Giochi per Reti Profonde

David Vencato

Università di Pisa

10 Ottobre 2023

Definizione 1

Un gioco simultaneo a una mossa è un gioco nel quale ci sono:

1. *un insieme di N giocatori;*
2. *un insieme Σ_i (finito o infinito) di azioni per ogni giocatore $i = 1, \dots, N$. Si definisce inoltre l'insieme delle azioni congiunte $\Sigma = \prod_{i=1}^{N} \Sigma_i$;*
3. *una funzione di utilità per ogni giocatore $u_i : \Sigma \longrightarrow \mathbb{R}$.*

In un gioco simultaneo a una mossa ogni giocatore deve fare la propria scelta senza conoscere quella degli altri giocatori.

Definizione 2

Il "supervised learning" o apprendimento supervisionato è una classe di problemi dove si ha un insieme di dati $\{(x_t, y_t)\}_{t=1}^{T} \in \mathcal{X} \times \mathcal{Y}$ *e si vuole "imparare" una funzione predittrice* $h : \mathcal{X} \longrightarrow \mathcal{Y}$ *(chiamato anche predittore).*

Ipotesi sul modello:

1. $\mathcal{X} = \mathbb{R}^m$ e $\mathcal{Y} = \mathbb{R}^n$;
2. lineare: il predittore è del tipo $h(x) = \phi(\theta x)$ dove θ è una matrice $n \times m$ che rappresenta i parametri "allenabili" del modello e $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ la funzione output;

Definizione 3

Il problema di apprendimento a un livello (OLP) *si basa su una funzione perdita* $l : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ *che è convessa e differenziabile nel primo argomento.*
Siano $l_t(z) = l(z, y_t)$ *e* $L_t(\theta) = l_t(\theta x_t)$.
Per allenare il modello si deve minimizzare la funzione $L(\theta) = T^{-1} \sum_{t=1}^{T} L_t(\theta)$.

Definizione 4

Il gioco di apprendimento a un livello (OLG) è un gioco simultaneo a una mossa con le seguenti caratteristiche:

1. *ci sono due giocatori: un protagonista p e un antagonista a;*
2. *p sceglie la matrice $\theta \in \mathbb{R}^m \times \mathbb{R}^n$; a sceglie $\{a_t, b_t\}_{t=1}^T$ con $a_t \in \mathbb{R}^n$ e $b_t \in \mathbb{R}$ tali che $a_t^T z + b_t \leq l_t(z)$ $\forall z \in \mathbb{R}^n$;*
3. *data un'azione congiunta $(\theta, \{a_t, b_t\}_{t=1}^T)$, la funzione di utilità dell'antagonista è $U^a = T^{-1} \sum_{t=1}^T a_t^T \theta x_t + b_t$ mentre quella del protagonista è $U_p = -U_a$.*

Caratteristiche del gioco:

1. somma zero con azioni continue;
2. se $\tilde{\sigma}^p = \theta$ denota l'azione del protagonista e $\tilde{\sigma}^a = \{a_t, b_t\}_{t=1}^T$ quella dell'antagonista, allora l'azione congiunta $(\tilde{\sigma^p}, \tilde{\sigma^a})$ è un equilibrio di Nash se $U_p(\tilde{\sigma}^p, \tilde{\sigma}^a) \geq U_p(\sigma^p, \tilde{\sigma}^a)\ \forall \sigma^p$ e $U_a(\tilde{\sigma}^p, \tilde{\sigma}^a) \geq U_p(\tilde{\sigma^p}, \sigma^a)\ \forall \sigma^a$.

Lemma 1

L'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^{T})$ *è un equilibrio di Nash per OLG se e soltanto se* $l_t(\theta x_t) = a_t^T \theta x_t + b_t$, $a_t = \nabla l_t(g)|_{g=\theta x_t}$ *(miglior risposta dell'antagonista) e* $T^{-1} \sum_{t=1}^{T} a_t x_t^T = 0$ *(miglior risposta del protagonista).*

Dimostrazione Lemma

Vediamo $\Longrightarrow$ (l'altro senso è analogo):

- Per ipotesi $a_t^T z + b_t \leq l_t(z)\ \forall z \in \mathbb{R}^n$ dunque $l_t(\theta x_t) = a_t^T \theta x_t + b_t$ è l'utilità massima per a. l_t è convessa e differenziabile allora $\exists!\ h(g)$ affine che è uguale a l_t in un punto e minore o uguale altrove, cioè $h(g) = \nabla l_t(\theta x_t)(g - \theta x_t) + l_t(\theta x_t)$, $a_t = \nabla l_t(g)|_{g=\theta x_t}$.
- $U^p = -T^{-1} \sum_{t=1}^T a_t^T \theta x_t + b_t$. Facendo $\nabla_\theta U^p = -T^{-1} \sum_{t=1}^T a_t x_t^T = 0$ otteniamo il massimo di U^p. $\square$

