

Il teorema di dualità forte

Ricordiamo la formulazione del problema di programmazione lineare nella sua forma primale e duale,

$$\begin{array}{ll} \max & \mathbf{c}^\top \mathbf{x} \\ & A \mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{array} \quad \begin{array}{ll} \min & \mathbf{b}^\top \mathbf{y} \\ & A^\top \mathbf{y} \geq \mathbf{c} \\ & \mathbf{y} \geq \mathbf{0} \end{array} \quad (1)$$

dove A è una matrice reale $m \times n$, $\mathbf{b} \in \mathbb{R}^m$ e $\mathbf{c} \in \mathbb{R}^n$.

Formuliamo ora delle condizioni sufficienti per l'ottimalità note come condizioni di complementarità. Data una coppia di soluzioni ammissibili \mathbf{x}, \mathbf{y} di (1) possiamo scrivere

$$\sum_{j=1}^n c_j x_j \leq \sum_{j=1}^n \left(\sum_{i=1}^m A_{i,j} y_i \right) x_j \quad (2)$$

$$= \sum_{i=1}^m \left(\sum_{j=1}^n A_{i,j} x_j \right) y_i \quad (3)$$

$$\leq \sum_{i=1}^m b_i y_i \quad (4)$$

Corollario 1 (Condizioni di complementarità) *Sia (P) un programma lineare con duale (D) , dove (P) e (D) hanno forma (1). Se \mathbf{x} e \mathbf{y} sono soluzioni ammissibili di (P) e (D) e inoltre valgono le condizioni*

1. ogni volta che $x_j \neq 0$, \mathbf{y} soddisfa il j -esimo vincolo con uguaglianza,
2. ogni volta che $y_i \neq 0$, \mathbf{x} soddisfa l' i -esimo vincolo con uguaglianza,

allora \mathbf{x} e \mathbf{y} sono entrambe ottime.

DIMOSTRAZIONE. La condizione 1 può essere riscritta in forma compatta come

$$c_j x_j = \left(\sum_{i=1}^m A_{i,j} y_i \right) x_j$$

e implica che la (2) vale con l'eguaglianza. Analogamente, la condizione 2 può essere riscritta come

$$\left(\sum_{j=1}^n A_{i,j} x_j \right) y_i = b_i y_i$$

e implica che pure la (4) vale con l'eguaglianza. Quindi concludiamo che $\mathbf{c}^\top \mathbf{x} = \mathbf{b}^\top \mathbf{y}$. Ma dato che $\mathbf{c}^\top \mathbf{x} \leq \mathbf{b}^\top \mathbf{y}$ vale per ogni coppia di soluzioni ammissibili, le due soluzioni devono essere entrambe

ottime. □

Le condizioni di complementarità stabiliscono che in una coppia di soluzioni ottime primale e duale non è possibile che una variabile primale sia non nulla e, simultaneamente, il vincolo duale ad essa associato sia soddisfatto in modo lasco (cioè con $<$ invece che con $=$). Possiamo usare le condizioni di complementarità operativamente sviluppando algoritmi che cercano di soddisfare simultaneamente le seguenti tre proprietà:

1. \mathbf{x} è ammissibile per (P),
2. \mathbf{y} è ammissibile per (D),
3. \mathbf{x} e \mathbf{y} soddisfano le condizioni di complementarità.

Le condizioni di complementarità stabiliscono che soluzioni primale e duale che soddisfano particolari condizioni devono essere ottime e, in particolare, avere valore uguale. Enunciamo ora il teorema di dualità forte, che come conseguenza ha il fatto che le condizioni di complementarità sono anche necessarie, oltre che sufficienti. Per semplicità, useremo una forma particolare di LP ricordando però che ogni LP può essere riscritto nella forma (1).

