Appunti sul Teorema di Dualità e sul Lemma di Farkas

Gabriele Rastello

27 aprile 2020

Indice

1 Problemi lineari **2**

2 Dualità e Lemma di Farkas **5**

2.1 Problemi lineari duali e Teorema di Dualità 5

2.2 Il Lemma di Farkas 7

2.3 Dimostrazione del Teorema di Dualità dal Lemma di Farkas 10

1 Problemi lineari

> Linear programming, surprisingly, is not directly related to computer programming.
>
> *Jiri Matousek, Bernd Garter*

Sono problemi lineari tutti quei problimi in cui ci si prefigge di trovare il valore massimo (o minimo) che una certa funzione lineare di n variabili può assumere, dato un qualche numero di vincoli (anche essi lineari) su queste variabili. Prima di definire formalmente un problema lineare consideriamo un esempio.

Esempio 1.1.

$$\begin{array}{ll} \text{Massimizza} & x_1 + x_2 \\ \text{rispetto ai vincoli} & x_1, x_2 \geq 0 \\ & x_2 - x_1 \leq 1 \\ & x_1 + 6x_2 \leq 15 \\ & 4x_1 - x_2 \leq 10 \end{array}$$

In $\mathbb{R}^2$ ogni vincolo individua un semipiano. La zona di $\mathbb{R}^2$ su cui vogliamo massimizzare $x_1 + x_2$ è dunque l'intersezione di tutti questi semipiani ed è rappresentata in Figura 1.1. Osserviamo che quest'area non è vuota e che è un poligono convesso. Esiste dunque una coppia (x_1^*, x_2^*) che massimizza $x_1 + x_2$; la coppia in questione può essere ottenuta cercando quale punto del poligono si trova "più distante" nella direzione di massima crescita della funzione (data dal suo gradiente $(1, 1)$). Otteniamo così $x_1^* = 3, x_2^* = 2$ e infine che il valore massimo di $x_1 + x_2$ rispetto ai vincoli dati è 5.

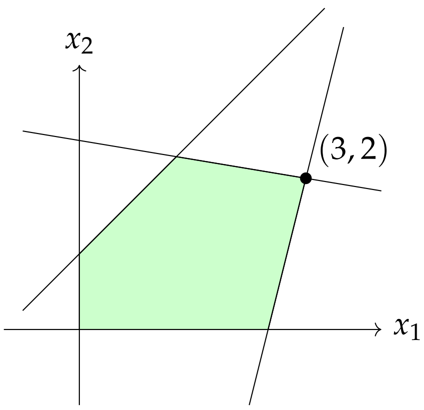

Figura 1.1

Definizione 1.2. Un **problema lineare** consiste in una funzione lineare di n variabili detta **funzione obiettivo** (o **funzione di costo**) e in un insieme di m vincoli lineari. La funzione obiettivo ha la forma $\mathbf{c}^T\mathbf{x} = c_1x_1 + \ldots + c_nx_n$ per qualche $\mathbf{c} \in \mathbb{R}^n$; lo stesso si applica ai vincoli. Dare un problema lineare è allora equivalente a dare un vettore $\mathbf{c} \in \mathbb{R}^n$, una matrice $A \in \mathbb{R}^{m\times n}$ e un vettore $\mathbf{b} \in \mathbb{R}^m$. Scriveremo compattamente

$$\begin{array}{ll} \text{Massimizza} & \mathbf{c}^T\mathbf{x} \\ \text{rispetto ai vincoli} & A\mathbf{x} \leq \mathbf{b}. \end{array}$$

Osservazione 1.3. La Definizione 1.2 è del tutto generale. Infatti un problema di minimizzazione può essere trasformato in uno di massimizzazione cambiando segno alla funzione obiettivo. I vincoli espressi tramite un'uguaglianza $\mathbf{a}^T\mathbf{x} = b$ sono equivalenti alla coppia di disuguaglianze $\mathbf{a}^T\mathbf{x} \geq b$, $\mathbf{a}^T\mathbf{x} \leq b$. Ed infine le disuguaglianze possono essere espresse tutte quante nella forma $\mathbf{a}^T\mathbf{x} \leq b$.

Definizione 1.4. Un vettore $\mathbf{x} \in \mathbb{R}^n$ che soddisfa tutti i vincoli di un problema lineare è una **soluzione possibile** per il problema. Un problema è **soddisfacibile** se ammette una soluzione possibile ed è **insoddisfacibile** altrimenti. Una soluzione possibile $\mathbf{x}^* \in \mathbb{R}^n$ è una **soluzione ottimale** se $\mathbf{c}^T\mathbf{x}^*$ è massimo tra i valori $\mathbf{c}^T\mathbf{x}$ con $\mathbf{x}$ soluzione possibile.

