

Ottimizzazione Combinatoria 2

A.A. 2004/05

Stefano Smriglio
Dipartimento di Informatica
Università di L'Aquila
smriglio@di.univaq.it

+

+

Indipendenza lineare ed affine

- un vettore $y \in \mathbb{R}^n$ è combinazione lineare dei vettori $\{x^1, \dots, x^k\}$ se esistono k moltiplicatori reali $\lambda_1, \dots, \lambda_k$ tali che $y = \sum_{i=1}^k \lambda_i x^i$. Una combinazione lineare con $\lambda_i \geq 0$, $i = 1, \dots, k$ è detta *conica*. Una combinazione lineare con $\sum_{i=1}^k \lambda_i = 1$ si dice *affine*. Una combinazione conica ed affine si dice *convessa*
- Un sottoinsieme $X \subset \mathbb{R}^n$ si dice *linearmente (affinemente) indipendente* se nessun vettore di X può essere espresso come combinazione lineare (affine) degli altri.

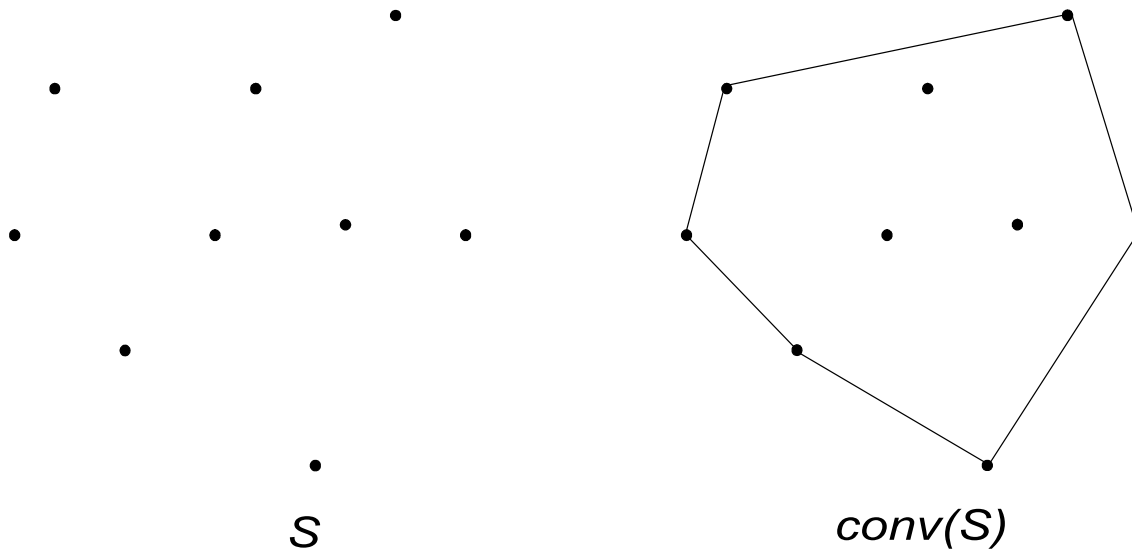
+

+

+

Involucro convesso

Dato un insieme $S \subseteq \mathbb{R}^n$, l'*inviluppo convesso* (conico) $conv(S)$ ($cone(S)$) di S è l'insieme di tutti i vettori ottenibili per combinazione convessa di sottoinsiemi finiti di vettori di S .



- Un insieme $X \subseteq \mathbb{R}^n$ è convesso se e solo se $X = conv(X)$; $conv(S)$ è il più piccolo insieme convesso che contiene S

+

+

+

Ottimizzare su $\text{conv}(S)$

Proposizione 1 Sia $S = \{x^1, \dots, x^k\} \subseteq \mathbb{R}^n$ un insieme finito e sia $w \in \mathbb{R}^n$. Allora, $\max\{w^T x : x \in S\} = \max\{w^T y : y \in \text{conv}(S)\}$

Dimostrazione. Sia $y = \lambda_1 x^1 + \dots + \lambda_k x^k$, con $\lambda_1, \dots, \lambda_k$ moltiplicatori non negativi tali che $\sum_{i=1, \dots, k} \lambda_i = 1$. Sia inoltre $r = \operatorname{argmax}_{i=1, \dots, k} (w^T x^i)$. Allora,

$$w^T y = \lambda_1 w^T x^1 + \dots + \lambda_k w^T x^k \leq \lambda_1 w^T x^r + \dots + \lambda_k w^T x^r = w^T x^r$$