Teorema 1

Se $(\theta^*, \{a_t, b_t\}_{t=1}^T)$ *è un equilibrio di Nash di OLG, allora* θ^* *è un punto di minimo globale di OLP. Viceversa, se* θ^* *è un punto di minimo globale di OLP, allora esiste una strategia dell'antagonista* $\{a_t, b_t\}_{t=1}^T$ *tale che* $(\theta^*, \{a_t, b_t\}_{t=1}^T)$ *è un equilibrio di Nash di OLG.*

Dimostrazione Teorema

$\Longrightarrow$: Per quanto visto nel lemma , si deve avere $L(\theta^*) = T^{-1}\sum_{t=1}^{T} a_t^T \theta^* x_t + b_t$, $a_t = \nabla l_t(g)|_{g=\theta^* x_t}$ e $T^{-1}\sum_{t=1}^{T} a_t x_t^T = 0$. Ultime due equazioni:

$$T^{-1}\sum_{t=1}^{T} \nabla l_t(g)|_{g=\theta^* x_t} x_t^T = 0.$$

Ma il membro di sinistra è $\nabla L(\theta^*)$ e dunque $\nabla L(\theta^*) = 0$. L convessa e differenziabile, allora θ^* minimo globale.

Dimostrazione Teorema

$\Longleftarrow$: L'antagonista per fare la risposta migliore a θ^* deve scegliere $a_t := \nabla l_t(g)|_{g=\theta^* x_t}$ e $b_t = l_t(\theta^* x_t) - a_t^T \theta^* x_t$. Allora, dato che θ^* è minimo globale:

$$0 = \nabla L(\theta^*) = T^{-1} \sum_{t=1}^{T} \nabla L_t(\theta^*) = \\ T^{-1} \sum_{t=1}^{T} (\nabla l_t(g)|_{g=\theta^* x_t}) x_t^T = T^{-1} \sum_{t=1}^{T} a_t x_t^T .$$

Per il lemma, anche θ^* è la risposta migliore e allora si ha un equilibrio di Nash. $\square$

Fino ad ora abbiamo ignorato la complessità del modello. Per questo introduciamo un vincolo $\theta \in \Theta$ per un qualche insieme convesso Θ.

Fino ad ora abbiamo ignorato la complessità del modello. Per questo introduciamo un vincolo $\theta \in \Theta$ per un qualche insieme convesso Θ.

Definizione 5

Si definisce il problema di apprendimento a un livello vincolato (OCP) un OLP al quale si aggiunge un vincolo di ottimizzazione $\theta \in \Theta$.

Definizione 6

Si definisce il gioco di apprendimento a un livello vincolato (OCG) un OLG al quale si aggiunge un vincolo di ottimizzazione $\theta \in \Theta$.

Supponiamo che Θ sia un politopo cioè un'intersezione finita di semispazi. Allora, Θ è esprimibile con un insieme finito J di funzioni affini, tale per cui $\theta \in \Theta$ se e soltanto se $j(\theta) \leq 0 \ \forall j \in J$.

Supponiamo che Θ sia un politopo cioè un'intersezione finita di semispazi. Allora, Θ è esprimibile con un insieme finito J di funzioni affini, tale per cui $\theta \in \Theta$ se e soltanto se $j(\theta) \leq 0 \ \forall j \in J$.

Lemma 2

L è convessa e differenziabile, dunque condizioni necessari e sufficienti (condizioni KKT) affinché $\theta^* \in \operatorname{argmin}_{\theta\in\Theta} L(\theta)$ è che esista $\{\mu_j\}_{j\in J}$ tale che $\forall j \in J$:

$$\begin{cases} \mu_j \geq 0 \\ \mu_j j(\theta^*) = 0 \\ j(\theta^*) \leq 0 \\ \sum_{j\in J} \mu_j \nabla j(\theta^*) = -\nabla L(\theta^*) \end{cases}$$

Lemma 3

L'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T)$ *è un equilibrio di Nash per OCG se e soltanto se* $l_t(\theta x_t) = a_t^T \theta x_t + b_t$, $a_t = \nabla l_t(g)|_{g=\theta x_t}$ *(miglior risposta dell'antagonista), e esiste* $\{\mu_j\}_{j \in J}$ *tale che* $\forall j \in J$:

$$\begin{cases} \mu_j \geq 0 \\ \mu_j j(\theta) = 0 \\ j(\theta) \leq 0 \\ \sum_{j \in J} \mu_j \nabla j(\theta) = -T^{-1} \sum_{t=1}^T a_t x_t^T \text{(miglior risposta protagonista)} \end{cases}$$

