Complementi di Algoritmi e Strutture Dati

Il teorema minimax di Von Neumann

Docente: *Nicolò Cesa-Bianchi* versione 31 maggio 2017

Consideriamo un gioco come la morra cinese (carta, forbice e sasso) oppure pari-e-dispari, dove il giocatore e il suo avversario rivelano simultaneamente le proprie mosse. Questi giochi sono detti a somma zero, in quanto ciò che vince il giocatore lo perde l'avversario e viceversa; possiamo rappresentarli con una matrice reale G di dimensioni $m \times n$, dove le m righe rappresentano le mosse del giocatore e le n colonne rappresentano le mosse dell'avversario. Se il giocatore sceglie la mossa i e l'avversario sceglie la mossa j, allora il giocatore guadagna $G_{i,j}$ e l'avversario perde $-G_{i,j}$. Se $G_{i,j} > 0$, allora l'avversario paga $G_{i,j}$ al giocatore. Se invece $G_{i,j} < 0$ è il giocatore a pagare $G_{i,j}$ all'avversario. Per esempio, la matrice G di morra cinese è indicata qui sotto a sinistra

	carta	forbice	sasso
carta	0	-1	$+1$
forbice	$+1$	0	-1
sasso	-1	$+1$	0

	alto	basso	centro
sinistra	$+3$	-1	$+2$
destra	-1	$+2$	-2

Ragionamo un istante su quello che può fare il giocatore nel gioco a destra. La mossa sinistra gli dà un guadagno massimo di $+3$ e una perdita massima di -1. La mossa destra gli dà un guadagno massimo di $+2$ e una perdita massima di -2. Quindi il giocatore preferirà giocare sinistra. L'avversario, usando un argomento del tutto simile, preferisce giocare centro. Se entrambi i giocatori seguono il proprio ragionamento, ovvero il giocatore sceglie sinistra e l'avversario centro, abbiamo che il giocatore vince due punti all'avversario. D'altra parte, l'avversario può immaginare che il giocatore faccia il ragionamento che lo porti a giocare sinistra e di conseguenza potrebbe scegliere basso invece di centro, in modo da prendere un punto al giocatore. A sua volta, però, il giocatore può prevedere la contromossa dell'avversario e prevenirla, e così via. Per spezzare questo circolo vizioso e neutralizzare il ragionamento dell'avversario il giocatore può usare la randomizzazione.

Supponiamo che il giocatore riveli all'avversario la distribuzione $\boldsymbol{p}$ da cui estrarrà la propria mossa $I \in \{1, \dots, m\}$. L'avversario sceglie quindi la propria mossa $J \in \{1, \dots, n\}$, dopodiché la mossa del giocatore viene estratta da $\boldsymbol{p}$ e il gioco termina. In questo modo il giocatore delega la scelta effettiva della sua mossa alla randomizzazione usata per estrarre I (il che non vuol dire che ogni mossa venga scelta con la stessa probabilità). Il circolo vizioso è così spezzato in quanto l'avversario conosce $\boldsymbol{p}$ e può quindi calcolare la propria mossa migliore *a meno della randomizzazione usata per estrarre I* sulla quale il giocatore non ha però più alcun controllo. Rivelando la propria strategia, sembra però ovvio che il giocatore abbia avvantaggiato l'avversario. Il teorema minimax ci mostra, sorprendentemente, che questo non è vero.

L'avversario può utilizzare la conoscenza di $\boldsymbol{p}$ per giocare la mossa J che minimizza il valore atteso del guadagno del giocatore,

$$J = \operatorname*{argmin}_{j=1,\dots,n} \sum_{i=1}^{m} G_{i,j}\, p_i \tag{1}$$

dove $p_i = \mathbb{P}\big(I = i\big)$.

A questo punto, il giocatore sceglierà la distribuzione $\boldsymbol{p}^*$ che massimizza il proprio guadagno atteso,

$$\boldsymbol{p}^* = \operatorname*{argmax}_{\boldsymbol{p}} \left(\min_{j=1,\dots,n} \sum_{i=1}^{m} G_{i,j}\, p_i \right) .$$

Si noti che l'avversario che conosce $\boldsymbol{p}$ non ha alcun vantaggio a estrarre la propria mossa J da una distribuzione $\boldsymbol{q}$. Ovvero,

$$\min_{j=1,\dots,n} \sum_{i=1}^{m} G_{i,j}\, p_i = \min_{\boldsymbol{q}} \sum_{j=1}^{n} \left(\sum_{i=1}^{m} G_{i,j}\, p_i \right) q_j = \min_{\boldsymbol{q}} \boldsymbol{p}^\top G \boldsymbol{q} \tag{2}$$

dove l'ultimo termine utilizza la notazione matriciale per denotare la doppia somma che lo precede. Sfruttando questa identità, possiamo allora scrivere il guadagno atteso del giocatore come

$$\max_{\boldsymbol{p}} \left(\min_{\boldsymbol{q}} \boldsymbol{p}^\top G \boldsymbol{q} \right) .$$

Simmetricamente, possiamo pensare che sia l'avversario a rivelare $\boldsymbol{q}$ al giocatore, che quindi può calcolare la propria mossa migliore I per massimizzare il valore atteso della propria vincita,

$$I = \operatorname*{argmax}_{i=1,\dots,n} \sum_{j=1}^{n} G_{i,j}\, q_j .$$

La strategia migliore dell'avversario porta allora ad un guadagno atteso per il giocatore pari a

$$\min_{\boldsymbol{q}} \left(\max_{i=1,\dots,m} \sum_{j=1}^{n} G_{i,j}\, q_j \right) = \min_{\boldsymbol{q}} \left(\max_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q} \right) . \tag{3}$$

Il teorema minimax dice che il guadagno atteso del giocatore non cambia a seconda di chi sia il primo a rivelare la propria strategia randomizzata.