Osservazione 1.5. Va osservato che, generalmente, un sistema lineare può avere più di una soluzione ottimale; come esempio si considerino i vincoli dell'Esercizio 1.1 applicati però alla funzione obiettivo $\frac{1}{6}x_1 + x_2$. È inoltre vero che, anche se un sistema è soddisfacibile, possono non esistere soluzioni ottimali; come esempio basta rimuovere i vincoli $x_1 + 6x_2 \leq 15$ e $4x_1 - x_2 \leq 10$ dall'Esercizio 1.1.

Definizione 1.6. Se un problema lineare ammette (almeno) una soluzione ottimale allora è detto **limitato**; se non ne ammette viene detto **illimitato**.

Concludiamo la sezione con un esempio di applicazione dei problemi lineari all'analisi numerica (in particolare alla regressione lineare) e con un'osservazione sulla difficoltà computazionale della ricerca di soluzioni (possibili e ottimali) ad un dato problema lineare.

Esempio 1.7. Dato un insieme di punti $\{(x_i, y_i) : i = 1, \ldots, n\}$ di $\mathbb{R}^2$ è possibile ottenere l'equazione di una retta che si avvicina il più possibile ai punti dati usando il *metodo dei minimi quadrati*. L'intera tecnica è basata sul cercare $a, b \in \mathbb{R}$ tali che

$$\sum_{i=1}^{n} (ax_i + b - y_i)^2$$

sia minimo. Un metodo alternativo (e per certi lati migliore) consiste nel minimizzare direttamente la somma degli errori in valore assoluto:

$$\sum_{i=1}^{n} |ax_i + b - y_i|. \tag{$*$}$$

Seppure questa quantità non sia lineare il problema può essere ridotto ad un problema lineare. Si consideri infatti il problema

$$\begin{array}{ll} \text{Minimizza} & e_1 + \ldots + e_n \\ \text{rispetto ai vincoli} & e_i \geq ax_i + b - y_i \\ & e_i \geq -(ax_i + b - y_i) \quad , \text{con } i = 1, \ldots, n. \end{array}$$

Le variabili qui sono $a, b, e_1, \ldots, e_n$. Ogni e_i è una variabile ausiliaria che rappresenta l'errore relativo al punto (x_i, y_i); infatti i vincoli ci assicurano che

$$e_i \geq \max\big(ax_i + b - y_i, -(ax_i + b - y_i)\big) = |ax_i + b - y_i|.$$

Nel caso di una soluzione ottimale tutte queste disuguaglianze devono valere come uguaglianze, altrimenti sarebbe possibile ridurre ulteriormente il corrispettivo e_i. Ne consegue che una soluzione ottimale del problema fornisce una retta che minimizza $(*)$.

Osservazione 1.8. Consideriamo un problema lineare generico

$$\begin{array}{ll} \text{Massimizza} & \mathbf{c}^T\mathbf{x} \\ \text{rispetto ai vincoli} & A\mathbf{x} \le \mathbf{b}. \end{array}$$

e supponiamo di sapere che $0 \le \mathbf{c}^T\mathbf{x} \le M$ per un $M \in \mathbb{R}$ e ogni soluzione possibile $\mathbf{x}$. Supponiamo inoltre di avere una procedura che ci permetta di sapere quando un arbitrario problema lineare è soddisfacibile. Allora possiamo approssimare il valore massimo di $\mathbf{c}^T\mathbf{x}$ con una tecnica di ricerca binaria:

(1) aggiungiamo al problema iniziale il vincolo $\mathbf{c}^T\mathbf{x} \ge \frac{M}{2}$ e determiniamo se questo nuovo problema è soddisfacibile,

(2) se lo è allora il massimo di $\mathbf{c}^T\mathbf{x}$ si trova tra $\frac{M}{2}$ e M, altrimenti si trova tra 0 e $\frac{M}{2}$,

(3) ripetendo i passi (1) e (2) su questo nuovo intervallo possiamo molto velocemente localizzare il massimo.

Questo ci dice che, computazionalmente, il problema di ricerca di una soluzione ottimale è tanto difficile quanto quello di ricerca di una soluzione qualunque.