Teorema 2

Se $(\theta^*, \{a_t, b_t\}_{t=1}^T)$ *è un equilibrio di Nash di OCG, allora* θ^* *è un punto di minimo globale con restrizione di OCP. Viceversa, se* θ^* *è un punto di minimo globale con restrizione di OCP, allora esiste una strategia dell'antagonista* $\{a_t, b_t\}_{t=1}^T$ *tale che* $(\theta^*, \{a_t, b_t\}_{t=1}^T)$ *è un equilibrio di Nash di OCG.*

Definizione 7

Un grafo G *è una coppia* (V, E) *dove* V *è l'insieme dei vertici ed* $E \subset V \times V$ *quello degli archi. Se, inoltre, gli archi sono orientati allora si parla di grafo diretto. (Ben definiti vertice iniziale e finale).*

Definizione 7

Un grafo G è una coppia (V, E) *dove V è l'insieme dei vertici ed* $E \subset V \times V$ *quello degli archi. Se, inoltre, gli archi sono orientati allora si parla di grafo diretto. (Ben definiti vertice iniziale e finale).*

Definizione 8

Sia G un grafo diretto. Allora un percorso in G è una sequenza di archi $(e_1, .., e_k)$ *tale che il vertice finale di* e_i *è anche il vertice iniziale di* e_{i+1} $\forall i = 1, .., k-1$*. Se vale anche che il vertice finale di* e_k *coincide col vertice iniziale di* e_1 *allora si parla di ciclo. Un grafo diretto senza cicli è detto aciclico.*

Definizione 9

Una rete neurale feed-forward è una quintupla $N = (V, E, I, O, F)$ *dove* (V, E) *è un grafo diretto aciclico,* $I = \{i_1, \ldots, i_m\} \subset V$ *è l'insieme dei vertici di input,* $O = \{o_1, \ldots, o_n\} \subset V$ *è l'insieme dei vertici di output e* $F = \{f_v : \mathbb{R} \longrightarrow \mathbb{R} \mid v \in V\}$ *è l'insieme delle funzioni di attivazione. I parametri allenabili sono dati da* $\theta : E \longrightarrow \mathbb{R}$.

Ipotesi aggiuntive:

- *non ci sono archi entranti nei vertici di input;*
- *non ci sono archi uscenti dai vertici di output.*

Considerazioni:

- un grafo diretto aciclico induce un ordine parziale $\leq$ sui vertici: $u \leq v$ se e soltanto se esiste un percorso da u a v; $\forall v \in V$, $E_v := \{(u, u') \in E : u' = v\}$
- La rete neurale è "collegata" al set dei dati di allenamento assumendo che $|I| = m$ (il numero dei vertici di input corrisponde al numero di caratteristiche degli input dell'apprendimento supervisionato) e $|O| = n$ (il numero di vertici di output è uguale alla dimensione degli output del supervised learning).

Funzionamento della rete neurale

Sia $x_t \in R^m$ un valore input di allenamento. La rete neurale feed-forward lavora creando funzioni c_t che assegnano valori a ogni vertice nel seguente modo:

$$c_t(i_k, \theta) = f_{i_k}((x_t)_k), \qquad i_k \in I$$

$$c_t(v, \theta) = f_v\left(\sum_{u:(u,v)\in E} c_t(u, \theta)\theta(u, v)\right), \qquad v \in V - I.$$

Denotiamo con $c_t(o, \theta)$ il vettore dei valori dei vertici di output (i.e. $(c_t(o, \theta))_k := c_t(o_k, \theta)$ con $o_k \in O$).

Per imporre vincoli a θ lo facciamo in questo modo: $\forall v \in V - I$ i parametri θ ristretti a E_v devono stare in un insieme $\Theta_v \subset \mathbb{R}^{E_v}$, ponendo $\Theta := \prod_{v \in V-I} \Theta_v$.

Per imporre vincoli a θ lo facciamo in questo modo: $\forall v \in V - I$ i parametri θ ristretti a E_v devono stare in un insieme $\Theta_v \subset \mathbb{R}^{E_v}$, ponendo $\Theta := \prod_{v \in V-I} \Theta_v$.

Definizione 10

Data una funzione di perdita $l(z, y)$ che è convessa nel primo argomento e soddisfa $0 \leq l(z, y) \leq \infty$ $\forall z \in R^n$, si definisce $l_t(z) = l(z, y_t)$ e $L_t(\theta) = l_t(c_t(o, \theta))$. Il "problema di allenamento" è trovare $\theta \in \Theta$ che minimizza $L(\theta) = T^{-1} \sum_{t=1}^{T} L_t(\theta)$ (si parla di "Deep Learning Problem").