Teorema 1 (Minimax) *In qualsiasi gioco G a somma zero vale*

$$\max_{\boldsymbol{p}} \left(\min_{\boldsymbol{q}} \boldsymbol{p}^\top G \boldsymbol{q} \right) = \min_{\boldsymbol{q}} \left(\max_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q} \right) .$$

Un altro modo di interpretare teorema minimax è il seguente. Sia V_G il valore comune delle formule al membro sinistro e destro nell'enunciato del teorema. Allora il giocatore ha una strategia

$$\boldsymbol{p}^* = \operatorname*{argmax}_{\boldsymbol{p}} \left(\min_{\boldsymbol{q}} \boldsymbol{p}^\top G \boldsymbol{q} \right)$$

che gli garantisce un guadagno atteso di almeno V_G qualunque sia la strategia $\boldsymbol{q}$ dell'avversario. Viceversa, l'avversario ha una strategia

$$\boldsymbol{q}^* = \operatorname*{argmin}_{\boldsymbol{q}} \left(\max_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q} \right)$$

che garantisce un guadagno atteso del giocatore pari ad al più V_G qualunque sia la strategia $\boldsymbol{p}$ del giocatore. La dimostrazione del teorema minimax ci mostra come $\boldsymbol{p}^*$ e $\boldsymbol{q}^*$ siano le soluzioni ottime di una coppia di LP duali.

Dimostrazione del teorema minimax. Per prima cosa riscriviamo l'enunciato tenendo conto di (2) e (3),

$$\max_{\boldsymbol{p}} \left(\min_{j=1,\dots,n} \sum_{i=1}^{m} G_{i,j}\, p_i \right) = \min_{\boldsymbol{q}} \left(\max_{i=1,\dots,m} \sum_{j=1}^{n} G_{i,j}\, q_j \right) \ . \tag{4}$$

Possiamo scrivere il valore definito del membro sinistro di (4) come la soluzione ottima del seguente programma lineare con variabili di decisione $(v, \boldsymbol{p})$

$$\begin{aligned}
&\max v \\
&\sum_{i=1}^{m} G_{i,j}\, p_i \geq v \quad j = 1, \dots, n \\
&\sum_{i=1}^{m} p_i = 1 \\
&p_i \geq 0 \quad i = 1, \dots, m \\
&v \in \mathbb{R}
\end{aligned}$$

Definiamo ora $A = \big[\mathbf{1},\, -G^\top\big]$, $\boldsymbol{c} = \big(1, \mathbf{0}\big)$ e $\boldsymbol{x} = \big(v, \boldsymbol{x}\big)$. Riscriviamo il programma sopra in modo equivalente qui sotto a sinistra,

$$\begin{array}{llcll}
\max & \boldsymbol{c}^\top \boldsymbol{x} & \qquad & \min & \boldsymbol{b}^\top \boldsymbol{y} \\
 & A\,\boldsymbol{x} \leq \mathbf{0} & & & A^\top \boldsymbol{y} \geq \mathbf{0} \\
 & \mathbf{1}^\top \boldsymbol{p} = 1 & & & \mathbf{1}^\top \boldsymbol{q} = 1 \\
 & \boldsymbol{p} \geq \mathbf{0} & & & \boldsymbol{q} \geq \mathbf{0} \\
 & v \in \mathbb{R} & & & w \in \mathbb{R}
\end{array}$$

Come è facile verificare, il programma nella parte destra è il duale della parte sinistra, con $\boldsymbol{b} = \big(1, \mathbf{0}\big)$ e $\boldsymbol{y} = \big(w, \boldsymbol{q}\big)$. In particolare, il vincolo di uguaglianza $\mathbf{1}^\top \boldsymbol{p} = 1$ corrisponde alla variabile $w \in \mathbb{R}$ nel duale, e viceversa. Riscriviamo ora il duale in modo equivalente,

$$\begin{aligned}
&\min w \\
&\sum_{j=1}^{n} G_{i,j}\, q_j \leq w \quad i = 1, \dots, m \\
&\sum_{j=1}^{m} q_j = 1 \\
&q_j \geq 0 \quad j = 1, \dots, n \\
&w \in \mathbb{R}
\end{aligned}$$

Si noti ora come il valore della soluzione ottima di questo programma corrisponda al valore definito dal membro destro di (4). Quindi i valori dei membri sinistro e destro di (4) sono i valori delle soluzioni ottime di programmi lineari duali. Per il teorema di dualità forte questi valori coincidono. $\square$