

L'algoritmo del simplesso

Ricerca operativa

Giovanni Righini



UNIVERSITÀ DEGLI STUDI
DI MILANO

Forma canonica

Dalla forma standard di un problema di PL

$$\begin{aligned} \text{minimize } z &= c^T x \\ \text{subject to } Ax &= b \\ x &\geq 0 \end{aligned}$$

scegliendo una base (e permutando le colonne di conseguenza) si ha

$$\begin{aligned} \text{minimize } z &= c_B^T x_B + c_N^T x_N \\ \text{subject to } Bx_B + Nx_N &= b \\ x_B, x_N &\geq 0. \end{aligned}$$

Moltiplicando a sinistra per B^{-1} :

$$\begin{aligned} \text{minimize } z &= c_B^T x_B + c_N^T x_N \\ \text{subject to } Ix_B + (B^{-1}N)x_N &= B^{-1}b \\ x_B, x_N &\geq 0. \end{aligned}$$

da cui si ha la soluzione di base $x_B = B^{-1}b - (B^{-1}N)x_N$.

Forma canonica

Sostituendo $x_B = B^{-1}b - (B^{-1}N)x_N$ in z si ha:

$$\begin{aligned} \text{minimize } z &= c_B^T B^{-1}b + (c_N^T - c_B^T B^{-1}N)x_N \\ \text{subject to } Ix_B + (B^{-1}N)x_N &= B^{-1}b \\ x_B, x_N &\geq 0. \end{aligned}$$

che si può riscrivere in modo più compatto

$$\begin{aligned} \text{minimize } z &= z_B + \bar{c}_N^T x_N \\ \text{subject to } Ix_B + \bar{N}x_N &= \bar{b} \\ x_B, x_N &\geq 0. \end{aligned}$$

Quando si pone $x_N = 0$ si ha $x_B = B^{-1}b = \bar{b}$. Se $\bar{b} \geq 0$, allora la soluzione di base è ammissibile.

Forma canonica

Esistono un numero combinatorio di forme canoniche: tante quante le possibili scelte della base.

$$\begin{aligned} \text{minimize } z &= z_B + \bar{c}_N^T x_N \\ \text{subject to } Ix_B + \bar{N}x_N &= \bar{b} \\ x_B, x_N &\geq 0. \end{aligned}$$

Un problema di PL è in forma canonica se e solo se:

- i coefficienti delle variabili di base x_B formano una matrice identità $m \times m$;
- le variabili di base x_B non compaiono nella funzione obiettivo.

Inoltre la forma canonica è forte se e solo se:

- i termini noti dei vincoli sono non-negativi ($\bar{b} \geq 0$).

Una forma canonica debole corrisponde ad una soluzione di base non ammissibile.

Il *tableau*

Una volta posto in forma canonica, un problema di PL si può rappresentare in una matrice, detta *tableau*, che è la struttura dati fondamentale sulla quale opera l'algoritmo del simplesso.

Per esempio:

$$\begin{array}{llllll}
 \text{minimize } z = & -x_1 & -2x_2 & & & \\
 \text{s.t.} & -x_1 & +2x_2 & +x_3 & & = 8 \\
 & x_1 & +x_2 & & +x_4 & = 10 \\
 & x_1 & & & & +x_5 = 7 \\
 & x \geq 0 & & & &
 \end{array}$$

è in forma canonica con le variabili di slack in base.

0	-1	-2	0	0	0
8	-1	2	1	0	0
10	1	1	0	1	0
7	1	0	0	0	1

$-z_B$	\bar{c}_N^T	0
\bar{b}	\bar{N}	I

L'algoritmo

```
while ( $\neg \text{Infeasible}(b, c)$ )  $\wedge$  ( $\neg \text{FeasibleBase}(b)$ ) do
    Pivot( $A, b, c$ )
if  $\text{Infeasible}(b, c)$  then
    Stop: problema inammissibile
else
    while ( $\neg \text{Optimal}(c)$ )  $\wedge$  ( $\neg \text{Unbounded}(A, c)$ ) do
        Pivot( $A, b, c$ )
    if  $\text{Optimal}(c)$  then
        Stop: soluzione ottima
    else
        Stop: problema illimitato
```

$Pivot(A, b, c)$

Ogni iterazione (*pivoting*) consiste in un cambio di base: una variabile in base esce dalla base e una variabile fuori base entra in base.

