Optimization problem: standard form

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

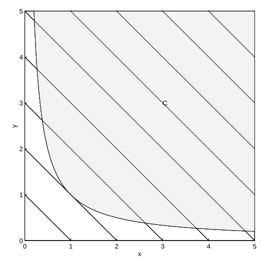
where $f_i, h_i : \mathbf{R}^n \to \mathbf{R}$

- x is *optimization variable*
- f_0 is *objective* or *cost function*;
  $f_i(x) \le 0$ are the *inequality constraints*;
  $h_i(x) = 0$ are the *equality constraints*
- x is *feasible* if it satisfies the constraints;
  the *feasible set* C is the set of all feasible points;
  problem is *feasible* if there are feasible points
- problem is *unconstrained* if $m = p = 0$
- *optimal value* is $f^\star = \mathbf{inf}_{x \in C} f_0(x)$
  (can be $-\infty$; convention: $f^\star = +\infty$ if infeasible);
  optimal point: $x \in C$ s.t. $f(x) = f^\star$;
  optimal set: $X_{\rm opt} = \{x \in C \mid f(x) = f^\star\}$

example:

$$\begin{array}{ll} \text{minimize} & x_1 + x_2 \\ \text{subject to} & -x_1 \le 0 \\ & -x_2 \le 0 \\ & 1 - x_1 x_2 \le 0 \end{array} \qquad (1)$$

- feasible set C is half-hyperboloid
- optimal value is $f^\star = 2$
- only optimal point is $x^\star = (1, 1)$



Implicit and explicit constraints

explicit constraints: $f_i(x) \le 0$, $h_i(x) = 0$

implicit constraint: $x \in \mathbf{dom}\, f_i$, $x \in \mathbf{dom}\, h_i$

$D = \mathbf{dom}\, f_0 \cap \ldots \cap \mathbf{dom}\, f_m \cap \mathbf{dom}\, h_1 \cap \ldots \cap \mathbf{dom}\, h_p$

is called *domain of the problem*

example

$$\begin{array}{ll} \text{minimize} & -\log x_1 - \log x_2 \\ \text{subject to} & x_1 + x_2 - 1 \le 0 \end{array}$$

has an implicit constraint

$$x \in D = \{x \in \mathbf{R}^2 \mid x_1 > 0, x_2 > 0\}$$

Feasibility problem

suppose objective $f_0 = 0$, so

$$f^\star = \begin{cases} 0 & \text{if } C \ne \emptyset \\ +\infty & \text{if } C = \emptyset \end{cases}$$

thus, problem is really to

- either find $x \in C$
- or determine that $C = \emptyset$

i.e., solve the inequality / equality system

$$\begin{array}{l} f_i(x) \le 0, \quad i = 1, \dots, m \\ h_i(x) = 0, \quad i = 1, \dots, p \end{array}$$

or determine that it is inconsistent

Convex optimization problem

convex optimization problem in standard form:

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

- $f_0, f_1, \ldots, f_m$ convex
- affine equality constraints
- feasible set is convex

often written as

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

where $A \in \mathbf{R}^{p \times n}$

example. problem (1)

- has convex objective and feasible set
- is **not** a standard form convex optimization problem since $f_3(x) = 1 - x_1 x_2$ is not convex

can be easily cast as standard form convex optimization problem:

$$\begin{array}{ll}\text{minimize} & x_1 + x_2\\ \text{subject to} & -x_1 \le 0\\ & -x_2 \le 0\\ & 1 - \sqrt{x_1 x_2} \le 0\end{array}$$

($1 - \sqrt{x_1 x_2}$ is convex on $\mathbf{R}_+^2$)

many different ways, *e.g.*,

$$\begin{array}{ll}\text{minimize} & x_1 + x_2\\ \text{subject to} & -x_1 \le 0\\ & -x_2 \le 0\\ & -\log x_1 - \log x_2 \le 0\end{array}$$

example. f_i all affine yields *linear program*

$$\begin{array}{ll}\text{minimize} & c_0^T x + d_0\\ \text{subject to} & c_i^T x + d_i \le 0, \ i = 1, \ldots, m\\ & Ax = b\end{array}$$

which is a convex optimization problem

example. minimum norm approximation with limits on variables

$$\begin{array}{ll}\text{minimize} & \|Ax - b\|\\ \text{subject to} & l_i \le x_i \le u_i, \ i = 1, \ldots, n\end{array}$$

is convex

example. maximum entropy with linear equality constraints

$$\begin{array}{ll}\text{minimize} & \sum_i x_i \log x_i\\ \text{subject to} & x_i \ge 0, \ i = 1, \ldots, n\\ & \sum_i x_i = 1\\ & Ax = b\end{array}$$

is convex

(more on these later)

