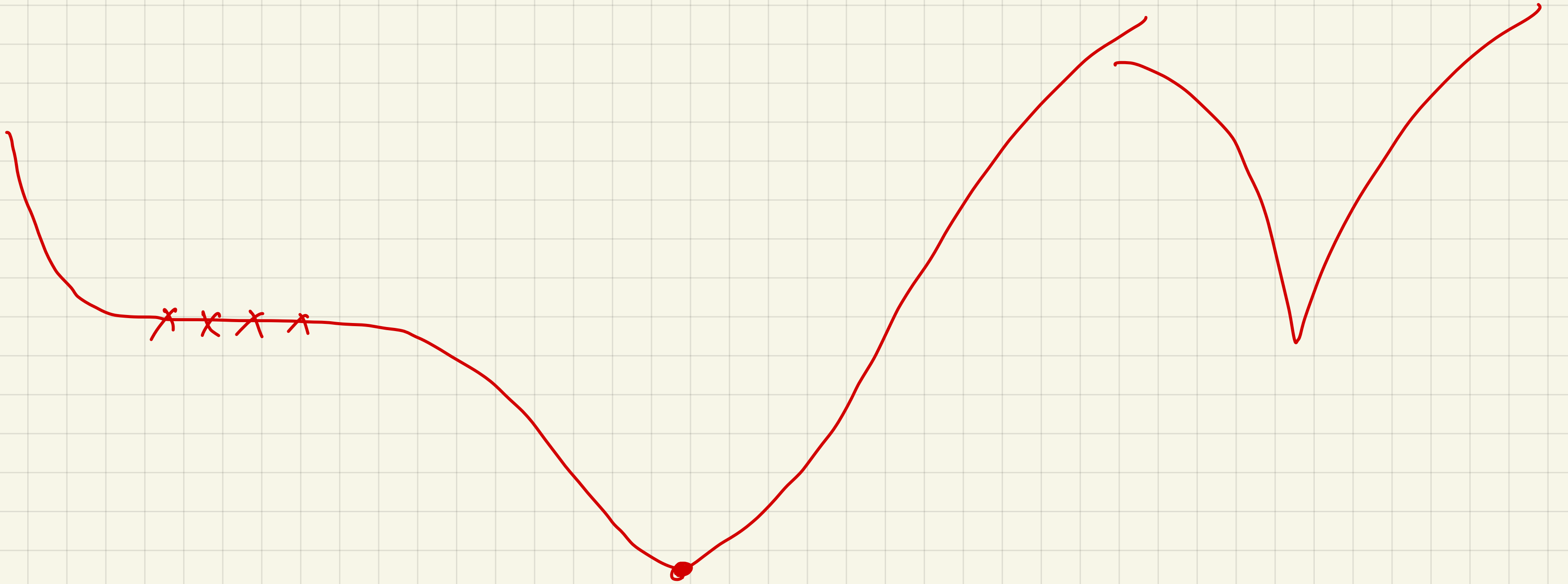


4. Convex optimization problems

- optimization problem in standard form
- convex optimization problems
- quasiconvex optimization
- linear optimization
- quadratic optimization
- geometric programming
- generalized inequality constraints
- semidefinite programming
- vector optimization

Optimization problem in standard form

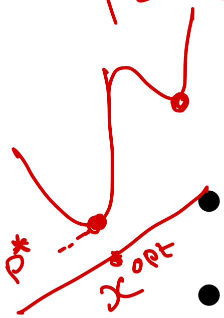
$f_0 \; f_1 \; f_2 \; f_3$

$h(x) = 0$

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\
& h_i(x) = 0, \quad i = 1, \ldots, p
\end{array}
$$

$h_1(x) \le 0$

$h_1(x) \ge 0$



- $x \in \mathbf{R}^n$ is the optimization variable
- $f_0 : \mathbf{R}^n \to \mathbf{R}$ is the objective or cost function
- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, m$, are the inequality constraint functions
- $h_i : \mathbf{R}^n \to \mathbf{R}$ are the equality constraint functions

optimal value:

$$p^\star = \inf\{f_0(x) \mid f_i(x) \le 0,\ i = 1, \ldots, m,\ h_i(x) = 0,\ i = 1, \ldots, p\}$$

- $p^\star = \infty$ if problem is infeasible (no x satisfies the constraints)

$\inf(\{\ \}) = \infty$

- $p^\star = -\infty$ if problem is unbounded below

$\hookrightarrow -\log x$
$x > 0$

minimize $f(x) = x$

$2 < x \leq 10$

$p^* = 2$ ✓

x

log x

$x > -2$

Optimal and locally optimal points

x is **feasible** if $x \in \mathbf{dom}\, f_0$ and it satisfies the constraints

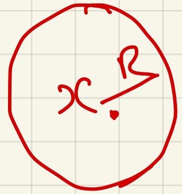
a feasible x is **optimal** if $f_0(x) = p^\star$; X_{opt} is the set of optimal points

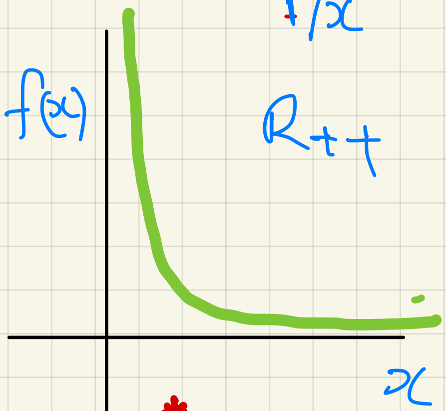
x is **locally optimal** if there is an $R > 0$ such that x is optimal for

$$
\begin{array}{ll}
\text{minimize (over } z\text{)} & f_0(z) \\
\text{subject to} & f_i(z) \le 0, \quad i = 1, \ldots, m, \quad h_i(z) = 0, \quad i = 1, \ldots, p \\
 & \|z - x\|_2 \le R
\end{array}
$$

examples (with $n = 1$, $m = p = 0$)

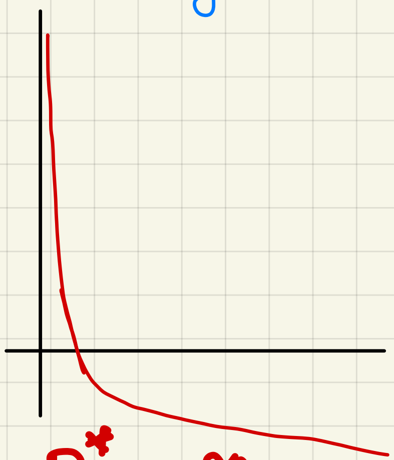
- $f_0(x) = 1/x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = 0$, no optimal point
- $f_0(x) = -\log x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = -\infty$
- $f_0(x) = x \log x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = -1/e$, $x = 1/e$ is optimal
- $f_0(x) = x^3 - 3x$, $p^\star = -\infty$, local optimum at $x = 1$





$p^* = 0$

No optimal point

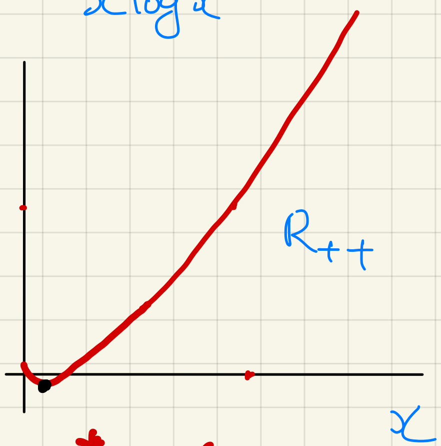


$p^* = -\infty$

No optimal point

$f'(x) = \frac{x}{x} + \log x = 0$

$\Rightarrow x = 1/e$



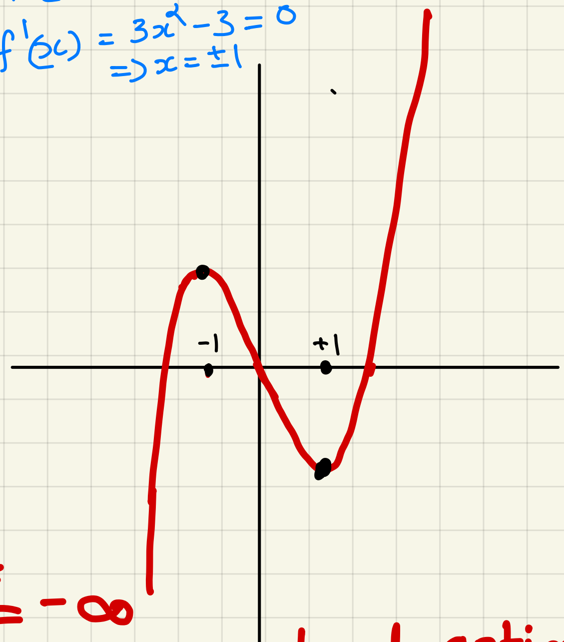
$p^* = -1/e$

$x = 1/e$ is optimal

$f(x) = x^3 - 3x$

$f'(x) = 3x^2 - 3 = 0$

$\Rightarrow x = \pm 1$



$p^* = -\infty$

$x = +1$ is a local optimal

Implicit constraints

the standard form optimization problem has an **implicit constraint**

$$x \in \mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom}\, f_i \;\cap\; \bigcap_{i=1}^{p} \mathbf{dom}\, h_i,$$

- we call $\mathcal{D}$ the **domain** of the problem
- the constraints $f_i(x) \leq 0$, $h_i(x) = 0$ are the explicit constraints
- a problem is **unconstrained** if it has no explicit constraints ($m = p = 0$)

example:

$$\text{minimize} \quad f_0(x) = -\sum_{i=1}^{k} \log(b_i - a_i^T x)$$

is an unconstrained problem with implicit constraints $a_i^T x < b_i$

Implicit and Explicit Constraints

minimize $-\log x$

EC. $-2 \le x \le 10$

IC. $x > 0$

$$\text{minimize} \quad f_0(x)$$
$$\text{subject to} \quad f_i(x) \le 0, \quad i = 1, 2, \cdots m$$
$$h_i(x) = 0, \quad i = 1, 2, \cdots p$$

Explicit Constraints

$$f_i(x) \le 0, \quad i = 1, 2, \cdots m$$
$$h_i(x) = 0, \quad i = 1, 2, \cdots p$$

No explicit constraints

$\Rightarrow$ Unconstrained problem

Implicit Constraints

$$x \in \bigcap_{i=0}^{m} \operatorname{dom} f_i \ \bigcap \ \bigcap_{i=1}^{p} \operatorname{dom} h_i$$

Feasibility problem

$$
\begin{array}{ll}
\text{find} & x \\
\text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& h_i(x) = 0, \quad i = 1, \ldots, p
\end{array}
$$

can be considered a special case of the general problem with $f_0(x) = 0$:

$$
\begin{array}{ll}
\text{minimize} & 0 \\
\text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& h_i(x) = 0, \quad i = 1, \ldots, p
\end{array}
$$

- $p^\star = 0$ if constraints are feasible; any feasible x is optimal
- $p^\star = \infty$ if constraints are infeasible

Convex optimization problem

$x \in \mathbf{R}^n.$

standard form convex optimization problem

$h_i(x) = a^T x + b$

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& a_i^T x = b_i, \quad i = 1, \ldots, p
\end{array}
$$

$h_i(x) = 0$

$x^2 = 1$

$a^T x = -b$

$a^T x + b = 0$

- $f_0, f_1, \ldots, f_m$ are convex; equality constraints are affine
- problem is *quasiconvex* if f_0 is quasiconvex (and $f_1, \ldots, f_m$ convex)

A

$\begin{bmatrix} & \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = a_1$

often written as

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
$$

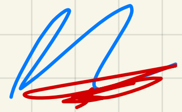
$A_{p \times n}$

important property: feasible set of a convex optimization problem is convex

Abstract form convex optimization problem:

minimize convex objective function

over convex set.



example

abstract form → Convex

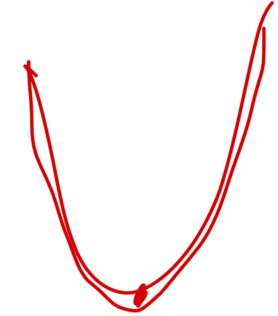
$$\begin{array}{ll} \text{minimize} & f_0(x) = x_1^2 + x_2^2 \\ \text{subject to} & f_1(x) = x_1/(1+x_2^2) \leq 0 \\ & h_1(x) = (x_1+x_2)^2 = 0 \end{array}$$

$\Rightarrow x_1 \leq 0$

Standard form → Non Convex

- f_0 is convex; feasible set $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ is convex
- not a convex problem (according to our definition): f_1 is not convex, h_1 is not affine
- equivalent (but not identical) to the convex problem

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 \\ \text{subject to} & x_1 \leq 0 \\ & x_1 + x_2 = 0 \end{array}$$



Local and global optima

any locally optimal point of a convex problem is (globally) optimal

proof: suppose x is locally optimal, but there exists a feasible y with $f_0(y) < f_0(x)$

x locally optimal means there is an $R > 0$ such that

$$z \text{ feasible}, \quad \|z - x\|_2 \leq R \quad \Longrightarrow \quad f_0(z) \geq f_0(x)$$

consider $z = \theta y + (1 - \theta)x$ with $\theta = R/(2\|y - x\|_2)$ $< \frac{R}{2 \cdot R}$

Convexity of feasible set

- $\|y - x\|_2 > R$, so $0 < \theta < 1/2$
- z is a convex combination of two feasible points, hence also feasible
- $\|z - x\|_2 = R/2$ and $\leq R$

$$f_0(z) \leq \theta f_0(y) + (1 - \theta) f_0(x) < f_0(x)$$

$\theta y + (1-\theta) y$

Convexity of f_0

which contradicts our assumption that x is locally optimal

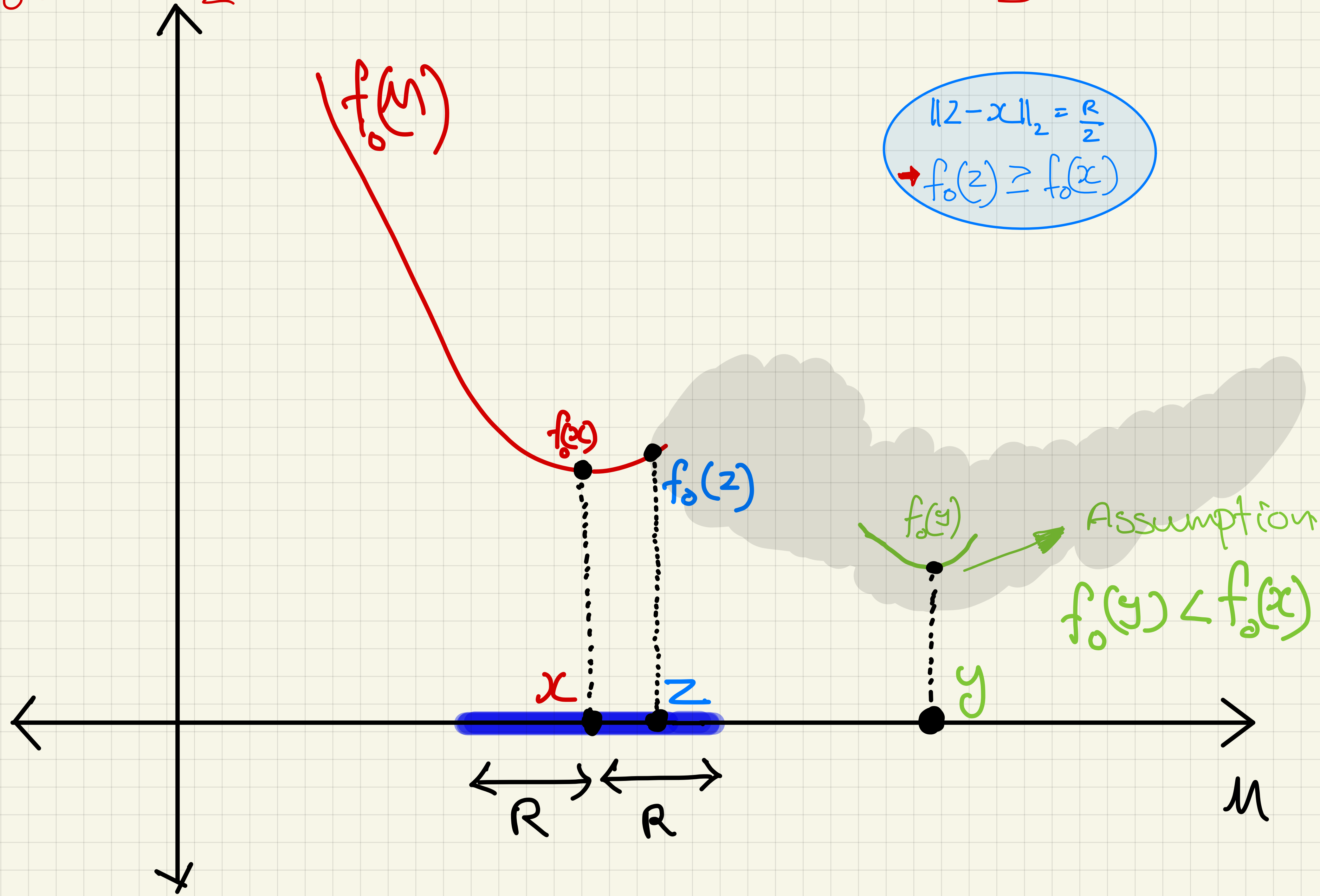
$\|z - x\|_2$

$= \|\theta y + (1-\theta)x - x\|_2$

$= \theta \|y - x\|_2$

$= \frac{R}{2\cancel{\|y-x\|_2}} \cdot \cancel{\|y - x\|_2}$

$$f_0(x) = \inf_z \{ f_0(z) \mid z \text{ is feasible}, \|z-x\|_2 \le R \}$$



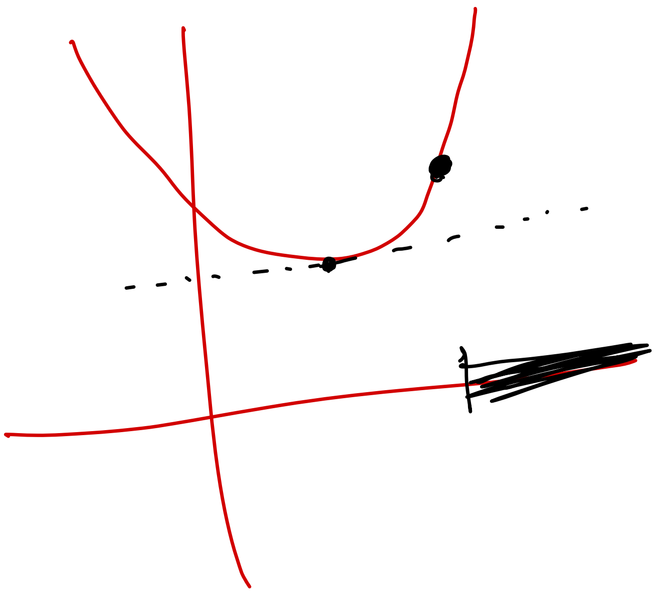
Optimality criterion for differentiable f_0

