# Il rilassamento Lagrangiano nella soluzione di problemi di programmazione lineare intera

Alessandro Agnetis, Paolo Detti *

January 24, 2012

## 1 La tecnica Lagrangiana

L'applicazione di algoritmi di enumerazione implicita a problemi di programmazione lineare intera (PLI) è basata sulla possibilità di avere, ad ogni nodo dell'albero di enumerazione, un bound sul valore della soluzione ottima. Il metodo più immediato per calcolare tale bound consiste nel risolvere un problema di programmazione lineare (PL) ottenuto rilassando i vincoli di interezza delle variabili. Tuttavia, esistono altri approcci. In questa dispensa vogliamo presentare l'approccio Lagrangiano, che è impiegato con successo in molti algoritmi di soluzione per un numero molto ampio di problemi di PLI.

L'intuizione alla base dell'approccio Lagrangiano consiste nel vedere un problema di programmazione lineare intera "difficile" come un problema "facile", complicato però da un insieme, in genere di dimensione ridotta, di vincoli ulteriori. Il *rilassamento Lagrangiano* consiste nell'eliminazione di tali vincoli "indesiderati" e nell'inserimento di questi nella funzione obiettivo, in modo tale da tenerne comunque conto nella soluzione del problema rilassato.

Si consideri il seguente problema di Programmazione Lineare Intera $PLI_1$:

$$\min c^T x \tag{1}$$
$$Ax \geq b \tag{2}$$
$$Cx \geq d \tag{3}$$
$$x \in Z_+^n \tag{4}$$

dove $A$ e $C$ sono matrici di dimensioni, rispettivamente, $m \times n$ e $m_1 \times n$, $b$ e $d$ sono vettori di dimensioni, rispettivamente, $m$ e $m_1$, mentre $c$ e $x$ sono vettori di dimensione $n$, e $Z_+^n$ denota l'insieme dei vettori a valori interi non negativi. Supponiamo che i vincoli $Ax \geq b$ siano quelli che complicano la risoluzione del problema, nel senso che, mentre $PLI_1$ è un problema di difficile risoluzione, risulta "facile" risolvere il problema ottenuto "inserendo" nella funzione obiettivo i vincoli (2) nel seguente modo:

*Dipartimento di Ingegneria dell'Informazione - Università di Siena, via Roma, 56 - 53100 Siena (Italy), e-mail: detti@dii.unisi.it

$$\begin{aligned} \min c^T x &- \lambda^T(Ax-b) && (5) \\ Cx &\geq d && (6) \\ x &\in Z_+^n && (7) \end{aligned}$$

Nella (5), $\lambda \in R_+^m$ è denominato vettore dei *moltiplicatori di Lagrange*. Il problema (5)–(7) è detto *problema Lagrangiano* e la funzione

$$L(\lambda) = \min\{c^T x - \lambda^T(Ax-b) : Cx \geq d, x \in x \in Z_+^n\} \quad (8)$$

è denominata *funzione Lagrangiana*. Il problema Lagrangiano ha la proprietà che *per ogni* $\lambda \geq 0$ *fissato*, la soluzione ottima costituisce un lower bound sull'ottimo del problema originario (un *upper bound* se il problema originario è di massimo):

Proposizione 1 *Sia $z(PLI_1)$ il valore della soluzione ottima di $PLI_1$. Per ogni vettore non negativo $\lambda \in R_+^m$, si ha*

$$L(\lambda) \leq z(PLI_1) \quad (9)$$

Dim.– Poiché per ogni soluzione ammissibile di $PLI_1$ si ha $Ax \geq b$ e $\lambda \geq 0$, allora $\lambda^T(Ax-b) \geq 0$. Quindi, per ogni fissato vettore di moltiplicatori, si ha:

$$\begin{aligned} z(PLI_1) &= \min\{c^T x : Ax \geq b, Cx \geq d, x \in Z_+^n\} \\ &\geq \min\{c^T x - \lambda^T(Ax-b) : Ax \geq b, Cx \geq d, x \in Z_+^n\} \\ &\geq \min\{c^T x - \lambda^T(Ax-b) : Cx \geq d, x \in Z_+^n\} = L(\lambda) \end{aligned}$$

dove l'ultimo passaggio è dovuto al fatto che, rimuovendo i vincoli $Ax \geq b$, si effettua un allargamento della regione ammissibile del problema. □

