

Il rilassamento Lagrangiano nella soluzione di problemi di Ottimizzazione Combinatoria

P. Detti *

7 Novembre 2000

1 La tecnica Lagrangiana

L'applicazione della tecnica Lagrangiana ai problemi di Ottimizzazione Combinatoria e più in generale ai problemi di programmazione matematica intera assume importanza negli anni 70, quando Held e Karp [6, 7] propongono un approccio Lagrangiano per il problema del TSP, ma l'idea originaria probabilmente risale al 1955 ed è dovuta a Lorie e Savage [9]. Riferimenti classici dalla letteratura sulla tecnica di rilassamento Lagrangiano sono i lavori di Geoffrion [4], Shapiro [12] e Fisher [2, 3]. Tutt'oggi la tecnica Lagrangiana è impiegata con successo nella messa a punto di algoritmi di soluzione per uno svariato numero di problemi di Programmazione Lineare Intera Mista.

L'intuizione alla base dell'approccio Lagrangiano consiste nel vedere molti dei problemi di Ottimizzazione Combinatoria "difficili" come problemi "facili", complicati da un insieme, in genere di dimensione ridotta, di vincoli. Il *Rilassamento Lagrangiano*, consiste nell'eliminazione dei vincoli "indesiderati" e nell'inserimento di questi nella funzione obiettivo in modo tale che, nella soluzione del problema rilassato, se ne tenga conto in qualche misura.

Si consideri ad esempio il seguente problema di Programmazione Lineare Intera PLI_1 :

$$\min cx \tag{1}$$

tale che

$$Ax \geq b \tag{2}$$

$$Cx \geq d \tag{3}$$

$$x \in \{0, 1\}^n \tag{4}$$

*Dipartimento di Informatica e Automazione - Università di Roma Tre, via della vasca navale, 79 - 00146 Roma (Italy), e-mail: detti@dia.uniroma3.it

dove A e C sono matrici di dimensioni, rispettivamente, $m \times n$ e $m_1 \times n_1$, b e d sono vettori di dimensioni, rispettivamente, m e m_1 , x è un vettore di dimensione n e c è un vettore riga di dimensione n . Si supponga che PLI_1 sia un problema di difficile risoluzione, mentre sia facile risolvere il problema ottenuto “inserendo” nella funzione obiettivo i vincoli (2) nel seguente modo:

$$\min cx + \lambda(b - Ax) \quad (5)$$

tale che

$$Cx \geq d \quad (6)$$

$$x \in \{0, 1\}^n \quad (7)$$

dove λ è un vettore riga non negativo di numeri reali di dimensione m , denominato vettore dei *moltiplicatori di Lagrange*. Il problema (5)–(7) è detto *problema Lagrangiano* e la funzione

$$L(\lambda) = \min\{cx + \lambda(b - Ax) : Cx \geq d, x \in \{0, 1\}^n\} \quad (8)$$

è denominata *funzione Lagrangiana*. Il problema Lagrangiano risultante deve essere naturalmente più semplice da risolvere, ed ha la proprietà che la soluzione ottima costituisce un bound sull’ottimo del problema originario: rispettivamente un *lower bound* se il problema considerato è di minimo, un *upper bound* se il problema è di massimo.

PROPOSIZIONE 1.1 *Sia $z(PLI_1)$ il valore della soluzione ottima di PLI_1 , per ogni vettore non negativo $\lambda \in \mathbf{R}^{m+}$ si ha*

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

Dimostrazione. Poiché $Ax \geq b$ e $\lambda \geq 0$, allora $\lambda(b - Ax) \leq 0$ per ogni soluzione ammissibile di PLI_1 e quindi per ogni fissato vettore di moltiplicatori si ha:

$$\begin{aligned} z(PLI_1) &= \min\{cx : Ax \geq b, Cx \geq d, x \in \{0, 1\}^n\} \geq \\ &\min\{cx + \lambda(b - Ax) : Ax \geq b, Cx \geq d, x \in \{0, 1\}^n\}. \end{aligned} \quad (10)$$