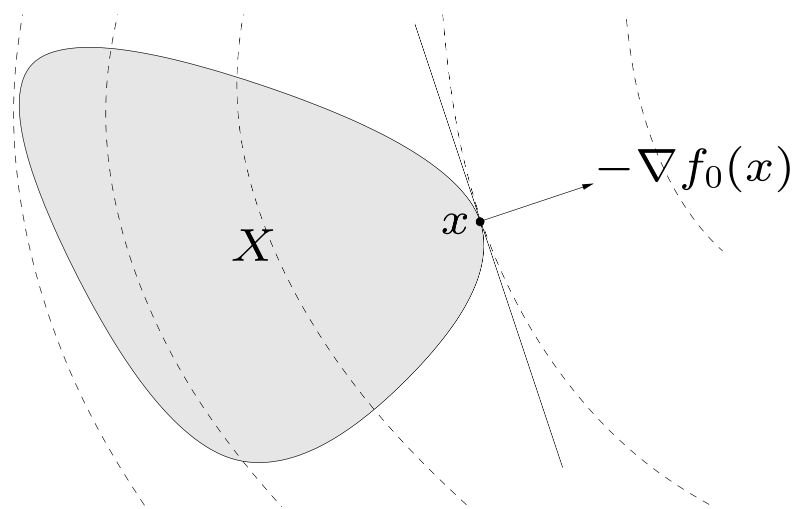
x is optimal if and only if it is feasible and

$$\nabla f_0(x)^T (y - x) \geq 0 \quad \text{for all feasible } y$$

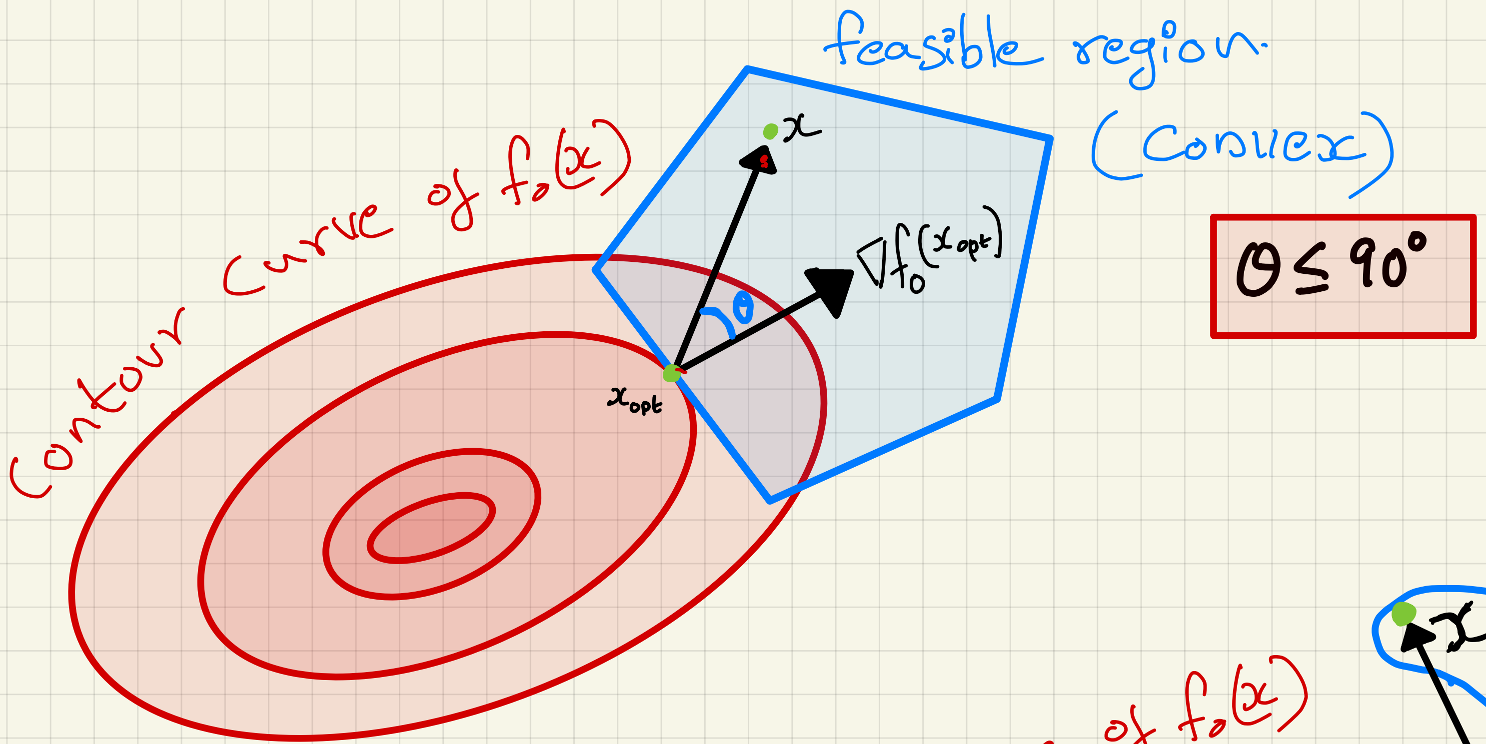
y x_{opt}

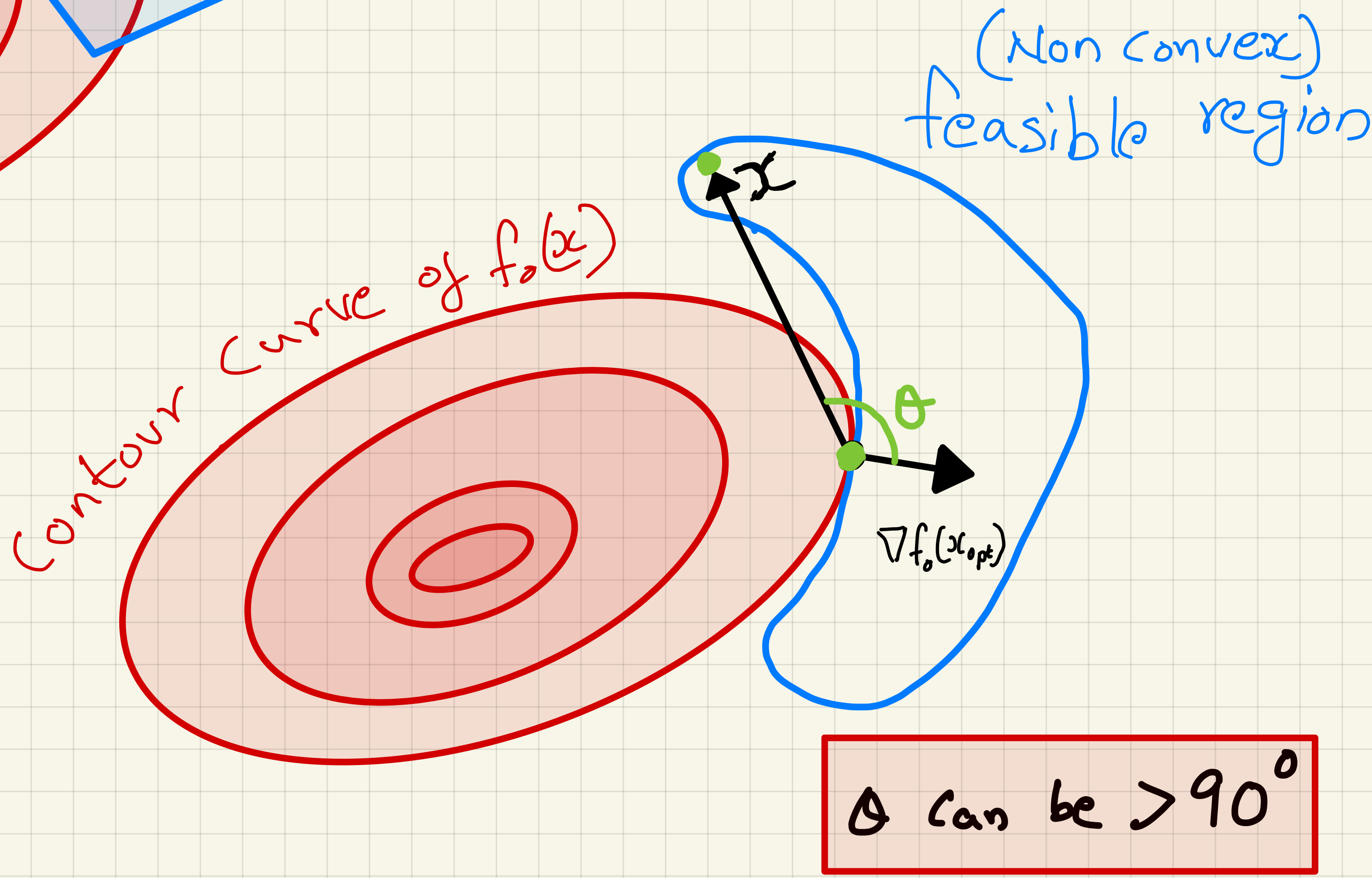
$\nabla f_0(x)^T y \geq \nabla f_0(x)^T x.$





if nonzero, $\nabla f_0(x)$ defines a supporting hyperplane to feasible set X at x


feasible region.
(Convex)
$\theta \le 90^\circ$
Contour curve of $f_0(x)$
x
$\nabla f_0(x_{opt})$
θ
x_{opt}


(Non convex)
feasible region
x
Contour curve of $f_0(x)$
θ
$\nabla f_0(x_{opt})$
θ can be $> 90^\circ$

$\nabla f_0(x)^T (y - x) \geq 0 \quad \forall y \in \mathbb{R}^n$

- **unconstrained problem**: x is optimal if and only if

Let $(y - x) = -\nabla f_0(x)$

$$x \in \mathbf{dom}\, f_0, \qquad \nabla f_0(x) = 0$$

$\Rightarrow -(\nabla f_0(x)^T \nabla f_0(x)) \geq 0$

$-\|\nabla f_0(x)\|_2^2 \geq 0$

$\Rightarrow \nabla f_0(x) = 0$

- **equality constrained problem**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad Ax = b$$

x is optimal if and only if there exists a ν such that

$$x \in \mathbf{dom}\, f_0, \qquad Ax = b, \qquad \nabla f_0(x) + A^T\nu = 0$$

(Lagrange multiplier)

- **minimization over nonnegative orthant**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad x \succeq 0$$

x is optimal if and only if

$\nabla f_0(x)_i \, x_i = 0$

$$x \in \mathbf{dom}\, f_0, \qquad x \succeq 0, \qquad \begin{cases} \nabla f_0(x)_i \geq 0 & x_i = 0 \\ \nabla f_0(x)_i = 0 & x_i > 0 \end{cases}$$

Complimentarity

$$\nabla f_0(\underline{x})^T(\underline{y}-\underline{x}) \geq 0 \quad \forall y$$

$$A\underline{y} = \underline{b}$$

$$A\underline{x} = \underline{b}$$

$$A(\underline{y}-\underline{x}) = 0 \Rightarrow \underline{y}-\underline{x} \in \mathcal{N}(A)$$

$$\nabla f_0(x)^T \underline{v} \geq 0 \quad \forall v \in \mathcal{N}(A)$$

$$\nabla f_0(x)^T(-v) \geq 0 \qquad -v \in \mathcal{N}(A)$$

$$A \in \mathbb{R}^{p \times n}$$

$$\nabla f_0(x) \perp \mathcal{N}(A)$$

orthogonal

$$\Rightarrow \nabla f_0(x) \in \mathcal{R}(A^T) \qquad \nabla f_0(x) = \underline{A^T}(-v)$$

$$\nabla f_0(x) + A^T\underline{v} = 0, \text{ there exists a } v \in \mathbb{R}^p$$

$$\mathcal{N}(A)^\perp = \mathcal{R}(A^T)$$

Revise

1. column space
2. Row space
3. Null space.

$$\nabla f_0(\underline{x})^T(\underline{y}-\underline{x}) \geq 0 \quad , \quad x, y \geq 0$$

$$\nabla f = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}^{\boxed{T}} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \leq 0$$

$[2, 1, 1]$

$\underline{y} = \underline{0}$

$$\Rightarrow \nabla f_0(x)^T(-\underline{x}) \geq 0$$

OR

$$\nabla f_0(\underline{x})^T \underline{x} \leq 0 \quad \text{———} \quad (1)$$

$$\nabla f_0(\underline{x}) \geq \underline{0} \quad \text{———} \quad (2)$$

(1), (2) $\Rightarrow \nabla f_0(x)_i \, x_i = 0$

Equivalent convex problems

two problems are (informally) **equivalent** if the solution of one is readily obtained from the solution of the other, and vice-versa

some common transformations that preserve convexity:

- **eliminating equality constraints**

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & Ax = b \end{array}$$

  is equivalent to

$$\begin{array}{ll} \text{minimize (over } z\text{)} & f_0(Fz + x_0) \\ \text{subject to} & f_i(Fz + x_0) \leq 0, \quad i = 1, \dots, m \end{array}$$

  where F and x_0 are such that

$$Ax = b \quad \Longleftrightarrow \quad x = \underline{Fz + x_0} \text{ for some } z$$

$$Ax = \underbrace{AFz}_{0} + \underbrace{Ax_0}_{b}.$$

① $x_0 \rightarrow$ A particular solution to $Ax = b$ (ie, $Ax_0 = b$)

② Columns of F span the $\mathcal{N}(A) \Rightarrow AF\underline{z} = 0$

$$F\underline{z} = \begin{bmatrix} \underline{c}_1 & \underline{c}_2 & \underline{c}_3 & \cdots & \underline{c}_n \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = z_1 \underline{c}_1 + z_2 \underline{c}_2 + \ldots z_n \underline{c}_n$$

$$\in \mathcal{N}(A)$$

- **introducing equality constraints**

$$\begin{array}{ll} \text{minimize} & f_0(A_0x + b_0) \\ \text{subject to} & f_i(A_ix + b_i) \leq 0, \quad i = 1, \ldots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize (over } x,\ y_i) & f_0(y_0) \\ \text{subject to} & f_i(y_i) \leq 0, \quad i = 1, \ldots, m \\ & y_i = A_ix + b_i, \quad i = 0, 1, \ldots, m \end{array}$$

- **introducing slack variables for linear inequalities**

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize (over } x,\ s) & f_0(x) \\ \text{subject to} & a_i^T x + s_i = b_i, \quad i = 1, \ldots, m \\ & s_i \geq 0, \quad i = 1, \ldots m \end{array}$$

- **epigraph form**: standard form convex problem is equivalent to

$f_0(x)$

$$\begin{array}{ll} \text{minimize (over } x, t) & t \\ \text{subject to} & f_0(x) - t \leq 0 \quad \rightarrow \quad f_0(x) \leq t \\ & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & Ax = b \end{array}$$

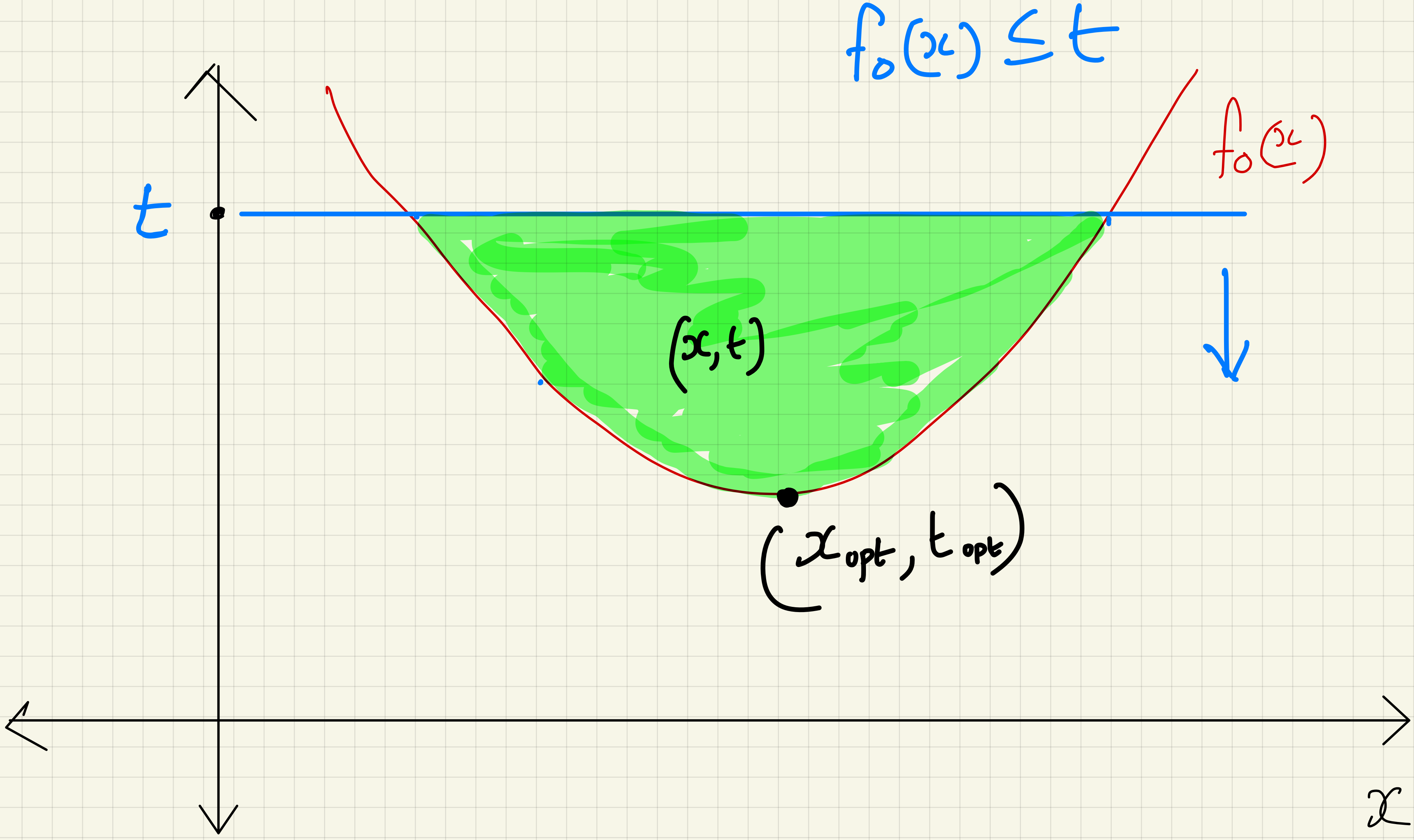
- **minimizing over some variables**

$$\begin{array}{ll} \text{minimize} & f_0(x_1, x_2) \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \dots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(x_1) \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \dots, m \end{array}$$

where $\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$



$f_0(x) \le t$
$f_0(x)$
t
(x, t)
(x_{opt}, t_{opt})
x

Recap

Standard form convex optimization

minimize $f_0(\underline{x}) \longrightarrow$ convex

subject to $f_i(\underline{x}) \le 0, \quad i = 1, 2, \dots m.$

$A\underline{x} = \underline{b}$

$h_i(x) = 0$

Abstract form convex optimization

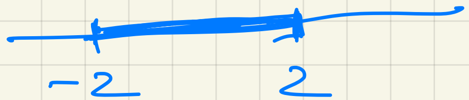
minimize convex function

over convex set

case-1 : $x^2 \le 4$

CVX $x^2 - 4 \le 0$

$\Rightarrow -2 \le x \le 2$



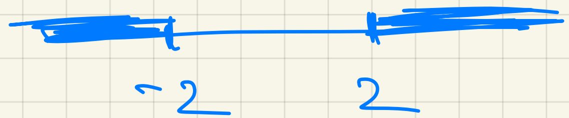
case 4 : $x^2 \ge -5$

$-x^2 + 5 \le 0$

case 2 : $x^2 \ge 4$

Not CVX $-x^2 + 4 \le 0$

$\Rightarrow$



minimize $f_0(x)$

$-x^2 - 5 \le 0$

$\Updownarrow$

minimize $f_0(x)$

case 3 : $x^2 \ge 0$

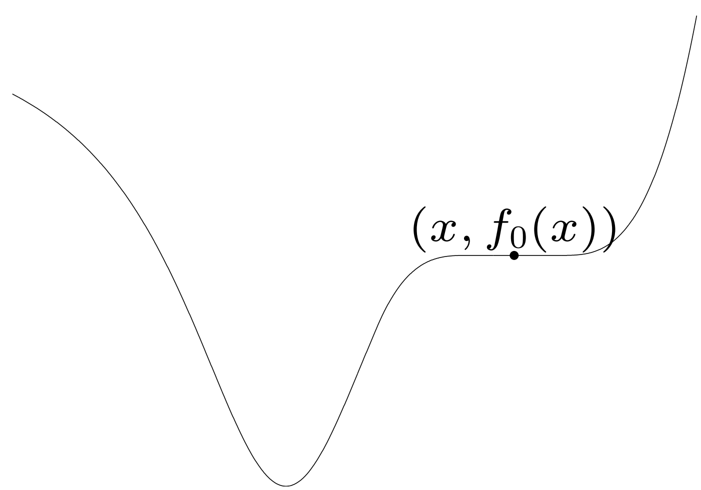
CVX $-x^2 \le 0$

Quasiconvex optimization

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

with $f_0 : \mathbf{R}^n \to \mathbf{R}$ quasiconvex, $f_1, \ldots, f_m$ convex

can have locally optimal points that are not (globally) optimal



convex representation of sublevel sets of f_0

if f_0 is quasiconvex, there exists a family of functions ϕ_t such that:

