262B-Lecture 23

Date created: 2021.04.20
N. of Pages: 19

SDP: $\min_{X \in \mathbb{R}^{n\times n}} \langle D, X \rangle$

inner product: $\langle A,B \rangle = \mathrm{trace}(AB)$

linear

s.t. $\langle A_i, X \rangle = b_i \qquad i = 1, \dots, m$

Conic constraint $X \in C \qquad (X \succeq 0)$

$D, A_1, \dots, A_m$ = symmetric matrices

feasible set = intersection of an affine set

& a PSD cone

Dual: $\max_{\mu \in \mathbb{R}^m} \sum_{i=1}^{m} b_i \mu_i$

s.t. $D - \sum_{i=1}^{m} A_i \mu_i \in C$

Thm: $\begin{cases} \text{- primal : finite optimal value} \\ \text{- } \exists \bar{X} : \text{feasible}, \ \bar{X} \succ 0 \end{cases} \Rightarrow$ no duality gap + existence of dual solution

$\begin{cases} \text{- dual: finite optimal value} \\ \text{- } \exists \bar{\mu} : D - \sum_{i=1}^{m} A_i \bar{\mu}_i \succ 0 \end{cases}$ $\Rightarrow$ no duality gap + existence of primal solution

---

$LP \subseteq QP \subseteq SOCP \overset{?}{\subseteq} SDP$

$$\sqrt{x_1^2 + \cdots + x_{n-1}^2} \leq x_n$$

$$\Longleftrightarrow \quad [x_1 \cdots x_{n-1}] \begin{bmatrix} x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} \leq x_n^2 \quad , \quad x_n \geq 0$$

$$\left[\begin{array}{c|c} A & B \\ \hline B^T & C \end{array}\right] \succ 0 \quad \Longleftrightarrow \quad A \succ 0 \, , \quad A - BC^{-1}B^T \succ 0$$

symmetric 2x2 block matrix

(generalize to $\succeq 0$)

---

schur complement

$$\Longleftrightarrow \quad x_n \geq 0 \, , \quad x_n - [x_1 \cdots x_{n-1}] (x_n I)^{-1} \begin{bmatrix} x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} \geq 0$$

$$\Longleftrightarrow \quad \begin{bmatrix} x_n & [x_1 \cdots x_{n-1}] \\ \begin{bmatrix} x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} & x_n I \end{bmatrix} \succeq 0 \rightarrow \text{PSD constraint}$$

Dual of SDP = SDP

Dual : $LP \subseteq QP \subseteq SOCP \subseteq SDP$

$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow$

$LP \quad QP \quad SOCP \quad SDP$

Dual : $\max\limits_{\mu} \sum_{i=1}^{m} b_i \mu_i \quad$ s.t. $D - \sum_{i=1}^{m} A_i \mu_i \succeq 0$

Variable: vector μ rather than matrix X.

But there is a matrix constraint.

So, there are two ways to write SDPs:

vector variable & matrix variable.

LMI: Linear matrix inequality

LMI is a matrix inequality in the PSD sense

that is linear in the variable.

Ex: $\underbrace{D - \sum_{i=1}^{m} A_i \mu_i}_{\text{linear in } \mu} \succeq 0$

Ex: $x \in \mathbb{R}^2$

LMI: $\begin{bmatrix} 1+x_1 & x_1-2x_2 \\ x_1-2x_2 & 3+x_1+x_2 \end{bmatrix} \preceq 0$

two LMIs can be merged into one LMI:

$$\begin{cases} A_0 + \sum_{i=1}^{m} A_i x_i \succeq 0 \\ B_0 + \sum_{i=1}^{m} B_i x_i \succeq 0 \end{cases}$$

$$\iff \begin{bmatrix} A_0 & 0 \\ 0 & B_0 \end{bmatrix} + \sum_{i=1}^{m} \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix} x_i \succeq 0$$

SDP: minimize a linear objective subject to LMIs.

How to formulate problems as SDP?

Ex: Given $A_0, \dots, A_n \in \underset{\text{symmetric } n\times n}{S^n}$, solve

$$\min_x \lambda_{max}\big(\underbrace{A_0 + A_1 x_1 + \dots + A_n x_n}_{\text{parametrized matrix}}\big)$$

$$\Longleftrightarrow \quad \min_{\substack{x \in \mathbb{R}^n \\ \alpha \in \mathbb{R}}} \ \alpha \qquad \text{(SDP)}$$

$$\text{s.t.} \quad \underbrace{A_0 + A_1 x_1 + \cdots + A_n x_n \preceq \alpha I}_{\text{LMI in } (x,\alpha)}$$