□

+

+

+

Appartenenza a $\text{conv}(S)$

Dato un insieme $S = \{x^1, \dots, x^k\} \subseteq \mathbb{R}^n$, è possibile certificare facilmente se un vettore $y \in \mathbb{R}^n$ appartiene a $\text{conv}(S)$. Infatti, ciò è equivalente a determinare se il sistema lineare:

$$\begin{aligned} \sum_{i=1}^k \lambda_i &= 1 \\ \sum_{i=1}^k \lambda_i \cdot x^i &= y \\ \lambda_i &\geq 0, \quad i = 1, \dots, k \end{aligned} \tag{1}$$

ammette soluzione oppure no.

+

+

+

Separazione

Proposizione 2 Sia $S = \{x^1, \dots, x^k\} \subseteq \mathbb{R}^n$ un insieme finito e sia $v \in \mathbb{R}^n \setminus \text{conv}(S)$. Allora esiste una disuguaglianza $\alpha^T x \leq \beta$ che separa v da $\text{conv}(S)$, cioè risulta $\alpha^T y \leq \beta$ per ogni $y \in \text{conv}(S)$, ma $\alpha^T v > \beta$

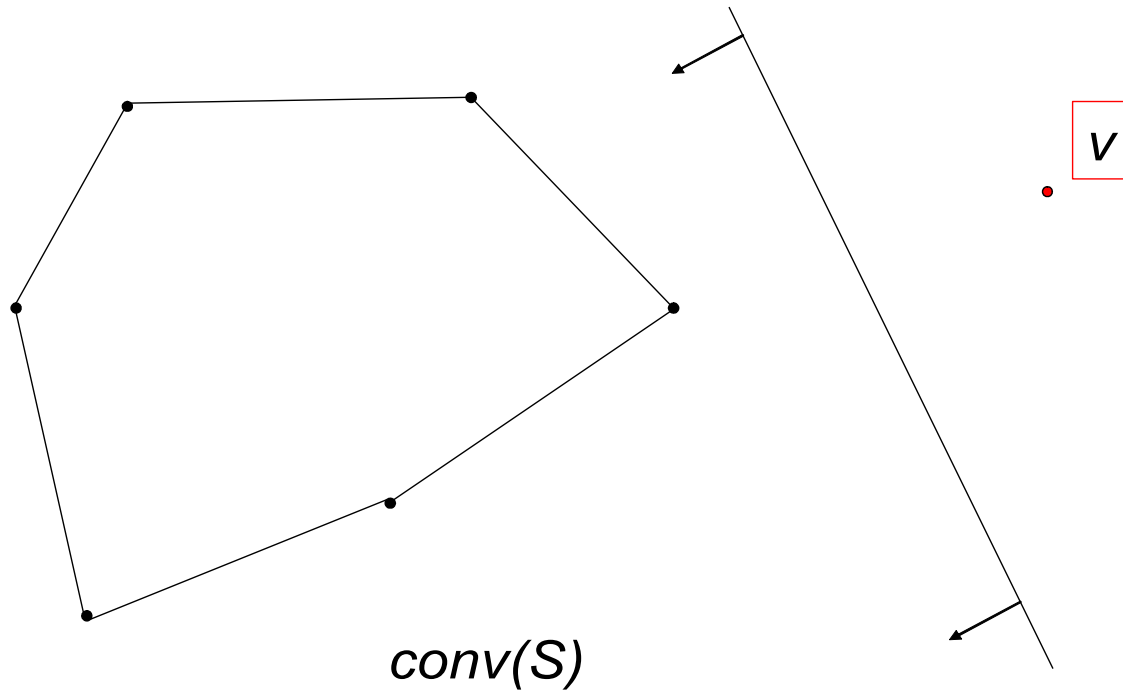
Dimostrazione. Per ipotesi, il sistema (1) non ammette soluzione. Quindi, per il lemma di Farkas, esiste $z \in \mathbb{R}$ e $p \in \mathbb{R}^n$ tali che

$$\begin{aligned} p^T x^i + z &\leq 0, & i = 1, \dots, k \\ p^T v + z &> 0. \end{aligned}$$

Si ponga $\alpha = p$ e $\beta = -z$. Essendo $\alpha^T x^i \leq \beta$, per $i = 1, \dots, k$, la Proposizione 1 implica $\alpha^T y \leq \beta$, per ogni $y \in \text{conv}(S)$. $\square$

