262B-Lecture 11

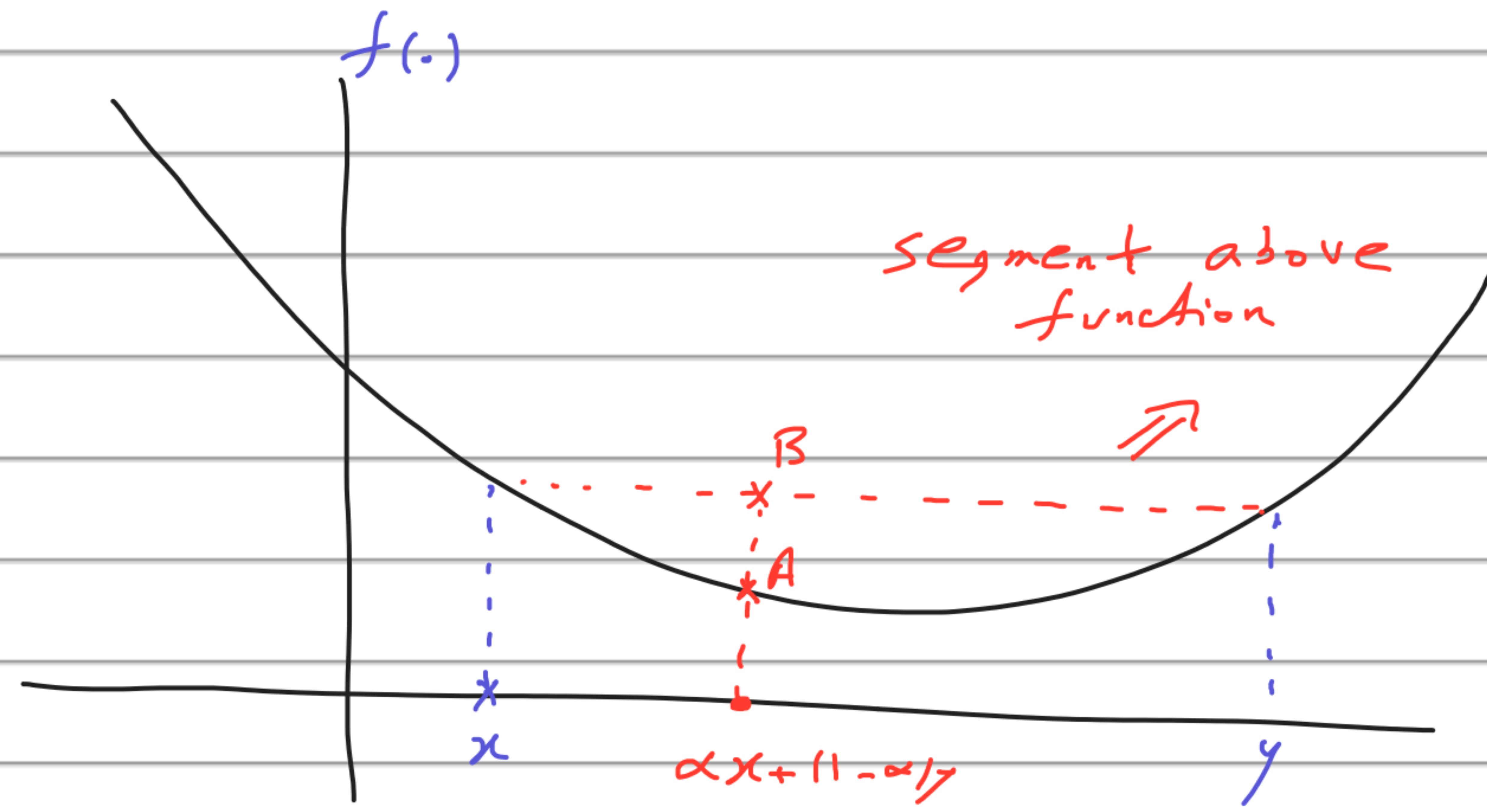
Date created: 2021.02.25
N. of Pages: 15

Algorithms for Constrained optimization:

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t.} \quad x \in X$$

$\Rightarrow$ Convex functions,
Convex sets

Convex function:



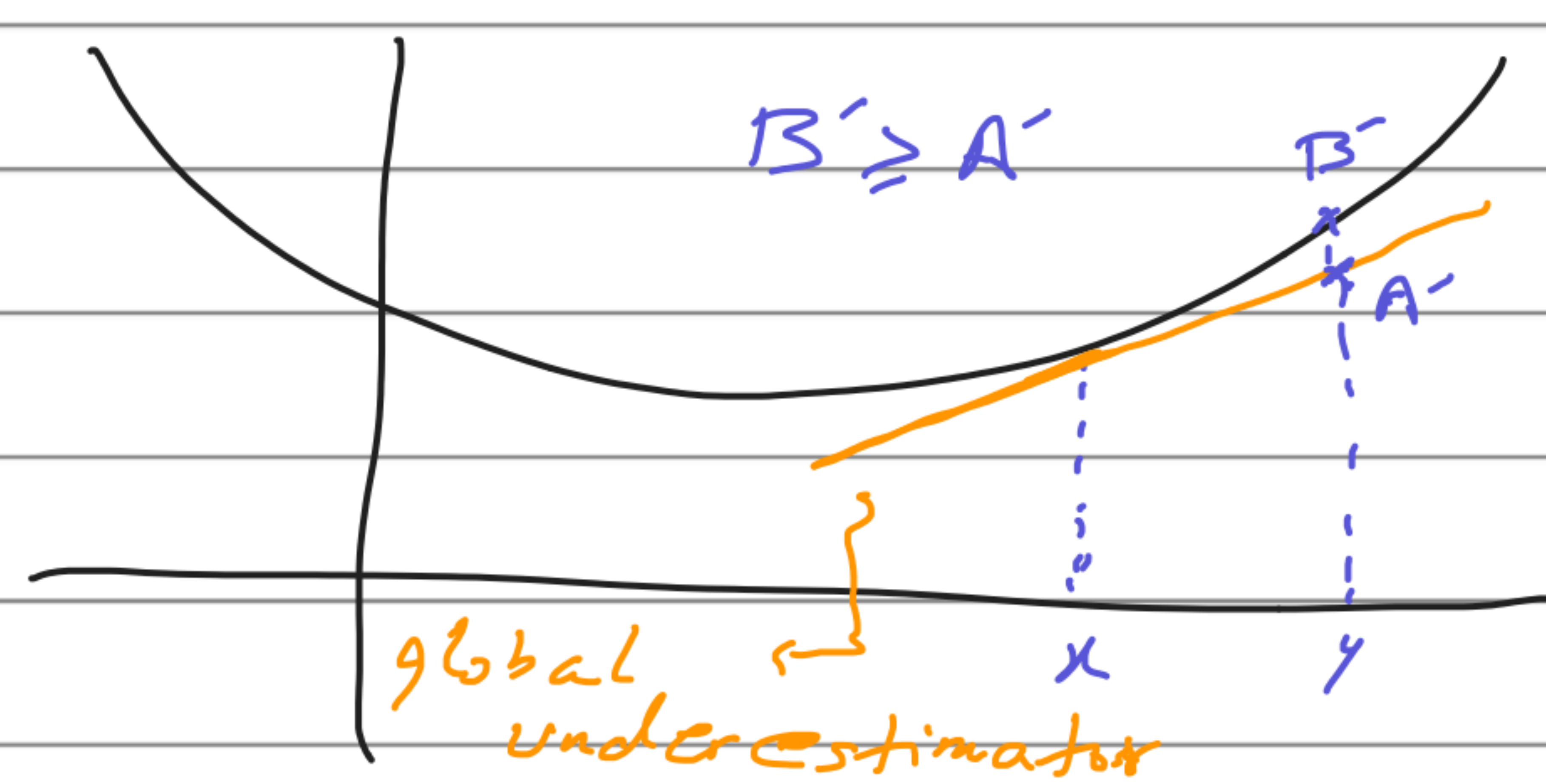
$B \geq A$

$\Rightarrow$

Zero-th order Convexity Condition : $f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$
$\forall x, y \in \mathbb{R}^n,$
$\forall \alpha \in [0, 1]$

first-order Convexity : $f(y) \geq f(x) + \nabla f(x)^T (y - x) \quad \forall x, y$

