Manos
Nguyen Tan Thai Hung 1001986
Zhangsheng Lai 1002554

Algorithmic Game Theory: HW 2

1. Let G be a cost-minimization game which has the function Φ such that

$$C_i(s_i', s_{-i}) < C_i(s)$$
$$\Phi(s_i', s_{-i}) < \Phi(s)$$

As the game is finite, we can use Φ and best-response dynamics such that the cost decreases with every beneficial deviation made by a particular agent until no more beneficial deviation exists for every agent and call this strategy s^*. Then this s^* is a pure Nash since for any agent i and any deviation s_i', we cannot have

$$C_i(s_i', s_{-i}) < C_i(s^*)$$

since it implies $\Phi(s_i', s_{-i}) < \Phi(s^*)$ and contradicts $\Phi(s^*) < \Phi(s)$ for all pure strategies $s \neq s^*$. Thus this game has at least one PNE.

2. (a) Consider the utility maximizing game below starting with the the initial outcome (A_1, B_1), from which best-response dynamics cycles forever, avoiding the pure Nash of (A_3, B_2).

		P1		
		A_1	A_2	A_3
	B_1	$1, 4$	$2, 3$	$0, 0$
P2	B_2	$0, 0$	$0, 0$	$5, 5$
	B_3	$4, 1$	$3, 2$	$0, 0$

(b) Consider the cost minimization game below, where for player P1, weights are decreased whenever A_1 or A_2 is played. Thus the weights are concentrated on A_3 in the long run and hence the time-averaged history of joint play would point to playing A_3 with close to 1 probability. Similar argument goes for P2 and thus average history of joint play will point to A_3, B_3 being the PNE.

		P1		
		A_1	A_2	A_3
	B_1	$1, 1$	$1, 1$	$0, 0$
P2	B_2	$1, 1$	$1, 1$	$0, 0$
	B_3	$0, 0$	$0, 0$	$0, 0$

3. For a fixed t' such that i is the smallest integer such that $t' \leq 2^i$. Then $\epsilon = \sqrt{\frac{\ln n}{t'}} \geq \sqrt{\frac{\ln n}{2^i}}$ and the regret is at most $2\sqrt{2^i \ln n}$ up till time t, i.e.

$$\sum_{t=1}^{t'} \nu^t \leq OPT + 2\sqrt{2^i \ln n}$$

Let $kt' \geq T$ for some integer k, then

$$\sum_{t=1}^{T} \nu^t \leq \sum_{t=1}^{kt'} \nu^t \leq OPT + k \cdot 2\sqrt{2^i \ln n} = OPT + (2\sqrt{k^2 \cdot 2^i \ln n}) \leq OPT + 2\sqrt{k}\sqrt{T \ln n}$$

and hence the expected regret is still $O(\sqrt{\ln n/T})$ with respect to every fixed action.

4. Let $f_\epsilon(x) = (1-\epsilon)^x$ and $g_\epsilon(x) = 1 - \epsilon x$, then

$$\begin{aligned} f_\epsilon(0) &= 1 = g_\epsilon(0) \\ f_\epsilon(1) &= 1 - \epsilon = g_\epsilon(1) \end{aligned}$$

$$\left.\begin{aligned} f'_\epsilon(x) &= (1-\epsilon)^x \ln(1-\epsilon) \\ g'_\epsilon(x) &= -\epsilon \end{aligned}\right\} f'_\epsilon(0) = \ln(1-\epsilon) < -\epsilon = g'_\epsilon(0)$$

The last inequality holds since

$$\ln(1-\epsilon) = -\epsilon - \frac{\epsilon^2}{2!(1-\xi)^2} \quad \text{by Taylor's theorem in Lagrange form}$$

for some $\xi \in (0, \epsilon)$. Also f_ϵ is a convex function as $f''_\epsilon(x) = (1-\epsilon)^x [\ln(1-\epsilon)]^2 > 0$ for $\epsilon \in (0, 1/2]$ on the interval $[0, 1]$. This this proves $f_\epsilon(x) \leq g_\epsilon(x)$.

Let $\hat{f}_\epsilon(x) = (1+\epsilon)^x$, then by Taylor's expansion

$$(1+\epsilon)^x = 1 + \epsilon x + \frac{x(x-1)(1-\xi)^{x-2}}{2!}\epsilon^2$$

for some $\xi \in (0, \epsilon)$. We observe that $\frac{x(x-1)(1-\xi)^{x-2}}{2!}\epsilon^2 \leq 0$ since $x \in [0, 1]$ we have proved that $(1+\epsilon)^x \leq 1 + \epsilon x$.

