Giochi per Reti Profonde

David Vencato

Settembre 2023

Sommario

In questa relazione vogliamo collegare la teoria dei giochi con il concetto di apprendimento supervisionato. In particolare, per problemi convessi a un livello, dimostriamo un'equivalenza tra i punti di minimo globali del problema di allenamento e gli equilibri di Nash in un gioco costruito ad hoc. In seguito, generalizziamo il gioco per una rete neurale aciclica. Sotto opportune ipotesi, anche in questo caso, riusciamo a trovare una bigezione tra gli equilibri di Nash del gioco e i punti KKT del problema di apprendimento.

1 Gioco di apprendimento a un livello

Definizione 1.1. 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 $i = 1, ..., N$. Si definisce inoltre l'insieme delle azioni congiunte $\Sigma = \prod_{i \in \mathbb{N}} \Sigma_i$;
3. una funzione di utilità per ogni giocatore $u_i : \Sigma \longrightarrow \mathbb{R}$.

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

Definizione 1.2. 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*).

Nel corso della trattazione si assumerà $\mathcal{X} = \mathbb{R}^m$ e $\mathcal{Y} = \mathbb{R}^n$. Nel modello lineare standard dell'apprendimento supervisionato 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 1.3. 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; definiamo $l_t(z) = l(z, y_t)$ e $L_t(\theta) = l_t(\theta x_t)$. Per allenare il modello si deve minimizzare la funzione $L(\theta) = T^{-1} \sum_{t=1}^T L_t(\theta)$.

Si costruisce quindi un gioco simultaneo a una mossa basato su OLP come segue:

Definizione 1.4. 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. il protagonista sceglie la matrice $\theta \in \mathbb{R}^m \times \mathbb{R}^n$ mentre l'antagonista sceglie un insieme di T vettori e T scalari $\{a_t, b_t\}_{t=1}^T$ con $a_t \in \mathbb{R}^n$ e $b_n \in \mathbb{R}$ tali che $a_t^T z + b_t \leq l_t(z)\ \forall z \in \mathbb{R}^n$. Essendo un gioco simultaneo, la scelta di un giocatore viene fatta senza sapere quella dell'altro;
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 $U_p = -U_a$.

Si nota subito che il gioco descritto è a somma zero e con azioni continue. Inoltre, ricordiamo che si ha un *equilibrio di Nash* quando l'azione congiunta dei giocatori è tale che per nessuno dei due è conveniente cambiare la propria posizione. Formalmente, 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 $U_a(\tilde{\sigma}^p, \tilde{\sigma}^a) \geq U_p(\tilde{\sigma^p}, \sigma^a)\ \forall \sigma^a$.

Il passo successivo della trattazione è dimostrare che c'è effettivamente una bigezione tra gli equilibri di Nash di OLG e i punti di minimo globale di OLP. Per fare questo abbiamo bisogno di un lemma preliminare:

Lemma 1.5. *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. Basta provare che un equilibrio di Nash deve soddisfare le condizioni del lemma, il viceversa è analogo. Per ipotesi $a_t^T z + b_t \leq l_t(z)\ \forall z \in \mathbb{R}^n$ dunque $l_t(\theta x_t) = a_t^T \theta x_t + b_t$ rappresenta l'utilità massima per l'antagonista. Inoltre, dato che l_t è convessa e differenziabile, esiste un'unica funzione affine che è uguale a l_t in un punto e minore o uguale altrove. Questa funzione è proprio $h(g) = \nabla l_t(\theta x_t)(g - \theta x_t) + l_t(\theta x_t)$. Allora, $a_t = \nabla l_t(g)|_{g=\theta x_t}$.
Scriviamo esplicitamente la funzione di utilità del protagonista $U^p = -T^{-1}\sum_{t=1}^T a_t^T \theta x_t + b_t$. Facendo il gradiente rispetto a θ e ponendo uguale a zero otteniamo il massimo di U^p e dunque p sta giocando la sua risposta migliore.

□

Teorema 1.6. *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. $\Longrightarrow$: Per quanto visto nel Lemma 1.5, 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 $T^{-1}\sum_{t=1}^T a_t x_t^T = 0$. In particolare, mettendo insieme le ultime due equazioni si ha:

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

Notiamo che il membro di sinistra dell'equazione è proprio $\nabla L(\theta^*)$ e dunque $\nabla L(\theta^*) = 0$. Per definizione di $L(\theta)$ (convessa e differenziabile perché lo é l_t), allora θ^* deve essere un punto di minimo globale ($\nabla L(\theta^*) = 0$ condizione necessaria e sufficiente perché θ^* sia minimo globale).

