Il metodo dei “Piani di Taglio” (Cutting Planes Method)

E’ un metodo di soluzione dei problemi (IP) di tipo generale.

L’idea di base:

Se la soluzione di (RL) non è intera allora la soluzione ottima intera è interna al poliedro P.

Si aggiungono vincoli a P cercando di “restringerlo”, in particolare eliminando solamente parti di P che non contengono soluzioni intere.

Si risolve una sequenza di problemi rilassati sempre più vincolati.

Sia

$$
\text{(IP)} \quad \begin{aligned} \max\ x_0 &= \underline{c}^T \underline{x} \\ A\underline{x} &= \underline{b} \\ \underline{x} &\in \mathbf{Z}_+^n \end{aligned} \quad \Longrightarrow \quad P = \left\{ \underline{x} \in \mathbf{Z}^n : A\underline{x} = \underline{b}, \underline{x} \geq \underline{0} \right\}
$$

$\underline{x}^*$ la soluzione ottima di (IP)

Il primo problema considerato è il problema (RL) ottenuto rilassando (IP)

$$
\text{(RL)} \quad \begin{aligned} \max\ x_0 &= \underline{c}^T \underline{x} \\ A\underline{x} &= \underline{b} \\ \underline{x} &\in \mathbf{R}_+^n \end{aligned} \quad \Longrightarrow \quad P_0 = \left\{ \underline{x} \in \mathbf{R}^n : A\underline{x} = \underline{b}, \underline{x} \geq \underline{0} \right\}
$$

$\underline{x}^\circ$ la soluzione ottima di (RL)

Si può costruire una sequenza di poliedri, detta sequenza di Gomory, tale che:

- $P_0 \supset P_1 \supset \ldots \supset P_t$
- $P_i \cap \mathbf{Z}^n = P$
- $\underline{x}^\circ_{i-1} \notin P_i$
- $\underline{x}^\circ_t \equiv \underline{x}^*$

La sequenza è costruita aggiungendo via via a P_0 un insieme di vincoli detti <u>tagli</u>.

Definizione.

Una disuguaglianza $\underline{a}^T \underline{x} \geq a_0$ è un **taglio** per un poliedro P' associato al (RL) di un problema (IP) se, detta $\underline{x}°$ la soluzione ottima non intera del (RL), si ha:

1) $\underline{a}^T \underline{y} \geq a_0$ $\quad\forall \underline{y}$ soluzione ammissibile di (IP) (la disuguaglianza si dice **valida**);

2) $\underline{a}^T \underline{x}^\circ < a_0$ $\quad$(non è soddisfatta da $\underline{x}°$)

Il metodo dei piani di taglio determina la soluzione ottima intera introducendo <u>un numero finito</u> di tagli.
Ogni taglio <u>separa</u> la soluzione non intera del (RL) corrente dalle soluzioni ammissibili per (IP).

Il taglio di Gomory (Taglio Frazionario)

Supponiamo di avere determinato la soluzione ottima di un (RL) di (IP). Le m variabili in base possono essere espresse come:

$$x_{B_i} = y_{i0} - \sum_{j \in R} y_{ij} x_j \qquad i = 1, \ldots, m \qquad (1)$$

dove R è l'insieme degli indici delle variabili fuori base.
La soluzione corrispondente è quindi:

$$x_{B_i} = y_{i0} \quad i = 1, \ldots, m; \qquad x_j = 0 \qquad \forall j \in R$$

Supponiamo che non tutti i y_{i0} siano interi. Scegliamo una componente della base con valore non intero e cerchiamo di definire le condizioni che devono essere rispettate perchè essa sia invece intera.

Sia i la componente non intera. Allora

$$y_{i0} = \lfloor y_{i0} \rfloor + f_{i0} \qquad 0 < f_{i0} < 1$$
$$y_{ij} = \lfloor y_{ij} \rfloor + f_{ij} \qquad 0 \leq f_{ij} < 1$$

dove $\lfloor a \rfloor$ rappresenta il più grande intero che non supera a.

Si può riscrivere il primo membro della i-esima equazione (1)

$$y_{i0} = x_{B_i} + \sum_{j\in R} \lfloor y_{ij} \rfloor x_j + \underbrace{\sum_{j\in R} f_{ij} x_j}_{\geq 0} \geq x_{B_i} + \sum_{j\in R} \lfloor y_{ij} \rfloor x_j$$

cioè $$\underbrace{x_{B_i} + \sum_{j\in R} \lfloor y_{ij} \rfloor x_j}_{(*)} \leq y_{i0} = \lfloor y_{i0} \rfloor + f_{i0}$$

Poichè si vuole imporre che x_{B_i} sia intero, (*) risulta intero.
Ma allora se (*) è intero non può essere superiore a $\lfloor y_{i0} \rfloor$
poichè f_{i0}<1.
Quindi

$$x_{B_i} + \sum_{j\in R} \lfloor y_{ij} \rfloor x_j \leq \lfloor y_{i0} \rfloor$$

