Sull'algoritmo di ascesa duale per il problema della localizzazione di impianti

A. Agnetis*

In queste note presentiamo l'algoritmo di ascesa duale per la generazione di lower bound di buona qualità per il problema di plant location. Tale esposizione seguirà la stessa impostazione e le stesse notazioni del testo di Sassano (*Modelli e Algoritmi della Ricerca Operativa*, Franco Angeli, 2004, §6.3), ma sarà leggermente semplificata.

Il problema di plant location (senza capacità) è il seguente. Sono dati un insieme U di possibili impianti, e un insieme V di clienti da servire. Attivare l'impianto $u \in U$ costa f_u, connettere il cliente $v \in V$ ad u costa c_{vu}. Ciascun cliente deve essere connesso a un impianto attivato. Il problema consiste nel decidere l'insieme ottimale di impianti da attivare, nonché a quale impianto connettere ciascun cliente. È facile vedere comunque che, una volta decisi gli impianti da attivare, decidere l'assegnamento dei clienti agli impianti è banale, in quanto il cliente v sarà assegnato all'impianto u, tra quelli attivati, per cui è minimo il costo di assegnamento. Introducendo variabili binarie x_u pari a 1 nel caso l'impianto u sia attivato, e y_{vu} pari a 1 nel caso in cui il cliente v sia assegnato all'impianto u, si ha la formulazione:

$$\begin{array}{rcll}
\min\ z = \sum_{u\in U}\sum_{v\in V} c_{vu}y_{vu} + \sum_{u\in U} f_u x_u & & & \\
\sum_{v\in V} y_{vu} & = & 1 & v \in V \\
x_u - y_{vu} & \geq & 0 & u \in U, v \in V \\
x_u \in \{0,1\}, y_{vu} \in \{0,1\} & & & u \in U, v \in V
\end{array} \tag{1}$$

il cui rilassamento lineare è:

$$\min\ z = \sum_{u\in U}\sum_{v\in V} c_{vu}y_{vu} + \sum_{u\in U} f_u x_u \tag{2}$$

$$\sum_{v\in V} y_{vu} = 1 \quad v \in V \tag{3}$$

*Dipartimento di Ingegneria dell'Informazione - Università di Siena

$$x_u - y_{vu} \geq 0 \quad u \in U, v \in V \tag{4}$$
$$x_u \geq 0, \quad y_{vu} \geq 0 \quad u \in U, v \in V \tag{5}$$

si noti che nel rilassamento lineare non abbiamo incluso alcun vincolo di upper bound sulle variabili ($x_u \leq 1$, $y_{vu} \leq 1$), in quanto implicati. Infatti, i vincoli (3) obbligano ciascuna variabile di attivazione y_{vu} a essere minore o uguale a 1, e inoltre non vi è alcun motivo perché, in una soluzione ottima del rilassamento, una variabile x_u debba assumere un valore superiore a 1 (assumendo ovviamente $f_u > 0$).

Il problema di plant location è NP-completo, per cui è di interesse costruire algoritmi efficienti di enumerazione implicita. La formulazione (1) è anche nota come formulazione *forte*, e il suo rilassamento consente tipicamente di ottenere buoni lower bound, ma al prezzo di uno sforzo computazionale piuttosto elevato (è necessario fare uso dell'algoritmo del simplesso dinamico), spesso incompatibile con il suo utilizzo in un algoritmo di branch and bound. D'altro canto, la formulazione che si ottiene sostituendo gli $|U||V|$ vincoli (1) con gli $|U|$ vincoli

$$\sum_{v \in V} y_{vu} \leq |V| x_u \quad u \in U$$

è detta formulazione *debole*, e ha la proprietà che il suo rilassamento è risolto in modo molto efficiente in forma chiusa, ma il lower bound è molto scadente. L'algoritmo DUALOC di ascesa duale è un algoritmo euristico che consente di produrre un lower bound sul valore ottimo del problema rilassato (2)–(5) in modo molto efficiente, e di qualità accettabile. In altre parole, lo scopo è di rinunciare in parte alla qualità che si otterrebbe risolvendo all'ottimo il problema (2)–(5), in cambio di una sostanziale riduzione dei tempi di calcolo.

L'idea di fondo è di applicare il teorema della dualità debole: il valore della funzione obiettivo di una qualunque soluzione ammissibile per il problema duale è un lower bound per il valore della soluzione ottima del problema primale. Dunque, l'idea è di riuscire, in modo efficiente, a produrre una soluzione ammissibile per il problema duale di qualità "buona".

