You against the world

Route-picking

betting on stock market

Repeatedly make decisions, experience consequences

Adaptive decision making

day t $\longrightarrow$ choose among set of possible actions

find out $\longleftarrow$
your cost/loss for day
(possibly losses for all alternatives)

How well can you do?

Formally, each day have n possible actions to choose from

choose $i_t \in [n]$ $\longrightarrow$

$\longleftarrow$ learn $\ell^t = (\ell_1^t, \dots \ell_n^t)$

$$\text{Loss} = \sum_{t=1}^{T} \ell_{i_t}^t$$

Benchmarks:

$$OPT_d = \sum_{t=1}^{T} \min_i \ell_i^t$$

$$L^* = \min_i \underbrace{\sum_{t=1}^{T} \ell_i^t}_{L_i^T}$$

We'll assume each $\ell_i^t \in [0,1]$

consider $\max_{\text{algs}} \min_{\{\ell^t\}}$ performance of alg on $\{\ell^t\}$

Can't compete with OPT_d even with $n=2$

ℓ_1^t equally likely to be 0 or 1 $\quad \ell_2^t = 1 - \ell_1^t$

Alg has loss $= \frac{T}{2}$ $\quad OPT_d = 0$

Instead consider $\boxed{\text{regret}}$

$$\text{Regret} = \text{Regret}(A, \{\ell^t\}) = \sum_{t=1}^{T} \text{Loss of } A \text{ on day } t - \min_i L_i^T$$

Multiplicative Weight Updates Alg

Fix $\varepsilon < \frac{1}{2}$; n possible actions. on each day t, w_i^t = weight of expert i

initially $w_i^0 = 1 \quad \forall i$

on day t, use mixed strategy $\vec{p}^t$ where $p_i^t = \frac{w_i^{t-1}}{\sum_k w_k^{t-1}}$

for each action i, observe the loss $\ell_i^t \in [0,1]$

and update the weight as follows:

$$w_i^t := w_i^{t-1} e^{-\varepsilon \ell_i^t}$$

Thm: Given any loss sequence $\{\ell^t\}_{t=1}^T$, let $L^T = \sum_{t=1}^T p^t \cdot \ell^t$

where $p^{(t)}$ are the prob distns used by MWU

Then $\forall i$

$$L^T \leq L_i^T + \underbrace{\frac{T\varepsilon}{8}}_{\text{we'll prove } \frac{T\varepsilon}{2}} + \frac{\log n}{\varepsilon} \qquad \text{where } L_i^T = \sum_{t=1}^T \ell_i^t$$

Corollary: with $\varepsilon = \sqrt{\frac{8 \log n}{T}}$, $\forall \{\ell^t\}_{t=1}^T$

$$\text{Regret (MWU)} \leq \sqrt{\frac{T \log n}{2}}$$

Proof of thm:

Let $W^t = \sum_{i=1}^n w_i^t = \sum_{i=1}^n w_i^{t-1} e^{-\varepsilon \ell_i^t}$

$$\frac{W^t}{W^{t-1}} = \frac{\sum w_i^{t-1}}{W^{t-1}} = \sum_{i=1}^n p_i^t e^{-\varepsilon \ell_i^t} = E\left(e^{-\varepsilon X_t}\right) = e^{-\varepsilon p^t \cdot \ell^t} \underbrace{E\left(e^{-\varepsilon (X_t - p^t \cdot \ell^t)}\right)}_{\leq e^{\varepsilon^2/2}}$$

where $X_t = \ell_i^t$ with prob p_i^t

[actually $\leq e^{\varepsilon^2/8}$]

$$\leq e^{-\varepsilon p^t \cdot \ell^t} e^{\varepsilon^2/2}$$

$$\Rightarrow W^T \le e^{-\varepsilon p^t \ell^t} e^{\varepsilon^2/2} W^{T-1}$$

$$\le e^{-\varepsilon L^T} e^{\frac{T\varepsilon^2}{2}} \cdot n$$

OTOH $\quad W^T \ge w_i^T = e^{-\varepsilon L_i^T}$

So $\quad e^{-\varepsilon L_i^T} \le e^{-\varepsilon L^T} e^{\frac{T\varepsilon^2}{2}} n$