Cambiando segno alla disuguaglianza e sostituendo x_{B_i} dalla i-esima equazione (1)

$$\sum_{j\in R} \lfloor y_{ij} \rfloor x_j + \sum_{j\in R} f_{ij} x_j - \lfloor y_{i0} \rfloor - f_{i0} - \sum_{j\in R} \lfloor y_{ij} \rfloor x_j \geq -\lfloor y_{i0} \rfloor$$

quindi semplificando, si ottiene il <u>taglio di Gomory</u> (taglio frazionario)

$$\boxed{\sum_{j\in R} f_{ij} x_j \geq f_{i0}}$$

6. Teorema

Per ogni componente i non intera della soluzione di un (RL) la disequazione

$$\sum_{j\in R} f_{ij}x_j \geq f_{i0} \qquad (2)$$

è un taglio rispetto al poliedro P.

Dim. Dimostriamo che (2) è soddisfatta da ogni soluzione ammissibile di (IP).

Supponiamo per assurdo che esista una soluzione $\underline{w}$ ammissibile per (IP) tale che

$$\sum_{j\in R} f_{ij}w_j < f_{i0} \qquad (\#)$$

Poichè $\underline{w}$ è ammissibile ogni sua componente soddisfa il sistema

$$A\underline{x} = B\underline{x}_B + N\underline{x}_N = \underline{b}$$

in particolare la componente w_i soddisfa l'i-esima equazione

$$w_i = y_{i0} - \sum_{j\in R} y_{ij}w_j = \lfloor y_{i0} \rfloor + f_{i0} - \sum_{j\in R} \lfloor y_{ij} \rfloor w_j - \sum_{j\in R} f_{ij}w_j$$

poichè $\underline{w}$ è intero $f_{i0} - \sum_{j\in R} f_{ij}w_j = \beta$ deve essere intero, e poichè

$f_{ij} > 0$ e $w_j \geq 0 \quad \Rightarrow \quad \beta < f_{i0} < 1$ allora β intero è negativo o nullo

$\beta \leq 0 \Rightarrow -\beta \geq 0 \Rightarrow \sum_{j\in R} f_{ij}w_j \geq f_{i0}$ che contraddice l'ipotesi (#).

Dimostriamo che (2) non è soddisfatta dalla soluzione di base non intera $\underline{x}°$ da cui la disequazione (2) è stata generata ((2) separa x° dal nuovo poliedro ottenuto aggiungendo (2) al poliedro di (RL)).

Sia $x_i°$ la componente i-esima frazionaria della soluzione, verifichiamo che $\underline{x}°$ soddisfa

$$\sum_{j \in R} f_{ij} x_j^{\circ} < f_{i0} \qquad (*)$$

Poichè la componente i-esima della soluzione di base è frazionaria si ha

$$x_i^{\circ} > 0 \Rightarrow x_i^{\circ} = \lfloor y_{i0} \rfloor + f_{i0} > 0 \Rightarrow f_{i0} > 0$$

