Ricerca Operativa

Richiami di Algebra Vettoriale

Def. Prende il nome di *vettore ad n componenti reali* una n-pla ordinata di numeri reali.
Es. La coppia (-1, 4) è un esempio di vettore a 2 componenti: la prima è -1, la seconda è 4.

Def. Prende il nome di *vettore colonna* il vettore le cui componenti sono disposte lungo una linea verticale (colonna). Lo si indica con la seguente notazione:

$\underline{x} = \begin{pmatrix} 3 \\ -1 \\ 7 \end{pmatrix}$, un vettore colonna di dimensione n = 3.

Def. Prende il nome di *vettore riga* il vettore le cui componenti sono disposte lungo una linea orizzontale (riga). Lo si indica con la seguente notazione:
$\underline{x}^T = (3, -1, 7)$, un vettore riga di dimensione n = 3.

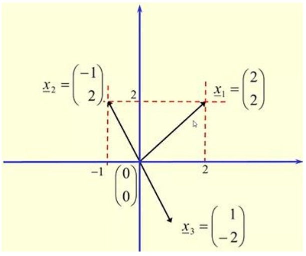
Def. Si chiama *trasposizione* l'operazione unaria che trasforma un vettore riga (colonna) in un vettore colonna (riga).

Es. $\underline{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ -4 \\ -6 \\ 7 \end{pmatrix} \Rightarrow \underline{x}^T = (1, 2, 3, -4, -6, 7)$.

Def. Prende il nome di *vettore nullo*, e lo si indica con $\underline{0}^T = (0, 0, \dots, 0)$, il vettore le cui componenti sono tutte nulle.

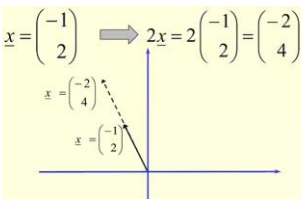
Def. Prende il nome di *scalare* un qualsiasi numero reale.

Es. (Vettori di dimensione 2) Ogni vettore può essere rappresentato tramite un punto o da una linea che connette l'origine al punto.
Nota: Noi useremo come nomi degli assi x_1 e x_2, anziché x ed y.

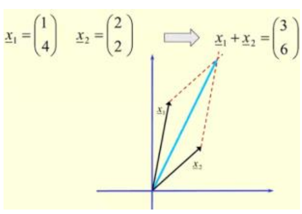


Ci sono tre caratteristiche associate ad un vettore:
- la **direzione**, che rappresenta la retta su cui poggia;
- il **modulo**, che rappresenta la lunghezza del vettore;
- il **verso**, che è indicato dalla sua freccia.

La **moltiplicazione** di uno scalare s per un vettore $\underline{x}$ ha come risultato un vettore di cui ogni componente è moltiplicata per s.
Da un punto di vista geometrico, la stessa operazione:
- se s > 1, allora viene incrementato il modulo del vettore;
- se 0 < s < 1, allora viene decrementato il modulo del vettore;
- se s < 0, allora vengono modificati modulo e/o verso del vettore.



L'**addizione** di due vettori $\underline{x}_1$ e $\underline{x}_2$ di dimensione n sarà un nuovo vettore, avente la stessa dimensione dei vettori di partenza, le cui componenti sono il risultato della somma delle componenti di $\underline{x}_1$ e $\underline{x}_2$. Graficamente, la *regola del parallelogramma* dice che, prendendo $\underline{x}_1$ e $\underline{x}_2$, si tracciano la retta parallela a $\underline{x}_1$ passante per x_2 e la retta parallela a $\underline{x}_2$ passante per $\underline{x}_1$; il punto in cui le due rette si incrociano rappresenta il modulo del nuovo vettore $\underline{x}_1 + \underline{x}_2$.



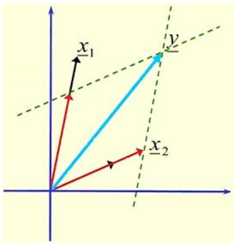
Il **prodotto interno** (o prodotto scalare) tra due vettori ha come risultato uno scalare e si ottiene facendo la somma dei prodotti delle componenti dei due vettori.
$\underline{x}^T \underline{y} = x_1y_1 + x_2y_2 + \ldots + x_ny_n$

Es. $\underline{x}^T = (0, 2)$, $\underline{y} = (3, 4) \Rightarrow \underline{x}^T \underline{y} = 8$

Il prodotto scalare è un'operazione fondamentale nella Ricerca Operativa.

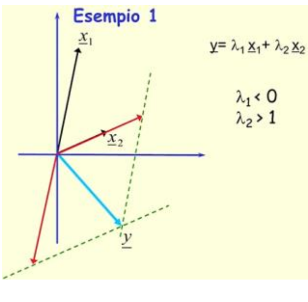
Esercizio. Supponiamo di andare in un centro commerciale per comprare delle pen drive. Le penne acquistabili sono da 4, 8, 16 e 32 GB, che costano rispettivamente 3€, 7€, 20€ e 29€. Dato un budget a disposizione di 100€, vogliamo acquistare un numero di pen drive tali che nel carrello siano presenti almeno tre penne da 16 GB, con l'obiettivo di ottimizzare i GB.
→ Siano x_1, x_2, x_3, x_4 variabili decisionali rappresentanti il numero di pen drive rispettivamente da 4, 8, 16 e 32 GB che si stanno acquistando.
Ponendo $x_1 = 4$, $x_2 = 0$, $x_3 = 0$, $x_4 = 3$; notiamo che tale soluzione è inammissibile per il problema, in quanto abbiamo il vincolo per cui x_3 dev'essere ≥ 3.
Ponendo $x_1 = 4$, $x_2 = 0$, $x_3 = 3$, $x_4 = 3$; notiamo che tale soluzione è inammissibile per il problema, in quanto è vero che $x_3 \geq 3$, ma viene sforato il budget; eseguendo il prodotto scalare, infatti, notiamo che il totale dell'acquisto sarebbe tot = $x_1 \cdot 3 + x_2 \cdot 7 + x_3 \cdot 20 + x_4 \cdot 29 = 4 \cdot 3 + 0 \cdot 7 + 3 \cdot 20 + 3 \cdot 29 = 159$€.
Ponendo $x_1 = 13$, $x_2 = 0$, $x_3 = 3$, $x_4 = 0$; notiamo che tale soluzione è ammissibile, in quanto il prodotto scalare dà come risultato tot = $x_1 \cdot 3 + x_2 \cdot 7 + x_3 \cdot 20 + x_4 \cdot 29 = 13 \cdot 3 + 0 \cdot 7 + 3 \cdot 20 + 0 \cdot 29 = 99$€; inoltre, la *funzione obiettivo* è pari a 100GB, eseguendo il prodotto scalare dei GB delle pen drive.
L'obiettivo dell'esercizio non era trovare la soluzione ottima del problema, ma è stato introdotto semplicemente per evidenziare quanto il prodotto scalare e l'ottimizzazione influenzi anche la quotidianità. □

Un vettore $\underline{y}$ è **combinazione lineare** dei vettori $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ se esistono $\lambda_1, \lambda_2, \ldots, \lambda_n$ scalari tali che: $\underline{y} = \lambda_1\underline{x}_1 + \lambda_2\underline{x}_2 + \ldots + \lambda_n\underline{x}_n$.
Ad esempio, il vettore $\underline{y}$ nell'esempio mostrato nella figura posta a lato, ponendo $\lambda_1 < 1$ e $\lambda_2 > 1$, possiamo modificare il modulo dei due vettori in modo tale da ottenere il parallelogramma e, di conseguenza, $\underline{y}$.

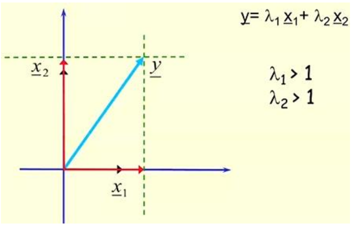


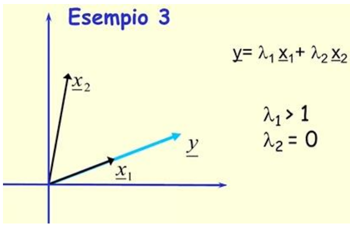
Es. Nella figura a lato, il vettore $\underline{y}$ è combinazione lineare di $\underline{x}_1$ e $\underline{x}_2$. Si procede come segue:
- si traccia la retta parallela ad $\underline{x}_2$ passante per $\underline{y}$;
- si traccia la retta parallela ad $\underline{x}_1$ passante per $\underline{y}$;
- si cercano i valori dei vari $\lambda_1, \lambda_2, \ldots, \lambda_n$.

In questo caso, bisogna scegliere opportuni $\lambda_1 < 0$ e $\lambda_2 > 1$, in quanto bisogna cambiare anche il verso di x_1 ed incrementare il modulo di x_2.

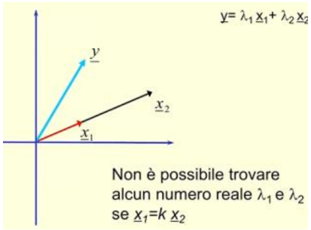


Es. Seguono ulteriori esempi sulla combinazione lineare.



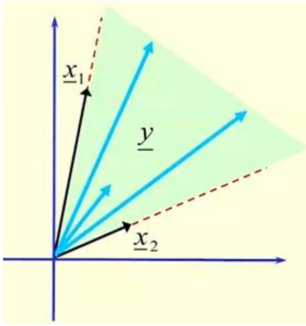


In quest'ultimo caso, non è possibile trovare alcun numero reale λ_1 e λ_2 in quanto $\underline{x}_1 = k\cdot\underline{x}_2$ (cioè, i due vettori sono sovrapposti).



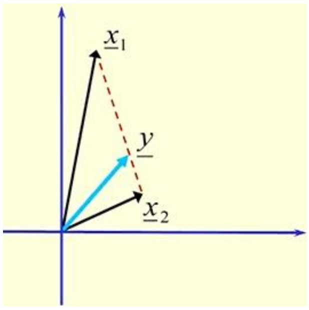
Un vettore $\underline{y}$ è **combinazione conica** dei vettori $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ se esistono $\lambda_1, \lambda_2, \dots, \lambda_n$ scalari tali che:
1. $\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$;
2. $\underline{y} = \lambda_1\underline{x}_1 + \lambda_2\underline{x}_2 + \dots + \lambda_n\underline{x}_n$.

Di conseguenza, è banale dire che una combinazione conica è anche una combinazione lineare; non vale, tuttavia, il viceversa.
Geometricamente, se impongo tutti i lambda maggiori o uguali a 0, posso rappresentare come combinazione conica i soli vettori all'interno del *cono* formato dai vettori $\underline{x}_1$ e $\underline{x}_2$; ciò avviene in quanto porre gli scalari ≥ 0 fa sì che non sia possibile modificare il verso di $\underline{x}_1$ e $\underline{x}_2$.



Un vettore $\underline{y}$ è **combinazione convessa** dei vettori $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ se esistono $\lambda_1, \lambda_2, \dots, \lambda_n$ scalari tali che:
1. $\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$;
2. $\lambda_1 + \lambda_2 + \dots + \lambda_n = 1$;
3. $\underline{y} = \lambda_1\underline{x}_1 + \lambda_2\underline{x}_2 + \dots + \lambda_n\underline{x}_n$.

Questo equivale a considerare tutti e soli i vettori che si trovano sul segmento che unisce x_1 ad x_2.
Di conseguenza, è banale dire che una combinazione convessa è anche una combinazione conica; non vale, tuttavia, il viceversa.



I vettori $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ si dicono **linearmente indipendenti** se:

$$\lambda_1\underline{x}_1 + \lambda_2\underline{x}_2 + \dots + \lambda_n\underline{x}_n = \underline{0} \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

I vettori $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ si dicono **linearmente dipendenti** se esistono $\lambda_1, \lambda_2, \dots, \lambda_n$ <u>non tutti nulli</u>, tali che:

$$\lambda_1\underline{x}_1 + \lambda_2\underline{x}_2 + \dots + \lambda_n\underline{x}_n = \underline{0}.$$

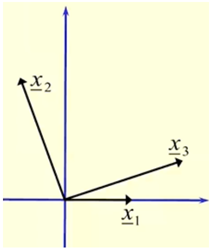
Es. I vettori $\underline{x}_1^T = (1, 2, 3)$, $\underline{x}_2^T = (-1, 1, -1)$, $\underline{x}_3^T = (0,3,2)$ sono linearmente dipendenti, perché l'equazione $\lambda_1\underline{x}_1 + \lambda_2\underline{x}_2 + \dots + \lambda_n\underline{x}_n = \underline{0}$ è verificata quando $\lambda_1 = \lambda_2 = 1$ e $\lambda_3 = -1$.
Infatti: $1 \cdot \begin{pmatrix}1\\2\\3\end{pmatrix} + 1 \cdot \begin{pmatrix}-1\\1\\-1\end{pmatrix} - 1 \cdot \begin{pmatrix}0\\3\\2\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$.
Se i tre vettori fossero stati linearmente indipendenti, allora l'unico modo per ottenere il vettore nullo sarebbe stato quello di porre tutti gli scalari lambda pari a 0.

I vettori $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ si dicono **linearmente dipendenti** se uno di essi può essere espresso come combinazione lineare degli altri. Nell'esempio precedente: $\underline{x}_1^T = (1, 2, 3)$, $\underline{x}_2^T = (-1, 1, -1)$, $\underline{x}_3^T = (0,3,2) \Rightarrow \underline{x}_1 + \underline{x}_2 = \underline{x}_3$.

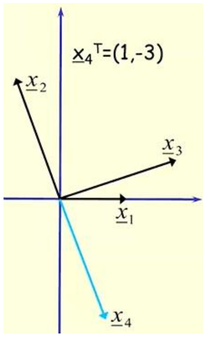
Un insieme di vettori $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$ di dimensione *n* **genera** l'insieme di vettori E^n se ogni vettore in E^n può essere rappresentato come combinazione lineare dei vettori $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$. Quindi, E^n è detto *spazio generato*.

Def. Un insieme $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ in E^n è una *base* di E^n se valgono le due seguenti condizioni:
1. $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ generano E^n;
2. se uno solo dei vettori è rimosso, allora i rimanenti k-1 vettori non generano E^n.
Alternativamente, è possibile utilizzare la seguente proprietà.
Proprietà. Un insieme di vettori $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ in E^n è una base di E^n se e solo se:
1. $k = n$;
2. $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ sono linearmente indipendenti.
Def. Il numero di vettori che formano una base per E^n è detto *dimensione* dello spazio E^n.

Es. Cerchiamo una base per lo spazio E^2 (dei vettori di dimensione 2).
Dobbiamo, quindi, cercare 2 vettori in E^2 linearmente indipendenti.
Siano $\underline{x}_1^T = (1, 0)$, $\underline{x}_2^T = (-1, 3)$, $\underline{x}_3^T = (2, 1)$.



1) Notiamo che essi generano $\mathbb{R}^2$, in quanto anche se essi sono 3 e non 2, possiamo porre (ad esempio) $\lambda_3 = 0$, e lavorare solo sui vettori $\underline{x}_1^T$ e $\underline{x}_2^T$; tramite combinazione lineare di questi due vettori, infatti, possiamo rappresentare <u>qualsiasi</u> vettore in $\mathbb{R}^2$.
2) Essi, tuttavia, non sono una base in $\mathbb{R}^2$ in quanto per esserlo deve avere esattamente 2 vettori linearmente indipendenti, ma in questo caso ne sono 3. Infatti, essi non sono linearmente indipendenti in quanto $\underline{x}_3^T$ può essere espresso come combinazione conica di $\underline{x}_1^T$ e $\underline{x}_2^T$.
3) $\underline{x}_1^T$ e $\underline{x}_2^T$ sono, presi singolarmente, una base in $\mathbb{R}^2$.
4) $\underline{x}_2^T$ e $\underline{x}_3^T$ sono, presi singolarmente, una base in $\mathbb{R}^2$.
5) Sia $\underline{x}_4^T = (1, -3)$. Allora $\underline{x}_2^T$ e $\underline{x}_4^T$ non sono una base in $\mathbb{R}^2$, dato che essi non sono linearmente indipendenti.



Esercizio. Dati i vettori $\underline{x}_1^T = (1, 3, 0)$, $\underline{x}_2^T = (2, 0, 1)$, $\underline{x}_3^T = (0, 1, 0)$ in $\mathbb{R}^3$:
1. Verificare che costituiscono una base;
2. Determinare le coordinate del vettore $\underline{y}^T = (2, 4, 1)$ in termini della base.
Nota: Per il punto due, una volta verificato il punto 1, occorre esprimere i valori di $\lambda_1, \lambda_2, \lambda_3$ che diano come risultato il vettore $\underline{y}$.

Matrici e Sistemi di Equazioni Lineari

Def. Prende il nome di *matrice* di ordine *mxn* una tabella di elementi ordinatamente disposti su *m* righe ed *n* colonne.

Indicheremo le matrici con lettere maiuscole A, B, … o per esteso con la seguente notazione:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} \qquad A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}.$$

La matrice A è una matrice (3x4). Indichiamo con a_{ij} un generico elemento della matrice nella riga *i* e colonna *j*.
Useremo anche le seguenti notazioni:
- una riga di una matrice può essere indicata con $\underline{a}^i$, ossia come un vettore riga, pertanto la precedente matrice A può essere indicata anche come insieme di vettori riga A = $\begin{pmatrix} \underline{a}^1 \\ \underline{a}^2 \\ \underline{a}^3 \end{pmatrix}$;
- la colonna di una matrice può essere indicata con $\underline{a}_j$, ossia come un vettore colonna, pertanto la precedente matrice A può essere indicata anche come insieme di vettori colonna A = $(\underline{a}_1, \underline{a}_2, \underline{a}_3, \underline{a}_4)$.

Se $m \neq n$, allora la matrice si dice *rettangolare*; si dice *quadrata* se $m = n$:
$A = \begin{pmatrix} 1 & 4 & 5 & 2 \\ 2 & 7 & 2 & 3 \\ 3 & 3 & 3 & 1 \end{pmatrix}$ è una matrice rettangolare, mentre $B = \begin{pmatrix} 1 & 4 & 5 \\ 2 & 7 & 2 \\ 3 & 3 & 3 \end{pmatrix}$ è una matrice quadrata.
In una matrice quadrata di ordine *n*, gli elementi a_{ii} (per i = 1, …, n) costituiscono la *diagonale principale*.

La **moltiplicazione per uno scalare** *k* di una matrice con *m* righe ed *n* colonne $A = \{a_{ij}\}$ dà come risultato la nuova matrice con *m* righe ed *n* colonne $k{\cdot}A = \{ka_{ij}\}$, ossia la matrice in cui ogni suo elemento è moltiplicato per tale scalare.

Es. Per k = 2 ed A = $\begin{pmatrix} 1 & 4 & 5 & 2 \\ 2 & 7 & 2 & 3 \\ 3 & 3 & 3 & 1 \end{pmatrix}$, abbiamo che k·A = $\begin{pmatrix} 2 & 8 & 10 & 4 \\ 4 & 14 & 4 & 6 \\ 6 & 6 & 6 & 2 \end{pmatrix}$.

L'**addizione** tra due matrici aventi *m* righe ed *n* colonne $A = \{a_{ij}\}$, $B = \{b_{ij}\}$, dà come risultato la nuova matrice con *m* righe ed *n* colonne $A + B = C$ avente elementi $c_{ij} = a_{ij} + b_{ij}$ (per ogni $i = 1, \dots, m$ e per ogni $j = 1, \dots, n$).
Le matrici devono avere le stesse dimensioni.
Es. Per A = $\begin{pmatrix} 1 & 4 \\ 2 & 1 \end{pmatrix}$, B = $\begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$ abbiamo che A + B = $\begin{pmatrix} 0 & 5 \\ 2 & 2 \end{pmatrix}$.

Sia $A = \{a_{ij}\}$ una matrice *mxn*, e sia B = $\{b_{ij}\}$ una matrice *nxp*. La **moltiplicazione** tra due matrici, dove il numero di colonne della prima è uguale al numero di righe della seconda, dà come risultato la matrice *mxp* di *m* righe e *p* colonne $AB = C$ avente elementi $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$ (per ogni $i = 1, \dots, m$ e per ogni $j = 1, \dots, p$). Si noti che tale operazione coincide con il prodotto scalare tra la riga i-esima e la colonna j-esima. Il numero di colonne della matrice *A* deve essere uguale al numero di righe della matrice *B*.
Es. Per A = $\begin{pmatrix} 1 & -1 & 1 \\ 4 & -2 & 5 \\ 2 & 0 & 1 \end{pmatrix}$, B = $\begin{pmatrix} 5 & 0 \\ 3 & 0 \\ 1 & 1 \end{pmatrix}$ abbiamo che A·B = $\begin{pmatrix} 3 & 1 \\ 19 & 5 \\ 11 & 1 \end{pmatrix}$.
Quindi, per la moltiplicazione tra due matrici A *mxn* e B *qxp* va ricordato che:
1. Il prodotto AB è definito solo se $n = q$. AB è allora una matrice *mxp*;
2. Il prodotto BA è definito solo se $m = p$. BA è allora una matrice *qxn*;
3. NON necessariamente vale la proprietà commutativa.

Alcune matrici particolari sono le seguenti. La matrice quadrata, di *n* righe ed *n* colonne, $I = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$ è detta *matrice identità*. In particolare, il prodotto tra una matrice *A* con *m* righe ed *n* colonne e la matrice identità dà come risultato A, cioè $A \cdot I = A$. Inoltre, vale che il prodotto tra la matrice identità ed una matrice A con n righe ed m colonne dà come risultato ancora A, cioè $I \cdot A = A$.

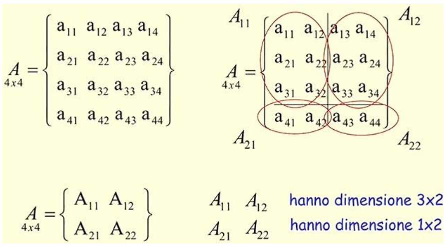
La matrice quadrata, di n righe ed n colonne, $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ & & \ddots & & \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}$ è detta *matrice triangolare superiore*, ed è la matrice avente tutti gli elementi al di sotto della diagonale principale pari a 0; viceversa, la *matrice triangolare inferiore* è la matrice avente tutti gli elementi al di sopra della diagonale principale pari a 0.

Data una matrice A = $\{a_{ij}\}$ di m righe ed n colonne, la sua matrice **trasposta** A^T è una matrice di n righe ed m colonne ottenuta invertendo le righe con le colonne.

Es. $A = \begin{pmatrix} 5 & 0 \\ 3 & 0 \\ 1 & 1 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 5 & 3 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Si osservano le seguenti proprietà:
1. $(A^T)^T = A$;
2. $(A + B)^T = A^T + B^T$ (quando la somma è definita);
3. $(AB)^T = B^T A^T$ (quando il prodotto è definito).

Una matrice A, di m righe ed n colonne, possiamo anche vederla **partizionata** in sottomatrici, come suggerisce la figura a lato.



Data una matrice A, di m righe ed n colonne, è possibile definire alcune **operazioni elementari** sulle righe e sulle colonne utili a risolvere un sistema di equazioni lineari:
- con lo <u>scambio</u> si effettua uno scambio di una riga i con una riga j;
- con la <u>moltiplicazione</u> si effettua una moltiplicazione di una riga per uno scalare (diverso da 0);
- con la <u>sostituzione</u> si effettua una sostituzione di una riga i con la somma della riga i e della riga j moltiplicata per uno scalare.

Sia A una matrice quadrata di n righe ed n colonne; allora se esiste una matrice quadrata B di n righe ed n colonne tale che $AB = I$ e $BA = I$, allora B è detta **matrice inversa** di A.
Occorre ricordare che:
- l'inversa di una matrice A (se esiste) è unica, ed è indicata con A^{-1};
- se una matrice ammette l'inversa, allora essa è detta matrice *non singolare*;
- una matrice quadrata è non singolare se e solo se le righe sono linearmente indipendenti o, equivalentemente, se e solo se le colonne sono linearmente indipendenti.

L'inversa di una matrice quadrata A può essere calcolata, utilizzando il **metodo con operazioni elementari**, attraverso un numero finito di operazioni elementari nel seguente modo:
1. Si considera la nuova matrice (A, I), ottenuta affiancando alla matrice A la matrice identità;
2. Si effettuano una serie di operazioni elementari sulle righe e sulle colonne di questa nuova matrice, in modo tale che:
   - A diventi la matrice identità I;
   - I diventi la matrice inversa A^{-1}.

Es. Data la matrice $A = \begin{pmatrix} 2 & 1 & 1 \\ -1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix}$, voglio calcolare la sua matrice inversa A^{-1}.

1. Affianchiamo ad A la matrice identità: $(A, I) = \begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 & 1 & 0 \\ 1 & -1 & 2 & 0 & 0 & 1 \end{pmatrix}$.
2. Effettuiamo le operazioni elementari. Si procede, essenzialmente, per colonne: la prima colonna, ora, deve quindi diventare la prima colonna della matrice identità; in sostanza, bisogna prima ottenere l'1, poi successivamente si lavora per ottenere gli 0. Una volta ottenuta la prima colonna, si procede con la seconda:

nella seconda riga della seconda colonna deve, quindi, comparire 1; successivamente, si cerca di ottenere 0 nei restanti due campi. Infine, si lavora sulla terza colonna per far comparire 1 al suo ultimo campo, e gli 0 nei restanti. Dunque:

$$(A, I) = \begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 & 1 & 0 \\ 1 & -1 & 2 & 0 & 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 3 & 2 & 1 & 1 & 0 \\ -1 & 2 & 1 & 0 & 1 & 0 \\ 1 & -1 & 2 & 0 & 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 3 & 2 & 1 & 1 & 0 \\ 0 & 1 & 3 & 0 & 1 & 1 \\ 1 & -1 & 2 & 0 & 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 3 & 2 & 1 & 1 & 0 \\ 0 & 1 & 3 & 0 & 1 & 1 \\ 0 & -4 & 0 & -1 & -1 & 1 \end{pmatrix} \Rightarrow$$

$$\Rightarrow \begin{pmatrix} 1 & 0 & -7 & 1 & -2 & -3 \\ 0 & 1 & 3 & 0 & 1 & 1 \\ 0 & -4 & 0 & -1 & -1 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & -7 & 1 & -2 & -3 \\ 0 & 1 & 3 & 0 & 1 & 1 \\ 0 & 0 & -12 & -1 & -5 & -3 \end{pmatrix} \Rightarrow \dots \Rightarrow \begin{pmatrix} 1 & 0 & 0 & \frac{5}{12} & -\frac{3}{12} & -\frac{1}{12} \\ 0 & 1 & 0 & \frac{3}{12} & \frac{3}{12} & -\frac{3}{12} \\ 0 & 0 & 1 & -\frac{1}{12} & \frac{3}{12} & \frac{5}{12} \end{pmatrix}$$

. Quindi, la matrice inversa in questo caso esiste. Questo è il primo metodo del calcolo della matrice inversa.

Commentato [DE1]: sostituendo la prima riga con la somma tra la seconda riga e la prima riga

Commentato [DE2]: sostituendo la seconda riga con la somma tra la seconda riga e la terza riga

Commentato [DE3]: sostituendo la terza riga con la differenza tra la terza riga e la prima riga

Commentato [DE4]: sostituendo la prima riga con la differenza tra la prima riga e la seconda riga moltiplicata per 3

Commentato [DE5]: sostituendo la terza riga con la differenza tra la terza riga e la seconda riga moltiplicata per 4

Il seguente è un metodo alternativo per il calcolo dell'inversa.
Il **determinante** di una matrice quadrata A è uno scalare che ne sintetizza alcune proprietà algebriche. Viene denotato con det(A) e si calcola, fissata una riga *i*, con la seguente formula:

$$\boxed{det(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot minor(A_{ij})}$$

dove *minor*(A_{ij}) è il determinante della sottomatrice di A ottenuta cancellando la *i*-esima riga e la *j*-esima colonna di A. Si ricordi che il determinante di una matrice A (2x2) si calcola effettuando la differenza tra il prodotto dei termini sulla diagonale principale e il prodotto dei termini sulla diagonale secondaria, cioè ($a_{11} \cdot a_{22}$) – ($a_{12} \cdot a_{21}$).

Es. $A = \begin{pmatrix} 2 & 1 & 1 \\ -1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix} \Rightarrow$

Commentato [DE6]: scegliendo di lavorare sulla prima riga

$$\det(A) = (-1)^{1+1} \cdot a_{11} \cdot minor(A_{11}) + (-1)^{1+2} \cdot a_{12} \cdot minor(A_{12}) + (-1)^{1+3} \cdot a_{13} \cdot minor(A_{13}) =$$
$$= (-1)^2 \cdot 2 \cdot \det\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} + (-1)^3 \cdot 1 \cdot \det\begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix} + (-1)^4 \cdot 1 \cdot \det\begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} = 2 \cdot 5 + (-1) \cdot (-3) + 1 \cdot (-1) =$$
$$= 12.$$

Il calcolo del determinante è importante, in quanto:
- se il determinante è diverso da 0, allora la matrice è invertibile ed esiste la matrice inversa. Ciò significa che le colonne e le righe di tale matrice sono linearmente indipendenti;
- se il determinante è uguale a 0, allora la matrice non è invertibile.

Una volta calcolato il determinante, si vuole calcolare la matrice inversa A^{-1}. La formula è la seguente, dove il *cofattore* $\boxed{cof(A_{ij}) = (-1)^{i+j} \cdot minor(A_{ij})}$:

$$\boxed{A^{-1} = \frac{1}{\det(A)} \cdot \begin{pmatrix} cof(A_{11}) & cof(A_{21}) & \cdots & cof(A_{n1}) \\ \vdots & & \ddots & \vdots \\ cof(A_{1n}) & cof(A_{2n}) & \cdots & cof(A_{nn}) \end{pmatrix}}$$

cioè, il rapporto tra la *matrice trasposta dei cofattori* ed il determinante di A.
Es. Considerata la matrice per l'esempio precedente, dove det(A) = 12, otteniamo la matrice A^{-1} come segue:
1. $cof(A_{ij}) = (-1)^{i+j} \cdot minor(A_{ij}) \Rightarrow cof(A_{11}) = (-1)^{1+1} \cdot \det\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} = 5$,
pertanto l'elemento A^{-1}_{11} della matrice inversa (rapporto tra cofattore e determinante) sarà 5/12;
2. $cof(A_{ij}) = (-1)^{i+j} \cdot minor(A_{ij}) \Rightarrow cof(A_{12}) = (-1)^{1+2} \cdot \det\begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix} = 3$,
pertanto l'elemento A^{-1}_{12} della matrice inversa (rapporto tra cofattore e determinante) sarà 3/12;
3. … dopo aver calcolato ogni cofattore della trasposta, e rapportato ogni elemento di tale matrice al determinante, otteniamo la matrice inversa A^{-1}.

Il **rango di riga** è il numero massimo di righe linearmente indipendenti. Analogamente, il **rango di colonna** è il numero massimo di colonne linearmente indipendenti.
Teorema. *Data una qualsiasi matrice, il rango per righe e per colonne coincide.*
Il precedente teorema implica che, data una matrice A di *m* righe ed *n* colonne, Rango(A) ≤ min(m, n). Di conseguenza, se rango(A) = min(m, n), allora A è detta matrice *a rango pieno*.

La seguente regola consente di calcolare in modo rapido l'inversa di una matrice *A* 2x2:

$$\boxed{A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightarrow A^{-1} = \frac{1}{\det(A)} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}.$$

Data una matrice A con *m* righe ed *n* colonne, cercare una soluzione ad un **sistema di equazioni lineari** $A\underline{x} = \underline{b}$ significa cercare quei valori $x_1, x_2, \ldots, x_n$ tali che il vettore $\underline{b}$ può essere espresso come combinazione lineare delle colonne della matrice.
Si noti che quando si moltiplica la matrice A *mxn* per il vettore $\underline{x}$ (che può essere considerato una matrice *nx1*) si può osservare che ognuna delle variabili $x_1, x_2, \ldots, x_n$ è associata ad una colonna della matrice A.

Es. $A = \begin{pmatrix} 2 & 1 & 1 \\ -1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 + x_3 \\ -x_1 + 2x_2 + x_3 \\ x_1 - x_2 + 2x_3 \end{pmatrix}$. Si osservi che la variabile x_i è associata alla *i*-esima colonna della matrice A. Quindi, il risultato $A \cdot \underline{x}$ può essere scritto anche come segue (dove, si ricorda, che $\underline{a}_i$ è la colonna *i*-esima della matrice A): $\underline{a}_1 x_1 + \underline{a}_2 x_2 + \underline{a}_3 x_3$; quando si risolve il sistema di equazioni, si sta scrivendo che la quantità precedente deve essere uguale al vettore $\underline{b}$: alla fine, quando si chiede di risolvere il sistema di equazioni si chiede esattamente di trovare i coefficienti per poter ottenere il vettore $\underline{b}$ in funzione di $\underline{a}_1$, $\underline{a}_2$ e $\underline{a}_3$ (cioè esprimere $\underline{b}$ come combinazione lineare delle colonne della matrice A, se non ci sono vincoli sulle x).

Il **metodo di Gauss-Jordan** afferma che per la risoluzione di un sistema di equazioni lineari valgono le seguenti:
1. Rango(A, $\underline{b}$) > Rango(A) ⇒ il sistema non ha soluzione, in quanto con l'aggiunta di $\underline{b}$ si aggiunge un vettore linearmente indipendente rispetto agli altri, cioè non rappresentabile come loro combinazione lineare;
2. Rango(A, $\underline{b}$) = Rango(A) ⇒ il sistema ha soluzione.

se m < n, esistono infinite soluzioni

Inoltre, stabilito che il sistema ammette soluzione, se il numero di righe *m* è più piccolo del numero di colonne *n*, allora esistono *infinite* soluzioni. Ad esempio, se si ha una matrice 3x10 e si suppone che rango(A) = 3 (cioè, la matrice è a rango pieno), allora per poter ottenere una soluzione al sistema si è costretti a fissare arbitrariamente i valori delle variabili x (bisogna fissarne 7), e una volta fatto ciò rimane la matrice quadrata 3x3 a rango pieno che ammette una sola soluzione, calcolata in funzione dei valori assegnati alle 7 variabili.
In generale, il numero di soluzioni di un sistema di equazioni è pari a $\boxed{\infty^{n-rango(A)}}$, dove *n* è il numero di variabili del sistema. Ad esempio, se si studia una matrice quadrata A *nxn* a rango pieno (cioè, rango(A) = n), avremo che il numero di soluzioni sarà 1: esiste una ed una sola soluzione.

Problemi di Programmazione Matematica

Data una funzione $f: \mathbb{R}^n \rightarrow \mathbb{R}$ e $X \subseteq \mathbb{R}^n$, un **Problema di Ottimizzazione** (PO) può essere formulato come:

$$\begin{aligned} & \min f(\underline{x}) \quad \text{(o max)} \\ & s.t. \\ & \underline{x} \in X, \end{aligned}$$

cioè si vuole *minimizzare la funzione f soggetta al fatto che le x debbano essere scelte all'interno dell'insieme X.* In particolare:

- la funzione min $f(\underline{x})$ è detta **funzione obiettivo**, ed indica l'obiettivo che si vuole raggiungere nel PO. Essa valuta la qualità della soluzione costruita per paragonare le varie soluzioni, per valutare qual è la migliore;
- s.t. significa "*subject to*", cioè "soggetto a", ed indica ciò a cui è soggetta la funzione;
- l'insieme X è detta **regione ammissibile** (l'insieme delle soluzioni ammissibili), cioè tutte le soluzioni che possono essere costruite rispettando i vincoli imposti. La variabile $\underline{x}$ è il vettore delle **variabili decisionali**, cioè le variabili del problema.

Quando l'insieme delle soluzioni ammissibili di un problema di ottimizzazione viene espresso attraverso un sistema di equazioni e disequazioni, tale problema prende il nome di **problema di Programmazione Matematica** (PM). I problemi di PM sono un sottoinsieme dei PO: la differenza tra PM e PO risiede nel modo in cui viene definito l'insieme delle soluzioni ammissibili. Infatti, il problema di PM può essere espresso come:

$$\begin{aligned} & \min f(\underline{x}) \\ & s.t. \\ & g_i(\underline{x}) \geq b_i \ (i = 1, \ldots, m), \end{aligned}$$

dove tutti i punti dello spazio considerato che soddisfano *contemporaneamente* tutte queste condizioni sono i punti della regione ammissibile; in questo caso, l'indice *i* è un indice che va da 1 a m, dove *m* è il numero di vincoli che si possono avere nel sistema, e rappresenta l'i-esima componente del vettore dei termini noti.

Un problema di PM è **lineare** quando:

- la funzione obiettivo è lineare: $f(\underline{x}) = \underline{c}^T\underline{x}$, cioè il prodotto scalare tra il vettore $\underline{c}$ (di **numeri reali**) e $\underline{x}$ (il vettore delle **variabili decisionali**);
- l'insieme X è espresso in termini di relazioni (uguaglianze e disuguaglianze) lineari.

La funzione $f(\underline{x})$ è lineare quando: 1) $f(\underline{x}+\underline{y}) = f(\underline{x}) + f(\underline{y})$; 2) $f(\lambda\underline{x}) = \lambda f(\underline{x})$. Noi lavoreremo solo su problemi di programmazione lineare, cioè la funzione obiettivo ed i vincoli del problema sono prodotti scalari (di grado 1).

Scrivendo la regione ammissibile del PM ($g_i(\underline{x}) \geq b_i$, i = 1, …, m) in *forma esplicita*, otteniamo il risultato di cui a destra: abbiamo la funzione obiettivo come il prodotto scalare tra $\underline{c}$ ed $\underline{x}$, mentre il sistema di vincoli è sempre (presa in considerazione la prima disuguaglianza) il prodotto scalare tra il vettore $\underline{a}_{11}, \underline{a}_{12}, \ldots, \underline{a}_{1n}$ e il vettore $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ in disuguaglianza con $\underline{b}_1$ (quest'ultimo è il vettore dei termini noti, cioè degli scalari spesso legati alle soluzioni del problema).

Forma esplicita

$$\begin{array}{llllll} \min & c_1x_1 & + c_2x_2 & +\ldots+ & c_nx_n & \\ \text{s.t.} & a_{11}x_1 & + a_{12}x_2 & +\ldots+ & a_{1n}x_n & \geq b_1 \\ & a_{21}x_1 & + a_{22}x_2 & +\ldots+ & a_{2n}x_n & \geq b_2 \\ & a_{m1}x_1 & + a_{m2}x_2 & +\ldots+ & a_{mn}x_n & \geq b_m \end{array}$$

La *forma compatta* della regione ammissibile è, invece, il prodotto scalare tra il vettore $\underline{c}^T$ ed il vettore $\underline{x}$: $A\underline{x} \geq \underline{b}$ include A (la matrice dei **coefficienti tecnologici**), e $\underline{c}$ (il vettore dei **coefficienti di costo**).

Forma compatta

$$\begin{array}{ll} \min & \underline{c}^T\underline{x} \\ \text{s.t.} & A\underline{x} \geq \underline{b} \end{array}$$

I nostri problemi di programmazione lineare si distinguono in due categorie:

1. problemi di *Programmazione Lineare Continua* (PL), dove le variabili $\underline{x}$ sono reali;
2. problemi di *Programmazione Lineare Intera* (PLI), dove le variabili $\underline{x}$ sono intere.

$$X \begin{cases} \{\underline{x} \in \mathbf{R}^n : A\underline{x} \geq \underline{b}\} \\ \{\underline{x} \in \mathbf{Z}^n : A\underline{x} \geq \underline{b}\} \end{cases}$$

Dati i problemi di PL e PLI, la difficoltà di risoluzione è maggiore per i problemi di PLI. Infatti, partendo dal presupposto che, anche se i problemi di programmazione lineare continua hanno un numero infinito di soluzioni, la matematica consente di individuare comunque il sottoinsieme delle soluzioni ottimali: questo in quanto i problemi di PL fanno parte dei problemi "P" (facili); i problemi di PLI, invece, fanno tendenzialmente parte dei "NP" (non facili), dato che bisogna fornire un algoritmo in base a delle euristiche, che sono sì buoni ma non garantiscono di individuare con precisione la soluzione ottima.

Es. Nella figura di cui a destra è posto un modello matematico di un problema di PL in forma esplicita e compatta. Si osservi che:

Forma esplicita

$$\begin{aligned} \min \quad & 500x_1 + 700x_2 + 350x_3 + 400x_4 + 200x_5 \\ s.t. \quad & \\ & 8x_1 + 10x_2 + 5x_4 + 7x_5 = 96 \\ & 5x_1 + 12x_2 + 4x_3 = 96 \\ & 20x_1 + 20x_2 + 20x_3 + 20x_4 + 20x_5 = 384 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0 \end{aligned}$$

Forma compatta

$$\begin{aligned} \min \quad & \underline{c}^T \underline{x} \\ \text{s.t.} \quad & A\underline{x} = \underline{b} \\ & \underline{x} \geq 0 \end{aligned}$$

$$\underline{c}^T = [500 \quad 700 \quad 350 \quad 400 \quad 200] \quad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

$$A = \begin{bmatrix} 8 & 10 & 0 & 5 & 7 \\ 5 & 12 & 4 & 0 & 0 \\ 20 & 20 & 20 & 20 & 20 \end{bmatrix} \quad \underline{b} = \begin{bmatrix} 96 \\ 96 \\ 384 \end{bmatrix}$$

- il vettore $\underline{c}^T$ dei coefficienti di costo è uguale a $\underline{c}^T = [500 \;\; 700 \;\; 350 \;\; 400 \;\; 200]$;
- il vettore $\underline{x}$ delle variabili decisionali è uguale a $\underline{x}^T = [x_1 \;\; x_2 \;\; x_3 \;\; x_4 \;\; x_5]$, quindi il problema ha 5 variabili;
- la matrice A dei coefficienti tecnologici è $A = \begin{pmatrix} 8 & 10 & 0 & 5 & 7 \\ 5 & 12 & 4 & 0 & 0 \\ 20 & 20 & 20 & 20 & 20 \end{pmatrix}$, dove alla colonna i-esima sono associati i coefficienti della variabile x_i;
- il vettore $\underline{b}$ dei termini noti è uguale a $\underline{b}^T = [96 \;\; 96 \;\; 384]$.

Effettuando il prodotto scalare, si ottiene esattamente il problema in forma esplicita.

Un modo alternativo per vedere lo stesso sistema di equazioni è il seguente: per colonne si esprimono le variabili del problema per cui moltiplicare i vettori colonna della matrice A; occorre esprimere, cioè, il vettore $\underline{b}$ come combinazione conica delle colonne di A.

Combinazione lineare delle colonne di A

$$\begin{bmatrix} 8 \\ 5 \\ 20 \end{bmatrix} x_1 + \begin{bmatrix} 10 \\ 12 \\ 20 \end{bmatrix} x_2 + \begin{bmatrix} 0 \\ 4 \\ 20 \end{bmatrix} x_3 + \begin{bmatrix} 5 \\ 0 \\ 20 \end{bmatrix} x_4 + \begin{bmatrix} 7 \\ 0 \\ 20 \end{bmatrix} x_5 = \begin{bmatrix} 96 \\ 96 \\ 384 \end{bmatrix}$$

Non basta, tuttavia, risolvere un sistema di equazioni. Si vuole, infatti, trovare la soluzione che minimizzi la funzione obiettivo.

Dato un problema di programmazione lineare, un vettore $\underline{x}'$ di $\mathbb{R}^n$:
- **soddisfa** il vincolo $g_i(\underline{x}) \geq b_i$ se $g_i(\underline{x}') \geq b_i$, cioè se andando a sostituire i valori delle componenti di $\underline{x}'$ il vincolo è soddisfatto. Nell'esempio precedente, presa l'equazione $8x_1 + 10x_2 + 5x_4 + 7x_5 = 96$, il vettore $\underline{x}' = [12 \;\; 0 \;\; 0 \;\; 0 \;\; 0]$ soddisfa il primo vincolo, mentre non soddisfa il secondo;
- **viola** il vincolo $g_i(\underline{x}) \geq b_i$ se $g_i(\underline{x}') < b_i$;
- **satura** (o rende attivo) il vincolo $g_i(\underline{x}) \geq b_i$ se $g_i(\underline{x}') = b_i$.

Un vettore $\underline{x}$ di $\mathbb{R}^n$ si dice **soluzione ammissibile** per il problema di PL se e solo se soddisfa tutti i vincoli del problema.

Un problema di programmazione lineare risulta:
1. **Inammissibile** se la regione ammissibile è vuota, ossia $X = \emptyset$. Preso un insieme di vincoli dato dall'azienda, si scopre che non esiste alcuna soluzione ammissibile al problema in tali condizioni;
2. **Illimitato** (inferiormente) se, scelto un qualsiasi scalare k, esiste sempre un punto $\underline{x} \in X$ tale che $f(\underline{x}) < k$. Il valore della soluzione ottima del problema è $-\infty$ se sto minimizzando e $+\infty$ se sto massimizzando; inoltre, se l'ottimo del problema è illimitato, non esiste un punto di ottimo;
3. **Ammettere soluzione ottima finita** se esiste un punto $\underline{x}^* \in X$ tale che $f(\underline{x}^*) \leq f(\underline{x})$ per ogni $\underline{x} \in X$. La funzione f, nel punto $\underline{x}^*$, è minore o uguale alla stessa funzione in un qualsiasi altro punto della regione.

Es (Piano di produzione aziendale). Un'azienda produce tre tipi di elettrodomestici: lavatrici, frigoriferi e forni. Per produrre una lavatrice occorrono 9 ore di lavorazione sulla macchina M1 e 8 ore di lavorazione sulla macchina M2; mentre per produrre un frigorifero occorrono 11 ore di lavorazione sulla macchina M2; infine per produrre un forno occorrono 4 ore sulla macchina M1 e 6 sulla macchina M2.
La macchina M1 è disponibile per 137 ore settimanali, mentre la macchina M2 è disponibile per 149 ore settimanali. Il numero di forni prodotti non può essere superiore alla somma dei frigoriferi e delle lavatrici prodotte. Tuttavia devono essere prodotti almeno 20 forni. Inoltre il numero di lavatrici prodotte non può essere superiore al numero di frigoriferi prodotti per al più 5 unità.
Il guadagno ottenuto dalla vendita di una lavatrice è di 375 euro, quello ottenuto per un frigorifero è 320 euro e quello per un forno è 170 euro. Si vuole conoscere la quantità di lavatrici, frigoriferi e forni da produrre settimanalmente per massimizzare il guadagno totale nel rispetto dei vincoli di produzione.
a) Si formuli il corrispondente modello di programmazione.

a) La prima cosa da fare per poter formulare un problema è individuare le variabili decisionali. Poiché il nostro obiettivo è quello di definire il numero di lavatrici, frigoriferi e forni da produrre, associamo ad ogni tipo di elettrodomestico una variabile distinta. Una volta definite le variabili, bisogna definire la funzione obiettivo e i vincoli.
Siano x_1 il numero di lavatrici da produrre settimanalmente, x_2 il numero di frigoriferi ed x_3 il numero di forni. Quindi, $\underline{x} = [x_1 \ \ x_2 \ \ x_3]$.
La *funzione obiettivo* è massimizzare il guadagno totale nel rispetto dei vincoli di produzione, cioè:
max $z = 375x_1 + 320x_2 + 170x_3$, cioè il prodotto scalare tra i *coefficienti di costo* (in questo caso tale vettore è $\underline{c}^T = [375 \ \ 320 \ \ 170]$) e le tre variabili del problema.

Ora, definiamo i *vincoli*: $\begin{cases} 9x_1 + \qquad\quad 4x_3 \leq 137 \ (M_1) \\ 8x_1 + 11x_2 + 6x_3 \leq 149 \ (M_2) \\ x_1 + x_2 \leq x_3 \\ x_3 \geq 20 \\ x_1 \leq x_2 + 5 \\ \underline{x} \geq \underline{0}, intere \ (PLI) \end{cases}$. Notiamo che le variabili sono *intere*, di conseguenza il problema è un problema di PLI, quindi difficile da risolvere.
Nota: Scritto questo problema in Excel (che possiede i risolutori), questo ci restituisce la soluzione ottima. □

Una volta scritto il modello, è buona norma spostare tutte le variabili sulla parte sinistra del vincolo, in quanto sulla parte destra devono esserci solo i termini noti. Ad esempio, il vincolo $x_1 + x_2 \leq x_3$ dovrebbe essere in realtà $x_1 + x_2 - x_3 \leq 0$.
Se si volesse risolvere un dato problema con l'algoritmo del Simplesso (applicabile solo a problemi di PL), noteremmo che esso richiede una forma del problema in input particolare: la forma canonica.