Se dal secondo termine della (10) si rimuovono i vincoli $Ax \geq b$, poiché si effettua un allargamento della regione ammissibile del problema, si ottiene:

$$\begin{aligned} &\min\{cx + \lambda(b - Ax) : Ax \geq b, Cx \geq d, x \in \{0, 1\}^n\} \geq \\ &\min\{cx + \lambda(b - Ax) : Cx \geq d, x \in \{0, 1\}^n\} = L(\lambda). \end{aligned} \quad (11)$$

□

Il miglior lower bound ottenibile da $L(\lambda)$ è ottenuto risolvendo il seguente problema di ottimizzazione ed è denominato *duale Lagrangiano*:

$$L(\lambda^*) = \max\{L(\lambda) : \lambda \geq 0\}. \quad (12)$$

Tale denominazione è giustificata dalle seguenti considerazioni. Sia dato il problema di Programmazione Lineare $LP = \min\{cx : Ax \geq b, x \geq 0\}$. Se si effettua il rilassamento Lagrangiano dei vincoli $Ax \geq b$ il duale Lagrangiano è:

$$L_{LP}(\lambda^*) = \max\{\min\{(c - \lambda A)x + \lambda b : x \geq 0\}, \lambda \geq 0\}. \quad (13)$$

Per λ fissato, nel calcolo della (13), ogni componente di x tale che la corrispondente componente del vettore $(c - \lambda A)$ è negativa deve essere posta a $+\infty$ e quindi $L(\lambda) = -\infty$. Nel problema di massimo, ha senso considerare, quindi, solamente quei vettori λ per cui $c - \lambda A \geq 0$, ottenendo il problema:

$$L_{LP}(\lambda^*) = \max\{\lambda b : c - \lambda A \geq 0, \lambda \geq 0\} \quad (14)$$

corrispondente al duale di LP .

Con riferimento al problema PLI_1 , tra rilassamento Lineare e rilassamento Lagrangiano vale il seguente risultato.

PROPOSIZIONE 1.2 *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 del duale Lagrangiano. Sia $z(PL_1)$ il valore ottenuto risolvendo il rilassamento Lineare del problema PLI_1 . Vale la condizione:*

$$L(\lambda^*) \geq z(PL_1) \quad (15)$$

Dimostrazione.

$$\begin{aligned} L(\lambda^*) &= \max\{\min\{cx + \lambda(b - Ax) : Cx \geq d, x \in \{0, 1\}^n, \lambda \geq 0\} \\ &\geq \max\{\min\{cx + \lambda(b - Ax) : Cx \geq d, 0 \leq x \leq 1\} : \lambda \geq 0\} \\ &\geq \max\{\min\{cx + \lambda(b - Ax) + \mu(d - Cx) : x \geq 0, -x \geq -1\} : \lambda \geq 0, \mu \geq 0\} \\ &\geq \max\{\min\{cx + \lambda(b - Ax) + \mu(d - Cx) + \nu(x - 1) : x \geq 0\} : \lambda \geq 0, \mu \geq 0, \nu \geq 0\} \\ &= \max\{\min\{(c - \lambda A - \mu C + \nu)x + \lambda b + \mu d - \nu \mathbf{1} : x \geq 0\} : \lambda \geq 0, \mu \geq 0, \nu \geq 0\} \\ &= \max\{\lambda b + \mu d - \nu \mathbf{1} : c - \lambda A - \mu C + \nu \geq 0, \lambda \geq 0, \mu \geq 0, \nu \geq 0\} \\ &=^* \min\{cx : Ax \geq b, Cx \geq d, 0 \leq x \leq 1\} = z(PL_1) \end{aligned} \quad (16)$$

L'ultimo passaggio ($=^*$) si ottiene calcolando il problema duale. \square

La Proposizione 1.2 stabilisce, quindi, che il rilassamento Lagrangiano è almeno tanto buono quanto il rilassamento Lineare. Tale risultato giustifica, almeno in linea teorica, l'impiego dei bound Lagrangiani nella messa a punto di algoritmi di enumerazione implicita di tipo branch and bound.

Tra rilassamento Lineare e rilassamento Lagrangiano vale, inoltre, il seguente risultato.