Local and global optimality

$x \in C$ is *locally optimal* if it satisfies

$$y \in C, \ \|y - x\| \le R \implies f_0(y) \ge f_0(x)$$

for some $R > 0$

c.f. (globally) optimal, which means $x \in C$,

$$y \in C \implies f_0(y) \ge f_0(x)$$

for cvx opt problems, any local solution is also global

proof:

- suppose x is locally optimal, but $y \in C$, $f_0(y) < f_0(x)$
- take small step from x towards y, *i.e.*, $z = \lambda y + (1 - \lambda)x$ with $\lambda > 0$ small
- z is near x, with $f_0(z) < f_0(x)$; contradicts local optimality

Quasiconvex optimization problem

quasiconvex optimization problem in standard form:

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

- f_0 is **quasiconvex**
- $f_1, \ldots, f_m$ are convex
- affine equality constraints
- feasible set, all sublevel sets are convex

example. linear-fractional programming

$$\begin{array}{ll} \text{minimize} & (a^Tx + b)/(c^Tx + d) \\ \text{subject to} & Ax = b, \quad Fx \preceq g, \quad c^Tx + d > 0 \end{array}$$

Solving quasiconvex problems via bisection

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

f_i convex, f_0 quasiconvex

idea: express sublevel set $f_0(x) \le t$ as sublevel set of *convex* function:

$$f_0(x) \le t \;\Leftrightarrow\; \phi(x, t) \le 0$$

where $\phi : \mathbf{R}^{n+1} \to \mathbf{R}$ is convex in x for each t

now solve quasiconvex problem by bisection on t, solving convex feasibility problem

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

(with variable x) at each iteration

bisection method for quasiconvex problem:

```
given l < p*; feasible x; ε > 0
u := f0(x)
repeat
   t := (u + l)/2
   solve convex feasibility problem
      φ(x, t) ≤ 0, fi(x) ≤ 0, Ax = b
   if feasible,
      u := t
      x := any solution of feas. problem
   else l := t
until u − l ≤ ε
```

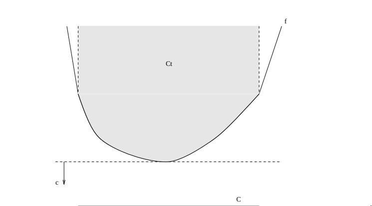
- reduces quasiconvex problem to sequence of convex feasibility problems
- finds ϵ-suboptimal solution in $\log_2(1/\epsilon)$ iterations

Epigraph form

write standard form problem as

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

- variables are (x, t)
- $m + 1$ inequality constraints
- objective is *linear*: $t = e_{n+1}^T(x, t)$
- if original problem is cvx, so is epigraph form



linear objective is 'universal' for convex optimization

Standard form with generalized inequalities

convex optimization problem in *standard form with generalized inequalities*:

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

where:

- $f_0 : \mathbf{R}^n \to \mathbf{R}$ convex
- $\preceq_{K_i}$ are generalized inequalities on $\mathbf{R}^{m_i}$
- $f_i : \mathbf{R}^n \to \mathbf{R}^{m_i}$ are K_i-convex

example. *semidefinite programming*

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & A_0 + x_1 A_1 + \cdots + x_n A_n \preceq 0 \end{array}$$

where $A_i \in \mathbf{R}^{p \times p}$

How f_i, h_i are described

analytical form

functions can have analytical form, *e.g.*,

$$f(x) = x^T P x + 2q^T x + r$$

f is specified by giving the problem *data, coefficients, or parameters*, *e.g.*

$$P = P^T \in \mathbf{R}^{n \times n}, \quad q \in \mathbf{R}^n, \quad r \in \mathbf{R}$$

oracle form

functions can be given by *oracle* or *subroutine* that, given x, computes $f(x)$ (and maybe $\nabla f(x)$, $\nabla^2 f(x)$, ...)

- oracle model can be useful even if f has analytic form, *e.g.*, linear but sparse
- how f given affects choice of algorithm, storage required to specify problem, etc.