Teorema 2 (Dualità forte) *Per ogni programma lineare (P) con duale (D) della forma*

$$\begin{array}{ll} \min & \mathbf{c}^\top \mathbf{x} \\ & A\mathbf{x} \geq \mathbf{b} \\ & \mathbf{x} \in \mathbb{R}^n \end{array} \quad \begin{array}{ll} \max & \mathbf{b}^\top \mathbf{y} \\ & A^\top \mathbf{y} = \mathbf{c} \\ & \mathbf{y} \geq \mathbf{0} \end{array} \quad (5)$$

Se (P) e (D) sono entrambi limitati con soluzioni ottime \mathbf{x}^ e \mathbf{y}^* , allora $\mathbf{c}^\top \mathbf{x}^* = \mathbf{b}^\top \mathbf{y}^*$.*

La dimostrazione del teorema di dualità forte utilizza il lemma di Farkas. A sua volta, la dimostrazione del lemma di Farkas si basa sul seguente risultato geometrico.

Lemma 3 (Iperpiano separatore) *Sia $C \subset \mathbb{R}^d$ chiuso e convesso e $\mathbf{z} \notin C$. Allora esiste un iperpiano che separa \mathbf{z} da C . Ovvero, esiste $\mathbf{w} \in \mathbb{R}^d$ tali che*

$$\mathbf{w}^\top \mathbf{z} > \max_{\mathbf{x} \in C} \mathbf{w}^\top \mathbf{x} . \quad (6)$$

Per verificare che l'esistenza dell'iperpiano equivale a (6) si noti che l'enunciato del lemma può essere riscritto come:

$$\exists \gamma > 0 \quad \text{tale che} \quad \mathbf{w}^\top \mathbf{z} > \gamma \quad \text{e} \quad \forall \mathbf{x} \in C \quad \mathbf{w}^\top \mathbf{x} \leq \gamma . \quad (7)$$

Dato che $\mathbf{w}^\top \mathbf{x} = \gamma$ è l'equazione dell'iperpiano (\mathbf{w}, γ) che intercetta il vettore perpendicolare \mathbf{w} a distanza $\gamma / \|\mathbf{w}\|$ dall'origine, la condizione (7) identifica un iperpiano (\mathbf{w}, γ) tale che \mathbf{z} appartiene al semispazio $\{\mathbf{x} \in \mathbb{R}^d : \mathbf{w}^\top \mathbf{x} > \gamma\}$, che è il semispazio complementare a $\{\mathbf{x} \in \mathbb{R}^d : \mathbf{w}^\top \mathbf{x} \leq \gamma\}$ entro il quale C è contenuto.

L'intuizione dietro la dimostrazione del lemma dell'iperpiano separatore è semplice. Dato che C è chiuso e convesso, la proiezione ortogonale \mathbf{z}_0 di \mathbf{z} su C esiste ed è unica. Si consideri il segmento che unisce \mathbf{z} e \mathbf{z}_0 . L'iperpiano passante per \mathbf{z}_0 e perpendicolare al vettore $\mathbf{z} - \mathbf{z}_0$ soddisfa le condizioni

stabilite nell'enunciato del lemma. L'equazione di questo iperpiano è $\mathbf{w}^\top \mathbf{x} = \gamma$ con $\mathbf{w} = \mathbf{z} - \mathbf{z}_0$ e $\gamma = \mathbf{w}^\top \mathbf{z}_0$. DIMOSTRAZIONE DEL LEMMA DELL'IPERPIANO SEPARATORE. Sia

$$\mathbf{z}_0 = \operatorname{argmin}_{\mathbf{x} \in C} \|\mathbf{x} - \mathbf{z}\|$$

la proiezione di \mathbf{z} su C . Questa esiste ed è unica perché C è chiuso e convesso. Poniamo $\mathbf{w} = \mathbf{z} - \mathbf{z}_0$ e $\gamma = \mathbf{w}^\top \mathbf{z}_0$. La prima disuguaglianza dell'enunciato si dimostra facilmente,

$$\mathbf{w}^\top \mathbf{z} = \mathbf{w}^\top \mathbf{z} + \mathbf{w}^\top \mathbf{z}_0 - \mathbf{w}^\top \mathbf{z}_0 = \mathbf{w}^\top (\mathbf{z} - \mathbf{z}_0) + \mathbf{w}^\top \mathbf{z}_0 = \|\mathbf{w}\|^2 + \gamma > \gamma .$$

