CSCE 658: Randomized Algorithms

Lecture 18

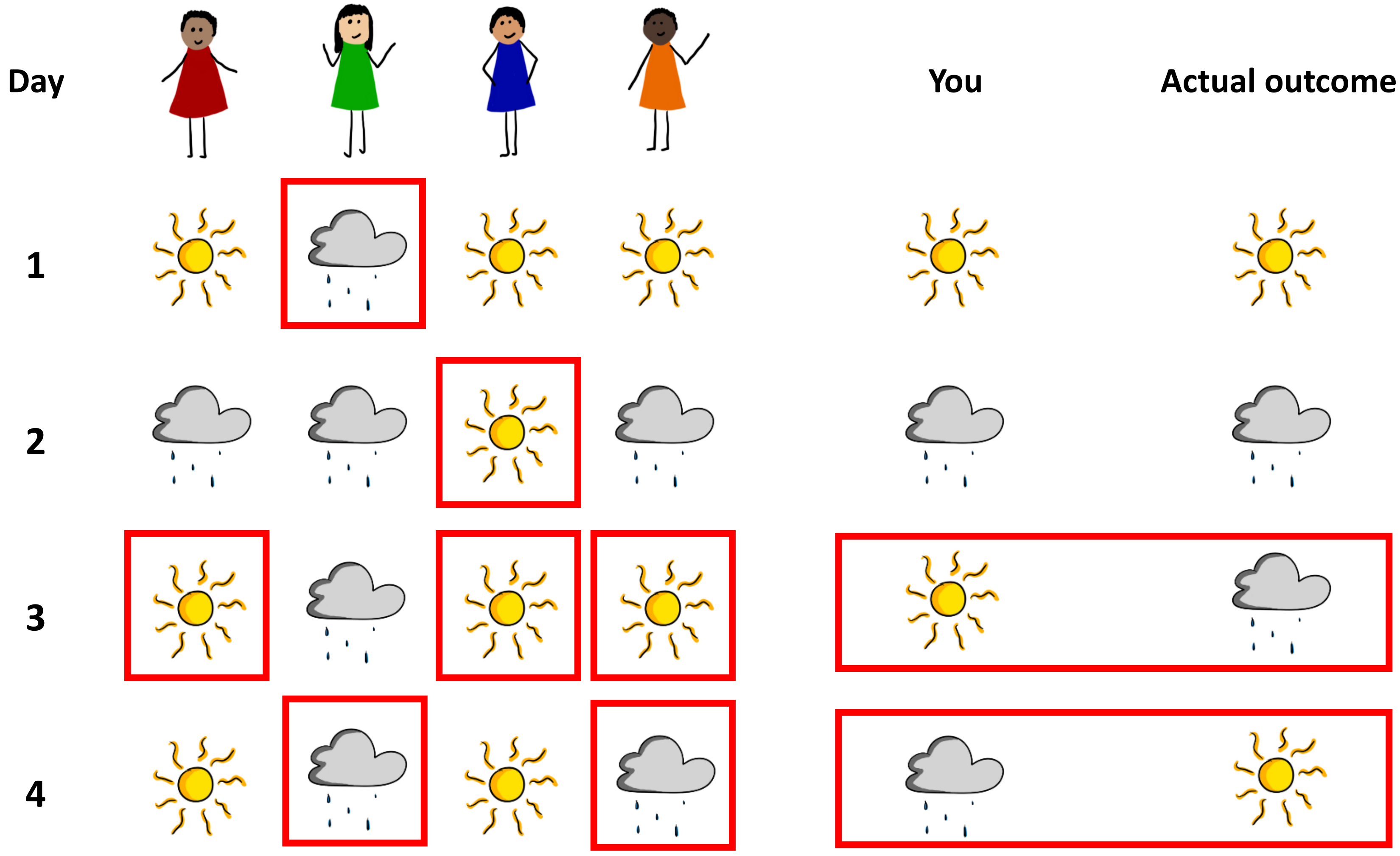
Samson Zhou

Relevant Supplementary Material

- Lecture 14 of "Advanced Algorithms" Course Notes (http://www.cs.cmu.edu/afs/cs.cmu.edu/academic/class/15850-f20/www/notes/lec15.pdf), by Anupam Gupta