COROLLARIO 1.3 *Sia $P = \min\{cx : x \in X\}$ un problema di Programmazione Lineare Intera, se il poliedro $Q = \text{conv}\{x \in X\}$ ha tutti i vertici interi, se, cioè, i vincoli di interezza sono ridondanti, allora il valore del rilassamento Lineare z_{R_p} è pari al duale Lagrangiano.*

Se vale il Corollario 1.3 si dice che il poliedro Q gode della *proprietà di interezza*.

1.1 Funzioni concave

Una funzione $f : \mathbf{R}^n \rightarrow \mathbf{R}$ si dice concava se, per ogni coppia di vertici $x^1, x^2 \in \mathbf{R}^n$ e per ogni reale α , con $0 \leq \alpha \leq 1$, si ha:

$$\alpha f(x^1) + (1 - \alpha)f(x^2) \leq f(\alpha x^1 + (1 - \alpha)x^2). \quad (17)$$

Si consideri l'insieme F definito nello spazio \mathbf{R}^{n+1} nel seguente modo:

$$F = \left\{ y \equiv \begin{bmatrix} x \\ z \end{bmatrix} : z \leq f(x) \right\}, \quad (18)$$

La cui frontiera è costituita dal grafico di f , cioè dall'insieme dei punti y tali che $z = f(x)$. Una conseguenza della definizione di funzione concava è enunciata dalla seguente proposizione.

PROPOSIZIONE 1.4 *Data una funzione concava $f : \mathbf{R}^n \rightarrow \mathbf{R}$ l'insieme F definito in (18) è convesso.*

Dimostrazione. Siano $y^1 \equiv \begin{bmatrix} x^1 \\ z^1 \end{bmatrix}$ e $y^2 \equiv \begin{bmatrix} x^2 \\ z^2 \end{bmatrix}$ due generici punti in F e α un reale tale che $0 \leq \alpha \leq 1$. Definiamo $y^3 \equiv \begin{bmatrix} x^3 \\ z^3 \end{bmatrix} = \alpha y^1 + (1 - \alpha)y^2$. Abbiamo :

$$\begin{aligned} x^3 &= \alpha x^1 + (1 - \alpha)x^2 \\ z^3 &= \alpha z^1 + (1 - \alpha)z^2 \leq \\ &\alpha f(x^1) + (1 - \alpha)f(x^2) \leq \\ &f(\alpha x^1 + (1 - \alpha)x^2) = f(x^3), \end{aligned} \quad (19)$$

e quindi, per la definizione dell'insieme F , il punto y^3 appartiene ad F , dimostrando la tesi. \square

La seguente proposizione descrive una ben nota proprietà degli insiemi convessi.

PROPOSIZIONE 1.5 Sia $F \subseteq \mathbf{R}^m$ un insieme convesso e y' un punto sulla sua frontiera. Allora esiste un vettore $w \in \mathbf{R}^m$ non nullo, definito a meno di una costante moltiplicativa positiva, tale che l'iperpiano $w^T(y - y') = 0$ passante per y' è di supporto per l'insieme F , cioè soddisfa la seguente relazione:

$$w^T(y - y') = 0 \quad \forall y \in F. \quad (20)$$

Usando la precedente proposizione, possiamo dimostrare la seguente importante proprietà.

PROPOSIZIONE 1.6 Sia $f : \mathbf{R}^n \rightarrow \mathbf{R}$ una funzione concava e x' un generico punto di \mathbf{R}^n . Allora esiste un vettore $s \in \mathbf{R}^n$ tale che per ogni $x \in \mathbf{R}^n$ risulta:

$$s^T(x - x') \geq f(x) - f(x'). \quad (21)$$

Dimostrazione. Si consideri l'insieme F precedentemente definito nello spazio \mathbf{R}^{n+1} e poniamo $y' \equiv \begin{bmatrix} x' \\ z' \end{bmatrix}$ con $z' = f(x')$. Dalla Proposizione 1.5 segue che esiste un vettore non nullo $w \equiv \begin{bmatrix} s \\ t \end{bmatrix}$ ($s \in \mathbf{R}^n, t \in \mathbf{R}$) tale che $w^T(y - y') \geq 0 \quad \forall y \in F$. Dimostriamo che $t < 0$.

