CSCE 631 — Algorithmic Game Theory

Lecture 7: Regret Minimization I

Alan Kuhnle

Summer 2026

Texas A&M University

Today's Plan

Online Decision Making

A New Perspective

- So far: games played *once*, solution concepts for strategic equilibrium.
- Now: **repeated** decisions over time.
- Key shift: from *equilibrium* to *learning*.
- Questions:
  - How should an agent make decisions when facing uncertainty?
  - Can an agent guarantee good performance relative to the best fixed action *in hindsight*?
  - What does learning imply about game-theoretic equilibria?

The Online Decision-Making Model

Setup

- A decision maker chooses among n **actions** (or "experts") over T rounds.
- At each round $t = 1, 2, \ldots, T$:
  1. Decision maker selects action $i_t \in [n]$ (or distribution $x^t \in \Delta_n$).
  2. Environment reveals loss vector $\ell^t \in [0, 1]^n$.
  3. Decision maker incurs loss $\ell^t_{i_t}$ (or expected loss $\langle x^t, \ell^t \rangle$).
- **No stochastic assumption:** losses can be adversarial!

Prediction with Expert Advice

Motivating scenario:

- You want to predict whether a stock goes up or down each day.
- You have access to n "experts" (analysts, models, etc.)
- Each day, each expert makes a prediction.
- After the day, the true outcome is revealed.
- **Goal:** perform nearly as well as the *best expert in hindsight*.

This is the **expert problem** — a foundational model in online learning.

Regret

Defining Regret

Definition (External Regret)

The **regret** of an online algorithm after T rounds is:

$$R_T \;=\; \sum_{t=1}^{T} \langle x^t, \ell^t \rangle \;-\; \min_{i \in [n]} \sum_{t=1}^{T} \ell_i^t$$

- Compares total loss of the algorithm to the total loss of the **best fixed action in hindsight**.
- The benchmark is chosen *after* seeing all losses — very strong!
- We want $R_T = o(T)$, equivalently $R_T / T \to 0$.

Regret Minimization as a Goal

No-Regret Algorithm

An algorithm is **no-regret** (or Hannan consistent) if:

$$\frac{R_T}{T} \to 0 \qquad \text{as } T \to \infty$$

regardless of the loss sequence (even adversarial).

Why is this a reasonable goal?

- We do not try to beat the best expert — just match it asymptotically.
- No stochastic assumptions on losses.
- The $o(T)$ regret bound means per-round regret vanishes.

Lower Bound on Regret

Theorem

For any randomized algorithm over n actions and T rounds, there exists an adversarial loss sequence such that:

$$R_T \;\geq\; \Omega\Big(\sqrt{T \log n}\Big)$$

Proof sketch:

- Adversary draws losses i.i.d. uniform from $\{0, 1\}$.
- Best expert in hindsight beats average by $\Theta(\sqrt{T \log n})$ (coupon collector / max of Gaussians argument).
- Any online algorithm gets expected loss $T/2$ (same as average).

$\Rightarrow$ $O(\sqrt{T \log n})$ regret is **optimal** (up to constants).

Weighted Majority Algorithm

Warm-Up: Halving Algorithm

Assumption: one expert is perfect (zero loss).

Halving Algorithm

1. Maintain a set S of "surviving" experts (initially $S = [n]$).
2. Each round: predict with the majority vote of S.
3. Eliminate every expert that was wrong.

- Each mistake eliminates $\geq$ half of S.
- Total mistakes $\leq \lfloor \log_2 n \rfloor$.
- **Problem:** requires a *perfect* expert; fails gracefully otherwise.

Weighted Majority Algorithm

Generalize halving: don't eliminate, just **downweight**.

Weighted Majority (Littlestone & Warmuth, 1994)

1. Initialize weights $w_i^1 = 1$ for each expert $i \in [n]$.
2. At round t: predict with the weighted majority vote.
3. Update: for each expert i that was wrong,

$$w_i^{t+1} = w_i^t \cdot (1 - \eta)$$

where $\eta \in (0, 1)$ is the learning rate.

- Wrong experts are penalized but not eliminated.
- Appropriate for settings without a perfect expert.

Weighted Majority: Regret Bound

Theorem

With $\eta \in (0, \frac{1}{2})$, the number of mistakes M_T of weighted majority satisfies:

$$M_T \;\leq\; \frac{2(1+\eta)}{1+2\eta}\, m^* + \frac{2\ln n}{\eta}$$

where m^* is the number of mistakes of the best expert.

