Complementi di Algoritmi e Strutture Dati

Il teorema minimax di Von Neumann

Docente: *Nicolò Cesa-Bianchi* versione 17 maggio 2026

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,

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

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, sia $\boldsymbol{g}$ il vettore di componenti

$$g_j = \sum_{i=1}^{m} G_{i,j}\, p_i \qquad \text{allora} \qquad \min_{\boldsymbol{q}} \sum_{j=1}^{n} g_j q_j = \min_{j=1,\dots,n} g_j$$

ovvero $\boldsymbol{g}^\top \boldsymbol{q}$ è minimizzato concentrando $\boldsymbol{q}$ sulla componente minima di $\boldsymbol{g}$. Questo implica che

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

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

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

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 coppia $(\boldsymbol{p}^*, \boldsymbol{q}^*)$ è un equilibrio di Nash del gioco a somma zero. Ovvero, vale simultaneamente che

$$(\boldsymbol{p}^*)^\top G \boldsymbol{q}^* \geq \boldsymbol{p}^\top G \boldsymbol{q}^* \quad \text{e} \quad (\boldsymbol{p}^*)^\top G \boldsymbol{q}^* \leq (\boldsymbol{p}^*)^\top G \boldsymbol{q} \qquad \text{per ogni } \boldsymbol{p}, \boldsymbol{q}$$

Queste due condizioni ci dicono che né il giocatore né l'avversario sono incentivati a deviare dalla propria strategia di equilibrio se anche l'altro non devia. Si noti anche che

$$\boldsymbol{p}^* = \operatorname*{argmax}_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q}^* \qquad \text{e} \qquad \boldsymbol{q}^* = \operatorname*{argmin}_{\boldsymbol{q}} (\boldsymbol{p}^*)^\top G \boldsymbol{q}$$

ovvero $\boldsymbol{p}^*$ e $\boldsymbol{q}^*$ sono l'una la risposta migliore all'altra.

Nel seguito, assumiamo senza perdita di generalità che gli elementi di G siano riscalati nell'intervallo $[-1, 1]$. Per comodità notazionale, nel seguito usiamo $p(i)$ e $q(j)$ per denotare, rispettivamente, le componenti di $\boldsymbol{p}$ e $\boldsymbol{q}$.

Lemma 2 *Per qualsiasi intero positivo T e per qualsiasi sequenza $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_T$ di strategie dell'avversario, l'algoritmo Hedge con parametro $\eta > 0$ garantisce*

$$\sum_{t=1}^{T} \boldsymbol{p}_t^\top G \boldsymbol{q}_t \geq \max_{\boldsymbol{p}} \sum_{t=1}^{T} \boldsymbol{p}^\top G \boldsymbol{q}_t - \frac{2 \ln m}{\eta} - \eta T$$

Dimostrazione. Per applicare Hedge, riscaliamo gli elementi di G nell'intervallo $[0, 1]$. Definiamo quindi il vettore delle perdite delle azioni al tempo t come $\boldsymbol{\ell}_t = (\mathbf{1} - G\boldsymbol{q}_t)/2 \in [0, 1]$ dove $\mathbf{1} = (1, \ldots, 1)$. L'analisi di Hedge garantisce che le distribuzioni $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_T$ calcolate dall'algoritmo soddisfano

$$\sum_{t=1}^{T} \boldsymbol{\ell}_t^\top \boldsymbol{p}_t \leq \min_{i=1,\ldots,m} \sum_{t=1}^{T} \ell_t(i) + \frac{\ln m}{\eta} + \frac{\eta}{2} T \tag{2}$$

Dato che la funzione lineare di $\boldsymbol{p}$

$$\sum_{t=1}^{T} \boldsymbol{\ell}_t^\top \boldsymbol{p}$$

è minimizzata quando $\boldsymbol{p}$ è concentrata sulla componente più piccola di $\boldsymbol{\ell}_t$, abbiamo che

$$\min_{i=1,\ldots,m} \sum_{t=1}^{T} \ell_t(i) = \min_{\boldsymbol{p}} \sum_{t=1}^{T} \boldsymbol{\ell}_t^\top \boldsymbol{p}$$

Ricordando che $\boldsymbol{\ell}_t = (\mathbf{1} - G\boldsymbol{q}_t)/2$ abbiamo che, per un $\boldsymbol{p}$ arbitrario,

$$\boldsymbol{p}^\top \boldsymbol{\ell}_t = \boldsymbol{p}^\top \frac{\mathbf{1} - G\boldsymbol{q}_t}{2} = \frac{\mathbf{1} - \boldsymbol{p}^\top G \boldsymbol{q}_t}{2}$$

Infine, moltiplicando per 2 entrambi i membri di (2) per cancellare il fattore $\frac{1}{2}$, otteniamo la tesi. □

Siamo ora pronti per la dimostrazione del teorema minimax.

