LECTURE 3

2/1/21

- Homework I
- Scribing Notes

PLAN: 1) Proof of Mirror Descent

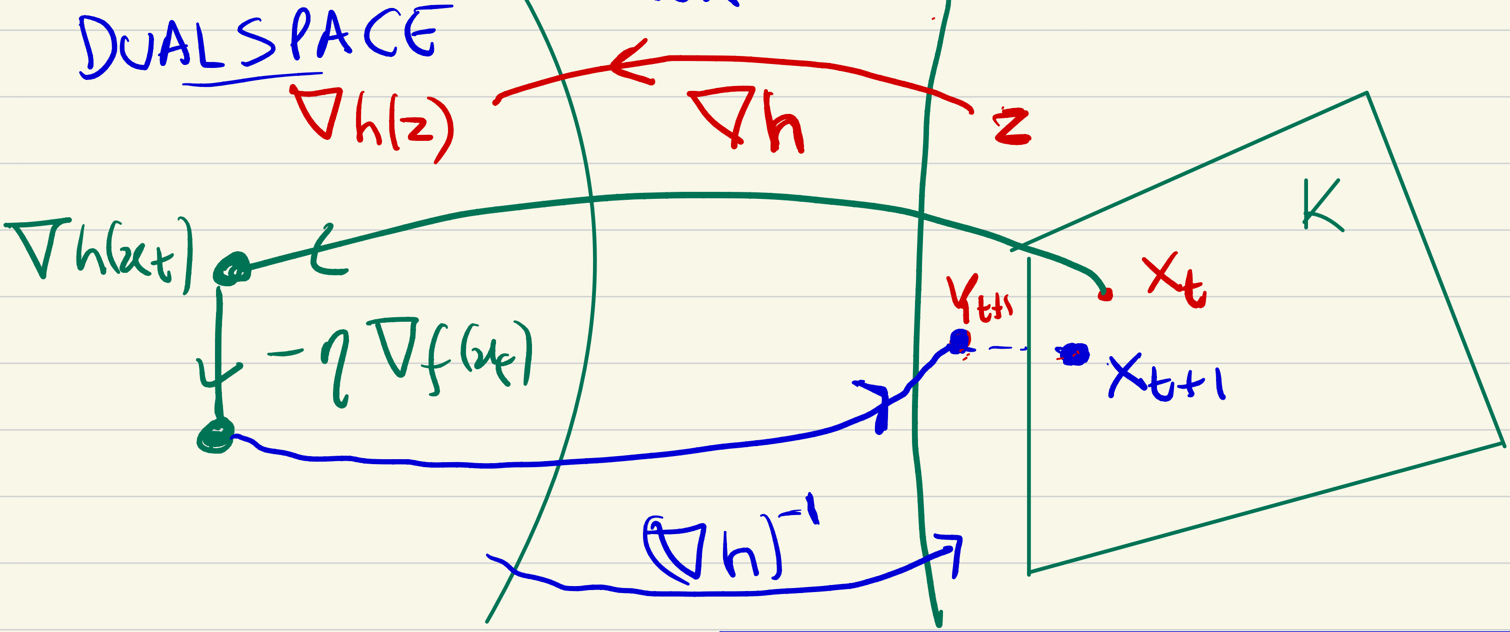
2) Online convex optimisation

- Solve LPs.

MIRROR DESCENT

$\min_{x \in K} f(x)$

$h : \Omega \to \mathbb{R}$



1) $$Y_{t+1} \leftarrow \operatorname*{argmin}_{y \in \Omega} \boxed{\eta \left[f_t(x_t) + \langle \nabla f_t(x_t), y - x_t \rangle \right] + D[y \| x_t]}$$

$$\Downarrow$$

$$\eta \cdot \nabla f_t(x_t) + \nabla h(y) - \nabla h(x_t) = 0$$

$$\Downarrow$$

$$\boxed{\nabla h(y) = \nabla h(x_t) - \eta \nabla f(x_t)}$$

$$\boxed{y = (\nabla h)^{-1} \left[\nabla h(x_t) - \eta \nabla f_t(x_t) \right]}$$