+

Separazione

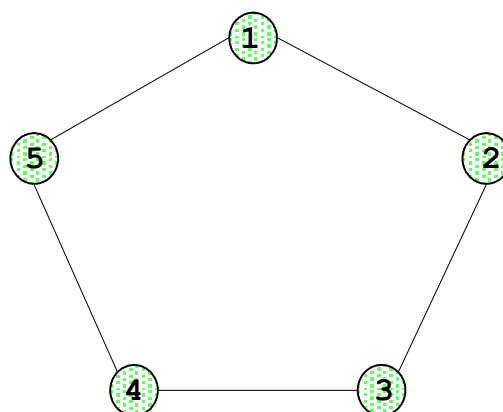


+

+

Esempio

Sia $S \subseteq \mathbb{R}^5$ la collezione di tutti gli insiemi stabili massimali del seguente grafo:



$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

+

+

+

Esempio

Verifichiamo se il vettore $y^T = (0.5, 0.5, 0.5, 0.5, 0.5)$ appartiene a $\text{conv}(S)$.

$$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 = 1$$

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \lambda_5 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$$

$$\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5 \geq 0$$

+

+

+

Esempio

il sistema

$$z] \quad \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 = 1$$

$$p_1] \quad \lambda_1 + \lambda_2 = 1/2$$

$$p_2] \quad \lambda_3 + \lambda_4 = 1/2$$

$$p_3] \quad \lambda_1 + \lambda_5 = 1/2$$

$$p_4] \quad \lambda_2 + \lambda_3 = 1/2$$

$$p_5] \quad \lambda_4 + \lambda_5 = 1/2$$

$$\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5 \geq 0$$

non ammette soluzione

+

+

+

Esempio

il sistema alternativo

$$p_1 + p_2 + p_3 + p_4 + p_5 + 2z > 0$$

$$p_1 + p_3 + z \leq 0$$

$$p_1 + p_4 + z \leq 0$$

$$p_2 + p_4 + z \leq 0$$

$$p_2 + p_5 + z \leq 0$$

$$p_3 + p_5 + z \leq 0$$

ammette la soluzione $(1, 1, 1, 1, 1, z = -2)$

quindi, la disuguaglianza $x_1 + x_2 + x_3 + x_4 + x_5 \leq 2$ separa y da $\text{conv}(S)$

+

Poliedri

- l'insieme delle soluzioni di un sistema finito di disuguaglianze lineari $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ è detto *poliedro*. Un poliedro limitato si dice *politopo*.
- La dimensione $\dim(P)$ di un poliedro P è il massimo numero di punti affinementemente indipendenti in P (*rango affine*) meno uno.
- Un punto $y \in P$ è detto *vertice* di P se non può essere ottenuto come combinazione convessa di altri punti di P .
- Un poliedro ha un insieme finito di vertici; Un poliedro si dice *puntato* se ha almeno un vertice; ogni politopo è puntato.