5. Consider the online decision-making setting where every time step t the adversary chooses a payoff vector $\pi^t : A \to [0, 1]$ where the time-averaged regret is defined as $\frac{1}{T} \max_{a \in A} \sum_{t=1}^{T} \pi^t(a) - \frac{1}{T} \sum_{t=1}^{T} \pi^t(a^t)$. Let $\Gamma^t = \sum_{a \in A} w^t(a)$ and define $OPT = \sum_{t=1}^{T} c^t(a^*)$ as the cumulative cost for the best fixed action a^*. Then

$$\begin{aligned} \Gamma^T &\leq w^T(a^*) \\ &= w^1(a^*) \prod_{t=1}^{T} (1+\epsilon)^{\pi^t(a^*)} = (1+\epsilon)^{OPT} \end{aligned}$$

Te expectec cost of the MW algorithm at time t is

$$\sum_{a \in A} p^t(a).\pi^t(a) = \sum_{a \in A} \frac{w^t(a)}{\Gamma^t} \pi^t(a)$$

We can rewrite Γ^{t+1} in terms of Γ^t in the following manner

$$\begin{aligned}
\Gamma^{t+1} &= \sum_{a\in A} w^{t+1}(a) \\
&= \sum_{a\in A} w^t(a)\cdot(1+\epsilon)^{\pi^t(a)} \\
&\le \sum_{a\in A} w^t(a)\cdot(1+\epsilon\pi^t(a)) \quad \text{by Q4} \\
&\le \Gamma^t \sum_{a\in A} p^t(a)\cdot(1+\epsilon\pi^t(a)) \\
&\le \Gamma^t \sum_{a\in A} p^t(a) + p^t(a)\epsilon\pi^t(a) \\
&\le \Gamma^t(1+\epsilon\nu^t) \quad \text{where } \nu^t \text{ is the expected utility at time } t.
\end{aligned}$$

Combining the results obtained from before,

$$(1+\epsilon)^{OPT} \le \Gamma^T \le \Gamma^1 \prod_{t=1}^{T}(1+\epsilon\nu^t)$$

$$OPT\ln(1+\epsilon) \le \ln n + \sum_{t=1}^{T}\ln(1+\epsilon\nu^t)$$

$$OPT(\epsilon-\epsilon^2) \le OPT\ln(1+\epsilon) \le \ln n + \sum_{t=1}^{T}\ln(1+\epsilon\nu^t) \le \ln n + \sum_{t=1}^{T}\epsilon\nu^t$$

$$OPT(\epsilon-\epsilon^2) \le \ln n + \sum_{t=1}^{T}\epsilon\nu^t$$

$$OPT \le (\ln n)/\epsilon + \sum_{t=1}^{T}\nu^t + \epsilon OPT \le (\ln n)/\epsilon + \epsilon T + \sum_{t=1}^{T}\nu^t$$

Hence equalizing the two error terms, we get $\epsilon = \sqrt{\ln n/T}$, similar to the case of the cost vector. Thus it has regret $O(\sqrt{\ln n/T})$

6. (a) Let $\hat{x}, \hat{y}$ be a mixed Nash equilibrium then,

$$\hat{x}^T A\hat{y} \ge x^T A\hat{y} \quad \text{for all mixed distributions } x \tag{1}$$
$$\hat{x}^T A\hat{y} \le \hat{x}^T Ay \quad \text{for all mixed distributions } y \tag{2}$$

if and only if,

$$\hat{x} \in \arg\max_x \left(x^T A\hat{y}\right) \subseteq \arg\max_x \left(\min_y x^T Ay\right)$$
$$\hat{y} \in \arg\min_y \left(\hat{x}^T Ay\right) \subseteq \arg\min_x \left(\max_x x^T Ay\right)$$

(b) Let x_1, y_1 and x_2, y_2 be the given mixed Nash equilibria of a two-player zero-sum game. Thus by the above result, for $i = 1, 2$

$$x_i \in \arg\max_x \left(\min_y x^T Ay\right)$$
$$y_{3-i} \in \arg\min_x \left(\max_x x^T Ay\right)$$

thus

$$x_i^T A y_{3-i} \geq x^T A y_{3-i} \quad \text{for all mixed distributions } x$$
$$x_i^T A y_{3-i} \leq x_i^T A y \quad \text{for all mixed distributions } y$$

which by definition tells us that (x_i, y_{3-i}) is a mixed Nash equilibrium for $i = 1, 2$.

7. (a) Let (A, B) be a given general bimatrix game with player one and player two having n and m different strategies to choose from their strategy set denoted by $\mathbf{x}$ and $\mathbf{y}$ respectively, thus A and B are $n \times m$ and $m \times n$ matrices respectively. The expected payoffs of player one is then $\mathbf{x}^T A\mathbf{y}$ and the expected payoffs of player two is $\mathbf{y}^T B\mathbf{x}$.

   (b)