Taking logs $\quad -\varepsilon L_i^T \le -\varepsilon L^T + \frac{T\varepsilon^2}{2} + \ln n$

$$\Rightarrow L^T \le L_i^T + \frac{T\varepsilon}{2} + \frac{\ln n}{\varepsilon}$$

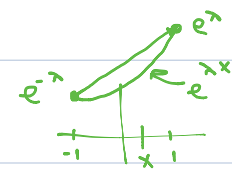
$$\frac{T\varepsilon}{2} = \frac{\ln n}{\varepsilon} \Rightarrow \varepsilon^2 = \frac{2\ln n}{T} \equiv \varepsilon = \sqrt{\frac{2\ln n}{T}}$$

$$\Rightarrow L^T \le L_i^T + \sqrt{2T\ln n}$$

Hoeffding Lemma: Let X be r.v. with $E(X)=0$ and $|X| \le 1$

Then $E(e^{\lambda X}) \le e^{\lambda^2/2}$

Proof: $e^{\lambda x}$ is convex $\Rightarrow e^{\lambda x} \le \frac{(1+x)e^{\lambda} + (1-x)e^{-\lambda}}{2}$ for $x \in [-1,1]$



Since $|X| \le 1$ & $E(X)=0$ $\quad E\left[\frac{(1+X)e^{\lambda} + (1-X)e^{-\lambda}}{2}\right] = \frac{e^{\lambda}+e^{-\lambda}}{2} = \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{(2k)!}$

$$\le \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{2^k k!} = e^{\lambda^2/2} \qquad (2k)! > 2^k k!$$

$\sqrt{T}$ unavoidable

Say $n=2$ flip coin each day

$H \to \ell_1^t=1 \quad \ell_2^t=0$
$T \to \ell_1^t=0, \ \ell_2^t=1$

Any alg has exp loss $\frac{T}{2}$

$L_1^T = \#$ H's at end
$L_2^T = \#$ T's at end

$E[\min(\#H's, \#T's)] = \frac{T}{2} - c\sqrt{T}$

In fact above bound $\sqrt{\frac{T \log n}{2}}$ is tight up to lower order terms

(analyse case where each ℓ_i^t indep uniform $\{0,1\}$

Applications: (started with game playing)

① Route-picking

② Learning: choosing prediction from class of fns $\{f_1, \ldots, f_r\}$

given x_t (data for day t), choose f_i, incur loss $\ell(f_i(x_t), y_t)$

low regret means that in hindsight do nearly as well as best fn in class

③ Pricing: sequence of customers, each with value v_i for items you're selling

Application to setting with gains & where $\forall t$ $\max_i g_i^t - \min_j g_j^t \le \rho$

Let $g_{max}^t = \max_k g_k^t$

Define $\ell_i^t = \frac{1}{\rho}(g_{max}^t - g_i^t)$

Regret (MWU) $\le \sqrt{\frac{T \log n}{2}}$

$$\Rightarrow L_{MWU}^T \le L_j^T + \sqrt{\frac{T \log n}{2}}$$

$$\sum_t (g_{max}^t - g^t \cdot p^t) \le \sum_t (g_{max}^t - g_j^t) + \rho \sqrt{\frac{T \log n}{2}}$$

$$\Rightarrow \text{MW}\overset{\text{exp.}}{\text{gain}} \ge G_j^t - \rho \sqrt{\frac{T \log n}{2}} \quad \forall j$$

Application: portfolio selection

· investing money in stocks/commodities/currencies

each day reinvest

$\dfrac{\text{closing price}}{\text{opening price}}$

		proportions
Stock 1	1.5	$\frac{1}{2}$
Stock 2	1	0
Stock 3	1	$\frac{1}{2}$
Stock 4	0.5	0

at end of day

$\$1 \Rightarrow \frac{1}{2}1.5 + \frac{1}{2}1 = \$1.25$

p_i^t fraction of wealth you put into stock i on day t

$$r_i^t = \frac{\text{closing price of stock } i \text{ on day } t}{\text{opening price on day } t}$$

