Min Max
Game Theory On-line Prediction and Boosting

CS 602

Keerthana Reddy
Paidi Venkata Ganesh

October 25, 2023

Game Theory

- Two-person games in normal form.
- Players: Row and Column player.
- Game defined by a loss matrix **M**.
- Row player chooses a row i, column player chooses a column j.
- Loss is represented by $M(i, j)$.
- Loss matrix for "Rock, Paper, Scissors":

	R	P	S
R	$\frac{1}{2}$	1	0
P	0	$\frac{1}{2}$	1
S	1	0	$\frac{1}{2}$

Game Objectives and Generalization

- Row player's goal: Minimize loss.
- Zero-sum game: Column player aims to maximize loss.
- Assumptions: Losses in the range $[0, 1]$ for simplicity.
- Finite choices for each player.

Randomized Play

- Players choose strategies randomly.
- Row player: $\mathbf{P}$ over rows, Column player: $\mathbf{Q}$ over columns.
- Row player's expected loss: $\mathbf{P}^{\mathrm{T}}\mathbf{MQ}$.
- Pure strategies vs. Mixed strategies.
- Number of rows denoted by n.

Sequential Play and Minmax Strategy

- Play is sequential, column player chooses $\mathbf{Q}$ after row player's $\mathbf{P}$.
- Column player aims to maximize the row player's loss.
- Row player minimizes $\max_{\mathbf{Q}} \mathbf{M}(\mathbf{P}, \mathbf{Q})$.
- Minmax strategy: $\mathbf{P}^*$.

The Minmax Theorem

- The player playing last doesn't matter.
- Von Neumann's minmax theorem:

$$\max_{\mathbf{Q}} \min_{\mathbf{P}} \mathbf{M}(\mathbf{P}, \mathbf{Q}) = \min_{\mathbf{P}} \max_{\mathbf{Q}} \mathbf{M}(\mathbf{P}, \mathbf{Q})$$

- Value of the game: v.
- Minmax strategy $\mathbf{P}^*$ and maxmin strategy $\mathbf{Q}^*$ are optimal.

Repeated Play

- Model: Learner vs. Environment
- Learner's strategy $\mathbf{P}_t$, Environment's strategy $\mathbf{Q}_t$
- Learner's goal: Minimize cumulative loss
- Cumulative loss: $\sum_{t=1}^{T} \mathbf{M}(\mathbf{P}_t, \mathbf{Q}_t)$
- Best strategy in hindsight: $\min_{\mathbf{P}} \sum_{t=1}^{T} \mathbf{M}(\mathbf{P}, \mathbf{Q}_t)$

Algorithm LW for Repeated Play

- Learner maintains nonnegative weights on rows of $\mathbf{M}$
- Weight update: $w_{t+1}(i) = w_t(i) \cdot \beta^{\mathbf{M}(i,\mathbf{Q}_t)}$

$$\mathbf{P}_t(i) = \frac{w_t(i)}{\sum_i w_t(i)}$$

- Theoretical bound on loss (Theorem 1):

$$\sum_{t=1}^{T} \mathbf{M}\left(\mathbf{P}_t, \mathbf{Q}_t\right) \leq a_\beta \min_{\mathbf{P}} \sum_{t=1}^{T} \mathbf{M}\left(\mathbf{P}, \mathbf{Q}_t\right) + c_\beta \ln n$$

where

$$a_\beta = \frac{\ln(1/\beta)}{1-\beta} \quad c_\beta = \frac{1}{1-\beta}.$$

Average Loss (Corollary 2)

- Under the conditions of Theorem 1 and with β set to

$$\frac{1}{1+\sqrt{\frac{2\ln n}{T}}}$$

the average per-trial loss suffered by the learner is

$$\frac{1}{T}\sum_{t=1}^{T}\mathbf{M}\left(\mathbf{P}_t,\mathbf{Q}_t\right) \leq \min_{\mathbf{P}}\frac{1}{T}\sum_{t=1}^{T}\mathbf{M}\left(\mathbf{P},\mathbf{Q}_t\right)+\Delta_T$$

where

$$\Delta_T=\sqrt{\frac{2\ln n}{T}}+\frac{\ln n}{T}=O\left(\sqrt{\frac{\ln n}{T}}\right)$$

