CSC 665: Homework 4

Chicheng Zhang

December 8, 2019

Please complete the following set of problems. **You are free to discuss with your classmates on your solutions, but only at a high level; if that is the case, please mention your collaborators.** The exercise is due **on Dec 3, 12:30pm, on Gradescope**. You are free to cite existing theorems from the textbooks and course notes.

Problem 1

In this exercise we will prove a special case of von Neumann's minimax theorem using online learning.

Theorem 1 (von Neumann's minimax theorem)**.** *For any matrix $M \in [0,1]^{n \times n}$,*

$$\min_{p \in \Delta^{n-1}} \max_{q \in \Delta^{n-1}} p^\top M q = \max_{q \in \Delta^{n-1}} \min_{p \in \Delta^{n-1}} p^\top M q. \tag{1}$$

1. (Optional) Show that for any function $f(x,y)$ and domains $\mathcal{X}$ and $\mathcal{Y}$, we always have

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} f(x,y) \geq \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} f(x,y), \tag{2}$$

   and use it to conclude that the left hand side is always at least the right hand side in Equation (1).

2. Consider two players R and C (denoting Row and Column respectively) playing a repeated game of T rounds against each other. At time t, R (resp. C) selects a probability distribution of rows $p_t \in \Delta^{n-1}$ (resp. $q_t \in \Delta^{n-1}$). For each player, it is associated with a DTOL game: for R (resp. C), its loss vector at time t is defined as $\ell_{R,t} = M q_t$ (resp. $\ell_{C,t} = (\mathbf{1} - M)^\top p_t$, where $\mathbf{1}$ is the $n \times n$ matrix of all 1's). R and C applies the Hedge algorithm with learning rate $\sqrt{\frac{8 \ln n}{T}}$ on their respective loss vectors.

   (a) Write down the regret guarantees provided by Hedge for both players (your answer should be in terms of M, p_t, q_t's.)

   (b) Define $\bar{p} = \frac{1}{T} \sum_{t=1}^T p_t$ and $\bar{q} = \frac{1}{T} \sum_{t=1}^T q_t$. Show that

$$\max_{q \in \Delta^{n-1}} \bar{p}^\top M q - \min_{p \in \Delta^{n-1}} p^\top M \bar{q} \leq \sqrt{\frac{2 \ln n}{T}}, \tag{3}$$

   and use this to conclude Equation (1).

3. Suppose we have a modified rock-paper-scissor game where the game matrix M is defined as follows:

	R	P	S
R	0.5	0.7	0
P	0.2	0.5	1
S	1	0	0.5

   Write a piece of code that simulates the learning process of both players in item 2, and plot the left hand side of Equation (3) as a function of T, for $T = 10^i$, $i = 1, 2, \ldots, 6$. Use this to experimentally verify the correctness of Equation (3). What are the $\bar{p}$ and $\bar{q}$'s for each T?

Solution

1. Let $x^\star = \arg\min_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} f(x,y)$.

$$\min_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} f(x,y) = \max_{y\in\mathcal{Y}} f(x^\star, y) \geq \max_{y\in\mathcal{Y}} \min_{x\in\mathcal{X}} f(x,y),$$

where the inequality uses the fact that $f(x^\star, y) \geq \min_{x\in\mathcal{X}} f(x,y)$ for all y.

2. (a) The regret of R is defined as:

$$\mathrm{Reg}_R(T) = \sum_{t=1}^{T} \langle p_t, Mq_t \rangle - \min_{i=1}^{n} \sum_{t=1}^{T} \langle e_i, Mq_t \rangle .$$

By the guarantee of Hedge, we have,

$$\sum_{t=1}^{T} \langle p_t, Mq_t \rangle - \min_{i} \sum_{t=1}^{T} \langle e_i, Mq_t \rangle \leq \sqrt{\frac{T \ln n}{2}}. \tag{4}$$

Similarly, the regret of C is defined as:

$$\mathrm{Reg}_R(T) = \sum_{t=1}^{T} \left\langle q_t, (\mathbf{1} - M)^\top p_t \right\rangle - \min_{j=1}^{n} \sum_{t=1}^{T} \left\langle e_j, (\mathbf{1} - M)^\top p_t \right\rangle ,$$

by the guarnatee of Hedge, we have,

$$\sum_{t=1}^{T} \left\langle q_t, (\mathbf{1} - M)^\top p_t \right\rangle - \min_{j=1}^{n} \sum_{t=1}^{T} \left\langle e_j, (\mathbf{1} - M)^\top p_t \right\rangle \leq \sqrt{\frac{T \ln n}{2}}. \tag{5}$$