Dimostrazione del teorema minimax. Sia

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

Allora,

$$\max_{\boldsymbol{p}} \min_{\boldsymbol{q}} \boldsymbol{p}^\top G \boldsymbol{q} \leq \max_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q}^* = \min_{\boldsymbol{q}} \max_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q}$$

Per dimostrare l'altra direzione, supponiamo che $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_T$ siano generati da Hedge rispetto alle strategie $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_T$ dell'avversario definite come

$$\boldsymbol{q}_t = \operatorname*{argmin}_{\boldsymbol{q}} \boldsymbol{p}_t^\top G \boldsymbol{q}$$

ovvero $\boldsymbol{q}_t$ è la risposta migliore dell'avversario rispetto alla $\boldsymbol{p}_t$ del giocatore. Definiamo anche

$$\overline{\boldsymbol{p}}_T = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{p}_t \qquad \text{e} \qquad \overline{\boldsymbol{q}}_T = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{q}_t$$

Possiamo quindi scrivere:

$$\begin{aligned}
\max_{\boldsymbol{p}} \min_{\boldsymbol{q}} \boldsymbol{p}^\top G \boldsymbol{q} &\geq \min_{\boldsymbol{q}} \overline{\boldsymbol{p}}_T^\top G \boldsymbol{q} \\
&= \min_{\boldsymbol{q}} \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{p}_t^\top G \boldsymbol{q} \\
&\geq \frac{1}{T} \sum_{t=1}^{T} \min_{\boldsymbol{q}} \boldsymbol{p}_t^\top G \boldsymbol{q} \\
&= \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{p}_t^\top G \boldsymbol{q}_t \\
&\geq \max_{\boldsymbol{p}} \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{p}^\top G \boldsymbol{q}_t - \frac{2 \ln m}{\eta T} - \eta \\
&= \max_{\boldsymbol{p}} \boldsymbol{p}^\top G \, \overline{\boldsymbol{q}}_T - \frac{2 \ln m}{\eta T} - \eta \\
&\geq \min_{\boldsymbol{q}} \max_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q} - \frac{2 \ln m}{\eta T} - \eta
\end{aligned}$$

dove abbiamo applicato il Lemma 2. Scegliendo

$$\eta = \sqrt{\frac{2 \ln m}{T}}$$

otteniamo

$$\max_{\boldsymbol{p}} \min_{\boldsymbol{q}} \boldsymbol{p}^\top G \boldsymbol{q} \geq \min_{\boldsymbol{q}} \max_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q} - \sqrt{\frac{8 \ln m}{T}}$$

Dato che la disuguaglianza vale per qualsiasi T, per $T \to \infty$ abbiamo che

$$\max_{\boldsymbol{p}} \min_{\boldsymbol{q}} \boldsymbol{p}^\top G \boldsymbol{q} \geq \min_{\boldsymbol{q}} \max_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q}$$

che conclude la dimostrazione. □

Si noti che la dimostrazione rivela che

$$\min_{\boldsymbol{q}} \max_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q} \geq \min_{\boldsymbol{q}} \overline{\boldsymbol{p}}_T^\top G \boldsymbol{q} \geq \min_{\boldsymbol{q}} \max_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q} - o(1)$$

dove $o(1) \to 0$ per $T \to \infty$. Questo implica che $\overline{\boldsymbol{p}}_T$ converge a $\boldsymbol{p}$ per $T \to \infty$. In altre parole, la distribuzione empirica delle azioni del giocatore riga converge alla distribuzione ottima.