Setting $\eta = \sqrt{\ln n / T}$ and bounding more carefully:

$$R_T \;\leq\; O\Big(\sqrt{T \ln n}\Big)$$

This is **optimal** up to constants!

Multiplicative Weights Update (MWU)

From Weighted Majority to MWU

Generalization: Allow fractional losses $\ell_i^t \in [0,1]$ and randomized predictions.

Multiplicative Weights Update (MWU)

1. Initialize weights $w_i^1 = 1$ for each $i \in [n]$.
2. At round t:
   - Play distribution $x_i^t = w_i^t / W^t$ where $W^t = \sum_i w_i^t$.
   - Observe loss vector $\ell^t \in [0,1]^n$.
3. Update: $w_i^{t+1} = w_i^t \cdot (1-\eta)^{\ell_i^t}$.

Equivalent formulation: $w_i^{t+1} = w_i^t \cdot e^{-\eta\,\ell_i^t}$ (exponential weights).

MWU: The Main Theorem

Theorem (MWU Regret Bound)

For any loss sequence $\ell^1, \ldots, \ell^T \in [0,1]^n$ and learning rate $\eta \in (0, \frac{1}{2})$:

$$\sum_{t=1}^{T} \langle x^t, \ell^t \rangle \;\leq\; \min_{i \in [n]} \sum_{t=1}^{T} \ell_i^t \;+\; \eta \sum_{t=1}^{T} \langle x^t, \ell^t \rangle \;+\; \frac{\ln n}{\eta}$$

Rearranging (dividing both sides' first term):

$$R_T \;\leq\; \frac{\eta T}{1-\eta} + \frac{\ln n}{\eta(1-\eta)}$$

Setting $\eta = \sqrt{\ln n / T}$:

$$\boxed{R_T \;\leq\; 2\sqrt{T \ln n} + \ln n}$$

MWU Regret Bound: Proof (1/3)

Potential function: $\Phi^t = \ln W^t = \ln(\sum_i w_i^t)$.

Step 1: Upper bound on Φ^{T+1}.

$$W^{t+1} = \sum_i w_i^t (1-\eta)^{\ell_i^t} \leq \sum_i w_i^t (1 - \eta\, \ell_i^t) = W^t (1 - \eta \langle x^t, \ell^t \rangle)$$

using $(1-\eta)^z \leq 1 - \eta z$ for $z \in [0,1]$.

Taking logs:

$$\Phi^{t+1} \leq \Phi^t + \ln(1 - \eta \langle x^t, \ell^t \rangle) \leq \Phi^t - \eta \langle x^t, \ell^t \rangle$$

using $\ln(1-x) \leq -x$. Telescoping:

$$\Phi^{T+1} \leq \ln n - \eta \sum_{t=1}^{T} \langle x^t, \ell^t \rangle$$

MWU Regret Bound: Proof (2/3)

Step 2: Lower bound on Φ^{T+1}.

For any fixed action i:

$$W^{T+1} \geq w_i^{T+1} = (1-\eta)^{\sum_{t=1}^{T} \ell_i^t}$$

Taking logs:

$$\Phi^{T+1} \geq \sum_{t=1}^{T} \ell_i^t \cdot \ln(1-\eta) \geq -\sum_{t=1}^{T} \ell_i^t \cdot (\eta + \eta^2)$$

using $\ln(1-\eta) \geq -\eta - \eta^2$ for $\eta \in (0, \frac{1}{2})$.

MWU Regret Bound: Proof (3/3)

Step 3: Combine bounds.

$$-(\eta+\eta^2)\sum_{t=1}^{T}\ell_i^t \;\le\; \Phi^{T+1} \;\le\; \ln n - \eta\sum_{t=1}^{T}\langle x^t,\ell^t\rangle$$

Rearranging:

$$\eta\sum_{t=1}^{T}\langle x^t,\ell^t\rangle \;\le\; (\eta+\eta^2)\sum_{t=1}^{T}\ell_i^t + \ln n$$

Dividing by η:

$$\sum_{t=1}^{T}\langle x^t,\ell^t\rangle \;\le\; (1+\eta)\sum_{t=1}^{T}\ell_i^t + \frac{\ln n}{\eta}$$

Since this holds for *all* i, it holds for the best i. With $\eta=\sqrt{\ln n/T}$:
$R_T \le 2\sqrt{T\ln n}+\ln n$. □

MWU: Choosing the Learning Rate

- Optimal $\eta = \sqrt{\ln n / T}$ requires knowing T in advance.
- **Doubling trick:** run MWU in epochs of length $1, 2, 4, 8, \ldots$, resetting weights each epoch. Total regret: $O(\sqrt{T \ln n})$ (geometric sum).
- **Adaptive learning rate:** set $\eta_t = \sqrt{\ln n / t}$ at round t. Similar guarantee with more careful analysis.