(b) Writing Equation (4) in matrix form, we have

$$\sum_{t=1}^{T} p_t^\top M q_t - \min_{i=1}^{n} \sum_{t=1}^{T} e_i^\top M q_t \leq \sqrt{\frac{T \ln n}{2}}. \tag{6}$$

Observe that for any probability vector a and b in Δ^{n-1}, $a^\top \mathbf{1} b = 1$. Therefore, Equation (5) implies that,

$$\sum_{t=1}^{T} (1 - \left\langle q_t, -M^\top p_t \right\rangle) - \min_{j=1}^{n} (1 - \left\langle e_j, (\mathbf{1} - M)^\top p_t \right\rangle) \leq \sqrt{\frac{T \ln n}{2}}, \tag{7}$$

which is equivalent to

$$\max_{j=1}^{n} \sum_{t=1}^{T} p_t^\top M e_j - \sum_{t=1}^{T} p_t^\top M q_t \leq \sqrt{\frac{T \ln n}{2}}. \tag{8}$$

Now, summing over Equations (6) and (8), we have

$$\max_{j=1}^{n} (\sum_{t=1}^{T} p_t)^\top M e_j - \min_{i=1}^{n} e_i^\top M (\sum_{t=1}^{T} q_t) \leq \sqrt{2T \ln n},$$

which is equivalent to

$$\max_{q\in\Delta^{n-1}} (\sum_{t=1}^{T} p_t)^\top M q - \min_{p\in\Delta^{n-1}} p^\top M (\sum_{t=1}^{T} q_t) \leq \sqrt{2T \ln n}.$$

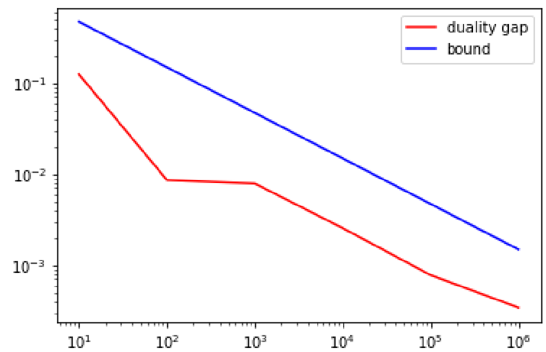

Figure 1: Duality gap and its upper bound against time horizon T.

Dividing both sides by T and using the definitions of $\bar{p}$ and $\bar{q}$, we get Equation (3).
As a consequence,

$$\begin{aligned} & \min_{p\in\Delta^{n-1}} \max_{q\in\Delta^{n-1}} p^\top M\bar{q} \\ \le \quad & \max_{q\in\Delta^{n-1}} \bar{p}^\top Mq \\ \le \quad & \min_{p\in\Delta^{n-1}} p^\top M\bar{q} + \sqrt{\frac{2\ln n}{T}} \\ \le \quad & \max_{q\in\Delta^{n-1}} \min_{p\in\Delta^{n-1}} p^\top M\bar{q} + \sqrt{\frac{2\ln n}{T}} \end{aligned}$$

$$\min_{p\in\Delta^{n-1}} \max_{q\in\Delta^{n-1}} p^\top Mq - \max_{q\in\Delta^{n-1}} \min_{p\in\Delta^{n-1}} p^\top Mq \le \sqrt{\frac{2\ln n}{T}}.$$

Observe that the left hand side does not depend on T. Therefore, let $T \to \infty$, we get

$$\min_{p\in\Delta^{n-1}} \max_{q\in\Delta^{n-1}} p^\top Mq \le \max_{q\in\Delta^{n-1}} \min_{p\in\Delta^{n-1}} p^\top Mq.$$

On the other hand, by Equation (2), we also have:

$$\min_{p\in\Delta^{n-1}} \max_{q\in\Delta^{n-1}} p^\top Mq \ge \max_{q\in\Delta^{n-1}} \min_{p\in\Delta^{n-1}} p^\top Mq.$$

Equation (1) is obtained by combining the above two inequalities.