Consideriamo un problema di Programmazione Lineare (PL) con *m* vincoli ed *n* variabili in **forma canonica di minimo**:

$$\begin{aligned} \min z = \underline{c}^T\underline{x} \\ A\underline{x} \geq \underline{b} \\ \underline{x} \geq \underline{0} \\ \underline{x} \in \mathbb{R}^n. \end{aligned}$$

Si ha, quindi, un problema di minimo con *z* valore che la funzione obiettivo assume nel punto $\underline{x}$; inoltre, tutti i vincoli sono di maggiore o uguale; infine, abbiamo $\underline{x} \geq \underline{0}$.
- $\underline{x}$ è il vettore nx1 delle **variabili decisionali**, dove il numero di tali variabili (*n*) dipende dal problema;
- $\underline{c}$ è il vettore nx1 dei **coefficienti di costo** della funzione obiettivo, cioè i coefficienti delle variabili $\underline{x}$ all'interno della funzione obiettivo;
- $\underline{b}$ è il vettore mx1 dei **termini noti** dei vincoli, dove il numero di tali vincoli (*m*) dipende dal problema;
- A è la matrice mxn dei coefficienti dei vincoli ($A = [a_{ij}]$, i = 1, …, n; j = 1, …, m).
Il modello scritto nell'esempio precedente non è in forma canonica: infatti, i vincoli non sono tutti espressi in ≥.
La forma canonica viene utilizzata per la Teoria della dualità.

Consideriamo un problema di Programmazione Lineare (PL) con *m* vincoli ed *n* variabili in **forma standard di minimo**:

$$\begin{aligned} \min z = \underline{c}^T\underline{x} \\ A\underline{x} = \underline{b} \ (1) \\ \underline{x} \geq \underline{0} \quad (2) \\ \underline{x} \in \mathbb{R}^n. \end{aligned}$$

Condizione: $\underline{b} \geq 0$

Questa forma è l'input preteso dall'algoritmo del Simplesso per trovare la soluzione ottima. Il problema deve essere di minimo, tutti i vincoli devono essere di uguaglianza e le variabili decisionali devono essere tutte non-negative; infine, il vettore $\underline{b}$ deve essere sempre maggiore o uguale a 0.
- I valori di $\underline{x}$ che soddisfano i vincoli (1) sono detti **soluzioni** del problema del PL;
- i valori di $\underline{x}$ che soddisfano anche i vincoli (2) sono detti **soluzioni ammissibili** del problema di PL.

Si assumono soddisfatte le seguenti ipotesi:

- $m < n$
- $m = \text{rango}(A)$.

L'ipotesi $m < n$ (più variabili che vincoli) non rappresenta una perdita di generalità. Infatti, è noto che il sistema di equazioni lineari (1) (avente una matrice quadrata a rango pieno):
- può ammettere una soluzione unica se $m = n$;
- può ammettere ∞^{n-m} soluzioni se $m < n$. => se ci sono più variabili che vincoli, il problema ammette infinite soluzioni

Solo il secondo caso è significativo dal punto di vista dei problemi di ottimizzazione.

È possibile effettuare delle *trasformazioni* per rendere un modello matematico un modello equivalente compatibile con uno tra Algoritmo del Simplesso e Teoria della dualità.
Def. Due problemi di programmazione lineare di minimo (massimo) (P) e (P') sono *equivalenti* se, per ogni soluzione ammissibile di (P), possiamo costruire una soluzione ammissibile di (P') con lo stesso valore, e viceversa.
Oss. Se due problemi di programmazione lineare sono equivalenti, allora i valori delle rispettive soluzioni ottime coincidono.
Oss. Qualunque problema di PL può essere trasformato in un problema equivalente in forma canonica o standard.

Per trasformare una *funzione obiettivo* max in min: $\boxed{max\ \underline{c}^T \underline{x} \Leftrightarrow -min\ -\underline{c}^T \underline{x}}$.

Es. max $3x_1 + 5x_2 \Leftrightarrow$ - min $-3x_1 -5x_2$. Va posto anche il min in negativo perché, ponendo $x_1 = x_2 = 1$, senza di esso otterremmo 8 e -8.

Per trasformare i *vincoli di uguaglianza* e di *maggiore o uguale*: $\boxed{\begin{array}{c} A\underline{x} \geq \underline{b} \Leftrightarrow -A\underline{x} \leq -\underline{b} \\ A\underline{x} = \underline{b} \Leftrightarrow \begin{cases} A\underline{x} \leq \underline{b} \\ A\underline{x} \geq \underline{b} \end{cases} \end{array}}$.

Per trasformare i *vincoli di minore o uguale* in vincoli di uguaglianza:
$\boxed{\sum_{j=1}^{n} a_{ij}x_j \leq b_i \Leftrightarrow \sum_{j=1}^{n} a_{ij}x_j + x_{n+1} = b_i}$, con $x_{n+1} \geq 0$.
La nuova variabile x_{n+1} introdotta prende il nome di **variabile di slack** (scarto): $x_{n+1} = b_i - \sum_{j=1}^{n} a_{ij}x_j$. Questa variabile non compare nella funzione obiettivo, per cui può assumere qualunque valore non negativo. Quindi, in sintesi, tale variabile dovrà assumere il valore della differenza tra il termine noto e il valore assunto dalle variabili $x_j, \ldots, x_n$.
Es. $4x_1 + 6x_2 \leq 15 \Leftrightarrow 4x_1 + 6x_2 + x_3 = 15$.
- $x_1 = x_2 = 1$ è un assegnamento di valori che soddisfa il vincolo di disuguaglianza ($4 + 6 \leq 15$). In corrispondenza di tale soluzione, è possibile soddisfare il vincolo di uguaglianza assegnando i medesimi valori ad x_1 e x_2 ed un valore non negativo ad x_3: $x_3 = 5 \Rightarrow 4 + 6 + 5 = 15$.
- $x_1 = 1$, $x_2 = 2$ è un assegnamento di valori che <u>non</u> soddisfa il vincolo di disuguaglianza ($4 + 12 > 15$). <u>Non</u> è possibile soddisfare il vincolo di uguaglianza utilizzando lo stesso assegnamento di valori per x_1 e x_2 ed assegnando un valore non negativo ad x_3.

Per trasformare i *vincoli di maggiore o uguale* in vincoli di uguaglianza:
$\boxed{\sum_{j=1}^{n} a_{ij}x_j \geq b_i \Leftrightarrow \sum_{j=1}^{n} a_{ij}x_j - x_{n+1} = b_i}$, con $x_{n+1} \geq 0$.
La nuova variabile x_{n+1} introdotta prende il nome di **variabile di surplus** (eccedenza): $x_{n+1} = \sum_{j=1}^{n} a_{ij}x_j - b_i$.
Questa variabile non compare nella funzione obiettivo, per cui può assumere qualunque valore non negativo. Quindi, in sintesi, tale variabile dovrà assumere il valore della differenza tra il valore assunto dalle variabili $x_j, \ldots, x_n$ e il termine noto.

Per trasformare le *variabili del problema* minori o uguali a 0: $\boxed{x_j \leq 0 \Leftrightarrow -x_j \geq 0 \Rightarrow x_j' = -x_j}$, con $x_j' \geq 0$.
Siccome si necessita di variabili maggiori o uguali a 0, si rimpiazza la variabile x_j con una nuova variabile x'_j che assume esattamente il suo opposto (quindi, se $x_j = 3$ allora $x'_j = 3$).
Da questa trasformazione consegue la necessità di sostituire ovunque (vincoli e funzione obiettivo) la variabile x_j

con la variabile $-x'_j$.

Es. $7x_1 + 2x_2 = 5$, $x_1 \geq 0$, $x_2 \leq 0$.

$x_1 = 1$, $x_2 = -1$ è un assegnamento di valori che soddisfa il vincolo. Tuttavia, essendo $x_2 \leq 0$ non viene rispettata la forma standard. Quindi, si introduce la nuova variabile.

$7x_1 - 2x'_2 = 5$, $x_1 \geq 0$, $x'_2 \geq 0$, $x'_2 = -x_2$.

$x_1 = 1$, $x'_2 = -x_2 = 1$ è un assegnamento di valori che soddisfa il vincolo.

Per trasformare le *variabili non vincolate*: $\boxed{x_j \text{ n.v.} \Rightarrow x_j = (x'_j - x''_j)}$, con $x'_j, x''_j \geq 0$.

Quindi, data una variabile non vincolata (che può assumere qualsiasi valore) x_j, è sempre possibile ottenerne il valore attraverso la differenza tra due valori maggiori o uguali a 0. Quindi, si rimpiazzerà la variabile non vincolata, ovunque appaia nel modello, con la differenza tra queste due nuove variabili.

Es. $7x_1 - 2x_2 = 5$, $x_1 \geq 0$, x_2 n.v.

$x_1 = 1$, $x_2 = -3$ è un assegnamento di valori che soddisfa il vincolo. Ora, fissati due valori per x'_2 e x''_2 tali che $x'_2 - x''_2 = -3$ (ad esempio, per $x'_2 = 6$ e $x''_2 = 0$) si ha:

$7x_1 - 2(x'_2 - x''_2) = 7\cdot 1 - 2(-3) = 13 > 5$, $x_1 \geq 0$, $x'_2 \geq 0$ e $x''_2 \geq 0$.

Esercizio. Scrivere la forma standard per il seguente problema di programmazione lineare.

$$
\begin{array}{rcl}
\max\ z = x_1 - x_2 - x_3 & & \\
-3x_1 - x_2 + x_3 & <= & -3 \\
2x1 - 3x_2 - 2x_3 & >= & 4 \\
x_1 - x_3 & = & 2 \\
x_1 > & = & 0 \\
x_2 & <= & 0 \\
x_3 & n.v. &
\end{array}
$$

→Il primo passo è trasformare la funzione obiettivo, quindi essa diventa *- min (-x₁ + x₂ + x₃).*

Il primo vincolo, essendo un vincolo di disuguaglianza ≤, va trasformato prima in vincolo di ≥, quindi diventa $3x_1 + x_2 - x_3 \geq 3$. Per essere standard deve essere trasformato in vincolo di uguaglianza, quindi diventa $3x_1 + x_2 - x_3 - x_4 = 3$.

Il secondo vincolo va trasformato in vincolo di uguaglianza, quindi diventa $2x_1 - 3x_2 - 2x_3 - x_5 = 4$ (si faccia attenzione a non utilizzare le stesse variabili di slack/surplus).

Il terzo vincolo, essendo già un vincolo di uguaglianza, va bene così; resta, cioè, $x_1 - x_3 = 2$.

Il quarto vincolo, sulla variabile x_1, essendo già maggiore o uguale di 0, va bene così; resta, cioè, $x_1 \geq 0$.

Il quinto vincolo va trasformato in vincolo di ≥, quindi si introduce una nuova variabile al posto di x_2 e la si rimpiazza nel resto del modello: quindi, $x_2 = -x'_2 \Rightarrow x'_2 \geq 0$ (e la si sostituisce anche nella funzione obiettivo, nel primo vincolo e nel secondo vincolo).

Il sesto vincolo, sulla variabile n.v. x_3, va trasformato in vincolo di ≥, quindi si rimpiazza x_3 con la differenza di due variabili x'_3 e x''_3, e si effettua la sostituzione in tutti i punti del modello in cui compare x_3; quindi, il sesto vincolo si scinde in due nuovi vincoli $x'_3 \geq 0$ e $x''_3 \geq 0$.

Ovviamente, oltre alle variabili x_1, x_2 e x_3, anche le variabili x_4 e x_5 devono essere maggiori o uguali a 0.

Questo di cui a sx è il modello in forma standard risultante.

$$
\begin{array}{l}
-\min\ -x_1 - x'_2 + (x'_3 - x''_3) \\
3x_1 - x'_2 - (x'_3 - x''_3) - x_4 = 3 \\
2x_1 + 3x'_2 - 2(x'_3 - x''_3) - x_5 = 4 \\
x_1 - (x'_3 - x''_3) = 2 \\
x_1 \geq 0 \quad x'_2 \geq 0 \quad x_4 \geq 0 \\
\boxed{x_2 = -x'_2} \\
x_3 = (x'_3 - x''_3) \quad x'_3 \geq 0 \quad x''_3 \geq 0 \quad x_5 \geq 0
\end{array}
$$

Nuovo Sistema

$$-\min\left(-x_1 + x_2' + (x_3' - x_3'')\right)$$

$$
\begin{cases}
3x_1 - 2x_2' - (x_3' - x_3'') - x_4 = 3 \\
2x_1 + 3x_2' - 2(x_3' - x_3'') - x_5 = 4 \\
x_1 - (x_3' - x_3'') = 2 \\
x_1, x_2', x_3', x_3'', x_4, x_5 \geq 0
\end{cases}
$$

$x_2 = -x_2' \longleftrightarrow x_2' = -x_2$

$x_3 = (x_3' - x_3'')$

Vettori linearmente dipendenti e indipendenti

Verificare se i seguenti vettori:

$\underline{x}_1^T=(4, 1, 2)$, $\underline{x}_2^T=(7, 3, 1)$ e $\underline{x}_3^T = (3, 2, 0)$

sono linearmente indipendenti o dipendenti.

I vettori $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ sono LINEARMENTE INDIPENDENTI se $\lambda_1\underline{x}_1+\lambda_2\underline{x}_2+\dots+\lambda_n\underline{x}_n=\underline{0}$ implica che $\lambda_1=0, \lambda_2=0, \dots, \lambda_n=0$

$$\lambda_1\begin{pmatrix}4\\1\\2\end{pmatrix}+\lambda_2\begin{pmatrix}7\\3\\1\end{pmatrix}+\lambda_3\begin{pmatrix}3\\2\\0\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix} \Rightarrow \begin{cases}4\lambda_1+7\lambda_2+3\lambda_3=0\\ \lambda_1+3\lambda_2+2\lambda_3=0\\ 2\lambda_1+\lambda_2=0\end{cases}$$

$$\begin{cases}4\lambda_1-14\lambda_1+3\lambda_3=0\\ \lambda_1-6\lambda_1+2\lambda_3=0\\ \lambda_2=-2\lambda_1\end{cases} \Rightarrow \begin{cases}-10\lambda_1+3\lambda_3=0\\ -5\lambda_1+2\lambda_3=0\\ \lambda_2=-2\lambda_1\end{cases}$$

$$\begin{cases}-10\lambda_1+3\lambda_3=0\\ \lambda_1=2/5\,\lambda_3\\ \lambda_2=-2\lambda_1\end{cases} \Rightarrow \begin{cases}-10\,(2/5\,\lambda_3)+3\lambda_3=0\\ \lambda_1=2/5\,\lambda_3\\ \lambda_2=-2\lambda_1\end{cases}$$

$$\begin{cases}-4\lambda_3+3\lambda_3=0\\ \lambda_1=2/5\,\lambda_3\\ \lambda_2=-2\lambda_1\end{cases} \Rightarrow \begin{cases}-\lambda_3=0\\ \lambda_1=2/5\,\lambda_3\\ \lambda_2=-2\lambda_1\end{cases} \Rightarrow \begin{cases}\lambda_3=0\\ \lambda_1=0\\ \lambda_2=0\end{cases}$$

Sono linearmente indipendenti

- - - - - - - - - -

Costruiamo la matrice $A=(\underline{x}_1, \underline{x}_2, \underline{x}_3)$

$$A=\begin{bmatrix}4&7&3\\1&3&2\\2&1&0\end{bmatrix}$$

Metodo alternativo

1. A è invertibile sse le sue righe (le sue colonne) sono linearmente indipendenti
2. A è invertibile sse il suo determinante è diverso da zero.

$$det(A)=2\cdot det\begin{bmatrix}7&3\\3&2\end{bmatrix}-1\cdot det\begin{bmatrix}4&3\\1&2\end{bmatrix}+0\cdot det\begin{bmatrix}4&7\\1&3\end{bmatrix}$$
$$=2\cdot(14-9)-1\cdot(8-3)+0\cdot(12-7)$$
$$=10-5=5\neq 0$$

Sono linearmente indipendenti

Se det(A)≠0 le righe (le colonne) di A sono linearmente indipendenti.

Cambiare, ora, il vettore $\underline{x}_1^T = (4, 1, 2)$ con $\underline{x}_1^T = (4, 1, 1)$ e verifichiamo se:

$\underline{x}_1^T=(4, 1, 1)$, $\underline{x}_2^T=(7, 3, 1)$ e $\underline{x}_3^T = (3, 2, 0)$

sono linearmente indipendenti o dipendenti.

$$\lambda_1\begin{pmatrix}4\\1\\1\end{pmatrix}+\lambda_2\begin{pmatrix}7\\3\\1\end{pmatrix}+\lambda_3\begin{pmatrix}3\\2\\0\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix} \Rightarrow \begin{cases}4\lambda_1+7\lambda_2+3\lambda_3=0\\ \lambda_1+3\lambda_2+2\lambda_3=0\\ \lambda_1+\lambda_2=0\end{cases}$$

$$\begin{cases}4\lambda_1+7\lambda_2+3\lambda_3=0\\ \lambda_1+3\lambda_2+2\lambda_3=0\\ \lambda_1=-\lambda_2\end{cases} \Rightarrow \begin{cases}-4\lambda_2+7\lambda_2+3\lambda_3=0\\ -\lambda_2+3\lambda_2+2\lambda_3=0\\ \lambda_1=-\lambda_2\end{cases} *$$

$$* \begin{cases}-4\lambda_2+7\lambda_2+3\lambda_3=0\\ -\lambda_2+3\lambda_2+2\lambda_3=0\\ \lambda_1=-\lambda_2\end{cases} \Rightarrow \begin{cases}3\lambda_2+3\lambda_3=0\\ 2\lambda_2+2\lambda_3=0\\ \lambda_1=-\lambda_2\end{cases} \Rightarrow \begin{cases}0=0\\ \lambda_2=-\lambda_3\\ \lambda_1=-\lambda_2\end{cases}$$

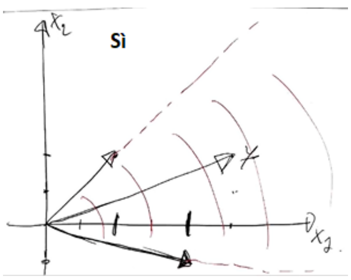
Fissiamo $\lambda_3=1$ ottenendo $\lambda_2=-1$ e $\lambda_1=1$.

$$1\begin{pmatrix}4\\1\\1\end{pmatrix}-1\begin{pmatrix}7\\3\\1\end{pmatrix}+1\begin{pmatrix}3\\2\\0\end{pmatrix}=\begin{pmatrix}4\\1\\1\end{pmatrix}+\begin{pmatrix}-7\\-3\\-1\end{pmatrix}+\begin{pmatrix}3\\2\\0\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

Poichè la combinazione lineare dei tre vettori con questi coefficienti restituisce il vettore nullo, $\underline{x}_1^T$, $\underline{x}_2^T$, $\underline{x}_3^T$ non sono lineamente indipendenti.

Combinazioni lineari, coniche e convesse

- *Un vettore $\underline{y}$ è combinazione LINEARE dei vettori $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ se esistono $\lambda_1, \lambda_2, \dots, \lambda_n$ numeri reali tali che: $\underline{y}=\lambda_1\underline{x}_1+\lambda_2\underline{x}_2+\dots+\lambda_n\underline{x}_n$*
- *Un vettore $\underline{y}$ è combinazione CONICA dei vettori $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ se esistono $\lambda_1, \lambda_2, \dots, \lambda_n$ numeri reali tali che: $\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$ e $\underline{y}=\lambda_1\underline{x}_1+\lambda_2\underline{x}_2+\dots+\lambda_n\underline{x}_n$*
- *Un vettore $\underline{y}$ è combinazione CONVESSA dei vettori $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ se esistono $\lambda_1, \lambda_2, \dots, \lambda_n$ numeri reali tali che: $\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$ e $\lambda_1+\lambda_2+\dots+\lambda_n=1$ e $\underline{y}=\lambda_1\underline{x}_1+\lambda_2\underline{x}_2+\dots+\lambda_n\underline{x}_n$*



Esercizi:

Determinare **geometricamente** se:

1)il vettore $\underline{y}^T=(8/3, 1)$ è combinazione **conica** dei vettori $\underline{x}_1^T=(1, 1)$ e $\underline{x}_2^T = (2, -1)$.

Determinare che tipo di combinazione (lineare, conica o convessa) è il vettore $\underline{y}^T = (5/3, 2/3, 4)$ rispetto ai vettori:

$\underline{x}_1^T=(2, 1, 8)$, $\underline{x}_2^T=(1, 0, 3)$ e $\underline{x}_3^T = (2, 1, 1)$

$$\lambda_1\begin{pmatrix}2\\1\\8\end{pmatrix}+\lambda_2\begin{pmatrix}1\\0\\3\end{pmatrix}+\lambda_3\begin{pmatrix}2\\1\\1\end{pmatrix}=\begin{pmatrix}5/3\\2/3\\4\end{pmatrix} \Rightarrow \begin{cases}2\lambda_1+\lambda_2+2\lambda_3=5/3\\ \lambda_1+\quad\lambda_3=2/3\\ 8\lambda_1+3\lambda_2+\lambda_3=4\end{cases}$$

$$\begin{cases}2\lambda_1+\lambda_2+2\lambda_3=5/3\\ \lambda_1=2/3-\lambda_3\\ 8\lambda_1+3\lambda_2+\lambda_3=4\end{cases} \Rightarrow \begin{cases}2(2/3-\lambda_3)+\lambda_2+2\lambda_3=5/3\\ \lambda_1=2/3-\lambda_3\\ 8(2/3-\lambda_3)+3\lambda_2+\lambda_3=4\end{cases} *$$

$$* \begin{cases}2(2/3-\lambda_3)+\lambda_2+2\lambda_3=5/3\\ \lambda_1=2/3-\lambda_3\\ 8(2/3-\lambda_3)+3\lambda_2+\lambda_3=4\end{cases} \Rightarrow \begin{cases}4/3-2\lambda_3+\lambda_2+2\lambda_3=5/3\\ \lambda_1=2/3-\lambda_3\\ 16/3-8\lambda_3+3\lambda_2+\lambda_3=4\end{cases}$$

$$\begin{cases}\lambda_2=5/3-4/3=1/3\\ \lambda_1=2/3-\lambda_3\\ -7\lambda_3+3\lambda_2=4-16/3=-4/3\end{cases} \Rightarrow \begin{cases}\lambda_2=1/3\\ \lambda_1=2/3-\lambda_3\\ -7\lambda_3+1=-4/3\end{cases}$$

$$\begin{cases}\lambda_2=1/3\\ \lambda_1=2/3-\lambda_3\\ -7\lambda_3+1=-4/3\end{cases} \Rightarrow \begin{cases}\lambda_2=1/3\\ \lambda_1=2/3-\lambda_3\\ -7\lambda_3=-1-4/3=-7/3\end{cases}$$

$$\begin{cases}\lambda_2=1/3\\ \lambda_1=2/3-\lambda_3\\ \lambda_3=1/3\end{cases} \Rightarrow \begin{cases}\lambda_2=1/3\\ \lambda_1=1/3\\ \lambda_3=1/3\end{cases}$$

$$\frac{1}{3}\begin{pmatrix}2\\1\\8\end{pmatrix}+\frac{1}{3}\begin{pmatrix}1\\0\\3\end{pmatrix}+\frac{1}{3}\begin{pmatrix}2\\1\\1\end{pmatrix}=\begin{pmatrix}5/3\\2/3\\4\end{pmatrix}$$

Combinazione convessa

- *Un vettore $\underline{y}$ è combinazione LINEARE dei vettori $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ se esistono $\lambda_1, \lambda_2, \ldots, \lambda_n$ numeri reali tali che: $\underline{y}=\lambda_1\underline{x}_1+\lambda_2\underline{x}_2+\ldots+\lambda_n\underline{x}_n$*
- *Un vettore $\underline{y}$ è combinazione CONICA dei vettori $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ se esistono $\lambda_1, \lambda_2, \ldots, \lambda_n$ numeri reali tali che: $\lambda_1, \lambda_2, \ldots, \lambda_n \geq 0$ e $\underline{y}=\lambda_1\underline{x}_1+\lambda_2\underline{x}_2+\ldots+\lambda_n\underline{x}_n$*
- *Un vettore $\underline{y}$ è combinazione CONVESSA dei vettori $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ se esistono $\lambda_1, \lambda_2, \ldots, \lambda_n$ numeri reali tali che: $\lambda_1, \lambda_2, \ldots, \lambda_n \geq 0$ e $\lambda_1+\lambda_2+\ldots+\lambda_n=1$ e $\underline{y}=\lambda_1\underline{x}_1+\lambda_2\underline{x}_2+\ldots+\lambda_n\underline{x}_n$*

□

Si determini un vettore che sia combinazione conica dei seguenti tre vettori:

$\underline{x}_1^T=(3, 0, 1)$, $\underline{x}_2^T=(5, 4, 1)$ e $\underline{x}_3^T = (1, 3, 8)$

$$\lambda_1\begin{pmatrix}3\\0\\1\end{pmatrix}+\lambda_2\begin{pmatrix}5\\4\\1\end{pmatrix}+\lambda_3\begin{pmatrix}1\\3\\8\end{pmatrix}=\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$$

Combinazione conica implica che $\lambda_1, \lambda_2, \ldots, \lambda_n \geq 0$

$$\frac{1}{3}\begin{pmatrix}3\\0\\1\end{pmatrix}+1\begin{pmatrix}5\\4\\1\end{pmatrix}+0\begin{pmatrix}1\\3\\8\end{pmatrix}=\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix} \Rightarrow \begin{cases}1+5=y_1\\ 4=y_2\\ 1/3+1=y_3\end{cases} \Rightarrow \begin{cases}6=y_1\\ 4=y_2\\ 4/3=y_3\end{cases}$$

□

Esempio: Pianificazione della produzione (formulazione)

Un'industria fabbrica 4 tipi di prodotti, **P1**, **P2**, **P3**, **P4**, la cui lavorazione è affidata a due reparti dell'industria: il reparto produzione e il reparto confezionamento. Per ottenere i prodotti pronti per la vendita è necessaria naturalmente la lavorazione in entrambi i reparti. La tabella che segue riporta, per ciascun tipo di prodotto i tempi (in ore) necessari di lavorazione in ciascuno dei reparti per avere una tonnellata di prodotto pronto per la vendita.

	P1	P2	P3	P4
Reparto produzione	2	1.5	0.5	2.5
Reparto confezionamento	0.5	0.25	0.25	1

Ciascuna tonnellata di prodotto dà i profitti (prezzi espressi in Euro per tonnellata) secondo la tabella di cui in basso a destra.

	P1	P2	P3	P4
Profitto	250	230	110	350

Determinare le quantità che si devono produrre settimanalmente di ciascun tipo di prodotto in modo da massimizzare il profitto complessivo, sapendo che ogni settimana, il reparto produzione e il reparto confezionamento hanno una capacità lavorativa massima rispettivamente di 100 e 50 ore.

→

Variabili di decisione. È naturale introdurre le variabili reali $\mathbf{x_1}$, $\mathbf{x_2}$, $\mathbf{x_3}$ e $\mathbf{x_4}$ rappresentanti rispettivamente le quantità di prodotto P1, P2, P3, P4 da fabbricare in una settimana.
Funzione obiettivo. Ciascuna tonnellata di prodotto contribuisce al profitto totale secondo la tabella data. Quindi il profitto totale sarà:

$$\max 250x_1 + 230x_2 + 110x_3 + 350x_4.$$

Vincoli. Ovviamente la capacità produttiva della fabbrica (risorsa «scarsa») limita i valori che possono assumere le variabili; infatti si ha una capacità massima lavorativa in ore settimanali di ciascun reparto. In particolare per il reparto produzione si hanno a disposizione al più 100 ore settimanali e poiché ogni tonnellata di prodotto P1 utilizza il reparto produzione per 2 ore, ogni tonnellata di prodotto P2 utilizza il reparto produzione per 1.5 ore e così via per gli altri tipi di prodotti si dovrà avere:

$$2x_1 + 1.5x_2 + 0.5x_3 + 2.5x_4 \leq 100.$$

Ragionando in modo analogo per il reparto confezionamento si ottiene:

$$0.5x_1 + 0.25x_2 + 0.25x_3 + x_4 \leq 50.$$

Vincolo di non negatività delle variabili:

$$x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0,\ x_4 \geq 0.$$

La formulazione finale quindi può essere scritta in questa forma:

$\max 250x_1 + 230x_2 + 110x_3 + 350x_4$
$2x_1 + 1.5x_2 + 0.5x_3 + 2.5x_4 \leq 100$ (utilizzo reparto produzione)
$0.5x_1 + 0.25x_2 + 0.25x_3 + x_4 \leq 50$ (utilizzo reparto confezionamento)
$x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0,\ x_4 \geq 0.$

Consideriamo ulteriori richieste (vincoli):
1. P1 non può utilizzare per più di 20 ore settimanali il reparto di produzione.
2. La produzione di P2 non può superare quella di P3.
3. P4 può utilizzare complessivamente 30 ore settimanali di lavorazione (tra il reparto di produzione e quello di confezionamento).
4. La produzione di P1 non può superare il doppio della produzione di P2 e P3.

Segue che:
1. P1 non può utilizzare per più di 20 ore settimanali il reparto di produzione:
   $2x_1 \leq 20 \Leftrightarrow x_1 \leq 10.$
2. La produzione di P2 non può superare quella di P3:
   $x_2 \leq x_3.$
3. P4 può utilizzare complessivamente 30 ore settimanali di lavorazione:
   $2.5x_4 + x_4 \leq 30 \Leftrightarrow 3.5x_4 \leq 30.$
4. La produzione di P1 non può superare il doppio della produzione di P2 e P3:
   $x_1 \leq 2(x_2 + x_3).$

□

Esempio: Stuccare, imbiancare e levigare.
Supponiamo che ci siano tre lavori da svolgere: **stuccare**, **imbiancare** e **levigare**. Abbiamo a disposizione tre persone **Mario**, **Luca** ed **Andrea** che sanno svolgere questi tre lavori ma con differenti tempistiche come indicato nella seguente tabella (i valori rappresentano le ore necessarie ad ogni persona per portare a termine il rispettivo lavoro).

	STUCCA	IMBIANCA	LEVIGA
MARIO	3	1	2
LUCA	2	1.5	1.5
ANDREA	3	1.5	3

Il nostro obiettivo è quello di assegnare ad ogni persona un lavoro e ad ogni lavoro una persona al fine di minimizzare le ore totali necessarie per svolgere i tre lavori.

$$x_{ij} = \begin{cases} 1 & \text{se assegnamo alla persona } i \text{ il lavoro } j \\ 0 & \text{altrimenti} \end{cases}$$

$$\text{Min } 3X_{11}+1X_{12}+2X_{13}+3X_{21}+1.5X_{22}+1.5X_{23}+3X_{31}+1.5X_{32}+3X_{33}$$

Min $3X_{11}+1X_{12}+2X_{13}+3X_{21}+1.5X_{22}+1.5X_{23}+3X_{31}+1.5X_{32}+3X_{33}$

$X_{11}+X_{12}+X_{13}=1$ (Mario)

$X_{21}+X_{22}+X_{23}=1$ (Luca)

$X_{31}+X_{32}+X_{33}=1$ (Andrea)

$X_{11}+X_{21}+X_{31}=1$ (Levigare)

$X_{12}+X_{22}+X_{32}=1$ (Stuccare)

$X_{13}+X_{23}+X_{33}=1$ (Imbiancare)

$$\begin{array}{l} X_{11}+X_{12}+X_{13}=1 \\ X_{21}+X_{22}+X_{23}=1 \\ X_{31}+X_{32}+X_{33}=1 \end{array} \Rightarrow \sum_{j=1}^{3} x_{ij} = 1 \quad \text{per ogni } i = 1,2,3$$

$$\begin{array}{l} X_{11}+X_{21}+X_{31}=1 \\ X_{12}+X_{22}+X_{32}=1 \\ X_{13}+X_{23}+X_{33}=1 \end{array} \Rightarrow \sum_{i=1}^{3} x_{ij} = 1 \quad \text{per ogni } j = 1,2,3$$

Min $3X_{11}+1X_{12}+2X_{13}+3X_{21}+1.5X_{22}+1.5X_{23}+3X_{31}+1.5X_{32}+3X_{33}$

c_{11} c_{22}

$$\min \sum_{i=1}^{3}\sum_{j=1}^{3} c_{ij}x_{ij}$$

$$\min \sum_{i=1}^{3}\sum_{j=1}^{3} c_{ij}x_{ij}$$

$$\sum_{j=1}^{3} x_{ij} = 1 \quad \text{per ogni } i = 1,2,3$$

$$\sum_{i=1}^{3} x_{ij} = 1 \quad \text{per ogni } j = 1,2,3$$

Generalizzando…

$$\min \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij}$$

$$\sum_{j=1}^{n} x_{ij} = 1 \quad \text{per ogni } i = 1,2,\dots n$$

$$\sum_{i=1}^{n} x_{ij} = 1 \quad \text{per ogni } j = 1,2,\dots n$$

□

Esempio: scommesse

Il signor Rossi è uno scommettitore incallito ma poco fortunato. Dopo aver perso tutte le scommesse della giornata ha deciso di puntare gli ultimi 100 euro sulla vincitrice della coppa Italia. Questa volta però vuole essere assolutamente sicuro di vincere e per farlo chiede aiuto ad un amico che ha studiato un po' di ricerca operativa. Le 4 squadre rimaste in gara per la vittoria finale sono Milan, Juventus, Napoli e Siena quotate rispettivamente 3:1, 4:1, 7:1, 10:1. Una di queste 4 squadre sarà sicuramente la vincitrice del torneo. Quanto deve scommettere il signor Rossi su ogni squadra per massimizzare la vincita nel caso peggiore?

→ Le variabili del problema sono: x_J (# euro su Juventus), x_M (# euro su Milan), x_N (# euro su Napoli), x_S (# euro su Siena); ovviamente tutte le variabili devono essere maggiori o uguali di 0. Stabilite le variabili, possiamo definire il vincolo per cui $x_J + x_M + x_N + x_S = budget$.

La tabella a destra ci consente di farci un'idea sulla funzione obiettivo.

I valori sulla riga *i* rappresentano la puntata sulla squadra sulla colonna corrispondente, ed i valori sulla riga *i+1* (colorata sul bordo) sono le vincite corrispondenti. Noi cerchiamo il valore che massimizzi la vincita nel caso peggiore.

M	J	N	S
3	4	7	10
20€	30€	50€	0€
60€	120€	350€	0€
30€	20€	40€	10€
90€	80€	280€	100€

Il valore minimo tra i valori delle possibili vincite dipende dalla puntata, pertanto dev'essere anch'esso una variabile *y* del problema: questa sarà utilizzata dal valore obiettivo, che possiamo formulare in ***max y - budget***.

Possiamo definire, inoltre, i vincoli per cui $y \le 3x_M$, così come $y \le 4x_J$, $y \le 7x_N$, $y \le 3x_S$.

Questo di cui a destra è il modello matematico risultante.

$$\begin{array}{l} \max \quad y - budget \\ s.t. \\ x_m + x_j + x_n + x_s = budget \\ y \le 3x_m \\ y \le 4x_j \\ y \le 7x_n \\ y \le 10x_s \\ x_j, x_m, x_n, x_s, y \ge 0 \end{array}$$

Calcolando l'ottimo su Excel, scopriremmo che esso sarebbe

OPT: profitto = 20, con x_M = 40, x_J = 30, x_N, x_S = 12.

Questo problema è "*facile*" da risolvere, se le puntate sono formulate con valori reali, *difficile* da risolvere altrimenti.

□

Risoluzione grafica dei problemi di PL, iperpiani e semispazi

Dato un Problema di Ottimizzazione (PO) del tipo

$$\min f(\underline{x})$$
$$s.t.$$
$$\underline{x} \in X \subseteq \mathbb{R}^n,$$

diamo la definizione di ottimo globale ed ottimo locale.

Def. Un punto $\underline{x}^* \in X$ è un *ottimo globale* per la funzione $f(\underline{x})$ se e solo se: $f(\underline{x}^*) \leq f(\underline{x})\ \forall \underline{x} \in X$.
Quindi, l'ottimo globale indica che, tra gli infiniti punti della regione ammissibile, esso è il punto in cui la funzione assume il valore minimo.

Def. Un punto $\underline{x}' \in X$ è un *ottimo locale* per la funzione $f(\underline{x})$ se e solo se: $f(\underline{x}') \leq f(\underline{x})\ \forall \underline{x} \in N(\underline{x}', \varepsilon)$, con $\varepsilon > 0$.

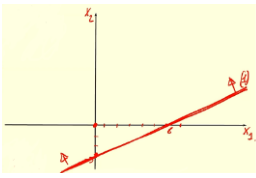
Commentato [DE7]: intorno

Quindi, l'ottimo locale indica che, ottenuta una soluzione nel punto $\underline{x}'$, ci sarà un insieme di punti $\underline{x}_i$ ($i = 1, 2, \dots$) nell'intorno di $\underline{x}'$ tale che la funzione $f(\underline{x}_i)$ è più grande o al più uguale a $f(\underline{x}')$. Dunque, localmente, quella è la migliore soluzione possibile.
Dalle definizioni, segue che un ottimo globale è anche un ottimo locale; generalmente, non vale il viceversa: esistono, però, casi particolari in cui tutti gli ottimi locali sono anche ottimi globali.

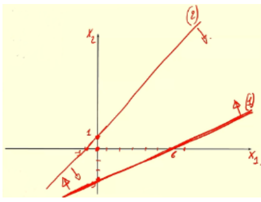
Es. L'azienda Rossi &C. ha vinto una gara d'appalto per la produzione di due tipologie di leghe di acciaio L1 ed L2. Il contratto prevede il pagamento di 10 milioni di euro a condizione che siano rispettate le seguenti proporzioni tra le tonnellate delle due leghe prodotte:
(1) la metà delle tonnellate di L1 prodotte non devono superare, per al più 3 unità, le tonnellate di L2 prodotte;
(2) le tonnellate di L2 possono essere al più di uno superiori a quelle di L1;
(3) le tonnellate di L2 prodotte non devono mai superare il doppio delle tonnellate di L1 decrementate di 2.
Sapendo che l'azienda spende 3 milioni di euro per produrre una tonnellata della lega L1 ed un milione di euro per la lega L2, individuare un piano di produzione che rispetti i vincoli di produzione minimizzando però i costi di produzione.
L'attuale piano di produzione individuato prevede la produzione di 2 tonnellate di L1 e mezza tonnellata di L2 per una spesa totale di 6,5 milioni di euro e un profitto finale pari a 10 - 6,5 = 3,5 milioni. Si può fare di meglio?
→ Il seguente è il modello matematico del problema.

$$\min z = 3x_1 + x_2$$
$$(1)\ \frac{1}{2}x_1 - x_2 \leq 3$$
$$(2)\ -x_1 + x_2 \leq 1$$
$$(3)\ 2x_1 - x_2 \geq 2$$
$$(4)\ x_1, x_2 \geq 0$$

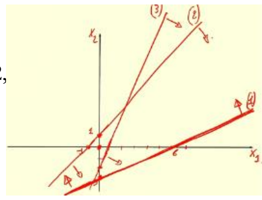
a) Risolvere graficamente il problema
→ *Quando un problema ha soltanto due variabili* (pertanto si può lavorare in $\mathbb{R}^2$) *è possibile risolvere il problema graficamente.* In primis, si costruisce la regione ammissibile: si individuano, cioè, gli infiniti punti che soddisfano il sistema di disequazioni; dalla (4), sappiamo che lavoriamo solo sul primo quadrante.



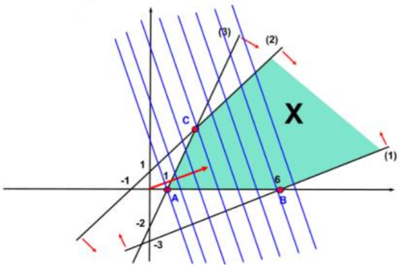
Per la (1), si considera la retta $\frac{1}{2}x_1 - x_2 = 3$: per disegnarla, azzeriamo prima x_1 (per trovare x_2), poi x_2 (per trovare x_1); troviamo, di conseguenza, i punti (0, -3) e (6, 0) e disegniamo la retta. Siccome consideriamo un vincolo ≤, bisogna considerare tutta la regione che lo soddisfa, per cui: scegliamo un qualsiasi punto (in questo esempio, (0, 0) va bene) che non sia sull'iperpiano (vedremo in seguito), se esso soddisfa la disequazione allora possiamo considerare tutti i punti che si trovano nella parte di semipiano (definito dalla retta disegnata) in cui è presente tale punto.



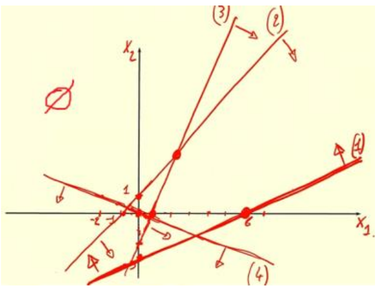
Per la (2), procediamo come il caso precedente ed otteniamo i punti (0, 1) e (-1, 0), per cui disegniamo la retta; in questo caso, notiamo che il vincolo viene soddisfatto per tutti i punti presenti nella parte del semipiano inferiore definito dalla nuova retta (quindi, la nostra regione per i primi due vincoli è quella compresa tra le due rette, nel primo quadrante).



Per la (3), otteniamo i punti (0, -2) e (1, 0), quindi tracciamo la retta: in questo caso, notiamo che il vincolo viene soddisfatto per tutti i punti presenti nella parte del semipiano inferiore (in quanto, ad esempio, preso (0, 0) che, sostituendo, non è ≥ 2, ci troviamo nella parte superiore del semipiano, per cui l'insieme dei punti che rispetta questo vincolo è l'insieme dei punti alla parte inferiore) creato dalla nuova retta.
La regione ammissibile è quella in cui vengono rispettati tutti i vincoli, cioè graficamente l'insieme dei punti che soddisfa tutti i semispazi contemporaneamente. Da ciò consegue che la regione ammissibile è la X mostrata nella figura seguente.



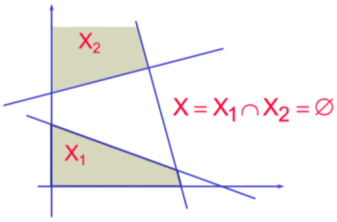
Ora, tra gli infiniti punti si vuole individuare il punto che vada a minimizzare la funzione obiettivo; il punto <u>x</u> sarà la *soluzione ottima*, mentre z sarà il *valore* della soluzione ottima.
Si osservi che, senza il 'min', $z = 3x_1 + x_2$ è una retta di cui non si conosce il termine noto *z*. Se nel piano in $\mathbb{R}^2$ si modifica il termine noto di una retta, si sta calcolando il fascio di rette parallele; la funzione obiettivo definisce, al variare di z, un fascio di rette parallele dette *curve di livello*. Ciò è importante, in quanto per individuare la soluzione ottima del problema si deve utilizzare il **gradiente** della funzione obiettivo, ossia un vettore le cui componenti sono i coefficienti di costo della funzione obiettivo. In questo esempio, il gradiente della funzione obiettivo è g = (3, 1). Il vettore disegnato in rosso nella figura precedente è il gradiente per questa funzione obiettivo. Una volta disegnato il gradiente, va tenuto a mente che:
1) *il gradiente è perpendicolare alle curve di livello*, quindi al fascio di rette parallele definito dalla funzione obiettivo;
2) *il verso del gradiente indica sempre la direzione di crescita della funzione obiettivo*.
Stabilite queste due informazioni, quando va trovata la soluzione ottima non si deve far altro che disegnare il fascio di rette parallele che siano perpendicolari al gradiente, nella direzione *opposta (se sto minimizzando) o concorde (se sto massimizzando)* indicata da quest'ultimo. Il fascio di rette va tracciato fin quando non si tocca <u>l'ultimo</u> punto (se esiste) della regione ammissibile, <u>mai</u> al primo punto (anche se sto minimizzando). Per questo problema, quindi, il punto di ottimo è il punto **A = <u>x</u>* = (1, 0)**, che implica **z* =** 3·1 + 1·0 **= 3**.
Se avessimo avuto una funzione di massimo, in questo caso la soluzione ottima non ci sarebbe stata, in quanto, fissato un punto, si riesce sempre a trovare un altro punto in cui la funzione migliora strettamente. Il valore di z* sarebbe stato, quindi, ∞. □



Supponiamo, ora, che ci sia un nuovo vincolo: $(4)\ x_1 + 2x_2 \leq 0$. Quando c'è un vincolo con termine noto 0, non possiamo azzerare una volta x_1 e una volta x_2 per trovare i punti, dato che troveremmo due volte il punto (0, 0), ed un punto non è sufficiente per disegnare una retta. A tal scopo, poniamo $x_2 = 1$ e troviamo anche il punto (-2, 1), quindi disegniamo la retta. Il semispazio definito da questo vincolo è quello al semipiano inferiore, in quanto se sostituiamo (-1, -1) all'equazione otteniamo -3 ≤ 0, e tale punto si trova al semipiano inferiore.
Se avessimo avuto anche questo vincolo (4) nell'esempio in questione, avremmo ottenuto una regione ammissibile *vuota*.

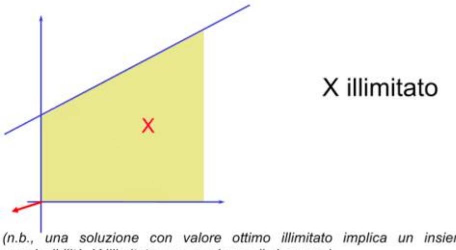
In questo caso non ci sono soluzioni al problema

Sappiamo che un problema di PL può essere *non ammissibile*, *ammissibile con valore ottimo illimitato* e **ammissibile con valore ottimo finito**; per quest'ultimo vanno distinti due casi:
1) quando si ha **unico punto di ottimo**;
2) quando si hanno **infiniti punti di ottimo**.
Va notato che avere infiniti punti di ottimo è diverso da avere una soluzione ammissibile con valore ottimo illimitato. Infatti, nel primo caso si ha che nella regione ammissibile esistono infiniti punti in cui la funzione assume il valore ottimo; invece, nel secondo caso non esiste un punto di ottimo finito, giacché è $\pm\infty$ a seconda se sto massimizzando o minimizzando.
Def. Un problema di ottimizzazione si dice *inammissibile* se $X = \emptyset$, cioè non esistono soluzioni ammissibili.
Graficamente:



$$X = \emptyset \Rightarrow \nexists \underline{x} \in \mathbf{R}^n : A\underline{x} \geq \underline{b},\ \underline{x} \geq 0$$

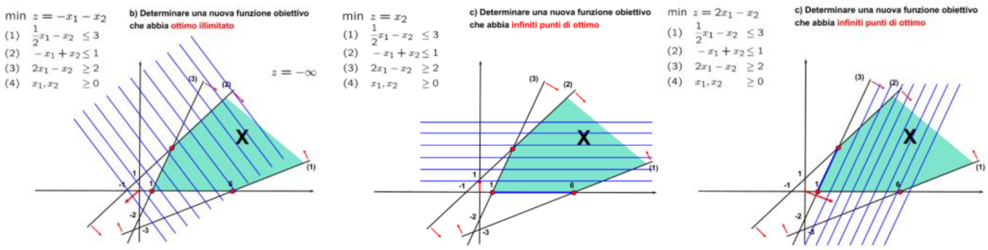
Def. Un problema di ottimizzazione si dice *illimitato* *(inferiormente)* se, scelto un qualsiasi valore M > 0, esiste sempre un punto $x \in X$ tale che $f(x) < -M$.
Graficamente:



(n.b., una soluzione con valore ottimo illimitato implica un insieme di ammissibilità X illimitato, ma non è vero il viceversa)

Es. Dato l'esempio precedente, se avessimo avuto una funzione obiettivo $\max z = 3x_1 + x_2$, allora il problema sarebbe stato illimitato (dato che bisognava tracciare il fascio di rette nella direzione concorde a quella del gradiente).

Quando possibile, si possono sfruttare gli assi cartesiani. Ad esempio, se avessimo avuto $\max z = x_1$, allora avremmo avuto gradiente (1, 0), per cui tracciando il fascio di rette nella direzione concorde a quella del gradiente si avrebbe un numero infinito di rette perpendicolari ad x_2.
Seguono ulteriori esempi.



Es. Una multinazionale produce due versioni di una bevanda energetica: normale e super. Per ogni quintale di bevanda venduta, l'azienda ha un profitto pari ad 1000 euro per il tipo normale e 1200 euro per il tipo super. Nella produzione è necessario utilizzare in sequenza tre tipi di macchinari, A, B, C, che ogni giorno possono lavorare un numero di ore massimo come riportato nella tabella seguente:

	ORE	NORMALE	SUPER
A	4	1	0.4
B	6	0.75	1
C	3.5	1	0

Per produrre un quintale di bevanda (normale o super) è richiesto l'utilizzo delle macchine per il tempo indicato nella stessa tabella. L'obiettivo del signor Rossi è quello di pianificare la produzione giornaliera dei due tipi di bevande al fine di massimizzare il profitto(supponendo che l'intera produzione verrà venduta).

→ Il nostro obiettivo è decidere **quanti quintali** produrre per ogni tipologia di bevanda; assegniamo ad ogni **tipologia di bevanda** una **variabile** (x_1 = normale, x_2 = super).

I **vincoli** del problema devono modellare il rispetto del numero massimo di ore di lavorazione per ogni **macchinario**.

