# Lecture 14: Boosting, Online Learning, and Games

Han Shao

Reading: *Algorithmic Game Theory*, Section 4.4.2
Background: Sections 4.2–4.3

In Lecture 13, we introduced AdaBoost and analyzed its training and generalization errors. In this lecture, we develop the connection between boosting, online learning, and zero-sum games. We first prove the minimax theorem using the Hedge regret guarantee established earlier. We then use minimax to show that the weak-learning condition implies the existence of a weighted vote with zero training error. Finally, we construct such a vote by running a no-regret learner over the training examples.

## 1 Learning and Games

### 1.1 Finite Zero-Sum Games

Consider a game with a row player and a column player. The row player has $N$ actions, the column player has $K$ actions, and $L \in \mathbb{R}^{N \times K}$ is the row player's loss matrix. If the players choose row $i$ and column $j$, their losses are $L_{ij}$ and $-L_{ij}$, respectively. Thus the row player wants to minimize $L_{ij}$, while the column player wants to maximize it.

**Example 1** (Rock, paper, scissors)**.** *A win has loss* $-1$*, a tie has loss* $0$*, and a loss has loss* $1$*. The row player's loss matrix is*

| | *Rock* | *Paper* | *Scissors* |
|---|---|---|---|
| *Rock* | 0 | 1 | $-1$ |
| *Paper* | $-1$ | 0 | 1 |
| *Scissors* | 1 | $-1$ | 0 |

*The column player's loss matrix is* $-L$*. If the row player announces one action before the column player acts, the column player can always win. In contrast, choosing each action with probability* $1/3$ *gives expected loss zero against every column action.*

**Mixed strategies.** For a positive integer $k$, write

$$\Delta_k = \left\{ p \in [0,1]^k : \sum_{i=1}^{k} p_i = 1 \right\}$$

for the probability simplex. A *mixed strategy* is a distribution over actions. Let $P \in \Delta_N$ and $Q \in \Delta_K$ be the row and column strategies. When the actions are sampled independently from these distributions, the expected loss of the row player is

$$L(P,Q) := P^\intercal L Q = \sum_{i=1}^{N} \sum_{j=1}^{K} P_i Q_j L_{ij}.$$

The column player's expected loss is $-L(P,Q)$. For fixed $P$, a best response can be a single column, and for fixed $Q$, a best response can be a single row:

$$\max_{Q\in\Delta_K} P^\intercal LQ = \max_{j\in[K]} (P^\intercal L)_j, \qquad \min_{P\in\Delta_N} P^\intercal LQ = \min_{i\in[N]} (LQ)_i. \tag{1}$$

Indeed, an average of finitely many numbers cannot exceed their maximum or fall below their minimum, and a point mass attains the relevant extreme.

**The order of play.** We distinguish three situations. In deterministic sequential play, a player commits to a pure action, which the other player observes before responding. In simultaneous mixed play, each player chooses a distribution without seeing the other's current choice. In randomized sequential play, the first player announces a distribution and the second player responds to that distribution, *before the first player's action is sampled*. Seeing the distribution is different from seeing the realized action.
The two possible orders of commitment to mixed strategies give

$$v_{\text{row}} = \min_{P\in\Delta_N} \max_{Q\in\Delta_K} L(P,Q), \qquad v_{\text{col}} = \max_{Q\in\Delta_K} \min_{P\in\Delta_N} L(P,Q).$$

The first is the row player's optimal worst-case loss when it commits first; the second is the loss when the column player commits first.

**Theorem 1** (Minimax theorem for finite zero-sum games)**.** *For every $L \in \mathbb{R}^{N\times K}$,*

$$\min_{P\in\Delta_N} \max_{Q\in\Delta_K} P^\intercal LQ = \max_{Q\in\Delta_K} \min_{P\in\Delta_N} P^\intercal LQ.$$

