

Capitolo 17

Algoritmi ai punti interni

Lo sviluppo degli algoritmi ai punti interni per la risoluzione di problemi di programmazione lineare ha origine con l'articolo di Karmarkar [1984], in cui veniva proposto un algoritmo polinomiale per la programmazione lineare che, a differenza del precedente algoritmo di Khacyan [1979], era anche praticamente efficiente. Come spesso succede, il nuovo approccio gettò luce diversa su metodi di calcolo precedenti che, apparentemente, non avevano suscitato sufficiente 'fiducia' nell'ambiente scientifico e quindi non erano stati studiati a fondo. La rivitalizzazione di questi metodi, conseguente al successo del metodo di Karmarkar, ha creato un nuovo campo di indagine, i cosiddetti *metodi ai punti interni*. In modo semplificato i metodi ai punti interni producono una successione di punti ammissibili e interni al poliedro che definisce l'istanza. L'iterazione da un lato tende verso il minimo (ovviamente un punto di frontiera), ma dall'altro è respinta dalla frontiera del poliedro e quindi si muove verso l'ottimo, tenendosi lontano il più possibile dalla frontiera. Questa paretofobia ha l'effetto di non far cadere l'iterazione nella trappola combinatoria dei vertici, in cui cade invece il simplesso, e quindi è, alla lunga, responsabile della polinomialità del metodo.

L'algoritmo di Karmarkar si basa su trasformazioni proiettive dello spazio che garantiscono un decremento geometrico della funzione obiettivo ed elevano una barriera sufficientemente alta verso la frontiera. Metodi di barriera erano stati proposti precedentemente da Fiacco e McCormick [1968] nel contesto della programmazione non lineare. Non si realizzò evidentemente che le condizioni di complementarità, che sono non lineari, potevano essere risolte con il metodo di Newton, previa una loro riscrittura che le rendesse adatte ai metodi di barriera.

I metodi ai punti interni che si sono dimostrati più efficaci sono quelli che si basano su una formulazione primale-duale. Sull'argomento sono disponibili approfondite monografie quali Nesterov e Nemirovskij [1994] e Wright [1997]. In particolare la seconda presenta in modo accessibile lo stato dell'arte nella nuova disciplina. Gli algoritmi ai punti interni primali-duali possono competere alla pari con il metodo del simplesso al punto che i maggiori codici commerciali ormai offrono anche questa opzione. La strategia algoritmica migliore sembra essere quella di cominciare a risolvere un problema con un metodo ai punti interni, in modo da arrivare abbastanza velocemente in una zona del poliedro vicina al vertice ottimo, e poi spostare l'iterazione su un vertice e da lì continuare con il metodo del simplesso.

L'algoritmo di Karmarkar, iniziatore delle tecniche ai punti interni, è stato sopravanzato, in termini di efficienza, dagli algoritmi primali-duali. Si ritiene comunque utile presentarlo sia per motivi storici sia per i concetti nuovi a cui fa riferimento. Le tecniche ai punti interni non si limitano inoltre alla programmazione lineare. Anche i problemi di complementarità lineare per matrici positive semidefinite si prestano in modo naturale ad essere risolti con i metodi ai punti interni. Da questi si è sviluppata la programmazione semidefinita che verrà presentata brevemente in conclusione.

17.1. L'algoritmo di Karmarkar

Per poter essere risolto con l'algoritmo di Karmarkar, un problema di programmazione lineare deve essere espresso in una forma particolare, che definiamo *forma di Karmarkar*, e che è caratterizzata dalle seguenti ipotesi:

- 1) i vincoli sono del seguente tipo

$$\begin{aligned} v &:= \min \quad cx \\ Ax &= 0 \\ \mathbf{1}x &= 1 \quad x \geq 0 \end{aligned} \tag{17.1}$$

- 2) il valore ottimo deve essere nullo ($v = 0$);
 3) deve valere $A\mathbf{1} = 0$;
 4) le righe di A sono linearmente indipendenti.

Vedremo più avanti come sia possibile trasformare una generica istanza nella forma di Karmarkar. Si ponga $\Delta := \{x \geq 0 : \mathbf{1}x = 1\}$ e $\mathbf{e} := \mathbf{1}/n$. Si noti che $\mathbf{e} \in \Delta$ e ne rappresenta il centro. Inoltre $A\mathbf{e} = 0$ e quindi \mathbf{e} è ammissibile per ipotesi.

Nell'analisi dell'algoritmo entrano in gioco il raggio R della minima sfera S_R in $\text{aff}(\Delta)$ di centro \mathbf{e} e circoscritta a Δ , e il raggio r della massima sfera S_r in $\text{aff}(\Delta)$ di centro \mathbf{e} e inscritta in Δ . La minima sfera circoscritta passa per tutti i vertici di Δ e quindi R si ottiene da

$$R = \|(\mathbf{e} - \{1, 0, \dots, 0\}) \cdot (\mathbf{e} - \{1, 0, \dots, 0\})\| = \sqrt{\left(\frac{n-1}{n}\right)^2 + (n-1)\left(\frac{1}{n}\right)^2} = \sqrt{\frac{n-1}{n}} < 1$$

La massima sfera inscritta tocca la frontiera in punti che hanno una coordinata uguale a zero (altrimenti non potrebbero essere di frontiera) e tutte le altre uguali fra loro e quindi uguali a $1/(n-1)$. Quindi

$$r = \|(\mathbf{e} - \left\{0, \frac{1}{n-1}, \dots, \frac{1}{n-1}\right\}) \cdot (\mathbf{e} - \left\{0, \frac{1}{n-1}, \dots, \frac{1}{n-1}\right\})\| = \sqrt{\frac{1}{n(n-1)}} > \frac{1}{n}$$

L'algoritmo, a partire dal punto $x^0 := \mathbf{e}$, genera una successione x^1, x^2, \dots che converge verso l'ottimo. Sia \bar{x} un generico punto dell'iterazione, ovviamente ammissibile per (17.1). L'idea centrale del metodo consiste nel trasformare Δ (operando un automorfismo) in modo che i vertici di Δ siano trasformati su se stessi e che il punto \bar{x} sia trasformato nel centro \mathbf{e} di Δ . Aniché trovare una direzione di discesa a partire da \bar{x} nello spazio originale, il metodo la cerca a partire da \mathbf{e} nello spazio trasformato. Una volta trovata la direzione e il nuovo punto d'iterazione il punto viene riportato con la trasformazione inversa nello spazio originale. È come se, per la ricerca della direzione di discesa, si operasse una 'zoomata' dell'insieme ammissibile, ingrandendo la zona locale intorno al punto corrente d'iterazione e rimpicciolendo tutto il resto e quindi deformando lo spazio. La direzione viene appunto determinata nello spazio deformato.

La seconda idea chiave del metodo consiste nel valutare il miglioramento della funzione obiettivo, non direttamente sulla funzione obiettivo, ma tramite una cosiddetta funzione potenziale che, oltre a tener conto della minimizzazione, penalizza anche l'avvicinamento alle pareti del poliedro d'ammissibilità. Il fatto importante è che il miglioramento della funzione potenziale è il medesimo sia nello spazio originale che in quello trasformato e questo giustifica le operazioni che si svolgono nello spazio trasformato.

La trasformazione richiesta non può essere lineare, in quanto solo l'identità, fra le trasformazioni lineari, lascia inalterati tutti i vertici di Δ . Le trasformazioni non lineari più 'vicine' a quelle lineari sono le trasformazioni proiettive. Le trasformazioni proiettive sono lineari in uno spazio più ampio. Sia definita la seguente mappa $\mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$

$$(u_0, u_1, \dots, u_{n-1}, u_n) \mapsto (x_1 := \frac{u_1}{u_0}, \dots, x_n := \frac{u_n}{u_0})$$

Si noti che vi sono punti in \mathbb{R}^{n+1} , cioè tutti quelli con $u_0 = 0$, non direttamente mappabili in \mathbb{R}^n ma identificabili con direzioni in \mathbb{R}^n . Un operatore lineare in \mathbb{R}^{n+1} può essere scritto come

$$\begin{pmatrix} \alpha & a \\ b & A \end{pmatrix} \begin{pmatrix} v_0 \\ v \end{pmatrix} = \begin{pmatrix} v_0 \alpha + a v \\ b v_0 + A v \end{pmatrix}$$

e quindi

$$x := \frac{v}{v_0} \quad \text{e} \quad y := \frac{b v_0 + A v}{v_0 \alpha + a v} = \frac{b + A x}{\alpha + a x}$$

Per far sì che i vertici ricadano sui vertici e che $\bar{x} \mapsto \mathbf{e}$ bisogna assegnare i seguenti valori ai parametri della trasformazione proiettiva: $A := (\text{diag}(\bar{x}))^{-1}$, $b := 0$, $\alpha := 0$, $a := A \mathbf{1}$. Allora la trasformazione si può scrivere anche come:

$$(x_1 \ x_2 \ \dots \ x_n) \mapsto \frac{1}{\sum_i \frac{x_i}{\bar{x}_i}} \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ \bar{x}_1 & \bar{x}_2 & \dots & \bar{x}_n \end{pmatrix} =: (y_1 \ y_2 \ \dots \ y_n)$$

La trasformazione, ristretta a Δ , è biettiva e quindi possiamo definire la sua inversa come

$$(y_1 \ y_2 \ \dots \ y_n) \mapsto \frac{1}{\sum_i y_i \bar{x}_i} (y_1 \bar{x}_1, \ y_2 \bar{x}_2, \ \dots \ y_n \bar{x}_n) =: (x_1 \ x_2 \ \dots \ x_n)$$

o anche in modo sintetico

$$y \mapsto \frac{\text{diag}(\bar{x}) y}{\mathbf{1} \text{diag}(\bar{x}) y} = x$$