2 Dualità e Lemma di Farkas

> The Farkas Lemma is not exactly a difficult theorem, but it is not trivial either.
>
> *Jiri Matousek, Bernd Garter*

2.1 Problemi lineari duali e Teorema di Dualità

Osservazione 2.1. Consideriamo un problema lineare

$$\begin{array}{ll} \text{Massimizza} & \mathbf{c}^T\mathbf{x} \\ \text{rispetto ai vincoli} & A\mathbf{x} \leq \mathbf{b}. \end{array} \tag{P}$$

Per ogni x_i si considerino due variabili y_i, y'_i e si sostituisca ogni occorrenza di x_i nel problema con $y_i - y'_i$. Aggiungendo i vincoli $\mathbf{y}, \mathbf{y}' \geq 0$ si ottiene un secondo problema lineare

$$\begin{array}{ll} \text{Massimizza} & \mathbf{c}'^T\mathbf{z} \\ \text{rispetto ai vincoli} & A'\mathbf{z} \leq \mathbf{b}' \\ & \mathbf{z} \geq 0 \end{array} \tag{P^*}$$

nelle variabili $\mathbf{y}, \mathbf{y}'$ (qui rappresentate come un unico gruppo $\mathbf{z}$ per semplicità) che è del tutto equivalente al sistema di partenza. Infatti una soluzione del primo fornisce una soluzione del secondo perché ogni numero reale può essere scritto come differenza di due reali positivi; mentre ogni soluzione del secondo fornisce una soluzione del primo ottenuta come $x_i = y_i - y'_i$. È dunque sempre possibile supporre che un problema lineare sia dato nella forma (P^*).

Osservazione 2.2. Consideriamo ora un generico problema lineare con m vincoli e n variabili

$$\begin{array}{ll} \text{Massimizza} & \mathbf{c}^T\mathbf{x} \\ \text{rispetto ai vincoli} & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0. \end{array}$$

Combinando linearmente gli m vincoli con altrettanti coefficienti *non negativi* $y_1, \ldots, y_m$ si ottiene una nuova disuguaglianza

$$d_1x_1 + \ldots + d_nx_n \leq y_1b_1 + \ldots + y_mb_m$$

dove $d_i = y_1a_{1,i} + \ldots + y_ma_{m,i}$. Se inoltre gli y_i vengono scelti in modo tale che $c_j \leq d_j$ per ogni $j = 1, \ldots, n$ si ottiene il seguente limite superiore alla funzione obiettivo $\mathbf{c}^T\mathbf{x}$

$$c_1x_1 + \ldots + c_nx_n \leq d_1x_1 + \ldots + d_nx_n \leq y_1b_1 + \ldots + y_mb_m.$$

Siamo ovviamente interessati a *minimizzare* questo limite superiore il che ci porta alla prossima definizione.

Definizione 2.3. Dato un problema lineare

$$\begin{array}{ll} \text{Massimizza} & \mathbf{c}^T\mathbf{x} \\ \text{rispetto ai vincoli} & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array} \tag{P}$$

definiamo il suo **problema lineare duale**

$$\begin{array}{ll} \text{Minimizza} & \mathbf{b}^T\mathbf{y} \\ \text{rispetto ai vincoli} & A^T\mathbf{y} \geq \mathbf{c} \\ & \mathbf{y} \geq 0. \end{array} \tag{P^{op}}$$

Osservazione 2.4. Riscrivendo (P^{op}) nella forma descritta nell'Osservazione 2.1 e computandone il duale si verifica che il problema "biduale" $(P^{\mathrm{op}})^{\mathrm{op}}$ non è altro che il problema originale (P). Possiamo dunque dire che (P) e (P^{op}) sono duali l'uno rispetto all'altro.

Proposizione 2.5 (Teorema di Dualità debole). Consideriamo ancora il problema lineare (P). Per ogni soluzione possibile $\mathbf{x}$ di (P) e $\mathbf{y}$ di (P^{op}) si ha che

$$\mathbf{c}^T\mathbf{x} \leq \mathbf{b}^T\mathbf{y}.$$

Questo ci dice che

(1) se (P) è illimitato allora (P^{op}) è insoddisfacibile,

(2) se (P^{op}) è illimitato [1] allora (P) è insoddisfacibile.

Dimostrazione. Ovvia dalla discussione fatta nell'Osservazione 2.2 e dalla Definizione 2.3 di problema duale. □

Esempio 2.6. Consideriamo il problema lineare

$$\begin{array}{ll} \text{Massimizza} & 2x_1 + 3x_2 \\ \text{rispetto ai vincoli} & 4x_1 + 8x_2 \leq 12 \\ & 2x_1 + x_2 \leq 3 \\ & 3x_1 + 2x_2 \leq 4 \\ & x_1, x_2 \geq 0 \end{array}$$

e il suo duale

$$\begin{array}{ll} \text{Minimizza} & 12y_1 + 3y_2 + 4y_3 \\ \text{rispetto ai vincoli} & 4y_1 + 2y_2 + 3y_3 \geq 2 \\ & 8y_1 + y_2 + 2y_3 \geq 3 \\ & y_1, y_2, y_3 \geq 0. \end{array}$$

Si può vedere che il problema duale è limitato ed ha soluzione ottimale 4.75. Anche il problema originale è limitato e, sorprendentemente, ha come soluzione ottimale anch'esso 4.75. Il Teorema di Dualità, che ora enunceremo, mostra che questa situazione particolare è in realtà necessaria.

