Advanced Algorithms / Fall 2026

Linear Programming and Optimization

Multiplicative Weights

Learning with experts and approximate optimization

Lecture 7: 100 minutes

These notes cover only the linear-programming strand taught by Shi Li. This module comprises nine 100-minute lectures, numbered consistently with the slides. Parts I and II are reading divisions within this lecture; optional appendices are for further study.

Instructor: Shi Li

Nanjing University

Prepared from the instructor's course slides, with AI-assisted expansion of explanations, proofs, and independent-study exercises. The English and Chinese editions have matching mathematical content.

Learning goals and reading route

Part I develops one potential-function argument for online learning. Part II uses exactly the same inequality to compute approximate strategies in zero-sum games and approximately feasible points of linear programs. The appendix gives the deterministic prediction algorithms that motivate the method.

- Distinguish the learning rate η, target error δ, and width bound ρ.
- Prove the regret bound by bounding the total weight from above and one expert's weight from below.
- Translate a best-response or aggregated-constraint oracle into an approximate optimization algorithm.

Contents

1 Part I: Learning with experts **1**

1.1 The online decision problem . . . 1
1.2 The multiplicative-weights rule . . . 1
1.3 The potential argument . . . 1
1.4 Rescaling the loss range . . . 2

2 Part II: Games and approximate LP feasibility **2**

2.1 Zero-sum games: keep the signs straight . . . 2
2.2 Approximate feasibility through one aggregated constraint . . . 3
2.3 Independent-study exercises . . . 4

A Optional: halving and weighted majority **5**

A.1 A perfect expert . . . 5
A.2 No perfect expert is needed . . . 5
A.3 Why deterministic prediction stops at two . . . 5

1 Part I: Learning with experts

1.1 The online decision problem

There are $m \geq 1$ experts and $T \geq 1$ rounds. Before round t reveals its losses, the learner chooses a distribution $p^t \in \Delta_m$, where $\Delta_m = \{p \geq 0 : \sum_i p_i = 1\}$. A loss vector $\ell^t \in [-1,1]^m$ is then revealed. Expert i incurs loss ℓ_i^t, and the learner incurs the fractional loss $L_t = \langle p^t, \ell^t \rangle$. Negative losses represent gains. The benchmark is the best single expert in hindsight, not the expert that performs best separately in every round.

Definition 1.1 (Regret)**.**

$$R_T = \sum_{t=1}^{T} L_t - \min_{i \in [m]} \sum_{t=1}^{T} \ell_i^t.$$

A sublinear bound $R_T = o(T)$ means that the average loss approaches that of the best fixed expert.

The fractional model also describes expected loss when an expert is sampled from p^t. That interpretation requires the current loss vector not to depend on the sampled expert. The algebraic bound below holds for every realized sequence of loss vectors, even when the sequence depends on the past or on the displayed distributions. A bound on the fractional loss is not automatically a pathwise bound on the learner's sampled loss.

1.2 The multiplicative-weights rule

Exponential multiplicative weights

Choose $0 < \eta \leq 1$ and initialize $w_i^0 = 1$ for all i. For $t = 1, \ldots, T$:

1. Set $\Phi^{t-1} = \sum_i w_i^{t-1}$ and choose $p_i^t = w_i^{t-1}/\Phi^{t-1}$.
2. Observe ℓ^t, incur $L_t = \sum_i p_i^t \ell_i^t$, and update

$$w_i^t = w_i^{t-1} \exp(-\eta \ell_i^t).$$

A smaller loss gives a larger relative weight. In particular, the algorithm does not discard an expert permanently after a bad round. The closed form $w_i^t = \exp(-\eta \sum_{s=1}^{t} \ell_i^s)$ shows that the distribution is a smooth preference for experts with smaller cumulative loss.

Example 1.2. For two experts and $\eta = \ln 2$, start with $p^1 = (1/2, 1/2)$. If $\ell^1 = (1, 0)$, the new weights are $(1/2, 1)$, so $p^2 = (1/3, 2/3)$. If the next loss vector is $(0, 1)$, the weights become $(1/2, 1/2)$ and the distribution returns to uniform. The cumulative losses of the two experts are now equal.