2) $$X_{t+1} = \operatorname*{argmin}_{z \in K} D(z \| Y_{t+1})$$

[ρ- strongly convex] $h : \Omega \to \mathbb{R}$ such that

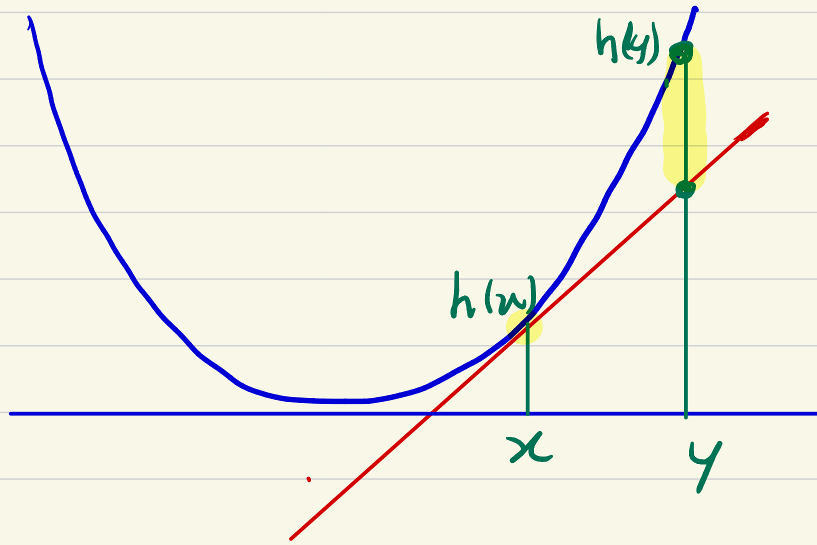
$$\forall x, y : \quad h(y) \geq \underbrace{\left[h(x) + \langle \nabla h(x), y - x \rangle \right]} + \frac{\rho}{2} \underbrace{\|x - y\|^2}$$

$\| \; \|_\infty \quad \| \; \|_1$

[Bregman Divergence]

$$D_h[y \| x] = h(y) - \left[h(x) + \langle \nabla h(x), y - x \rangle \right]$$

asymetric



$$\boxed{h(x) = \|x\|^2}$$

THEOREM: 1) h is ρ-strongly convex with respect to $\|\ \|_p$

2) $\nabla h : \Omega \to \mathbb{R}^n$ is a bijection

3) $\|\nabla f_s(x)\|_* \leq L$

Then
$$\frac{1}{t}\sum_{s=1}^{t} f_s(x_s) - f_s(x_*) \leq O\left[L\underbrace{\sqrt{\frac{D(x_*\|x_0)}{\rho t}}}_{O(1/\sqrt{t})}\right]$$

Proof: $\forall s$
$$f_s(x_s) - f(x_*) \leq \langle \nabla f_s(x_s), x_s - x_* \rangle$$

[Definition of convexity]

$$= \frac{1}{\eta}\langle \underbrace{\nabla h(x_s) - \nabla h(y_{s+1})}_{\|}, x_s - x_* \rangle$$

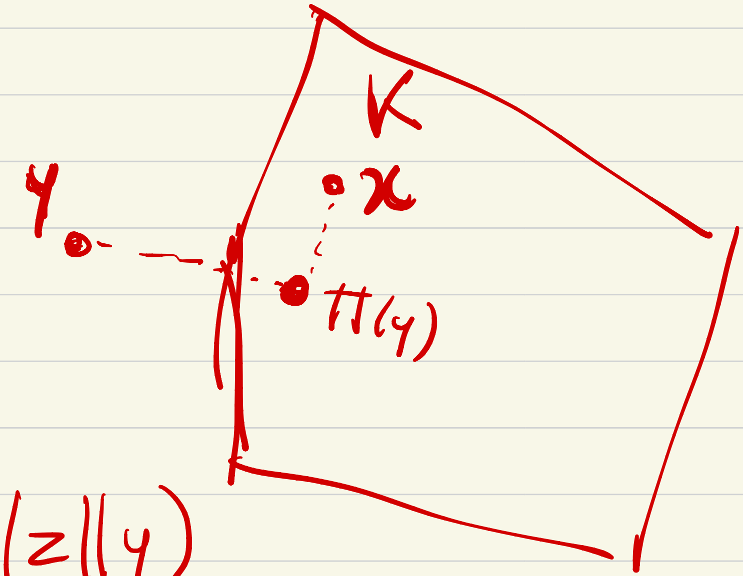
$$= \frac{1}{\eta}\left[D(x_*\|x_s) + D(x_s\|y_{s+1}) - \underbrace{D(x_*\|y_{s+1})}\right]$$

[law of cosine]