- $\phi_t(x)$ is convex in x for fixed t
- t-sublevel set of f_0 is 0-sublevel set of ϕ_t, *i.e.*, $S_t = \left\{ x \mid \frac{p(x)}{q(x)} \leq t \right\}$

$$f_0(x) \leq t \quad \Longleftrightarrow \quad \phi_t(x) \leq 0$$

$$\underbrace{p(x)}_{\text{cvx}} \underbrace{- t\, q(x)}_{\text{cvx}} \leq 0$$

concave

example

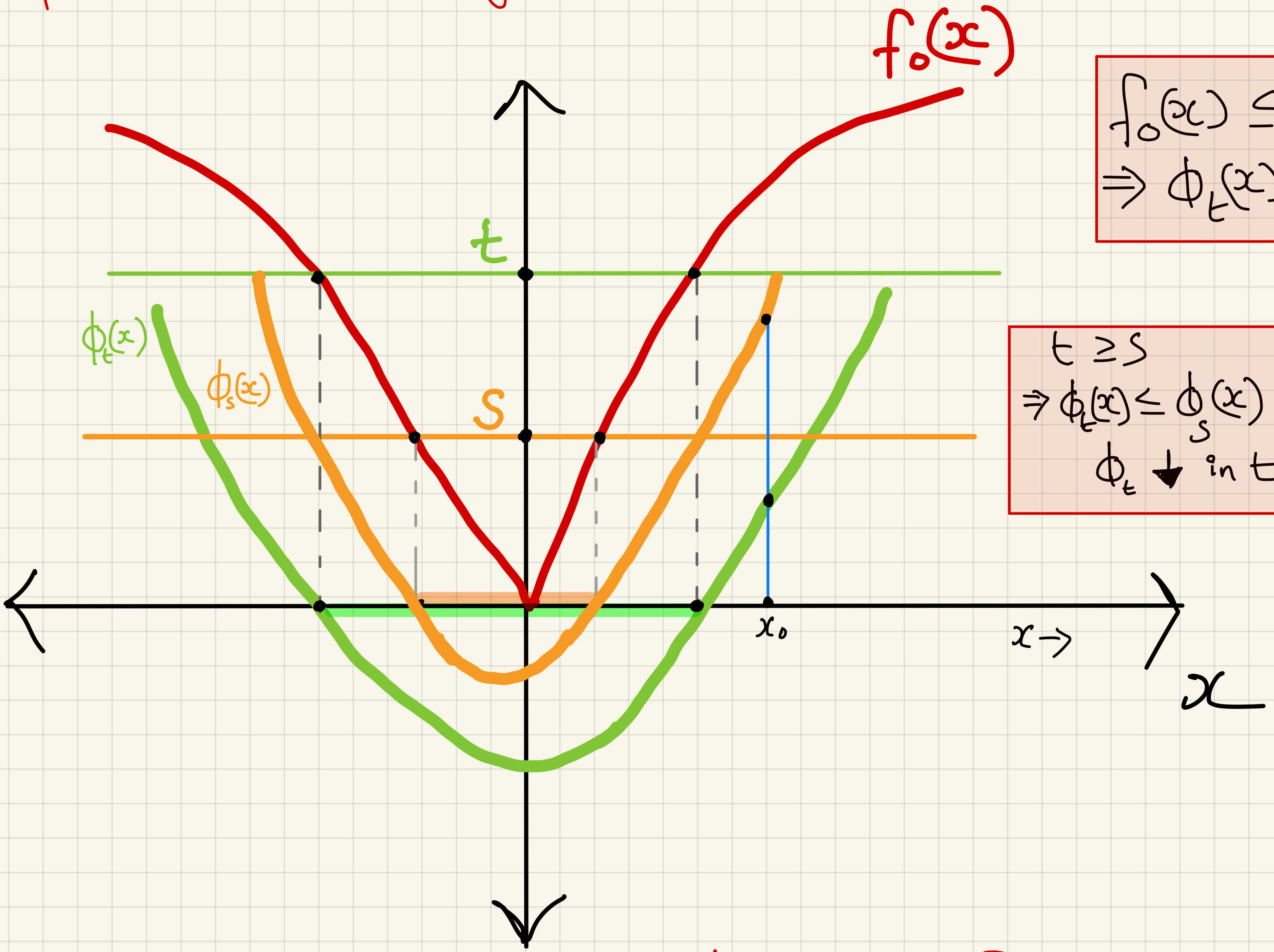
$$f_0(x) = \frac{p(x)}{q(x)} \quad \rightarrow \text{quasi cvx}$$

with p convex, q concave, and $p(x) \geq 0$, $q(x) > 0$ on $\mathbf{dom}\, f_0$

can take $\phi_t(x) = p(x) - tq(x)$:

- for $t \geq 0$, ϕ_t convex in x
- $p(x)/q(x) \leq t$ if and only if $\phi_t(x) \leq 0$

Convex representation of sublevel set of a quasi-convex function.



$$\boxed{\begin{array}{l} f_0(x) \le t \\ \Rightarrow \phi_t(x) \le 0 \end{array}}$$

$$\boxed{\begin{array}{l} t \ge s \\ \Rightarrow \phi_t(x) \le \phi_s(x) \\ \phi_t \downarrow \text{ in } t \end{array}}$$

$$\{ x \mid f_0(x) \le t \} = \{ x \mid \phi_t(x) \le 0 \}$$

Quasi-convex optimization

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \le 0, \quad i = 1, \dots, m \\
 & Ax = b
\end{array}
$$

with $f_0 : \mathbf{R}^n \to \mathbf{R}$ quasiconvex, $f_1, \dots, f_m$ convex

$f_0 \to$ Quasi-convex

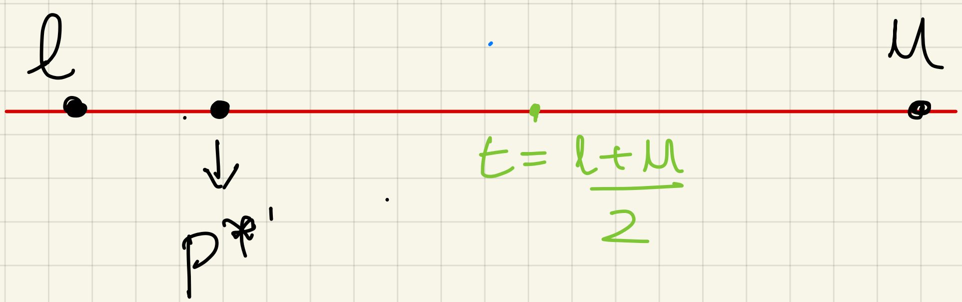
convex feasibility problem

$$
\begin{array}{ll}
\text{find} & x \\
\text{subject to} & \phi_t(x) \le 0 \\
 & f_i(x) \le 0, \quad i = 1, \dots, m \\
 & Ax = b,
\end{array}
$$

CHECK FEASIBILITY for a t

x is feasible	x is not feasible
$\downarrow$	$\downarrow$
$\phi_t(x) \le 0$	$\phi_t(x) \ge 0$
$\downarrow$	$\downarrow$
$f_0(x) \le t$	$f_0(x) \ge t$
$\downarrow$	$\downarrow$
$t \ge p^*$	$t \le p^*$

Bisection method



quasiconvex optimization via convex feasibility problems

$$\phi_t(x) \leq 0, \qquad f_i(x) \leq 0, \quad i = 1, \dots, m, \qquad Ax = b \tag{1}$$

- for fixed t, a convex feasibility problem in x
- if feasible, we can conclude that $t \geq p^\star$; if infeasible, $t \leq p^\star$

Bisection method for quasiconvex optimization

given $l \leq p^\star$, $u \geq p^\star$, tolerance $\epsilon > 0$.
repeat
1. $t := (l + u)/2$.
2. Solve the convex feasibility problem (1).
3. **if** (1) is feasible, $u := t$; **else** $l := t$.

until $u - l \leq \epsilon$.

$\frac{u-l}{2^k} \leq \epsilon$

$k \geq \log_2\left(\frac{u-l}{\epsilon}\right)$

requires exactly $\lceil \log_2((u - l)/\epsilon) \rceil$ iterations (where u, l are initial values)

length of the interval

$\hookrightarrow u-l \rightarrow \frac{u-l}{2} \rightarrow \frac{u-l}{2 \times 2} \rightarrow \cdots \rightarrow \frac{u-l}{2^k}$

Linear program (LP)

g_1

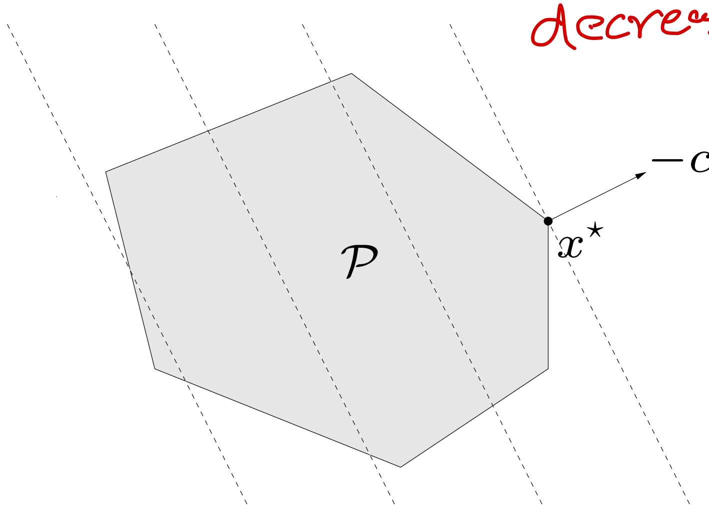
$$\begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \rightarrow \quad g_1^T x \leq h_1$$

$f_i(x) \leq 0$

$$\begin{array}{ll} \text{minimize} & c^T x + d \\ \text{subject to} & \left.\begin{array}{l} Gx \preceq h \\ Ax = b \end{array}\right\} \text{polyhedron.} \end{array}$$

- convex problem with affine objective and constraint functions
- feasible set is a polyhedron

$\longrightarrow$ decrease.

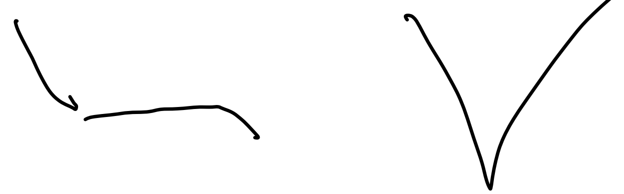


Examples

1 - apple
2 - orange

diet problem: choose quantities $x_1, \ldots, x_n$ of n foods

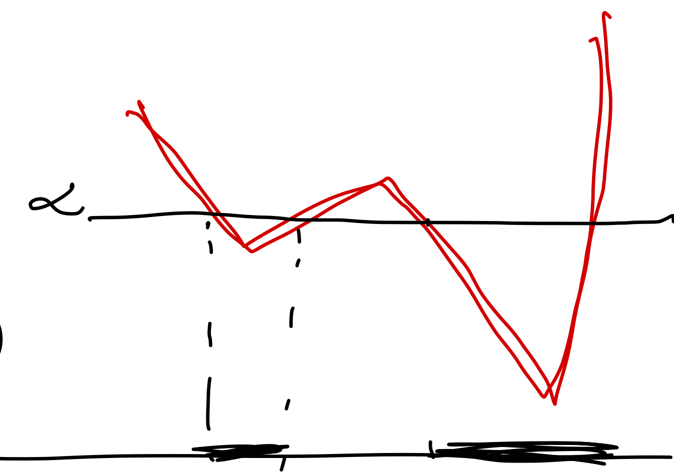
- one unit of food j costs c_j, contains amount a_{ij} of nutrient i
- healthy diet requires nutrient i in quantity at least b_i



to find cheapest healthy diet,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \succeq b, \quad x \succeq 0 \end{array}$$

piecewise-linear minimization



$$\text{minimize} \quad \max_{i=1,\ldots,m}(a_i^T x + b_i)$$

equivalent to an LP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & a_i^T x + b_i \le t, \quad i = 1, \ldots, m \end{array}$$

Diet Problem

cost $\rightarrow c_1 x_1 + c_2 x_2 + c_3 x_3 + \ldots c_n x_n = \underline{c}^T \underline{x}$

food-1 food 2. food n.

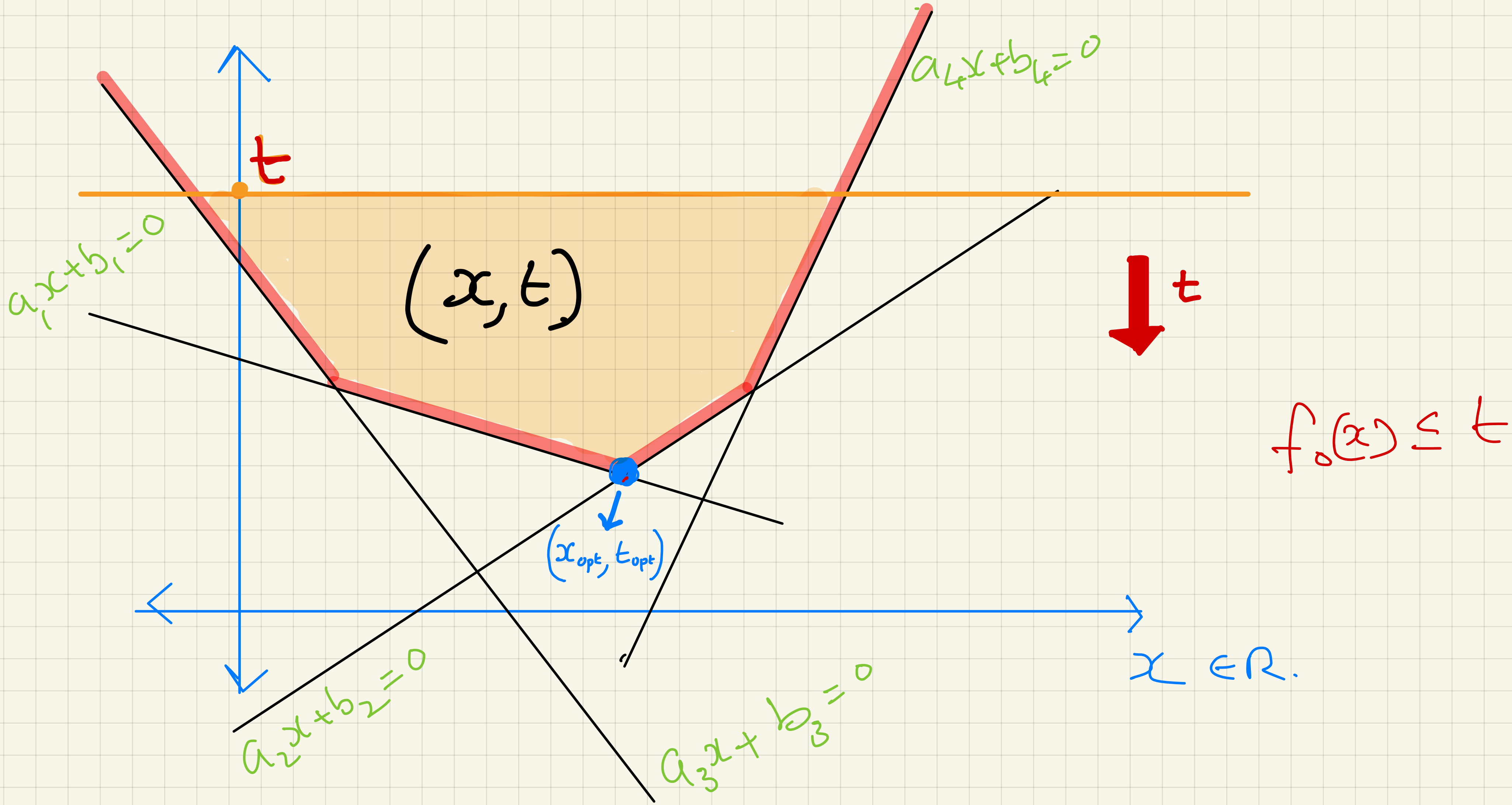
nutrient-1 $\rightarrow a_{11} x_1 + a_{12} x_2 + \ldots a_{1n} x_n \geq b_1$

nutrient-2 $\rightarrow a_{21} x_1 + a_{22} x_2 + \ldots a_{2n} x_n \geq b_2$

$\cdot \quad \cdot \quad \cdot \quad \cdot$

nutrient-n $\rightarrow a_{n1} x_1 + a_{n2} x_2 + \ldots a_{nn} x_n \geq b_n$

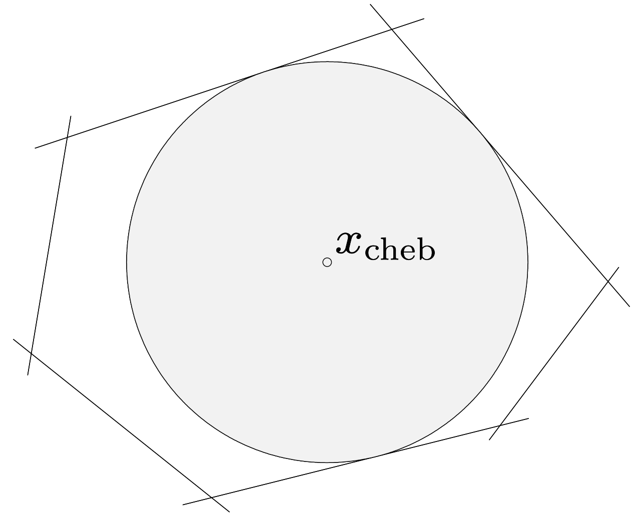
$$A \underline{x} \geq \underline{b}$$



$a_1x+b_1=0$
$a_2x+b_2=0$
$a_3x+b_3=0$
$a_4x+b_4=0$
t
(x,t)
(x_{opt}, t_{opt})
t
$f_0(x) \le t$
$x \in \mathbb{R}$.

Chebyshev center of a polyhedron

Chebyshev center of



$$\mathcal{P} = \{x \mid a_i^T x \le b_i,\ i = 1, \ldots, m\}$$

is center of largest inscribed ball

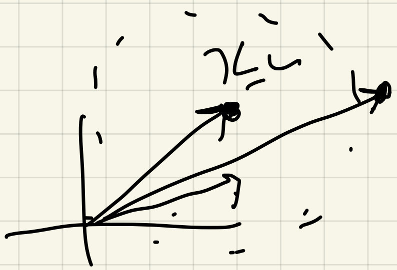
$$\mathcal{B} = \{x_c + u \mid \|u\|_2 \le r\}$$

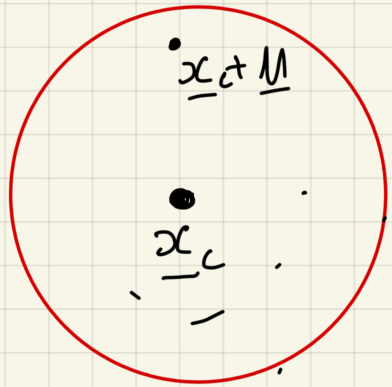
- $a_i^T x \le b_i$ for all $x \in \mathcal{B}$ if and only if

$$\sup\{a_i^T(x_c + u) \mid \|u\|_2 \le r\} = a_i^T x_c + r\|a_i\|_2 \le b_i$$

- hence, x_c, r can be determined by solving the LP

$$\begin{array}{ll} \text{maximize} & r \\ \text{subject to} & a_i^T x_c + r\|a_i\|_2 \le b_i, \quad i = 1, \ldots, m \end{array}$$





$$\sup\{ a_i^T (x_c + u) \mid \|u\|_2 \le r \}$$

$$= \sup\{ a_i^T x_c + a_i^T u \mid \|u\|_2 \le r \}$$

$$= a_i^T x_c + r\|a_i\|_2$$

$$\sup\{ a_i^T u \mid \|u\|_2 \le r \}$$

$$= \|a_i\|_2 \cdot r$$

$$a_i^T u = \|a\|_2 \|u\|_2 \cos\theta$$

Linear-fractional program

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & Gx \preceq h \\
& Ax = b
\end{array}
$$

linear-fractional program

quasi convex function

$$
f_0(x) = \frac{c^T x + d}{e^T x + f}, \qquad \textbf{dom}\, f_0(x) = \{x \mid e^T x + f > 0\}
$$

- a quasiconvex optimization problem; can be solved by bisection
- also equivalent to the LP (variables y, z)

$$
\begin{array}{ll}
\text{minimize} & c^T y + dz \\
\text{subject to} & Gy \preceq hz \\
& Ay = bz \\
& e^T y + fz = 1 \\
& z \geq 0
\end{array}
$$