Teorema 2.7 (Teorema di Dualità). Dato un problema lineare

$$\begin{array}{ll} \text{Massimizza} & \mathbf{c}^T\mathbf{x} \\ \text{rispetto ai vincoli} & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array} \tag{P}$$

e il suo duale

$$\begin{array}{ll} \text{Minimizza} & \mathbf{b}^T\mathbf{y} \\ \text{rispetto ai vincoli} & A^T\mathbf{y} \geq \mathbf{c} \\ & \mathbf{y} \geq 0 \end{array} \tag{P^{op}}$$

una e solo una delle seguenti affermazioni è vera:

[1] Inferiormente, trattandosi di un problema di minimizzazione.

(1) (P) e (P^{op}) sono insoddisfacibili,

(2) (P) è illimitato e (P^{op}) è insoddisfacibile,

(3) (P^{op}) è illimitato e (P) è insoddisfacibile,

(4) (P) e (P^{op}) sono entrambi soddisfacibili e limitati. Inoltre se $\mathbf{x}^*$ e $\mathbf{y}^*$ sono due soluzioni ottimali (rispettivamente di (P) e di (P^{op})) allora

$$\mathbf{c}^T\mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*.$$

Esistono diverse dimostrazioni di questo risultato. Quella che presenteremo noi (si veda §2.3) utilizza il Lemma di Farkas, che analizzeremo nella prossima sezione.

2.2 Il Lemma di Farkas

Lemma 2.8 (Lemma di Farkas). Data una matrice $A \in \mathbb{R}^{m\times n}$ e un vettore $\mathbf{b} \in \mathbb{R}^m$ allora una e solo una delle seguenti affermazioni è vera:

(1) esiste un vettore $\mathbf{x} \in \mathbb{R}^n$ tale che $A\mathbf{x} = \mathbf{b}$ e $\mathbf{x} \geq 0$,

(2) esiste un vettore $\mathbf{y} \in \mathbb{R}^m$ tale che $\mathbf{y}^T A \geq \mathbf{0}^T$ e $\mathbf{y}^T\mathbf{b} < 0$.

Il Lemma di Farkas può essere interpretato geometricamente come un teorema di separazione per insiemi convessi. Per farlo introduciamo la nozione di cono convesso generato da un insieme di vettori.

Definizione 2.9. Dati vettori $\mathbf{a}_1, \dots, \mathbf{a_n} \in \mathbb{R}^m$ il **cono convesso** da loro generato è l'insieme

$$\left\{\lambda_1\mathbf{a}_1 + \dots + \lambda_n\mathbf{a}_n\colon\ \lambda_1, \dots, \lambda_n \geq 0\right\}$$

delle loro combinazioni lineari a coefficienti non negativi.

Da questa definizione si ottiene facilmente che il Lemma di Farkas è equivalente alla seguente affermazione geometrica (si veda anche la Figura 2.1).

Lemma 2.10 (Lemma di Farkas (versione geometrica)). Se $\mathbf{a}_1, \dots, \mathbf{a_n} \in \mathbb{R}^m$ sono vettori, C è il loro cono convesso e $\mathbf{b} \in \mathbb{R}^m$ è un ulteriore vettore allora una e solo una delle seguenti affermazioni è vera:

(1) $\mathbf{b}$ appartiene a C,

(2) esiste un iperpiano $H = \{\mathbf{x} \in \mathbb{R}^M\colon\ \mathbf{y}^T\mathbf{x} = 0\}$ di $\mathbb{R}^m$ che "separa" gli $\mathbf{a}_i$ da $\mathbf{b}$; cioè tale che $\mathbf{y}^T\mathbf{a}_i \leq 0$ per ogni $i = 1, \dots, n$ e $\mathbf{y}^T\mathbf{b} > 0$.

La seguente proposizione verrà utilizzata nella dimostrazione del Teorema di Dualità.

Proposizione 2.11. Data una matrice $A \in \mathbb{R}^{m\times n}$ e un vettore $\mathbf{b} \in \mathbb{R}^m$ il sistema $A\mathbf{x} \leq \mathbf{b}$ ammette una soluzione *non negativa* se e solo se ogni $\mathbf{y} \in \mathbb{R}^m$ *non negativo* che soddisfa $\mathbf{y}^T A \geq \mathbf{0}^T$ soddisfa anche $\mathbf{y}^T\mathbf{b} \geq 0$.