Abbiamo il modello matematico di cui a destra. Il problema è facile, e poiché questo modello ha soltanto due variabili, è possibile risolverlo graficamente.

$$\begin{aligned}&\max \;\; 1000x_1 + 1200x_2\\ &x_1 + 0.4x_2 \le 4\\ &x_1 \le 3.5\\ &0.75x_1 + x_2 \le 6\\ &\underline{x} \ge \underline{0}\end{aligned}$$

Es. Il cuoco del ristorante dove lavoriamo ci ha assegnato il compito di andare a comprare le mele e le arance con 20 euro in tasca. Il costo di ogni kg di mele è pari a 5 euro mentre ogni kg di arance costa 2 euro. Inoltre il cuoco non vuole che acquistiamo più di 3.5 kg di mele. Infine il fruttivendolo questa settimana offre un buono sconto da 1 euro su ogni kg di mele e di 1.2 euro su ogni kg di arance acquistato. Questi buoni sconto sono però offerti a condizione che il numero di kg di mele, moltiplicato per 3, più il numero di kg di arance, moltiplicato per 4, non superi i 24 kg. L'obiettivo da raggiungere è quello di ottenere il massimo sconto, rispettando però le indicazioni sia del cuoco che del fruttivendolo.

→

- x_1 = chili di mele da acquistare, x_2 = chili di arance da acquistare;
- Funzione obiettivo: Massimizzare il **valore totale** dei **buoni sconto** ottenuti;
- Vincolo 1: rispetto del **limite di spesa**;
- Vincolo 2: rispetto della **richiesta del cuoco**;
- Vincolo 3: rispetto della **condizione imposta dal fruttivendolo** per avere accesso ai buoni sconto.

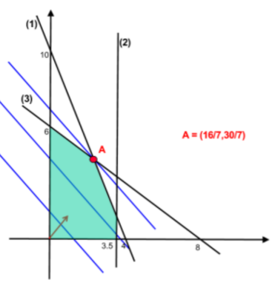
Abbiamo il modello matematico di cui a destra. Il problema è facile, e poiché questo modello ha soltanto due variabili, è possibile risolverlo graficamente.

$$\begin{aligned}&\max \;\; x_1 + 1.2x_2\\ &5x_1 + 2x_2 \le 20\\ &x_1 \le 3.5\\ &3x_1 + 4x_2 \le 24\\ &\underline{x} \ge \underline{0}\end{aligned}$$

Notiamo che i modelli matematici di questi ultimi due esempi sono *equivalenti*. In realtà, il sistema di vincoli è identico: se moltiplichiamo il primo vincolo del primo problema per 5, otteniamo il primo vincolo del secondo problema; inoltre, risolvendo questi due problemi graficamente, notiamo che la regione ammissibile è la stessa. Infine, la funzione obiettivo dei due problemi ha esattamente lo stesso gradiente: ne stiamo, cioè, cambiando solo il modulo. Il valore della funzione obiettivo all'ottimo per P1 ($\underline{c}^T\underline{x}$) sarà pari a 1000 volte quello di P2.

$$\begin{aligned}&\max \;\; 1000x_1 + 1200x_2\\ &x_1 + 0.4x_2 \le 4\\ &x_1 \le 3.5\\ &0.75x_1 + x_2 \le 6\\ &\underline{x} \ge \underline{0}\end{aligned}$$

$$\begin{aligned}&\max \;\; x_1 + 1.2x_2\\ &5x_1 + 2x_2 \le 20 && (1)\\ &x_1 \le 3.5 && (2)\\ &3x_1 + 4x_2 \le 24 && (3)\\ &\underline{x} \ge \underline{0}\end{aligned}$$

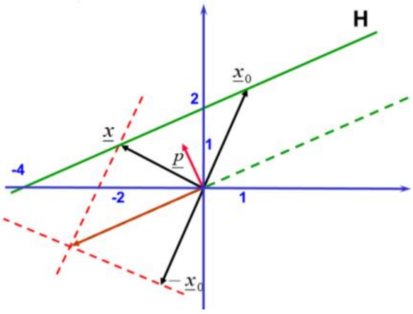


è come se fosse una retta a più dimensioni

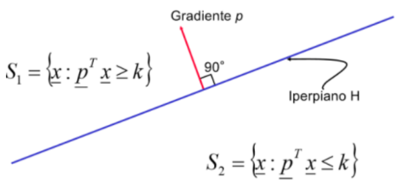
Def. Un insieme geometrico H è un *iperpiano* se e solo se $H = \{\underline{x} \mid \underline{p}^T\underline{x} = k\}$, o equivalentemente $H = \{\underline{x} \mid p_1x_1 + p_2x_2 + \ldots + p_nx_n = k\}$.
Informalmente, l'iperpiano H è il luogo geometrico dei punti $\underline{x}$ tali che il prodotto scalare tra i due vettori $\underline{p}^T$ e $\underline{x}$ è uguale ad uno scalare k.
Si noti che $\underline{p}^T\underline{x} = k$ è l'equazione della retta, dove $\underline{p}$ è un vettore e k è uno scalare. Il vettore $\underline{p}$ è il gradiente dell'iperpiano, il cui verso ne indica la direzione di crescita.

Consideriamo un punto $\underline{x}_0$ di H ed il gradiente $\underline{p}$. L'iperpiano H è l'insieme dei vettori $\underline{x}$ tali che il vettore $\underline{x} - \underline{x}_0$ è perpendicolare a $\underline{p}$. Supponiamo di avere due punti $\underline{x}, \underline{x}_0 \in H$, quindi $\underline{p}^T\underline{x} = k$ e $\underline{p}^T\underline{x}_0 = k$; sottraendo, otteniamo che $\underline{p}^T(\underline{x} - \underline{x}_0) = 0$. Se due vettori hanno prodotto interno nullo, allora sono perpendicolari.

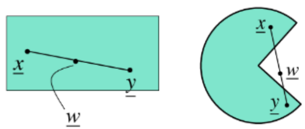
Es. Sia $H = \{(x_1, x_2) \mid p_1x_1 + p_2x_2 = k\} = -(1/2)x_1 + x_2 = 2$.
Sia $\underline{x}_0 = (1, 5/2)$ un punto di H, e verifichiamo che qualunque altro punto $\underline{x} \in H$ (ad esempio (-2, 1)) è tale che $\underline{x} - \underline{x}_0$ è perpendicolare a $\underline{p}$.
Disegniamo la retta dell'iperpiano, e tracciamo i due vettori considerati $\underline{x}$ e $\underline{x}_0$. Il gradiente è $\underline{p} = (-1/2, 1)$. Se si disegna $\underline{x} - \underline{x}_0$ (la bisettrice dei vettori $\underline{x}$ e $-\underline{x}_0$, in rosso), si nota che questo è perpendicolare a $\underline{p}$. Inoltre, questo vettore può essere rappresentato dalla diagonale minore (in quanto è risultato di una differenza, altrimenti maggiore) del parallelogramma ottenuto con $\underline{x}$ e $\underline{x}_0$.



Un iperpiano H divide lo spazio $\mathbb{R}^n$ cui appartiene in due semispazi, definiti come $S_1 = \{\underline{x} \mid \underline{p}^T\underline{x} \geq k\}$ e $S_2 = \{\underline{x} \mid \underline{p}^T\underline{x} \leq k\}$.



Def. Un insieme X è *convesso* se e solo se, dati due punti $\underline{x}, \underline{y} \in X$, ogni punto $\underline{w}$ generato come loro combinazione convessa ($\underline{w} = \lambda\underline{x} + (1 - \lambda)\underline{y}$, con $\lambda \in [0, 1]$) è tale che $\underline{w} \in X$.
Geometricamente, la combinazione convessa di due vettori in $\mathbb{R}^2$ corrisponde ai punti della retta che unisce i vertici dei due vettori.
Se un insieme X è convesso, la definizione ci dice che, presi due punti $\underline{x}$ e $\underline{y}$ di X e tracciato il segmento che li unisce, ogni punto del segmento si trova ancora all'interno dell'insieme X. Nella figura a destra, l'insieme X a sinistra è convesso, mentre quello a destra non lo è.



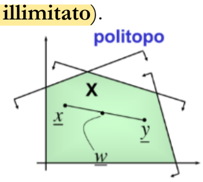
Per dimostrare che un insieme X è convesso, dobbiamo dimostrare (dalla def. precedente) che, presi due punti qualsiasi di X e tracciato il segmento che li unisce, ogni punto del segmento si trova ancora all'interno dell'insieme X.
- L'insieme $X = \{\underline{x} \mid A\underline{x} = \underline{b}\}$ è un insieme convesso.
DIM. Se $\underline{x}' \in X$, allora $A\underline{x}' = \underline{b}$. Analogamente, se $\underline{x}'' \in X$, allora $A\underline{x}'' = \underline{b}$.
Sia $\underline{w} = \lambda\underline{x}' + (1 - \lambda)\underline{x}''$, con $\lambda \in [0, 1]$. Allora $A\underline{w} = \lambda A\underline{x}' + (1 - \lambda)A\underline{x}'' = \lambda\underline{b} + (1 - \lambda)\underline{b} = \underline{b}$. □

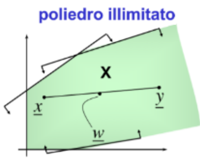
Commentato [DE8]: premoltiplicando per la matrice A

Commentato [DE9]: poiché $\underline{x}'$ e $\underline{x}''$ appartengono a X

- Un iperpiano è un insieme convesso.
- Un semispazio è un insieme convesso.
- Dalle due precedenti, si ricava che **l'intersezione** di iperpiani/semispazi produce un insieme convesso.

Def. Un *poliedro* è l'intersezione di un numero finito di semispazi.
Ciò implica che un poliedro X è un insieme convesso. Un poliedro può essere chiuso e limitato (**politopo**) o illimitato (**poliedro illimitato**).





Def. Una funzione $f(\underline{x})$ si dice *convessa* su insieme X se, presi comunque due punti $\underline{x}_1$, $\underline{x}_2 \in X$ risulta che:
$f(\lambda\underline{x}_1 + (1-\lambda)\underline{x}_2) \leq \lambda f(\underline{x}_1) + (1-\lambda)f(\underline{x}_2)$, con $\lambda \in [0,1]$.
Teorema (FUNZIONE CONVESSA)**.** *Una funzione lineare del tipo* $\underline{c}^T\underline{x}$ *è una funzione convessa.*
DIM. Dalla definizione di funzione convessa, sostituendo la $f(\underline{x})$ con $\underline{c}^T\underline{x}$ si ha:
- $f(\lambda\underline{x}_1 + (1-\lambda)\underline{x}_2) \rightarrow \underline{c}^T\lambda\underline{x}_1 + \underline{c}^T(1-\lambda)\underline{x}_2$;
- $\lambda f(\underline{x}_1) + (1-\lambda)f(\underline{x}_2) \rightarrow \lambda\underline{c}^T\underline{x}_1 + (1-\lambda)\underline{c}^T\underline{x}_2$.

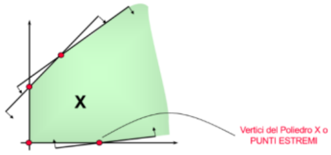
I due precedenti risultati sono uguali. Poiché $f\left(\lambda\underline{x}_1 + (1-\lambda)\underline{x}_2\right) = \lambda f(\underline{x}_1) + (1-\lambda)f(\underline{x}_2)$, la funzione $\underline{c}^T\underline{x}$ è convessa. □

Questo teorema è importante, in quanto il metodo del Simplesso può garantire l'ottimalità della soluzione trovata grazie ad esso, dato che esso si fermerà sull'ottimo locale individuale.

Punti e Direzioni estreme. Teorema della rappresentazione

Def. Un punto di un poliedro X è un *punto estremo* se e solo se non può essere espresso come combinazione convessa STRETTA di altri punti di X.

Geometricamente, questa definizione afferma che se prendiamo due punti qualsiasi in X, il vertice non si trova sulla retta che congiunge i due punti: $\underline{w} = \lambda\underline{x} + (1 - \lambda)\underline{y}$.



Teorema (PROPRIETÀ DEI PUNTI ESTREMI DI UN POLIEDRO LIMITATO). *Dato un poliedro X non vuoto e limitato con punti estremi $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$, ogni punto $\underline{y} \in X$ può essere espresso come combinazione convessa dei punti estremi di X, cioè*

$$\underline{y} = \sum_{j=1}^{k} \lambda_j \underline{x}_j \text{ , con } \sum_{j=1}^{k} \lambda_j = 1 \text{ e } \lambda_j \geq 0 \ \forall j=1, \ldots, k.$$

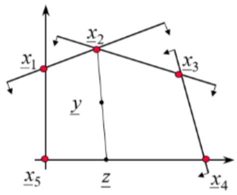
Es. Voglio esprimere il vettore $\underline{y}$ come combinazione convessa dei vertici del politopo a lato.

Scelto $\underline{y}$, per dimostrare ciò si traccia da uno qualsiasi dei vertici un segmento passante per $\underline{y}$; sappiamo, per definizione, che $\underline{y} = \lambda\underline{x}_2 + (1-\lambda)\underline{z}$, con $\lambda \in (0, 1)$.

Sappiamo che $\underline{z}$ non è un vertice del poliedro, perché è combinazione convessa di $\underline{x}_5$ e $\underline{x}_4$. Quindi $\underline{z} = \mu\underline{x}_5 + (1-\mu)\underline{x}_4$, con $\mu \in (0, 1)$.

Sostituendo $\underline{z}$ nell'equazione di $\underline{y}$, otteniamo che

$\underline{y} = \lambda\underline{x}_2 + (1-\lambda)(\mu\underline{x}_5 + (1-\mu)\underline{x}_4) = \lambda\underline{x}_2 + \mu(1-\lambda)\underline{x}_5 + (1-\lambda)(1-\mu)\underline{x}_4$.

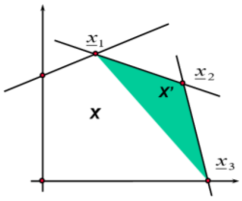


Si noti che:

1. $\lambda \geq 0 \qquad \mu(1-\lambda) \geq 0 \qquad (1-\lambda)(1-\mu) \geq 0$;
2. $\lambda + \mu(1-\lambda) + (1-\lambda)(1-\mu) = \lambda + \mu - \mu\lambda + 1 - \mu - \lambda + \lambda\mu = 1$.

Questo dimostra che il punto $\underline{y}$ è rappresentabile come combinazione convessa dei vertici del poliedro.

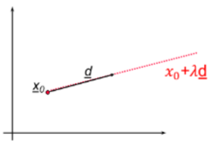
In generale, una combinazione convessa di $\underline{x}_1, \underline{x}_2, \underline{x}_3$ permette di ottenere tutti i punti di X' ⊂ X.



Nei problemi che affronteremo, spesso avremo a che fare con regioni illimitate. In questo caso, non si possono utilizzare i soli vertici, bensì bisogna considerare anche le sue *direzioni estreme.*

Def. Un *raggio* R di vertice $\underline{x}_0$ e di direzione $\underline{d}$ è un insieme di punti della forma $R = \{\underline{x}_0 + \lambda\underline{d} \mid \lambda \geq 0\}$.

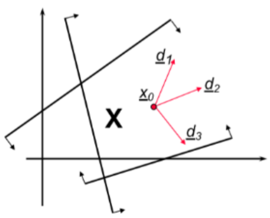
Informalmente, $\underline{x}_0 + \lambda\underline{d}$ è la semiretta che parte dal punto $\underline{x}_0$ e si sposta all'infinito lungo il vettore $\underline{d}$.



Def. Dato un poliedro X, il vettore $\underline{d}$ è una *direzione* di X se e solo se, per ogni punto $\underline{x}_0 \in X$, il raggio $\underline{x}_0 + \lambda\underline{d}$ (con $\lambda \geq 0$) appartiene a X.

Es. Dato il poliedro a lato, $\underline{d}_1$ e $\underline{d}_3$ non sono direzioni, mentre $\underline{d}_2$ è una direzione.

Ovviamente, è possibile avere dei vettori direzione solo se si ha una regione ammissibile come poliedro illimitato; per definizione, quindi, il politopo non ha direzione.



Di seguito è posto il procedimento algebrico per calcolare le direzioni di un poliedro.
Sia $X = \{\underline{x} \mid A\underline{x} \le \underline{b}, \underline{x} \ge \underline{0}\}$ un poliedro. Affinché il raggio $\underline{x} + \lambda\underline{d}$ appartenga a X, si può sostituire $\underline{x} + \lambda\underline{d}$ al vettore $\underline{x}$. Dato un qualsiasi punto $\underline{x} \in X$, il vettore $\underline{d}$ è una direzione del poliedro X se:

$$(i)\, A(\underline{x} + \lambda\underline{d}) \le \underline{b};$$
$$(ii)\quad \underline{x} + \lambda\underline{d} \ge \underline{0};$$
$$(iii)\qquad \underline{d} \ne \underline{0}.$$

(i) poiché $\underline{x} \in X$, $A\underline{x} + \lambda A\underline{d} \le \underline{b} \Leftrightarrow \lambda A\underline{d} \le \underline{0} \Leftrightarrow A\underline{d} \le \underline{0}$.
(ii) $\underline{x} + \lambda\underline{d} \ge \underline{0} \Leftrightarrow \underline{d} \ge \underline{0}$.
Quindi, le direzioni $\underline{d}$ del poliedro X sono tutti e soli i vettori tali che:

$$A\underline{d} \le \underline{0}$$
$$\underline{d} \ge \underline{0}$$
$$\underline{d} \ne \underline{0}. \qquad \text{Sistema omogeneo}$$

Ora, sia $X' = \{\underline{x} \mid A\underline{x} \ge \underline{b}, \underline{x} \ge \underline{0}\}$. Allora, le direzioni $\underline{d}'$ del poliedro X' sono tutti e soli i vettori tali che:

$$A\underline{d} \ge \underline{0}$$
$$\underline{d} \ge \underline{0}$$
$$\underline{d} \ne \underline{0}.$$

Quindi: se il vincolo del poliedro è di maggiore o uguale, va posto $A\underline{d} \ge \underline{0}$; se, invece, il vincolo del poliedro è di minore o uguale, va posto $A\underline{d} \le \underline{0}$. Questo risultato in quanto il raggio, al variare di λ all'infinito, deve comunque rimanere all'interno della regione, quindi deve soddisfare tutti i vincoli del poliedro.

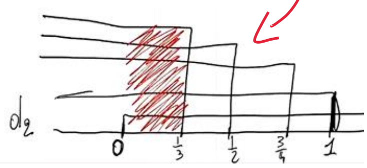
Es. Sia $\underline{X} = \begin{cases} -3x_1 + x_2 \le -2 \\ -x_1 + x_2 \le 2 \\ -x_1 + 2x_2 \le 8 \\ \underline{x} \ge \underline{0} \end{cases}$. Calcolare le direzioni del poliedro $\underline{X}$.

→ Siccome lavoriamo in $\mathbb{R}^2$, dobbiamo calcolare il vettore $\underline{d}' = (d_1, d_2)$. Una volta trovata la direzione del vettore, non ha importanza il valore del suo modulo. A questo scopo, "normalizziamo" il vettore, cioè ne dividiamo le componenti per la loro somma: il risultato è che si imporrà come condizione $d_1 + d_2 = 1$. Cioè, si sta chiedendo che, preso il vettore $\underline{d}'$, non ci interessa conoscere $2\underline{d}'$, $3\underline{d}'$, $4\underline{d}'$, …, bensì la sola direzione: prenderemo solo un vettore che rappresenta tutta la semiretta, cioè quello la cui somma delle componenti è uguale a 1. Quindi, poniamo il sistema:

$-3(1-d_2)+d_2 \le 0$

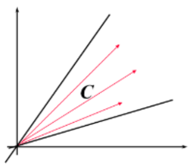
$$\begin{cases} -3d_1 + d_2 \le 0 \\ -d_1 + d_2 \le 0 \\ -d_1 + 2d_2 \le 0 \\ d_1 + d_2 = 1 \\ d_1, d_2 \ge 0 \end{cases} \Leftrightarrow \begin{cases} -3 + 3d_2 + d_2 \le 0 \\ -1 + d_2 + d_2 \le 0 \\ -1 + d_2 + 2d_2 \le 0 \\ d_1 = 1 - d_2 \\ d_1, d_2 \ge 0 \end{cases} \Leftrightarrow \begin{cases} 4d_2 \le 3 \\ 2d_2 \le 1 \\ 3d_2 \le 1 \\ d_1 = 1 - d_2 \\ d_1, d_2 \ge 0 \end{cases} \Leftrightarrow \begin{cases} d_2 \le 3/4 \\ d_2 \le 1/2 \\ d_2 \le 1/3 \\ d_1 = 1 - d_2 \\ d_1, d_2 \ge 0 \end{cases}$$

Siccome è un sistema, dobbiamo trovare i valori (in questo caso di d_2) soddisfatti per tutte le disequazioni. Inoltre, sapendo che lavoriamo su un intervallo di valori [0, 1], costruiamo la seguente tabella, da cui si ricava che $0 \le d_2 \le 1/3$.



Dunque, scegliamo il valore di d_2 e calcoliamo automaticamente il valore di d_1, in quanto abbiamo posto che $d_1 = 1 - d_2$. Ad esempio, possiamo usare i seguenti valori per $\underline{d}'$: $\underline{d}' = (1, 0)$, $\underline{d}' = (2/3, 1/3)$, …
Questi valori creano le infinite direzioni. *Siccome a noi interessano le sole direzioni estreme, queste si ottengono proprio prendendo gli estremi dell'intervallo calcolato*, cioè [0, 1/3]. Le precedenti $\underline{d}'$ e $\underline{d}'$ sono proprio le direzioni estreme del poliedro X.
Per esercizio, disegnare la regione ammissibile per X e verificare che $\underline{d}'$ e $\underline{d}''$ sono delle direzioni per X. □

Def. Un *cono convesso* C è un insieme convesso tale che se $\underline{x} \in C$ allora anche $\lambda\underline{x} \in C$ $\forall\lambda \ge 0$.
Un cono convesso è un insieme convesso che contiene raggi che partono dall'origine, in quanto possiamo scegliere $\lambda = 0$ per azzerare il vettore $\underline{x}$. Inoltre, possiamo scegliere $0 < \lambda < 1$ per diminuirne il modulo.
Nota. Alcuni raggi possono essere espressi come combinazione conica di altri raggi. Ciò significa, in generale, che un cono convesso può essere espresso in funzione dei suoi raggi.

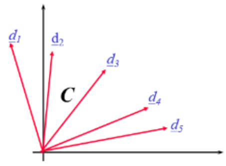


Commentato [DE10]: in quanto $A\underline{x} \le \underline{b}$
Commentato [DE11]: in quanto $\lambda \ge 0$
Commentato [DE12]: in quanto $A\underline{x} \le \underline{b}$
Commentato [DE13]: in quanto $A\underline{x} \le \underline{b}$
Commentato [DE14]: in quanto $d_1 = 1 - d_2$
Commentato [DE15]: in quanto $A\underline{x} \le \underline{b}$
Commentato [DE16]: in quanto $d_1 = 1 - d_2$
Commentato [DE17]: in quanto $A\underline{x} \le \underline{b}$
Commentato [DE18]: in quanto $d_1 = 1 - d_2$

Solo alcuni raggi sono sufficienti (detti *raggi estremi*), perché gli altri sono espressi come combinazione conica di questi.
Dato un insieme di vettori $\underline{d}_1, \underline{d}_2, \ldots, \underline{d}_k$ il **cono convesso** generato da questi vettori è dato da:

$$C = \left\{ \sum_{j=1}^{k} \lambda_j \underline{d}_j : \lambda_j \geq 0 \quad j = 1,2,\ldots,k \right\}.$$

Ad esempio, data la figura a lato, otteniamo che il cono C è tutto il cono compreso tra $\underline{d}_1$ e $\underline{d}_5$: queste sono le "direzioni estreme" del cono.



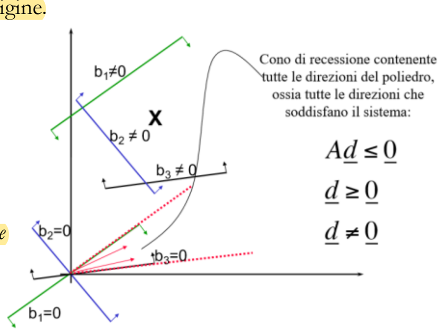
Def. Una direzione $\underline{d}$ di un poliedro X, è una *direzione estrema* di X se e solo se non è esprimibile come combinazione conica di altre direzioni di X.

Di seguito è posto il procedimento geometrico per calcolare le direzioni di un poliedro.
Sia $X = \{\underline{x} \mid A\underline{x} \leq \underline{b}, \underline{x} \geq \underline{0}\}$ un poliedro. Abbiamo visto che $\underline{d}$ è una direzione del poliedro se

$$A\underline{d} \leq \underline{0}$$
$$\underline{d} \geq \underline{0}$$
$$\underline{d} \neq \underline{0}.$$

Questo è un sistema omogeneo che definisce un cono poliedrico (detto *cono di recessione*) ottenuto traslando gli iperpiani che definiscono X parallelamente a sé stessi fino all'origine.

Es. Nella figura a destra abbiamo la regione ammissibile X. Per calcolare le direzioni estreme di questo poliedro, effettuiamo la traslazione di questi vincoli sull'origine degli assi, e ciò lo si fa semplicemente ponendo i termini noti uguali a 0.



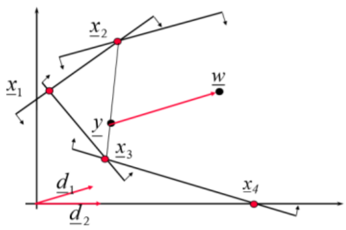
Teorema (DI RAPPRESENTAZIONE DI POLIEDRI).
Dato un poliedro X non vuoto con punti estremi $\underline{x}_i$ (con $i = 1, \ldots, k$) e direzioni estreme $\underline{d}_j$ (con $j = 1, \ldots, t$), ogni punto $\underline{x} \in X$ può essere espresso come combinazione convessa dei punti estremi di X e combinazione lineare non negativa (conica) delle sue direzioni estreme:

$$\underline{x} = \sum_{i=1}^{k} \lambda_i \underline{x}_i + \sum_{j=1}^{t} \mu_j \underline{d}_j$$

$$\sum_{i=1}^{k} \lambda_i = 1 \quad \lambda_i \geq 0 \quad i = 1,2,\ldots,k$$

$$\mu_j \geq 0 \quad j = 1,2,\ldots,t$$

Es. Si osservi la figura a lato. Se si considera il punto $\underline{w}$, per rappresentarlo si sfruttano le direzioni estreme. Prendiamo un punto $\underline{y}$ lungo il quale ci spostiamo sfruttando le direzioni estreme $\underline{d}_1$ e $\underline{d}_2$ (nell'esempio, scegliamo di usare la prima direzione): infatti, $\underline{w} = \underline{y} + \mu\underline{d}_1$, con $\mu \geq 0$. Se si prende il punto $\underline{y}$ sul segmento tra $\underline{x}_2$ e $\underline{x}_3$, si ottiene che $\underline{y} = \lambda\underline{x}_2 + (1 - \lambda)\underline{x}_3$.
Nulla ci vieta di spostarci lungo $\underline{y}$ usando la direzione estrema $\underline{d}_2$.



$\underline{y} = \lambda\underline{x}_2 + (1-\lambda)\underline{x}_3 \qquad \lambda \in (0,1)$

$\underline{w} = \underline{y} + \mu\underline{d}_1 \qquad \mu \geq 0$

quindi: $\underline{w} = \lambda\underline{x}_2 + (1-\lambda)\underline{x}_3 + \mu\underline{d}_1$

Dunque, qualunque sia il punto $\underline{w}$ della regione ammissibile, esso è raggiungibile utilizzando la combinazione convessa dei vertici del poliedro e la combinazione conica delle direzioni estreme.

Consideriamo il problema di Programmazione Lineare (PL) in forma standard:

$$\min z = \underline{c}^T\underline{x}$$
$$\text{s.t. } A\underline{x} = \underline{b}$$
$$\underline{x} \geq \underline{0}$$
$$\underline{x} \in \mathbb{R}^n.$$

Se si calcolano i vertici $\underline{x}_i$ ($i = 1, 2, \ldots, k$) e le direzioni estreme $\underline{d}_j$ ($j = 1, 2, \ldots, t$) del poliedro X, si è in grado di rappresentare questo sistema di vincoli con un modo alternativo, cioè usando il teorema della rappresentazione.
Ogni punto $\underline{x} \in X$ può essere espresso come combinazione convessa dei punti estremi di X e combinazione

conica delle sue direzioni estreme. Possiamo trasformare il problema di PL in un nuovo problema, con incognite: $\lambda_1, \lambda_2, \ldots \lambda_k$ e $\mu_1, \mu_2, \ldots, \mu_t$

$$\min \quad z = \sum_{i=1}^{k} (\underline{c}^T \underline{x}_i)\, \lambda_i \; + \sum_{j=1}^{t} (\underline{c}^T \underline{d}_j) \mu_j$$

$$\sum_{i=1}^{k} \lambda_i = 1 \quad \lambda_i \geq 0 \quad i = 1,2,\ldots,k \quad \mu_j \geq 0 \quad j = 1,2,\ldots,t$$

Infatti:

$$\min/\max \underline{c}^T \underline{x} = \min/\max z = \sum_{i=1}^{k} \lambda_i (\underline{c}^T \underline{x}_i) + \sum_{j=1}^{t} \mu_j (\underline{c}^T \underline{d}_j).$$

Commentato [DE19]: dal Teorema di rappresentazione di poliedri

Commentato [DE20]: posto $\sum_{i=1}^{k} \lambda_i = 1, \lambda_i \geq 0, \mu_j \geq 0$.

Considerando prima $\underline{c}^T \underline{d}_j$ (se stiamo *minimizzando*):
- per ogni j = 1, …, t, se esso è ≥ 0, allora $\mu_j = 0$, il che implica che cancelliamo la seconda sommatoria (cioè, non ha ottimo illimitato);
- se esiste < 0, allora $\mu_j = \infty$, il che implica che la soluzione ottima è Z* = -∞.
Se eliminiamo la seconda sommatoria, ci concentriamo sulla prima.
Ad esempio, se abbiamo $\min z = 5\lambda_1 + 3\lambda_2 + 6\lambda_3 + 12\lambda_4$, allora per ottenere la soluzione ottima poniamo $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 0, \lambda_4 = 0$; scegliamo, cioè, il minimo $\underline{c}^T \underline{x}_p$ e fissiamo $\lambda_p = 1$, in quanto il nostro obiettivo è minimizzare, rispettando i vincoli del problema (compreso che la somma dei λ_i dev'essere 1).

Considerando prima $\underline{c}^T \underline{d}_j$ (se stiamo *massimizzando*):
- se esso è > 0, allora $\mu_j = +\infty$, il che implica che la soluzione ottima è Z* = +∞;
- se esso è ≤ 0, allora $\mu_j = 0$, il che implica che cancelliamo la seconda sommatoria (cioè, non ha ottimo illimitato).
Se eliminiamo la seconda sommatoria, ci concentriamo sulla prima.

Riassumendo:
1. La soluzione ottima di un problema di minimo è finita se e solo se $\underline{c}^T \underline{d}_j \geq 0$, per ogni $\underline{d}_j$;
2. In questo caso, una soluzione ottima si trova su uno dei vertici del poliedro;
3. Se esistono più vertici ottimi, allora ogni combinazione convessa di questi punti è una soluzione ottima.

Esercizio. *Slide 23-27*.

Soluzioni di base ammissibili, teorema fondamentale della PL

Consideriamo un problema di Programmazione Lineare (PL) con *m* vincoli ed *n* variabili in forma standard:

$$\min z = \underline{c}^T\underline{x}$$
$$A\underline{x} = \underline{b} \quad (1)$$
$$\underline{x} \geq \underline{0} \quad (2)$$
$$\underline{x} \in \mathbb{R}^n.$$

Condizione: $\underline{b} \geq 0$

Poiché m = rango(A) ed m < n, si può partizionare A come **$A = [A_B \mid A_N]$**, dove:

- A_B è una matrice non singolare mxm ($\det(A_B) \neq 0$, per cui questa matrice è una base di $\mathbb{R}^m$);
- A_N è una matrice mx(n-m).

La matrice A_B è composta da m colonne linearmente indipendenti di A. Tali colonne (viste come vettori) sono quindi una base dello spazio vettoriale ad m dimensioni delle colonne di A. La matrice A_B è detta **matrice di base**.

In corrispondenza di una scelta di A_B ed A_N si può partizionare anche il vettore delle $\underline{x}$:

$$\underline{x} = \begin{bmatrix} \underline{x}_B \\ \underline{x}_N \end{bmatrix} \begin{matrix} \text{m componenti} \\ \text{n-m componenti} \end{matrix}$$

Ad esempio, se le colonne sono da 1 a 5 ed abbiamo 2 vincoli, allora la base ha dimensione 2x2, quindi se A_B ha gli indici 1, 5 e A_N ha indici 2, 3, 4 ciò significa che il vettore $\underline{x}$ partizionato avrà in $\underline{x}_B$ le componenti x_1 e x_5, mentre avrà in $\underline{x}_N$ le componenti x_2, x_3, x_4.

Il vettore $\underline{x}_B$ è detto **vettore delle variabili in base** (o "vettore di base").

Il vettore $\underline{x}_N$ è detto **vettore delle variabili fuori base**.

Il sistema di equazioni lineari $A\underline{x} = \underline{b}$ può essere riscritto come:

$$A_B\underline{x}_B + A_N\underline{x}_N = \underline{b} \rightarrow A_B\underline{x}_B = \underline{b} - A_N\underline{x}_N \rightarrow \underline{x}_B = A^{-1}{}_B\underline{b} - A^{-1}{}_BA_N\underline{x}_N.$$

Commentato [DE21]: dividendo tutto per A_B, e sappiamo che $1/A_B = A^{-1}{}_B$.

A questo punto, una soluzione del sistema di equazioni (1) corrisponde a determinare il valore per m variabili ($\underline{x}_B$) avendo fissato arbitrariamente il valore per le restanti n-m variabili ($\underline{x}_N$).

Es. Sia definito il sistema di equazioni $\begin{cases} x_1 + x_2 + x_3 + 2x_4 - 5x_5 = 5 \\ -x_1 + x_2 - 3x_3 + x_4 + 2x_5 = 3 \\ 6x_1 + 2x_2 + x_3 - x_4 + x_5 = 21 \end{cases}$. Dunque, la matrice è definita come

$$A = \begin{array}{c} \begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 1 & 2 & -5 \\ -1 & 1 & -3 & 1 & 2 \\ 6 & 2 & 1 & -1 & 1 \end{bmatrix} \end{array}$$

. Poiché questo è un sistema con 3 equazioni (cioè, vincoli) e 5 incognite (cioè, variabili), allora la dimensione della base è 3. Avremo come base una matrice quadrata 3x3: assumiamo che questa matrice sia

$$A_B = \begin{array}{c} \begin{matrix} x_1 & x_3 & x_5 \end{matrix} \\ \begin{bmatrix} 1 & 1 & -5 \\ -1 & -3 & 2 \\ 6 & 1 & 1 \end{bmatrix} \end{array}$$

(nel Simplesso avremo uno strumento che ci calcolerà questa base), da cui segue che

$\underline{x}_B = \begin{bmatrix} x_1 \\ x_3 \\ x_5 \end{bmatrix}$; inoltre, si ha che $A_N = \begin{array}{c} \begin{matrix} x_2 & x_4 \end{matrix} \\ \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 2 & -1 \end{bmatrix} \end{array}$, $\underline{x}_N = \begin{bmatrix} x_2 \\ x_4 \end{bmatrix}$. Infine, il vettore dei termini noti è $\underline{b} = \begin{bmatrix} 5 \\ 3 \\ 21 \end{bmatrix}$.

Dal precedente risultato, sappiamo che il precedente sistema può essere riscritto come $A_B\underline{x}_B + A_N\underline{x}_N = \underline{b}$.

Una scelta particolarmente importante è porre $\underline{x}_N = 0$ da cui si ottiene $\boxed{\underline{x} = \begin{bmatrix} \underline{x}_B \\ \underline{x}_N \end{bmatrix} = \begin{bmatrix} A^{-1}{}_B\underline{b} \\ \underline{0} \end{bmatrix}}$, che rappresenta una **soluzione di base**, associata ad una base di $\mathbb{R}^m$.

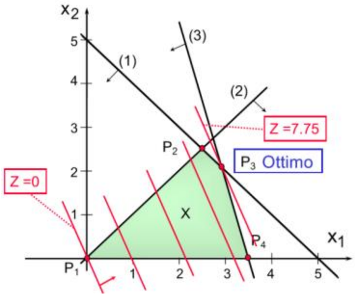
Se $\underline{x}_B = A^{-1}{}_B\underline{b} \geq \underline{0}$, allora si ottiene una soluzione di base **ammissibile**.

Le soluzioni di base sono importanti poiché vale il seguente teorema.

Teorema. *Dato* $X = \{\underline{x} \mid A\underline{x} = \underline{b}, \underline{x} \geq \underline{0}\}$ *insieme convesso, dove A è una matrice mxn di rango m (con m < n),* $\underline{x}_e$ *è un punto estremo di X se e solo se* $\underline{x}_e$ *è una soluzione di base ammissibile.*

Tra le infinite soluzioni del poliedro, si individua una soluzione di base ammissibile perché i vertici del poliedro corrispondono alle soluzioni di base ammissibili.

Es. Consideriamo il problema di PL posto a lato. Se si disegna la regione ammissibile, si ottiene il risultato seguente. Il punto P_3 è il punto di ottimo, dato che si vuole massimizzare, e sostituendo i valori di $P_3(x_1, x_2)$ nella funzione obiettivo si ottiene che il valore ottimo è Z = 7.75.



$$\begin{array}{lll} \max & z = 2x_1 + x_2 & \\ & x_1 + x_2 \le 5 & (1) \\ & -x_1 + x_2 \le 0 & (2) \\ & 6x_1 + 2x_2 \le 21 & (3) \\ & x_1 \ge 0\; x_2 \ge 0 & \end{array}$$

$$P_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad P_3 = \begin{bmatrix} 11/4 \\ 9/4 \end{bmatrix} \quad P_2 = \begin{bmatrix} 5/2 \\ 5/2 \end{bmatrix} \quad P_4 = \begin{bmatrix} 7/2 \\ 0 \end{bmatrix}$$

Per il teorema appena visto, ad ogni vertice del poliedro corrispondono *una o più* soluzioni di base ammissibili; viceversa, ad ogni soluzione di base ammissibile corrisponde *uno ed un solo* vertice del poliedro. Dunque, vogliamo sapere qual è la base associata ad ognuno di questi vertici. Ad esempio, sia B_i l'insieme degli indici delle variabili che formano la base associata al vertice P_i; per ora, affermiamo che $B_3 = \{1, 2, 4\}$, cioè le variabili x_1, x_2, x_4 formano una base associata al vertice P_3.
Prima di tutto, per ragionare **in termini delle basi**, bisogna trasformare il problema in forma standard. La dimensione della base per questo problema sarà 3, in quanto esso è costituito da *3 vincoli* e 5 variabili.

$$\begin{array}{l} \max\ z = 2x_1 + x_2 \\ x_1 + x_2 \le 5 \\ -x_1 + x_2 \le 0 \\ 6x_1 + 2x_2 \le 21 \\ x_1 \ge 0,\quad x_2 \ge 0 \end{array} \Rightarrow \begin{array}{l} -\min\ z = -2x_1 - x_2 \\ x_1 + x_2 + x_3 = 5 \\ -x_1 + x_2 + x_4 = 0 \\ 6x_1 + 2x_2 + x_5 = 21 \\ x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0, x_5 \ge 0 \end{array}$$

Il **massimo numero di basi possibili** corrisponde al numero di possibili estrazioni di *m* colonne su *n* colonne di A: $\binom{n}{m} = \frac{n!}{m!(n-m)!}$. In questo esempio, $\binom{5}{3} = 10$.
In generale, non tutte le possibili sottomatrici mxm sono non-singolari (quindi invertibili). Inoltre, non tutte le matrici di base danno luogo a soluzioni ammissibili (ossia, con tutte le componenti positive). Per questo motivo, il numero delle possibili combinazioni corrisponde ad un *limite superiore*.

Nell'esempio, solo 6 combinazioni danno luogo a basi ammissibili: infatti, presa la matrice A con 5 colonne (x_1, x_2, …, x_5), proviamo a selezionare in tutti i possibili modi tre di queste colonne; una volta selezionate, calcoliamo il determinante per verificare se esso è diverso da 0, e una volta stabilito che lo è (quindi, che lavoriamo su una base) calcoliamo le variabili in base per poter stabilire se si è trovata una soluzione di base ammissibile. Ovviamente, solo per scopi didattici (per ora) andremo a tentativi, ma esistono algoritmi che consentono di calcolare velocemente le informazioni necessarie.

$$A = \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 6 & 2 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

$$A_{B_1} = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ 6 & 2 & 0 \end{bmatrix} \end{array} \quad A_{B_2} = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_4 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \\ 6 & 2 & 0 \end{bmatrix} \end{array} \quad A_{B_3} = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_5 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 6 & 2 & 1 \end{bmatrix} \end{array} \quad \dots$$

Abbiamo detto che al punto P_3 corrispondeva la base $B_3 = \{1, 2, 4\}$. Infatti, come si nota nella figura a destra, preso il vettore $\underline{x}_{B2}$, presa la matrice A_{B2} e calcolata l'inversa, siccome $\underline{x}_B = A^{-1}{}_B\underline{b}$, fatti i vari calcoli otteniamo il vettore in basso a destra, che implica $x_1 = 11/4$ e $x_2 = 9/4$ (cioè, il punto P_3).

$$\underline{x}_{B_2} = \begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix} \quad A_{B_2} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \\ 6 & 2 & 0 \end{bmatrix} \quad A_{B_2}^{-1} = \begin{bmatrix} -1/2 & 0 & 1/4 \\ 3/2 & 0 & -1/4 \\ -2 & 1 & 1/2 \end{bmatrix}$$

$$\underline{x}_{B_2} = \begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix} = A_{B_2}^{-1}\underline{b} = \begin{bmatrix} -1/2 & 0 & 1/4 \\ 3/2 & 0 & -1/4 \\ -2 & 1 & 1/2 \end{bmatrix} \begin{bmatrix} 5 \\ 0 \\ 21 \end{bmatrix} = \begin{bmatrix} 11/4 \\ 9/4 \\ 1/2 \end{bmatrix} \Rightarrow P_3$$

Ora, notiamo che (in P_3) $x_3 = x_5 = 0$, per definizione di soluzione di base (cioè, le variabili fuori base sono uguali a 0). Come consiglio, durante gli esercizi è bene scrivere sulla rappresentazione grafica del vincolo (la retta) la variabile che vale 0 su quel vincolo. In pratica, la variabile da aggiungere sul vincolo dev'essere quella di slack/surplus aggiunta per ottenere il vincolo di

uguaglianza (che corrisponde alla retta del vincolo). In questo caso, sul vincolo (1) abbiamo $x_3 = 0$, sul vincolo (2) abbiamo $x_4 = 0$, sul vincolo (3) abbiamo $x_5 = 0$.
Allora, se consideriamo il punto P_3 per stabilire graficamente qual è la base B_3 bisogna prima stabilire la dimensione della base; ora, notiamo (si guardi il piano disegnato all'inizio della pagina precedente) che sul punto P_3 le variabili uguali a 0 sono x_3 e x_5 (cioè, *le variabili dei vincoli le cui rette passano per P_3*); inoltre, $B_2 = \{1, 2, 5\}$, $B_4 = \{1, 3, 4\}$.

Per il punto P_1 abbiamo, invece, più basi possibili: $B_5 = \{3, 4, 5\}$, $B_6 = \{1, 3, 5\}$, $B_7 = \{2, 3, 5\}$; infatti, per questo punto $x_1 = x_2 = x_4 = 0$. In queste soluzioni particolari (a cui sono associate più soluzioni di base ammissibili), dette **soluzioni degenere**, almeno una variabile $\underline{x}_{Bi}$ in base è uguale a 0. Questo è un problema, in quanto l'algoritmo del Simplesso potrebbe andare in loop infinito.

$$\underline{x}_{B_5} = \begin{bmatrix} x_1 \\ x_3 \\ x_5 \end{bmatrix} = A_{B_5}^{-1}\underline{b} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 6 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 0 \\ 21 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 21 \end{bmatrix} \Longrightarrow P_1$$

$$\underline{x}_{B_6} = \begin{bmatrix} x_2 \\ x_3 \\ x_5 \end{bmatrix} = A_{B_6}^{-1}\underline{b} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 0 \\ 21 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 21 \end{bmatrix} \Longrightarrow P_1$$

$$\underline{x}_{B_7} = \begin{bmatrix} x_3 \\ x_4 \\ x_5 \end{bmatrix} = A_{B_7}^{-1}\underline{b} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 0 \\ 21 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 21 \end{bmatrix} \Longrightarrow P_1$$

$$\underline{x} = \begin{array}{cc} \begin{bmatrix} 0 \\ 0 \\ 5 \\ 0 \\ 21 \end{bmatrix} & \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array} \end{array}$$

soluzione **degenere**

□

La ricerca delle soluzioni di un problema di PL si può effettuare esaminando solamente un numero finito di soluzioni corrispondenti alle soluzioni di base associate al poliedro dei vincoli.
A ciascuna matrice di base B (ammissibile) corrisponde una sola soluzione di base (ammissibile).
Viceversa, ad una soluzione di base (ammissibile) possono corrispondere più matrici di base. Questi casi sono associati a soluzioni dette *degeneri*, ovvero soluzioni per cui qualche componente del vettore di base $\underline{x}_B$ risulta nullo.

Dalla corrispondenza delle soluzioni di base ammissibili con i punti estremi del poliedro X deriva il seguente teorema.
Teorema (FONDAMENTALE DELLA PL). *Dato un problema di PL in forma standard, dove A è una matrice mxn con rango(A) = m ed m < n, allora:*
1. *esiste una soluzione ammissibile ⇔ esiste una soluzione ammissibile di base;*
2. *esiste una soluzione ottima finita ⇔ esiste una soluzione ottima finita che è anche di base.*

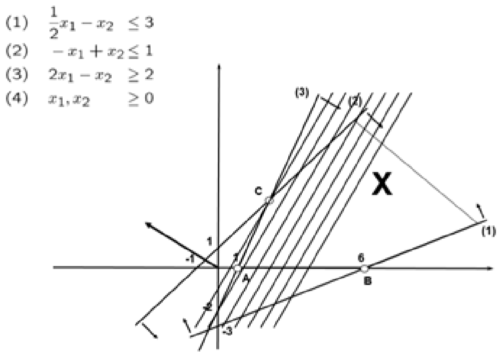
In particolare, il punto 2 afferma che se esiste all'interno della regione ammissibile un punto di ottimo, allora esisterà anche una soluzione di base ammissibile dove la funzione obiettivo assume il valore ottimo: algoritmicamente, mi devo concentrare *solo sulle soluzioni di base ammissibili del problema.*

Poiché il massimo numero di possibili basi di un problema di PL è finito, tali problemi hanno una struttura discreta.
I problemi di ottimizzazione corrispondenti alla selezione tra un numero finito di alternative si dicono problemi combinatorici.
La PL è quindi un problema combinatorico.

Un possibile algoritmo (*naive*) per determinare la soluzione ottima potrebbe consistere nella generazione esplicita di tutte le soluzioni ammissibili di base, quindi nella scelta di quella soluzione che rende massimo l'obiettivo.
Tale strategia non è conveniente poiché il numero massimo delle possibili basi cresce in maniera esponenziale col crescere delle dimensioni del problema(numero di variabili e vincoli).
Algoritmi che richiedono in generale un numero di passi che cresce in maniera esponenziale con le dimensioni del problema non sono efficienti.

Lezione 9 - Esercitazione

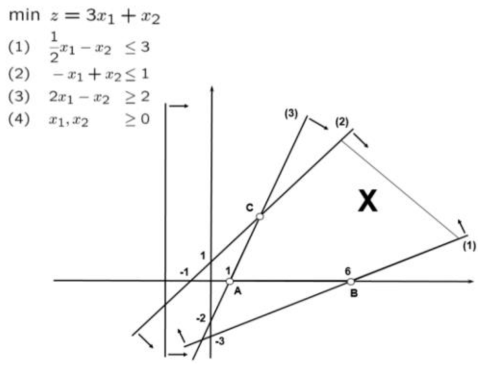
Es. Si osservi la figura a lato.
Dato tale problema di PL, si disegna la regione ammissibile.



d) Determinare una nuova funzione obiettivo che renda *C punto di ottimo unico*.
→ Per fare ciò, essa dovrebbe essere parallela al vincolo 2: poniamo $z = -x_1 + x_2$. Per stabilire se una retta è parallela ad un'altra, verifichiamo se i loro coefficienti angolari sono uguali ($ax_1 + bx_2 + c = 0 \Rightarrow m = -(a/b)$). Quindi il coefficiente angolare, per quanto riguarda il vincolo 2, è m = -(-1)/1 = 1, che è uguale a quello della funzione obiettivo definita all' inizio del punto a).
Questa funzione obiettivo, quindi, ci dà curve di livello parallele al vincolo 2. Ora, quando vogliamo una funzione obiettivo, bisogna anche sapere se stiamo massimizzando o massimizzando: come si ragiona per stabilire se qui abbiamo un minimo o un massimo? Si disegna il gradiente di questa funzione (-1, 1), e si nota che le curve di livello sono parallele al vincolo 2, quindi la funzione obiettivo dev'essere di massimo (infatti, tracciando le curve di livello, l'ultimo punto toccato è C). Se si scegliesse, invece, una funzione di minimo, otterremmo che l'ottimo sarebbe -∞.