Si noti che abbiamo supposto $\lambda \geq 0$ in quanto nel problema originario i vincoli che rendono il problema difficile sono disuguaglianze. Se invece i vincoli da rilassare sono di uguaglianza, il problema originario $(PLI_1')$ è:

$$\begin{aligned} \min c^T x & && (10) \\ Ax &= b && (11) \\ Cx &\geq d && (12) \\ x &\in Z_+^n && (13) \end{aligned}$$

e indicando con $z(PLI_1')$ il valore della soluzione ottima di tale problema, si ha il seguente risultato, del tutto analogo al precedente:

Proposizione 2 *Per ogni vettore $\lambda \in R^m$, si ha*

$$L(\lambda) \leq z(PLI_1') \quad (14)$$

Dim.– Pressoché identica a quella precedente, viene lasciata per esercizio. $\square$

Poiché, per ogni valore del vettore $\lambda$, $L(\lambda)$ costituisce un lower bound sul valore di $z(PLI)$, siamo interessati a cercare quel vettore $\lambda^*$ che dà luogo al migliore lower bound ottenibile con questo approccio. Si tratta quindi di risolvere il seguente problema di ottimizzazione, denominato *duale Lagrangiano*:

$$L(\lambda^*) = \max\{L(\lambda) : \lambda \geq 0\}. \tag{15}$$

Con riferimento al problema $PLI_1$, è interessante andare a valutare quale relazione sussiste tra i bound ottenibili con il rilassamento lineare e con quello Lagrangiano. Indicando con $PL_1$ il rilassamento lineare del problema $PLI_1$, e con $z(PL_1)$ il valore della sua soluzione ottima, vale il seguente risultato.

Teorema 1 *Sia $L(\lambda)$ la funzione Lagrangiana ottenuta da $PLI_1$ effettuando il rilassamento Lagrangiano dei vincoli $Ax \geq b$, e sia $L(\lambda^*)$ il valore ottimo del duale Lagrangiano. Vale la condizione:*

$$L(\lambda^*) \geq z(PL_1) \tag{16}$$

Dim.–

$$\begin{aligned}
L(\lambda^*) &= \max\{\min\{c^T x - \lambda^T(Ax - b) : Cx \geq d, x \in Z^n_+\} : \lambda \geq 0\} \\
&\geq \max\{\min\{c^T x - \lambda^T(Ax - b) : Cx \geq d, x \geq 0\} : \lambda \geq 0\} \\
&= \max\{\min\{(c^T - \lambda^T A)x + \lambda^T b : Cx \geq d, x \geq 0\} : \lambda \geq 0\} \\
&\geq \max\{\min\{(c^T - \lambda^T A)x - \mu^T(Cx - d) + \lambda^T b : x \geq 0\} : \lambda, \mu \geq 0\} \\
&= \max\{\min\{(c^T - \lambda^T A - \mu^T C)x + \mu^T d + \lambda^T b : x \geq 0\} : \lambda, \mu \geq 0\}
\end{aligned} \tag{17}$$

ora, nel problema di minimo in (17), $x$ non è soggetto ad altri vincoli oltre a quelli di non negatività. Dunque, se $\lambda$ e $\mu$ sono tali che almeno una componente, e.g. la $j$-esima, del vettore $(c^T - \lambda^T A - \mu^T C)$ risulta negativa, il minimo della funzione obiettivo più interna sarebbe $-\infty$ (per $x_j \to +\infty$), e dunque certamente il max non si ottiene per quei valori di $\lambda$ e $\mu$. Quindi, gli unici valori di $\lambda$ e $\mu$ per i quali ha senso calcolare il minimo di $(c^T - \lambda^T A - \mu^T C)x + \mu^T d + \lambda^T b$ sono quelli per cui $c^T - \lambda^T A - \mu^T C \geq 0$. Inoltre, in tal caso il problema di minimo è banalmente risolto ponendo $x = 0$, con il che la funzione obiettivo vale $\mu^T d + \lambda^T b$. In definitiva, si ha:

$$= \max\{\mu^T d + \lambda^T b : c^T - \lambda^T A - \mu^T C \geq 0 : \lambda, \mu \geq 0\}$$

quest'ultimo è un problema di programmazione lineare. Scrivendone il duale, per la dualità forte si ha:

$$= \min\{c^T x : Ax \geq b, Cx \geq d, x \geq 0\} = z(PL_1).$$

$\square$

La Proposizione 1 stabilisce, quindi, che il bound ottenibile con il rilassamento Lagrangiano è almeno tanto buono quanto quello ottenibile con il rilassamento lineare, e giustifica quindi, almeno in linea teorica, l'impiego dei bound Lagrangiani nella messa a punto di algoritmi di enumerazione implicita di tipo branch and bound. L'efficacia pratica del rilassamento Lagrangiano è legata ovviamente alla efficienza con cui si riesce a risolvere il duale Lagrangiano.