Il problema (17.1) diventa nello spazio trasformato

$$\begin{aligned} \min & \frac{c \text{diag}(\bar{x}) y}{\mathbf{1} \text{diag}(\bar{x}) y} \\ & \frac{A \text{diag}(\bar{x}) y}{\mathbf{1} \text{diag}(\bar{x}) y} = 0 \\ & y \in \Delta \end{aligned}$$

che, essendo il valore ottimo di (17.1) nullo, è equivalente a:

$$\begin{aligned} \min & c \text{diag}(\bar{x}) y \\ & A \text{diag}(\bar{x}) y = 0 \\ & y \in \Delta \end{aligned}$$

e, ponendo $\bar{A}_i^j := \bar{x}_j A_i^j$, $\bar{c}^j := \bar{x}_j c^j$, possiamo riscrivere come

$$\begin{aligned} \min \quad & \bar{c} y \\ & \bar{A} y = 0 \\ & y \in \Delta \end{aligned} \tag{17.2}$$

Una direzione di discesa per (17.2) è data dall'opposto della proiezione di \bar{c} sullo spazio nullo dei vincoli $\bar{A} y = 0$ e $\mathbf{1} y = 0$. È noto che una tale proiezione \bar{c}' si ottiene come

$$\bar{c}' = \bar{c} - (\bar{A}^T \quad \mathbf{1}^T) \left(\begin{pmatrix} \bar{A} \\ \mathbf{1} \end{pmatrix} (\bar{A}^T \quad \mathbf{1}^T) \right)^{-1} \begin{pmatrix} \bar{A} \\ \mathbf{1} \end{pmatrix} \bar{c}$$

cioè risolvendo con il metodo di Choleski un sistema lineare e sostituendo:

$$\left(\begin{pmatrix} \bar{A} \\ \mathbf{1} \end{pmatrix} (\bar{A}^T \quad \mathbf{1}^T) \right) w = \begin{pmatrix} \bar{A} \\ \mathbf{1} \end{pmatrix} \bar{c} \quad \bar{c}' = \bar{c} - (\bar{A}^T \quad \mathbf{1}^T) w$$

Si indichi con d la direzione di discesa normalizzata, quindi

$$d := -\frac{\bar{c}'}{\|\bar{c}'\|}$$

Si tratta ora di valutare il passo $\hat{\alpha}$ per determinare il nuovo punto d'iterazione $\hat{y} := \mathbf{e} + \hat{\alpha} d$. A questo fine introduciamo la *funzione potenziale*:

$$\varphi(x) := n \ln c x - \sum_i \ln x_i$$

Siccome il valore ottimo è nullo la funzione potenziale scende verso $-\infty$ al tendere di x all'ottimo. Inoltre, al tendere verso punti non ottimi di frontiera (che è data solo dai vincoli di non negatività) la funzione tende a $+\infty$. Nello spazio trasformato la funzione diventa:

$$\begin{aligned} \varphi(x) &= n \ln \bar{c} y - n \ln \left(\sum_i \bar{x}_i y_i \right) - \sum_i \left(\ln(\bar{x}_i y_i) - \ln \left(\sum_j \bar{x}_j y_j \right) \right) = \\ &= n \ln \bar{c} y - \sum_i \ln y_i - \sum_i \ln \bar{x}_i = \bar{\varphi}(y) - \sum_i \ln \bar{x}_i \end{aligned}$$

dove si è posto $\bar{\varphi}(y) := n \ln \bar{c} y - \sum_i \ln y_i$. Allora la differenza fra $\varphi(x)$ e $\bar{\varphi}(y)$ per valori trasformati x e y è costante e quindi il decremento di funzione potenziale $\bar{\varphi}$ misurato nello spazio trasformato è il medesimo del decremento di funzione potenziale φ nello spazio originario. Questo significa che per calcolare il passo $\hat{\alpha}$ possiamo valutare il decremento di $\bar{\varphi}$, ovvero minimizzare, rispetto ad α , la funzione

$$\bar{\varphi}(\mathbf{e} + \alpha d)$$

Nel calcolo effettivo questa minimizzazione si effettua con una ricerca di linea. Ora, volendo analizzare teoricamente il comportamento dell' algoritmo, fissiamo un valore che, a posteriori, permetterà di limitare il numero di iterazioni del metodo. Si noti che un valore di $\alpha < r$ garantisce che il successivo punto d'iterazione sia ancora interno. Quindi possiamo prendere $\hat{\alpha} = 1/n$ e calcolare

$$\bar{\varphi}(\mathbf{e}) - \bar{\varphi}(\hat{y}) = n \ln(\bar{c}\mathbf{e}) - \sum_i \ln(1/n) - n \ln(\bar{c}(\mathbf{e} + d/n)) + \sum_i \ln\left(\frac{1+d_i}{n}\right)$$

Ora si noti che $\bar{c}d = \bar{c}'d = -\bar{c}'\bar{c}'/\|\bar{c}'\| = -\|\bar{c}'\|$. Inoltre la sfera S_R contiene Δ e quindi anche l'insieme ammissibile di (17.2) e allora

$$\min \{\bar{c}y : y \in S_R\} \leq \min \{\bar{c}y : \bar{A}y = 0, y \in \Delta\} = 0 \quad (17.3)$$

Siccome $\min \{\bar{c}y : y \in S_R\} = \bar{c}\mathbf{e} + R\bar{c}d$ si ha, da (17.3), $\bar{c}\mathbf{e} - R\|\bar{c}'\| \leq 0$ e, tenendo conto che $R < 1$, si ha $\bar{c}\mathbf{e} < \|\bar{c}'\|$. Quindi

$$\ln\left(1 - \frac{\|\bar{c}'\|}{n\bar{c}\mathbf{e}}\right) < \ln\left(1 - \frac{1}{n}\right) < -\frac{1}{n}$$

da cui

$$\begin{aligned} n \ln(\bar{c}\mathbf{e}) - n \ln(\bar{c}(\mathbf{e} + d/n)) &= n \ln(\bar{c}\mathbf{e}) - n \ln(\bar{c}\mathbf{e} - \|\bar{c}'\|/n) = \\ n \ln(\bar{c}\mathbf{e}) - n \ln\left(\bar{c}\mathbf{e}\left(1 - \frac{\|\bar{c}'\|}{n\bar{c}\mathbf{e}}\right)\right) &= -n \ln\left(1 - \frac{\|\bar{c}'\|}{n\bar{c}\mathbf{e}}\right) > 1 \end{aligned}$$

L'espressione $\sum_i \ln(1 + d_i)$, con d vincolato da $\sum_i d_i = 0$ e $\sum_i d_i^2 = 1$, assume il minimo valore quando due coordinate, ad esempio d_1 e d_2 , valgono $d_1 = 1/\sqrt{2}$ e $d_2 = -1/\sqrt{2}$, e tutte le altre valgono zero, e in questo caso l'espressione vale $-\ln 2$. Quindi

$$-\sum_i \ln(1/n) + \sum_i \ln\left(\frac{1+d_i}{n}\right) = \sum_i \ln(1 + d_i) \geq -\ln 2$$

Mettendo assieme le due limitazioni si ottiene

$$\bar{\varphi}(\mathbf{e}) - \bar{\varphi}(\hat{y}) > 1 - \ln 2 =: \beta > 0$$

Quindi anche

$$\varphi(\bar{x}) - \varphi(\hat{x}) > \beta$$

ovvero, dopo k iterazioni

$$n \ln c x^k - \ln \sum_i x_i^k - n \ln c x^0 + \ln \sum_i x_i^0 < -k\beta$$

Siccome $\ln \sum_i x_i^0 \geq \ln \sum_i x_i^k$ (si tenga presente che $x^0 = \mathbf{e}$) possiamo scrivere

$$n \ln c x^k - n \ln c x^0 < -k\beta$$

da cui

$$c x^k \leq e^{-\frac{k\beta}{n}} c x^0$$

Questa relazione insieme ai risultati del capitolo 6 dimostra che il metodo di Karmarkar è polinomiale. L'algoritmo di Karmarkar è sintetizzato nel riquadro. Il calcolo più oneroso per ogni iterazione è la risoluzione del sistema lineare. Anche se la matrice è positiva definita il sistema può presentare instabilità numerica al progredire dell'iterazione, dovuta al fatto che le colonne di A sono moltiplicate per dei valori, alcuni dei quali rimangono finiti ed

Algoritmo di Karmarkar

```

input( $A, c, \varepsilon$ );
sia  $x := \mathbf{e}$ 
while  $cx > \varepsilon$  do
  begin
     $\bar{c} := c \operatorname{diag}(\bar{x}); \bar{A} := A \operatorname{diag}(\bar{x});$ 
     $B^T := (\bar{A}^T, \mathbf{1}^T);$ 
    si risolva  $(B B^T) w = B \bar{c};$ 
     $c' = \bar{c} - B w; d := -c' / \|c'\|;$ 
     $\hat{\alpha} := \operatorname{argmin}_{\alpha} \varphi(\mathbf{e} + \alpha d); y := \mathbf{e} + \hat{\alpha} d;$ 
     $x' := \operatorname{diag}(x) y / (\mathbf{1} \operatorname{diag}(x) y); x := x';$ 
  end
output( $x$ );

```

altri tendono a 0. In letteratura sono stati proposti degli accorgimenti per ovviare a questo problema. Il lettore interessato può trovare documentazione in Goldfarb e Todd [1989].