Per $(\boldsymbol{p}', \boldsymbol{q}')$ arbitrari, definiamo la funzione di scarto

$$\gamma(\boldsymbol{p}', \boldsymbol{q}') = \max_{\boldsymbol{p}} \boldsymbol{p}^\top G \boldsymbol{q}' - \min_{\boldsymbol{q}} (\boldsymbol{p}')^\top G \boldsymbol{q}$$

Allora $(\boldsymbol{p}', \boldsymbol{q}')$ è un equilibrio di Nash se e solo se $\gamma(\boldsymbol{p}', \boldsymbol{q}') = 0$.

Teorema 3 *Sia G un gioco a due persone a somma zero e siano $(\boldsymbol{p}_1, \boldsymbol{q}_1), (\boldsymbol{p}_2, \boldsymbol{q}_2), \ldots$ generate dalle giocate ripetute di due istanze indipendenti di Hedge. Allora*

$$\lim_{T \to \infty} \gamma\big(\overline{\boldsymbol{p}}_T, \overline{\boldsymbol{q}}_T\big) = 0$$

Dimostrazione.

$$\max_{\boldsymbol{p}} \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{p}^\top G \boldsymbol{q}_t \leq \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{p}_t^\top G \boldsymbol{q}_t + o(1)$$

$$\frac{1}{T} \sum_{t=1}^{T} \boldsymbol{p}_t^\top G \boldsymbol{q}_t \leq \min_{\boldsymbol{q}} \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{p}_t^\top G \boldsymbol{q} + o(1)$$

Sommando le disuguaglianze e semplificando,

$$\max_{\boldsymbol{p}} \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{p}^\top G \boldsymbol{q}_t \leq \min_{\boldsymbol{q}} \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{p}_t^\top G \boldsymbol{q} + o(1)$$

Per linearità, questo è equivalente a

$$\max_{\boldsymbol{p}} \boldsymbol{p}^\top G \overline{\boldsymbol{q}}_T \leq \min_{\boldsymbol{q}} \overline{\boldsymbol{p}}_T^\top G \boldsymbol{q} + o(1)$$

Riordinando l'espressione si ottiene la tesi. □

Corollario 4 *Sia G un gioco a due persone a somma zero con un unico equilibrio di Nash $(\boldsymbol{p}^*, \boldsymbol{q}^*)$ e siano $(\boldsymbol{p}_1, \boldsymbol{q}_1), (\boldsymbol{p}_2, \boldsymbol{q}_2), \ldots$ generate dalle giocate ripetute di due istanze indipendenti di Hedge. Allora $(\overline{\boldsymbol{p}}_T, \overline{\boldsymbol{q}}_T) \to (\boldsymbol{p}^*, \boldsymbol{q}^*)$ per $T \to \infty$.*

Dimostrazione. Se $(\overline{\boldsymbol{p}}_T, \overline{\boldsymbol{q}}_T)$ converge, allora dato che γ è continua

$$0 = \lim_{T\to\infty} \gamma(\overline{\boldsymbol{p}}_T, \overline{\boldsymbol{q}}_T) = \gamma\left(\lim_{T\to\infty} \left(\overline{\boldsymbol{p}}_T, \overline{\boldsymbol{q}}_T\right)\right)$$

Dato che l'equilibrio è unico,

$$\lim_{T\to\infty} \left(\overline{\boldsymbol{p}}_T, \overline{\boldsymbol{q}}^*_T\right) = \left(\boldsymbol{p}^*, \boldsymbol{q}^*\right)$$

Se, per assurdo, $(\overline{\boldsymbol{p}}_T, \overline{\boldsymbol{q}}_T)$ non convergesse, allora $\gamma(\overline{\boldsymbol{p}}_T, \overline{\boldsymbol{q}}_T) \to 0$ necessariamente implica che ci siano più coppie $(\boldsymbol{p}', \boldsymbol{q}')$ tali che $\gamma(\boldsymbol{p}', \boldsymbol{q}') = 0$. Ma questo contraddice l'unicità dell'equilibrio. □

Si noti che ci sono giochi a somma zero con equilibri di Nash multipli, anche se il guadagno atteso in tutti gli equilibri deve corrispondere al valore del gioco. Per esempio, nel gioco qui sotto $\boldsymbol{p}^* = (1, 0)$ e $\boldsymbol{q}^* = (q, 1-q)$ per $q \in [0, 1]$ sono tutti equilibri di Nash con guadagno atteso pari a 1,

	L	R
U	1	1
D	0	1