1.3 The potential argument

Lemma 1.3. *For $z \in [-1, 1]$, $e^z \leq 1 + z + z^2$.*

Proof. For $z \in [0, 1]$, the power series gives $e^z - 1 - z \leq z^2 \sum_{k \geq 2} 1/k! = (e-2)z^2 \leq z^2$. For $z \in [-1, 0]$, Taylor's theorem gives $e^z = 1 + z + e^\xi z^2/2$ for some $\xi \in [z, 0]$, so the remainder is at most $z^2/2$. □

Theorem 1.4 (MWU regret bound)**.** *For every expert i, the preceding algorithm satisfies*

$$\frac{1}{T} \sum_{t=1}^{T} L_t \leq \frac{1}{T} \sum_{t=1}^{T} \ell_i^t + \eta + \frac{\ln m}{\eta T}. \tag{1.1}$$

Proof. Because $|\eta\ell_i^t| \le 1$, the preceding lemma yields

$$\begin{aligned}\Phi^t &= \Phi^{t-1}\sum_i p_i^t e^{-\eta\ell_i^t}\\ &\le \Phi^{t-1}\sum_i p_i^t(1-\eta\ell_i^t+\eta^2(\ell_i^t)^2)\\ &\le \Phi^{t-1}(1-\eta L_t+\eta^2) \le \Phi^{t-1}\exp(-\eta L_t+\eta^2).\end{aligned}$$

The last step uses $1+u \le e^u$. Iterating from $\Phi^0 = m$ gives an upper bound; keeping any one summand gives a lower bound:

$$e^{-\eta\sum_t \ell_i^t} = w_i^T \le \Phi^T \le m\exp\Big(-\eta\sum_t L_t + T\eta^2\Big).$$

Take logarithms, rearrange, and divide by ηT to obtain (1.1). □

Thus $T \ge \ln m/\eta^2$ gives average regret at most 2η. For $m > 1$ and a known horizon $T \ge \ln m$, setting $\eta = \sqrt{\ln m/T}$ gives total regret at most $2\sqrt{T\ln m}$. If $m = 1$, the learner equals the only expert and regret is zero.

1.4 Rescaling the loss range

Corollary 1.5. *Suppose $\ell_i^t \in [-\rho,\rho]$, where $\rho > 0$. For target error $0 < \delta \le 2\rho$, use*

$$T = \max\left\{1, \left\lceil\frac{4\rho^2\ln m}{\delta^2}\right\rceil\right\}, \qquad w_i^t = w_i^{t-1}\exp\left(-\frac{\delta\ell_i^t}{2\rho^2}\right).$$

Then $T^{-1}\sum_t L_t \le \min_i T^{-1}\sum_t \ell_i^t + \delta$.

Proof. Apply the theorem to losses ℓ_i^t/ρ with learning rate $\eta = \delta/(2\rho) \le 1$. The normalized error is at most $2\eta = \delta/\rho$. Multiply the bound by ρ. □

What Part I establishes

The same update handles positive and negative losses. The guarantee compares an average of the learner's losses with every fixed expert. The price of a larger loss range is quadratic in ρ/δ in the number of rounds.

2 Part II: Games and approximate LP feasibility

2.1 Zero-sum games: keep the signs straight

Let $M \in [-1,1]^{m\times n}$ be the payment from the row player to the column player. The row player minimizes, and the column player maximizes. Write $M(x,y) = x^{\mathsf{T}}My$, where $x \in \Delta_m$ and $y \in \Delta_n$. The minimax theorem from Lecture 4 gives the common value

$$\lambda^* = \min_x\max_j M(x,j) = \max_y\min_i M(i,y).$$

Use rows as experts. In round t, MWU supplies a row distribution x^t. Choose a best-response column $j^t \in \arg\max_j M(x^t,j)$, and reveal losses $\ell_i^t = M_{i,j^t}$. For $0 < \delta \le 2$, choose the following parameters:

$$\eta = \frac{\delta}{2}, \qquad T = \max\left\{1, \left\lceil\frac{4\ln m}{\delta^2}\right\rceil\right\}.$$

Extracting two approximate strategies

Let $\hat{t}$ minimize $M(x^t, j^t)$ over all rounds, and set $\hat{x} = x^{\hat{t}}$. Let $\hat{y}$ be the empirical distribution of the selected columns:

$$\hat{y}_j = \frac{|\{t : j^t = j\}|}{T}.$$

Repeated occurrences of a column count repeatedly.