Bisogna ancora far vedere come un generico problema di programmazione lineare può essere posto nella forma di Karmarkar. Il primo passo consiste nel trasformare l'ottimalità in ammissibilità. Quindi dalla coppia di problemi primale-duale

$$\begin{array}{ll}
 \max & cx \\
 & Ax \leq b \\
 & x \geq 0 \\
 \min & ub \\
 & uA \geq c \\
 & u \geq 0
 \end{array} \quad (17.4)$$

che possiamo riscrivere, introducendo le variabili di scarto, come

$$\begin{array}{ll}
 \max & cx \\
 & Ax + Is = b \\
 & x \geq 0 \quad s \geq 0 \\
 \min & ub \\
 & uA - rI = c \\
 & u \geq 0 \quad r \geq 0
 \end{array} \quad (17.5)$$

otteniamo la condizione d'ammissibilità

$$\begin{pmatrix} A & \mathbf{0} & I & \mathbf{0} & 0 \\ \mathbf{0} & -A^T & \mathbf{0} & I & 0 \\ -c & b^T & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ u^T \\ s \\ r^T \\ t \end{pmatrix} = \begin{pmatrix} b \\ -c^T \\ 0 \end{pmatrix} \quad (x, u, s, r, t) \geq 0 \quad (17.6)$$

dove t è una variabile di scarto relativo allo scarto di dualità (si noti che t è l'opposto dello scarto di dualità). Riscriviamo sinteticamente (17.6) come

$$My = q \quad y \geq 0 \quad (17.7)$$

Poi si rende l'espressione omogenea introducendo la variabile w

$$Mv - qw = 0 \quad y, w \geq 0 \quad (17.8)$$

Ovviamente ogni soluzione y in (17.7) si trasforma banalmente nella soluzione $(v, 1)$ in (17.8). Affinché ogni soluzione (v, w) in (17.8) possa essere trasformata in una soluzione v/w di (17.7) è ovviamente richiesto $w > 0$.

Se $w = 0$ allora esistono soluzioni non nulle di $My = 0$. Queste soluzioni ristrette alle variabili x e u definiscono due coni $C_x := \{x : Ax \leq 0, x \geq 0\}$ e $C_u := \{u : uA \geq 0, u \geq 0\}$ tali che, per ogni $h_x \in C_x$ e $h_u \in C_u$ e ogni x e u ammissibili, anche $x + \alpha h_x$ e $u + \alpha h_u$ sono ammissibili per ogni $\alpha > 0$. Inoltre anche $c(x + \alpha h_x) \geq (u + \alpha h_u)b$ (da $My = 0$) e $c(x + \alpha h_x) \leq (u + \alpha h_u)b$ (dalla dualità debole) e quindi $c(x + \alpha h_x) = (u + \alpha h_u)b$, da cui $(x + \alpha h_x)$, $(u + \alpha h_u)$ sono ottimi per ogni $\alpha \geq 0$. Allora deve valere $ch_x = h_u b = 0$ e ogni soluzione ammissibile è anche ottima. Pertanto c è ortogonale all'affine che contiene l'insieme ammissibile primale e analogamente per b e il problema duale.

Se il poliedro primale o quello duale sono corpi convessi la situazione descritta non può verificarsi e quindi l'ipotesi $w = 0$ è falsa. Però la situazione può verificarsi in generale, come si vede dal problema primale $\max \{-x_1 : x_1 \leq 0, -x_2 \leq 1, x \in \mathbb{R}_+^2\}$ il cui duale è $\min \{u_2 : u_1 \geq -1, u_2 \geq 0, u \in \mathbb{R}_+^2\}$. Quindi se si vuole evitare che w assuma il valore 0 è necessario limitare le variabili (primali o duali) in base a limitazioni note del problema. Allora $w = 0$ implica la non ammissibilità del problema (primale o duale).

Bisogna ancora soddisfare il vincolo $Ae = 0$ e l'ipotesi di valore ottimo nullo. Questi requisiti vengono soddisfatti entrambi aggiungendo una variabile artificiale z che deve tendere a 0. Sia $p := M\mathbf{1} - q\mathbf{1}$. Allora il problema (17.1) è

$$\begin{aligned} \min \quad & z \\ & Mv - qw - pz = 0 \\ & (v, w, z) \in \Delta \end{aligned} \tag{17.9}$$

Allora, se n è il numero di variabili del problema primale originario e m è il numero dei suoi vincoli espliciti, il problema (17.9) presenta $2n + 2m + 3$ variabili e $n + m + 1$ vincoli (oltre al vincolo di appartenenza a Δ).

17.1 ESERCIZIO. Si operi la trasformazione secondo la forma di Karmarkar per un problema di programmazione lineare in forma standard $\min \{cx : Ax = b, x \geq 0\}$. ■

17.2 ESEMPIO. Consideriamo dapprima il semplice problema:

$$\begin{aligned} \max \quad & 2x_1 + x_2 \\ & x_1 \leq 2 \\ & x_2 \leq 2 \\ & x_1 + x_2 \leq 3 \\ & x_1, x_2 \geq 0 \end{aligned}$$

L'inizializzazione porta alla seguente espressione numerica per la matrice in (17.9)

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -2 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ -2 & -1 & 2 & 2 & 3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -5 \end{pmatrix}$$

Il punto iniziale d'iterazione è allora $e \in \Delta \subset \mathbb{R}^{13}$, cioè $(1/13, \dots, 1/13)$. La funzione potenziale lungo la direzione di discesa a partire dal punto iniziale è rappresentata in figura 17.1 (in ascissa il valore del passo normalizzato a $1/n$). Come si vede, il calcolo

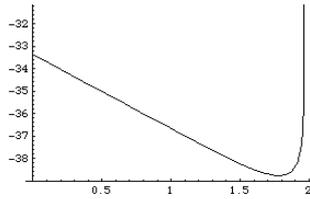


FIGURA 17.1

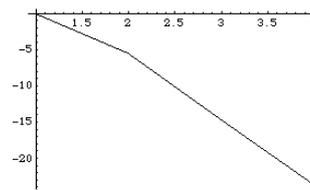


FIGURA 17.2

effettivo del passo porta a dei valori superiori al valore 1 precedentemente fissato teoricamente e quindi l'iterazione è più veloce della stima teorica. È da notare come l'effetto di barriera della frontiera si faccia sentire solo in prossimità della stessa.

I valori delle prime tre iterazioni, per le variabili primali x e duali u (dove $(x, u, s, r, t) := v/w$) e per le variabili ausiliarie w e z , sono i seguenti

x_1	x_2	u_1	u_2	u_3	w	z
0.07692	0.07692	0.07692	0.07692	0.07692	0.07692	0.07692
1.37359	1.06227	0.81320	0.50187	0.87546	0.11092	0.04185
1.94422	1.03348	0.93022	0.10570	1.04518	0.13281	0.01110
1.96418	1.01663	1.01914	0.03097	0.97468	0.14084	0.00273

riportati anche, solo per le variabili primali x , in figura 17.3. In figura 17.2 si vede il decremento globale della funzione potenziale.

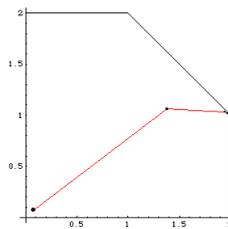


FIGURA 17.3

A questo punto si può operare un arrotondamento a zero dei valori sufficientemente piccoli (ad esempio se minori di 0.1) e verificare se le condizioni di complementarità sono soddisfatte. Si hanno i seguenti valori:

	1	2
x	1.96418	1.01663
r	0.01320	0.00566

	1	2	3
s	0.0358	0.9833	0.01918
u	1.0191	0.0309	0.97468

che, arrotondati danno

	1	2
x	1.96418	1.01663
r	0	0

	1	2	3
s	0	0.9833	0
u	1.0191	0	0.97468

Come si vede le condizioni $\hat{x} \hat{r} = 0$ e $\hat{u} \hat{s} = 0$ sono soddisfatte e si tratta di verificare se da queste condizioni si ottengono valori ammissibili (si riveda a questo proposito la sezione 6.4). Non essendoci degenerazione si ottengono due sistemi lineari, uno primale e l'altro duale:

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad (u_1 \quad u_3) \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = (2 \quad 1)$$

le cui soluzioni sono $\hat{x}_1 = 2$, $\hat{x}_2 = 1$, $\hat{u}_1 = 1$, $\hat{u}_3 = 1$. Essendo tutti i valori non negativi, si tratta della soluzione ottima. ■

17.3 ESEMPIO. Risolviamo ora il rilassamento d'interrezza di un problema d'accoppiamento non bipartito su un grafo completo di 8 nodi con tabella dei costi data da:

$$\begin{pmatrix} 0 & 3 & 6 & 8 & 1 & 9 & 5 & 4 \\ 3 & 0 & 2 & 3 & 9 & 5 & 4 & 2 \\ 6 & 2 & 0 & 9 & 6 & 3 & 7 & 5 \\ 8 & 3 & 9 & 0 & 5 & 5 & 7 & 2 \\ 1 & 9 & 6 & 5 & 0 & 4 & 2 & 7 \\ 9 & 5 & 3 & 5 & 4 & 0 & 3 & 1 \\ 5 & 4 & 7 & 7 & 2 & 3 & 0 & 8 \\ 4 & 2 & 5 & 2 & 7 & 1 & 8 & 0 \end{pmatrix}$$

Il problema viene modellato come $\max \{cx : Ax \leq \mathbf{1}, x \geq 0\}$ con A matrice d'incidenza nodi-archi. Dopo la trasformazione in forma di Karmarkar il problema (17.9) ha 75 variabili e 37 equazioni. Quindi ad ogni iterazione sarà richiesta la risoluzione di un sistema lineare di ordine 38. Dopo solo 8 iterazioni la soluzione primale è data da:

$$x = (0.000481268, 0.000497828, 0.00143907, 0.000104681, 1.00012, 0.000376396, 0.000685837, \\ 0.000302908, 0.000576117, 1.00063, 0.000912273, 0.000162024, 0.000768598, 0.998537, \\ 0.000583199, 0.000662876, 0.00199235, 0.000505158, 0.000535562, 0.000636173, 0.0017685, \\ 0.000235228, 0.000377514, 0.0000179621, 0.00156814, 0.000435709, 0.000324903, 0.998745)$$

da cui si riconosce che la soluzione è molto vicina all'ottimo $\hat{x}_{16} = \hat{x}_{25} = \hat{x}_{34} = \hat{x}_{78} = 1$. Tuttavia, mentre l'ottimo duale vale

$$\hat{u} = (6, 2, 7, 2, 7, 3, 8, 0)$$

la soluzione duale fornita dall'algoritmo è

$$u = (5.15987, 3.74536, 4.34179, 4.65637, 5.25191, 3.84004, 4.52337, 3.47434)$$

e, delle 28 variabili di scarto del duale, solo 4 sono prossime allo 0, mentre la minima delle altre vale 1.72978. Per determinare un vertice duale servirebbero altre 4 variabili nulle. È chiaro che è la degenerazione primale, ovvero la non unicità dell'ottimo duale, responsabile di una tale 'impossibilità' dell'algoritmo di avvicinarsi ad un vertice duale. Continuare con le iterazioni non è molto utile in quanto il sistema lineare da risolvere diventa mal condizionato al punto da non garantire più l'accuratezza della soluzione. In figura 17.4 si vedono le iterazioni relative a x_{16} (in ascissa) e x_{34} (in ordinata). ■