e sostituendo la soluzione in (*) si ottiene

$$x_j^{\circ} = 0 \qquad \forall j \in R \qquad \Rightarrow 0 < f_{i0}$$

quindi x° soddisfa la (*). ■

Un esempio

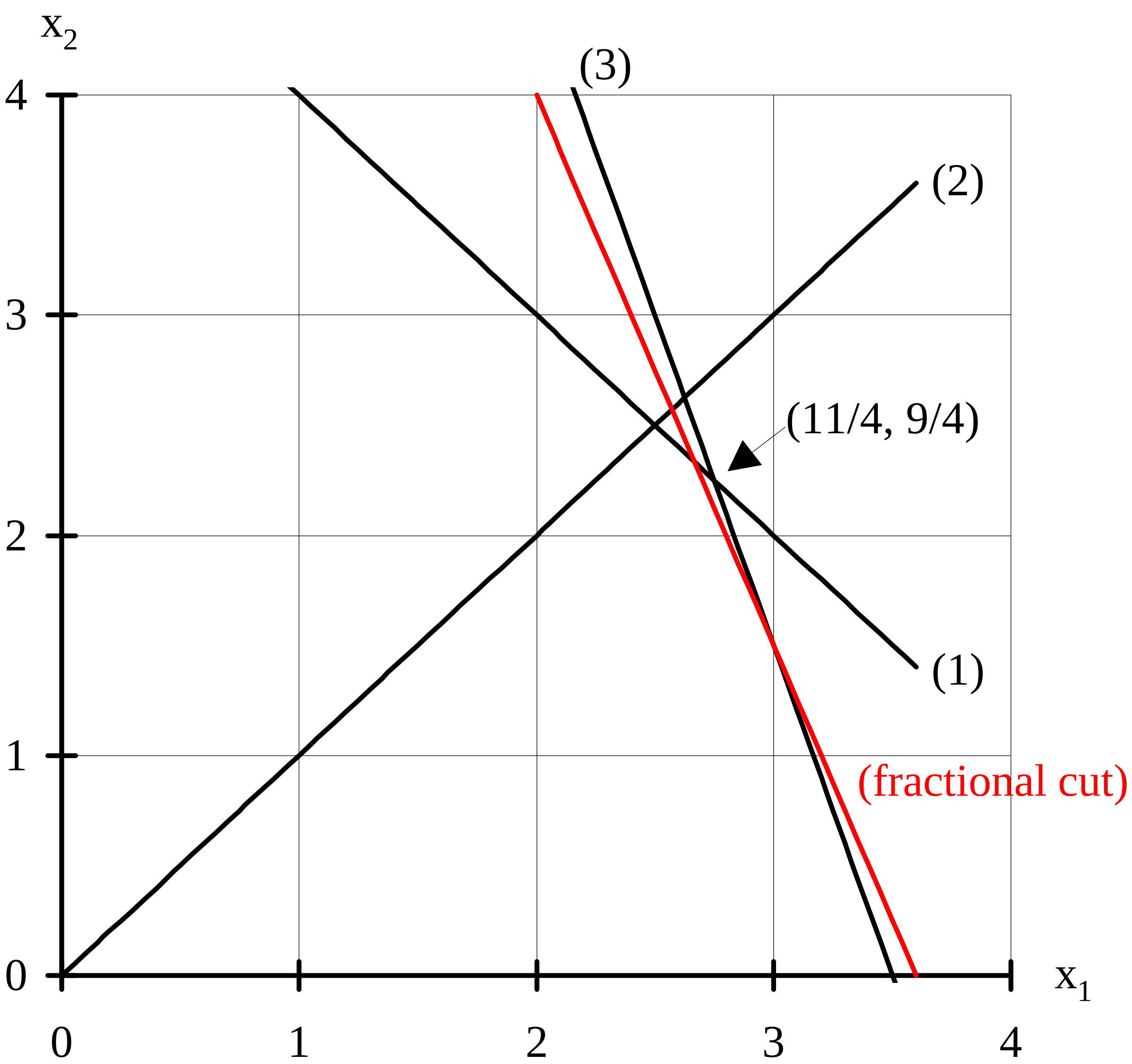
$$\begin{aligned} \max x_0 = 2x_1 + \ & x_2 & \\ x_1 + \ & x_2 \le 5 & (1) \\ -x_1 + \ & x_2 \le 0 & (2) \\ 6x_1 + 2 & x_2 \le 21 & (3) \\ x_1, x_2 \ge 0 & & \\ x_1, x_2 \in \mathbf{Z} & & \end{aligned} \quad \Longrightarrow \quad \begin{aligned} \max x_0 &= 2x_1 + x_2 \\ x_1 + x_2 + x_3 &= 5 \\ -x_1 + x_2 + x_4 &= 0 \\ 6x_1 + 2x_2 + x_5 &= 21 \\ x_1, x_2, x_3, x_4, x_5 &\ge 0 \\ x_1, x_2 &\in \mathbf{Z} \end{aligned}$$

Costruiamo un taglio dalla seconda riga del tableau ottimo di (RL)

	x_3	x_4	x_5	x_1	x_2	
x_0	$1/2$	0	$1/4$	0	0	$31/4$
x_2	$3/2$	0	$-1/4$	0	1	$9/4$
x_4	-2	1	$1/2$	0	0	$1/2$
x_1	$-1/2$	0	$1/4$	1	0	$11/4$

i coefficienti
$$\frac{3}{2} = 1 + \frac{1}{2}$$
$$-\frac{1}{4} = -1 + \frac{3}{4}$$
$$\frac{9}{4} = 2 + \frac{1}{4}$$

il taglio $\frac{1}{2}x_3 + \frac{3}{4}x_5 \ge \frac{1}{4}$



sostituendo x_1 ed x_2 in $\frac{1}{2}x_3 + \frac{3}{4}x_5 \geq \frac{1}{4}$

si ottiene $5x_1 + 2x_2 \leq 18$

che non è soddisfatta dal punto (11/4, 9/4)

Il Metodo del Branch-and-Bound

E’ una strategia di esplorazione dello spazio delle soluzioni basata su l’enumerazione implicita delle soluzioni:

il metodo esamina sottinsiemi disgiunti di soluzioni (branching) e li valuta sulla base di una stima della funzione obiettivo (bounding), eliminando quegli insiemi di soluzioni che non contengono la soluzione ottima.

L’esplorazione viene effettuata risolvendo una sequenza di (RL) associati a sottinsiemi disgiunti di soluzioni.

Consideriamo il problema a numeri interi

$$
\begin{aligned}
\text{(IP)} \quad \max\ x_0 &= \underline{c}^T \underline{x} \\
A\underline{x} &= \underline{b} \\
\underline{x} &\in \mathbf{Z}_+^n
\end{aligned}
$$

si risolve il (RL) associato, determinando $\underline{x}^0$.

