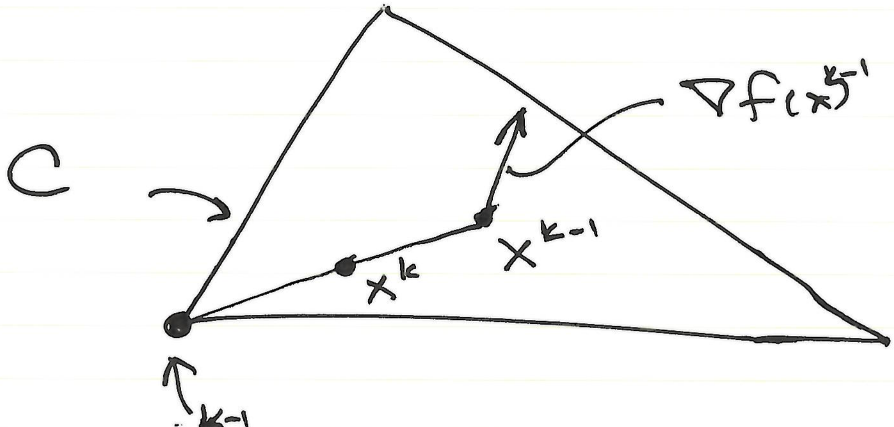


$$s^{k-1} = \operatorname{argmin}_{s \in C} \langle \nabla f(x), s \rangle$$

Consider $\min_x \; f(x)$

$x \in \operatorname{conv}(A)$

$A \subseteq \mathbb{R}^n$ set of atoms s.t.

$g \mapsto \operatorname{argmin}_{a \in A} \langle g, a \rangle$ is computable

---

Recall

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

$$I_C^*(s) = \max_x \{\langle s, x \rangle - I_C(x)\}$$

$$= \max_{x \in C} \langle s, x \rangle = \text{support funct of } C \text{ at } s$$

Observe: $\forall x, y \in C$

$$f(y) - f(x) - \langle \nabla f(x), y - x \rangle$$

Proof of Thm: by induction.

Assume $\quad f(x^{(k-1)}) - f^* \leq \dfrac{2M}{k+1}$

Apply basic ineq:

$$f(x^{(k)}) \leq f(x^{(k-1)}) - \gamma_k g(x^{(k-1)}) + \frac{\gamma_k^2}{2} M$$

$$g(x^{(k-1)}) \geq f(x^{(k-1)}) - f^* \quad \& \quad \gamma_k = \frac{2}{k+1}$$

Thus $\quad f(x^{(k)}) \leq f(x^{(k-1)}) - \dfrac{2}{k+1}\left(f(x^{(k-1)}) - f^*\right) + \dfrac{4}{2(k+1)^2} M$

$$f(x^{(k)}) - f^* \leq \left(1 - \frac{2}{k+1}\right)\left(f(x^{(k-1)}) - f^*\right) + \frac{2M}{(k+1)^2}$$

$$\leq \frac{k-1}{k+1} \cdot \frac{2M}{k+1} + \frac{2M}{(k+1)^2} =$$

$$= \frac{2M}{\cancel{}} \underbrace{\left(\frac{k}{(k+1)^2}\right)}_{\leq \frac{1}{k+2}} \leq \frac{2M}{k+2}$$

Recall:

Gradient Descent $\quad x_+ = x - t\nabla f(x)$

Newtons Method $\quad x_+ = x - \nabla^2 f(x)^{-1} \nabla f(x)$

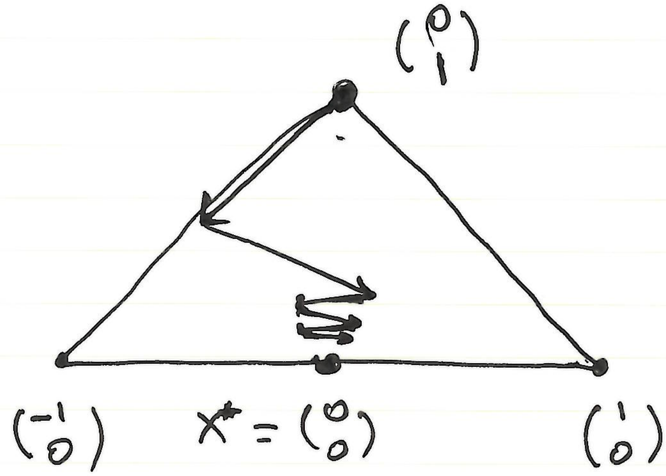
Gradient Descent is <u>NOT</u> affine invariant

Newton's Method <u>IS</u> affine invariant.

affine invariance of conditional grad.

$$\min_{x \in C} f(x) \iff \min_{x' \in A^{-1}C} h(x')$$

where $h(x') = f(Ax')$



$$\min \ \frac{1}{2}\|x\|^2$$
$$x \in \operatorname{conv}\left\{\begin{pmatrix}-1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$$