

Tecniche di rilassamento

1 Introduzione

Consideriamo un problema di ottimizzazione P in forma di minimo

$$(P) \quad \min f(x), \quad x \in F(P)$$

Possiamo definire un nuovo problema R nel seguente modo:

$$(R) \quad \min \Phi(x), \quad x \in F(R)$$

R si dice *problema rilassato* rispetto a P se valgono le seguenti condizioni:

- (a) $F(P) \subseteq F(R)$
- (b) $\Phi(x) \leq f(x) \quad \forall x \in F(P)$

Quindi un rilassamento è una qualunque tecnica che allarga la regione ammissibile (condizione (a)) e/o migliora la funzione obiettivo in tutti i punti della regione ammissibile originale (condizione (b)).

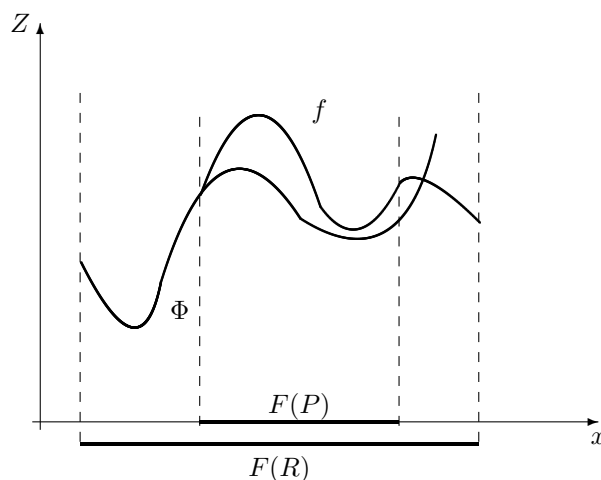


Figura 1: Esempio di rilassamento.

È da notare che la condizione $\Phi(x) \leq f(x)$ è richiesta solo per $x \in F(P)$ mentre non vi è nessuna condizione per i punti di $F(R) \setminus F(P)$ (dove la funzione f potrebbe anche non essere definita).

Inoltre non è detto che una soluzione ottima del problema rilassato sia una soluzione ottima del problema di partenza, nè (come nell'esempio di figura 1) che questa soluzione sia ammissibile per P . Infine, non è detto che il valore della soluzione ottima del problema rilassato coincida con il valore della soluzione ottima del problema originale.

Ovviamente quella cui vorremmo tendere è una situazione nella quale il problema rilassato ha le seguenti proprietà:

- è più facile da risolvere rispetto al problema originale;

- fornisce un “buon“ lower bound sul valore ottimo del problema di partenza;
- (almeno) una soluzione ottima del problema rilassato è ammissibile per il problema originale;
- (almeno) una soluzione ottima del problema rilassato è ottima anche per il problema originale.

Vedremo sotto quali condizioni ed in quale modo è possibile sperare di raggiungere questi obiettivi.

L'utilizzo di tecniche di rilassamento risulta essere estremamente utile nella soluzione di problemi di ottimizzazione combinatoria e, in generale, per i problemi reali per vari motivi:

- permette di avere una stima sul valore della soluzione ottima;
- consente di accelerare la convergenza di algoritmi di tipo enumerativo;
- serve a prevedere quale potrebbe essere l'impatto di una certa scelta (ad esempio rimozione di vincoli) sul costo della soluzione ottima

Tutte queste considerazioni motivano l'esigenza di studiare tecniche sofisticate ed attendibili per il calcolo di bound. In generale, la scelta del tipo di rilassamento da adottare per la soluzione di un problema di ottimizzazione combinatoria non è immediata; a volte potrebbe essere conveniente avere bound meno forti ma calcolabili in minor tempo di calcolo piuttosto che non bound migliori che richiedano tempi di calcolo elevati.

Nel seguito supporremo che il problema P sia definito nel seguente modo

$$v(P) = \min_x f(x) \tag{1}$$

$$g_i(x) \leq b_i \quad (i = 1, \dots, m) \tag{2}$$

$$x \in X \tag{3}$$

ossia che sia $F(P) = \{x \in X : g_i(x) \leq b_i, i = 1, \dots, m\}$. Immaginiamo anche che il problema di ottimizzare la funzione $f(x)$ sull'insieme X sia “facile“ da risolvere.

2 Rilassamento per eliminazione

Il rilassamento per eliminazione consiste nel definire un problema rilassato R nel quale alcuni vincoli sono stati cancellati. Eliminando i vincoli (2) si ottiene il seguente problema rilassato:

$$(R) \quad \min f(x), \quad x \in X$$

Poichè i problemi P e R hanno la stessa funzione obiettivo siamo certi che la condizione (b) sia verificata. Inoltre, è banale verificare che $F(P) \subseteq F(R)$ in quanto $F(R)$ è ottenuto da $F(P)$ rimuovendo alcuni vincoli.

In generale, la soluzione rilassata non è ammissibile per il problema originale quindi il rilassamento fornisce solo un lower bound sul valore della soluzione ottima (caso (a) in figura 2). È facile verificare che, se una soluzione ottima del problema rilassato è ammissibile per il problema originale, allora questa è anche una soluzione ottima (ed i due valori coincidono), come accade nel caso (b) in figura 2.

Un caso particolare di rilassamento per eliminazione si ha quando vengono eliminati i vincoli di interezza delle variabili. In tal caso si ottiene il *rilassamento continuo* del problema.

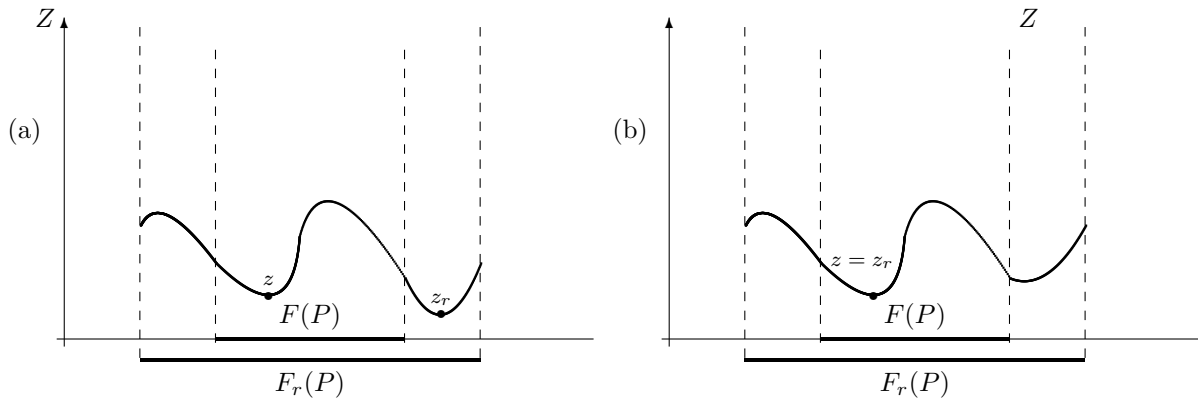


Figura 2: Possibilità nel rilassamento per eliminazione, surrogato e lagrangiano (uguaglianza).

3 Rilassamento surrogato

L'idea di base di questo tipo di rilassamento è quello di sostituire un insieme di vincoli con una loro combinazione lineare, lasciando inalterata la funzione obiettivo.

Preso un vettore $(u_1, \dots, u_m) \in R_+^m$ di *moltiplicatori surrogati*, una combinazione lineare degli m vincoli (2) è una disuguaglianza del tipo

$$\sum_{i=1}^m u_i g_i(x) \leq \sum_{i=1}^m u_i b_i$$

e quindi il problema ottenuto per rilassamento surrogato è definito nel seguente modo:

$$S_u(P) = \min_{\substack{x \in X \\ \sum_{i=1}^m u_i g_i(x) \leq \sum_{i=1}^m u_i b_i}} f(x)$$

È immediato verificare che qualunque vettore x che soddisfa tutti i vincoli (2) soddisfa anche la loro combinazione lineare per qualunque vettore di moltiplicatori surrogati (mentre, in generale, non è vero il contrario); da questo consegue che $F(P) \subseteq F(R)$. Anche per questo rilassamento, vale la proprietà secondo la quale l'ammissibilità della soluzione rilassata implica l'ottimalità della soluzione stessa.

Ciascun moltiplicatore serve per "pesare" l'importanza del vincolo nel nuovo vincolo che definisce il problema rilassato. Un rilassamento nel quale ci sia un vincolo (o più di uno) con $u_i = 0$ equivale ad un rilassamento per eliminazione del vincolo stesso. Nel caso in cui si rilassi un vincolo di uguaglianza, il corrispondente moltiplicatore non è vincolato a dover assumere valori non negativi.