Scelta x_j^0 una componente non intera della soluzione, la regione di ammissibilità del (RL) costituita dal poliedro

$$P' = \left\{ \underline{x} \in \mathbf{R}^n : A\underline{x} = \underline{b}, \underline{x} \geq \underline{0} \right\}$$

può essere partizionata in due regioni disgiunte, dando luogo a due nuovi problemi rilassati ottenuti aggiungendo a P rispettivamente uno tra i due seguenti vincoli.

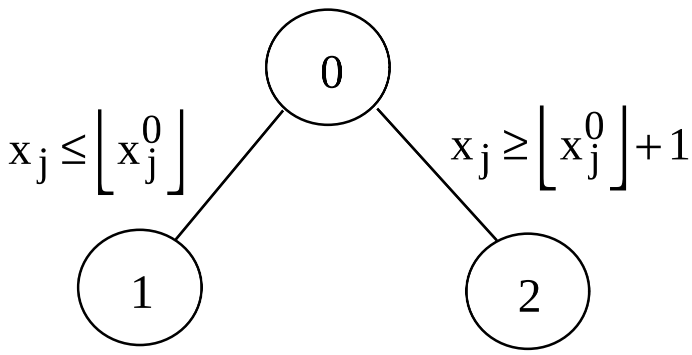
$$x_j \leq \left\lfloor x_j^0 \right\rfloor \qquad\qquad x_j \geq \left\lfloor x_j^0 \right\rfloor + 1$$

In questo modo si separa il problema originale in due problemi disgiunti. Quindi si considerano i nuovi due problemi iterando il procedimento.

L'esplorazione dello spazio delle soluzioni effettuato dal metodo del Branch-and-Bound si può rappresentare per mezzo di un albero ("enumeration tree").

Ad ogni nodo è associato un (RL)

nodo 0 $\Rightarrow$ (RL) associato al (IP) originale

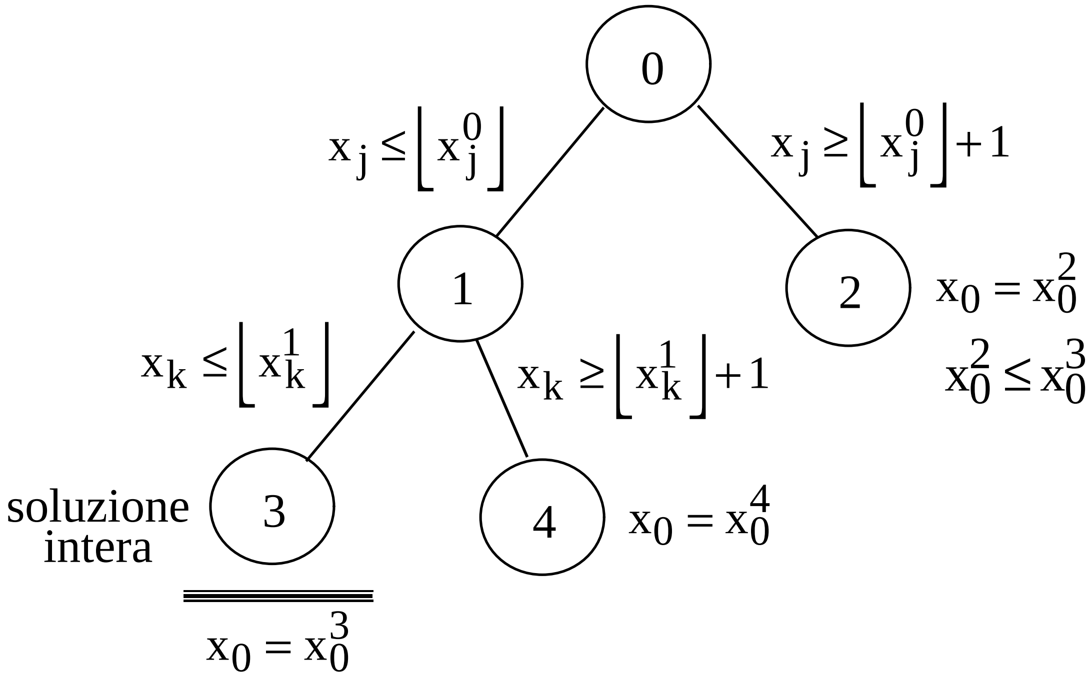


Si procede risolvendo gli (RL) associati ai nuovi nodi 1 e 2. Se la soluzione non è intera si ramifica (branching) ulteriormente l'albero. Perchè l'esplorazione non sia totale, il branching deve essere fermato quando:

- il (RL) associato ad un nodo non ammette soluzione;
- la soluzione di (RL) associato al nodo è intera;
- la soluzione di (RL) associato al nodo non è intera, ma è possibile stabilire che l'esplorazione della porzione di albero al di sotto del nodo non conduce alla soluzione ottima (bounding).