Deep Learning Game

Vogliamo definire un gioco simultaneo a una mossa che richiami il DLP. Chiamiamo questo gioco "Deep learning game" (DLG). Definiamo i giocatori, l'insieme delle azioni e le funzioni di utilità:

Deep Learning Game

Vogliamo definire un gioco simultaneo a una mossa che richiami il DLP. Chiamiamo questo gioco "Deep learning game" (DLG). Definiamo i giocatori, l'insieme delle azioni e le funzioni di utilità:

- *Giocatori*: per ogni vertice $v \in V - I$ c'è un protagonista p; per ogni vertice $v \in V$ c'è uno zanni s_v (un agente che opera per il proprio interesse); un antagonista a.

Deep Learning Game

Vogliamo definire un gioco simultaneo a una mossa che richiami il DLP. Chiamiamo questo gioco "Deep learning game" (DLG). Definiamo i giocatori, l'insieme delle azioni e le funzioni di utilità:

- *Giocatori*: per ogni vertice $v \in V - I$ c'è un protagonista p; per ogni vertice $v \in V$ c'è uno zanni s_v (un agente che opera per il proprio interesse); un antagonista a.
- *Azioni*: il protagonista nel vertice v sceglie una funzione parametro $\theta_v \in \Theta_v$. L'antagonista sceglie un insieme di vettori T e scalari $\{a_t, b_t\}_{t=1}^{T}$ con $a_t \in \mathbb{R}^n$ e $b_t \in \mathbb{R}$ tali che $a_t^T z + b_t \leq l_t(z)$ $\forall z \in \mathbb{R}^n$. Allo stesso modo, ogni zanni s_v sceglie un insieme di $2T$ scalari $\{q_{vt}, d_{vt}\}_{t=1}^{T}$ con $q_{vt} \in \mathbb{R}$ e $d_{vt} \in \mathbb{R}$ tali che $q_{vt} z + d_{vt} \leq f_v(z)$ $\forall z \in \mathbb{R}$.

Deep Learning Game

- *Funzioni di utilità*: consideriamo un'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$.
  Utilità per gli zanni: t-esimo elemento del training set è:
  $$U_{it}^s(\sigma) = d_{it} + q_{it} x_{it}, \qquad i \in I$$
  $$U_{vt}^s(\sigma) = d_{vt} + q_{vt} \sum\nolimits_{u:(u,v)\in E} U_{tu}^s(\sigma)\theta(u, v), \qquad v \in V - I.$$
  Dunque, fissato $v \in V$, la funzione utilità totale per lo zanni s_v è data da $U_v^s(\sigma) = \sum_{t=1}^T U_{vt}^s(\sigma)$.
  Utilità per l'antagonista: $U^a = T^{-1} \sum_{t=1}^T U_t^a$ dove $U_t^a(\sigma) = b_t + \sum_{k=1}^n a_{kt} U_{o_k t}^s(\sigma)$.
  Utilità per i protagonisti: $U^p(\sigma) = -U^a(\sigma)$.

Lemma 4

Data un'azione dei protagonisti θ, esiste un'unica azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^{T}, \{q_{vt}, d_{vt}\})$ nella quale gli zanni e l'antagonista giocano le loro migliori risposte a θ (σ è la scelta congiunta espansa per θ).
Inoltre, $U_p(\sigma) = -L(\theta)$, $\nabla_\theta U^p(\sigma) = -\nabla L(\theta)$, e dato un protagonista in $v \in V - I$, se teniamo le scelte di ogni altro agente fisse, $U^p(\sigma)$ è una funzione affine rispetto alla strategia del protagonista in v.

Lemma 4

Data un'azione dei protagonisti θ, esiste un'unica azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$ nella quale gli zanni e l'antagonista giocano le loro migliori risposte a θ (σ è la scelta congiunta espansa per θ).
Inoltre, $U_p(\sigma) = -L(\theta)$, $\nabla_\theta U^p(\sigma) = -\nabla L(\theta)$, e dato un protagonista in $v \in V - I$, se teniamo le scelte di ogni altro agente fisse, $U^p(\sigma)$ è una funzione affine rispetto alla strategia del protagonista in v.

Teorema 3

L'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$ è un equilibrio di Nash del DLG se e soltanto se è un'espansione per θ e θ è un punto KKT del DLP.

Generalizziamo: consideriamo una partizione P di $V - I$ tale che per ogni $\rho \in P$, se $u, v \in \rho$ e $u \leq v$ allora $v \leq u$.
Sia $E_\rho = \bigcup_{v \in \rho} E_v$, $\Theta_\rho \subset \mathbb{R}^{E_\rho}$ e $\Theta = \prod_{\rho \in P} \Theta_\rho$.
Lasciamo invariati lo spazio delle azioni degli zanni e dell'antagonista, ma ora ogni protagonista controlla una parte del nodo ρ.
Se la partizione è quella dei singoletti, si torna al caso base.

