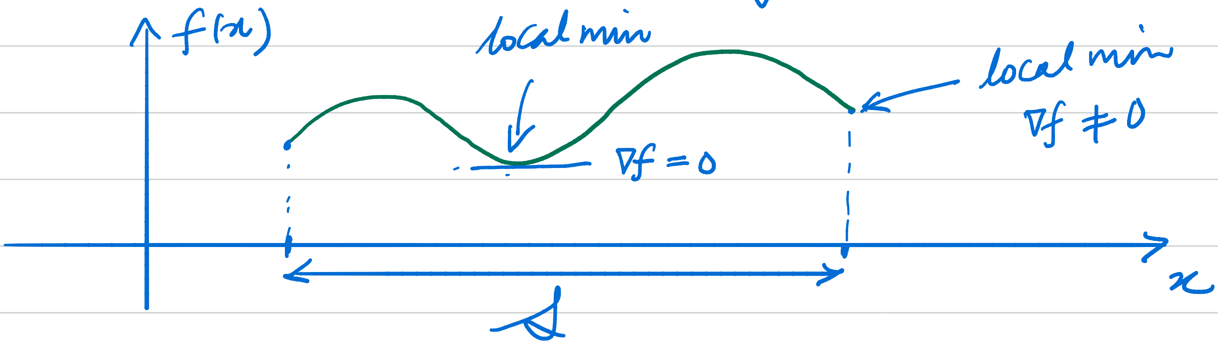
Constrained Optimization & Gradient Projection

$$\min_{x \in \mathcal{S}} f(x)$$

where $\mathcal{S}$ is a (non-empty) closed and convex subset of $\mathbb{R}^n$

- Assume that f is continuously differentiable on $\mathcal{S}$.



Definition x^* is a local min. of f over $\mathcal{S}$ if $\exists\, \varepsilon > 0$ s.t. $f(x^*) \leq f(x)$ $\forall\, x \in \mathcal{S}$ with $\|x - x^*\| < \varepsilon$.

x^* is global min if $f(x^*) \leq f(x)$ $\forall\, x \in \mathcal{S}$.

Proposition (Optimality Conditions)

(a) If x^* is a local min. of f over $\mathcal{S}$, then

$$\nabla f(x^*)^T (x - x^*) \geq 0 \quad \forall\, x \in \mathcal{S} \quad \text{--- } (*)$$

(b) If f is convex over $\mathcal{S}$, then $(*)$ is also sufficient for x^* to be a (global) min.

Proof (a) Suppose x^* is a local min. and

$\nabla f(x^*)^T(x-x^*) < 0$ for some $x \in \mathcal{S}$.

Let $g(\varepsilon) = f(x^* + \varepsilon(x-x^*))$.

Then $g'(\varepsilon) = \nabla f(x^* + \varepsilon(x-x^*))^T (x-x^*)$.

By MVT, $g(\varepsilon) = g(0) + \varepsilon g'(\beta\varepsilon)$ for some $\beta \in [0,1]$

$\Rightarrow f(x^* + \varepsilon(x-x^*)) = f(x^*) + \varepsilon \nabla f(x^* + \beta\varepsilon(x-x^*))^T (x-x^*)$

for some $\beta \in [0,1]$.

Since ∇f is continuous, we have that for all suff. small $\varepsilon > 0$, $\nabla f(x^* + \beta\varepsilon(x-x^*))^T (x-x^*) < 0$

$\Rightarrow f(x^* + \varepsilon(x-x^*)) < f(x^*)$

Since $x^* + \varepsilon(x-x^*) = \varepsilon x + (1-\varepsilon)x^* \in \mathcal{S}$ (convexity)

x^* cannot be a local min over $\mathcal{S}$ $\Rightarrow\Leftarrow$.