17.4 ESEMPIO. In questo esempio si fa vedere il comportamento dell'algoritmo quando il problema non è ammissibile. Sia dato $\min \{x_1 + 2x_2 : -x_1 - x_2 \geq 1, x_1 \geq 0, x_2 \geq 0\}$. Alla prima iterazione del metodo la funzione potenziale presenta un andamento anomalo lungo la direzione di discesa, come si vede dalla figura 17.5: non c'è l'effetto barriera della frontiera! Le variabili duali (normali e di scarto) valgono 1171 mentre le variabili primali valgono 1. Alla seconda iterazione la variabile w vale $1.82 \cdot 10^{-7}$, e le variabili duali hanno valori dell'ordine di 10^6 . ■

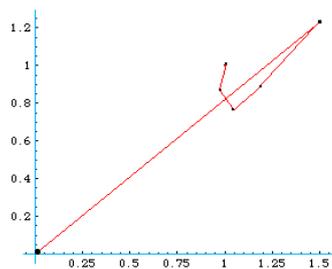


FIGURA 17.4

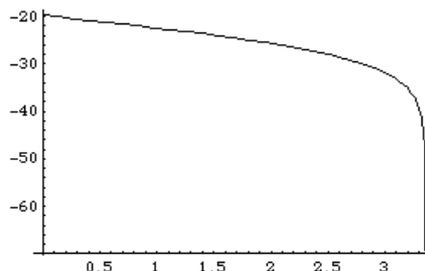


FIGURA 17.5

17.2. Metodi primali-duali

I metodi primali-duali (da non confondere con gli omonimi metodi basati sul semplice) sono basati sulla risoluzione con il metodo di Newton delle equazioni non lineari di complementarità. Sia data la coppia primale-duale (il primale è in forma standard, ma è chiaro che un'analoga analisi può essere svolta con una coppia primale-duale in forma canonica):

$$\begin{array}{ll} \min & cx \\ & Ax = b \\ & x \geq 0 \end{array} \qquad \begin{array}{ll} \max & ub \\ & uA + r = c \\ & r \geq 0 \end{array}$$

Come è noto (vedi sezione 6.4), se una soluzione (x, u, r) è ammissibile ed inoltre $x_j r_j = 0, \forall j$, la soluzione è anche ottima in base alle condizioni di complementarità. Si può allora riformulare la condizione separando i vincoli di eguaglianza da quelli di disequaglianza e dire che una soluzione (x, u, r) è ottima se e solo se $x \geq 0, r \geq 0$, e $F(x, u, r) = 0$, dove

$$F(x, u, r) := \begin{cases} uA + r - c \\ Ax - b \\ XR\mathbf{1} \end{cases} \quad (17.10)$$

e $X = \text{diag}(x)$ e $R = \text{diag}(r)$. Lo Jacobiano di $F(x, u, r)$ è

$$J(x, r) := \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ R & 0 & X \end{pmatrix}$$

e quindi la risoluzione di (17.10) con il metodo di Newton richiede ad ogni passo d'iterazione di risolvere il sistema lineare

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ R & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta u \\ \Delta r \end{pmatrix} = -F(x, u, r) \quad (17.11)$$

e di aggiornare i valori secondo $x := x + \Delta x, u := u + \Delta u, r := r + \Delta r$.

In connessione con questa applicazione del metodo di Newton vi sono tuttavia alcuni problemi da chiarire. Come primo fatto notiamo che non c'è alcuna garanzia che la soluzione finale che si ottiene dal metodo di Newton soddisfi il vincolo di non negatività per (x, r) .

L'idea proposta per ottenere questa garanzia consiste nel modificare la funzione F imponendo $x_i r_i = \tau > 0$ anziché $x_i r_i = 0$. Sia allora

$$F_\tau(x, u, r) := \begin{cases} u A + r - c \\ A x - b \\ X R \mathbf{1} - \tau \mathbf{1} \end{cases} \quad (17.12)$$

e si definisca l'insieme

$$\mathcal{C} := \{(x, u, r) : F_\tau(x, u, r) = 0, \tau > 0\}$$

che viene chiamato *cammino centrale*. Le proprietà del cammino centrale si possono riassumere nel seguente teorema, la cui dimostrazione alquanto complessa, non viene data (si veda Wright [1997]):

17.5 TEOREMA. *Se esistono soluzioni ammissibili strettamente positive (cioè $x_i > 0, \forall i$, e $r_i > 0, \forall i$), allora per ogni $\tau > 0$ esiste una soluzione di $F_\tau(x, u, r) = 0$. Inoltre tale soluzione ristretta a (x, r) è unica.* ■

Ha senso allora, in base al teorema, la denominazione di cammino per l'insieme \mathcal{C} , quantomeno per le variabili non negative. Si noti inoltre che, se la coppia primale-duale è in forma canonica, allora l'unicità della soluzione è garantita per tutte le variabili. Se l'ipotesi del teorema non è verificata, ovvero non esistono soluzioni strettamente positive, tutto il ragionamento si può ripetere nell'affine che contiene l'insieme ammissibile.

17.6 ESEMPIO. Si vuole rappresentare il cammino centrale per un piccolo esempio in \mathbb{R}^2 . Sia dato $\max \{x_1 : x_1 \geq x_2, x_1 + x_2 \leq 1, x \in \mathbb{R}_+^2\}$. Il problema trasformato in forma standard dà luogo alla coppia primale duale:

$\begin{aligned} \min \quad & -x_1 \\ & -x_1 + x_2 + x_3 = 0 \\ & x_1 + x_2 + x_4 = 1 \\ & x_i \geq 0 \forall i \end{aligned}$	$\begin{aligned} \max \quad & u_2 \\ & -u_1 + u_2 + r_1 = -1 \\ & u_1 + u_2 + r_2 = 0 \\ & u_1 + r_3 = 0 \\ & u_2 + r_4 = 0 \\ & r_i \geq 0, \forall i \end{aligned}$
--	---

In figura 17.6 è rappresentato l'insieme ammissibile primale insieme con il cammino centrale calcolato per valori di $\tau = 2^k, k := 6, 5, \dots, -5$. ■

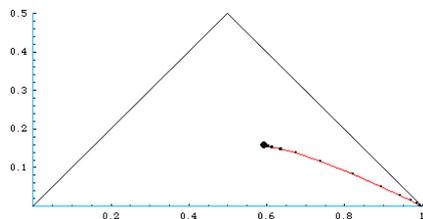


FIGURA 17.6

Il cammino centrale può anche essere caratterizzato alternativamente nel seguente modo. Si consideri la famiglia di problemi convessi per valori diversi di $\tau > 0$

$$\min \left\{ cx - \tau \sum_j \ln x_j : Ax = b, x > 0 \right\} \quad (17.13)$$

Una semplice applicazione delle condizioni duali di ottimalità e la sostituzione $\tau/x_j =: r_j$ conducono direttamente a $F_\tau(x, u, r) = 0$. Quindi la proiezione del cammino centrale su x coincide con i minimi di (17.13). Dato un poliedro $\{x : Ax \leq b\}$ il punto che massimizza $\sum_i \ln(b_i - A_i x)$ viene definito *centro analitico* del poliedro. Quindi il cammino centrale tende, per $\tau \rightarrow \infty$, verso il centro analitico del poliedro.

Il metodo primale-duale allora cerca di trovare con il metodo di Newton un punto sul cammino centrale per un valore sufficientemente elevato di τ . In questo modo si ha la garanzia della positività delle soluzioni. Dopodiché si varia τ e si continua l'iterazione di Newton in modo che la traiettoria delle soluzioni segua da vicino il cammino centrale. Si noti che $F(x, u, r)$ e $F_\tau(x, u, r)$ hanno lo stesso Jacobiano, quindi si tratta di risolvere ad ogni passo d'iterazione

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ R & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta u \\ \Delta r \end{pmatrix} = -F_\tau(x, u, r) \quad (17.14)$$

Un secondo problema connesso con (17.14) riguarda l'ovvia condizione che lo Jacobiano sia non singolare per tutti i valori dell'iterazione. Inoltre, per avere la convergenza superlineare tipica del metodo di Newton, è richiesto che lo Jacobiano calcolato negli ottimi sia non singolare. Per la prima condizione, considerato che l'iterazione opera con valori di x e r strettamente positivi, basta che le righe di A siano linearmente indipendenti.

La condizione invece sullo Jacobiano calcolato negli ottimi può dare luogo a qualche problema. Se l'ottimo è unico e regolare (nel senso della complementarità) allora x_j e r_j non sono entrambi negativi e lo Jacobiano è non singolare. Però soluzioni degeneri e non uniche sono all'ordine del giorno in programmazione lineare e in questi casi lo Jacobiano è singolare. Tuttavia si è visto che ottimi regolari (ma non di vertice) esistono sempre (teorema 6.27) e il cammino centrale tende proprio a questi ottimi regolari. Questo fatto risolve un problema e ne crea un altro. Il problema risolto è la garanzia della non singolarità dello Jacobiano. Il problema che si crea è dovuto alla difficoltà di ottenere una soluzione di vertice a partire dall'ottimo fornito dall'iterazione che è situato in modo baricentrico rispetto ai vertici ottimi. Questo è un problema di tutti i metodi ai punti interni e può essere risolto in modo analogo a quanto fatto per il metodo dell'ellissoide, cioè perturbando i dati.