0	-1	-2	0	0	0
8	-1	2	1	0	0
10	1	1	0	1	0
7	1	0	0	0	1

variabili fuori
base (?)

$x = \begin{bmatrix} 0 \\ 0 \\ 8 \\ 10 \\ 7 \end{bmatrix}$

$z = 0$

variabili di
surplus

1. Scegliere un elemento *pivot* positivo su una colonna fuori base.

0	-1	-2	0	0	0
8	-1	2	1	0	0
10	1	1	0	1	0
7	1	0	0	0	1

$r = 1 \quad c = 2$

$Pivot(A, b, c)$

0	-1	-2	0	0	0
8	-1	2	1	0	0
10	1	1	0	1	0
7	1	0	0	0	1

$$r = 1 \quad c = 2$$

2. Dividere la riga r per il pivot.

0	-1	-2	0	0	0
4	-1/2	1	1/2	0	0
10	1	1	0	1	0
7	1	0	0	0	1

$Pivot(A, b, c)$

0	-1	-2	0	0	0
4	-1/2	1	1/2	0	0
10	1	1	0	1	0
7	1	0	0	0	1

3. Sottrarre ad ogni riga $i \neq r$ la riga r moltiplicata per a_{ic} .

8	-2	0	1	0	0
4	-1/2	1	1/2	0	0
6	3/2	0	-1/2	1	0
7	1	0	0	0	1

$$x = \begin{bmatrix} 0 \\ 4 \\ 0 \\ 6 \\ 7 \end{bmatrix} \quad z = -8$$

Entra in base la colonna del pivot ($c = 2$).

Esce di base la colonna corrispondente alla riga del pivot ($r = 1$) nella matrice identità.

Interpretazione algebrica

Dal punto di vista algebrico, l'iterazione corrisponde a riformulare in modo equivalente il sistema di m equazioni in n variabili.

$$\begin{array}{c|ccccc} 0 & -1 & -2 & 0 & 0 & 0 \\ 8 & -1 & 2 & 1 & 0 & 0 \\ 10 & 1 & 1 & 0 & 1 & 0 \\ 7 & 1 & 0 & 0 & 0 & 1 \end{array} \quad \left\{ \begin{array}{lcl} z & = & -x_1 - 2x_2 \\ x_3 & = & 8 + x_1 - 2x_2 \\ x_4 & = & 10 - x_1 - x_2 \\ x_5 & = & 7 - x_1 \end{array} \right.$$

Facendo pivot sull'elemento in riga 1 e colonna 2, entra in base x_2 ed esce di base x_3 .

$$x_3 = 8 + x_1 - 2x_2 \Rightarrow x_2 = 4 + \frac{1}{2}x_1 - \frac{1}{2}x_3$$

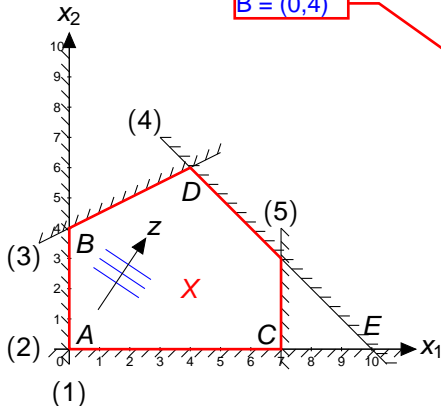
da cui, per sostituzione, si ottiene

$$\begin{array}{c|ccccc} 8 & -2 & 0 & 1 & 0 & 0 \\ 4 & -1/2 & 1 & 1/2 & 0 & 0 \\ 6 & 3/2 & 0 & -1/2 & 1 & 0 \\ 7 & 1 & 0 & 0 & 0 & 1 \end{array} \quad \left\{ \begin{array}{lcl} z & = & -8 - 2x_1 + x_3 \\ x_2 & = & 4 + \frac{1}{2}x_1 + \frac{1}{2}x_3 \\ x_4 & = & 10 - x_1 - x_2 \\ x_5 & = & 7 - x_1 \end{array} \right.$$

Interpretazione geometrica

$$A = (0, 0)$$

$$B = (0, 4)$$



L'iterazione corrisponde allo spostamento da A a B.