Theorem 2.1.

$$\max_j M(\hat{x}, j) \le \min_i M(i, \hat{y}) + \delta.$$

Consequently, $\hat{x}$ guarantees payment at most $\lambda^* + \delta$, and $\hat{y}$ guarantees payment at least $\lambda^* - \delta$.

Proof. Best responses, the choice of $\hat{t}$, and the regret inequality give

$$\begin{aligned}\max_j M(\hat{x}, j) = M(x^{\hat{t}}, j^{\hat{t}}) &\le \frac{1}{T}\sum_t M(x^t, j^t) \\ &\le \min_i \frac{1}{T}\sum_t M(i, j^t) + \delta = \min_i M(i, \hat{y}) + \delta.\end{aligned}$$

Every row strategy has worst-case payment at least λ^*, and every column strategy has guaranteed payment at most λ^*. Thus

$$\lambda^* \le \max_j M(\hat{x}, j) \le \min_i M(i, \hat{y}) + \delta \le \lambda^* + \delta.$$

□

The average row strategy $\bar{x} = T^{-1}\sum_t x^t$ also works: convexity of $x \mapsto \max_j M(x, j)$ bounds its worst-case payment by the average of the round values. The slides use the best iterate, which avoids storing this average. With an explicit matrix, one best response costs $O(mn)$ arithmetic operations; when columns encode exponentially many choices, the relevant resource is a best-response oracle.

2.2 Approximate feasibility through one aggregated constraint

Let $K \subseteq \mathbb{R}^n$ be a nonempty, compact convex polytope on which linear optimization is easy. We seek $x \in K$ with $Ax \ge b$, where A has $m \ge 1$ rows. Assume a known width bound

$$|A_i x - b_i| \le \rho \quad (x \in K,\ i \in [m]), \qquad \rho > 0.$$

For example, if $K = [0,1]^n$, one valid bound is $\rho = \max_i(\sum_j |a_{ij}| + |b_i|)$ when this quantity is positive. Any larger positive bound also works. Given $y \in \Delta_m$, solve the single optimization problem

$$\max_{x \in K} \langle y, Ax - b \rangle.$$

For the box, choose each coordinate according to the sign of the corresponding coefficient of $A^{\mathsf{T}}y$. In general, the algorithm assumes access to this optimization oracle.

MWU for approximate LP feasibility

Choose $0 < \delta \le 2\rho$ and the number of rounds from Corollary 1.5. Initialize $w_i^0 = 1$. In round t:

1. Normalize the weights to obtain $y^t \in \Delta_m$.
2. Find $x^t \in K$ with $\langle y^t, Ax^t - b \rangle \ge 0$. If the maximum over K is negative, return "infeasible."

3. Set $\ell_i^t = A_i x^t - b_i$ and update

$$w_i^t = w_i^{t-1} \exp\left(-\frac{\delta(A_i x^t - b_i)}{2\rho^2}\right).$$

If all rounds succeed, return $\hat{x} = T^{-1} \sum_t x^t$.

Theorem 2.2. *The algorithm either correctly certifies that no exact feasible point exists or returns $\hat{x} \in K$ satisfying $A\hat{x} \geq b - \delta \mathbf{1}$.*

Proof. If x^* were exactly feasible, then $y \geq 0$ would imply $\langle y, Ax^* - b \rangle \geq 0$ for every distribution y. Thus a negative oracle optimum is a valid infeasibility certificate. Otherwise, the learner's loss in every round is nonnegative. For every row i, the scaled regret bound gives

$$0 \leq \frac{1}{T} \sum_t \langle y^t, Ax^t - b \rangle \leq \frac{1}{T} \sum_t (A_i x^t - b_i) + \delta = A_i \hat{x} - b_i + \delta.$$

Hence every constraint is violated by at most δ. Convexity of K implies $\hat{x} \in K$. □

The sign of the update can look surprising: a well-satisfied constraint receives a positive loss and loses weight. This is exactly what is wanted. A violated constraint has negative slack, so its weight increases and the next aggregated constraint pays more attention to it.