LFP

$$\underset{}{\text{minimize}} \quad \frac{\underline{c}^T \underline{x} + d}{\underbrace{\underline{e}^T \underline{x} + f}_{>0}}$$

$$\text{subject to} \quad G\underline{x} \preceq \underline{h}$$

$$A\underline{x} = \underline{b}$$

$$\underset{\underline{y},\, z}{\text{minimize}} \quad \underline{c}^T \underline{y} + d\,z$$

subject to

$$G\underline{y} \preceq z\underline{h}$$

$$A\underline{y} = \underline{b}\,z$$

$$\underline{e}^T \underline{y} + f z = 1$$

$$z > 0$$

LP

$$\underline{y} = \left(\frac{1}{\underline{e}^T \underline{x} + f}\right) \underline{x} \quad \Rightarrow \quad \frac{\underline{y}}{z} = \underline{x}$$

$$z = \frac{1}{\underline{e}^T \underline{x} + f} > 0$$

$$f_0(x) = \frac{\underline{c}^T \underline{x} + d}{\underline{e}^T \underline{x} + f} = \underline{c}^T \underline{y} + d\,z$$

$$G\underline{x} \preceq \underline{h} \rightarrow G\frac{\underline{y}}{z} \preceq \underline{h} \rightarrow G\underline{y} \preceq z\underline{h}$$

$$A\underline{x} = \underline{b} \rightarrow A\frac{\underline{y}}{z} = \underline{b} \rightarrow A\underline{y} = \underline{b}\,z$$

$$z > 0$$

$$\underline{e}^T \underline{y} + f z = \frac{\underline{e}^T \underline{x}}{\underline{e}^T \underline{x} + f} + \frac{f \times 1}{\underline{e}^T \underline{x} + f} = 1$$

generalized linear-fractional program

$$f_0(x) = \max_{i=1,\ldots,r} \frac{c_i^T x + d_i}{e_i^T x + f_i}, \qquad \mathbf{dom}\, f_0(x) = \{x \mid e_i^T x + f_i > 0,\ i = 1, \ldots, r\}$$

a quasiconvex optimization problem; can be solved by bisection

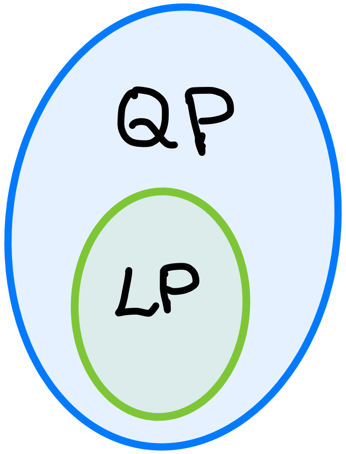
example: Von Neumann model of a growing economy

$$\begin{array}{ll} \text{maximize (over } x,\ x^+) & \min_{i=1,\ldots,n} x_i^+/x_i \\ \text{subject to} & x^+ \succeq 0, \quad Bx^+ \preceq Ax \end{array}$$

- $x, x^+ \in \mathbf{R}^n$: activity levels of n sectors, in current and next period
- $(Ax)_i$, $(Bx^+)_i$: produced, resp. consumed, amounts of good i
- x_i^+/x_i: growth rate of sector i

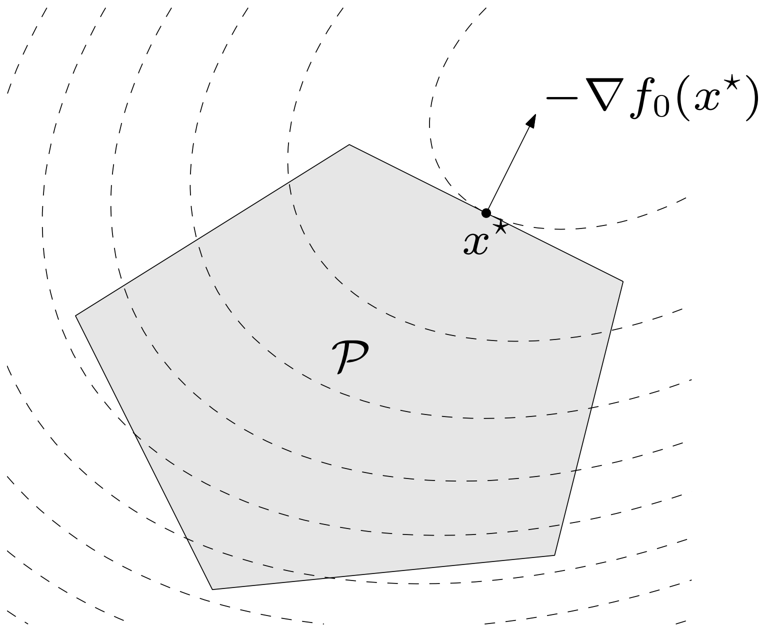
allocate activity to maximize growth rate of slowest growing sector

Quadratic program (QP)



$$\begin{array}{ll} \text{minimize} & (1/2)x^T P x + q^T x + r \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array}$$

- $P \in \mathbf{S}_+^n$, so objective is convex quadratic
- minimize a convex quadratic function over a polyhedron



Examples

least-squares

$$\text{minimize} \quad \|Ax - b\|_2^2$$

$A^\dagger = (A^TA)^{-1}A^T$

- analytical solution $x^\star = A^\dagger b$ ($A^\dagger$ is pseudo-inverse)
- can add linear constraints, *e.g.*, $l \preceq x \preceq u$

LP

linear program with random cost

c^Tx

$$\begin{array}{ll} \text{minimize} & \bar{c}^Tx + \gamma x^T\Sigma x = \mathbf{E}\, c^Tx + \gamma\, \mathbf{var}(c^Tx) \\ \text{subject to} & Gx \preceq h, \quad Ax = b \end{array}$$

- c is random vector with mean $\bar{c}$ and covariance Σ
- hence, c^Tx is random variable with mean $\bar{c}^Tx$ and variance $x^T\Sigma x$
- $\gamma > 0$ is risk aversion parameter; controls the trade-off between expected cost and variance (risk)

$c^Tx \dots$

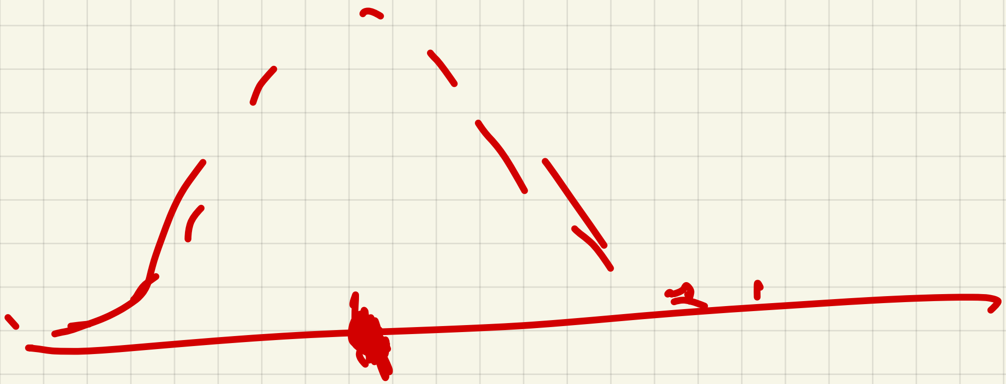
$$\Sigma = E\left[(c-\bar{c})(c-\bar{c})^T\right] = E\left[c\,c^T\right] - \bar{c}\,\bar{c}^T$$

$$\mathrm{Var}(c^T x) = E\left[(\underbrace{c^T x}_{x^T c})(\underbrace{c^T x}_{x^T c})^T\right] - (\bar{c}^T x)(\bar{c}^T x)^T$$

$$= E\left[x^T c\, c^T x\right] - x^T \bar{c}\,\bar{c}^T x$$

$$= x^T\left(\Sigma + \bar{c}\,\bar{c}^T\right)x - x^T \bar{c}\,\bar{c}^T x$$

$$= x^T \Sigma x$$

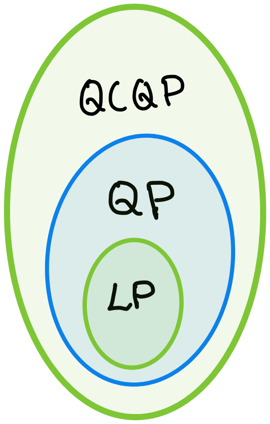


$E[c] = \bar{c}$

$$\min\ (Ec)^T x$$

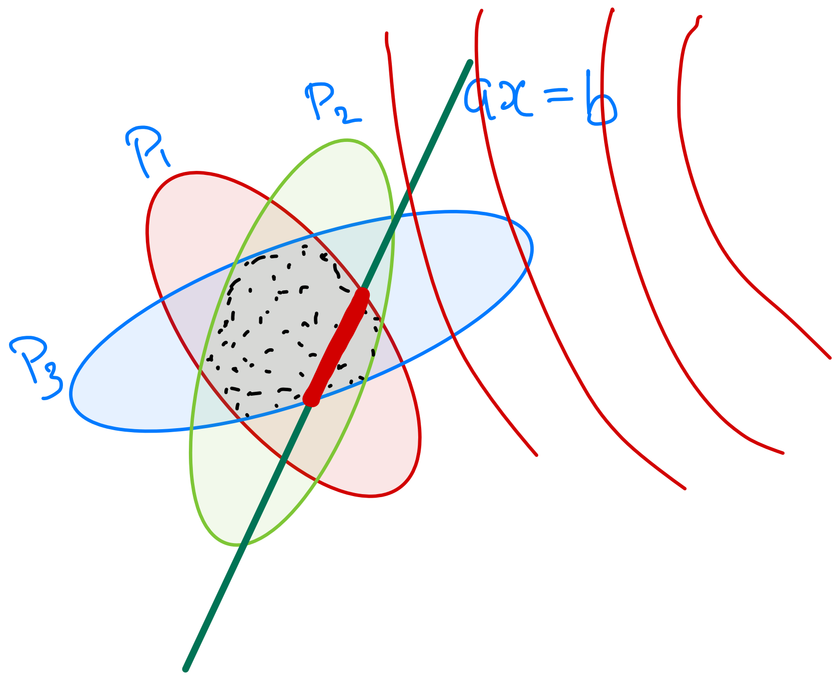
Quadratically constrained quadratic program (QCQP)

$$
\begin{array}{ll}
\text{minimize} & (1/2)x^T P_0 x + q_0^T x + r_0 \\
\text{subject to} & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
$$



- $P_i \in \mathbf{S}_+^n$; objective and constraints are convex quadratic

- if $P_1, \ldots, P_m \in \mathbf{S}_{++}^n$, feasible region is intersection of m ellipsoids and an affine set



Second-order cone programming

$$
\begin{array}{ll}
\text{minimize} & f^T x \\
\text{subject to} & \|A_i x + b_i\|_2 \le c_i^T x + d_i, \quad i = 1, \dots, m \\
 & Fx = g
\end{array}
$$

$|x| \le 2$

$-2 \le x \le 2$

$\| \underset{1\times n}{A_i} \underset{n\times 1}{x_i} \|_2 \ge |A_i x|$

($A_i \in \mathbf{R}^{n_i \times n}$, $F \in \mathbf{R}^{p \times n}$)

$1 \times n$

- inequalities are called second-order cone (SOC) constraints:

$$
(\underbrace{A_i x + b_i}_{\mathbf{R}^{np}}, \underbrace{c_i^T x + d_i}_{\mathbf{R}^1}) \in \text{second-order cone in } \mathbf{R}^{n_i+1}
$$

- for $\cancel{n_i = 0}$ ($A_i = 0$, $n_i = 1$), reduces to an LP; if $c_i = 0$, reduces to a QCQP

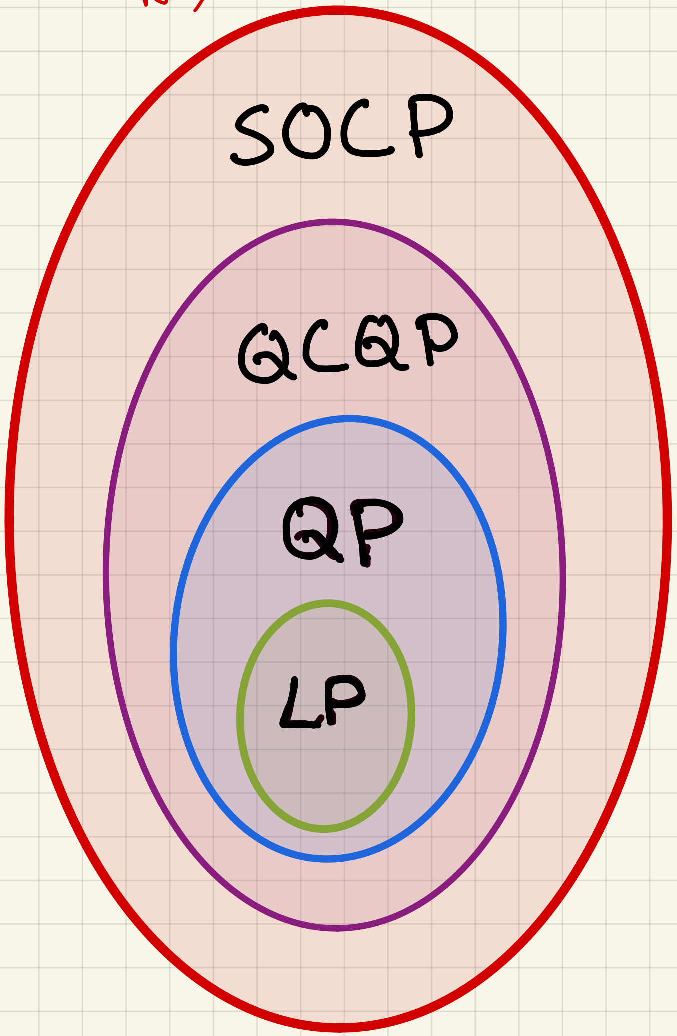
- more general than QCQP and LP

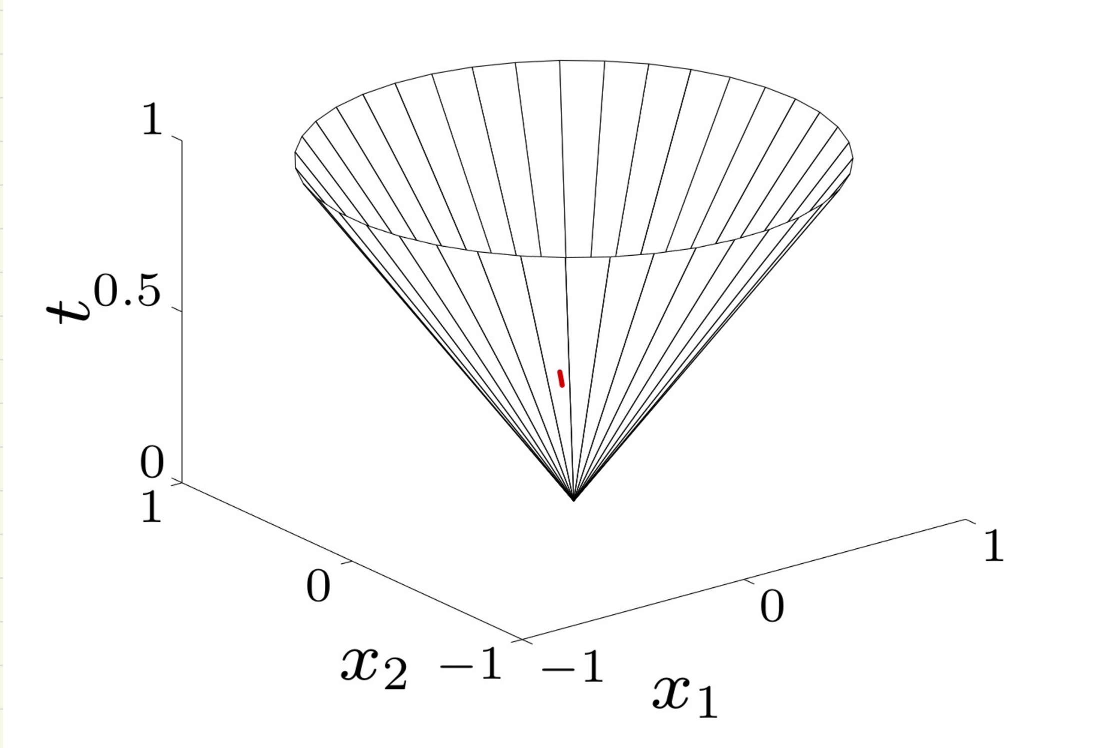
Second order Cone (Euclidean norm cone)

$(\underline{x}, t) \in SOC$

$$\{(\underline{x}, t) \mid \|x\|_2 \le t\}$$

$\mathbb{R}^n, \mathbb{R}$





$\mathbb{R}^{n+1}$

2. Problems that can be cast as SOCPs

In this section we describe some general classes of problems that can be formulated as SOCPs.

2.1. Quadratically constrained quadratic programming

We have already seen that an LP is readily expressed as an SOCP with one-dimensional cones (i.e., $n_i = 1$). Let us now consider the general *convex quadratically constrained quadratic program* (QCQP)

$$\begin{array}{ll} \text{minimize} & x^{\mathrm{T}} P_0 x + 2q_0^{\mathrm{T}} x + r_0 \\ \text{subject to} & x^{\mathrm{T}} P_i x + 2q_i^{\mathrm{T}} x + r_i \leqslant 0, \quad i = 1 \ldots p, \end{array} \tag{4}$$

where $P_0, P_1, \ldots, P_p \in \mathbb{R}^{n\times n}$ are symmetric and positive semidefinite. We will assume for simplicity that the matrices P_i are positive definite, although the problem can be reduced to an SOCP in general. This allows us to write the QCQP (4) as

$$\begin{array}{ll} \text{minimize} & \left\| P_0^{1/2} x + P_0^{-1/2} q_0 \right\|^2 + r_0 - q_0^{\mathrm{T}} P_0^{-1} q_0 \\ \text{subject to} & \left\| P_i^{1/2} x + P_i^{-1/2} q_i \right\|^2 + r_i - q_i^{\mathrm{T}} P_i^{-1} q_i \leqslant 0, \quad i = 1, \ldots, p, \end{array}$$

which can be solved via the SOCP with $p+1$ constraints of dimension $n+1$

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \| P_0^{1/2} x + P_0^{-1/2} q_0 \| \leqslant t, \\ & \| P_i^{1/2} x + P_i^{-1/2} q_i \| \leqslant \left(q_i^{\mathrm{T}} P_i^{-1} q_i - r_i \right)^{1/2}, \quad i = 1, \ldots, p, \end{array} \tag{5}$$

where $t \in \mathbb{R}$ is a new optimization variable. The optimal values of problems (4) and (5) are equal up to a constant and a square root. More precisely, the optimal value of the QCQP (4) is equal to $p^{*^2} + r_0 - q_0^{\mathrm{T}} P_0^{-1} q_0$, where p^* is the optimal value of the SOCP (5).