second-order Convexity : $\nabla^2 f(x) \succeq 0$
$\forall x$

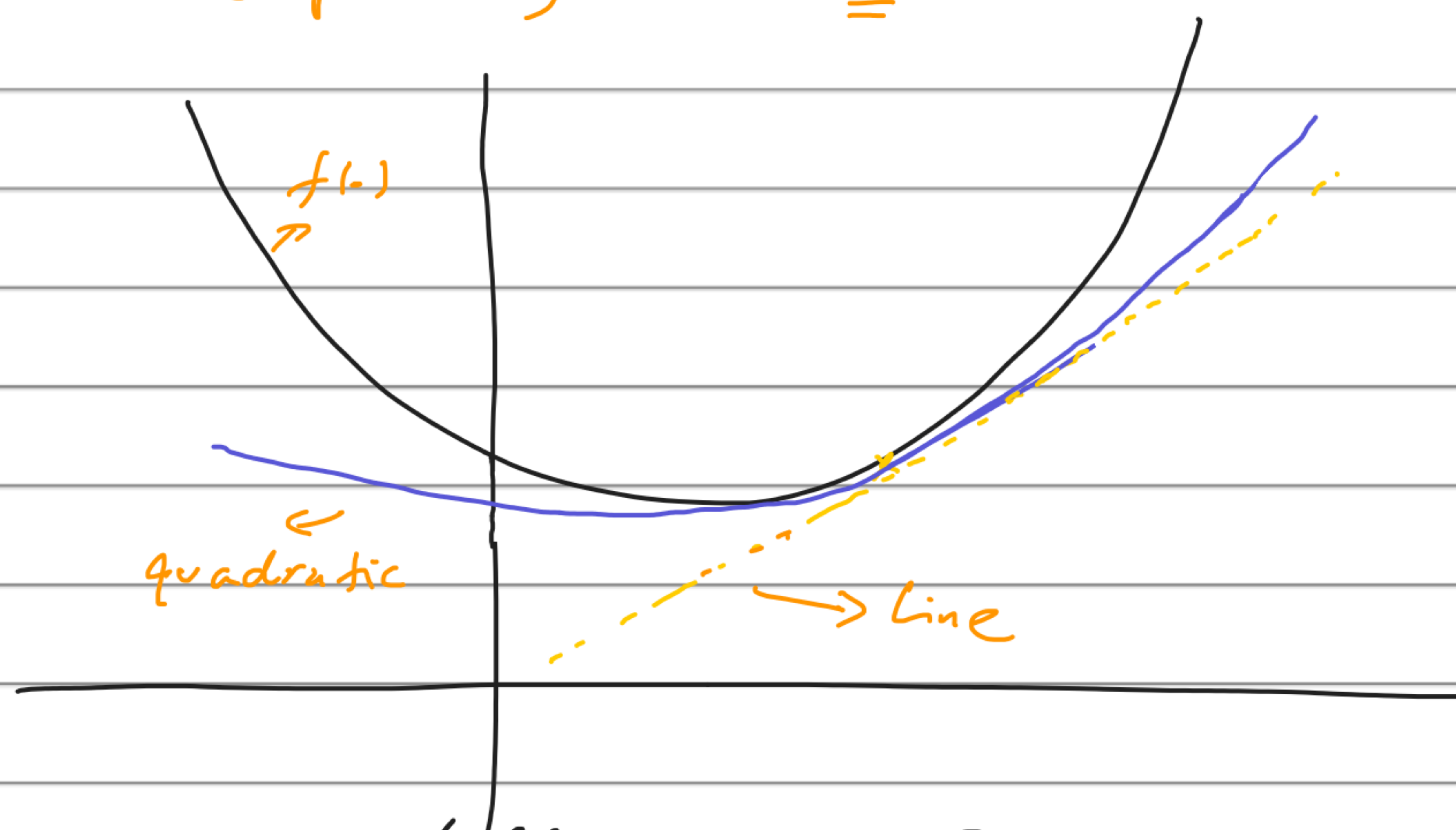


Strong convexity: (equivalent definitions)

1- $\nabla^2 f(x) \succeq mI \quad \forall x \in \mathbb{R}^n \qquad (m > 0)$

2- $f(y) \geq \underbrace{f(x) + \nabla f(x)^T (y - x) + \frac{m}{2} \|x - y\|_2^2}_{\text{global quadratic underestimator}}$

$\Rightarrow$ function has some curvature depending on $\underline{\underline{m}}$



(what if $f(.)$ is not differentiable?)

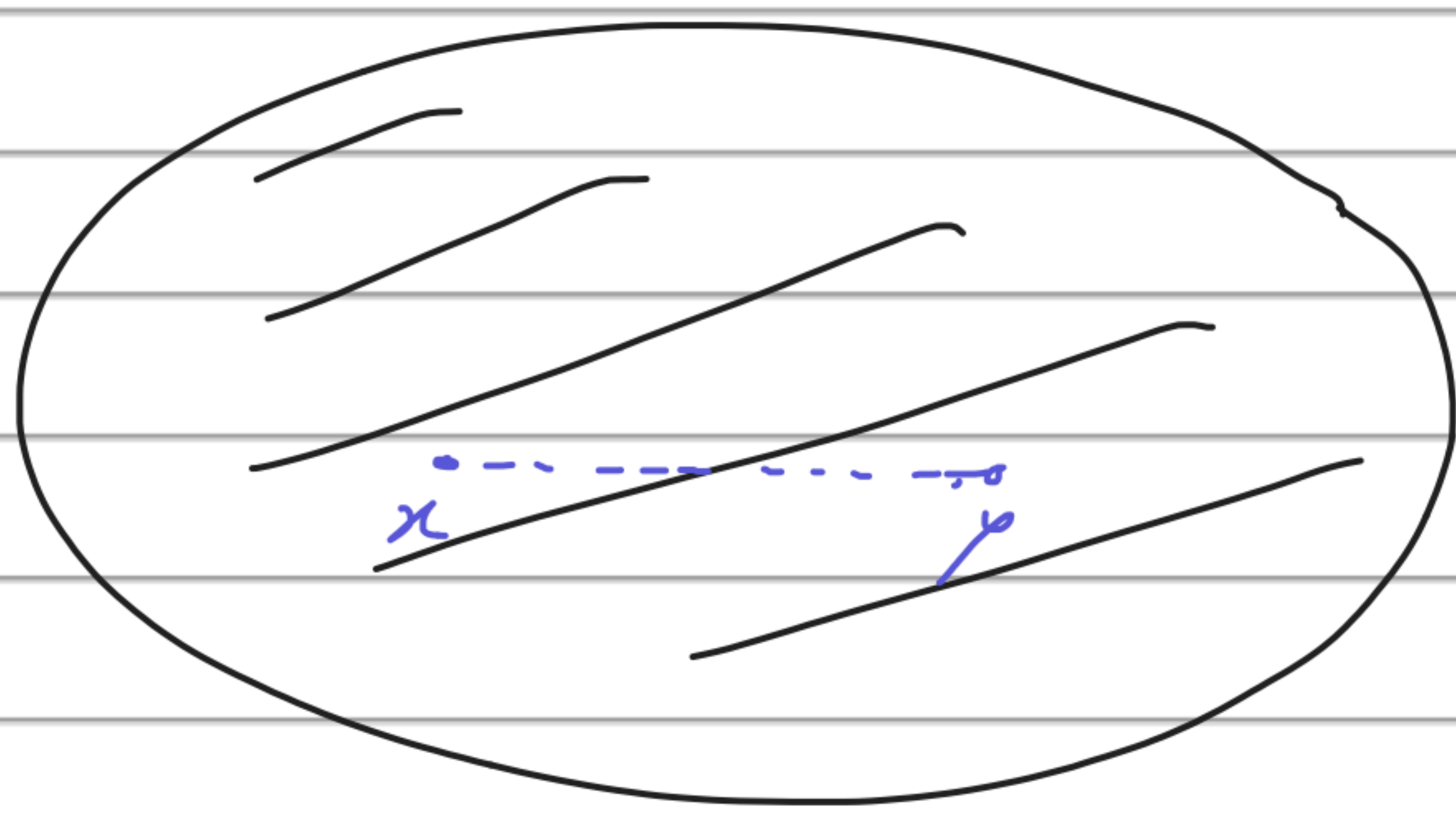
3- $f(\alpha x + (1-\alpha) y) + \frac{1}{2} m \alpha (1-\alpha) \|x - y\|^2 \leq \alpha f(x) + (1-\alpha) f(y)$

quadratic function

4- $\underbrace{(\nabla f(x) - \nabla f(y))^T}_{\nabla f(.) \text{ is a monotone operator}} (x - y) \geq m \|x - y\|^2 \quad \forall x, y$

$\Longrightarrow$ $f(x)$ is called m-strongly convex.