$$D[x_*|x_{s+1}] + D(x_{s+1}\|y_{s+1})$$

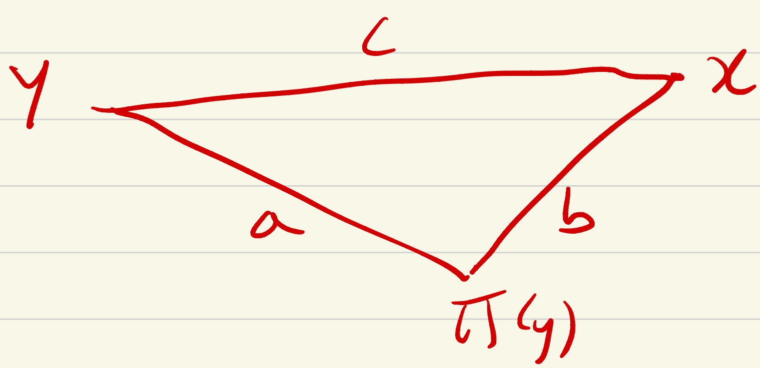
$\forall x \in K$

$y \in \Omega$



$$\Pi(y) = \min_{z \in K} D(z \| y)$$

$$\boxed{D(x \| y) \geq D(x \| \pi(y)) + D(\pi(y) \| y)}$$



$c^2 \geq a^2 + b^2$

$f_s(x_s) - f_s(x_*)$

$$\leq \frac{1}{\eta}\Big[D(x_* \| x_s) + D(x_s \| y_{s+1}) - D(x_* \| x_{s+1}) - D(x_{s+1} \| y_{s+1}) \Big]$$

[Pythagorean identity]

$$\sum_{s=1}^{t} f_s(x_s) - f_s(x_*) \leq \frac{1}{\eta}\Big[D[x_* \| x_1] - D(x_* \| x_{t+1}) \Big]$$

$+$

$$D[x_s \| y_{s+1}] - D[x_{s+1} \| y_{s+1}]$$

$$= \Big[\eta \underline{\langle \nabla f_s(x_s), x_s - x_{s+1} \rangle} - \rho/2 \, \|x_s - x_{s+1}\|^2 \Big]$$

$$= \eta \, \|\underline{\underline{\nabla f_s}}(x)\|_* \|x_s - x_{s+1}\|_p - \frac{\rho}{2} \|x_s - x_{s+1}\|_p^2$$

$$\leq \frac{(\eta L)^2}{2\rho}$$

Exercise: $h(x) = -\sum x_i \log x_i$

$\Downarrow$

$D(y\|x) = $ KL divergence

$\Downarrow$

h is strongly convex w.r.t $\underline{1}$-norm

Multiplicative weights.

Online Convex Optimization

Algorithm: At time t,

it plays $x_t \in K$

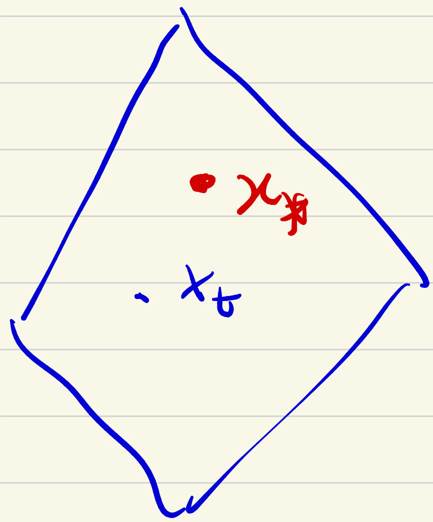
Adversary/Cost: $f_t : K \to \mathbb{R}$

convex cost function

[Can choose any function]

could depend on x_t

$\boxed{\text{Cost} = f_t(x_t) \text{ in round } t}$



After T:

$$\text{Regret}_T = \sum_{s=1}^{T} f_s(x_s) - \min_{x \in K} \sum_{s=1}^{T} f_s(x)$$

