

ECONOMICS & COMPUTATION

Spring 2026 | Lecture 19

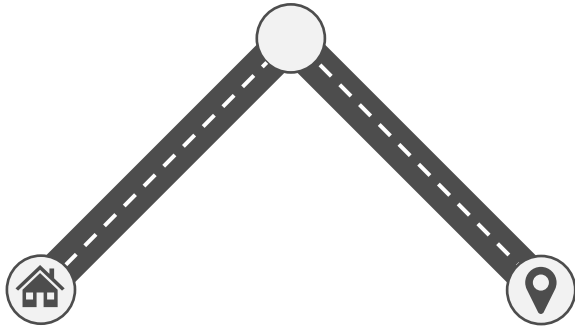
Minimax Theorem via No-Regret Learning

Ariel Procaccia | Harvard University

THE MINIMAX THEOREM: REMINDER

- **Theorem [von Neumann 1928]:** Every 2-player zero-sum game has a unique value v such that:
 - Player 1 can guarantee utility at least v
 - Player 2 can guarantee utility at least $-v$
- I claimed that “we will prove the theorem from scratch later in the course” — now is the time!

NO-REGRET LEARNING: MOTIVATION



Day 1: 53 minutes



Day 2: 47 minutes

Each morning pick one of n possible routes from home to work, then find out how long it took. Is there a strategy for picking routes that does almost as well as the best fixed route **in hindsight?**

THE MODEL

- View the interaction as a matrix

							Adversary	
Algorithm							0.2	
							0.7	
							0.8	

- Algorithm picks row, adversary column
- Alg pays cost of (row,column) and gets column as feedback
- Assume costs are in $[0,1]$

THE MODEL

- Define **average regret** in T time steps as (average per-day cost of alg) – (average per-day cost of best fixed row in hindsight)

		Adversary		
Algorithm	0.1	0.2	0.6	
	0.5	0.7	0.2	
	0.9	0.8	0.1	

$t = 1$

		Adversary		
Algorithm	0.1	0.2	0.6	
	0.5	0.7	0.2	
	0.9	0.8	0.1	

$t = 2$

		Adversary		
Algorithm	0.1	0.2	0.6	
	0.5	0.7	0.2	
	0.9	0.8	0.1	

$t = 3$

Average per day cost of the algorithm: $\frac{(0.7+0.2+0.9)}{3} = 0.6$

Up is the best in hindsight: $\frac{(0.2+0.6+0.1)}{3} = 0.3$

THE MODEL

- Define **average regret** in T time steps as
(average per-day cost of alg) – (average per-day cost of best fixed row in hindsight)
- **No-regret algorithm:** regret $\rightarrow 0$ as $T \rightarrow \infty$
- Not competing with adaptive strategy, just the best **fixed** row

EXAMPLE

1	0
0	1

Poll 1

Consider an alg that alternates between U and D.
What is its worst-case average regret?

- $\Theta(1/T)$
- $\Theta(1)$
- $\Theta(T)$
- ∞



EXAMPLE

1	0
0	1

Poll 2

Consider an alg that chooses action that has lower cost so far. What is its worst-case average regret?

- $\Theta(1/T)$
- $\Theta(1/\sqrt{T})$
- $\Theta(1/\log T)$
- $\Theta(1)$



EXAMPLE

1	0
0	1






Question

Building on this example, what can we say more generally about deterministic algorithms?



USING EXPERT ADVICE

- Want to predict the weather
- Solicit advice from n experts
 - Expert = someone with an opinion







					
Day 1	+	-	-	+	+
Day 2	-	-	+	-	-
Day 3	+	-	-	-	-

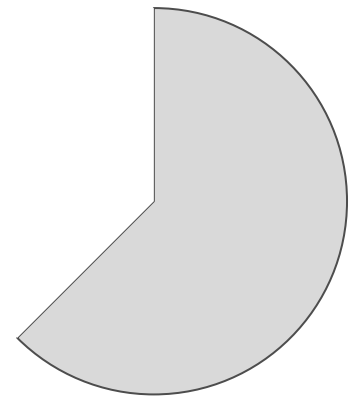
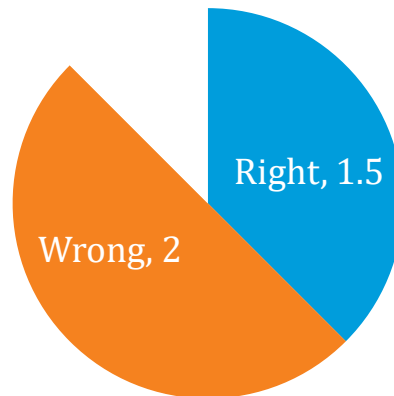
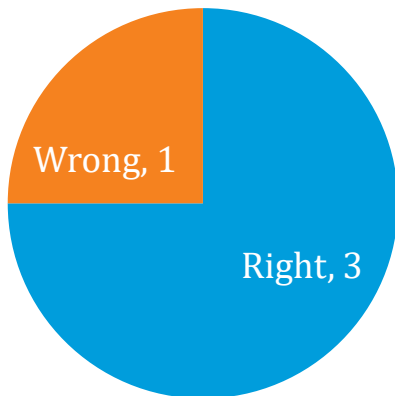
- Can we do as well as the best in hindsight?