$\Longleftarrow$: 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, sfruttando che θ^* è minimo globale si ha:

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

Dunque, sempre il Lemma 1.5, anche θ^* è la risposta migliore e dunque si ha un equilibrio di Nash.

□

Fino ad ora abbiamo ignorato la complessità del modello, che, però, è un aspetto importante del problema, al pari di cercare di minimizzare la funzione perdita. Per questo introduciamo un vincolo $\theta \in \Theta$ per un qualche insieme convesso Θ. Possiamo, dunque, introdurre le rispettive definizioni con questa nuova ipotesi.

Definizione 1.7. 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 1.8. Si definisce il *gioco di apprendimento a un livello vincolato* (OCG) un OLG al quale si aggiunge un vincolo di ottimizzazione $\theta \in \Theta$.

In maniera del tutto analoga al caso non vincolato, vediamo che otteniamo la stessa bigezione adattata a questo contesto. 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 1.9. *Dato che l_t nella definizione di L è convessa e differenziabile, allora L è convessa e differenziabile, dunque condizioni necessari e sufficienti (chiamate condizioni KKT) affinché $\theta^* \in argmin_{\theta\in\Theta}L(\theta)$ 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 1.10. *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 esiste $\{\mu_j\}_{j\in J}$ tale che $\forall j \in J$:*

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

Dimostrazione. Per quanto riguarda l'antagonista, la dimostrazione è la stessa del Lemma 1.5 dato che non c'è nessuna nuova restrizione sulle scelte di a. Ora, dato che la funzione di utilità del protagonista è una funzione affine rispetto a θ, un punto che soddisfa le condizioni KKT è un punto di massimo globale e viceversa. Dunque applicando il Lemma 1.9 al gradiente rispetto a θ di U^p si ottiene la tesi.

□

Teorema 1.11. *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.*

Dimostrazione. $\Longrightarrow$ Usando i risultati per l'antagonista del Lemma 1.10, otteniamo che $\nabla L(\theta^*) = T^{-1} \sum_{t=1}^T \nabla L_t(\theta^*) = T^{-1} \sum_{t=1}^T (\nabla l_t(g)|_{g=\theta^* x_t}) x_t^T = T^{-1} \sum_{t=1}^T a_t x_t^T = -\nabla U^p$. Allo stesso tempo, devono essere verificate le condizioni KKT nel Lemma 1.10 e dunque, usando in particolare l'ultima equazione del sistema, si ha: $\sum_{j \in J} \mu_j \nabla j(\theta) = -\nabla L(\theta^*)$. Questa è proprio la condizione del Lemma 1.9 il quale ci assicura che θ^* è un punto di minimo globale con restrizione di OCP.
$\Longleftarrow$ Costruiamo la risposta dell'antagonista come nel Teorema 1.6. Questo implica sempre $\nabla L(\theta^*) = -\nabla U^p$ e concludiamo come nel punto precedente usando i Lemmi 1.9 e 1.10 i quali garantiscono che anche θ^* è la risposta ottimale del protagonista.

□

2 Gioco di apprendimento per reti neurali

Definizione 2.1. 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.*

Notiamo che in un grafo diretto ha senso parlare, fissato un arco, di vertice iniziale e vertice finale.

Definizione 2.2. 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, inoltre, 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 2.3. Una *rete neurale feed-forward* è una quintupla $N = (V, E, I, O, F)$ dove (V, E) è un grafo diretto aciclico, $I = \{i_1, ..., i_m\} \subset V$ è l'insieme dei *vertici di input,* $O = \{o_1, ..., o_n\} \subset V$ 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 $\theta : E \longrightarrow \mathbb{R}$. Supponiamo inoltre che i vertici di input non siano vertici finali di nessun arco (i.e. non ci sono archi entranti nei vertici di input) e analogamente che i vertici di output non siano vertici iniziali di nessun arco (i.e. non ci sono archi uscenti dai vertici di output).

Notiamo che 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. Dunque, $v = u$ se e soltanto se esiste un percorso da v in u e viceversa. Definiamo, quindi, $\forall v \in V$ l'insieme $E_v = \{(u, u') \in E : u' = v\}$. 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).
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:

$$c_t(i_k, \theta) = f_{i_k}((x_t)_k), \quad i_k \in I$$
$$c_t(v, \theta) = f_v(\textstyle\sum_{u:(u,v)\in E} c_t(u, \theta)\theta(u, v)), \quad v \in V - I.$$

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 $o_k \in O$). Si noti che data la differenziabilità di f_v, lo è anche $c_t(o, \theta)$ rispetto a θ.
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 $\Theta := \prod_{v \in V-I} \Theta_v$. Se $\Theta = \mathbb{R}^E$, consideriamo la rete neurale senza restrizioni. Analogamente, se Θ è limitata , allora diremo che la rete neurale è limitata.