Definizione 11

Sia $u : \mathbb{R}^{E_\rho} \to \mathbb{R}$ *una funzione affine. Trovare l'insieme* $\operatorname{argmax}_{\theta_\rho \in \Theta_\rho} u(\theta_\rho)$ *è chiamato problema parziale in* $\rho \in P$.

Definizione 11

Sia $u : \mathbb{R}^{E_\rho} \to \mathbb{R}$ una funzione affine. Trovare l'insieme $\operatorname{argmax}_{\theta_\rho \in \Theta_\rho} u(\theta_\rho)$ è chiamato problema parziale in $\rho \in P$.

Per ogni $\rho \in P$, siano $H_\rho \subset \mathbb{R}^{\mathbb{R}^{E_\rho}}$ e $J_\rho \subset \mathbb{R}^{\mathbb{R}^{E_\rho}}$ insiemi finiti di funzioni differenziabili continue. Allora, possiamo porre che Θ_ρ sia l'insieme di tutti $\theta_\rho \in \mathbb{R}^{E_\rho}$ tale che per ogni $h \in H_\rho$, $h(\theta_\rho) = 0$, e per ogni $j \in J_\rho$, $j(\theta_\rho) \leq 0$.

Definizione 12

Diciamo che θ_ρ è un punto KKT per il problema parziale in $\rho \in P$ se $\theta_\rho \in \Theta_\rho$ e esistono moltiplicatori $\mu_j \geq 0$ e $\lambda_h \in \mathbb{R}$ tali che:

$$\nabla u(\theta_\rho) = \sum_{j \in J_\rho} \mu_j \nabla j(\theta_\rho) + \sum_{h \in H_\rho} \lambda_h \nabla h(\theta_\rho)$$

$$\mu_j j(\theta_\rho) = 0 \text{ per ogni } j \in J_\rho$$

Definizione 13

Se per ogni $\rho \in P$, gli elementi di H_ρ e J_ρ sono affini allora si parla di restrizione parziale affine. Se, invece, per ogni $\rho \in P$, gli elementi di H_ρ sono affini e quelli di J_ρ sono convessi allora si parla di restrizione parziale di Slater.

Definizione 13

Se per ogni $\rho \in P$, gli elementi di H_ρ e J_ρ sono affini allora si parla di restrizione parziale affine. Se, invece, per ogni $\rho \in P$, gli elementi di H_ρ sono affini e quelli di J_ρ sono convessi allora si parla di restrizione parziale di Slater.

Teorema 4

Mettiamoci nelle ipotesi della restrizione parziale affine o di Slater. Allora un punto è di minimo globale se e soltanto se è un punto KKT.

Definizione 14

Sia $\sigma = (\theta, \{a_t, b_t\}_{t=1}^{T}, \{q_{vt}, d_{vt}\}_{t=1}^{T})$ *un'azione congiunta e* $v \in V$. *Se* f_v *è convessa e differenziabile, si dice che lo zanni di* v *è ragionevole per* σ *se* $\forall t \in \{1, ..., T\}$, *vale che:*

$$\begin{cases} q_{vt} = f_v'(\sum_{u:(u,v)\in E} c_t(u,\theta)\theta(u,v)) \\ f_v(\sum_{u:(u,v)\in E} c_t(u,\theta)\theta(u,v)) = d_{vt} + q_{vt}(\sum_{u:(u,v)\in E} c_t(u,\theta)\theta(u,v)). \end{cases}$$

Definizione 14

Sia $\sigma = (\theta, \{a_t, b_t\}_{t=1}^{T}, \{q_{vt}, d_{vt}\}_{t=1}^{T})$ *un'azione congiunta e* $v \in V$. *Se* f_v *è convessa e differenziabile, si dice che lo zanni di* v *è ragionevole per* σ *se* $\forall t \in \{1, ..., T\}$, *vale che:*

$$\begin{cases} q_{vt} = f_v'(\sum_{u:(u,v)\in E} c_t(u,\theta)\theta(u,v)) \\ f_v(\sum_{u:(u,v)\in E} c_t(u,\theta)\theta(u,v)) = d_{vt} + q_{vt}(\sum_{u:(u,v)\in E} c_t(u,\theta)\theta(u,v)). \end{cases}$$

Definizione 15

In maniera analoga, se la funzione perdita l *è convessa e differenziabile nel primo termine, allora l'antagonista è ragionevole se* $\forall t \in \{1, ..., T\}$*:*

$$\begin{cases} a_t = \nabla l_t(z)|_{z=c_t(o,\theta)} \\ a_t^T c_t(o,\theta) + b_t = l_t(c_t(o,\theta)). \end{cases}$$

Teorema 5

Siano dati un insieme finito S, un ordinamento parziale $\leq$ su S e $X \subset S$. Se $\forall s \in S$ si ha che $\{s' \in S : s' < s\} \subset X \implies s \in X$, allora $X = S$. Questa, in teoria degli insiemi, è chiamata induzione forte su un insieme parzialmente ordinato.