Per dimostrare la seconda disuguaglianza dell'enunciato, scegliamo un $\mathbf{x} \in C$ arbitrario. Dato che C è convesso tutti i punti del segmento che unisce \mathbf{x} con \mathbf{z}_0 appartengono a C . Ovvero, $\mathbf{x}_\alpha \in C$ per ogni $0 < \alpha \leq 1$, dove $\mathbf{x}_\alpha = (1 - \alpha)\mathbf{z}_0 + \alpha\mathbf{x}$. Inoltre, dato che \mathbf{z}_0 è il punto di C più vicino a \mathbf{z} , vale $\|\mathbf{z} - \mathbf{z}_0\| \leq \|\mathbf{z} - \mathbf{x}_\alpha\|$. Quindi,

$$\begin{aligned} \|\mathbf{z}_0 - \mathbf{z}\|^2 &\leq \|(1 - \alpha)\mathbf{z}_0 + \alpha\mathbf{x} - \mathbf{z}\|^2 = \|(1 - \alpha)(\mathbf{z}_0 - \mathbf{z}) + \alpha(\mathbf{x} - \mathbf{z})\|^2 \\ &= (1 - \alpha)^2 \|\mathbf{z}_0 - \mathbf{z}\|^2 + 2\alpha(1 - \alpha)(\mathbf{z}_0 - \mathbf{z})^\top (\mathbf{x} - \mathbf{z}) + \alpha^2 \|\mathbf{x} - \mathbf{z}\|^2 \end{aligned}$$

da cui, osservando che $(1 - \alpha)^2 = 1 - 2\alpha + \alpha^2$, otteniamo

$$\alpha(2 - \alpha) \|\mathbf{z}_0 - \mathbf{z}\|^2 \leq 2\alpha(1 - \alpha)(\mathbf{z}_0 - \mathbf{z})^\top (\mathbf{x} - \mathbf{z}) + \alpha^2 \|\mathbf{x} - \mathbf{z}\|^2 .$$

Dividendo entrambi i membri per $2\alpha > 0$, la disuguaglianza diventa

$$\left(1 - \frac{\alpha}{2}\right) \|\mathbf{z}_0 - \mathbf{z}\|^2 \leq (1 - \alpha)(\mathbf{z}_0 - \mathbf{z})^\top (\mathbf{x} - \mathbf{z}) + \frac{\alpha}{2} \|\mathbf{x} - \mathbf{z}\|^2 .$$

Si noti ora che scegliendo α abbastanza piccolo otteniamo che

$$\|\mathbf{z}_0 - \mathbf{z}\|^2 \leq (\mathbf{z}_0 - \mathbf{z})^\top (\mathbf{x} - \mathbf{z}) . \quad (8)$$

Dato che $\|\mathbf{z}_0 - \mathbf{z}\|^2 = (\mathbf{z}_0 - \mathbf{z})^\top (\mathbf{z}_0 - \mathbf{z})$, la disuguaglianza (8) è equivalente a $(\mathbf{z} - \mathbf{z}_0)(\mathbf{z}_0 - \mathbf{x}) \geq 0$. Ricordando che $\mathbf{w} = \mathbf{z} - \mathbf{z}_0$ e $\gamma = \mathbf{w}^\top \mathbf{z}_0$ concludiamo che $\gamma \geq \mathbf{w}^\top \mathbf{x}$ e la dimostrazione è conclusa. \square

Siamo ora pronti per enunciare e dimostrare il lemma di Farkas.

Lemma 4 (Farkas) *Sia A una matrice $m \times n$ e $\mathbf{b} \in \mathbb{R}^m$ un vettore. Allora il sistema*

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

ha soluzione se e solo se il sistema

$$\begin{aligned} A^\top \mathbf{y} &\leq \mathbf{0} \\ \mathbf{b}^\top \mathbf{y} &> 0 \end{aligned}$$

non ha soluzioni.