Scriviamo dunque il duale di (2)–(5):

$$\max \; g(z) = \sum_{v \in V} z_v \tag{6}$$
$$z_v - w_{vu} \leq c_{vu} \quad u \in U, v \in V \tag{7}$$
$$\sum_{v \in V} w_{vu} \leq f_u \quad u \in U \tag{8}$$
$$w_{vu} \geq 0 \quad u \in U, v \in V \tag{9}$$

L'osservazione fondamentale è la seguente. Supponiamo di fissare *a priori* i valori delle variabili z_v. Chiaramente, non è detto che vi sia soluzione ammissibile per quei valori z_v, in quanto devono poi essere soddisfatti i vincoli (7)–(9). Questo però può essere facilmente verificato. Infatti, dalle (8) si vede che conviene far assumere a ciascun variabile w_{vu} il valore più basso possibile che rispetti gli altri vincoli (7) e (9), e dunque si ha che, dato un vettore $\bar{z}$, dev'essere

$$\bar{w}_{vu} = \max\{0, \bar{z}_v - c_{vu}\} \tag{10}$$

Come detto, scegliendo $\bar{z}$ a caso, non si ha la garanzia che i valori $\bar{w}_{vu}$ calcolati tramite le (10) siano poi tali da soddisfare le (8), e dunque potrebbe non esserci soluzione duale ammissibile. A questo proposito, osserviamo però che una scelta che dà sicuramente luogo a una soluzione ammissibile è quella di porre

$$\bar{z}_v = \min_{u \in U}\{c_{vu}\}, \qquad v \in V \tag{11}$$

infatti, dalla (10) si verifica immediatamente che a questa scelta corrisponde $\bar{w}_{vu} = 0$ per tutti i v, u (e dunque le (8) sono sicuramente soddisfatte). D'altro canto, dalla (7) si vede che il valore della soluzione duale – che noi desideriamo sia il più alto possibile – dipende esclusivamente dalla somma dei valori $\bar{z}_v$. Come ora vedremo, l'algoritmo DUALOC *migliora* in modo euristico la soluzione di partenza (11), aumentando le singole componenti del vettore z_v, senza bisogno di effettuare conti particolarmente complessi.

L'idea di fondo dell'algoritmo DUALOC è la seguente. Partendo dalla soluzione data dalle (11), l'algoritmo fissa un ordinamento delle variabili duali z_v, dopo di che, seguendo questo ordinamento, si procede ad aumentare ciascuna variabile (senza toccare le altre) della *massima quantità* che consente di mantenere l'ammissibilità duale della soluzione. Alla fine, avremo una soluzione duale, e dunque un lower bound, di qualità senz'altro superiore a quella di partenza.

Tale massima quantità può essere facilmente calcolata in forma chiusa come segue. Dato un vettore di partenza $\bar{z}_v$, e i corrispondenti valori $\bar{w}_{vu}$, consideriamo una variabile z_s, e poniamoci il problema di determinare il massimo valore z_s^{max} che essa può assumere, ferme restando le altre $z_v = \bar{z}_v$, senza perdere l'ammissibilità della soluzione. Dalle (8), si ha che per ogni impianto $u \in U$, il nuovo vettore w_{su} deve soddisfare:

$$w_{su} \le f_u - \sum_{v \neq s} \bar{w}_{vu} \tag{12}$$

di conseguenza, tenendo presente la (10), avremo che deve essere, per ogni $u \in U$:

$$z_s^{\max} \le c_{su} + f_u - \sum_{v \neq s} \bar{w}_{vu}, \qquad u \in U \tag{13}$$

e dunque

$$z_s^{\max} = \min_{u \in U} \{c_{su} + f_u - \sum_{v \neq s} \bar{w}_{vu}\} \tag{14}$$

Una volta aggiornato il valore di z_s, questo risulterà fissato e si passerà a tentare di aumentare (in quanto potrebbe anche non essere possibile alcun incremento) la successiva variabile, in base all'ordinamento prescelto. Nei successivi conti, chiaramente, sarà utilizzato il nuovo valore $z_s^{\max}$ al posto di quello iniziale $\bar{z}_s$. Facciamo un esempio che dovrebbe chiarire il procedimento, indicando con $\hat{z}$ il vettore dei valori correnti delle variabili z_v. inizialmente, $\hat{z}$ è posto pari a $\bar{z}$.