Loss vs. Game Value (Corollary 3)

- Under the conditions of Corollary 2,

$$\frac{1}{T}\sum_{t=1}^{T}\mathbf{M}(\mathbf{P}_t,\mathbf{Q}_t) \le v + \Delta_T$$

where v is the value of the game $\mathbf{M}$.
Proof: Let $\mathbf{P}^*$ be a minmax strategy for $\mathbf{M}$ so that for all column strategies $\mathbf{Q}, \mathbf{M}(\mathbf{P}^*,\mathbf{Q}) \le v$. Then, by Corollary 2,
$\frac{1}{T}\sum_{t=1}^{T}\mathbf{M}(\mathbf{P}_t,\mathbf{Q}_t) \le \frac{1}{T}\sum_{t=1}^{T}\mathbf{M}(\mathbf{P}^*,\mathbf{Q}_t) + \Delta_T \le v + \Delta_T$.

Proof of the Minmax Theorem

- Proof of von Neumann's minmax theorem
- Key inequality:

$$\min_{\mathbf{P}} \max_{\mathbf{Q}} \mathbf{M}(\mathbf{P}, \mathbf{Q}) \leq \max_{\mathbf{Q}} \min_{\mathbf{P}} \mathbf{M}(\mathbf{P}, \mathbf{Q})$$

Proof of the Minmax Theorem

Let $\overline{\mathbf{P}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{P}_t$ and $\overline{\mathbf{Q}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{Q}_t$

$$\begin{aligned}
\min_{\mathbf{P}} \max_{\mathbf{Q}} \mathbf{P}^{\mathrm{T}}\mathbf{MQ} \\
&\le \max_{\mathbf{Q}} \overline{\mathbf{P}}^{\mathrm{T}}\mathbf{MQ} \\
&= \max_{\mathbf{Q}} \frac{1}{T}\sum_{t=1}^{T} \mathbf{P}_t^{\mathrm{T}}\mathbf{MQ} && \text{by definition of } \overline{\mathbf{P}} \\
&\le \frac{1}{T}\sum_{t=1}^{T} \max_{\mathbf{Q}} \mathbf{P}_t^{\mathrm{T}}\mathbf{MQ} \\
&= \frac{1}{T}\sum_{t=1}^{T} \mathbf{P}_t^{\mathrm{T}}\mathbf{MQ}_t \\
&\le \min_{\mathbf{P}} \frac{1}{T}\sum_{t=1}^{T} \mathbf{P}^{\mathrm{T}}\mathbf{MQ}_t + \Delta_T \\
&= \min_{\mathbf{P}} \mathbf{P}^{\mathrm{T}}\mathbf{M}\overline{\mathbf{Q}} + \Delta_T \\
&\le \max_{\mathbf{Q}} \min_{\mathbf{P}} \mathbf{P}^{\mathrm{T}}\mathbf{MQ} + \Delta_T.
\end{aligned}$$

Approximately Solving a Game

- Algorithm LW can find an approximate minmax or maxmin strategy.
- $\max_{\mathbf{Q}} \mathbf{M}(\overline{\mathbf{P}}, \mathbf{Q}) \leq \max_{\mathbf{Q}} \min_{\mathbf{P}} \mathbf{M}(\mathbf{P}, \mathbf{Q}) + \Delta_T = v + \Delta_T$
- $\overline{\mathbf{P}}$ is an approximate minmax strategy within Δ_T of the game value v.
- $\min_{\mathbf{P}} \mathbf{M}(\mathbf{P}, \overline{\mathbf{Q}}) \geq v - \Delta_T$
- $\overline{\mathbf{Q}}$ is an approximate maxmin strategy within Δ_T of the game value v.