Disuguaglianze valide

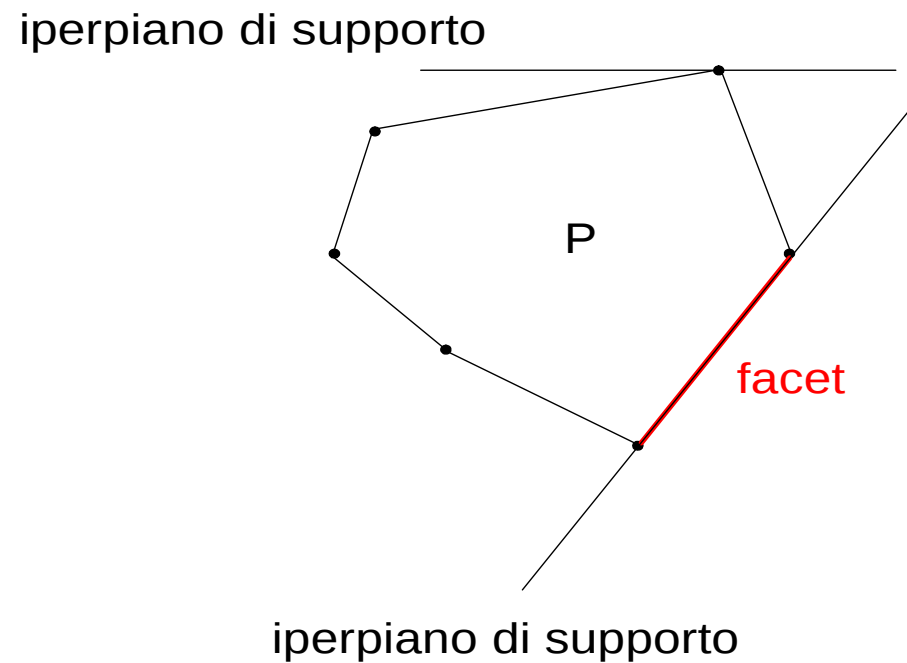
- Una disuguaglianza $\alpha x \leq \beta$ è *valida* per un poliedro P se $P \subseteq \{x : \alpha x \leq \beta\}$;
- una disuguaglianza $\alpha x \leq \beta$ valida per P definisce una *faccia* $F = P \cap \{x \in \mathbb{R}^n : \alpha x = \beta\}$ di P ; l'insieme $\{x \in \mathbb{R}^n : \alpha x = \beta\}$ è detto *iperpiano di supporto* di F ;
- Data una faccia F di P risulta ovviamente $\dim(F) \leq \dim(P)$. Se $\dim(F) = \dim(P)$, il poliedro P è contenuto in F ed F è detta faccia *impropria*. Una faccia F ha dimensione 0 se e solo se è un vertice di P .

+

+

Facce massimali

Una faccia F di P per cui $\dim(F) = \dim(P) - 1$, F è detta *faccia massimale* o *facet*;



+

Caratterizzazione di un politopo

Teorema 1 *Un politopo coincide con l'involucro convesso dei suoi vertici*

Dimostrazione. Sia P un politopo non vuoto e siano $v^1, \dots, v^k$ i suoi vertici (P è puntato). Ovviamente $\{v^1, \dots, v^k\} \subseteq P$. Supponiamo che esista $u \in P \setminus \text{conv}(\{v^1, \dots, v^k\})$. Allora, per la Proposizione 2, esiste una disuguaglianza $\alpha x \leq \beta$ che separa u da $\text{conv}(\{v^1, \dots, v^k\})$. Sia $\beta^* = \max\{\alpha^T x : x \in P\}$. Essendo $u \in P$ risulta $\beta^* > \beta$. Ma allora la faccia $F = \{x \in P : \alpha^T x = \beta^*\}$ non contiene vertici di P , contraddizione. $\square$

Caratterizzazione di un politopo

Teorema 2 *Un insieme P è un politopo se e solo se esiste un insieme finito V tale che P coincide con l'involucro convesso di V .*

Dimostrazione. Una direzione deriva immediatamente dal Teorema 1. Per l'altra direzione, sia $V = \{v^1, \dots, v^k\} \subseteq \mathbb{R}^n$ un insieme finito e sia $P = \text{conv}(V)$.

Definiamo un politopo che contiene tutte le disuguaglianze valide per P e, quindi, ne scegliamo un insieme finito utilizzando il Teorema 1.

Si consideri l'insieme $Q \subseteq \mathbb{R}^{n+1}$ definito da

$$\{(\alpha, \beta) : \alpha \in \mathbb{R}^n, \beta \in \mathbb{R}, -e \leq \alpha \leq e, -1 \leq \beta \leq 1, \alpha^T v \leq \beta \text{ per ogni } v \in V\}$$

+

+

Dimostrazione

Essendo V finito, l'insieme $Q = \{(\alpha, \beta) : \alpha \in \mathbb{R}^n, \beta \in \mathbb{R}, -e \leq \alpha \leq e, -1 \leq \beta \leq 1, \alpha^T v \leq \beta \text{ per ogni } v \in V\}$ è un politopo.

Quindi, per il Teorema 1, coincide con l'involucro convesso dei suoi vertici $\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \dots, \begin{pmatrix} a_m \\ b_m \end{pmatrix}$.

Dimostriamo che il sistema lineare

$$\begin{aligned} a_1^T x &\leq b_1 \\ a_2^T x &\leq b_2 \\ &\dots \\ a_m^T x &\leq b_m \end{aligned} \tag{2}$$

definisce P .