Last Time: Prediction with Expert Advice

a fundamental problem of **sequential prediction**



Algorithm makes 2 mistakes

Best expert makes 1 mistake

Last Time: Hedge Algorithm

- Initially give each expert weight 1
- On day t, randomly follows expert i with probability $p_i^t = \frac{w_i^t}{w_1^t + \cdots + w_n^t} = \frac{w_i^t}{\sum_j w_j^t}$
- Each weight is updated by $w_i^{t+1} = w_i^t \cdot \exp(-\varepsilon \ell_i^t)$

Last Time: Hedge Algorithm

- If we set $\varepsilon = \sqrt{\frac{\ln n}{T}}$, we have $\sum_t \langle p^t, \ell^t \rangle \leq \sum_t \ell_{m^*}^t + \varepsilon T + \frac{\ln n}{\varepsilon} \leq \sum_t \ell_{m^*}^t + 2\sqrt{T \ln N}$

- Total regret is $2\sqrt{T \ln N}$

- Amortized regret is $2\sqrt{\frac{\ln N}{T}}$

Two-Player Zero Sum Games

- Two players in a game, the "row player" and the "column player"
- Each player has some set of actions: row player has actions $[m]$, column player has actions $[n]$
- Payoff matrix $M \in \mathbb{R}^{m \times n}$

Two-Player Zero Sum Games

- In each round of the game, the row player chooses a row $i \in [m]$, and the column player chooses a row $j \in [n]$

- The row player gets $M_{i,j}$ and the column player loses $M_{i,j}$ (which could be good for the column player if $M_{i,j}$ is negative) so the payoff is from column player to row player

- Winnings of the two players sum to zero (hence the name)

Two-Player Zero Sum Games

- Row player wants to maximize $M_{i,j}$ and the column player wants to minimize $M_{i,j}$

Two-Player Zero Sum Games

	A	B	C
1	30	-10	20
2	-10	20	-20

- What would you play as the row player?
- What would you play as the column player?

Two-Player Zero Sum Games

	A	B	C
1	30	-10	20
2	-10	20	-20

- What would you play as the row player if column player played B?
- What would you play as the column player if row player played 1?

Two-Player Zero Sum Games

	A	B	C
1	30	-10	20
2	-10	20	-20

- What would you play as the row player if you went first (and your choice is known)?
- What would you play as the column player if you went first (and your choice is known)?

Two-Player Zero Sum Games

	A	B	C
1	30	-10	20
2	-10	20	-20

- What would you play as the row player if you went first? 1
- What would you play as the column player if you went first? B or C

Best Strategies

- Each player plays simultaneously and can have randomized strategies

- Are there two-player games where randomized strategies are better than deterministic strategies?

Rock-Paper-Scissors

	Rock	Paper	Scissors
Rock	0	-1	1
Paper	1	0	-1
Scissors	-1	1	0

- Any deterministic strategy can have value -1
- Best randomized strategy has value 0

Best Strategies

- Each player plays simultaneously and can have randomized strategies
- Suppose row player plays p and the column player plays q, what is the expected payoff to the row player?

Best Strategies

- Each player plays simultaneously and can have randomized strategies

- Suppose row player plays p and the column player plays q, what is the expected payoff to the row player?

$$E[\text{Payoff to row player}] = p^\top M q = \sum_{i,j} p_i q_j M_{i,j}$$

Best Strategies

- Row player wants to maximize $\sum_{i,j} p_i q_j M_{i,j}$ and column player wants to minimize $\sum_{i,j} p_i q_j M_{i,j}$

- Suppose strategy p of row player is known, then the column player computes $C(p) = \min_q p^\top M q = \min_{j \in [n]} p^\top M e_j$.
- Suppose strategy q of column player is known, then the row player computes $R(q) = \max_p p^\top M q = \max_{i \in [m]} e_i^\top M q$

Best Strategies

- Suppose strategy p of row player is known, then the column player computes $C(p) = \min_q p^\top M q = \min_{j\in[n]} p^\top M e_j$.
- Suppose strategy q of column player is known, then the row player computes $R(q) = \max_p p^\top M q = \max_{i\in[m]} e_i^\top M q$
- How do $C(p)$ and $R(q)$ relate?

