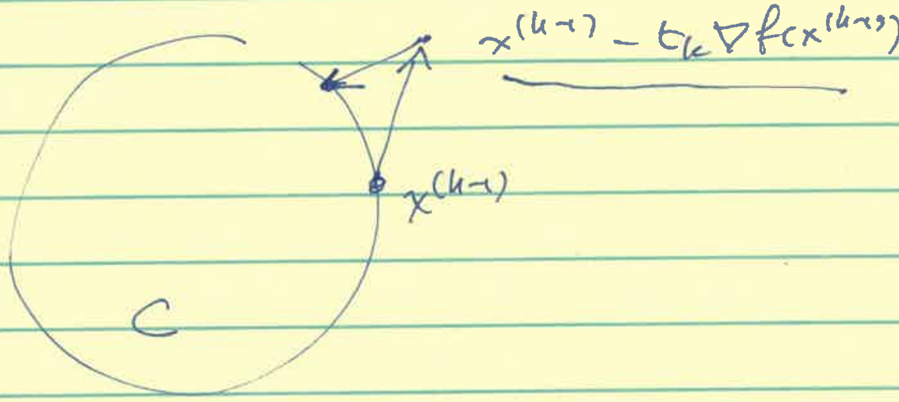
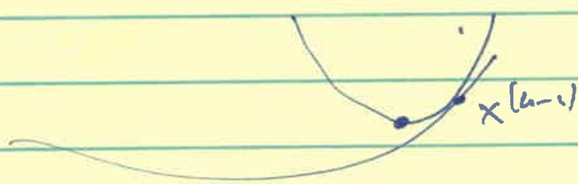
$$\|Ax + Bz - c + w\|_2^2$$



$C = \{x : Ax \leq b\}$. general A, b. P_C is hard!

$C = \{x : 1^T x, x \geq 0\}$ P_C is linear-time



$$\min \; a^T x$$

$$\cancel{P_C \left(\underset{y}{\operatorname{argmin}} \; \nabla f(x^{(k-1)})^T (y - x^{(k-1)}) \right)}$$

doesn't make sense

$$\underset{y \in C}{\operatorname{arg\,min}} \; \nabla f(x^{(k-1)})^T (y - x^{(k-1)}).$$

"linear minimization oracle"

$$s^{(k-1)} = \underset{s \in C}{\operatorname{argmin}} \; \nabla f(x^{(k-1)})^T s$$

$$x^{(k)} = (1 - \gamma_k) x^{(k-1)} + \gamma_k s^{(k-1)}$$

$$\min_{\|s\|\le t} \nabla f(x)^T s$$

$$= - \max_{\|s\|\le t} - \nabla f(x)^T s$$

$$= - t \cdot \max_{\|z\|\le 1} - \nabla f(x)^T z$$

$$= -t \max_{\|z\|\le 1} \nabla f(x)^T z$$

$$\operatorname{argmin}_{\|s\|\le t} \nabla f(x)^T s = -t \underbrace{\operatorname{argmax}_{\|s\|\le 1} \nabla f(x)^T s}$$

$$= \underline{-t \cdot \partial \|\nabla f(x)\|_*}$$

$$\|z\|_* = \max_{\|s\|\le 1} z^T s \quad \text{dual norm}$$

rule for subgrads of max.

$$\left(\|\cdot\|_p\right)_* = \|\cdot\|_q \quad \text{where } \frac{1}{p} + \frac{1}{q} = 1.$$

$$f^* \geq f(x^{(k)}) + \nabla f(x^{(k)})^T (s^{(k)} - x^{(k)})$$

$$\nabla f(x^{(k)})^T (x^{(k)} - s^{(k)}) \geq f(x^{(k)}) - f^* \quad \checkmark$$

∇f being Lipschitz $\Rightarrow$

$$f(y) \leq f(x) + \nabla f(x)^T (y-x) + \frac{L}{2} \|y-x\|_2^2$$

$$\max_{\substack{x, s \in C \\ \gamma \in [0,1] \\ y = (1-\gamma)x + \gamma s}} \frac{1}{\gamma^2} \|y - x\|_2^2$$

$$= \max_{\substack{x, s \in C \\ \gamma \in [0,1]}} \frac{1}{\gamma^2} \|(1-\gamma)x + \gamma s - x\|_2^2$$

$$= \max_{s, x \in C} \|x - s\|_2^2.$$

$x \in C$, $x = Ax' \iff x' \in A^{-1}C$.

$$(x')^+ = (1-\gamma)x' + \gamma s'$$

$$A(x')^+ = (1-\gamma)\underbrace{Ax'}_{=x} + \gamma \underbrace{As'}$$

if this's = s
then we would
get back usual
FW update on f

$$As' = A \cdot \operatorname*{argmin}_{z \in A^{-1}C} \nabla F(x')^T z$$

$$= A \cdot \operatorname*{argmin}_{z \in A^{-1}C} \nabla f(Ax')^T Az$$

$F(x') = f(Ax')$

$\nabla F(x') = A^T \nabla f(Ax')$

$$= AA^{-1} \operatorname*{argmin}_{w \in C} \nabla f(x)^T w$$

$$= \operatorname*{argmin}_{w \in C} \nabla f(x)^T w$$