Bounding: il principio base

Supponiamo di aver trovato una soluzione intera per (IP) risolvendo il (RL) associato al nodo 3.



Il valore dell'obiettivo $x_0 = x_0^3$ rappresenta la miglior soluzione corrente. Poichè stiamo massimizzando, è il limite inferiore (<u>lower bound</u>) corrente per la soluzione di (IP).

Quando si ramifica un nodo, il valore dell'obiettivo associato ai nodi successori non può essere migliore dell'obiettivo associato al nodo padre, poichè i (RL) dei nodi successori sono più vincolati del (RL) del nodo padre.

Il valore dell'obiettivo associato ad un nodo il cui (RL) ha soluzione non intera rappresenta quindi il limite superiore (<u>upper bound</u>) rispetto qualsiasi soluzione determinabile ramificando tale nodo.

Quindi se, ad esempio, $x_0^2 \le x_0^3$ non ha più senso esplorare l'albero sotto il nodo 2 poichè le soluzioni associate ai suoi nodi non possono migliorare la soluzione intera corrente.

In generale, dati:

- il $(RL)^i$ associato ad un nodo i (dove $A^i\underline{x}=\underline{b}^i$ è l'insieme dei vincoli originali più i vincoli via via aggiunti ad ogni branching)

$$
\begin{aligned}
(RL)^i \quad \max x_0 &= \underline{c}^T \underline{x} \\
A^i \underline{x} &= \underline{b}^i \\
\underline{x} &\in \mathbf{R}^n_+
\end{aligned}
$$

- $\underline{x}^{lb}$, la migliore soluzione intera sino ad ora trovata esplorando l'albero
- $\underline{x}^{ub}$, la soluzione non intera di $(RL)^i$

se vale $\underline{c}^T \underline{x}^{ub} \leq \underline{c}^T \underline{x}^{lb}$, poichè è noto che $\underline{c}^T \underline{x}^{int} \leq \underline{c}^T \underline{x}^{ub}$

dove $\underline{x}^{int}$ è soluzione ottima intera di

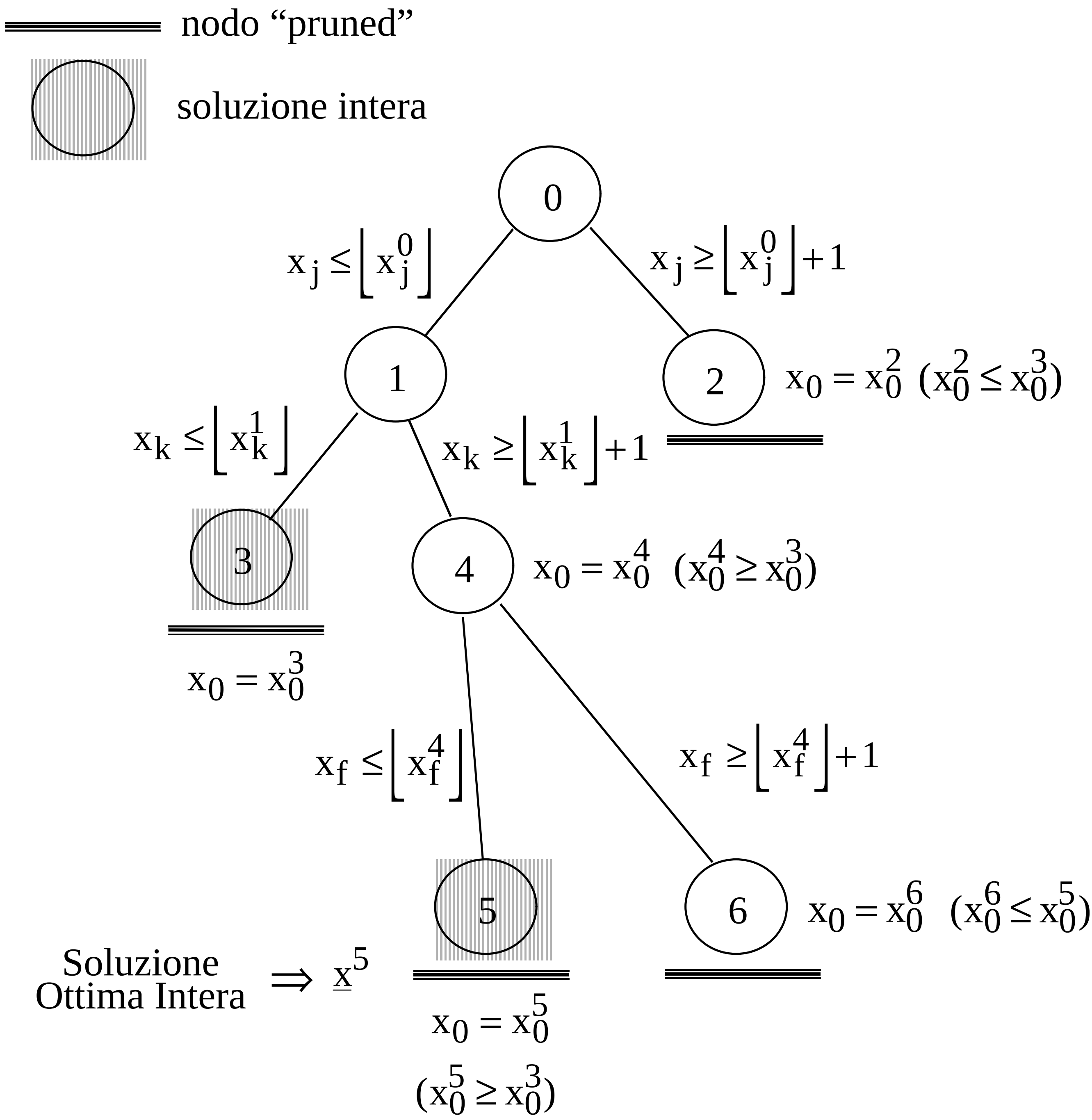
$$
\begin{aligned}
(IP)^i \quad \max x_0 &= \underline{c}^T \underline{x} \\
A^i \underline{x} &= \underline{b}^i \\
\underline{x} &\in \mathbf{Z}^n_+
\end{aligned}
$$