La funzione $z = -x_1 + x_2$ non risponde all'esercizio, perché non è solo C il punto di ottimo, bensì *tutti* i punti sulla semiretta che parte da C e si sposta sul vincolo (2) all'infinito lo sono.
Data la retta $ax_1 + bx_2 + c = 0$, i coefficienti angolari dei vincoli (2) e (3) sono $m_2 = -\frac{-1}{1} = 1$ e $m_3 = -\frac{2}{-1} = 2$. Per risolvere l'esercizio, basta costruire una retta avente coefficiente angolare $1 < m < 2$: ad esempio, $z = -3x_1 + 2x_2$, in quanto il suo coefficiente angolare è 3/2.
Ora, usando il gradiente (-3, 2) stabiliamo se essa è una funzione di minimo o massimo: vale quest'ultimo. ✓

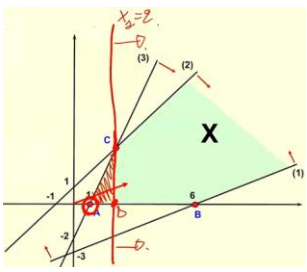
e) Data la figura a lato, aggiungere un vincolo *ridondante*.



→ $x_1 \geq -2$ è ridondante.
$x_1 \geq 1000$ non è ridondante.
In generale, per aggiungere un vincolo ridondante basta prendere un vincolo (anche con una variabile, ad esempio $x_1 \geq 0$) e cambiarne il termine noto, purché copra per intero la regione ammissibile del problema. Ad esempio: $2x_1 - x_2 \geq -4$. ✓

f) Data la figura a lato, aggiungere un vincolo che renda il sistema *inammissibile.*
→ $x_2 \leq -2$.
$2x_1 - x_2 \leq -4$.
In generale, per aggiungere un vincolo che renda il sistema inammissibile basta prendere un vincolo e cambiare il verso della disuguaglianza. ✓

g) Data la figura precedente, aggiungere un vincolo affinché il punto *C diventi punto di ottimo.*
→ $x_2 \geq 4$.
Bisogna fare particolare attenzione alla scelta dei vincoli sugli assi. Infatti, potrebbe venire spontaneo scegliere $x_1 \geq 2$; tuttavia, tracciando la retta per questo vincolo, otterremmo che alla regione ammissibile farebbe parte un nuovo punto, sull'asse x₁, che sarà punto di ottimo al posto di C (nella figura a destra, D). ✓

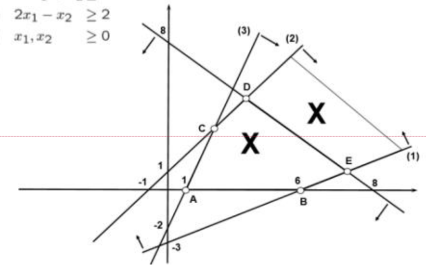


h) Aggiungere un vincolo che renda la regione ammissibile un *politopo*.

→ $x_1 + x_2 \leq 8$.

In questo caso, basta aggiungere una qualsiasi retta parallela all'asse x_2 passante per B: ad esempio, $x_1 \leq 6$ (se B ha come x_1 6). ✓

$$\min\ z = 3x_1 + x_2$$
$$\begin{array}{ll} (1) & \frac{1}{2}x_1 - x_2 \leq 3 \\ (2) & -x_1 + x_2 \leq 1 \\ (3) & 2x_1 - x_2 \geq 2 \\ (4) & x_1, x_2 \geq 0 \end{array}$$



i) Riscrivere il problema applicando il *teorema della rappresentazione* e risolverlo.

→ A = (1, 0), B = (6, 0). C = ?

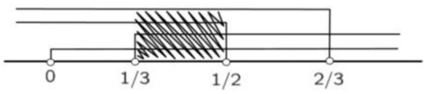
$$\begin{cases} -x_1 + x_2 & = 1 \\ 2x_1 - x_2 & = 2 \end{cases} \Rightarrow \begin{cases} -x_1 + x_2 & = 1 \\ x_2 & = 2x_1 - 2 \end{cases}$$

$$\begin{cases} -x_1 + 2x_1 - 2 & = 1 \\ x_2 & = 2x_1 - 2 \end{cases} \Rightarrow \begin{cases} x_1 & = 3 \\ x_2 & = 4 \end{cases} \quad \textbf{C= (3,4)}$$

Di conseguenza, C = (3, 4). Dato un poliedro $X = \{\underline{x}\,|A\underline{x} \leq \underline{b}, \underline{x} \geq \underline{0}\}$, abbiamo visto che $\underline{d}$ è una sua direzione se: $A\underline{d} \leq 0$, $\underline{d} \geq 0$, $\underline{d} \neq 0$. Di conseguenza:

$$\begin{cases} \frac{1}{2}d_1 - d_2 & \leq 0 \\ -d_1 + d_2 & \leq 0 \\ 2d_1 - d_2 & \geq 0 \\ d_1 + d_2 & = 1 \\ d_1 \geq 0, d_2 & \geq 0 \end{cases} \Rightarrow \begin{cases} \frac{1}{2}d_1 - d_2 & \leq 0 \\ -d_1 + d_2 & \leq 0 \\ 2d_1 - d_2 & \geq 0 \\ d_1 & = 1 - d_2 \\ d_1 \geq 0, d_2 & \geq 0 \end{cases} \Rightarrow \begin{cases} \frac{1}{2} - \frac{1}{2}d_2 - d_2 & \leq 0 \\ -1 + d_2 + d_2 & \leq 0 \\ 2 - 2d_2 - d_2 & \geq 0 \\ d_1 & = 1 - d_2 \\ d_1 \geq 0, d_2 & \geq 0 \end{cases}$$

$$\begin{cases} -\frac{3}{2}d_2 & \leq -\frac{1}{2} \\ d_2 & \leq \frac{1}{2} \\ -3d_2 & \geq -2 \\ d_1 & = 1 - d_2 \\ d_1 \geq 0, d_2 & \geq 0 \end{cases} \Rightarrow \begin{cases} d_2 & \geq 1/3 \\ d_2 & \leq 1/2 \\ d_2 & \leq 2/3 \\ d_1 & = 1 - d_2 \\ d_1 \geq 0, d_2 & \geq 0 \end{cases}$$



$$\Rightarrow \quad \underline{d}' = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} \quad \underline{d}'' = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

Di conseguenza, applicando il Teorema della rappresentazione, otteniamo che

$$\min\ z = \sum_{i=1}^{k} (\underline{c}^T \underline{x}_i)\lambda_i + \sum_{j=1}^{t} (\underline{c}^T \underline{d}_j)\mu_j$$

$$\sum_{i=1}^{k} \lambda_i = 1, \quad \lambda_i \geq 0 \quad i = 1,2,\dots,k$$

$$\mu_j \geq 0 \quad j = 1,2,\dots,t$$

$$\min\ z = \lambda_1 (3\ 1)\begin{pmatrix}1\\0\end{pmatrix} + \lambda_2 (3\ 1)\begin{pmatrix}6\\0\end{pmatrix} + \lambda_3 (3\ 1)\begin{pmatrix}3\\4\end{pmatrix} +$$

$$+ \mu_1 (3\ 1)\begin{pmatrix}2/3\\1/3\end{pmatrix} + \mu_2 (3\ 1)\begin{pmatrix}1/2\\1/2\end{pmatrix}$$

$$\min\ z = 3\lambda_1 + 18\lambda_2 + 13\lambda_3 + \frac{7}{3}\mu_1 + 2\mu_2$$

$$\lambda_1 + \lambda_2 + \lambda_3 = 1$$
$$\lambda_1, \lambda_2, \lambda_3 \geq 0$$
$$\mu_1, \mu_2 \geq 0$$

$$\min\ z = 3\lambda_1 + 18\lambda_2 + 13\lambda_3 + \cancel{\frac{7}{3}\mu_1 + 2\mu_2}$$
$$\lambda_1 + \lambda_2 + \lambda_3 = 1, \quad \lambda_1, \lambda_2, \lambda_3, \mu_1, \mu_2 \geq 0$$
$$\lambda_1 = 1,\ \lambda_2 = 0,\ \lambda_3 = 0\ \mu_1 = 0\ \mu_2 = 0$$
$$z = 3$$

Poiché i coefficienti 7/3 e 2 sono maggiori di 0, poniamo $\mu_1 = \mu_2 = 0$. Per minimizzare, poniamo $\lambda_1 = 1$, $\lambda_2 = 0$, $\lambda_3 = 0$. Di conseguenza, otteniamo come *valore ottimo* $z^* = 3$. ✓

Commentato [DE22]: Poiché i coefficienti 7/3 e 2 sono maggiori di 0, poniamo μ_1, $\mu_1 = 0$

j) Si determinino le *basi associate ad ogni vertice* della regione ammissibile, in base al problema di cui a destra.
→ Si ricordi il teorema secondo cui:
Dato $X = \{\underline{x} \mid A\underline{x} = \underline{b}, \underline{x} \geq \underline{0}\}$ *insieme convesso, dove A è una matrice mxn di rango m (con m < n),* $\underline{x}_e$ *è un punto estremo di X se e solo se* $\underline{x}_e$ *è una soluzione di base ammissibile.*

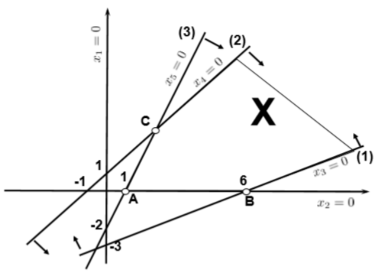
$$\min z = 3x_1 + x_2$$

$$\begin{array}{ll} (1) & \frac{1}{2}x_1 - x_2 \leq 3 \\ (2) & -x_1 + x_2 \leq 1 \\ (3) & 2x_1 - x_2 \geq 2 \\ (4) & x_1, x_2 \geq 0 \end{array}$$

Forma Standard

$$\min z = 3x_1 + x_2$$

$$\begin{array}{ll} (1) & \frac{1}{2}x_1 - x_2 + x_3 = 3 \\ (2) & -x_1 + x_2 + x_4 = 1 \\ (3) & 2x_1 - x_2 - x_5 = 2 \\ (4) & x_1, x_2, x_3, x_4, x_5 \geq 0 \end{array}$$



La prima informazione fondamentale è individuare la dimensione della base in questo problema: 3. Dal teorema precedente, sappiamo che ad ogni vertice della regione ammissibile sono associate una o più basi ammissibili. Il numero di componenti di queste basi è dato dal numero *m* di righe della matrice dei vincoli A. Per individuare la base associata ad un vertice *x* è sufficiente trovare le variabili che assumono il valore 0 sui vincoli la cui intersezione individua *x* sul piano.
Ad esempio, la base per il punto C è $B_C = \{1, 2, 3\}$, in quanto le variabili che assumono valore 0 sui vincoli la cui intersezione individua *x* sul piano sono x_4 e x_5. Analogamente: $B_A = \{1, 3, 4\}$, $B_B = \{1, 4, 5\}$. ✓

k) Si individui (geometricamente) una soluzione *ammissibile non basica* per il problema precedente.
→ Va bene qualsiasi punto della regione ammissibile ad eccezione dei punti estremi A, B, C. ✓

l) Si individui (geometricamente) una soluzione *ammissibile basica* per il problema precedente.
→ Per il teorema studiato, ad ogni vertice del poliedro corrispondono *una o più* soluzioni di base ammissibili; viceversa, ad ogni soluzione di base ammissibile corrisponde *uno ed un solo* vertice del poliedro. Di conseguenza, va bene uno qualsiasi dei punti estremi della regione ammissibile (A, B, C). ✓

m) Si individui (geometricamente) una soluzione *non ammissibile non basica* per il problema precedente.
→ Va bene un qualsiasi punto al di fuori della regione ammissibile diverso da quelli ottenuti dall'intersezione di due o più vincoli del problema. Ad esempio: (0, 3), (1, 5), (-3, 0), …
Preso (1, 5), ad esempio, viola i vincoli 3 e 2. ✓

n) Si individui (geometricamente) una soluzione *non ammissibile basica* per il problema precedente.
→ Va bene un qualsiasi punto al di fuori della regione ammissibile ottenuto dall'intersezione di due o più vincoli del problema. Ad esempio: (0, 0), (-1, 0), (0, 1), (0, -2), (0, -3), (-2/3, -10/3). ✓

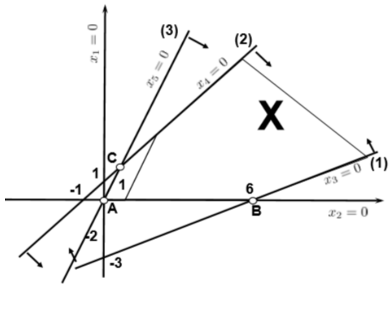
o) Si individui *una soluzione di base ammissibile degenere*, se esiste, per il problema precedente.
→ Bisogna individuare un punto estremo su cui passano almeno n – m + 1 vincoli. Questa condizione garantisce che almeno una variabile in base sia nulla. Nell'esempio in questione non ci sono soluzioni di base degeneri. ✓

p) Per il problema precedente, modificare il vincolo 3 al fine di generare una *soluzione di base ammissibile degenere*.
→ Effettuando una traslazione del vincolo 3 sull'origine degli assi (cioè, ponendo a 0 il termine noto del vincolo 3), cambiando la regione ammissibile, otterremmo che $B'_A = \{1, 3, 4\}$, $B''_A = \{2, 3, 4\}$, $B'''_A = \{3, 4, 5\}$.

Ora, considerando $B'''_A = \{3, 4, 5\}$:

$$A_{B_1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad x_{N_1} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$A_{B_1}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad x_{B_1} = \begin{bmatrix} x_3 \\ x_4 \\ x_5 \end{bmatrix} = A_{B_1}^{-1} b = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$



✓

q) Dal punto precedente, verificare algebricamente che $x_{B''_A}$ e $x_{B'_A}$ sono *soluzioni di base ammissibile degeneri.*
→

r) I seguenti vettori *sono soluzioni di base ammissibili* per il problema dato?

$$\underline{x}_1 = \begin{bmatrix} \frac{1}{4} \\ 0 \\ 2 \\ 4 \\ 1 \end{bmatrix}, \underline{x}_2 = \begin{bmatrix} 6 \\ 0 \\ 0 \\ 7 \\ 12 \end{bmatrix}, \underline{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \\ 1 \\ 0 \end{bmatrix}, \underline{x}_4 = \begin{bmatrix} \frac{1}{2} \\ 4 \\ -2 \\ 0 \\ 0 \end{bmatrix}.$$

$$\begin{aligned} &\min z = 3x_1 + x_2 \\ &(1) \quad \frac{1}{2}x_1 - x_2 + x_3 = 3 \\ &(2) \quad -x_1 + x_2 + x_4 = 1 \\ &(3) \quad 2x_1 - x_2 - x_5 = 0 \\ &(4) \quad x_1, x_2, x_3, x_4, x_5 \geq 0 \end{aligned}$$

→ Il *primo passo* da fare è contare il numero di 0 necessari almeno n – m;
il *secondo passo* è controllare per ogni vettore quali sono le variabili in base, e verificare nel sistema che le colonne corrispondenti agli indici delle variabili in base del vettore formino una base di $\mathbb{R}^m$; il *terzo passo* è controllare l'ammissibilità (o per sostituzione o calcolando $A^{-1}{}_B\underline{b}$).
Tenendo presente la definizione di soluzione di base ammissibile, possiamo escludere $\underline{x}_1$ perché c'è un solo 0 (in questo caso, sono necessari almeno due 0 per avere una soluzione di base).
Il resto è per esercizio.

Metodo del simplesso: condizioni di ottimalità e di illimitatezza

Introduciamo una serie di condizioni che sono gli step che verranno eseguiti ad ogni iterazione dell'algoritmo del simplesso, e che "risolveranno" i problemi dell'algoritmo naive per l'individuazione delle soluzioni ottime dei problemi di PL.

Calcolo della soluzione ottima di un problema di PL.
Consideriamo il problema di PL in Forma Standard

$$\min z = \underline{c}^T\underline{x}$$
$$A\underline{x} = \underline{b}$$
$$\underline{x} \geq \underline{0}$$
$$\underline{x} \in \mathbb{R}^n.$$

Condizione: $\underline{b} \geq 0$

Data una base B ammissibile, partizioniamo sia A come $\mathbf{A = [A_B \mid A_N]}$, sia il vettore delle $\underline{x}$ come $\underline{x} = \begin{bmatrix} \underline{x}_B \\ \underline{x}_N \end{bmatrix}$.

Il sistema di equazioni lineari $A\underline{x} = \underline{b}$ può essere riscritto come:

$$A_B\underline{x}_B + A_N\underline{x}_N = \underline{b} \rightarrow A_B\underline{x}_B = \underline{b} - A_N\underline{x}_N \rightarrow \underline{x}_B = A^{-1}{}_B\underline{b} - A^{-1}{}_B A_N\underline{x}_N.$$

Commentato [DE23]: dividendo tutto per A_B, e sappiamo che $1/A_B = A^{-1}{}_B$.

Ora, con un ragionamento analogo riscriviamo anche la funzione obiettivo come:

$$z = \underline{c}^T\underline{x} = \begin{bmatrix} \underline{c}^T{}_B & \underline{c}^T{}_N \end{bmatrix} \begin{bmatrix} \underline{x}_B \\ \underline{x}_N \end{bmatrix} = \underline{c}^T{}_B\underline{x}_B + \underline{c}^T{}_N\underline{x}_N \quad (\mathbf{1}).$$

Sostituiamo in (1) l'espressione delle variabili di base:

$$\underline{x}_B = A^{-1}{}_B\underline{b} - A^{-1}{}_B A_N\underline{x}_N \quad (\mathbf{2}),$$

ottenendo:

$$z = \underline{c}^T{}_B A^{-1}{}_B\underline{b} - \underline{c}^T{}_B A^{-1}{}_B A_N\underline{x}_N + \underline{c}^T{}_N\underline{x}_N \quad (\mathbf{3}).$$

Il **valore della funzione obiettivo** corrispondente alla base B è:

$$z = \underline{c}^T{}_B A^{-1}{}_B\underline{b},$$

in quanto per definizione $\underline{x}_N = 0$.

Le equazioni (2) e (3) esprimono rispettivamente i vincoli e la funzione obiettivo in funzione delle variabili fuori base.

$z = \underline{c}^T{}_B A^{-1}{}_B\underline{b} - \underline{c}^T{}_B A^{-1}{}_B A_N\underline{x}_N + \underline{c}^T{}_N\underline{x}_N$

$\underline{x}_B = A^{-1}{}_B\underline{b} - A^{-1}{}_B A_N\underline{x}_N$

Indichiamo con $\underline{\overline{\boldsymbol{b}}} = \boldsymbol{A^{-1}{}_B\underline{b}}$, e poiché $\boldsymbol{A_N\underline{x}_N} = \sum_{j\in N}\boldsymbol{\underline{a}_j x_j}$ otteniamo:

$$z = \underline{c}^T{}_B\underline{\overline{b}} - \sum_{j\in N}\underline{c}^T{}_B A^{-1}{}_B\underline{a}_j x_j + \sum_{j\in N} c_j x_j$$
$$\underline{x}_B = \underline{\overline{b}} - \sum_{j\in N} A^{-1}{}_B\underline{a}_j x_j\,,$$

dove $\underline{a}_j$ è la colonna di N che moltiplica la j-esima variabile fuori base.

Se $\boldsymbol{z_0} = \boldsymbol{\underline{c}^T{}_B\underline{\overline{b}}}$ e $\boldsymbol{z_j} = \boldsymbol{\underline{c}^T{}_B A^{-1}{}_B\underline{a}_j}$, allora la funzione obiettivo diventa:

$$z = z_0 - \sum_{j\in N} z_j x_j + \sum_{j\in N} c_j x_j = z_0 - \sum_{j\in N}(z_j - c_j)x_j\,.$$

I coefficienti $(z_j - c_j)$ vengono detti **coefficienti di costo ridotto** perché sono i coefficienti di costo associati alle sole variabili fuori base; vedremo che essi sono fondamentali per verificare se la base è ottima o meno.

Infine, fissato $\boldsymbol{\underline{y}_j = A^{-1}{}_B\underline{a}_j}$ i vincoli diventano:

$$\underline{x}_B = \underline{\overline{b}} - \sum_{j\in N} A^{-1}{}_B\underline{a}_j x_j = \underline{x}_B = \underline{\overline{b}} - \sum_{j\in N}\underline{y}_j x_j.$$

Forma canonica in funzione di una base B.

Consideriamo il problema di PL in Forma Standard

$$\begin{aligned} \min z &= \underline{c}^T\underline{x} \\ A\underline{x} &= \underline{b} \\ \underline{x} &\geq \underline{0} \\ \underline{x} &\in \mathbb{R}^n. \end{aligned}$$

Condizione: $\underline{b} \geq 0$

Data una base B ammissibile, riscriviamo il problema in funzione di B come segue:

$$\min z = z_0 - \sum_{j \in N} (z_j - c_j)x_j$$

$$\underline{x}_B = \overline{\underline{b}} - \sum_{j \in N} \underline{y}_j x_j$$

$$\underline{x} \geq \underline{0}.$$

$$z_0 = \underline{c}_B^T A_B^{-1} \underline{b}$$

$$z_j = \underline{c}_B^T A_B^{-1} \underline{a}_j$$

$$\overline{\underline{b}} = A_B^{-1} \underline{b}$$

$$\underline{y}_j = A_B^{-1} \underline{a}_j$$

Siccome $x_j = 0$ per definizione, abbiamo che $\min z = z_0$ e $\underline{x}_B = \overline{\underline{b}}$ (dove, si ricordi, $z_0 = \underline{c}^T{}_B \overline{\underline{b}}$ e $\overline{\underline{b}} = A^{-1}{}_B \underline{b}$).

Ora, scelta una variabile fuori base $x_j > 0$, con $j \in N$, e la faccio incrementare, la funzione obiettivo *cambia*? Sì, e al crescere di x_j la funzione obiettivo cresce, decresce o non varia in base al segno di $(z_j - c_j)$:

- se $(z_j - c_j) < 0$, allora il valore della funzione obiettivo cresce al crescere di x_j;
- se $(z_j - c_j) = 0$, allora il valore della funzione obiettivo non cambia;
- se $(z_j - c_j) > 0$, allora il valore della funzione obiettivo diminuisce al crescere di x_j.

Osservazione. In un problema di minimo, $(z_j - c_j) \leq 0 \; \forall j \in \mathrm{N} \Rightarrow$ *base ottima.*

Se calcoliamo i coefficienti di costo ridotto $(z_j - c_j)$ e verifichiamo che sono tutti ≤ 0, allora incrementare una qualsiasi variabile fuori base è dannoso, in quanto può o lasciare la funzione obiettivo così com'è o peggiorarla: di conseguenza, il valore minimo è z_0, ed abbiamo trovato la base ottima.

Osservazione. In un problema di minimo, $\exists k \mid (z_k - c_k) > 0 \Rightarrow$ *base non ottima.*

Da questo risultato, capito che conviene incrementare x_k, ma siccome ciò influisce sui vincoli del problema (cioè $\underline{x}_B = \overline{\underline{b}} - \sum_{j \in N} \underline{y}_j x_j$) bisogna incrementarla *al massimo* ma purché messa a sistema con i vincoli $\underline{x}_B$.

Analogamente, per un problema di massimo ricaviamo la seguente osservazione.

Osservazione. In un problema di massimo, $(z_j - c_j) \geq 0 \; \forall j \in \mathrm{N} \Rightarrow$ *base ottima.*

Verifichiamo se la soluzione di Base corrente è ottima o può essere migliorata.

Consideriamo l'obiettivo $\min z = z_0 - \sum_{j \in N}(z_j - c_j)x_j$. Supponiamo che esista un coefficiente $k \in \mathrm{N}$ tale che $z_j - c_j > 0$, e consideriamo come varia l'obiettivo facendo diventare positiva la variabile fuori base $\mathbf{x_k}$, attualmente nulla.

$$z = z_0 - \overbrace{(z_k - c_k)}^{>0} \underbrace{x_k}_{>0}$$

L'obiettivo migliora !

Teorema (CONDIZIONE DI OTTIMALITÀ). *Una soluzione di base non degenere di un problema di PL è ottima se e solo se:*

1) $\overline{b}_i \geq 0$ (ammissibile);
2) $(z_j - c_j) \leq 0 \; \forall j \in \mathrm{N}$ (non migliorabile). N sono le variabili fuori base

È possibile iterare il procedimento fino a che esiste qualche variabile fuori base che può migliorare l'obiettivo se portata in base.

Nel caso di soluzione degenere possono esistere soluzioni ottime in cui il punto 2) del teorema precedente non è soddisfatto. Tuttavia, se un problema ammette soluzione ottima finita allora ammette una soluzione di base ottima che soddisfa le condizioni 1) e 2).

Si osservi che, nel calcolo dei z_j, l'unico termine a variare è $\underline{a}_j$, per cui è consigliabile sempre calcolare prima $\underline{c}^T{}_B A^{-1}{}_B$ per poterla riciclare successivamente nei calcoli.

Questo teorema è importante perchè se vogliamo verificare se una soluzione di base è ottima oppure no, possiamo farlo.

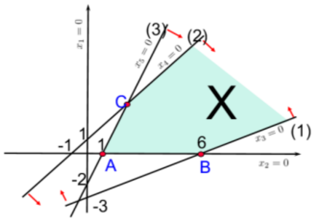
Calcoliamo Ab^-1
Calcoliamo il valore delle variabili xb.
Se sono tutti > 0, significa che la soluzione è ammissibile.

Per verificare se non può essere migliorata calcolo i coefficienti di costo ridotto. Se anch'essi sono <= 0, allora la soluzione di base è ottima e ammissibile.

é importante perchè senza il teorema (per verificare se una soluzione di base è ottima) l'unico modo per vedere se il punto è ottimo sarebbe quello di controllare il valore obiettivo in tutti i punti estremi

Es. Si consideri la figura a lato.
La base associata al punto A è B_A = {1, 3, 4}. Da ciò consegue che N = {2, 5}.
Ora, *verificare algebricamente che B_A è la base ottima che stiamo cercando.*
→Prima di tutto, trasformiamo il problema in forma standard. Per il teorema precedente, si vuole verificare la 2). Ricordando che $z_j = \underline{c}^T{}_B A^{-1}{}_B \underline{a}_j$, procediamo: abbiamo la matrice di base $A_{B_A} = \begin{bmatrix} \frac{1}{2} & 1 & 0 \\ -1 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix} \rightarrow A^{-1}{}_{B_A} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 1 & 0 & -\frac{1}{4} \\ 0 & 1 & \frac{1}{2} \end{bmatrix}$.

$\min z = 3x_1 + x_2$
(1) $\frac{1}{2}x_1 - x_2 \leq 3$
(2) $-x_1 + x_2 \leq 1$
(3) $2x_1 - x_2 \geq 2$
(4) $x_1, x_2 \geq 0$



Forma Standard
$\min z = 3x_1 + x_2$
(1) $\frac{1}{2}x_1 - x_2 + x_3 = 3$
(2) $-x_1 + x_2 + x_4 = 1$
(3) $2x_1 - x_2 - x_5 = 2$
(4) $x_1, x_2, x_3, x_4, x_5 \geq 0$

Ora, verifichiamo la 2):

- $z_2 - c_2 = \left([3 \;\; 0 \;\; 0] \cdot A^{-1}{}_{B_A} \cdot \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \right) - (1) = \begin{bmatrix} 0 & 0 & \frac{3}{2} \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} - 1 = -\frac{3}{2} - 1 = -\frac{5}{2}$. Siccome $z_2 - c_2 \leq 0$, allora i coefficienti di costo ridotto, associati alla variabile x_2, è negativa, e siccome si sta minimizzando non conviene incrementare la variabile x_2 fuori base, perché si andrebbe a peggiorare il valore della funzione obiettivo;
- $z_5 - c_5 = \begin{bmatrix} 0 & 0 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} - 0 = -\frac{3}{2}$. Siccome $(z_j - c_j) \leq 0 \; \forall j \in$ N, allora abbiamo verificato che B_A = {1, 3, 4} è la base ottima. ✓

Poiché incrementando il valore della variabile fuori base x_k il valore della funzione obiettivo migliora, si potrebbe pensare di aumentare indefinitivamente x_k.
Tuttavia, aumentando x_k anche le equazioni (5) corrispondenti ai vincoli variano, modificando i valori delle variabili di base:

$$\underline{x}_B = \underline{\overline{b}} - \sum_{j \in N} \underline{y}_j x_j \quad (5)$$
$$\underline{x} \geq \underline{0}.$$

Dal momento che per j ∈ N le x_j sono uguali a zero per j ≠ k, la relazione (5) diventa:

$$\underline{x}_B = \underline{\overline{b}} - \underline{y}_k x_k.$$

Deciso che bisogna far crescere la sola variabile x_k con un valore maggiore di 0, se la si fa crescere allora le componenti del vettore $\underline{x}_B$ cambiano – in forma vettoriale – come segue (ricordando che $\underline{y}_k = A^{-1}{}_B \underline{a}_k$):

$$\begin{bmatrix} x_{B_1} \\ x_{B_2} \\ x_{B_3} \\ \vdots \\ x_{B_r} \\ \vdots \\ x_{B_m} \end{bmatrix} = \begin{bmatrix} \overline{b}_1 \\ \overline{b}_2 \\ \overline{b}_3 \\ \vdots \\ \overline{b}_r \\ \vdots \\ \overline{b}_m \end{bmatrix} - \begin{bmatrix} y_{1k} \\ y_{2k} \\ y_{3k} \\ \vdots \\ y_{rk} \\ \vdots \\ y_{mk} \end{bmatrix} x_k$$

- se $\boldsymbol{y_{ik} \leq 0} \; \forall \boldsymbol{i} \in \boldsymbol{B}$, allora x_{B_i} cresce al crescere di x_k, e così x_{B_i} continua a essere non negativo (**ottimo illimitato**);
- se $\exists \boldsymbol{i} : \; \boldsymbol{y_{ik} > 0}$, allora x_{B_i} decresce al crescere di x_k. Il valore di x_k verrà incrementato finché una delle variabili in base assumerà valore zero. Infatti, noi vogliamo che

$$x_{B_1} = \overline{b}_1 - y_{1k} x_k \geq 0$$
$$x_{B_2} = \overline{b}_2 - y_{2k} x_k \geq 0$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$x_{B_m} = \overline{b}_m - y_{mk} x_k \geq 0,$$

ed in questo caso la variabile x_{B_i} che si azzererà per prima verrà tolta dalle variabili di base e sarà rimpiazzata dalla variabile x_k.

Possiamo scrivere

$$x_{B_1} = \overline{b}_1 - y_{1k}x_k \geq 0 \Leftrightarrow sempre \text{ (siccome } y_{1k} \leq 0)$$
$$x_{B_2} = \overline{b}_2 - y_{2k}x_k \geq 0 \Leftrightarrow x_k \leq \frac{\overline{b}_2}{y_{2k}} \text{ (siccome } y_{2k} > 0)$$
$$\vdots$$
$$x_{B_m} = \overline{b}_m - y_{mk}x_k \geq 0 \Leftrightarrow x_k \leq \frac{\overline{b}_m}{y_{mk}} \text{ (siccome } y_{mk} > 0).$$

Dobbiamo considerare solo i rapporti in cui $y_{ik} > 0$.
Quindi, considerando quei rapporti in cui $y_{jk} > 0$ la variabile x_k assumerà il seguente valore (**test dei minimi rapporti**):

$$x_k = \frac{\overline{b}_r}{y_{rk}} = \min_{i \in B} \left\{ \frac{\overline{b}_i}{y_{ik}} : y_{ik} > 0 \right\}.$$

Così:

$$x_{B_1} = \overline{b}_1 - y_{1k}x_k \geq 0 \Leftrightarrow \overline{b}_1 - y_{1k}\frac{\overline{b}_r}{y_{rk}} \geq 0$$
$$\vdots$$
$$x_{B_r} = \overline{b}_r - y_{rk}x_k \geq 0 \Leftrightarrow \overline{b}_r - y_{rk}\frac{\overline{b}_r}{y_{rk}} = 0$$
$$\vdots$$
$$x_{B_m} = \overline{b}_m - y_{mk}x_k \geq 0 \Leftrightarrow \overline{b}_m - y_{mk}\frac{\overline{b}_r}{y_{rk}} \geq 0.$$

Fare assumere a x_k un valore positivo significa portare la variabile x_k in base.
Nello stesso tempo il valore delle altre variabili di base per cui $y_{ik} > 0$ diminuisce.
Il valore che x_k assume in base è quello corrispondente all'annullamento della prima variabile di base, cioè

$$x_k = \frac{\overline{b}_r}{y_{rk}} = \min_{i \in B} \left\{ \frac{\overline{b}_i}{y_{ik}} : y_{ik} > 0 \right\}.$$

La variabile x_k entra in base, con tale valore, mentre la variabile x_{Br} esce dalla base.
Il coefficiente y_{rk} è detto **Pivot** (l'aggiornamento della base si dice **Pivoting**), e viene usato per aggiornare i valori delle variabili in base dopo l'ingresso in base di x_k. In sintesi, la *nuova soluzione di base è*

$$\boxed{\begin{aligned} x_{B_i} &= \overline{b}_i - y_{ik}\frac{\overline{b}_r}{y_{rk}} \\ x_k &= \frac{\overline{b}_r}{y_{rk}} \end{aligned}}.$$

Inoltre, *le nuove variabili fuori base sono* $x_j = 0 \ \forall j \in N'$, con $N' = \{B_r\} \cup N - \{k\}, x_{B_r} = 0$.

Con il cambio delle variabili in base, la nuova matrice di base risulta composta delle stesse colonne della vecchia base ad eccezione del fatto che la colonna associata a x_{B_r} è stata sostituita dalla colonna associata a x_k.

La nuova soluzione di base ha migliorato il valore della funzione obiettivo *strettamente*:

$$z = z_0 - (z_k - c_k)\frac{\overline{b}_r}{y_{rk}} < z_0,$$

in quanto $(z_k - c_k) > 0$ e, complessivamente, $(z_k - c_k)\frac{\overline{b}_r}{y_{rk}} > 0$.

Pertanto, il metodo del Simplesso si sposta da una base ad una ad essa adiacente se e solo se la funzione obiettivo migliora strettamente, *tranne in un caso*: quello in cui $\overline{b}_r = 0$, dove si ha una soluzione degenere, che (vedremo) può creare problemi al Simplesso.

Metodo del simplesso: passi ed esempio

Consideriamo il problema di PL in Forma Standard

$$\begin{aligned} \min z = \underline{c}^T\underline{x} \\ A\underline{x} = \underline{b} \\ \underline{x} \geq \underline{0} \\ \underline{x} \in \mathbb{R}^n. \end{aligned}$$

Condizione: $\underline{b} \geq 0$

Data una base B ammissibile, riscriviamo il problema in funzione di B come segue:

$$\min z = z_0 - \sum_{j \in N} (z_j - c_j) x_j$$

$$\underline{x}_B = \underline{\overline{b}} - \sum_{j \in N} \underline{y}_j x_j$$

$$\underline{x} \geq \underline{0},$$

dove: $z_0 = \underline{c}^T{}_B \underline{\overline{b}}$, $z_j = \underline{c}^T{}_B A^{-1}{}_B \underline{a}_j$, $\underline{\overline{b}} = A^{-1}{}_B \underline{b}$, $\underline{y}_j = A^{-1}{}_B \underline{a}_j$.

Abbiamo, poi, introdotto il seguente teorema.
Teorema (CONDIZIONE DI OTTIMALITÀ). *Una soluzione di base non degenere di un problema di PL è ottima se e solo se:*

1) $\overline{b}_i \geq 0$ (ammissibile);
2) $(z_j - c_j) \leq 0 \ \forall j \in \mathrm{N}$ (non migliorabile).

<u>Scelta della variabile entrante.</u>

Quando le condizioni di ottimalità non sono verificate è sempre possibile scegliere una variabile fuori base x_k da portare in base per migliorare l'obiettivo.
Quando esistono più alternative la scelta non preclude il raggiungimento della soluzione ottima, ma può al peggio aumentare il tempo necessario per la sua ricerca.

Informalmente, cosa succede se ci si ritrova con $z_j - c_j > 0$ e $z_r - c_r > 0$, con k, r ∈ N? Cioè, se esistono più variabili il cui incremento *migliora* la funzione obiettivo, quale bisogna scegliere?
Il **metodo del gradiente** sceglie la variabile fuori base x_k che localmente fa aumentare più rapidamente l'obiettivo:

$x_k - c_k$

$$\boxed{\cancel{z_j - c_j}} = \max_{j \in N} \{z_j - c_j\}.$$

Determinata la variabile fuori base x_k da portare in base, si deve scegliere la variabile uscente. Esistono due alternative:

$$\begin{bmatrix} x_{B_1} \\ x_{B_2} \\ x_{B_3} \\ \vdots \\ x_{B_r} \\ \vdots \\ x_{B_m} \end{bmatrix} = \begin{bmatrix} \overline{b}_1 \\ \overline{b}_2 \\ \overline{b}_3 \\ \vdots \\ \overline{b}_r \\ \vdots \\ \overline{b}_m \end{bmatrix} - \begin{bmatrix} y_{1k} \\ y_{2k} \\ y_{3k} \\ \vdots \\ y_{rk} \\ \vdots \\ y_{mk} \end{bmatrix} x_k$$

a) $y_{ik} \leq 0 \ \forall i = 1, \dots, m$, ed osservando il sistema di equazioni precedente si nota che **la soluzione del problema è illimitata** (non esiste un punto di ottimo). In questo caso facendo aumentare x_k il valore di nessuna variabile di base diminuisce:
$z = z_0 - (z_k - c_k)x_k \leq z_0$, per qualsiasi valore di x_k;
b) $y_{rk} > 0 \ per\ almeno\ un\ r$, ed in questo caso la soluzione di base non è ottima, e bisogna quindi passare alla base successiva.

Nota. Se una variabile di base x_{Bi} è nulla (soluzione degenere) e $y_{ik} > 0$, allora x_k entra in base con valore nullo. In questo caso la soluzione non cambia, ed in particolare rimane degenere. Per questa ragione la ricerca della soluzione potrebbe rimanere bloccata generando sempre la medesima soluzione (questo fatto è detto **cycling**). Il cycling è piuttosto raro e comunque esistono strategie per evitarlo.

Algoritmo del Simplesso (Pseudocodice)

INPUT: Problema di PL (in forma standard) e una soluzione di base ammissibile.

Commentato [DE24]: Un errore comune è cercare casualmente una soluzione di base ammissibile del problema. Esiste, infatti, un algoritmo che viene in aiuto al Simplesso che calcolerà la soluzione di base di partenza e la fornirà al simplesso.

1. **Test di ottimalità**:
   Se $z_j - c_j \leq 0\ \forall j \in N$, allora la soluzione corrente è ottima e l'algoritmo termina.
   Altrimenti andare al passo 2.
2. **Scelta della variabile entrante in base**:
   Scegliere una variabile fuori base x_k tale che $z_k - c_k > 0$ ed andare al passo 3.
3. **Test di illimitatezza**:
   Se $y_{ik} \leq 0\ \forall i = 1, \dots m$, allora la soluzione del problema è illimitata (non esiste ottimo finito), e l'algoritmo termina.
   Altrimenti vai al Passo 4.
4. **Scelta della variabile uscente dalla base (Test dei minimi rapporti)**:
   Scegliere la variabile x_r tale che $\frac{\bar{b}_r}{y_{rk}} = \min_{i \in B} \left\{ \frac{\bar{b}_i}{y_{ik}} : y_{ik} > 0 \right\}$.
   x_r è la variabile uscente e la variabile entrante x_k assume valore pari a $\frac{\bar{b}_r}{y_{rk}}$.
5. **Aggiornamento della base:**
   Aggiornare gli indici delle variabili in base (B) e quelli delle variabili fuori base (N). Tornare al passo 1.

L'algoritmo Naive per la ricerca della soluzione ottima presenta i seguenti tre problemi:
1) inefficienza (molto lento per valori grandi), in quanto di forza bruta;
2) se la regione ammissibile è vuota, esso prova a costruire in tutti i possibili modi una soluzione di base ammissibile, e soltanto alla fine poteva stabilirlo (grazie al Teorema fondamentale);
3) non è in grado di capire quando la soluzione ottima ha valore illimitato. Infatti, essendo un algoritmo di forza bruta trovava la migliore soluzione di base ammissibile del problema.

In base allo pseudocodice dell'algoritmo del Simplesso, si nota che esso ha risolto i problemi dell'ottimo illimitato (3) e dell'ottimalità di una soluzione senza provarle tutte (1). Vedremo che il problema della regione ammissibile vuota (2) viene risolto dall'algoritmo che calcola la soluzione di base di partenza, che verrà accostato al Simplesso: se questo algoritmo non è in grado di trovarla, allora la regione ammissibile è vuota.

La complessità dell'algoritmo del Simplesso, nel *caso peggiore* (cioè, si parte da un vertice del poliedro e soltanto l'ultimo vertice è la soluzione ottima), è **esponenziale**. Tuttavia, anche se in linea teorica la complessità è esponenziale, il *caso medio* afferma che l'algoritmo è molto veloce (solitamente lineare rispetto al numero di vincoli del problema). Fino al 1972, non esistevano problemi non risolvibili in modo inefficiente dal Simplesso: in quell'anno fu creato il "Cubo" (cioè, un problema dove la regione ammissibile aveva un numero di vertici esponenziale rispetto al numero di variabili), che lo mette in difficoltà soltanto teoricamente.

Es (Applicazione del Simplesso)**.** Dato il problema di PL posto a destra, applichiamo il metodo del Simplesso. Prima di tutto, trasformiamo il problema in forma standard.

$$\begin{array}{l} \min z = -x_1 - 2x_2 \\ -2x_1 + x_2 \leq 1 \\ x_1 - x_2 \geq -2 \\ x_1 + x_2 \leq 5 \\ \underline{x} \geq 0 \end{array} \Rightarrow \begin{array}{l} \min z = -x_1 - 2x_2 \\ -2x_1 + x_2 + x_3 = 1 \\ -x_1 + x_2 + x_4 = 2 \\ x_1 + x_2 + x_5 = 5 \\ \underline{x} \geq 0 \quad x_3 \geq 0, x_4 \geq 0, x_5 \geq 0 \end{array}$$

Quando bisogna applicare il metodo del Simplesso, si necessita la base di partenza. In questo caso, la dimensione della base è 3. Data la matrice $A = \begin{bmatrix} -2 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}$, bisogna quindi individuare tre vettori colonna linearmente indipendenti, perché formano una base di $\mathbb{R}^3$. Non bisogna *mai* sceglierli in modo casuale: bisogna avere un metodo per farlo. Il **caso banale** individua questi tre vettori *se è presente la matrice identità come sottomatrice di A*: i vettori che la costituiscono rappresenteranno la base di partenza. In questo caso, A contiene la matrice identità, per cui la base di partenza sarà B = {3, 4, 5}. Infatti, data una base B costruita a partire dalla matrice identità, è possibile costruire la soluzione di base ammissibile $\underline{x} = \begin{bmatrix} \underline{x}_B \\ \underline{x}_N \end{bmatrix} = \begin{bmatrix} A^{-1}{}_B \underline{b} \\ \underline{0} \end{bmatrix}$, in quanto vale certamente che $A^{-1}{}_B \underline{b} \geq 0$, dato che $\underline{b} \geq 0$ e $A^{-1}{}_B$ è sempre la stessa matrice, per cui $\underline{x}_B = A^{-1}{}_B \underline{b} \Leftrightarrow \underline{x}_B = I\underline{b} \Leftrightarrow \underline{x}_B = \underline{b}$.

Quindi, $A_B = A_B^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Ora, bisogna capire se B = {3, 4, 5} è la base ottima che stiamo cercando, e per farlo occorre calcolare i coefficienti di costo delle variabili fuori base. Siccome N = {1, 2}:

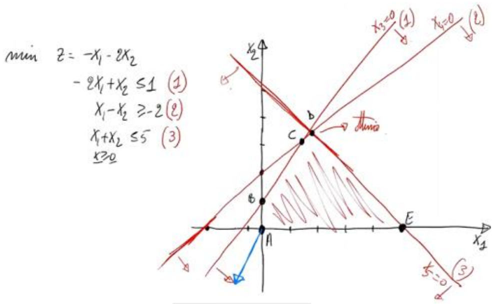
- $z_1 - c_1 = [0 \ \ 0 \ \ 0] \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} - (-1) = 1 > \mathbf{0}$;
- $z_2 - c_2 = [0 \ \ 0 \ \ 0] \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - (-2) = 2 > \mathbf{0}$.

Entrambe le variabili fuori base danno come coefficiente di costo ridotto un valore maggiore di 0, quindi si può scegliere arbitrariamente una delle due e trovare la soluzione ottima del problema. Di seguito, useremo il Metodo del gradiente.

A destra è posta la regione ammissibile per il problema. Useremo la simbologia B_i per indicare la base associata al punto *i*. Dunque:

- B_A = {3, 4, 5}; - B_B = {2, 4, 5}; - B_C = {1, 2, 5};
- B_D = {1, 2, 3}; - B_E = {1, 3, 4}.

Tornando al problema iniziale, abbiamo costruito la matrice A e notato che essa contiene come sottomatrice la matrice identità, associata alle variabili x_3, x_4, x_5.
La base B = {3, 4, 5} è la base associata al punto A, quindi siamo partiti dal punto A.



$$\begin{bmatrix} x_3 \\ x_4 \\ x_5 \end{bmatrix} = A_B^{-1}\underline{b} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} \rightarrow x_3 = 1, x_4 = 2, x_5 = 5.$$

$$z_0 = \underline{c}^T{}_B A_B^{-1}\overline{\underline{b}} \rightarrow z_0 = [0 \ \ 0 \ \ 0]\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} = 0.$$

A questo punto, fatti i calcoli, si osserva che la funzione obiettivo, nel punto (0, 0), vale 0: questo è il valore da cui siamo partiti.

Calcolati $z_1 - c_1 = 1$ e $z_2 - c_2 = 2$, se decidiamo di far entrare x_2 in base, allora x_3 esce dalla base e la nuova base sarà B = {2, 4, 5}, la quale corrisponde geometricamente al punto B. La variabile x_1 è bloccata a 0 (fuori base), quindi possiamo spostarci solo sull'asse x_2, ed è arrivata fino al punto B. A partire dalla base {2, 4, 5}, reiteriamo.

$$A_B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \rightarrow A_B^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

$$\underline{x_B} = A_B^{-1}\underline{b} \rightarrow \begin{bmatrix} x_2 \\ x_4 \\ x_5 \end{bmatrix} = A_B^{-1}\underline{b} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}.$$

Ora, $z_0 = \underline{c}^T{}_B A_B^{-1}\overline{\underline{b}} = [-2 \ \ 0 \ \ 0]\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} = [-2 \ \ 0 \ \ 0]\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} = -2$. Eseguendo il test di ottimalità (cioè, eseguendo il calcolo di $z_j - c_j \ \forall j \in N$, con B = {2, 4, 5} e N = {1, 3}), si ottiene:

- $z_1 - c_1 = [-2 \ \ 0 \ \ 0]\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} - (-1) = [-2 \ \ 0 \ \ 0]\begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} + 1 = 5$;
- $z_3 - c_3 = [-2 \ \ 0 \ \ 0]\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 0 = [-2 \ \ 0 \ \ 0]\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = -2$.

Facendo entrare x_1 in base, per il Test dei minimi rapporti la nuova base sarà {1, 2, 5} = B_C. Reiteriamo, si eseguono tutti i calcoli e la nuova base sarà {1, 2, 3} = B_D: in questo caso, scopriamo che D è il valore ottimo. ✓

Metodo delle due fasi e Big-M

Introduciamo i metodi che ci forniranno la base di partenza per consentire al Simplesso di risolvere il problema.

Consideriamo il problema di PL in Forma Standard

$$\min z = \underline{c}^T\underline{x}$$
$$A\underline{x} = \underline{b} \quad (1)$$
$$\underline{x} \geq \underline{0} \quad (2)$$
$$\underline{x} \in \mathbb{R}^n,$$

Condizione: $\underline{b} \geq 0$

con A(m,n) e $\underline{b} \geq 0$. Per sfruttare l'algoritmo del Simplesso, occorre una base iniziale. Abbiamo visto il caso banale, ossia il caso in cui la matrice A contiene come sottomatrice la matrice identità. Nel caso in cui non ci si ritrova in questo caso, <u>non</u> bisogna andare per tentativi.