*The common number is called the* value *of the game.*

**The no-regret guarantee we use.** Recall the Hedge guarantee from the earlier lecture. For loss vectors $\ell_t \in [0,1]^N$ and learning rate $\eta > 0$, its distributions $p_t \in \Delta_N$ satisfy

$$\sum_{t=1}^{T} \langle p_t, \ell_t\rangle \le \min_{p\in\Delta_N} \sum_{t=1}^{T} \langle p, \ell_t\rangle + R_T, \qquad R_T = \frac{\ln N}{\eta} + \frac{\eta T}{8}. \tag{2}$$

The minimum over fixed distributions equals the minimum over fixed actions because the objective is linear in $p$. For $N \ge 2$, setting $\eta = \sqrt{8\ln N/T}$ gives $R_T = \sqrt{T\ln N/2}$; for $N = 1$, the regret is zero. We use the bound directly for *mixed losses* $\langle p_t, \ell_t\rangle$, without sampling actions. It holds for every realized interaction, including when $\ell_t$ depends on the current distribution $p_t$. This is the form needed for the best-response constructions below.

## 1.2 Proving Minimax by No-Regret Learning

*Proof of theorem 1.* We establish the two inequalities separately.

**The easy direction.** For every pair $P, Q$,

$$\min_{P'} L(P',Q) \le L(P,Q) \le \max_{Q'} L(P,Q').$$

The leftmost expression does not depend on $P$, and the rightmost does not depend on $Q$. Taking the maximum over $Q$ on the left and the minimum over $P$ on the right yields

$$v_{\text{col}} \le v_{\text{row}}.$$

**Normalizing the losses.** It suffices to prove the other direction for $L \in [0,1]^{N\times K}$. Indeed, let $a = \min_{i,j} L_{ij}$ and $b = \max_{i,j} L_{ij}$. If $a = b$, the result is immediate. Otherwise, replace each entry by $(L_{ij} - a)/(b - a)$. For every $P, Q$, this transforms the expected loss by the same positive affine map, so equality for the normalized game implies equality for the original one.

**A no-regret player against best responses.** For the normalized game, repeat the following interaction for $T$ rounds:

1. The row player uses Hedge to choose $P_t \in \Delta_N$ from the preceding loss vectors.

2. After seeing $P_t$, the column player chooses

$$Q_t \in \operatorname*{argmax}_{Q\in\Delta_K} P_t^\intercal LQ.$$

3. The row player receives the loss vector $\ell_t = LQ_t \in [0,1]^N$ and updates its Hedge weights.

This is an allowed sequence for (2), even though $Q_t$ depends on the current distribution $P_t$. No realized row action is revealed to the column player.
Define the average strategies

$$\overline{P}_T = \frac{1}{T}\sum_{t=1}^{T} P_t, \qquad \overline{Q}_T = \frac{1}{T}\sum_{t=1}^{T} Q_t.$$

They lie in the corresponding simplices. Linearity, best responses, and the regret guarantee give

$$\begin{aligned}
\max_Q \overline{P}_T^\intercal LQ &= \max_Q \frac{1}{T}\sum_{t=1}^{T} P_t^\intercal LQ \\
&\le \frac{1}{T}\sum_{t=1}^{T} \max_Q P_t^\intercal LQ \\
&= \frac{1}{T}\sum_{t=1}^{T} P_t^\intercal LQ_t \\
&\le \min_P \frac{1}{T}\sum_{t=1}^{T} P^\intercal LQ_t + \frac{R_T}{T} \\
&= \min_P P^\intercal L\overline{Q}_T + \frac{R_T}{T}. \qquad (3)
\end{aligned}$$

The first inequality lets the column player choose a different best response in each round instead of one fixed $Q$ for all rounds. The next equality uses the definition of $Q_t$. The second inequality is Hedge's regret bound against the best fixed row distribution.
Consequently,

$$v_{\text{row}} \le \max_Q \overline{P}_T^\intercal LQ \le \min_P P^\intercal L\overline{Q}_T + \frac{R_T}{T} \le v_{\text{col}} + \frac{R_T}{T}.$$

For $N \ge 2$, take $R_T = \sqrt{T \ln N/2}$; for $N = 1$, take $R_T = 0$. In either case, $R_T/T \to 0$. The game values do not depend on $T$, so $v_{\text{row}} \le v_{\text{col}}$. Combined with the easy direction, this proves equality. All extrema are attained: by (1), the relevant objectives are finite maxima or minima of continuous linear functions on compact simplices. □

**Optimal strategies and the order of play.** Let $v$ be the common value, and let $P^\star$ and $Q^\star$ attain the minimax and maximin values. Then, for every $P, Q$,

$$L(P^\star, Q) \le v \le L(P, Q^\star), \qquad L(P^\star, Q^\star) = v.$$

Thus each player can guarantee the value using its own optimal mixed strategy. Observing the other player's distribution before responding does not improve the optimal guaranteed value. This statement does not apply when the responder observes the other player's sampled action.