Best Strategies

- Suppose strategy p of row player is known, then the column player computes $C(p) = \min_q p^\top M q = \min_{j \in [n]} p^\top M e_j$.
- Suppose strategy q of column player is known, then the row player computes $R(q) = \max_p p^\top M q = \max_{i \in [m]} e_i^\top M q$
- How do $C(p)$ and $R(q)$ relate?
- $C(p) \leq R(q)$

Best Strategies

- Intuitively, if the column player plays a strategy q first, then the row player has more power

- Formally, the row player could always play strategy p in response to q and achieve value $C(p)$

- $R(q)$ is the best response, which can be even higher, so $C(p) \leq R(q)$

Von Neumann's Minimax Theorem

- For any finite zero-sum game $M \in \mathbb{R}^{m \times n}$,

$$\max_p C(p) = \min_q R(q)$$

- Suppose $\min_q R(q) - \max_p C(p) = \delta$ for some $\delta > 0$
- Set initial row player strategy $p^1 = \left(\frac{1}{m}, \frac{1}{m}, \dots, \frac{1}{m}\right)$
- At each time t, column player plays the best response j_t to p^t

Von Neumann's Minimax Theorem

- Row player experiences gain $g_t = Me_{j_t}$
- Update the weights to get p^{t+1} via Hedge algorithm

- By the Hedge algorithm, after $T \geq O\left(\frac{\log m}{\varepsilon^2}\right)$ steps, we have

$$\frac{1}{T}\sum_i \langle p^t, g^t \rangle \geq \max_i \frac{1}{T}\sum_i \langle e_i, g^t \rangle - \varepsilon$$

Von Neumann's Minimax Theorem

$$\frac{1}{T}\sum_i \langle p^t, g^t\rangle \geq \max_i \frac{1}{T}\sum_i \langle e_i, g^t\rangle - \varepsilon$$

$$= \max_i \left\langle e_i, \frac{1}{T}\sum_i g^t \right\rangle - \varepsilon$$

(definition of g^t) $$= \max_i \left\langle e_i, M\left(\frac{1}{T}\sum_i e_{j_t}\right)\right\rangle - \varepsilon$$

$$= R(\hat{q}) - \varepsilon$$ (for $\hat{q} = \frac{1}{T}\sum_i e_{j_t}$)

Von Neumann's Minimax Theorem

- We have $\frac{1}{T}\sum_i \langle p^t, g^t \rangle \geq R(\hat{q}) - \varepsilon$
- Can also show $\frac{1}{T}\sum_i \langle p^t, g^t \rangle \leq C(\hat{p})$, so it follows there exists strategies with $C(\hat{p}) \geq R(\hat{q}) - \varepsilon$

- We know $C(\hat{p}) \leq R(\hat{q})$
- By setting $\varepsilon \leq \delta$, the claim follows

Previously: Linear Programming (Standard Form)

- Maximize a linear objective function:

$$c^\top x = \langle c, x \rangle, \;\; c, x \in \mathbb{R}^n$$

- Subject to constraints:

$$Ax \leq b \text{ for } A \in \mathbb{R}^{m \times n}, \; b \in \mathbb{R}^m$$
$$x \geq 0 \text{ (entry-wise non-negativity)}$$

Online Learning for Solving LP's

- Use learning with experts to solve LP

- Assume we have any oracle that finds a point in feasible region that does not violate a certain constraint

Online Learning for Solving LP's

- Use learning with experts to solve LP
- Have m experts, one corresponding to each constraint
- The loss of an expert in a round is based on how badly the constraint was violated by the current solution
- Intuition: Greater violation means more loss, and hence larger weight decrease in the next iteration, which forces us to not violate the constraint
- After $T = O\left(\frac{\log n}{\varepsilon^2}\right)$ iterations, approximately solves LP