# 2 Condizioni di ottimalità per il problema Lagrangiano

È noto che se, andando a risolvere il rilassamento lineare di un problema intero, la soluzione ottima risulta intera, questa è anche ammissibile per il problema intero e quindi ottima. Ci si può chiedere se un'analoga situazione si può presentare nel caso del rilassamento Lagrangiano.

Se, per un dato valore di $\lambda$, la soluzione del problema Lagrangiano è ammissibile anche per il problema originario (se, cioè, rispetta i vincoli rilassati (2)) non si può affermare che tale soluzione sia ottima per $PLI_1$, in quanto le funzioni obiettivo dei due problemi sono diverse. Vale però la seguente condizione.

Proposizione 3 *Sia $L(\lambda)$ la funzione Lagrangiana ottenuta da $PLI_1$ effettuando il rilassamento Lagrangiano dei vincoli $Ax \geq b$. Se, per un fissato valore $\lambda$ dei moltiplicatori, si ha che:*

*i. La soluzione ottima $x^*(\lambda)$ del problema Lagrangiano è ammissibile per $PLI_1$;*

*ii. $\lambda$ e $x^*(\lambda)$ verificano la condizione di complementarità*

$$\lambda^T(Ax^*(\lambda) - b) = 0 \tag{18}$$

*allora $x^*(\lambda)$ è una soluzione ottima anche per il problema $PLI_1$.*

Dim.– Per definizione, $L(\lambda) = \min\{c^T x - \lambda^T(Ax - b)\}$. Dalla condizione $ii$, si ha che $L(\lambda) = c^T x^*(\lambda)$. Inoltre poiché $L(\lambda) \leq c^T x$ per ogni soluzione ammissibile $x$ di $PLI_1$, e poiché $x^*(\lambda)$ è ammissibile per $PLI_1$, si deduce che $x^*(\lambda)$ è una soluzione ottima per $PLI_1$. $\square$

La proposizione precedente vale, in generale, per il rilassamento Lagrangiano di vincoli di disuguaglianza. Se per un dato problema si effettua, invece, il rilassamento Lagrangiano di vincoli di uguaglianza (problema $PLI_1'$), poiché la condizione $ii$ è sempre soddisfatta, condizione sufficiente di ottimalità per il problema originario è che la soluzione $x^*(\lambda)$ del problema Lagrangiano sia ammissibile per $PLI_1'$.

# 3 Soluzione del duale Lagrangiano

Si consideri il problema Lagrangiano (8), e sia $X$ l'insieme (supponiamo finito) di punti interi che soddisfano i vincoli del problema, ossia $X \equiv \{x^1, \ldots, x^K\} \equiv \{Cx \geq d, x \in Z_+^n\}$. Per un dato $\lambda$, in ciascun punto $x^k$ il valore della funzione lagrangiana è evidentemente dato da $(c^T - \lambda^T A)x^k + \lambda^T b$, e quindi $L(\lambda)$ sarà dato dal più piccolo di tali valori, vale a dire:

$$L(\lambda) = \max \quad v \tag{19}$$
$$v \leq (c^T - \lambda^T A)x^k + \lambda^T b, \quad k = 1, \ldots, K \tag{20}$$
$$\lambda \geq 0 \tag{21}$$

Dunque, la funzione Lagrangiana $L : \lambda \rightarrow L(\lambda)$ è definita dalla *frontiera inferiore* di un insieme composto da un insieme finito di funzioni lineari (in $\lambda$), e risulta quindi continua, lineare a tratti e concava. A scopo illustrativo, la figura 3 mostra l'andamento di $L(\lambda)$ per il caso monodimensionale ($m = 1$), in cui quindi vi è un solo vincolo $a^T x \geq b$, e $\lambda$ è scalare. In questo esempio, $X$ contiene 5 punti $(x^1, \ldots, x^5)$.

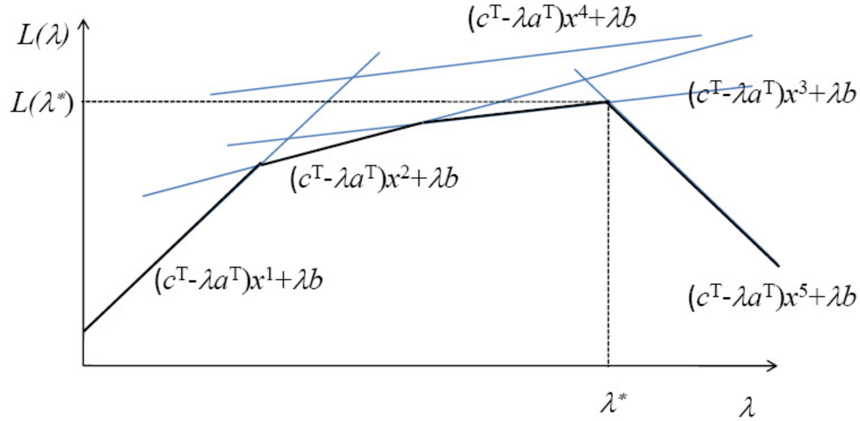


Figura 1: Andamento di $L(\lambda)$.

