EECS598: Prediction and Learning: It's Only a Game Fall 2013

Lecture 6: Lower Bounds and Game Theory I

Prof. Jacob Abernethy *Scribe: Raymundo Navarrete*

6.1 Lower Bounds and Minimax Regret

In the context of prediction with expert advice, the concept of minimax regret is introduced and lower bounds on this quantity are obtained using a randomization argument.

6.1.1 Minimax Regret

Here we introduce the concept of minimax regret.

Consider the problem of prediction with expert advice where a loss function l and the number of experts N have been fixed. By an *algorithm* we are referring to a rule or prescription that at each time t makes a prediction based on previous outcomes at times $1,\ldots,t-1$ and the current expert advice at time t. For a given algorithm and set of experts, the cumulative regret at time T (compared to the best expert) is given by

$$\sum_{t=1}^{T} \mathbf{p}^t \cdot \mathbf{l}^t - \min_{i\in[N]} \sum_{t=1}^{T} l_i^t \,.$$

We would like an algorithm that minimizes the cumulative regret regardless of what the expert advice and outcomes happen to be. That is, we want an algorithm that achieves the *Minimax regret* given by

$$\min_{\text{algorithms}} \max_{\mathbf{l}^1,\ldots,\mathbf{l}^T} \left(\sum_{t=1}^{T} \mathbf{p}^t \cdot \mathbf{l}^t - \min_{i\in[N]} \sum_{t=1}^{T} l_i^t \right) . \tag{6.1}$$

For a given algorithm, the max in the expression finds the maximum cumulative regret over all possible expert advice and outcomes (that is, over all possible losses $\mathbf{l}^1,\ldots,\mathbf{l}^T$). The min in the expression minimizes such maximum cumulative regret over all possible algorithms.

6.1.2 Randomization Argument for Two Experts

Here we obtain a lower bound over the minimax regret in the case of two experts using a randomization argument.

Consider a set of two experts making opposite decisions at random. The loss vector is

$$\hat{l}^t := \langle X^t, 1-X^t \rangle$$

where $X^1,\ldots,X^T$ are random variables defined by

$$X^t := \begin{cases} 0 & \text{w.p. } 1/2 \\ 1 & \text{w.p. } 1/2 \end{cases} .$$

Then

$$\begin{aligned}
\min_{\text{algorithms}} \max_{\mathbf{l}^1,\dots,\mathbf{l}^T}\left(\sum_{t=1}^{T}\mathbf{p}^t\cdot\mathbf{l}^t - \min_{i\in[N]}\sum_{t=1}^{T} l_i^t\right) &\geq \min_{\text{algorithms}} \mathop{\mathbb{E}}_{X^1,\dots,X^T}\left[\sum_{t=1}^{T}\mathbf{p}^t\cdot\hat{\mathbf{l}}^t - \min_{i\in[N]}\sum_{t=1}^{T}\hat{l}_i^t\right] \\
&= \min_{\text{algorithms}} \mathop{\mathbb{E}}_{X^t}\left[\sum_{t=1}^{T}\mathbf{p}^t\cdot \mathop{\mathbb{E}}_{X^t|X^1,\dots,X^{t-1}}\left[\hat{l}^t|\hat{l}^1,\dots,\hat{l}^{t-1}\right] - \min_{i\in[N]}\sum_{t=1}^{T}\hat{l}_i^t\right] \\
&= \min_{\text{algorithms}} \mathop{\mathbb{E}}_{X^t}\left[\sum_{t=1}^{T}\mathbf{p}^t\cdot\langle 1/2,1/2\rangle - \min_{i\in[N]}\sum_{t=1}^{T}\hat{l}_i^t\right] \\
&= T/2 - \mathop{\mathbb{E}}_{X^1,\dots,X^T}\left[\min_i \sum_{t=1}^{T}\langle X^t, 1-X^t\rangle\right] \\
&= \mathop{\mathbb{E}}_{X^1,\dots,X^T}\left[\max_i \sum_{t=1}^{T}\langle 1/2-X^t, X^t-1/2\rangle\right] \\
&= \mathop{\mathbb{E}}_{X^1,\dots,X^T}\left|\sum_{t=1}^{T}(X^t-1/2)\right| \\
&= \frac{1}{2}\mathop{\mathbb{E}}_{Z^1,\dots,Z^T}\left|\sum_{t=1}^{T}Z^t\right| \\
&= \sqrt{\frac{T}{2\pi}} + o(1)\,.
\end{aligned}$$