If invested all your money in i

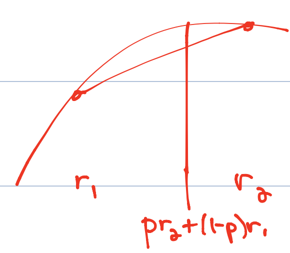
$$\frac{\text{final wealth}}{\text{initial}} = r_i^1 r_i^2 \cdots r_i^t$$

$$\log \frac{W^T}{W^0} = \sum \log r_i^t$$

if we choose portfolio each day using MW

$$\sum_{t=1}^{T} \underbrace{\sum_i p_i^t \log r_i^t}_{\leq \log\left(\sum_i p_i^t r_i^t\right)} \geq \sum_t \log r_i^t - \varepsilon \quad \forall i$$

log concave fn



$$\Rightarrow \sum_{t=1}^{T} \log\left(\sum_i p_i^t r_i^t\right) \geq \log \prod_t r_i^t - \varepsilon$$

$$= \log \prod \sum_i p_i^t r_i^t$$

$$\frac{\text{Final wealth}_{MW}}{W_0} \geq \frac{W_i^T}{W_0} e^{-\varepsilon}$$

Application to playing 2-player zero-sum games

game defined by matrix

row player is R, col player is C; entry is payoff to R

Example 1: Penalty Kicks

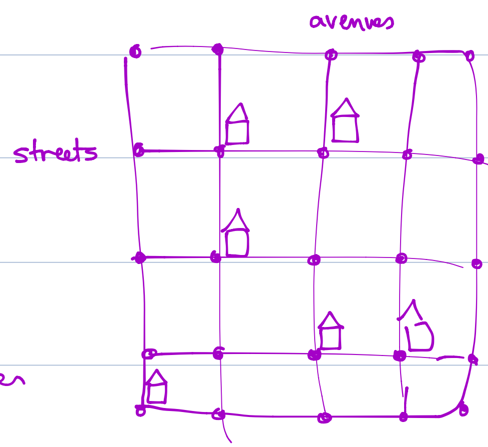
		goalee L	goalee R
kicker	L	1	0
	R	0	1

Example 2 Hide & Seek

Cop picks street or avenue

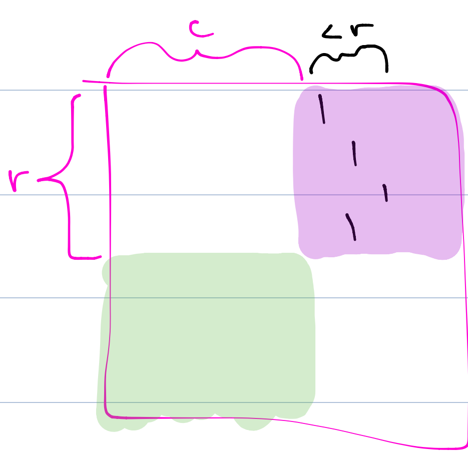
Robber picks a safehouse

Cop's payoff is 1 if picks street or avenue where robber is hiding



König's Lemma: max matching = minimum line (vertex) cover

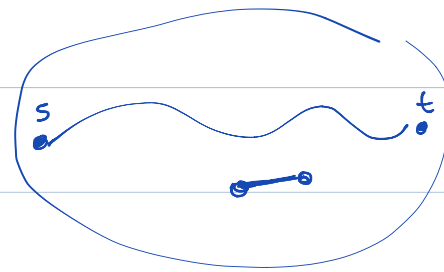
Pf: max matching $\le$ min vertex cover



if no matching of size r in purple part, can replace some rows in r by columns in purple

Example 3 "Smuggler vs border guard" (Blum)

Graph G, source s, sink t



smuggler chooses path

border guard chooses edge to watch

If edge is in path, guard wins, else smuggler wins

Menger's Thm: Min s-t cut (edges) = max # edge disjoint s-t paths

minimax opt strategy for row player

① $V_R = \max_p \min_j \sum_i p_i a_{ij}$

what R would play if she had to reveal strategy first

minimax opt strategy for col player

② $V_C = \min_q \max_i \sum_j a_{ij} q_j$

what C would play if he had to reveal strategy first

Clearly ① < ②

von Neumann 1928 $\qquad V_R = V_C = V$

no loss to publish strategy

Proof: Suppose $V_C > V_R + \varepsilon$ $\qquad$ Scale payoffs to $[0,1]$

Have C play RWM against R responding optimally

In T steps: C's loss $\leq \sum_{t=1}^{T} a_{i_t j} + c\sqrt{T \log n} \quad \forall j$

$$\leq T V_R + \sqrt{T \log n}$$

C's loss $=$ R's gain $\geq T V_C$ $\qquad \Rightarrow$ $T V_C \leq T V_R + \sqrt{T \log n}$

