# Advanced ML & AGT

Class 10

Swap regret

# Swap Regret and Correlated Equilibrium

1. Correlated Equilibrium
2. Swap Regret
3. אלפיון

פעולות נשלטות

קלינברגיה

# שיווי משקל מתואם (Correlated)

מודל:

שחקנים $N$ קבוצה $[N]$

פעולות שחקן $i$ פעולות $A_i$

פעולה משותפת $A = A_1 \times \cdots \times A_N$

הפונ' $l_i(a_1 \cdots a_N) \in [0,1]$ $i$ ppref

$a = (a_1 \cdots a_N)$

$a_{-i} = (a_1 \cdots a_{i-1}\, a_{i+1} \cdots a_N)$

# בסדרה: שיווי משקל מתואם Correlated Equilibrium (CE)

התפלגות $Q$ מעל $A$ היא CE

אם

$$\underbrace{\forall i \in [N]}_{\text{שחקן}}\ \underbrace{\forall b, c \in A_i}_{\text{פעולות}}$$

$$E_{a \sim Q}\left[ l_i(a) - l_i(a_{-i}, c) \mid a_i = b \right] \le 0$$

שיווי משקל Nash

לכל שחקן $i$

יש התפלגות $q_i$

בלתי תלויים $a_i \sim q_i$ באופן בלתי

לכל התפלגות $q_i'$

$$\mathop{E}_{a\sim Q}[\ell_i(a)] \le \mathop{E}_{\substack{a_{-i}\sim Q\\ a_i'\sim q_i'}}[\ell_i(a_{-i},a_i')]$$

$$Q = q_1 \times \cdots \times q_N$$

הגדרה switch function

$$\text{switch}(a_i,b,c) = \begin{cases} c & a_i = b \\ a_i & a_i \neq b \end{cases}$$

הגדרה שיווי משקל CE

$$\underbrace{\forall i \in [N]}_{\text{שחקן}}\ \underbrace{\forall b,c \in A_i}_{\text{פעולות}}$$

$$E_{a\sim Q}[\ell_i(a) - \ell_i(a_{-i},\text{switch}(a_i,b,c))] \le 0$$

מה ההבדל משיווי משקל Nash?

**משפט: לכל משחק סופי יש לפחות CE אחד.**

**הוכחה:** נבנה משחק סכום אפס

שורות $a \in A$ (פעולה משותפת)

עמודות $(i,b,c)$ $i \in [N]$, $b,c \in A_i$

ערך

$$M(a,(i,b,c)) = \ell_i(a) - \ell_i(a_{-i}, switch(a_i,b,c))$$

קיום CE שקול

ערך המשחק לאל תיבי. (בלא שהוא אפס)

שתקן השורות (min)

שתקן עמודות (max)

אסטרטגיה לשתקן העמודות שמבטיחה אפס

$(i,b,b)$ כי: $switch(a_i,b,b) = a_i$

נראה אסטרטגיה לשתקן השורות

המבוססת על התפלגות $D$ על העמודות

נגדיר מטריצה $D_i$ לכל שתקן

$$D_i(b,c) = Pr[a'_i = c \mid a_i = b, (i,*,*) \text{ נבחר}]$$

זה הפעולה אחר ה switch שנבחר

$D_i$ מטריצה סטוכסטית

**טענה:**

לכל מטריצה סטוכסטית $D_i$ יש

התפלגות סטציונרית $q_i$:

$$q_i^T D_i = q_i^T$$

נגדיר לשחקן i שורות התפלגות Q

$$Q = q_1 \times \cdots \times q_N$$

מתקיים

① $a \sim Q$

② $a \sim Q$ , $(i,b,c) \sim D$

$$a' = (a_{-i}, switch(i,b,c))$$

הפילוס ב ① וב ② נהיה

$$\mathop{E}_{\substack{a \sim Q \\ (i,b,c) \sim D}} [\ell(a') - \ell(a)] = 0$$

ערך המשחק אפס!