The guarantee and its limits

If the exact system is infeasible but a δ-feasible point exists, either output is allowed. The bound is additive and depends on the scaling of each row. The iteration count depends on $(\rho/\delta)^2$, so this is not, by itself, a polynomial-time exact LP algorithm in the binary input length. The analysis uses exact oracle calls and ideal exponential updates; an implementation must budget numerical and oracle errors.

2.3 Independent-study exercises

1. Verify the two-expert example for a symbolic learning rate η. What happens if both loss vectors are $(1, 0)$?
2. Show that adding the same scalar to every expert's loss in a round leaves the next normalized distribution unchanged. Explain why the stated width bound may nevertheless change.
3. Prove the approximate game guarantee for the average row strategy. Does an arbitrary iterate have the same guarantee?
4. If a best-response column is within γ of optimal in each round, show that the game saddle gap is at most $\delta + \gamma$.
5. Let $K = [0, 1]^2$ and require $x_1 \geq 3/4$, $x_2 \geq 3/4$, and $x_1 + x_2 \leq 1$. Find a nonnegative combination of these constraints certifying infeasibility.
6. Suppose successful oracle calls guarantee only $\langle y^t, Ax^t - b \rangle \geq -\gamma$. Show that the returned average is $(\delta + \gamma)$-feasible. State what additional evidence is needed before reporting exact infeasibility.

A Optional: halving and weighted majority

A.1 A perfect expert

For binary prediction, suppose at least one expert never makes a mistake. Keep only experts that have always been correct, and predict their majority opinion. Every learner mistake eliminates at least half of the surviving experts. After M mistakes, at most $m/2^M$ remain, while the perfect expert survives. Thus $M \le \log_2 m$ (and in particular $M \le \lceil \log_2 m \rceil$).

A.2 No perfect expert is needed

Start with unit weights and follow the weighted majority. After a round, halve each incorrect expert's weight. If the learner is wrong, experts holding at least half the total weight are wrong, so the total weight falls by at least one quarter. If expert i makes M_i mistakes and the learner makes M mistakes,

$$2^{-M_i} \le \Phi^T \le m(3/4)^M, \qquad M \le \frac{\ln 2}{\ln(4/3)} M_i + \frac{\ln m}{\ln(4/3)} \le 2.41 M_i + 3.48 \ln m.$$

More generally, multiply an incorrect expert's weight by $1-\eta$, with $0 < \eta < 1$. A learner mistake reduces the potential by a factor at most $1 - \eta/2$, giving

$$M \le \frac{\ln(1-\eta)}{\ln(1-\eta/2)} M_i - \frac{\ln m}{\ln(1-\eta/2)} = (2 + O(\eta)) M_i + O(\eta^{-1} \ln m).$$

A.3 Why deterministic prediction stops at two

Consider two experts, one always predicting zero and the other always predicting one. For any deterministic learner, choose each outcome opposite to its prediction. The learner makes T mistakes; one expert makes at most $T/2$. No bound of the form $M \le (2-\gamma) \min_i M_i + f(m)$, with fixed $\gamma > 0$ and a finite additive term independent of T, can hold for all T. Randomized choice changes the conclusion under the usual condition that the outcome cannot react to the learner's current random sample. For losses in $\{0, 1\}$, the sharper bound $e^{-\eta\ell} = 1 - (1 - e^{-\eta})\ell$ gives

$$\sum_t L_t \le \frac{\eta}{1 - e^{-\eta}} \min_i \sum_t \ell_i^t + \frac{\ln m}{1 - e^{-\eta}} = (1 + O(\eta)) \min_i \sum_t \ell_i^t + O(\eta^{-1} \ln m).$$

This is the multiplicative $(1+\varepsilon)$-type guarantee alluded to in the slides. It concerns fractional or appropriately interpreted expected loss, so it does not contradict the deterministic lower bound.

Source and reading notes

Primary source: the instructor's slide folder

`7.MWU`

`Topics.tex` and `Optional.tex`.

Optional sections follow the optional slides where present. Explanatory examples and exercises are added to support independent reading; they do not change the lecture allocation.

Course website. The early-course notes linked there were used as examples of organization and level of detail, not as the source of this module's subject matter.