.

Innanzitutto $\sum_{i\in C} c_i = \sum_{i\in C}\sum_{[i,j]\in E} y_{ij}$ perché ogni volta che un nodo è inserito in C è perché rende attivo un vincolo (1).

Abbiamo quindi:

$$\sum_{i\in C} c_i = \sum_{i\in C}\sum_{[i,j]\in E} y_{ij} = \sum_{[i,j]\in E} |C \cap \{i,j\}| y_{ij} \le \sum_{[i,j]\in E} 2y_{ij}$$

ma siccome y è soluzione ammissibile per il duale e la funzione obiettivo del duale valutata in y è un lower bound al valore ottimo z* del primale, si ha che

$$\sum_{[i,j]\in E} 2y_{ij} \le 2z^*$$

Abbiamo quindi dimostrato che $\sum_{i\in C} c_i \le 2z^*$.