On-line Prediction

Formally, let X be a finite set of instances, and let $\mathcal{H}$ be a finite set of hypotheses $h : X \to \{0, 1\}$. Let $c : X \to \{0, 1\}$ be an unknown target concept, not necessarily in $\mathcal{H}$.

In the on-line prediction model, learning takes place in a sequence of rounds. On round $t = 1, \dots, T$:

1. the learner observes an example $x_t \in X$;
2. the learner makes a randomized prediction $\hat{y}_t \in \{0, 1\}$ of the label associated with x_t;
3. the learner observes the correct label $c(x_t)$. It is straightforward now to reduce the on-line prediction problem to a special case of the repeated game problem.

$$\mathbf{M}(h, x) = \begin{cases} 1 & \text{if } h(x) \neq c(x) \\ 0 & \text{otherwise} \end{cases}$$

ON-LINE PREDICTION

$\mathbf{M}(h,x)$ is 1 if and only if h disagrees with the target c on instance x. We call this a mistake matrix.

$$\begin{aligned}\mathbf{M}\left(\mathbf{P}_t, x_t\right) &= \sum_{h\in\mathcal{H}} \mathbf{P}_t(h)\mathbf{M}\left(h, x_t\right) \\ &= \Pr_{h\sim\mathbf{P}_t}\left[h\left(x_t\right) \neq c\left(x_t\right)\right] \\ &= \Pr\left[\hat{y}_t \neq c\left(x_t\right)\right].\end{aligned}$$

$$\sum_{t=1}^{T} \mathbf{M}\left(\mathbf{P}_t, x_t\right) \leq \min_{h\in\mathcal{H}} \sum_{t=1}^{T} \mathbf{M}\left(h, x_t\right) + O(\sqrt{T \ln |\mathcal{H}|})$$

Boosting

- Boosting converts a "weak" learning algorithm into one that performs with arbitrarily good accuracy.
- Dual connection: On-line prediction and boosting are closely related.
- Boosting algorithms can be derived from on-line prediction algorithms through this connection.
  For $\gamma > 0$, we say that algorithm WL is a γ-weak learning algorithm for $(\mathcal{H}, c)$ if, for any distribution $\mathbf{Q}$ over the set X, the algorithm takes as input a set of labeled examples distributed according to $\mathbf{Q}$ and outputs a hypothesis $h \in \mathcal{H}$ with error at most $1/2 - \gamma$, i.e., $\Pr_{x \sim Q}[h(x) \neq c(x)] \leq \frac{1}{2} - \gamma$.
  Given a weak learning algorithm, the goal of boosting is to run the weak learning algorithm many times on many distributions, and to combine the selected hypotheses into a final hypothesis with arbitrarily small error rate

Boosting

Boosting proceeds in rounds. On round $t = 1, \ldots, T$:
1. the booster constructs a distribution D_t on X which is passed to the weak learner;
2. the weak learner produces a hypothesis $h_t \in \mathcal{H}$ with error at most $1/2 - \gamma$:

$$\Pr_{x \sim D_t}[h_t(x) \neq c(x)] \leq \frac{1}{2} - \gamma$$

After T rounds, the weak hypotheses $h_1, \ldots, h_T$ are combined into a final hypothesis h_{fin} .
The important issues for designing a boosting algorithm are: (1) how to choose distributions D_t, and (2) how to combine the h_t 's into a final hypothesis.

