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

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

max is a duality gap

ie. $f(x) - f^* \leq$ this max

$$I_C(x) = \begin{cases} 0 & x \in C \\ \infty & x \notin C \end{cases}$$

$$I_C^*(x) = \max_{y \in C} y^T x$$

Fenchel: $f(x) + f^*(u) \geq x^T u$

$$x^T \nabla f(x) + \max_{s \in C} -\nabla f(x)^T s$$

$$= \max_{s \in C} \nabla f(x)^T (x - s)$$

$$f(y) - f(x) - \nabla f(x)^T (y - x) \overset{?}{\geq} 0$$

$$\frac{2M}{k+2} \approx \frac{2\,\mathrm{diam}^2(C) L}{k+2}$$ cond grad

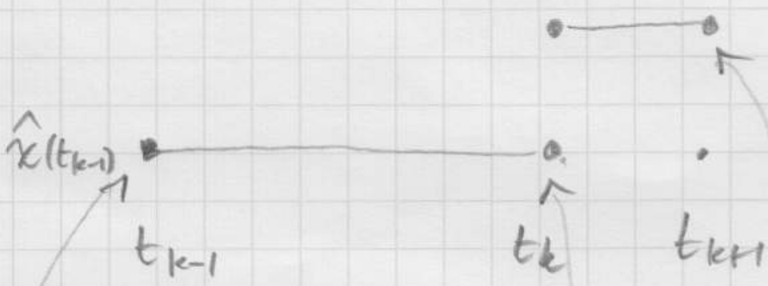
$$\frac{2 \|x^{(0)} - x^*\|_2^2 L}{k+2}$$ proj grad

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

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

$$x^{(k)} = \underset{\substack{x \in \operatorname{conv}\{ \\ \{x^{(0)}, \dots s^{(k-1)}\}}}{\operatorname{argmin}} \; f(x)$$

Backward steps.



i.e. set $\hat{x}(t) = \hat{x}(t_{k-1})$ all $t \in (t_{k-1}, t_k)$

then subopt gap is $\le \varepsilon$ along this segment

Frank-Wolfe w. accuracy $\frac{\varepsilon}{m}$

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

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

$\|\nabla f(x)\|_*$