Infatti, scegliendo y tale che $y \equiv \begin{bmatrix} x' \\ z \end{bmatrix}$, con $z < z'$, abbiamo:

$$w^T(y - y') = \begin{bmatrix} s \\ t \end{bmatrix}^T \left(\begin{bmatrix} x' \\ z \end{bmatrix} - \begin{bmatrix} x' \\ z' \end{bmatrix} \right) = t(z - z') \geq 0$$

che implica $t \leq 0$. Se poi fosse $t = 0$ per ogni $y = \begin{bmatrix} x \\ z \end{bmatrix} \in F$ avremmo:

$$w^T(y - y') = \begin{bmatrix} s \\ t \end{bmatrix}^T \left(\begin{bmatrix} x \\ z \end{bmatrix} - \begin{bmatrix} x' \\ z' \end{bmatrix} \right) = s^T(x - x') \geq 0.$$

Tale equazione, soddisfatta da ogni $x \in \mathbf{R}^n$, implicherebbe $s = \mathbf{0}_n$ contro l'ipotesi che $w \equiv \begin{bmatrix} s \\ t \end{bmatrix}$ è un vettore non nullo. Abbiamo quindi $t < 0$ e, dal momento che il vettore w è definito a meno di una costante moltiplicativa positiva, possiamo assumere senza perdita di generalità $t = -1$. Abbiamo quindi, scegliendo y sulla frontiera di F , ponendo cioè $y = \begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} x \\ f(x) \end{bmatrix}$:

$$w^T(y - y') = \begin{bmatrix} s \\ t \end{bmatrix}^T \left(\begin{bmatrix} x \\ z \end{bmatrix} - \begin{bmatrix} x' \\ z' \end{bmatrix} \right) = s^T(x - x') - (z - z') = s^T(x - x') - (f(x) - f(x')) \geq 0.$$

da cui segue la tesi. \square

1.2 Condizioni di ottimalità del problema Lagrangiano

Se, per un dato valore di λ , 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 . Vale però la seguente condizione di ottimalità.

PROPOSIZIONE 1.7 *Sia $L(\lambda)$ la funzione Lagrangiana ottenuta da PLI_1 effettuando il rilassamento Lagrangiano dei vincoli $Ax \geq b$. Si supponga che per un fissato valore λ^* dei moltiplicatori siano verificate le due condizioni:*

- i. *la soluzione x^* del problema Lagrangiano sia ammissibile per PLI_1 ;*
- ii. *la condizione di complementarità*

$$\lambda^*(Ax^* - b) = 0 \tag{22}$$

sia soddisfatta.

Allora x^ è una soluzione ottima anche per il problema PLI_1 .*

Dimostrazione. Per definizione $L(\lambda) = \min\{cx + \lambda(b - Ax)\}$. Dalla condizione ii. si ottiene che $L(\lambda^*) = cx^*$. Inoltre poiché $L(\lambda^*) \leq cx$ e x^* è ammissibile per il problema originario si deduce che x^* 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, poiché la condizione ii. è sempre soddisfatta, condizione sufficiente di ottimalità per il problema originario è che la soluzione x^* del problema Lagrangiano sia ammissibile per PLI_1 . In generale, il duale Lagrangiano è minore della soluzione ottima del problema originario ($L(\lambda^*) < z(PLI_1)$, per un problema di minimo). La differenza $z(PLI_1) - L(\lambda^*)$ è detta *gap di dualità*.

Qualora le condizioni della proposizione precedente non siano verificate le soluzioni del problema Lagrangiano possono essere utilizzate comunque nella ricerca della soluzione ottima: le soluzioni ammissibili per il problema originario, ma non dimostrate ottime, costituiscono un upper bound dell'ottimo (problema di minimo) e sono candidate ad essere soluzioni ottime, mentre le soluzioni Lagrangiane non ammissibili per il problema originario possono essere *modificate* e rese ammissibili dando luogo spesso a *buone* soluzioni.

1.3 Calcolo del Duale Lagrangiano

Si consideri la funzione Lagrangiana (8), e sia X l'insieme di punti $X \equiv \{x^1, \dots, x^T\} \equiv \{Cx \geq d, x \in \{0, 1\}^n\}$. Si supponga che X sia costituito da un insieme finito di punti. Si può dimostrare che la funzione (8) è una funzione continua, concava, non differenziabile.