Soluzione A: $x = [0 \ 0 \ 8 \ 10 \ 7]$

$B = \{3, 4, 5\}$, $\mathcal{N} = \{1, 2\}$, $z = 0$.

Vincoli attivi: $x_1 \geq 0$, $x_2 \geq 0$.

Soluzione B: $x = [0 \ 4 \ 0 \ 6 \ 7]$

$B = \{2, 4, 5\}$, $\mathcal{N} = \{1, 3\}$, $z = -8$.

Vincoli attivi: $x_1 \geq 0$, $x_3 \geq 0$.

Per determinare l'elemento pivot sono necessarie una **regola di scelta della colonna** e una **regola di scelta della riga**.

Test di ottimalità

Il test di ottimalità, indicato con *Optimal(c)* nello pseudocodice, riguarda i **coefficienti di costo ridotto**.

$$z = z_B + \bar{c}_N^T x_N.$$

I coefficienti \bar{c}_N^T indicano di quanto aumenterebbe la funzione obiettivo da minimizzare se le variabili fuori base, x_N , aumentassero di valore anziché valere 0, cioè entrassero in base.

Quando tutti i **coefficienti di costo ridotto** \bar{c}_N^T sono non-negativi, non esistono direzioni ammissibili miglioranti e questo garantisce **l'ottimalità della soluzione corrente**, se è ammissibile. In tal caso $z^* = z_B$.

N.B.: Possono esistere soluzioni degeneri nelle quali la condizione di ottimalità risulta verificata o no, a seconda della base scelta.

Regole di scelta della colonna

Perciò per ogni iterazione dell'algoritmo del simplesso si sceglie sempre una colonna (variabile entrante in base) che abbia **costo ridotto negativo**.

negative reduced cost

Nell'esempio precedente, partendo dal punto A esistono due colonne con costo ridotto negativo e quindi due possibili modi di fare pivot, entrambi corretti:

- facendo entrare in base x_1 la soluzione si sposta nel punto C;
- facendo entrare in base x_2 la soluzione si sposta nel punto B.

In entrambi i casi z migliora.

Nel punto B invece esiste un solo modo corretto di fare pivot (colonna x_1).

Regole di scelta della colonna

Ferma restando la regola suddetta, possono essere utilizzate diverse strategie. Ad es.:

- la colonna col minimo coefficiente di costo ridotto;
- la colonna che produce il maggior miglioramento di z ;
- la prima colonna con costo ridotto negativo, secondo un ordinamento fissato (regola di Bland);
- una colonna scelta a caso tra quelle con costo ridotto negativo.

La regola di scelta della colonna deve garantire che non si possano provocare cicli infiniti nel caso di soluzioni degeneri.

La regola di scelta della colonna garantisce che l'algoritmo del simplesso raggiunga l'ottimalità (quando una soluzione ottima esiste).

Regole di scelta della riga

La scelta della riga (variabile uscente dalla base) deve garantire l'ammissibilità. A questo scopo, una volta determinata la variabile entrante x_j , cioè lo spigolo del poliedro lungo il quale la soluzione cambia, l'unica possibilità corretta è

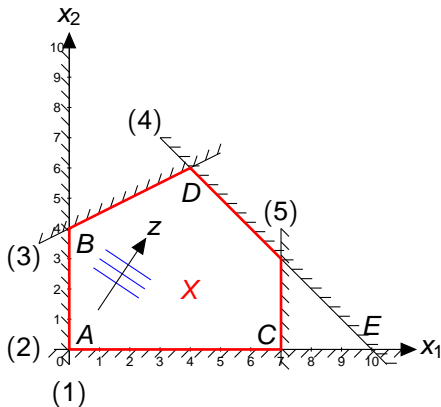
- muoversi verso l'interno del poliedro
- fermarsi appena si incontra una soluzione di base.

