Sept 13:

learning in 2-person

2-person game

matrix A row & column

e.g

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

reward row player = loss of column player

<u>Nash equilibrium</u>: row, column prob dist x & y such that

Player 1 expected to get:

$$\sum_{i,j} x_i y_j a_{ij} = x^T A y$$

x & y Nash if:

$$x^T A y \geq \bar{x}^T A y \quad \text{all prob } \bar{x}$$

$$x^T A y \leq x^T A \bar{y} \quad \text{all prob } \bar{y}$$

going second to choose strategy
is ~~better~~ no worse

Slight check: no-regret learning

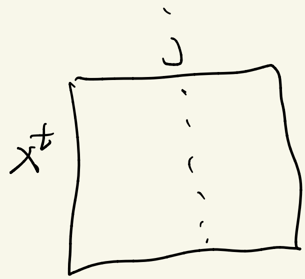
no regret at all

prob vectors played

$x^1 \ldots x^t \qquad y^1 \ldots y^t$

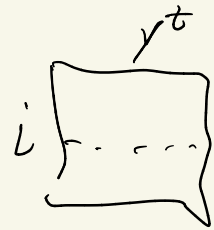
column player's loss:

$$\ell^c_t(j) = \sum_i x_i^t a_{ij}$$



row player's loss

$$\ell^r_t(i) = -\sum_j a_{ij} y_j^t$$



Recall both players have no regret.

loss for column player over T periods (on average)

$$\frac{1}{T}\sum_{t=1}^{T}(x^t)^T A y^t \le \frac{1}{T}\min_j \sum_{t=1}^{T} \ell_t^c(j)$$

best strategy with hind sight

$$= \frac{1}{T}\min_j \sum_t \sum_i x_i^t a_{ij}$$

$$\boxed{\bar{x} = \frac{1}{T}\sum_t x^t}$$

$$= \min_j \sum_i \bar{x}_i a_{ij}$$

---

loss row player (average)

$$-\frac{1}{T}\sum_{t=1}^{T}(x^t)^T A y^t \le \frac{1}{T}\min_i \sum_{t=1}^{T} \ell_t^r(i)$$

$$= \frac{1}{T}\min_i -\sum_t \sum_j a_{ij} y_j^t$$

$$\boxed{\bar{y} = \frac{1}{T}\sum_t y^t}$$

$$= \min_i -\sum_j a_{ij}\bar{y}$$

Made with Goodnotes