Robust linear programming

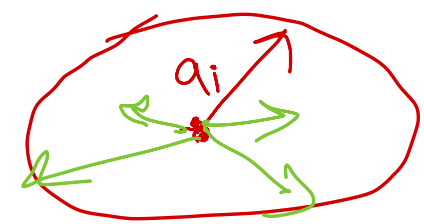
the parameters in optimization problems are often uncertain, *e.g.*, in an LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m, \end{array}$$

there can be uncertainty in c, a_i, b_i

two common approaches to handling uncertainty (in a_i, for simplicity)

- deterministic model: constraints must hold for all $a_i \in \mathcal{E}_i$



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \text{ for all } a_i \in \mathcal{E}_i, \quad i = 1, \ldots, m, \end{array}$$

- stochastic model: a_i is random variable; constraints must hold with probability η → chance constraint

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}(a_i^T x \leq b_i) \geq \eta, \quad i = 1, \ldots, m \end{array}$$

deterministic approach via SOCP

- choose an ellipsoid as $\mathcal{E}_i$:

$$\mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\} \qquad (\bar{a}_i \in \mathbf{R}^n, \quad P_i \in \mathbf{R}^{n \times n})$$

center is $\bar{a}_i$, semi-axes determined by singular values/vectors of P_i

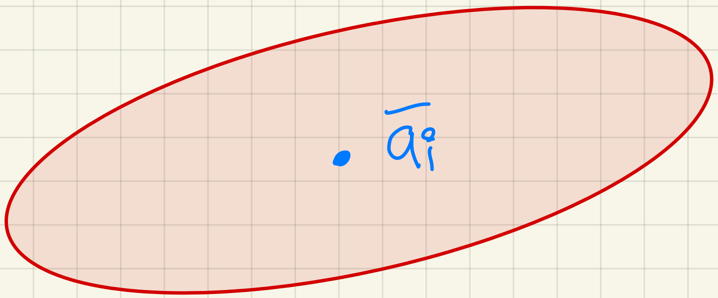
- robust LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \quad \forall a_i \in \mathcal{E}_i, \quad i = 1, \ldots, m \end{array}$$

is equivalent to the SOCP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \|P_i^T x\|_2 \leq b_i, \quad i = 1, \ldots, m \end{array}$$

(follows from $\sup_{\|u\|_2 \leq 1} (\bar{a}_i + P_i u)^T x = \bar{a}_i^T x + \|P_i^T x\|_2$)



$$\mathcal{E}_i = \{ \underline{a}_i + P_i \underline{u} \mid \|\underline{u}\|_2 \le 1 \}$$

robust LP: $\underline{a}_i^T \underline{x} \le b_i$

$$\sup_{\|u\|_2 \le 1} \left\{ (\bar{a}_i + P_i u)^T x \right\} = \bar{a}_i^T x + \sup_{\|u\|_2 \le 1} \left\{ u^T P_i^T x \right\}$$

$$= \bar{a}_i^T x + \|P_i^T x\|_2 .$$

stochastic approach via SOCP

- assume a_i is Gaussian with mean $\bar{a}_i$, covariance Σ_i ($a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i)$)
- $a_i^T x$ is Gaussian r.v. with mean $\bar{a}_i^T x$, variance $x^T \Sigma_i x$; hence

$$\mathbf{prob}(a_i^T x \le b_i) = \Phi\left(\frac{b_i - \bar{a}_i^T x}{\|\Sigma_i^{1/2} x\|_2}\right)$$

$$\frac{x - \mu}{\sigma}$$

where $\Phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^{x} e^{-t^2/2}\, dt$ is CDF of $\mathcal{N}(0,1)$

- robust LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}(a_i^T x \le b_i) \ge \eta, \quad i = 1, \ldots, m, \end{array}$$

with $\eta \ge 1/2$, is equivalent to the SOCP

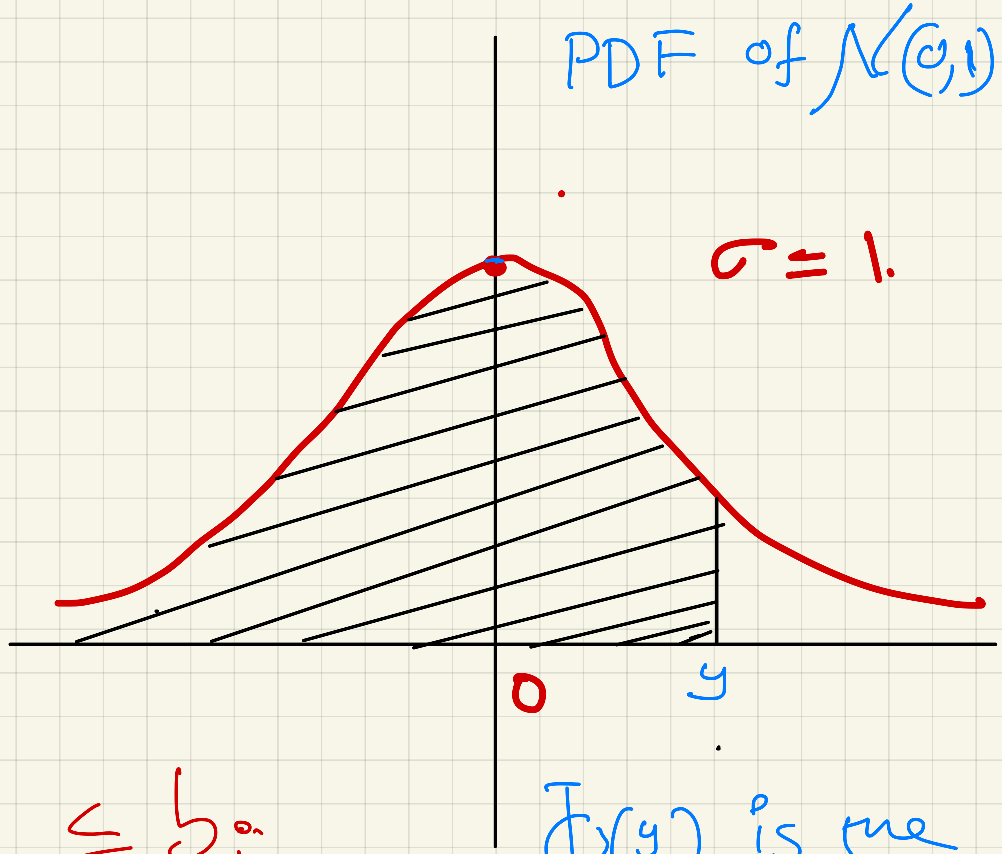
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \Phi^{-1}(\eta) \|\Sigma_i^{1/2} x\|_2 \le b_i, \quad i = 1, \ldots, m \end{array}$$

$$\text{prob}\left(a_i^T x \le b_i\right) \ge \eta$$

$$\Phi\left(\frac{b_i - \bar{a}_i^T x}{\|\Sigma_i^{1/2} x\|_2}\right) \ge \eta$$

$$\frac{b_i - a_i^T x}{\|\Sigma_i^{1/2} x\|_2} \ge \Phi^{-1}(\eta)$$

$$\Rightarrow \quad a_i^T x + \Phi^{-1}(\eta)\,\|\Sigma_i^{1/2} x\|_2 \le b_i$$



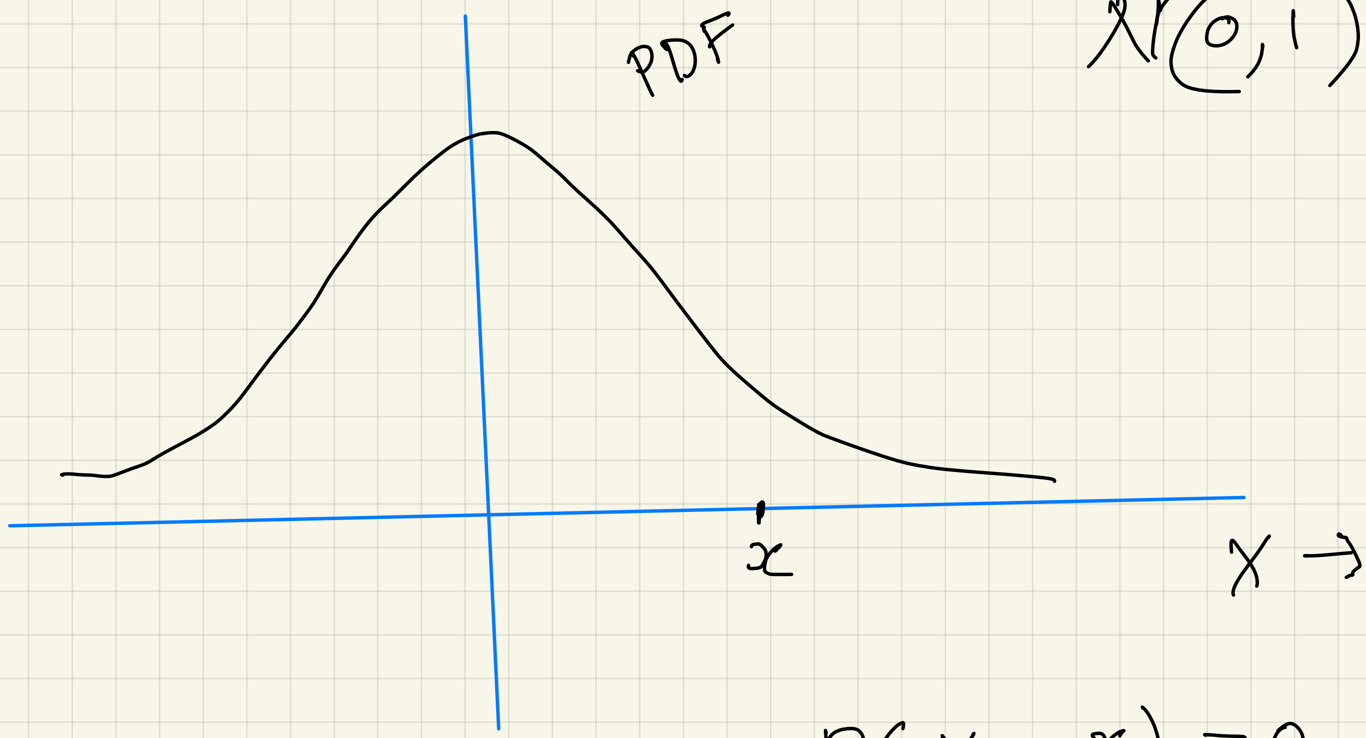
$\Phi(y)$ is the area of the shaded region.

$$\text{prob}\left(\frac{u - \overset{\text{mean}}{\bar{u}}}{\underset{\text{std}}{\sigma}} \le \frac{b_i - \bar{u}}{\sigma}\right) = \Phi\left(\frac{b_i - \bar{u}}{\sigma}\right)$$

$$u = \bar{a}_i^T x$$

$$\sigma^2 = x^T \Sigma_i x = x^T (\Sigma_i^{1/2})^T \Sigma_i^{1/2} x = (\Sigma_i^{1/2} x)^T (\Sigma_i^{1/2} x) = \|\Sigma_i^{1/2} x\|_2^2$$

$\mathcal{N}(0,1)$



$P(X = x) = 0$

CDF → Φ

$P(X \leq y) = \Phi(y)$

$\Phi(y)$

Geometric programming

monomial function

$$f(x) = cx_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}, \qquad \mathbf{dom}\, f = \mathbf{R}_{++}^n$$

$x_1^{1/2}$

x_2^{-10}

with $c > 0$; exponent a_i can be any real number

posynomial function: sum of monomials

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}, \qquad \mathbf{dom}\, f = \mathbf{R}_{++}^n$$

geometric program (GP)

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 1, \quad i = 1, \ldots, m \\ & h_i(x) = 1, \quad i = 1, \ldots, p \end{array}$$

with f_i posynomial, h_i monomial

Geometric program in convex form

change variables to $y_i = \log x_i$, and take logarithm of cost, constraints

- monomial $f(x) = c x_1^{a_1} \cdots x_n^{a_n}$ transforms to

$$\log f(e^{y_1}, \ldots, e^{y_n}) = a^T y + b \qquad (b = \log c)$$

- posynomial $f(x) = \sum_{k=1}^K c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}$ transforms to

$$\log f(e^{y_1}, \ldots, e^{y_n}) = \log \left(\sum_{k=1}^K e^{a_k^T y + b_k} \right) \qquad (b_k = \log c_k)$$

- geometric program transforms to convex problem

$$\begin{array}{ll} \text{minimize} & \log \left(\sum_{k=1}^K \exp(a_{0k}^T y + b_{0k}) \right) \\ \text{subject to} & \log \left(\sum_{k=1}^K \exp(a_{ik}^T y + b_{ik}) \right) \leq 0, \quad i = 1, \ldots, m \\ & Gy + d = 0 \end{array}$$

$$f(\underline{x}) = c\, x_1^{a_1} x_2^{a_2} \cdots\cdots x_n^{a_n}$$

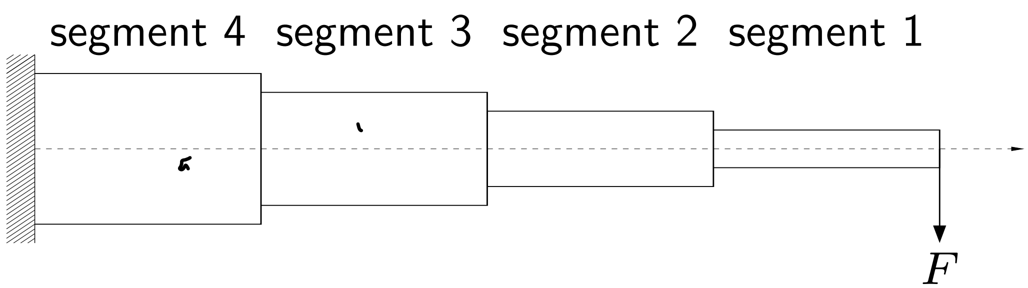
$$\log f(\underline{x}) = \log c + a_1 \log x_1 + a_2 \log x_2 + \cdots\cdots a_n \log x_n$$

$$= \underbrace{\log c}_{b} + \underbrace{\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}}_{\underline{a}^T} \begin{bmatrix} \log x_1 \\ \log x_2 \\ \vdots \\ \log x_n \end{bmatrix} \rightarrow \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \rightarrow \underline{y}$$

$$= b + \underline{a}^T \underline{y}$$

$$\log\left(\sum_{k=1}^{K} \underbrace{e^{a_k^{(1)} \log x_1} \cdot e^{a_k^{(2)} \log x_2} \cdots e^{a_k^{(n)} \log x_n} \cdot e^{\log c_k}}_{x_1^{a_k^{(1)}}\, x_2^{a_k^{(2)}} \cdots x_n^{a_k^{(n)}}\, c_k} \right)$$

Design of cantilever beam



- N segments with unit lengths, rectangular cross-sections of size $w_i \times h_i$
- given vertical force F applied at the right end

design problem

$$
\begin{array}{ll}
\text{minimize} & \text{total weight} \\
\text{subject to} & \text{upper \& lower bounds on } w_i,\ h_i \\
& \text{upper bound \& lower bounds on aspect ratios } h_i/w_i \\
& \text{upper bound on stress in each segment} \\
& \text{upper bound on vertical deflection at the end of the beam}
\end{array}
$$

variables: w_i, h_i for $i = 1, \ldots, N$

objective and constraint functions

- total weight $w_1 h_1 + \cdots + w_N h_N$ is posynomial
- aspect ratio h_i/w_i and inverse aspect ratio w_i/h_i are monomials
- maximum stress in segment i is given by $6iF/(w_i h_i^2)$, a monomial
- the vertical deflection y_i and slope v_i of central axis at the right end of segment i are defined recursively as

$$\begin{aligned} v_i &= 12(i-1/2)\frac{F}{E w_i h_i^3} + v_{i+1} \\ y_i &= 6(i-1/3)\frac{F}{E w_i h_i^3} + v_{i+1} + y_{i+1} \end{aligned}$$

for $i = N, N-1, \ldots, 1$, with $v_{N+1} = y_{N+1} = 0$ (E is Young's modulus)

v_i and y_i are posynomial functions of w, h

formulation as a GP

$$
\begin{array}{ll}
\text{minimize} & w_1 h_1 + \cdots + w_N h_N \\
\text{subject to} & w_{\max}^{-1} w_i \leq 1, \quad w_{\min} w_i^{-1} \leq 1, \quad i = 1, \ldots, N \\
& h_{\max}^{-1} h_i \leq 1, \quad h_{\min} h_i^{-1} \leq 1, \quad i = 1, \ldots, N \\
& S_{\max}^{-1} w_i^{-1} h_i \leq 1, \quad S_{\min} w_i h_i^{-1} \leq 1, \quad i = 1, \ldots, N \\
& 6iF\sigma_{\max}^{-1} w_i^{-1} h_i^{-2} \leq 1, \quad i = 1, \ldots, N \\
& y_{\max}^{-1} y_1 \leq 1
\end{array}
$$

note

- we write $w_{\min} \leq w_i \leq w_{\max}$ and $h_{\min} \leq h_i \leq h_{\max}$

$$
w_{\min}/w_i \leq 1, \qquad w_i/w_{\max} \leq 1, \qquad h_{\min}/h_i \leq 1, \qquad h_i/h_{\max} \leq 1
$$

- we write $S_{\min} \leq h_i/w_i \leq S_{\max}$ as

$$
S_{\min} w_i/h_i \leq 1, \qquad h_i/(w_i S_{\max}) \leq 1
$$

Minimizing spectral radius of nonnegative matrix

Perron-Frobenius eigenvalue $\lambda_{\rm pf}(A)$

- exists for (elementwise) positive $A \in \mathbf{R}^{n\times n}$
- a real, positive eigenvalue of A, equal to spectral radius $\max_i |\lambda_i(A)|$
- determines asymptotic growth (decay) rate of A^k: $A^k \sim \lambda_{\rm pf}^k$ as $k \to \infty$
- alternative characterization: $\lambda_{\rm pf}(A) = \inf\{\lambda \mid Av \preceq \lambda v \text{ for some } v \succ 0\}$

minimizing spectral radius of matrix of posynomials

- minimize $\lambda_{\rm pf}(A(x))$, where the elements $A(x)_{ij}$ are posynomials of x
- equivalent geometric program:

$$\begin{array}{ll} \text{minimize} & \lambda \\ \text{subject to} & \sum_{j=1}^n A(x)_{ij} v_j/(\lambda v_i) \le 1, \quad i = 1, \ldots, n \end{array}$$

variables λ, v, x

skip

$\mathbb{R}$ $\quad 2 < 3$

$\mathbb{R}^n$ vector $\quad x \preceq_{\mathbb{R}^n_+} y$

matrix. $X \preceq_{S^n_+} Y$

Generalized inequality constraints

convex problem with generalized inequality constraints

Can often be solved as easily as ordinary convex problem

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \preceq_{K_i} 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
$$

$f(\theta x + (1-\theta) y) \preceq_K \theta f(x) + (1-\theta) f(y)$

- $f_0 : \mathbf{R}^n \to \mathbf{R}$ convex; $f_i : \mathbf{R}^n \to \mathbf{R}^{k_i}$ K_i-convex w.r.t. proper cone K_i
- same properties as standard convex problem (convex feasible set, local optimum is global, etc.)

conic form problem: special case with affine objective and constraints

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Fx + g \preceq_K 0 \\
& Ax = b
\end{array}
$$