Definizione 2.4. 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").

Definizione 2.5. Vogliamo definire un gioco simultaneo a una mossa che richiami il DLP. Chiamiamo questo gioco "Deep learning game" (DLG). Per fare questo dobbiamo specificare chi sono i giocatori, l'insieme delle azioni e le funzioni di utilità:

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

2. *Azioni*: il protagonista nel vertice v sceglie una funzione parametro $\theta_v \in \Theta_v$. L'antagonista sceglie un insieme di T vettori e T 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 $q_{vt}z + d_{vt} \leq f_v(z)$ $\forall z \in \mathbb{R}$. Essendo un gioco simultaneo, tutti i giocatori fanno le proprie scelte senza sapere quelle degli altri.

3. *Funzioni di utilità*: consideriamo un'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$. Le funzioni degli zanni sono definite ricorsivamente, in particolare la funzione utilità per lo zanni s_i o s_v sul t-esimo elemento del training set è:

$$U^s_{it}(\sigma) = d_{it} + q_{it}x_{it}, \qquad i \in I$$
$$U^s_{vt}(\sigma) = d_{vt} + q_{vt} \textstyle\sum_{u:(u,v)\in E} U^s_{tu}(\sigma)\theta(u, v), \qquad v \in V - I.$$

Dunque, fissato $v \in V$, la funzione utilità totale per lo zanni s_v è data da $U^s_v(\sigma) = \sum_{t=1}^T U^s_{vt}(\sigma)$. La funzione utilità dell'antagonista è, invece, definita come $U^a = T^{-1}\sum_{t=1}^T U^a_t$ dove $U^a_t(\sigma) = b_t + \sum_{k=1}^n a_{kt}U^s_{o_k t}(\sigma)$. In maniera analoga ai casi precedenti, le funzioni utilità dei protagonisti sono tutte uguali ponendo $U^p(\sigma) = -U^a(\sigma)$.

A questo punto, cerchiamo di capire come gli zanni e l'antagonista devono comportarsi in risposta a una fissata scelta dei protagonisti e vediamo, anche in questo caso, come si legano gli equilibri di Nash del DLG e i punti KKT del DLP (Sezione 3). Per ora enunceremo solo i risultati, usando le prossime due sezioni per dimostrare passo per passo le seguenti proposizioni.

Lemma 2.6. *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. Non solo, $U^p(\sigma) = -L(\theta)$, $\nabla_\theta U^p(\sigma) = -\nabla L(\theta)$, 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. Chiamiamo σ la scelta congiunta espansa per θ.*

Teorema 2.7. *L'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$ è un equilibrio di Nash di DLG se e soltanto se è un'espansione per θ (nel senso del Lemma 2.6) e θ è un punto KKT di DLP.*

Corollario 2.8. *Se la rete neurale è illimitata, l'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$ è un equilibrio di Nash di DLG se e soltanto se è un'espansione per θ e θ è un punto critico di DLP.*

3 Problema parziale e punti KKT

In questa sezione daremo una serie di risultati legati alle condizioni KKT e alle reti neurali, la maggior parte dei quali senza dimostrazione essendo risultati standard di teoria dell'ottimizzazione e funzioni affini. Riprendendo, dunque, il concetto di rete neurale, consideriamo una partizione P di $V - I$, dove per ogni $\rho \in P$ definiamo $E_\rho = \bigcup_{v\in\rho} E_v$ (ricordiamo che $E_v = \{(u, u') \in E : u' = v\}$). Sia, inoltre, $\Theta_\rho \subset \mathbb{R}^{E_\rho}$ e $\Theta = \prod_{\rho\in P}\Theta_\rho$. Un'importante restrizione è che $\forall\rho \in P$, $\forall u, v \in \rho$ se $u \leq v$ allora $v \leq u$. Lasciamo invariati lo spazio delle azioni degli zanni e dell'antagonista, ma ora ogni protagonista controlla una parte del nodo ρ. Notiamo che è una generalizzazione di quanto detto fino ad ora dato che possiamo prendere come partizione quella data da tutti i singoletti.

Definizione 3.1. Sia $u : \mathbb{R}^{E_\rho} \to \mathbb{R}$ una funzione affine. Trovare l'insieme $argmax_{\theta_\rho \in \Theta_\rho} u(\theta_\rho)$ 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 $j \in J_\rho$, $j(\theta_\rho) \leq 0$.