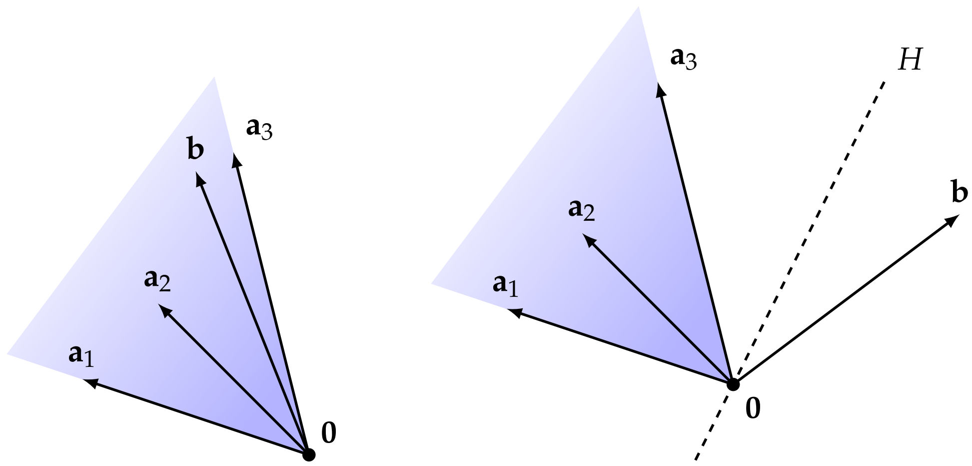

Figura 2.1

Dimostrazione. Il Lemma di Farkas 2.8 ci dice che un sistema $A\mathbf{x} = \mathbf{b}$ ha una soluzione non negativa se e solo se per ogni $\mathbf{y}$ tale che $\mathbf{y}^T A \geq \mathbf{0}^T$ si ha anche $\mathbf{y}^T\mathbf{b} \geq 0$. Per poter applicare questo risultato costruiamo, data una matrice $A \in \mathbb{R}^{m\times n}$, $\overline{A} = (A|I_m)$ dove $I_m \in \mathbb{R}^{m\times m}$ è la matrice identità. Si ha ora che $A\mathbf{x} \leq \mathbf{b}$ ammette una soluzione non negativa se e solo se $\overline{A}\overline{\mathbf{x}} = \mathbf{b}$ ha una soluzione non negativa.

Per quanto detto prima questo è equivalente a chiedere che ogni $\mathbf{y} \in \mathbb{R}^m$ che soddisfa $\mathbf{y}^T\overline{A} \geq \mathbf{0}^T$ soddisfi anche $\mathbf{y}^T\mathbf{b} \geq 0$. Ma richiedere che $\mathbf{y}^T\overline{A} \geq \mathbf{0}^T$ è equivalente a richiedere $\mathbf{y}^T A \geq \mathbf{0}^T$ e $\mathbf{y} \geq \mathbf{0}$ (i.e. la non negatività di $\mathbf{y}$); dunque abbiamo dimostrato l'equivalenza richiesta. $\square$

Procederemo ora a dimostrare il Lemma di Farkas nella sua versione geometrica. Occorre prima dimostrare il seguente ulteriore lemma.

Lemma 2.12. Se C è un cono convesso generato dai vettori $\mathbf{a}_1, \ldots, \mathbf{a}_n$ e $\mathbf{b}$ è un punto non appartenente a C allora esiste un punto $\mathbf{z} \in C$ tale che $\|\mathbf{b} - \mathbf{z}\| \leq \|\mathbf{b} - \mathbf{z}'\|$ per ogni $\mathbf{z}' \in C$. Si veda anche la Figura 2.2.

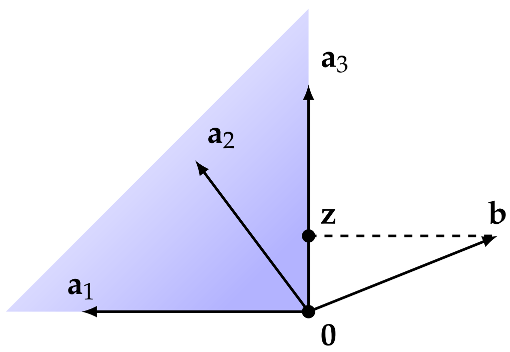

Figura 2.2

Per dimostrare questo lemma sono necessari un altro paio di risultati.

Proposizione 2.13. Sia $X \subseteq \mathbb{R}^m$ un chiuso non vuoto e $\mathbf{b}$ un punto arbitrario di $\mathbb{R}^m$. Allora esiste un punto $\mathbf{z} \in X$ tale che $\|\mathbf{b} - \mathbf{z}\| \leq \|\mathbf{b} - \mathbf{z}'\|$ per ogni $\mathbf{z}' \in X$.