Il valore della soluzione ottima del problema rilassato prima definito fornisce un bound valido sul valore della soluzione ottima del problema originale per qualunque vettore $u \in R_+^m$ di moltiplicatori surrogati. È ovvio che il valore di questo bound dipende dai moltiplicatori utilizzati e che il bound è tanto migliore quanto più alto è il suo valore. Nasce quindi il problema di identificare il vettore di moltiplicatori cui corrisponde il miglior (più alto) lower bound; questo problema si chiama il *problema surrogato duale*.

$$S_{u^*}(P) = \max_{u \geq 0} S_u(P)$$

4 Rilassamento lagrangiano

L'idea di base di questo tipo di rilassamento è quello di eliminare un insieme di vincoli ma di tenere conto di tali vincoli nella funzione obiettivo.

Preso un vettore $(\lambda_1, \dots, \lambda_m) \in R_+^m$ di *moltiplicatori lagrangiani*, il problema ottenuto per rilassamento lagrangiano dei vincoli (2) è definito nel seguente modo:

$$L_\lambda(P) = \min_{x \in X} f(x) + \sum_{i=1}^m \lambda_i [g_i(x) - b_i] \quad (4)$$

$$x \in X \quad (5)$$

Avendo eliminato i vincoli (2), la regione ammissibile del problema rilassato comprende la regione ammissibile del problema di partenza, per cui $F(P) \subseteq F(R)$.

Per quel che riguarda la nuova funzione obiettivo, questa è della forma

$$\Phi_\lambda(x) = - \sum_{i=1}^m \lambda_i b_i + \left[f(x) + \sum_{i=1}^m \lambda_i g_i(x) \right]$$

ossia comprende un termine costante (che dipende dai moltiplicatori) ed un termine da minimizzare (fissati i moltiplicatori, dipende dalle variabili decisionali x). La funzione obiettivo del problema rilassato è stata ottenuta sommando alla funzione obiettivo di partenza un termine migliorante (non positivo) in qualunque punto ammissibile per cui, per ogni $x \in F(P)$ si ha $\Phi_\lambda(x) \leq f(x)$. Infatti, per ogni $x \in F(P)$ si ha $g_i(x) \leq b_i$ per cui il termine aggiunto alla funzione obiettivo è ≤ 0 (ricordando che $\lambda_i \geq 0 \forall i = 1, \dots, m$). Nel caso in cui la funzione obiettivo fosse stata da massimizzare avremmo dovuto aggiungere un termine non negativo per qualunque punto ammissibile. Nel caso in cui i vincoli da rilassare fossero stati di uguaglianza, i corrispondenti moltiplicatori lagrangiani non sarebbero stati vincolati in segno ed il contributo alla funzione obiettivo sarebbe stato nullo $\forall x \in F(P)$. Un rilassamento nel quale ci sia un vincolo (o più di uno) con $u_i = 0$ equivale ad un rilassamento per eliminazione del vincolo stesso, per cui il corrispondente vincolo non è assolutamente preso in esame (nè nella definizione della regione ammissibile, nè nei vincoli).

• Esempio: Multiple KP-01

- estensione di KP-01 al caso in cui ci sono a disposizione m contenitori;
- il generico contenitore i ha capacità pari a c_i ($i = 1, \dots, m$);
- n oggetti;
- il generico oggetto j ha profitto p_j e peso w_j ($j = 1, \dots, n$)

Introducendo le seguenti variabili binarie

$$x_{ij} = \begin{cases} 1 & \text{se l'oggetto } j \text{ viene inserito nel contenitore } i \\ 0 & \text{altrimenti} \end{cases} \quad (i = 1, \dots, m; j = 1, \dots, n)$$

si ottiene un possibile modello di PLI per il problema in esame:

$$\begin{aligned} \text{MKP-01} \quad \max \quad & \sum_{i=1}^m \sum_{j=1}^n p_j x_{ij} \\ & \sum_{j=1}^n w_j x_{ij} \leq c_i \quad i = 1, \dots, m \\ & \sum_{i=1}^m x_{ij} \leq 1 \quad j = 1, \dots, n \\ & x_j \in \{0, 1\} \quad j = 1, \dots, n \end{aligned}$$

Rilassamento surrogato dei vincoli di capacità con moltiplicatori con $u \in R_+^m$:

$$\begin{aligned}
 S_u(\text{MKP-01}) \quad & \max \sum_{i=1}^m \sum_{j=1}^n p_j x_{ij} \\
 & \sum_{i=1}^m u_i \sum_{j=1}^n w_j x_{ij} \leq \sum_{i=1}^m u_i c_i \\
 & \sum_{i=1}^m x_{ij} \leq 1 \quad j = 1, \dots, n \\
 & x_j \in \{0, 1\} \quad j = 1, \dots, n
 \end{aligned}$$

I vincoli di capacità per gli m contenitori sono sostituiti dall'unico vincolo di "capacità" pari a

$$C := \sum_{i=1}^m u_i c_i$$

Quindi ci si riconduce ad un problema KP-01 dove per ogni oggetto j esistono m esemplari, l' i -esimo dei quali ha peso $u_i w_j$ e profitto p_j , e al massimo un esemplare di ciascun oggetto è inseribile nel contenitore. Per ogni oggetto j conviene eventualmente inserire l'esemplare i tale che u_i è minimo, per cui per ogni oggetto basta considerare un solo esemplare. Quindi il problema si riconduce ad un KP-01 con tutti gli oggetti originali ed un unico contenitore di capacità C .

Si può dimostrare che i moltiplicatori surrogati che forniscono l'upper bound più basso sono definiti da $u_i = 1$ per $i = 1, \dots, m$.

Rilassamento lagrangiano dei vincoli di cardinalità con moltiplicatori con $\lambda \in R_+^n$:

$$\begin{aligned}
 L_\lambda(\text{MKP-01}) \quad & \max \sum_{i=1}^m \sum_{j=1}^n p_j x_{ij} + \sum_{j=1}^n \lambda_j (1 - \sum_{i=1}^m x_{ij}) \quad (\text{con } \lambda_j \geq 0 \forall j) \\
 = \sum_{j=1}^n \lambda_j + & \max \sum_{i=1}^m \sum_{j=1}^n \tilde{p}_j x_{ij} \quad (\text{con } \tilde{p}_j = p_j - \lambda_j) \\
 & \sum_{j=1}^n w_j x_{ij} \leq c_i \quad (i \in M) \\
 & x_{ij} \in \{0, 1\} \quad (i \in M, j \in N)
 \end{aligned}$$

Nel problema lagrangiano, ogni oggetto j ha profitto $p_j - \lambda_j$, peso w_j , e può essere inserito anche in più di un contenitore (non c'è alcun vincolo che leghi le variabili associate a contenitori differenti). Quindi il problema rilassato equivale ad m KP-01 indipendenti:

$$\begin{aligned}
 \text{per } i = 1, \dots, m : \quad & z_i = \max \sum_{j=1}^n \tilde{p}_j x_{ij} \\
 & \sum_{j=1}^n w_j x_{ij} \leq c_i \\
 & x_{ij} \in \{0, 1\} \quad (j \in N) \\
 \Rightarrow z(L_\lambda(\text{MKP-01})) = & \sum_{i=1}^m z_i + \sum_{j=1}^n \lambda_j
 \end{aligned}$$

4.1 Proprietà della soluzione rilassata

Generalmente una soluzione ottima del rilassamento lagrangiano non soddisfa i vincoli rilassati (caso (a) in figura 3). Il fatto di aver modificato la funzione obiettivo non consente, in generale, di affermare

che una soluzione ottima del rilassamento lagrangiano, se ammissibile per P , è anche ottima per il problema di partenza (come nel caso (b) di figura 3). Indicando infatti con \tilde{x}_λ una soluzione ottima del rilassamento lagrangiano, si ha che $L_\lambda(P) = \Phi_\lambda(\tilde{x}_\lambda)$ è un lower bound sul valore della soluzione ottima, cioè $\Phi_\lambda(\tilde{x}_\lambda) \leq v(P)$. Nel caso un cui la soluzione \tilde{x}_λ sia ammissibile per P si ha che $f(\tilde{x}_\lambda) \geq v(P)$, ossia il valore della soluzione, valutata con la funzione obiettivo originale, è un upper bound sul valore della soluzione ottima. Ne deriva la seguente relazione

$$\Phi_\lambda(\tilde{x}_\lambda) = f(\tilde{x}_\lambda) + \sum_{i=1}^m \lambda_i [g_i(\tilde{x}_\lambda) - b_i] \leq v(P) \leq f(\tilde{x}_\lambda)$$

e quindi una soluzione $\tilde{x}_\lambda \in P$ è ottima se vale la *condizione di complementarità*

$$\sum_{i=1}^m \lambda_i [g_i(\tilde{x}_\lambda) - b_i] = 0$$

Sotto tali condizioni il lower bound fornito dal rilassamento lagrangiano coincide con il valore della soluzione ottima (caso (c) in figura 3). Quindi \tilde{x}_λ , qualora ammissibile per P , è ottima se ha moltiplicatore nullo per per ogni vincolo lasco, ossia se $\lambda_i = 0$ per ogni vincolo i tale che $g_i(\tilde{x}_\lambda) < b_i$ e $g_i(\tilde{x}_\lambda) = b_i$ per ogni vincolo i per cui $\lambda_i > 0$.