Se tra le colonne di A non è presente la matrice identità, non è facile individuare m colonne di A linearmente indipendenti. Possiamo utilizzare il **metodo delle due fasi** per costruire "artificialmente" la matrice identità; si modifica il sistema dei vincoli come segue,

$$\min \underline{c}^T \underline{x}$$
$$A\underline{x} = \underline{b} \rightarrow A\underline{x} + I\underline{y} = \underline{b} \quad (1)$$
$$\underline{x} \geq 0 \rightarrow \underline{x} \geq 0, \underline{y} \geq 0 \quad (2)$$

con l'aggiunta di una variabile artificiale y_i ad ogni vincolo del sistema. Nel nuovo sistema sarà presente una matrice identità (associata alle variabili y).

Una soluzione ($\underline{x}'$, $\underline{y}'$) del nuovo sistema sarà soluzione anche del sistema di partenza <u>se e solo</u> se $\underline{y}' = 0$. Per ottenere una tale soluzione (se esiste) risolviamo il seguente problema di PL:

$$\min g = \sum_{i=1}^{m} y_i$$
$$A\underline{x} + I\underline{y} = \underline{b} \quad (1)$$
$$\underline{x} \geq 0, \underline{y} \geq 0 \quad (2).$$

$2x_1 + 3x_2 = 10$ $\qquad$ $2x_1 + 3x_2 + y_1 = 10$

$\begin{array}{cc} x_1 & x_2 \\ 2, & 2 \end{array}$ $\quad$ $[4, 0]$ $\qquad$ $\begin{array}{ccc} x_1 & x_2 & y_1 \\ 2, & 2, & 0 \end{array}$ $\quad$ $\begin{array}{ccc} x_1 & x_2 & y_1 \\ 4, & 0, & 2 \end{array}$

A destra è posto un esempio; notiamo che:
- ponendo $x_1 = x_2 = 2$, e $y_1 = 0$, sono soddisfatti entrambi i vincoli del vecchio e del nuovo sistema;
- ponendo $x_1 = 4$, $x_2 = 0$, e $y_1 = 2$, è soddisfatto il vincolo del nuovo sistema ma non quello del vecchio sistema.

Formulando il nuovo sistema è stata costruita la *prima fase* del metodo delle due fasi. Per risolvere tale problema possiamo utilizzare il simplesso utilizzando come base iniziale le colonne della matrice (A, I) associate alle variabili artificiali $\underline{y}$. Inizialmente tutte le variabili $\underline{y}$ sono in base, mentre tutte le variabili $\underline{x}$ sono fuori base.

Per il modello matematico costruito nella prima fase, possiamo avere le seguenti condizioni:
1) poiché minimizziamo la somma delle $\underline{y}$ (che sono ≥ 0), il valore più piccolo possibile nella funzione obiettivo è 0, quindi sicuramente il problema non avrà mai ottimo illimitato;
2) se si considerano le colonne della matrice identità, allora già si ha una soluzione di base ammissibile.

Il succo è: sia nel caso in cui ci ritroviamo la matrice iniziale *A* che contiene la matrice identità, sia nel caso in cui dobbiamo applicare il metodo delle due fasi per ottenerla (e, quindi, fare due simplessi), bisogna comunque partire dalla matrice identità.

Alla fine della prima fase possono verificarsi due casi, per l'ottimo della funzione obiettivo:
1. $g^* = \sum_{i=1}^{m} y_i > 0$. Questo significa che almeno una variabile artificiale è maggiore di 0. Questo implica che $A\underline{x} = \underline{b}$ *non ammette soluzione*. Supponiamo per assurdo che $\exists \underline{x}': A\underline{x}' = \underline{b}$, con $\underline{x}' \geq 0$. Costruito il vettore $[\underline{x}', \underline{y}'] = [\underline{x}', \underline{0}]$, notiamo che esso è una soluzione ammissibile. Questo ci consente di costruire una soluzione ammissibile nel nuovo sistema, dove $g^* = 0$, ma questo è assurdo in quanto avevamo supposto $g^* > 0$.
Di conseguenza, non si passa alla seconda fase;
2. $g^* = \sum_{i=1}^{m} y_i = 0$. Questo implica che $A\underline{x} = \underline{b}$ *ammette soluzione* (a meno di soluzioni degeneri). Si passa alla seconda fase, cioè si risolve il problema iniziale utilizzando la base ottima della prima fase come base iniziale della seconda fase.

Es. Si vuole risolvere il problema in figura utilizzando il metodo del simplesso.

$$\min z = -x_1 - 2x_2 \quad x_1 + x_2 \geq 1 \quad x_1 + 2x_2 \leq 4 \quad \underline{x} \geq 0$$

$$\Rightarrow$$

$$\min z = -x_1 - 2x_2 \quad x_1 + x_2 - x_3 = 1 \quad x_1 + 2x_2 + x_4 = 4 \quad \underline{x} \geq 0$$

Prima di tutto, portiamo il problema in forma standard.
Ora, verifichiamo se esiste la matrice identità come sottomatrice di A: in questo caso non c'è.
Infatti: $\underline{x}_B = A^{-1}{}_B\underline{b} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$, ed abbiamo la componente $\underline{x}_3 = -1$ negativa. Dunque, applichiamo il metodo delle due fasi per costruire la matrice identità artificialmente.

<u>Prima fase.</u> Otteniamo

$$x_1 + x_2 - x_3 + x_5^a = 1$$
$$x_1 + 2x_2 + x_4 + x_6^a = 4$$
$$x_1, x_2, \dots, x_6^a \geq 0$$

Osservazione importante: in questo caso, la variabile x_6^a non ci interessa, in quanto già abbiamo la variabile x_4 positiva. In questo modo risparmiamo tempo computazionale, in quanto il numero minimo di iterazioni per ottenere il min g* è *almeno pari* al numero di variabili artificiali.
Procedendo, abbiamo ottenuto la base di partenza (data dagli indici della matrice identità) B = {5, 4}; quindi, $A_B = I = A_B^{-1}$.
Utilizziamo il metodo del simplesso. Prima di tutto, verifichiamo se questa è la base ottima:
- $z_1 - c_1 = z_j = \underline{c}^T{}_B A^{-1}{}_B \underline{a}_j - c_j = [1 \;\; 0]\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} - 0 = 1 > 0$;
- $z_2 - c_2 = [1 \;\; 0]\begin{bmatrix} 1 \\ 2 \end{bmatrix} - 0 = 1 > 0$;
- $z_3 - c_3 = [1 \;\; 0]\begin{bmatrix} -1 \\ 0 \end{bmatrix} = -1 < 0$.

La base non è ottima, per cui possiamo scegliere una delle variabili x_1 o x_2 in modo arbitrario: nel nostro caso, facciamo entrare in base x_2.
Il passaggio successivo è verificare chi esce dalla base, quindi: $\underline{y}_2 = A_B^{-1}\underline{a}_2 = I\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Per quanto detto in precedenza, non possiamo mai trovare il vettore $\underline{y}$ con tutte le componenti negative: questo indicherebbe che l'ottimo è illimitato, e questo non può accadere nella prima fase. Procedendo: $\underline{x}_B = A^{-1}{}_B\underline{b} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$. Per il test dei minimi rapporti: $\min\left\{\frac{1}{1}, \frac{4}{2}\right\} = \frac{1}{1}$, cioè quello associato alla variabile x_5, che uscirà dalla base. Di conseguenza, la nuova base sarà B = {2, 4}.
Reiterando, verifichiamo se quest'ultima è la base ottima: $(z_1 - c_1) = \ldots = 0$, $(z_3 - c_3) = \ldots = 0$, $(z_5 - c_5) = \ldots = -1$. Due osservazioni importanti:
- il coefficiente di costo ridotto della variabile che abbiamo fatto uscire dalla base, all'iterazione precedente, è sicuramente negativo. Questo perché se mi spostassi lungo quella direzione (cioè, tornassi indietro), allora la funzione obiettivo peggiora;
- tutti i coefficienti di costo ridotto sono ≤ 0, quindi il test di ottimalità è soddisfatto. Tuttavia, se esiste un coefficiente di costo ridotto uguale a 0, allora esiste più di un ottimo.

Una volta stabilito che B = {2, 4} è la base ottima della prima fase, questa sarà la base di partenza. Infatti, siccome $g^* = [0 \;\; 0]\begin{bmatrix} 1 & 0 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix} = 0$ si può passare alla seconda fase.
<u>Seconda fase.</u> Si risolve con il simplesso il problema iniziale con B = {2, 4} (e, quindi, N = {1, 3}):

$$\min z = -x_1 - 2x_2$$
$$x_1 + x_2 - x_3 = 1$$
$$x_1 + 2x_2 + x_4 = 4 \quad \ldots\checkmark$$
$$\underline{x} \geq 0$$

Esiste un metodo alternativo per calcolare la base di partenza: il **metodo del Big-M**. Questo metodo costruisce la matrice identità artificialmente esattamente come il precedente ma, a differenza del metodo delle due fasi: alla funzione obiettivo originale vengono sommate (nel caso di un problema di minimo) le variabili artificiali *moltiplicate per un coefficiente* M molto grande.

$$\min \underline{c}^T\underline{x} \rightarrow \min \underline{c}^T\underline{x} + M\underline{y}$$
$$A\underline{x} = \underline{b} \rightarrow A\underline{x} + I\underline{y} = \underline{b} \quad (1)$$
$$\underline{x} \geq 0 \rightarrow \underline{x} \geq 0, \underline{y} \geq 0 \quad (2)$$

In questo caso si applica il simplesso una sola volta, perché l'obiettivo è sempre lo stesso: le variabili artificiali devono essere poste a 0, dunque si penalizzano le $\underline{y}$ assegnandogli un valore molto grande; infatti, il simplesso assegnerà un valore $\underline{y}>0$ solo se strettamente necessario.
Se si avesse una funzione di massimo, siccome si devono penalizzare le $\underline{y}$, si assegna un valore *negativo* ad M:

$$\max \underline{c}^T \underline{x} \to \max \underline{c}^T \underline{x} - M\underline{y}$$

$$A\underline{x} = \underline{b} \to A\underline{x} + I\underline{y} = \underline{b} \quad (1)$$

$$\underline{x} \geq 0 \to \underline{x} \geq 0, \underline{y} \geq 0 \quad (2)$$

Dei due metodi visti utilizzeremo il metodo delle due fasi.

Teoria della Dualità: Parte 1

Dall'inizio del corso, per i modelli matematici ci concentriamo su due loro forme: standard di minimo, e canonica (da non confondere con la forma canonica in funzione di una base B). Quest'ultima è utilizzata per la Teoria della dualità:

$$\begin{aligned} \min z = \underline{c}^T\underline{x} & \\ A\underline{x} &\geq \underline{b} \\ \underline{x} &\geq \underline{0} \\ \underline{x} &\in \mathbb{R}^n. \end{aligned}$$

Dato il problema di PL continua in forma canonica, definiamo un numero reale $lb \in \mathbb{R}$ come il **lower bound** del problema se e solo se $\forall \underline{x} \in X$ vale che $lb \leq \underline{c}^T\underline{x}$. Cioè, qualunque sia il punto della regione ammissibile X, il valore della funzione obiettivo in questo punto sarà sempre $\geq$ lb.
Si tenga presente che se, per ipotesi, $\underline{x}^*$ ottimo $\Rightarrow z^* = \underline{c}^T\underline{x}^*$ valore ottimo, allora $lb \leq z^*$ per definizione.

Il lower bound è un concetto importante, in quanto può fungere da *stima* nel caso in cui è più complesso trovare la soluzione ottima di un problema di PL.

Es. Si consideri il generico problema di PL in forma canonica. Si vuole creare un lower bound per $\underline{c}^T\underline{x}$.
• Se $\underline{w} \in \mathbb{R}^m$ è un vettore avente tutte le componenti positive ($\underline{w} \geq \underline{0}$), allora, andando a moltiplicare $\underline{w}$ per il sistema $A\underline{x} \geq \underline{b}$, si ottiene come risultato che $\underline{w}^T A\underline{x} \geq \underline{w}^T\underline{b}$.
• Se $\underline{w}^T A \leq \underline{c}^T \Leftrightarrow \underline{c}^T \geq \underline{w}^T A$, allora $\underline{c}^T\underline{x} \geq \underline{w}^T A\underline{x} \geq \underline{w}^T\underline{b}$. In questo caso, $\underline{w}^T\underline{b}$ è un lower bound della funzione obiettivo $\underline{c}^T\underline{x}$.

Dall'ultimo risultato dell'esempio precedente, quindi, ponendo come condizioni $\underline{w}^T A \leq \underline{c}^T$ e $\underline{w}^T \geq \underline{0}$, otteniamo sicuramente $\underline{w}^T\underline{b}$ come lower bound di $\underline{c}^T\underline{x}$.
A questo punto, tra tutti i possibili modi per costruire il vettore $\underline{w}$ come lower bound, rispettando i due vincoli precedenti, il migliore possibile (rispetto alla soluzione ottima del problema) è ovviamente quello più vicino al *valore* della soluzione ottima; dunque, vogliamo il $\max \underline{w}^T\underline{b}$. Sintetizzando, il problema seguente

$$\begin{aligned} &\max \underline{w}^T\underline{b} \\ &\underline{w}^T A \leq \underline{c}^T \Leftrightarrow A^T\underline{w} \leq \underline{c} \\ &\underline{w}^T \geq 0 \end{aligned}$$

è detto **problema duale (D)** del problema *primale* **(P)** di partenza.

Ad ogni problema di PL (Primale) è associato un problema Duale:

(n variabili, m vincoli)	(m variabili, n vincoli)
Problema Primale (P)	Problema Duale (D)
$\min \ c_1x_1 + \cdots + c_nx_n$	$\max \ b_1w_1 + \cdots + b_mw_m$
$s.t.$	$s.t.$
$a_{11}x_1 + \cdots + a_{1n}x_n \geq b_1$	$a_{11}w_1 + \cdots + a_{m1}w_m \leq c_1$
$\vdots$	$\vdots$
$a_{m1}x_1 + \cdots + a_{mn}x_n \geq b_m$	$a_{1n}w_1 + \cdots + a_{mn}w_m \leq c_n$
$\underline{x} \geq 0$	$\underline{w} \geq 0$

Notiamo alcune **caratteristiche** che differenziano il problema primale P ed il problema duale D:
- il problema primale è di *minimo*, mentre il problema duale è di *massimo*;
- i termini noti del problema duale sono i *coefficienti di costo* del problema primale. Inoltre, i termini noti del problema primale diventano i coefficienti di costo del problema duale;
- nel problema duale la matrice *A* è trasposta, quindi (date *n* variabili e *m* vincoli del problema primale) questo problema avrà *m* variabili e *n* vincoli. Di conseguenza, il vettore $\underline{w}$ contiene *m* componenti.
- nel problema duale i vincoli diventano di $\leq$, mentre nel primale sono vincoli di $\geq$.

Es. Si vuole costruire il problema duale D del problema primale P in figura.
Si può effettuare ciò utilizzando la seguente tabella, dove:
- la prima riga contiene le variabili di P;
- all'interno della tabella inseriamo i coefficienti tecnologici, cioè la matrice A;
- l'ultima colonna contiene i termini noti di P;
- l'ultima riga contiene i coefficienti di costo della funzione obiettivo di P;
- sulla prima colonna *si associa*, ad ogni vincolo di P, una variabile di D (per convenzione le chiamiamo $w_1, w_2, \dots, w_m$).

La tabella sarà, quindi, la seguente.

$$\begin{aligned} &\min\ c_1x_1 + c_2x_2 + c_3x_3 \\ &s.t. \\ &a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \ge b_1 \\ &a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \ge b_2 \\ &a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \ge b_3 \\ &\underline{x} \ge 0 \end{aligned}$$

	x_1	x_2	x_3	
w_1	a_{11}	a_{12}	a_{13}	b_1
w_2	a_{21}	a_{22}	a_{23}	b_2
w_3	a_{31}	a_{32}	a_{33}	b_3
	c_1	c_2	c_3	

Esaminate le caratteristiche individuate in precedenza, costruiamo D:
- la funzione obiettivo di D si ottiene moltiplicando l'ultima colonna della tabella (dei termini noti) per la prima colonna della tabella (delle variabili di D). Ovviamente, questa sarà una funzione di max;
- costruiamo i vincoli di D per colonne. Otteniamo il vincolo i-esimo (x_i) moltiplicando la colonna i della matrice per il vettore delle variabili di D, e ogni vincolo dovrà terminare con una disuguaglianza con c_i;
- indichiamo il segno delle variabili $w_1, w_2, \dots, w_m$.

Otteniamo il problema D, posto a destra. ✓

$$\begin{aligned} &\max\ b_1w_1 + b_2w_2 + b_3w_3 \\ &s.t \\ &a_{11}w_1 + a_{21}w_2 + a_{31}w_3 \le c_1 \\ &a_{12}w_1 + a_{22}w_2 + a_{32}w_3 \le c_2 \\ &a_{13}w_1 + a_{23}w_2 + a_{33}w_3 \le c_3 \\ &w_1 \ge 0, w_2 \ge 0, w_3 \ge 0 \end{aligned}$$

In generale, possiamo basarci sulla **forma matriciale**:

$$\begin{array}{llll} (P) & \min\ \underline{c}^T \underline{x} & (D) & \max\ \underline{b}^T \underline{w} \\ & A\underline{x} \ge \underline{b} & & A^T \underline{w} \le \underline{c} \\ & \underline{x} \ge \underline{0} & & \underline{w} \ge \underline{0} \\ & \underline{x} \in \mathbf{R}^n & & \underline{w} \in \mathbf{R}^m \end{array}$$

La corrispondenza costruita dalle frecce è detta **forma di riferimento**, e ci permette di calcolare il duale di un problema primale *senza* dover trasformare il primale in forma canonica. Le frecce blu nella figura precedente ci dicono che:
- se si considera P e i suoi vincoli sono di ≥, allora le variabili di D saranno ≥ 0. Quindi, i vincoli di P indicano il segno delle variabili di D (*proporzionalità diretta*);
- se si considera P e le sue variabili sono ≥ 0, allora i vincoli saranno di ≤. Le variabili di P indicano il segno dei vincoli di D (*proporzionalità indiretta*).

Es. Dato il problema primale a destra, si vuole costruirne il duale.
Prima di tutto, una <u>nota</u> importantissima. Quando si eseguono queste operazioni, bisogna tener sempre presente la struttura del problema di minimo e di massimo: infatti, il problema di partenza potrebbe anche essere di massimo; in questo caso, il primale della figura precedente è (D), ed il duale da costruire corrisponde a (P).

$$\begin{aligned} &\min\ 3x_1 + 4x_2 \\ &s.t. \\ &2x_1 + 1/2\,x_2 \ge 3 \\ &4x_1 + x_2 \ge 2 \\ &1/5\,x_1 \ge 7 \\ &x_1, x_2 \ge 0 \end{aligned}$$

Tornando a questo esempio, otteniamo la matrice seguente, che ci consente di formulare il problema duale come segue.

	x_1	x_2	
w_1	2	1/2	3
w_2	4	1	2
w_3	1/5	0	7
	3	4	

$$\Rightarrow \quad \begin{aligned} &\max\ 3w_1 + 2w_2 + 7w_3 \\ &s.t \\ &2w_1 + 4w_2 + 1/5\,w_3 \le 3 \\ &1/2\,w_1 + w_2 \le 4 \\ &w_1 \ge 0, w_2 \ge 0, w_3 \ge 0 \end{aligned}$$

✓

Notiamo che *il duale del problema duale corrisponde al problema primale*. Infatti:

$$\begin{array}{ll} (P)\ \max & \underline{b}^T \underline{w} \\ & A^T \underline{w} \le \underline{c} \\ & \underline{w} \ge \underline{0} \\ & \underline{w} \in \mathbf{R}^m \end{array} \quad \Rightarrow \quad \begin{array}{ll} -\min & -\underline{b}^T \underline{w} \\ & -A^T \underline{w} \ge -\underline{c} \\ & \underline{w} \ge \underline{0} \\ & \underline{w} \in \mathbf{R}^m \end{array}$$

Il **duale** di questo problema è:

	$\underline{w}$	
$\underline{x}$	$-A^T$	$-\underline{c}$
	$-\underline{b}^T$	

$$\begin{array}{ll} -\max & -\underline{c}^T \underline{x} \\ & -A\underline{x} \le -\underline{b} \\ & \underline{x} \ge \underline{0} \\ & \underline{x} \in \mathbf{R}^n \end{array} \quad \Rightarrow \quad \begin{array}{ll} (D)\ \min & \underline{c}^T \underline{x} \\ & A\underline{x} \ge \underline{b} \\ & \underline{x} \ge \underline{0} \\ & \underline{x} \in \mathbf{R}^n \end{array}$$

Ma come costruiamo il duale di un primale con **vincoli di uguaglianza**? In linea teorica, si può sempre trasformare il primale in forma canonica per poi costruire il duale. Quindi, trasformiamo i vincoli di uguaglianza in vincoli di maggiore o uguale come segue:

$$\begin{array}{ll} (P) \quad \min & \underline{c}^T \underline{x} \\ & A\underline{x} = \underline{b} \\ & \underline{x} \ge \underline{0} \\ & \underline{x} \in R^n \end{array}$$

$$A\underline{x} = \underline{b} \text{ equivale a } \begin{array}{c} A\underline{x} \ge \underline{b} \\ A\underline{x} \le \underline{b} \Rightarrow -A\underline{x} \ge -\underline{b} \end{array}.$$

Otteniamo, quindi, il risultato seguente. Quindi, si introducono $2m$ **variabili duali** $\underline{u}$, $\underline{v}$.

$$\begin{array}{ll} (P) \quad \min & \underline{c}^T \underline{x} \\ & A\underline{x} \ge \underline{b} \\ & -A\underline{x} \ge -\underline{b} \\ & \underline{x} \ge \underline{0} \\ & \underline{x} \in R^n \end{array} \quad \Rightarrow \quad \begin{array}{c|c|c} & \underline{x} & \\ \hline \underline{u} & A & \underline{b} \\ \underline{v} & -A & -\underline{b} \\ \hline & \underline{c} & \end{array} \quad \Rightarrow \quad \begin{array}{ll} (D) & \max \ (\underline{b}^T \underline{u} - \underline{b}^T \underline{v}) \\ & A^T \underline{u} - A^T \underline{v} \le \underline{c} \\ & \underline{u} \ge \underline{0}, \underline{v} \ge \underline{0} \\ & \underline{u}, \underline{v} \in R^m \end{array}$$

Sostituendo $\underline{w} = \underline{u} - \underline{v}$ in D, otteniamo il seguente duale, con $\underline{w}$ *non vincolato* (in quanto dipende da $\underline{u}$ e $\underline{v}$).

$$\boxed{\begin{array}{ll} (D) \quad \max & \underline{b}^T \underline{w} \\ & A^T \underline{w} \le \underline{c} \\ & \underline{w}\ n.v. \\ & \underline{w} \in R^m \end{array}}$$

Dato un generico problema di PL, sarebbe possibile trasformarlo in uno equivalente in forma canonica di min/max per calcolarne il duale. In realtà questo non è necessario, in quanto è *sempre* possibile calcolare direttamente il duale del problema dato con delle opportune **regole di trasformazione generali**.

Es. Dato il problema seguente, si vuole costruirne il duale in modo diretto (cioè, senza trasformarlo prima in forma canonica). A tal scopo, costruiamo prima la seguente tabella.

$$\begin{array}{ll} \max & 5x_1 + 3x_2 \\ s.t. & \\ x_1 + 4x_2 - 8x_3 & \ge 2 \\ 7x_2 + 12x_3 & \le 4 \\ 9x_1 - 8x_2 + x_3 & = 7 \\ x_1, x_2, x_3 & \ge 0 \end{array}$$

	x_1	x_2	x_3	
w_1	1	4	-8	2
w_2	0	7	12	4
w_3	9	-8	1	7
	5	3	0	

Osservata la tabella, tenendo presente la forma di riferimento strutturiamo D come segue.
L'errore da non fare quando si osserva la forma di riferimento è considerare sempre il primale come il problema alla sinistra. Infatti:
- se il problema è di minimo, bisogna considerare come primale quello a sx;
- se il problema è di massimo, bisogna considerare come primale quello a dx.
Siccome il nostro primale è di massimo, dato che le sue variabili sono tutte ≥ 0 allora tutti i vincoli del duale risultante saranno vincoli di ≥. Inoltre:
- $w_1 \leq 0$, in quanto il primo vincolo di P è di ≥;
- $w_2 \geq 0$, in quanto il secondo vincolo di P è di ≤;
- $w_3\ n.v.$, in quanto il terzo vincolo di P è di =. ✓

$$\begin{array}{c} \min 2w_1 + 4w_2 + 7w_3 \\ w_1 + 0w_2 + 9w_3 \geq 5 \\ 4w_1 + 7w_2 - 8w_3 \geq 3 \\ -8w_1 + 12w_2 + w_3 \geq 0 \\ w_1 \leq 0, w_2 \geq 0, w_3\ n.v. \end{array}$$

Segue una tabella in cui sono poste le regole di trasformazione generali.

	min		max	
Coefficienti f.o.	$\underline{c}^T$	$\Leftrightarrow$	$\underline{c}$	Termini noti vincoli
Termini noti vincoli	$\underline{b}$	$\Leftrightarrow$	$\underline{b}^T$	Coefficienti f.o.
Coefficienti vincoli	A	$\Leftrightarrow$	A^T	Coefficienti vincoli
i-mo vincolo (*i*=1…m)	≥		≥ 0	*i*-ma variabile (*i*=1…m)
	≤	$\Leftrightarrow$	≤ 0	
	=		n.v.	
j-ma variabile (*j*=1...n)	≥ 0		≤	*j*-mo vincolo (*j*=1...n)
	≤ 0	$\Leftrightarrow$	≥	
	n.v.		=	

La teoria della Dualità è importante perché:
- le soluzioni di (P) e (D) sono legate tra loro;
- *la soluzione ottima del duale è un bound sulla soluzione ottima del primale*;
- le soluzioni duali hanno un'interpretazione economica utile per l'analisi di sensitività (*post-ottimalità*). Vedremo maggiori dettagli su quest'aspetto in seguito (in sintesi, dopo aver dato la soluzione ottima all'azienda si inizia la fase della post-ottimalità, ossia si analizza quanto cambia il profitto al variare di una o più risorse nel tempo);
- sulla teoria della dualità sono basati algoritmi, quali il Simplesso Duale, l'Algoritmo Primale-Duale, il Delayed Column Generation alternativi al Simplesso (Primale) utili per certe classi di problemi;
- può in certi casi essere conveniente risolvere D al posto di P (generalmente, conviene risolvere il problema con il minor numero di vincoli).

Siano dati il problema primale **di minimo** (P) $\begin{array}{c} \min \underline{c}^T\underline{x} \\ A\underline{x} \geq \underline{b} \\ \underline{x} \geq \underline{0} \end{array}$ ed il suo duale di massimo (D) $\begin{array}{c} \max \underline{b}^T\underline{w} \\ A^T\underline{w} \leq \underline{c} \\ \underline{w} \geq \underline{0} \end{array}$.

Teorema (DEBOLE DELLA DUALITÀ). *Siano $\underline{\hat{x}}$ e $\underline{\hat{w}}$ soluzioni ammissibili rispettivamente per (P) e (D). Allora*
$$\underline{c}^T\underline{\hat{x}} \geq \underline{b}^T\underline{\hat{w}}.$$
DIM. Sappiamo che $\underline{\hat{x}}$ è una soluzione ammissibile di (P), quindi $A\underline{\hat{x}} \geq \underline{b}$. (1)
Poiché $\underline{\hat{w}} \geq \underline{0}$, possiamo premoltiplicare (1) per $\underline{\hat{w}}$, ottenendo $\underline{\hat{w}}^T A\underline{\hat{x}} \geq \underline{\hat{w}}^T\underline{b}$. (2)
Poiché $\underline{\hat{w}}$ è una soluzione ammissibile di (D), essa soddisfa il vincolo $A^T\underline{w} \leq \underline{c}$, che implica $\underline{c}^T \geq \underline{w}^T A$. (3)
Infine, sfruttando le disequazioni (2) e (3), sapendo che $\underline{\hat{x}} \geq \underline{0}$ si può premoltiplicare (3) per $\underline{\hat{x}}$, ottenendo che:
$$\underline{c}^T\underline{\hat{x}} \geq \underline{w}^T A\underline{\hat{x}} \geq \underline{\hat{w}}^T\underline{b} \quad \rightarrow \quad \underline{c}^T\underline{\hat{x}} \geq \underline{\hat{w}}^T\underline{b}.$$ □

Il Teorema debole della dualità dice che il duale è il lower bound del primale, in particolar modo facendo riferimento alle soluzioni $\underline{\hat{x}}$ e $\underline{\hat{w}}$. In realtà, se noi non consideriamo i precedenti due problemi come primale e duale, bensì semplicemente come problemi di *minimo* e di *massimo*, questo teorema: scelta una qualsiasi soluzione $\underline{\hat{x}}$ del problema di minimo, e scelta una soluzione $\underline{\hat{w}}$ del problema di massimo, si ottiene che $\underline{c}^T\underline{\hat{x}} \geq \underline{b}^T\underline{\hat{w}}$ sempre; infatti, nel problema di minimo si tende a scegliere la soluzione più piccola possibile, mentre nel problema di massimo si tende a scegliere la soluzione più grande possibile, e affinché siano l'una il bound

dell'altra deve valere questa disequazione.
Capito questo risultato più generico, se nello specifico poniamo il problema di minimo *duale* ed il problema di massimo *primale*, semplicemente cambia il verso della disequazione; cioè: $\underline{c}^T\underline{\hat{x}} \leq \underline{b}^T\underline{\hat{w}}$.

Corollario 1. Se x è una soluzione ammissibile per il primale (P) e w una soluzione ammissibile per il duale (D) tali che $\underline{c}^T\underline{x} = \underline{b}^T\underline{w}$, allora x e w sono soluzioni ottime dei rispettivi problemi.
DIM. Supponiamo per assurdo che $\exists \underline{x}' \in X$: $\underline{c}^T\underline{x}' < \underline{c}^T\underline{x}^* = \underline{b}^T\underline{w}^*$; siccome vale quest'ultima uguaglianza, allora $\underline{c}^T\underline{x}' < \underline{b}^T\underline{w}^*$: questo è un assurdo, poiché va contro la tesi del Teorema debole della dualità. □

Corollario 2. Se il problema primale (P) è illimitato inferiormente, allora il duale (D) è inammissibile. Viceversa, se il duale (D) è illimitato superiormente, allora il primale (P) è inammissibile.
DIM. Supponiamo che il valore ottimo del primale (P) sia $z^* = -\infty$ e che il problema duale ammetta una soluzione w. Dal teorema della dualità debole si ha che $\underline{c}^T\underline{x} \geq \underline{b}^T\underline{w}$ per una qualsiasi soluzione ammissibile x di (P), dove $\underline{b}^T\underline{w}$ è ovviamente uno scalare. Questo implica che $\underline{b}^T\underline{w} \leq -\infty$: questo è assurdo. □

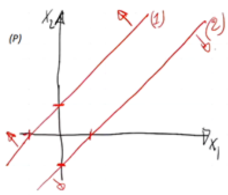
Il corollario 2 stabilisce che l'illimitatezza di un problema implica l'inammissibilità del suo duale. Tuttavia, questa non è una proprietà simmetrica: cioè, se un problema è inammissibile non è detto che il suo duale sia illimitato.
Per esempio:

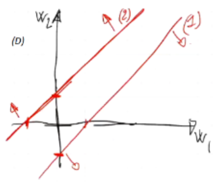
$$(P)\quad \min -x_1 - x_2 \qquad -x_1 + x_2 \geq 1 \qquad x_1 - x_2 \geq 1 \qquad x_1, x_2 \geq 0$$

$$(D)\quad \max w_1 + w_2 \qquad -w_1 + w_2 \leq -1 \qquad w_1 - w_2 \leq -1 \qquad w_1 \geq 0,\ w_2 \geq 0$$

	x_1	x_2	
w_1	-1	1	1
w_2	1	-1	1
	-1	-1	

Risolvendo geometricamente i problemi precedenti, notiamo che il problema primale (P) è inammissibile, così come il problema duale (D).





Quindi, riassumendo, se il problema primale (P) è inammissibile, allora ci sono due possibilità per il duale (D): o quest'ultimo ha ottimo illimitato, oppure è anch'esso inammissibile (come nell'ultimo esempio).

Teoria della Dualità: Parte 2

Teorema (FORTE DELLA DUALITÀ)**.** *Data una coppia di problemi primale-duale, (P) e (D), se uno dei due problemi ammette una soluzione ottima finita, allora anche l'altro problema ammette una soluzione ottima finita ed i valori ottimi delle funzioni obiettivo coincidono, i.e.*

$$\underline{c}^T\underline{x}^* = \underline{b}^T\underline{w}^*.$$

N.B.: Si ricordi che il vettore $\underline{x}^*$ ha spesso una dimensione diversa dal vettore $\underline{w}^*$. Questo teorema non afferma che sono uguali le soluzioni dei due problemi, bensì sono uguali i valori di quelle soluzioni all'ottimo.

DIM. Sia $\underline{x}^*$ la soluzione ottima del primale, e sia B la base ad esso associata:

$$\underline{x}^* = \begin{bmatrix} \underline{x}^*{}_B \\ \underline{x}^*{}_N \end{bmatrix} = \begin{bmatrix} A^{-1}{}_B\underline{b} \\ \underline{0} \end{bmatrix} \Rightarrow \underline{c}^T\underline{x}^* = \underline{c}_B^T\underline{x}_B^* + \underline{c}_N^T\underline{x}_N^* = \underline{c}_B^T A_B^{-1}\underline{b}.$$

Sia $\underline{w}^{*T} = \underline{c}_B^T A_B^{-1}$; vogliamo dimostrare che questo vettore è una soluzione ammissibile (1) ed ottima (2) per (D). Si ricordi che la dimensione del vettore $\underline{w}^*$ è 1x*m*, in quanto $\underline{c}_B^T$ ha dimensione 1x*m* e A_B^{-1} ha dimensione *m*x*m*, quindi il loro prodotto interno avrà dimensione 1x*m*.

(1) Dal duale, il vincolo che dev'essere soddisfatto è che $A^T\underline{w}^* \leq \underline{c} \Rightarrow \underline{w}^{*T}A \leq \underline{c}^T$. Partizioniamo la matrice *A* (e il vettore $\underline{c}^T$) in colonne che appartengono e non alla base: $\underline{w}^{*T}A \leq \underline{c}^T \Rightarrow \underline{w}^{*T}[A_B|A_N] - [\underline{c}_B|\underline{c}_N] \leq \underline{0}$. A questo punto, sostituiamo a $\underline{w}^{*T}$ il valore deciso all'inizio della dimostrazione:

$$\underline{c}_B^T A_B^{-1}[A_B|A_N] - [\underline{c}_B|\underline{c}_N] \leq \underline{0} \Rightarrow [\underline{c}_B^T A_B^{-1}A_B - \underline{c}_B \mid \underline{c}_B^T A_B^{-1}A_N - \underline{c}_N] \leq \underline{0} \Rightarrow$$
$$\Rightarrow [\underline{c}_B^T I - \underline{c}_B \mid \underline{c}_B^T A_B^{-1}A_N - \underline{c}_N] \leq \underline{0} \Rightarrow [\underline{0} \mid z_j - c_j] \leq \underline{0}.$$

Infatti, per definizione $z_j - c_j \leq 0\ \forall j$ (soddisfano la condizione di ottimalità). Questo dimostra che il vettore costruito è una soluzione ammissibile per il duale.

(2) Il valore della funzione obiettivo duale in $\underline{w}^{*T}$ è $\underline{w}^{*T}\underline{b} = \underline{c}_B^T A_B^{-1}\underline{b} = \underline{c}_B^T\overline{b} = \underline{c}_B^T\underline{x}_B + \underline{c}_N^T\underline{x}_N = \underline{c}^T\underline{x}$. Dal Corollario 1 del teorema della dualità debole sappiamo che, essendo $\underline{c}^T\underline{x}^* = \underline{b}^T\underline{w}^*$, anche $\underline{w}^{*T}$ è ottima. □

Dal teorema della dualità forte ricaviamo che, data la base ottima B del primale, è possibile calcolare velocemente la soluzione ottima del duale (D) tramite l'equazione:

$$\boxed{\underline{w}^{*T} = \underline{c}_B^T A_B^{-1}}.$$

Riassumendo:
- (P) è illimitato ⇒ (D) non è ammissibile;
- (P) ha soluzione ottima finita ⇔ (D) ha soluzione ottima finita (ed i valori delle loro f.o. coincidono);
- (P) è inammissibile ⇒ (D) illimitato o inammissibile.

Per il teorema seguente (degli "scarti complementari" – *Complementary Slackness Theorem*), consideriamo la coppia di problemi (P) e (D) in forma canonica e trasformiamoli nella seguente forma: trasformiamo i vincoli dei due problemi in vincoli di uguaglianza, pertanto aggiungiamo le opportune variabili di slack ($\underline{s}$) e surplus ($\underline{v}$).
Sappiamo che ad ogni variabile di (P) è associato un vincolo di (D), e quindi la corrispondente variabile di slack/surplus e viceversa.

$$(P)\quad \begin{array}{l} \min \underline{c}^T\underline{x} \\ A\underline{x} \geq \underline{b} \\ \underline{x} \geq \underline{0} \end{array} \longrightarrow \begin{array}{ll} \min \underline{c}^T\underline{x} & \\ A\underline{x} - I\underline{s} = \underline{b} & \\ \underline{x} \geq \underline{0} & n \text{ var.} \\ \underline{s} \geq \underline{0} & m \text{ var. } di\ surplus \end{array}$$

$$(D)\quad \begin{array}{l} \max \underline{b}^T\underline{w} \\ A^T\underline{w} \leq \underline{c} \\ \underline{w} \geq \underline{0} \end{array} \longrightarrow \begin{array}{ll} \max \underline{b}^T\underline{w} & \\ A^T\underline{w} + I\underline{v} = \underline{c} & \\ \underline{w} \geq \underline{0} & m \text{ var.} \\ \underline{v} \geq \underline{0} & n \text{ var. } di\ slack \end{array}$$

Teorema (DEGLI SCARTI COMPLEMENTARI)**.** *Data la coppia di soluzioni* $\underline{x}$ *e* $\underline{w}$ *rispettivamente ammissibili per (P) e (D),* $\underline{x}$ *e* $\underline{w}$ *sono ottime per (P) e (D) se e solo se*

$$s_j w_j = (\underline{a}^j\underline{x} - b_j)w_j = 0 \quad (j = 1, \dots, m)$$
$$v_i x_i = (c_i - \underline{a}_i^T\underline{w})x_i = 0 \quad (i = 1, \dots, n),$$

dove $\underline{a}^j$ *è la j-esima colonna di A, e* $\underline{a}_i$ *è la i-esima colonna di A.*

Introduciamo il teorema solo per scopi didattici.

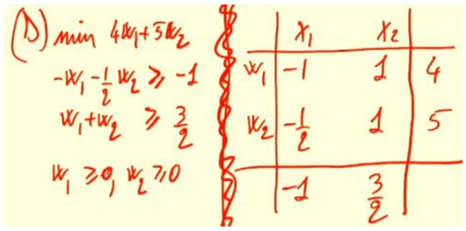
Esercizio. Dato il problema posto a destra:

1. Scrivere il duale del problema e determinare una coppia di soluzioni (P)-(D) ammissibile;
2. Verificare che le soluzioni trovare soddisfano il teorema debole della dualità;
3. Verificare se le soluzioni trovate sono ottime;
4. Risolvere graficamente il primale ed individuare il punto di ottimo e la base B;
5. Calcolare la soluzione ottima del duale a partire dalla base ottima B;
6. Verificare, utilizzando gli scarti complementari, che le soluzioni trovate nei due punti precedenti siano effettivamente ottime.

$$\max \; -x_1 + \frac{3}{2}x_2$$
$$-x_1 + x_2 \leq 4$$
$$-\frac{1}{2}x_1 + x_2 \leq 5$$
$$x_1 \geq 0 \;\; x_2 \geq 0$$

→

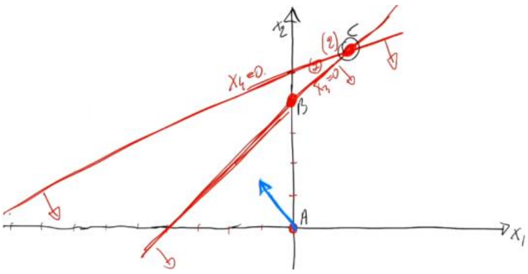
1. Otteniamo il problema duale, posto a destra.
Inoltre, osserviamo che la coppia $\underline{x}' = [1 \;\; 1]$ e $\underline{w}' = [½ \;\; 1]$ è una coppia di soluzioni primale-duale ammissibile.



2. Occorre far attenzione sul Teorema debole: esso lavora su un primale di *minimo*, mentre in questo caso il nostro è un primale di massimo. Infatti, in questo caso bisogna provare una sua conseguenza, cioè che $\underline{c}^T\underline{\hat{x}} \leq \underline{b}^T\underline{\hat{w}}$: questo è vero, in quanto $½ \leq 7$.

3. Per verificare se le soluzioni trovate sono ottime, bisogna verificare se esse sono uguali (Corollario 1): in questo caso non sono ottime, in quanto $½ \neq 7$.

4. Usando la regola del gradiente, tracciando le curve di livello perpendicolari al gradiente nella direzione giusta (in questo caso quella del gradiente, in quanto massimizziamo) notiamo che l'ottimo è il punto C.
Siccome il punto C è l'ottimo, allora $B_C = \{1, 2\}$.
Il punto di ottimo è C = $\underline{x}$ = (2, 6): quindi, $x_1 = 2$ e $x_2 = 6$.



5. $A_B = \begin{bmatrix} -1 & 1 \\ -\frac{1}{2} & 1 \end{bmatrix} \rightarrow A_B^{-1} = \begin{bmatrix} -2 & 2 \\ -1 & 2 \end{bmatrix}$.

Conoscendo le coordinate di C, abbiamo che il valore ottimo del problema è $z^* = -2 + \frac{3}{2} \cdot 6 = 7$. Dal Teorema forte della dualità, sappiamo che il duale ammette soluzione ottima finita; inoltre, il valore ottimo del duale è 7. Calcoliamo, quindi, la soluzione ottima del duale: $\underline{w}^{*T} = \underline{c}_B^T A_B^{-1} = \begin{bmatrix} -1 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} -2 & 2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 1 \end{bmatrix}$: quindi, $w_1 = \frac{1}{2}$ e $w_2 = 1$.

6. Per il Teorema degli scarti complementari, dobbiamo impostare il sistema di equazioni e fare la sostituzione dei valori per verificare se sono soddisfatti. Bisogna ricordare che ogni equazione del sistema deve verificare che il prodotto della variabile del duale per la slack/surplus del vincolo del primale ad esso associato è uguale a 0:

$$\begin{cases} w_1x_3 = w_1(4 + x_1 - x_2) = 0 \\ w_2x_4 = w_2\left(5 + \frac{1}{2}x_1 - x_2\right) = 0 \\ x_1w_3 = x_1\left(-w_1 - \frac{1}{2}w_2 + 1\right) = 0 \\ x_2w_4 = x_2\left(w_1 + w_2 - \frac{3}{2}\right) = 0 \end{cases} \Leftrightarrow \begin{cases} 1 \cdot (4 + 2 - 6) = 0 \;\; \text{(OK)} \\ 1 \cdot \left(5 + \frac{1}{2} \cdot 2 - 6\right) = 0 \;\; \text{(OK)} \\ 2 \cdot \left(-\frac{1}{2} - \frac{1}{2} \cdot 1 + 1\right) = 0 \;\; \text{(OK)} \\ 6 \cdot \left(\frac{1}{2} + 1 - \frac{3}{2}\right) = 0 \;\; \text{(OK)} \end{cases}.$$

Poiché il sistema degli scarti complementari è soddisfatto, allora le due soluzioni trovate ($[2 \;\; 6]$ per il primale e $\begin{bmatrix} \frac{1}{2} & 1 \end{bmatrix}$ per il duale) sono le rispettive soluzioni ottime di (P) e (D). ✓

Interpretazione Economica della Teoria della Dualità

Esercizio. Dato il problema primale (P) a lato:

a) Data la base B = {2, 1} verificare se è ammissibile;
b) Verificare se è ottima;
c) Risolvere graficamente il problema primale;
d) Calcolare la soluzione ottima del problema duale;
e) Risolvere graficamente il problema duale;
f) Verificare attraverso il corollario 1 del teorema debole della dualità che la coppia di soluzioni primale/duale calcolate sono ottime;
g) Verificare attraverso il teorema degli scarti complementari se la coppia di soluzioni primale/duale è ottima.

$$(P) \quad \min \; x_1 - \frac{3}{2}x_2$$
$$-x_1 + x_2 \le 4$$
$$-\frac{1}{2}x_1 + x_2 \le 5$$
$$\underline{x} \ge \underline{0}$$

→

a) Trasformando il problema in forma standard: $\begin{matrix} \min x_1 - \frac{3}{2}x_2 \\ -x_1 + x_2 + x_3 = 4 \\ -\frac{1}{2}x_1 + x_2 + x_4 = 5 \\ \underline{x} \ge \underline{0} \end{matrix}$. Dunque, $A_B = \begin{bmatrix} 1 & -1 \\ 1 & -1/2 \end{bmatrix}$. Otteniamo che

$\underline{x}_B = A_B^{-1}\underline{b} = \begin{bmatrix} -1 & 2 \\ -2 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix} \ge \underline{0}$, pertanto B è una soluzione di base ammissibile (con valori $x_2 = 6$, $x_1 = 2$). ✓

b) Per verificare che essa è ottima, bisogna avere $z_j - c_j \le 0 \quad \forall j \in N$, con $z_j = \underline{c}_B^T A_B^{-1} \underline{a}_j$. Otteniamo:

- $z_3 - c_3 = \begin{bmatrix} -\frac{3}{2} & 1 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ -2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} - 0 = \begin{bmatrix} -\frac{1}{2} & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = -\frac{1}{2} \le 0$ (OK), quindi x_3 non deve entrare in base;
- $z_4 - c_4 = \begin{bmatrix} -\frac{1}{2} & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} - 0 = -1 \le 0$, quindi x_4 non deve entrare in base.

Poiché tutti i coefficienti di costo ridotto sono ≤ 0, la base B = {2, 1} è la base ottima del problema. ✓

c) *Lasciato per esercizio.*

d) La soluzione ottima del primale è $[2 \;\; 6 \;\; 0 \;\; 0]$, pertanto il suo valore ottimo è $z^* = -7$.
Ora, calcoliamo la soluzione ottima del duale: $\underline{w}^{*T} = \underline{c}_B^T A_B^{-1} = \begin{bmatrix} -\frac{3}{2} & 1 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & -1 \end{bmatrix}$. ✓

e) *Lasciato per esercizio.*

f) Per il Corollario 1, dev'essere $\underline{c}^T \underline{x} = \underline{b}^T \underline{w}$, quindi costruiamo il duale.

Abbiamo ottenuto $\underline{x}^{*T} = [x_2 \quad x_1] = [6 \quad 2]$ e $\underline{w}^{*T} = \left[-\frac{1}{2} \quad -1\right]$, con $z^* = -7$.

$$
\begin{array}{ll}
(P) & \min\ x_1 - \frac{3}{2} x_2 \\
(w_1) & -x_1 + x_2 \leq 4 \\
(w_2) & -\frac{1}{2} x_1 + x_2 \leq 5 \\
 & \underline{x} \geq \underline{0}
\end{array}
\qquad
\begin{array}{l}
(D)\ \max\ 4w_1 + 5w_2 \\
-w_1 - \frac{1}{2} w_2 \leq 1 \\
w_1 + w_2 \leq -\frac{3}{2} \\
w_1 \leq 0,\ w_2 \leq 0
\end{array}
$$

Ovviamente, il valore ottimo del duale coincide con z*:
$g^* = 4 \cdot \left(-\frac{1}{2}\right) + 5 \cdot (-1) = -7$. ✓

g) Per utilizzare il teorema degli scarti complementari, conviene osservare (P) e (D) e scrivere esplicitamente quali sono le variabili di slack/surplus. Quindi, semplicemente trasformiamo i vincoli di ≤/≥ in vincoli di uguaglianza anche per il duale. In questo caso, sappiamo che per (P) aggiungiamo x_3 e x_4, mentre per (D) ricaviamo il seguente.

	x_1	x_2	
w_1	-1	1	4
w_2	$-\frac{1}{2}$	1	5
	1	$-\frac{3}{2}$	

$$
\begin{array}{ll}
(D) & \max\ 4w_1 + 5w_2 \\
(x_1) & -w_1 - \frac{1}{2} w_2 + w_3 = 1 \\
(x_2) & w_1 + w_2 + w_4 = -\frac{3}{2} \\
 & w_1 \leq 0,\ w_2 \leq 0,\ w_3, w_4 \geq 0
\end{array}
$$

A questo punto, il teorema degli scarti complementari ci dice che, preso il primale, se per ogni vincolo si moltiplica la variabile w_i per la variabile di slack/surplus del vincolo da esso associata, si deve ottenere l'uguaglianza con 0; lo stesso si deve ottenere per il duale. Quindi:

$$
\begin{cases}
w_1 x_3 = w_1(4 + x_1 - x_2) = 0 \\
w_2 x_4 = w_2\left(5 + \frac{1}{2}x_1 - x_2\right) = 0 \\
x_1 w_3 = x_1\left(1 + w_1 + \frac{1}{2}w_2\right) = 0 \\
x_2 w_4 = x_2\left(-\frac{3}{2} - w_1 - w_2\right) = 0
\end{cases}
\Leftrightarrow
\begin{cases}
-\frac{1}{2} \cdot (4 + 2 - 6) = 0 \quad (\text{OK}) \\
-1 \cdot \left(5 + \frac{1}{2} \cdot 2 - 6\right) = 0 \quad (\text{OK}) \\
2 \cdot \left(1 - \frac{1}{2} - \frac{1}{2}\right) = 0 \quad (\text{OK}) \\
6 \cdot \left(-\frac{3}{2} + \frac{1}{2} + 1\right) = 0 \quad (\text{OK})
\end{cases}.
$$

Le due soluzioni, quindi, sono ottime per (P) e (D). ✓

———

Esempio (Piano di produzione)**.** Una azienda produce due tipi di mangime A e B, entrambi costituiti da una miscela di carne e cereali. La seguente tabella mostra quanti kg di carne e di cereali sono necessari per produrre un kg di A e B (come posto nella tabella a destra).