Boosting and the minmax theorem

$$\begin{aligned}\min_{\mathbf{P}} \max_{x} \mathbf{M}(\mathbf{P}, x) &= \min_{\mathbf{P}} \max_{\mathbf{Q}} \mathbf{M}(\mathbf{P}, \mathbf{Q}) \\ &= v \\ &= \max_{\mathbf{Q}} \min_{\mathbf{P}} \mathbf{M}(\mathbf{P}, \mathbf{Q}) \\ &= \max_{\mathbf{Q}} \min_{h} \mathbf{M}(h, \mathbf{Q}).\end{aligned}$$

It is straightforward to show that, for any $\mathbf{Q}$, $\min_{\mathbf{P}} \mathbf{M}(\mathbf{P}, \mathbf{Q})$ is realized at a pure strategy h. Similarly for $\mathbf{P}$ and x

$$\mathbf{M}(h, \mathbf{Q}) = \Pr_{x \sim \mathbf{Q}}[h(x) \neq c(x)]$$

There exists a distribution $\mathbf{Q}^*$ on X such that for every hypothesis h, $\mathbf{M}\left(h, \mathbf{Q}^*\right) = \Pr_{x \sim \mathbf{Q}^*}[h(x) \neq c(x)] \geq v$.

Boosting and the MinMax theorem

Because we assume γ-weak learnability, there must exist a hypothesis h such that

$$\Pr_{x \sim \mathbf{Q}^*}[h(x) \neq c(x)] \leq \frac{1}{2} - \gamma$$

Combining these facts gives that $v \leq 1/2 - \gamma$.
There exists a distribution $\mathbf{P}^*$ over the hypothesis space $\mathcal{H}$ such that for every $x \in X$:

$$\mathbf{M}\left(\mathbf{P}^*, x\right) = \Pr_{h \sim \mathbf{P}^*}[h(x) \neq c(x)] \leq v \leq \frac{1}{2} - \gamma < \frac{1}{2}.$$

That is, every instance x is misclassified by less than $1/2$ of the hypotheses (as weighted by $\mathbf{P}^*$).

Idea of Boosting

Recall that on each round, algorithm LW computes a distribution over the rows of the game matrix (hypotheses, in the case of matrix **M**). However, in the boosting model, we want to compute on each round a distribution over instances (columns of **M**).
The dual $\mathbf{M}'$ of **M** is simply

$$\mathbf{M}' = \mathbf{1} - \mathbf{M}^{\mathrm{T}}$$

$$\mathbf{M}'(x,h) = 1 - \mathbf{M}(h,x) = \begin{cases} 1 & \text{if } h(x) = c(x) \\ 0 & \text{otherwise.} \end{cases}$$

Note that any minmax strategy of the game **M** becomes a maxmin strategy of the game $\mathbf{M}'$.

Idea of Boosting

The reduction proceeds as follows: On round t of boosting
1. algorithm LW computes a distribution $\mathbf{P}_t$ over rows of $\mathbf{M}'$ (i.e., over X);
2. the boosting algorithm sets $D_t = \mathbf{P}_t$ and passes D_t to the weak learning algorithm;
3. the weak learner returns a hypothesis h_t satisfying

$$\Pr_{x \sim D_t}[h_t(x) = c(x)] \geq \frac{1}{2} + \gamma$$

4. the weights maintained by algorithm LW are updated where $\mathbf{Q}_t$ is defined to be the pure strategy h_t.
In other words, h_t should have maximum accuracy with respect to distribution $\mathbf{P}_t$.

Idea of Boosting

Finally, this method suggests that $\overline{\mathbf{Q}} = (1/T)\sum_{t=1}^{T} \mathbf{Q}_t$ is an approximate maxmin strategy, and we know that the target c is equivalent to a majority of the hypotheses if weighted by a maxmin strategy of $\mathbf{M}'$. Since $\mathbf{Q}_t$ is in our case concentrated on pure strategy (hypothesis) h_t, this leads us to choose a final hypothesis h_{fin} which is the (simple) majority of $h_1, \ldots, h_T$.