Da questo lemma, riusciamo a dimostrare uno dopo l'altro risultati cruciali riguardanti le azioni e le utilità degli zanni e antagonista. Supponiamo che $\forall v \in V$, f_v sia convessa e differenziabile e la funzione perdita l sia convessa differenziabile nel primo termine. Sia $\leq$ l'ordinamento parziale generato dal grafo diretto aciclico della rete neurale. Fissiamo $v \in V$.

1. Data un'azione congiunta σ dove $\forall u \leq v$, lo zanni di u è ragionevole per σ, allora $U^s_{tv}(\sigma) = c_t(v, \theta)$;
2. Data un'azione congiunta σ dove tutti gli zanni e l'antagonista sono ragionevoli, allora $U^a_t(\sigma) = l_t(c_t(o, \theta))$, per ogni dato di allenamento x_t;
3. data un'azione congiunta σ dove $\forall u \leq v$ (tranne al più v), lo zanni di u è ragionevole per σ, allora l'unica migliore risposta per lo zanni in v è di essere ragionevole.
4. Data un'azione congiunta σ dove tutti gli zanni sono ragionevoli, allora l'unica migliore risposta per l'antagonista è di essere ragionevole.

In queste ipotesi, quindi, gli zanni e l'antagonista agiscono in maniera ragionevole se e soltanto se giocano la loro migliore risposta. Vediamo i protagonisti.

Lemma 5

Assumiamo che $\forall v \in V$, f_v *sia convessa e differenziabile e la funzione perdita* l *sia convessa differenziabile nel primo termine. Data un'azione congiunta* $\sigma = (\theta, \{a_t, b_t\}_{t=1}^{T}, \{q_{vt}, d_{vt}\}_{t=1}^{T})$ *dove tutti gli zanni e l'antagonista sono ragionevoli, allora se* U^p *è l'utilità del protagonista, vale:*

$$\nabla_\theta U^p(\sigma) = -\nabla_\theta L(\theta).$$

Dimostrazione Lemma

1. Mostriamo che $\nabla_\theta U_t^p(\sigma) = -\nabla_\theta l_t(c_t(o,\theta))$.
2. Partiamo vedendo $\frac{\partial U_t^p(\sigma)}{\partial U_{tv}^s(\sigma)} = -\frac{\partial l_t(c_t(o,\theta))}{\partial c_t(v,\theta)}$:
    - per $o \in O$ viene subito dalla definizione di utilità ($U_t^p(\sigma) = -b_t - \sum_{k=1}^n a_{kt}(U_t^s(o_k))$) e ragionevolezza dell'antagonista ($a_{kt} = \frac{\partial l_t(c_t(o,\theta))}{\partial c_t(o_k,\theta)}$).
    - Applichiamo induzione forte non su $\leq$ ma sull'ordine parziale opposto $\sqsubseteq$ dato che siamo partiti dagli output. Usiamo ipotesi induttiva e la ragionevolezza degli zanni.

Dimostrazione Lemma

Ora, consideriamo $(u, v) \in E$, allora:

$$\frac{\partial U_t^p(\sigma)}{\partial \theta(u, v)} = \frac{\partial U_t^p(\sigma)}{\partial U_{tv}^s(\sigma)} q_{tv} U_{tu}^s(\sigma) = -\frac{\partial l_t(c_t(o, \theta))}{\partial c_t(v, \theta)} q_{tv} U_{tu}^s(\sigma),$$

Usando sempre che gli zanni sono ragionevoli, si ha $q_{vt} = f_v'(\sum_{u':(u',v)\in E} c_t(u', \theta)\theta(u', v))$ e $U_{tu}^s(\sigma) = c_t(u, \theta)$:

$$\frac{\partial U_t^p(\sigma)}{\partial \theta(u, v)} = -\frac{\partial l_t(c_t(o, \theta))}{\partial c_t(v, \theta)} f_v'\Big(\sum_{u':(u',v)\in E} c_t(u', \theta)\theta(u', v)\Big) c_t(u, \theta).$$

Dalla definizione di $c_t(v, \theta)$ si ottiene che $\frac{\partial c_t(v,\theta)}{\partial \theta(u,v)} = f_v'(\sum_{u':(u',v)\in E} c_t(u', \theta)\theta(u', v))c_t(u, \theta)$:

$$\frac{\partial U_t^p(\sigma)}{\partial \theta(u, v)} = -\frac{\partial l_t(c_t(o, \theta))}{\partial c_t(v, \theta)} \frac{\partial c_t(v, \theta)}{\partial \theta(u, v)} = -\frac{\partial l_t(c_t(o, \theta))}{\partial \theta(u, v)}.$$

Definizione 16

Sia $P(u, v)$ l'insieme di tutti i percorsi da u in v, e per ogni percorso p sia $|p|$ il numero di nodi del percorso.