Vi è ancora il problema tipico del metodo di Newton di poter partire con l'iterazione da un punto interno alla regione di convergenza. Questo viene risolto dalla strategia generale di modificare il passo d'iterazione per ottenere risultati di convergenza globale e superlineare. Fra le strategie più usate vi è il metodo cosiddetto *predictor-corrector* dovuto a Mehrotra [1992]. Si veda a questo proposito anche Wright [1997]. Non analizzeremo in questa sede tutti i dettagli necessari per assicurare teoricamente il buon funzionamento del metodo. Ad esempio, si tratta di definire in modo esatto un intorno del cammino centrale e di definire i passi dell'iterazione in modo che tutti i valori siano contenuti nell'intorno. Diverse definizioni di intorno sono state proposte a seconda del tipo di norma usata. Negli esempi seguiremo un semplice approccio in cui non appena il punto è all'interno dell'intorno $\max |x_j r_j - \tau| < \theta \sum_j x_j r_j$ il valore di τ viene dimezzato.

Infine, a conclusione dell'iterazione, c'è il problema, già anticipato, di ottenere un ottimo da un ottimo approssimato. Non avendo la garanzia di avere in generale una soluzione vicina

ad un vertice non si può adottare l'approccio di ricavare dalle condizioni di complementarità abbastanza equazioni da risolvere un sistema lineare. Bisogna ricorrere ad una stima basata sui minimi quadrati. Quindi, poste a zero quelle componenti di x e r che hanno un valore al di sotto di una soglia prefissata (siano Z_x e Z_r tali insiemi di indici) e verificato che $|Z_x| + |Z_r| = n$ si risolve il problema (dove $(\bar{x}, \bar{u}, \bar{r})$ è la soluzione fornita dall'iterazione)

$$\begin{aligned} \min \quad & \|x - \bar{x}\|^2 + \|r - \bar{r}\|^2 \\ & Ax = b \\ & uA + r = c \\ & x_j = 0, \quad j \in Z_x, \quad r_j = 0, \quad j \in Z_r \end{aligned} \quad (17.15)$$

Se la soluzione $(\hat{x}, \hat{u}, \hat{r})$ di (17.15) è tale che $\hat{x}_j > 0, j \notin Z_x, \hat{r}_j > 0, j \notin Z_r$, allora si ha una soluzione che soddisfa in modo regolare la complementarità e quindi è ottima. Altrimenti si prosegue con l'iterazione.

Risolvere (17.15) è meno difficile di quanto sembra. A ben guardare (17.15) si spezza in due problemi indipendenti, uno per il problema primale e l'altro per il problema duale. Inoltre un generico problema di minimi quadrati del tipo $\min \{\|\bar{x} - x\|_2 : Ax = b\}$ si risolve proiettando \bar{x} sull'affine dato da $Ax = b$. Quindi $\bar{x} - \hat{x}$ è ortogonale al sottospazio generatore dell'affine, ovvero esistono moltiplicatori y tali che $\bar{x} - \hat{x} = A^T y$. Siccome \hat{x} deve appartenere all'affine, bisogna imporre $A\hat{x} = b$, cioè $A(\bar{x} - A^T y) = b$ da cui $y = (AA^T)^{-1}(A\bar{x} - b)$ e quindi $\hat{x} = \bar{x} - A^T(AA^T)^{-1}(A\bar{x} - b)$. In definitiva si risolve il sistema lineare $(AA^T)z = A\bar{x} - b$ e si calcola $\hat{x} := \bar{x} - A^T z$. Per il sottoproblema duale si può ragionare trovando le condizioni di minimo per il problema

$$\begin{aligned} \min \quad & \sum_{j \notin Z_r} (r_j - \bar{r}_j)^2 \\ & uA^j = c_j \quad j \in Z_r \\ & uA^j + r_j = c_j \quad j \notin Z_r \end{aligned} \quad (17.16)$$

che corrispondono alla risoluzione del sistema lineare

$$\begin{pmatrix} \mathbf{0} & A_0 & A_1 \\ A_0^T & \mathbf{0} & \mathbf{0} \\ A_1^T & \mathbf{0} & I \end{pmatrix} \begin{pmatrix} u \\ \xi \\ r_1 \end{pmatrix} = \begin{pmatrix} A_1 \bar{r}_1 \\ c_B \\ c_N \end{pmatrix} \quad (17.17)$$

dove A_0 è formata con le colonne in Z_x e A_1 è formata con le colonne complementari e analogamente per r_1 . La variabile ξ è la variabile duale del vincolo $uA^j = c_j, j \in Z_r$. Non deve sfuggire il fatto che se A_0 è quadrata allora è la matrice di base ottima del vertice e (17.16) ha un'unica soluzione ammissibile e necessariamente ottima.

17.7 ESEMPIO. Sia

$$A = \begin{pmatrix} 2 & 1 & 3 & 1 \\ 0 & 1 & 2 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 7 \\ 4 \end{pmatrix} \quad c = (1 \quad -2 \quad 3 \quad -1)$$

e si scelgano $\tau = 2^6, x = \mathbf{1}, u = 0$ e $r = \mathbf{1}$ come valori iniziali. Ogni qualvolta $\sigma := \max |x_i r_i - \tau| < 0.001$ il valore di τ viene dimezzato. L'iterazione viene interrotta quando $\tau < 0.001$. I valori ottenuti sono riportati in figura 17.7. Si noti che nelle prime due iterazioni vengono prodotti quattro valori x_i negativi. Questo è tuttavia dovuto al fatto che

τ	x_1	x_2	x_3	x_4	u_1	u_2	r_1	r_2	r_3	r_4	σ
	1.	1.	1.	1.	0.	0.	1.	1.	1.	1.	63.
2^6	16.62	32.73	-30.23	31.73	-23.69	-10.58	48.38	32.27	95.23	33.27	2942.9
2^6	-0.76	-2.64	4.53	-2.42	-26.73	-12.09	54.46	36.82	107.38	37.82	422.23
2^6	1.38	1.78	0.25	1.73	-33.89	-5.48	68.78	37.37	115.63	38.37	35.34
2^6	1.23	1.47	0.55	1.44	-26.5	-18.01	53.99	42.5	118.5	43.5	2.21
2^6	1.23	1.48	0.54	1.45	-25.49	-19.79	51.97	43.27	119.04	44.27	0.01
2^6	1.23	1.48	0.54	1.45	-25.49	-19.78	51.98	43.27	119.03	44.27	0.
2^5	1.24	1.5	0.53	1.45	-12.4	-10.71	25.79	21.11	61.61	22.11	0.54
2^5	1.24	1.52	0.52	1.45	-12.4	-10.72	25.8	21.12	61.64	22.12	0.
2^4	1.25	1.55	0.5	1.45	-5.86	-6.2	12.71	10.06	32.97	11.06	0.52
2^4	1.26	1.59	0.48	1.44	-5.86	-6.22	12.72	10.08	33.02	11.08	0.
2^3	1.27	1.66	0.45	1.44	-2.6	-3.99	6.19	4.58	18.77	5.58	0.47
2^3	1.29	1.73	0.42	1.42	-2.61	-4.02	6.21	4.63	18.86	5.63	0.
2^3	1.29	1.73	0.42	1.42	-2.61	-4.02	6.21	4.63	18.86	5.63	0.
2^2	1.32	1.87	0.37	1.39	-0.99	-2.95	2.97	1.94	11.86	2.94	0.38
2^2	1.33	2.	0.33	1.33	-1.	-3.	3.	2.	12.	3.	0.01
2^2	1.33	2.	0.33	1.33	-1.	-3.	3.	2.	12.	3.	0.
2	1.37	2.26	0.26	1.23	-0.21	-2.53	1.42	0.74	8.69	1.74	0.32
2	1.39	2.46	0.23	1.09	-0.22	-2.6	1.44	0.82	8.86	1.82	0.02
2	1.39	2.45	0.23	1.1	-0.22	-2.6	1.44	0.82	8.86	1.82	0.
1	1.43	2.83	0.15	0.87	0.16	-2.44	0.68	0.28	7.41	1.28	0.2
1	1.43	2.99	0.13	0.74	0.15	-2.49	0.7	0.34	7.52	1.34	0.01
1	1.43	2.98	0.13	0.75	0.15	-2.49	0.7	0.34	7.52	1.34	0.
2^{-1}	1.46	3.35	0.08	0.49	0.33	-2.46	0.34	0.13	6.92	1.13	0.08
2^{-1}	1.46	3.42	0.07	0.44	0.33	-2.48	0.34	0.15	6.96	1.15	0.
2^{-1}	1.46	3.42	0.07	0.44	0.33	-2.48	0.34	0.15	6.96	1.15	0.
2^{-2}	1.48	3.67	0.04	0.25	0.42	-2.48	0.17	0.06	6.71	1.06	0.02
2^{-2}	1.48	3.69	0.04	0.23	0.42	-2.48	0.17	0.07	6.72	1.07	0.
2^{-3}	1.49	3.84	0.02	0.13	0.46	-2.49	0.08	0.03	6.6	1.03	0.01
2^{-3}	1.49	3.84	0.02	0.12	0.46	-2.49	0.08	0.03	6.61	1.03	0.
2^{-4}	1.5	3.92	0.01	0.06	0.48	-2.49	0.04	0.02	6.55	1.02	0.
2^{-4}	1.5	3.92	0.01	0.06	0.48	-2.5	0.04	0.02	6.55	1.02	0.
2^{-5}	1.5	3.96	0.	0.03	0.49	-2.5	0.02	0.01	6.53	1.01	0
2^{-6}	1.5	3.98	0.	0.02	0.49	-2.5	0.01	0.	6.51	1.	0.
2^{-7}	1.5	3.99	0.	0.01	0.5	-2.5	0.01	0.	6.51	1.	0.
2^{-8}	1.5	3.99	0.	0.	0.5	-2.5	0.	0.	6.5	1.	0.
2^{-9}	1.5	4.	0.	0.	0.5	-2.5	0.	0.	6.5	1.	0.