Convex set : set $\underline{\underline{S}}$



segment is in the set

$\Longrightarrow$ $\alpha x + (1-\alpha) y \in S$

$\forall\, x, y \in S \;,\; \forall\, \alpha \in [0,1]$

m-strongly convex set :

$$\alpha x + (1-\alpha) y + \underbrace{\frac{m}{2}\alpha(1-\alpha)\|x-y\|^2}_{\text{in definition of strongly convex function}} \underbrace{z}_{\text{different directions}} \in S$$

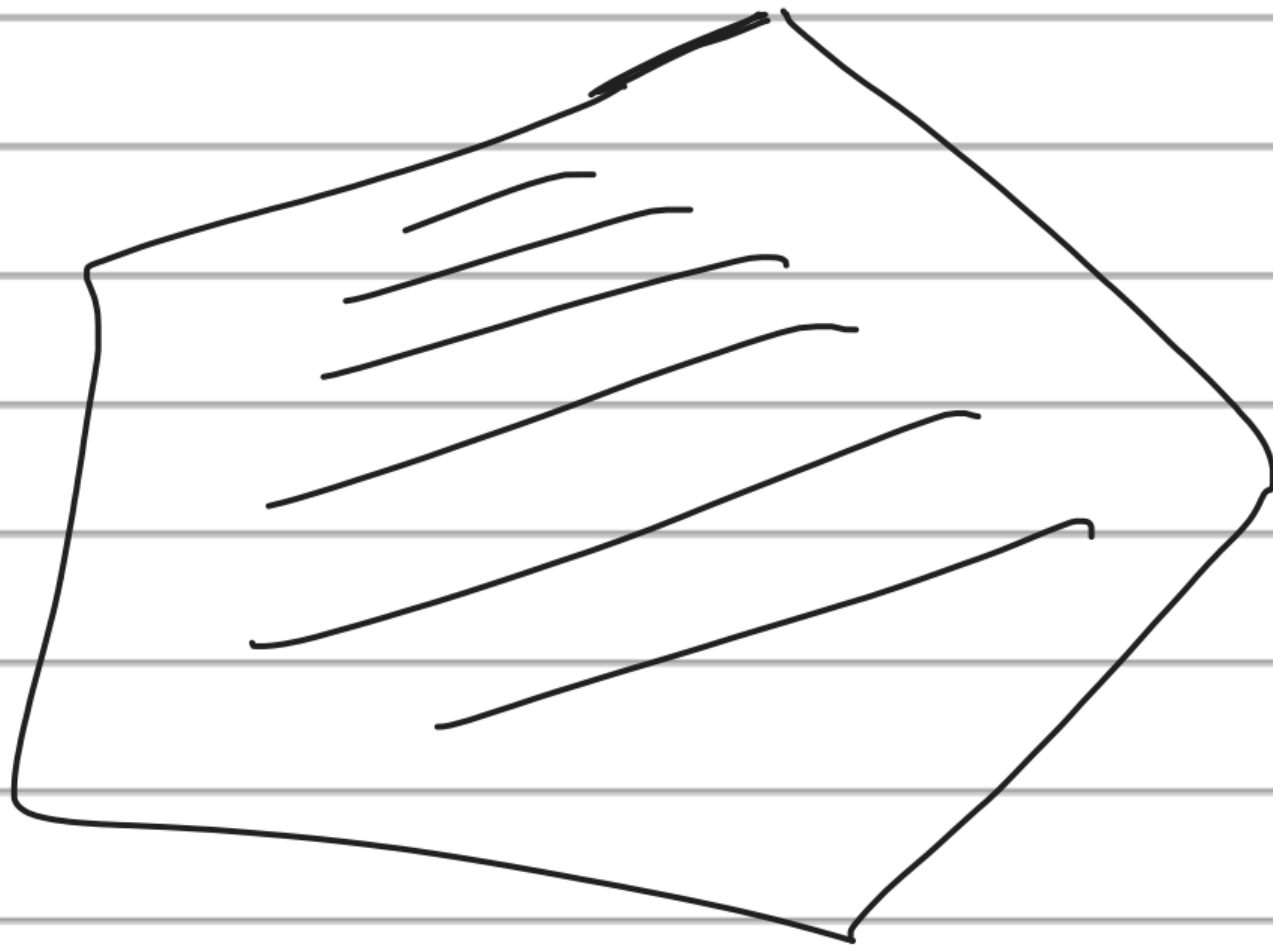
$\forall\, x, y \in S,$
$\forall\, \alpha \in [0,1],$
$\forall\, z : \|z\| \le 1$

segment

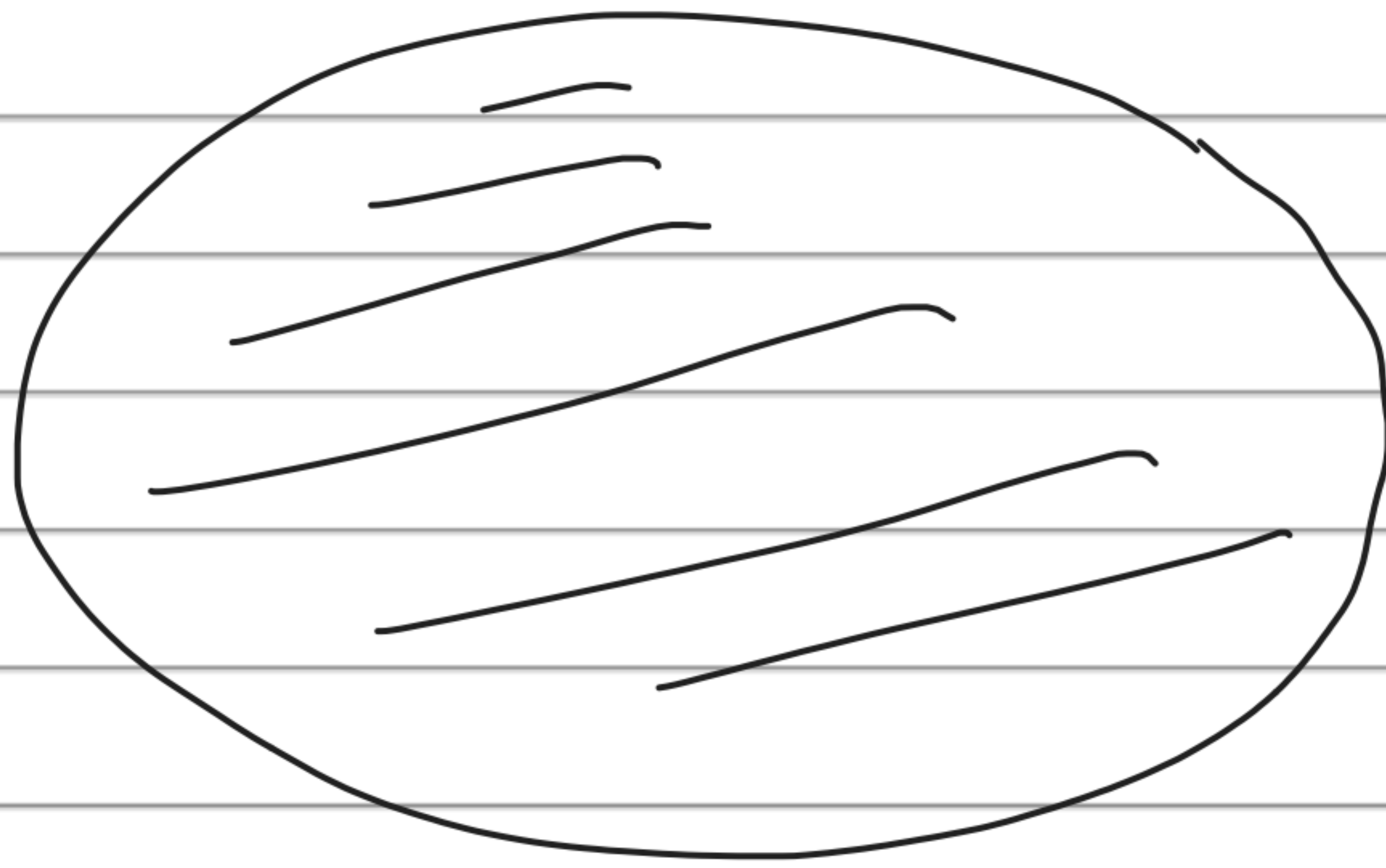
$\Longrightarrow$

is in $\underline{\underline{S}}$

$\Longrightarrow$ Boundary of $\underline{\underline{S}}$ has curvature.