(b) Convexity of f over $\mathcal{S}$

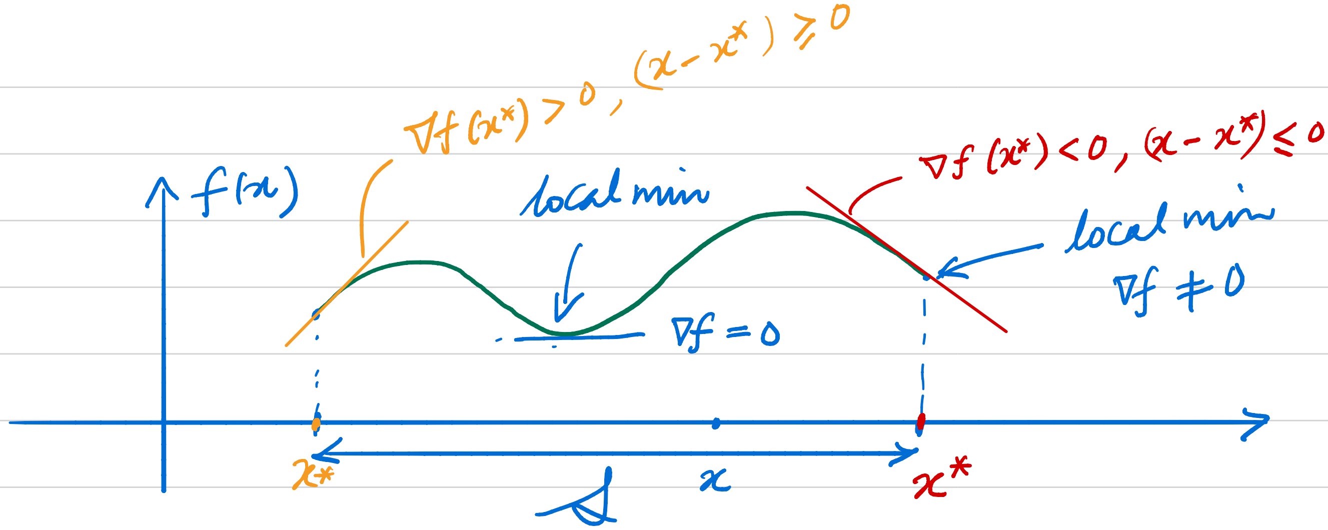
$\Rightarrow f(x) \geq f(x^*) + \nabla f(x^*)^T(x-x^*)$, $\forall x \in \mathcal{S}$.

Thus,

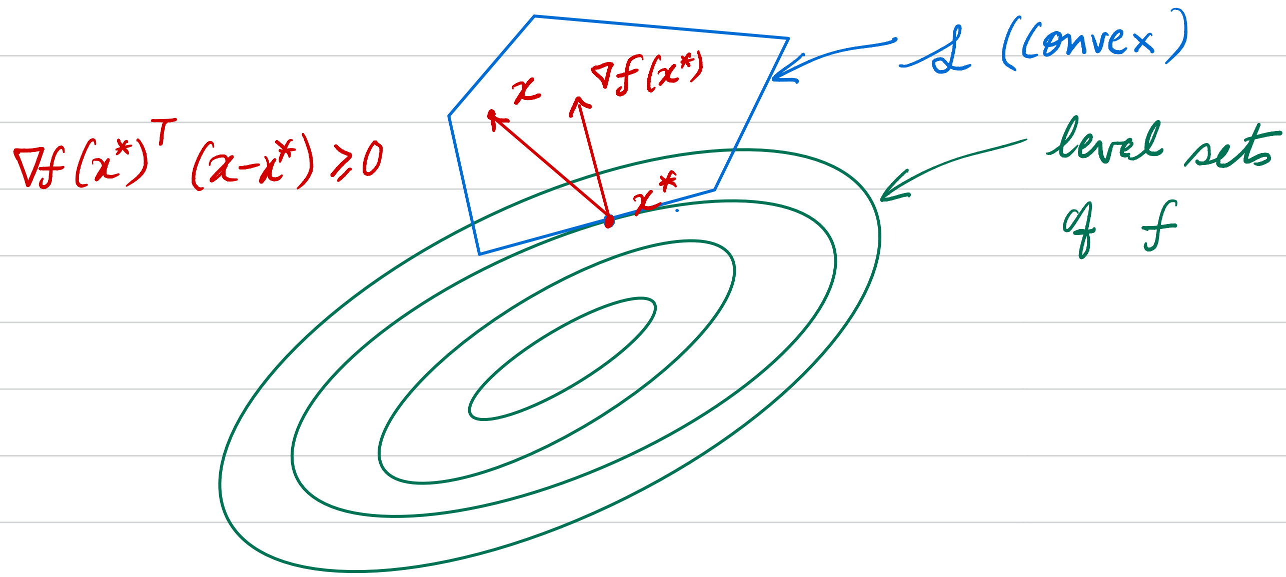
$\nabla f(x^*)^T(x-x^*) \geq 0 \quad \forall x \in \mathcal{S}$

