4. Convex optimization problems (part 1: general)

- optimization problem in standard form
- convex optimization problems
- quasiconvex optimization

Optimization problem in standard form

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

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

optimal value:

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

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

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

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

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

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 $x = 1$

$3x^2 - 3 = 0 \quad x = \pm 1$

$6x$

Implicit constraints

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

domain of problem including $f_0(x)$

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

- 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 ($m = p = 0$)

$b_i - a_i^T x > 0$

example:

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

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

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

standard form convex optimization problem

$$\left\{\begin{array}{ll}\text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \dots, m \\ & \underbrace{a_i^T x = b_i,} \quad i = 1, \dots, p\end{array}\right.$$

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

often written as

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

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

example

$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

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

- f_0 is convex; feasible set $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ is convex
- not a convex problem as written (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

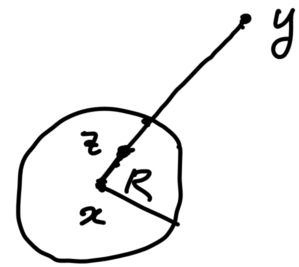
$$\begin{cases} \text{min.} & f_0(x) \\ \text{s.t.} & \underbrace{f_i(x) \le 0, \quad a_i^T x = b_i} \end{cases}$$

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

proof: suppose x is locally optimal and y is optimal with $f_0(y) < f_0(x)$

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

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



consider $z = \theta y + (1 - \theta)x$ with $\theta = R/(2\|y - x\|_2)$

$z = \theta x + ((1-\theta)y$

$\theta = \frac{R/2}{\|y-x\|}$

$z = \theta(x-y)+y$

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

$$f_0(z) \le \underbrace{\theta f_0(x) + (1 - \theta)\overset{< f_0(x)}{f_0(y)}}_{< \cancel{\theta f_0(x)} + f_0(x) - \cancel{\theta f_0(x)}} < f_0(x)$$

which contradicts our assumption that x is locally optimal

Optimality criterion for differentiable f_0

differentiable convex

$$\begin{cases} \min. & f_0(x) \\ \text{s.t.} & x \in X \to \text{convex} \end{cases}$$

x is optimal if and only if it is feasible and

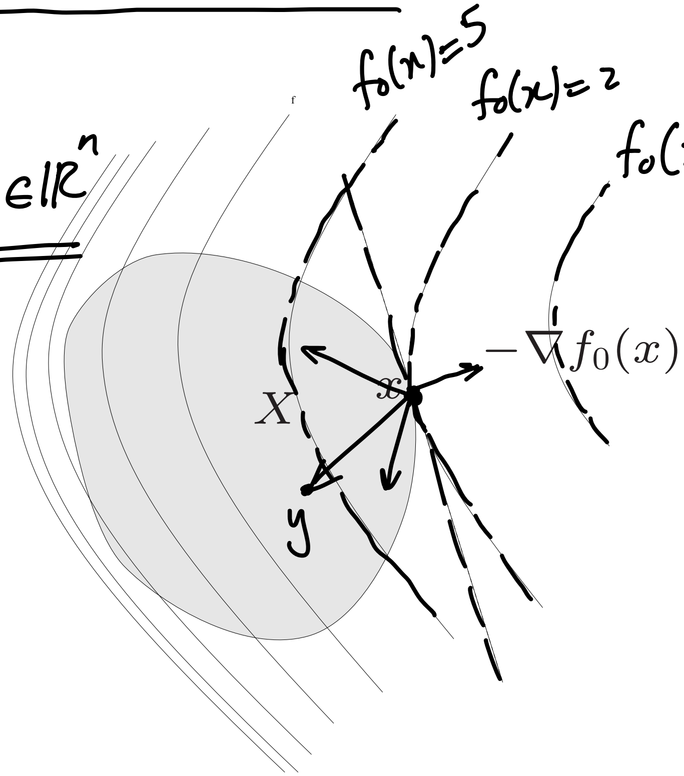
$$\nabla f_0(x)^T (y - x) \geq 0 \quad \text{for all feasible } y$$

x^*

special case: $X = \mathbb{R}^n$

$\nabla f_0(x)^T y \geqslant \nabla f_0(x)^T x, \forall y \in \mathbb{R}^n$

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



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

- **unconstrained problem**: x is optimal if and only if $X = \mathbb{R}^n$

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

- **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

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

- **minimization over nonnegative orthant**

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

x is optimal if and only if

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

$$\begin{cases} \min\limits_{x} & f_0(x) \\ \text{s.t.} & Ax = b \end{cases}$$

x is optimal iff $A\underline{x} = b$ and $x \in \operatorname{dom} f_0$ and

$\nabla f_0(x)^T \underbrace{(y - x)} \geq 0 \quad \forall\ y \in \{y \mid Ay = b\}$ → nullspace