Ci sono due particolare restrizioni, spesso utilizzate, che sono:

Definizione 3.2. 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 3.3. 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:

$$\begin{aligned} \nabla u(\theta_\rho) &= \sum_{j \in J_\rho} \mu_j \nabla j(\theta_\rho) + \sum_{h \in H_\rho} \lambda_h \nabla h(\theta_\rho) \\ \mu_j j(\theta_\rho) &= 0 \text{ per ogni } j \in J_\rho \end{aligned}$$

Teorema 3.4. *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.*

Osservazione 3.5. Per il problema parziale stiamo assumendo di avere una funzione di utilità affine (altrimenti un punto KKT non è necessariamente un punto di minimo globale). Non faremo la stessa assunzione per il problema globale.

Nel problema di deep learning, vogliamo trovare $\theta^* \in \Theta$ tale che per ogni $\theta \in \Theta, L(\theta^*) \leq L(\theta)$. Questo minimo globale non deve necessariamente esistere, e nemmeno essere unico. Prendiamo come distanza tra funzioni su $\mathbb{R}^E$, dove per ogni $\theta, \bar{\theta} \in \Theta, d(\theta, \bar{\theta}) = \left(\sum_{e \in E}(\theta(e) - \bar{\theta}(e))^2\right)^{1/2}$. Definiamo un intorno di θ $N \subseteq \Theta$ se esiste un $r > 0$ tale che per ogni $\bar{\theta} \in \Theta, d(\theta, \bar{\theta}) < r$. Dunque, θ è un minimo locale se esiste un intorno N di θ tale che per ogni $\bar{\theta} \in N, L(\theta) \leq L(\bar{\theta})$. Evidentemente, essere minimo globale implica essere minimo locale.

Definizione 3.6. Diciamo che $\theta \in \mathbb{R}^E$ è un punto KKT se $\theta \in \Theta$ e esistono moltiplicatori $\mu_j \geq 0$ e $\lambda_h \in \mathbb{R}$ tali che:

$$\begin{aligned} -\nabla L(\theta) &= \sum_{j \in J} \mu_j \nabla j(\theta) + \sum_{h \in H} \lambda_h \nabla h(\theta) \\ \mu_j j(\theta) &= 0 \text{ per ogni } j \in J. \end{aligned}$$

4 Approfondimento teorico per DLG

Restiamo sempre nel contesto del deep learning e introduciamo nuovi concetti che ci permetteranno di dimostrare i risultati della Sezione 2.

Definizione 4.1. Sia $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$ 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:

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

In altre parole, stiamo richiedendo che i valori e le derivate delle funzioni f_v e $d_{vt}+q_{vt}x$ coincidano per le energie di attivazione presenti nel grafo. In maniera analoga, se la funzione perdita l è convessa e differenziabile nel primo termine, allora l'antagonista è *ragionevole* se $\forall t \in \{1, ..., T\}$:

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

Proposizione 4.2. *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.*

Lemma 4.3. *Assumiamo che $\forall v \in V$, f_v sia convessa e differenziabile. Sia $\leq$ l'ordinamento parziale generato dal grafo diretto aciclico della rete neurale. Fissiamo $v \in V$. Data un'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$ dove $\forall u \leq v$, lo zanni di u è ragionevole per σ, allora $U_{tv}^s(\sigma) = c_t(v,\theta)$.*

Dimostrazione. Sia $v \in V$. Usando l'induzione forte bisogna provare che se $\forall u < v$ $U_{tu}^s(\sigma) = c_t(u,\theta)$ allora $U_{tv}^s(\sigma) = c_t(v,\theta)$. Dunque:

- *Passo base*: sia $v \in I$ allora per definizione $U_{tv}^s(\sigma) = d_{vt} + q_{vt}x_{vt}$ ma lo zanni è ragionevole per ipotesi allora $d_{vt} + q_{vt}x_{vt} = f_v(x_{tv}) = c_t(v,\theta)$.

- *Passo induttivo*: sia $v \in V - I$. Per definizione $U_{vt}^s(\sigma) = d_{vt} + q_{vt}\sum_{u:(u,v)\in E} U_{tu}^s(\sigma)\theta(u,v)$; ma se $(u,v) \in E$ allora $u < v$ dunque, per ipotesi induttiva, $U_{tu}^s(\sigma) = c_t(u,\theta)$. Allora, si ha $U_{vt}^s(\sigma) = d_{vt} + q_{vt}\sum_{u:(u,v)\in E} c_t(u,\theta)\theta(u,v)$. D'altra parte, $v \leq v$ e dunque lo zanni in v è ragionevole cioè $d_{vt} + q_{vt}\sum_{u:(u,v)\in E} c_t(u,\theta)\theta(u,v) = f_v(\sum_{u:(u,v)\in E} c_t(u,\theta)\theta(u,v)) = c_t(v,\theta)$.