Ex: $\max_{x} \ \lambda_{min}(A_0 + A_1 x_1 + \cdots + A_n x_n)$

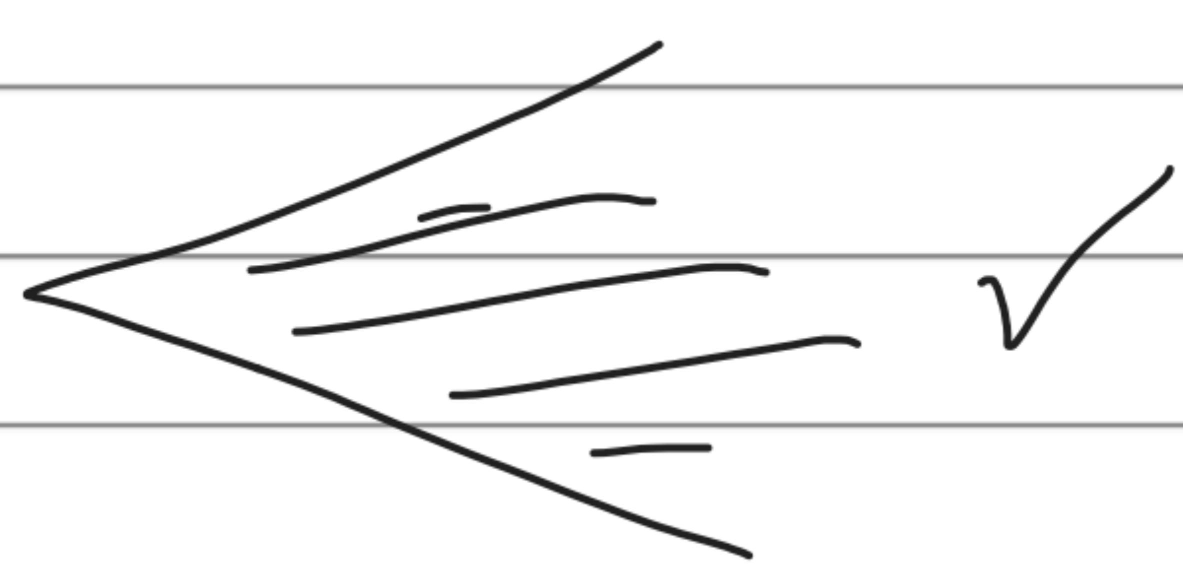
$$\Updownarrow$$

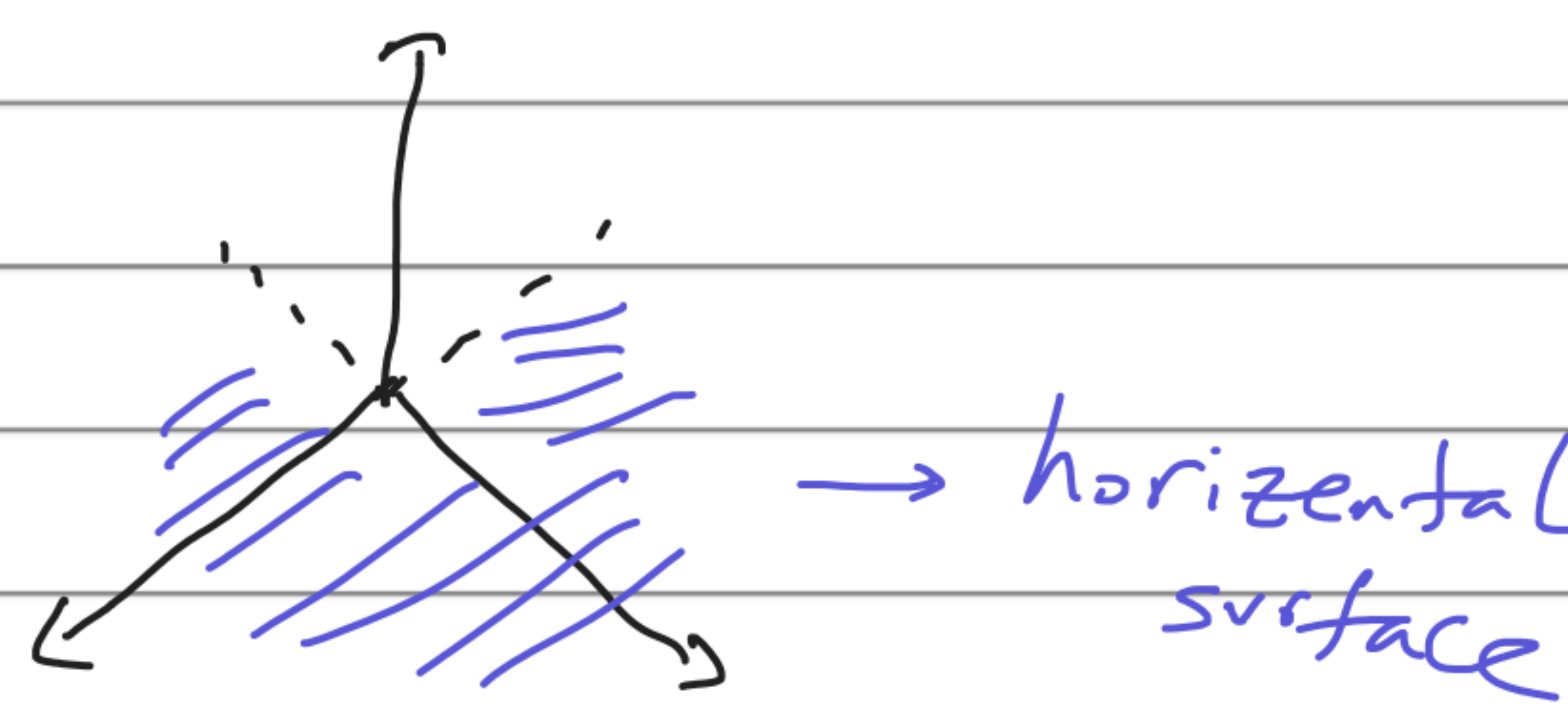
$$\max_{x,\alpha} \ \alpha \qquad \text{(SDP)}$$