Dimostrazione. Fissiamo un punto $\mathbf{x}_0 \in X$ e consideriamo l'insieme

$$K = \{\mathbf{x} \in X\colon\ \|\mathbf{b} - \mathbf{x}\| \leq \|\mathbf{b} - \mathbf{x}_0\|\}.$$

Chiaramente se il risultato vale per K vale anche per X. Inoltre K è l'intersezione di X (chiuso) con la palla chiusa di centro $\mathbf{b}$ e raggio $\|\mathbf{b} - \mathbf{x}_0\|$ dunque è un insieme chiuso e limitato e quindi compatto. Consideriamo ora la funzione

$$\begin{array}{rccc} f\colon & K & \to & \mathbb{R} \\ & \mathbf{x} & \mapsto & \|\mathbf{b} - \mathbf{x}\| \end{array}$$

che sappiamo essere continua. Siccome K è compatto f ha un minimo; esiste cioè un punto $\mathbf{z} \in K$ tale che $\|\mathbf{b} - \mathbf{z}\| \leq \|\mathbf{b} - \mathbf{z}'\|$ per ogni $\mathbf{z}' \in K$. □

Definizione 2.14. Un **cono primitivo** di $\mathbb{R}^m$ è un cono convesso generato da un insieme *linearmente indipendente* di vettori.

Proposizione 2.15. Ogni cono primitivo P di $\mathbb{R}^m$ è chiuso.

Dimostrazione. Supponiamo che $P \subseteq \mathbb{R}^m$ sia generato dai vettori (linearmente indipendenti) $\mathbf{a}_1, \ldots, \mathbf{a}_k$ e sia P_0 il cono chiuso di $\mathbb{R}^k$ generato dai vettori $\mathbf{e}_1, \ldots, \mathbf{e}_k$ della base canonica; P_0 è chiaramente chiuso. Consideriamo ora la mappa lineare

$$\begin{array}{rccc} f\colon & \mathbb{R}^k & \to & \mathbb{R}^m \\ & \mathbf{x} & \mapsto & x_1\mathbf{a}_1 + \ldots + x_k\mathbf{a}_k \end{array}$$

che è iniettiva, per l'indipendenta degli $\mathbf{a}_i$, e tale che $P = f(P_0)$. Per iniettività f è un isomorfismo lineare tra $\mathbb{R}^k$ e $\mathrm{im}(f)$; possiamo dunque considerare la mappa lineare inversa f^{-1}. Siccome le mappe lineari sono continue si ha che P, in quanto controimmagine di P_0 mediante f^{-1}, è chiuso in $\mathrm{im}(f)$. Ma $\mathrm{im}(f)$ è un sottospazio lineare di $\mathbb{R}^m$, dunque è chiuso in $\mathbb{R}^m$ e così anche P lo è[2]. □

Proposizione 2.16. Sia C un cono convesso di $\mathbb{R}^m$ generato dai vettori $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Allora C si può esprimere come unione di un numero finito di coni primitivi.

Dimostrazione. Per ogni $\mathbf{x} \in C$ verifichiamo che esiste un cono principale generato da un sottoinsieme linearmente indipendente dei vettori $\mathbf{a}_i$ che contiene $\mathbf{x}$.

Sia $I \subseteq \{1, \ldots, n\}$ l'insieme più piccolo tale che $\mathbf{x}$ appartenga al cono generato da $A_I = \{\mathbf{a}_i\colon\ i \in I\}$. Abbiamo dunque che esistono coefficienti non negativi α_i tali che $\mathbf{x} = \sum_{i \in I} \alpha_i \mathbf{a}_i$. Ma possiamo addirittura richiedere che gli α_i siano strettamente positivi; infatti se $\alpha_j = 0$ si ha che $\mathbf{x}$ è ancora contenuto nel cono generato da $A_{I \setminus \{j\}}$. È ora sufficiente mostrare che A_I è un insieme linearmente indipendente di vettori.

Per assurdo supponiamo che esistano coefficienti β_i tali che $\sum_{i \in I} \beta_i \mathbf{a}_i = \mathbf{0}$ con almeno un $\beta_j \neq 0$. Allora esiste un $t \in \mathbb{R}$ tale che tutte le espressioni $\alpha_i - t\beta_i$ sono non negative e almeno una di queste è nulla (se esistono dei β_j strettamente positivi basta porre t uguale al minimo del valore $\frac{\alpha_j}{\beta_j}$ computato su questi β_j; se non esistono β_j strettamente positivi si può prima sostituire ogni β_i con $-\beta_i$)). Si ottiene quindi

$$\mathbf{x} = \mathbf{x} - t \cdot \mathbf{0} = \sum_{i \in I} (\alpha_i - t\beta_i)\mathbf{a}_i$$

che esprime $\mathbf{x}$ come combinazione lineare a coefficienti non negativi di meno di $|I|$ vettori (perché una delle $\alpha_i - t\beta_i$ è nulla) in contraddizione con la definizione di I. □

[2] Un sottoinsieme di un sottospazio topologico chiuso è chiuso se e solo se è chiuso nello spazio ambiente.