Nel caso in cui tutti i vincoli rilassati siano di uguaglianza, la condizione di complementarità vale automaticamente per ogni soluzione rilassata che risulti ammissibile per il problema di partenza (vedi figura 2).

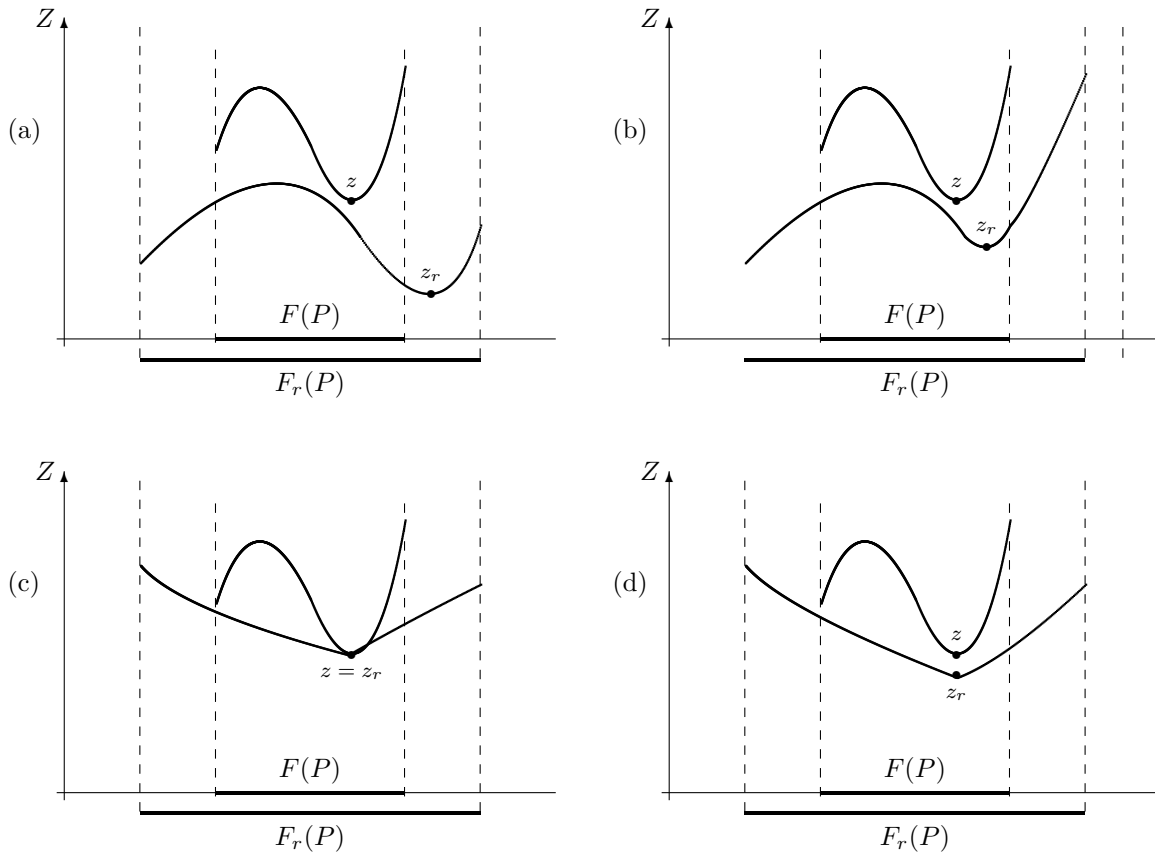


Figura 3: Possibilità nel rilassamento lagrangiano (disuguaglianza).

Nota: la condizione di complementarità è condizione sufficiente ma non necessaria. Consideriamo infatti il seguente esempio:

$$\begin{aligned} \min z = & 3x_1 + 7x_2 + 10x_3 \\ & x_1 + 3x_2 + 5x_3 \geq 7 \\ & x_1, x_2, x_3 \in \{0, 1\} \end{aligned}$$

per il quale si vede facilmente che una soluzione ottima vale 17 ed è data dal vettore $x^* = (0, 1, 1)$.

Consideriamo il rilassamento lagrangiano del problema con moltiplicatore $\lambda \geq 0$

$$(L_\lambda) \begin{cases} 7\lambda + \min(3 - \lambda)x_1 + (7 - 3\lambda)x_2 + (10 - 5\lambda)x_3 \\ x_1, x_2, x_3 \in \{0, 1\} \end{cases}$$

Prendendo il moltiplicatore $\lambda = \frac{7}{3}$ una possibile soluzione ottima del rilassamento lagrangiano è $\widetilde{x_{7/3}} = (0, 1, 1)$ cui corrisponde un lower bound pari a $\frac{44}{3}$. (Si noti che anche la soluzione $x^* = (0, 0, 1)$ è ottima per il problema rilassato). La soluzione trovata è ammissibile per il problema originale ma il contributo lagrangiano in corrispondenza di tale soluzione vale $\lambda(7 - x_1 - 3x_2 - 5x_3) = 7/3 > 0$. Nonostante ciò la soluzione trovata è proprio la soluzione ottima del problema originale (si veda il caso (d) in figura 3).

4.2 Scelta dei moltiplicatori lagrangiani

Come avveniva nel caso del rilassamento surrogato, il valore della soluzione ottima del problema rilassato fornisce un lower bound sul valore della soluzione ottima per ogni vettore $\lambda \in R_+^m$ di moltiplicatori lagrangiani (si noti che la condizione $\lambda_i \geq 0$ non serve se il corrispondente vincolo è di uguaglianza).

Il problema di identificare il vettore di moltiplicatori cui corrisponde il miglior lower bound è noto come *problema lagrangiano duale*.

$$L_{\lambda^*}(P) = \max_{\lambda \geq 0} L_\lambda(P) = \max_{\lambda \geq 0} \left\{ \min_{x \in X} \left[f(x) + \sum_{i=1}^m \lambda_i (g_i(x) - b_i) \right] \right\}$$

Supponiamo che l'insieme X su cui è definito il problema rilassato sia un insieme discreto (cosa molto frequente nelle applicazioni descritte da problemi di ottimizzazione combinatoria) definito nel seguente modo:

$$X = \{x^1, x^2, \dots, x^M\}$$

Allora è possibile scrivere

$$L_\lambda(P) = \min_{k=1, \dots, M} \left[f(x^k) + \sum_{i=1}^m \lambda_i (g_i(x^k) - b_i) \right]$$

Consideriamo un generico moltiplicatore λ_i ; fissato l'indice k , l'andamento di $L_\lambda(P)$ in funzione di λ_i è lineare. Dato che $L_\lambda(P)$ è definito come minimo tra le M soluzioni del problema rilassato, il suo andamento in funzione di λ_i sarà quello di un minimo tra M rette distinte; ne consegue una funzione con andamento lineare a tratti e concava (anche se non sempre differenziabile), come mostrato in figura 4.

La funzione $L_\lambda(P)$ sarà quindi data dall'intersezione di una serie di iperpiani di dimensioni ed inclinazioni diverse; data la difficoltà di determinare il minimo di tale funzione, spesso ci si accontenta di una soluzione subottima in grado comunque di fornire un buon lower bound (anche se non quello ottimo).

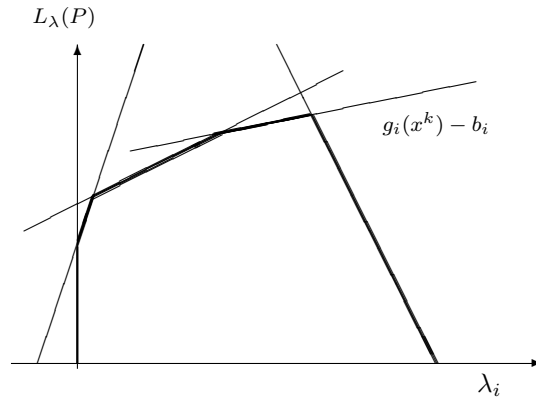


Figura 4: Valore del rilassamento lagrangiano in funzione di un moltiplicatore λ_i .