Consideriamo allora il problema di calcolare il massimo della funzione $L(\lambda)$. Si noti che $L(\lambda)$ è una funzione continua ma *non differenziabile*. Per risolvere tale problema, l'approccio più utilizzato è un metodo molto usato per funzioni continue ma non differenziabili, noto come *metodo del subgradiente*.

Com'è noto, se $f(\lambda)$ è una funzione concava differenziabile in $\lambda$, e si indica con $\nabla f(\lambda)$ il suo gradiente, per qualsiasi altro punto $\bar{\lambda}$ vale la

$$f(\bar{\lambda}) - f(\lambda) \leq \nabla f(\lambda)^T (\bar{\lambda} - \lambda) \tag{22}$$

Se invece $f$ non è differenziabile in $\lambda$, tale proprietà evidentemente non sussiste più. Data quindi una $f(\lambda)$ concava, e un punto $\lambda \in R^m$ in cui la $f$ non è differenziabile, un *subgradiente di $f$ in $\lambda$* è un vettore $s \in R^m$ tale che per qualsiasi altro punto $\bar{\lambda} \in R^n$ si ha

$$f(\bar{\lambda}) - f(\lambda) \leq s^T (\bar{\lambda} - \lambda) \tag{23}$$

Si noti che vi possono essere molti vettori $s$ che soddisfano la (23). Se $f$ fosse differenziabile in $\lambda$, il suo unico subgradiente in $\lambda$ sarebbe il gradiente $\nabla f(\lambda)$. Dalla definizione di subgradiente si può peraltro dedurre che $\lambda^*$ è un punto di massimo per la funzione $f(\lambda)$ se $\mathbf{0}$ è un subgradiente di $f$ in $\lambda^*$. Dalla (22) segue infatti che $f(\bar{\lambda}) - f(\lambda^*) \leq \mathbf{0}^T (\bar{\lambda} - \lambda^*)$ per tutti i $\bar{\lambda} \in R^n$, e quindi $f(\lambda^*) \geq f(\bar{\lambda})$ per qualsiasi altro $\bar{\lambda} \in R^n$.

Il *metodo del subgradiente* per il calcolo di $\lambda^*$ consiste nel generare una successione di valori $\lambda^{(0)}, \lambda^{(1)}, \ldots, \lambda^{(h)}, \ldots$ convergente a $\lambda^*$, in cui

$$\lambda^{(h+1)} = \lambda^{(h)} + \theta^{(h)} d^{(h)} \tag{24}$$

dove $d^{(h)}$ è un subgradiente di $L(\lambda)$ in $\lambda^{(h)}$, e $\theta^{(h)}$ un passo opportunamente specificato. Si può dimostrare che questo semplice algoritmo converge a $\lambda^*$, purché il passo sia scelto con una certa attenzione.

## 3.1 Scelta del subgradiente

La seguente proposizione determina un subgradiente per la funzione Lagrangiana (8).

TEOREMA 2 *Siano* $\lambda^{(h)} \in R^m$ *e* $x^{(h)} \in Z_+^n$ *due vettori tali che*

$$\begin{aligned} L(\lambda^{(h)}) &= \min\{c^T x - \lambda^{(h)^T}(Ax - b) : Cx \geq d, x \in \{0,1\}^n\} = && (25)\\ &= c^T x^{(h)} - \lambda^{(h)^T}(Ax^{(h)} - b) && (26) \end{aligned}$$

*allora il vettore* $(b - Ax^{(h)})$ *è un subgradiente di* $L(\lambda)$ *in* $\lambda^{(h)}$.