$$\text{Avg Regret} = \frac{1}{T} \text{Regret}_T$$

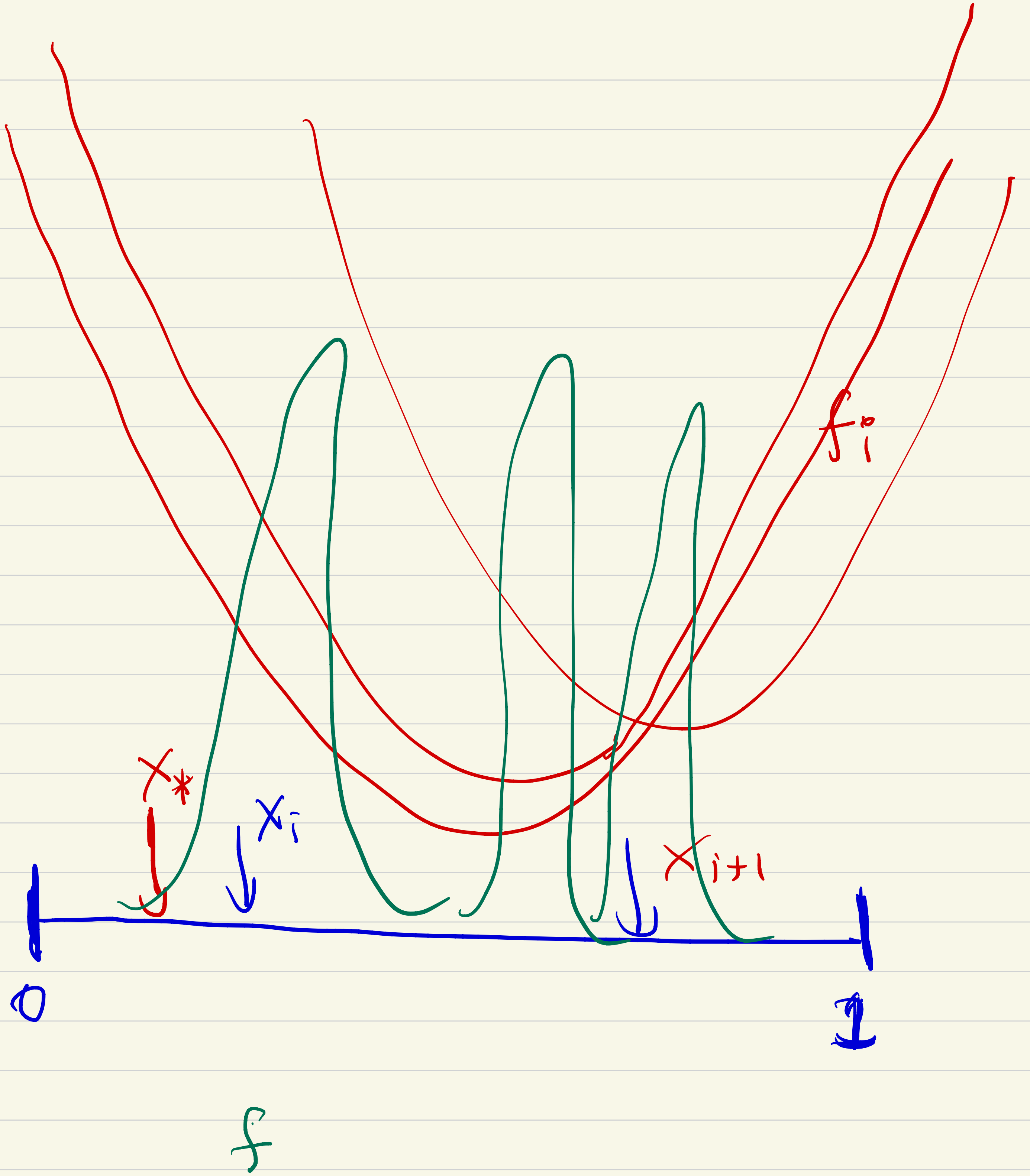
PARADOX:

Action

$$\underline{\underline{X_{t+1}}} \in \left\{ \begin{array}{l} \text{function of the} \\ \text{past } f_1, \dots f_t \\ \quad x_1, \dots x_t \end{array} \right\}$$

Reward / Cost: f_{t+1} independent of the past

Avg Regret : $O\left(\frac{1}{\sqrt{t}}\right) \to 0$ as $t \to \infty$



f_i
x_*
x_i
x_{i+1}
0
1
f

Solving LP

$$x_1 \ldots x_n \in [-B, B]$$

$$\forall i = 1..m \quad \boxed{l_i(x) = \langle w_i, x\rangle - b_i \leq 0}$$

Find $x \in$ feasible (satisfies all constraints)

ALG TO SOLVE LPS

Online Convex Optimization

(Online GD)

candidate solution x_t

$f_t = l_i$

Adversary:

1) Pick a constraint violated by x_t

$l_i(x_t) \geq \varepsilon$

If Adversary Fails: No constraint violated by ε, Return X_t

Suppose $\exists\, \underline{\underline{x_*}} \in [-B, B]^n$

that satisfies all the constraints:

$$\sum_{s=1}^{T} f_s(x_*) = \sum_{s=1}^{T} \underbrace{l_{i_s}(x_*)}$$

$$= \sum_{s=1}^{T} (<0) < \underline{\underline{0}}$$

On the other hand, if (Adv-OCO)

runs for T steps,

$$\sum_{s=1}^{T} f_s(x_s) \geq \sum_{s=1}^{T} \varepsilon = \underline{\underline{\varepsilon T}}$$

$$\boxed{\text{Regret}_T \geq \varepsilon T} \Rightarrow \underline{\underline{\text{Avg Regret} = \varepsilon}}$$

Runtime = $\text{poly}(1/\varepsilon)$

Solution with $\underline{\underline{b}}$-bits of precision $\to \text{poly}(b)$

$\to$ Time = $2^{O(b)}$ $\quad \boxed{\varepsilon = 2^{-b}}$

Gold standard $\to$ $\boxed{\text{poly}\left(\log\left(1/\varepsilon\right)\right)}$ $= \text{poly}(b)$

Khachiyan

$$\boxed{F(x) = \sum -\log l_i(x)}$$