• **Esempio: KP-01**

$$\begin{aligned} \max \quad & \sum_{j=1}^n p_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^n w_j x_j \leq c \\ & x_j \in \{0, 1\} \quad j = 1, \dots, n \end{aligned}$$

Per ogni $\lambda \geq 0$, il rilassamento lagrangiano del problema è

$$\lambda c + \max \sum_{j=1}^n \tilde{p}_j x_j \quad x_j \in \{0, 1\} \quad j = 1, \dots, n$$

con $\tilde{p}_j = p_j - \lambda w_j$ ($j = 1, \dots, n$). La soluzione del problema rilassato è immediata:

$$x_j = \begin{cases} 1 & \text{se } \tilde{p}_j > 0 \\ 0 & \text{se } \tilde{p}_j \leq 0 \end{cases}$$

ed il corrispondente lower bound vale

$$z(L_\lambda(\text{KP})) = \lambda c + \sum_{j: \tilde{p}_j > 0} \tilde{p}_j$$

Consideriamo la seguente istanza di KP-01:

$n = 3, c = 4, p_j = (7, 12, 11), w_j = (1, 3, 2)$, per la quale $X = \{x^0, \dots, x^7\}$ con

k	x^k	z^k
0	\emptyset	4λ
1	$\{1\}$	$7 + 3 \lambda$
2	$\{2\}$	$12 + \lambda$
3	$\{1, 2\}$	19
4	$\{3\}$	$11 + 2 \lambda$
5	$\{1, 3\}$	$18 + \lambda$
6	$\{2, 3\}$	$23 - \lambda$
7	$\{1, 2, 3\}$	$30 - 2 \lambda$

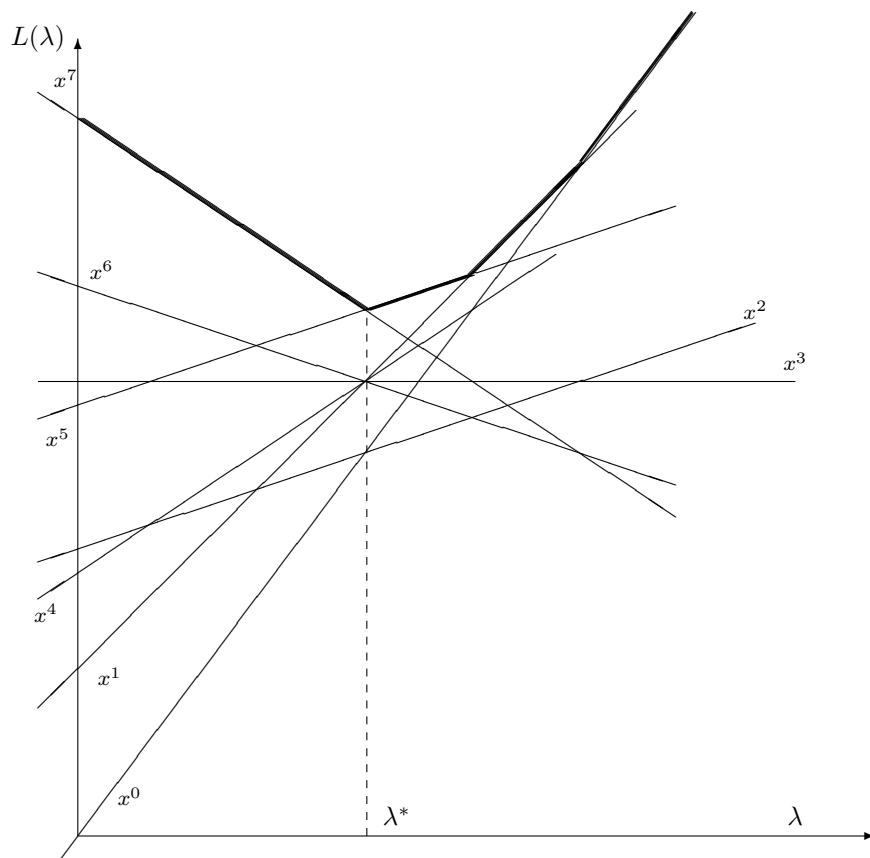


Figura 5: Upper bound lagrangiano per KP-01.

L'andamento dell'upper bound in funzione di λ è mostrato in figura 5; il valore λ^* cui corrisponde il miglior upper bound è dato dall'intersezione delle seguenti rette

$$\begin{cases} (x^5) & L(\lambda) = 18 + \lambda \\ (x^7) & L(\lambda) = 30 - 2\lambda \end{cases}$$

e vale quindi $\lambda^* = 4$.

In generale, per il problema KP-01, il moltiplicatore ottimo è pari al rapporto tra profitto e peso dell'oggetto critico (in questo caso l'oggetto 2).

Tecnica subgradiente

Si tratta di una tecnica algoritmica per la determinazione di “buoni” moltiplicatori lagrangiani basata considerando ciascun moltiplicatore come uno strumento per “pesare” l'importanza del vincolo associato nella funzione obiettivo. Di fatto la situazione è la seguente: ci dimentichiamo dei vincoli nella definizione della regione ammissibile del problema rilassato, ma teniamo conto del soddisfacimento o della violazione dei vincoli stessi nella funzione obiettivo. Il termine aggiunto alla funzione obiettivo originale “premia” il soddisfacimento dei vincoli e ne penalizza la violazione, per cui la funzione obiettivo va nella direzione di soddisfare tutti i vincoli. L'entità di questo premio/penalizzazione dipende dai moltiplicatori lagrangiani utilizzati. Se un moltiplicatore è troppo piccolo, allora il contributo alla funzione obiettivo del vincolo associato è modesto e la funzione obiettivo non spinge abbastanza nella direzione di soddisfacimento di questo vincolo; viceversa, se il moltiplicatore associato ad un vincolo è troppo grande, la funzione obiettivo è troppo “deviata” da questo vincolo e tende a soddisfarlo in modo eccessivo. Il procedimento converge quando ciascun vincolo è soddisfatto in modo stretto oppure il suo moltiplicatore è nullo, così che il corrispondente contributo alla funzione obiettivo è nullo.

Un vettore $s \in R^m$ è un *subgradiente* per la funzione concava $L_\lambda(P)$ in $\lambda^0 \geq 0$ se

$$L_\lambda(P) \leq L_{\lambda^0}(P) + s(\lambda - \lambda^0), \forall \lambda \geq 0$$

L'insieme dei subgradienti di $L_\lambda(P)$ in λ^0 è detto *subdifferenziale* ed è indicato con $\partial L_{\lambda^0}(P)$; si può mostrare che il subdifferenziale è un insieme convesso.

Dato un vettore $\lambda^0 \geq 0$ di moltiplicatori lagrangiani, il vettore $s = g(\widetilde{x}_{\lambda^0}) - b$, calcolato in corrispondenza di una soluzione ottima $\widetilde{x}_{\lambda^0}$ del rilassamento lagrangiano, è sempre un vettore subgradiente. Infatti, il fatto che $\widetilde{x}_{\lambda^0}$ sia una soluzione ottima del rilassamento lagrangiano con moltiplicatori λ^0 significa che

$$L_{\lambda^0}(P) = \Phi_{\lambda^0}(\widetilde{x}_{\lambda^0}) = f(\widetilde{x}_{\lambda^0}) + \lambda^0 [g(\widetilde{x}_{\lambda^0}) - b]$$

D'altro canto, preso un qualunque altro vettore di moltiplicatori $\lambda \geq 0$ si ha

$$L_\lambda(P) = \min_{x \in X} \Phi_\lambda(x) \leq \Phi_\lambda(\widetilde{x}_{\lambda^0}) = f(\widetilde{x}_{\lambda^0}) + \lambda [g(\widetilde{x}_{\lambda^0}) - b]$$

per cui

$$L_\lambda(P) \leq L_{\lambda^0}(P) + (\lambda - \lambda^0) [g(\widetilde{x}_{\lambda^0}) - b]$$

In generale, il subdifferenziale in λ^0 contiene anche dei vettori subgradiente che non hanno la forma vista; quelli della forma $g(\widetilde{x}_{\lambda^0}) - b$ sono detti subgradienti *attivi*. La conoscenza completa di tutti i subgradienti permette una completa caratterizzazione di $L_\lambda(P)$ nell'intorno di λ^0 . Poichè però non è possibile conoscere completamente i subgradienti associati al punto corrente, si utilizzano tecniche di tipo approssimato per la determinazione dei moltiplicatori.

I metodi di tipo subgradiente sono basati su relazioni di tipo iterativo nelle quali i moltiplicatori dell'iterazione successiva sono definiti a partire da quelli dell'iterazione corrente e dalla soluzione ottima del rilassamento lagrangiano identificata all'iterazione corrente. Lo schema iterativo è del tipo

$$\begin{cases} \lambda_i^0 \in R_+^m \\ \lambda_i^{k+1} := \max\{0, \lambda_i^k + t^k d_i^k\} \end{cases} \quad (i = 1, \dots, m)$$

dove il vettore $d^k \in \partial L_{\lambda^k}(P)$ e $t^k > 0$ per $k \geq 1$. Non necessariamente d^k è una direzione di crescita per la funzione $L_\lambda(P)$ nel punto λ^k se questo è un punto di non differenziabilità; questo implica che i valori di lower bound trovati non sono monotoni.