Definizione 16

Sia $P(u, v)$ l'insieme di tutti i percorsi da u in v, e per ogni percorso p sia $|p|$ il numero di nodi del percorso.

Lemma 6

$$\frac{\partial U^p(\sigma)}{\partial \theta(u, v)} = -\frac{1}{T} \sum_{t=1}^{T} \sum_{k=1}^{n} \sum_{p \in P(v, o_k)} U^s_{tu}(\sigma) q_{t, p_{|p|}} a_{kt} \prod_{j=1}^{|p|-1} \theta\left(p_j, p_{j+1}\right) q_{t, p_j}$$

Definizione 17

Sia $\rho \in P$*, definiamo* $U^{p}_{\rho,\sigma} : \mathbb{R}^{E_\rho} \to \mathbb{R}$ *tale che* $U^{p}_{\rho,\sigma}\left(\theta_\rho\right)$ *è l'utilità del protagonista su* ρ *se decide unilateralmente di giocare* θ_ρ *invece che* σ*.*

Definizione 17

Sia $\rho \in P$, definiamo $U^{p}_{\rho,\sigma} : \mathbb{R}^{E_\rho} \to \mathbb{R}$ tale che $U^{p}_{\rho,\sigma}(\theta_\rho)$ è l'utilità del protagonista su ρ se decide unilateralmente di giocare θ_ρ invece che σ.

Lemma 7

$U^{p}_{\rho,\sigma}$ è una funzione affine.

Dimostrazione Lemma

Sia $\sigma|_\rho : \Theta_\rho \to \Sigma$ t.c. $\sigma|_\rho(\tilde{\theta})$ è lo stesso di σ eccetto l'azione del protagonista su ρ è rimpiazzata da $\tilde{\theta}$. Vale quindi $\forall (u, v) \in E_\rho$:

$$U^p_{\rho,\sigma}(\tilde{\theta}) = U^p\left(\sigma|_\rho(\tilde{\theta})\right) \implies \frac{\partial U^p_{\rho,\sigma}(\tilde{\theta})}{\partial \tilde{\theta}_{(u,v)}} = \frac{\partial U^p\left(\sigma|_\rho(\tilde{\theta})\right)}{\partial \tilde{\theta}_{(u,v)}}$$

Dunque:

$$\frac{\partial U^p\left(\sigma|_\rho(\tilde{\theta})\right)}{\partial \tilde{\theta}(u,v)} = -\frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{n}\sum_{p\in P(v,o_k)} U^s_{tu}\left(\sigma|_\rho(\tilde{\theta})\right) q_{t,p_{|p|}}\, a_{kt} \prod_{j=1}^{|p|-1} \theta\left(p_j, p_{j+1}\right) q_{t,p_j}$$

Per ipotesi di partizione, né u né un nodo "antenato" sono in ρ. Allora, $U^s_{tu}\left(\sigma|_\rho(\tilde{\theta})\right) = U^s_{tu}(\sigma)$. Quindi, la derivata parziale è una funzione solo di $\sigma \implies$ derivata parziale costante lungo ogni coordinata $\implies$ affine.

Lemma 8

Data un'azione dei protagonisti θ, esiste un'unica azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^{T}, \{q_{vt}, d_{vt}\})$ (la scelta congiunta espansa) nella quale gli zanni e l'antagonista giocano le loro migliori risposte a θ. Non solo, $U_p(\sigma) = -L(\theta)$, $\nabla_\theta U^p(\sigma) = -\nabla L(\theta)$, e dato un protagonista in $\rho \in P$, se teniamo le scelte di ogni altro agente fisse, $U^p(\sigma)$ è una funzione affine rispetto alla strategia del protagonista in ρ.

Dimostrazione Lemma

- **Scelta congiunta espansa**: Fissiamo azione arbitraria σ_0. Sia $\sqsubseteq$ un'estensione lineare di $\leq$. Così si ha $v_1, \ldots, v_{|V|}$ tali che $v_k \sqsubseteq v_{k+1}$. Usiamo ricorsione: σ_k è uguale a σ_{k-1}, tranne per lo zanni in v_k che gioca la risposta migliore a σ_{k-1}. Usando lemma precedente, per ogni zanni al passo k esiste la risposta ragionevole ed è unica. Al passo $|V|$ abbiamo un'azione congiunta dove tutti gli zanni sono ragionevoli.
  Per lemma esiste ed è unica la risposta ottima per l'antagonista ed è quella nella quale gioca in maniera ragionevole.
- $U_p(\sigma) = -L(\theta)$: gli zanni e l'antagonista sono ragionevoli allora $U_t^a(\sigma) = l_t(c_t(o, \theta))$. Se ne fa la media ricordando che $U^p = -U^a$.