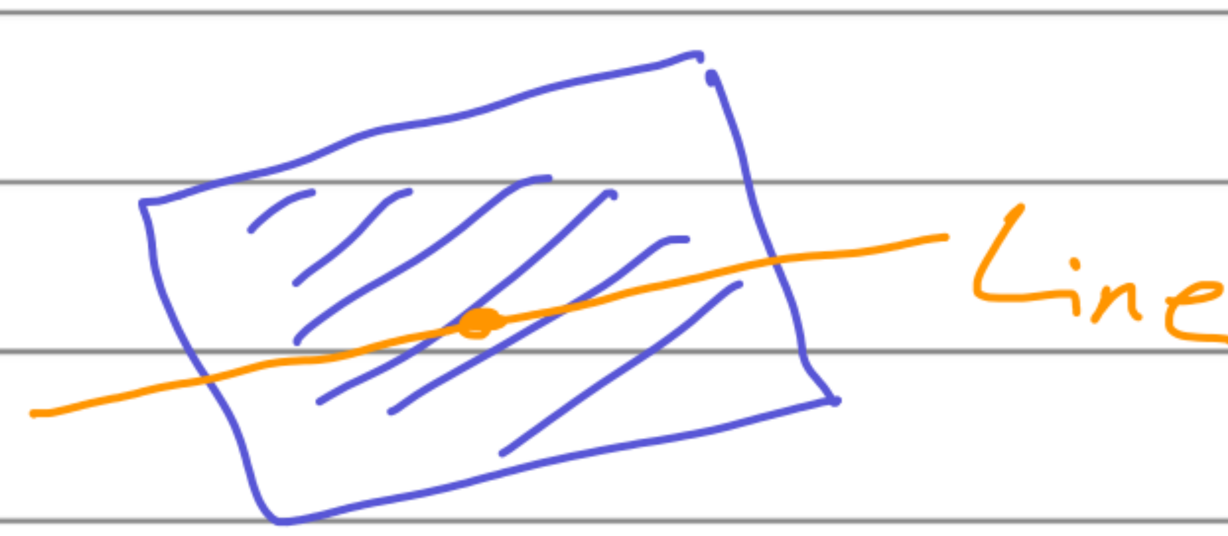
$$\text{s.t.} \quad \alpha I \preceq A_0 + A_1 x_1 + \cdots + A_n x_n$$

Generalized inequality :

Proper Cone : Convex, closed, non-empty interior, contains no line.







Consider a proper cone C:

Define the notion of partial ordering:

- say $x \preccurlyeq_C y$ if $y - x \in C$

$(0 \preccurlyeq_C y \implies y \in C)$

- say $x \prec_C y$ if $y - x \in \underbrace{\text{int}}_{\text{interior}} (C)$

This is not perfect ordering and is partial ordering since it could happen that $\exists\, x, y$ s.t. $x \not\preccurlyeq_C y$, $y \not\preccurlyeq_C x$

$C = PSD$: $x = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

If $C = \mathbb{R}_+ \implies$ then $x \preccurlyeq_C y$ means $x \leq y$ or $y - x \geq 0$ in a regular sense properties:

1 - addition:

$$x \preceq_C y \ , \ u \preceq_C v \implies x + u \preceq_C y + v$$

2 - transitive

$$x \preceq_C y \ , \ y \preceq_C z \implies x \preceq_C z$$

3 - non-negative scaling:

$$x \preceq_C y \ , \ \alpha \geq 0 \implies \alpha x \preceq_C \alpha y$$

4 - reflexive : $x \preceq_C x$

5 - antisymmetric: $x \preceq_C y \ , \ y \preceq_C x$

$$\implies x = y$$

Optimization with generalized inequalities:

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$

$$\text{s.t.} \quad h_i(x) = 0 \qquad i = 1, \dots, m$$

$$g_j(x) \leq 0 \qquad j = 1, \dots, r$$

scalar

standard opt

generalized

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$

$$\text{s.t.} \quad h_i(x) = 0 \qquad i = 1, \dots, m$$

$$g_j(x) \preceq_{C_j} 0 \qquad j = 1, \dots, r$$

$\in \mathbb{R}^{n_j}$

$C_1, \dots, C_r$ = proper cones

f, h_i, g_j = non-convex

Lagrange multipliers:

$h_i(x) = 0 \longrightarrow \lambda_i \in \mathbb{R}$

$g_j(x) \preceq_{C_j} 0 \longrightarrow \mu_j \in C_j^*$ (dual cone)

Lagrangian:

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{r} \underbrace{\underbrace{\mu_j^T}_{\in \mathbb{R}^{n_j}} \underbrace{g_j(x)}_{\in \mathbb{R}^{n_j}}}_{\text{inner product}}$$

Dual function:

$$q(\lambda, \mu) = \inf_x L(x, \lambda, \mu)$$

Dual optimization: $\max_{\lambda, \mu} q(\lambda, \mu)$

s.t. $\mu_j \succeq_{C_j^*} 0 \quad j = 1, \ldots, r$

Define: $f_* =$ optimal primal value

$q_* =$ optimal dual value

weak duality : $q_* \leq f_*$

strong duality : ?

function : convexity $\longrightarrow$ Cone - Convexity

$(g_j(x) \leq 0)$ $\qquad$ $(g_j(x) \preceq_{C_j} 0)$

C : proper cone $\in \mathbb{R}^l$

$g : \mathbb{R}^n \longrightarrow \mathbb{R}^l$

$g(x)$ is called C- cone if

$$g(\alpha x + (1-\alpha) y) \preceq_C \alpha g(x) + (1-\alpha) g(y)$$

$\forall x, y$, $\forall \alpha \in [0,1]$ $\qquad$ scalar, vector, matrix

- If $l = 1$, $C = \mathbb{R}_+$ $\Rightarrow$ Cone - Convexity
= regular convexity

$g(x)$: Linear $\Rightarrow$ g : Cone - Convex

$$\begin{cases} \min & f(x) \\ \text{s.t} & h_i(x) = 0 \quad i = 1, \dots, m \\ & g_j(x) \preceq_{C_j} 0 \quad j = 1, \dots, r \end{cases}$$

strong duality: If

- $f(x)$ = convex
- $h_i(x)$ = Linear
- $g_j(x)$ = C_j - convex

and - slater holds ($\exists\ \bar{x}$ s.t.

$h_i(\bar{x}) = 0 \quad i = 1, \dots, m, \quad g_j(\bar{x}) \prec_{C_j} 0 \quad j = 1, \dots, r$)

$\Rightarrow$ strong duality holds ($f_* = q_*$)

+ existence of dual solution

If: f, h_i, g_j = Linear functions

$\Longrightarrow$ Linear conic optimization

SDP:

$$\min_X \ \langle D, X \rangle$$

$$\text{s.t.} \quad \langle A_i, X \rangle = b_i \qquad i = 1, \dots, m$$

$$X \in C \Longleftrightarrow X \succeq 0$$

$$\Updownarrow$$

$$-X \preceq 0$$

Define Lagrange multipliers:

$$\langle A_i, X \rangle = b_i \longrightarrow \lambda_i$$

$$-X \preceq 0 \longrightarrow Y \in C_* \longrightarrow Y \succeq 0$$

$$L(X, \lambda, Y) = \langle D, X \rangle + \sum_{i=1}^{m} (\langle A_i, X \rangle - b_i)\lambda_i$$

$$+ \underbrace{\langle -X, Y \rangle}_{\text{inner product of constraint \& dual variable}}$$

$$\Rightarrow \quad L(X, \lambda, Y) = \left\langle D + \sum_{i=1}^{m} A_i \lambda_i - Y \,,\; X \right\rangle + \left(- \sum_{i=1}^{m} b_i \lambda_i \right)$$

$$q(\lambda, Y) = \min_{X} L(X, \lambda, Y)$$

$$\min_{X} \underbrace{\left\langle D + \sum_{i=1}^{m} A_i \lambda_i - Y \,,\; X \right\rangle}_{\text{Linear function, } X = \text{arbitrary}}$$

$$= \begin{cases} 0 & \text{if } D + \sum_{i=1}^{m} A_i \lambda_i - Y = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Dual opt:

$$\max_{\lambda, Y} \; - \sum_{i=1}^{m} b_i \lambda_i$$

$$\text{s.t.} \quad D + \sum_{i=1}^{m} A_i \lambda_i - Y = 0$$

$$Y \succeq 0$$

$$D + \sum_{i=1}^{m} A_i \lambda_i \succeq 0 \iff$$

Convexification:

non-convex opt $\longrightarrow$ $\underbrace{\text{FOC, SOC, algorithms}}$

$\downarrow$

local min or saddle

how do find a global min $\longleftarrow$

extreme case:

$$\min_x \quad x^T Q x$$

$$\text{s.t.} \quad x_i^2 = 1 \qquad i = 1, \dots, n$$

2^n feasible points

$\rightarrow$ 2^n local minima $\longrightarrow$ so, notion of local minima is useless here

$$\min_x \quad f(x) \longrightarrow \text{non-convex}$$

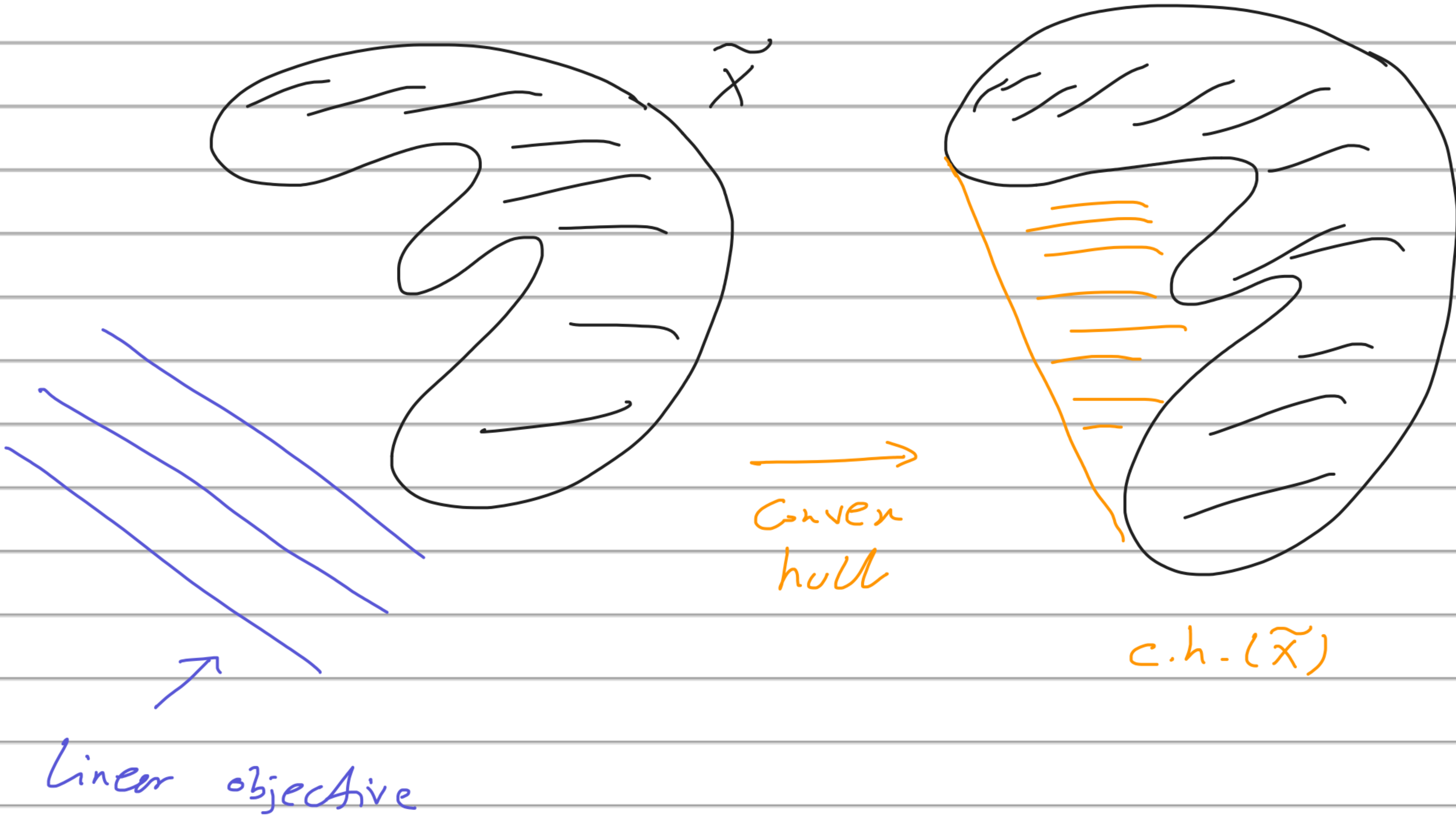
$$\text{s.t.} \quad x \in X$$

non-convex feasible set

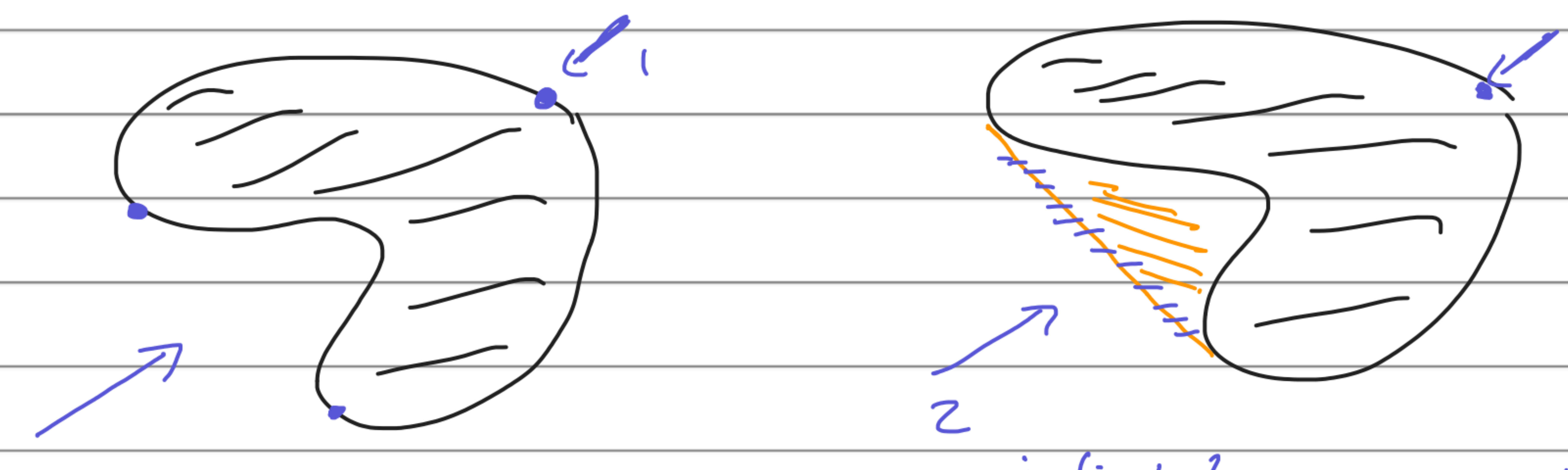
$\longrightarrow$ make the objective linear

$$\min_{x,\alpha} \quad \alpha$$

$$\text{s.t.} \quad (x,\alpha) \in \tilde{X}, \quad \tilde{X} = \{(x,\alpha) \mid x \in X, f(x) \le \alpha\}$$



minimization of a linear function over a non-convex set $\tilde{X}$ and its convex hull c.h.($\tilde{X}$) gives the same optimal objective value.