Il significato dello schema per il generico vincolo i è il seguente. Se $d_i^k > 0 \rightarrow g_i(x_{\lambda^k}) > b_i$ ossia il vincolo i -esimo è violato; conviene quindi aumentare λ_i affinché la funzione obiettivo cerchi di soddisfare il vincolo alla prossima iterazione. Viceversa, Se $d_i^k < 0 \rightarrow g_i(x_{\lambda^k}) < b_i$ ossia il vincolo i -esimo è lasco; conviene quindi calare λ_i affinché la funzione obiettivo cerchi di soddisfare meno il vincolo alla prossima iterazione. Infine, se $d_i^k = 0$ allora vuol dire che il vincolo è soddisfatto in modo stringente, ossia il moltiplicatore associato λ_i è corretto e non deve essere modificato nella successiva iterazione.

Si può dimostrare che se

$$\sum_{k=1}^{\infty} t^k < \infty$$

e

$$\lim_{k \rightarrow \infty} t^k \rightarrow 0$$

allora il metodo converge, ossia

$$\forall \epsilon > 0, \exists k : |L_{\lambda^*}(P) - L_{\lambda^k}(P)| < \epsilon$$

La condizione richiede che la *lunghezza del passo* t^k decresca al crescere del numero di iterazioni. Dal punto di vista pratico, spesso si utilizza il seguente criterio, proposto da Held e Karp. L'idea di base è quella di definire

$$t^k := \alpha^k \frac{\bar{v} - L_{\lambda^k}(P)}{\sum_{i=1}^m s_i^2(\lambda^k)}$$

dove α^k è un parametro positivo e \bar{v} è una stima per eccesso del valore della soluzione ottima. Se il denominatore vale 0 vuol dire che il vettore subgradiente è nullo, per cui abbiamo trovato una soluzione ottima. Si dimostra che se $\bar{v} = L_{\lambda^*}(P)$ e $\alpha^k \in [0, 2]$ allora il vettore λ^k converge al vettore ottimo λ^* ; nella pratica il problema è che generalmente \bar{v} viene posto pari al valore della miglior soluzione conosciuta e non è mai pari a $L_{\lambda^*}(P)$. Generalmente si parte con $\alpha^k = 2$ e se ne dimezza il valore se il bound non è migliorato nelle ultime T iterazioni (es. $T = 10$). Quando il valore di α^k diventa troppo piccolo, i vettori subgradienti successivi cambiano di poco, per cui il bound si stabilizza ed il procedimento può terminare; altrimenti si attende il verificarsi di una qualche condizione di arresto (numero massimo di iterazioni o time limit).

Le tecniche di tipo subgradiente sono molto utilizzate per la loro semplicità e facilità di implementazione, anche se la convergenza del metodo ad una soluzione duale ottima è garantita solo in un numero infinito di iterazioni.

Esistono comunque altre tecniche per la soluzione del problema lagrangiano duale. I più utilizzati sono i metodi di tipo *bundle*, basati sulla soluzione di problemi quadratici per la generazione della direzione di spostamento, in modo da tener conto di tutte le soluzioni lagrangiane calcolate nelle precedenti iterazioni. Anche in questo modo non è garantito che la direzione scelta sia di crescita per la funzione $L_{\lambda}(P)$, per cui deve esistere un criterio per decidere se accettare la direzione proposta; l'idea di base è di valutare se la funzione obiettivo ha un incremento sufficiente e di generare una nuova direzione nel caso in cui la direzione precedente sia stata ritenuta inefficiente. Questi metodi generalmente portano a soluzioni migliori ma richiedono tempi di calcolo più elevati, che possono essere inaccettabili per istanze di grandi dimensioni.

4.3 Rilassamento lagrangiano e programmazione lineare

Consideriamo ora il comportamento del rilassamento lagrangiano di un problema di programmazione lineare.

$$v(P) = \min c^T x \tag{6}$$

$$A x \geq b \tag{7}$$

$$x \geq 0 \tag{8}$$

I vincoli (7) corrispondono ai vincoli $g_i(x) \leq b_i$ da rilassare, mentre i vincoli (8) corrispondono ai vincoli impliciti $x \in X$. Dato un vettore $u \geq 0$ di moltiplicatori lagrangiani per i vincoli (7), si ottiene il rilassamento

$$L_u(P) = \min_{x \geq 0} [c^T x + u^T (b - A x)] = u^T b + \min_{x \geq 0} [(c^T - u^T A) x]$$

La funzione obiettivo del problema lagrangiano è data da un termine costante più un contributo corrispondente ai costi ridotti nel problema di programmazione lineare.

È ovvio che in una qualunque soluzione ottima del problema

$$\min_{x \geq 0} (c^T - u^T A) x$$

sarà della forma

$$x_j = \begin{cases} 0 & \text{se } (c^T - u^T A)_j > 0 \\ -\infty & \text{se } (c^T - u^T A)_j \leq 0 \end{cases}$$

e che quindi

$$\min_{x \geq 0} = \begin{cases} 0 & \text{se } (c^T - u^T A) \geq 0 \\ -\infty & \text{altrimenti} \end{cases}$$

Nel problema lagrangiano duale vogliamo scegliere un vettore $u^* \geq 0$ tale che la quantità

$$u^T b + \min_{x \geq 0} [(c^T - u^T A) x]$$

sia massima; ovviamente questo non si può avere se $\min_{x \geq 0} [(c^T - u^T A) x] = -\infty$, per cui il problema si riconduce alla determinazione del vettore u in modo da massimizzare $u^T b$ con il vincolo che la quantità $(c^T - u^T A)$ sia non negativa. Questo problema può essere riscritto come

$$v(D) = \max u^T b \tag{9}$$

$$c^T - u^T A \geq 0 \tag{10}$$

$$u \geq 0 \tag{11}$$

ossia questo è il problema duale del problema primale (6)–(8).

La condizione di complementarità equivale alle condizioni di ottimalità di una coppia di soluzioni primale-duale (nel caso in cui le soluzioni primale e duale siano ammissibili per i rispettivi problemi). Quindi una soluzione ottima del problema lagrangiano duale non è altro che una soluzione ottima del duale del problema stesso.

Quindi, per un problema di programmazione lineare, il rilassamento lagrangiano non fornisce un lower bound, ma fornisce addirittura il valore della soluzione ottima.

5 Confronto tra i vari rilassamenti

rilassamento per eliminazione vs. rilassamento surrogato e lagrangiano

È immediato vedere che il rilassamento per eliminazione di vincoli (2) è dominato (in termini di qualità del bound prodotto) sia dal rilassamento surrogato duale che dal rilassamento lagrangiano duale (relativamente agli stessi vincoli). Infatti il rilassamento per eliminazione dei vincoli (2) può essere ottenuto come caso particolare di rilassamento surrogato (o lagrangiano) dei vincoli (2) con moltiplicatori nulli; quindi, a parità di vincoli rilassati, il bound prodotto dal rilassamento surrogato (e lagrangiano) duale è non peggiore di quello prodotto dal rilassamento per eliminazione. Spesso infatti il rilassamento per eliminazione produce bound scarsi; il suo utilizzo nella pratica è motivato dal fatto che talvolta è impossibile (o estremamente difficile) modellare alcuni vincoli che si presentano nelle applicazioni reali, per cui l'unico modo per avere una stima sul valore della soluzione è quello di trascurare i vincoli associati.

rilassamento surrogato vs. rilassamento lagrangiano

In maniera del tutto analoga si può osservare che, a parità di vincoli rilassati, il rilassamento surrogato duale domina il rilassamento lagrangiano duale. Infatti, dato un qualunque vettore λ di moltiplicatori lagrangiani associati ai vincoli (2), il corrispondente problema rilassato $L_\lambda(P)$ può essere visto come rilassamento lagrangiano con moltiplicatore unitario del problema $S_\lambda(P)$ ottenuto per rilassamento surrogato dei vincoli (2) con moltiplicatori λ (il problema ottenuto per rilassamento surrogato ha un solo vincolo - rilassando tale vincolo in modo lagrangiano con moltiplicatore pari a 1 si ottiene il rilassamento lagrangiano del problema di partenza). Di conseguenza, per qualunque vettore λ di moltiplicatori, il rilassamento lagrangiano è un rilassamento del rilassamento surrogato (ossia è più debole del rilassamento surrogato). Questo vale anche prendendo i moltiplicatori λ^* lagrangiani duali ottimi, per i quali si ha

$$L_{\lambda^*}(P) \leq S_{\lambda^*}(P)$$

D'altro canto, in corrispondenza dei moltiplicatori u^* surrogati ottimi si ha

$$S_{u^*}(P) \geq S_{\lambda^*}(P)$$

per cui il rilassamento surrogato duale domina il rilassamento lagrangiano duale (a parità di vincoli rilassati). Nella pratica però spesso si utilizza il rilassamento lagrangiano anziché il surrogato; questo fatto è essenzialmente dovuto a due motivi:

- il problema lagrangiano è più semplice da risolvere del problema surrogato (ha un vincolo in meno);

- esistono tecniche efficienti per la determinazione di “buoni” moltiplicatori lagrangiani, mentre lo stesso non si può dire relativamente ai moltiplicatori surrogati (in quanto i moltiplicatori non compaiono nella funzione obiettivo). Generalmente anche per la determinazione dei moltiplicatori surrogati si utilizzano tecniche di tipo subgradiente, ma in questo caso la convergenza del metodo al valore ottimo avviene più lentamente.

rilassamento continuo vs. rilassamento lagrangiano

Consideriamo ora la relazione che intercorre tra il rilassamento lagrangiano duale ed il rilassamento continuo per un problema P di Programmazione Lineare Intera definito nel seguente modo:

$$v(P) = \min f(x) \quad (12)$$

$$g_i(x) \leq b_i \quad (i = 1, \dots, m) \quad (13)$$

$$x \in Y \quad \text{intere} \quad (14)$$

Questo problema è del tutto equivalente al problema (1)–(3), salvo per il fatto che è stato esplicitamente specificato il vincolo di interezza sulle variabili (che nel problema (1)–(3) compariva implicitamente nella definizione dell’insieme X).