		Prodotti	
		A	B
materie prime	Cereali	1 Kg	1.5 Kg
	Carne	2 Kg	1 Kg

Ogni giorno l'azienda ha a disposizione 240 kg di cereali e 180 Kg di carne. Inoltre, essendo il mangime A più pregiato, prima di poter essere venduto deve essere raffinato attraverso un macchinario in grado di raffinare al più 110 kg di mangime al giorno. Il ricavo per kg è pari a 560€ euro per il mangime A e 420€ per il mangime B.

PROBLEMA: determinare quanti kg di mangime A e B devono essere prodotti giornalmente per massimizzare il ricavo dell'azienda.

→ Decise le variabili x_1 = # kg di mangime A, x_2 = # kg di mangime B, scriviamo il seguente modello matematico:
$$
\begin{array}{l}
\max\ 560x_1 + 420x_2 \\
x_1 + 1.5x_2 \leq 240 \quad (\text{cereali}) \\
2x_1 + x_2 \leq 180 \quad (\text{carne}) \\
x_1 \leq 110 \quad (\text{raffinatrice}) \\
x_1 \geq 0, x_2 \geq 0
\end{array}
$$
. Trasformiamo il modello in forma standard:
$$
\begin{array}{l}
\min\ -560x_1 - 420x_2 \\
x_1 + 1.5x_2 + x_3 = 240 \\
2x_1 + x_2 + x_4 = 180 \\
x_1 + x_5 = 110 \\
\underline{x} \geq \underline{0}
\end{array}
$$
. Siccome il sistema dei vincoli è un sistema con 3 vincoli e 5 variabili, si sa già che la soluzione ottima sarà una soluzione di base con 3 variabili in base e 2 variabili fuori base: siccome quest'ultime sono per definizione = 0, allora avremo al più 3 variabili > 0; vogliamo capire quali sono queste variabili.

Supponiamo di aver dato il problema in pasto al Simplesso, e che quest'ultimo ci ha fornito la soluzione ottima.

Ora, supponiamo che un'altra azienda ci proponga di vendergli dei cereali: a quanto bisogna vendergli i cereali per non avere decrementato il profitto?

Prima di vedere ciò, analizziamo un problema più semplice. Eugenio suona il pianoforte e canta, quindi egli compra 100 CD che costano €1,00 cad, su cui incide la sua musica e vende ognuno di questi CD a €10,00 cad. Inoltre, se vende tutti i CD, ottiene un ricavo pari a €1000,00. Sara ha bisogno di un CD, su cui vuole registrare un album di foto dell'ultima vacanza fatta con il suo fidanzato. Sara non può comprare un CD, per cui chiede al suo amico Eugenio un CD: *a quanto Eugenio deve vendere il CD a Sara*? Distinguiamo i seguenti casi:

1) 0 euro. Sara è una sua amica, ed Eugenio è un signore;
2) 1 euro. Sara è una sua amica, ma Eugenio non vuole rimetterci il prezzo di acquisto;
3) 10 euro. Sara è appena una conoscente, che male c'è a guadagnarci qualcosa;
4) 11 euro. Ma quale amica?! La verità è che gli è sempre stata antipatica.
Delle quattro risposte, la giusta è la 3), in quanto Eugenio deve mantenere intatto il ricavo di €10,00.

Tornando al problema dei cereali, da un punto di vista di Ricerca Operativa dobbiamo venderli alla nuova azienda ad un prezzo (detto **prezzo ombra**) tale che la funzione obiettivo non cambi all'ottimo; quindi, come calcoliamo questo valore? Possiamo farlo attraverso la dualità: utilizzando questa Teoria, scopriamo che *i prezzi ombra non sono altro che le variabili del problema duale*.

Quindi, per rispondere a questa domanda risolviamo il problema duale:
$$\begin{array}{c} \min g=240w_1+180w_2+110w_3 \\ w_1+2w_2+w_3 \geq 560 \\ 1.5w_1+w_2 \geq 420 \\ w_1 \geq 0. w_2 \geq 0, w_3 \geq 0 \end{array}$$
. Si tenga presente che, quando risolviamo il primale, calcoliamo automaticamente anche la soluzione del duale: il valore della funzione obiettivo già lo conosciamo, la soluzione ottima del duale la calcoliamo con $\underline{w}^{*T} = \underline{c}_B^T A_B^{-1}$, quindi già conosciamo il vettore $\underline{w}^{*T}$ (ossia, i prezzi ombra associati alle risorse del problema primale).

Tornando all'esempio precedente:
- la soluzione ottima del problema primale (P) è
$$x_1^* = 15, x_2^* = 150, x_3^* = x_4^* = 0, x_5^* = 95.$$
Tenendo presente il problema primale (primo vincolo sui *cereali*, secondo vincolo sulla carne e terzo vincolo sulla raffinatrice), notiamo che il numero di variabili maggiori di 0 è 3. Analizzati i valori delle $\underline{x}^*$, e controllato il modello matematico, da un punto di vista dell'azienda – all'ottimo – le variabili di slack/surplus aggiunte al sistema ci indicano se quel vincolo è soddisfatto all'uguaglianza (cioè, se il vincolo è *attivo*) o meno. Questa è un'informazione molto importante, in quanto ci dice ad esempio che – all'ottimo – sono stati utilizzati i cereali (variabile x_3) interamente (risorsa **scarsa**). Per la carne, $x_4 = 0$, quindi anch'essa è stata interamente utilizzata, pertanto è una risorsa scarsa. Per quanto riguarda x_5, invece, poiché essa è > 0 otteniamo che la potenzialità della raffinatrice non è stata sfruttata al massimo (risorsa **abbondante**). Questo fatto implica che l'azienda deve investire sulle risorse scarse, in quanto sono quelle a vincolare di non poter migliorare l'attuale piano di produzione. Cambiando i termini noti del modello (associati a risorse scarse o abbondanti), va informata l'azienda di quanto cambia il profitto: cioè, indicare all'azienda, per ogni risorsa scarsa, di quanto aumenterà il suo profitto se si investe in quella risorsa;
- la soluzione ottima del problema duale (D) è
$$w_1^* = 140, w_2^* = 210, w_3^* = w_4^* = w_5^* = 0.$$
L'informazione dettagliata in precedenza ce la forniscono i prezzi ombra. Se si considerano i prezzi ombra, in particolare: w_1 e w_2 sono associati rispettivamente ai vincoli su cereali e carne, mentre w_3 è associata alla macchina raffinatrice. Si osserva che per la macchina raffinatrice $w_3 = 0$, dato che essa è una risorsa abbondante: se si modifica quel termine noto, non varia il profitto dell'azienda. Per i cereali e per la carne, siccome $w_1 = 140$ e $w_2 = 210$, al variare di questi valori varia anche il profitto dell'azienda, dato che queste sono risorse scarse. Infatti, se si incrementano di 1kg i cereali, allora si incrementa di 140 il profitto dell'azienda; analogamente, se si incrementa di 1kg la carne, allora si incrementa di 210 il profitto dell'azienda.

Vediamo adesso il significato delle $\underline{x}$ e delle $\underline{w}$, considerando la tabella a lato. Sulla prima riga è posta la soluzione ottima trovata, e sulle righe successive sono posti i valori ottimi al variare di cereali/carne.

Sull'ultima riga, si stimava che il profitto fosse di 71400 + 10 · 140 + 170 · 210 = 108500; in realtà, il profitto è di 100800 in quanto la variabile x_5 è saturata, e di conseguenza è stata cambiata la base ottima (cioè si è scelto x_1, x_2, x_4 anziché x_1, x_2, x_5); cambiata la base ottima, i $\underline{w}$ precedentemente calcolati non valgono più.

Cereali	Carne	x_1	x_2	x_3	x_4	x_5	w_1	w_2	w_3	w_4	w_5	$z^* = g^*$
240	**180**	15	150	0	0	95	140	210	0	0	0	71400
239	**180**	15,5	149	0	0	94,5	140	210	0	0	0	71260
241	**180**	14,5	151	0	0	95,5	140	210	0	0	0	71540
240	**179**	14,25	150,5	0	0	95,75	140	210	0	0	0	71190
240	**181**	15,75	149,5	0	0	94,25	140	210	0	0	0	71610
250	**200**	25	150	0	0	85	140	210	0	0	0	77000
250	**250**	110	93,3	0	36,6	0	280	0	280	0	0	100800

Il precedente esempio suggerisce che il prezzo ombra $\underline{w}$ è valido fin quando la base ottima del problema non cambia. Prima di poter fare queste stime, quindi, bisogna sapere con certezza *di quanto* si può cambiare un termine noto senza cambiare la base ottima. Useremo l'*analisi della sensitività* per calcolare questi range.

Riassumendo:
• Le variabili duali $\underline{w}$ rappresentano i "prezzi ombra", ovvero i prezzi minimi a cui bisogna vendere le risorse per mantenere invariato il valore ottimo della funzione obiettivo;
• I prezzi ombra sono validi fino a quando non viene cambiata la base ottima (quando ciò avviene devono essere ricalcolati);
• Quando un vincolo è attivo, la risorsa ad esso associata è scarsa. La variabile duale corrispondente, a meno di degenerazione, sarà diversa da zero. Se invece la risorsa è abbondante sicuramente la variabile duale ad essa associata è nulla.

Ma perché i prezzi ombra indicano il tasso di crescita/decrescita della funzione obiettivo al variare del termine noto b_i? Supponiamo di avere la soluzione di base ammissibile ottima $\underline{x}^* = \begin{bmatrix} \underline{x}_B^* \\ \underline{x}_N^* \end{bmatrix} = \begin{bmatrix} A^{-1}{}_B \underline{b} \\ \underline{0} \end{bmatrix}$, e il valore ottimo $\underline{c}^T \underline{x}^* = \underline{c}_B^T \underline{x}_B^* + \underline{c}_N^T \underline{x}_N^* = \underline{c}_B^T A_B^{-1} \underline{b}$, e per il Teorema forte della dualità $\underline{c}_B^T A_B^{-1} \underline{b} = \underline{w}^T \underline{b}$. Notiamo che $\underline{w}^T \underline{b} = w_1 b_1 + w_2 b_2 + \ldots + w_n b_n$. Da quest'ultima espressione si può notare che, incrementando un termine noto w_i di 1, la funzione obiettivo del primale varia di w_i.

È possibile avere un valore $w_i < 0$ associato ad un vincolo del problema, cioè in cui se si incrementa il termine noto b_i il valore della funzione obiettivo diminuisce?

Analisi della sensitività

Al termine dell'argomento precedente ci siamo chiesti se è possibile avere un prezzo ombra $w_i < 0$, in modo tale che incrementando di un'unità il termine noto b_i la funzione obiettivo diminuirebbe di w_i unità. Questo dipende dalla tipologia di vincolo: per i vincoli associati ad una risorsa (di ≤) questo non accade, ma per i vincoli *di domanda* (richiesti dall'azienda, ad esempio in cui questa vuole garantire che la produzione di una risorsa x_1 sia ≥ ad una certa quantità) questo sussiste.
Ad esempio, data x_1 = # kg di mangime da produrre, l'azienda vuole $x_1 \geq 30$. Se si ritrova il vincolo soddisfatto all'uguaglianza (cioè, è attivo), allora $w_1 \neq 0$: in particolare, se $w_1 < 0$ (es. $w_1 = -3$), allora se si incrementa il termine noto di questo vincolo di una unità (es. $x_1 \geq 31$), questo implica che la funzione obiettivo viene *decrementata* di 3 unità.

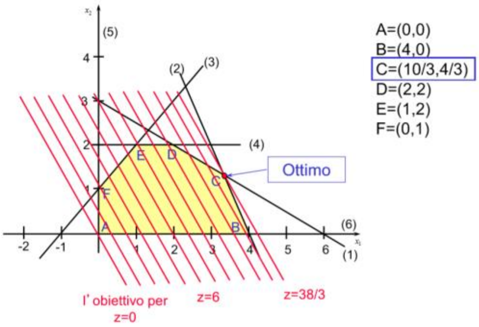
Le stime dei prezzi ombra restano valide fin quando la base ottima non cambia. L'**analisi della sensitività** studia i range entro i quali si può modificare un parametro del problema senza far variare la base ottima. Per farlo, introduciamo un esempio.

Esempio (Pianificare la produzione di una piccola azienda)**.** L'azienda produce due tipi di prodotti, il prodotto P_1 ed il prodotto P_2, usando due materie prime indicate con A e B. La disponibilità al giorno di materia prima A è pari a 6 ton, mentre quella di materia prima B è di 8 ton. La quantità di A e B consumata per produrre una ton di prodotto P_1 e P_2 è riportata nella tabella a lato.
Si ipotizza che tutta la quantità prodotta venga venduta. Il prezzo di vendita per tonnellata è Euro 3000 per P_1 e Euro 2000 per P_2. L'azienda ha effettuato un'indagine di mercato con i seguenti esiti:

		Prodotti	
		P_1	P_2
materie prime	A	1	2
	B	2	1

- la domanda giornaliera di prodotto P_2 non supera mai di più di 1 ton quella di prodotto P_1;
- la domanda massima giornaliera di prodotto P_2 è di 2 ton.

Problema: determinare le quantità dei due prodotti che debbono essere fabbricati giornalmente in modo da rendere massimo il guadagno.
→
<u>Definizione delle variabili.</u> Si introducono due variabili che rappresentano le quantità prodotte (e vendute) al giorno per P_1 e P_2 (ton): definiamo x_1 = # tonnellate del prodotto P_1 e x_2 = # tonnellate del prodotto P_2. Le due variabili sono continue.
<u>Definizione dell'obiettivo.</u> Il guadagno giornaliero (K€) è dato da $z = 3x_1 + 2x_2$. L'obiettivo è rappresentato da un'equazione lineare, in cui massimizziamo z: $\max z = 3x_1 + 2x_2$.
<u>Definizione dei vincoli.</u> Vincoli (tecnologici) sull'uso delle materie prime (l'uso giornaliero delle materie prime non può eccedere la disponibilità): (A) $x_1 + 2x_2 \leq 6$ e (B) $2x_1 + x_2 \leq 8$. Vincoli conseguenti le indagini di mercato: $-x_1 + x_2 \leq 1$ e $x_2 \leq 2$. Non negatività delle variabili: $x_1 \geq 0$ e $x_2 \geq 0$.
Questo problema si può risolvere graficamente.



Analizziamo la sensitività della soluzione. Notiamo che possono esserci le seguenti variazioni rispetto la disponibilità delle risorse:
(a) come aumentare le risorse per migliorare la soluzione ottima;
(b) come ridurre le risorse disponibili lasciando invariata la soluzione ottima.

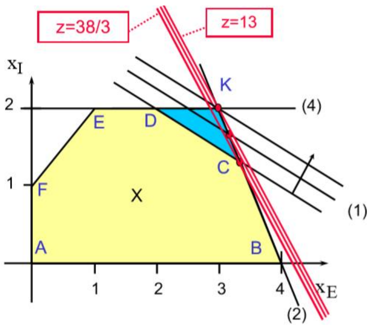
I vincoli del problema hanno tutti la forma quantità di risorsa usata ≤ disponibilità di risorsa, anche se solamente i vincoli (1) e (2) rappresentano effettivamente il consumo delle materie prime A e B.
Poiché i vincoli (1) e (2) sono soddisfatti all'uguaglianza dalla soluzione ottima corrispondente al punto C = (10/3, 4/3), il livello ottimo di produzione per i due prodotti è tale da utilizzare tutte le materie prime disponibili. I vincoli (1) e (2) sono **saturi**, quindi le materie prime A e B sono utilizzate completamente, ovvero sono risorse scarse. Dunque, *per capire quali sono le risorse scarse* basta verificare quali vincoli del primale sono soddisfatti all'uguaglianza. Più nello specifico, trasformando il problema in forma standard aggiungiamo le variabili di slack/surplus: se nel vincolo *i*-esimo la sua (eventuale) variabile di slack/surplus è nulla, allora il vincolo è saturo (o attivo), pertanto la risorsa associata al vincolo *i*-esimo è scarsa.

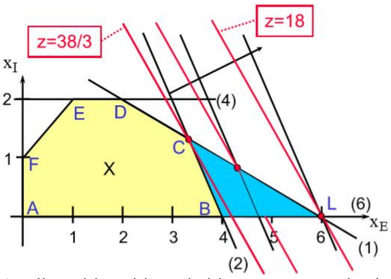
Vincolo saturo → Risorsa scarsa

Vincolo non saturo → Risorsa abbondante

In questo caso, i vincoli per le indagini di mercato (di domanda) rappresentano risorse abbondanti.
- È possibile *aumentare* la disponibilità di una risorsa *scarsa* per migliorare la soluzione ottima (caso (a));
- È possibile *diminuire* la disponibilità di una risorsa *abbondante* senza variare la soluzione ottima (caso (b)).

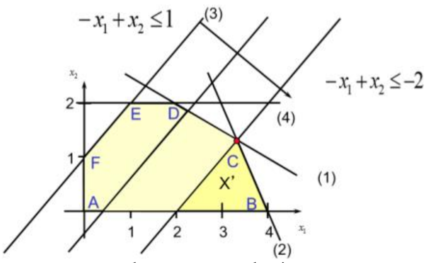
Verifichiamo graficamente sino a che livello ha senso aumentare la materia prima A.
In $\mathbb{R}^2$, incrementare il termine noto di un vincolo equivale ad effettuare la traslazione del vincolo nel verso del vettore costituito dai coefficienti associati alle sue variabili x_1 e x_2; decrementare il termine noto, invece, equivale ad effettuare la traslazione nel verso opposto.
Aumentando la risorsa A, il vincolo (1) si sposta e di conseguenza varia il punto di ottimo. Oltre K = (3, 2) (intersezione di (2) e (4)) non ha più senso aumentare la risorsa A. Il nuovo valore di A è 7.

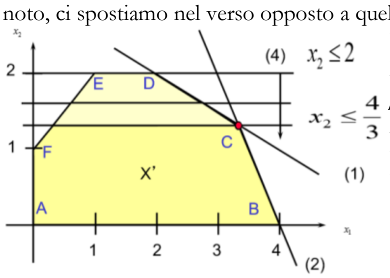




Analoga verifica può essere fatta per la materia prima B.
Aumentando la risorsa B, il vincolo (2) si sposta e di conseguenza varia il punto di ottimo. Oltre L =(6, 0) (intersezione di (1) e (6)) non ha più senso aumentare la risorsa B. Il nuovo valore di B è 12.

Analizzati i casi in cui si incrementano le risorse scarse, analizziamo ora i casi in cui si decrementano le risorse abbondanti. Supponendo i vincoli (3) e (4) relativi al consumo di due ulteriori risorse abbondanti, è possibile verificare di quanto diminuirne la disponibilità senza modificare la soluzione ottima. Per il vincolo (3), a destra, siccome decrementiamo il termine noto, ci spostiamo nel verso opposto a quello del vettore (-1, 1).





Possiamo effettuare la traslazione *fino al punto di ottimo* (C), ed in questo caso decrementare la risorsa non influisce sulla risoluzione del problema: si risparmiano risorse non utilizzate.
Per il vincolo (4), a sinistra, ci spostiamo verso il basso, anche in questo caso fino al punto C.

La traslazione del vincolo va effettuata fin quando tale vincolo non diventa ridondante, non annulla la regione ammissibile o fin quando non interseca il punto di ottimo calcolato.

Dopo aver verificato la convenienza di una possibile maggiore disponibilità delle risorse A e B, è interessante determinare quale sia la risorsa che di più convenga aumentare. Nell'esempio, l'azienda potrebbe avere una limitata disponibilità finanziaria che vorrebbe far fruttare al meglio, acquisendo un'ulteriore quantità di una delle risorse in modo da incrementare maggiormente i propri profitti. Questa informazione è ottenibile per mezzo della Programmazione Lineare.

Si può calcolare il **valore di una unità di risorsa** w_i attraverso la formula

$$\boxed{w_i = \frac{\text{massima variazione di } z}{\text{massima variazione della risorsa } i}},$$

dove il termine "massima" è indicato al fatto che la base ottima non deve cambiare.

Per la risorsa A: $w_A = \frac{13-\frac{38}{3}}{7-6} = \frac{39-38}{3} = \frac{1}{3}$ (K€/ton).

Per la risorsa B: $w_B = \frac{18-\frac{38}{3}}{12-8} = \frac{\frac{(54-38)}{3}}{4} = \frac{4}{3}$ (K€/ton).

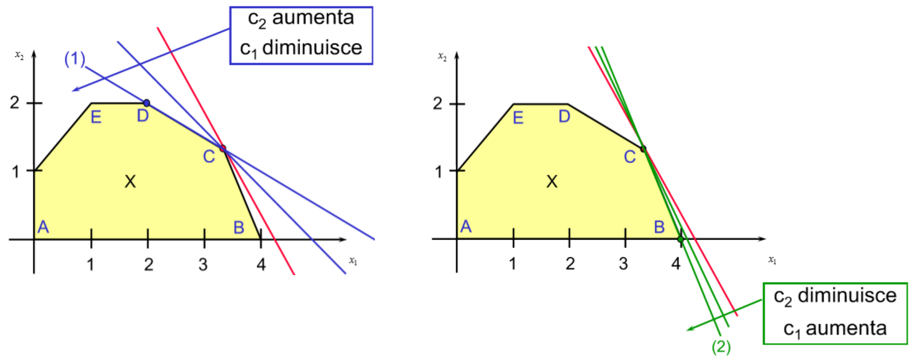
La quantità w_i indica di quanto aumenta l'obiettivo in corrispondenza dell'acquisizione di un'ulteriore unità di risorsa. È evidente come nell'esempio l'incremento unitario migliore è associato alla risorsa B.

Cosa succede, invece, nel caso in cui non modifichiamo la regione ammissibile e semplicemente variamo il prezzo di vendita dei prodotti? Verifichiamo cosa succede geometricamente. Si tratta di analizzare entro quali limiti di tolleranza possono variare i prezzi di vendita senza alterare la base ottima (la produzione associata al punto C).

Sappiamo che X è la regione ammissibile, e che C è il punto di ottimo; inoltre, sappiamo che la funzione obiettivo è $\max z = 3x_1 + 2x_2$, quindi $c_1 = 3$ e $c_2 = 2$ (vettore $\underline{c}$).

Decrementando c_1 di 2 unità ($c_1' = c_1 - 2 = 1$), il gradiente è modificato e le curve di livello cambiano la loro inclinazione. In questo caso, il nuovo punto di ottimo sarà D.

In generale, variando c_1 e/o c_2 cambia la pendenza della funzione obiettivo. Quando si decide di modificare i coefficienti di costo (prezzi di vendita, in questo esempio) bisogna sempre far attenzione a non modificare la base ottima. In questo esempio, si può effettuare la rotazione in senso antiorario fin quando le curve di livello non si sovrappongono al segmento CD. Analogamente, si può effettuare la rotazione in senso orario fin quando le curve di livello non si sovrappongono al segmento CB.



✓

Formalmente, dato un problema di programmazione lineare

$$\min z = \underline{c}^T \underline{x}$$
$$A\underline{x} = \underline{b}$$
$$\underline{x} \geq \underline{0},$$

e data la soluzione ottima $\underline{x}^*$ e la base ottima associata B, determinare come sia possibile variare certe caratteristiche del problema lasciando invariata la base ottima.

Distinguiamo cinque casi: 1) variazione nel vettore dei costi $\underline{c}$; 2) variazione nel vettore dei termini noti $\underline{b}$; 3) variazione nella matrice di vincoli A; 4) aggiunta di una nuova variabile; 5) aggiunta di un nuovo vincolo. Concentriamoci sui primi due casi.

Caso 1: *variazione nel vettore dei costi $\underline{c}$*. Data una soluzione di base ottima $\underline{x}^*$ (sia B la base associata a tale soluzione), supponiamo che il coefficiente di una delle variabili sia cambiato da c_k a c_k'. L'effetto di questo cambio si ripercuoterà solo sui coefficienti di costo ridotto.
Bisogna considerare i seguenti due casi:
caso 1.1) variazione di un coefficiente di costo relativo ad una variabile non in base;
caso 1.2) variazione di un coefficiente di costo relativo ad una variabile in base.

Caso 1.1. Teniamo a mente lo specchietto a destra. Cosa succede se si modifica una componente del vettore $\underline{c}$? Sia c_k $(k \in N)$, il coefficiente viene modificato come segue:

$$c_k' = c_k + \delta.$$

B base ottima
- $\underline{x}_B = A_B^{-1}\underline{b} \geq 0$ $\boxed{\text{AMM.}}$
- $z_j - c_j \leq 0 \ \forall j \in N$ $\boxed{\text{OTT.}}$

$\underline{c}_B^T A_B^{-1} \underline{a}_j - c_j \leq 0$

In questo caso, $\underline{c}_B^T$ non subisce variazioni, e quindi $z_j = \underline{c}_B^T A_B^{-1} \underline{a}_j$ rimane inalterato per ogni j in N; viene modificato, quindi, il solo c_k. La nuova formula per il calcolo di $z_k - c_k$ sarà

$$z_k - c_k' = z_k - (c_k + \delta) = (z_k - c_k) - \delta \leq 0 \Leftrightarrow \delta \geq z_k - c_k.$$

Dunque, preso il coefficiente di costo c_k di una variabile fuori base, lo si può modificare considerando un δ (da sommare c_k) maggiore o uguale di $z_k - c_k$.
Se $z_k - c_k' \leq 0$, allora $\underline{x}^*$ è ancora la soluzione ottima. Se invece $z_k - c_k' > 0$, allora $\underline{x}^*$ non è più la soluzione ottima e quindi occorre effettuare un'iterazione del simplesso per far entrare in base la variabile x_k.

Caso 1.2. Sia c_{B_i} $(i \in B)$ il coefficiente di costo che viene modificato in

$$c_{B_i}' = c_{B_i} + \delta.$$

In questo caso, la variazione si ripercuote su z_j, cioè implica la variazione di tutti i coefficienti di costo ridotto associati alle variabili fuori base. In particolare, si ha che:

$c_{B_i}' = c_{B_i} + \delta \Rightarrow \underline{c}_B' = \underline{c}_B + \delta \underline{e}_i$, dove $\underline{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, con l'*i*-esimo elemento uguale a 1 ed i restanti uguali a 0.

La nuova formula sarà

$$z_j' - c_j = (\underline{c}_B^T + \delta \underline{e}_i^T) A_B^{-1} \underline{a}_j - c_j = \underline{c}_B^T A_B^{-1} \underline{a}_j + \delta \underline{e}_i^T A_B^{-1} \underline{a}_j - c_j = (z_j - c_j) + \delta (A_B^{-1})^i \underline{a}_j \leq 0,$$

in quanto $\underline{e}_i^T A_B^{-1} = (A_B^{-1})^i$ è la riga *i*-esima di A_B^{-1}, in quanto $[0 \ \cdots \ 1 \ \cdots \ 0] A_B^{-1} = (A_B^{-1})^i$.

Caso 2: *variazione del termine noto di un vincolo*. Sia b_i il termine noto dell'*i*-esimo vincolo che viene variato in

$$b_i' = b_i + \delta \Rightarrow \underline{b}' = \underline{b} + \delta \underline{e}_i.$$

A causa di tale variazione si modificano i valori delle variabili di base:

$$\underline{x}_B' = A_B^{-1} \underline{b}' = A_B^{-1}(\underline{b} + \delta \underline{e}_i) = A_B^{-1} \underline{b} + \Rightarrow \underline{x}_B' = \underline{x}_B + \delta (A_B^{-1})_i,$$

dove $A_B^{-1} \underline{e}_i = (A_B^{-1})_i$ è la colonna *i*-esima di A_B^{-1}.
Le condizioni su δ, quindi, si ottengono imponendo che

$$\underline{x}_B' = \underline{x}_B + \delta (A_B^{-1})_i \geq \underline{0}.$$

In sintesi, il caso 1 influisce sull'*ottimalità* della soluzione, mentre il caso 2 influisce sull'*ammissibilità* della stessa.

Nella pagina seguente è posto un esempio sull'analisi della sensitività.

Esempio (Analisi di sensitività)**.**
Dato il problema di PL posto a destra, calcoliamo la base ottima, la matrice A_B^{-1}, i coefficienti di costo ridotto e il valore della soluzione ottima.

$$\begin{aligned} &\min -2x_1 + x_2 - x_3 \\ &x_1 + x_2 + x_3 \le 6 \\ &-x_1 + 2x_2 \le 4 \\ &\underline{x} \ge 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} &\min -2x_1 + x_2 - x_3 \\ &x_1 + x_2 + x_3 + x_4 = 6 \\ &-x_1 + 2x_2 + x_5 = 4 \\ &\underline{x} \ge 0 \end{aligned}$$

Base ottima: $B = \{1,5\}$; $N = \{2,3,4\}$; $A_B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$; $A_B^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

Infatti:

$$z_2 - c_2 = -3; \quad z_3 - c_3 = -1; \quad z_4 - c_4 = -2$$

$$\underline{x}_B = A_B^{-1}\underline{b} = \begin{bmatrix} x_1 \\ x_5 \end{bmatrix} = \begin{bmatrix} 6 \\ 10 \end{bmatrix}; \quad z^* = \underline{c}_B^T A_B^{-1} \underline{b} = -12$$

Caso 1.1. Di quanto può variare il coefficiente c_2 prima di cambiare la base ottima?
$z_2 - c_2' = (z_2 - c_2) - \delta \le 0 \rightarrow -3 - \delta \le 0$
$3 - \delta \le 0 \rightarrow \delta \ge -3$.
Verifichiamo cosa succede se scegliamo $\delta = -4$.
Poiché x_2 non è in base, il valore $z_j = \underline{c}_B^T A_B^{-1} \underline{a}_j$ non cambia per nessun indice j in N. L'unico coefficiente di costo ridotto che cambia è $z_2 - c_2' = (z_2 - c_2) - \delta = -3 + 4 = 1 > 0$. Poiché vale questo, la soluzione non è più ottima. Bisogna far entrare in base x_2.

Caso 1.2. Di quanto può variare il coefficiente c_1 prima di cambiare la base ottima?

$$z'_j - c_j = (z_j - c_j) + \delta (A_B^{-1})^i \underline{a}_j \le 0 \quad \forall j \in N \qquad A_B^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$z'_2 - c_2 = (z_2 - c_2) + \delta \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \le 0 \Rightarrow -3 + \delta \le 0 \Rightarrow \delta \le 3$$

$$z'_3 - c_3 = (z_3 - c_3) + \delta \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \le 0 \Rightarrow -1 + \delta \le 0 \Rightarrow \delta \le 1$$

$$z'_4 - c_4 = (z_4 - c_4) + \delta \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \le 0 \Rightarrow -2 + \delta \le 0 \Rightarrow \delta \le 2$$

Verifichiamo, ora, cosa succede se scegliamo un $\delta > 1$, per esempio $\delta = 2$.
Poiché x_1 è in base, il valore z_j cambia per ciascun j in N secondo la relazione: $z_j' - c_j = (z_j - c_j) +$
$+ \delta (A_B^{-1})^i \underline{a}_j \le 0$. Dunque:

$$z'_2 - c_2 = -3 + 2\begin{bmatrix} 1 , 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = -3 + 2 \times 1 = -1$$

$$z'_3 - c_3 = -1 + 2\begin{bmatrix} 1 , 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = -1 + 2 \times 1 = 1$$

$$z'_4 - c_4 = -2 + 2\begin{bmatrix} 1 , 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = -2 + 2 \times 1 = 0$$

Caso 2. Di quanto può variare al più il termine noto b_1 prima di rendere inammissibile la base ottima?

$$\underline{x}_B' = A_B^{-1} \underline{b}' = A_B^{-1} (\underline{b} + \delta \underline{e}_i) = A_B^{-1} \underline{b} + \delta (A_B^{-1})_i \Rightarrow \underline{x}_B' = \underline{x}_B + \delta (A_B^{-1})_i$$

Dunque, otteniamo:

$$\underline{x}'_B = \begin{bmatrix} x_1 \\ x_5 \end{bmatrix} + \delta \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 10 \end{bmatrix} + \begin{bmatrix} \delta \\ \delta \end{bmatrix} = \begin{cases} 6 + \delta \ge 0 \Rightarrow \delta \ge -6 \\ 10 + \delta \ge 0 \Rightarrow \delta \ge -10 \end{cases}$$

Se scegliamo $\delta = -7$:

$$\underline{x}'_B = \begin{bmatrix} x_1 \\ x_5 \end{bmatrix} - 7 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 10 \end{bmatrix} + \begin{bmatrix} -7 \\ -7 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

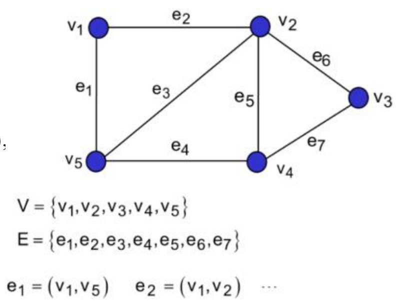
✓

Teoria dei grafi e problema del flusso a costo minimo

Un **grafo non orientato** G = (V, E) è dato da una coppia di insiemi finiti, dove:
- $V = \{v_1, ..., v_n\}$ è l'insieme degli *n* **nodi** di G;
- $E = \{e_1, ..., e_m\} \subseteq V \times V$ è l'insieme degli *m* **archi non orientati** di G.

Ogni arco non orientato di G corrisponde ad una *coppia non ordinata* di nodi di G $e_k = (v_i, v_j)$. I nodi v_i e v_j sono gli estremi dell'arco e_k.
La presenza di un arco tra una coppia di nodi indica una relazione tra i nodi stessi.

Seguono alcune definizioni di base:
- un arco (v, v) è detto **loop**;
- un arco e = (u, v) ∈ E si dice **incidente** su u e v;
- due nodi u, v ∈ V sono detti **adiacenti** ⇔ (u, v) ∈ E;
- due archi $e_1, e_2 \in E$ sono detti **adiacenti** ⇔ $e_1 = (u, v)$ e $e_2 = (v, w)$, cioè se hanno un nodo in comune;
- l'insieme di nodi **N(u)** = {v ∈ V: v adiacente a u} è detto **intorno** di u in G;
- l'insieme di archi **δ(u)** = {e ∈ E: e incide su u} è detto **stella** di u in G;
- |δ(u)| è detto **grado** del nodo u.

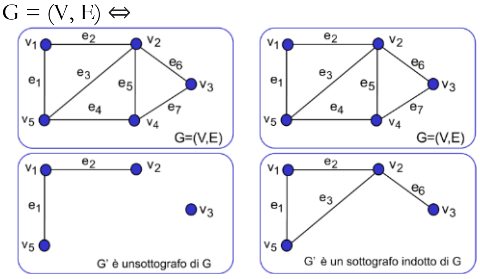


I grafi sono un mezzo per rappresentare relazioni binarie. Ad esempio:
- due città connesse da una strada;
- due calcolatori connessi in una rete telematica;
- due persone legate da una relazione di parentela (come, padre-figlio);
- due persone che condividono una stanza;
- il collegamento tra due componenti elettronici;
- un'operazione che deve essere eseguita da una certa macchina;
- ...

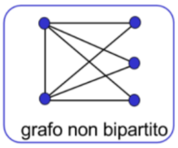
I grafi possono essere usati come strumento per modellare in maniera schematica un vastissimo numero di problemi decisionali. Ad esempio:
- determinare il percorso più breve che connette due città;
- determinare come connettere nella maniera più economica (più efficiente) un insieme di calcolatori in una rete telematica;
- assegnare un insieme di operazioni ad un insieme di macchine;
- determinare il percorso più conveniente da far percorrere ad una flotta di veicoli commerciali per effettuare delle consegne e quindi rientrare al deposito;
- ...

Un grafo è detto **semplice** se non esistono archi paralleli (al più un arco per ogni coppia di nodi) o "loop".
Un grafo G' = (V', E') è detto **sottografo** di G = (V, E) ⇔
- V' ⊆ V;
- E' ⊆ E, e $(v_i, v_j) \in E' \Rightarrow v_i, v_j \in V'$.

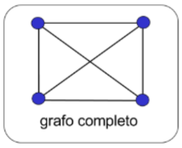
Un grafo G' = (V', E') è detto **sottografo indotto** da V' in G = (V, E) ⇔
- V' ⊆ V;
- ∀u, v ∈ V', se (u, v) ∈ E allora (u, v) ∈ E'.



Un grafo G è detto **bipartito** se esiste una partizione di V in due sottoinsiemi V_1 e V_2 tali che:
- $V_1 \cap V_2 = \emptyset$;
- $V_1 \cup V_2 = V$;
- $\forall e = (u, v) \in E$, se $u \in V_1$ allora $v \in V_2$ oppure se $u \in V_2$ allora $v \in V_1$.

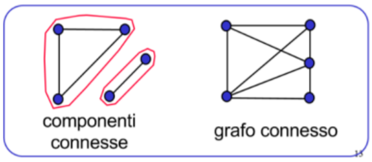


Un grafo G è detto **completo** se e solo se contiene tutti i possibili archi, ovvero $|\delta(v)| = n - 1 \ \forall v \in V$.
Il numero di archi in un grafo completo è $\binom{n}{2} = \frac{n(n-1)}{2}$.

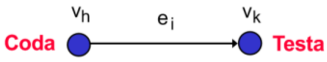


Dato un grafo G = (V, E), un nodo $v \in V$ si dice **connesso** ad un nodo $u \in V$ se esiste un cammino tra u e v in G. Si hanno le seguenti proprietà:
- $v \in V$ è connesso a v (riflessività);
- $v \in V$ è connesso a $u \in V \Rightarrow u \in V$ è connesso a $v \in V$ (simmetria);
- $v \in V$ è connesso a $u \in V$ e $u \in V$ è connesso a $w \in V \Rightarrow v \in V$ è connesso a $w \in V$ (transitività);
- un grafo G = (V, E) è connesso se e solo se tutti i suoi nodi sono connessi tra loro.

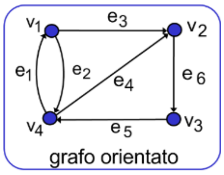
L'insieme V può essere partizionato in sottoinsiemi $C_i = \{v \in V: v \text{ è connesso a } u, \forall u \in C_i\}$. Il sottografo indotto da C_i in G è detto **componente connessa** di G.
G è connesso se e solo se possiede una sola componente connessa (v è connesso a u, $\forall u, v \in V$).



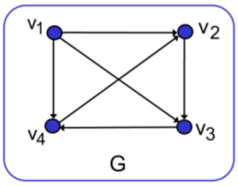
Un grafo G = (V, E) è detto **orientato** se, dato $V = \{v_1, ..., v_n\}$, l'insieme degli archi $E = \{e_1, ..., e_m\}$ è formato da coppie ordinate di nodi.
Per un grafo orientato, si ha che $e_i = (v_h, v_k) \neq e_j = (v_k, v_h)$, con $e_i, e_j \in E$.
L'arco e_i si dice **uscente** da v_h ed **entrante** in v_k.

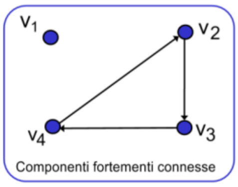


Es. Dato il grafo posto a destra:
- $Fs(v) = \{u \in V: (v, u) \in E\}$ è detto **stella uscente** di v;
- $Bs(v) = \{u \in V: (u, v) \in E\}$ è detto **stella entrante** di v;
- $S(v) = Fs(v) \cup Bs(v)$ è detto **stella** di v;
- le definizioni di sottografo, sottografo indotto e componente connessa di un grafo orientato sono analoghe a quelle date per i grafi non orientati.



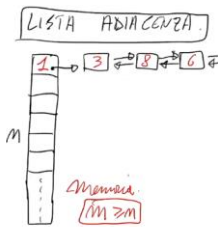
Dato G = (V, E), un nodo $v \in V$ si dice **fortemente connesso** ad un nodo $u \in V$ se esiste una *path* (cammino orientato) tra v e u in G. Si hanno le seguenti proprietà:
- $v \in V$ è connesso a v (riflessività);
- $v \in V$ è fortemente connesso a $u \in V$ e $u \in V$ è fortemente connesso a $w \in V \Rightarrow v \in V$ è fortemente connesso a $w \in V$ (transitività);
- un grafo G = (V, E) è fortemente connesso se e solo se tutti i suoi nodi sono fortemente connessi tra loro.

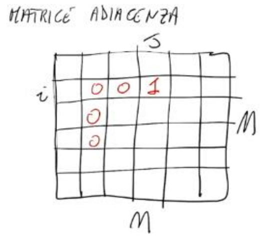




Un grafo *non orientato* è rappresentabile in tre modi:
- **lista di adiacenza**, ad ogni vertice è associata la lista dei vertici adiacenti (può essere una tabella o una lista concatenata);
- **matrice di adiacenza** (nxn), dove $a_{ij} = 1$ se $(v_i, v_j) \in E$, $a_{ij} = 0$ altrimenti;
- **matrice di incidenza** (nxm), dove $a_{ij} = 1$ se $v_i \in e_j$, $a_{ij} = 0$ altrimenti.

La lista di adiacenza è costituita da un insieme di *n* nodi che contengono: informazioni circa il nodo del grafo (grado, peso, colore, ...), e un puntatore all'insieme di vertici adiacenti al vertice considerato.
Per grafi non orientati, nella matrice di incidenza ogni colonna rappresenta un arco del grafo. Siccome gli archi del grafo hanno due estremi, vuol dire che su una colonna ci saranno *2* uni (sulla riga degli estremi corrispondenti) e *n – 2* zeri.

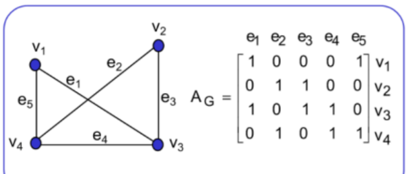




Nella scelta del modo in cui rappresentare un grafo figurano i seguenti vantaggi e svantaggi:
- *lista di adiacenza*, memoria O(n).
  - Vantaggi: permette di scorrere i nodi adiacenti a v in O(grado(v));
  - Svantaggi: inserimenti e cancellazioni su liste concatenate in O(grado(v)).
- *matrice di adiacenza*, memoria O(n^2).
  - Vantaggi: inserimenti e cancellazioni in O(1);
  - Svantaggi: permette di scorrere i nodi adiacenti a v in O(n).
- *matrice di incidenza*, memoria O(n · m). [Fonte: *https://it.wikipedia.org/wiki/Grafo_(tipo_di_dato_astratto)#Implementazioni*]
  - Vantaggi: inserimenti in O(n · m), in quanto basta aggiungere una colonna alla fine;
  - Svantaggi: cancellazioni in O(n) (copiando l'ultima colonna nella colonna da eliminare, ed eliminando l'ultima colonna). Permette di scorrere gli archi incidenti a v in O(m).

Delle tre utilizzeremo la matrice di incidenza, in quanto ci dà delle informazioni fondamentali.
Sia G = (V, E) un grafo *non orientato* con |V| = n ed |E| = m.
Denotiamo con A = [a_{ij}], con i = 1, ..., n e j = 1, ..., m, la **matrice di incidenza** di G, dove:

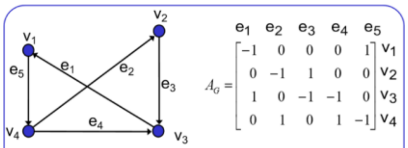
$$a_{ij} = \begin{cases} 1 & \text{se } v_i \text{ è un estremo di } e_j \\ 0 & \text{altrimenti} \end{cases}.$$



matrice di incidenza di un grafo non orientato

Sia G = (V, E) un grafo *orientato* con |V| = n ed |E| = m.
Denotiamo con A = [a_{ij}], con i = 1, ..., n e j = 1, ..., m, la **matrice di incidenza** di G, dove:

$$a_{ij} = \begin{cases} 1 & \text{se } v_i \text{ è coda di } e_j \text{ (arco uscente da } v_i) \\ -1 & \text{se } v_i \text{ è testa di } e_j \text{ (arco entrante in } v_i) \\ 0 & \text{altrimenti} \end{cases}.$$



matrice di incidenza di un grafo orientato

Esempio (Problema del Flusso a Costo Minimo)**.** Sia G = (V, E) un grafo connesso e orientato in cui:
• Ad ogni arco (i, j) è associato un costo c_{ij} che rappresenta il costo da pagare per ogni unità di flusso che transita sull'arco (i, j).
• Ad ogni vertice v ∈ V è associato un valore <u>intero</u> b_v dove:
  - $b_v > 0$ indica che il nodo v è un nodo di offerta;
  - $b_v < 0$ indica che il nodo v è un nodo di domanda;
  - $b_v = 0$ indica che il nodo v è un nodo di passaggio (non è né un nodo di offerta né un nodo di domanda).
• La somma di tutti i b_v deve essere uguale a zero (condizione di bilanciamento). Ciò che viene prodotto dalle sorgenti viene consumato dalle destinazioni.

Nel problema del flusso a costo minimo bisogna far giungere la merce prodotta (dai nodi di offerta) alle destinazioni (nodi di domanda) minimizzando i costi di trasporto.

Abbiamo bisogno di scrivere il modello matematico di questo problema.
Bisogna decidere *quanto flusso far transitare sugli archi*, quindi le **variabili decisionali** di questo problema sono le x_{ij} = quantità di flusso sull'arco (i, j). N.B.: Per quasi tutti i problemi di ottimizzazione che vedremo fino al termine del corso, le variabili sono sempre queste.
La **funzione obiettivo** sarà $\min \sum_{(i,j)\in E} c_{ij}x_{ij}$. Bisogna effettuare il prodotto $c_{ij}x_{ij}$ in quanto, se decido che sull'arco (i, j) deve passare un "veicolo", pago una sola volta il casello c_{ij}; se decido che sull'arco (i, j) devono passare dieci "veicoli", allora pago dieci volte il casello c_{ij}.
Per quanto riguarda i **vincoli**, bisogna garantire ogni mattina al cliente che arrivi la merce. Per imporre questo, usando x_{ij}, bisogna garantire che la somma del flusso degli archi uscenti sia uguale a b_v: $\sum_{j\in FS(i)} x_{ij} = b_v$. Tuttavia, il nodo *i* potrebbe ricevere altre informazioni oltre a queste che deve far uscire, per cui il vincolo dev'essere costruito in modo tale che $\sum_{j\in FS(i)} x_{ij} - \sum_{k\in BS(i)} x_{ki} = b_i \quad \forall i \in V$. I vincoli saranno sempre legati ai vertici del grafo, e servono per stabilire il ruolo che quel nodo ha sul grafo. In sintesi:
- se $b_v > 0$, allora la differenza tra il flusso che esce e quello che entra dev'essere positiva, cioè sto introducendo merce sulla rete (quindi sono un nodo di offerta);
- se $b_v < 0$, allora la differenza tra il flusso che esce e quello che entra dev'essere negativa, cioè sto acquisendo merce dalla rete (quindi sono un nodo di domanda);
- se $b_v = 0$, allora la quantità di flusso che entra dev'essere esattamente pari alla quantità di flusso che esce (quindi sono un nodo di passaggio).