Dim.– Per un generico vettore $\lambda$ si ha

$$\begin{aligned} L(\lambda) &= \min\{c^T x - \lambda^T(Ax - b) : Cx \geq d, x \in \{0,1\}^n\} && (27)\\ &\leq c^T x^{(h)} - \lambda^T(Ax^{(h)} - b) = && (28)\\ &= c^T x^{(h)} - \lambda^{(h)^T}(Ax^{(h)} - b) - (\lambda - \lambda^{(h)})^T(Ax^{(h)} - b) = && (29)\\ &= L(\lambda^{(h)}) + (\lambda - \lambda^{(h)})^T(b - Ax^{(h)}) && (30) \end{aligned}$$

cioè

$$(\lambda - \lambda^{(h)})^T(b - Ax^{(h)}) \geq L(\lambda) - L(\lambda^{(h)})$$

Ricordando la definizione di subgradiente (23), si ha dunque che $(b - Ax^{(h)})$ è un subgradiente di $L(\lambda)$ in $\lambda^{(h)}$. □

Questo teorema mostra dunque che un subgradiente è dato dal valore dei (termini di sinistra dei) vincoli che abbiamo rilassato, ed è quindi calcolabile molto facilmente.

## 3.2 Scelta del passo

La difficoltà principale nella messa a punto del metodo del subgradiente riguarda la scelta del *passo* $\theta^{(h)}$. Il modo con cui tale scelta è realizzata determina la convergenza o meno della sequenza di punti $\lambda^{(h)}$ ad un punto di massimo $\lambda^*$, e, nel caso in cui si abbia tale convergenza, ne influenza notevolmente la velocità. In particolare, si può dimostatre [2] che se la sequenza $\theta^{(h)}$ soddisfa le condizioni:

$$\begin{aligned} \lim_{t\to\infty} \theta^{(h)} &= 0 && (31)\\ \sum_{h=0}^{\infty} \theta^{(h)} &= +\infty && (32) \end{aligned}$$

allora la successione di punti $\{\lambda^{(h)}\}$ generata scegliendo a ogni passo la direzione $s^{(h)}/||s^{(h)}||$, converge a $\lambda^*$. In pratica, si richiede che la lunghezza del passo tenda a zero (31), ma non troppo rapidamente (32).

**Algoritmo del subgradiente**

**Input:** funzione da massimizzare $L(\lambda)$, punto iniziale $\lambda^0$, massimo numero di iterazioni senza miglioramento $K$

1. Inizializzazione $h := 1$.
2. Per un dato valore $\lambda^{(h)}$, calcola la soluzione $x^{(h)}$ del problema Lagrangiano (5)–(7).
3. $s^{(h)} := b - Ax^{(h)}$. Se $s^{(h)} = 0$, STOP
4. $\theta^{(h)} := \frac{B-L(\lambda^{(h-1)})}{||s^{(h)}||^2}$
5. $\lambda^{(h+1)} := \lambda^{(h)} + \theta^{(h)} \frac{s^{(h)}}{||s^{(h)}||}$
6. $h := h + 1$
7. Se il lower bound non è migliorato nelle ultime $K$ iterazioni, STOP, altrimenti vai al passo 2.

Figura 2: Algoritmo del subgradiente

In pratica, dato un subgradiente $s^{(h)}$, si può dimostrare che una scelta della lunghezza del passo $\theta^{(h)}$ che soddisfa le condizioni (31) e (32) è dato dalla relazione [2, 1][1]

$$\theta^{(h)} = \frac{B - L(\lambda^{(h)})}{||s^{(h)}||^2} \tag{33}$$

dove per $B$ si può scegliere il valore di una soluzione ammissibile del problema originario (meglio se di buona qualità).In Figura 2 è riportato un possibile schema dell'algoritmo del subgradiente. L'algoritmo del subgradiente è garantito convergere al vettore ottimo dei moltiplicatori di Lagrange $\lambda^*$. I maggiori incovenienti del metodo sono il comportamento oscillatorio dei lower bound generati (la serie di lower bound generati non è, infatti, monotonicamente crescente) e la bassa velocità di convergenza (lineare). Tuttavia, per la sua facilità di implementazione e per la buona qualità dei bound prodotti (anche in termini di velocità di calcolo della singola iterazione), tale metodo è uno dei più utilizzati.

[1] Una formula più generale della (33) è

$$\theta^{(h)} = \alpha_h \frac{B - L(\lambda^{(h)})}{||s^{(h)}||^2}$$

dove $\alpha_h$ è uno scalare che nelle prime iterazioni vale 2, e poi viene fatto decrescere. Questo rende la convergenza tipicamente più veloce.