Il rilassamento lagrangiano di tutti i vincoli (13) con dei generici moltiplicatori lagrangiani $\lambda \in R_+^m$ è definito da

$$L_\lambda(P) = \min \Phi_\lambda(x) \quad (15)$$

$$x \in Y \quad \text{intere} \quad (16)$$

dove al solito la funzione obiettivo è definita da

$$\Phi_\lambda(x) = - \sum_{i=1}^m \lambda_i b_i + \left[f(x) + \sum_{i=1}^m \lambda_i g_i(x) \right]$$

Il rilassamento continuo del problema lagrangiano (15)–(16) è il seguente problema:

$$C(L_\lambda(P)) = \min \Phi_\lambda(x) \quad (17)$$

$$x \in Y \quad (18)$$

Per ogni vettore di moltiplicatori lagrangiani il problema (17)–(18) è un rilassamento del problema (15)–(16), per cui si ha

$$C(L_\lambda(P)) \leq L_\lambda(P) \quad \forall \lambda \in R_+^m \quad (19)$$

Consideriamo ora il rilassamento continuo del problema originale (12)–(14):

$$C(P) = \min f(x) \quad (20)$$

$$g_i(x) \leq b_i \quad (i = 1, \dots, m) \quad (21)$$

$$x \in Y \quad (22)$$

Il rilassamento lagrangiano dei vincoli (21) utilizzando i moltiplicatori lagrangiani λ^C ottimi per il problema (20)–(22) è il seguente:

$$L_{\lambda^C}(C(P)) = \min \Phi_{\lambda^C}(x) \quad (23)$$

$$x \in Y \quad (24)$$

È immediato vedere che il problema (23)–(24) non è altro che il problema (17)–(18) quando i moltiplicatori scelti sono pari a λ^C .

Dalla discussione del paragrafo 4.3, dato che (23)–(24) è il rilassamento lagrangiano con moltiplicatori ottimi di un problema di programmazione lineare si ha che

$$C(P) = L_{\lambda^C}(C(P)) = C(L_{\lambda^C}(P))$$

In virtù della (19) si ha che

$$C(P) = C(L_{\lambda^C}(P)) \leq L_{\lambda^C}(P)$$

In particolare, prendendo i moltiplicatori lagrangiani λ^I ottimi per il problema (12)–(14), si ha

$$L_{\lambda^I}(P) \geq L_{\lambda^C}(P) \geq C(P)$$

cioè il rilassamento lagrangiano duale domina il rilassamento continuo qualunque sia il sottoinsieme dei vincoli rilassati. Si noti che, in generale, i moltiplicatori λ^I ottimi per il problema originale non coincidono con i moltiplicatori λ^C ottimi per il suo rilassamento continuo.

Esistono dei problemi per i quali il rilassamento lagrangiano produce lo stesso bound prodotto dal rilassamento continuo; questo si verifica qualora il problema lagrangiano possieda la *proprietà di interezza*, ossia sia tale che per qualunque istanza, il suo rilassamento continuo ha soluzione ottima intera.

Ovviamente, le stesse proprietà di dominanza valgono anche per il rilassamento surrogato duale, in quanto questo domina il lagrangiano duale. Invece non esiste, in generale, nessuna relazione di dominanza tra il rilassamento continuo ed il rilassamento per eliminazione.

• **Esempio**

Consideriamo il seguente problema di PLI

$$\begin{aligned} z = \max \quad & 11x_1 + 4x_2 + 3x_3 \\ & 3x_1 + 2x_2 + x_3 \leq 7 \\ & x_1 + x_2 + 2x_3 \leq 3 \\ & x_1, x_2, x_3 \geq 0 \quad (\text{intere}) \end{aligned}$$

Soluzione ottima intera: (tramite branch-and-bound):

$$x_1^* = 2, x_2^* = x_3^* = 0 \quad z^* = 22.$$

Rilassamento continuo:

$$\begin{aligned} UB^C = \max \quad & 11x_1 + 4x_2 + 3x_3 \\ & 3x_1 + 2x_2 + x_3 \leq 7 \\ & x_1 + x_2 + 2x_3 \leq 3 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

soluzione ottima (tramite algoritmo del simplesso): $x_1^* = 7/3, x_2^* = x_3^* = 0 \rightarrow UB^C = 77/3 (\simeq 25.66)$

Rilassamento per eliminazione del secondo vincolo:

$$\begin{aligned} UB^E = \max \quad & 11x_1 + 4x_2 + 3x_3 \\ & 3x_1 + 2x_2 + x_3 \leq 7 \\ & x_1, x_2, x_3 \geq 0 \quad (\text{intere}) \end{aligned}$$

Soluzione ottima (tramite branch-and-bound): $x_1^* = 2, x_2^* = 0, x_3^* = 1 \rightarrow UB^E = 25$

Rilassamento lagrangiano del secondo vincolo con $\lambda = 1$:

$$\begin{aligned} UB^L = 3 + \max \quad & 10x_1 + 3x_2 + x_3 \\ & 3x_1 + 2x_2 + x_3 \leq 7 \\ & x_1, x_2, x_3 \geq 0 \quad (\text{intere}) \end{aligned}$$

Soluzione ottima (tramite branch-and-bound): $x_1^* = 2, x_2^* = 0, x_3^* = 1 \rightarrow UB^L = 24$

Rilassamento surrogato con moltiplicatori con $u_1 = u_2 = 1$:

$$UB^S = \max \begin{array}{l} 11x_1 + 4x_2 + 3x_3 \\ 4x_1 + 3x_2 + 3x_3 \leq 10 \\ x_1, x_2, x_3 \geq 0 \text{ (intere)} \end{array}$$

Soluzione ottima (tramite branch-and-bound): $x_1^* = 2, x_2^* = x_3^* = 0 \rightarrow UB^S = 22$. (ammissibile per il problema originale \rightarrow ottima)

• **Esempio: KP-01**

Consideriamo una istanza di KP-01 nella quale gli oggetti sono ordinati secondo valori non crescenti del rapporto $\frac{p_j}{w_j}$. Indichiamo con s l'oggetto critico e con \bar{c} la corrispondente capacità residua

$$s := \min\{i : \sum_{j=1}^i w_j > c\}$$

$$\bar{c} := c - \sum_{j=1}^{s-1} w_j$$

L'upper bound dato dal rilassamento continuo vale

$$UB^C = \sum_{j=1}^{s-1} p_j + \bar{c} \frac{p_s}{w_s}$$

Abbiamo visto in precedenza che una soluzione ottima del rilassamento lagrangiano per un generico moltiplicatore $\lambda > 0$ è sempre intera:

$$x_j = \begin{cases} 1 & \text{se } p_j > \lambda w_j \\ 0 & \text{altrimenti} \end{cases}$$

Prendendo il moltiplicatore $\lambda = \frac{p_s}{w_s}$ si ha

$$x_j = \begin{cases} 1 & \text{se } \frac{p_j}{w_j} > \frac{p_s}{w_s} \\ 0 & \text{altrimenti} \end{cases}$$

ed il valore del corrispondente upper bound è

$$z(L_\lambda(KP)) = \frac{p_s}{w_s} c + \sum_{\frac{p_j}{w_j} > \frac{p_s}{w_s}} \tilde{p}_j = \frac{p_s}{w_s} c + \sum_{\frac{p_j}{w_j} > \frac{p_s}{w_s}} (p_j - \frac{p_s}{w_s} w_j) = \sum_{j=1}^{s-1} p_j + \frac{p_s}{w_s} \left(c - \sum_{j=1}^{s-1} w_j \right)$$

Questo valore è pari al bound del rilassamento continuo. Poichè vale la proprietà di interezza del rilassamento lagrangiano, questo è il miglior upper bound che possiamo trovare con il rilassamento lagrangiano; ciò dimostra che il moltiplicatore ottimo λ^* è pari al rapporto profitto su peso dell'oggetto critico.