Ciò si traduce in

- considerare solo candidati pivot a_{ij} positivi;
- tra le righe i ad essi corrispondenti, scegliere quella che rende minimo il rapporto tra il termine noto b_i ed il candidato pivot a_{ij} .

In caso di ex-aequo, la regola di Bland impone di scegliere la riga di indice minimo, secondo una numerazione prefissata arbitrariamente.

Regole di scelta della riga: esempio



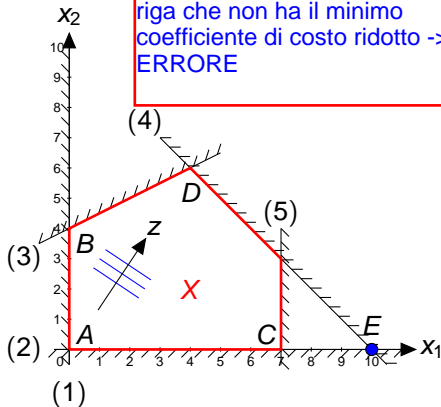
0	-1	-2	0	0	0
8	-1	2	1	0	0
10	1	1	0	1	0
7	1	0	0	0	1

-8	0	-4	-1	0	0
-8	1	-2	-1	0	0
18	0	3	1	1	0
15	0	2	1	0	1

Scegliendo un pivot negativo si ottiene una soluzione non ammissibile (direzione sbagliata).

Regole di scelta della riga: esempio

qui ho scelto come pivot la
riga che non ha il minimo
coefficiente di costo ridotto ->
ERRORE

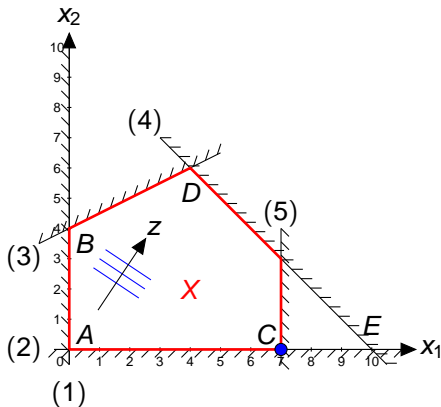


0	-1	-2	0	0	0
8	-1	2	1	0	0
10	1	1	0	1	0
7	1	0	0	0	1

10	0	-1	0	1	0
18	0	3	1	1	0
10	1	1	0	1	0
-3	0	-1	0	-1	1

La direzione è giusta ma il passo è troppo lungo: soluzione inammissibile.

Regole di scelta della riga: esempio



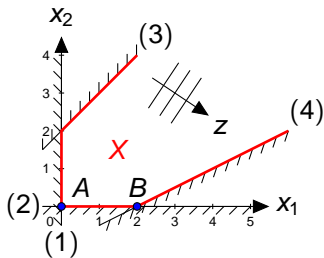
0	-1	-2	0	0	0
8	-1	2	1	0	0
10	1	1	0	1	0
7	1	0	0	0	1

7	0	-2	0	0	1
15	0	2	1	0	1
3	0	1	0	1	-1
7	1	0	0	0	1

Fermandosi sulla prima soluzione di base che si incontra, si ottiene una soluzione ammissibile.

Test di illimitatezza

Se non esistono candidati pivot positivi su una colonna con costo ridotto negativo il problema è **illimitato**.



$$\begin{aligned}
 \text{minimize } z &= -3x_1 + 2x_2 \\
 \text{s.t. } & x_1 - x_2 \geq -2 \\
 & -x_1 + 2x_2 \geq -2 \\
 & x \geq 0
 \end{aligned}$$

$$\begin{array}{c|cccc}
 0 & -3 & 2 & 0 & 0 \\
 \hline
 2 & -1 & 1 & 1 & 0 \\
 2 & 1 & -2 & 0 & 1
 \end{array}$$