Si consideri un'istanza del problema di plant location con $|U| = 5$, $|V| = 6$, e i seguenti dati. I costi di attivazione $\{f_u\}$ sono dati dal vettore

$$f = (4 \quad 3 \quad 1 \quad 4 \quad 7)$$

mentre la matrice dei costi di connessione $\{c_{vu}\}$ e quindi il vettore dei valori iniziali $\{\bar{z}_v\}$ (calcolato tramite le (11)) sono dati da:

$$\begin{pmatrix} 12 & 13 & 6 & 0 & 1 \\ 8 & 4 & 9 & 1 & 2 \\ 2 & 6 & 6 & 0 & 1 \\ 3 & 5 & 2 & 10 & 8 \\ 8 & 0 & 5 & 10 & 8 \\ 2 & 0 & 3 & 4 & 1 \end{pmatrix} \qquad \bar{z} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$$

Si noti che la soluzione duale corrente ha un valore pari a 3 (somma delle componenti di $\bar{z}$). Andiamo allora a considerare le variabili nell'ordine indotto dalla loro numerazione, partendo dunque con z_1, attualmente al valore 0. Dalle (14) e (10) ci calcoliamo i valori $\bar{w}_{1u}$ per $u = 1, \ldots, 5$, il che può farsi per semplice ispezione (è sottolineato il termine in corrispondenza del quale si ha il minimo):

$$\begin{aligned} z_1^{\max} = \min\{ \quad & 12 + 4 - (\max\{0, 1-8\} + \max\{0, 0-2\} + \max\{0, 2-3\} \\ & \qquad + \max\{0, 0-8\} + \max\{0, 0-2\}); \\ & 13 + 3 - (\max\{0, 1-4\} + \max\{0, 0-6\} + \max\{0, 2-5\} \\ & \qquad + \max\{0, 0-0\} + \max\{0, 0-0\}); \\ & 6 + 1 - (\max\{0, 1-9\} + \max\{0, 0-6\} + \max\{0, 2-2\} \\ & \qquad + \max\{0, 0-5\} + \max\{0, 0-3\}); \\ & \underline{0 + 4 - (\max\{0, 1-1\} + \max\{0, 0-0\} + \max\{0, 2-10\}} \\ & \qquad \underline{+ \max\{0, 0-10\} + \max\{0, 0-4\})}; \\ & 1 + 7 - (\max\{0, 1-2\} + \max\{0, 0-1\} + \max\{0, 2-8\} \end{aligned}$$

$$+\max\{0,0-8\}+\max\{0,0-1\})\} = 4$$

Dunque, z_1 passa dal valore $\bar{z}_1 = 0$ al valore 4, e si ha:

$$\hat{z} = \begin{pmatrix} 4 \\ 1 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$$

Si noti che corrispondentemente anche la funzione obiettivo aumenta della stessa quantità, passando al valore 7.

Passiamo ora al calcolo delle altre variabili. Partendo da z_2, si ha:

$$\begin{aligned} z_2^{\max} = \min\{ \quad & 8+4-(\max\{0,4-12\}+\max\{0,0-2\}+\max\{0,2-3\} \\ & \quad +\max\{0,0-8\}+\max\{0,0-2\}); \\ & 4+3-(\max\{0,4-13\}+\max\{0,0-6\}+\max\{0,2-5\} \\ & \quad +\max\{0,0-0\}+\max\{0,0-0\}); \\ & 9+1-(\max\{0,4-6\}+\max\{0,0-6\}+\max\{0,2-2\} \\ & \quad +\max\{0,0-5\}+\max\{0,0-3\}); \\ & \underline{1+4-(\max\{0,4-0\}+\max\{0,0-0\}+\max\{0,2-10\}} \\ & \quad \underline{+\max\{0,0-10\}+\max\{0,0-4\})}; \\ & 2+7-(\max\{0,4-1\}+\max\{0,0-1\}+\max\{0,2-8\} \\ & \quad +\max\{0,0-8\}+\max\{0,0-1\})\} = 1 \end{aligned}$$

Poiché $z_2^{\max} = \hat{z}_2 = 1$, la variabile z_2 non viene incrementata e dunque $\hat{z}$ non subisce variazioni. Passando a z_3, si ha:

$$\begin{aligned} z_3^{\max} = \min\{ \quad & 2+4-(\max\{0,4-12\}+\max\{0,1-8\}+\max\{0,2-3\} \\ & \quad +\max\{0,0-8\}+\max\{0,0-2\}); \\ & 6+3-(\max\{0,4-13\}+\max\{0,1-4\}+\max\{0,2-5\} \\ & \quad +\max\{0,0-0\}+\max\{0,0-0\}); \\ & 6+1-(\max\{0,4-6\}+\max\{0,1-9\}+\max\{0,2-2\} \\ & \quad +\max\{0,0-5\}+\max\{0,0-3\}); \\ & \underline{0+4-(\max\{0,4-0\}+\max\{0,1-1\}+\max\{0,2-10\}} \end{aligned}$$

$$
\begin{aligned}
&\underline{+\max\{0,0-10\}+\max\{0,0-4\}});\\
&2+7-(\max\{0,4-1\}+\max\{0,1-2\}+\max\{0,2-8\}\\
&\quad+\max\{0,0-8\}+\max\{0,0-1\})\}=0
\end{aligned}
$$