FIGURA 17.7

la soluzione iniziale è lontana dal cammino centrale. I valori successivi sono tutti vicini al cammino centrale e vengono prodotte soltanto soluzioni non negative. Inoltre le soluzioni già alla seconda iterazione soddisfano i vincoli $Ax = b$ e $uA + r = c$ a causa della loro linearità. Tutte le cifre riportate in figura sono state arrotondate alla seconda cifra decimale per comodità di lettura. Dalla soluzione dell'ultima riga si ottiene l'ottimo ponendo uguale a zero le variabili con valore al di sotto di una prefissata tolleranza, risolvendo i sistemi lineari rispetto alle restanti variabili e verificando le relazioni di complementarità. ■

17.8 ESEMPIO. Si consideri un problema di assegnamento pesato su un grafo bipartito 10×10 con la tabella dei costi riportata in figura 17.8. Il problema è modellato come

$$\begin{aligned} \min \quad & cx \\ & Ax = \mathbf{1} \\ & x \geq 0 \end{aligned}$$

dove A è la matrice d'incidenza nodi archi di un grafo bipartito, privata però dell'ultima riga, altrimenti le righe di A non sarebbero linearmente indipendenti. I valori iniziali d'iterazione

sono $x := \mathbf{1}$, $u := 0$ e $r := \mathbf{1}$. Come valore iniziale di τ si è scelto 2^{10} . In 35 iterazioni il valore di τ è stato ridotto a 2^{-9} . Dalle soluzioni di fine iterazione si sono ottenuti i valori finali u^1 , x^1 e r^1 tramite (17.15). Tali valori sono riportati in figura 17.8 (valori arrotondati alla seconda cifra decimale). La matrice in alto a destra è l'ottimo x (i valori sono stati ottenuti razionalizzando ogni componente con errore al massimo 0.1 e denominatore minimo). Come si vede la soluzione non è un vertice! Si verifica immediatamente che è la combinazione convessa di due vertici ottimi (si noti che in corrispondenza dei valori 0.5 i valori dei costi danno luogo a due assegnamenti di pari costo). In figura 17.9 si vede la convergenza di $x_{1,9}$ e in figura 17.10 quella di $x_{1,10}$.

Per calcolare una base ottima si perturbano i valori di c aggiungendo ad ogni costo un valore casuale distribuito uniformemente fra 0 e 0.1. L'iterazione, con i nuovi valori dei costi, converge verso un vertice. In figura 17.11 si vedono l'iterazione iniziale e quella con i costi perturbati proiettate sul piano $x_{2,1}$, $x_{2,8}$. Con i valori scelti di errore la soluzione è ancora distante dal vertice. Infatti si ottiene $x_{2,1} = 0.922$, $x_{2,8} = 0.078$. Applicando (17.16) si ottiene tuttavia una soluzione non ammissibile! I valori u^2 , x^2 e r^2 , che si ottengono da (17.16) sono riportati in figura 17.8 e si nota che $r_{9,1} < 0$. Quindi bisognerebbe continuare con l'iterazione. ■

17.3. Programmazione quadratica

I metodi ai punti interni sono particolarmente indicati a risolvere problemi di programmazione quadratica del seguente tipo (dove Q è simmetrica positiva semidefinita)

$$\begin{aligned} \min \quad & \frac{1}{2} x^T Q x + c x \\ & A x \geq b \\ & x \geq 0 \end{aligned}$$

in quanto, in base alle condizioni di Kuhn-Tucker (vedi esempio 5.43), si traducono nel problema di complementarità lineare monotona

$$\begin{aligned} \begin{pmatrix} Q & -A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ u^T \end{pmatrix} - \begin{pmatrix} r^T \\ s \end{pmatrix} &= \begin{pmatrix} -c^T \\ b \end{pmatrix} \\ (x, u, r, s) &\geq 0 \\ x_j r_j = 0, \forall j, \quad u_i s_i = 0, \forall i \end{aligned}$$

Ponendo allora

$$F_\tau(x, u, s, r) := \begin{cases} x^T Q - u A - r + c \\ A x - s - b \\ Y T \mathbf{1} - \tau \mathbf{1} \end{cases}$$

dove $Y := \text{diag}\{x_1, \dots, x_n, u_1, \dots, u_m\}$ e $T := \text{diag}\{r_1, \dots, r_n, s_1, \dots, s_m\}$, e siccome lo Jacobiano di F_τ è

$$\begin{pmatrix} M & -I \\ T & Y \end{pmatrix} \quad \text{con} \quad M := \begin{pmatrix} Q & -A^T \\ A & \mathbf{0} \end{pmatrix}$$

si tratta di risolvere

$$\begin{pmatrix} M & -I \\ T & Y \end{pmatrix} \begin{pmatrix} \Delta y \\ \Delta t \end{pmatrix} = -F_\tau(y, t)$$

dove naturalmente $y := (x, u)$ e $t := (r, s)$ e aggiornare come $y := y + \Delta y$, $t := t + \Delta t$.

$$c = \begin{pmatrix} 8 & 5 & 8 & 3 & 4 & 6 & 1 & 7 & 1 & 8 \\ 2 & 1 & 1 & 9 & 5 & 6 & 6 & 8 & 2 & 6 \\ 2 & 6 & 4 & 6 & 6 & 9 & 8 & 3 & 3 & 6 \\ 7 & 9 & 1 & 5 & 5 & 3 & 8 & 5 & 9 & 4 \\ 6 & 9 & 2 & 5 & 7 & 6 & 7 & 6 & 5 & 5 \\ 1 & 3 & 8 & 4 & 7 & 9 & 7 & 3 & 8 & 2 \\ 1 & 4 & 8 & 9 & 5 & 5 & 1 & 6 & 4 & 9 \\ 7 & 8 & 4 & 1 & 6 & 6 & 6 & 5 & 2 & 5 \\ 3 & 3 & 5 & 4 & 9 & 5 & 6 & 4 & 9 & 6 \\ 3 & 3 & 8 & 2 & 9 & 1 & 6 & 1 & 1 & 5 \end{pmatrix} \quad x^1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$r^1 = \begin{pmatrix} 7.55 & 4.69 & 8.51 & 2.37 & 0.22 & 4.38 & 0.24 & 5.55 & 0 & 6.16 \\ 0.85 & 0 & 0.82 & 7.68 & 0.52 & 3.68 & 4.55 & 5.85 & 0.31 & 3.46 \\ 0 & 4.15 & 2.96 & 3.83 & 0.67 & 5.83 & 5.69 & 0 & 0.45 & 2.61 \\ 5.33 & 7.48 & 0.3 & 3.16 & 0 & 0.16 & 6.03 & 2.33 & 6.78 & 0.94 \\ 3.04 & 6.18 & 0 & 1.86 & 0.7 & 1.87 & 3.73 & 2.04 & 1.49 & 0.64 \\ 0.39 & 2.54 & 8.36 & 3.22 & 3.06 & 7.22 & 6.09 & 1.39 & 6.84 & 0 \\ 0.31 & 3.45 & 8.27 & 8.13 & 0.97 & 3.14 & 0 & 4.31 & 2.76 & 6.91 \\ 6.17 & 7.32 & 4.14 & 0 & 1.84 & 4. & 4.87 & 3.17 & 0.63 & 2.78 \\ 0 & 0.15 & 2.96 & 0.83 & 2.67 & 0.83 & 2.69 & 0 & 5.45 & 1.61 \\ 3.17 & 3.32 & 9.13 & 2. & 5.84 & 0 & 5.86 & 0.17 & 0.62 & 3.78 \end{pmatrix}$$

$$u^1 = \{-0.17, -0.32, -1.13, 0, 3.16, 1., 0.14, 0.83, 0.38, 1.22, 0.62, 1.32, 2.17, 1.84, 3.13, 0.78, 0.86, 1., 3.17\}$$

$$u^2 = \{-0.12, -0.32, -1.09, 0.07, 3.21, 1.01, 0.18, 0.85, 0.41, 1.26, 0.69, 1.36, 2.2, 1.89, 3.18, 0.82, 0.87, 1.02, 3.19\}$$

$$x^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.922 & 0 & 0 & 0 & 0 & 0 & 0 & 0.078 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.078 & 0 & 0 & 0 & 0 & 0 & 0 & 0.922 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}; r^2 = \begin{pmatrix} 7.53 & 4.7 & 8.52 & 2.35 & 0.21 & 4.36 & 0.22 & 5.54 & 0 & 6.15 \\ 0.83 & 0 & 0.82 & 7.65 & 0.52 & 3.66 & 4.52 & 5.84 & 0.3 & 3.45 \\ 0 & 4.17 & 2.99 & 3.82 & 0.69 & 5.84 & 5.69 & 0.01 & 0.47 & 2.62 \\ 5.31 & 7.48 & 0.3 & 3.13 & 0 & 0.15 & 6. & 2.32 & 6.79 & 0.93 \\ 3.01 & 6.18 & 0 & 1.83 & 0.7 & 1.85 & 3.7 & 2.02 & 1.48 & 0.63 \\ 0.38 & 2.55 & 8.37 & 3.2 & 3.07 & 7.22 & 6.07 & 1.39 & 6.85 & 0 \\ 0.31 & 3.48 & 8.3 & 8.13 & 1. & 3.14 & 0 & 4.32 & 2.78 & 6.93 \\ 6.18 & 7.35 & 4.17 & 0 & 1.87 & 4.02 & 4.87 & 3.19 & 0.65 & 2.8 \\ -0.01 & 0.16 & 2.98 & 0.81 & 2.68 & 0.83 & 2.68 & 0 & 5.46 & 1.61 \\ 3.16 & 3.34 & 9.15 & 1.98 & 5.85 & 0 & 5.86 & 0.17 & 0.64 & 3.78 \end{pmatrix}$$