A: $B = \{3, 4\}$ $x = (0, 0, 2, 2)$

$$\begin{array}{c|cccc}
 6 & 0 & -4 & 0 & 3 \\
 \hline
 4 & 0 & -1 & 1 & 1 \\
 2 & 1 & -2 & 0 & 1
 \end{array}$$

B: $B = \{1, 3\}$ $x = (2, 0, 4, 0)$

Variabili limitate

Può capitare che alcune variabili siano limitate non solo inferiormente ma anche superiormente:

$$0 \leq l \leq x \leq u.$$

Naturalmente i vincoli $x \geq l$ e $x \leq u$ possono essere trattati come vincoli qualsiasi, ma ciò aumenta inutilmente le dimensioni del modello.

In alternativa è possibile estendere la definizione di soluzione di base, considerando due diversi modi in cui una variabile può essere fuori base: una **soluzione di base estesa** è una soluzione nella quale n variabili hanno un valore pari al loro limite **inferiore o superiore** e le altre m formano un sistema lineare indipendente, cioè una base.

Con opportuni accorgimenti, l'algoritmo del simplesso può lavorare con soluzioni di base estese.

Inizializzazione

L'algoritmo del simplesso **mantiene l'ammissibilità** e **cerca l'ottimalità**, quando è inizializzato con una soluzione di base ammissibile.

Tuttavia può accadere che la soluzione di base iniziale sia inammissibile.

In questo caso, si può procedere in diversi modi per inizializzare l'algoritmo.

Metodo delle variabili artificiali

Data un modello di PL in forma standard (anche non canonica) con n variabili e m vincoli,

$$z = \min\{c^T x : Ax = b, x \in \mathbb{R}_+^n\}, \quad (1)$$

si introduce una variabile artificiale $u_i \geq 0$ con coefficiente unitario in ogni vincolo $i = 1, \dots, m$, definendo così un problema artificiale

$$z^a = \min\{e^T u : Ax + Iu = b, x, u \in \mathbb{R}_+^n\}, \quad (2)$$

dove $e^T = [1, \dots, 1]$.

Vale la prima condizione per la forma canonica (i coefficienti di u formano una matrice identità), ma non la seconda (le variabili u non hanno coefficienti nulli nell'obiettivo).

Per ottenere una forma canonica, si effettua la sostituzione $u_i = b_i - \sum_{j=1}^n a_{ij}x_j$ nell'obiettivo, ottenendo la riga 0 del tableau in forma canonica.

Metodo delle variabili artificiali

La soluzione iniziale $x = 0$, $u = b$ è ammissibile per costruzione e quindi si può eseguire l'algoritmo del simplesso sul problema artificiale (2).

Se il valore ottimo del problema artificiale è nullo, cioè si ottiene $z^a = 0$, allora si è trovata una soluzione ammissibile per il problema originario (1).

Altrimenti si è dimostrata l'inammissibilità del problema originario (1).

Vantaggio: si può sempre applicare.

Svantaggio: può comportare nella prima fase molti passi di pivot; lavora su un tableau più grande di quello originario.

Metodo delle variabili artificiali

Per evitare di dover eseguire m passi di pivot (come minimo) per espellere dalla base le m variabili artificiali, si possono usare alcune delle variabili di slack o surplus del problema originario, scrivendo la forma alle disuguaglianze in modo che $b \geq 0$.

$$\begin{aligned} \min z &= \sum_{j=1}^n c_j x_j \\ \text{s.t. } \sum_{j=1}^n a_{ij} x_j &\leq b_i \quad \forall i \in I_1 \\ \sum_{j=1}^n a_{ij} x_j &\geq b_i \quad \forall i \in I_2 \\ \sum_{j=1}^n a_{ij} x_j &= b_i \quad \forall i \in I_3 \\ x &\geq 0 \end{aligned}$$

$$\begin{aligned} \min z &= \sum_{j=1}^n c_j x_j \\ \text{s.t. } \sum_{j=1}^n a_{ij} x_j + \hat{x}_i &= b_i \quad \forall i \in I_1 \\ \sum_{j=1}^n a_{ij} x_j - \hat{x}_i &= b_i \quad \forall i \in I_2 \\ \sum_{j=1}^n a_{ij} x_j &= b_i \quad \forall i \in I_3 \\ x, \hat{x} &\geq 0 \end{aligned}$$

Metodo delle variabili artificiali

Le variabili \hat{x}_i per i vincoli $i \in I_1$ soddisfano già la prima condizione per la forma canonica.