## 2 Boosting and the Minimax Theorem

We now interpret the weak-learning condition as a statement about a finite zero-sum game. Fix a labeled sample

$$S = ((x_1, y_1), \dots, (x_m, y_m)), \qquad y_i \in \{-1, +1\},$$

and a nonempty base class $\mathcal{H} \subseteq \{-1, +1\}^{\mathcal{X}}$. As in Lecture 13, distributions on the sample are distributions on its $m$ indexed occurrences. For $Q \in \Delta_m$, write

$$\mathrm{err}_Q(h) = \sum_{i=1}^{m} Q_i \mathbb{1}\{h(x_i) \ne y_i\}.$$

The finite-sample weak-learning condition, with edge $\gamma \in (0, 1/2)$, is

$$\text{for every } Q \in \Delta_m, \qquad \text{there exists } h \in \mathcal{H} \text{ with } \mathrm{err}_Q(h) \le \frac{1}{2} - \gamma. \tag{4}$$

**The loss matrix.** Choose one representative $h_j$ for each distinct prediction vector induced by $\mathcal{H}$ on $S$, and write these representatives as $h_1, \dots, h_N$. There are at most $2^m$ such vectors, so this gives a finite game even when the base class is infinite. Define its loss matrix by

$$M_{ji} = \mathbb{1}\{h_j(x_i) \ne y_i\}, \qquad j \in [N], \quad i \in [m].$$

The rows are hypotheses and the columns are sample occurrences:

| | $x_1$ | $x_2$ | $\cdots$ | $x_i$ | $\cdots$ | $x_m$ |
|---|---|---|---|---|---|---|
| $h_1$ | $M_{11}$ | $M_{12}$ | $\cdots$ | $M_{1i}$ | $\cdots$ | $M_{1m}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ | | $\vdots$ |
| $h_j$ | $M_{j1}$ | $M_{j2}$ | $\cdots$ | $\boxed{M_{ji}}$ | $\cdots$ | $M_{jm}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ | | $\vdots$ |
| $h_N$ | $M_{N1}$ | $M_{N2}$ | $\cdots$ | $M_{Ni}$ | $\cdots$ | $M_{Nm}$ |

The row player chooses a distribution $P \in \Delta_N$ over hypotheses; the column player chooses a distribution $Q \in \Delta_m$ over sample occurrences. The expected error is

$$P^\intercal M Q = \sum_{j=1}^{N} \sum_{i=1}^{m} P_j Q_i \mathbb{1}\{h_j(x_i) \ne y_i\}.$$

The row player minimizes error, while the column player maximizes it.

**The two orders of play.** For this game, define

$$v_{\text{row}} = \min_{P \in \Delta_N} \max_{Q \in \Delta_m} P^\intercal M Q, \qquad v_{\text{col}} = \max_{Q \in \Delta_m} \min_{P \in \Delta_N} P^\intercal M Q.$$

Let $P^\star$ attain the first minimum and $Q^\star$ attain the second maximum. For a fixed $P$, a maximizing column player can concentrate all its mass on a single column, so

$$v_{\text{row}} = \max_{Q \in \Delta_m} (P^\star)^\intercal M Q = \max_{i \in [m]} \sum_{j=1}^{N} P_j^\star \mathbb{1}\{h_j(x_i) \neq y_i\}.$$

Thus $v_{\text{row}}$ is the smallest possible worst-example error of a randomized hypothesis drawn from a fixed mixture.
Similarly, for fixed $Q$, a minimizing row player can concentrate on a single hypothesis. Thus

$$v_{\text{col}} = \min_{P \in \Delta_N} P^\intercal M Q^\star = \min_{h \in \mathcal{H}} \text{err}_{Q^\star}(h).$$

Representatives preserve these errors because their predictions on the sample are identical to those of the hypotheses they represent.