3. We plot the values of $\max_{q\in\Delta^{n-1}} \bar{p}^\top Mq - \min_{p\in\Delta^{n-1}} p^\top M\bar{q}$ (called "duality gap") and $\sqrt{\frac{2\ln n}{T}}$ (called "bound") against T in Figure 1 (both axes use log scales).

   See Table 1 for the values of $\bar{p}$ and $\bar{q}$ for different T's.

Problem 2 (Optional)

Show that in realizable online classification with a finite hypothesis class $\mathcal{H} \subset (\mathcal{X} \to \{0,1\})$, if at time t, one predicts label 1 with probability $\frac{|V_t^+|}{|V_t|}$ (in other words, $\hat{y}_t = \frac{|V_t^+|}{|V_t|}$), the algorithm has a mistake bound of

T	$\bar{p}$	$\bar{q}$
10^1	$(0.477455, 0.27333962, 0.24920537)$	$(0.45804817, 0.27061733, 0.2713345)$
10^2	$(0.41813745, 0.3839381, 0.19792445)$	$(0.37575671, 0.42357713, 0.20066616)$
10^3	$(0.4192868, 0.37519701, 0.20551619)$	$(0.38338588, 0.41521, 0.20140412)$
10^4	$(0.42258522, 0.37879681, 0.1986179)$	$(0.38673172, 0.41394257, 0.19932571)$
10^5	$(0.41671345, 0.38469113, 0.19859542)$	$(0.38535522, 0.41513073, 0.19951405)$
10^6	$(0.41697589, 0.38462693, 0.19839717)$	$(0.38483394, 0.41655439, 0.19861167)$

Table 1: The average strategies of row and column players versus the number of rounds.

$\ln|\mathcal{H}|$, that is,

$$\sum_{t=1}^{T} |\hat{y}_t - y_t| \leq \ln|\mathcal{H}|.$$

Solution

We will first show the following key inequality: for every round t,

$$|\hat{y}_t - y_t| \leq \ln|V_t| - \ln|V_{t+1}|, \tag{9}$$

where V_{t+1} is the updated version space after iteration t: $V_{t+1} = \left\{h \in V_t : h(x_t) \in y_t\right\}$. We use $V_t^+ = \left\{h \in V_t : h(x_t) = 1\right\}$ and $V_t^- = \left\{h \in V_t : h(x_t) = 0\right\}$ respectively.

By definition, $\hat{y}_t = \frac{|V_t^+|}{|V_t|}$. Let us consider two cases of y_t:

1. $y_t = 1$. In this case, $V_{t+1} = V_t^+$. Consequently, $|\hat{y}_t - y_t| = 1 - \hat{y}_t = 1 - \frac{|V_t^+|}{|V_t|} = 1 - \frac{|V_{t+1}|}{|V_t|}$.

2. $y_t = 0$. In this case, $V_{t+1} = V_t^+$. Consequently, $|\hat{y}_t - y_t| = \hat{y}_t = \frac{|V_t^+|}{|V_t|} = 1 - \frac{|V_{t+1}|}{|V_t|}$.

So what remains is to show that

$$1 - \frac{|V_{t+1}|}{|V_t|} \leq \ln\frac{|V_t|}{|V_{t+1}|}.$$

The above inequality is equivalent to

$$\ln\frac{|V_{t+1}|}{|V_t|} \leq \frac{|V_{t+1}|}{|V_t|} - 1,$$

which is true as $\ln x \leq x - 1$ for all positive x.

Summing over Equation (9) from $t = 1$ to T, we have

$$\sum_{t=1}^{T} |\hat{y}_t - y_t| \leq \ln|V_1| - \ln|V_{T+1}|.$$

The result is proved by observing that $V_1 = \mathcal{H}$ and $|V_{T+1}| \geq 1$ by realizability assumption.

Problem 3 (Optional)

Consider realizable online classification with hypothesis class $\mathrm{Ldim}(\mathcal{H}) = \infty$. If the learner is allowed to randomly predict a label at every timestep, can it achieve a finite mistake bound? Why or why not?