# 4 Esempio: il problema del commesso viaggiatore simmetrico

Vediamo ora una delle applicazioni più efficaci della tecnica lagrangiana: quella alla soluzione del problema del commesso viaggiatore simmetrico (Symmetric Traveling Salesperson Problem, STSP). Ricordiamo che il problema è il seguente: dato un grafo completo $G(V,E)$, per ogni arco $e \in E$ è definito un costo $c_e$ per andare da un estremo all'altro di tale arco. Il problema consiste nel cercare il ciclo hamiltoniano (ossia, il ciclo passante esattamente una volta per ciascun nodo) di costo totale minimo. Nella formulazione che segue, indichiamo con $n = |V|$ il numero di nodi, $S$ un sottoinsieme di nodi, $E(S)$ l'insieme degli archi con entrambi gli estremi in $S$, e con $\delta(S)$ l'insieme degli archi aventi un estremo in $S$ e l'altro in $V \setminus S$ (quando $S$ consiste di un solo nodo $i$, si ha $\delta(i)$, ossia la *stella incidente* il nodo $i$). Il problema può essere formulato come segue:

$$z = \min \sum_{e \in E} c_e x_e \tag{34}$$

$$\sum_{e \in \delta(i)} x_e = 2, \quad i \in V \tag{35}$$

$$\sum_{e \in E(S)} x_e \leq |S| - 1, \quad 2 \leq |S| \leq n-1 \tag{36}$$

$$x \in \{0,1\}^{|E|} \tag{37}$$

dove i vincoli (35) indicano che ciascun nodo deve essere incidente esattamente due archi del ciclo, mentre i vincoli (36) sono i *subtour elimination constraints*: ciascun sottografo con $S$ nodi non può contenere più di $|S| - 1$ archi del ciclo, e dunque non si possono formare sottocicli in $S$.

È importante fare due osservazioni. Anzitutto, bastano i vincoli (35) a garantire il fatto che il ciclo conterrà esattamente $n$ archi: infatti, sommando le (35) per $i = 1, \ldots, n$, considerando che ciascun arco $\{i,j\}$ appartiene a due sole stelle incidenti, ossia $\delta(i)$ e $\delta(j)$, si ottiene

$$\sum_{i=1}^{n} \sum_{e \in \delta(i)} x_e = 2 \sum_{e \in E} x_e = 2n \tag{38}$$

Un'altra osservazione riguarda il numero di vincoli (36). Tali vincoli sono chiaramente $O(2^n)$. Tuttavia, la metà di essi è superflua. Infatti, consideriamo un nodo arbitrario, sia esso il nodo 1. Chiaramente, metà dei possibili sottoinsiemi di nodi contengono 1 e l'altra metà non lo contengono. Se scriviamo i vincoli (36) soltanto per i sottoinsiemi $S$ che *non* contengono 1, riusciamo ancora a impedire la formazione dei sottocicli. Infatti, immaginiamo di avere una soluzione ammissibile per il problema (34–37) che rispetti solo questi vincoli, e supponiamo per assurdo che essa non sia un ciclo hamiltoniano. Se non è un ciclo hamiltoniano, per via dei vincoli (35) è comunque una collezione di almeno 2 sottocicli, di cui quindi almeno uno *non* contiene il nodo 1: ma allora il corrispondente subtour elimination constraint risulterebbe violato, un assurdo.

Dunque, possiamo riscrivere la formulazione del problema di STSP in questo modo:

$$z = \min \sum_{e \in E} c_e x_e \tag{39}$$

$$
\begin{aligned}
\sum_{e \in \delta(1)} x_e &= 2 & (40)\\
\sum_{e \in \delta(i)} x_e &= 2, \quad i \in V \setminus \{1\} & (41)\\
\sum_{e \in E(S)} x_e &\leq |S| - 1, \quad 2 \leq |S| \leq n-1, 1 \notin S & (42)\\
\sum_{e \in E} x_e &= n & (43)\\
x &\in \{0,1\}^{|E|} & (44)
\end{aligned}
$$

Si noti che il vincolo (43) non sarebbe necessario, ma il suo ruolo sarà chiaro tra poco.

Prima di applicare il rilassamento Lagrangiano al problema (39–44), introduciamo una definizione. Si consideri un grafo non orientato $G(V, E)$, con i nodi numerati 1, 2,..,$n$. Definiamo *1-albero* un sottoinsieme $T$ di archi avente le seguenti proprietà:

- $T$ contiene $n$ archi
- esattamente due archi di $T$ incidono il nodo 1, siano essi $e$ ed $f$
- L'insieme $T \setminus \{e, f\}$ non contiene cicli.

In altre parole, un 1-albero è un albero ricoprente per il sottografo costituito dai nodi $\{2, 3, \ldots, n\}$, cui si aggiungono due archi di $\delta(1)$ (questi archi quindi necessariamente formeranno un ciclo con gli altri). Se gli archi del grafo sono pesati, si può considerare il problema di determinare l'1-albero di peso minimo. Questo problema può risolversi facilmente applicando un qualsiasi algoritmo per il calcolo dell'albero di costo minimo sul sottografo costituito dai nodi $\{2, 3, \ldots, n\}$, e aggiungendo poi semplicemente i due archi meno costosi incidenti il nodo 1.