allora ne consegue che $\underline{c}^T \underline{x}^{int} \leq \underline{c}^T \underline{x}^{ub} \leq \underline{c}^T \underline{x}^{lb}$

Quindi la migliore soluzione intera contenuta in

$$P^i = \left\{ \underline{x} \in \mathbf{R}^n : A^i \underline{x} = \underline{b}^i, \underline{x} \geq \underline{0} \right\}$$

non potrà mai essere migliore della soluzione intera corrente $\underline{x}^{lb}$.



I nodi che non vengono tagliati (pruned) sono detti **nodi attivi** e vengono esplorati.

L’esplorazione termina quando non ci sono più nodi attivi.

L'algoritmo del Branch-and-Bound

1. **Inizializzazione**

   Sia (0) il nodo attivo e P_0 il poliedro associato al (RL).

   Sia $z^{lb} = -\infty$ il valore corrente dell'obiettivo (lower bound), e $z^0 = \infty$ il valore iniziale dell'obiettivo del nodo (0) (upper bound di (0)).

2. **Branching**

   Se non esiste un nodo attivo andare al passo 7, altrimenti scegliere il nodo attivo (j).

   Se il (RL) di (j) è stato risolto andare al passo 3, altrimenti al 4.

3. **Separazione**

   Scegliere una variabile frazionaria di base $x_{B_i} = y_{i0}^j$ e partizionare P_j in

   $$P_j \cap \left\{\underline{x}: x_{B_i} \le \left\lfloor y_{i0}^j \right\rfloor\right\} \qquad P_j \cap \left\{\underline{x}: x_{B_i} \ge \left\lfloor y_{i0}^j \right\rfloor + 1\right\}$$

   creando due nuovi nodi con lo stesso upper bound del nodo j.

   Vai a 2.

4. **Soluzione di (RL)**

   Risolvere il (RL) associato a (j). Se non esiste soluzione tagliare il nodo (il nodo non è più attivo) ed andare a 2.

   Se esiste soluzione ottima $\underline{x}^j$ porre $z^j = x_0^j$ ed andare al passo 5.

5. **Pruning per integrità**

   Se $\underline{x}^j$ non è intera andare al passo 6. Altrimenti tagliare il nodo (j) e porre $z^{lb} = \max\{z^{lb}, z^j\}$. Se il lower bound viene aggiornato, la nuova soluzione corrente è quella associata al nodo (j).

   Andare al passo 6.

6. **Pruning per bound**

   Tagliare ogni nodo attivo (k) tale che $z^k \le z^{lb}$

   Andare al passo 2.

7. **Terminazione**

   L'algoritmo termina.

   Se $z^{lb} = -\infty$ allora (IP) non ammette soluzione.

   Se $z^{lb} \ge -\infty$ la soluzione corrente associata a z^{lb} è ottima.

Osservazioni sull'algoritmo di Branch-and-Bound

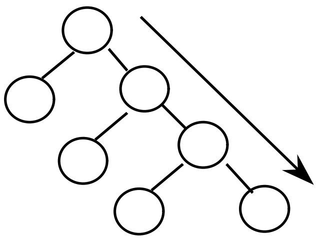
- La convergenza in un numero finito di passi è garantita se il problema ha soluzioni ottime finite.

- Le prestazioni dell'algoritmo sono influenzate dalle diverse politiche di scelta del nodo da ramificare al passo 3.