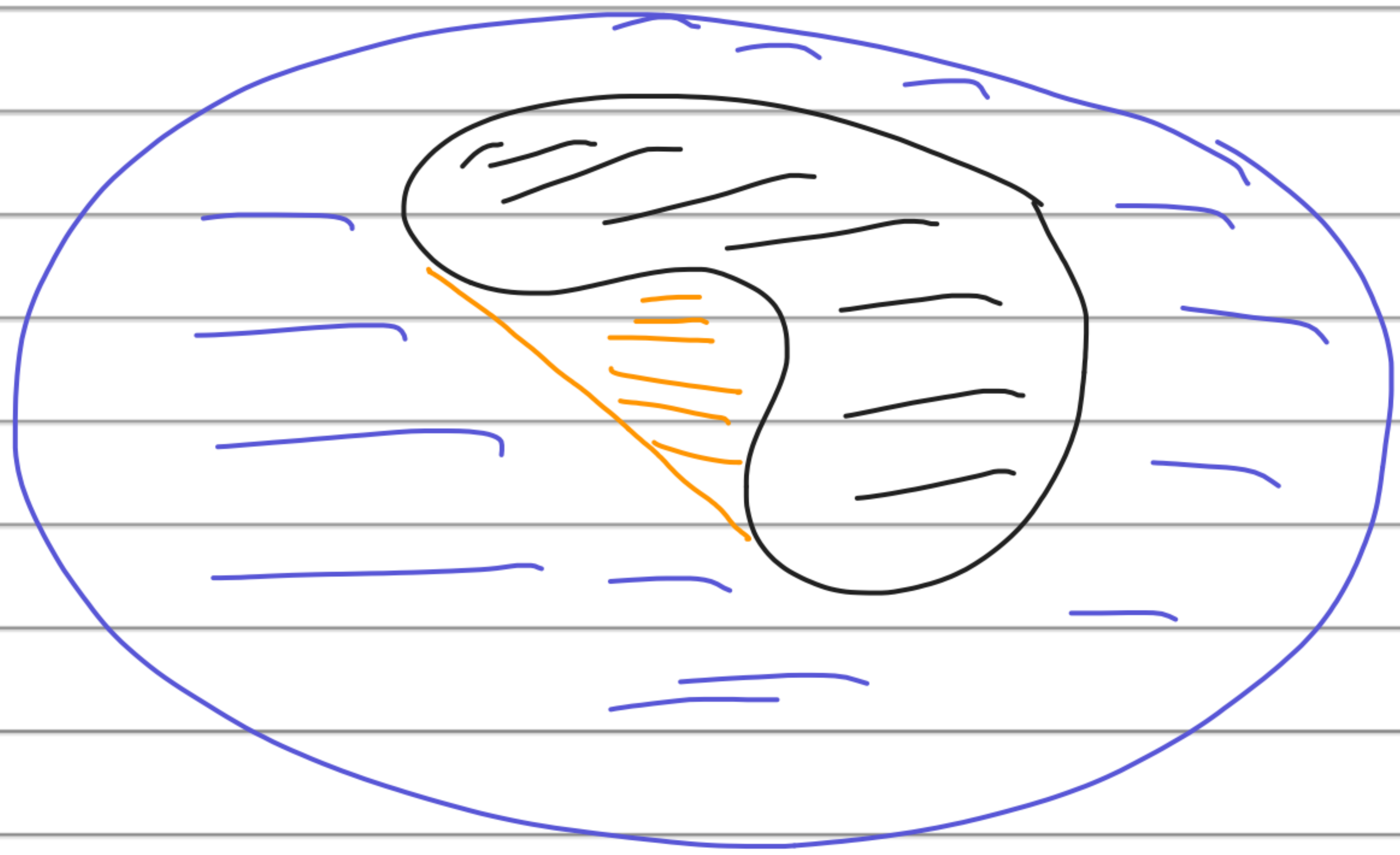


infinitely many solutions
(including the correct solutions)

So, to avoid finding a non-global minimum, we should find convex hull.

