

4EK421 – zadání úkolu č. 4

Zadání

Uvažujme bimaticovou hru s následujícími výplatními maticemi

$$X_1 = \begin{pmatrix} 2 & 0 & -2 \\ 3 & 1 & -1 \\ 4 & 2 & 0 \end{pmatrix}, \text{ a } X_2 = X_1^T = \begin{pmatrix} 2 & 3 & 4 \\ 0 & 1 & 2 \\ -2 & -1 & 0 \end{pmatrix}$$

Řekněme, že se hra bude hrát opakovaně. Hráči se snaží maximalizovat současnou hodnotu výhry k okamžiku, kdy se hraje nultá hra, přičemž pro diskontování budoucích výher se používá určitý neznámý diskontní faktor $\delta < 1$ (přesněji viz definice ve slidech). Cílem je formulovat předpis, podle kterého má řádkový hráč volit strategii v t -tém opakování hry v závislosti na strategiích v minulých hrách, pro $t = 0, 1, \dots$. Pro označení strategií v předpisech použijeme čísla 0, 1 a 2, kde 0 značí 1. řádek, 1 značí 2. řádek a 2 značí 3. řádek. V předpisech označujeme symbolem a_t strategii řádkového hráče v t -té hře a symbolem b_t strategii sloupcového hráče v t -té hře.

Příklad předpisu, který v prvních 11 hrách (hry jsou indexované od 0) vybere 1. řádek (nebo sloupec) a dále střídá 2. a 3. řádek nebo sloupec:

$$a_t = \begin{cases} 0 & \text{pokud } t \leq 10, \\ 1 & \text{pokud } t > 10 \wedge t \bmod 2 = 1, \\ 2 & \text{pokud } t > 10 \wedge t \bmod 2 = 0, \end{cases}$$

kde \bmod je infixový operátor vracející zbytek po dělení.

Příklad předpisu, který v první hře zahraje 2. řádek nebo sloupec a v každé další kopíruje bezprostředně předcházející strategii protihráče:

$$a_t = \begin{cases} 0 & \text{pokud } t = 0, \\ b_{t-1} & \text{pokud } t > 0. \end{cases}$$

Další příklady předpisů (pro hru s 2×2 strategiemi) naleznete na slidech zde.

Odevzdání a hodnocení

Úkol se odevzdává do příslušné odevzdávací v InSIS do 25. 4. 23:59. Za zaslání úkolu lze získat 5 řádných bodů. Získání řádných bodů není závislé na výplatě, kterou předpis zajistí (tj. body lze získat za jakýkoli předpis, který bude formálně správně). Nicméně, čím lepších výplat bude předpis dosahovat, tím lépe, viz níže.

Mimo řádné body lze získat ještě bonusové body. Na cvičení 26. 4. předpisy porovnáme. Kritériem pro porovnání bude současná hodnota výplat v prvních T hrách, kde T bude před porovnáváním náhodně zvoleno. Výplaty budou diskontovány diskontním faktorem $\delta = 0,99$. Několik nejlepších předpisů (v závislosti na počtu odevzdaných úkolů) získá bonusové body. Definice „nejlepšího“ bude zvolena operativně (může jít např. o počet vyhraných párových srovnání předpisů, o součet výher přes všechna párová srovnání, o součet výher soupeřů, atp.).