Complementi di Algoritmi e Strutture Dati

Programmazione lineare

Docente: *Nicolò Cesa-Bianchi* versione 18 aprile 2018

Moltissimi problemi reali si prestano ad essere modellati in termini di ottimizzazione lineare (detta anche programmazione lineare, o LP). A dispetto di questa generalità, i problemi di LP possono essere risolti in computazionalmente efficiente. Inoltre, la LP è concettualmente utile, in quanto fornisce uno strumento potente e versatile per l'analisi teorica di svariati problemi.

Per introdurre la LP possiamo partire dalla soluzione di sistemi di equazioni lineari

$$
\begin{array}{ccc}
a_{1,1}x_1 + & \cdots & + a_{1,n}x_n = b_1 \\
 & \vdots & \\
a_{m,1}x_1 + & \cdots & + a_{m,n}x_n = b_m
\end{array}
$$

dove $\{a_{i,j} \,:\, i = 1,\dots,m,\ j = 1,\dots,m\}$ e $b_1,\dots,b_m$ sono coefficienti reali noti e lo scopo è trovare un assegnamento a $x_1,\dots,x_n$ che soddisfi le m equazioni lineari. Un semplice esempio con $n = m = 2$ è

$$
\begin{aligned}
2x_1 + x_2 &= 1 \\
x_1 + 2x_2 &= 1
\end{aligned}
$$

È utile pensare al sistema in termini geometrici: $2x_1 + x_2 = 1$ e $x_1 + 2x_2 = 1$ sono le equazioni di due rette nel piano e la soluzione del sistema (se esiste) è il punto di intersezione delle rette. In n dimensioni, le equazioni lineari $a_{k,1}x_1 + \cdots + a_{k,n}x_n = b_k$ definiscono degli iperpiani e la soluzione del sistema (sempre se questa esiste) è data da un qualunque punto all'intersezione degli iperpiani. Si rammenta infatti che un sistema di m equazioni lineari in n incognite può avere zero, una, o infinite soluzioni (teorema di Rouché-Capelli) a seconda della relazione fra n e i ranghi delle matrici A e B, dove

$$
A = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \vdots & \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} \qquad B = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n},\ b_1 \\ & \vdots & \\ a_{m,1} & \cdots & a_{m,n},\ b_m \end{bmatrix}
$$

In LP le equazioni lineari possono essere sostituite da disequazioni della forma $a_{k,1}x_1 + \cdots + a_{k,n}x_n \le b_k$. Ogni disequazione lineare definisce un semispazio e un sistema di disequazioni lineari definisce una regione —chiamata regione ammissibile— determinata dall'intersezione di semispazi (cioè un poliedro). La soluzione del problema di LP è il "miglior" punto $(x_1,\dots,x_n)$ all'interno della regione ammissibile (se essa contiene almeno un punto), dove migliore è riferito al valore di una funzione obiettivo $f(x_1,\dots,x_n)$ che è anch'essa lineare nelle sue variabili.

Gli ingredienti di un programma lineare sono quindi:

1. Variabili di decisione $x_1,\dots,x_n \in \mathbb{R}$.
2. Vincoli lineari

$$
\sum_{j=1}^{n} a_{i,j}x_j \le b_i \quad \text{oppure} \quad \sum_{j=1}^{n} a_{i,j}x_j = b_i \quad \text{oppure} \quad \sum_{j=1}^{n} a_{i,j}x_j \ge b_i
$$

3. Funzione obiettivo lineare

$$\min \sum_{j=1}^{n} c_j x_j \quad \text{oppure} \quad \max \sum_{j=1}^{n} c_j x_j$$

In pratica, possiamo sempre ridurre un problema LP in un problema equivalente (cioè, con la stessa soluzione) della forma

$$\begin{aligned} &\max \sum_{j=1}^{n} c_j x_j \\ &\sum_{j=1}^{n} a_{i,j} x_j \leq b_i \qquad i = 1, \ldots, m \end{aligned}$$

Infatti, un vincolo $\sum_{j=1}^{n} a_{i,j} x_j \geq b_i$ può essere equivalentemente riscritto come $\sum_{j=1}^{n} (-a_{i,j}) x_j \leq -b_i$ semplicemente cambiando segno ai coefficienti. Un vincolo $\sum_{j=1}^{n} a_{i,j} x_j = b_i$ può essere sostuito da due vincoli $\leq$ e $\geq$. Un problema $\min \sum_{j=1}^{n} c_j x_j$ può essere equivalentemente riscritto come $\max \sum_{j=1}^{n} (-c_j) x_j$ sempre cambiando segno ai coefficienti.

