Seminář z kooperativní teorie her

Filip Úradník

9. března 2023

1 Kooperativní hry

Definice 1.1 (Kooperativní hra): *Kooperativní hra* je dvojice (N, v), kde

- $N = \{1, \dots, n\}$ je množina *hráčů*,
- $v : 2^N \to \mathbb{R}$ je *charakteristická funkce*, kde $v(\emptyset) = 0$.

Množina všech her o n hráčích je Γ^n.

Definice 1.2 (Koalice): *Koalice* je množina $C \subseteq N$.

Definice 1.3 (Třídy): Hra (N, v) je

- *monotonní* $\equiv (\forall S \subseteq T \subseteq N)(v(S) \leq v(T))$,
- *superaditivní* $\equiv (\forall S, T \subseteq N : S \cap T = \emptyset)(v(S) + v(T) \leq v(S \cup T))$,
- *konvexní* $\equiv (\forall S, T \subseteq N)(v(S) + v(T) \leq v(S \cap T) + v(S \cup T))$.

Definice 1.4 (Výplatní vektor): Nechť (N, v) je kooperativní hra. *Výplatní vektor* je $x \in \mathbb{R}^n$, kde x_i je výplata hráče i. Vektor je

- *eficientní* $\equiv x(N) := \sum_{i \in N} x_i = v(N)$,
- *individuálně racionální* $\equiv (\forall i)(x_i \geq v(i))$.

S ohledem na vlastnosti vektorů definujeme množiny

- preimputace $\mathcal{I}^*(v) := \{x \in \mathbb{R}^n \mid x(N) = v(N)\}$,
- imputace $\mathcal{I}(v) := \{x \in \mathcal{I}^*(v) \mid (\forall i)(x_i \geq v(i))\}$.

Definice 1.5 (Koncept řešení): Nechť (N, v) je hra. *Koncept řešení* je funkce $\mathcal{S} : \Gamma^n \to 2^{\mathbb{R}^n}$.

Definice 1.6 (Jádro): Nechť (N, v) je kooperativní hra. *Jádro* je více-bodový koncept řešení definován jako

$$\mathcal{C}(v) := \{x \in \mathcal{I}(v) \mid (\forall S \subseteq N)(x(S) \geq v(S))\}$$

Definice 1.7 (Shapleyho hodnota): Nechť (N, v) je kooperativní hra. *Shapleyho hodnota* pro hráče i je jednobodový koncept řešení definován jako

$$\phi_i(v) := \sum_{S \subseteq N \setminus \{i\}} \frac{s!\,(n - s - 1)!}{n!} \left(v(S \cup i) - v(S)\right).$$

2 Neúplné hry

Definice 2.1 (Neúplná hra): *Neúplná hra* je hra, ve které nejsou známy hodnoty všech koalic. Formálně je to trojice $(N, \mathcal{K}, v)$, kde

- $N = \{1, \ldots, n\}$ je množina *hráčů*,
- $\mathcal{K} \subseteq 2^N$ je množina *koalic se známými hodnotami* v, $\emptyset \in \mathcal{K}$,
- $v : \mathcal{K} \to \mathbb{R}$ je *charakteristická funkce*, $v(\emptyset) = 0$.

Neúplná hra je *minimální* $\equiv \mathcal{K} = \{N\} \cup \bigcup_i \{i\}$.

Definice 2.2 (C-extenze): Nechť $C \subseteq \Gamma^n$ je třída kooperativních her o n lidech. Kooperativní hra (N, w) je C-*extenzí* hry $(N, \mathcal{K}, v) \equiv \forall S \in \mathcal{K} : v(S) = w(S)$.

Množina všech C-extenzí hry $(N, \mathcal{K}, v)$ je $C(v)$. Pokud je $C(v)$ neprázdná, pak je hra $(N, \mathcal{K}, v)$ C-*extenzibilní*. Množinu všech C-extenzibilních neúplných her s pevným $\mathcal{K}$ značíme $C(\mathcal{K})$.

Definice 2.3 (Horní/dolní hra C-extenzí): Nechť $C \subseteq \Gamma^n$ je třída kooperativních her, $(N, \mathcal{K}, v)$ je C-extenzibilní neúplná hra. *Horní hra* je hra $(N, \overline{v})$, tž.

$$(\forall w \in C(v))(\forall S \subseteq N)(w(S) \leq \overline{v}(S)),$$

navíc však $(\forall S)(\exists w \in C(v))(w(S) = \overline{v}(S))$. Identicky lze definovat i *dolní hru* $(N, \underline{v})$.

Definice 2.4 (Slabé řešení): Nechť $(N, \mathcal{K}, v)$ je neúplná hra.

- *Slabé řešení* je funkce
$$\cup\mathcal{S}(C, \mathcal{K})(v) := \bigcup_{w \in C(v)} \mathcal{S}(w).$$
- *Silné řešení* je funkce
$$\cap\mathcal{S}(C, \mathcal{K})(v) := \bigcap_{w \in C(v)} \mathcal{S}(w).$$

3 Markov Decision Process

Definice 3.1 (MDP): MDP je $\langle \varphi, \Sigma, \tau, r \rangle$, kde

- $\varphi \subseteq \mathbb{R}^k$ je množina *stavů*,
- Σ je množina *akcí*,
- $\tau : \varphi \times \Sigma \to \varphi$ je *transition function*,
- $r : \varphi \times \varphi \times \Sigma \to \mathbb{R}$ je *reward function*.

Definice 3.2 (Policy): *Policy* je $\pi : \varphi \to \Delta^{|\Omega(s)|}$.

Definice 3.3 (Return): *Return* je $R_T := \sum_{t=1}^{T} r_t$.

Předpokládáme, že za T kroků hra končí. Lze to rozšířit i na nekonečné hry. Chceme najít $\pi \in \operatorname{argmax} \mathbb{E}_\pi R_T$. Nejde o to, aby v každém kroku se zvýšil reward optimálně, ale aby po celkovém hraní celé hry byl reward co největší.

Definice 3.4 (POMDP): POMDP je MDP, kde ale máme observation funkci $o : \varphi \to \mathbb{R}^l$, která nám předá jen částečnou informaci o stavu, kterou můžu pozorovat.

Definice 3.5 (Value funkce): Nechť jsme ve stavu s', do kterého jsme se dostali v čase t'. Chceme odhadnout, jaký return dostaneme při aktuální policy π. Proto definujeme tzv. value funkci jako

$$\mu(s') := \mathbb{E}_\pi \left[\sum_{t=t'}^{T} r_t \right]$$

Algoritmus se bude snažit vylepšovat policy tak, že pro zvolenou akci bude minimalizovat

$$\alpha = (R_T - \mu)\,\frac{\pi_i}{\pi_{i,old}}$$

DEFINICE 3.6 (q-VALUE): Nechť τ je tranistion funkce. Nechť μ je value funkce. Chceme-li ve stavu s udělat akci a, předpokládaná hodnota pak bude tzv. q-value:

$$q(s,a) := \sum_{s'} P(\tau(s,a) = s')\,\mu(s')\,.$$