Indeed, the resulting boosting procedure will compute a final hypothesis h_{fin} identical to c for sufficiently large T.
As noted earlier, for all t,

$$\mathbf{M}'(\mathbf{P}_t, h_t) = \Pr_{x \sim \mathbf{P}_t}[h_t(x) = c(x)] \geq \frac{1}{2} + \gamma$$

$$\frac{1}{2} + \gamma \leq \frac{1}{T}\sum_{t=1}^{T} \mathbf{M}'(\mathbf{P}_t, h_t) \leq \min_x \frac{1}{T}\sum_{t=1}^{T} \mathbf{M}'(x, h_t) + \Delta_T$$

Therefore, for all x,

$$\frac{1}{T}\sum_{t=1}^{T} \mathbf{M}'(x, h_t) \geq \frac{1}{2} + \gamma - \Delta_T > \frac{1}{2}$$

Analysis

Note that, by definition of $\mathbf{M}'$, $\sum_{t=1}^{T} \mathbf{M}'(x, h_t)$ is exactly the number of hypotheses h_t which agree with c on instance x. In words says that more than half the hypotheses h_t are correct on x. Therefore, by definition of h_{fin}, we have that $h_{fin}(x) = c(x)$ for all x.

The algorithm is actually quite intuitive in this form: after each hypothesis h_t is observed, the weight associated with each instance x is decreased if h_t is correct on that instance and otherwise is increased. Thus, each distribution focuses on the examples most likely to be misclassified by the preceding hypotheses.

A proof of theorem

For $t = 1, \ldots, T$, we have that

$$\begin{aligned}
\sum_{i=1}^{n} w_{t+1}(i) &= \sum_{i=1}^{n} w_t(i) \cdot \beta^{\mathbf{M}(i, \mathbf{Q}_t)} \\
&\leq \sum_{i=1}^{n} w_t(i) \cdot (1 - (1-\beta)\mathbf{M}(i, \mathbf{Q}_t)) \\
&= \left(\sum_{i=1}^{n} w_t(i) \right) \cdot (1 - (1-\beta)\mathbf{M}(\mathbf{P}_t, \mathbf{Q}_t))
\end{aligned}$$

The first line uses the definition of $w_{t+1}(i)$. The second line follows from the fact that $\beta^x \leq 1 - (1-\beta)x$ for $\beta > 0$ and $x \in [0, 1]$. The last line uses the definition of $\mathbf{P}_t$.

A proof of theorem

Unwrapping this simple recurrence gives

$$\sum_{i=1}^{n} w_{T+1}(i) \leq n \cdot \prod_{t=1}^{T} \left(1 - (1-\beta)\mathbf{M}\left(\mathbf{P}_t, \mathbf{Q}_t\right)\right)$$

(Recall that $w_1(i) = 1$.) Next, note that, for any j,

$$\sum_{i=1}^{n} w_{T+1}(i) \geq w_{T+1}(j) = \beta^{\sum_{t=1}^{T} \mathbf{M}(j,,\mathbf{Q}_t})$$

Combining with Eq. and taking logs gives

$$(\ln \beta) \sum_{t=1}^{T} \mathbf{M}\left(j, \mathbf{Q}_t\right)$$

$$\leq \ln n + \sum_{t=1}^{T} \ln\left(1 - (1-\beta)\mathbf{M}\left(\mathbf{P}_t, \mathbf{Q}_t\right)\right)$$

A proof of theorem

$$\leq \ln n - (1-\beta)\sum_{t=1}^{T} \mathbf{M}(\mathbf{P}_t, \mathbf{Q}_t)$$

since $\ln(1-x) \leq -x$ for $x < 1$. Rearranging terms, and noting that this expression holds for any j gives

$$\sum_{t=1}^{T} \mathbf{M}(\mathbf{P}_t, \mathbf{Q}_t) \leq a_\beta \min_j \sum_{t=1}^{T} \mathbf{M}(j, \mathbf{Q}_t) + c_\beta \ln n.$$

Since the minimum (over mixed strategies $\mathbf{P}$) in the bound of the theorem must be achieved by a pure strategy j, this implies the theorem.