not strong convex

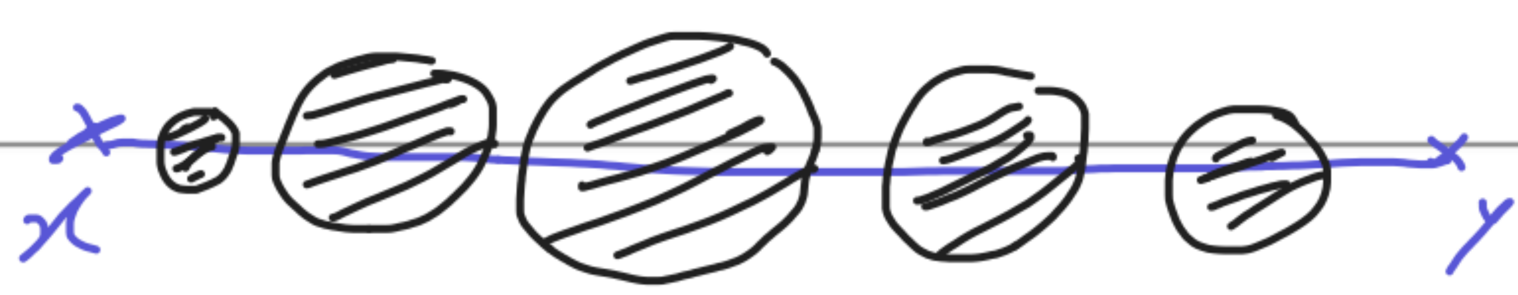


strongly convex

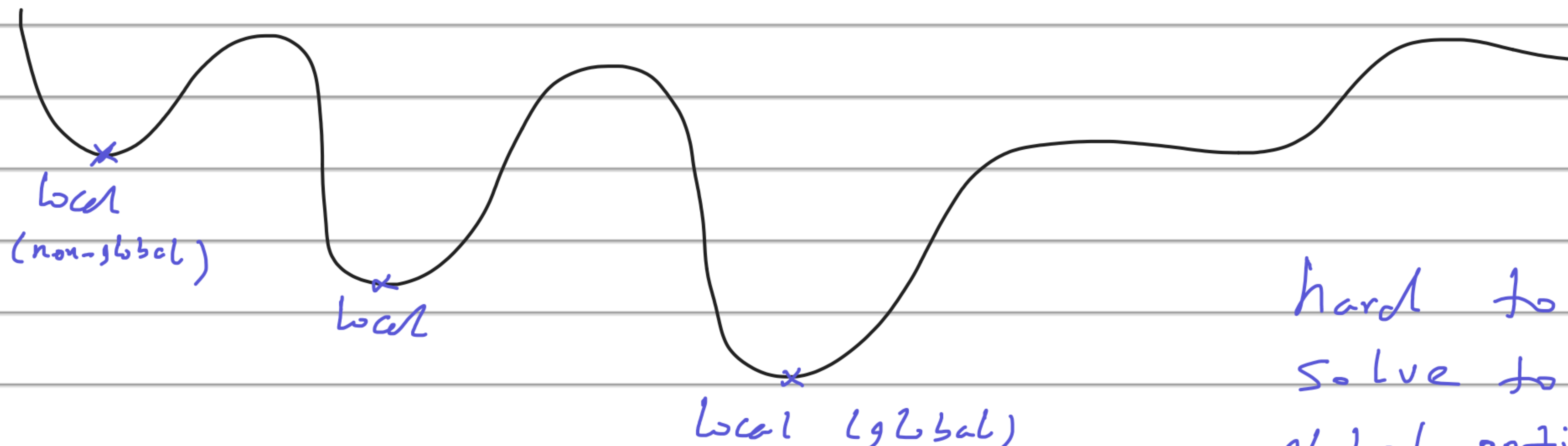
equivalent definition :

$$\forall x,y \in S, \quad \forall \alpha \in [0,1] \implies \underbrace{B}_{\text{Ball}}\Big(\underbrace{\alpha x + (1-\alpha) y}_{\text{Center}},\ \overbrace{\tfrac{m}{2}\,\alpha(1-\alpha)\,\|x-y\|^2}^{\text{radius}}\Big) \in S$$

So, $\underline{\underline{z}}$ defines a ball.

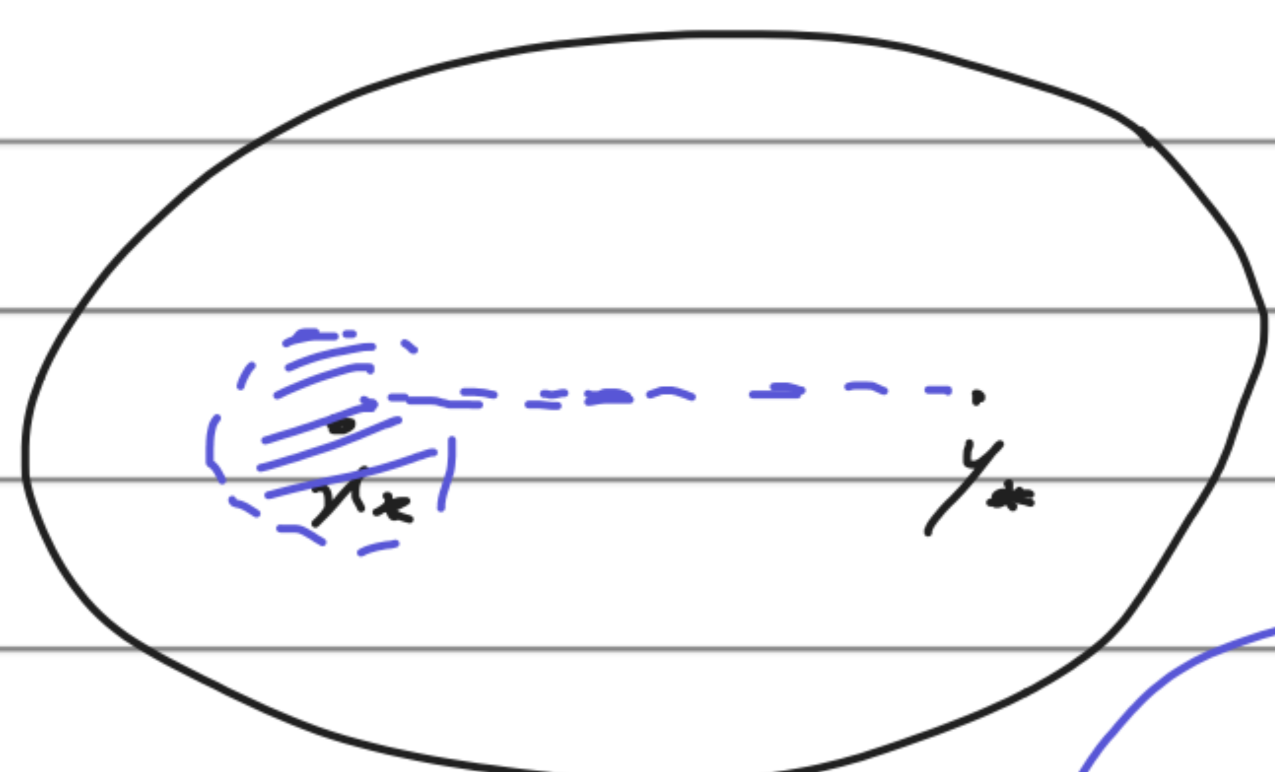


Thm: Consider $\min f(x)$ s.t. $x \in X$ where $f(\cdot)$ is a convex function & X is a convex set. Every local min is a global min



hard to solve to global optimality

Proof: By contradiction, assume x_* is a non-global local min & y_* is a global min $\Rightarrow f(y_*) < f(x_*)$



If $\alpha \neq 0$:

$$f(\alpha x_* + (1-\alpha) y_*) \leq \alpha f(x_*) + (1-\alpha) f(y_*)$$

$$< \alpha f(x_*) + (1-\alpha) f(x_*) = f(x_*)$$

If $\alpha \to 1 \Rightarrow f(\text{points close to } x_*) < f(x_*)$

$\Rightarrow x_* \neq$ local min

Assumptions: – X : convex, closed

– $f(x)$ is differentiable on an open set containing X.

Ex: X : • $\Longrightarrow$  open set

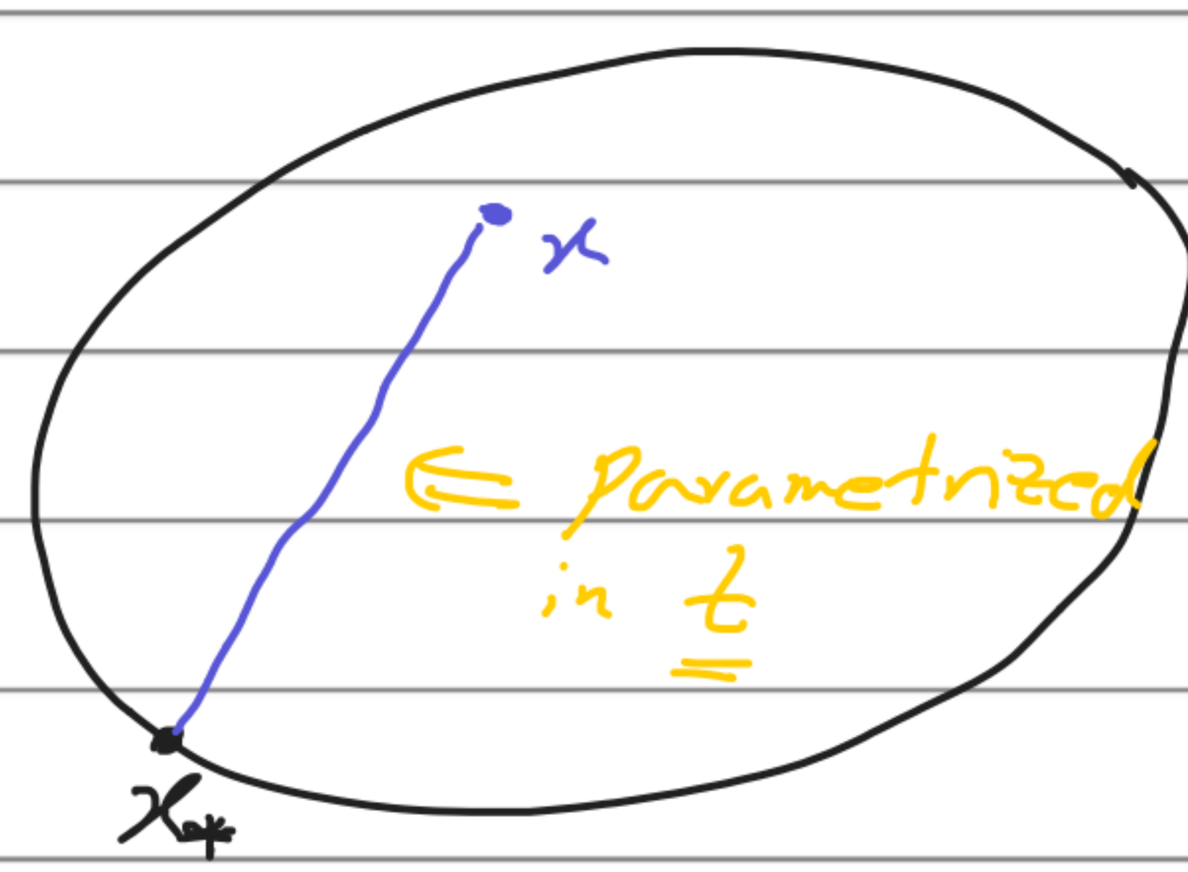
This chapter (3) : X = convex, $f(\cdot)$ = arbitrary