The first line follows from definition of max. The second line is an application of the tower rule. the third line uses the independence of the losses. The fourth line uses the fact that $\mathbf{p}^t$ is a distribution. The fifth line distributes the term $T/2$ over the sum and reverses min to max. The sixth term uses that the expectancies between the two terms will be the negatives of each other. The seventh line defines the random variables $Z^t = 2(X^t - 1/2)$. The last line uses that $\sum_{t=1}^{T} Z^t$ is a random walk with steps of size ± 1.

6.1.3 Randomization Argument for N Expert

Here we generalize the results for N experts. We assume N is even.

Consider $N/2$ pairs of experts making decisions as before. The loss vector is

$$\hat{\mathbf{l}}^t := \langle X_1^t, 1-X_1^t, \dots, X_{N/2}^t, 1-X_{N/2}^t\rangle$$

where

$$X_i^t := \begin{cases} 0 & \text{w.p. } 1/2 \\ 1 & \text{w.p. } 1/2 \end{cases}.$$

With a similar chain of reasoning as before, the minimax regret is bounded below by

$$\frac{1}{2}\mathop{\mathbb{E}}_{Z^1,\dots,Z^T}\left|\max_{i=1,\dots,N/2}\sum_{t=1}^{T}Z_i^t\right|.$$

It is a fact that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} Z_i^t \to G_i \sim N(0,1)$$

where the convergence is in distribution, G is the binomial distribution, and $N(0,1)$ is the normal distribution with mean 0 and variance 1.

The minimax regret is bounded below by

$$\gtrapprox \frac{1}{2}\mathbb{E}\left[\sqrt{T}\max_i |G_i|\right] \approx \frac{1}{2}\sqrt{T}\sqrt{\log N/2} = O(\sqrt{T\log N})\,.$$

6.2 Game Theory

Game definitions, examples, mixed strategies.

6.2.1 Definitions and Examples

A game is defined by two matrices $A, B \in [0,1]^{n\times m}$. Player 1 makes a choice $i \in [n]$ and player 2 makes a choice $j \in [m]$. Player 1 ears A_{ij} and Player 2 ears B_{ij}.

Example: Prisoners dilemma.

Here two players can each betray the other player and cooperate with the authorities (option 1) or remain loyal to the other player and deny any complicity (option 2). The players do not know in advance what the other player is going to do. The playoff matrices may look as follows

$$A = \begin{pmatrix} -8 & 0 \\ -10 & -1 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} -8 & -10 \\ 0 & -1 \end{pmatrix}.$$

so that if Player 1 betrays and Player 2 denies the crime, then Player 1 gets a payoff of $A_{1,2} = 0$ (that is, no time in jail) and Player 2 gets a payoff of $B_{1,2} = -10$ (that is, 10 years in jail), etc.

It may seem that the best outcome would be for both players to deny the crime and get only 1 years each in jail. But each player can always do better by confessing regardless of what the other player does. If both players think like that, the outcome will be that both players get 8 years in jail. This is the *Nash equilibrium* of this game.

Example: Penalty Kicks.

Here Player 1 is the penalty kicker and Player 2 is the goal keeper. The playoff matrices are as follows

$$A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Since $A + B = 0$, this is called a *zero sum game*.

6.2.2 Mixed Strategy

In the context of game theory, suppose Player 1 plays according to a distribution $x \in \Delta_n$ and Player 2 plays according to a distribution $y \in \Delta_m$.

The expected payoffs to Player 1 and Player 2, respectively, are

$$x^\top A y = \sum_{i,j} x_i y_j A_{ij} \qquad \text{and} \qquad x^\top B y = \sum_{i,j} x_i y_j B_{ij} \, .$$

Given $x \in \Delta_n$, Player 2's best response to x is

$$\arg\max_{y \in \Delta_m} x^\top B y = \arg\max_{j \in [m]} x^\top B e_j \, .$$

Similarly, given $y \in \Delta_m$, Player 1's best response to y is

$$\arg\max_{x \in \Delta_m} x^\top A y = \arg\max_{i \in [n]} e_i^\top B y \, .$$