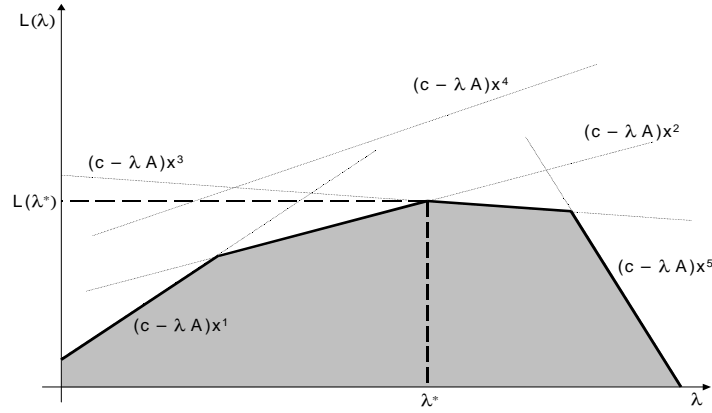


Figura 1: Andamento di $L(\lambda)$.

Il duale Lagrangiano (12) può essere calcolato, quindi, risolvendo un problema di ottimizzazione non differenziabile. Un altro metodo di soluzione consiste nel formulare (12) come problema di Programmazione Lineare. Il problema del calcolo del duale Lagrangiano (12) è infatti equivalente al problema:

$$\max v \quad (23)$$

tale che

$$v \leq (c - \lambda A)x^t + \lambda b, \quad t = 1, \dots, T \quad (24)$$

$$\lambda \geq 0 \quad (25)$$

Notare che i vincoli (24) sono pari al numero di elementi dell'insieme X . Da questo punto di vista, è facile verificare come la funzione Lagrangiana $L : \lambda \rightarrow L(\lambda)$ sia definita dalla frontiera inferiore di un insieme composto da un insieme finito di funzioni lineari. Nel caso monodimensionale ($|\lambda| = 1$) e per $T = 5$, l'andamento della funzione Lagrangiana, il valore ottimo dei moltiplicatori λ^* ed il corrispondente duale Lagrangiano sono mostrati in Figura 1.

Una delle tecniche per risolvere il problema (23)–(25) è il metodo di *Decomposizione di Dantzig-Wolfe* [10, 1]. Uno degli svantaggi di tale metodo è quello di richiedere la risoluzione di una serie di problemi di Programmazione Lineare che possono essere onerosi dal punto di vista computazionale.

Un altro approccio di soluzione per il problema (12) deriva dalle tecniche di ottimizzazione non differenziabili. Nei problemi di ottimizzazione con funzione obiettivo non lineare $f(\lambda)$, con λ vettore n -dimensionale, la maggior parte dei metodi utilizzati si

basano sul concetto di *gradiente* $\nabla f(\lambda)$ della funzione f , definito come il vettore riga $(\partial f(\lambda)/\partial \lambda_1, \partial f(\lambda)/\partial \lambda_2, \dots, \partial f(\lambda)/\partial \lambda_n)$. La *derivata direzionale* di f lungo la direzione d soddisfa la relazione:

$$\lim_{\theta \rightarrow 0} \frac{f(\lambda + \theta d) - f(\lambda)}{\theta} = \nabla f(\lambda) d. \quad (26)$$

Ciò significa che se si sceglie una direzione d di “spostamento”, tale che $\nabla f(\lambda) d > 0$ con θ (detto *lunghezza del passo*) opportunamente piccolo, lo spostamento avviene in una direzione di avvicinamento di un punto di ottimo della funzione. I metodi basati su tali considerazioni sono conosciuti in letteratura come *metodi del gradiente*. In generale, tali metodi, data una funzione concava non differenziabile $f(\lambda)$, con λ^* un punto di massimo per $f(\lambda)$, consistono nel generare una successione di punti $\lambda^{(0)}, \lambda^{(1)}, \dots, \lambda^{(k)}, \dots$ ottenuta da una relazione di ricorrenza del tipo $\lambda^{(t+1)} = \lambda^{(t)} + \theta^{(t)} d^{(t)}$, tale che

$$\lim_{t \rightarrow \infty} \lambda^{(t)} = \lambda^*. \quad (27)$$

Una direzione di spostamento d può essere calcolata sfruttando le 2 seguenti proposizioni.

