9/18: no-regret learning in general games

Learning outcome

same game each iteration

player i $\quad s_i^1 \dots \quad s_i^t \dots \quad \forall i$

loss player i $\quad \ell_t^i(x) = c_i(x, s_{-i}^t)$

no regret for player i

her cost: $\sum_{t=1}^{T} c_i(s^t) \le \min_x \sum_{t=1}^{T} c_i(x, s_{-i}^t)$ + error

vector of strategies choosen at time t

Consider distribution

$\sigma \quad s^1 \dots s^T$ all prob $\frac{1}{T}$

Special case RPS

$$\begin{array}{c|c} & \begin{matrix} R & P & S \end{matrix} \\ \begin{matrix} R \\ P \\ S \end{matrix} & \begin{bmatrix} 0 & -1 & +1 \\ +1 & 0 & -1 \\ -1 & +1 & 0 \end{bmatrix} \end{array}$$

row player play R too much

arrows = resulting dynamic

limit: close to no prob on diagonal

resulting average

	R	P	S	$\bar{x}$
R	0	1/6	1/6	1/3
P	1/6	0	1/6	1/3
S	1/6	1/6	0	1/3
$\bar{y}$	1/3	1/3	1/3	

· marginal distribution for each player

$\bar{x}$ & $\bar{y}$ Nash

but play is correlated

Shapley game

	R	P	S
R	0	1/6	1/6
P	1/6	0	1/6
S	1/6	1/6	0

diagonal ~~p~~ both loose 3 each

1/3 1/3 1/3 is Nash

learning outcome in limit

back to general case

$$\sigma = \{s^1 \dots s^t \dots s^T\} \quad \text{all prob } \frac{1}{T}$$

no regret means

$$\underset{s\sim\sigma}{E}(c_i(s)) \le \underset{s\in\sigma}{E}(c_i(x, s_{-i})) + \text{error}$$

all i & all x

Def: coarse correlated equilibrium

dist. of vector of strategies

$$\underset{s\sim\sigma}{E}(c_i(s)) \le \underset{s\sim\sigma}{E}(c_i(x_i, s_{-i}))$$

all x & all i

Coarse correlated equilibrium is Nash if & only if

$$\sigma = \sigma_1 \times \dots \times \sigma_n$$

for the n player

<u>Note:</u> learners correlate due to shared history.

Correlated equibrium
(2nd lecture)

σ prob dist on strategy vectors s

no-regret condition for player i

$$\mathop{E}_{s \in \sigma}(c_i(s) \mid s_i = x) \leq \mathop{E}_{s \sim \sigma}(c_i(y, s_{-i}) \mid s_i = x)$$

all i all x all y