• **Esempio: Set Covering**

Consideriamo il problema del set covering

$$\begin{aligned}
 \text{(SCP)} \quad \min \quad & \sum_{j=1}^n c_j x_j \\
 & \sum_{j=1}^n a_{ij} x_j \geq 1 \quad (i = 1, \dots, m) \\
 & x_j \in \{0, 1\} \quad (j = 1, \dots, n)
 \end{aligned}$$

dove le variabili x_j hanno il seguente significato: $x_j = \begin{cases} 1 & \text{se colonna } j \text{ selezionata} \\ 0 & \text{altrimenti} \end{cases}$

Rilassamento surrogato di tutti i vincoli

$$\begin{aligned}
 S_\pi(\text{SCP}) \quad \min \quad & \sum_{j=1}^n c_j x_j \\
 & \sum_{i=1}^m \pi_i \sum_{j=1}^n a_{ij} x_j \geq \sum_{i=1}^m \pi_i \quad (\text{con } \pi_i \geq 0 \quad \forall i) \\
 & x_j \in \{0, 1\} \quad (j = 1, \dots, n)
 \end{aligned}$$

Quindi nel rilassamento surrogato i vincoli di copertura sono sostituiti dall'unico vincolo:

$$\sum_{j=1}^n \left(\sum_{i=1}^m \pi_i a_{ij} \right) x_j \geq \sum_{i=1}^m \pi_i$$

e pertanto il problema equivale ad un problema di KP-01 in forma di minimo in cui sono dati n oggetti, il j -esimo dei quali ha peso $\sum_{i=1}^m \pi_i a_{ij}$ e "costo" c_j , e si vuole determinare un sottoinsieme di oggetti di

costo minimo e di peso complessivo almeno pari a $\sum_{i=1}^m \pi_i$.

Rilassamento lagrangiano di tutti i vincoli

$$\begin{aligned}
 L_\lambda(\text{SCP}) \quad \min \quad & \sum_{j=1}^n c_j x_j + \sum_{i=1}^m \lambda_i \left(1 - \sum_{j=1}^n a_{ij} x_j \right) \quad (\text{con } \lambda_i \geq 0 \quad \forall i) \\
 = \sum_{i=1}^m \lambda_i + \min \quad & \sum_{j=1}^n \tilde{c}_j x_j \quad (\text{con } \tilde{c}_j = c_j - \sum_{i=1}^m a_{ij} \lambda_i) \\
 & x_j \in \{0, 1\} \quad (j = 1, \dots, n)
 \end{aligned}$$

Nel problema lagrangiano, ogni colonna j ha costo $\tilde{c}_j = c_j - \sum_{i=1}^m a_{ij} \lambda_i$ e non c'è alcun vincolo che leghi le variabili associate a colonne differenti

La soluzione del problema rilassato è immediata:

$$x_j = \begin{cases} 1 & \text{se } \tilde{c}_j < 0 \\ 0 & \text{se } \tilde{c}_j \geq 0 \end{cases}$$

ed il corrispondente lower bound vale

$$z(L_\lambda(\text{SCP})) = \sum_{\tilde{c}_j < 0} \tilde{c}_j + \sum_{i=1}^m \lambda_i$$

Es: $n = 10$; $m = 4$;

$$(c_j) = (50, 60, 45, 45, 50, 45, 70, 70, 30, 70)$$

$$(a_{ij}) = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Soluzione ottima intera: (tramite branch-and-bound):

$$(x_j^*) = (0, 0, 0, 1, 0, 0, 0, 0, 1, 0)$$
$$z^* = 75.$$

Rilassamento continuo:

$$(x_j^*) = (\frac{1}{3}, 0, 0, \frac{1}{3}, \frac{1}{3}, 0, 0, 0, \frac{2}{3}, 0)$$
$$z^{LP} = \frac{205}{3} \simeq 68.33.$$

Rilassamento surrogato: si considerino ad esempio i seguenti moltiplicatori:

$$(\pi_i) = (1, 1, 1, 1).$$

Soluzione ottima (tramite branch-and-bound per KP-01):

$$(x_j^*) = (0, 0, 0, 0, 0, 1, 0, 0, 1, 0)$$

(non ammissibile per SCP).

$$z(S_\pi(\text{SCP})) = 75.$$

L'ammissibilità della soluzione surrogata è facile da verificare.

Rilassamento lagrangiano: si considerino ad esempio i seguenti moltiplicatori:

$$(\lambda_i) = (13, 38, 7, 12)$$
$$(\tilde{c}_j) = (-1, 40, 20, 0, 0, 26, 12, 7, -2, 13).$$

Soluzione ottima (immediata):

$$(x_j^*) = (1, 0, 0, 0, 0, 0, 0, 0, 1, 0)$$

(ammissibile per SCP ma non soddisfa la condizione sufficiente per l'ottimalità).

$$z(L_\lambda(\text{SCP})) = 13 + 38 + 7 + 12 - 3 = 67.$$

L'ammissibilità della soluzione lagrangiana è facile da verificare.

6 Problema del commesso viaggiatore

6.1 Eliminazione dei SECs

Nel caso del problema del commesso viaggiatore asimmetrico, eliminando i vincoli di subtour elimination (SECs) dal modello matematico, il problema risultante è un problema di assegnamento. Questo problema è polinomiale in quanto la matrice dei vincoli è totalmente unimodulare; esistono algoritmi efficienti per la soluzione del problema dell'assegnamento che richiedono tempo di calcolo $O(n^3)$.

- Numero di possibili assegnamenti distinti $n!$
- Di questi $(n - 1)!$ sono dei tour

- \rightarrow Mediamente un assegnamento su n è un tour
- Modificando gli elementi della matrice dei costi, possiamo eliminare gli assegnamenti che contengono elementi diagonali, in quanto questi non sono di interesse per ATSP: in tal caso, il numero di possibili assegnamenti distinti si riduce a $\lfloor n!/e + 1/2 \rfloor$
- \rightarrow rilassamento molto forte sul valore della soluzione ottima (su alcune classi di istanze il lower bound fornito è circa pari al 99% del valore della soluzione ottima).

Nella versione simmetrica del problema, l'eliminazione dei vincoli di subtour fornisce un problema rilassato di 2-matching (anche questo polinomiale).

6.1.1 Possibile schema di branching

La soluzione del rilassamento per eliminazione dei vincoli di subtour porta ad un problema di assegnamento, la cui soluzione ottima contiene, in generale, dei sottocicli: un algoritmo branch-and-bound basato sul rilassamento per eliminazione dei SECs deve prevedere uno schema di branching che consenta di "rompere" i sottocicli.

Per eliminare una eventuale soluzione rilassata non ammissibile si può considerare il seguente schema di branching: si prende il ciclo più corto (di lunghezza k) e si generano k figli vietando, a turno, uno degli archi del ciclo. Nell'esempio di figura 6 il ciclo più corto coinvolge 3 archi, per cui dal nodo corrente vengono generati 3 figli nei quali, a turno, ciascun arco è imposto in soluzione. Si può inoltre rafforzare il branching imponendo che, nel generico nodo figlio j ($j = 1, \dots, k$), gli archi di posizione $1, \dots, j - 1$ nel ciclo siano in soluzione. Ovviamente si evita di generare i nodi corrispondenti a condizioni che vietano archi fissati in soluzione ai livelli precedenti.

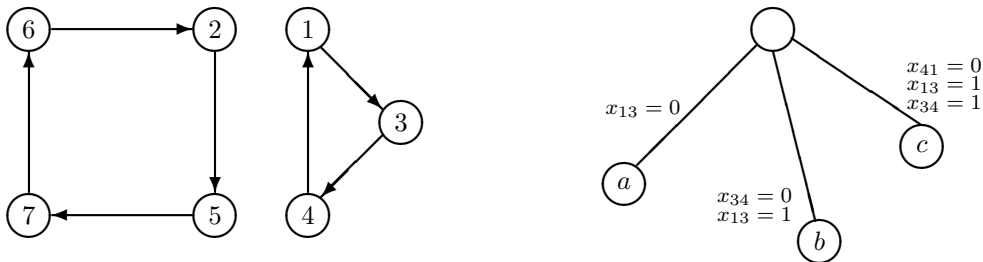


Figura 6: Schema di branching per eliminazione di subtour.

Il sottoproblema da risolvere ad ogni nodo è sempre un problema di assegnamento con alcuni archi fissati ed altri archi proibiti. Per poter tenere in conto di questo fatto basta modificare la matrice dei costi del problema di assegnamento nel seguente modo: se $x_{ij} = 0 \rightarrow c_{ij} = \infty$ (in modo da penalizzare la scelta di un arco proibito), se $x_{ij} = 1 \rightarrow c_{ik} = \infty \forall k \neq j$ (in modo da penalizzare la mancata scelta di un arco fissato).