Il **modello matematico** risultante è posto a destra. Nei problemi che affronteremo, i vincoli saranno quasi sempre questi.
In tale modello:
- x_{ij} = quantità di flusso sull'arco (i, j);
- c_{ij} = costo di trasporto di un'unità di flusso sull'arco (i, j);
- b_i = intero associato al nodo i:
  - se $b_i > 0$: nodo offerta
  - se $b_i < 0$: nodo domanda
  - se $b_i = 0$: nodo di passaggio

$$\min \sum_{(i,j)\in A} c_{ij}x_{ij}$$
$$\text{s.t.}$$
$$\sum_{j\in FS(i)} x_{ij} - \sum_{k\in BS(i)} x_{ki} = b_i \qquad i = 1,\dots n$$
$$x_{ij} \geq 0 \qquad \mathrm{i} \in \mathrm{A}$$

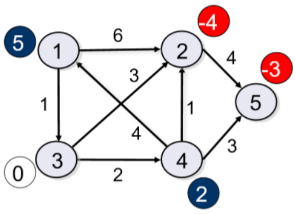
Possiamo sempre scrivere il problema in forma matriciale. Si presti attenzione sulla matrice A: siccome per i grafi specifichiamo *n* vertici e *m* archi, tale matrice non sarà mxn, bensì **nxm** (cioè, *n* vincoli ed *m* variabili). Useremo la matrice di incidenza, pertanto abbiamo tante colonne quanti sono gli archi del grafo. Inoltre, la variabile x_{ij} compare esattamente 2 volte: una volta come arco uscente dal nodo i (prima sommatoria), e una volta come arco entrante del nodo j (seconda sommatoria).
Osservazioni:
1. La matrice A è la matrice di incidenza nodo-arco con dimensione nxm. Ogni colonna $\underline{a}_{ij}$ è associata all'arco (i, j), ed in particolare abbiamo che: $\underline{a}_{ij} = \underline{e}_i - \underline{e}_j$ (si ricordi che $\underline{e}_i$ è il vettore colonna con tutti 0 eccetto un 1 in posizione *i*-esima).
2. Il rango di questa matrice è: $r(A) = n - 1$.

L'osservazione 1 è importante. Esistono alcune eccezioni di problemi di PLI (variabili intere, quindi teoricamente difficili) risolvibili mediante Simplesso. Supponiamo di avere il problema $\begin{array}{c} \max \frac{1}{2}x_1 + \frac{1}{2}x_2 \\ x_1 \leq \frac{3}{2} \\ x_2 \leq \frac{3}{2} \\ x_1, x_2 \geq 0, \text{ intere} \end{array}$. Se lo si disegna graficamente, si ottiene come regione ammissibile il quadrato costituito dai punti (0, 0), (0, 1), (1, 0), (1, 1), in quanto bisogna considerare i soli valori interi. Se fosse un problema di PL, allora il punto di ottimo sarebbe (3/2, 3/2), ma in questo caso è il punto (1, 1). Analizzando la regione ammissibile, l'ottimo del problema di PL coincide con l'ottimo del PLI quando *tutti* i vertici del poliedro hanno coordinate intere, e non solo quello ottimo. Il problema del flusso a costo minimo gode di questa proprietà. Di conseguenza, il problema è comunque facile. Questa operazione viene chiamata informalmente *rilassamento del vincolo di interezza sulle* x_{ij}.

Per ora, sull'osservazione 2 ci limitiamo a dire che la matrice di incidenza A sicuramente non è a rango pieno: infatti, sommandone le righe si ottiene il vettore nullo, per cui esse non sono linearmente indipendenti.

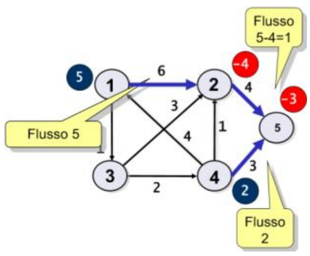
Come sarà fatta una soluzione di base del nostro problema? Sapendo che $\underline{x} = \begin{array}{c c} \begin{bmatrix} A_B^{-1}\underline{b} \\ \underline{0} \end{bmatrix} & \begin{array}{c} m \\ n-m \end{array} \end{array}$, si può dimostrare che $A_B^{-1}\underline{b}$ (se $\underline{b}$ è costituito da componenti intere) dà sempre come risultato un vettore di componenti intere. Per farlo, si osservi che A (matrice di incidenza) gode della *totale unimodularità*, cioè presa una qualsiasi sottomatrice quadrata di A di qualunque dimensione allora il suo determinante è uguale a 1, -1 o 0.

Segue un esempio.
Consideriamo un grafo orientato G = (V, E) rappresentante una rete di trasporto. L'obiettivo è quello di far viaggiare ("fluire"), al minimo costo, determinate quantità di merce (unità di flusso) dai nodi di **offerta** a quelli di **domanda** (eventualmente passando per dei nodi di **passaggio**). Abbiamo:
- una quantità $\mathbf{b_i} > \mathbf{0}$ per i nodi offerta, $< \mathbf{0}$ per i nodi domanda, $= \mathbf{0}$ per i nodi di passaggio (quantità di offerta/domanda);
- un costo $\mathbf{c_{ij}} \geq \mathbf{0}$ per ogni arco (costo per il trasporto di una unità di merce).



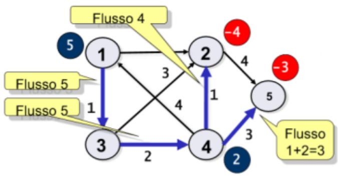
Nel grafo posto a destra abbiamo, quindi:
- i nodi 1 ($b_1 = 5 > 0$) e 4 ($b_4 = 2 > 0$) sono nodi di offerta;
- il nodo 3 ($b_3 = 0$) è un nodo di passaggio;
- i nodi 2 ($b_2 = -4 < 0$) e 5 ($b_5 = -3 < 0$) sono nodi di domanda.

Una soluzione può essere la seguente. Supponiamo di inviare 5 unità di flusso dal nodo 1 al nodo 2, in quanto il nodo 1 ha come capacità produttiva 5. Dopodiché, poiché il nodo 2 riceve 5 unità di flusso (ed esso ne richiede solo 4), allora ci sarà un'altra unità di flusso che dal nodo 2 dev'essere instradata nella rete: in questo caso, tale unità deve passare forzatamente al nodo 5. Successivamente, vengono inviate due unità di flusso dal nodo 4 al nodo 5, e il risultato è una soluzione ammissibile per il problema: infatti, tutti i nodi di offerta hanno inviato la merce prodotta, tutti i nodi di domanda hanno ricevuto la merce richiesta, e il costo di questa soluzione è dato da $costo = (6 \cdot 5) + (4 \cdot 1) + (3 \cdot 2) = 40$.



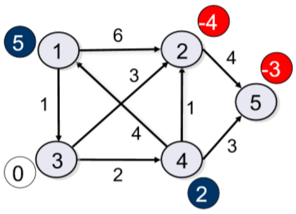
Una soluzione può essere la seguente:
- 5 unità di flusso dal nodo 1 al nodo 3;
- 5 unità di flusso dal nodo 3 al nodo 4;
- 1+2 = 3 unità di flusso dal nodo 4 al nodo 5;
- 4 unità di flusso dal nodo 4 al nodo 2.

Il costo sarà $costo = (5 \cdot 1) + (5 \cdot 2) + (3 \cdot 3) + (4 \cdot 1) = 28$.



Modelliamo il problema: consideriamo una variabile $x_{ij} \geq 0 \quad \forall (i,j) \in E$, rappresentante la *quantità di flusso* che attraverserà tale arco nella soluzione.

$$\min 6x_{12} + x_{13} + 4x_{25} + 3x_{32} + 2x_{34} + 4x_{41} + x_{42} + 3x_{45}$$
$$s.t.$$
$$x_{12} + x_{13} - x_{41} = 5$$
$$x_{25} - x_{12} - x_{32} - x_{42} = -4$$
$$x_{32} + x_{34} - x_{13} = 0$$
$$x_{41} + x_{42} + x_{45} - x_{34} = 2$$
$$-x_{25} - x_{45} = -3$$



Funzione obiettivo. $\min \sum_{(i,j)\in E} c_{ij}x_{ij}$, ossia minimizzare la somma tra tutti i prodotti costo per quantità di flusso associata all'arco (x_{ij}), considerando tutti gli archi del grafo (i, j) ∈ E.
Vincoli. Il numero di vincoli è pari al numero di nodi; in questo caso, 5. Il vincolo *i*-esimo (quindi, associato al vertice *i*) è dato dall'equazione $\sum_{j\in Fs(i)} x_{ij} - \sum_{k\in Bs(i)} x_{ki} = b_i$, dove semplicemente va sommata (se esiste) la quantità di flusso uscente da *i* verso ogni altro vertice *j*, rimossa (se esiste) la quantità di flusso entrante in *i* da ogni altro vertice *j*, il tutto uguagliato all'unità di flusso b_i associata al vertice *i*. Non vanno utilizzati i costi associati agli archi all'interno dei vincoli.

Ad esempio, $x_{12} + x_{13} - x_{41} = 5$ in quanto (1, 2) e (1, 3) sono archi uscenti da 1, (4, 1) è un arco entrante in 1, e 5 è la quantità di flusso associata al vertice 1.

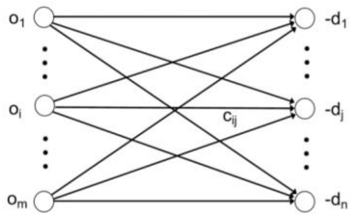
Se si osserva la struttura del sistema dei vincoli, ogni variabile x_{ij} compare *una* volta con coefficiente **1** in un certo vincolo e *una* volta con coefficiente **-1** in un altro vincolo. La struttura del sistema dei vincoli, in effetti, è costituita dai coefficienti di costo di una matrice di incidenza.
Rappresentiamo, infatti, il grafo mediante una matrice di incidenza nodo-arco A; per ogni nodo v ed arco e, la corrispondente entrata a_{ve} varrà: 1 se e esce da v (v è la coda di e), -1 se e entra in v (v è la testa di e), 0 altrimenti.

A	**(1,2)**	**(1,3)**	**(2,5)**	**(3,2)**	**(3,4)**	**(4,1)**	**(4,2)**	**(4,5)**
1	1	1	0	0	0	-1	0	0
2	-1	0	1	-1	0	0	-1	0
3	0	-1	0	1	1	0	0	0
4	0	0	0	0	-1	1	1	1
5	0	0	-1	0	0	0	0	-1

Problema del trasporto

Esempio (Problema del trasporto)**.** In questo problema, si hanno a disposizione *m* fornitori, i quali producono $o_1, \dots, o_m$ quantità di un certo prodotto. Si hanno anche *n* clienti, i quali richiedono $d_1, \dots, d_n$ quantità di prodotto. Il prodotto può essere trasportato da ogni fornitore ad ogni cliente. Il grafo sottostante a questo problema è un grafo *bipartito*, dove i nodi di origine (fornitori) hanno solo archi uscenti (per definizione di grafo bipartito) ed i nodi di destinazione (clienti) hanno solo archi entranti. Ad ogni arco (i, j) è associato un costo c_{ij} positivo.

Si vuole determinare la quantità di merce da trasportare su ogni arco (i, j) (fornitore-cliente) affinché ogni fornitore *i* invii la merce o_i prodotta, ogni cliente *j* riceva la quantità d_j richiesta, ed il costo complessivo di trasporto sia minimizzato.



Le **variabili** del problema sono le x_{ij} = quantità di flusso sull'arco (i, j).
La **funzione obiettivo** è $\min \sum_{(i,j)\in E} c_{ij}x_{ij}$.
I **vincoli** sono del tipo $\sum_{j\in Fs(i)} x_{ij} = o_i$ per i fornitori,
e/o $-\sum_{k\in Bs(i)} x_{ki} = b_i = -d_i$ per i clienti.

L'unico cambiamento dal problema del flusso a costo minimo ha riguardato i vincoli. Nel problema del trasporto, va identificato il ruolo di ogni nodo all'interno della rete.
Se si considera un nodo di offerta (fornitore), per come è costruita tale rete in quel nodo non si hanno archi entranti, per cui la sommatoria $\sum_{k\in Bs(i)} x_{ki}$ è nulla. Ecco perché $\sum_{j\in Fs(i)} x_{ij} = o_i$.
Se si ragiona in termini del cliente (nodo di domanda), per come è costruita tale rete in quel nodo non si hanno archi uscenti, per cui la sommatoria $\sum_{j\in Fs(i)} x_{ij}$ è nulla.

A destra è posta la sua **formulazione**.
In realtà, il problema del trasporto è un *caso particolare* del problema del flusso a costo minimo, in quanto si hanno solo nodi di offerta e nodi di domanda: in questo problema non esistono nodi di passaggio. In generale, una volta compresa la formulazione del problema del flusso a costo minimo, è possibile ricavare la formulazione di un altro problema dalla definizione di quest' ultimo.

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

$$\sum_{j=1}^{n} x_{ij} = o_i \qquad i = 1, \dots, m; \qquad (1)$$

$$-\sum_{i=1}^{m} x_{ij} = -d_j \qquad j = 1, \dots, n; \qquad (2)$$

$$x_{ij} \geq 0 \qquad i = 1, \dots, m; \quad j = 1, \dots, n \qquad (3)$$

$$x_{ij} \in R \qquad i = 1, \dots, m; \quad j = 1, \dots, n$$

Possiamo cercare di risolvere il problema utilizzando sì il Simplesso, ma in modo differente rispetto a quello visto, cioè senza dover calcolare le matrici inverse. Una proprietà che serve per risolvere il problema è la seguente.
Ipotesi di ammissibilità (Condizioni di bilanciamento)**.** Affinché il problema possa ammettere una soluzione deve essere verificata la seguente condizione sui dati:

$$\sum_{i=1}^{m} o_i - \sum_{j=1}^{n} d_j = 0,$$

ossia la quantità totale di prodotto disponibile deve essere uguale alla richiesta totale del prodotto stesso.

Tuttavia, esiste un modo per "aggirare" questo ostacolo. Supponiamo che $\sum_{i=1}^{m} o_i = 120$ e che $\sum_{j=1}^{n} d_j = 100$. Ciò significa che la rete non è in grado di assorbire tutto quanto prodotto, quindi questa condizione di bilanciamento non è soddisfatta. Questo non ci vieta, però, di aggiungere un nodo fittizio ai d_j, cui associamo l'eccedenza $\sum_{i=1}^{m} o_i - \sum_{j=1}^{n} d_j = 120 - 100 = 20$, e di collegarlo a tutti i nodi di offerta; i costi degli archi saranno posti a 0. Questo è fattibile, in quanto non influisce sulla funzione obiettivo.

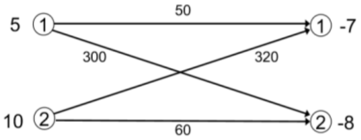
Nel problema del trasporto non si può avere un ottimo illimitato, dato che bisognerebbe avere tutti i $c_{ij} < 0$ (assurdo). Infatti, siccome per ipotesi $c_{ij} \geq 0$ e $x_{ij} \geq 0$, il valore minimo che può assumere la funzione obiettivo è 0. Dunque, sicuramente esiste una soluzione ottima finita.

Dimostriamo l'esistenza di una soluzione ammissibile. Sia $x_{ij} = \frac{o_i d_j}{\Delta}$, con $i = 1, \dots, m, j = 1, \dots, n$ e $\Delta = \sum_{i=1}^{m} o_i = \sum_{j=1}^{n} d_j$. Per farlo bisogna dimostrare che i vincoli (1) e (2) del sistema siano soddisfatti.

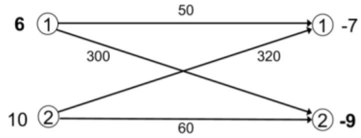
$$(1)\quad \sum_{j=1}^{n} x_{ij} = \sum_{j=1}^{n} \frac{o_i d_j}{\Delta} = \frac{o_i}{\Delta} \sum_{j=1}^{n} d_j = \frac{o_i}{\Delta} \Delta = o_i.$$

$$(2)\quad -\sum_{i=1}^{m} x_{ij} = -\sum_{i=1}^{m} \frac{o_i d_j}{\Delta} = -\frac{d_j}{\Delta} \sum_{i=1}^{m} o_i = -\frac{d_j}{\Delta} \Delta = -d_j.$$

Questo conclude la dimostrazione. □



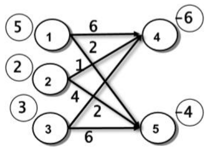
Segue un esempio.
A sinistra sono posti i nodi di offerta, e a destra sono posti i nodi di domanda. Le domande sono -7 e -8, e le offerte sono 5 e 10.
Ovviamente, una soluzione ammissibile del problema significa stabilire la quantità di flusso che passa su questi archi, per fare in modo che dalle offerte la merce arrivi ai nodi di domanda, cercando di minimizzare i costi (scritti sugli archi).
La **soluzione ottima** si ottiene con $x_{11} = 5, x_{12} = 0, x_{21} = 2, x_{22} = 8$, che implica $z^* = 1370$. Dunque, inviamo: *5* unità di flusso dal nodo di offerta 1 al nodo di domanda 1, ...



Data la soluzione ottima precedentemente calcolata, è possibile migliorare tale soluzione se si incrementano la domanda e l'offerta? *In alcuni casi*, sì. Questo fatto è chiamato "paradosso del trasporto", e dipende dalla struttura della rete (e, in particolare, dai costi degli archi).
Supponiamo che incrementiamo di 1 l'offerta del nodo 1 ($5 + 1 = \mathbf{6}$), e incrementiamo di 1 la domanda del nodo 2 ($-8 - 1 = \mathbf{-9}$) per rispettare il bilanciamento. La soluzione ottima si ottiene, ora, con $x_{11} = 6, x_{12} = 0, x_{21} = 1, x_{22} = 9$, che implica $z^* = 1160$.

Esistono teoremi che permettono di stabilire, in base alla matrice dei costi, se quest'ultima è immune al paradosso del trasporto.

Abbiamo detto che il problema del trasporto è un **sottocaso** particolare del flusso a costo minimo:
• non esistono nodi di passaggio;
• è possibile andare da ogni nodo offerta (insieme O) a ogni nodo richiesta (insieme D);
• il grafo del problema è un grafo bipartito.



$$\min 6x_{14} + 2x_{15} + x_{24} + 4x_{25} + 2x_{34} + 6x_{35}$$

soggetto ai vincoli

$$\begin{aligned} x_{14} + x_{15} &= 5 \\ x_{24} + x_{25} &= 2 \\ x_{34} + x_{35} &= 3 \\ -x_{14} - x_{24} - x_{34} &= -6 \\ -x_{15} - x_{25} - x_{35} &= -4 \qquad x_{ij} \geq 0 \end{aligned}$$

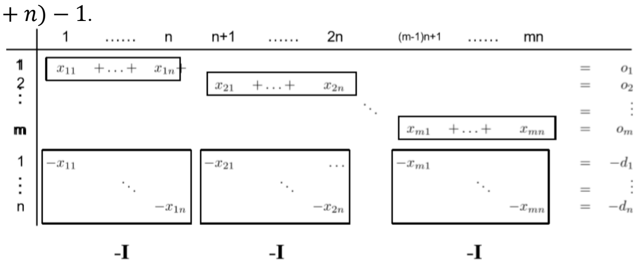
A	(1,4)	(1,5)	(2,4)	(2,5)	(3,4)	(3,5)
1	1	1	0	0	0	0
2	0	0	1	1	0	0
3	0	0	0	0	1	1
4	-1	0	-1	0	-1	0
5	0	-1	0	-1	0	-1
	-I		**-I**		**-I**	

Ora, bisogna risolvere il problema: anziché usare il Simplesso classico, come detto in precedenza ne vediamo una variante più veloce, sfruttando le caratteristiche del problema (in particolare la struttura del grafo).
Consideriamo la matrice di incidenza nodo-arco A per il problema del trasporto precedente; notiamo che:
- per quanto riguarda i nodi di offerta (1, 2, 3) abbiamo sempre una sequenza di 1 verso tutti i nodi clienti;
- per quanto riguarda i nodi di domanda (4, 5) abbiamo delle matrici identità moltiplicate per -1 (che indichiamo con **-I**).
Tale struttura è dovuta al fatto che tale grafo è bipartito.

Nell'argomento precedente, abbiamo osservato che la matrice di incidenza A non è a rango pieno, in quanto la somma delle sue righe dà come risultato il vettore nullo. Dimostriamo, ora, direttamente l'osservazione per cui $r(A) = n - 1$; cioè, che il numero di righe di A linearmente indipendenti è uguale a n – 1.
Nell'esempio precedente, sappiamo di avere *m* = 3 fornitori ed *n* = 2 clienti, per cui la matrice A è di dimensione

(m+n)x(m·n). Ricordando di non confondere gli indici (in quanto usiamo come indicizzazione della matrice nxm), si può dimostrare che $r(A) = (m+n) - 1$.

A partire da questo caso particolare, generalizziamo come segue. Da una parte si hanno le *m* righe associate ai fornitori e le *n* righe associate ai clienti; dall'altra, sono state scritte le colonne enumerandole in base agli archi.



Vogliamo dimostrare che il rango di questa matrice è uguale a $(m+n) - 1$.

Abbiamo già stabilito che il rango della matrice non è pieno, quindi si può eliminare una riga: per semplicità, eliminiamo l'ultima. Ora, quindi, abbiamo una matrice avente $(m+n) - 1$ righe. Per dimostrare che questa nuova matrice è a rango pieno, *cerchiamo di costruire una sottomatrice quadrata* <u>*triangolare superiore*</u>, di dimensione (n+m-1)x(n+m-1), che sia a rango pieno; il motivo di questa scelta è che, dopo averla trovata, il determinante di questa si potrà calcolare banalmente moltiplicando gli elementi sulla diagonale.

Per ottenere un 1 in alto a sinistra, scegliamo la colonna *n*; per ottenere un 1 nella seconda riga della posizione successiva, scegliamo la colonna *2n*; per ottenere un 1 nella terza riga della posizione successiva, scegliamo la colonna *3n*; procediamo in questo modo fino alla colonna *mn*, ed avremo coperto la nuova sottomatrice fino alla riga m.

Per completare la sottomatrice, scegliamo le colonne *1*, *2*, ..., *n-1*, dato che esse hanno sulla diagonale principale il termine -1. La sottomatrice ottenuta è posta a destra.

$$\begin{array}{c|ccccccccc}
 & n & 2n & 3n & \dots\dots & mn & 1 & 2 & \dots\dots & n\text{-}1 \\
\hline
1 & 1 & 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 \\
2 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
3 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\
\vdots & \vdots & & & & \vdots & & & & \vdots \\
m & 0 & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \\
1 & 0 & 0 & 0 & \dots & 0 & -1 & 0 & \dots & 0 \\
\vdots & 0 & 0 & 0 & \dots & 0 & 0 & -1 & \dots & 0 \\
\vdots & \vdots & & & & \vdots & & & & \vdots \\
n\text{-}1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & -1
\end{array}$$

Come risultato, è stata costruita la sottomatrice quadrata di dimensione (n+m-1)x(n+m-1) sulla cui diagonale si hanno tutti elementi diversi da 0.

Essendo una matrice triangolare superiore, si può applicare il teorema per cui "*Se A è una matrice diagonale o triangolare superiore o triangolare inferiore allora il determinante di A è dato dal prodotto degli elementi sulla diagonale principale*".

Questo dimostra che la matrice dei vincoli del problema del trasporto ha rango $r(A) = (m+n) - 1$.

Questa informazione è importante, in quanto sappiamo che <u>il problema del trasporto ha una base ammissibile di dimensione $m + n - 1$</u>, dato che questo è il numero di righe linearmente indipendenti di A. □

La **risoluzione** del problema del trasporto si effettua utilizzando due tabelle, aventi un numero di righe pari ai fornitori (*m*) ed un numero di colonne pari ai clienti (*n*):
- la prima tabella, relativa alle variabili;
- la seconda tabella, relativa ai costi.

	1	2	...	...	*n*	
1	x_{11}	x_{12}	...	...	x_{1n}	O_1
2	x_{21}	x_{22}	...	...	x_{2n}	O_2
...	...	...	...	...	...	...
...	...	...	...	...	...	...
m	x_{m1}	x_{m2}	...	...	x_{mn}	O_m
	d_1	d_2	...	...	d_n	

	1	2	...	...	*n*
1	c_{11}	c_{12}	...	...	c_{1n}
2	c_{21}	c_{22}	...	...	c_{2n}
...	...	...	...	...	...
...	...	...	...	...	...
m	c_{m1}	c_{m2}	...	...	c_{mn}

Costruiamo innanzitutto una <u>soluzione di base ammissibile</u> per il problema. Sappiamo che la sua dimensione è $m + n - 1$, dunque bisogna:

1. Scegliere (al più, se soluzione degenere) $m + n - 1$ variabili decisionali che abbiano un valore ≥ 0.

2. Verificare l'ammissibilità della soluzione, cioè se i vincoli sono soddisfatti:

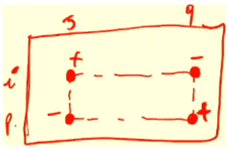
  2.1. La somma dei valori delle x_{ij} (sulla riga *i*) dev'essere uguale a uguale a O_i;

  2.2. La somma dei valori delle x_{ij} (sulla colonna *j*) dev'essere uguale a d_j.

3. Garantire che gli $\underline{a}_{ij} = \underline{e}_i - \underline{e}_j$, associati alle variabili scelte per inserirle in base, siano linearmente indipendenti.

Per garantire la 3, non bisogna <u>mai costruire dei cicli</u> nella tabella. Questo è importante.

Ad esempio, supponiamo di avere una matrice del tipo a destra. Se scegliessi delle variabili che formano un ciclo, allora quelle colonne sarebbero linearmente dipendenti. Dimostriamolo, usando valori alternati per gli elementi sulla diagonale (principale +, secondaria -). Quindi:



$\underline{a}_{ij} - \underline{a}_{iq} + \underline{a}_{pq} - \underline{a}_{pj} = (\underline{e}_i - \underline{e}_j) - (\underline{e}_i - \underline{e}_q) + (\underline{e}_p - \underline{e}_q) - (\underline{e}_p - \underline{e}_j) = \underline{0}$.

Per risolvere il problema usiamo questi due "algoritmi", che evitano il calcolo della matrice inversa:
1. Trovare una soluzione di base ammissibile di partenza, con il *Metodo del Nord-Ovest*.
2. Migliorare la soluzione ammissibile trovata fino a soddisfare le condizioni di ottimalità, con la *Regola del ciclo*.

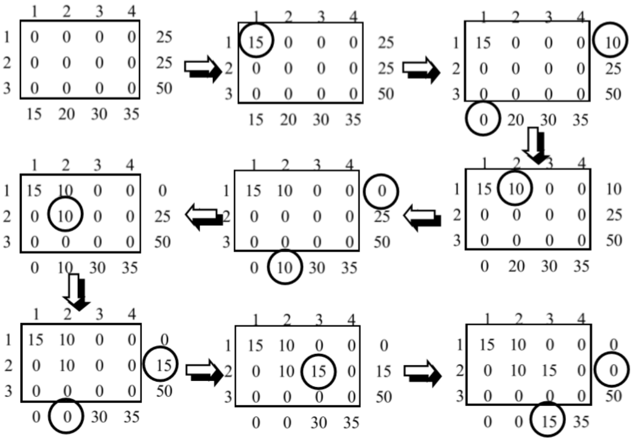
Metodo dell'Angolo di Nord-Ovest. Di seguito è posto lo pseudocodice, ma è molto più semplice introdurlo con un esempio. All'inizio, si annullano le variabili x_{ij} e si parte dall'elemento in altro a sinistra.

Passo 0: Poni $x_{ij}=0$ per ogni i e per ogni j
Passo 1: $i=1$, $j=1$
Passo 2: $x_{ij}=minimo\ \{o_i, d_j\}$
    Se il minimo è uguale a o_i allora vai al passo 3
    Se il minimo è uguale a d_j allora vai al passo 4
Passo 3: Poni $i = i+1$; $d_j = d_j - o_i$ e vai al passo 2
Passo 4: Poni $j = j+1$; $o_i = o_i - d_j$ e vai al passo 2

Consideriamo la seguente tabella dei costi, avente $m=3$ fornitori e $n=4$ clienti.

	1	2	3	4	
1	10	5	6	7	25
2	8	2	7	6	25
3	9	3	4	8	50
	15	20	30	35	

Ovviamente, si parte col verificare che domanda e offerta coincidono: $25 + 25 + 50 = 15 + 20 + 30 + 35$.
All'interno della matrice sono posti i costi. Ad esempio, l'arco (1, 1) ha costo 10.
Detto ciò, sappiamo che le variabili in base per questo problema sono $m + n - 1 = 3 + 4 - 1 = 6$.
L'algoritmo itera come segue:
1) si annulla la matrice dei costi;
2) si sceglie l'elemento (1, 1), e si nota che in questa posizione si ha offerta 25 e domanda 15. Si assegna alla variabile x_{11} il minimo tra questi due, pertanto si scrive nella posizione (1, 1) il valore 15;
3) si aggiornano offerta e domanda sottraendovi il valore scelto. Quindi, l' offerta avrà valore 25 – 15 = 10, mentre la domanda avrà valore 15 – 15 = 0;
4) ci si sposta verso destra, o verso il basso. In particolare, se il fornitore ha ancora della merce da inviare, allora ci si sposta verso destra; se invece è il cliente a non avere merce da chiedere, allora ci si sposta verso il basso. In questo caso ci si sposta verso destra, e si ripete il passaggio 2). Si sceglie, quindi, l'elemento (1, 2) e gli si assegna il valore min{10, 20} = 10;

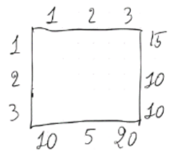


5) si aggiornano offerta e domanda sottraendovi il valore scelto. Quindi, l'offerta avrà valore 10 – 10 = 0, mentre la domanda avrà valore 20 – 10 = 10;
6) in questo caso, ci si sposta verso il basso, e si reitera sull'elemento (2, 2).
...
Si ripete fin quando non si arriva all'angolo Sud-Est. In questo caso, il risultato sarà questo posto a destra. Notiamo che le variabili maggiori di 0, nella tabella, sono esattamente 6. Inoltre, per come è stata costruita la tabella sono stati rispettate domanda e offerta. Infine, non si creeranno mai cicli, in quanto con questo metodo ci si sposta sempre o verso destra o verso il basso, senza mai tornare indietro.

	1	2	3	4	
1	15	10	0	0	25
2	0	10	15	0	25
3	0	0	15	35	50
	15	20	30	35	

Al termine, quindi, le variabili di base della soluzione ammissibile iniziale sono $x_{11}, x_{12}, x_{22}, x_{23}, x_{33}, x_{34}$.

Es. Se avessimo avuto una matrice del tipo a destra, avremmo ottenuto ad una certa iterazione un caso in cui domanda e offerta coincidono. In questo caso avremo sicuramente una soluzione ammissibile degenere, ed in questo caso ci si può spostare sia a destra, sia in basso, indifferentemente. L'importante è rispettare il Metodo del Nord-Ovest, quindi non bisogna mai pensare di spostarsi in obliquo.



Con il metodo del Nord-Ovest è stata trovata una soluzione di base ammissibile iniziale. Il passaggio successivo è verificare se questa soluzione di base è la soluzione di base *ottima* che sto cercando. Ricordiamo le condizioni di ottimalità del simplesso: $(z_j - c_j) \leq 0\ \forall j \in \mathrm{N}$, dove $z_j = \underline{c}_B^T A_B^{-1} \underline{a}_j$.
Ricordiamo che $\underline{c}_B^T A_B^{-1}$ è il vettore $\underline{w}$ dei moltiplicatori del simplesso, ed inoltre sappiamo che, quando la base B è ottima, il vettore $\underline{w}$ diventa la base ottima del problema duale.
Per la condizione di ottimalità del problema del trasporto, vogliamo calcolare i coefficienti di costo ridotto in un modo alternativo, senza cioè calcolare la matrice inversa A_B^{-1}: la chiave è costruire il *duale del problema del trasporto*.
In particolare, denoteremo con $\boldsymbol{u_i}$ e $\boldsymbol{v_j}$ le componenti del vettore $\underline{w}$ associate ai vincoli del problema primale. Pertanto, si avrà $\underline{w}^T = [u_1 \ \ u_2 \ \ \cdots \ \ u_m \ \ v_1 \ \ v_2 \ \ \cdots \ \ v_n]$. Per ora, abbiamo soltanto denotato le componenti di questo vettore con nomi diversi per poterle distinguere.
Non utilizzeremo il duale del problema del trasporto, bensì ci serve soltanto sapere quali sono le componenti del vettore $\underline{w}$.

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

$$(u_i) \qquad \sum_{j=1}^{n} x_{ij} = o_i \qquad i = 1, \ldots, m; \qquad (1)$$

$$(v_j) \qquad -\sum_{i=1}^{m} x_{ij} = -d_j \qquad j = 1, \ldots, n; \qquad (2)$$

$$x_{ij} \geq 0 \qquad i = 1, \ldots, m; \ \ j = 1, \ldots, n \qquad (3)$$

$$x_{ij} \in R \qquad i = 1, \ldots, m; \ \ j = 1, \ldots, n$$

$$\Downarrow$$

$$\max \sum_{i=1}^{m} o_i u_i - \sum_{j=1}^{n} d_j v_j$$

$$u_i - v_j \leq c_{ij} \qquad i = 1, \ldots, m; \ \ j = 1, \ldots, n;$$

Le **condizioni di ottimalità** del problema del trasporto si basano sul calcolo dei coefficienti di costo ridotto $(z_{ij} - c_{ij}) \leq 0$, per ogni x_{ij} non in base. Possiamo ricavare che:

$$\boldsymbol{z_{ij} - c_{ij}} = \underline{w}^T \underline{a}_{ij} - c_{ij} = \boldsymbol{u_i - v_j - c_{ij}}.$$

L'obiettivo è calcolare u_i e v_j per ottenere il coefficiente di costo ridotto, per verificare se la base è ottima. Dunque:
- u_i è la variabile duale associata all'*i*-simo vincolo di offerta;
- v_j è la variabile duale associata al *j*-simo vincolo di domanda.

Per calcolare queste variabili in modo rapido, dobbiamo considerare sia la matrice dei costi iniziali, sia la matrice delle variabili corrispondente alla soluzione di base trovata con il metodo del Nord-Ovest.
L'idea è sfruttare l'informazione secondo cui il coefficiente di costo ridotto di una variabile in base è nullo; quindi

$$z_{ij} - c_{ij} = 0 \ \ \forall [ij] \in \mathrm{B}.$$

	1	2	3	4	
1	10	5	6	7	25
2	8	2	7	6	25
3	9	3	4	8	50
	15	20	30	35	

	1	2	3	4	
1	15	10	0	0	25
2	0	10	15	0	25
3	0	0	15	35	50
	15	20	30	35	

Le variabili in base le conosciamo, cioè (a meno di soluzione degenere) quelle diverse da 0 nella matrice delle variabili di base di partenza. Di conseguenza, scriviamo un sistema di equazioni per ricavare i valori di u_i e v_j:

$$\begin{cases} u_1 - v_1 - c_{11} = 0 \\ u_1 - v_2 - c_{12} = 0 \\ u_2 - v_2 - c_{22} = 0 \\ u_2 - v_3 - c_{23} = 0 \\ u_3 - v_3 - c_{33} = 0 \\ u_3 - v_4 - c_{34} = 0 \end{cases} \Leftrightarrow \begin{cases} 0 - v_1 - 10 = 0 \ \ \Rightarrow \ \ v_1 = -10 \\ 0 - v_2 - 5 = 0 \ \ \Rightarrow \ \ v_2 = -5 \\ u_2 - (-5) - 2 = 0 \ \ \Rightarrow \ \ u_2 = -3 \\ -3 - v_3 - 7 = 0 \ \ \Rightarrow \ \ v_3 = -10 \\ u_3 - (-10) - 4 = 0 \ \ \Rightarrow \ \ u_3 = -6 \\ -6 - v_4 - 8 = 0 \ \ \Rightarrow \ \ v_4 = -14 \end{cases}.$$

Si noti che il sistema ha, in questo caso, 6 equazioni in quanto le associamo alle variabili in base; in generale, il numero di equazioni sarà sempre uguale a $n + m - 1$. Le variabili del problema u_i e v_j, però, sono 7, quindi il sistema ha infinite soluzioni. Per trovare una soluzione del sistema, annulliamo una variabile del sistema e ricaviamo i valori delle restanti. Nel sistema precedente, ad esempio, abbiamo posto $u_1 = 0$.

Dopo aver calcolato questi valori, possiamo finalmente verificare le condizioni di ottimalità ($z_{ij} - c_{ij} = u_i - v_j - c_{ij} \leq 0 \ \ \forall [ij] \in \mathrm{N}$); i calcoli sono posti alla pagina seguente.

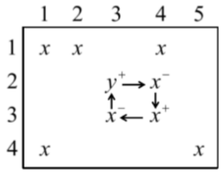
- (x_{13}) $z_{13} - c_{13} = u_1 - v_3 - c_{13} = 0 + 10 - 6 = 4 > 0;$
- (x_{14}) $z_{14} - c_{14} = u_1 - v_4 - c_{14} = 0 + 14 - 7 = 7 > 0;$
- (x_{21}) $z_{21} - c_{21} = u_2 - v_1 - c_{21} = -3 + 10 - 8 = -1 < 0;$
- (x_{24}) $z_{24} - c_{24} = u_2 - v_4 - c_{24} = -3 + 14 - 6 = 5 > 0;$
- (x_{31}) $z_{31} - c_{31} = u_3 - v_1 - c_{31} = -6 + 10 - 9 = -5 < 0;$
- (x_{32}) $z_{32} - c_{32} = u_3 - v_2 - c_{32} = -6 + 5 - 3 = -4 < 0.$

	1	2	3	4	
1	10	5	6	7	25
2	8	2	7	6	25
3	9	3	4	8	50
	15	20	30	35	

	1	2	3	4	
1	15	10	0	0	25
2	0	10	15	0	25
3	0	0	15	35	50
	15	20	30	35	

La soluzione non è ottima, quindi dobbiamo scegliere una variabile non in base da introdurre in base. Per la regola del gradiente, entra in base la variabile con coefficiente di costo maggiore, ossia x_{14}.

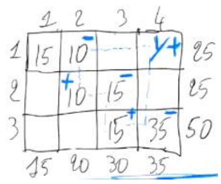
Nel simplesso classico, a questo punto dovremmo applicare il test di illimitatezza per verificare se la soluzione al problema è infinito. Per il problema del trasporto non serve, in quanto non si può mai verificare l'ottimo illimitato. Ora, bisogna capire come scegliere la variabile da far uscire dalla base: la logica è sempre la stessa, ma applichiamo la **regola del ciclo** per ottenerla.



Supponiamo di avere la tabella generica posta a destra, in cui le x rappresentano le variabili di base e la y <u>la nuova variabile entrante</u> (nell'esempio precedente, x_{14}). Una volta inserita la variabile y all'interno della tabella, si creerà un ciclo: a partire da y, bisogna cercare – tracciando segmenti orizzontali e verticali – di creare un ciclo tra la variabile y e le altre variabili in base; si può cambiare direzione soltanto quando stiamo considerando una x, cioè una variabile in base.

Nell'esempio precedente, data la matrice delle variabili corrispondente alla soluzione di base trovata con il metodo del Nord-Ovest, inseriamo la y nella cella con indici della variabile che deve entrare in base, cioè (1, 4), ed un ciclo è *(1, 4), (1, 2), (2, 2), (2, 3), (3, 3), (3, 4), (1, 4).*

A questo punto, trovato il ciclo, dobbiamo far crescere la y (per cui, la denotiamo con y^+), e per bilanciare l'offerta del nodo 1 dobbiamo far decrescere la variabile x_{12} (per cui, la denotiamo con x^-), e per bilanciare la domanda del nodo 2 dobbiamo far crescere la variabile x_{22} (per cui, la denotiamo con x^+), e per bilanciare l'offerta del nodo 2 dobbiamo far decrescere la variabile x_{23} (per cui, la denotiamo con x^-), e così via, fin quando non si torna alla y.



In questo esempio, il valore massimo che la y può assumere è 10, in quanto si può decrementare al più di 10 la variabile x_{12}. In generale, *tra tutte le variabili <u>decrementate</u>* (in questo esempio x_{12}, x_{23}, x_{33}) *bisogna scegliere la più piccola*: questa sarà la variabile uscente dalla base. In basso a destra è posta la nuova matrice, ottenuta incrementando o decrementando tutte le variabili del ciclo del valore della variabile uscente dalla base.

	1	2	3	4	
1	15	0	0	10	25
2	0	20	5	0	25
3	0	0	25	25	50
	15	20	30	35	

In sintesi, in questa iterazione: il bilanciamento è comunque stato rispettato, la variabile entrante in base è x_{14} e la variabile uscente dalla base è x_{12}. Dopo aver fatto ciò, proseguiamo con un'altra iterazione usando la matrice di base a destra.

Formalmente, la Regola del Ciclo è costituita dai seguenti step:
1. Consideriamo le variabili in base che formano un ciclo con la variabile entrante.
2. I segni + e – sono assegnati in modo alternato <u>negli angoli del ciclo</u> partendo dalla variabile fuori base y a cui viene assegnato il + perché deve essere incrementata di un nuovo valore $\Delta \geq 0$.
3. Le variabili di base che si trovano negli angoli del ciclo verranno incrementate di Δ, se hanno segno positivo, e decrementate di Δ, se hanno segno negativo.
4. La variabile uscente sarà quella che si azzererà per prima.

Di seguito si sintetizza l'algoritmo del simplesso per il problema del trasporto:
1. Trova una soluzione di base ammissibile con la regola dell'angolo di Nord-Ovest.
2. Calcola $z_{ij} - c_{ij}$ per ogni variabile non in base ($z_{ij} - c_{ij} = u_i - v_j - c_{ij}$):
- Se $z_{ij} - c_{ij} \leq 0$ per ogni variabile non in base, allora la base è ottima;
- Altrimenti, seleziona la variabile entrante con il massimo $z_{ij} - c_{ij}$.

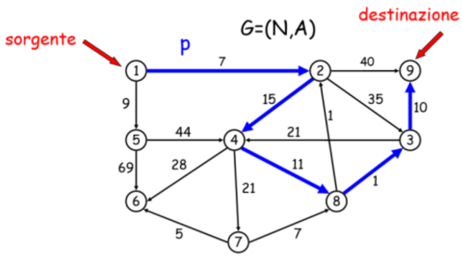
3. Determina la variabile uscente applicando la Regola del ciclo.
4. Ricalcola la nuova soluzione di base ammissibile e torna al passo 2.

Problema dei cammini minimi

Esistono tre versioni del problema dei cammini minimi: uno-a-uno, uno-a-tutti, tutti-a-tutti.

Versione uno-a-uno.

Sia G = (N, A) un grafo orientato su cui sia definito un vettore <u>c</u> = [c_{ij}] dei costi associati agli archi del grafo; inoltre, siano s e t due nodi distinti, detti rispettivamente *origine* e *destinazione*. Il problema dei cammini minimi **1 a 1** consiste nel determinare il <u>percorso di costo minimo</u> (più corto) da s a t in G.



In breve, si ha un nodo sorgente ed un nodo di destinazione. Stabilito qual è il problema da risolvere, bisogna scrivere il modello matematico; detto p il percorso, in questo problema:

- c_{ij} è il <u>costo</u> dell'arco (i, j);
- x_{ij} è una <u>variabile decisionale</u>, con $x_{ij} = \begin{cases} 1 & \text{se } x_{ij} \in p \\ 0 & \text{se } x_{ij} \notin p \end{cases}$. Quindi $x_{ij} = 1$ se selezioniamo (i, j), 0 altrimenti;
- l'<u>obiettivo</u> è minimizzare i costi degli archi da attraversare, quindi $\min \sum_{(i,j)\in A} c_{ij}x_{ij}$;
- i <u>vincoli</u> del problema sono $\sum_{j\in Fs(i)} x_{ij} - \sum_{k\in Bs(i)} x_{ki} = \begin{cases} 1 & \text{se } i = s \\ 0 & \text{se } i \in N \setminus \{s,t\} \\ -1 & \text{se } i = t \end{cases}$. Quindi la differenza tra somma del flusso uscente dal nodo *i* e la somma del flusso entrante nel nodo *i* dev'essere uguale a quel valore, in base al ruolo.

La matrice di questo problema è *di incidenza*, la quale è unimodulare: questa proprietà e il fatto che le componenti del vettore <u>*b*</u> sono intere ci garantiscono che tutte le soluzioni di base ammissibili di questo problema hanno coordinate intere; quindi, si può risolvere il problema ponendo $x_{ij} \geq 0$ senza aggiungere il vincolo di interezza. Anche in questo caso, sebbene si tratti di un problema di PLI è possibile "rilassare" il vincolo di interezza, quindi risolverlo come problema di PL.

Di seguito si sintetizza il modello matematico **generale** del problema dei cammini minimi uno-a-uno:

$$\min z = \sum_{(i,j)\in A} c_{ij}x_{ij}$$
$$\sum_{j\in Fs(i)} x_{ij} - \sum_{k\in Bs(i)} x_{ki} = \begin{cases} 1 & \text{se } i = s \\ 0 & \text{se } i \in N \setminus \{s,t\} \\ -1 & \text{se } i = t \end{cases}.$$
$$x_{ij} \in \{0,1\}$$

Es. Di seguito è posto il modello matematico **esplicito** del problema precedente, con *s* = 1 e *t* = 9.

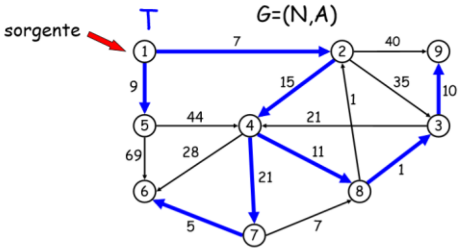
$$\min z = 7x_{12} + 40x_{29} + 9x_{15} + 35x_{23} + 15x_{24} + 10x_{39} + 21x_{34} + \ ...$$

$$x_{12} + x_{15} - 0 = 1 \quad (i = 1)$$
$$0 - x_{29} - x_{39} = -1 \quad (i = 9)$$

$$x_{48} + x_{47} + x_{46} - x_{54} - x_{24} - x_{34} = 0 \quad (i = 4)$$
$$x_{23} + x_{24} + x_{29} - x_{12} - x_{82} = 0 \quad (i = 4)$$
$$x_{39} + x_{34} - x_{23} - x_{83} = 0 \quad (i = 3)$$
$$x_{54} + x_{56} - x_{15} = 0 \quad (i = 5)$$
$$0 - x_{46} - x_{56} - x_{76} = 0 \quad (i = 6)$$
$$x_{76} + x_{78} - x_{47} = 0 \quad (i = 7)$$
$$x_{82} + x_{83} - x_{48} - x_{78} = 0 \quad (i = 8)$$

La soluzione ottima del problema è $x_{12} = 1, x_{24} = 1, x_{48} = 1, x_{83} = 1, x_{39} = 1$, con $z^* = 44$.

Versione uno-a-tutti.



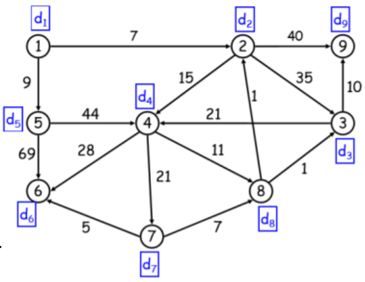
Sia G = (N, A) un grafo orientato su cui sia definito un vettore <u>c</u> = [c_{ij}] dei costi associati agli archi del grafo; inoltre, sia s il nodo *origine*. Il problema dei cammini minimi **1 a tutti** consiste nel determinare l'<u>albero dei cammini minimi</u> da s a tutti gli altri nodi di G.
In breve, da un nodo sorgente *s* ci si può spostare ad ogni altro nodo in N. Quindi, si vuole calcolare il percorso minimo da un nodo sorgente a tutti gli altri nodi della rete.
Il nodo *s* è la radice dell'albero dei cammini minimi da calcolare. Nell'esempio in figura, l'albero T è descritto come segue: il nodo *1* è la radice avente figli *5* (foglia) e *2*; il nodo *2* ha come figlio *4*; il nodo *4* ha come figli *7* e *8*; il nodo *7* ha come figlio *6* (foglia), mentre il nodo *8* ha come figlio *3*; il nodo *3* ha come figlio *9* (foglia).

In questo modo, è possibile raggiungere ogni nodo della rete a patto che esista un percorso da una sorgente ad una determinata destinazione.

Questa versione differisce dalla precedente – definendo *n* come numero di nodi del grafo – in:
- <u>vincoli</u>, siccome non esistono nodi di passaggio, saranno $\sum_{j\in Fs(i)} x_{ij} - \sum_{k\in Bs(i)} x_{ki} = \begin{cases} n-1 & \text{se } i = s \\ -1 & \text{se } i \neq s \end{cases}$;
- <u>variabili decisionali</u>, con $x_{ij} \in Z^+ \cup \{0\}$.

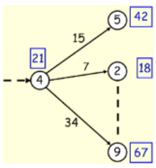
Di seguito si sintetizza il modello matematico **generale** del problema dei cammini minimi uno-a-tutti:

$$\min z = \sum_{(i,j)\in A} c_{ij} x_{ij}$$
$$\sum_{j\in Fs(i)} x_{ij} - \sum_{k\in Bs(i)} x_{ki} = \begin{cases} n-1 & \text{se } i = s. \\ -1 & \text{se } i \neq s \end{cases}$$
$$x_{ij} \in Z^+ \cup \{0\}$$



Seguono alcune idee sulla risoluzione del problema dei cammini minimi uno-a-tutti. Esistono diverse tipologie di algoritmi che lo risolvono, ma essenzialmente si basano tutti sul continuo aggiornamento delle **etichette** che vengono associate ai nodi del grafo. Non va confuso l' identificativo del nodo (cioè, il numero all'interno del cerchio) con l' etichetta: quest'ultimo è un valore che associamo al nodo, praticamente il costo della path più breve trovata *fino a quel momento* dalla sorgente a quel nodo; questo costo, se il nodo è raggiungibile dalla sorgente, viene via via aggiornato (decrementato) fin quando non diventa quello ottimo. Se il nodo <u>non</u> è raggiungibile dalla sorgente, la sua etichetta varrà ∞.