Sia $h \in I_2$ l'indice del vincolo col massimo valore del termine noto:

$$h = \operatorname{argmax}_{i \in I_2} \{b_i\}.$$

Sottraendo ogni riga $i \in I_2 \setminus \{h\}$ dalla riga h , si ottengono vincoli equivalenti con termine noto non-negativo e variabile \hat{x}_i con coefficiente unitario.

Bisogna quindi inserire variabili artificiali solo per il vincolo h e per i vincoli in I_3 .

Metodo “big M ”

Anziché eliminare dalla formulazione del problema artificiale le variabili x , è possibile mantenerle e penalizzare nell'obiettivo le variabili u con coefficienti molto grandi.

Si risolve quindi il problema

$$z^a = \min\{c^T x + w^T u : Ax + Iu = b, x, u \in \mathbb{R}_+^n\},$$

dove $w^T = [M, \dots, M]$ è un vettore di coefficienti “abbastanza grandi”, cioè tali che ogni soluzione con una variabile $u_i > 0$ abbia costo maggiore del valore ottimo z^* del problema (1).

Vantaggio: non serve una fase di inizializzazione.

Svantaggi:

- non è detto che sia facile determinare il valore appropriato per M (potrebbe esistere una base con $u_i = \epsilon$ per qualche $i = 1, \dots, m$, con ϵ molto piccolo);
- il valore di M potrebbe provocare instabilità numerica.

Metodo di Balinski-Gomory

L'algoritmo trascura temporaneamente la funzione obiettivo e minimizza una misura dell'inammissibilità rispetto ad un vincolo violato, ripetendo l'operazione per tutti i vincoli violati, finché non raggiunge una base ammissibile oppure dimostra che il problema è inammissibile.

Per ogni vincolo violato una misura della violazione è data dal valore assoluto della corrispondente variabile (che ha valore negativo).

Esempio

$$\text{minimize } z = 2x_1 + x_2$$

$$\text{s.t. } x_1 - x_2 \leq 2$$

$$x_2 \leq 5$$

$$x_1 + x_2 \geq 4$$

$$x \geq 0$$

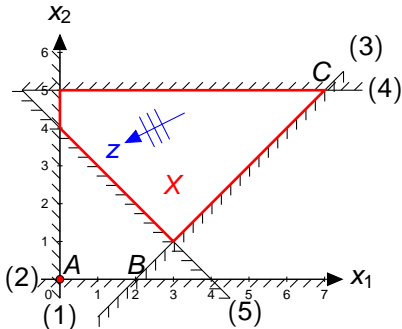


Figura: Base iniziale inammissibile: il vincolo $x_5 \geq 0$ è violato.

0	2	1	0	0	0
2	1	-1	1	0	0
5	0	1	0	1	0
-4	-1	-1	0	0	1

Solution A (infeasible)

$$B = \{3, 4, 5\}$$

$$x^T = [0 \ 0 \ 2 \ 5 \ -4]$$

$$z = 0$$

Il problema ausiliario

0	2	1	0	0	0
2	1	-1	1	0	0
5	0	1	0	1	0
-4	-1	-1	0	0	1

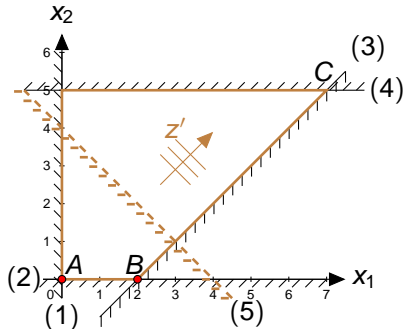
Il problema originale

-4	-1	-1	0	0	1
2	1	-1	1	0	0
5	0	1	0	1	0
0	2	1	0	0	0

Il problema ausiliario

N.B.: La forma canonica si conserva rispetto ai vincoli originali.