**Using the weak-learning condition.** By (4), for every distribution $Q$ over the sample there is a hypothesis with error at most $1/2 - \gamma$. In particular, this holds for $Q^\star$, so $v_{\text{col}} \leq 1/2 - \gamma$. Applying theorem 1,

$$v_{\text{row}} = v_{\text{col}} \leq \frac{1}{2} - \gamma.$$

Consequently, for every $i \in [m]$,

$$\sum_{j=1}^{N} P_j^\star \mathbb{1}\{h_j(x_i) \neq y_i\} \leq \frac{1}{2} - \gamma.$$

The total weight of hypotheses that predict $y_i$ correctly is therefore at least $1/2 + \gamma$. Their weighted majority is correct at every sample point. Equivalently,

$$y_i \sum_{j=1}^{N} P_j^\star h_j(x_i) = 1 - 2 \sum_{j=1}^{N} P_j^\star \mathbb{1}\{h_j(x_i) \neq y_i\} \geq 2\gamma > 0.$$

Hence the classifier

$$h_{P^\star}(x) = \text{sign}\left( \sum_{j=1}^{N} P_j^\star h_j(x) \right)$$

has zero empirical error. This gives an existence explanation for boosting: being able to beat random guessing for every distribution over the examples implies the existence of one weighted vote that classifies all the examples correctly.
Our remaining goal is constructive: find such a vote using calls to the weak learner, without enumerating the possibly exponentially many rows of $M$.

## 3 Boosting and No-Regret Algorithms

We now reverse the roles of hypotheses and examples. Define the accuracy matrix

$$M'_{ij} = 1 - M_{ji} = \mathbb{1}\{h_j(x_i) = y_i\}.$$

Its rows are examples and its columns are hypotheses. The example player minimizes accuracy, while the hypothesis player maximizes it. For any $h \in \mathcal{H}$, write $M'(x_i, h) = \mathbb{1}\{h(x_i) = y_i\}$.
At round $t$, the example player uses a no-regret algorithm to choose $P_t \in \Delta_m$, placing weight on examples that earlier hypotheses classified poorly. The weak learner responds with $h_t$ satisfying

$$M'(P_t, h_t) := \sum_{i=1}^{m} P_t(i) M'(x_i, h_t) = 1 - \operatorname{err}_{P_t}(h_t) \geq \frac{1}{2} + \gamma.$$

Here the example player's loss vector is $\ell_t(i) = M'(x_i, h_t)$, and $R_T$ denotes its cumulative mixed-loss regret. Combining the weak-learning guarantee with the regret bound gives

$$\begin{aligned} \frac{1}{2} + \gamma &\leq \frac{1}{T} \sum_{t=1}^{T} M'(P_t, h_t) \\ &\leq \min_{i \in [m]} \frac{1}{T} \sum_{t=1}^{T} M'(x_i, h_t) + \frac{R_T}{T}. \end{aligned}$$

Consequently, for every $i \in [m]$,

$$\frac{1}{T} \sum_{t=1}^{T} \mathbb{1}\{h_t(x_i) = y_i\} \geq \frac{1}{2} + \gamma - \frac{R_T}{T}.$$

If $R_T/T < \gamma$, more than half of the hypotheses classify each training example correctly. Therefore,

$$h_{\text{final}}(x) = \operatorname{sign}\left( \frac{1}{T} \sum_{t=1}^{T} h_t(x) \right)$$

has zero empirical error.
For Hedge over the $m \geq 2$ examples, the earlier regret bound gives

$$\frac{R_T}{T} \leq \sqrt{\frac{\ln m}{2T}}.$$

Taking $T \geq 2 \ln m / \gamma^2$ makes this at most $\gamma/2$. Thus $O(\ln m / \gamma^2)$ weak-learner calls suffice for zero training error, matching the order obtained from the AdaBoost analysis.