Per continuare a sviluppare l'intuizione geometrica, consideriamo il seguente problema LP,

$$\begin{aligned} \max \quad & x_1 + x_2 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \\ & 2x_1 + x_2 \leq 1 \\ & x_1 + 2x_2 \leq 1 \end{aligned}$$

Le curve di livello $\{x_1 + x_2 = v \,:\, v \in \mathbb{R}\}$ della funzione obiettivo sono rette con coefficiente angolare -1. Dato che i primi due vincoli impongono di cercare la soluzione nel quadrante positivo, possiamo pensare di risolvere il problema considerando la retta $x_1 + x_2 = v$ per valori di v crescenti a partire da $v = 0$. L'ottimo della funzione obiettivo corrisponde quindi all'ultimo punto di intersezione della retta $x_1 + x_2 = v$ (per $v \geq 0$ crescente) con la regione ammissibile.

In generale possiamo riassumere come segue.

- La regione ammissibile di un problema LP è l'intersezione dei semispazi corrispondenti ai vincoli lineari. Tale regione può essere limitata, vuota o illimitata (pensate, per esempio, a un caso con due soli vincoli $x_1 \geq 0$ e $x_2 \geq 0$). Se la regione ammissibile è vuota allora non esistono soluzioni (ammissibili) al problema.
- Le curve di livello della funzione obiettivo $c_1 x_1 + \cdots + c_n x_n$ sono iperpiani ortogonali al vettore $\boldsymbol{c} = (c_1, \ldots, c_n)$.
- La soluzione ottima è il punto della regione ammissibile più lontano nella direzione di $\boldsymbol{c}$ (per un problema di massimo).
- Se la regione ammissibile è illimitata, allora il massimo della funzione obiettivo in tale regione può essere infinito. Questo avviene quando il prolungamento di $\boldsymbol{c}$ non esce mai dalla regione ammissibile.
- Se la soluzione ottima esiste ed ha valore finito, allora è uno dei vertici del poliedro che definisce la regione ammissibile. Può capitare che la soluzione ottima si trovi su un vertice che sta su una faccia del poliedro ortogonale a $\boldsymbol{c}$. In questo caso tutti i punti della faccia del poliedro sono soluzioni ottime.

Ricapitolando ulteriormente, un problema di LP può essere:

1. limitato (almeno una soluzione ottima finita),
2. illimitato (la funzione obiettivo ha un massimo illimitato nella regione ammissibile),
3. non ammissibile (la regione ammissibile è vuota).

Vediamo ora primo problema che possiamo formulare come LP. Consideriamo un poliedro definito da un sistema di disequazioni lineari,

$$\sum_{j=1}^{n} a_{i,j} x_j \leq b_i \qquad i = 1, \ldots, m \tag{1}$$

Ricordiamo che un punto $\boldsymbol{x} \in \mathbb{R}^n$ appartiene al poliedro se e solo se $\boldsymbol{x}$ soddisfa il sistema di disequazioni (1).

Vogliamo determinare il centro della più grande sfera completamente contenuta nel poliedro (detto centro di Chebyshev). Le variabili di decisione sono quindi il centro $\boldsymbol{x} \in \mathbb{R}^n$ e il raggio $r \in \mathbb{R}$ e vogliamo massimizzare r sotto il vincolo che la sfera di centro $\boldsymbol{x}$ sia contenuta nel poliedro. I punti sulla sfera sono della forma $\boldsymbol{x} + \boldsymbol{u}$, dove $\boldsymbol{u}$ è un vettore di norma r, $\|\boldsymbol{u}\| = r$. Quindi la sfera $(\boldsymbol{x}, r)$ è contenuta nel poliedro se e solo se

$$\max_{\boldsymbol{u} : \|\boldsymbol{u}\| = r} \sum_{j=1}^{n} a_{i,j} \big(x_j + u_j \big) \leq b_i \qquad i = 1, \ldots, m$$

Notiamo ora che, detto $\boldsymbol{a}_i = (a_{i,1}, \ldots, a_{i,n})$,

$$\begin{aligned}
\max_{\boldsymbol{u} : \|\boldsymbol{u}\| = r} \sum_{j=1}^{n} a_{i,j} \big(x_j + u_j \big) &= \sum_{j=1}^{n} a_{i,j} x_j + \max_{\boldsymbol{u} : \|\boldsymbol{u}\| = r} \sum_{j=1}^{n} a_{i,j} u_j \\
&= \sum_{j=1}^{n} a_{i,j} x_j + \frac{r}{\|\boldsymbol{a}_i\|} \|\boldsymbol{a}_i\|^2 && (\text{per } \boldsymbol{u} = (r/\|\boldsymbol{a}_i\|)\boldsymbol{a}_i) \\
&= \sum_{j=1}^{n} a_{i,j} x_j + r \|\boldsymbol{a}_i\|
\end{aligned}$$

Per capire la seconda uguaglianza ricordiamo che, dati due vettori $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ che formano un angolo θ fra loro, vale che