Tornando alla formulazione (39–43), rilassiamo gli $n-1$ vincoli (41), ottenendo così il problema Lagrangiano:

$$
\begin{aligned}
L(\lambda) = \min \quad & \sum_{e \in E} c_e x_e - \sum_{i=2}^{n} \lambda_i \Big( \sum_{e \in \delta(i)} x_e - 2 \Big) & (45)\\
& \sum_{e \in \delta(1)} x_e = 2 & (46)\\
& \sum_{e \in E(S)} x_e \leq |S| - 1, \qquad 2 \leq |S| \leq n-1, 1 \notin S & (47)\\
& \sum_{e \in E} x_e = n & (48)\\
& x \in \{0,1\}^{|E|} & (49)
\end{aligned}
$$

è facile rendersi conto che le soluzioni ammissibili per questo problema sono precisamente tutti gli 1-alberi del grafo. Dunque, in questo caso il problema Lagrangiano consiste nel determinare l'1-albero di costo minimo, laddove, dalla (45), il peso di un arco $e = \{i, j\}$ è dato da

$$c_e - \lambda_i - \lambda_j. \tag{50}$$

Dato $\bar{\lambda}$, se $\bar{x}$ è la soluzione ottima del corrispondente problema Lagrangiano, in base al Teorema 2 come subgradiente $s$ possiamo prendere un vettore la cui $i$-esima componente è data da

$$s_i = 2 - \sum_{e \in \delta(i)} \bar{x}_e \tag{51}$$

dove, si noti, $\sum_{e \in \delta(i)} \bar{x}_e$ è il numero di archi dell'1-albero incidenti il nodo $i$ ($i = 2, \ldots, n$). Si noti che

$$||s||^2 = \sum_{i=2}^{n} \left( \sum_{e \in \delta(i)} \bar{x}_e - 2 \right)^2 \tag{52}$$

Esempio 1 *(tratto da [3]). Si consideri un grafo $G = (N, E)$ costituito da 5 nodi, e con matrice dei costi data da*

$$\begin{pmatrix} - & 30 & 26 & 50 & 40 \\ 30 & - & 24 & 40 & 50 \\ 26 & 24 & - & 24 & 26 \\ 50 & 40 & 24 & - & 30 \\ 40 & 50 & 26 & 30 & - \end{pmatrix}$$

*Una soluzione (trovata euristicamente, ad esempio) è quella mostrata in figura 3, di valore 148.*

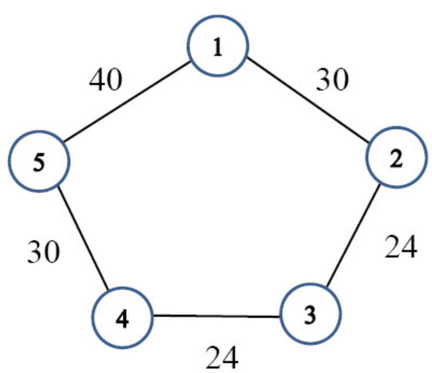


Figura 3: Un ciclo hamiltoniano di costo 148.

*Consideriamo ora, come vettore iniziale di moltiplicatori, $\lambda^{(0)^T} = (0\ 0\ 0\ 0)$. (Poiché il vincolo relativo al grado del nodo 1 non viene rilassato, per mantenere la corrispondenza coi nodi numeriamo i moltiplicatori da 2 a 5.) In tal caso, si noti, il problema Lagrangiano equivale al problema originario dal quale sono cancellati i vincoli (41). Risolvendo il problema Lagrangiano, si ottiene l'1-albero in figura 4, di costo pari a 130. Dunque, questo è un lower bound per il valore della soluzione ottima del problema (34–37).*

*Ora, applichiamo il metodo del subgradiente per aggiornare il vettore dei moltiplicatori. Poniamo $B = 148$, valore della soluzione nota. Dalla (51) e dalla figura 4, si vede che il*

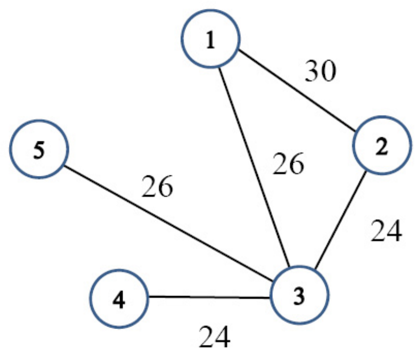


Figura 4: Soluzione del problema Lagrangiano con ${\lambda^{(0)}}^T = (0\ 0\ 0\ 0)$.