L'**algoritmo prototipo** è un algoritmo generico da cui vengono ricavate le implementazioni per il problema dei cammini minimi. Nell'inizializzazione, denotiamo con: d_s l'etichetta del nodo *s*, P_s il padre del nodo *s*.
Segue un esempio di *relaxing* (aggiornamento delle etichette); si suppone di avere la situazione seguente.



Passo 1: *Inizializzazione.*
d_s=0, P_s=NULL, d_k=∞, P_k=s ∀k∈N\{s},
Q={s};

Passo 2: *Estrai un vertice x da Q (Q= Q \{x}) ed aggiorna quando possibile le etichette dei vertici in FS(x):*
∀y∈FS(x) se $d_x + c_{xy} < d_y$ allora
$d_y = d_x + c_{xy}$, P_y= x e se y ∉Q inseriscilo in Q *(Q= Q ∪ {y})* *(Relaxing)*

Passo 3: *Fino a quando Q ≠ ∅ ripeti il passo 2 ;*

Le etichette sono poste in blu. Se si effettua il test:
1) x=4, y=5 ⇒ $d_4 + c_{45} < d_5$, quindi si aggiorna l'etichetta in $d_5 = d_4 + c_{45} = 36$ e $P_5 = 4$;
2) x=4, y=2 ⇒ $d_4 + c_{42} > d_2$, quindi non conviene arrivare al nodo 2 passando per il nodo 4;
3) x=4, y=9 ⇒ $d_4 + c_{49} < d_9$, quindi si aggiorna l'etichetta in $d_9 = d_4 + c_{49} = 55$ e $P_9 = 4$.

Nel passo 2, aggiornare (decrementare) l'etichetta di un noto *t* e porlo come figlio di un altro nodo *s* coincide con l'individuazione di un percorso più breve (rispetto a quello computato fino a quel momento) per raggiungere *t*.

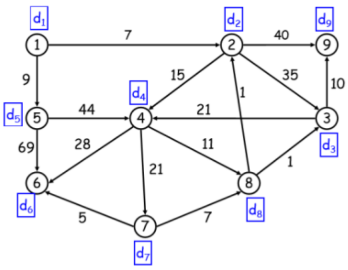
Gli algoritmi per l'SPT (Shortest Path First) si distinguono per la politica di estrazione del nodo da Q (*label setting* e *label correcting*) e per la struttura dati utilizzata per implementare Q.
Negli algoritmi **label correcting** [Bellman-Ford]:
- i nodi vengono estratti dalla coda Q in ordine FIFO (è una delle possibili implementazioni dell'algoritmo);
- le etichette dei nodi sono temporanee per tutta la durata della computazione. Solo al termine dell'algoritmo tali etichette rappresenteranno le distanze minime;
- l'algoritmo è in grado di risolvere il problema dei cammini minimi su un qualsiasi grafo che non presenta cicli di peso negativo. Infatti, se il grafo presentasse cicli di peso negativo (percorrendo il ciclo, si diminuisce la lunghezza della path) allora l'algoritmo andrebbe in loop infinito;

Negli algoritmi **label setting** [Dijkstra]:
- ad ogni iterazione viene estratto dalla coda Q il nodo x con etichetta minima;
- l'etichetta del nodo x estratto rappresenta la distanza minima dalla sorgente al nodo stesso. Tale etichetta viene fissata in modo permanente e non viene più aggiornata (quindi una volta estratto un nodo non può essere reinserito in Q);
- gli algoritmi label setting sono più efficienti dei label correcting, ma possono essere applicati solo su grafi dove $c_{ij} \geq 0$.

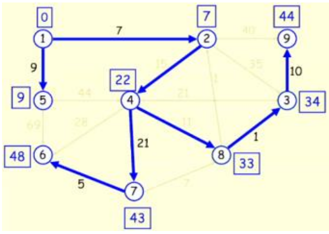
Di seguito è posto l'**algoritmo di Dijkstra**.

```
Dijkstra (G,s)

  Inizializzazione; ( ds=0, Ps=NULL,   dk=∞, Pk=s ∀k∈N\{s} , Q={s})

  while ( Q ≠ ∅){
            x = Extract_min(Q);
            Relax(x,y);     con y∈FS(x);
  }
```

Es. Dato il grafo posto a destra, utilizzare l'algoritmo di Dijkstra per individuare l'albero dei cammini minimi, partendo dal nodo 1.

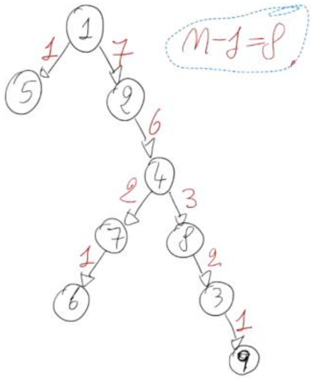


→
INIZIALIZZAZIONE. $d_1 = 0, P_1 = NULL, d_i = \infty \quad \forall i \in N \setminus \{1\}$.
Nella coda, è consigliabile inserire un elemento costituito da: identificativo del nodo, etichetta del nodo in un quadrato, padre del nodo in un rombo. Queste sono le informazioni che servono durante la computazione. Dunque, $Q = \{1\,\boxed{0}\,\langle NULL\rangle\}$;
ITERAZIONE 1. Estrazione del nodo dalla coda con etichetta più piccola: 1. Una volta estratto, si effettua il relaxing su tutti gli archi uscenti dal nodo 1, quindi:
- (1, 2): $d_1 + c_{12} < d_2 \Leftrightarrow 0 + 7 < \infty \Rightarrow d_2 = d_1 + c_{12} = 7, P_2 = 1, Q = \{2\,\boxed{7}\,\langle 1\rangle\}$;
- (1, 5): $d_1 + c_{15} < d_5 \Leftrightarrow 0 + 9 < \infty \Rightarrow d_5 = d_1 + c_{15} = 9, P_5 = 1, Q = \{2\,\boxed{7}\,\langle 1\rangle, 5\,\boxed{9}\,\langle 1\rangle\}$.
ITERAZIONE 2. Estrazione del nodo dalla coda con etichetta più piccola: 2. Una volta estratto, lo si aggiunge all'albero e si effettua il relaxing su tutti gli archi uscenti dal nodo 2, quindi:
- (2, 9): $d_2 + c_{29} < d_9 \Leftrightarrow 7 + 40 < \infty \Rightarrow d_9 = d_2 + c_{29} = 47, P_9 = 2, Q = \{5\,\boxed{9}\,\langle 1\rangle, 9\,\boxed{47}\,\langle 2\rangle\}$;
- (2, 3): $d_2 + c_{23} < d_3 \Leftrightarrow 7 + 35 < \infty \Rightarrow d_3 = d_2 + c_{23} = 42, P_3 = 2, Q = \{5\,\boxed{9}\,\langle 1\rangle, 9\,\boxed{47}\,\langle 2\rangle, 3\,\boxed{42}\,\langle 2\rangle\}$;
- (2, 4): $d_2 + c_{24} < d_4 \Leftrightarrow 7 + 15 < \infty \Rightarrow d_4 = d_2 + c_{24} = 22, P_4 = 2$,
$Q = \{5\,\boxed{9}\,\langle 1\rangle, 9\,\boxed{47}\,\langle 2\rangle, 3\,\boxed{42}\,\langle 2\rangle, 4\,\boxed{22}\,\langle 2\rangle\}$.
ITERAZIONE 3. Estrazione del nodo dalla coda con etichetta più piccola: 5. Una volta estratto, lo si aggiunge all'albero e si effettua il relaxing su tutti gli archi uscenti dal nodo 5, quindi: …



Una volta eseguite tutte le iterazioni del While, cioè quando la coda sarà vuota, l'albero risultante sarà questo posto a destra. L'etichetta associata ad ogni nodo rappresenta la distanza dalla sorgente e, siccome questo è un albero, scelto un qualsiasi nodo esiste uno ed un solo percorso dalla sorgente fino a quel nodo: quello minimo. Ad esempio, il percorso minimo dalla sorgente 1 al nodo 7 del grafo G di partenza è 1 – 2 – 4 – 7.

Riprendendo il modello matematico generale del problema dei cammini minimi uno-a-tutti, per questo esempio i valori delle variabili x_{ij} nella soluzione ottima sono: $x_{76} = 1, x_{39} = 1, x_{15} = 1, x_{47} = 2, x_{83} = 2$, ... Per ricavare velocemente questi valori, l'idea è sfruttare l'albero con una strategia bottom-up, partendo dalle foglie dell'albero; quest'ultime non hanno archi uscenti, di conseguenza le variabili x_{ij}, con *j* foglia, valgono 1. A partire da un qualsiasi livello dell'albero, preso un nodo *i*, la quantità di flusso che deve entrare in *i* è data dalla quantità di flusso richiesta dai suoi sottoalberi sommata ad una unità di flusso (per il nodo *i*-esimo). Quindi, se un nodo *i* ha tre sottoalberi, allora la quantità di flusso che deve entrare in *i* sarà data da $f_1 + f_2 + f_3 + 1$.
Nell'albero seguente, sugli archi (i, j) sono posti i valori delle variabili x_{ij} corrispondenti (quindi, <u>non</u> i costi).



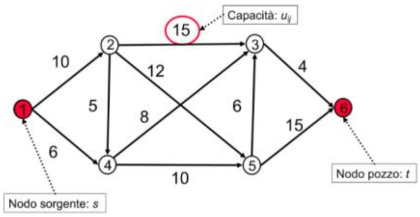
Versione tutti-a-tutti.
In breve, si vuole calcolare il percorso minimo da ogni nodo sorgente ad ogni altro nodo della rete. Questa variante consiste nell'eseguire la versione uno-a-tutti per ogni elemento della rete. Esistono algoritmi ad-hoc: una possibilità è utilizzare l'algoritmo di Dijkstra *n* volte (quindi, per ogni nodo del grafo), oppure l'algoritmo di Floyd-Warshall, facilmente implementabile con un triplo ciclo For, che calcola il cammino minimo per tutte le coppie di un grafo pesato e orientato.

Osservazione finale. Per la risoluzione di un problema del tipo uno-a-uno, in realtà si applica sempre l'algoritmo di Dijkstra (per calcolare l'albero dei cammini minimi uno-a-tutti), e per conoscere il percorso minimo da un nodo *i* a un nodo *j* – durante la computazione – non appena si estrae il nodo *j* dalla coda, si arresta l'algoritmo.

Problema del massimo flusso

Sia G = (V, A) un grafo orientato su cui sia definito un vettore <u>u</u> = [u_{ij}] delle capacità associate agli archi del grafo; inoltre, siano s e t due nodi distinti, detti rispettivamente sorgente (o *origine*) e pozzo (o *destinazione*). Il problema del flusso massimo consiste nel determinare la massima quantità di flusso che è possibile inviare da s a t attraverso G.

N.B.: In questo problema, i valori sugli archi rappresentano delle *capacità*, non dei costi. La **capacità** è la quantità di flusso che può passare lungo quell'arco.



In particolare:
- il nodo sorgente *fornisce* flusso (f);
- il nodo destinazione *assorbe* flusso (-f);
- tutti gli altri nodi sono nodi di transito.

Dunque, voglio spedire dalla sorgente s al pozzo t la <u>massima</u> quantità di flusso f senza violare i vincoli di capacità degli archi.

Il **modello matematico** del problema è posto a destra:
- le <u>variabili decisionali</u> x_{ij} rappresentano la quantità di flusso che passa sull'arco (i, j), e devono essere $0 \le x_{ij} \le u_{ij} \quad \forall (i,j) \in A$, dato che la quantità di flusso non può superare la capacità dell'arco;
- la <u>funzione obiettivo</u> è massimizzare f, cioè la quantità di flusso che esce dalla sorgente s ($f = \sum_{j \in Fs(s)} x_{sj}$);
- i <u>vincoli</u> sono

$$\sum_{j \in Fs(i)} x_{ij} - \sum_{k \in Bs(i)} x_{ki} = \begin{cases} 0 & \forall i \in V,\ \ i \neq s,t \\ f & \text{se } i = s \\ -f & \text{se } i = t \end{cases}.$$

max f

con vincoli :

Vincoli di **bilanciamento del flusso**

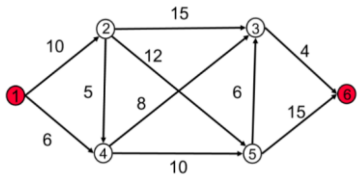
$$\sum_{j \in FS(i)} x_{ij} - \sum_{k \in BS(i)} x_{ki} = \begin{cases} 0 & \forall i \in V\ i \neq s,t \\ f & se\ \ i = s \\ -f & se\ \ i = t \end{cases} \qquad (1)$$

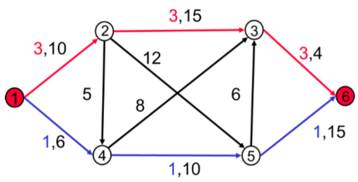
$$0 \le x_{ij} \le u_{ij} \qquad \forall\ (i, j) \in A \qquad (2)$$

Vincoli di **capacità**

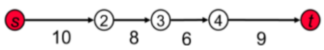
Si può risolvere questo problema utilizzando il metodo del Simplesso, ma anche in questo caso esistono delle proprietà interessanti che consentono di risolvere questo problema in modo più efficiente, utilizzando algoritmi ad-hoc.

Es. Si consideri il grafo nella figura a destra.
Nella figura in basso, con "x, y" s'intende x = quantità di flusso sull' arco (i, j) e y = capacità dell'arco (i, j).
Una possibile soluzione è data da $x_{12} = 3, x_{23} = 3, x_{36} = 3, x_{14} = 1$, $x_{45} = 1, x_{56} = 1$, mentre le altre x_{ij} sono ovviamente nulle. Questo appena descritto è un **flusso ammissibile** per il grafo G, dove si divide il flusso in due path da s a t, con valore $f = 4$.



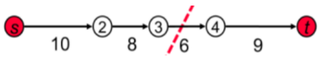


Vediamo alcuni *concetti fondamentali* del Problema del Massimo Flusso.



Si consideri il grafo precedente. Notiamo che la quantità massima di flusso che si può inviare è data da $\min_{(i,j) \in A} u_{ij}$, cioè dal minimo delle capacità di ogni arco (in questo caso è 6, in quanto $u_{34} = 6 < u_{23} = 8 < \ldots$).

Il segmento tratteggiato in figura mostra un **taglio s-t** del grafo, ossia un **partizionamento dei vertici** del grafo in due sottoinsiemi V_1 = {s, 2, 3} e V_2 = {4, t} tali che:



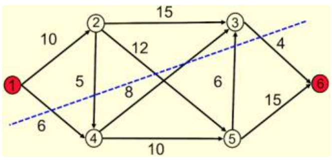
- il nodo sorgente appartiene a V_1;
- il nodo pozzo appartiene a V_2;
- $V_1 \cup V_2 = V$;
- $V_1 \cap V_2 = \emptyset$.

Un partizionamento che non soddisfi una delle prime due condizioni è detto semplicemente *taglio* del grafo. Per i nostri scopi, utilizzeremo solo tagli s-t.

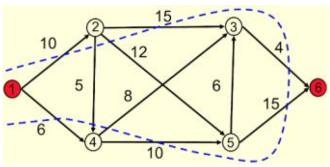
Def. Gli *archi diretti* del taglio [V_1, V_2] appartengono ad un insieme {(i, j) : i ∈ V_1, j ∈ V_2}. Sono, cioè, tutti gli archi del grafo in cui la coda appartiene al primo sottoinsieme e la testa appartiene al secondo sottoinsieme.

Def. Gli *archi inversi* del taglio [V_1, V_2] appartengono ad un insieme {(p, q) : p ∈ V_2, q ∈ V_1}. Sono, cioè, tutti gli archi del grafo in cui la coda appartiene al secondo sottoinsieme e la testa appartiene al primo sottoinsieme.

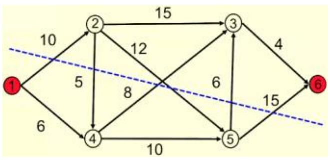
Nell'esempio precedente, l'arco (3, 4) è diretto, mentre non esistono archi inversi.



Taglio 1: V_1 = {1, 2, 3}, V_2 = {4, 5, 6} ⇒ ⇒ archi diretti del taglio = {(1, 4), (2, 4), (2, 5), (3, 6)}.



Taglio 2: V_1 = {1, 3, 5}, V_2 = {2, 4, 6} ⇒ ⇒ archi diretti del taglio = {(1, 2), (1, 4), (3, 6), (5, 6)}.



Taglio 3: V_1 = {1, 4, 5}, V_2 = {2, 3, 6} ⇒ ⇒ archi diretti del taglio = {(1, 2), (4, 3), (5, 3), (5, 6)}.

Il concetto di "arco diretto" è importante, in quanto ci serve per definire il concetto di "capacità di taglio". Dato il taglio s-t [V_1, V_2], la **capacità del taglio u[V_1, V_2]** è pari alla somma delle <u>capacità degli archi diretti</u> del taglio. Cioè, $u[V_1, V_2] = \sum_{(i,j)\in[V_1,V_2]} u_{ij}$.

Taglio 1: V_1 = {1, 2, 3}, V_2 = {4, 5, 6} ⇒ ⇒ archi diretti del taglio = {(1, 4), (2, 4), (2, 5), (3, 6)} ⇒ ⇒ *Capacità u*[V_1, V_2] = 6+5+12+4 = 27.

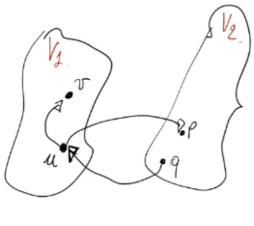
Taglio 2: V_1 = {1, 3, 5}, V_2 = {2, 4, 6} ⇒ ⇒ archi diretti del taglio = {(1, 2), (1, 4), (3, 6), (5, 6)} ⇒ ⇒ *Capacità u*[V_1, V_2] = 10+6+4+15 = 35.

Taglio 3: V_1 = {1, 4, 5}, V_2 = {2, 3, 6} ⇒ ⇒ archi diretti del taglio = {(1, 2), (4, 3), (5, 3), (5, 6)} ⇒ ⇒ *Capacità u*[V_1, V_2] = 10+8+6+15 = 39.

Sorge, ora, un dubbio. Data la rete in input, e dati dei tagli s-t della rete con le rispettive capacità, il flusso massimo f (dalla sorgente s al pozzo t), a quanto può equivalere *al più*? Si può dimostrare che f equivale al minimo delle capacità dei tagli.
Consideriamo il modello matematico, ed in particolare i vincoli.

$$\sum_{j\in Fs(i)} x_{ij} - \sum_{k\in Bs(i)} x_{ki} = \begin{cases} 0 & \forall i \in V,\ i \neq s,t \\ f & \text{se } i = s \\ -f & \text{se } i = t \end{cases}$$

Proprietà 1. *Il valore di un qualunque flusso ammissibile è minore o uguale alla capacità di un qualunque taglio.*
DIM. Sia <u>x</u> un flusso ammissibile e [V_1, V_2] un qualunque taglio s-t del grafo. Sommando i vincoli di bilanciamento del flusso relativi ai nodi *i* in V_1 (siccome *i* ≠ t) otteniamo $f = \sum_{i\in V_1}[\sum_{j\in Fs(i)} x_{ij} - \sum_{k\in Bs(i)} x_{ki}]$.
Ad esempio, infatti, considerando x_{uv} essa comparirà nella sommatoria una volta come flusso uscente dal nodo *u* (segno +) ed una volta come flusso entrante nel nodo *v* (segno –), pertanto si annulla. Questo accade ad ogni nodo in V_1.
Il risultato della sommatoria sarà dato da tutti gli archi che:
- dal nodo *u* in V_1 vanno verso i nodi in V_2;
- da tutti i nodi di V_2 vanno verso il nodo *u*.



Dalla precedente si può ricavare che:

$$f = \sum_{i\in V_1}\left[\sum_{j\in Fs(i)} x_{ij} - \sum_{k\in Bs(i)} x_{ki}\right] = \sum_{(i,j)\in[V_1,V_2]} x_{ij} - \sum_{(p,q)\in[V_2,V_1]} x_{pq}.$$

Inoltre:

$$\sum_{(i,j)\in[V_1,V_2]} x_{ij} - \sum_{(p,q)\in[V_2,V_1]} x_{pq} \le \sum_{(i,j)\in[V_1,V_2]} x_{ij} \le \sum_{(i,j)\in[V_1,V_2]} u_{ij} = u[V_1,V_2]. \qquad \square$$

Commentato [DE25]: Dai vincoli del problema, siccome le variabili x_{ij} sono limitate superiormente dalla capacità degli archi u_{ij} del taglio.

La capacità di un taglio s-t fornisce un limite superiore al valore del flusso f che posso spedire dalla sorgente al pozzo. Se ho un flusso ammissibile di valore f e riesco a trovare un taglio s-t la cui capacità è uguale ad f, allora posso concludere che il flusso che ho trovato è massimo.

Teorema (MAX FLOW – MIN CUT)**.** *Il flusso massimo che può essere spedito dalla sorgente al pozzo su un grafo orientato G è uguale alla capacità del taglio s-t minimo di G.*

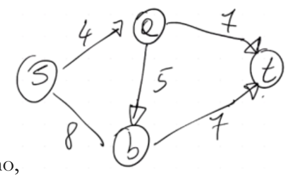
Per risolvere il problema utilizziamo i seguenti concetti.

Dato un grafo G = (V, A) ed un flusso ammissibile x̲ su G, il **grafo ausiliario** G(x̲) = (V', A') è così costruito:

- V' = V;
- per ogni arco (i, j) in A, A' contiene gli archi (i, j) e (j, i). La capacità u_{ij} di ogni arco (i, j) $\notin$ A è pari a 0;
- ad ogni arco di A' è associata una **capacità residua** $\boxed{r_{ij} = u_{ij} - x_{ij} + x_{ji}}$.

Si può utilizzare la seguente notazione.

Es. Si consideri l'arco in figura.



La capacità residua è un concetto che potrebbe essere fuorviante. Se poniamo $x_{sa} = 3$, allora si potrebbe pensare che la capacità residua è $4 - 3 = 1$; tuttavia, la reale capacità residua è $r_{sa} = u_{sa} - x_{sa} + x_{as}$, ed in questo caso x_{as} neanche esiste (è come se la sua capacità fosse 0).

Per quale motivo serve x_{as}? Durante la computazione dell'algoritmo che vedremo, individueremo nella rete una serie di path (chiamati *path aumentanti*, perché sfruttano le reti residue per continuare ad inviare flusso dalla sorgente al pozzo). Potrebbe succedere che, in qualche iterazione dell'algoritmo (in cui bisogna fare delle scelte), inviare del flusso su un determinato arco non è stata una scelta corretta; in sintesi, bisogna poter tornare indietro su quella scelta per poter inviare più flusso.

Considerando questa rete:

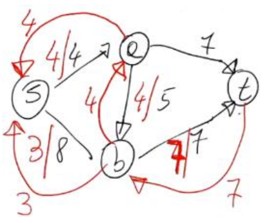
<u>path 1.</u> se poniamo $x_{sa} = 4|4$, $x_{ab} = 4|5$, $x_{bt} = 4|7$, otteniamo s – a – b – t, $\Delta = 4$, $f = 4$;

<u>path 2.</u> se poniamo $x_{sb} = 3|8$, $x_{bt} = (3+4)|7 = 7|7$, otteniamo s – b – t, $\Delta = 3$, $f = 7$.

Con quest'ultima scelta la rete viene saturata. A questo punto non si può più inviare altro flusso, in quanto sulla path 1 l'arco (s, a) ha raggiunto la capacità massima, mentre sulla path 2 l'arco (b, t) ha raggiunto la capacità massima, e non esistono altri cammini possibili per *t*.

Teoricamente dovremmo fermarci, ma in realtà il flusso massimo su questa rete non è 7 (quello raggiunto con la path 2), ed è qui che entrano in gioco le capacità residue:

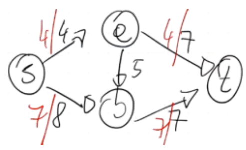
$$r_{sa} = u_{sa} - x_{sa} + x_{as} = 4 - 4 + 0 = 0$$
$$r_{as} = u_{as} - x_{as} + x_{sa} = 0 - 0 + 4 = 4$$



La capacità residua coincide con il costruire l'arco (a, s) avente capacità 4. Analogamente, possiamo costruire l'arco (b, a) avente capacità 4. Questo consente di restituire il flusso all'indietro (fino a *s*) se non conviene.

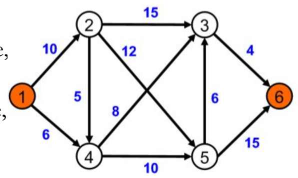
A questo punto, costruita la **rete residua** secondo questa formula, si scopre che si può ancora inviare del flusso al pozzo *t* aggiungendo un'altra path:

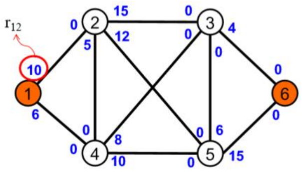
<u>path 3.</u> se poniamo $x_{sb} = (4+3)|8 = 7|8$, $x_{ba} = 4|4$, $x_{at} = 4|7$, otteniamo s – b – a – t, con $\Delta = 4$, $f = 11$.



Il risultato finale è posto a destra. L'arco (b, a) ha consentito di inviare la quantità massima di flusso.

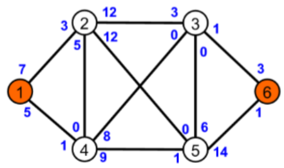
Es. Si consideri il grafo in figura. Siccome $u_{12} = 10$, possiamo accostare questo valore, sull'arco (1, 2), al nodo 1. Poiché (2, 1) ∉ A, $u_{21} = 0$ e possiamo accostare questo valore, sull'arco (1, 2), al nodo 2.





Calcoliamo la rete residua prodotta dal seguente flusso ammissibile x: $\underline{x}^T = [x_{12}\ \ x_{14}\ \ x_{23}\ \ x_{24}\ \ x_{25}\ \ x_{36}\ \ x_{43}\ \ x_{45}\ \ x_{53}\ \ x_{56}] = [3\ \ 1\ \ 3\ \ 0\ \ 0\ \ 3\ \ 0\ \ 1\ \ 0\ \ 1]$.
Vediamo come cambiano le capacità residue in base a questi valori:
- $x_{12} = 3$, allora $r_{12} = 10 - 3 + 0 = 7$ e $r_{21} = 0 - 0 + 3 = 3$;
- $x_{23} = 3$, allora $r_{23} = 15 - 3 + 0 = 12$ e $r_{32} = 0 - 0 + 3 = 3$;
- ...

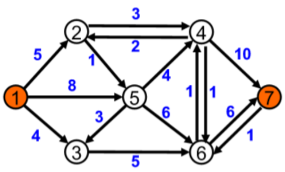
La rete residua risultante è posta a destra.



Ora, passiamo al metodo di **risoluzione** di questo problema, utilizzando grafo ausiliario e capacità residue:
- se un arco ha capacità residua maggiore di zero, significa che posso spedire ancora del flusso attraverso quell'arco;
- se riesco ad individuare un cammino da s a t sul grafo ausiliario, allora posso spedire del flusso addizionale dalla sorgente al pozzo;
- un cammino da s a t sul grafo ausiliario viene definito **cammino aumentante**;
- fino a quando nel grafo ausiliario sono presenti cammini aumentanti è sempre possibile incrementare il flusso da s a t.

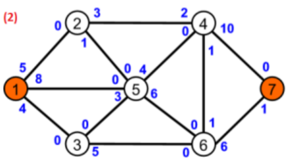
L'**algoritmo dei cammini aumentanti** risolve il problema del flusso massimo utilizzando il grafo ausiliario (o delle capacità residue) per stabilire come instradare il flusso sulla rete.
Consideriamo un grafo G = (V, A) ed un flusso ammissibile x (inizialmente il metodo considera il flusso nullo, ossia $x_{ij} = 0 \quad \forall (i, j) \in A$. I passi principali dell'algoritmo dei cammini aumentanti sono:
1. Individuare nel grafo ausiliario un qualsiasi cammino p dal nodo sorgente al nodo pozzo su cui è possibile far transitare una quantità di flusso Δ > 0 (cammino aumentante). Se non esiste tale cammino, l'algoritmo si arresta;
2. Il valore del flusso da inviare lungo il cammino p è pari alla capacità residua minima degli archi in p (i.e. $\Delta = \min\{r_{ij}: (i, j) \in p\}$);
3. Incrementare di Δ il valore del flusso f corrente, quindi $f = f + \Delta$, e aggiornare le capacità residue degli archi lungo il cammino p nel seguente modo: $r_{ij} = r_{ij} - \Delta$, e $r_{ji} = r_{ji} + \Delta$.

Esempio (Applicazione dell'algoritmo dei cammini aumentanti)**.**
Abbiamo il grafo G = (V, A) in input, con *s* = 1 nodo sorgente e *t* = 7 nodo pozzo. Il flusso ammissibile iniziale è $x_{ij} = 0 \quad \forall (i, j) \in A$.



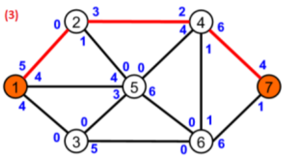
Prima di tutto, si trasforma questo grafo in un grafo non orientato, ponendo le capacità degli archi agli estremi di ognuno di essi. Otteniamo il grafo **(2)**, con $f = 0$.

(2)



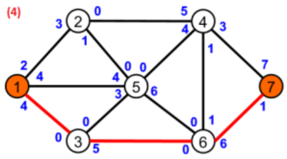
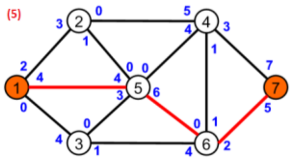
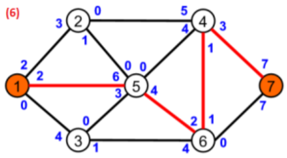
Iniziamo, poi, a scegliere i path aumentanti. Un *path aumentante* è un path da s a t sul grafo ausiliario. Viene chiamato "aumentante" perché permette di aumentare il flusso sul grafo da s a t utilizzando gli archi del path. Il flusso che posso spedire è uguale alla minima capacità residua degli archi del path.
Ad esempio, scegliendo *p* = 1 – 5 – 4 – 7, otteniamo $\Delta = 4$ e $f = 4$. Fatta questa scelta, come cambiano le capacità? Sottraendo Δ al lato sinistro e aggiungendo al lato destro di ogni arco, otterremo il grafo **(3)**.

(3)

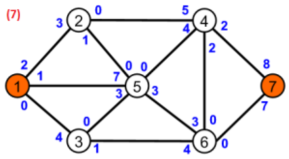


Ora, scegliendo *p* = 1 – 2 – 4 – 7, otteniamo $\Delta = 3$ e $f = f + \Delta = 7$. Fatta questa scelta, otterremo il grafo **(4)**, nella pagina seguente.
Ora, scegliendo *p* = 1 – 3 – 6 – 7, otteniamo $\Delta = 4$ e $f = f + \Delta = 11$. Fatta questa scelta, otterremo il grafo **(5)**, nella pagina seguente.
Ora, scegliendo *p* = 1 – 5 – 6 – 7, otteniamo $\Delta = 2$ e $f = f + \Delta = 13$. Fatta

questa scelta, otterremo il grafo **(6)**, di seguito.

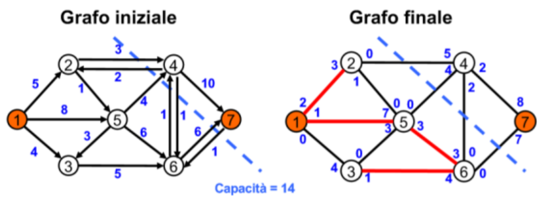
Ora, scegliendo p = 1 – 5 – 6 – 4 – 7, otteniamo $\Delta = 1$ e $f = f + \Delta = 14$. Fatta questa scelta, otterremo il grafo **(7)**. Siccome non riesco ad individuare un cammino aumentante, allora il flusso che ho individuato è ottimo.



Il <u>valore delle variabili decisionali</u>, per ogni arco del grafo di partenza, è pari alla differenza tra la capacità originale dell'arco meno quella residua nell'ultimo grafo ausiliario (se tale valore è negativo, la variabile decisionale varrà zero). Quindi, $x_{ij} = \max\{0, u_{ij} - r_{ij}\}$.
Ad esempio, per l'arco (1, 2) abbiamo una capacità iniziale pari a 5 e una finale pari a 2: quindi $x_{12} = 5 - 2 = 3$. Inoltre, $x_{42} = 2 - 5 = -3 \rightarrow x_{42} = 0$.
Analogamente, abbiamo: $x_{15} = 8 - 1 = 7$, $x_{13} = 4 - 0 = 4$, $x_{25} = 1 - 1 = 0$, $x_{24} = 3 - 0 = 3$, ...

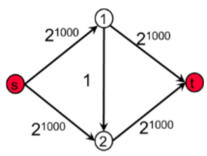
Si noti che nello schema generale del metodo dei cammini aumentanti ci sono dei dettagli che devono essere meglio chiariti:
• Come si individua un cammino aumentante? O come si mostra che non esiste un cammino aumentante?
• Come certificare che il flusso ottenuto è quello massimo?
La risposta a queste domande può essere ottenuta considerando una particolare implementazione dell'algoritmo del cammino aumentante che dà luogo al **Labeling Algorithm di Ford and Fulkerson**.

<u>Idea principale</u>:
- Tramite opportuni algoritmi di visita è possibile etichettare tutti i nodi che possono essere raggiunti dalla sorgente nel grafo ausiliario <u>tramite archi con capacità residua positiva</u>. Banalmente, quindi, prendiamo la nostra rete residua e costruiamo un nuovo grafo in cui sono presenti i soli archi con capacità residua positiva. Dopodiché, eseguendo la visita del grafo (es. DFS) si possono individuare tutti i nodi raggiungibili dalla sorgente, sfruttando la rete residua. Possiamo avere le seguenti due opzioni:
  - Se il pozzo viene etichettato durante la precedente visita, allora esiste un cammino aumentante da s a t tramite il quale è possibile inviare ulteriori unità di flusso;
  - Se il pozzo non viene etichettato, allora si costruisce un *taglio s-t* nel seguente modo:
    - in V_1 si inseriscono i nodi etichettati (ovvero raggiungibili da s);
    - in V_2 si inseriscono i nodi non etichettati.
- Poiché la capacità del taglio così costruito è pari al flusso f inviato fino a quel momento, dal teorema MinCut/MaxFlow tale flusso f è massimo.



Ad esempio, per il problema precedente si ha
- V_1 = {1, 2, 5, 3, 6};
- V_2 = {7, 4}.
A destra si può notare il taglio del grafo, avente capacità $u[V_1, V_2] = \sum_{(i,j)\in[V_1,V_2]} u_{ij} = 14$.

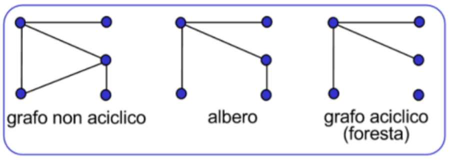
Per poter individuare il taglio minimo, la cui capacità sarà uguale al flusso massimo $f = 14$, è sufficiente controllare quali sono i nodi raggiungibili dalla sorgente 1 attraverso archi con capacità residua > 0 nell'ultimo grafo ausiliario.



La **complessità** dell'algoritmo dipende dal tempo impiegato ad individuare le path aumentanti (al più O(m), se m è il numero di archi del grafo, per la DFS/BFS). Nel caso peggiore, l'algoritmo ha complessità *pseudo-polinomiale*, dato che dipende sia dalla dimensione del problema che dalle capacità utilizzate sul grafo (oltre ad altri parametri).

Problema dell'albero di copertura minimo

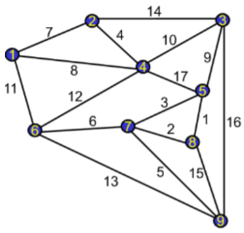
Sia G = (V, E) un grafo non orientato e connesso:
- G è **aciclico** se i suoi archi non formano cicli;
- un **albero** è un grafo connesso ed aciclico;
- ogni grafo aciclico è in generale l'unione di uno o più alberi e viene detto **foresta**.



Dato un grafo G = (V, E), le seguenti affermazioni sono equivalenti:
- G è un albero;
- ogni coppia di nodi di G è connessa da un unico cammino;
- G è aciclico e |E| = |V| - 1;
- G è connesso e |E| = |V| - 1.

Sia G = (V, E) un grafo non orientato e connesso. Un **albero ricoprente** (*spanning tree*) di G è un sottografo T di G tale che: T è un albero e contiene tutti i nodi di G.
Un grafo può avere più alberi ricoprenti.

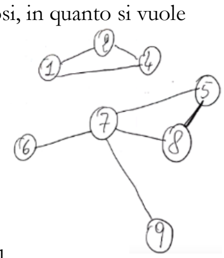
Esempio (Problema del Minimo Albero Ricoprente – Minimum Spanning Tree Problem)**.**
Sia G = (V, E) un grafo non orientato e connesso dove ad ogni arco $e_i \in E$ è associato un costo c_i. Il **costo** di un albero ricoprente T di G è dato dalla somma dei costi degli archi che lo compongono.
Problema: Determinare l'albero ricoprente di G di costo minimo.



Esistono numerose applicazioni di questo problema:
- determinare la rete di comunicazione più affidabile;
- determinare la connessione tra n centri a costo minimo (e.g., distribuzione del gas);
- progettare i circuiti elettronici per collegare fra loro le diverse componenti minimizzando la quantità di filo da utilizzare;
- ...

Il **modello matematico** di questo problema è diverso da quelli visti in precedenza: non basta, infatti, conoscere il modello del problema del flusso a costo minimo. Proviamo a formularlo, considerando il grafo a destra:
- le variabili decisionali sono $x_{ij} = \begin{cases} 1 & \text{se scelgo } (i,j) \\ 0 & \text{altrimenti} \end{cases}$, per ogni (i, j) in E;
- la funzione obiettivo è $\min \sum_{(i,j)\in E} c_{ij}x_{ij}$;
- il problema principale riguarda i vincoli, dato che si vuole costruire un albero di copertura del grafo. Teniamo in considerazione l'affermazione secondo cui, se G è un albero, allora *G è aciclico e* |E| = |V| - 1. Dunque, otteniamo che $\sum_{(i,j)\in E} x_{ij} = |V| - 1$. Ma come stabiliamo matematicamente che un grafo è aciclico?

Il modello matematico (incompleto) appena definito, dato il grafo in esempio, ci porta ad individuare la soluzione ottima scegliendo il numero (in questo caso, |V| - 1 = 8) di archi meno costosi, in quanto si vuole minimizzare il costo. Otteniamo il grafo seguente, non aciclico.
Per ricavare il modello matematico corretto del Problema, bisogna inserire dei vincoli lineari che consentano di evitare (o *eliminare*) cicli. Prendendo in considerazione il sottografo costituito da 1, 2 e 4, si può evitarne il ciclo se $x_{12} + x_{24} + x_{14} \leq 2$, cioè se non si scelgono tutti gli archi che collegano i suoi nodi.



In generale, per ogni sottografo con vertici S ⊂ V, |S| ≥ 3: $\sum_{\substack{(i,j)\in E \\ i\in S, j\in S}} x_{ij} \leq |S| - 1$.

Pertanto, possiamo eliminare l'arco (7, 5) – dato che ha il costo maggiore tra gli archi del ciclo – per scegliere l'arco (4, 5): otterremo un albero ricoprente.

Questa modalità di costruzione dei vincoli è detta *Subtour Elimination* (cioè, eliminazione dei sottocircuiti): essa incrementa la complessità, dato che aumenta considerevolmente il numero di vincoli del problema.

Un'altra modalità di costruzione dei vincoli è detta *Cut Formulation*, la quale considera l'affermazione secondo cui, se G è un albero, allora *G è connesso e* $|E| = |V| - 1$. Anche in questo caso $\sum_{(i,j)\in E} x_{ij} = |V| - 1$, ma anziché distruggere i sottocircuiti occorre lavorare sulla connettività. Allora, considerando l'ultimo grafo costruito, abbiamo ottenuto più componenti connesse: per inserire nel modello matematico un vincolo che evita ciò, poniamo $\sum_{\substack{(i,j)\in E \\ i\in S, j\in V\setminus S}} x_{ij} \geq 1$, per ogni $S \subset V$, $|S| \geq 1$; questo consente di aggiungere un arco che colleghi S a V\S.

Subtour Elimination

$$\min \sum_{(i,j)\in E} c_{ij} x_{ij}$$

$$\sum_{(i,j)\in E} x_{ij} = n - 1$$

$$\sum_{\substack{(i,j)\in E \\ i\in S, j\in S}} x_{ij} \leq |S| - 1 \qquad \forall S \subset V,\ |S| \geq 3$$

$$x_{ij} \in \{0,1\} \qquad \forall\ (i,j) \in E$$

Cut Formulation

$$\min \sum_{(i,j)\in E} c_{ij} x_{ij}$$

$$\sum_{(i,j)\in E} x_{ij} = n - 1$$

$$\sum_{\substack{(i,j)\in E \\ i\in S, j\in V\setminus S}} x_{ij} \geq 1 \qquad \forall S \subset V,\ |S| \geq 1$$

$$x_{ij} \in \{0,1\} \qquad \forall\ (i,j) \in E$$

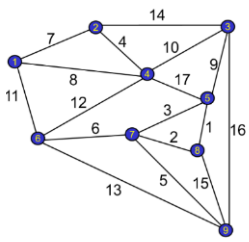
In entrambi i casi, si ottiene una complessità esponenziale a causa del numero di vincoli. La chiave è costruire dinamicamente il modello matematico ottimale, cioè si parte dal modello con il solo insieme di vincoli $\sum_{(i,j)\in E} x_{ij} = |V| - 1$, e man mano che si identificano cicli si aggiunge il vincolo per distruggerlo o per connettere le (sotto)componenti, risolvendo il problema in modo incrementale. Maggiori dettagli al corso di Ottimizzazione.

Algoritmi risolutivi

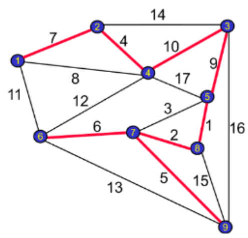
Algoritmo di Kruskal. Sia G = (V, E) un grafo non orientato con *n* nodi ed *m* archi.
1. Ordinare gli archi $e_1, e_2, ..., e_m$ in modo non decrescente ($c_1 \leq c_2 \leq ... \leq c_m$). Siano $E^0 = \varnothing$, $ST^0 = (V, \varnothing)$ e k=1.
2. Se $(V, E^{k-1} \cup \{e_k\})$ è un grafo aciclico, allora $ST^k = (V, E^k)$, con $E^k = E^{k-1} \cup \{e_k\}$, altrimenti e_k viene scartato, $E^k = E^{k-1}$ e $ST^k = ST^{k-1}$.
3. Se $|E^k| = n - 1$, allora l'algoritmo si arresta ed ST^k è l'albero ricoprente cercato, altrimenti k = k + 1 e si ritorna al passo (2).

Es. $n = 9$, $m = 17$.

(5,8)	(7,8)	(5,7)	(2,4)	(7,9)	(6,7)	(1,2)	(1,4)	(3,5)	(3,4)	(1,6)	(4,6)	(6,9)	(2,3)	(8,9)	(3,9)	(4,5)
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17



Nella tabella, nella prima riga sono posti gli archi e nella seconda riga i relativi costi.
Siccome il grafo ha $|V| = 9$, bisogna scegliere $|V| - 1 = 8$ archi.
Un costo cerchiato nella tabella indica che quell'arco è stato selezionato. Dunque:



- scelgo (5, 8). Crea cicli? No, quindi lo seleziono (1);
- scelgo (7, 8). Crea cicli? No, quindi lo seleziono (2);
- scelgo (5, 7). Crea cicli? Sì, quindi non lo seleziono;
- scelgo (2, 4). Crea cicli? No, quindi lo seleziono (3);
- scelgo (7, 9). Crea cicli? No, quindi lo seleziono (4);
- scelgo (6, 7). Crea cicli? No, quindi lo seleziono (5);
- scelgo (1, 2). Crea cicli? No, quindi lo seleziono (6);
- scelgo (1, 4). Crea cicli? Sì, quindi non lo seleziono;
- scelgo (3, 5). Crea cicli? No, quindi lo seleziono (7);
- scelgo (3, 4). Crea cicli? No, quindi lo seleziono (8).

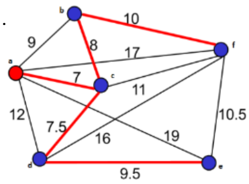
L'ultimo grafo ha come albero di copertura minimo quello costituito dagli archi in rosso. Segue la tabella.

~~(5,8)~~	~~(7,8)~~	~~(5,7)~~	~~(2,4)~~	~~(7,9)~~	~~(6,7)~~	~~(1,2)~~	~~(1,4)~~	~~(3,5)~~	~~(3,4)~~	(1,6)	(4,6)	(6,9)	(2,3)	(8,9)	(3,9)	(4,5)
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

• Sia 4+k il nuovo peso dell'arco (2, 4) nel grafo. Per quali valori di k l'albero di copertura minimo non cambia? Per come funziona l'algoritmo di Kruskal, al crescere di k, l'arco (2, 4) viene spostato verso destra nella lista. Per ottenere il massimo valore da assegnare a k prima che (2, 4) venga escluso, non serve arrivare all'ultimo arco selezionato: bisogna spostarsi verso destra e individuare quando quell'arco viene scartato perché crea un ciclo con gli archi precedentemente selezionati (in questo caso, fino a (1, 4)).
• Sia 8+k il nuovo peso dell'arco (1, 4) in G. Per quali valori di k l'arco (1,4) verrà inserito nella soluzione ottima?
In questo caso, bisogna decrementare il suo peso fino ad ottenere 6, in modo tale da selezionare questo arco al posto di (1, 2).

Algoritmo di Prim. Sia G = (V, E) un grafo non orientato con *n* nodi ed *m* archi.
1. Selezionare un qualsiasi vertice $v_s \in V$ e porre $V^0 = \{v_s\}$, $E^0 = \emptyset$, k=1.
2. Dato il taglio $[V^{k-1}, V \setminus V^{k-1}]$, selezionare l'arco del taglio (v_i, v_h) avente costo minimo e porre $V^k = V^{k-1} \cup \{v_h\}$ e $E^k = E^{k-1} \cup \{(v_i, v_h)\}$.
3. Se $|E^k| = n - 1$, allora l'algoritmo si arresta ed $ST = (V^k, E^k)$ è l'albero ricoprente cercato, altrimenti k = k+1 e si ritorna al passo (2).

L'algoritmo di Prim $O(|E|\log|V|)$ è più efficiente di quello di Kruskal $O(|E|\log|E|)$.



Es. Supponiamo di voler partire dal nodo *a*. Dunque, inizialmente abbiamo $V^0 = \{a\}$.
ITERAZIONE 1. Consideriamo il taglio tra V^0 e i restanti vertici del grafo: l'arco di costo minimo è (a, c), avente costo 7, quindi $V^1 = \{a, c\}$.
ITERAZIONE 2. Consideriamo il taglio tra V^1 e i restanti vertici del grafo: l'arco di costo minimo è (c, d), avente costo 7.5, quindi $V^2 = \{a, c, d\}$.
ITERAZIONE 3. Consideriamo il taglio tra V^2 e i restanti vertici del grafo: l'arco di costo minimo è (c, b), avente costo 8, quindi $V^3 = \{a, c, d, b\}$.
ITERAZIONE 4. Consideriamo il taglio tra V^3 e i restanti vertici del grafo: l'arco di costo minimo è (d, e), avente costo 9.5, quindi $V^4 = \{a, c, d, b, e\}$.
ITERAZIONE 5. Consideriamo il taglio tra V^4 e i restanti vertici del grafo: l'arco di costo minimo è (b, f), avente costo 10, quindi $V^5 = \{a, c, d, b, e, f\}$.
Al termine, otteniamo $E^5 = \{(a, c), (c, d), (c, b), (d, e), (b, f)\}$. Siccome $|E^5| = |V| - 1 = 5$, l'algoritmo si arresta ed otteniamo l'albero in figura (con archi in rosso).

Si noti che esiste una sottile differenza tra albero dei cammini minimi ed albero di copertura minimo:
- l'albero dei cammini minimi, data la coppia di nodi *sorgente* – destinazione, indica qual è il percorso minimo tra quei due nodi sul grafo;
- l'albero di copertura minimo è l'albero la cui somma degli archi è minima, ma non è detto che il percorso tra due nodi sia minimo.

Infatti, dato l'esempio seguente, nell'albero dei cammini minimi il cammino minimo da 1 a 3 è l'arco diretto avente costo 9; applicando l'algoritmo di Kruskal/Prim otteniamo l'albero di copertura minimo a destra, che consente di giungere da 1 a 3 passando per 2, con costo 13.