□

Lemma 4.4. *Assumiamo che $\forall v \in V$, f_v sia convessa e differenziabile e la funzione perdita l sia convessa e differenziabile nel primo termine. Data un'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$ dove tutti gli zanni e l'antagonista sono ragionevoli, allora $U_t^a(\sigma) = l_t(c_t(o,\theta))$, per ogni dato di allenamento x_t.*

Dimostrazione. La dimostrazione è analoga al Lemma 4.3. □

Lemma 4.5. *Assumiamo che $\forall v \in V$, f_v sia convessa e differenziabile. Sia $\leq$ l'ordinamento parziale generato dal grafo diretto aciclico della rete neurale. Fissiamo $v \in V$. Data un'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$ 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.*

Dimostrazione. Fissiamo uno specifico dato di allenamento x_t (dato che lo zanni può cambiare parametri $\forall t$ non è restrittivo farlo). Poniamo $z = x_{tv}$ se $v \in I$ oppure $z = \sum_{u:(u,v)\in E} U_{tu}^s(\sigma)\theta(u, v)$. Per definizione, l'utilità dello zanni in v è $d_{vt} + q_{vt}z$. Vediamo chi è la migliore risposta per lo zanni. Poniamo $q_{tv} = f_v'(z)$ e $d_{tv} = f_v(z) - f_v'(z)z$. La scelta è legale in quanto f_v è convessa e differenziabile dunque $f_v(x) \geq f_v'(z)(x-z)+f_v(z)$ e quindi $f_v(x) \geq d_{vt} + q_{vt}x$. È anche risposta che massimizza l'utilità ($f_v(z) = d_{vt} + q_{vt}z$ $(\star)$) e è unica sempre per ipotesi di convessità e differenziabilità di f_v. Infine, se $v \in I$, $z = x_{tv}$ e la risposta è ragionevole per definizione; se $v \in V - I$ per il Lemma 4.3 si ha $\forall u : (u, v) \in E$ che $U_{tu}^s(\sigma) = c_t(u, \theta)$ e dunque $z = \sum_{u:(u,v)\in E} c_t(u, \theta)\theta(u, v)$. Sostituendo in $(\star)$ si ottiene che lo zanni agisce in maniera ragionevole. □

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

Lemma 4.6. *Assumiamo che $\forall v \in V$, f_v sia convessa e differenziabile e la funzione perdita l sia convessa e differenziabile nel primo termine. Data un'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$ dove tutti gli zanni sono ragionevoli, allora l'unica migliore risposta per l'antagonista è di essere ragionevole.*

Dimostrazione. La dimostrazione è analoga al Lemma 4.5. □

Dunque, le scelte ragionevoli degli zanni e dell'antagonista sono semplici da determinare, analizziamo ora la ragionevolezza dei protagonisti.

Lemma 4.7. *Assumiamo che $\forall v \in V$, f_v sia convessa e differenziabile e la funzione perdita l sia convessa e differenziabile nel primo termine. Data un'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$ 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. Per prima cosa spezziamo U^p nella sua componente relativa a x_t:

$$U_t^p(\sigma) = -U_t^a(\sigma) = -b_t - \textstyle\sum_{k=1}^n a_{kt}(U_t^s(o_k)).$$

Per linearità, la tesi diventa dimostrare che $\nabla_\theta U_t^p(\sigma) = -\nabla_\theta l_t(c_t(o, \theta))$.
Dunque, considerando l'uguaglianza iniziale:

$$\frac{\partial U_t^p(\sigma)}{\partial U_t^s(o_k)} = -a_{kt}. \tag{1}$$

Usando semplicemente la definizione di ragionevolezza dell'antagonista, si ha che $a_{kt} = \frac{\partial l_t(c_t(o,\theta))}{\partial c_t(o_k,\theta)}$. Dunque la (1) diventa:

$$\frac{\partial U_t^p(\sigma)}{\partial U_t^s(o_k)} = -\frac{\partial l_t(c_t(o,\theta))}{\partial c_t(o_k,\theta)}. \tag{2}$$

Definiamo $X \subset V$ l'insieme di tutti i $v \in V$ tali che $\frac{\partial U_t^p(\sigma)}{\partial U_{tv}^s(\sigma)} = -\frac{\partial l_t(c_t(o,\theta))}{\partial c_t(v,\theta)}$. Abbiamo dimostrato che $O \subset X$. Dunque vorremmo applicare l'induzione forte sull'ordine parziale $\leq$ indotto dal grafico diretto aciclico della rete neurale. Notiamo, però, che adesso siamo partiti dagli output e non dagli input, dunque occorre considerare l'ordine parziale opposto $\sqsubseteq$. Sia $v \in V - O$, vogliamo mostrare che $v \in X$, assumendo che $\forall u \in V$ con $u \sqsubset v$ allora $u \in X$. Dunque, ricordiamo che $U_{vt}^s(\sigma) = d_{vt} + q_{vt} \sum_{u:(u,v)\in E} U_{tu}^s(\sigma)\theta(u,v)$. Dovendo però muoverci dagli output verso gli input, se $u \sqsubset v$ l'equazione utile diventa $U_{ut}^s(\sigma) = d_{ut} + q_{ut} \sum_{w:(w,u)\in E} U_{tw}^s(\sigma)\theta(w,u) = d_{ut} + q_{ut} \sum_{w\neq v:(w,u)\in E}(U_{tw}^s(\sigma)\theta(w,u)) + q_{ut}U_{tv}^s(\sigma)\theta(v,u)$. Allora $\frac{\partial U_{ut}^s(\sigma)}{\partial U_{tv}^s(\sigma)} = q_{ut}\theta(v,u)$. Si ottiene che:

$$\frac{\partial U_t^p(\sigma)}{\partial U_{tv}^s(\sigma)} = \sum_{u:(v,u)\in E} q_{ut}\theta(v,u)\frac{\partial U_t^p(\sigma)}{\partial U_{tu}^s(\sigma)} \tag{3}$$

Allora, per ipotesi induttiva:

$$\frac{\partial U_t^p(\sigma)}{\partial U_{tv}^s(\sigma)} = -\sum_{u:(v,u)\in E} q_{ut}\theta(v,u)\frac{\partial l_t(c_t(o,\theta))}{\partial c_t(u,\theta)} \tag{4}$$

Dato che gli zanni sono ragionevoli $\forall u \in V$ si ha $q_{ut} = f_v'(\sum_{u':(u',u)\in E} c_t(u',\theta)\theta(u',u))$. Quindi:

$$\frac{\partial U_t^p(\sigma)}{\partial U_{tv}^s(\sigma)} = -\sum_{u:(v,u)\in E} f_v'\Big(\sum_{u':(u',u)\in E} c_t(u',\theta)\theta(u',u))\theta(v,u)\Big)\frac{\partial l_t(c_t(o,\theta))}{\partial c_t(u,\theta)} \tag{5}$$

$$\frac{\partial U_t^p(\sigma)}{\partial U_{tv}^s(\sigma)} = -\frac{\partial l_t(c_t(o,\theta))}{\partial c_t(v,\theta)}. \tag{6}$$

Siamo quindi giunti a provare il claim iniziale, cioè $X = V$. Ora, consideriamo $(u,v) \in E$, allora:

$$\frac{\partial U_t^p(\sigma)}{\partial \theta(u,v)} = \frac{\partial U_t^p(\sigma)}{\partial U_{tv}^s(\sigma)} q_{tv}U_{tu}^s(\sigma) = -\frac{\partial l_t(c_t(o,\theta))}{\partial c_t(v,\theta)} q_{tv}U_{tu}^s(\sigma), \tag{7}$$

dove il primo uguale segue dalla definizione e il secondo per quanto dimostrato prima. 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 $U_{tu}^s(\sigma) = c_t(u,\theta)$, sostituendo in (7):

$$\frac{\partial U_t^p(\sigma)}{\partial \theta(u,v)} = -\frac{\partial l_t(c_t(o,\theta))}{\partial c_t(v,\theta)} f_v'\Big(\sum_{u':(u',v)\in E} c_t(u',\theta)\theta(u',v)\Big)c_t(u,\theta). \tag{8}$$

Dalla definizione di $c_t(v,\theta)$ si ottiene che $\frac{\partial c_t(v,\theta)}{\partial \theta(u,v)} = f'_v(\sum_{u':(u',v)\in E} c_t(u',\theta)\theta(u',v))c_t(u,\theta)$. E quindi sostituendo in (8):

$$\frac{\partial U_t^p(\sigma)}{\partial \theta(u,v)} = -\frac{\partial l_t(c_t(o,\theta))}{\partial c_t(v,\theta)}\frac{\partial c_t(v,\theta)}{\partial \theta(u,v)} = -\frac{\partial l_t(c_t(o,\theta))}{\partial \theta(u,v)}. \tag{9}$$

□

Fissiamo ora un po' di notazione. Per una rete neurale, definiamo $P(u,v)$ l'insieme di tutti i percorsi da u in v, e per ogni percorso p sia $|p|$ il numero di nodi del percorso. In particolare, si ha che se $p \in P(u,v)$ allora $p_1 = u$ e $p_{|p|} = v$.

Lemma 4.8.