FIGURA 17.8

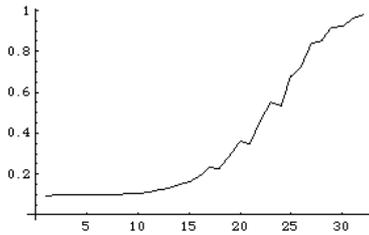


FIGURA 17.9

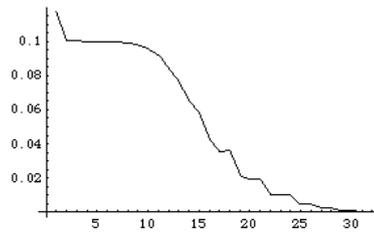


FIGURA 17.10

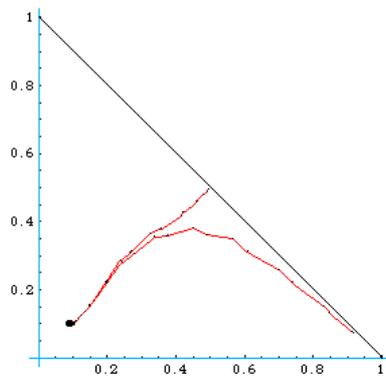


FIGURA 17.11

17.4. Programmazione semidefinita

Nella programmazione semidefinita gli elementi dell'insieme delle soluzioni sono matrici simmetriche, anziché vettori. I vincoli a cui le matrici sono soggette sono tipicamente lineari. Ciò che differenzia questi problemi dalla programmazione lineare è il vincolo di non negatività che, anziché imporre la non negatività elemento per elemento della matrice, richiede alla matrice di essere positiva semidefinita. Questo tipo di vincolo non è formulabile all'interno dei consueti modelli di programmazione lineare e non lineare. Si è invece visto che i metodi ai punti interni possono essere estesi anche a questi problemi che vengono così risolti in tempo polinomiale. Per una rassegna delle tematiche della programmazione semidefinita si veda ad esempio l'articolo di Vandenberghe e Boyd [1996] e, per una trattazione approfondita ed esauriente, si può consultare Wolkowicz et al. [2000].

La programmazione semidefinita si riduce alla programmazione lineare se le matrici sono diagonali. L'estensione degli algoritmi al caso generale non è tuttavia immediata e vi sono molti problemi tecnici da risolvere prima di ottenere dei metodi robusti di calcolo. Ci limitiamo pertanto a fornire gli strumenti iniziali della teoria.

Un problema di programmazione semidefinita viene formulato nello spazio lineare \mathcal{S}^n delle matrici simmetriche di ordine n (spazio di dimensione $n(n+1)/2$). Lo spazio viene reso di Hilbert introducendo il prodotto scalare $A \bullet B := \text{Tr}(AB)$ fra due matrici A e B di \mathcal{S}^n (ricordiamo che $\text{Tr } A := \sum_i A_{ii}$). Si verifica facilmente che valgono le proprietà del prodotto interno (cioè $A \bullet A \geq 0$, $A \bullet A = 0 \iff A = \mathbf{0}$, $A \bullet B = B \bullet A$ e $(\alpha A) \bullet B = \alpha(A \bullet B)$). Si noti ancora che $A \bullet B = \sum_{ij} A_{ij} B_{ij}$ (quindi il prodotto è assimilabile al prodotto vettoriale normale trasformando ogni matrice in un unico vettore).

In \mathcal{S}^n si introduce un ordine parziale \succeq definendo $A \succeq B$ se $A - B$ è positiva semidefinita (e $A \succ B$ se $A - B$ è positiva definita). L'ordine e il prodotto interno sono legati dalle proprietà: i) $A \succeq \mathbf{0}$, $B \succeq \mathbf{0}$ implicano $A \bullet B \geq 0$, ii) $A \succeq \mathbf{0}$, $B \succeq \mathbf{0}$, $A \bullet B = 0$ implicano $AB = \mathbf{0}$, iii) data una matrice $A \succeq \mathbf{0}$ esiste sempre una matrice $B \in \mathcal{S}^n$ tale che $A \bullet B < 0$.

Un problema di programmazione semidefinita viene definito nel seguente modo: sono assegnate matrici $C, A_1, A_2, \dots, A_m \in \mathcal{S}^n$ e scalari b_1, b_2, \dots, b_m . Si cerca una matrice $X \succeq \mathbf{0}$, $X \in \mathcal{S}^n$, che minimizza

$$\begin{aligned} \min \quad & C \bullet X \\ & A_i \bullet X = b_i \quad i := 1, \dots, m \\ & X \succeq \mathbf{0} \end{aligned} \quad (17.18)$$

Si vede subito che se le matrici C e A_i sono diagonali, allora anche X può essere diagonale e (17.18) è di fatto un problema di programmazione lineare. Vogliamo ribadire che, in generale, il vincolo $X \succeq \mathbf{0}$ è convesso ma non è lineare. A titolo illustrativo si consideri una matrice simmetrica di ordine 3 e la si normalizzi ponendo gli elementi diagonali uguali a 1. Quindi la matrice si può scrivere in forma generale come:

$$X = \begin{pmatrix} 1 & x_{12} & x_{13} \\ x_{12} & 1 & x_{23} \\ x_{13} & x_{23} & 1 \end{pmatrix}$$

e la condizione $X \succeq \mathbf{0}$ diventa

$$x_{ij}^2 \leq 1 \quad i \neq j \quad x_{12}^2 + x_{13}^2 + x_{23}^2 \leq 1 + 2x_{12}x_{13}x_{23} \quad (17.19)$$

I valori che soddisfano (17.19) formano l'insieme convesso rappresentato in figura 17.12, denominato *ellittopo*. Tale insieme 'assomiglia' ad un politopo e, anche se ha infiniti punti

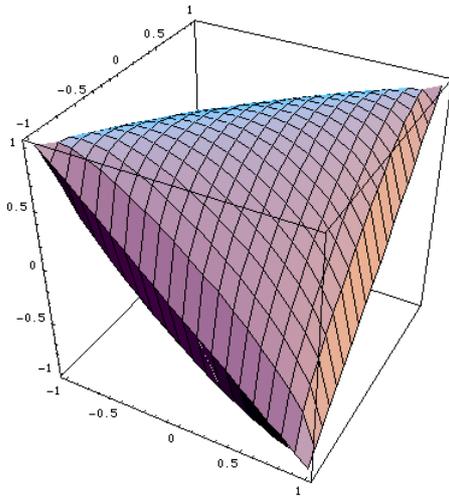


FIGURA 17.12

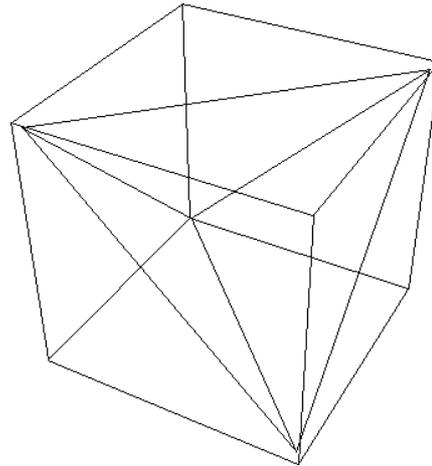


FIGURA 17.13

estremi, alcuni di questi sono di tipo particolare e possono essere assimilati a dei vertici. Si noti che il tetraedro di figura 17.13 è contenuto nell'ellittopo e che i suoi spigoli e vertici sono facce di dimensione 1 e 0 dell'ellittopo. Per maggiori approfondimenti sulle proprietà degli ellittopi si veda Laurent [1995].

Il duale di (17.18) può essere calcolato applicando le tecniche generali della dualità. Quindi

$$L(u) = \inf_{X \succeq \mathbf{0}} C \bullet X + \sum_{i=1}^m u_i (b_i - A_i \bullet X) = \sum_{i=1}^m u_i b_i + \inf_{X \succeq \mathbf{0}} (C - \sum_{i=1}^m u_i A_i) \bullet X$$

In base alle proprietà citate di \succeq , l'inf è limitato solo se $(C - \sum_{i=1}^m u_i A_i)$ è positiva semidefinita, quindi il duale è

$$\begin{array}{ccc} \max & u b & \\ & \sum_{i=1}^m u_i A_i \preceq C & \text{ovvero} & \max & u b \\ & & & & \sum_{i=1}^m u_i A_i + R = C \\ & & & & R \succeq \mathbf{0} \end{array}$$

Siccome le matrici positive semidefinite costituiscono un cono convesso e siccome i vincoli sono lineari, vale la dualità forte. Le condizioni di complementarità si ricavano nel seguente modo: siano $(\hat{X}, \hat{u}, \hat{R})$ ottimi. Allora

$$C \bullet \hat{X} = \sum_i \hat{u}_i b_i = \sum_i \hat{u}_i (A_i \bullet \hat{X}) \implies (C - \sum_i \hat{u}_i A_i) \bullet \hat{X} = 0 \implies R \bullet \hat{X} = 0$$

Viceversa sia $R \bullet X = 0$. Allora moltiplicando per X i vincoli duali si ha

$$\sum_i u_i A_i \bullet X + R \bullet X = C \bullet X \implies \sum_i u_i A_i \bullet X = C \bullet X \implies \sum_i u_i b_i = C \bullet X$$

cioè l'ottimalità. La condizione $R \bullet X = 0$, essendo anche $X \succeq \mathbf{0}$ e $R \succeq \mathbf{0}$ si può esprimere come $X R = \mathbf{0}$ (questa è analoga alla condizione vettoriale $x \geq 0, r \geq 0, x r = 0$, implicano

$x_i r_i = 0, \forall i$). Inoltre $X \succeq \mathbf{0}, R \succeq \mathbf{0}$ e $XR = \mathbf{0}$ implicano anche $XR + RX = \mathbf{0}$ (per dimostrarlo si provi il fatto quando una delle due matrici è diagonale, dopodiché si vede che con un'opportuna trasformazione ortogonale di coordinate una delle due matrici può sempre essere resa diagonale). L'uso di $XR + RX = \mathbf{0}$, anziché $XR = \mathbf{0}$, è dovuto al fatto che la prima espressione è simmetrica per costruzione.