Thm (FOC necessary): If X is convex & $f(x)$ is arbitrary, and if x_* is a local min, then

$$\boxed{\nabla f(x_*)^T (x - x_*) \geq 0 \quad \forall x \in X}$$

Proof: Pick arbitrary $x \in X$, and note

$x_* + t(x - x_*) \in X \quad \forall\, t \in [0,1]$



$f(x_* + t(x - x_*))$ is differentiable with respect to t at $t=0$ due to open set assumption.

$$\Rightarrow \quad 0 \le \lim_{t \downarrow 0} \frac{\overbrace{f(x_* + t(x - x_*)) - f(x_*)}^{\geq 0 \text{ local optimality}}}{\underbrace{t}_{>0}}$$

$$= \nabla f(x_*)^T (x - x_*)$$

Thm: If X = convex set & $f(x)$ = convex function,

then x_* is a local min $\underbrace{\text{if only if}}$

$$\nabla f(x_*)^T (x - x_*) \geq 0 \quad \forall x \in X$$

Proof of $\underline{\text{if}}$ part: Assume $\nabla f(x_*)^T (x - x_*) \geq 0$

Then by first-order convexity condition:

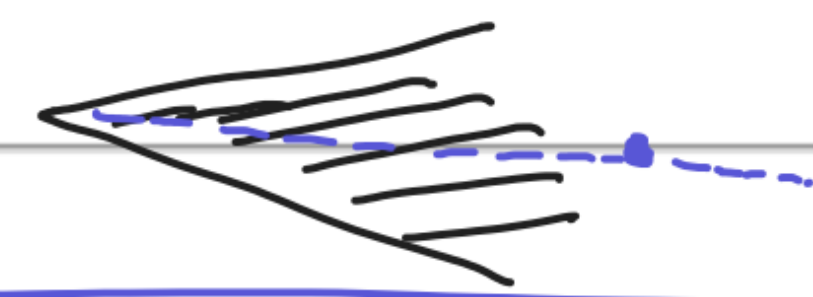
$$f(x) \geq f(x_*) + \underbrace{\nabla f(x_*)^T (x - x_*)}_{\geq 0} \geq f(x_*)$$

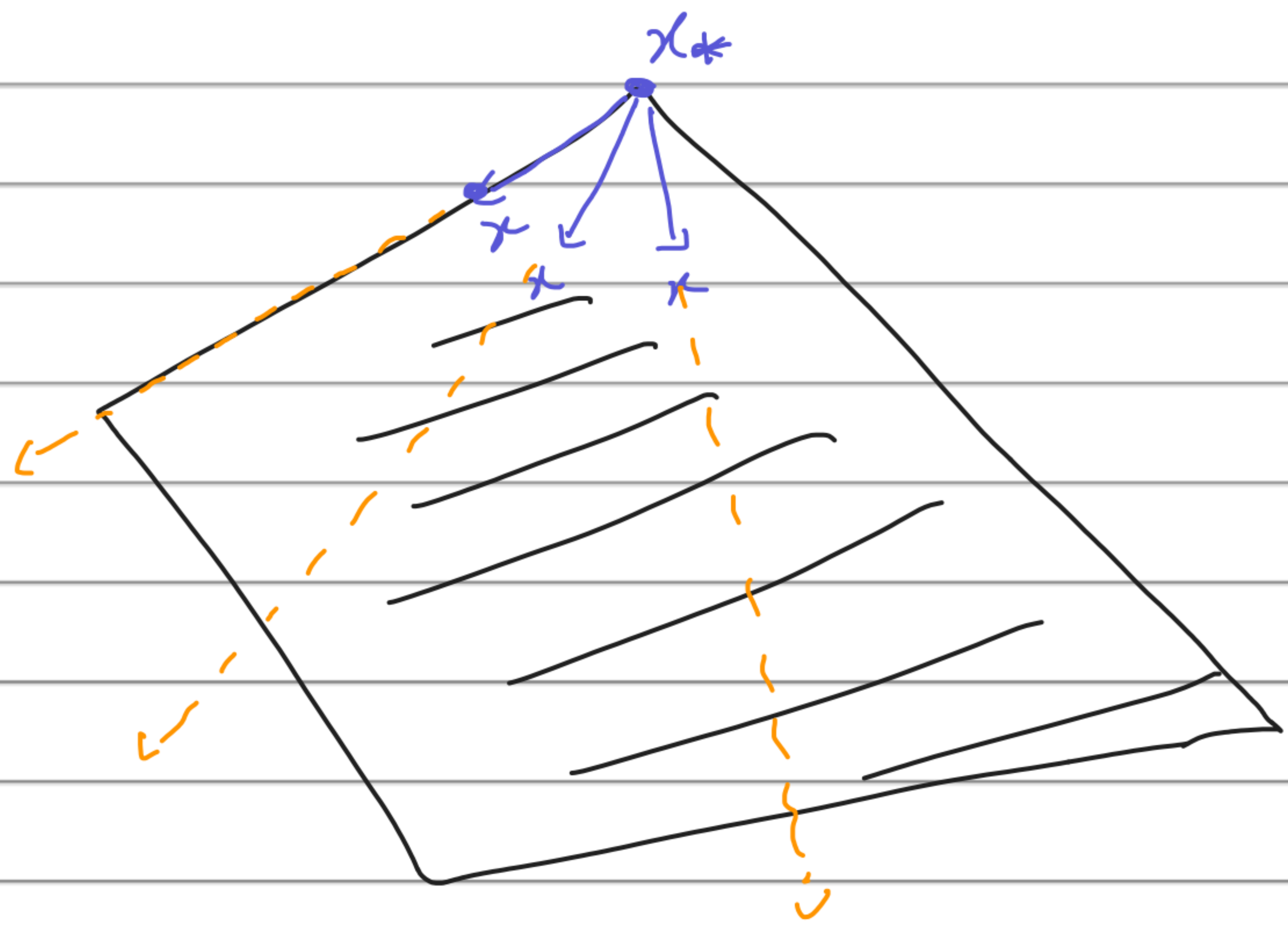
$$\Rightarrow f(x) \geq f(x_*) \quad \forall x \in X \Rightarrow x_* = \text{global min}$$

Focus on $x - x_*$ where $x \in X$

- Cone of feasible directions
- Tangent Cone
- Normal Cone

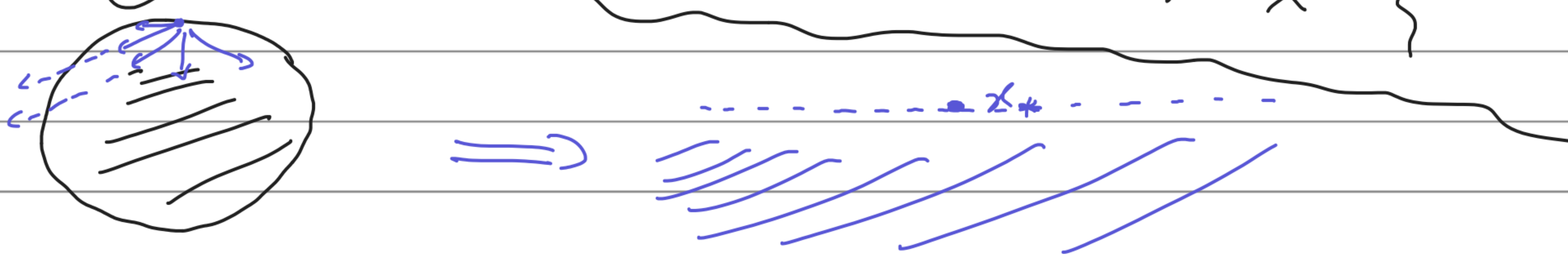
K is a cone if $\alpha x \in K$ for $\forall x \in K$, $\forall \alpha \geq 0$





Cone of feasible directions:

$$F_X(x_*) = \{ \Delta x \mid \Delta x = \alpha (x - x_*) \text{ for some } \alpha \geq 0, \ x \in X \}$$