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

Dimostrazione. Diamo solo un'idea della dimostrazione. Dalla definizione di $U_p(\sigma)$ segue che:

$$\frac{\partial U^p(\sigma)}{\partial \theta(u,v)} = \sum_{t=1}^{T} U_{tu}^s(\sigma) q_{t,v} \frac{\partial U^p(\sigma)}{\partial U_{tv}^s(\sigma)}$$

Dunque, il problema si riduce a dimostrare ricorsivamente, partendo da O:

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

□

Questo lemma tecnico ci permette di dimostrare qualcosa molto interessante. Sia $\rho \in P$, definiamo $U_{\rho,\sigma}^p : \mathbb{R}^{E_\rho} \to \mathbb{R}$ tale che $U_{\rho,\sigma}^p(\theta_\rho)$ è l'utilità del protagonista su ρ se decide unilateralmente di giocare θ_ρ invece che σ.

Lemma 4.9. *$U_{\rho,\sigma}^p$ è una funzione affine.*

Dimostrazione. Fissiamo $\sigma = \left(\theta, \{a_t, b_t\}_{t=1}^T, \{q_{v,t}, d_{v,t}\}\right)$. Definiamo $\sigma|_\rho : \Theta_\rho \to \Sigma$ tale che per ogni $\tilde{\theta} \in \Theta_\rho$, $\sigma|_\rho(\tilde{\theta})$ è lo stesso di σ eccetto l'azione del protagonista su ρ è rimpiazzata da $\tilde{\theta}$. Vale quindi:

$$U_{\rho,\mathbf{a}}^p(\tilde{\theta}) = U^p\left(\sigma|_\rho(\tilde{\theta})\right)$$

Segue dunque dalla definizione che per ogni $(u,v) \in E_\rho$:

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

Usando il Lemma 4.8:

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

Quindi, la funzione $U^p_{\rho,\sigma}$ è differenziabile dovunque. Non solo, consideriamo $U^s_{tu}\left(\sigma|_\rho(\tilde{\theta})\right)$. Preso $v\in\rho$, U^s_{tu} è l'output del nodo u. Quindi, dato che per ogni u' tali che $u'\in\rho\backslash\{v\}$ si ha $u'\not\leq v$, allora né u né un nodo "antenato" è in ρ. Dunque, U^s_{tu} non cambia al variare di θ_ρ. Formalmente, $U^s_{tu}\left(\sigma|_\rho(\tilde{\theta})\right) = U^s_{tu}(\sigma)$. Allora:

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

Abbiamo mostrato che la derivata parziale è una funzione solo di σ, e non di $\tilde{\theta}$. Ma una funzione con derivata parziale costante lungo ogni coordinata è affine.

□

Ora, mostriamo un lemma più forte del Lemma 2.6.

Lemma 4.10. *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. Per quanto visto nel corso della sezione, tutti gli agenti stanno giocando ragionevolmente. Dobbiamo dunque costruire l'azione congiunta mostrando che esiste ed è unica. Fissiamo $\theta\in\Theta$ e un'arbitraria azione congiunta $\sigma_0 = (\theta, \{a_t,b_t\}_{t=1}^T, \{q_{vt},d_{vt}\})$. Per prima cosa, dato l'ordinamento parziale $\leq$ su V, consideriamo $\sqsubseteq$ un'estensione lineare di $\leq$. Questo ci permette di ordinare tutti i vertici $v_1,...,v_{|V|}$ tali che $v_k\sqsubseteq v_{k+1}$. Definiamo ricorsivamente σ_k così: è uguale a σ_{k-1}, tranne per lo zanni in v_k che gioca la risposta migliore a σ_{k-1}. Ora, usando il Lemma 4.5, per ricorsione, sappiamo che per ogni zanni al passo k esiste la risposta ragionevole ed è unica. Di conseguenza al passo $|V|$ abbiamo un'azione congiunta dove tutti gli zanni sono ragionevoli. Allora, per 4.6 esiste ed è unica la risposta ottima per l'antagonista ed è quella nella quale gioca in maniera ragionevole. Dunque, esiste ed è unica l'azione congiunta σ^* nella quale tutti gli zanni e l'antagonista giocano la loro migliore risposta a θ.
Mostriamo le altre proprietà enunciate nel lemma. Dato che gli zanni e l'antagonista sono ragionevoli per il Lemma 4.4 si ha $U^a_t(\sigma) = l_t(c_t(o,\theta))$ e dunque facendone la media su t e ricordando che $U^p = -U^a$ si ha l'uguaglianza cercata. Poi, dal Lemma 4.7 abbiamo subito che $\nabla_\theta U^p(\sigma) = -\nabla L(\theta)$ e dal Lemma 4.9 da ogni protagonista abbiamo una funzione affine se decide (solo lui) di cambiare scelta.

□

Evidentemente, questo lemma implica il Lemma 2.6 poiché basta considerare come partizione P quella data dai singoli nodi.