DIMOSTRAZIONE. Sia \mathbf{x}_0 una soluzione del primo sistema. Ovvero, $A\mathbf{x}_0 = \mathbf{b}$ e $\mathbf{x}_0 \geq \mathbf{0}$. Allora il secondo sistema non ha soluzioni. Infatti, se \mathbf{y}_0 fosse una soluzione del secondo sistema, allora

$$\mathbf{b}^\top \mathbf{y}_0 = (A\mathbf{x}_0)^\top \mathbf{y}_0 = \mathbf{x}_0^\top (A^\top \mathbf{y}_0) \leq 0$$

che contraddice $\mathbf{b}^\top \mathbf{y}_0 > 0$.

Dimostriamo ora che se il primo sistema non ha soluzioni allora il secondo ha almeno una soluzione. Siano $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ le colonne della matrice A e sia C l'insieme dei punti $x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n$ con $\mathbf{x} \geq \mathbf{0}$. In altre parole, C contiene tutti i vettori della forma $A\mathbf{x}$ per ogni $\mathbf{x} \geq \mathbf{0}$. Geometricamente, C è un cono chiuso e convesso; ovvero, un prisma centrato nell'origine, senza base e con spigoli ottenuti prolungando i vettori $\mathbf{a}_1, \dots, \mathbf{a}_n$ nel verso positivo. Si noti ora che assumere che il primo sistema non abbia soluzioni è equivalente a $\mathbf{b} \notin C$. Dato che C è chiuso e convesso, per il lemma dell'iperpiano separatore esiste un \mathbf{y} tale che

$$\max_{\mathbf{x} \in C} \mathbf{y}^\top \mathbf{x} < \mathbf{y}^\top \mathbf{b} . \quad (9)$$

Dimostriamo ora che \mathbf{y} è soluzione del secondo sistema. Per prima cosa abbiamo che $\mathbf{y}^\top \mathbf{b} > 0$ perché $\mathbf{0} \in C$. Dimostriamo ora che vale $A^\top \mathbf{y} \leq \mathbf{0}$. Per assurdo, se per esempio $(A^\top \mathbf{y})_1 > 0$, allora esisterebbe un vettore $\mathbf{x}' = (\alpha, 0, \dots, 0)$ dove $\alpha > 0$ è scelto in modo che

$$(A^\top \mathbf{y})^\top \mathbf{x}' = (A^\top \mathbf{y})_1 \alpha > \mathbf{b}^\top \mathbf{y} .$$

Ma ciò non è possibile. Infatti

$$(A^\top \mathbf{y})^\top \mathbf{x}' = \mathbf{y}^\top (A\mathbf{x}') < \mathbf{y}^\top \mathbf{b}$$

per (9) dato che $A\mathbf{x}' \in C$. Questo conclude la dimostrazione. \square

Per dimostrare il teorema di dualità forte useremo una variante del Lemma di Farkas la cui dimostrazione (che omettiamo) è simile a quella del Lemma 4.

Lemma 5 (Farkas, seconda variante) *Sia A una matrice $m \times n$ e $\mathbf{b} \in \mathbb{R}^m$ un vettore. Allora il sistema $A\mathbf{x} \geq \mathbf{b}$ ha soluzione in \mathbb{R}^n se e solo se il sistema*

$$\begin{aligned} A^\top \mathbf{y} &= \mathbf{0} \\ \mathbf{b}^\top \mathbf{y} &> 0 \\ \mathbf{y} &\geq \mathbf{0} \end{aligned}$$

non ha soluzioni.

Siamo ora pronti a dimostrare il teorema di dualità forte.