$V_C - V_R \leq \sqrt{\frac{\log n}{T}}$

$V_C - V_R \leq \sqrt{\frac{\log n}{T}} \longrightarrow \leftarrow$ for T sufficiently large

Note: if opponent not playing optimally, do much better than V

MWU gives fast way to compute approx optimal strategies

Let $q^* = \frac{1}{T}\sum_{t=1}^{T} p^{(t)}$ and let i^* be best response of R to q^*

$$\left(\frac{1}{T}\sum_{t=1}^{T} p_j^t\right) a_{i^* j} \leq \frac{1}{T}\sum_{j=1}^{n} p_j^t a_{i_t j} \leq V + \varepsilon$$

i_t is best response at each step $\qquad$ MW guarantee

$$\leq \min_j \frac{1}{T}\sum_{j=1}^{n} a_{i_t j} + \varepsilon$$

Let $p_i^* = \frac{|\{t \mid i_t = i\}|}{T}$ empirical drstn

$\forall j \ \frac{1}{T} \sum_{t=1}^{T} a_{i_t j} + \varepsilon \geq$ MW

$$\frac{1}{T} \sum_{t=1}^{T} a_{i_t j} \geq \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{n} p_j^t a_{i_t j} - \varepsilon$$

$$\geq \underbrace{V}_{= \min_{\vec{q}} \max_{i} \sum_{j=1}^{n} q_j a_{ij}} - \varepsilon$$

Solving LP

$$\exists x \geq 0 \quad \text{s.t.} \quad \underbrace{A}_{m \times n} x \geq b$$

Use MWU to find $x \geq 0$ s.t. $\forall i \ \ a_i x \geq b_i - \varepsilon$ or prove system infeasible

Requires $\underset{\in \mathbb{R}^m}{\vec{p}} \rightarrow$ oracle

→ yes $\exists x \geq 0$ s.t. $p^T A x \geq p^T b$ and $|a_i x - b_i| \leq \rho$

→ no system infeasible

Thm: Assume $0 < \eta \le 2\rho$

Then $\exists$ alg that either finds $x \ge 0$ s.t. $\forall i \; a_i x \ge b_i - \varepsilon$

or correctly concludes that system is infeasible

$O\left(\rho^2 \frac{\log m}{\varepsilon^2}\right)$ calls to oracle with additional processing time of $O(m)$ per call

call oracle

get $x^{(t)}$

$$\text{MWU} \xrightarrow[\text{distn over } m \text{ constraints}]{p^{(t)}}$$

$$\xleftarrow{\quad} \; \ell_i^{(t)} = \frac{1}{\rho}\left[a_i x^{(t)} - b_i\right]$$

$$p^{(t)} \cdot \ell^{(t)} = p^{(t)} \cdot \frac{1}{\rho}\left[A x^{(t)} - b\right] \ge 0 \qquad \text{guarantee of oracle}$$

Thm about MWU guarantees

$$0 \le \frac{1}{\rho}\sum_{t=1}^{T}\left(a_i x^{(t)} - b_i\right) + \frac{\eta}{\rho}\sum_t \overbrace{\left|a_i x^{(t)} - b_i\right|}^{\le \rho} + \frac{\ln m}{\eta}$$

$$\le \sum_t \left(a_i x^{(t)} - b_i\right) + \eta \rho T + \frac{\ln m}{\eta}\rho$$

$$\Rightarrow \frac{1}{T}\sum_t\left(a_i x^{(t)} - b_i\right) \ge \eta\rho + \frac{\rho \ln m}{\eta T} \Rightarrow a_i\left(\frac{1}{T}\sum_t x^{(t)}\right) - b_i \ge \eta\rho + \frac{\rho \ln m}{\eta T}$$

choose $\eta = \frac{\varepsilon}{4\rho}$ and $T = \frac{8\rho^2 \ln m}{\varepsilon^2}$

and rhs becomes $\frac{\varepsilon}{4} + \frac{\rho \ln m (16 \eta^2 \rho^2)}{\eta \, 8 \rho^2 \ln m}$