$\Rightarrow f(x) \geq f(x^*) \quad \forall x \in \mathcal{S}$

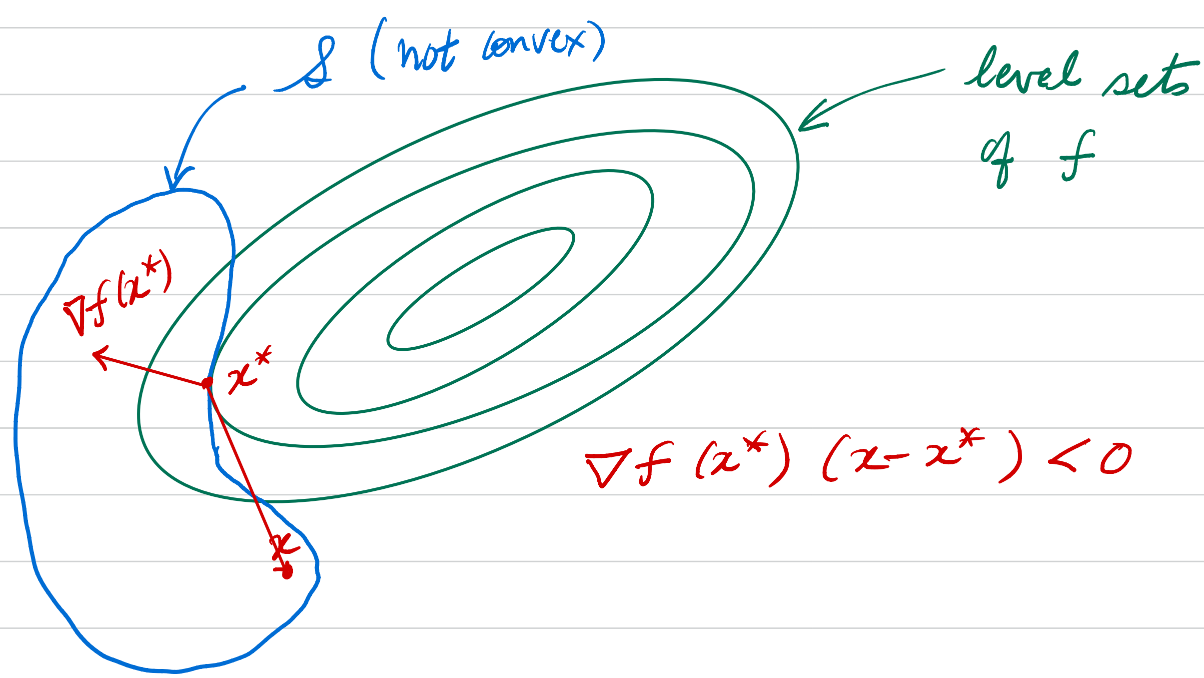
$\Rightarrow x^*$ is a global min of f over $\mathcal{S}$



∇f (x*) > 0 , (x - x*) ≥ 0
∇f (x*) < 0 , (x - x*) ≤ 0
f(x)
local min
local min
∇f ≠ 0
∇f = 0
x*
x
x*
S



S (convex)
x
∇f (x*)
∇f (x*)ᵀ (x - x*) ≥ 0
level sets of f
x*



S (not convex)
level sets of f
∇f (x*)
x*
∇f (x*) (x - x*) < 0
x

Definition y is an interior point of $\mathcal{S}$ if $\exists\, \varepsilon > 0$ s.t. $B_\varepsilon = \{x : \|y - x\| < \varepsilon\} \subset \mathcal{S}$.