$y \in \{x + z \mid z \in \mathcal{N}(A)\}$ $\qquad A \in \mathbb{R}^{m \times n}$

$y = x + z \qquad y - x = z$

$\Rightarrow$

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

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

$\nabla f_0(x)^T z = 0 \quad \forall\ z \in \mathcal{N}(A)$

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

$\nabla f_0(x) \in \underset{\text{range}}{\mathcal{R}}(A^T) \iff \exists\, \nu \in \mathbb{R}^m,\ \underbrace{\nabla f_0(x) = A^T \nu} \qquad \nabla f_0(x) - A^T \nu = 0$

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**

$$\left[\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & Ax = b \iff \{x_0 + y \mid \underbrace{y \in N(A)}_{y = Fz}\} \end{array}\right.$$

is equivalent to

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

$N(A) = \text{span}\{f_1, \dots, f_m\}$

where F and x_0 are such that

$$Ax = b \quad \Longleftrightarrow \quad x = \overset{[f_1 \cdots f_m]}{F}z + x_0 \text{ for some } z$$

- **introducing equality constraints**

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

$y_i = A_i x + b_i$

is equivalent to

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

- **introducing slack variables for linear inequalities**

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

is equivalent to

$s \in \mathbb{R}^m$

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

slack variable

$$\begin{cases} \min_x & f_0(x) \\ \text{s.t.} & f_i(x) \le 0 \\ & Ax = b \end{cases}$$

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

(x^*, t^*)

$f_0(x^*) = t^*$ why?

if $f_0(x^*) < t^*$

then t^* can be

further reduced

$$\begin{array}{ll} \text{minimize (over } x,\ t\text{)} \quad t \in \mathbb{R} & t \\ \text{subject to} & f_0(x) - t \le 0 \qquad f_0(x) \le t \\ & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- **minimizing over some variables**

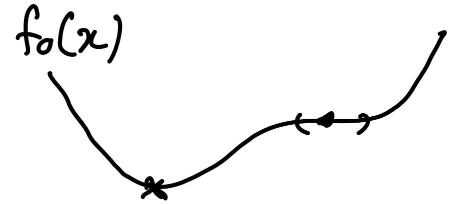
$$\begin{array}{ll} \text{minimize} & f_0(x_1, x_2) \\ \text{subject to} & f_i(x_1) \le 0, \quad i = 1, \ldots, m \end{array}$$

is equivalent to

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

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

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.*,

convex

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

example

$$f_0(x) = \frac{p(x)}{q(x)}$$

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

$f_0(x) \leq t$ $\quad \frac{p(x)}{q(x)} \leq t \implies p(x) \leq t\,q(x)$ $\quad \underbrace{p(x) - t\,q(x)}_{\phi_t(x)} \leq 0$

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

- for $t \geq 0$, ϕ_t convex in x (sum of 2 convex fcts)
- $p(x)/q(x) \leq t$ if and only if $\phi_t(x) \leq 0$

quasiconvex optimization via convex feasibility problems

for any fixed t, finding an x is a convex feas. problem:

$$\phi_t(x) \le 0, \qquad f_i(x) \le 0, \quad i = 1, \ldots, m, \qquad Ax = b \qquad (1)$$

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

Bisection method for quasiconvex optimization

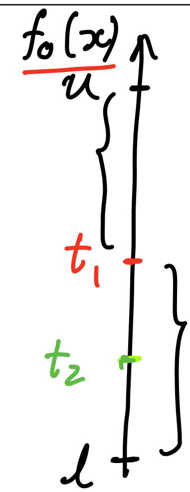
given $l \le p^\star$, $u \ge p^\star$, tolerance $\epsilon > 0$.

repeat

1. $t := (l + u)/2$.
2. Solve the convex feasibility problem (1).
3. **if** (1) is feasible, $u := t$; **else** $l := t$.

until $u - l \le \epsilon$.

accuracy in $f_0(x^*)$



fix t. can I find x that achieves that t?

outer

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