WEIGHTED MAJORITY

- **Idea:** Experts are penalized every time they make a mistake
- **Weighted Majority Algorithm:**
 - Start with all experts having weight 1
 - Predict based on weighted majority vote
 - Penalize mistakes by cutting weight in half

WEIGHTED MAJORITY: EXAMPLE

						
Weight	1	1	1	1	3 vs. 1	
Prediction	-	+	+	+	+	+
Weight	0.5	1	1	1	1.5 vs. 2	
Prediction	+	+	-	-	-	+
Weight	0.5	1	0.5	0.5		



WEIGHTED MAJORITY: ANALYSIS

- $M = \#$ mistakes we've made so far
- $m = \#$ mistakes of best expert so far
- $W =$ total weight (starts at n)
- For each mistake, W drops by at least 25%,
so after M mistakes: $W \leq n(3/4)^M$
- Weight of best expert is $(1/2)^m$
- It follows that $(1/2)^m \leq n(3/4)^M$, and
therefore $M \leq 2.5(m + \log n)$

BEYOND WEIGHTED MAJORITY

- **Modified Weighted Majority Algorithm:**
 - Start with all experts having weight 1
 - Predict based on weighted majority vote
 - Penalize mistakes by **removing ϵ fraction** of weight

Question

Is there an ϵ that would guarantee $M \leq (1 + \delta)m$
for a small $\delta > 0$?

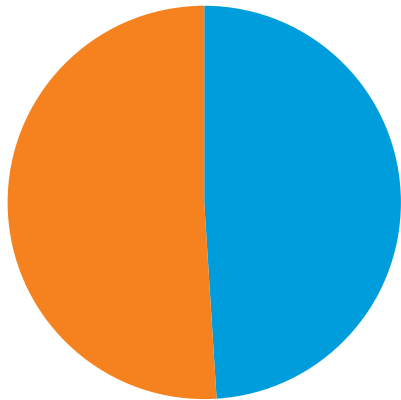


RANDOMIZED WEIGHTED MAJORITY

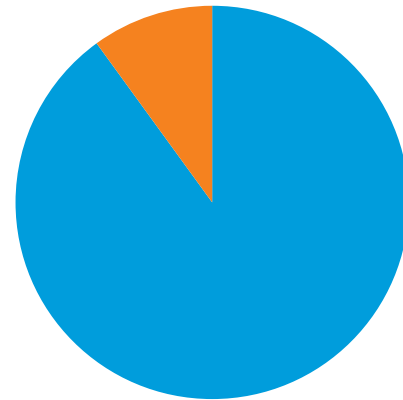
- **Idea:** Predict proportionally to weights
- **Randomized Weighted Majority Algorithm:**
 - Start with all experts having weight 1
 - If the total weight of + is w_+ and the total weight of - is w_- , predict + with probability $\frac{w_+}{w_+ + w_-}$ and - with probability $\frac{w_-}{w_+ + w_-}$
 - Penalize mistakes by removing ϵ fraction of weight

RANDOMIZED WEIGHTED MAJORITY

Idea: smooth out the worst case



The worst-case is
~50-50: now we
have a 50% chance
of getting it right



What about 90-10?
We're very likely to
agree with the
majority

RANDOMIZED WEIGHTED MAJORITY

- **Theorem:** For suitable ϵ , the randomized weighted majority algorithm has average regret at most $(2\sqrt{T \ln n})/T \rightarrow 0$
- More generally, Each **expert** is an **action** with cost in $[0,1]$
- Run Randomized Weighted Majority
 - Choose expert i with probability w_i/W
 - Update weights: $w_i \leftarrow w_i(1 - c_i\epsilon)$
- Same bound applies

THE MINIMAX THEOREM: PROOF

- In a zero-sum game G , denote:
 - V_C is the smallest reward (to row) the column player can guarantee if they commit first
 - V_R is the largest reward (to row) the row player can guarantee if they commit first
- Obviously $V_C \geq V_R$, and the theorem says equality holds
- Assume for contradiction that $V_C > V_R$
- Shift and scale matrix so that payoffs to row player are in $[-1,0]$, and let $V_C = V_R + \delta$

THE MINIMAX THEOREM: PROOF

- Suppose the game is played repeatedly; in each round the row player commits, and the column player responds
- Let the row player play RWM, and let the column player respond optimally to current mixed strategy
- After T steps
 - $\text{ALG} \geq \text{best row in hindsight} - 2\sqrt{T \log n}$
 - $\text{ALG} \leq T \cdot V_R$

THE MINIMAX THEOREM: PROOF

- **Claim:** Best row in hindsight $\geq T \cdot V_C$
 - Suppose the column player played s_t in round t
 - Define a mixed strategy y that plays each s_t with probability $1/T$ (multiplicities possible)
 - Let x be row's best response to y
 - $V_C \leq u_1(x, y) = \frac{1}{T} u_1(x, s_1) + \cdots + \frac{1}{T} u_1(x, s_T)$
 - $u_1(x, s_1) + \cdots + u_1(x, s_T) \leq$ best row in hindsight ■
- It follows that $T \cdot V_R \geq T \cdot V_C - 2\sqrt{T \log n}$
- $\delta T \leq 2\sqrt{T \log n}$ – contradiction for large T ■