$= \frac{\varepsilon}{4} + \frac{2\rho}{\eta}$

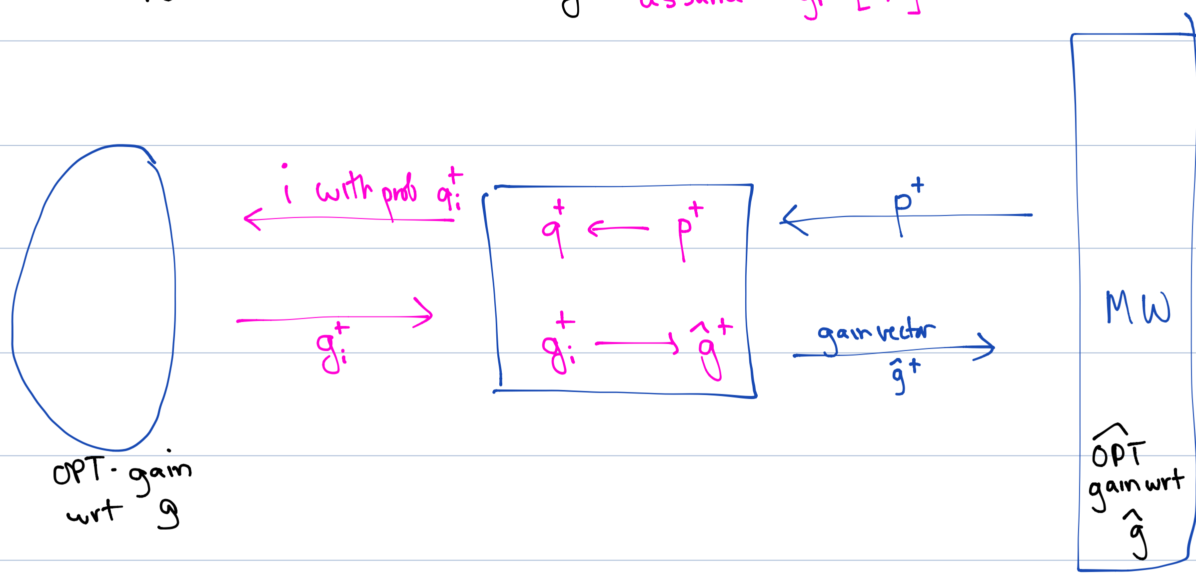
Example: flow $\quad\quad \max \sum_p y_p$

$$\sum_{p \mid e \in p} y_p \le c_e$$

$$-\sum_{p \mid e \in p} \frac{y_p}{c_e} \geq -1$$

Multi-armed bandit Problem [Auer, Cesa-Bianchi, Freund Schapire]

reduction to standard setting assume $g_i^t \in [0,h]$



$$q^t \longleftarrow p^t$$

$$q^t = (1-\gamma)p^t + \gamma(\text{uniform}) \qquad \gamma \text{ small}$$

$$\left(q_i^t \geq \frac{\gamma}{n}\right)$$

$$g_i^t \longrightarrow \hat{g}^t$$

$$\hat{g}^t = (0,0,\dots,\underbrace{\frac{g_i^t}{q_i^t}}_{\leq \frac{nh}{\gamma}},0,\dots,0) \qquad \hat{g}^t \text{ is r.v.} \qquad E(\hat{g}^t) = g^t$$

$$\text{MW} \Rightarrow \sum_t \hat{g}^t \cdot \hat{p}^t \geq E\left(\max_j \hat{G}_j\right) - \frac{nh}{\gamma}\sqrt{\frac{T \log n}{2}}$$

$$\sum_t q^t \cdot g^t \geq \sum_t (1-\gamma) \sum_i p_i^t g_i^t = \sum_t (1-\gamma) \sum_i p_i^t E(\hat{g}_i^t)$$

Claim: $E(\hat{G}_j^T) = G_j^T$ $\qquad E(\hat{g}_j^t) = (1-q_j^t)\,0 + q_j^t\left(\frac{g_j^t}{q_j^t}\right) = g_j^t$

$$E\left[\max_j \hat{G}_j^T\right] \geq \max_j E\left(\hat{G}_j^T\right) = \max_j G_j^T$$