+

+

+

Dimostrazione

1. P è contenuto nella soluzione del sistema (2).

Sia $\bar{x} \in P$. Allora, $\bar{x} = \lambda_1 v^1 + \dots, \lambda_k v^k$, con $\sum_{r=1, \dots, k} \lambda_r = 1$, $\lambda_r \geq 0$.
Quindi, per ogni $i = 1, \dots, m$ risulta

$$a_i^T \bar{x} = \lambda_1 a_i^T v^1 + \dots + \lambda_k a_i^T v^k \leq \lambda_1 b_i + \dots + \lambda_k b_i = b_i \quad (3)$$

+

+

+

Dimostrazione

2. Ogni soluzione del sistema (2) è contenuta in P .

Sia $\bar{x}$ una soluzione del sistema (2) e si assuma $\bar{x} \notin P$. Allora, per la Proposizione 2, esiste una disuguaglianza $w^T x \leq t$ che separa $\bar{x}$ da P . Si può inoltre assumere $-e \leq w \leq e$, $-1 \leq t \leq 1$ e, cioè, $\begin{pmatrix} w \\ t \end{pmatrix} \in Q$ (questo è sempre ottenibile dividendo w e t per una costante).

Ciò implica che $\begin{pmatrix} w \\ t \end{pmatrix}$ può essere scritto come combinazione convessa dei vertici di Q , $\begin{pmatrix} w \\ t \end{pmatrix} = \gamma_1 \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \dots, \gamma_m \begin{pmatrix} a_m \\ b_m \end{pmatrix}$, con $\sum_{r=1, \dots, m} \gamma_r = 1$, $\gamma_r \geq 0$.

Quindi,

$$w^T \bar{x} = \gamma_1 a_1^T \bar{x} + \dots, \gamma_m a_m^T \bar{x} \leq \gamma_1 b_1 + \dots + \gamma_m b_m = t \quad (4)$$

ma questa è una contraddizione a $w^T \bar{x} > t$. Quindi, $\bar{x} \in P$. □

+

Dominanza fra disuguaglianze

- Date due disuguaglianze $ax \leq \alpha$ e $cx \leq \gamma$, si dice che la prima *domina* la seconda se esiste un $\lambda > 0$ tale che $a = \lambda c$ e $\gamma \leq \lambda \alpha$. Inoltre risulta $\{x \in \mathbb{R}^n : a^T x \leq \alpha\} \subseteq \{x \in \mathbb{R}^n : c^T x \leq \gamma\}$. Se gli insiemi di soluzioni coincidono le disuguaglianze si dicono equivalenti;

Proposizione 3 *Una disuguaglianza $ax \leq \alpha$ è valida per un poliedro $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ se e solo se è equivalente o dominata da una disuguaglianza della forma $u^T Ax \leq u^T b$, con $u \geq 0_m$, ovvero da una combinazione conica delle disuguaglianze del sistema.*

Rappresentazione minimale di un poliedro

- la rappresentazione esterna $Ax \leq b$ di un poliedro non è univocamente determinata. Infatti, è sempre possibile aggiungere alla rappresentazione nuove disuguaglianze valide per P , senza modificare l'insieme delle sue soluzioni.
- la rappresentazione $Ax \leq b$ di P è *minimale* se rimuovendo una qualsiasi disuguaglianza si ottiene un sistema $A'x \leq b'$ tale che $P \subset \{x \in \mathbb{R}^n : A'x \leq b'\}$;

Proposizione 4 *La rappresentazione $Ax \leq b$ di un poliedro P di dimensione n è minimale se e solo se ciascuna disuguaglianza del sistema definisce una faccia massimale di P . Inoltre, ogni faccia massimale di P è definita da una disuguaglianza del sistema $Ax \leq b$.*

+

+

Programmazione Lineare $\{0, 1\}$

Dato un insieme di vettori $S \subseteq \{0, 1\}^n$ (*regione ammissibile*) ed un vettore di costo $c \in \mathbb{R}^n$, determinare una soluzione ammissibile $x^* \in S$ tale da minimizzare la funzione lineare $c^T x$.

PL01

$$\begin{array}{ll} \min & cx \\ & x \in S \end{array}$$

Commento 5 Essendo $S \subseteq \{0, 1\}^n$, S ha un numero finito di elementi

Commento 6 Non ci sono ipotesi sulla struttura di S .