Remark If x^* is an interior point of $\mathcal{S}$, then from lec 3, x^* is local min $\Rightarrow \nabla f(x^*) = 0$. Also, if f is convex,

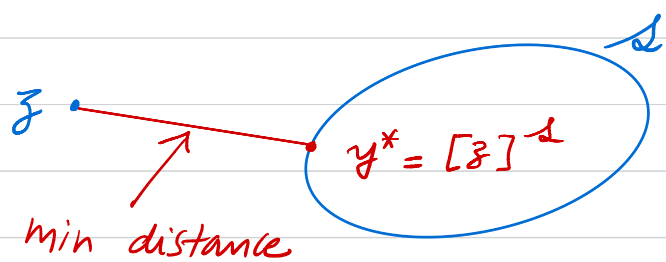
x^* is a global min $\Leftrightarrow \nabla f(x^*) = 0$

Projection onto Closed Convex Sets

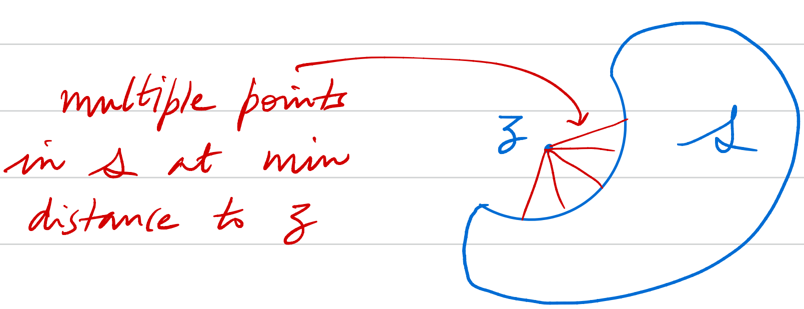
Definition Let $\mathcal{S}$ be a closed convex subset of $\mathbb{R}^n$. Then, for $z \in \mathbb{R}^n$, the projection of z on $\mathcal{S}$ is denoted by $[z]^{\mathcal{S}}$ and is given by

$$[z]^{\mathcal{S}} = \arg\min_{y \in \mathcal{S}} \|z - y\|^2$$

("=" justified since we will show that $[z]^{\mathcal{S}}$ exists and is unique.)



If $\mathcal{S}$ is not convex minimizer may not be unique



<u>Result</u> (Existence and Uniqueness of Projection)

Let S be a <u>closed convex</u> subset of $\mathbb{R}^n$.

Then, for every $z \in \mathbb{R}^n$, there exists a unique $[z]^S$.

<u>Proof</u> Need to show that $\min_{y \in S} \|y - z\|^2$ exists and is unique.

Let x be some element of S. Then

minimizing $\|z-y\|^2$ over all $y \in S$

$\equiv$ minimizing $\|z-y\|^2$ over the set

$$A = \{y \in S : \|z-y\|^2 \leq \|z-x\|^2\}$$

$g(y) = \|z-y\|^2$ is strictly convex on set S

$\Rightarrow$ A is convex set and g is convex on A

Also g is continuous $\Rightarrow$ A is closed

Finally, $y \in A \Rightarrow \|y\|^2 = \|y - z + z\|^2$

$$\leq \|y-z\|^2 + \|z\|^2$$

Triangle ineq

$$\leq \|z-x\|^2 + \|z\|^2$$

$\Rightarrow$ A is bounded

Thus, $g(y) = \|z-y\|^2$ is strictly convex over set A, which is compact. Therefore,

$$\min_{y \in S} \|z-y\|^2 = \min_{y \in A} \|z-y\|^2 \text{ exists and is unique}$$

(WT) (strict convexity)