This is NP-hard.

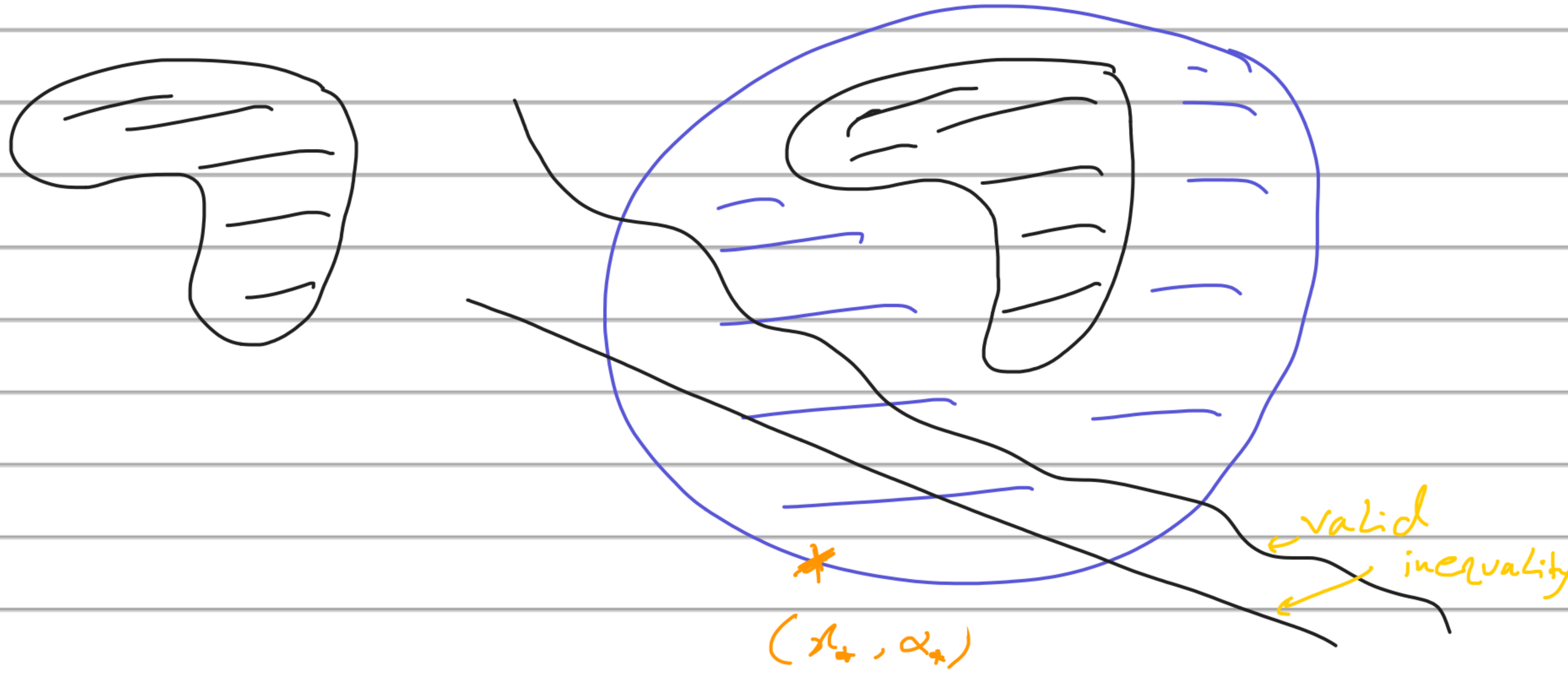
Remedy: Convex relaxation



minimize over a convex set that includes the original feasible set.