ע"י משפט ה minimax

קיימת התפלגות Q מעל A

כך שלכל שחקן i

פעולות $c,b \in A_i$

מתקיים

$$\mathop{E}_{a \sim Q}[M(a,(i,b,c))] \le 0$$

$$\mathop{E}_{a \sim Q}[\ell_i(a)] \le \mathop{E}_{a \sim Q}[\ell_i(a_{-i}, switch(i,b,c))]$$

☆

# מימוש שווי משקל מתואם (CE)

בעזרת תוכנית LP

$\forall i \in [N] \ \forall b,c \in A_i$

$$\sum_{a \in A} q(a)\left[\ell_i(a) - \ell_i(a_{-i}, switch(i,b,c))\right] \le 0$$

משתנים $q(a)$ $\forall a \in A$

אילוצים $\sum_{i=1}^{N} |A_i|^2$

האם אפשר לפרוץ מגילות?

# $\varepsilon$-Correlated Equilibrium

פונקציית deviation

$F(A_i) = \{f: A_i \to A_i\}$

התפלגות $Q$ היא $\varepsilon$-CE

$\forall i \in [N] \ \forall f \in F_i$

$$\mathop{E}_{a \sim Q}[\ell_i(a)] - \mathop{E}_{a \sim Q}[\ell_i(a_{-i}, f_i(a_i))] \le \varepsilon$$

דוגמא:

$f_i(1)=2$ $\quad$ $f_j(1)=2$

$f_i(2)=3$ $\quad$ $f_j(2)=2$

$f_i(3)=3$ $\quad$ $f_j(3)=2$

הוכחה: בשלילה, לכל $i$ ו- $f_i \in F_i$

$$\varepsilon < \mathop{E}_{a\sim Q}\left[\ell_i(a) - \ell_i(a_{-i}, f_i(a_i))\right]$$

$$= \sum_{a\in A} Q(a)\left[\ell_i(a) - \ell_i(a_{-i}, f_i(a_i))\right]$$

$$= \frac{1}{T}\sum_{t=1}^{T} \ell_i(a_t) - \ell_i(a_{t,-i}, f_i(a_{t,i}))$$

$$\leq \frac{R(T,|A|)}{T} = \varepsilon$$

☆

## Swap Regret (שתקן חדש)

בהינתן הפסדים $\ell_1 \dots \ell_T$

פעולות $a_1 \dots a_T$

$$SR = \sum_{t=1}^{T} \ell_t(a_t) - \min_{f\in F(A)} \sum_{t=1}^{T} \ell_t(f(a_t))$$

