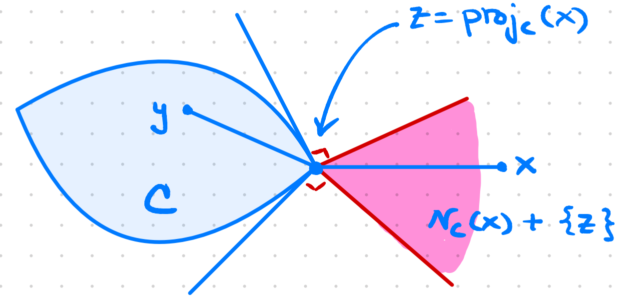
projected gradient descent

- projection onto a convex set
- gradient projection method

orthogonal projection onto convex set

for $C \subseteq \mathbb{R}^n$ closed convex,

$$\mathrm{proj}_C(x) := \underset{z \in C}{\mathrm{argmin}} \ \tfrac{1}{2}\|z-x\|^2$$



properties

- if $x \in C \implies \mathrm{proj}_C(x) = x$
- $\mathrm{proj}_C(z)$ is <u>unique</u> (because objective is strictly convex)
- $z = \mathrm{proj}_C(x) \iff z \in C$ and $(x-z)^T(y-z) \leq 0 \quad \forall y \in C \quad (*)$

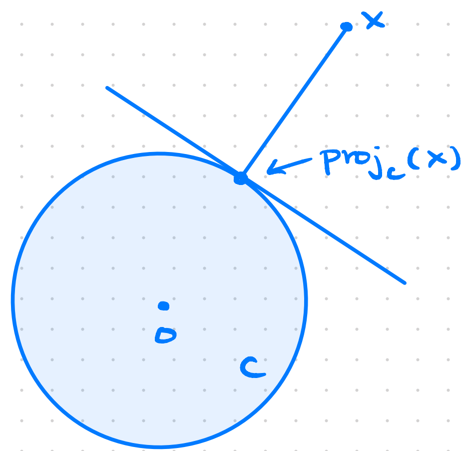
Proof: let $g(x) := \frac{1}{2}\|x-z\|^2$. By optimality,

$$z = \mathrm{proj}_C(x) \iff -\nabla g(x) = x - z \in N_C(x) \iff (*)$$

example : Projection onto 2-norm ball

$$C = \{ x \mid \|x\|_2 \le \alpha \} \equiv \alpha \mathbb{B}_2$$

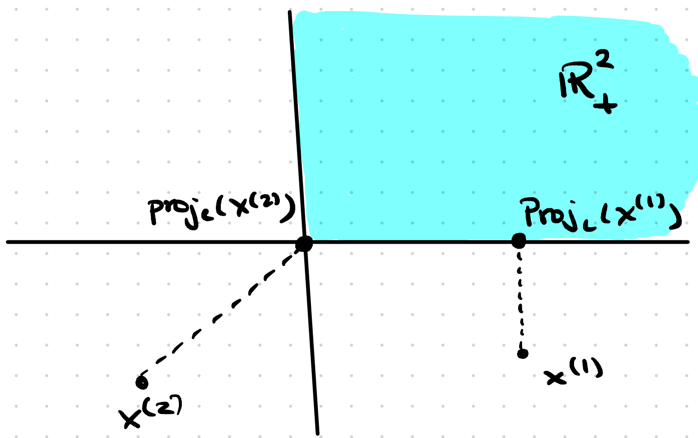
$$\mathrm{proj}_C(x) = \begin{cases} x & \text{if } \|x\|_2 \le \alpha \\ \alpha \dfrac{x}{\|x\|_2} & \text{otherwise} \end{cases}$$



example : projection onto nonnegative orthant

$$C = \{ x \mid x \geq 0 \} = \mathbb{R}^n_+$$

$$\text{proj}_C(x) = \left[\max\{0, x_j\} \right]_{j=1}^n$$



stationarity of projected gradient

$x^* \in \underset{x \in C}{\operatorname{argmin}} \; f(x)$ with $C \subseteq \mathbb{R}^n$ closed cvx, $f: \mathbb{R}^n \to \mathbb{R}$ cvx diff'l iff

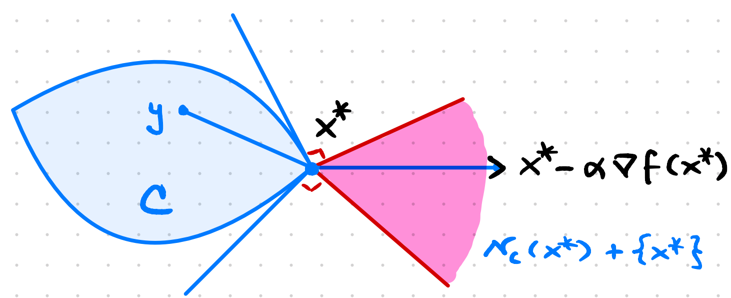
$$x^* = \operatorname{proj}_C(x^* - \alpha \nabla f(x^*)) \quad \forall \alpha > 0$$

by projection theorem:

$$(x^* - \alpha \nabla f(x^*) - x^*)^T (x - x^*) \leq 0 \quad \forall x \in C$$

$$\Longleftrightarrow \quad -\nabla f(x^*)^T (x - x^*) \leq 0 \quad \forall x \in C$$