Solution

No. We will show that given a mistake tree T of depth d, the adversary can use it to ensure that the learner makes at least $\frac{d}{2}$ mistakes in expectation.

Let $T_1 \leftarrow T$. For $t = 1, 2, \ldots, d$:

1. Let x_t be the root node of T_t. Show x_t to the learner.
2. See learner's prediction $\hat{y}_t \in [0, 1]$.
3. If $\hat{y}_t \geq \frac{1}{2}$, reveal label $y_t = 0$, and let T_{t+1} be the left subtree of T_t; else reveal label $y_t = 1$, and let T_{t+1} be the right subtree of T_t.

It can be seen that for the first d rounds, the value of $|\hat{y}_t - y_t|$ (instantaneous classification error) is at least $\frac{d}{2}$. Moreover, the realizability condition is satisfied by the definition of the mistake tree - the leaf we reach contains a classifier that is consistent with all examples $\left\{(x_t, y_t)\right\}_{t=1}^{d}$.

Now, suppose there exists an randomized online classification algorithm that has a mistake bound of M with respect to $\mathcal{H}$. We know that $\mathrm{Ldim}(\mathcal{H}) = \infty$, therefore $\mathcal{H}$ has a mistake tree of depth $2M + 1$. By the above reasoning, there exists an adversary that forces the learner to make at least $\frac{2M+1}{2} > M$ mistakes in expectation, contradition.

Problem 4 (Optional)

Show that Hedge with learning rate $\eta > 0$ has a regret as follows:

$$\sum_{t=1}^{T} \langle p_t, \ell_t \rangle - \min_{i=1}^{N} \sum_{t=1}^{T} \ell_{t,i} \leq \frac{\ln N}{\eta} + \eta \sum_{t=1}^{T} \sum_{i=1}^{N} p_{t,i} \ell_{t,i}^2.$$

You can use the fact that $e^x \leq 1 + x + x^2$ for $x \leq 1$.

(This bound has many useful applications, for example, adversarial multi-armed bandits.)

Solution

Note that in class, we have shown that

$$Z_t = \ln\left(\sum_{i=1}^{N} p_{t,i} e^{-\eta \ell_{t,i}}\right) \leq -\eta \langle p_t, \ell_t \rangle + \frac{\eta^2}{8},$$

which gives the familiar regret bound of the form $\frac{\ln N}{\eta} + \frac{\eta}{8} T$. (Please review the notes to refresh if you have forgotten this.)

For this exercise, it suffices to show that for every round t,

$$Z_t = \ln\left(\sum_{i=1}^{N} p_{t,i} e^{-\eta \ell_{t,i}}\right) \leq -\eta \langle p_t, \ell_t \rangle + \eta^2 \sum_{i=1}^{N} p_{t,i} \ell_{t,i}^2.$$

To show this, let us start with upper bounding the sum inside the logarithm.

$$\begin{aligned} & \sum_{i=1}^{N} p_{t,i} e^{-\eta \ell_{t,i}} \\ \leq \quad & \sum_{i=1}^{N} p_{t,i} (1 - \eta \ell_{t,i} + \eta^2 \ell_{t,i}^2) \\ = \quad & 1 - \eta (\sum_{i=1}^{N} p_{t,i} \ell_{t,i}) + \eta^2 (\sum_{t=1}^{T} p_{t,i} \ell_{t,i}^2). \end{aligned}$$

where the inequality uses the fact that $e^x \leq 1 + x + x^2$ for $x \leq 1$, and $\eta \ell_{t,i} \leq 0 \leq 1$; the equality is by algebra.

Taking natural logarithm on both sides, we have

$$\begin{aligned} Z_t = \ln \left(\sum_{i=1}^{N} p_{t,i} e^{-\eta \ell_{t,i}} \right) \quad & \leq \quad \ln \left(1 - \eta (\sum_{i=1}^{N} p_{t,i} \ell_{t,i}) + \eta^2 (\sum_{t=1}^{T} p_{t,i} \ell_{t,i}^2) \right) \\ & \leq \quad -\eta (\sum_{i=1}^{N} p_{t,i} \ell_{t,i}) + \eta^2 (\sum_{t=1}^{T} p_{t,i} \ell_{t,i}^2). \end{aligned}$$

where the second inequality uses the fact that $\ln(1 + x) \leq x$ for all x. □