Per-Round Regret

$$\frac{R_T}{T} \leq \frac{2\sqrt{T \ln n}}{T} = 2\sqrt{\frac{\ln n}{T}} \to 0$$

MWU is a **no-regret** algorithm.

MWU: Properties and Variants

MWU as Entropic Mirror Descent

MWU has a clean optimization interpretation:

$$x^{t+1} = \arg\min_{x \in \Delta_n} \left\{ \langle \ell^t, x \rangle + \frac{1}{\eta} D_{\mathrm{KL}}(x \| x^t) \right\}$$

where D_{KL} is the KL divergence.

- **Mirror descent** with the negative entropy as the mirror map.
- Balances *exploiting* low-loss actions with *staying close* to the previous distribution.
- The KL divergence is the "right" geometry for the simplex (gives $\log n$ dependence rather than n).

Follow the Regularized Leader (FTRL)

An equivalent perspective:

$$x^{t+1} = \arg\min_{x \in \Delta_n} \left\{ \sum_{s=1}^{t} \langle \ell^s, x \rangle + \frac{1}{\eta} \sum_{i=1}^{n} x_i \ln x_i \right\}$$

- "Follow the leader" (play the best action so far) but add an entropy regularizer to avoid instability.
- Without regularization: FTL can have linear regret on adversarial sequences.
- Negative entropy regularizer $\Rightarrow$ MWU.
- Euclidean regularizer $\frac{1}{2\eta}\|x\|^2 \Rightarrow$ Online Gradient Descent.

MWU: Universality

MWU appears across many fields under different names:

Field	Name
Machine learning	Exponential weights / Hedge
Game theory	Fictitious play variant
Optimization	Mirror descent
Statistics	Bayesian updating (with log loss)
Biology	Replicator dynamics (continuous limit)
Boosting	AdaBoost (dual view)

See Arora, Hazan, Kale (2012): "The Multiplicative Weights Update Method: A

Connection to Game Theory (Preview)

Why Regret Minimization Matters for Games

Key Connection (Preview)

If all players in a repeated game use no-regret algorithms, then their empirical distribution of play converges to an **equilibrium**.

More precisely:

- No-*external*-regret $\Rightarrow$ convergence to **coarse correlated equilibrium**.
- No-*swap*-regret $\Rightarrow$ convergence to **correlated equilibrium**.
- Zero-sum games + no-regret $\Rightarrow$ convergence to **Nash equilibrium**!

This gives a **computational** and **behavioral** justification for equilibria: they arise from simple learning dynamics.

Empirical Distribution of Play

Definition

If player i plays mixed strategies $x_i^1, x_i^2, \ldots, x_i^T$, the **empirical distribution of play** is:

$$\bar{x}_i = \frac{1}{T} \sum_{t=1}^{T} x_i^t$$

- $\bar{x}_i \in \Delta_{|A_i|}$ — a valid mixed strategy.
- Represents the "average behavior" over time.
- Convergence results: the empirical profile $(\bar{x}_1, \ldots, \bar{x}_n)$ converges to an equilibrium (in an appropriate sense).

Next two lectures: formal statements and proofs.

Summary

Summary

1. **Online decision making:** adversarial losses, no distributional assumptions.
2. **Regret:** difference between algorithm's total loss and best fixed action in hindsight.
3. **Lower bound:** $\Omega(\sqrt{T \log n})$ regret is unavoidable.
4. **Weighted Majority:** downweight wrong experts; achieves $O(\sqrt{T \log n})$ for binary prediction.
5. **MWU:** general version with fractional losses; $R_T \leq 2\sqrt{T \ln n}$; optimal rate.
6. **Proof technique:** potential function $\Phi^t = \ln W^t$; sandwich between upper/lower bounds.
7. **Preview:** no-regret dynamics in games converge to equilibria.