Proposizione 2.17. Ogni cono convesso generato da un insieme finito di vettori $\mathbf{a}_1, \dots, \mathbf{a}_n$ è chiuso.

Dimostrazione. Ovvia dalle Proposizioni 2.15, 2.16 e dal fatto che l'unione finita di chiusi è un chiuso. □

Osservazione 2.18. La Proposizione 2.17 non si applica ai coni generati da un numero infinito di vettori. Si consideri infatti un disco chiuso D il cui bordo contenga l'origine $\mathbf{0}$. Il cono generato da tutti i vettori che compongono D è un semipiano aperto unito al singoletto $\{\mathbf{0}\}$ e questo non è un insieme chiuso.

Dimostrazione del Lemma 2.12. Immediata dalle Proposizioni 2.17 e 2.13. □

Dimostrazione del Lemma di Farkas (2.10). Se $\mathbf{b} \in C$ non c'è nulla da dimostrare, dunque poniamoci nel caso $\mathbf{b} \notin C$. Sia quindi $\mathbf{z}$ il punto di C più vicino a $\mathbf{b}$ che esiste in virtù del Lemma 2.12. Definiamo $\mathbf{y} = \mathbf{z} - \mathbf{b}$ e controlliamo che l'iperpiano $\{\mathbf{x} \in \mathbb{R}^m : \ \mathbf{y}^T\mathbf{x} = 0\}$ "separa" gli $\mathbf{a}_i$ da $\mathbf{b}$.

Proviamo innanzitutto che $\mathbf{y}^T\mathbf{z} = 0$ i.e. che $\mathbf{y}$ e $\mathbf{z}$ sono perpendicolari. Se $\mathbf{z} = 0$ la cosa è ovvia; possiamo dunque assumere che $\mathbf{z} \neq 0$. Supponiamo per assurdo che $\mathbf{z}$ non sia perpendicolare a $\mathbf{y}$. Allora è possibile trovare un punto $\mathbf{z}'$ sulla semiretta $\{t\mathbf{z} : \ t \geq 0\}$ tale che $\|\mathbf{b} - \mathbf{z}'\| < \|\mathbf{b} - \mathbf{z}\|$; ma $\mathbf{z}$ è il punto di C più vicino a $\mathbf{b}$, dunque abbiamo una contraddizione.

Ora siccome $\mathbf{b} \notin C$ si ha $\mathbf{y} \neq \mathbf{0}$ e dunque

$$0 < \mathbf{y}^T\mathbf{y} = \mathbf{y}^T(\mathbf{z} - \mathbf{b}) = -\mathbf{y}^T\mathbf{b}.$$

Questo ci dice che $\mathbf{y}^T\mathbf{b} < 0$.

Infine consideriamo un generico $\mathbf{x} \in C, \mathbf{x} \neq \mathbf{z}$ e dimostriamo che $\mathbf{y}^T\mathbf{x} \geq 0$; una volta fatto questo, siccome gli $\mathbf{a}_i$ sono tutti elementi di C, la dimostrazione sarà completata. Si deve avere $(\mathbf{b} - \mathbf{z})^T(\mathbf{x} - \mathbf{z}) \leq 0$ perché altrimenti l'angolo $\sphericalangle\mathbf{bzx}$ sarebbe minore di $90°$ e allora ci sarebbe un punto $\mathbf{z}'$ del segmento $\mathbf{zx}$ tale che $\|\mathbf{b} - \mathbf{z}'\| < \|\mathbf{b} - \mathbf{z}\|$ contraddicendo ancora il fatto che $\mathbf{z}$ sia il punto di C più vicino a $\mathbf{b}$. Possiamo ora scrivere

$$0 \geq (\mathbf{b} - \mathbf{z})^T(\mathbf{x} - \mathbf{z}) = -\mathbf{y}^T\mathbf{x} + \mathbf{y}^T\mathbf{z} = -\mathbf{y}^T\mathbf{x}$$

e dunque $\mathbf{y}^T\mathbf{x} \geq 0$. □

2.3 Dimostrazione del Teorema di Dualità dal Lemma di Farkas

Il seguente teorema fornisce condizioni sufficienti per la limitatezza di un problema lineare.

Teorema 2.19. Si consideri il seguente problema lineare

$$\begin{array}{ll} \text{Minimizza} & \mathbf{c}^T\mathbf{x} \\ \text{rispetto ai vincoli} & A\mathbf{x} \geq \mathbf{b} \\ & \mathbf{x} \geq 0. \end{array}$$

Se il problema ha una soluzione possibile e la funzione obiettivo $\mathbf{c}^T\mathbf{x}$ è limitata *inferiormente* sull'insieme di tutte le soluzioni possibili allora il problema ammette una soluzione ottimale.