La soluzione del rilassamento del nodo corrente non è ammissibile per il nodo figlio in quanto un arco (correntemente nella soluzione rilassata) è stato proibito dal branching. Il nodo figlio quindi eredita dal padre un assegnamento ammissibile parziale di dimensione $n - 1$, cui deve aggiungere un solo cammino aumentante; pertanto, ad ogni nodo dell'albero decisionale (a parte il nodo radice), il rilassamento può essere calcolato in modo parametrico in tempo $O(n^2)$.

6.2 Rilassamento 1-albero

Consideriamo una qualunque soluzione ammissibile per un problema di TSP non orientato: il numero di lati incidenti su ciascun vertice è pari a 2 e, immaginando di rimuovere i due lati incidenti su un vertice

a piacere (ad esempio il vertice 1), quel che rimane è un albero ricoprente per il sottografo indotto da $V \setminus \{1\}$. Pertanto, qualunque soluzione ha la seguente struttura:

- (a) ci sono due lati incidenti nel vertice 1;
- (b) togliendo il vertice 1 rimane un albero;
- (c) l'insieme dei lati individuato forma un ciclo hamiltoniano.

Il vincolo (c) è difficile da imporre, per cui posso rimuoverlo ed ottenere una soluzione rilassata nel seguente modo:

- calcolando lo SST sul sottografo ottenuto eliminando il vertice 1;
- aggiungendo i due lati di costo minimo incidenti in 1.

La soluzione così ottenuta si chiama *1-albero*.

Avendo rimosso il vincolo (c) abbiamo ottenuto la possibilità di ottimizzare sulla famiglia degli 1-alberi anziché sulla famiglia dei cicli hamiltoniani; è evidente che la prima contiene la seconda come sottoinsieme stretto, per cui il valore della soluzione trovata sarà sempre minore o uguale al valore della soluzione ottima del problema di partenza.

Matematicamente questo rilassamento può essere ottenuto a partire dal seguente modello matematico, dove è stato aggiunto il vincolo (27) (che sarebbe ridondante), che impone di selezionare esattamente n lati in soluzione.

$$\min \sum_{e \in E} c_e x_e \quad (25)$$

$$\text{subject to} \quad \sum_{e \in \delta(i)} x_e = 2 \quad (i \in V) \quad (26)$$

$$\sum_{e \in E} x_e = n \quad (27)$$

$$\sum_{e \in E(S)} x_e \leq |S| - 1 \quad (S \subseteq V \setminus \{1\}, S \neq \emptyset) \quad (28)$$

$$x_e \in \{0, 1\} \quad (e \in E) \quad (29)$$

Eliminando i vincoli di grado (26) tranne che per il vertice 1 si ottiene il seguente problema: selezionare n lati di costo complessivo minimo, di cui 2 incidenti sul vertice 1, che non formino cicli nel sottografo indotto da $V \setminus \{1\}$. Il problema risultante è quindi l'1-albero.

La qualità (in termini di valore del bound) fornito da questo rilassamento è generalmente abbastanza scarsa per le istanze della letteratura. È possibile migliorare la qualità dei bound ottenuti tenendo conto in modo lagrangiano dei vincoli rilassati. Dato un vettore $(\lambda_1, \dots, \lambda_n)$ di moltiplicatori (non vincolati in segno) il problema lagrangiano associato è

$$L_\lambda(TSP) = \min \left[\sum_{e \in E} c_e x_e + \sum_{i=1}^n \lambda_i \left(2 - \sum_{e \in \delta(i)} x_e \right) \right]$$

con i vincoli (26) (solo per $i = 1$), (27), (28) e (29).

La funzione obiettivo del problema lagrangiano può essere riscritta nel seguente modo

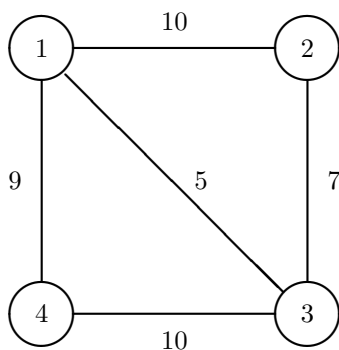
$$2 \sum_{i=1}^n \lambda_i + \sum_{e \in E} \bar{c}_e x_e$$

dove il *costo lagrangiano* di ciascun lato $e = (i, j) \in E$ è definito come $\bar{c}_e := c_e - \lambda_i - \lambda_j$.

Le stesse considerazioni possono essere fatte nel caso del problema del commesso viaggiatore asimmetrico: il corrispondente problema rilassato sarà quello di determinare una *1-arborescenza*.

• **Esempio**

Consideriamo il seguente grafo



La soluzione ottima del rilassamento 1-albero è composta dall'albero $(2, 3), (3, 4)$ di costo $7+10=17$ e dai lati $(1, 3)$ e $(1, 4)$, di costo pari a 5 e 9, rispettivamente. Il corrispondente lower bound è quindi pari a 31. La soluzione rilassata non è però ammissibile per il problema originale, in quanto non è soddisfatto il vincolo di grado per i vertici 2 e 3.

L'idea del rilassamento lagrangiano è di “pesare“ l'importanza dei vari vertici nella soluzione; in particolare, si dovrà cercare di aumentare il moltiplicatore del vertice 2 (in modo da favorire i lati incidenti su questo nodo) e di diminuire il moltiplicatore associato al vertice 3 (in modo da penalizzare l'utilizzo di lati incidenti sul nodo).

Considerando il seguente vettore di moltiplicatori lagrangiani $\lambda = (0, 1, -1, 0)$, il costo degli archi viene modificato nel seguente modo:

i	j	c_e	λ_i	λ_j	\bar{c}_e
1	2	10	0	1	9
1	3	5	0	-1	6
1	4	9	0	0	9
2	3	7	1	-1	7
3	4	10	-1	0	11

Una soluzione ottima del rilassamento è identica a quella ottenuta dal rilassamento per eliminazione, in quanto composta dall'albero $(2, 3), (3, 4)$ e dai lati $(1, 3)$ e $(1, 4)$ (si noti che la soluzione ottenuta prendendo il lato $(1, 2)$ anziché il lato $(1, 4)$ era altrettanto ottima). Essendo però cambiati i costi dei lati, il bound fornito dal rilassamento è pari a $2(0 + 1 - 1 + 0) + 18 + 6 + 9 = 33$.

Prendendo infine il seguente vettore di moltiplicatori $\lambda = (0, 2, -4, 1)$ la soluzione ottima è composta dai lati $(2, 3), (3, 4), (1, 2)$ e $(1, 4)$ ed il corrispondente bound è pari a $2(0 + 2 - 4 + 1) + 13 + 9 + 8 + 8 = 36$. Questa soluzione è ammissibile per il problema di partenza ed è quindi la soluzione ottima (si ricordi che sono stati rilassati solo dei vincoli di uguaglianza).

6.2.1 Possibile schema di branching

Una soluzione ottima del rilassamento 1-albero non soddisfa, in generale, i vincoli di grado per qualche nodo, per cui un algoritmo branch-and-bound basato sul rilassamento 1-albero deve prevedere uno schema di branching che consenta di ripristinare tali vincoli i sottocicli.

Generalmente un algoritmo branch-and-bound di questo tipo richiede di calcolare, al nodo radice, un "buon" lower bound tramite rilassamento lagrangiano 1-albero utilizzando diversi vettori moltiplicatori e di memorizzare il vettore di moltiplicatori cui corrisponde il miglior lower bound e la corrispondente soluzione rilassata. Se la soluzione rilassata non è ammissibile, si scelgono un vertice v con grado ≥ 3 e due lati a e b incidenti su v ; il problema è suddiviso in tre sottoproblemi ottenuti vietando, a turno, il lato a , il lato b , oppure tutti gli altri lati incidenti su v (vedi figura 7)

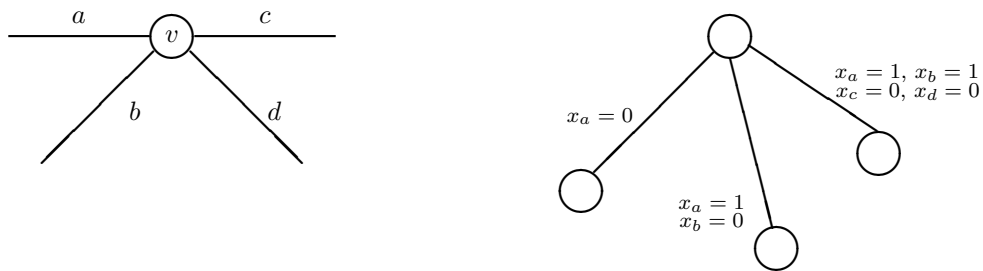


Figura 7: Schema di branching per rispettare i vincoli di grado.

Ai nodi successivi dell'albero decisionale il calcolo del lower bound deve essere effettuato in modo molto veloce, per cui il numero di iterazioni di subgradiente è generalmente molto ridotto; al limite si possono utilizzare ad ogni nodo i moltiplicatori ottimi trovati al nodo radice, in quanto i moltiplicatori ottimi del generico nodo non saranno troppo diversi dai moltiplicatori ottimi del nodo radice.