*subgradiente è dato da* $s^{(1)} = (0\ -2\ 1\ 1)^T$*, e dunque* $||s^{(1)}||^2 = 0 + 4 + 1 + 1 = 6$*. Dalla (24) si ha quindi*

$$\lambda^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \left(\frac{148 - 130}{6}\right) \begin{pmatrix} 0 \\ 2 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -6 \\ 3 \\ 3 \end{pmatrix}$$

*Si tratta ora di aggiornare i costi nella funzione Lagrangiana, tramite la (50). Si ottiene la nuova matrice*

$$\begin{pmatrix} - & 30 & 32 & 47 & 37 \\ 30 & - & 30 & 37 & 47 \\ 32 & 30 & - & 27 & 29 \\ 47 & 37 & 27 & - & 24 \\ 37 & 47 & 29 & 24 & - \end{pmatrix}$$

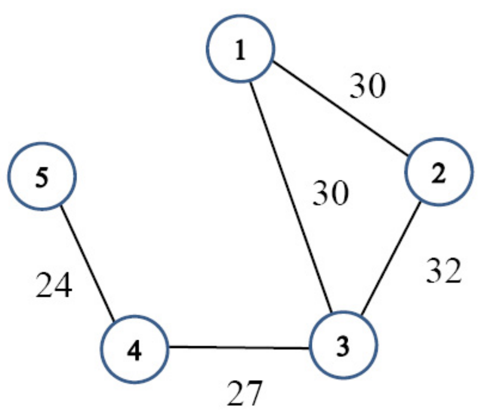


Figura 5: Soluzione del problema Lagrangiano con ${\lambda^{(1)}}^T = (0\ -6\ 3\ 3)$.

*Da questa possiamo calcolare dunque la soluzione ottima del nuovo problema Lagrangiano, ossia $L(\lambda^{(1)})$, in figura 5. Il valore di $L(\lambda^{(1)})$ è dato da $143+2\sum_{i=2}^{5}\lambda_i^{(1)} = 143$. Calcoliamo ora il nuovo vettore di moltiplicatori, tenendo presente che nel nuovo 1-albero i gradi dei nodi 2,3,4 e 5 sono dati rispettivamente da 2,3,2,1 e dunque il nuovo subgradiente è $s^{(2)} = (0\ -1\ 0\ 1)$ e $||s^{(2)}||^2 = 1+1 = 2$:*

$$\lambda^{(2)} = \begin{pmatrix} 0 \\ -6 \\ 3 \\ 3 \end{pmatrix} + \left(\frac{148-143}{2}\right) \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -17/2 \\ 3 \\ 11/2 \end{pmatrix}$$

*La nuova matrice dei costi diviene*

$$\begin{pmatrix} - & 30 & 34.5 & 47 & 34.5 \\ 30 & - & 32.5 & 37 & 44.5 \\ 34.5 & 32.5 & - & 29.5 & 29 \\ 47 & 37 & 29.5 & - & 21.5 \\ 34.5 & 44.5 & 29 & 21.5 & - \end{pmatrix}$$

*e la nuova soluzione ottima è indicata in figura 6. Si ottiene così $L(\lambda^{(2)}) = 147.5$, che, siccome tutti i dati sono interi, implica un lower bound di 148, il che dimostra che la soluzione iniziale di cui disponevamo era in realtà ottima.*

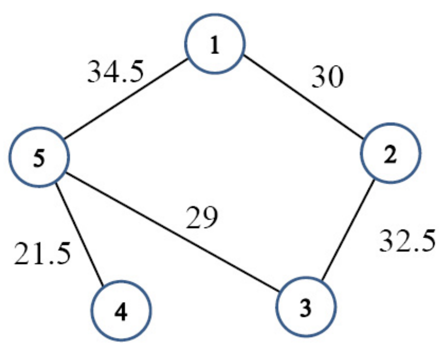


Figura 6: Soluzione del problema Lagrangiano con ${\lambda^{(2)}}^T = (0\ -17/2\ 3\ 11/2)$.

# References

[1] M. Held, R. M. Karp, *The traveling salesman problem and minimun spanning trees: Part II*, Mathematical Programming, 1, 62–88, 1971.

[2] B. T. Polyak, *Minimization of Unsmooth Functionals*, Computational Mathematics and Mathematical Physics, 9, 14–29, 1969.

[3] L. A. Wolsey, *Integer Programming*, Wiley, 1998.