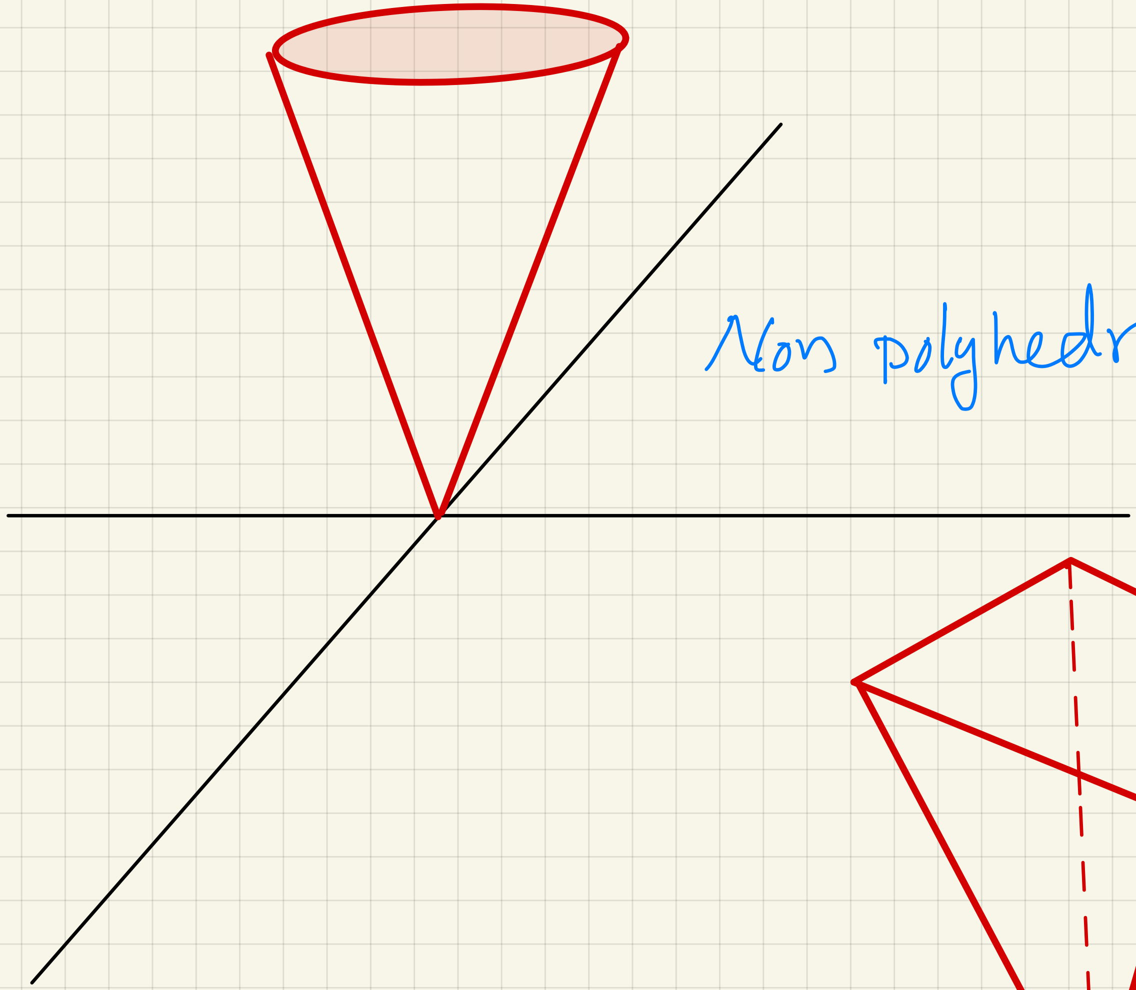
$(Fx+g)_i \leq 0 \quad \forall i$

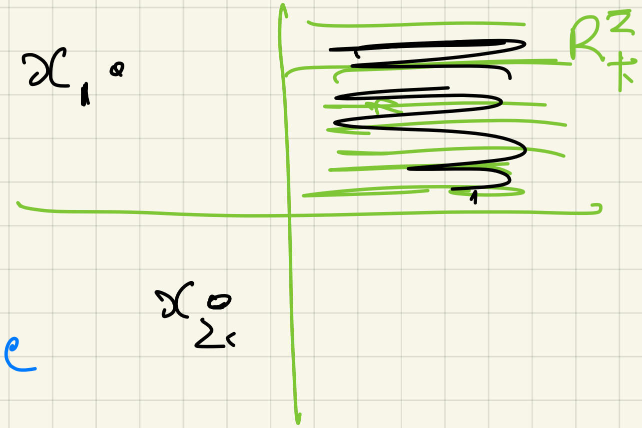
$y \preceq_{K = \mathbb{R}^n_+} 0$

extends linear programming ($K = \mathbf{R}^m_+$) to nonpolyhedral cones

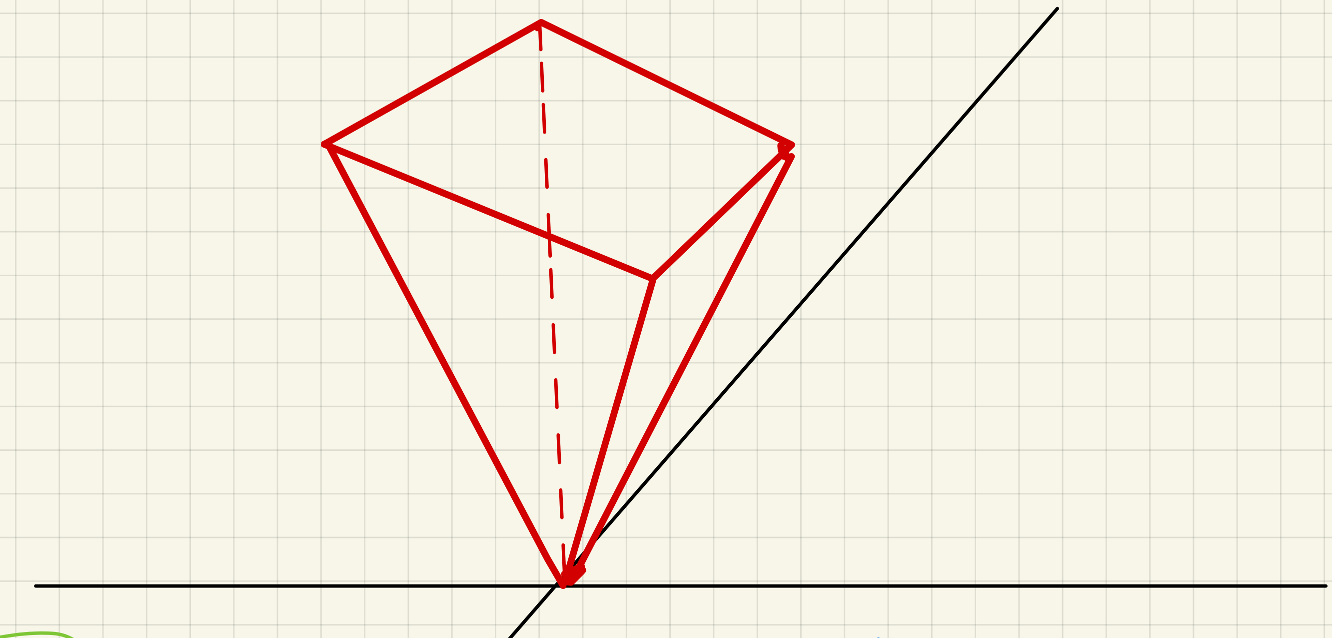
$y_1 \leq 0$

$y_2 \leq 0$





Non polyhedral cone



polyhedral cone

$$\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \preceq_{K = R^m_+} 0$$

$$\begin{bmatrix} f_{11} & f_{12} & f_{13} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \leq 0$$

RECAP

Generalized Inequality (w.r.t. a proper cone K)

2 commonly used generalized Inequalities

$K = \mathbb{R}^n_+$ | $K = S^n_+$

vectors.

Point wise inequality	Matrix inequality (PSD Inequality)
$K = \mathbb{R}^n_+$	$K = S^n_+$
$\underline{x} \preceq_{\mathbb{R}^n_+} \underline{y} \iff \underline{y} - \underline{x} \in \mathbb{R}^n_+$	$X \succeq_{S^n_+} Y \Rightarrow X - Y \in S^n_+$
$\Rightarrow y_i \geq x_i \ \forall i$	$\Rightarrow X - Y$ is PSD
$\underline{x} \preceq_{\mathbb{R}^n_+} \underline{y}$	$X \succeq_{S_n^+} Y$

Semidefinite program (SDP) $K = S_+^n$

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \\
& Ax = b
\end{array}
$$

$Fx + g \preceq_K 0$

$\underbrace{}_{A(x)} \preceq_{S_+} 0$

$A(x)$ matrix.

with F_i, $G \in \mathbf{S}^k$

- inequality constraint is called linear matrix inequality (LMI)
- includes problems with multiple LMI constraints: for example,

$$x_1 \hat{F}_1 + \cdots + x_n \hat{F}_n + \hat{G} \preceq 0, \qquad x_1 \tilde{F}_1 + \cdots + x_n \tilde{F}_n + \tilde{G} \preceq 0$$

is equivalent to single LMI

$$x_1 \begin{bmatrix} \hat{F}_1 & 0 \\ 0 & \tilde{F}_1 \end{bmatrix} + x_2 \begin{bmatrix} \hat{F}_2 & 0 \\ 0 & \tilde{F}_2 \end{bmatrix} + \cdots + x_n \begin{bmatrix} \hat{F}_n & 0 \\ 0 & \tilde{F}_n \end{bmatrix} + \begin{bmatrix} \hat{G} & 0 \\ 0 & \tilde{G} \end{bmatrix} \preceq 0$$

Block diagonal matrix is PSD iff the diagonal blocks are PSD

$$A(x) = \begin{bmatrix} x_1 + x_2 & x_2 + 1 \\ x_2 + 1 & x_3 \end{bmatrix} \preceq_{S^n_+} 0 \qquad \begin{bmatrix} (x_1 \hat{F}_1 + x_2 \hat{F}_2 + \dots) & 0 \\ 0 & (x_1 \tilde{F}_1 + \dots) \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x_1 + \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} x_2 + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} x_3 + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \preceq_{S^n_+} 0$$

$$\overbrace{\begin{bmatrix} \underbrace{x_1 \quad x_2 \quad x_3}_{y_1^T} & \underbrace{x_4 \quad x_5}_{y_2^T} \end{bmatrix}}^{x^T} \begin{bmatrix} F_1 & 0 \\ 0 & F_2 \end{bmatrix} \overbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}}^{\underline{x}} \leq 0 \qquad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underline{y_1} \qquad \begin{bmatrix} x_3 \\ x_4 \\ x_5 \end{bmatrix} = \underline{y_2}$$

$$\begin{bmatrix} \underline{y}_1^T & \underline{y}_2^T \end{bmatrix} \begin{bmatrix} F_1 & 0 \\ 0 & F_2 \end{bmatrix} \begin{bmatrix} y_1 \\ \underline{y_2} \end{bmatrix} = \underline{y}_1^T F_1 \underline{y_1} + \underline{y}_2^T F_2 \underline{y_2} \leq 0$$

LP and SOCP as SDP

LP and equivalent SDP

LP: minimize $c^T x$
subject to $Ax \underset{\mathbf{R}^n_+}{\preceq} b$

SDP: minimize $c^T x$
subject to $\mathbf{diag}(Ax - b) \underset{\mathbf{S}^n_+}{\preceq} 0$

(note different interpretation of generalized inequality $\preceq$)

SOCP and equivalent SDP

SOCP: minimize $f^T x$
subject to $\|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m$

x^*

SDP: minimize $f^T x$
subject to

$$\begin{bmatrix} \overset{P}{(c_i^T x + d_i)I} & \overset{Q}{A_i x + b_i} \\ \underset{Q^T}{(A_i x + b_i)^T} & \underset{R}{c_i^T x + d_i} \end{bmatrix} \succeq 0, \quad i = 1, \ldots, m$$

$$\operatorname{diag}(A\underline{x}-\underline{b}) = \begin{bmatrix} (Ax-b)_1 & & & 0 \\ & (Ax-b)_2 & & \\ & & \ddots & \\ 0 & & & (Ax-b)_m \end{bmatrix} \preceq_{S^n_+} 0$$

$$\Rightarrow (Ax-b)_i \leq 0$$

SOCP

$$\left.\begin{array}{l} \text{minimize } \underline{f}^T\underline{x} \\ \text{subject to } \|A\underline{x}+b_i\|_2 \leq \underline{c}_i^T\underline{x}+d_i \\ \qquad\qquad i=1,\dots m \end{array}\right\} \Rightarrow \begin{array}{l} \text{minimize } \underline{f}^T\underline{x} \\ \text{subject to} \\ -(A\underline{x}+\underline{b}_i,\ \underline{c}_i^T\underline{x}+d_i) \preceq_{K_i} 0 \end{array}$$

$$K_i = \{(y,t)\in\mathbb{R}^{n+1} \mid \|y\|_2 \leq t\}$$

2nd order cone.

$$\{(x,t) \mid \|x\|_2 \leq t\}$$

$(x,t) \in SOC$

Schur complement

Consider a matrix $X \in \mathbf{S}^n$ partitioned as

$$X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix},$$

where $A \in \mathbf{S}^k$. If $\det A \neq 0$, the matrix

$$S = C - B^T A^{-1} B$$

is called the *Schur complement* of A in X. Schur complements arise in several contexts, and appear in many important formulas and theorems. For example, we have

$$\det X = \det A \det S.$$

$X = \begin{bmatrix} P & Q \\ Q^T & R \end{bmatrix}$ is PSD iff

① P is PSD

② Schur complement of P in X is PSD.

① $\Rightarrow \quad P \succeq 0 \quad \Rightarrow \quad (\underline{c}_i^T \underline{x} + d_i) I \succeq 0$

$\Rightarrow \quad \underline{c}_i^T \underline{x} + d_i \geq 0$

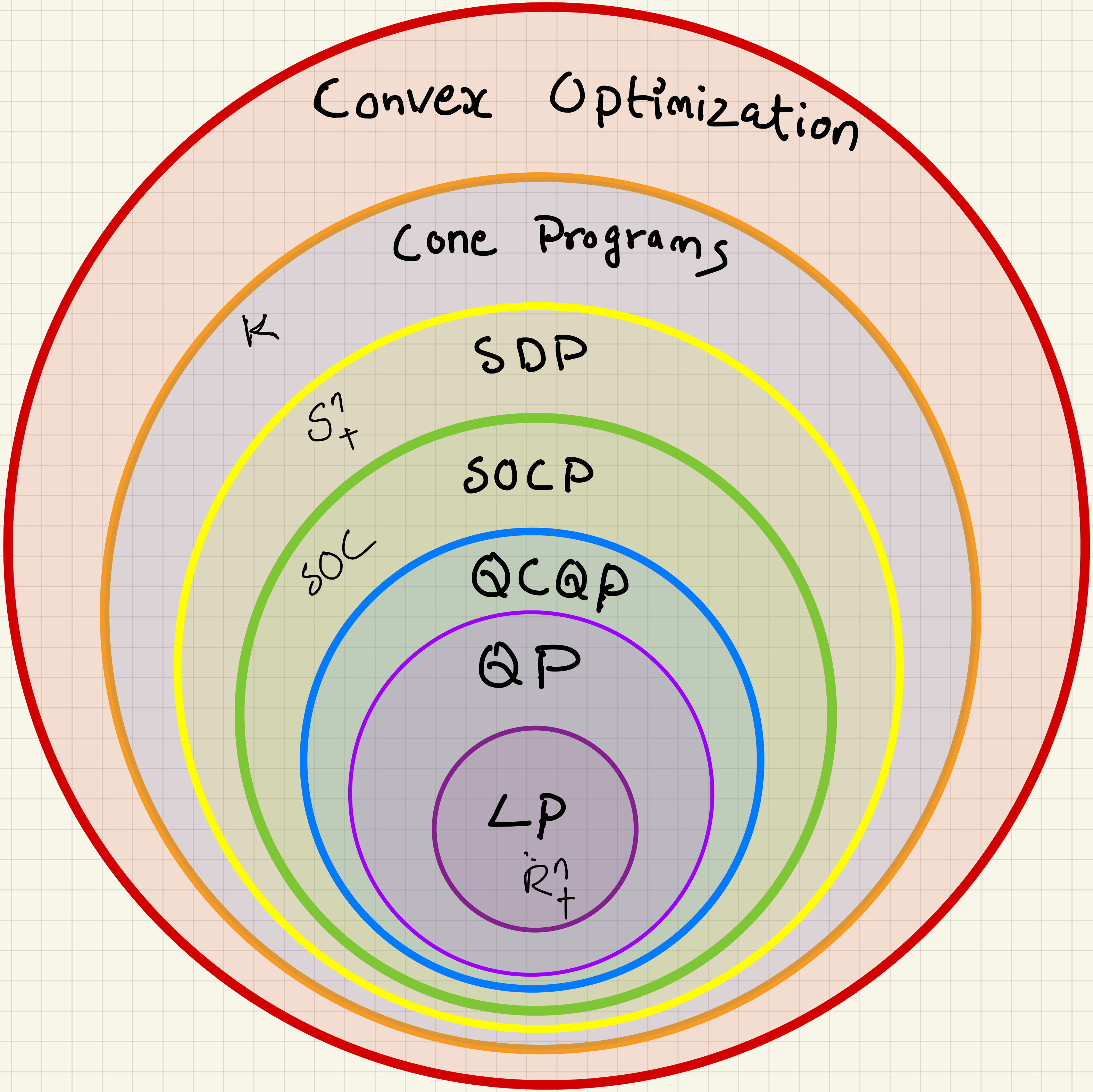
② $\Rightarrow \quad R - Q^T P^{-1} B \succeq 0$

$$\Rightarrow \quad \underline{c}_i^T \underline{x} + d_i - (A_i \underline{x} + \underline{b}_i)^T (\underline{c}_i^T \underline{x} + d_i)^{-1} I^{-1} (A_i \underline{x} + \underline{b}_i) \geq 0$$

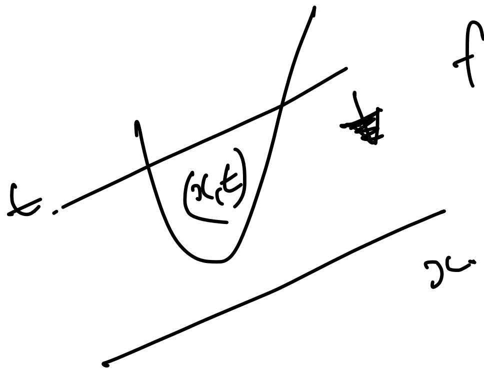
$$\Rightarrow \quad \| A_i \underline{x} + \underline{b}_i \|_2^2 \leq (\underline{c}_i^T \underline{x} + d_i)^2$$

$$\Rightarrow \quad \Rightarrow \quad \| A_i \underline{x} + \underline{b}_i \|_2 \leq \underline{c}_i^T \underline{x} + d_i$$

$\rightarrow$ SOCP inequality.



Convex Optimization
Cone Programs
K
SDP
S^n_+
SOCP
SOC
QCQP
QP
LP
$\mathbb{R}^n_+$



$f(x) \leq t.$

Eigenvalue minimization

Epigraph form

minimize t

subject to

$\lambda_{max}(A(x)) \leq t$

$$\text{minimize} \quad \lambda_{\max}(A(x))$$

where $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ (with given $A_i \in \mathbf{S}^k$)

equivalent SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & A(x) \preceq tI \end{array}$$

- variables $x \in \mathbf{R}^n$, $t \in \mathbf{R}$
- follows from

$$\lambda_{\max}(A) \leq t \quad \Longleftrightarrow \quad A \preceq tI$$

$Ax = \lambda x$

$Ax - tx = \lambda x - tx$

$Ax - tIx = (\lambda - t)x$

$(A - tI)x = (\lambda - t)x$

$A \preceq tI$

$\underbrace{(A - tI)} \preceq 0$

All eigen values are ≤ 0

$\Rightarrow \lambda_{max}(A) - t \leq 0$

$\lambda_{max}(A) \leq t$

Eigen value of $A = \lambda_i$, $i = 1, 2, \cdots k$.

Eigen value of $A - tI = \lambda_i - t$, $i = 1, 2, \cdots k$

Matrix norm minimization

- Revise: EVD
SVD

$$\text{minimize} \quad \|A(x)\|_2 = \left(\lambda_{\max}(A(x)^T A(x))\right)^{1/2}$$

where $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ (with given $A_i \in \mathbf{R}^{p \times q}$)

equivalent SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \left[\begin{array}{cc} tI & A(x) \\ A(x)^T & tI \end{array}\right] \succeq 0 \end{array}$$

- variables $x \in \mathbf{R}^n$, $t \in \mathbf{R}$
- constraint follows from

$$\begin{aligned} \|A\|_2 \leq t \quad &\Longleftrightarrow \quad A^T A \preceq t^2 I, \quad t \geq 0 \\ &\Longleftrightarrow \quad \left[\begin{array}{cc} tI & A \\ A^T & tI \end{array}\right] \succeq 0 \end{aligned}$$

minimize t

subject to $\|A(\underline{x})\|_2 \le t$

$$\left(\lambda_{max}\left(A(\underline{x})^T A(\underline{x})\right)\right)^{1/2} \le t$$

$$\lambda_{max}\left(A(\underline{x})^T A(\underline{x})\right) \le t^2$$

$$A(x)^T A(x) - t^2 I \preceq 0 \quad \text{— (1)}$$

Also $t \ge 0$ — (2)

$\Rightarrow$ minimize t

subject to

$$\begin{bmatrix} tI & A(x) \\ A(x)^T & tI \end{bmatrix} \succeq 0$$

SDP

(1), (2) $\Rightarrow$

$$\begin{bmatrix} tI & A(x) \\ A(x)^T & tI \end{bmatrix} \succeq 0$$

$$tI \succeq 0 \Rightarrow t \ge 0$$

$$tI - A^T(x)(tI)^{-1}A^T(x) \succeq 0$$

Vector optimization

general vector optimization problem

$f : R^n \to R^q$

$$\begin{array}{lll} \text{minimize (w.r.t. } K\text{)} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \ldots, m \\ & h_i(x) = 0, & i = 1, \ldots, p \end{array}$$

vector objective $f_0 : \mathbf{R}^n \to \mathbf{R}^q$, minimized w.r.t. proper cone $K \in \mathbf{R}^q$

convex vector optimization problem

$$\begin{array}{lll} \text{minimize (w.r.t. } K\text{)} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \ldots, m \\ & Ax = b & \end{array}$$

with f_0 K-convex, f_1, . . . , f_m convex

Issue with $f_0 : \mathbb{R}^n \to \mathbb{R}^q$

What we have seen so far $\Rightarrow$ $\boxed{f_0 : \mathbb{R}^n \to \mathbb{R}}$

$2 \leq 3 \leq 10$

$f_0 : \mathbb{R}^n \to \mathbb{R}$. $\rightarrow$ linear ordering exists.