Some hard problems

'slight' modification of convex problem can be very hard

- convex maximization, concave minimization, *e.g.*

$$\begin{array}{ll} \text{maximize} & \|x\| \\ \text{subject to} & Ax \preceq b \end{array}$$

- nonlinear equality constraints, *e.g.*

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x^T P_i x + q_i^T x + r_i = 0, \; i = 1, \ldots, K \end{array}$$

- minimizing over non-convex sets, *e.g.*, Boolean variables

$$\begin{array}{ll} \text{find} & x \\ \text{such that} & Ax \preceq b, \\ & x_i \in \{0, 1\} \end{array}$$

Multicriterion optimization

vector objective

$$F(x) = (F_1(x), \ldots, F_N(x))$$

$F_1, \ldots, F_N : \mathbf{R}^n \to \mathbf{R}$
(can include equality, inequality constraints)

F_i called *objective functions*: roughly speaking, want all F_i small

family of *specifications* indexed by $t \in \mathbf{R}^N$:

$$F(x) \preceq t$$

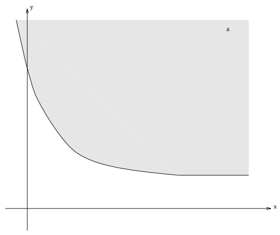
i.e., $F_i(x) \le t_i$, $i = 1, \ldots, N$.

achievable specification: t s.t. $F(x) \preceq t$ feasible

Achievable specifications

set of achievable objectives:

$$\mathcal{A} = \{t \in \mathbf{R}^N \mid \exists x \text{ s.t. } F(x) \preceq t\}$$



if F_i are convex then $\mathcal{A}$ is convex

$\mathcal{A}$ is projection of *vector function epigraph*

$$\mathbf{epi}\,(F) = \{(x,t) \in \mathbf{R}^n \times \mathbf{R}^N \,| F(x) \preceq t\}$$

on t-space.

boundary of $\mathcal{A}$ is called (optimal) *tradeoff surface*

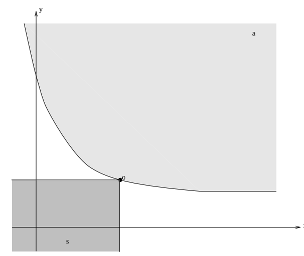
Pareto optimality

x *dominates* (is better than) $\tilde{x}$ if $F(x) \neq F(\tilde{x})$

$$F(x) \preceq F(\tilde{x})$$

i.e., x is no worse than $\tilde{x}$ in any objective, and better in at least one

x_0 is *Pareto optimal* if no x dominates it

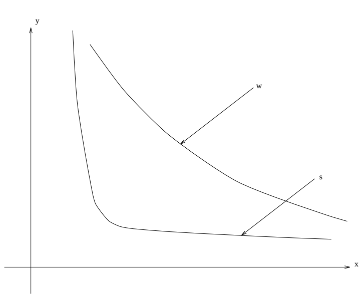


roughly, x_0 Pareto optimal means $F(x_0)$ is on tradeoff surface ($F(x_0) \in \partial\mathcal{A}$)

Pareto problem: find Pareto-optimal x

real (but more vague) engineering problem:
search/explore/characterize tradeoff surface, *e.g.*:

- 'can reduce F_5 below 0.1, but only at huge cost in F_4 and F_2'
- 'can pretty much minimize F_3 independently of other objectives'
- 'F_1 and F_2 tradeoff strongly for $F_1 \leq 1$, $F_2 \leq 2$'



Scalarization

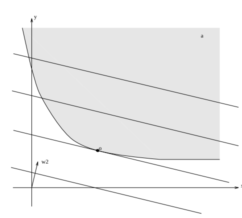
multicriterion problem with $F_1, \ldots, F_N$

minimize weighted sum of objectives: choose weights $w_i > 0$, and solve

$$\text{minimize } \sum_i w_i F_i(x)$$

which is the same as

$$\begin{array}{ll} \text{minimize} & w^T t \\ \text{subject to} & F(x) \leq t \quad (\textit{i.e.}, t \in \mathcal{A}) \end{array}$$



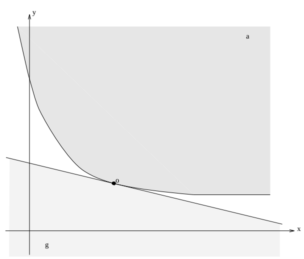
- solution x_0 is Pareto optimal
- for many cvx problems, all Pareto optimal points can be found this way, as weights vary

interpretation

- hyperplane $w^T t = w^T F(x_0)$ supports $\mathcal{A}$ at $F(x_0)$
- specifications in halfspace

$$\{t \mid w^T t < w^T F(x_0)\}$$

  are unachievable



Restriction and relaxation

original problem, with optimal value $f^\star$:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in C \end{array}$$

new problem, with optimal value $\tilde{f}^\star$:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in \tilde{C} \end{array}$$

new problem is

- *relaxation* (of original) if $\tilde{C} \supseteq C$
  (in which case $\tilde{f}^\star \leq f^\star$)
- *restriction* if $\tilde{C} \subseteq C$
  (in which case $\tilde{f}^\star \geq f^\star$)

Example. f is convex, C is nonconvex; $\tilde{C} = \mathbf{Co}\, C$

relaxation is convex problem that gives lower bound for original, nonconvex problem