  Due strategie limite:

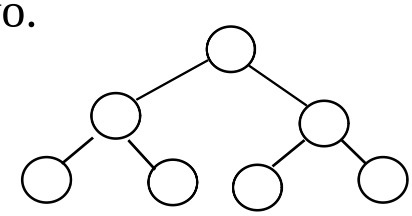
  - Depth First (esplorazione in profondità):

    se il nodo corrente non è tagliato generare i due nodi figli e proseguire ad esplorare uno di essi al livello successivo.



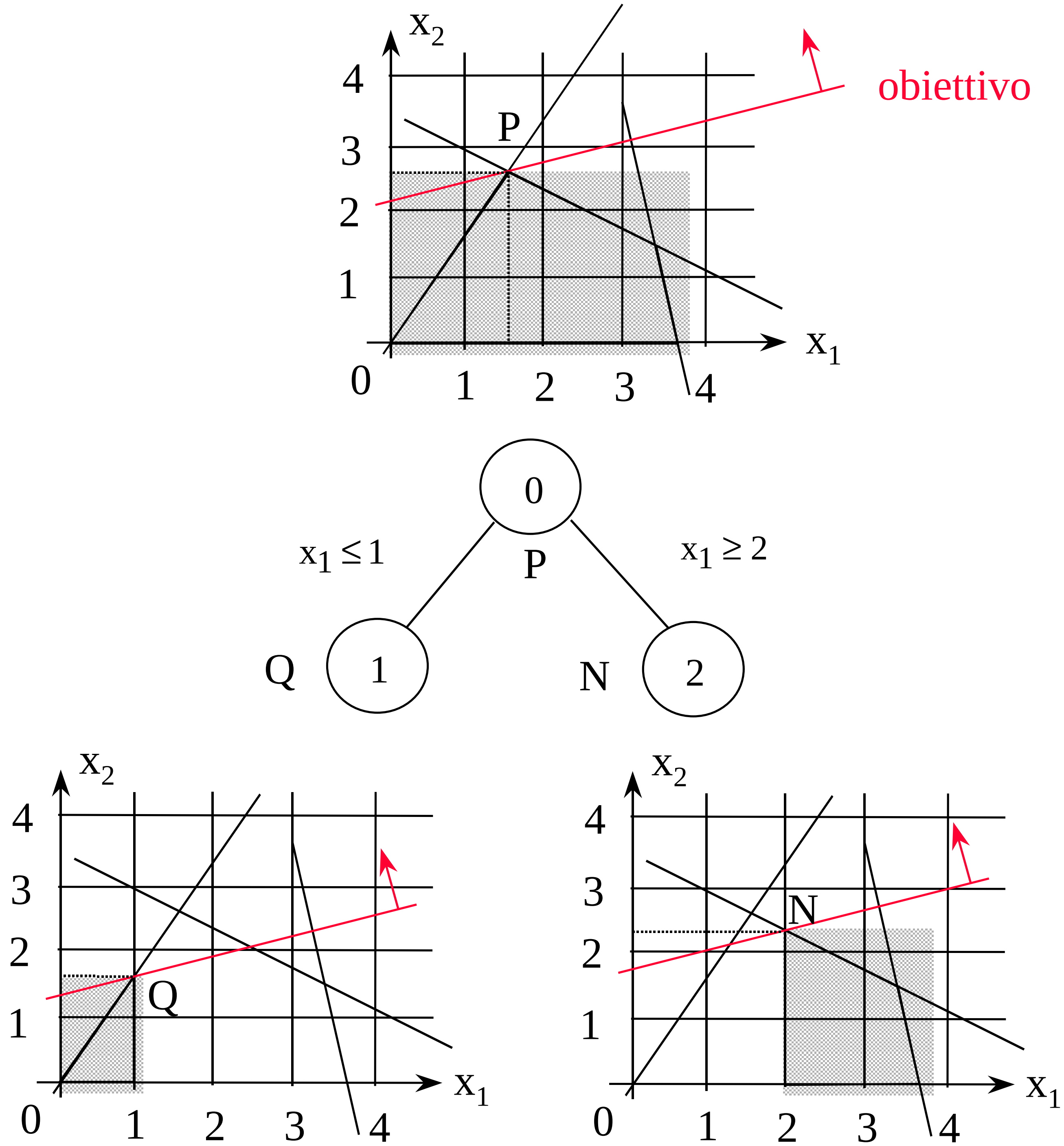
  - Breadth First (esplorazione in ampiezza):