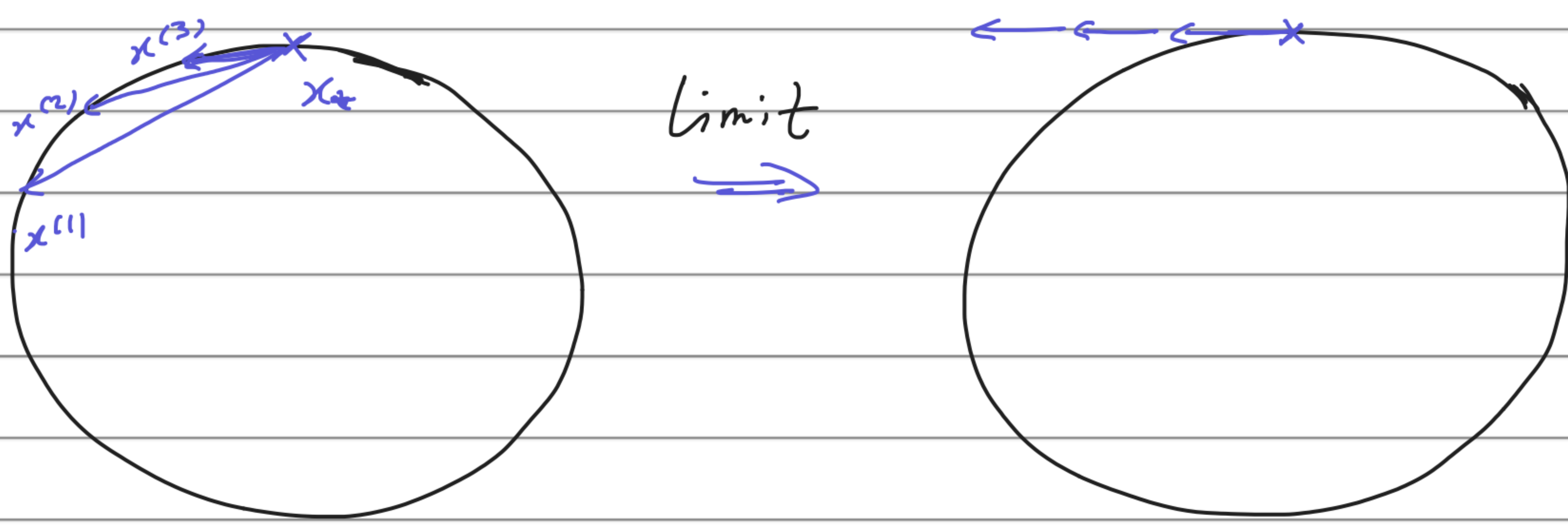
In that example, $F_X(x_*)$ doesn't contain its top boundary.

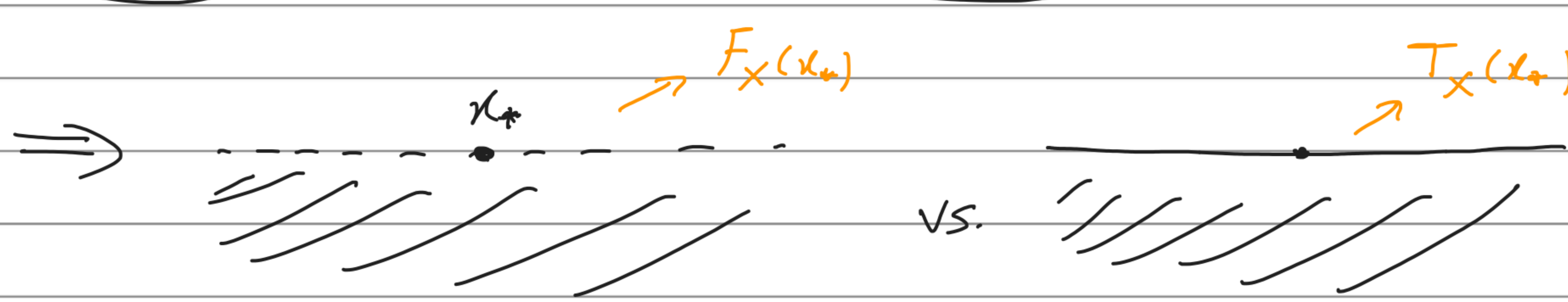
Tangent cone at $x_* = T_X(x_*)$

$= \{ \Delta x \mid \Delta x = 0$ or $\exists \{x^{(k)}\}_{k=1}^{\infty} \subset X$ s.t.

$x^{(k)} \neq x_*, \quad \lim_{k\to\infty} x^{(k)} = x_*,$

$\lim_{k\to\infty} \dfrac{x^{(k)} - x_*}{\|x^{(k)} - x_*\|} = \dfrac{\Delta x}{\|\Delta x\|} \}$

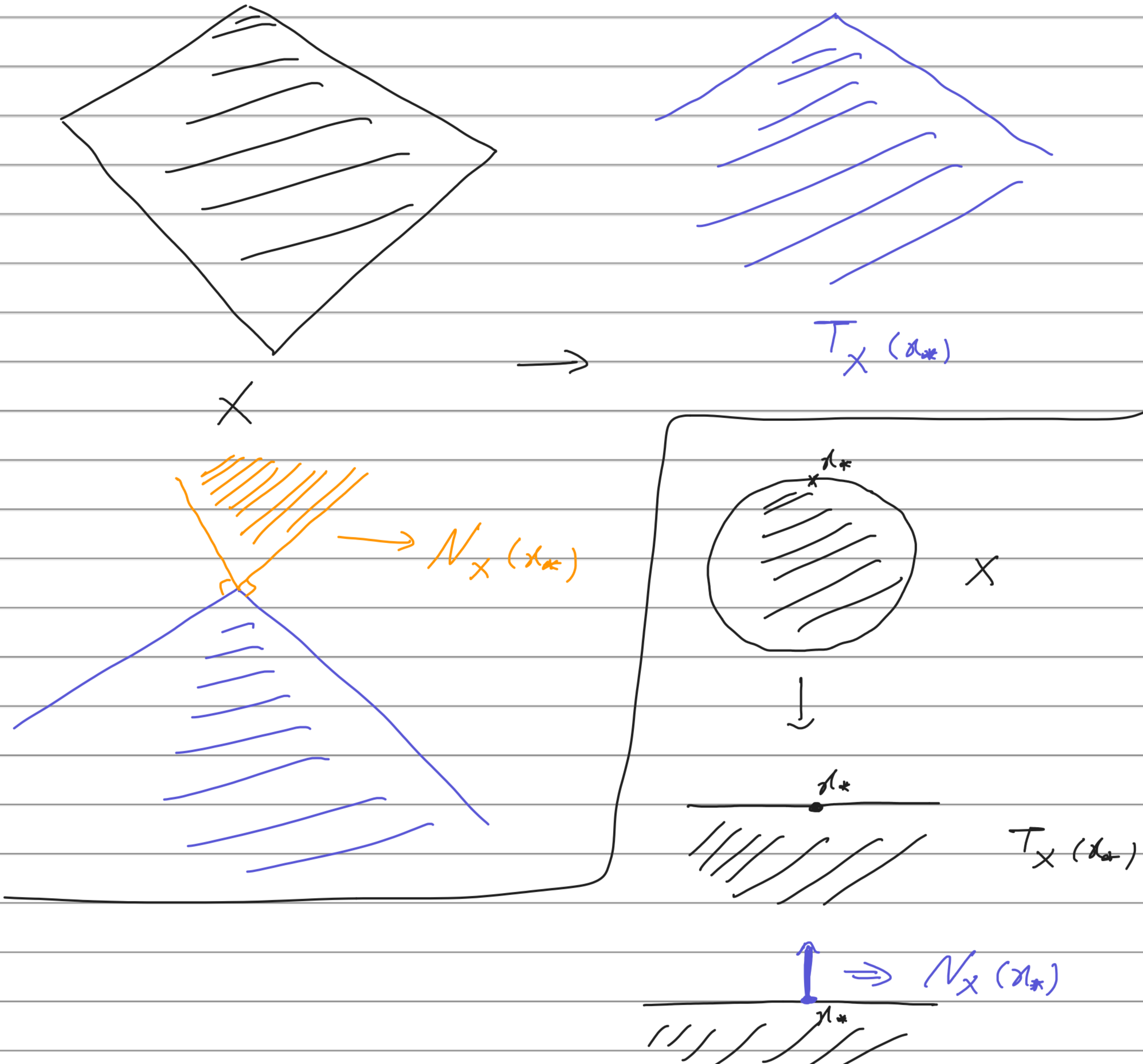




Normal cone at $x_* = N_X(x_*)$

$= \{ \Delta y \mid \Delta y^T \Delta x \le 0 \ , \ \forall \Delta x \in T_X(x_*) \}$

Ex:



FOC : $\nabla f(x_*)^T \underbrace{(x - x_*)}_{\text{multiply by } \alpha \geq 0} \geq 0 \quad \forall x \in X$