Il problema ausiliario



-4	-1	-1	0	0	1
2	1	-1	1	0	0
5	0	1	0	1	0
0	2	1	0	0	0

Soluzione A (inammissibile)

$$B = \{3, 4, 5\}$$

$$x^T = [0 \ 0 \ 2 \ 5 \ -4]$$

$$z = 0$$

-2	0	-2	1	0	1
2	1	-1	1	0	0
5	0	1	0	1	0
-4	0	3	-2	0	0

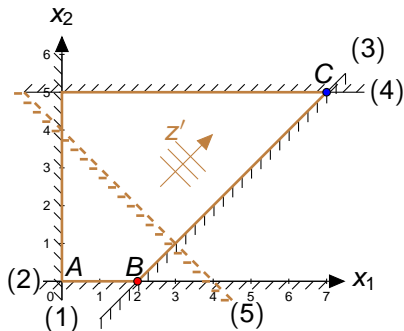
Soluzione B (inammissibile)

$$B = \{1, 4, 5\}$$

$$x^T = [2 \ 0 \ 0 \ 5 \ -2]$$

$$z = 4$$

Il problema ausiliario



-2	0	-2	1	0	1
2	1	-1	1	0	0
5	0	1	0	1	0
-4	0	3	-2	0	0

Soluzione B (inammissibile)

$$B = \{1, 4, 5\}$$

$$x^T = [2 \ 0 \ 0 \ 5 \ -2]$$

$$z = 4$$

8	0	0	1	2	1
7	1	0	1	1	0
5	0	1	0	1	0
-19	0	0	-2	-3	0

Soluzione C (ammissibile)

$$B = \{1, 2, 5\}$$

$$x^T = [7 \ 5 \ 0 \ 0 \ 8]$$

$$z = 19$$

Il problema originale

8	0	0	1	2	1
7	1	0	1	1	0
5	0	1	0	1	0
-19	0	0	-2	-3	0

Il problema ausiliario

-19	0	0	-2	-3	0
7	1	0	1	1	0
5	0	1	0	1	0
8	0	0	1	2	1

Il problema originale

Il valore della funzione obiettivo (trascurata finora) può essere peggiorato.

Test di inammissibilità

Se l'inammissibilità rispetto ad un vincolo violato è stata minimizzata ma il valore della corrispondente variabile resta negativo, questo dimostra che il problema è inammissibile e l'algoritmo termina.

Esempio: sostituendo $x_2 \leq 5$ con $x_2 \leq 1/2$ nell'ultimo esempio. Dopo due passi di pivot si ottiene il tableau seguente:

-1		0	0	1	2	1
5/2		1	0	1	1	0
1/2		0	1	0	1	0
-11/2		0	0	-2	-3	0

Il proble ausiliario è risolto all'ottimo, (tutti i costi ridotti sono non-negativi).

L'inammissibilità, quindi, non può essere ridotta ulteriormente: il problema è inammissibile.

$$0x_1 + 0x_2 + 1x_3 + 2x_4 + 1x_5 = -1$$

non ha soluzione per $x \geq 0$.

Problema ausiliario illimitato

Può capitare che il problema ausiliario sia illimitato (anche se il problema originale non lo è).

Esempio: cancellando il vincolo $x_2 \leq 5$ dall'esempio precedente.

0	2	1	0	0
2	1	-1	1	0
-4	-1	-1	0	1

Il problema originale

-4	-1	-1	0	1
2	1	-1	1	0
0	2	1	0	0

Il problema ausiliario

La seconda colonna è interamente composta da elementi negativi:
problema illimitato.

Problema ausiliario illimitato

In tal caso, il pivot da scegliere è l'elemento (negativo!) sulla riga del vincolo violato.

-4	-1	-1	0	1
2	1	-1	1	0
0	2	1	0	0

4	1	1	0	-1
6	2	0	1	-1
-4	1	0	0	1

La soluzione di base risultante è un punto sul vincolo violato.