+

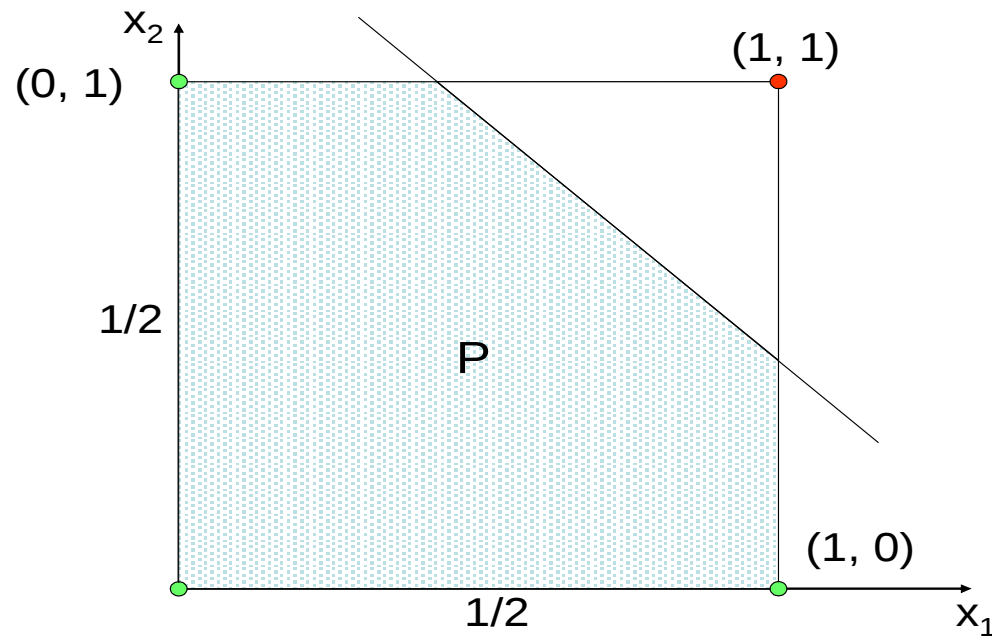
+

+

Formulazioni lineari

Definizione 7 Un poliedro $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ è una formulazione lineare del problema (S, c) se e solo se $S = P \cap \{0, 1\}^n$

Esempio: $S = \{(0, 0), (1, 0), (0, 1)\}$



+

Esistenza di una formulazione Lineare

Teorema 3 *Un problema (S, c) di PL01 ammette sempre la formulazione lineare $\text{conv}(S)$ con la proprietà che $x \in S$ se e solo se x è un vertice di $\text{conv}(S)$.*

Dimostrazione. Poiché l'insieme S è finito, il teorema 2 implica che $\text{conv}(S)$ è un politopo e, per il Teorema 1, i suoi vertici appartengono ad S . Inoltre, tutti i vettori in $\{0, 1\}^n$ sono vertici del cubo unitario $\{x \in \mathbb{R}^n : 0 \leq x_i \leq 1, \text{ per } i = 1, \dots, n\}$ e nessun vettore $x \in \{0, 1\}^n$ è esprimibile come combinazione convessa degli altri. Quindi,

- tutti e soli i punti di S sono vertici di $\text{conv}(S)$;
- tutti i punti $x \in \{0, 1\}^n \setminus S$ non appartengono a $\text{conv}(S)$

□

+

+

Rilassamento lineare

Sia P una formulazione lineare del problema (S, c) di PL01. Questo può essere riscritto come

$$\begin{aligned} \min \quad & c^T x \\ & x \in P \\ & x \in \{0, 1\}^n \end{aligned}$$

Il problema di PL

$$\begin{aligned} \min \quad & c^T x \\ & x \in P \end{aligned}$$

è detto *rilassamento lineare* di (S, c) .

+

Lower bound

La soluzione ottima $\hat{x}$ del rilassamento lineare fornisce una limitazione inferiore (per un problema di minimo) del valore ottimo intero:

$$c^T \hat{x} \leq \min\{c^T x : x \in S\}$$

- Se $\hat{x} \in S$, allora è anche soluzione ottima di (S, c) . Inoltre, se $\hat{x} \notin S$ ma esiste una soluzione $\bar{x} \in S$ per cui $c^T \hat{x} = c^T \bar{x}$, allora $\hat{x}$ è ottima di (S, c) .
- Dato un problema (S, c) di PL01 è possibile definire diverse formulazioni lineari di (S, c) . Diverse formulazioni producono diversi lower bound del valore ottimo.