Convex

$$\Updownarrow$$

$$\boxed{\nabla f(x_*)^T \Delta x \geq 0 \quad \forall \Delta x \in F_X(x_*)}$$

or $\quad \nabla f(x_*)^T \dfrac{x - x_*}{\|x - x_*\|} \geq 0$

$\Big\downarrow$ take a limit for a sequence $\{x^{(k)}\}$

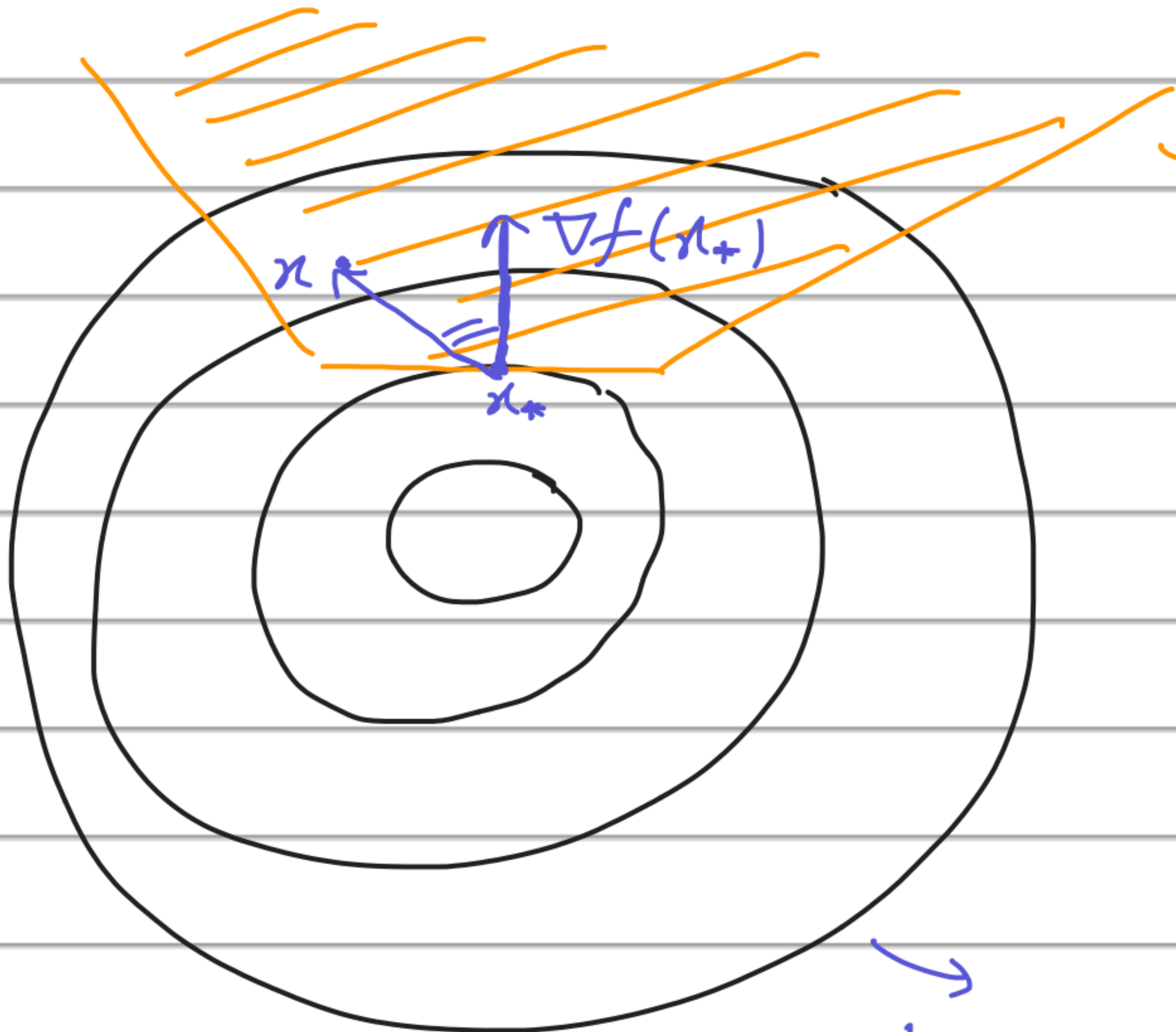
$$\boxed{\nabla f(x_*)^T \Delta x \geq 0 \quad \forall \Delta x \in T_X(x_*)}$$

$$\Updownarrow$$

$$\boxed{-\nabla f(x_*) \in N_X(x_*)}$$ FOC

$X = \mathbb{R}^n \implies N_X(x_*) = 0 \implies \nabla f(x_*) = 0$

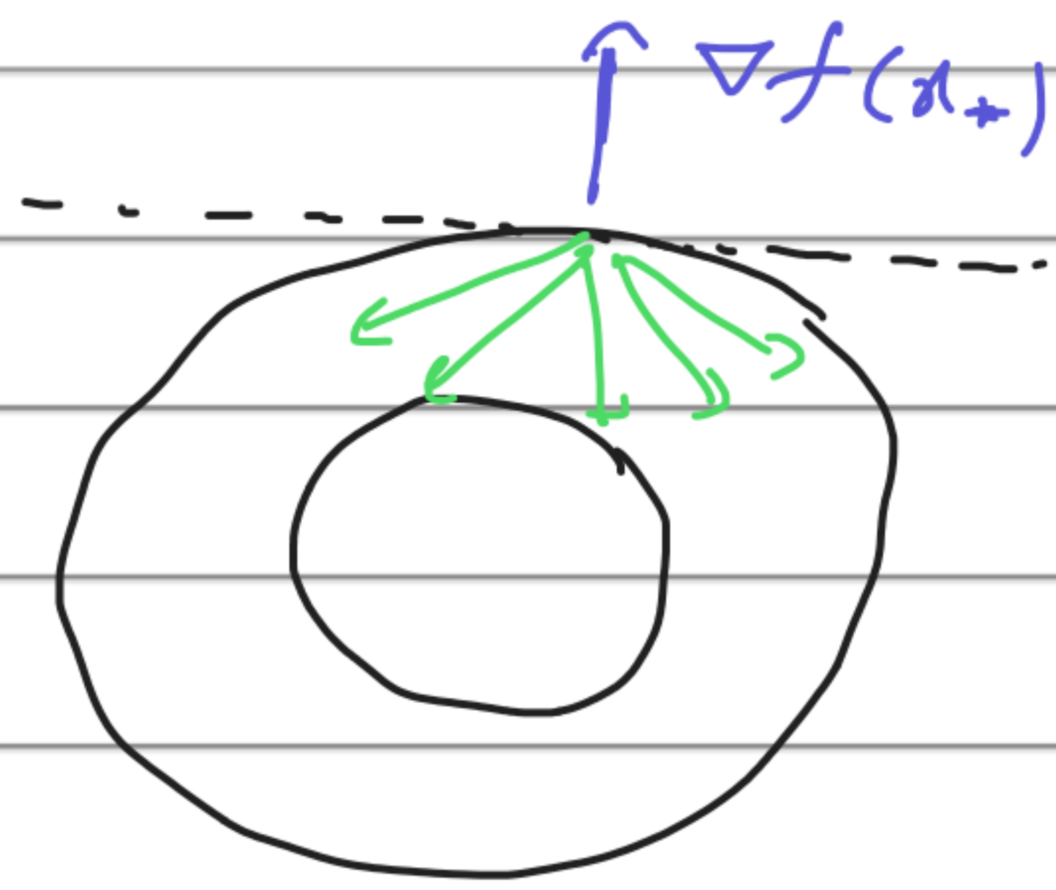
Geometric intuition:



feasible set

angle between $\nabla f(x_*)$ & $x - x_* \leq 90^\circ$

Level sets for $f(\cdot)$

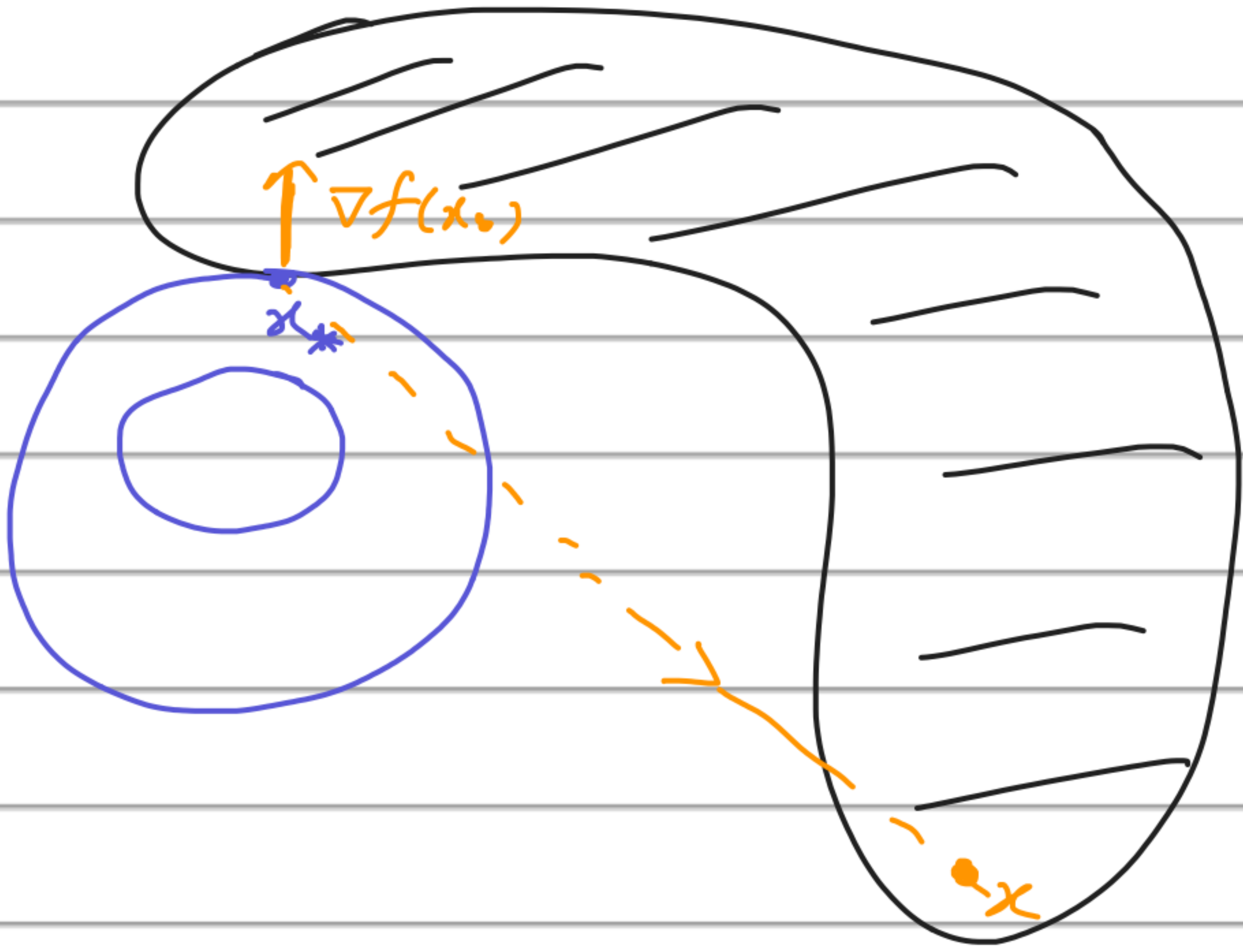


Descent direction:

angle between $\nabla f(x_*)$ & direction $> 90^\circ$

$\Longrightarrow$ This is no feasible descent direction.

This is not true for non-convex X:



$$\Longrightarrow \nabla f(x_*)^T(x - x_*) < 0$$

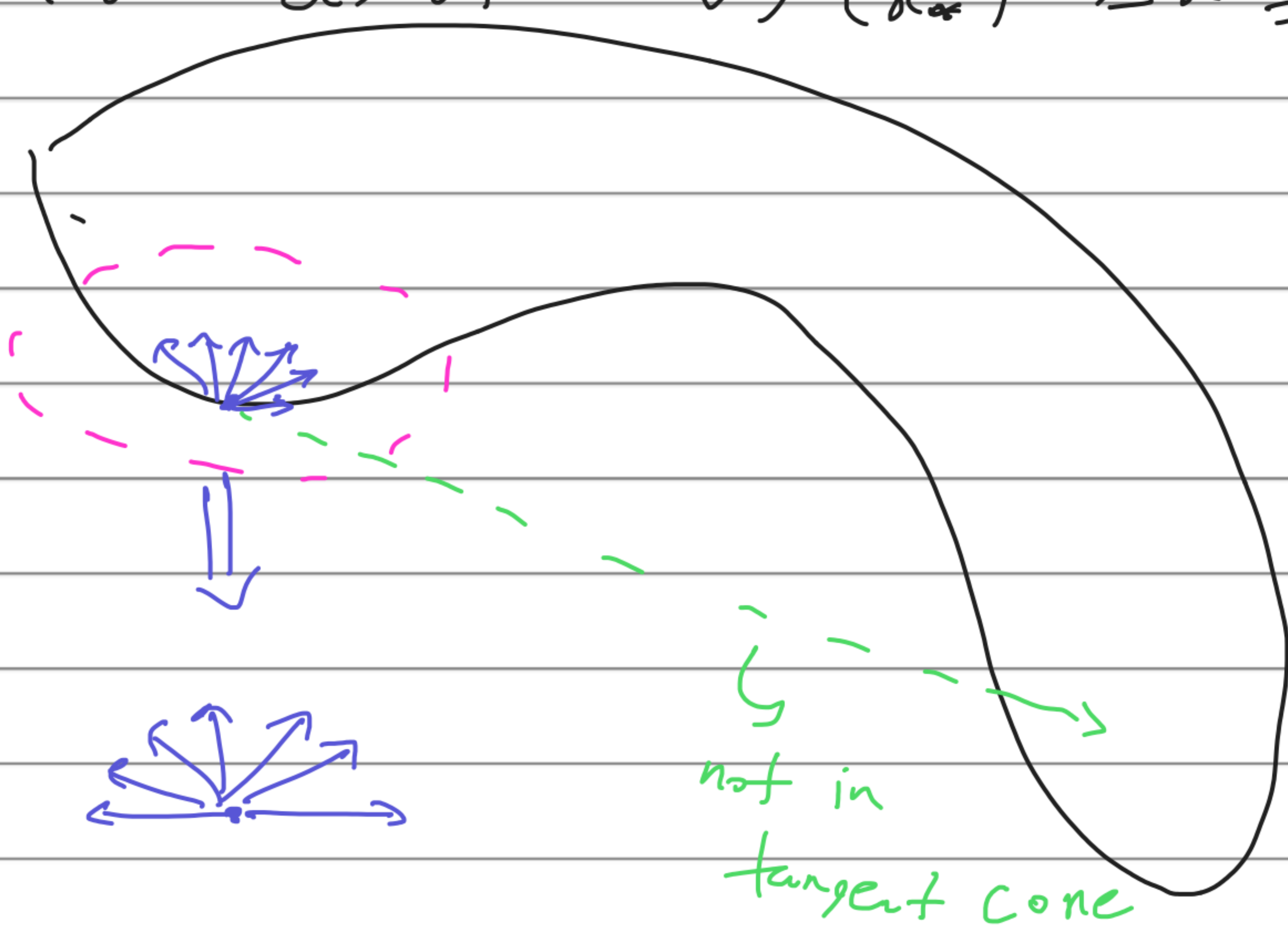
This violates FOC.

So $\nabla f(x_*)^T(x - x_*) \geq 0 \quad \forall x \in X$

(not working)

How about $\nabla f(x_*)^T \Delta x \geq 0 \quad \forall \Delta x \in T_X(x_*)$

works



Def: $\{x^{(k)}\} \subset X$

s.t. $x^{(k)} \to x_*$

SOC (necessary): Assume X = convex set

& $f(x)$ = arbitrary. If x_* is a local

min, then $\Delta x^T \nabla^2 f(x_*) \Delta x \geq 0$

$\forall \Delta x$: $\Delta x \in F_X(x_*)$

$\nabla f(x_*)^T \Delta x = 0$

restricted Hessian

$\boxed{\begin{array}{l} \min \quad f(x) \\ \qquad \downarrow \\ \nabla f(x_*) = 0 \\ \qquad \downarrow \\ \nabla^2 f(x_*) \geq 0 \end{array}}$

proof: Consider Δx : $\Delta x \in F_X(x_*)$,

$\nabla f(x_*)^T \Delta x = 0$

$\Rightarrow \exists \alpha \geq 0$: $\Delta x = \alpha(x - x_*)$, $x \in X$

$f(x_* + t(x - x_*)) - f(x_*) = \boxed{\nabla f(x_*)^T (t(x - x_*))}$

$t \in [0,1]$

$+ \frac{1}{2} (t(x - x_*))^T \nabla^2 f(z(t)) (t(x - x_*))$

converges to x_* as $t \to 0$

zero by assumption

≥ 0 when t is small

$\Rightarrow \Delta x^T \nabla^2 f(x_*) \Delta x \geq 0$

SOC (sufficient) : Assume X = Convex set & $f(x)$ = arbitrary. x_* is a local min if it satisfies:

1) FOC

2) $\Delta x^T \nabla^2 f(x_*) \Delta x > 0 \quad \forall \Delta x$:

s.t. $\Delta x \neq 0$, $\nabla f(x_*)^T \Delta x = 0$

, $\Delta x \in T_X(x_*)$

$\rightarrow$ not $F_X(x_*)$

$\min f(x)$

$\downarrow$

$\nabla f(x_*) = 0$

$\nabla^2 f(x_*) > 0$

$\min f(x)$ s.t. $x \in X$

Counterpart of gradient

Conditional

Projected

Is Convergence slower? $\leftarrow$

Counterpart of newton