

# Il Lemma di Farkas

Gabriele Rastello

20 aprile 2020

## Indice

<b>1</b>	<b>Problemi lineari</b>	<b>1</b>
<b>2</b>	<b>Dualità e Lemma di Farkas</b>	<b>4</b>
2.1	Problemi lineari duali e Teorema di Dualità . . . . .	4
2.2	Il Lemma di Farkas . . . . .	6
2.3	Dimostrazione del Teorema di Dualità dal Lemma di Farkas . . . . .	8

## 1 Problemi lineari

Linear programming, surprisingly, is not directly related to computer programming.

---

*Jiri Matousek, Bernd Gartner*

Sono problemi lineari tutti quei problemi 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. Osserviamo che quest'area non è vuota e che è un poligono convesso. Esiste dunque una

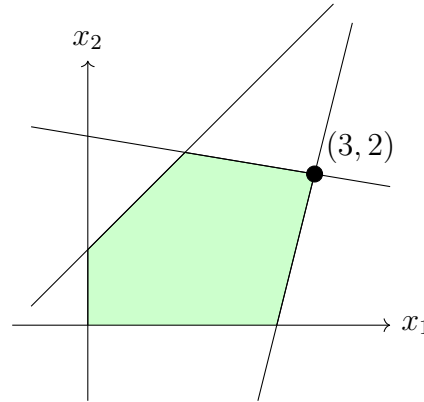


Figura 1

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.

**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  $\mathbf{c}^T \mathbf{x} = c_1 x_1 + \dots + c_n x_n$  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  $\{(x_i, y_i) : i = 1, \dots, n\}$  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:

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

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

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

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

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

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

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

e supponiamo di sapere che  $0 \leq \mathbf{c}^T \mathbf{x} \leq M$  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  $\mathbf{c}^T \mathbf{x} \geq \frac{M}{2}$  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

### 2.1 Problemi lineari duali e Teorema di Dualità

**Osservazione 2.1.** Consideriamo un problema lineare

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

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

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

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*  $y_1, \dots, y_m$  si ottiene una nuova disuguaglianza

$$d_1 x_1 + \dots + d_n x_n \leq y_1 b_1 + \dots + y_m b_m$$

dove  $d_i = y_1 a_{1,i} + \dots + y_m a_{m,i}$ . Se inoltre gli  $y_i$  vengono scelti in modo tale che  $c_j \leq d_j$  per ogni  $j = 1, \dots, n$  si ottiene il seguente limite superiore alla funzione obiettivo  $\mathbf{c}^T \mathbf{x}$

$$c_1 x_1 + \dots + c_n x_n \leq d_1 x_1 + \dots + d_n x_n \leq y_1 b_1 + \dots + y_m b_m.$$

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

**Definizione 2.3.** Dato un problema lineare

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

definiamo il suo **problema lineare duale**

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

**Osservazione 2.4.** Riscrivendo  $(P^{\text{op}})$  nella forma descritta nell'Osservazione 2.1 e computandone il duale si verifica che il problema “biduale”  $(P^{\text{op}})^{\text{op}}$  non è altro che il problema originale  $(P)$ . Possiamo dunque dire che  $(P)$  e  $(P^{\text{op}})$  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  $(P^{\text{op}})$  si ha che

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

Questo ci dice che

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

(2) se  $(P^{\text{op}})$  è illimitato <sup>1</sup> allora  $(P)$  è insoddisfacibile.

*Dimostrazione.* Ovvio dalla discussione fatta nell'Osservazione 2.2 e dalla Definizione 2.3 di problema duale.  $\square$

**Esempio 2.6.** Consideriamo il problema lineare

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

---

<sup>1</sup>Inferiormente, trattandosi di un problema di minimizzazione.

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 enuncieremo, mostra che questa situazione particolare è in realtà necessaria.

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

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

e il suo duale

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

una e solo una delle seguenti affermazioni è vera:

- (1)  $(P)$  e  $(P^{\text{op}})$  sono insoddisfacibili,
- (2)  $(P)$  è illimitato e  $(P^{\text{op}})$  è insoddisfacibile,
- (3)  $(P^{\text{op}})$  è illimitato e  $(P)$  è insoddisfacibile,
- (4)  $(P)$  e  $(P^{\text{op}})$  sono entrambi soddisfacibili e limitati. Inoltre se  $\mathbf{x}^*$  e  $\mathbf{y}^*$  sono due soluzioni ottimali (rispettivamente di  $(P)$  e di  $(P^{\text{op}})$ ) 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  $\mathbf{y}^T \mathbf{b} < 0$ .

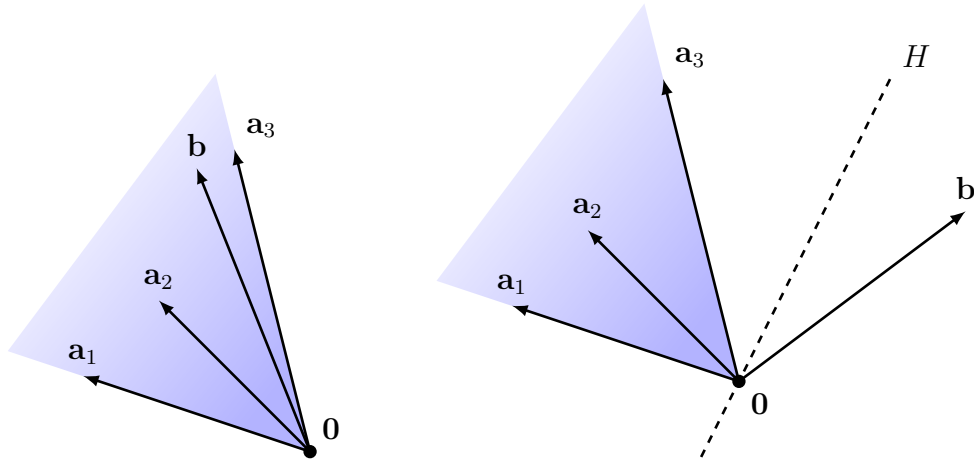


Figura 2

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

$$\{\lambda_1 \mathbf{a}_1 + \dots + \lambda_n \mathbf{a}_n : \lambda_1, \dots, \lambda_n \geq 0\}$$

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).

**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  $H = \{\mathbf{x} \in \mathbb{R}^m : \mathbf{y}^T \mathbf{x} = 0\}$  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  $\mathbf{y}^T \mathbf{b} \geq 0$ .

*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  $A \in \mathbb{R}^{m \times n}$ ,  $\bar{A} = (A|I_m)$  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  $\bar{A}\bar{\mathbf{x}} = \mathbf{b}$  ha una soluzione non negativa.

Per quanto detto prima questo è equivalente a chiedere che ogni  $\mathbf{y} \in \mathbb{R}^m$  che soddisfa  $\mathbf{y}^T \bar{A} \geq \mathbf{0}^T$  soddisfi anche  $\mathbf{y}^T \mathbf{b} \geq 0$ . Ma richiedere che  $\mathbf{y}^T \bar{A} \geq \mathbf{0}^T$  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.  $\square$

## 2.3 Dimostrazione del Teorema di Dualità dal Lemma di Farkas

### Riferimenti bibliografici

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