Result (Necessary and Sufficient Condition for Projection)

Let $\mathcal{S}$ be a closed convex subset of $\mathbb{R}^n$. Then

$$[z]^{\mathcal{S}} = y^* \iff (y^* - z)^T (y - y^*) \geq 0, \ \forall y \in \mathcal{S}.$$
$$\iff (z - y^*)^T (y - y^*) \leq 0, \ \forall y \in \mathcal{S}.$$

Proof

$$[z]^{\mathcal{S}} = \underset{y \in \mathcal{S}}{\arg\min} \ g(y), \quad \text{with } g(y) = \|z - y\|^2$$

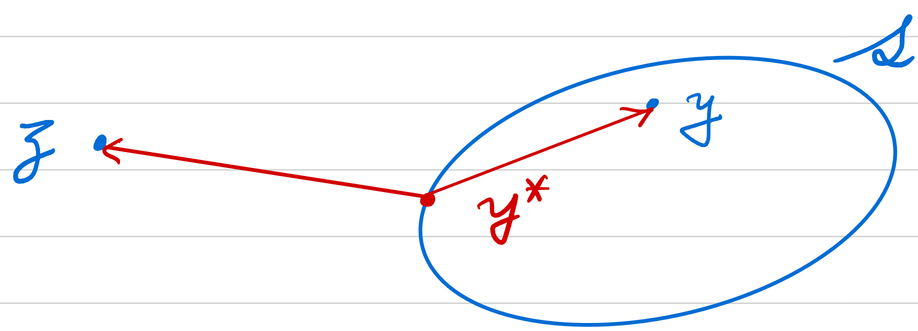
strictly convex

$$\nabla g(y) = 2(y - z)$$

From optimality condition for minimizing a convex function over closed convex set $\mathcal{S}$,

y^* is the unique minimizer of $g(y)$ over $\mathcal{S}$

$$\iff \nabla g(y^*)^T (y - y^*) \geq 0 \ \forall y \in \mathcal{S}$$
$$\iff (y^* - z)^T (y - y^*) \geq 0 \ \forall y \in \mathcal{S},$$
$$\iff (z - y^*)^T (y - y^*) \leq 0 \ \forall y \in \mathcal{S}$$



$$(z - y^*)^T (y - y^*) \leq 0.$$

<u>Result</u> (Projection is non-expansive)

Let $\mathcal{S}$ be a <u>closed convex</u> subset of $\mathbb{R}^n$. Then for $x, z \in \mathbb{R}^n$,

$$\|[x]^{\mathcal{S}} - [z]^{\mathcal{S}}\| \leq \|x - z\| \quad \forall\, x, z \in \mathbb{R}^n$$

<u>Proof</u> From previous result,

$$(1) \cdots \left([x]^{\mathcal{S}} - x\right)^T \left(y - [x]^{\mathcal{S}}\right) \geq 0 \quad \forall\, y \in \mathcal{S}$$

$$(2) \cdots \left([z]^{\mathcal{S}} - z\right) \left(y - [z]^{\mathcal{S}}\right) \geq 0 \quad \forall\, y \in \mathcal{S}$$

Setting $y = [z]^{\mathcal{S}}$ in (1) and $y = [x]^{\mathcal{S}}$ in (2), and adding, we get

$$\left([z]^{\mathcal{S}} - [x]^{\mathcal{S}}\right)^T \left([x]^{\mathcal{S}} - x + z - [z]^{\mathcal{S}}\right) \geq 0$$

$$\Rightarrow \left([z]^{\mathcal{S}} - [x]^{\mathcal{S}}\right)^T (z - x) \geq \|[z]^{\mathcal{S}} - [x]^{\mathcal{S}}\|^2$$

Applying Cauchy-Schwarz inequality to LHS,

$$\|[z]^{\mathcal{S}} - [x]^{\mathcal{S}}\|^{\cancel{2}} \leq \cancel{\|[z]^{\mathcal{S}} - [x]^{\mathcal{S}}\|}\, \|z - x\| \quad \blacksquare$$