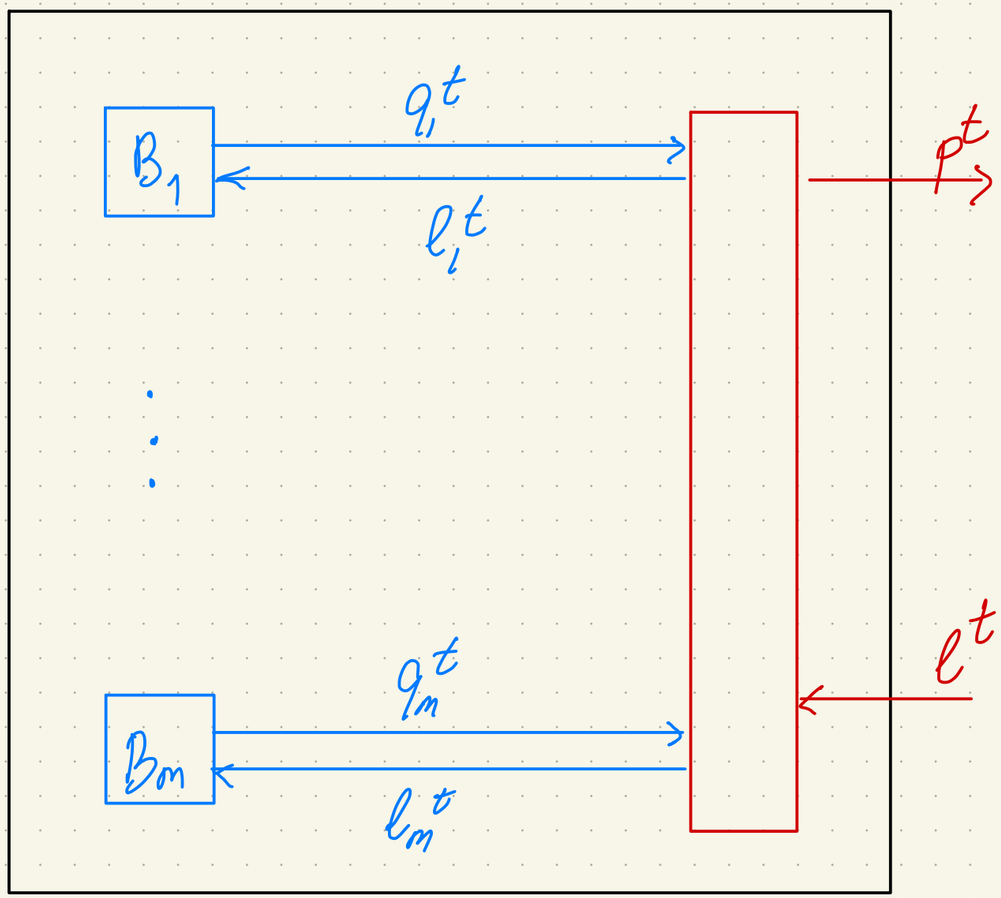
$$F(A) = \{f : A \to A\}$$

משפט: לכל משחק $G$

אם כל שתקן משחק עם $SR \leq R(T,|A|)$

אז ההתפלגות האמפירית $Q$

היא $\varepsilon$-CE, $\varepsilon = \frac{R(T,|A|)}{T}$




# Swap Regret אינו

רבוץ 3 מ SR-N External Regret δ

$|A| = m$

$B_1 \cdots B_m$ אלגוריתם regret

בזמן $t$:

כל $B_i$ מחזיר $q_i^t$

מחזירים $p^t$

מקבלים $\ell^t$

מחזירים $\ell_i^t$ לכל $B_i$

הבסדים $l_i^t$ בהינתן $l^t$ ו-$p^t$

$$l_i^t = p^t(i)\, l_t$$

$$l_i^t(j) = p^t(i)\, l^t(j)$$

התפלגות $p^t$ בהינתן $q_1^t \dots q_m^t$

נבנה מטריצה סטוכסטית $M$

$$M^t = \begin{pmatrix} - & q_1^t & - \\ & \vdots & \\ - & q_m^t & - \end{pmatrix}$$

נתבונן $(p^t)^T M^t = (p^t)^T$ (פילוג סטציונרי)

שתי דרכים שקולות לדגום פעולה:

① ישירות $a_t \sim p^t$

② בעקיפין $i \sim p^t$, $a_t \sim q_i^t$

מהדרך ש $p^t$ המסדר

$$\forall a \in A \qquad \sum_{i=1}^{m} p^t(i)\, q_i^t(a) = p^t(a)$$

לכן השקולות.

## ניתוח ה Swap Regret

אלגוריתם $B_i$ בזמן $t$:

$$\ell_{B_i}^t = (\ell_i^t)^T q_i^t = (p^t(i)\ell^t)^T q_i^t = p^t(i)((\ell^t)^T q_i^t)$$

ה Regret ביחס לפונקציה $j$:

$$L_{B_i}^T = \sum_{t=1}^T \ell_{B_i}^t = \sum_{t=1}^T p^t(i)((\ell^t)^T q_i^t)$$

$$\leq \sum_{t=1}^T \ell_i^t(j) + R_i = \sum_{t=1}^T p^t(i)\ell^t(j) + R_i$$

$j$ זה $i$ לאחר swap שבו הוא

סכום על כל $B_i$:

$$L_{ON}^T = \sum_{t=1}^T \ell_{ON}^t = \sum_{t=1}^T (p^t)^T \ell^t = \sum_{t=1}^T (p^t)^T [M^t] \ell^t$$

$$= \sum_{t=1}^T \sum_{i=1}^m p^t(i)((\ell_t)^T q_i^t) = \sum_{i=1}^m L_{B_i}^T$$

$$\leq \sum_{i=1}^m L_{i\to f(i)}^T + \sum_{i=1}^m R_i$$

$R_i \leq 2\sqrt{T\log m}$ לפי אלג' הבסיס

$$\text{SwapRegret} \leq 2m\sqrt{T\log m}$$

מכאן

$$\text{SwapRegret} \le \sum_{i=1}^{m} R_i$$

$$\le \frac{m \log m}{\eta} + \eta \sum_{i=1}^{m} L_{B_i}^{T}$$

$$\le \frac{m \log m}{\eta} + \eta T$$

עבור $\eta = \sqrt{\frac{m \log m}{T}}$ נקבל

$$\text{SwapRegret} \le 2\sqrt{m T \log m}$$

$$\sqrt{T \underbrace{\log m^{m}}_{\text{מס' פתק'}}}$$

הוכחנו SwapRegret מספר

$$\text{SwapRegret} \le 2\sqrt{m T \log m}$$

לכל i:

$$R_i \le \frac{\log m}{\eta} + \eta L_{B_i}^{T}$$

כמו כן

$$T \ge L_{ON}^{T} = \sum_{i=1}^{m} L_{B_i}^{T}$$

## חסם תחתון

עבור פרמטרים נקבע $m=2k$, $T=\alpha k$

נראה $S.R. = \Omega(\sqrt{Tm}) = \Omega(\sqrt{\alpha}k)$

נגדיר $N_t(i) = \sum_{\tau=1}^{t-1} p^\tau(i)$

נחלק פעולות לזוגות $(2j-1, 2j)$

בתירוץ ההבסדים בזמן $t$:

אם $N_t(2j)+N_t(2j-1) < \frac{1}{2}\alpha$ –

ההבסדים של $(2j-1,2j)$ הם

$$(\ell^t(2j-1), \ell^t(2j)) = \begin{cases} (1,0) & \text{הסת. } 1/2 \\ (0,1) & \text{הסת. } 1/2 \end{cases}$$

אחרת $\ell^t(2j-1) = \ell^t(2j) = 1$

## הבסד אלגו' online:

$$\sum_{i=1}^{m} N_t(i) = \sum_{i=1}^{m}\sum_{\tau=1}^{t-1} p^\tau(i) = \sum_{\tau=1}^{t-1}\underbrace{\sum_{i=1}^{m} p^\tau(i)}_{1} = t-1$$

חסם על ההבסד

$$k\cdot\frac{1}{2}\alpha\cdot\frac{1}{2} + \underbrace{\left(T-\frac{k\alpha}{2}\right)}_{\frac{k\alpha}{2}}\cdot 1 = \frac{3}{4}k\alpha = \frac{3}{4}T$$

הפעולה הטובה ביותר

אם יש צמד כך $N_T(2j-1)+N_T(2j) < \frac{\alpha}{2}$

ישנה פעולה עם הבסד $\geq T/2$

ה regret לפחות $T/4$

אחרת

לכל צמד $(2j-1,2j)$ : $N_T(2j-1)+N_T(2j) \geq \frac{\alpha}{2}$

עבור הצמד יש SR של $\Omega(\sqrt{\alpha})$

סכ"ה $\Omega(\sqrt{Tm}) = \Omega(\sqrt{\alpha}k)$

# אפיון קבוצות של Swap Regret

## פעולות נשלטות (Dominated Actions)

פעולה $a_{i,2}$ נשלטת ע"י $a_{i,1}$ אם

$$\forall a_{-i} \quad \ell_i(a_{-i}, a_{i,1}) \le \ell_i(a_{-i}, a_{i,2})$$

$\varepsilon$-dominated

$$\forall a_{-i} \quad \ell_i(a_{-i}, a_{i,1}) \le \ell_i(a_{-i}, a_{i,2}) - \varepsilon$$

**משפט:** אם נשתמש אלגו' SR ע"ג מספר הפעולות שנשתמש $\varepsilon$-dominated

תסתם ע"י $\varepsilon/SR$

**הוכחה:**

לכל פעולה $a_{i,2}$ ה $\varepsilon$-dominated ע"י $a_{i,1}$

לפי $a_{i,1} = f(a_{i,2})$

אם נשתמש $k$ פעולות $\varepsilon$-dominated

אז ה Swap Regret היחס $f$

לפחות $SR \ge k \cdot \varepsilon$

☆

## קליברציה (Calibration)

איך בודקים תחזית הסתברותית?

מחזי יום שבהם הסתברות 30%

הגורמה תלה בהסתברות 52%

## דרישת מינימום: קליברציה

הממוצע בכל הימים בה

ה תחזית לגשם 20%

קרובה ל 20%

## מודל

בזמן $t$

נותנים תחזית $p_t \in [0,1]$

רואים תוצאה $y_t \in \{0,1\}$

## מודל סטוכסטי

קיימת הסתברות $q$ קבועה

נבחר את הממוצע

$$\mu_t = \sum_{\tau=1}^{t-1} y_\tau \frac{1}{t-1}$$

כיוון

$$\mu_t \longrightarrow q$$

$$|\mu_t - q| \le \frac{1}{\sqrt{t}} \quad (\text{w.p. } 1-\delta)$$

## מודל adversarial

לכל $i$ שהתחזית הוא הסתברות $\frac{i}{m}$
($m+1$ אפשרויות)