Dimostrazione. Una dimostrazione di questo teorema (in una forma leggermente diversa, ma equivalente) si può trovare in [1, Theorem 4.2.3]. □

Dimostrazione del Teorema di Dualità 2.7. Come visto in §1 un problema lineare può essere insoddisfacibile, soddisfacibile limitato o soffisfacibile illimitato. Abbiamo dunque 3 possibilità per (P) e altrettante per (P^{op}) per un totale di 9 possibili combinazioni. La Proposizione 2.5 ci permette subito di escludere i seguenti casi

- (P) illimitato e (P^{op}) limitato,
- (P^{op}) illimitato e (P) limitato.
- (P) illimitato e (P^{op}) illimitato.

Mostreremo ora che se (P) è limitato allora anche (P^{op}) lo è e dunque, per dualità, che se (P^{op}) è limitato allora anche (P) lo è. Questo esclude i casi

- (P) limitato e (P^{op}) insoddisfacibile,
- (P^{op}) limitato e (P) insoddisfacibile.

Una volta fatto questo i casi rimasti sono solo quelli elencati nell'enunciato del Lemma di Farkas 2.8 e la dimostrazione è completa.

Supponiamo quindi che (P) sia limitato e siano: $\mathbf{x}^*$ una soluzione ottimale, $\gamma = \mathbf{c}^T\mathbf{x}^*$ il valore di massimo e $\varepsilon \geq 0$. Definiamo

$$\overline{A} = \begin{pmatrix} A \\ -\mathbf{c}^T \end{pmatrix}, \quad \overline{\mathbf{b}_\varepsilon} = \begin{pmatrix} \mathbf{b} \\ -\gamma - \varepsilon \end{pmatrix}$$

e osserviamo che il sistema di disequazioni $\overline{A}\mathbf{x} \leq \overline{\mathbf{b}_\varepsilon}$ è equivalente a

$$A\mathbf{x} \leq \mathbf{b}, \mathbf{c}^T\mathbf{x} \geq \gamma + \varepsilon$$

e dunque ammette una soluzione non negativa se e solo se $\varepsilon = 0$.

Siccome $\overline{A}\mathbf{x} \leq \overline{\mathbf{b}_\varepsilon}$ non ha soluzioni non negative per $\varepsilon > 0$ la Proposizione 2.11 ci dice che esiste un vettore $\overline{\mathbf{y}} = (\mathbf{u}, z) \in \mathbb{R}^{m+1}$ non negativo e tale che $\overline{\mathbf{y}}^T\overline{A} \geq \mathbf{0}^T$ ma $\overline{\mathbf{y}}^T\overline{\mathbf{b}_\varepsilon} < 0$. Queste disuguaglianze possono essere riscritte, più chiaramente, come

$$A^T\mathbf{u} \geq z\mathbf{c}, \ \mathbf{b}^T\mathbf{u} < z(\gamma + \varepsilon). \tag{$*$}$$

Applicando la Proposizione 2.11 al caso $\varepsilon = 0$ si ottiene che il suddetto vettore $\overline{\mathbf{y}}$ deve soddisfare $\overline{\mathbf{y}}^T\overline{\mathbf{b}_0} \geq 0$ e cioè che $\mathbf{b}^T\mathbf{u} \geq z\gamma$. Da questa ultima relazione e da $(*)$ segue che bisogna avere $z > 0$ ed è quindi possibile definire il vettore $\mathbf{v} = \frac{1}{z}\mathbf{u}$ che è chiaramente non negativo e tale che

$$A^T\mathbf{v} \geq \mathbf{c}, \ \mathbf{b}^T\mathbf{v} < \gamma + \varepsilon.$$

Abbiamo dunque ottenuto una soluzione di (P^{op}) con valore della funzione obiettivo minore di $\gamma + \varepsilon$.

Per la Proposizione 2.5 ogni soluzione di (P^{op}) deve avere valore della funzione obiettivo maggiore o uguale a γ. Si ha quindi che (P^{op}) è un problema lineare soddisfacibile e con funzione obiettivo limitata inferiormente sull'insieme delle soluzioni possibili, dunque ammette una soluzione ottimale $\mathbf{y}^*$ (Teorema 2.19). Inoltre si ha che $\gamma \leq \mathbf{b}^T\mathbf{y}^* < \gamma + \varepsilon$ per ogni $\varepsilon > 0$ e quindi si deve avere $\gamma = \mathbf{c}^T\mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$. □

Riferimenti bibliografici

[1] Jiri Matousek & Bernd Gartner, *Understanding and Using Linear Programming*, Springer, 2007.