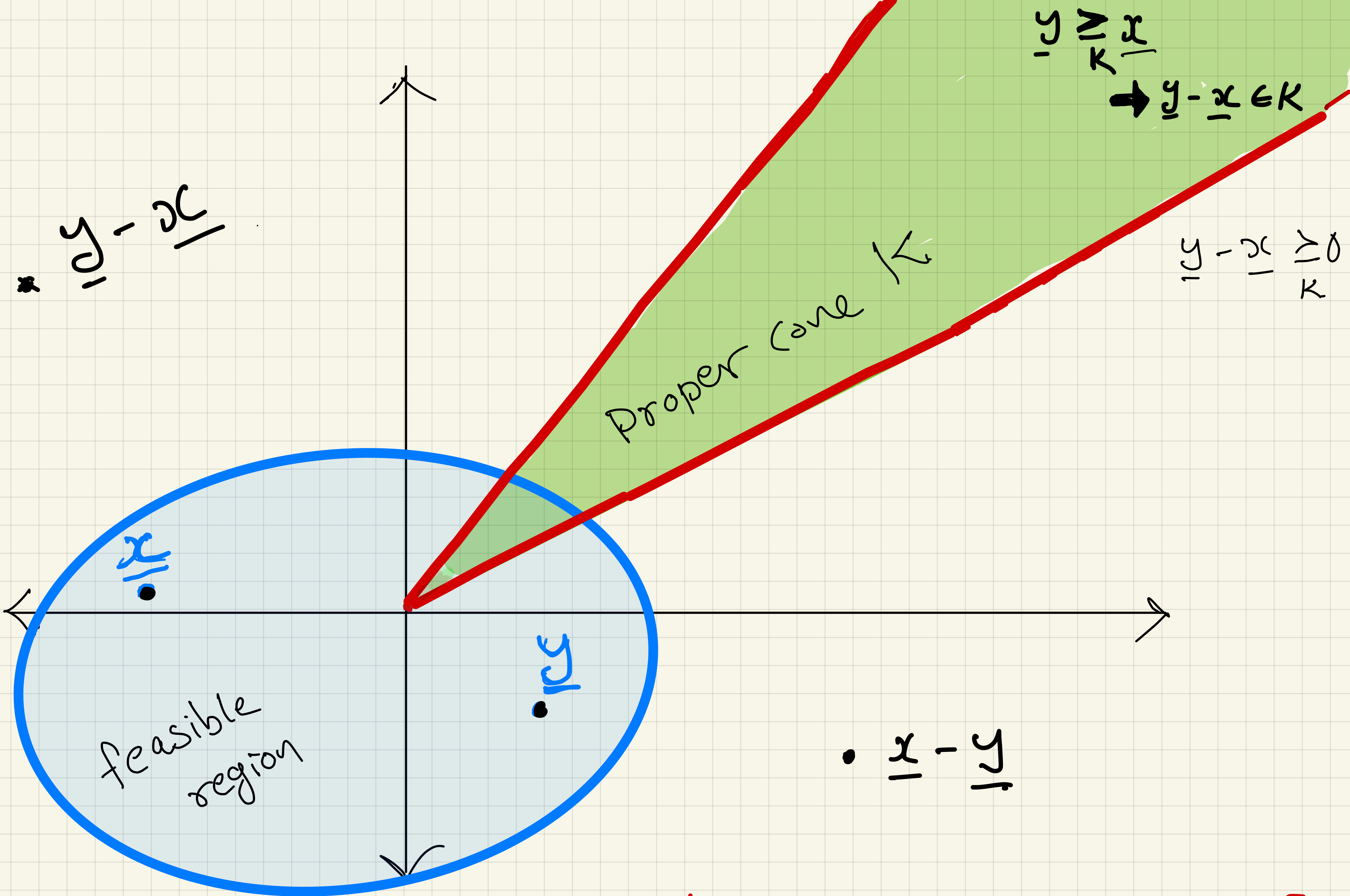
$$f_0(\underline{x}_1) \leq f_0(\underline{x}_2) \leq f_0(\underline{x}_3)$$

$\underline{x}_i$'s are feasible

But for convex vector optimization $\boxed{f_0 : \mathbb{R}^n \to \mathbb{R}^q}$

minimization is w.r.t. a proper cone K

$\Rightarrow$ Generalized inequalities $\Rightarrow$ Not a linear ordering



It can happen that

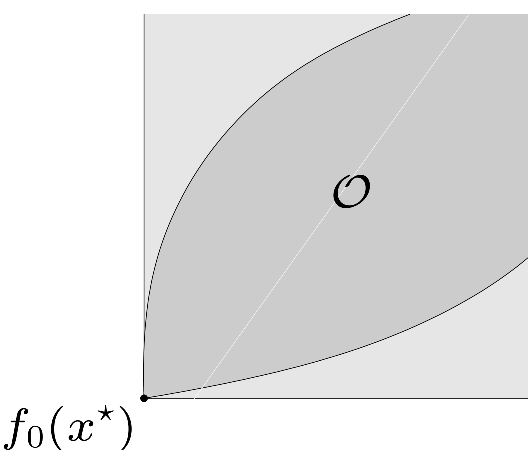
$$\left.\begin{aligned} \underline{y} &\preceq_K \underline{x} \\ \underline{x} &\preceq_K \underline{y} \end{aligned}\right\} \rightarrow \text{Both FALSE}$$

Optimal and Pareto optimal points

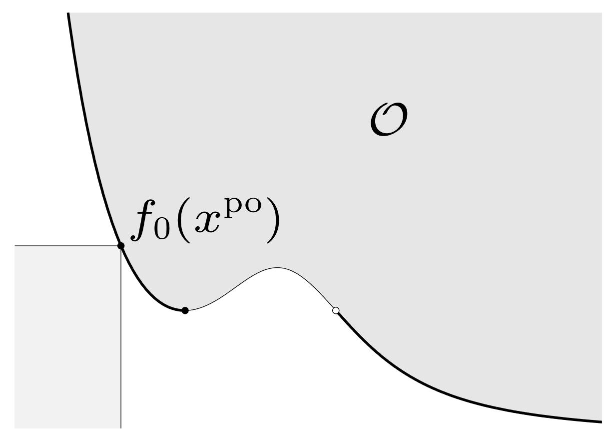
set of achievable objective values

$$\mathcal{O} = \{f_0(x) \mid x \text{ feasible}\}$$

- feasible x is **optimal** if $f_0(x)$ is the minimum value of $\mathcal{O}$
- feasible x is **Pareto optimal** if $f_0(x)$ is a minimal value of $\mathcal{O}$

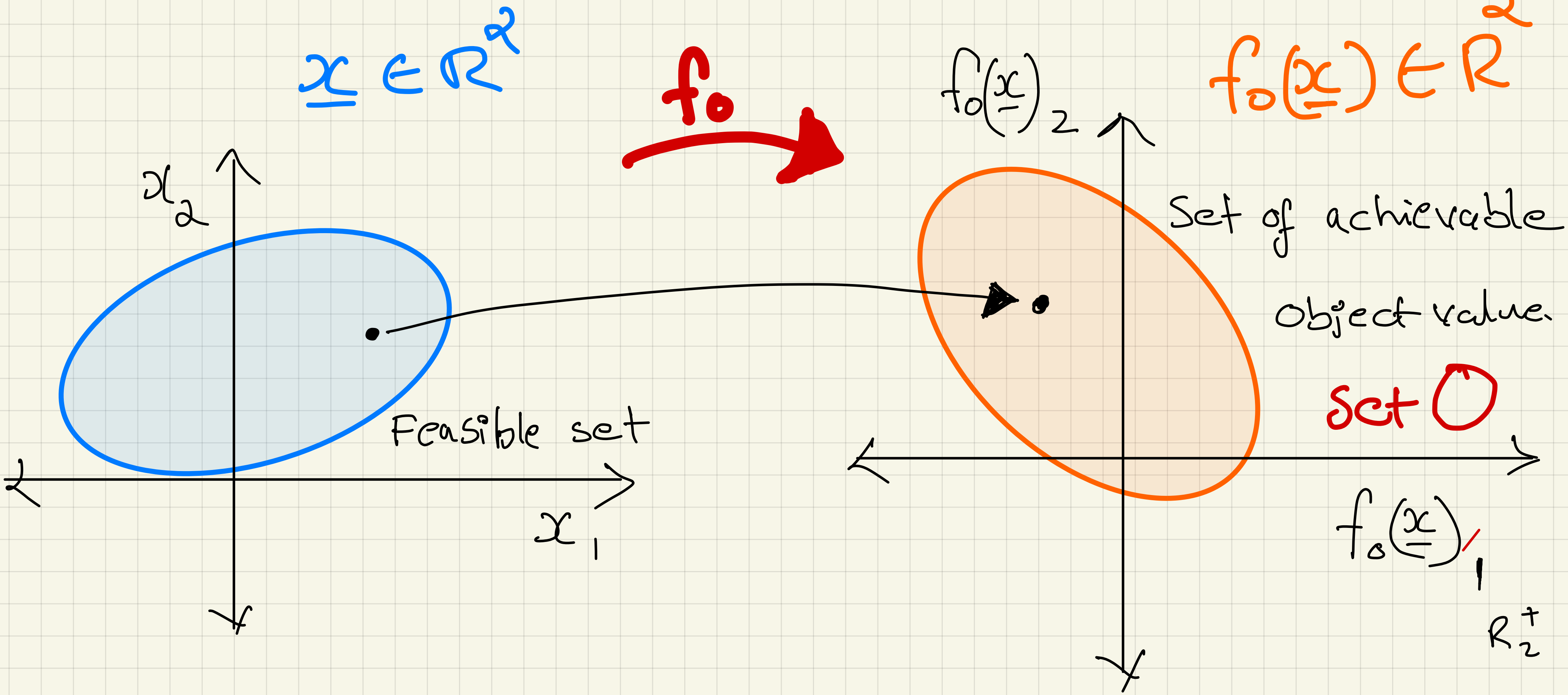


$x^\star$ is optimal



x^{po} is Pareto optimal

$f_0 : R^2 \to R^2$



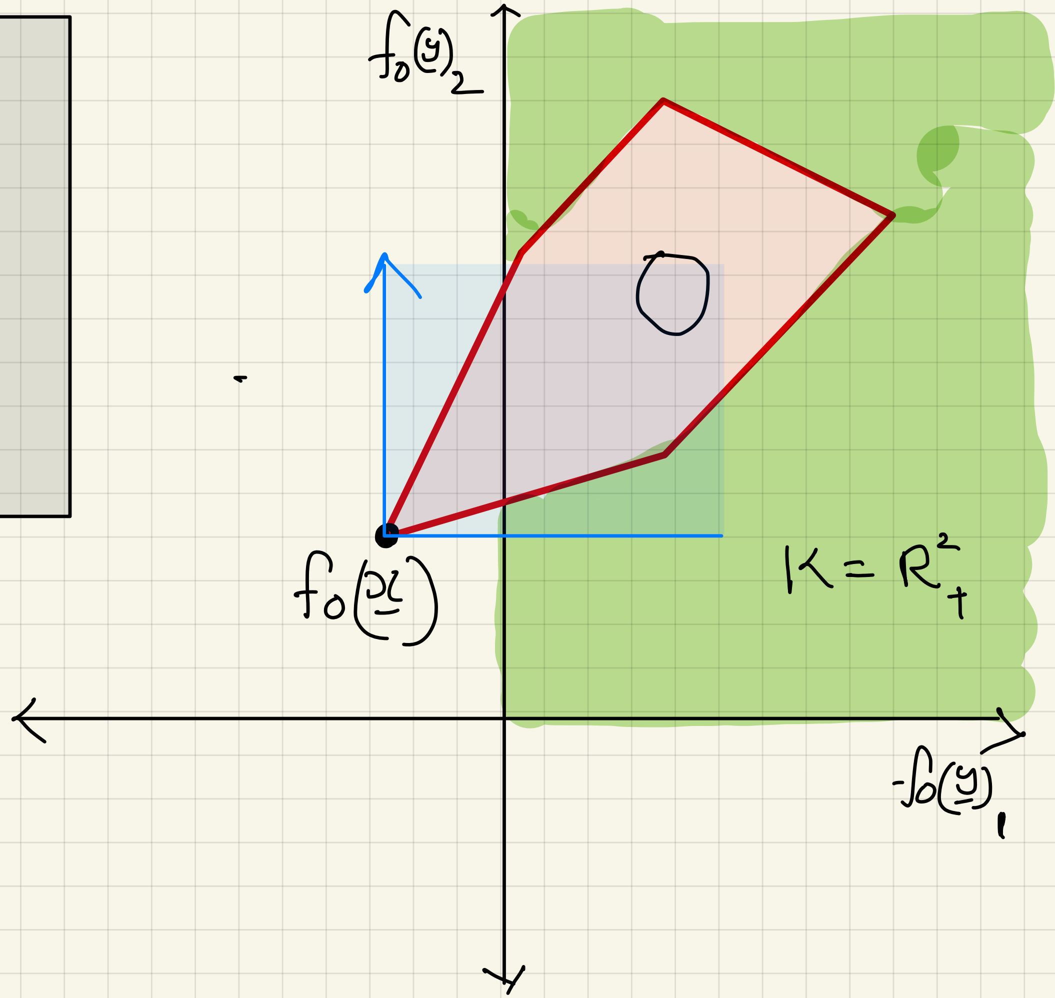
$$f_0(\underline{x}) = \begin{bmatrix} f_0(\underline{x})_1 \\ f_0(\underline{x})_2 \end{bmatrix}$$

Optimality of $f_0(\underline{x})$ is defined using the minimum and minimal values of the set O

$\underline{x}$ (feasible) is **optimum** if $f_0(\underline{x})$ is **the minimum** value of the set $\mathcal{O}$.

minimum: $f_0(\underline{x})$ is the minimum of $\mathcal{O}$ w.r.t. $\preceq_K$ iff

$$\boxed{f_0(\underline{x}) \preceq_K f_0(\underline{y}) \quad \forall f_0(\underline{y}) \in \mathcal{O} \qquad \text{OR} \qquad \mathcal{O} \subseteq f_0(\underline{x}) + K}$$



minimum is unique.

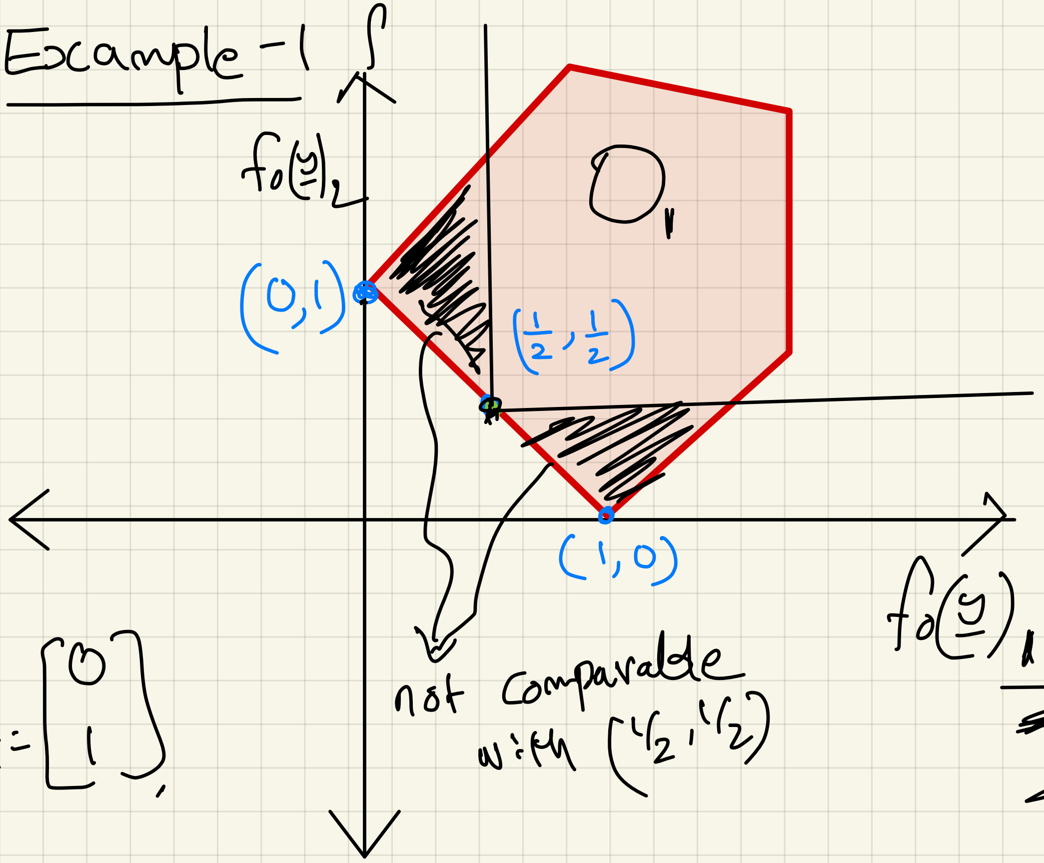
>

$\underline{x}$ (feasible) is **Pareto optimum** if $f_0(\underline{x})$ is **a minimal** value of the set $\mathcal{O}$

minimal : $f_0(\underline{x})$ is a minimal of $\mathcal{O}$ w.r.t. $\preceq_K$ iff

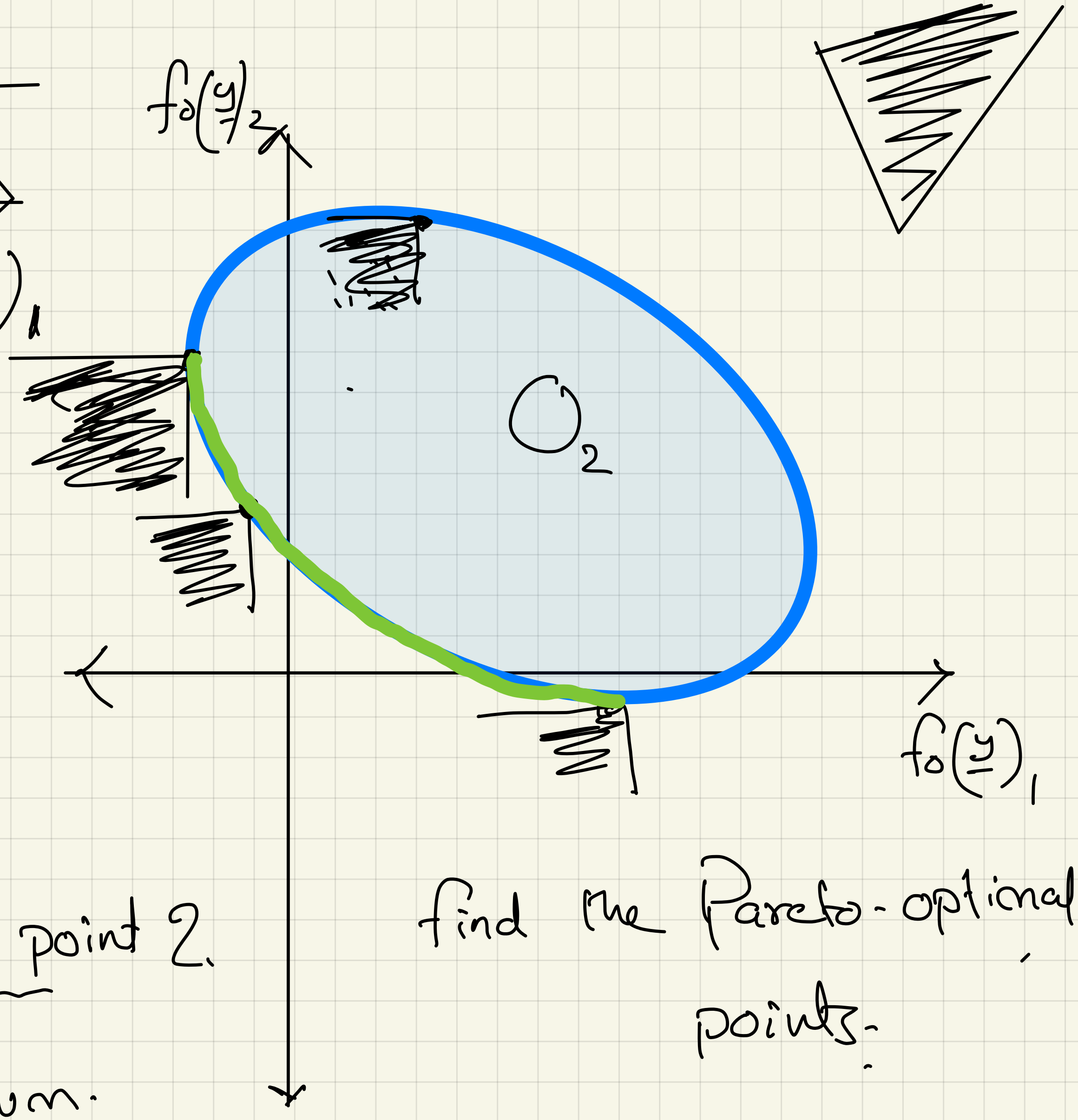
$$\boxed{\begin{array}{c} \text{for any } f_0(\underline{y}) \in \mathcal{O} \\ f_0(\underline{y}) \preceq_K f_0(\underline{x}) \Rightarrow f_0(\underline{y}) = f_0(\underline{x}) \\ \text{OR} \\ \left(f_0(\underline{x}) - K\right) \cap \mathcal{O} = \{f_0(\underline{x})\} \end{array}}$$

Example-1



$K = \mathbb{R}^2_+$

Example-2



$\underline{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \underline{y} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$

$\underline{x} - \underline{y} = \begin{bmatrix} -1/2 \\ +1/2 \end{bmatrix} \qquad \underline{x} \not\succeq_K \underline{y}$

$\underline{y} - \underline{x} = \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix} \qquad \underline{x} \not\preceq_K \underline{y}$

① Does $\mathcal{O}_2$ has an optimal point ?