נגדיר

$$S\left(\frac{i}{m}\right) = \left\{ t : p_t = \frac{i}{m} \right\}$$

$$\rho\left(\frac{i}{m}\right) = \frac{\sum_{t \in S(\frac{i}{m})} y_t}{\left|S\left(\frac{i}{m}\right)\right|}$$

המטרה: $(\alpha, \varepsilon)$-calibrated

עבור כל $\frac{i}{m}$ כך $\left|S\left(\frac{i}{m}\right)\right| \geq \alpha T$

מתקיים $\left|\rho\left(\frac{i}{m}\right) - \frac{i}{m}\right| \leq \varepsilon$

## תחזיות דטרמיניסטיות

אפשר להכריח כל אלגוריתם דטרמיניסטי
בניית הירים:

אם $\frac{1}{2} > p_t = \frac{i}{m}$ אז $y_t = 1$

אם $\frac{1}{2} \leq p_t = \frac{i}{m}$ אז $y_t = 0$

תמיד מתקיים $\frac{1}{2} \leq |p_t - y_t|$

לאחר $T$ צעדים, יש פעולה שמתקבלו $\frac{T}{m}$
עבורה שגיאת קליברציה גדולה.

משפט:

אין אלגוריתם דטרמיניסטי
שהוא $(\alpha = \frac{1}{m}, \varepsilon \leq 1/2)$ - calibrated

## תוצאות סטטיסטיות:

לכל $i$

$$S\left(\frac{i}{m}\right) = \sum_{t=1}^{T} p_t\left(\frac{i}{m}\right) \qquad p\left(\frac{i}{m}\right) = \frac{\sum_{t=1}^{T} p_t\left(\frac{i}{m}\right) y_t}{S\left(\frac{i}{m}\right)}$$

## משפט (קיום):

קיים אלגוריתם רנדומלי

שבהסתברות $1-\delta$, $\delta = m e^{-\alpha \varepsilon^2 T}$

הוא $(\alpha,\varepsilon)$-calibrated

## הוכחה: נבנה משחק סכום אפס

שורות: $y = (y_1 \dots y_T)$

עמודות: אלגוריתמים

$$f: \text{history} \to \Delta(\{0, \tfrac{1}{m}, \dots, 1\})$$

ערך: עבור $(y,f)$ הוא 0 אם $(\alpha,\varepsilon)$-calibrated אחרת 1.

מהנחת התפלגות $Q$ על $y$ (שורות) שמקיים העמודות

$$p_t = \Pr_Q[y_t = 1 \mid \text{history}]$$

לכל $\delta$ $\frac{i}{m} = \hat{p}_t$

## ערך המשחק:

אם מ"מ נבחר $> \alpha T$

אזי הוא $(\alpha,\varepsilon)$ - Calibrated

אם מ"מ נבחר $\leq \alpha T$

בהסתברות $1-$:

$$\left|\sum_{t\in S(h)} y_t - \frac{i}{m}\right| \leq \frac{1}{m} + \frac{1}{\sqrt{\alpha T}} = \varepsilon$$

ערך המשחק מסוג ע"י $m$

לשם כך העמדות יש לסוג הטיה

שמבטיחה שההסתברות שלנו

$(\alpha,\varepsilon)$-calibrated הוא לפי הוקטור $m$

☆

## תסריטים

① יעילות חישובית

② הסתברות שגוי

נראה רדוקציה ל SwapRegret

פותרת את שתי הבעיות

# רדוקציה אל SwapRegret

נגדיר פונק' הפסד

$$\ell(i, y_t) = \left(y_t - \frac{i}{m}\right)^2$$

פונקציית עלות

$$C = \sum_{t=1}^{T} \sum_{i=1}^{m} p_t(i)\, \ell(i, y_t)$$

$$\geq \sum_{i=1}^{m} \left(\rho\left(\frac{i}{m}\right) - \frac{i}{m}\right)^2 \underbrace{\left(\sum_{t=1}^{T} p_t(i)\right)}_{\text{מס' הפעמים פל מ/i}}$$