Lemma 4.11. *Assumiamo che per ogni $v \in V$, f_v è convessa e differenziabile, la funzione perdita l è convessa e differenziabile nella prima componente, e, data un'azione congiunta $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$, tutti gli zanni e l'antagonista sono ragionevoli. Se θ è un punto KKT, allora le azioni dei protagonisti sono le migliori risposte a σ, e σ è un equilibrio di Nash.*

Dimostrazione. Per provare che è un equilibrio di Nash, dobbiamo mostrare che per ogni $\rho \in P$, il protagonista in ρ sta giocando la sua risposta migliore (tutti gli zanni e l'antagonista sono ragionevoli, dunque sappiamo già che stanno giocando la loro migliore risposta). Formalmente, se vediamo $U^p(\sigma)$ come una funzione dei valori di θ su $(u, v) \in E_\rho$, allora il θ scelto in σ è un massimo globale. Faremo questo in due step:

1. Trasportiamo le condizioni KKT del problema completo nelle condizioni KKT per il problema parziale $U_{\rho,\sigma}$ (la funzione di utilità per il protagonista in v deviata).
2. Dato che $U_{\rho,\sigma}$ è affine, le condizioni KKT per un massimo implicano che il massimo sia globale.

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 per il Teorema 4.10 $\nabla_\theta U^p(\sigma) = -\nabla L(\theta)$, allora:

$$\nabla_\theta U^p(\sigma) = \sum_{\rho \in P} \sum_{j \in J_\rho} \mu_{j,\rho} \nabla j(\theta) + \sum_{\rho \in P} \sum_{h \in H_\rho} \lambda_{h,\rho} \nabla h(\theta)$$
$$\mu_{j,\rho} j(\theta) = 0 \;\; \forall \rho \in P, j \in J_\rho$$

Queste sono le condizioni KKT necessarie perchè θ sia un massimo locale, ma non sono sufficienti. Fissiamo $\rho \in P$. Sia $\theta_\rho \in \Theta_\rho$ l'azione del protagonista su ρ in θ. Adesso, se ci restringiamo su E_ρ, solo le restrizioni in J_ρ e H_ρ varieranno, ottenendo:

$$\nabla_{\theta_\rho} U^p(\sigma) = \sum_{j \in J_\rho} \mu_{j,\rho} \nabla j(\theta) + \sum_{h \in H_\rho} \lambda_{h,\rho} \nabla h(\theta)$$
$$\mu_{j,\rho} j(\theta) = 0 \;\; \forall j \in J_\rho$$

E dunque ha senso sostituire $U^p(\sigma)$ con $U^p_{\rho,\sigma}$. Per la strategia θ_ρ che è parte di θ, si ha:

$$\nabla_{\theta_\rho} U^p_{\rho,\sigma}(\theta_\rho) = \sum_{j \in J_\rho} \mu_{j,\rho} \nabla j(\theta_\rho) + \sum_{h \in H_\rho} \lambda_{h,\rho} \nabla h(\theta_\rho)$$
$$\mu_{j,\rho} j(\theta_\rho) = 0 \;\; \forall j \in J_\rho$$

Queste sono le condizioni KKT per θ_p per essere un massimo locale di $U_{\rho,\sigma}$ in Θ_ρ. E dunque, il protagonista su ρ non può guadagnare cambiando. Ora, per quanto visto, sappiamo che $U^p_{\rho,\sigma}$ è affine, allora se le condizioni KKT sono soddisfatte per un massimo locale, questo punto è un massimo globale. Questo ci dice che ciascun protagonista non può unilateralmente migliorare su σ, e dunque è un equilibrio di Nash.

$\square$

Teorema 4.12. *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. $\Longrightarrow$ Segue immediatamente dal Lemma 4.11.
$\Longleftarrow$ Percorriamo al contrario la dimostrazione del Lemma 4.11. Sia $\sigma = (\theta, \{a_t, b_t\}_{t=1}^T, \{q_{vt}, d_{vt}\})$, vogliamo mostrare che θ è un punto KKT. Per prima cosa osserviamo che in ogni equilibrio di Nash, gli zanni e l'antagonista giocano ragionevolmente (perché stanno giocando la loro miglior risposta). Quindi, σ è l'azione congiunta estesa di θ e dunque $\nabla_\theta U^p(\sigma) = -\nabla L(\theta)$. Poiché l'equilibrio è un valore ottimale per la funzione affine $U^p_{v,\sigma}$, le condizioni KKT devono valere per ogni protagonista. Combinando le condizioni KKT per ogni protagonista, otteniamo le condizioni KKT per massimizzare U^p su Θ. Dato che $\nabla_\theta U^p(\sigma) = -\nabla L(\theta)$ queste condizioni sono quelle per minimizzare $L(\theta)$.

$\square$

Questo dimostra anche il Teorema 2.7.

5 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.