DIMOSTRAZIONE DEL TEOREMA 2. Il teorema di dualità debole ci dice che se \mathbf{x}, \mathbf{y} sono soluzioni ammissibili di (P) e (D) della forma (5), allora vale sempre $\mathbf{c}^\top \mathbf{x} \geq \mathbf{b}^\top \mathbf{y}$. Sia \mathbf{y}^* la soluzione ottima di (D) con valore $\mathbf{b}^\top \mathbf{y}^* = \delta$. Si consideri il sistema

$$\begin{aligned} A\mathbf{x} &\geq \mathbf{b} \\ \mathbf{c}^\top \mathbf{x} &\leq \delta \end{aligned} \quad (10)$$

Se (10) ha soluzione, allora (P) deve avere una soluzione \mathbf{x}^* tale che $\mathbf{c}^\top \mathbf{x}^* \leq \delta$ e il teorema è dimostrato.

Per dimostrare che (10) ha soluzione, notiamo che il sistema (10) può essere equivalentemente riscritto come

$$\begin{bmatrix} A \\ -\mathbf{c}^\top \end{bmatrix} \mathbf{x} \geq \begin{bmatrix} \mathbf{b} \\ -\delta \end{bmatrix} \quad (11)$$

Supponiamo per assurdo che (11) non abbia soluzioni. Allora per il Lemma 5 esistono $\mathbf{y} \in \mathbb{R}^m$ e $z \in \mathbb{R}$ tali che

$$[A^\top, -\mathbf{c}] \begin{bmatrix} \mathbf{y} \\ z \end{bmatrix} = \mathbf{0} \quad \begin{bmatrix} \mathbf{b} \\ -\delta \end{bmatrix}^\top \begin{bmatrix} \mathbf{y} \\ z \end{bmatrix} > 0 \quad \begin{bmatrix} \mathbf{y} \\ z \end{bmatrix} \geq \mathbf{0} \quad (12)$$

Ciò implica $A^\top \mathbf{y} = z\mathbf{c}$ e $\mathbf{b}^\top \mathbf{y} > z\delta$. Consideriamo due casi.

Caso 1: $z \neq 0$. Dato che $(\mathbf{y}, z) \geq \mathbf{0}$ per (12), $\mathbf{y}' = \mathbf{y}/z$ soddisfa $\mathbf{y}' \geq \mathbf{0}$. Inoltre $A^\top \mathbf{y} = z\mathbf{c}$ implica $A^\top \mathbf{y}' = \mathbf{c}$ mentre $\mathbf{b}^\top \mathbf{y} > z\delta$ implica $\mathbf{b}^\top \mathbf{y}' > \delta$. Quindi \mathbf{y}' è una soluzione ammissibile di (D) con un valore maggiore di δ e abbiamo una contraddizione.

Caso 2: $z = 0$. Allora $A^\top \mathbf{y} = \mathbf{0}$ e $\mathbf{b}^\top \mathbf{y} > 0$. Dato che $A^\top \mathbf{y}^* = \mathbf{c}$ e $\mathbf{b}^\top \mathbf{y}^* = \delta$, vale $A^\top (\mathbf{y}^* + \varepsilon \mathbf{y}) = \mathbf{c}$ per ogni $\varepsilon > 0$. Inoltre, dato che $\mathbf{y} \geq \mathbf{0}$ per (12) e $\mathbf{y}^* \geq \mathbf{0}$, abbiamo che $(\mathbf{y}^* + \varepsilon \mathbf{y}) \geq \mathbf{0}$. Infine, $\mathbf{b}^\top (\mathbf{y}^* + \varepsilon \mathbf{y}) = \delta + \varepsilon \mathbf{b}^\top \mathbf{y} > \delta$. Quindi $\mathbf{y}^* + \varepsilon \mathbf{y}$ è una soluzione ammissibile di (D) con un valore maggiore di δ e abbiamo ancora una contraddizione.

Ma allora (10) ha soluzioni e il teorema è dimostrato. □

Corollario 6 *Sia (P) un programma lineare con duale (D), dove (P) e (D) hanno forma (1) e siano \mathbf{x} e \mathbf{y} soluzioni ammissibili di (P) e (D). Allora \mathbf{x} e \mathbf{y} sono ottime se e solo se le condizioni di complementarità valgono.*

DIMOSTRAZIONE. Da completare. □