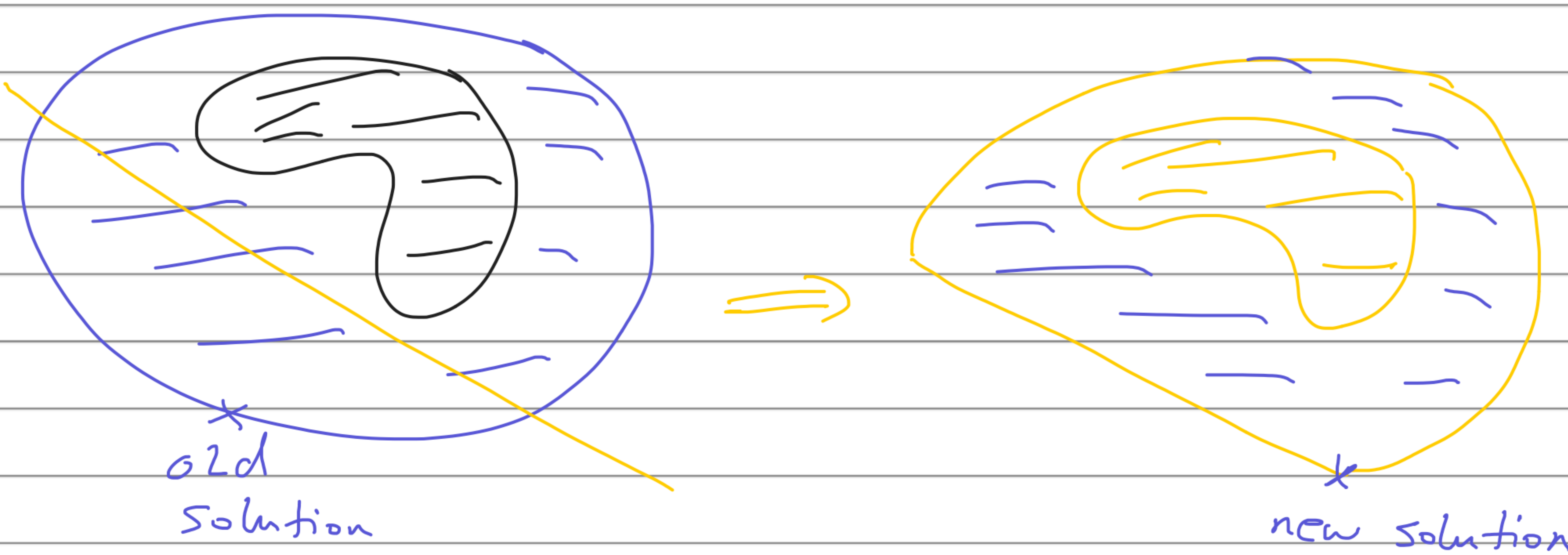
$$\begin{aligned} \min\ & \alpha \\ \text{s.t}\ & (x,\alpha) \in \tilde{X} \end{aligned} \longrightarrow \begin{aligned} \min\ & \alpha \\ \text{s.t}\ & (x,\alpha) \in \underbrace{\hat{X}}_{\text{convex } \& \supseteq \tilde{X}} \end{aligned}$$

optimal objective value reduces (serves as a lower bound) but solution could be infeasible for original problem.



what if we add a convex constraint $K(x) \leq 0$ that is satisfied for $\forall x \in \tilde{X}$ but $K(x_*) \not\leq 0$. This is called valid inequality.

If $K(x)$ = linear $\Rightarrow$ cutting plane.



- This generates a sequence of convex problems.
- It turns out that under mild conditions almost every smooth non-convex optimization can be solved to global optimality by

LP → LP → LP → .... ⇒ LP hierarchy

SOCP → SOCP → SOCP → ... ⇒ SOCP hierarchy

SDP → SDP → .... ⇒ SDP hierarchy

Usually: dimension of LP/SOCP/SDP grows as we take more iterations in the sequence (lifting).

$$\underbrace{LP}_{10 \text{ variables}} \rightarrow \underbrace{LP}_{100} \rightarrow \underbrace{LP}_{1000} \rightarrow \cdots$$

Complexities:	LP	SOCP	SDP
Complexity per iteration:	fast	OK	slow
iteration complexity:	many	OK	a few

state of the art: SDP-based relaxation for arbitrary non-convex (mixed-integer) problems.