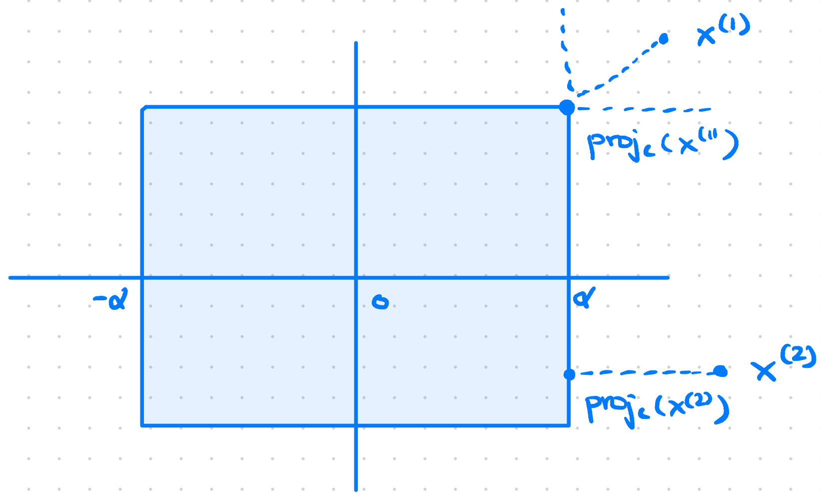
$$\Longleftrightarrow \quad -\nabla f(x^*)^T \in N_C(x^*) \quad \text{(optimality)}$$



example : projection onto inf-norm ball

$$C = \{ x \mid \|x\|_\infty \le \alpha \} \qquad \|x\|_\infty = \max_{j=1\cdots n} |x_j|$$

$$\text{proj}_C(x) = \left[\text{sign}(x_j) \cdot \min\{ \alpha, |x_j| \} \right]_{j=1}^{n}$$

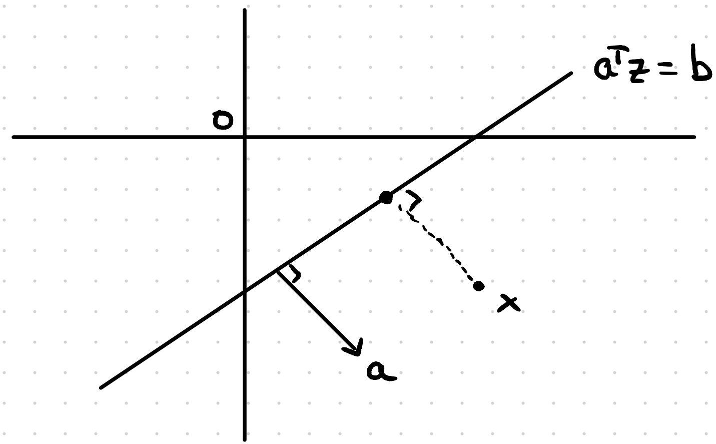


example : projection onto Affine set

$$C = \{ x \mid Ax = b \} \qquad A = \boxed{/ / /} \quad \text{full-row rank}$$

$$\mathrm{proj}_C(x) = \underset{z}{\mathrm{argmin}} \left\{ \tfrac{1}{2}\|z - x\|^2 \mid Az = b \right\}$$

Claim : $x - \mathrm{proj}_C(x) \in$?



example : de-biasing

$$C = \{ x \mid e^T x = 0 \}$$

$$\text{proj}_C(x) = \underset{z}{\text{argmin}} \left\{ \tfrac{1}{2} \| z - x \|^2 \mid e^T z = 0 \right\}$$

by optimality, $x - \text{proj}_C(x) \in \text{range}(e)$

$$\Rightarrow \quad x - \text{proj}_C(x) = \alpha e \quad \text{for some } \alpha \in \mathbb{R}$$

$$\Rightarrow \quad \text{proj}_C(x) = x - \alpha e$$

to find α: $\quad 0 = e^T \text{proj}_C(x) = e^T (x - \alpha e) = e^T x - \alpha n$

$$\Rightarrow \quad \alpha = \frac{1}{n} e^T x$$

projected gradient descent

solve the differentiable convex optimization problem

$$\min_{x \in C} f(x) \qquad f : \mathbb{R}^n \to \mathbb{R} \text{ diff'l convex}$$
$$C \subseteq \mathbb{R}^n \text{ convex}$$

iteration : $x_0 \in C$ arbitrary

for $k = 0, 1, 2, \dots$

choose α_k (eg, Armijo, constant, decreasing)

$$x_{k+1} = \operatorname{proj}_C \left(x_k - \alpha_k \nabla f(x_k) \right)$$

stop if $\| x_k - x_{k+1} \| \leq \varepsilon$