+

+

Gerarchia di formulazioni

Dato un problema (S, c) di PL01, siano P_1 e P_2 due formulazioni lineari distinte di (S, c) .

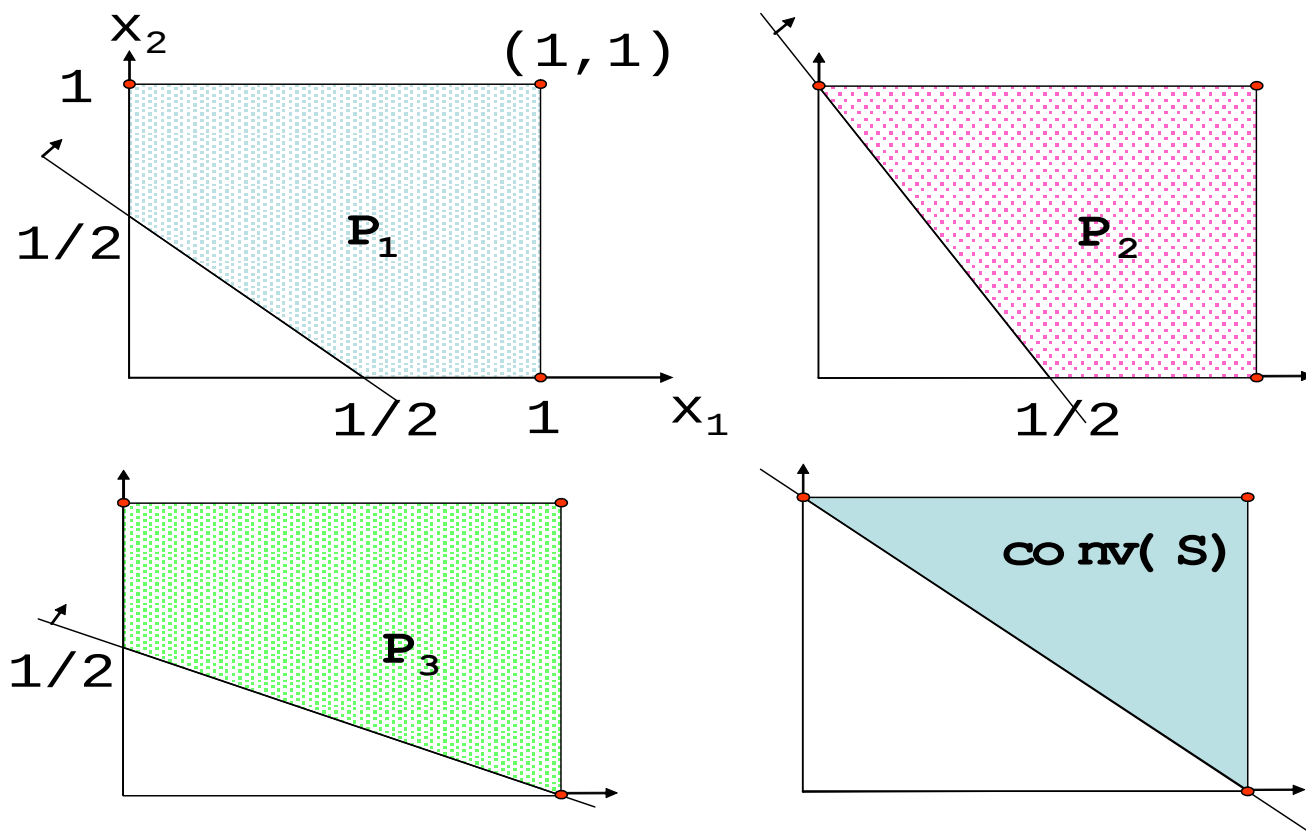
- consideriamo P_1 migliore di P_2 se $\min_{x \in P_1} c^T x > \min_{x \in P_2} c^T x$
- è importante definire un criterio di confronto delle formulazioni indipendente dalla funzione obiettivo;

Definizione 8 *Si dice che P_1 è migliore di P_2 se e solo se $P_1 \subset P_2$* $\square$

- Quindi, la migliore formulazione possibile è rappresentata dal poliedro contenuto in tutti i poliedri contenenti S , cioè da $\text{conv}(S)$

+

Gerarchia di formulazioni



P_2 migliore di P_1 ; P_3 migliore di P_1 ; P_2 e P_3 non confrontabili

$\text{conv}(S)$ formulazione "ottima"