Dunque, $z_3^{\max}=\hat{z}_3=0$, e di nuovo $\hat{z}$ non subisce variazioni. Passando a z_4, si ha:

$$
\begin{aligned}
z_4^{\max}=\min\{\quad &3+4-(\max\{0,4-12\}+\max\{0,1-8\}+\max\{0,0-2\}\\
&\quad+\max\{0,0-8\}+\max\{0,0-2\});\\
&5+3-(\max\{0,4-13\}+\max\{0,1-4\}+\max\{0,0-6\}\\
&\quad+\max\{0,0-0\}+\max\{0,0-0\});\\
&\underline{2+1-(\max\{0,4-6\}+\max\{0,1-9\}+\max\{0,0-6\}}\\
&\quad\underline{+\max\{0,0-5\}+\max\{0,0-3\}});\\
&10+4-(\max\{0,4-0\}+\max\{0,1-1\}+\max\{0,0-0\}\\
&\quad+\max\{0,0-10\}+\max\{0,0-4\});\\
&8+7-(\max\{0,4-1\}+\max\{0,1-2\}+\max\{0,0-1\}\\
&\quad+\max\{0,0-8\}+\max\{0,0-1\})\}=3
\end{aligned}
$$

Dunque, z_4 passa dal valore 2 al valore 3, e si ha il nuovo $\hat{z}$:

$$
\hat{z}=\begin{pmatrix}4\\1\\0\\3\\0\\0\end{pmatrix}
$$

La variabile successiva è z_5:

$$
\begin{aligned}
z_5^{\max}=\min\{\quad &8+4-(\max\{0,4-12\}+\max\{0,1-8\}+\max\{0,0-2\}\\
&\quad+\max\{0,3-3\}+\max\{0,0-2\});\\
&\underline{0+3-(\max\{0,4-13\}+\max\{0,1-4\}+\max\{0,0-6\}}\\
&\quad\underline{+\max\{0,3-5\}+\max\{0,0-0\}});\\
&5+1-(\max\{0,4-6\}+\max\{0,1-9\}+\max\{0,0-6\}\\
&\quad+\max\{0,3-2\}+\max\{0,0-3\});\\
&10+4-(\max\{0,4-0\}+\max\{0,1-1\}+\max\{0,0-0\}\\
&\quad+\max\{0,3-10\}+\max\{0,0-4\});
\end{aligned}
$$

$$8 + 7 - (\max\{0, 4-1\} + \max\{0, 1-2\} + \max\{0, 0-1\} + \max\{0, 3-8\} + \max\{0, 0-1\})\} = 3$$

Dunque anche z_5 viene portata al valore 3:

$$\hat{z} = \begin{pmatrix} 4 \\ 1 \\ 0 \\ 3 \\ 3 \\ 0 \end{pmatrix} \tag{15}$$

L'ultimo step è il calcolo di $z_6^{\max}$:

$$\begin{aligned} z_6^{\max} = \min\{ \quad & 2 + 4 - (\max\{0, 4-12\} + \max\{0, 1-8\} + \max\{0, 0-2\} + \max\{0, 3-3\} + \max\{0, 0-2\}); \\ & \underline{0 + 3 - (\max\{0, 4-13\} + \max\{0, 1-4\} + \max\{0, 0-6\} + \max\{0, 3-5\} + \max\{0, 3-0\})}; \\ & 3 + 1 - (\max\{0, 4-6\} + \max\{0, 1-9\} + \max\{0, 0-6\} + \max\{0, 3-2\} + \max\{0, 3-5\}); \\ & 4 + 4 - (\max\{0, 4-0\} + \max\{0, 1-1\} + \max\{0, 0-0\} + \max\{0, 3-10\} + \max\{0, 3-10\}); \\ & 1 + 7 - (\max\{0, 4-1\} + \max\{0, 1-2\} + \max\{0, 0-1\} + \max\{0, 3-8\} + \max\{0, 3-8\})\} = 0 \end{aligned}$$

Dunque z_6 non subisce incrementi, e in definitiva il vettore restituito dall'algoritmo di ascesa duale è (15), di valore 11, che, per inciso, in questo caso è anche il valore della soluzione ottima. Si noti che invece il valore del lower bound ottenuto dalla formulazione debole è pari ad appena $6 + 1/6$.

Un'osservazione conclusiva: si noti che il valore della soluzione ottenuta alla fine *dipende* dall'ordinamento scelto tra le variabili. Qui abbiamo usato un ordinamento casuale (abbiamo semplicemente seguito la numerazione delle variabili). L'algoritmo di Erlenkotter, per cui si rimanda al testo di Sassano, calcola un indice per ciascuna variabile, e utilizza questo indice per ordinarle. I risultati ottenuti sono mediamente molto buoni: in molti casi il lower bound così ottenuto coincide con il valore ottimo del rilassamento della formulazione forte, e spesso con l'ottimo del problema.