PROPOSIZIONE 1.8 *Data una funzione concava $f : \mathbf{R}^n \rightarrow \mathbf{R}$ e un punto $\lambda' \in \mathbf{R}^n$, un vettore $s \in \mathbf{R}^n$, che per ogni $\lambda \in \mathbf{R}^n$ soddisfa la relazione*

$$s^T(\lambda - \lambda') \geq f(\lambda) - f(\lambda') \quad (28)$$

è detto subgradiente di f in λ'

PROPOSIZIONE 1.9 *Data una funzione concava $f : \mathbf{R}^n \rightarrow \mathbf{R}$ da massimizzare, un punto $\lambda^{(t)} \in \mathbf{R}^n$ non di massimo per f e un subgradiente $s^{(t)}$ di f in $\lambda^{(t)}$, la direzione definita da $d^{(t)} = \frac{s^{(t)}}{\|s^{(t)}\|}$ è di avvicinamento ad un qualunque punto di massimo $\lambda^* \in \mathbf{R}^n$ per f , dove $\|s^{(t)}\| = \sqrt{s^{(t)T} s^{(t)}}$.*

Dimostrazione. Dalla (28) segue che, per ogni $\lambda \in \mathbf{R}^n$ tale che $f(\lambda) > f(\lambda^{(t)})$, risulta $s^{(t)T}(\lambda - \lambda^{(t)}) > 0$. Ciò vale in particolare per $\lambda = \lambda^*$ dimostrando la tesi. □

Dalla definizione di subgradiente si può dedurre che λ' è un punto di ottimo per la funzione $f(\lambda)$ se $\mathbf{0}$ è un subgradiente di f in λ' . Dalla (28) segue infatti che $\mathbf{0}(\lambda - \lambda') \geq f(\lambda) - f(\lambda') \forall \lambda \in \mathbf{R}^n$, e quindi $f(\lambda') \geq f(\lambda) \forall \lambda \in \mathbf{R}^n$.

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

PROPOSIZIONE 1.10 *Siano $\lambda^{(k)} \in \mathbf{R}^n$ e $x^{(k)} \in \{0, 1\}^n$ due vettori tali che*

$$L(\lambda^{(k)}) = \min\{cx + \lambda^k(b - Ax) : Cx \geq d, x \in \{0, 1\}^n\} = \quad (29)$$

$$cx^{(k)} + \lambda^{(k)}(b - Ax^{(k)}) \quad (30)$$

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

Dimostrazione.

Per un generico vettore λ si ottiene

$$L(\lambda) = \min\{cx + \lambda(b - Ax) : Cx \geq d, x \in \{0, 1\}^n\} \leq \quad (31)$$

$$cx^{(k)} + \lambda(b - Ax^{(k)}) = \quad (32)$$

$$cx^{(k)} + \lambda^{(k)}(b - Ax^{(k)}) + (\lambda - \lambda^{(k)})(b - Ax^{(k)}) = \quad (33)$$

$$L(\lambda^{(k)}) + s(\lambda - \lambda^{(k)}) \quad (34)$$

cioè $(b - Ax^{(k)})(\lambda - \lambda^{(k)}) \geq L(\lambda) - L(\lambda^{(k)})$. Ricordando la definizione di subgradiente si ha la tesi. \square

Algoritmo del subgradiente

Input: funzione da massimizzare $L(\lambda)$, punto iniziale λ^0 , massimo numero di iterazioni T .

1. Inizializzazione $t = 1$.
2. Per un dato valore $\lambda^{(t-1)}$, calcola la soluzione x^t del problema Lagrangiano (5)–(7).
3. Dato il punto $x^{(t)}$, calcola un subgradiente $s^{(t)}$ di $L(\lambda)$. Se $s^{(t)} = 0$ allora stop, altrimenti vai al passo 4.
4. $\lambda^{(t)} := \lambda^{(t-1)} + \theta^{(t)}s^{(t)}$, con $\theta^{(t)}$ scelto in modo tale che le condizioni (35), (36) siano soddisfatte.
5. Poni $t := t + 1$, se $t > T$ stop, altrimenti vai al passo 2.