נכיף אלגו' SR. על ההפסדים

**למה:** התחזית שמבטאה למינימום את העלות בזמנים בהם התחזית היא מ/i צה $\rho(i/m)$

**למה:** $C \leq S.R. + \frac{T}{m^2}$

**משפט:** עבור $m \geq \frac{2}{\sqrt{\alpha}\,\varepsilon}$ $\quad T \geq m^5 \log m^{-1}$

התחזית הן $(\alpha, \varepsilon)$- calibrated

הוכחת המשפט:

בשלילה: ננית שהסדרה אינה $(\alpha,\varepsilon)$-Calibrated

קיימת פעולה $i$ כך:

$$\sum_{t=1}^{T} p_t(i/m) \geq \alpha T \quad \& \quad \left|p\left(\tfrac{i}{m}\right) - \tfrac{i}{m}\right| \geq \varepsilon$$

אז $C \geq \varepsilon^2 \alpha T$

עבור האלגו' $SR \leq 2\sqrt{mT\lg m}$

לכן

$$\varepsilon^2 \alpha T \leq C \leq \overbrace{2\sqrt{mT\lg m} + \frac{T}{m^2}}^{\text{מבטיחה}}$$

$$\varepsilon^2 \alpha \leq 2\sqrt{\frac{m\lg m}{T}} + \frac{1}{m^2}$$

עבור $T \geq m^5 \log m$

$$\alpha \varepsilon^2 \leq \frac{3}{m^2}$$

לכן

$$m \leq \frac{\sqrt{3}}{\sqrt{\alpha}\,\varepsilon}$$

בסתירה לבחירה על $m$

☆

הוכחת למה: $C \le S.R. + \frac{T}{m^2}$

לכל ערך $z=\rho(\frac{i}{m})$ קיים $\frac{j}{m}$ כך

$$\left|\rho\left(\frac{i}{m}\right)-\frac{j}{m}\right| \le \frac{1}{m}$$

נגדיר $IR\left(\frac{i}{m} \longrightarrow \frac{j}{m}\right)$

מתלבטים כל מופע של מונה ב- $m/j$

$$\sum_{t=1}^{T} p_t\left(\frac{i}{m}\right)\left[\ell(i,y_t)-\ell(j,y_t)\right] = \sum_{t=1}^{T} p_t\left(\frac{i}{m}\right)\left[\left(y_t-\frac{i}{m}\right)^2-\left(y_t-\frac{j}{m}\right)^2\right]$$

$$= \sum_{t=1}^{T} p_t\left(\frac{i}{m}\right)\frac{j-i}{m}\left(2y_t-\frac{i+j}{m}\right)$$

$$= \left(\frac{j-i}{m}\right)\left(\sum_{t=1}^{T} p_t\left(\frac{i}{m}\right)\right)\left(2\rho\left(\frac{i}{m}\right)-\frac{i+j}{m}\right)$$

$$= \left(\sum_{t=1}^{T} p_t\left(\frac{i}{m}\right)\right)\left[\left(\rho\left(\frac{i}{m}\right)-\frac{i}{m}\right)^2-\left(\rho\left(\frac{i}{m}\right)-\frac{j}{m}\right)^2\right]$$

סכום על כל הפעולות

$$SR = \sum_{i=1}^{m}\left(\sum_{t=1}^{T} p_t\left(\frac{i}{m}\right)\right)\left[\left(\rho\left(\frac{i}{m}\right)-\frac{i}{m}\right)^2-\left(\rho\left(\frac{i}{m}\right)-\frac{j_i}{m}\right)^2\right]$$

$$\ge \sum_{i=1}^{m}\left(\sum_{t=1}^{T} p_t\left(\frac{i}{m}\right)\right)\left[\left(\rho\left(\frac{i}{m}\right)-\frac{i}{m}\right)^2-\frac{1}{m^2}\right]$$

$$= \underbrace{\sum_{i=1}^{m}\left(\sum_{t=1}^{T} p_t\left(\frac{i}{m}\right)\right)\left(\rho\left(\frac{i}{m}\right)-\frac{i}{m}\right)^2}_{C} - \frac{T}{m^2}$$

$$SR \ge C - T/m^2$$

☆

מקורות

סיכום הרצאות 2018/19

# סיכום הרצאה

1. Correlated Equilibrium
2. Swap Regret
3. אפליקציות

פעולות נשלטות

קלי ברצה