Lecture 5

1 Overview

We will discuss the first of two important applications of the expert framework: approximately solving zero sum games.

1.1 Zero sum games

A zero sum game is specified by an $m \times n$ matrix A, the matrix entry A_{ij} is the loss for the row player if the row player plays i and the column player plays j. For this lecture, we assume that the losses belong to $[0, 1]$. The row player can follow a mixed strategy by choosing a row i according to a probability distribution on the rows. The column player can similarly follow a mixed strategy y. The expected loss for the row player is $x^t Ay$ if the players follows mixed strategies (x, y).

If the row player plays first and chooses strategy x, the column player gets to choose an optimal response, the loss for the row player is:

$$C(x) = \max_y x^t Ay$$

If the column player plays first and chooses strategy y the row player gets to choose an optimal response, the loss for the row player is:

$$R(y) = \min_x x^t Ay$$

Von Neumann's minimax theorem states that $\max_y R(y) = \min_x C(x)$, the common value is called the value of the game and is denoted by V.

ϵ-optimal strategies: A pair of strategies (x, y) is optimal if they are best responses to each other. The column player's best response to a non-optimal strategy x has value more than V, we therefore have the inequality,

$$R(y) \leq V \leq C(x)$$

A pair of strategies is ϵ optimal if $C(x) - R(y) \leq \epsilon$, that is no player can gain gain more than ϵ by changing strategies.

2 Experts algorithm for zero sum games

Finding a pair of optimal strategies for zero sum games is equivalent to solving linear programs. Finding ϵ optimal strategies is a non trivial problem as it is like solving linear programs approximately.

We will assume that we can solve the simpler problem of finding the best response to a strategy played by the row player. If the payoff matrix is given explicitly, the best response is found by computing the payoff for every column and choosing the maximum.

$$\overline{C}(x) = \text{argmax}_{j \in [n]} (x^t A)_j$$

Even if the matrix A is implicitly specified and has an exponential number of columns, we will see that for some cases the optimal response can be computed in polynomial time.

2.1 Experts recap

Recall that the experts framework consists of n experts who each make predictions every day, each expert incurs a loss that is revealed at the end of the day.

We analyzed the algorithm where all experts have weight 1 initially, the algorithm chooses expert i with probability proportional to the weight w_i, and weights are updated as $w_i(t+1) = w_i(t)(1-\epsilon)^{\ell_i(t)}$. The expected loss L for this algorithm is close to of the loss L^* of the best expert in rerospect,

$$L \leq (1+\epsilon)L^* + \frac{\ln n}{\epsilon} \tag{1}$$

2.2 Algorithm

The m pure strategies of the row player will be the experts, the experts algorithm picks experts according to a probability distribution, for the game setting choosing an expert probabilistically is equivalent to playing a mixed strategy.

An ϵ approximate solution to a zero sum game can be found by iterating the following steps for $T = \frac{\log n}{\epsilon^2}$ rounds,

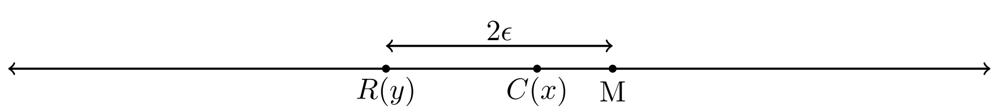
1. The m strategies of the row player are the experts, in round t the row player plays the mixed strategy x_t specified by the experts algorithm. 2. The column player plays $\overline{C}(x_t)$, the optimal response to the row player's mixed strategy.

The algorithm trains the row player by playing against a column player who always plays the best response. By following the experts algorithm, the row player ensures that the average over many rounds is close to an optimal strategy,

THEOREM 1
The average strategies $x = \frac{1}{T}\sum_t x(t)$ and $y = \frac{1}{T}\sum_t y(t)$ are a 2ϵ-optimal pair.

PROOF: Let $\overline{C}(x)$ denote the best response to the strategy x. The average gain over T rounds for the column player would be $C(x)$ if the column player played $\overline{C}(x)$ in each round.

The column player chooses the best strategy in each round, in particular for round t the column player's gain for the chosen strategy $y(t)$ is greater than the gain for $\overline{C}(x)$. The average gain M of the column player over T rounds is therefore greater than $C(x)$, justifying the order of points in the following picture:



For a zero sum game the column player's gain is the row player's loss, so M is also the average loss of the experts algorithm over T rounds.

The best expert in retrospect is the best response to y, so $R(y)$ is the average loss for the best expert. The analysis of the expert's algorithm (1) shows that the total loss MT is not much worse than the loss $R(y)T$ of the best expert in retrospect,

$$MT \leq (1+\epsilon)R(y)T + \frac{\ln n}{\epsilon} \Rightarrow M \leq R(y) + 2\epsilon \tag{2}$$

We used the assumption that the losses lie in $[0,1]$ for the inequality $\epsilon R(y) \leq \epsilon$ and the choice of T for $\frac{\ln n}{\epsilon T} \leq \epsilon$. From the picture it follows that $C(x) - R(y) \leq 2\epsilon$, hence the strategies (x, y) are 2ϵ optimal.

□