Figura 2: Algoritmo del subgradiente

Nella messa a punto di un metodo del gradiente un altro aspetto importantissimo riguarda la scelta del *passo* $\theta^{(t)}$. Il modo con cui tale scelta è realizzata determina la convergenza o meno della sequenza di punti $\lambda^{(t)}$ ad un punto di massimo λ^* , e, nel caso in cui si abbia tale convergenza, ne influenza enormemente la velocità. In particolare, si può dimostrare che se la sequenza $\theta^{(t)}$ per $k = 0, 1, \dots$ soddisfa le condizioni [11]:

$$\lim_{t \rightarrow \infty} \theta^{(t)} = 0 \quad (35)$$

$$\sum_{t=0}^k \theta^{(t)} \rightarrow \infty (t \rightarrow \infty) \quad (36)$$

allora la successione di punti generata attraverso la relazione $\lambda^{(t+1)} = \lambda^{(t)} + \theta^{(t)} s^{(t)} / \|s^{(t)}\|$ converge ad un punto di ottimo. In Figura 2 è riportato uno schema dell'algoritmo del subgradiente.

In pratica, un metodo per la scelta della lunghezza del passo $\theta^{(t)}$, dato un subgradiente $s^{(t)}$, soddisfacente le condizioni (35) e (36) è dato dalla relazione [11, 7]:

$$\theta^{(t)} = \alpha_t \frac{UB - L(\lambda^{(t)})}{\|s^{(t)}\|} \quad (37)$$

dove UB è un upper bound per il problema originario e α_k è uno scalare compreso tra 0 e 2.

Il metodo del subgradiente è un metodo di approssimazione non polinomiale, e, in linea teorica, è garantita la convergenza al vettore ottimo di moltiplicatori di Lagrange λ^* in un numero infinito di passi. Inoltre, uno dei maggiori inconvenienti del metodo risiede nel 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), tale metodo è uno dei più utilizzati.

Altri metodi di ottimizzazione, come ad esempio, i metodi di bundle, richiedono la soluzione di un problema quadratico per la generazione della direzione di spostamento d . Tali metodi hanno, in generale, migliori proprietà di convergenza locali, ma richiedono ad ogni iterazione un maggiore carico computazionale. Di conseguenza all'aumentare delle dimensioni del problema possono richiedere tempi di calcolo più lunghi.

2 Applicazione della tecnica Lagrangiana

Uno degli impieghi principali della tecnica Lagrangiana di rilassamento nella soluzione di problemi di Ottimizzazione Combinatoria, è legata al calcolo di bound da utilizzare in schemi di enumerazione implicita quali ad esempio algoritmi di branch and bound.

Tra i recenti sviluppi e applicazioni del rilassamento Lagrangiano vi sono i risultati di Hoogeveen e van de Velde [8] e van de Velde [13, 14] che presentano approcci Lagrangiani per la soluzione di vari problemi di *machine scheduling*. In particolare, in [8] viene presentata una tecnica che impiega variabili di slack per rafforzare i bound Lagrangiani. Guignard [5] presenta procedure per il miglioramento dei bound Lagrangiani attraverso l'individuazione ed il rilassamento di tagli validi attraverso uno schema di

“Relax-and-Cut”. Tali tagli sono individuati, durante il processo di approssimazione del duale, dall’analisi delle soluzioni del problema Lagrangiano.

Un’altra nota tecnica per il calcolo di bound è il Rilassamento Lineare. Data una formulazione lineare intera per un problema \mathcal{NP} –hard, la tecnica di rilassamento Lineare consiste, come è noto, nella rimozione dei vincoli di interezza sulle variabili. La soluzione del corrispondente problema Lineare è un bound (un lower bound se il problema è di minimo) alla soluzione ottima del problema originale.

La bontà di una tecnica di rilassamento dipende principalmente da due fattori:

- qualità del bound prodotto;
- velocità di calcolo del bound.

La scelta di un rilassamento, in generale, non è immediata: può essere preferibile, ad esempio, scegliere bound meno “forti”, ma computazionalmente meno onerosi, che bound migliori a fronte di tempi di calcolo elevati.