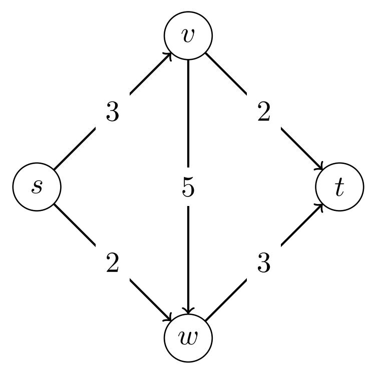
$$\sum_{j=1}^{n} u_j v_j = \boldsymbol{u}^\top \boldsymbol{v} = \|\boldsymbol{u}\| \, \|\boldsymbol{v}\| \cos\theta \ .$$

Quindi il massimo si ha quando $\cos\theta = 1$ ovvero quando i due vettori sono allineati. In tal caso $\boldsymbol{u} = \alpha \boldsymbol{v}$ per un qualche $\alpha > 0$ e quindi $\sum_{j=1}^{n} u_j v_j = \alpha \|\boldsymbol{v}\|^2$. Il vettore $\boldsymbol{u}$ di norma r che massimizza $\sum_{j=1}^{n} a_{i,j} u_j = \boldsymbol{a}_i^\top \boldsymbol{u}$ è quindi $\boldsymbol{u} = (r/\|\boldsymbol{a}_i\|)\boldsymbol{a}_i$.

Dato che $\sum_{j=1}^{n} a_{i,j} x_j + r \|\boldsymbol{a}_i\| \leq b_i$ è un vincolo lineare, il centro di Chebyshev è la soluzione del programma lineare

$$\begin{aligned}
&\max \ \ r \\
&\sum_{j=1}^{n} a_{i,j} x_j + r \|\boldsymbol{a}_i\| \leq b_i \qquad i = 1, \ldots, m
\end{aligned}$$

Vediamo ora un secondo esempio di problema che si presta ad una formulazione LP. Il problema di flusso massimo (max flow) consiste nel determinare qual è il massimo flusso che può essere fatto scorrere tra un vertice sorgente e un vertice destinazione in un grafo orientato i cui archi hanno una capacità limitata. Più formalmente, sia $G = (V, E)$ un grafo orientato tale che per ogni $u, v \in V$, $(u, v) \in E$ implica $(v, u) \notin E$ e, inoltre, $(v, v) \notin E$. Sia $s \in V$ un vertice sorgente e $t \in V$ un vertice destinazione (senza perdita di generalità, possiamo assumere che s non ha archi entranti e t non ha archi uscenti). Sia inoltre $c : E \to \mathbb{R}$ una funzione che assegna una capacità $c(e) \geq 0$ positiva ad ogni arco del grafo (si veda la figura qui sotto).



Introduciamo la notazione $E_v^+ \equiv E \cap \{(v, u) : u \in V\}$ e $E_v^- \equiv E \cap \{(u, v) : u \in V\}$ per indicare, rispettivamente, l'insieme degli archi uscenti ed entranti di v.

Un flusso è una funzione $f : E \to \mathbb{R}$ che assegna un flusso $f(e) \geq 0$ ad ogni arco di G. Un flusso ammissibile deve soddisfare le due seguenti condizioni:

1. Vincoli di capacità: $f(e) \leq c(e)$ per ogni $e \in E$.
2. Vincoli di conservazione: per ogni vertice $v \in V$ diverso da s e t, la somma dei flussi sugli archi entranti in v dev'essere uguale alla somma dei flussi sugli archi uscenti da v,

$$\sum_{e \in E_v^-} x_e = \sum_{e \in E_v^+} x_e \qquad v \in V \setminus \{s, t\}$$

Per esempio, $f(s, v) = 3 \quad f(s, w) = 2 \quad f(v, w) = 1 \quad f(v, t) = 2 \quad f(w, t) = 3$ è un flusso ammissibile rispetto al grafo in figura. L'obiettivo è calcolare un flusso massimo ammissibile. Ovvero, un flusso ammissibile che massimizza il flusso sugli archi uscenti da s (o, equivalentemente, il flusso sugli archi entranti in t).

Per dare una formulazione LP al problema di flusso massimo allochiamo una variabile di decisione x_e per ogni arco $e \in E$ del grafo. Indichiamo con c_e i valori $c(e)$ di capacità. Possiamo quindi scrivere il programma lineare

$$\begin{aligned} \max \quad & \sum_{e \in E_s^+} x_e & \\ x_e &\geq 0 & e \in E \\ x_e &\leq c_e & e \in E \\ \sum_{e \in E_v^-} x_e &= \sum_{e \in E_v^+} x_e & v \in V \setminus \{s, t\} \end{aligned}$$

I vincoli $x_e \geq 0$ ci dicono che i flussi non possono essere negativi. I vincoli $x_e \leq c_e$ implementano il vincolo di capacità, mentre i vincoli $\sum_{e \in E_v^-} x_e = \sum_{e \in E_v^+} x_e$ implementano il vincolo di flusso.