$\downarrow$ minimum.

No

find the Pareto-optimal points.

Optimal point $\longrightarrow$ the minimum.

Pareto optimal. $\longrightarrow$ a minimal.

Multicriterion optimization

vector optimization problem with $K = \mathbf{R}_+^q$

- return.
risk.

$$f_0(x) = (F_1(x), \ldots, F_q(x))$$

- q different objectives F_i; roughly speaking we want all F_i's to be small
- feasible $x^\star$ is optimal if

$$y \text{ feasible} \quad \Longrightarrow \quad f_0(x^\star) \preceq f_0(y)$$

minimum.

  if there exists an optimal point, the objectives are noncompeting
- feasible x^{po} is Pareto optimal if

$$y \text{ feasible}, \quad f_0(y) \preceq f_0(x^{\mathrm{po}}) \quad \Longrightarrow \quad f_0(x^{\mathrm{po}}) = f_0(y)$$

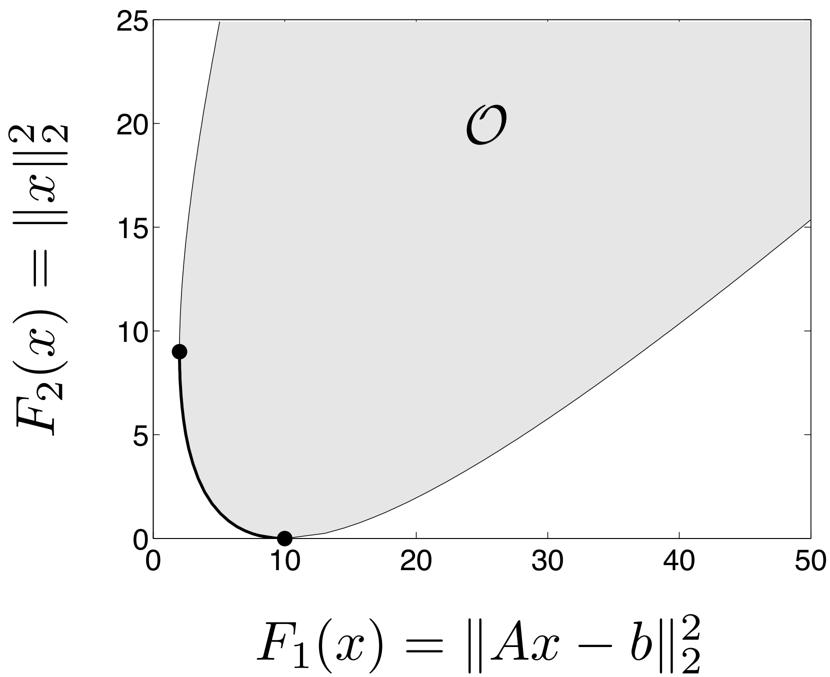
minimal.

  if there are multiple Pareto optimal values, there is a trade-off between the objectives

Regularized least-squares

$$\text{minimize (w.r.t. } \mathbf{R}^2_+) \quad (\|Ax-b\|_2^2, \|x\|_2^2)$$

$(\|Ax-b\|_2^2, \|x\|_1)$



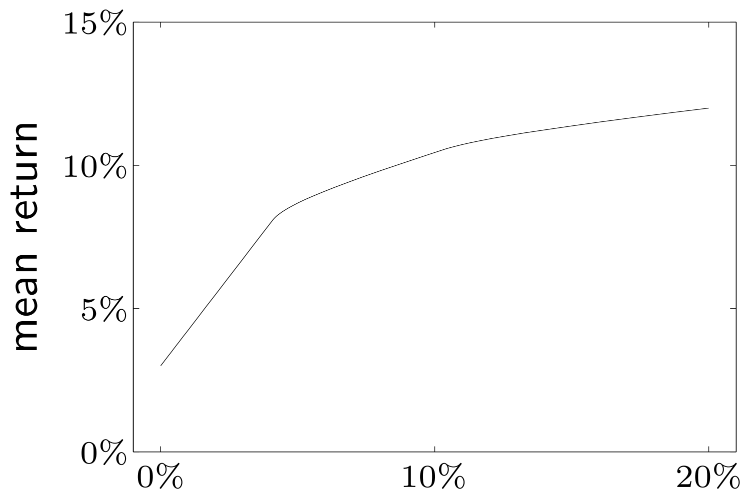
example for $A \in \mathbf{R}^{100\times 10}$; heavy line is formed by Pareto optimal points

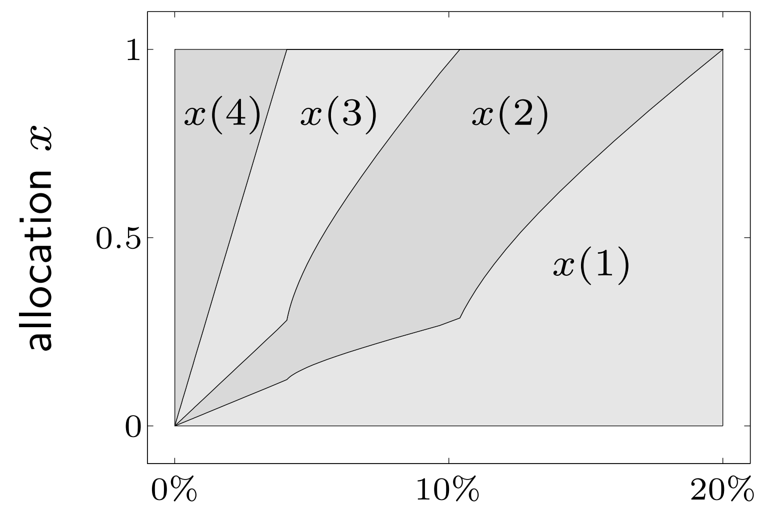
Risk return trade-off in portfolio optimization

$$\begin{array}{ll} \text{minimize (w.r.t. } \mathbf{R}^2_+) & (-\bar{p}^T x, x^T \Sigma x) \\ \text{subject to} & \mathbf{1}^T x = 1, \quad x \succeq 0 \end{array}$$

- $x \in \mathbf{R}^n$ is investment portfolio; x_i is fraction invested in asset i
- $p \in \mathbf{R}^n$ is vector of relative asset price changes; modeled as a random variable with mean $\bar{p}$, covariance Σ
- $\bar{p}^T x = \mathbf{E}\, r$ is expected return; $x^T \Sigma x = \mathbf{var}\, r$ is return variance

example





Scalarization

to find Pareto optimal points: choose $\lambda \succ_{K^*} 0$ and solve scalar problem

$f_0(x) = [f_{0,1}(x), f_{0,2}(x) \ldots]$

$$
\begin{array}{ll}
\text{minimize} & \lambda^T f_0(x) \\
\text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\
 & h_i(x) = 0, \quad i = 1, \ldots, p
\end{array}
$$

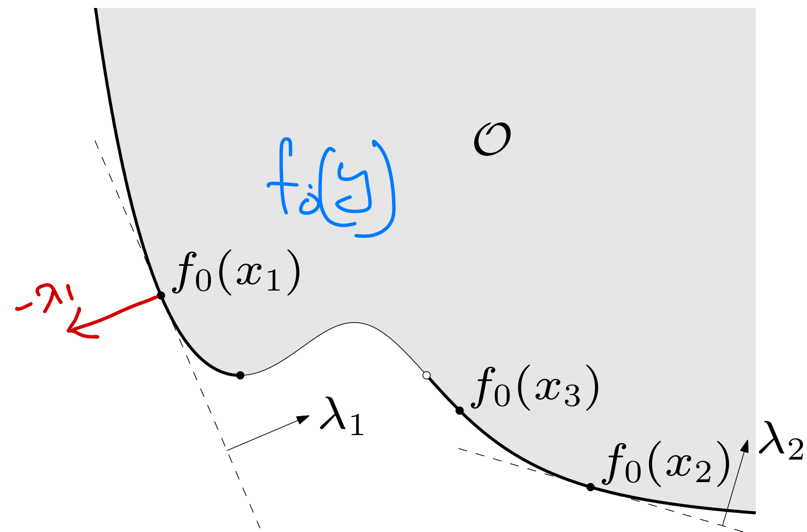
$\lambda_1^T f_0(x_1) \le \lambda_1^T f_0(y)$

$-\lambda_1^T f_0(x_1) \ge -\lambda_1^T f_0(y)$

$\lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$

if x is optimal for scalar problem, then it is Pareto-optimal for vector optimization problem

Proof below



for convex vector optimization problems, can find (almost) all Pareto optimal points by varying $\lambda \succ_{K^*} 0$

Background

Dual cones

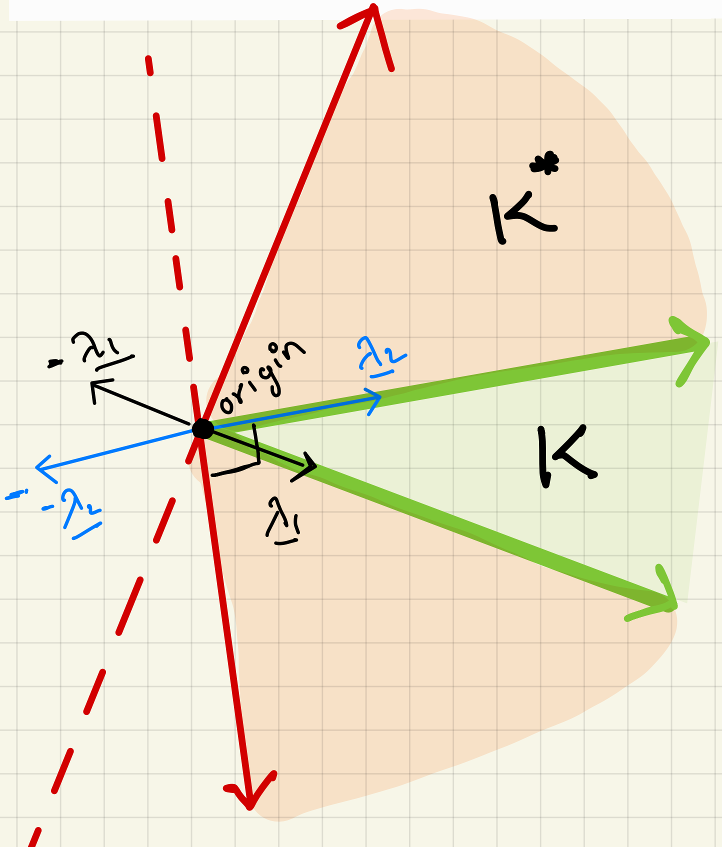
Let K be a cone. The set

$$K^* = \{y \mid x^T y \geq 0 \text{ for all } x \in K\} \tag{2.19}$$

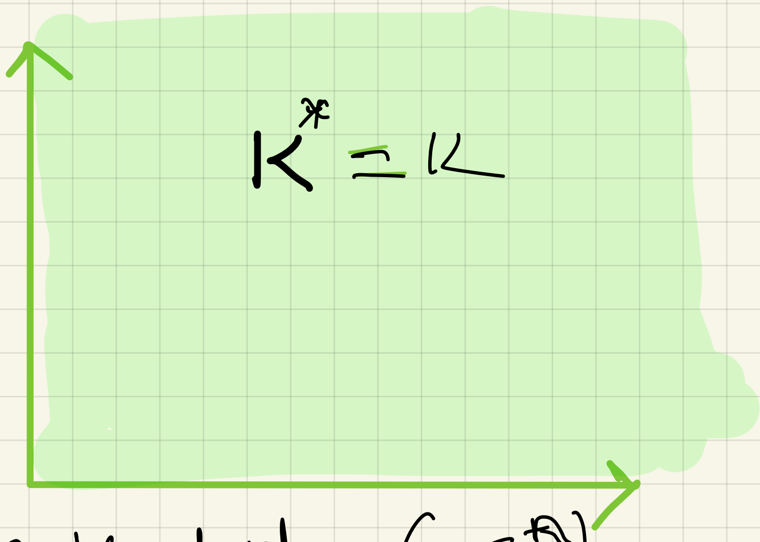
is called the *dual cone* of K. As the name suggests, K^* is a cone, and is always convex, even when the original cone K is not ([redacted] 2.31).

$x^T y \geq 0$

Geometrically, $y \in K^*$ if and only if $-y$ is the normal of a hyperplane that supports K at the origin. This is illustrated in figure [redacted].



Dual cone o $\mathbb{R}^n_+$



self dual. (S^n_+)

Background.

Dual cones satisfy several properties, such as:

- K^* is closed and convex. even when K is not convex
- $K_1 \subseteq K_2$ implies $K_2^* \subseteq K_1^*$.
- If K has nonempty interior, then K^* is pointed.
- If the closure of K is pointed then K^* has nonempty interior.
- K^{**} is the closure of the convex hull of K. (Hence if K is convex and closed, $K^{**} = K$.)

(~~See exercise 2.31~~.) These properties show that if K is a proper cone, then so is its dual K^*, and moreover, that $K^{**} = K$.

Dual generalized inequalities: Dual of $\succeq_K$ is $\succeq_{K^*}$

$\underline{x} \preceq_K \underline{y}$ iff $\underline{\lambda}^T \underline{x} \le \underline{\lambda}^T \underline{y}$ $\quad\quad \lambda \succeq_{K^*} 0$

$10 \times 2 < 3 \times 10$

Vector optimization (P_1)

$$\begin{array}{ll} \text{minimize (w.r.t. } K) & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \\ & h_i(x) = 0, \end{array}$$

scalar optimization (P_2)

$$\begin{array}{ll} \text{minimize} & \lambda^T f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \\ & h_i(x) = 0, \end{array}$$

$\underline{\lambda} \succeq_{K^*} 0$

it is Pareto optimal ← $\underline{x}$ is optimal

Proof Assume $\underline{x}$ is optimal for P_2 —— 1

Assume $\underline{x}$ are not Pareto optimal of P_1 —— 2

Then there exists a feasible $\underline{y}$ satisfying

$\underline{\lambda} \succeq_{K^*} 0$

$f_0(\underline{y}) \preceq_K f_0(\underline{x})$ and $f_0(\underline{y}) \neq f_0(\underline{x})$

$\Rightarrow \quad f_0(\underline{x}) - f_0(\underline{y}) \succ_K 0 \qquad \Rightarrow \underline{\lambda}^T (f_0(\underline{x}) - f_0(\underline{y})) \succ_K 0$

$\Rightarrow \underline{\lambda}^T f_0(\underline{x}) > \underline{\lambda}^T f_0(\underline{y})$ $\Rightarrow$ Constradicts Assumption 1

Scalarization for multicriterion problems

to find Pareto optimal points, minimize positive weighted sum

$$\lambda^T f_0(x) = \lambda_1 F_1(x) + \cdots + \lambda_q F_q(x)$$

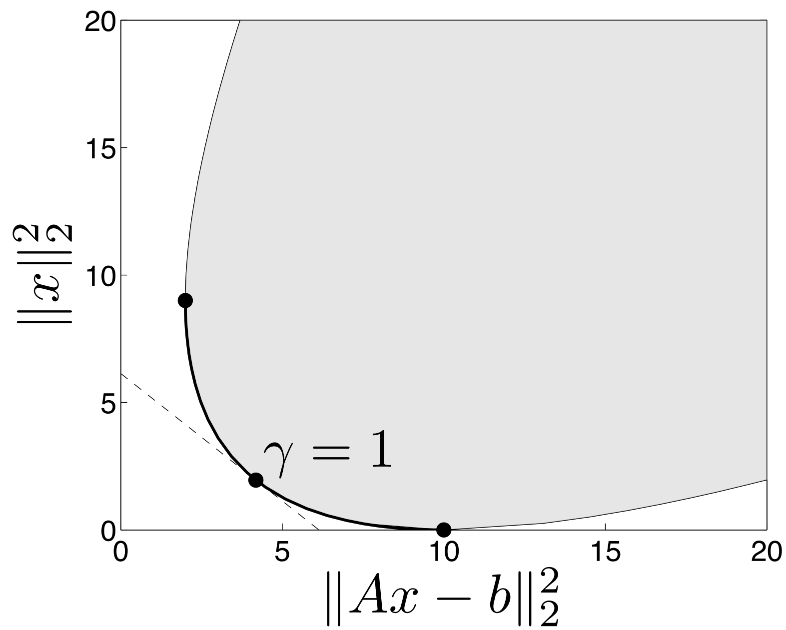
examples

- regularized least-squares problem of page 4–43

take $\lambda = (1, \gamma)$ with $\gamma > 0$

$$\text{minimize} \quad \underbrace{\|Ax - b\|_2^2}_{\lambda_1} + \underbrace{\gamma}_{\lambda_2}\|x\|_2^2$$

for fixed γ, a LS problem



$\underline{\lambda} = (1, r)$

① $r = 1 \Rightarrow \underline{\lambda} = (1, 1)$

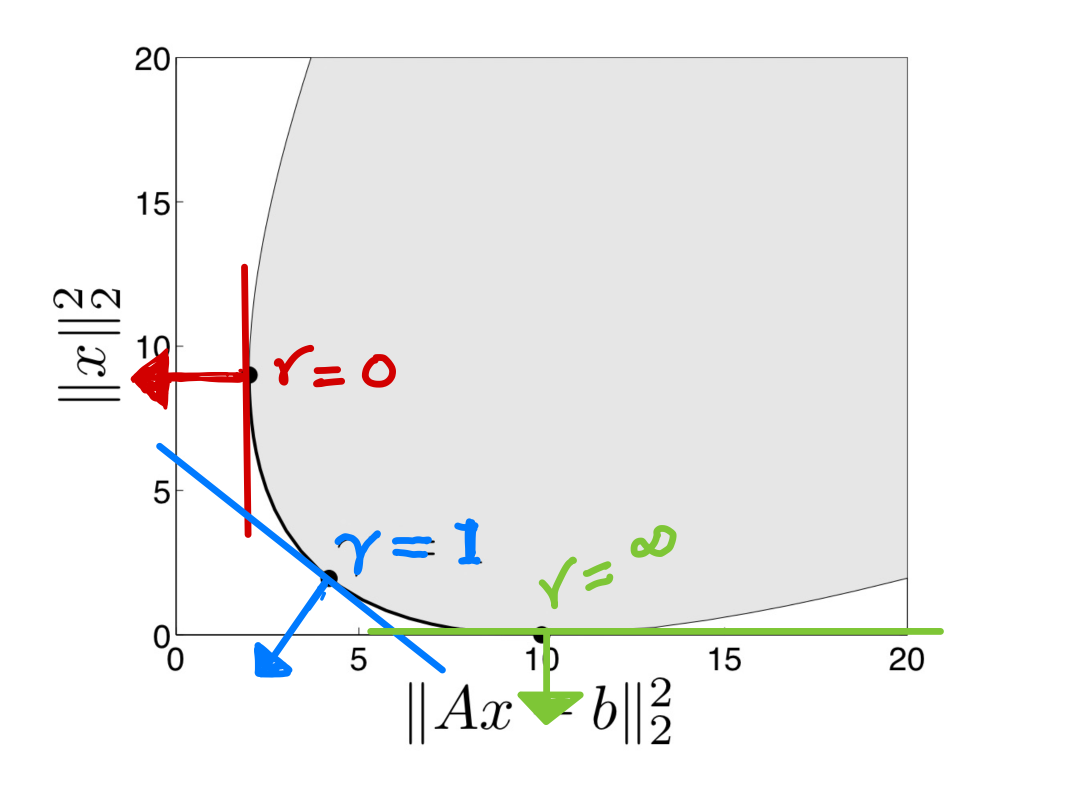
normal $(-\lambda) = (-1, -1)$

② $r = 0 \Rightarrow \underline{\lambda} = (1, 0)$

normal $(-\lambda) = (-1, 0)$

③ $r = \infty$

normal $= (0, -1)$



- risk-return trade-off of page 4–44

$$\begin{array}{ll} \text{minimize} & -\bar{p}^T x + \gamma x^T \Sigma x \\ \text{subject to} & \mathbf{1}^T x = 1, \quad x \succeq 0 \end{array}$$

for fixed $\gamma > 0$, a quadratic program