Come mostrato nel Paragrafo 1, per una data formulazione di PLI, il duale Lagrangiano è maggiore od uguale al valore del rilassamento Lineare, mentre i bound Lagrangiano e Lineare sono uguali se la formulazione gode della proprietà di interezza. Tuttavia, anche in quest’ultimo caso, se il problema Lagrangiano è risolvibile con poco sforzo computazionale, il calcolo del duale Lagrangiano può essere più rapido e, quindi, può preferirsi al calcolo del rilassamento Lineare.

D’altra parte, la tecnica Lagrangiana è una metodologia che trae vantaggio dalla specificità del problema considerato (quali, ad esempio, proprietà particolari che rendono il problema rilassato semplice da risolvere in modo combinatorio). L’applicazione dell’approccio Lagrangiano presuppone, quindi, non solo un’analisi di differenti formulazioni dello stesso problema, ma anche, per una stessa formulazione, un’analisi di diversi rilassamenti ottenibili dalla “dualizzazione” di diversi insiemi di vincoli. La scelta di una formulazione e di un rilassamento può effettuarsi sia con metodi analitici (Wolsey [15] considera differenti formulazioni e rilassamenti per il problema $1|r_i|\sum_i w_i C_i$, e stabilisce in modo analitico gerarchie tra rilassamenti), sia in modo empirico. In generale, il rilassamento Lagrangiano di un minor numero di vincoli è preferibile, sia perché il problema Lagrangiano, essendo di dimensioni ridotte, risulta essere più semplice da risolvere, sia perché meno sono i vincoli rimossi maggiore è la “vicinanza” del problema rilassato al problema originale.

*Riferimenti

- [1] G. B. Dantzig, P. Wolf, *Decomposition principle for linear programs*, Operations Research, 8, 101–111, 1960.
- [2] M. L. Fisher, *The Lagrangian relaxation method for solving integer programming problems*, Management Science, 27, 1–18, 1981.

- [3] M. L. Fisher, *An application oriented guide to the Lagrangian relaxation*, Interfaces, 15, 10–21, 1985.
- [4] A. Geoffrion, *Lagrangian Relaxation and its uses in Integer Programming*, Mathematical Programming Study, 2, 82–114, 1974.
- [5] M. Guignard, *Efficient cuts in Lagrangean ‘Relax-and-Cut’ schemes.*, European Journal of Operational Research, 105, 216–223, 1998.
- [6] M. Held, R. M. Karp, *The traveling salesman problem and minimum spanning trees*, Operations Research, 18, 1138–1162, 1970.
- [7] M. Held, R. M. Karp, *The traveling salesman problem and minimum spanning trees: Part II*, Mathematical Programming, 1, 62–88, 1971.
- [8] J.A. Hoogeveen, S.L. Van de Velde, *Stronger Lagrangian bounds by use of the slack variables: applications to machine scheduling problem*, Mathematical Programming, 70 (2), 173–190, 1995.
- [9] J. Lorie, L. I. Savage, *Three problems in capital rationing*, Journal of Business, 229–239, 1955
- [10] T. L. Magnanti, J. F. Shapiro, M. H. Wagner, *Generalized linear programming solves the dual*, Management Science, 22, 1195–1204, 1976.
- [11] B. T. Polyak, *Minimization of Unsmooth Functionals*, Computational Mathematics and Mathematical Physics, 9, 14–29, 1969.
- [12] J. F. Shapiro, *A survey of Lagrangean Techniques for Discrete Optimization*, Annals of Discrete Mathematics, 5, 113–138, 1979.
- [13] S.L. van de Velde, *Machine Scheduling and Lagrangian Relaxation*, PhD Thesis, CWI, Amsterdam (The Netherlands), 1991.
- [14] S.L. van de Velde, *Dual decomposition of a single-machine scheduling problem*, Mathematical Programming 69, 413–428, 1995.
- [15] L. A. Wolsey, *Formulating Single-Machine Scheduling Problems with Precedence Constraints*, CORE Discussion paper 8294, Catholic University of Louvain, Louvain-la-Neuve, 1991.