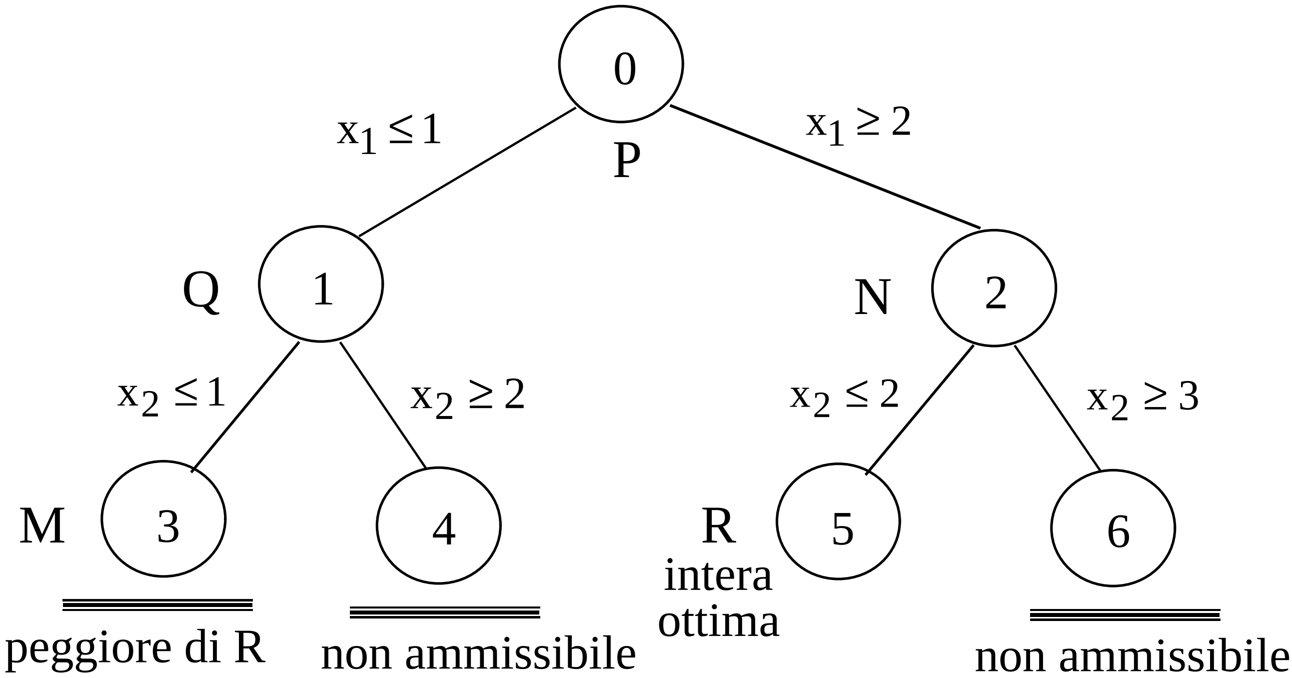
    si esplorano tutti i nodi allo stesso livello prima di passare al livello successivo.

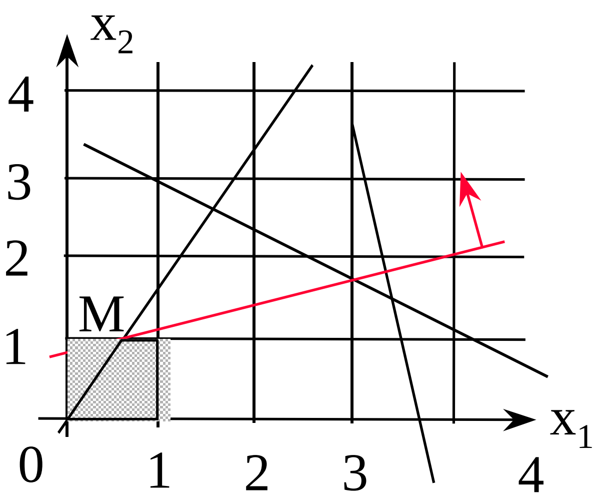


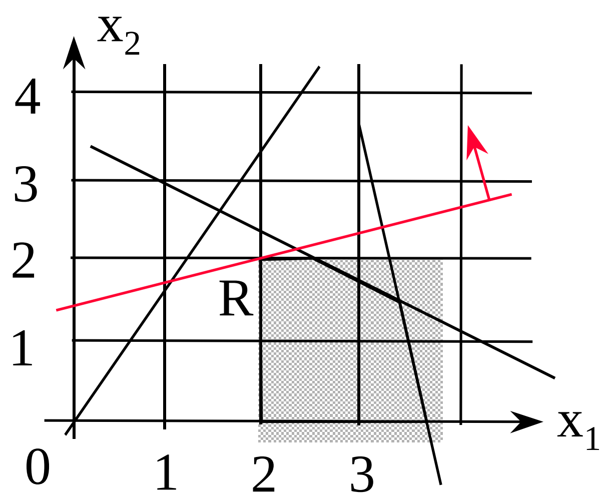
La strategia depth first richiede minor occupazione di memoria della breadth first ma può richiedere maggior tempo

Un esempio (grafico)




0
P
x1 ≤ 1
x1 ≥ 2
Q
1
N
2
x2 ≤ 1
x2 ≥ 2
x2 ≤ 2
x2 ≥ 3
M
3
4
R
5
6
intera
ottima
peggiore di R
non ammissibile
non ammissibile


x2
4
3
2
1
M
0
1
2
3
4
x1


x2
4
3
2
1
R
0
1
2
3
x1