Teorema 6

Assumiamo che per ogni $v \in V$, f_v è convessa e differenziabile e la perdita l è convessa e differenziabile rispetto alla prima componente. Per ogni punto KKT $\theta \in \Theta$, c'è un equilibrio di Nash dove l'azione congiunta dei protagonisti è θ. Viceversa, per ogni equilibrio di Nash dove l'azione congiunta dei protagonisti è θ , allora θ è un punto KKT.

Dimostrazione Teorema

Basta fare la freccia $\Longrightarrow$, l'altra si ottiene percorrendo la dimostrazione in senso opposto.
Scegliamo σ azione espansa di θ (tutti gli zanni e l'antagonista sono ragionevoli, dunque sappiamo già che stanno giocando la loro migliore risposta).
Mostriamo che che per ogni $\rho \in P$, il protagonista in ρ sta giocando la sua risposta migliore, cioè se $U^p(\sigma)$ è vista come una funzione dei valori di θ su $(u, v) \in E_\rho$, allora il θ scelto in σ è un massimo globale.
Strategia: trasportiamo le condizioni KKT del problema completo nelle condizioni KKT per il problema parziale $U_{\rho,\sigma}$.

Dimostrazione Teorema

Ora, le condizioni KKT sulla perdita L implicano che esistono i moltiplicatori KKT $\mu_{j,\rho}$ e $\lambda_{h,\rho}$ tali che:

$$-\nabla L(\theta) = \sum_{\rho \in P} \sum_{j \in J_\rho} \mu_{j,\rho} \nabla j(\theta) + \sum_{\rho \in P} \sum_{h \in H_\rho} \lambda_{h,\rho} \nabla h(\theta)$$

$$\mu_{j,\rho} j(\theta) = 0 \;\; \forall \rho \in P, j \in J_\rho$$

Ricordando che $\nabla_\theta U^p(\sigma) = -\nabla L(\theta)$:

$$\nabla_\theta U^p(\sigma) = \sum_{\rho \in P} \sum_{j \in J_\rho} \mu_{j,\rho} \nabla j(\theta) + \sum_{\rho \in P} \sum_{h \in H_\rho} \lambda_{h,\rho} \nabla h(\theta)$$

$$\mu_{j,\rho} j(\theta) = 0 \;\; \forall \rho \in P, j \in J_\rho$$

Dimostrazione Teorema

Fissiamo $\rho \in P$. Sia $\theta_\rho \in \Theta_\rho$ l'azione del protagonista su ρ in θ. Restringendoci su E_ρ, solo le restrizioni in J_ρ e H_ρ varieranno:

$$\nabla_{\theta_\rho} U^p(\sigma) = \sum_{j \in J_\rho} \mu_{j,\rho} \nabla j(\theta) + \sum_{h \in H_\rho} \lambda_{h,\rho} \nabla h(\theta)$$

$$\mu_{j,\rho} j(\theta) = 0 \ \ \forall j \in J_\rho$$

Ora, ha senso sostituire $U^p(\sigma)$ con $U^p_{\rho,\sigma}$:

$$\nabla_{\theta_\rho} U^p_{\rho,\sigma}(\theta_\rho) = \sum_{j \in J_\rho} \mu_{j,\rho} \nabla j(\theta_\rho) + \sum_{h \in H_\rho} \lambda_{h,\rho} \nabla h(\theta_\rho)$$

$$\mu_{j,\rho} j(\theta_\rho) = 0 \ \ \forall j \in J_\rho$$

Queste sono le condizioni KKT per essere massimo locale ma $U^p_{\rho,\sigma}$ è affine $\implies$ punto di massimo globale $\implies$ ciascun protagonista non può unilateralmente migliorare su σ.

Bibliografia

[1] Dale Schuurmans e Martin Zinkevich. *Deep Learning Games*, 2016.

[2] S.Boyd e L. Vandenberghe. *Convex Optimization*. Cambridge U. Press, 2004.

[3] N. Cesa-Bianchi e G. Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.

[4] W. Karush. *Minima of functions of several variables with inequalities as side constraints*. Master’ s thesis, Univ. of Chicago, Chicago, Illinois, 1939.

[5] H. Kuhn e A. Tucker. Nonlinear programming. *In Proceedings of 2nd Berkeley Symposium*, pages 481–492. University of California Press, 1951.

Grazie per l'attenzione!