Possiamo pertanto dire che (X, u, R) sono ottimi (nei rispettivi problemi) se e solo se $X \succeq \mathbf{0}, R \succeq \mathbf{0}$ e $F(X, u, R) = 0$ dove

$$F(X, u, R) := \begin{cases} \sum_{i=1}^m u_i A_i + R - C \\ A_i \bullet X - b_i \quad i := 1, \dots, m \\ XR + RX \end{cases} \quad (17.20)$$

Come si vede $F : \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n \rightarrow \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ e quindi lo Jacobiano è un operatore lineare fra due spazi di uguale dimensione e si può rappresentare come una matrice quadrata. L'espressione dello Jacobiano come matrice quadrata è alquanto complicata. Rappresentando una matrice simmetrica A come un vettore colonna di $n(n+1)/2$ componenti $\hat{A} := (a_{11}, a_{12}, \dots, a_{1n}, a_{22}, a_{23}, \dots, a_{2n}, a_{33}, \dots, a_{nn})$ e usando una diversa rappresentazione vettoriale $\tilde{A} := (a_{11}, 2a_{12}, \dots, 2a_{1n}, a_{22}, 2a_{23}, \dots, 2a_{2n}, a_{33}, \dots, a_{nn})$, possiamo scrivere lo Jacobiano come

$$J(X, R) = \begin{pmatrix} \mathbf{0} & \tilde{A}_1 & \tilde{A}_2 & \tilde{A}_m & I \\ \hat{A}_1^T & 0 & 0 & 0 & 0 \\ \hat{A}_2^T & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \hat{A}_m^T & 0 & 0 & 0 & 0 \\ \sigma(R) & 0 & 0 & 0 & \sigma(X) \end{pmatrix}$$

dove $\sigma(R) = \partial(XR + RX)/\partial X$ e $\sigma(X) = \partial(XR + RX)/\partial R$. Con questa impostazione si può progettare un algoritmo di cammino centrale, introducendo il parametro τ in modo del tutto simile al caso della programmazione lineare. È anche possibile estendere il metodo potenziale di Karmarkar alla programmazione semidefinita. Il nuovo metodo è molto complesso e il lettore può consultare Alizadeh [1995]. Citiamo solo alcuni esempi di problemi formulabili con un modello di programmazione semidefinita.

17.9 ESEMPIO. La separazione di due insiemi di punti in \mathbb{R}^n , $\{x^1, \dots, x^h\}$ e $\{y_1, \dots, y^k\}$ viene normalmente eseguita tramite piani, cioè si cercano un vettore $a \in \mathbb{R}^n$ ed uno scalare b tali che $a x^j \leq b, j := 1, \dots, h$, e $a y^i \geq b, i := 1, \dots, k$. Come noto, questo si può effettuare con la programmazione lineare risolvendo il problema

$$\max \{v - w : v \leq a y^i, \forall i, w \geq a x^j, \forall j, \|a\|_\infty \leq 1\}$$

Si può però, a seconda della natura dei dati, avere la necessità di separare i punti con oggetti geometrici diversi dai piani, ad esempio con ellissoidi, ovvero si cercano una matrice positiva semidefinita P , un vettore a ed uno scalare b (ammettendo quindi ellissoidi che degenerano in paraboloidi) tali che $x^{jT} P x^j + a x^j \leq b, \forall j$, e $y^{iT} P y^i + a y^i \geq b, \forall i$. Indicando $X^j := x^j x^{jT}$ e $Y^i := y^i y^{iT}$, si vede che $x^{jT} P x^j = X^j \bullet P$ e quindi il vincolo è esprimibile con la formalizzazione della programmazione semidefinita. Come esercizio si possono definire tutti i dettagli della modellizzazione. ■

17.10 ESEMPIO. Sono state inoltre individuate delle connessioni particolarmente significative fra la programmazione semidefinita e l'ottimizzazione combinatoria. Si veda ad esempio Alizadeh [1995] oppure Grötschel e Lovász [1995], Laurent [1995].

Come esempio di applicazione consideriamo l'approccio ai problemi di programmazione lineare 0-1 tramite la procedura d'innalzamento e proiezione, proposta in sezione 13.5. In tale approccio il vincolo $y_{ij} = y_{ii} y_{jj}$ ($= x_i x_j$) non può essere presente a causa della sua non linearità (e anche non convessità). Però possiamo sfruttare il fatto che una matrice definita dalla forma diadica $y := x x^T$ è positiva semidefinita e quindi, se pretendiamo dai valori y di formare una matrice positiva semidefinita, non perdiamo soluzioni ammissibili del problema originario. In generale questo vincolo restringe ulteriormente il rilassamento dato da \tilde{P}_x (notazione come in sezione 13.5). Sia dato

$$\begin{aligned} \max \quad & x_1 + x_2 + x_3 \\ & 2x_1 + 2x_2 + 2x_3 \leq 5 \\ & x_i \in \{0, 1\} \end{aligned}$$

(si tratta dell'esempio 13.5 in una dimensione maggiore, dato che in \mathbb{R}^2 non si può ottenere il restringimento voluto) e trascuriamo, a scopi illustrativi del metodo, l'ovvia possibilità di abbassare il termine destro dell'equazione. Applicando la procedura d'innalzamento e proiezione, dalla diseuguaglianza $2x_1 + 2x_2 + 2x_3 \leq 5$ si ottengono le 6 diseuguaglianze

$$\begin{aligned} -3y_{11} + 2y_{12} + 2y_{13} &\leq 0 & 5y_{11} + 2y_{22} + 2y_{33} - 2y_{12} - 2y_{13} &\leq 5 \\ -3y_{22} + 2y_{21} + 2y_{23} &\leq 0 & 5y_{22} + 2y_{11} + 2y_{33} - 2y_{21} - 2y_{23} &\leq 5 \\ -3y_{33} + 2y_{31} + 2y_{32} &\leq 0 & 5y_{33} + 2y_{11} + 2y_{22} - 2y_{31} - 2y_{32} &\leq 5 \end{aligned} \quad (17.21)$$

Risolvendo con la programmazione lineare si ottiene l'ottimo

$$\hat{Y} = \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{7}{12} & \frac{3}{8} \\ 1 & \frac{3}{8} & \frac{11}{12} \end{pmatrix} \quad (17.22)$$

di valore 2.5, il medesimo valore del semplice rilassamento d'interesse, e quindi senza un effettivo rafforzamento. Però gli autovalori di \hat{Y} sono $(2.19797, 0.357293, -0.0552676)$ e quindi \hat{Y} non è semidefinita positiva. Allora aggiungendo il vincolo $\hat{Y} \succeq \mathbf{0}$ la soluzione (17.22) viene 'tagliata' e, risolvendo il problema con la programmazione semidefinita, si può sperare di ottenere una limitazione superiore migliore di 2.5.

Tuttavia questo esempio è troppo elementare affinché l'ottimo dell'istanza di programmazione semidefinita ricavata dalle diseuguaglianze quadratiche linearizzate produca una migliore limitazione inferiore. In ogni caso è utile vedere come il problema viene affrontato. Dalle diseuguaglianze (17.21) si ricava l'istanza di programmazione semidefinita

$$\begin{aligned} \max \quad & C \bullet Y \\ & A_i \bullet Y \leq b_i, \quad i := 1, \dots, 6 \\ & Y \succeq \mathbf{0} \end{aligned} \quad (17.23)$$

con i seguenti dati:

$$C := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad A_1 := \begin{pmatrix} -3 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad A_2 := \begin{pmatrix} 5 & -1 & -1 \\ -1 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix}$$

$$A_3 := \begin{pmatrix} 0 & 1 & 0 \\ 1 & -3 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad A_4 := \begin{pmatrix} 2 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 2 \end{pmatrix} \quad A_5 := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & -3 \end{pmatrix} \quad A_6 := \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 5 \end{pmatrix}$$

e $b = (0, 5, 0, 5, 0, 5)^T$ (non sono stati inseriti i vincoli $y_{ij} \leq 1$; a posteriori si vedrà che sono soddisfatti dall'ottimo). Applicando l'iterazione descritta precedentemente si ottiene la soluzione ottima primale

$$\hat{Y} := \begin{pmatrix} \frac{5}{6} & \frac{5}{8} & \frac{5}{8} \\ \frac{5}{8} & \frac{5}{6} & \frac{5}{8} \\ \frac{5}{8} & \frac{5}{8} & \frac{5}{6} \end{pmatrix} \quad (17.24)$$

di autovalori $5/24$, $5/24$ e $25/12$ e quindi positiva definita, e la soluzione ottima duale $\hat{u} := (1/6) \mathbf{1}$. Il lettore può per esercizio verificare che le condizioni di complementarità sono soddisfatte e quindi si tratta veramente dell'ottimo. Si può anche notare che la soluzione (17.24) è combinazione convessa con coefficienti uguali dei tre vertici dati da (17.22) e dalle altre espressioni ottenute da (17.22) per rotazione degli indici, a conferma che, in presenza di soluzioni multiple, i metodi ai punti interni trovano la combinazione convessa dei vari ottimi.

Però il valore ottimo è 2.5, esattamente uguale al rilassamento d'interrezza e non abbiamo ottenuto il restringimento sperato.

Proviamo allora a modificare la funzione obiettivo in modo da evitare soluzioni multiple. Ad esempio poniamo $c := (1.2, 1, 1.1)$. Con questo valore l'ottimo (unico) del normale rilassamento d'interrezza è $\bar{x}_1 = 1$, $\bar{x}_2 = 1$, $\bar{x}_3 = 0.5$, con valore ottimo 2.8. Risolvendo il problema (17.21) si ottiene nuovamente (17.22) con valore ottimo 2.79167. La diminuzione della limitazione superiore è pressoché trascurabile. Però sappiamo già che questa soluzione non è positiva definita e possiamo tentare l'approccio con la programmazione semidefinita. Risolvendo (17.23) si ottiene

$$\hat{Y} = \begin{pmatrix} 0.970749 & 0.538849 & 0.917272 \\ 0.538849 & 0.638482 & 0.418871 \\ 0.917272 & 0.418871 & 0.890764 \end{pmatrix}$$

$$\hat{u} = (0.153419, 0.209273, 0.196956, 0.160791, 0.168436, 0.186582)$$

di valore ottimo 2.78322. Anche se di molto poco (ma molto meglio in confronto all'innalzamento e proiezione), c'è stato un restringimento nella limitazione superiore. Il valore ottimo del problema intero è 2.3. ■